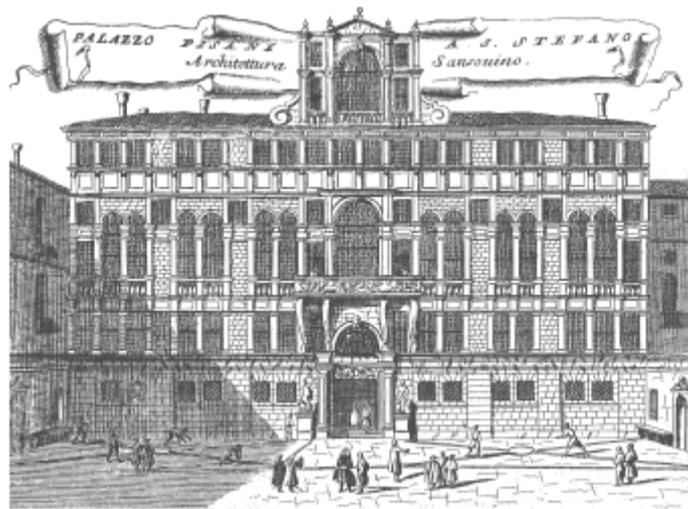


# XVII Colloquio di Informatica Musicale



Il Palazzo Pisani di Santo Stefano sede del Conservatorio di Musica di Venezia in una incisione del Cornelli, sec. XVII

**Sala Concerti del Conservatorio “Benedetto Marcello”  
Palazzo Pisani  
15 - 17 ottobre 2008**



Associazione di Informatica  
Musicale Italiana



Conservatorio “B. Marcello”  
di Venezia



Ex Novo Ensemble

I  
-  
U  
-  
A  
-  
V

Facoltà di Design e Arti  
dell’Università IUAV di  
Venezia



La Biennale di Venezia

**Coordinamento**

Davide Rocchesso, Aldo Orvieto, Alvise Vidolin.

**Redazione**

Maria Grazia Ballerano

**Informazioni**

<http://www.aimi-musica.org/>

<http://www.labiennale.org/it/musica/>

<http://www.conseve.it>

## **Comitato Scientifico**

Nicola Bernardini, Conservatorio di Padova

Roberto Bresin, KTH, Stoccolma

Antonio Camurri, Università di Genova

Sergio Canazza, Università di Udine

Giovanni Cospito, Conservatorio di Como

Giovanni De Poli, Università di Padova

Riccardo Dapelo, Conservatorio di Genova

Agostino Di Scipio, Conservatorio di Napoli

Roberto Doati, Conservatorio di Genova

Federico Fontana, Università di Verona

Goffredo Haus, Università di Milano

Vincenzo Lombardo, Università di Torino

Michelangelo Lupone, Conservatorio de L'Aquila

Aldo Orvieto, Conservatorio di Padova

Davide Rocchesso, Università IUAV di Venezia

Sylviane Sapis, Conservatorio di Como

Augusto Sarti, Politecnico di Milano

Davide Tiso, Accademia di Belle Arti di Venezia

Andrea Valle, Università di Torino

Alvise Vidolin, Conservatorio di Venezia

Gualtiero Volpe, Università di Genova

Laura Zattra, Università di Padova



## Indice

<b>Pag.</b>	<b>Contenuto</b> <b>Autore</b>
3.	Programma
7.	“NoaNoa” for Flute and Electronics: problemi e strategie di esecuzione della parte elettronica. <i>Simone Conforti</i>
17.	Il ruolo del movimento corporeo nell’esecuzione musicale in presenza di feedback uditivo ritardato <i>Marco Matteazzi</i>
23.	Dalla Computer Music al Sound and Music Computing: Contesti culturali e tecnologici <i>Nicola Bernardini, Alvise Vidolin</i>
29.	Voice Painter: un’interfaccia multimodale per dipingere con la voce <i>Amalia de Götzen, Riccardo Marogna, Federico Avanzini</i>
35.	Un modello per la generazione dinamica di paesaggi sonori <i>Mattia Schirosa, Andrea Valle, Vincenzo Lombardo</i>
43.	Del suonare l’elettronica <i>Gabriel Bacalov, Massimo Marchi, Francesco Vitale</i>
47.	Principi elementari per una sonificazione del gesto <i>Maurizio Goina, Pietro Polotti</i>
53.	Meta-Morfologie Musicali: un ambiente di composizione elettroacustica <i>Riccardo Dapelo, Roberto Doati, Emilio Pozzolini, Alessandro Quaranta, Martino Sarolli, Corrado Canepa, Barbara Mazzarino</i>
59.	Sostituzione Incrociata della Fase nel Dominio di Scala <i>Antonio De Sena</i>
65.	La Computer Vision per l’annotazione automatica di documenti audio <i>Lauro Snidaro, Sergio Canazza</i>
69.	Analisi del comportamento di fruitori di sistemi di ascolto attivo: il caso dell’Esploratore d’Orchestra <i>Francesca Cavallero, Antonio Camurri, Corrado Canepa, Gualtiero Volpe</i>
77.	Modelli VCF discreti a retroazione istantanea <i>Marco Civolani, Federico Fontana</i>
85.	La ricostruzione dello Studio di Fonologia della Rai di Milano <i>Maddalena Novati</i>
89.	Un approccio alle interfacce acustiche tangibili basato sulla trasformata wavelet di onde dispersive <i>Antonio Camurri, Corrado Canepa, Chiara Erra, Alberto Massari, Gualtiero Volpe, Luciano Fava</i>

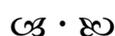
95. Un modello per la simulazione in tempo reale di un riverbero a piastra  
*Stefano Zambon*
103. Un campionatore per "Musica su due dimensioni" di Bruno Maderna  
*Marco Gasperini*
109. Sviluppo di un'interfaccia audio-aptica basata sulla profondità spaziale  
*Stefano Delle Monache, Delphine Devallez, Pietro Polotti, Davide Rocchesso*
115. Omaggio a Stockhausen - Technical set-up digitale per una performance di SOLO [Nr.19] für melodieninstrument mit rückkopplung  
*Enrico Francioni*
125. Comporre (con) oggetti audio-visivi: un ambiente di lavoro in via di sviluppo  
*Alessio Santini*
131. Generazione di suoni e musica mediante sistemi caotici  
*Costantino Rizzuti*
139. DepThrow: uno strumento di indagine sulla percezione uditiva della distanza in forma di gioco audio  
*Stefano Papetti, Delphine Devallez, Federico Fontana*
145. CHORALE - Compositional Hyper ORganized ALgorithmic Environment  
*Gianni Della Vittoria*
149. Suoni scolpiti e sculture sonore: alcuni esempi di installazioni d'arte elettroacustica  
*Silvia Lanzalone*
157. Oggetti sonori nella rete  
*Dante Tanzi*
161. Rendering acustico di liquidi in movimento rappresentati con modelli a particelle  
*Carlo Drioli, Davide Rocchesso*
165. Echi tra le Volte: un progetto di sound design rivolto agli edifici di culto  
*Andrea Taroppi*
171. Photos of GHOSTS (Photos of Grooves and HOles, Supporting Tracks Separation)  
Conservazione attiva di dischi fonografici per mezzo di immagini digitali  
*Sergio Canazza, Giovanni Ferrin, Lauro Snidaro*
177. Atomi distratti di Mario Garuti (2005) - Analisi musicale e tecnologica ed elaborazione di un algoritmo di sintesi in ambiente Max/MSP per l'esecuzione della parte di tastiera elettronica  
*Marco Marinoni*
191. AudioDrome  
*Alessandro Petrolati*
199. Score Following in Tempo Reale basato su Hidden Markov Models  
*Nicola Montecchio*

## Programma



### 15 Ottobre

- 10.30** Apertura lavori, saluto delle autorità. Giovanni Umberto Battel, Direttore del Conservatorio "B. Marcello" di Venezia; Sylviane Sapir, Presidente della Associazione di Informatica Musicale Italiana; Claudio Ambrosini, Presidente dell' Ex Novo Ensemble; Medardo Chiapponi, Preside della Facoltà di Design e Arti dell' Università IUAV di Venezia; Luca Francesconi, Direttore del Settore Musica de La Biennale di Venezia
- 11.00** Sessione scientifica 1. Presiede: Riccardo Dapelo  
Emanuele Casale *Musica elettroacustica*; Simone Conforti *Noa Noa: problemi e strategie di esecuzione della parte elettronica*; Marco Matteazzi *Il ruolo del movimento corporeo nell'esecuzione musicale in presenza di feedback uditivo ritardato*; Nicola Bernardini e Alvisè Vidolin *Dalla computer music al sound and music computing: contesti culturali e tecnologici*
- 12.30** Pausa pranzo
- 14.30** Sessione Demo / poster 1. Presiede: Sergio Canazza  
Amalia De Götzen, Riccardo Marogna e Federico Avanzini *Voice painter: un'interfaccia multimodale per dipingere con la voce*; Mattia Schirosa, Andrea Valle e Vincenzo Lombardo *Un modello per la generazione dinamica di paesaggi sonori*; Gabriel Bacalov, Massimo Marchi e Francesco Vitale *Del suonare l'elettronica*; Maurizio Goina e Pietro Polotti *Principi elementari per una sonificazione del gesto*; Riccardo Dapelo, Roberto Doati, Emilio Pozzolini, Alessandro Quaranta, Martino Sarolli, Corrado Canepa, Barbara Mazzarino *Meta-Morfologie musicali: un ambiente di composizione elettroacustica*; Antonio De Sena *Sostituzione incrociata della fase nel dominio di scala*
- 16.30** Pausa Caffè
- 17.00** Concerto  
Kaja Saariaho, *Noa noa* per flauto e live electronics (1992, 9')  
Emanuele Casale, *5* per flauto, clarinetto ed elettronica (2003, 9')  
Karlheinz Stockhausen, *Kontakte per pianoforte*, percussioni e nastro magnetico (1959/60, 35')  
flauto Daniele Ruggieri  
clarinetto Davide Teodoro  
percussioni Annunziata Dellisanti  
pianoforte Aldo Orvieto  
regia del suono Alvisè Vidolin  
**Ex Novo Ensemble**  
in collaborazione con scuola di Musica Elettronica del Conservatorio Benedetto Marcello
- 19.00** Fine lavori

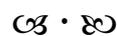


## 16 Ottobre

- 9.00** Sessione scientifica 2. Presiede: Nicola Bernardini  
Lauro Snidaro e Sergio Canazza *La computer vision per l'annotazione automatica di documenti audio*; Francesca Cavallero, Antonio Camurri, Corrado Canepa, Gualtiero Volpe *Analisi del comportamento di fruitori di sistemi di ascolto attivo: il caso dell'esploratore d'orchestra*; Marco Civolani e Federico Fontana *Modelli VCF discreti a retroazione istantanea*
- 10.30** Pausa caffè
- 11.00** Sessione scientifica 3. Presiede: Roberto Bresin  
Maddalena Novati, *La ricostruzione dello Studio di Fonologia della Rai di Milano*; Corrado Canepa, Antonio Camurri, Chiara Erra, Alberto Massari, Gualtiero Volpe e Luciano Fava, *Un approccio alle interfacce acustiche tangibili basato sulla trasformata wavelet di onde dispersive*; Stefano Zambon *Un modello per la simulazione in tempo reale di un riverbero a piastra*; Marco Gasperini *Un campionario per "Musica su due dimensioni" di Bruno Maderna*
- 12.30** Pausa pranzo
- 14.30** Sessione Demo / poster 2. Presiede: Sylviane Sapir  
Stefano Delle Monache, Delphine Devallez, Pietro Polotti, Davide Rocchesso *Sviluppo di un'interfaccia audio-aptica basata sulla profondità spaziale*; Enrico Francioni *Omaggio a Stockhausen: technical-setup digitale per una performance di "Solo"*; Alessio Santini *Comporre (con) oggetti audio-visivi: un ambiente di lavoro in via di sviluppo*; Costantino Rizzuti *Generazione di suoni e musica mediante sistemi caotici*; Stefano Papetti, Delphine Devallez e Federico Fontana *DepThrow: uno strumento di indagine sulla percezione uditiva della distanza in formadi gioco audio*; Gianni Della Vittoria *CHORALE: Compositional Hyper ORganized ALgorithmic Environment*
- 16.30** Pausa caffè
- 17.00** Concerto per Ipertastiere e live electronics  
composizioni per tastiere tradizionali ed elettroniche selezionate dal comitato scientifico del CIM – XVII Colloquio di Informatica Musicale  
Marco Marinoni, *Il grifo nelle perle nere* per pianoforte e ipertastiere (2008, 7') prima es. ass.  
Nicola Buso, *Last Transition* per pianoforte e live coding (2008, 8'), prima es. ass.  
Julian Scordato, *Cross Paths* per pianoforte e live electronics (2008, 5') prima es. ass.  
Roberto Girolin, *Piega di palpebra* per pianoforte, operatore sulla cordiera ed elettronica (2008, 12') prima es. ass.

Nicola Ferrari, *Lontano in alto mare, l'acqua è azzurra* canzone a due pianoforti e sistema EyesWeb  
(2008, 15') prima es. ass.  
Edoardo Lattes, *Lo specchio della regina* per pianoforte e live electronics  
(2008, 8') prima es. ass.  
Luca Richielli, *Lo specchio incantato* per disklavier e ipersampler  
(2008, 8') prima es. ass.  
tastiere tradizionali Aldo Orvieto  
tastiere digitali Davide Tiso  
live electronics e regia del suono Alvisè Vidolin  
in collaborazione con scuola di Musica Elettronica del Conservatorio Benedetto Marcello

**19.00** Fine lavori



## **17 Ottobre**

**9.00** Sessione scientifica 4. Presiede: Davide Rocchesso  
Antonio Rodà e Laura Zattra *Musica Endoscopica: la ricerca sonora di Teresa Rampazzi*; Silvia Lanzalone *Suoni scolpiti e sculture sonore: alcuni esempi di installazioni d'arte elettroacustica*; Dante Tanzi *Oggetti sonori nella rete*; Carlo Drioli e Davide Rocchesso *Rendering acustico di liquidi in movimento rappresentati con modelli a particelle*

**10.30** Pausa caffè

**11.00** Sessione Demo / poster 3. Presiede: Gualtiero Volpe  
Andrea Taroppi *Echi tra le Volte: un progetto di sound design rivolto agli edifici di culto*; Sergio Canazza, Giovanni Ferrin e Lauro Snidaro *Photos of GHOSTS: conservazione attiva di dischi fonografici per mezzo di immagini digitali*; Marco Marinoni *"Atomi distratti" di Mario Garuti (2005): analisi musicale e tecnologica ed elaborazione di un algoritmo di sintesi*; Alessandro Petrolati *Audio Drome*; Nicola Montecchio *Score following in tempo reale basato su Hidden Markov Models*

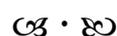
**12.30** Pausa pranzo

**14.30** Tavola rotonda *Sviluppi dell'informatica musicale fra università e conservatorio*.  
Moderatore Sylviane Sapir

**16.30** Pausa caffè

**17.00** *Installazione sonora*  
Carlo De Pirro, *Il tempo sospeso* per disklavier e suoni elettronici (2007)  
a cura di Marco Gasperini  
in collaborazione con scuola di Musica Elettronica  
del Conservatorio Benedetto Marcello

**19.00** Fine lavori





# “NoaNoa” for Flute and Electronics: problemi e strategie di esecuzione della parte elettronica.

**Simone Conforti**

MARTLab

(Conservatorio di Musica “L. Cherubini” di  
Firenze)

## SOMMARIO

Seguendo una prassi, o per meglio dire una consuetudine, sempre più diffusa, anche il brano *NoaNoa* di Kaija Saariaho, si va ad inserire in quella lunga lista di composizioni che fanno uso dei mezzi elettronici e che, col passare degli anni, rischiano e rischieranno sempre di più di finire nel limbo dell'ineseguibilità.

L'obsolescenza dei mezzi tecnologici porta infatti alla necessità di dover sviluppare uno studio e di conseguenza una prassi, capace di rendere l'elaborazione elettronica in musica meno soggetta al legame con la tecnologia con cui è stata realizzata.

Partendo da questi presupposti il presente lavoro cerca di risolvere le discrepanze, fra le varie versioni e descrizioni della parte elettronica del brano in esame, e cerca allo stesso tempo di proporre dei modelli su cui sviluppare una traducibilità dei processi effettuati dai mezzi elettronici, attraverso la ricerca di una scrittura che riesca ed essere completa, esaustiva ed universale sul modello della notazione tradizionale.

## PREMESSE AD UN LAVORO DI SALVAGUARDIA

Seguendo una prassi, o per meglio dire una consuetudine, sempre più spesso diffusa, anche il brano *NoaNoa* di Kaija Saariaho, si va ad inserire in quella lunga lista di composizioni che fanno uso dei mezzi elettronici e che col passare degli anni sono e saranno sempre più a rischio di finire nel limbo della ineseguibilità.

Ma da cosa dipende questa pratica che porterà, a meno che non venga fatta un'opera di preservazione mirata, alla scomparsa di molte opere con l'elettronica?

Troppo spesso potremmo dire dalla pigrizia, intesa come mancanza di lungimiranza e anche di investimento, da parte delle istituzioni che focalizzano l'attenzione sull'atto performativo e non su quello creativo nel suo insieme; altre volte invece la mancanza di una necessaria documentazione dell'opera in tutte le sue parti può dipendere dalla assenza di tempi adeguati per poter curare la composizione in ogni suo aspetto; altre volte è anche possibile ricondurre tale fenomeno ad una certa pratica, che si è venuta diffondendo particolarmente nell'area di produzione che gravita intorno all'IRCAM e alla quale appartiene anche il brano in esame, la quale è animata dalla convinzione che l'utilizzo di determinati softwares, più o meno condivisi da tutta la comunità, siano garanzia di esportabilità e replicabilità delle opere. Vedremo di seguito come tale convinzione possa rivelarsi errata.

Certo giustificare, solo demonizzando, un certo atteggiamento, potrebbe risultare assai riduttivo, dato che il problema della

scrittura della parte elettronica nelle composizioni che ne fanno uso è un quesito che ad oggi non trova soluzione. E molto spesso è condivisibile l'imbarazzo in cui il compositore può venirsi a trovare al momento di definire sulla carta le tecnologie usate.

Allo stesso tempo ci potremmo infatti domandare fino a che punto la strada tracciata da Luigi Nono, per citare un esempio illustre, nella stesura di una certa notazione della musica elettronica, sia sufficiente o adeguata al virtuosismo tecnico e musicale che i mezzi informatici hanno raggiunto anche nella prassi e nella scrittura della musica contemporanea.

Per quanto complesse e ricche di spunti in molteplici direzioni, nel caso del brano in esame, le problematiche sottolineate, potrebbero essere non prese in considerazione in funzione della peculiarità del pezzo, che si attesta fra quelli dove la tipologia d'intervento elettronico, potrebbe essere definita con un buon dettaglio anche sulla carta.

Qui giustificazioni tecniche che rendano giustizia alla mancanza di informazioni dettagliate, per ciò che riguarda la parte elettronica, possono essere difficilmente rintracciate, ma potremmo piuttosto ricondurre tale fenomeno a quella prassi-consuetudine che tende a caratterizzare sempre più spesso i lavori prodotti all'interno dell'IRCAM.

La logica che sta alla base di questo atteggiamento nel caso della Saariaho, prende vita dall'idea che debba essere fatto un continuo rinnovamento dei programmi elettronici per l'esecuzione dei suoi brani, tenendo così il passo con l'evoluzione dei sistemi informatici. Il problema è che tale atteggiamento, nonostante sia sostenuto da una volontà che è poco presente in altri compositori, non riesce ugualmente a soddisfare le funzioni di trasmissibilità dell'opera sia nel breve che nel lungo tempo.

Proprio nella mia personale e recente esperienza è venuto fuori questo vizio di forma, che si è verificato proprio a causa della necessità di aggiornare il sistema informatico che stavo adoperando; tale passaggio infatti mi ha messo nella condizione di non poter più utilizzare il patch scritto su Max/MSP distribuito dall'editore della compositrice, dato che molti degli oggetti che in esso erano contenuti, non erano più compatibili col processore presente sulla nuova macchina. Si è venuta così creando la necessità di realizzare un nuovo ambiente esecutivo che fosse compatibile con il nuovo sistema informatico.

## RACCOLTA DELLE INFORMAZIONI

In calce alla partitura è presente una descrizione tecnica, la quale fornisce però solo le tipologie dei mezzi per la trasformazione del suono, senza descrivere in alcun modo i settaggi e l'uso di questi all'interno della composizione. Ciò oltre a non guidare il regista del suono nella realizzazione informatica non lo aiuta neanche per ciò che riguarda l'esecuzione, perché non viene esaustivamente istruito circa le funzioni musicali dei trattamenti elettronici.

Inoltre la descrizione tecnica si limita a dare delle vaghe idee circa gli effetti utilizzati nel brano, indicando delle tipologie di macchine, e relegando quindi la scelta musicale non ad una scelta dettata dalla ricerca di un determinato suono, ma piuttosto come frutto delle possibilità offerte dai mezzi a disposizione. Vengono così alla mente alcune partiture della

prima era della musica elettronica dove invece dei dati tecnici precisi veniva riportato il nome della macchina utilizzata, evidenziando una certa mancanza di previsione nei confronti dello sviluppo tecnologico.

Entrando nel merito della partitura possiamo notare come la scelta di guidare, nell'elettronica, l'esecutore al flauto, venga effettuata segnando dei numeri che accanto riportano una generica descrizione dell'algoritmo che verrà attivato, tali numeri saranno da guida per lo stesso esecutore sul palco che gestirà la successione degli eventi del Live Electronics, mediante cambi di scena attivati da un pedale.

E' quindi possibile fare riferimento anche a queste, seppur minime, descrizioni nella fase di raccolta dei dati necessari alla ricostruzione (fig. 1).

Una volta messi insieme tutti i dati ricavabili dalla partitura, si rende necessaria una fase di studio approfondito dell'ambiente esecutivo presente sul sito dell'editore, che consiste in un lavoro di raccolta di dati e strategie, che sono presumibilmente quelle maggiormente certificabili, dato che il suddetto ambiente viene definito dalla compositrice stessa, garanzia per un'esecuzione ottimale.

Per affrontare tale fase è necessario però l'impiego di macchine che siano in grado di leggere correttamente il patch fornito dall'editore, in modo che dalle simulazioni sia possibile controllare che la realizzazione tecnica corrisponda al risultato sonoro, anche perché, per garantire di aver compreso correttamente il risultato perseguito, è necessario un continuo confronto uditivo con le registrazioni presenti in circolazione.

Analizzando il patch alcune incongruenze nascono subito, la più evidente è la diffusione del suono, in quanto nella descrizione tecnica sulla partitura non è mai specificata una gestione dello spazio sonoro in modo dettagliato, non solo, ma dalla realizzazione elettronica è possibile notare che il brano può essere eseguito in stereofonia o ottofonia ( si veda la fig. 5 che rappresenta uno scenario tipico di diffusione ad otto canali). Come è facile intuire la diversa scelta del sistema di riproduzione determina una notevole differenza interpretativa dal punto di vista della regia del suono, che rende poco definibile la direzione musicale in tal senso; indubbiamente una possibilità di scelta così ampia permette una esportabilità del brano molto vasta.

Un'altra incongruenza nasce nella gestione del riverbero che è segnato nelle note tecniche come dipendente in maniera inversamente proporzionale alla dinamica, ovvero, deve essere messa in relazione la lunghezza del riverbero con la dinamica captata dal microfono del flauto, facendo in modo che il tempo di riverberazione sia lungo qualora il flautista suoni piano e viceversa per il forte. Tutto questo processo non viene implementato nella versione resa pubblica dall'editore, ponendo quindi la domanda se l'omissione di tale processo sia da attribuirsi: ad una scelta estetica modificatasi nel tempo, ad una necessità determinata dalla mancanza di mezzi adeguati oppure ad una mancanza dovuta a disattenzione da parte del realizzatore della parte informatica.

Si comprende da quanto appena descritto che oltre all'impegno richiesto per la migrazione del patch, è necessario anche fare un'attenta valutazione estetica dell'elettronica impiegata nella composizione, decidendo la strada da percorrere nei confronti delle incongruenze sopra descritte.

Nella versione da me riproposta, come vedremo in seguito, c'è stata una precisa volontà di integrare anche quelle peculiarità che sono state omesse nella realizzazione informatica ufficiale.

## REALIZZAZIONE INFORMATICA

Il primo tipo di lavoro su cui concentrare l'attenzione passa dalla formulazione di un sistema logico in cui determinare il percorso del segnale, spiegato in fig. 2.

Si noti come la mia scelta proponga una certa anomalia rispetto alle indicazioni che il titolo del brano riporta, dove infatti si pone l'accento sul fatto che si tratti di una composizione per strumento ed elettronica, nello schema riportato in fig. 2 si pone invece su uno stesso livello di ordine logico il flauto e i suoni registrati mentre il trattamento in tempo reale viene posto come successivo a queste due fonti sonore. Questo, non per divergenza dall'idea della compositrice, ma piuttosto per sottolineare la forte relazione che esiste fra i files registrati e l'esecuzione dal vivo. Questi ultimi sono infatti realizzati con lo stesso materiale della partitura mettendo in stretta relazione le due fonti sonore. Tale scelta è dipesa anche dalla ricerca di una maggiore aderenza possibile con le idee compositive che le note all'opera permettono di estrapolare, l'appendice tecnica sottolinea appunto l'importanza di processare il flauto e suoni registrati, prima del trattamento vero e proprio, con il medesimo riverbero in modo da renderne il suono meno aspro e più vicino ad una percezione naturale rispetto a quella determinata da una microfonia ravvicinata.

Una volta applicato quel riverbero, che potremmo chiamare "correttivo", il segnale segue il percorso del trattamento vero e proprio, che come è possibile vedere si divide in quattro rami principali:

1. suono non trasformato agli altoparlanti 1 e 2
2. suono armonizzato agli altoparlanti 3 e 4
3. suono riverberato ad un algoritmo di spazializzazione sugli altoparlanti 5, 6, 7 e 8
4. suono con riverbero infinito ad un algoritmo di spazializzazione sugli altoparlanti 5, 6, 7 e 8.

La suddivisione appena effettuata non ha valore di priorità ma serve solo a dividere i rami del percorso del segnale.

La descrizione fatta, come si nota, non rende esplicita la presenza di un algoritmo di envelope following che invece è presente sullo schema di fig. 2 e che si inserisce all'interno dell'algoritmo di riverberazione, tale mancanza dipende solo ed esclusivamente dal fatto che non si tratta di un processore di modifica del segnale bensì un algoritmo di controllo.

E' però tale elemento ad essere motivo di forte interesse in questo lavoro di rielaborazione dell'ambiente esecutivo, dato che, come accennato in precedenza, è richiesto nella nota tecnica in appendice alla partitura la presenza del envelope follower, oggetto che non è stato integrato nella versione ufficiale del patch. La sua funzione è quella di mettere in relazione l'andamento dinamico dell'esecuzione flautistica con la durata temporale del riverbero, e quindi, astruendo in termini musicali, potremmo dire che tramite tale algoritmo c'è un cambio di spazio sonoro assai frequente che è dipendente dal gesto musicale del flautista.

Per poter rendere funzionale l'algoritmo di analisi della dinamica ai nostri scopi, è stata necessaria l'integrazione, all'interno dell'ambiente esecutivo, dell'external **gigaverb~**, che ha come caratteristica di non generare clicks durante il cambio dinamico della lunghezza del riverbero; la parte di analisi invece è stata realizzata ricavando i dati catturati dall'oggetto **peakamp~** a cui è stato mandato in ingresso una copia dell'ingresso microfonico, dati che poi sono stati opportunamente scalati e interpolati temporalmente al fine di rendere la ricezione della dinamica in ingresso meno soggetta a sbalzi istantanei dovuti a transitori d'attacco. La variazione della lunghezza dei tempi di riverberazione, è stata definita fra un minimo di un secondo ed un massimo di sei secondi, in modo da cercare di rendere evidente il tipo di processo senza che questo divenisse didascalico, inoltre è da sottolineare che il range dei tempi è stato mediato rispetto alle indicazioni a commento delle scene nel patch ufficiale, che come vedremo in seguito fungono da controllo per tutta l'elettronica.

Facendo un passo indietro e tornando al punto 1. della descrizione del flusso del segnale, si può notare come ancora una volta si sottolinei la relazione fra suoni registrati ed esecuzione dal vivo, questi escono entrambi dagli stessi diffusori che fungendo da rinforzo per il flauto creano un vertice di diffusione sonora compatto fra files e suono acustico, realizzando così quell'ottima ambiguità che rende a volte l'effetto così integrato da portare quasi alla sensazione di essere davanti ad un'esecuzione che supera i limiti fisici dello strumento. L'effetto si enfatizza non soltanto per la similitudine del materiale sonoro, ma anche grazie allo studio dello spettro del suono del flauto che è simulato in studio integrando in maniera molto fine anche quei suoni la cui natura è stata modificata dal trattamento in preproduzione.

Agli altoparlanti 3 e 4, come descritto al punto 2., viene inviato il suono processato da un armonizzatore che è uno degli strumenti più controversi in questo brano. Non ne sappiamo i parametri, o per meglio dire la compositrice non li descrive nello spartito e neanche sul sito internet dell'editore, l'unica informazione manifesta dice che questo processore è composto di due linee di armonizzazione. Occorre quindi affidarsi allo studio del patch ufficiale per cercare di ricavare i parametri di questo processore.

Durante l'operazione di recupero dei dati relativi all'armonizzatore è stato analizzato anche il processore utilizzato, che è un'operazione di primaria importanza per capire la tipologia di suono che deve essere generato. L'algoritmo in esame è stato riprodotto per mezzo dell'oggetto **harmv2~** che è contenuto nella libreria *Jimmies* sviluppata dall'IRCAM, ed è un armonizzatore basato sulla tecnologia delle linee di ritardo tempo variabili. La caratteristica di questa tipologia di traspositori, da un punto di vista sonoro, è quella di produrre una modulazione ciclica nel suono, ed in questo caso quello che potrebbe essere considerato un difetto viene utilizzato come una potenzialità del processore; questa caratteristica, tramite i dati immessi, viene qui espansa creando l'effetto di una moltitudine di glissandi ravvicinati che si generano tramite la ripetizione della lettura della memoria per mezzo di una certa quantità di feedback.

Risulta evidente che l'idea di considerare l'armonizzatore in veste tradizionale decade, e decade anche la possibilità di affidarsi ad un'altra tecnologia per realizzare l'effetto appena descritto, ciò che conviene invece sviluppare a questo punto, è l'integrazione di tale tecnologia corredandola dei valori presenti nel patch ufficiale, dai quali si può anche ricavare che la trasposizione effettuata consiste in una trasposizione di un semitono superiore in una voce dell'armonizzatore ed un semitono inferiore nell'altra.

Passando al punto 4., dato che il 3. è già stato evidenziato in precedenza, si nota la presenza di un algoritmo che deve generare, quello che è stato chiamato riverbero infinito, nome che contiene già al suo interno le caratteristiche di tale trattamento.

Il riverbero infinito svolge funzione di congelamento del suono, perché è organizzato in modo da captare un frammento sonoro di breve durata tenendolo poi in attività per tutta la durata della riverberazione, che è caratterizzata da tempi che hanno non meno di sessanta secondi di durata. In questo modo il suono processato con il riverbero infinito, quando attivato, crea la sensazione di una fascia sonora che accompagna l'esecuzione e che si trova su un piano spaziale differente.

La differenza spaziale appena sottolineata, introduce una tematica che ancora non è stata affrontata nel dettaglio, e che consiste nell'analisi del sistema di diffusione del suono. E' stato evidenziato dal diagramma del percorso del segnale la presenza di otto diffusori differenti e come questi siano stati coinvolti nel processo di separazione dei segnali, non è stata invece analizzata la precisa funzione e la disposizione in sala.

In effetti, ancora una volta non viene fatta descrizione circa tipologia e la logica di diffusione del suono, anche perché, come abbiamo visto, il brano è stato originariamente concepito per un sistema stereofonico; allo stesso tempo, facendo riferimento al patch ufficiale è possibile notare la presenza di un menù che permette di scegliere se utilizzare il sistema stereofonico o ottofonico. Di conseguenza, qual ora si decida, come in questo caso, di procedere ad una diffusione con otto altoparlanti, si rende necessario uno studio, e quindi un'interpretazione, della logica usata nel patch per tale scopo.

Si distinguono i vari indirizzamenti del segnale in maniera molto chiara mandando in esecuzione il patch ufficiale, comprendendo chiaramente la destinazione dei diffusori. Con la stessa logica e disposizione è organizzata la diffusione sonora nello schema da me ricostruito per la trascrizione dell'ambiente esecutivo. Si rende però necessario un lavoro più approfondito per quanto riguarda gli altoparlanti 5, 6, 7, e 8, dove è possibile notare come questi ricevano il segnale da un blocco di controllo che è stato chiamato **spatter**; questo svolge la funzione di indirizzamento del suono elaborato dai riverberi. Il problema consiste nello stabilire la modalità di tale indirizzamento: ancora una volta è necessario fare riferimento al patch ufficiale, cercando di studiare il comportamento di tale algoritmo, dal quale si nota come il percorso del suono non si svolga mai secondo una direttività precisa e distinguibile, ma invece alterni il suo baricentro spostandosi lievemente fra una zona e l'altra del quadrilatero sonoro che si viene a formare, senza mai addensarsi su una singola fonte sonora. Potremmo semplificare dicendo che il suono passa continuamente il centro dell'energia verso una direzione differente senza però mai escludere del tutto il contributo energetico del resto del sistema.

L'algoritmo che ha reso possibile la realizzazione di un tale movimento spaziale del suono, è stato costruito partendo dall'idea di posizionare i quattro altoparlanti posteriori come vertici di un ideale piano cartesiano, che ha nel centro dello spazio delimitato da tali diffusori, il punto d'incontro tra gli assi. La mobilità del suono è stata realizzata andando a controllare il movimento delle ascisse e delle ordinate tramite degli oscillatori in bassa frequenza che hanno lo stesso controllo per la frequenza, ma fase con differenza di 90°, un simile sfasamento permette di creare un andamento circolare del suono. E' stato poi aggiunto un altro oscillatore per modulare ciclicamente la fase dell'oscillatore incaricato di controllare le ordinate, che per mezzo di tale modulazione rende mobile e imprevedibile lo spostamento del suono sul piano, andando a ricreare il movimento dello spazio sonoro analizzato sul patch ufficiale.

La fase di esecuzione di tutta l'elettronica, come accennato in precedenza, avviene tramite l'attivazione di una playlist, nella quale sono salvati tutti i dati di controllo degli algoritmi che abbiamo analizzato fino ad ora. Questa lista di eventi viene mandata in esecuzione tramite un pedale di controllo che il flautista pilota durante la performance (questa volta l'indicazione è presente nelle note tecniche), e che permette di legare molto bene il gesto strumentale con quello generato dalla elaborazione elettronica.

Ultima nota riguardante questa rielaborazione dell'ambiente esecutivo, riguarda la strategia perseguita per rendere il più possibile indipendente l'esecutore al flauto dalla necessità di un tecnico specializzato. Tale strategia è basata sull'utilizzo di un pedale di controllo del volume generale, che mette il flautista nelle condizioni di abbassare fino a 6dB la dinamica nel caso di problemi derivanti da rientri acustici del sistema, e qual ora ce ne fosse necessità di rialzare il volume abbassato.

## CONCLUSIONI

Il lavoro d'indagine circa le problematiche legate alla rieseguitività dell'opera che fa uso di mezzi elettronici, che è stato messo in luce dall'analisi e dal lavoro di elaborazione nel brano *NoaNoa*, porta a soffermarsi ancora una volta su quelle tematiche che già hanno dato alla luce autorevoli indagini da parte di molti compositori e studiosi. Questi si sono dovuti confrontare, come in questo caso, con la difficoltà di trovare un mezzo capace di trasferire in un dominio, non soggetto a così facile obsolescenza, l'informazione riguardante la parte dell'elettronica dovendosi inoltre sempre relazionare con un codice, quello della notazione, che si è consolidato e ottimizzato attraverso secoli di storia. Il peso e la portata di tale lavoro è quindi sempre da affrontare su due fronti per il compositore, ovvero la ricerca di un linguaggio esaustivo e comune come quello della notazione tradizionale, come anche la ricerca di strategie di astrazione di concetti e gesti musicali che sono spesso profondamente legati alla tecnologia di cui fanno utilizzo.

Nel caso specifico, prendendo come riferimento le strategie riportate da Nono nella realizzazione delle partiture elettroniche, il percorso da fare non sembra così arduo, data la

tipologia degli algoritmi utilizzati, ed è possibile schematizzare un esempio di partitura elettronica che serva principalmente come guida nella ricostruzione e nello studio del brano. Il breve esempio a cui faccio riferimento (fig. 3) è corrispondente all'estratto della partitura di fig. 4, e ha la funzione di rendere più chiara la successione degli eventi nel dialogo fra parte acustica e parte elettronica.

È stato scelto perché rappresenta un momento di alternanza rapida di eventi e perché, analizzando attentamente lo spartito, è possibile notare come al punto 12 siano presenti due informazioni separate da una virgola, le quali stanno ad indicare con molta probabilità, la simultaneità dei processi descritti, e che corrispondono al lancio di un suono e all'attivazione del riverbero.

Questa ambiguità generata dalla mancanza di indicazioni per la corretta comprensione dello spartito, sottolinea ancora una volta l'urgenza di adottare, da parte dei compositori, un sistema per permettere la sopravvivenza delle composizioni che contengono elaborazioni elettroniche.

## BIBLIOGRAFIA

K. Saariaho, *"NoaNoa"* for flute and Electronics, Chester Music, Londra 2007.

L. Nono, *"Pierre. Dell'azzurro silenzio, inquietum"* flauto contrabbasso in Sol, per clarinetto contrabbasso in Sib e live electronics, Ricordi 1985.

L. Nono, *"Post-Prae-Ludium per Donau"* per tuba in Fa e live electronics, Ricordi 1987.

M. Stroppa, *"Un'Orchestra Sintetica: Appunti di una notazione Personale"*, in *Le Timbre: métaphore pour la composition*, J. B. Barrière ed., Parigi 1988.

K. Stockhausen, *"Electroacoustic Performance Practice"*, in *Perspectives of New Music*, Vol. 34, No 1, pp 74-105, 1996.

N. Bernardini, A. Vidolin, *"Sustainable Live Electro-Acoustic Music"*, in *Proceedings XV CIM*, Salerno 2005.

C. Roads, *"The computer music tutorial"*, The MIT Press, Massachusetts Institute of Technology 1996.

<http://www.chesternovello.com>

<http://www.akustische-kunst.org/maxmsp/>

<http://mustica.ircam.fr/mustica/index.php>

Fig. 1

for Camilla Hoitenga  
**NoaNoa**

Flute

Intenso, poco agitato  
tempo I (♩ = c.72)

tempo II (♩ = c.54)  
calmo  
molto vibr.

molto vibr. S.V.

con vibr. S.V. → molto vibr. → agitato

overblow into harmonics

1 infinite rev.

2 inf. rev.

3

9 tempo I dolce

tempo II calmo (♩ = c.54)

mp

mp

mf

mf

mp

poco sfz

mf

4 inf. rev.

5 stop rev.

Fig. 2

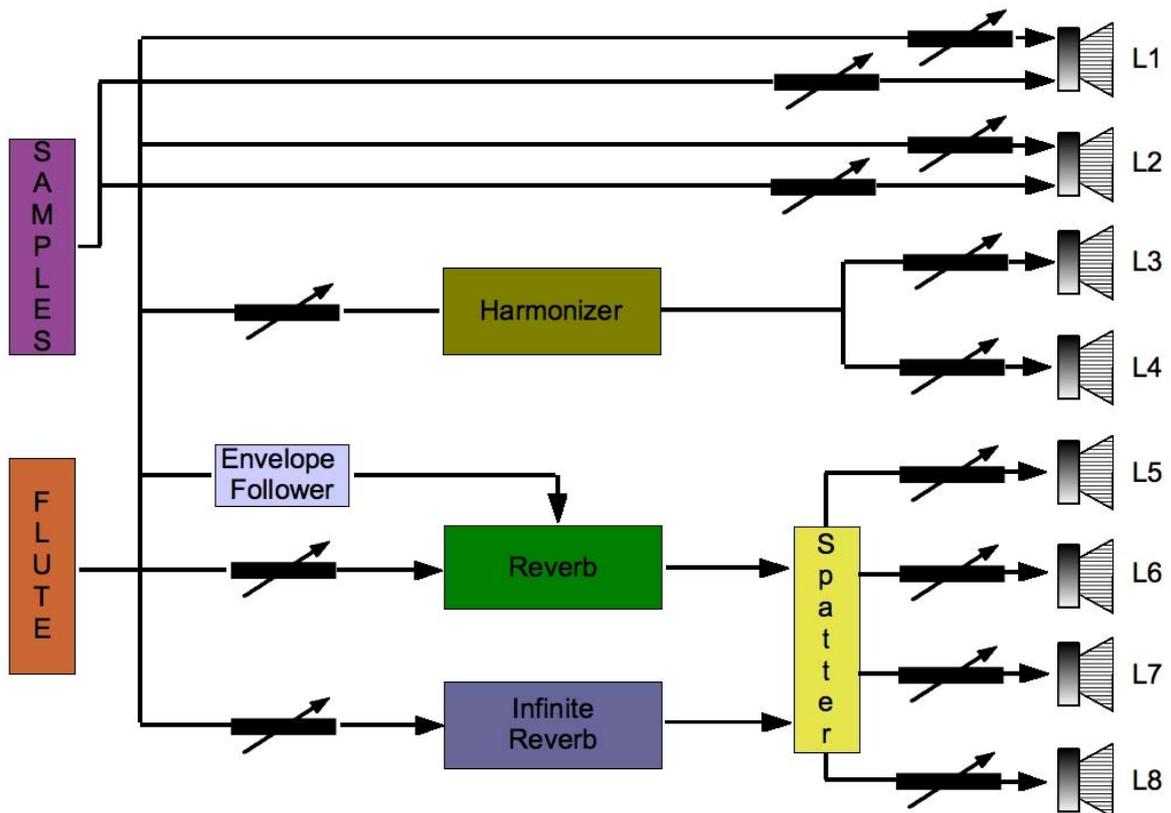
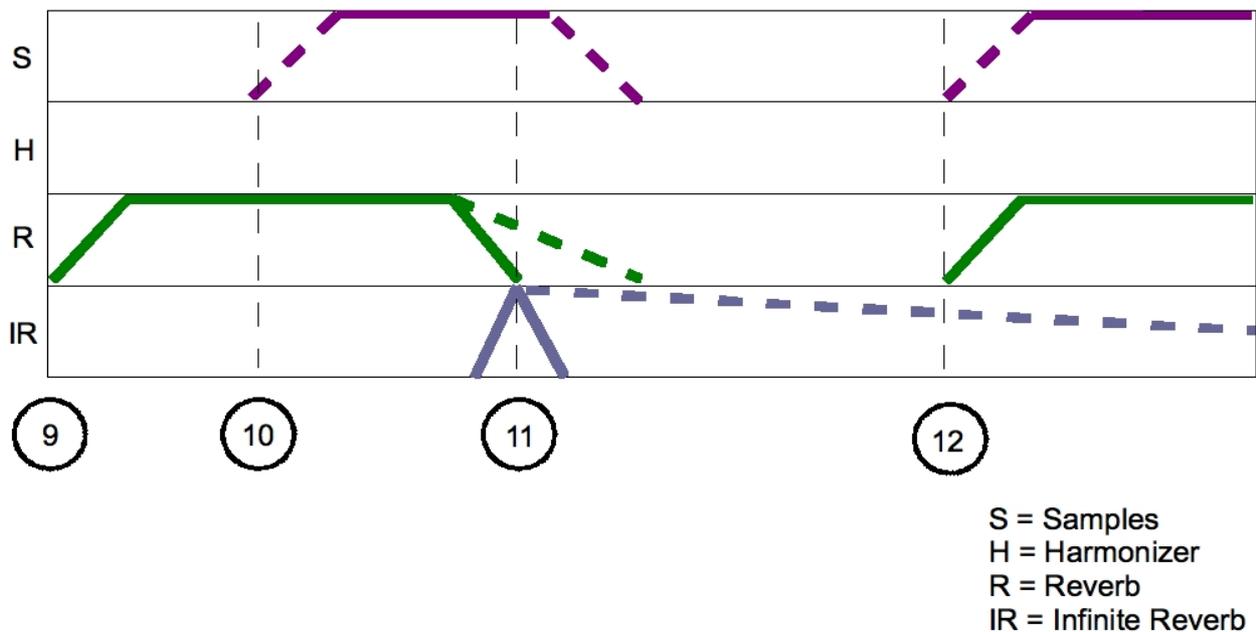


Fig. 3

The musical score shows a Flute (Fl) part and a Voice (V) part. The Flute part starts at measure 35 with a 'misterioso rubato' tempo. It features various dynamics: *p* (piano), *mf* (mezzo-forte), *p* (piano), and *mp* (mezzo-piano). Performance instructions include 'S.V.' (Sostenuto/Vibrato), 'gliss' (glissando), and 'calmato' (calmato). The Voice part has lyrics 'sen-tait la rose' and is marked with *p* and *mp*. Electronic processing annotations are numbered 9 through 12: 9 rev., 10 noise, 11 inf. rev., and 12 sampled flute, rev. The score also includes tempo changes: 'tempo II molto vibr.' and 'tempo I espr.'.

Fig.4

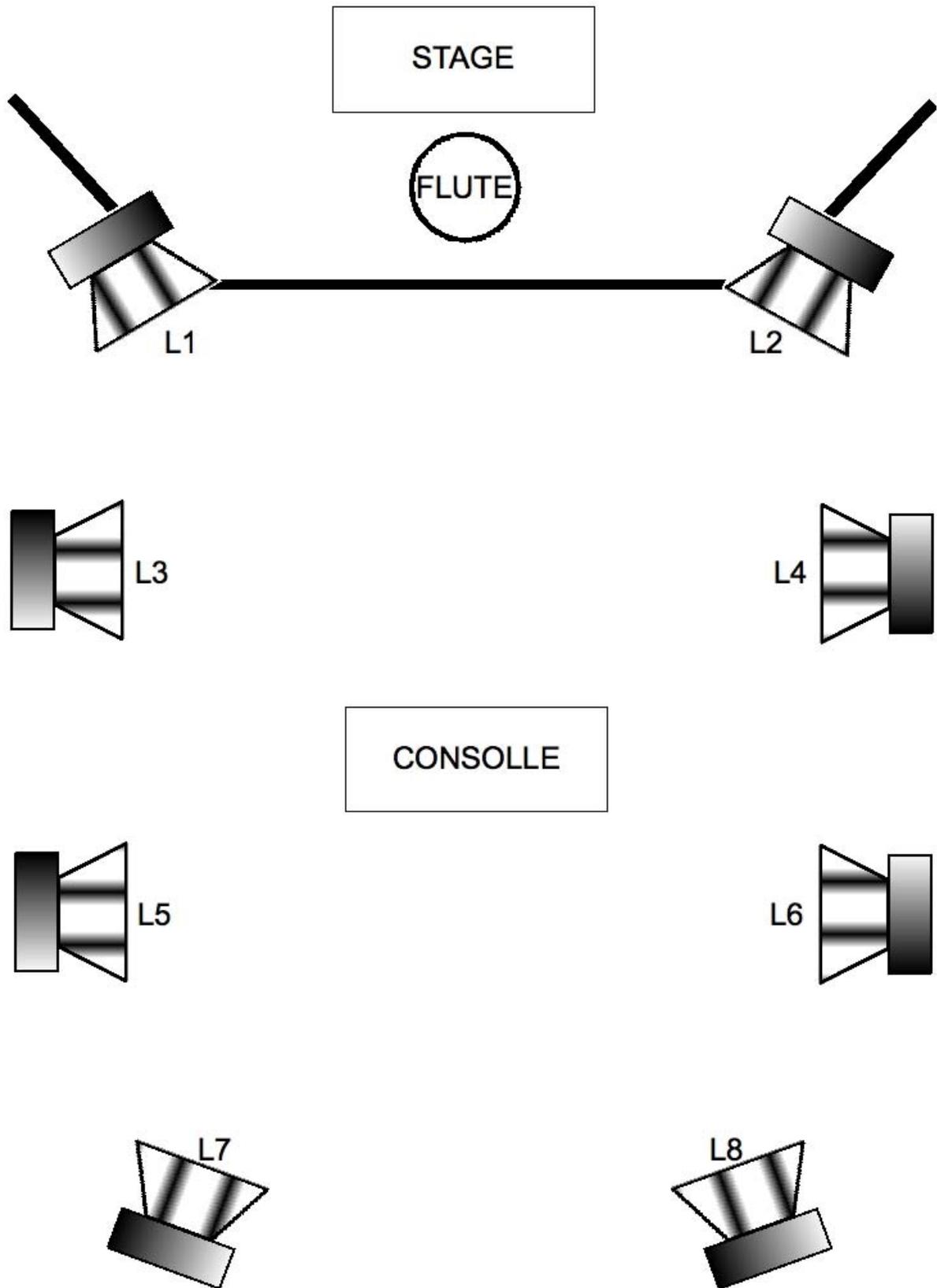
Esempio di possibile notazione della partitura elettronica del brano *Noa noa* di Kaija Saariaho



[Nella notazione della parte elettronica si noti la differenziazione nel tratto degli involucri, che sono stati distinti in: involucri propri dell'algoritmo, si pensi al decadimento del riverbero e involucri dinamici dei suoni registrati, che sono sottolineati per mezzo del tratto discontinuo; involucri determinati dall'elaborazione elettronica, che sono invece quelli creati all'interno dell'ambiente esecutivo e che sono segnati con il tratto continuo.]

Si noti inoltre la mancanza di indicazioni relative alla dinamica, dipendente da una scelta precisa, in quanto il livello dei trattamenti deve essere tarato in prova e deve avere una dinamica equivalente a quella dello strumentista, le variazioni dinamiche musicali saranno poi dettate dalla stimolazione degli algoritmi da parte dell'esecutore al flauto.]

Fig.5



# Partitura Elettronica da battuta 22 a 41

22 misterioso, poco rubato S.V. molto vibr. tr

Fl  
V

l'ar - bre sen - tait la rose la rose

⑥ changed rev. time

S  
H  
R  
IR

29

Fl

mf f mf p mp

tempo II tempo I energico tempo I

⑦ stop rev.

S  
H  
R  
IR

35 misterioso rubato S.V. gliss

Fl  
V

sen - tait la rose

⑨ rev. ⑩ noise ⑪ inf. rev. ⑫ sampled flute, rev.

S  
H  
R  
IR

*poco a poco più energico*

*tempo I intenso*

*mp*

très o-do-rant sen-tait rose rose sen-tait rose

s t r t s t s t s t t t t t

*f poss.*

*poco rit. ....*

*calma*

*tempo I espr.*

*p* *mp* *mf* *p*

8 sampled flute

125C

**B** *tempo II*

*S.V.*

*gliss* *overblow into harmonics* *mp* *(mp)* *mf* *p*

*dolce tr*

*molto vibr.* *S.V.* *gliss*

13 inf. rev.



# Il ruolo del movimento corporeo nell'esecuzione musicale in presenza di feedback uditivo ritardato

Marco Matteazzi

CSC - DEI

Università degli Studi di Padova

## SOMMARIO

Lo studio si propone di investigare il ruolo del movimento corporeo nell'esecuzione musicale in presenza di feedback uditivo ritardato. A sedici musicisti (nove pianisti e sette organisti) è stato richiesto di eseguire uno stesso brano al pianoforte elettronico in tre differenti situazioni di feedback uditivo: inizialmente in condizione di feedback normale, quindi in assenza di feedback, infine con un feedback ritardato di 300 ms. Sul busto e sulla testa dei musicisti sono stati posizionati dei sensori tramite i quali sono stati analizzati intensità e periodicità dei movimenti durante le esecuzioni, in relazione alla scansione in battute della partitura. All'analisi del movimento è stata associata un'analisi della qualità dell'esecuzione musicale: per ogni battuta sono stati calcolati il tempo in bpm e l'intensità media (utilizzando il parametro midi *key velocity*). I risultati preliminari, riferiti alle esecuzioni dei primi dieci soggetti (nove pianisti e un organista), sembrano confermare l'ipotesi che il movimento corporeo abbia un ruolo importante nell'esecuzione musicale in presenza di feedback uditivo ritardato, e che determinate strategie psico-motorie possano alleviare gli effetti negativi dello stesso sulla qualità dell'esecuzione.

## Parole Chiave

Delayed Auditory Feedback; Embodied Music Cognition; Church Organ.

## INTRODUZIONE

Lo studio si inquadra nell'ambito teorico dell'*embodied music cognition* (Leman, 2008), un approccio ecologico alla ricerca sulla musica che considera il corpo un mediatore tra la musica come esperienza mentale e la sua trasmissione a livello di energia fisica, e viceversa. In questa ottica il corpo ed il suo movimento giocano un ruolo fondamentale sia nella produzione sia nella fruizione di opere musicali, ruolo che non può essere trascurato dalla ricerca scientifica.

Nell'ambito dell'*embodied music cognition*, un concetto centrale è quello di accoppiamento di azione e percezione (Gibson, 1979): questo concetto è infatti indispensabile per la comprensione delle relazioni tra un soggetto e l'ambiente circostante (approccio ecologico). Nello specifico, la relazione tra il corpo, la mente e l'ambiente fisico circostante può essere descritta dal ciclo azione-reazione. Si tratta di un modello dinamico secondo il quale l'output di ogni azione eseguita, percepito dal soggetto (*feedback*), fornisce informazioni in base alle quali il soggetto può reagire per modificare opportunamente il suo comportamento. Si crea così un ciclo grazie al quale l'esecutore sperimenta una forma di controllo sull'esecuzione, avendo al tempo stesso l'impressione che l'ambiente in cui si trova ad operare abbia un alto grado di presenza e di realtà. Grazie all'accoppiamento di azione e percezione l'esecutore arriva a sentirsi immerso nell'ambiente al punto di poterlo modificare direttamente con i propri movimenti sullo strumento. Studi recenti sull'attività cerebrale

sembrano confermare come l'esecuzione musicale sia strettamente collegata all'accoppiamento di azione e percezione ed al coinvolgimento dell'esecutore rispetto all'ambiente fisico in cui la musica risuona.

In presenza di feedback uditivo ritardato (*delayed auditory feedback, DAF*) viene a stabilirsi un'asincronia tra azione e percezione per cui l'accoppiamento tra i due momenti non è più verificato. Il ciclo di azione-reazione viene spezzato, in quanto il feedback udito in un dato momento non è collegato all'azione appena compiuta ma ad un'azione precedente. L'effetto di tale feedback uditivo ritardato sulla produzione musicale e non è stato studiato approfonditamente. È stato dimostrato (Finney, 1997, 1999; MacKay, 1987) come il DAF interferisca negativamente sulla produzione di sequenze di parlato provocando interruzioni all'interno delle stesse. Numerosi studi documentano altri effetti negativi del DAF sulla realizzazione di obiettivi di diversa natura, tra cui l'esecuzione musicale e il muoversi ritmicamente. Alcuni ricercatori (MacKay, 1987; Smith, 1962) hanno sostenuto che lo sforzo di ignorare il DAF e un'apposita pratica possono ridurre l'effetto negativo che il DAF ha sull'esecuzione.

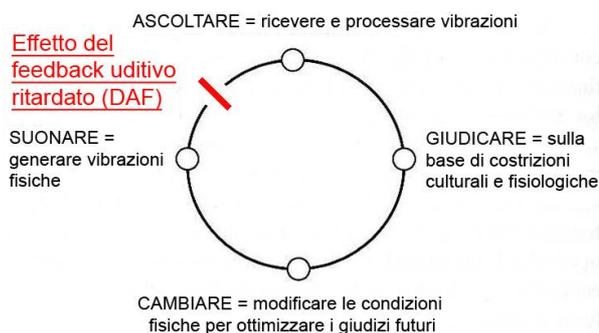


Figura 1. Il ciclo azione-reazione viene spezzato dall'effetto del DAF.

Lo sviluppo di nuove tecnologie multimediali per l'esecuzione musicale ha aperto nuovi scenari in cui una latenza non eliminabile può essere insita nelle tecnologie stesse. Si pensi ad esempio a strumenti e interfacce elettronici, in cui l'input (gesti del musicista) non è direttamente collegato all'output (generazione del suono), per cui possono intervenire fastidiosi ritardi temporali; o alla possibilità per più musicisti di suonare dal vivo tramite Internet, in analogia con quanto già accade per i videogiochi. Una sfida aperta riguarda quindi lo studio di tecniche, metodologie e tecnologie per attenuare l'effetto negativo del DAF sull'esecuzione musicale.

Lo studio si propone di indagare l'effetto del DAF sull'esecuzione musicale di pianisti e organisti. Questi ultimi, in particolare, sono per formazione abituati a suonare in presenza di DAF, in quanto negli organi da chiesa il suono viene prodotto con un ritardo sensibile rispetto alla pressione del tasto. L'idea è quindi che dallo studio delle esecuzioni di organisti in condizioni di DAF si possano trarre indicazioni sulle strategie opportune per mantenere alto il livello qualitativo dell'esecuzione anche in questo tipo di circostanze.

In particolare, si vuole investigare la relazione tra il DAF e il movimento corporeo durante l'esecuzione, e stabilire se l'ausilio di particolari strategie motorie può ridurre gli effetti negativi del DAF. Un'appendice della ricerca investigherà inoltre quale sia effettivamente l'ordine del ritardo del feedback negli organi da chiesa, considerando il caso particolare dell'organo pneumatico della chiesa di Sant'Anna a Gand, in Belgio.

## METODOLOGIA SPERIMENTALE

A due gruppi di soggetti, il primo costituito da pianisti ed il secondo da organisti, è stato richiesto di suonare le prime sedici battute della "Sarabande" di G. F. Handel, una prima volta con un feedback uditivo normale, una seconda senza feedback, infine con un feedback ritardato di 300 ms. In tutte e tre le condizioni prima dell'esecuzione è stato dato ai musicisti come riferimento un battito di metronomo di 56 bpm; non è stato tuttavia richiesto loro di eseguire il brano strettamente a tempo, quanto di cercare di conferirgli l'espressività che ritenevano appropriata, sia in termini di velocità dell'esecuzione (rallentamenti, accelerazioni), sia in termini di dinamiche (crescendo, diminuendo).

Il brano e il ritardo del feedback sono stati scelti sulla base di uno studio pilota nel quale un pianista ha eseguito tre brani dalle caratteristiche differenti a diversi tempi di ritardo del feedback. I brani utilizzati sono stati la "Sarabande", "33 Veränderungen" di A. Diabelli, e "Für Elise" di L. W. Beethoven, mentre i tempi di ritardo sono utilizzati, in ms, sono stati 90, 140, 190, 240, 290, 340, 490, 890. Un'analisi statistica dei risultati dello studio pilota ha consentito di valutare in 300 ms un tempo di ritardo sufficiente per avere un effetto significativo sulla qualità dell'esecuzione musicale, e al tempo stesso non così alto da poter essere considerato estremo. Per quanto riguarda il brano, è stato scelto il pezzo più facile e lento, alla portata di qualunque pianista di discreta esperienza, e relativamente semplice da analizzare. Lo studio pilota ha inoltre evidenziato come le differenze tra l'esecuzione in condizioni normali e quella in condizioni di feedback ritardato siano più rilevanti a bpm bassi, fatto che ci ha ulteriormente indirizzato verso la scelta della "Sarabande".

Per quanto riguarda la strumentazione, è stato utilizzato un pianoforte elettronico (Yamaha P-60 Electronic Piano) collegato a un computer tramite un'interfaccia MIDI (M-Audio MIDI USB Interface), e tre sensori di movimento (MT Xsens) in grado di calcolare in tempo reale gli angoli pratici (imbardata, rollio e beccheggio) e l'accelerazione lineare tridimensionale.

Se per realizzare la condizione di feedback normale si sono semplicemente utilizzati gli altoparlanti del pianoforte, e per la condizione di assenza di feedback è stato sufficiente abbassare il volume a zero, per la condizione di feedback ritardato è stato invece necessario scrivere un'applicazione per Pure Data che sintetizzasse le note MIDI in ingresso dopo un intervallo temporale impostabile a piacere. Si è quindi misurato il ritardo complessivo intrinseco del sistema costituito da pianoforte, interfaccia midi e scheda sonora del computer impostando il ritardo a 0 ms; al valore così misurato si è aggiunto un ulteriore ritardo calcolato in modo da ottenere le condizioni desiderate. Le note sintetizzate dalla scheda sonora sono quindi state fatte suonare da due altoparlanti posizionati sopra a quelli del pianoforte, questi ultimi tenuti a zero, e il volume è stato impostato in modo da essere uguale a quello utilizzato nella condizione normale.

Per ogni esecuzione sono stati registrati i dati provenienti dal pianoforte nel formato MIDI. Utilizzando questi dati e relazionandoli con la partitura è stato possibile calcolare per ogni battuta il tempo in bpm e l'intensità media (grazie al parametro key velocity). Il movimento del corpo dei musicisti è stato registrato tramite due sensori, uno posizionato sulla fronte, l'altro sul busto. Un terzo sensore è stato utilizzato per coordinare le registrazioni del movimenti con i dati MIDI: prima di ogni esecuzione si è premuto con il sensore un tasto del pianoforte, di modo che fosse poi possibile far corrispondere l'istante temporale di pressione registrato dal MIDI con il picco dell'accelerazione verticale registrata dal sensore. Tutti gli esperimenti sono stati ripresi con una videocamera, ed ulteriori indicazioni sono state ricavate dai questionari fatti compilare ai soggetti prima e dopo l'esecuzione.

## ANALISI DEI DATI

Allo stato corrente sono state analizzate le esecuzioni di tutti i nove pianisti e di un organista. Un'analisi qualitativa dei valori

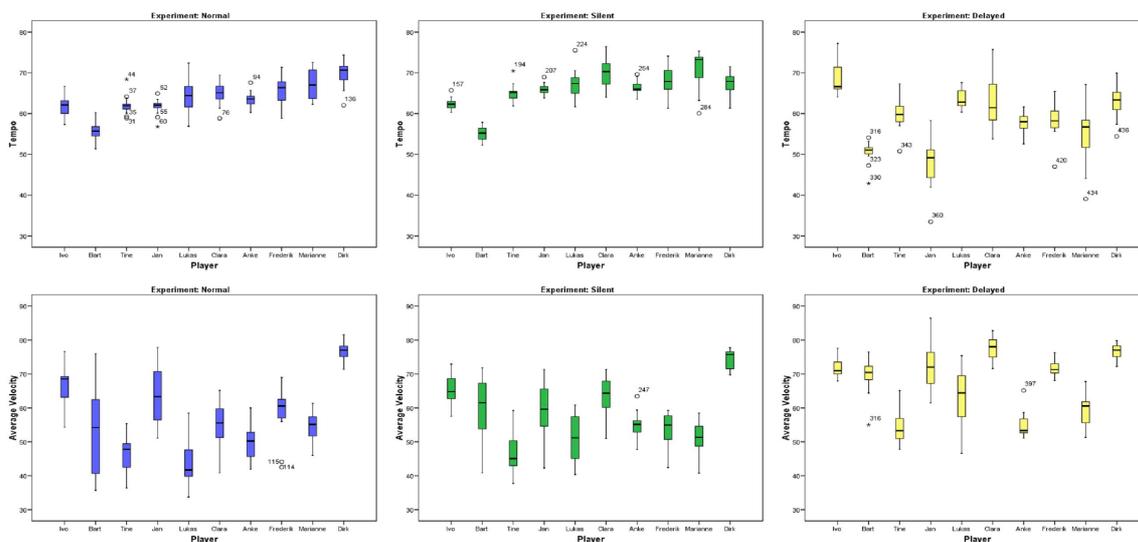


Figura 2. I box plot di tempo (prima riga) e key velocity (seconda riga) per i dieci soggetti. La prima colonna si riferisce all'esecuzione normale, la seconda all'esecuzione in assenza di feedback, la terza all'esecuzione in presenza di feedback ritardato.

per battuta di tempo e key velocity media mostra come i pianisti, pur in grado di mantenere il loro livello esecutivo in assenza di feedback, non riescano in generale a riprodurre l'esecuzione allo stesso modo in presenza di DAF. Le distribuzioni dei valori di tempo e key velocity, rappresentate nei box plot di figura 2, rimangono pressoché invariate nel passaggio dall'esecuzione normale all'esecuzione in assenza di feedback. Al contrario nell'esecuzione in presenza di DAF si può notare una diminuzione nella velocità d'esecuzione (tempo) e un incremento nelle dinamiche (key velocity). Variazioni significative si possono notare anche nella varianza dei valori, la quale aumenta per quanto concerne il tempo e diminuisce leggermente per quanto concerne la key velocity media. Sono evidenti tuttavia delle differenziazioni nel comportamento dei diversi soggetti. In particolare, il decimo soggetto (l'organista), a differenza degli altri, riesce a mantenere pressoché costanti i valori al variare delle condizioni di feedback. Un'altra eccezione è costituita dal primo pianista, il quale in presenza di DAF aumenta la velocità dell'esecuzione. Nell'analisi quantitativa seguente i valori di questi due soggetti non sono stati considerati, e ci si è concentrati sulle differenze tra l'esecuzione in condizioni normali e quella in presenza di DAF, tralasciando il caso di assenza di feedback. Ci riferiremo ai risultati dei singoli soggetti utilizzando i loro nomi di battesimo: Bart, Tine, Jan, Lukas, Clara, Anke, Frederik, Marianne.

Nella tabella di Figura 3 sono riportati i valori medi e la deviazione standard delle distribuzioni di key velocity media e tempo.

	Average Velocity Normal	Average Velocity Delayed	Tempo Normal	Tempo Delayed
Valid	120	120	120	120
Missing	0	0	0	0
Mean	53,149	65,432	63,09607	56,95310
Std. Dev.	9,6534	9,4983	4,321062	7,104677

Figura 3. I valori medi e la deviazione standard di key velocity media e tempo nelle esecuzioni normali e in presenza di DAF.

Dopo aver verificato la normalità delle distribuzioni utilizzando il test di Kolmogorov-Smirnov, si è stabilita la rilevanza delle variazioni dei valori medi utilizzando il T-test appaiato. I risultati del T-test hanno dimostrato come le distribuzioni in condizioni di DAF differiscano significativamente da quelle in condizioni normali (per entrambi i parametri sig. = 0,00).

Tramite un'analisi ANOVA è stata stimata l'influenza dei diversi fattori sui risultati dell'esperimento. Poiché il test di Levene ha evidenziato come le varianze dei dati nelle due condizioni non rispettino le condizioni di omogeneità, il test di Tukey è stato sostituito dal test post hoc di Tamhane. Per quanto riguarda il tempo, l'ANOVA mostra come l'incidenza del fattore Esperimento (condizioni normali o condizioni di DAF) sui risultati sia del 21,6%, contro il 32% di incidenza del fattore Musicista (diversi esecutori), mentre l'Interazione Musicista-Esperimento incide per il 12,3%, con un errore dovuto a fattori sconosciuti stimato nel 34,3%. Il coefficiente di correlazione è di 0,657, per cui il modello considerato è accettabile. Per quanto riguarda la key velocity media, il test mostra un'incidenza del fattore Esperimento del 29,3%, contro il 31,3% di Musicista e 7,8% di Interazione Musicista-Esperimento. In questo caso l'errore dovuto a fattori sconosciuti è stimato nel 31,6%. Il coefficiente di correlazione è 0,68, per cui anche questo modello è accettabile.

In Figura 4 sono rappresentate le medie marginali stimate per il parametro tempo. Dal grafico si può notare come, sebbene per tutti i musicisti considerati la velocità di esecuzione decresca nella condizione di DAF, per alcuni di essi (Clara, Lukas, Tine) questa variazione non è particolarmente accentuata, mentre per altri (Jan, Marianne) il rallentamento è molto significativo, dell'ordine dei 12-14 bpm.

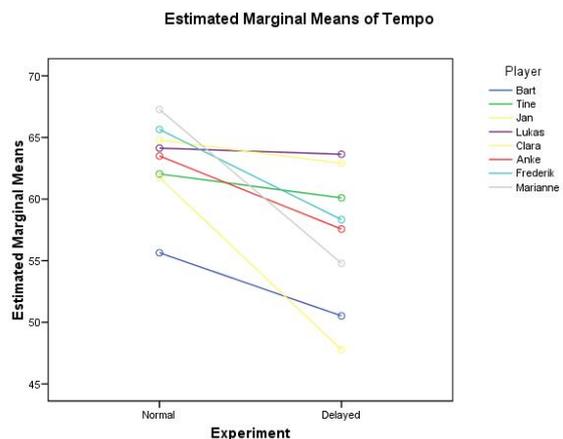


Figura 4. Le medie marginali stimate per il parametro tempo.

Le medie marginali stimate del parametro key velocity media, rappresentate in Figura 5, mostrano invece come sia presente in tutti i soggetti un incremento delle dinamiche nella condizione di DAF: l'aumento è in particolar modo rilevante per quanto riguarda Bart, Lukas e Clara.

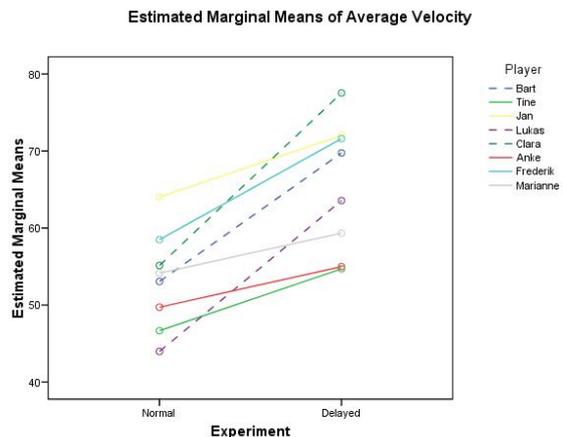


Figura 5. Le medie marginali stimate per il parametro key velocity media.

L'analisi effettuata finora ha dimostrato come nell'esecuzione con DFA siano significative una diminuzione del parametro tempo e un incremento del parametro key velocity media rispetto all'esecuzione in condizioni normali: c'è quindi la possibilità che i due parametri siano tra loro correlati. L'analisi della correlazione tra tempo e key velocity media per entrambi gli esperimenti mostra come questa non sia rilevante, e che quindi la diminuzione del tempo non è proporzionale all'aumento della key velocity media.

Per quanto concerne l'analisi del movimento corporeo, i risultati relativi al sensore posizionato sulla testa e quelli relativi al sensore posizionato sul busto sono risultati essere omogenei, per cui tratteremo qui solamente l'analisi statistica relativa al sensore posizionato sulla testa.

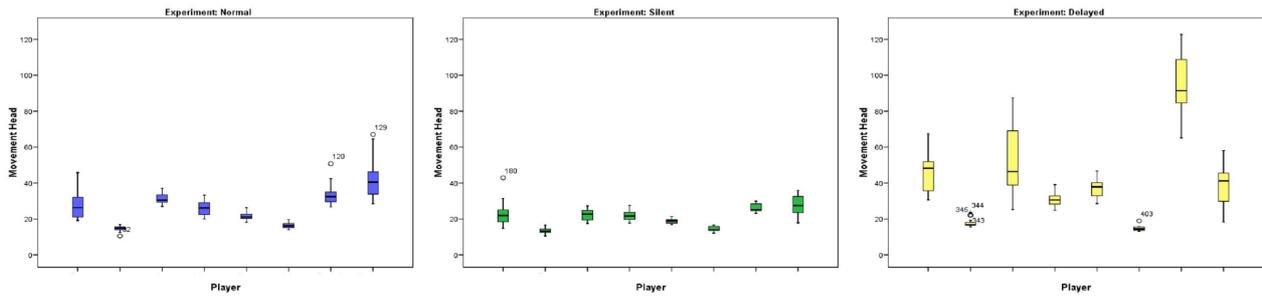


Figura 6. I box plot relativi all'intensità del movimento della testa.

Utilizzando i dati relativi all'accelerazione tridimensionale si è calcolata una misura dell'intensità di movimento per ogni battuta, intesa come media sui campioni della battuta del modulo della derivata dell'accelerazione (*jerk*). I box plot dei valori dell'intensità di movimento (Figura 6) mostrano come non siano presenti grandi differenze tra l'esecuzione normale e quella in assenza di feedback: anche in questo caso l'analisi di quest'ultima non verrà considerata. Sono evidenti invece, per alcuni soggetti, forti differenze tra l'esecuzione normale e quella in presenza di DAF. In particolare per Bart, Jan e Frederik è visibile un forte incremento dell'intensità di movimento. Proprio i valori estremi dell'esecuzione di Frederik in situazione di feedback ritardato causano la non accettabilità della condizione di normalità per la distribuzione dei valori del parametro, necessaria per effettuare l'analisi ANOVA: per questa ragione Frederik sarà escluso dall'analisi. Il test di Kolmogorov-Smirnov mostra che la distribuzione così ottenuta può essere considerata normale. Il test di Levene mostra invece come le varianze non siano omogenee nelle due situazioni: anche in questo caso il test post hoc di Tamhane sostituirà il test di Tukey.

L'analisi ANOVA mostra come il fattore Esperimento abbia incidenza per il 7,9%, il fattore Musicista per il 48,9%, e il fattore Interazione Esperimento-Musicista per il 10,7%, con un errore del 32,5% dovuto a fattori sconosciuti. Il coefficiente di correlazione vale in questo caso 0,68, per cui anche questo modello può essere considerato accettabile. Il valore di 48,9% del fattore Musicista significa che i risultati sono determinati in larga parte dal comportamento motorio caratteristico dei diversi soggetti. La stima delle medie marginali (Figura 7) conferma questa osservazione: in presenza di DAF, alcuni soggetti tendono a muoversi molto di più (Bart, Jan, Frederik), altri si muovono meno (Marianne), altri ancora tendono a muoversi poco sia nella condizione normale sia in quella di feedback ritardato (Anke, Tine).

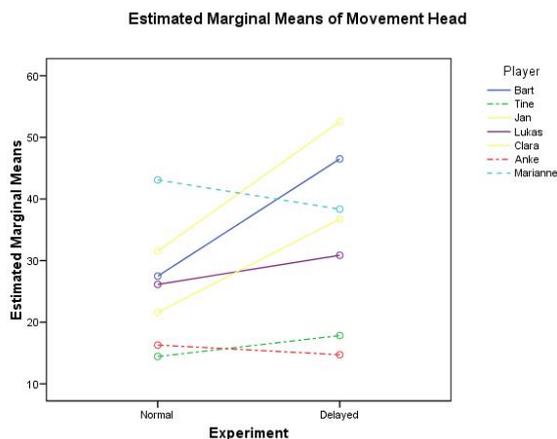


Figura 7. Medie marginali stimate dell'intensità di movimento per battuta della testa.

Un'ulteriore studio è stato effettuato sulla periodicità del movimento, in particolare sulla periodicità del beccheggio della testa. I valori del beccheggio (Figura 8) sono stati ricampionati in modo che ogni battuta avesse lo stesso numero di campioni; un'applicazione della Fast Fourier Transform ai dati così ottenuti ha evidenziato la presenza di diverse periodicità, in particolare in corrispondenza delle frequenze relative a periodi di una battuta e di due battute (Figura 9).

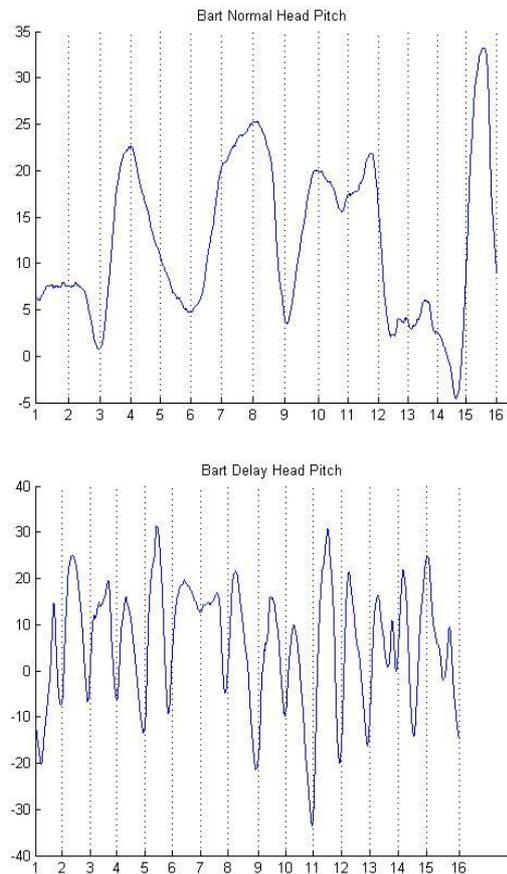


Figura 8. Il beccheggio della testa di uno dei musicisti nell'esecuzione in condizioni normali (sopra) e in condizioni di feedback ritardato (sotto).

Un'analisi qualitativa dei picchi principali mostra come per 5 soggetti (Bart, Tine, Jan, Lukas, Marianne) la frequenza del movimento corrispondente alla periodicità principale aumenti nell'esecuzione in presenza di DAF. Marianne, ad esempio, mostra una periodicità principale di due battute nell'esecuzione in condizioni normali, mentre in presenza di DAF la periodicità principale è di una battuta. Soltanto un soggetto (Clara) diminuisce la frequenza del movimento nel

passaggio da una condizione all'altra: in questo caso la periodicità principale passa da una battuta a due battute. In due soggetti (Anke, Lukas) la condizione di feedback ritardato causa un marcato smussamento della periodicità: i picchi della FFT non sono più particolarmente evidenti in presenza di DAF, mentre lo erano in condizioni normali. In un caso (Tine) la periodicità non era invece evidenziata nella condizione normale mentre lo è in quella di DAF. Il proseguimento degli studi quantificherà le osservazioni appena riportate.

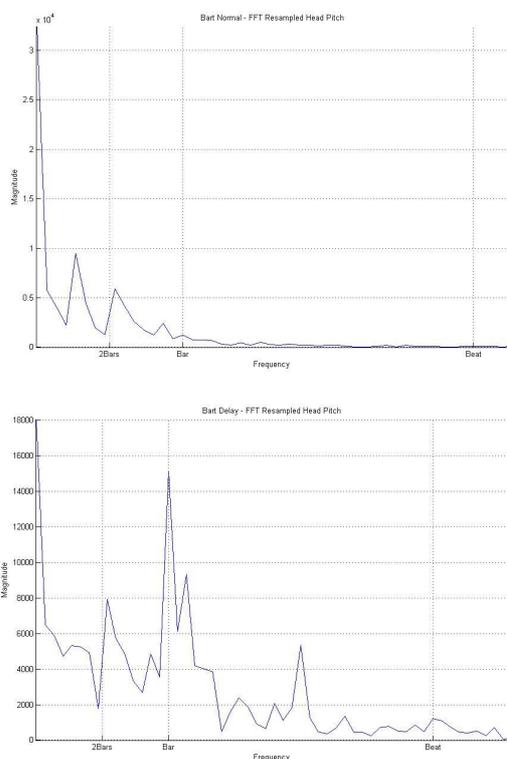


Figura 9. La FFT applicata alla versione ricampionata dei segnali di Figura 8. Le linee verticali tratteggiate indicano, da destra verso sinistra, le frequenze corrispondenti a due battute, una battuta, e un battito.

## COMMENTO DEI RISULTATI PRELIMINARI

I risultati preliminari non mostrano grandi differenziazioni tra le esecuzioni in condizioni normali e quelle in assenza di feedback: questo risultato è in linea con quanto riportato da numerosi studi in letteratura (Gates e Bradshaw, 1974; Banton, 1995; Finney, 1997; Repp, 1999; Finney e Palmer, 2003; Pfordresher, 2005), e può essere spiegato con il fatto che i musicisti dispongono di un'immagine uditiva e motoria del brano che viene utilizzata per sostituire le informazioni mancanti. Studi recenti (Highben and Palmer, 2004; Altmüller, 2004) sembrano confermare questa interpretazione. Il fatto che le due esecuzioni non differiscano è inoltre importante perché ci garantisce che i musicisti sono in grado di suonare due volte il brano alla stessa maniera: le differenziazioni che abbiamo notato nella terza condizione sono dunque dovute al DAF, come l'ANOVA conferma.

In base alle medie marginali stimate del parametro tempo è possibile suddividere i soggetti in tre gruppi: un primo gruppo (gruppo 1: Tine, Lukas, Clara) riesce a mantenere quasi costante la velocità dell'esecuzione anche in presenza di DAF; un secondo (gruppo 2: Jan, Marianne) mostra un forte rallentamento dovuto al DAF; infine un terzo (gruppo 3: Bart, Anke, Frederik) rallenta ma non vistosamente quanto i

musicisti del gruppo 2. Possibili spiegazioni di questi risultati sono state cercate nei dati riguardanti il movimento, nelle registrazioni video e nei questionari. Questa ricerca ha mostrato come i due estremi (gruppi 1 e 2) abbiano fatto ricorso a strategie diverse per contrastare l'effetto del DAF. Il gruppo 1 ha cercato esplicitamente di disaccoppiare l'esecuzione dal feedback uditivo, usando invece il proprio corpo come riferimento per il battito. Al contrario il gruppo 2 (Jan esplicitamente) ha continuato a fare riferimento al feedback: questo fatto, essendo il feedback percepito in ritardo, ha indotto un rallentamento molto accentuato nell'esecuzione.

Quasi tutti i soggetti, in presenza di DAF, hanno diminuito il periodo di beccheggio, passando da periodi di due battute a periodi di una battuta: la spiegazione va ricercata nel fatto che, in condizioni normali, il corpo risona con la struttura musicale, che nel caso della Sarabande è costituita da frasi di appunto due battute; in condizioni di DAF, invece, il disaccoppiamento di azione e percezione rende impossibile questa risonanza, per cui il musicista, confuso dal ritardo nel feedback, si concentra più sul suonare correttamente e sul mantenere il ritmo (periodo corto) che sul comunicare attraverso frasi musicali (periodo lungo). Questa perdita di espressività è confermata dall'analisi dei questionari: tutti i soggetti hanno affermato di non riuscire ad esprimere in presenza di DAF le stesse emozioni espresse in condizioni normali. La perdita di espressività sembrerebbe spiegare anche la diminuzione della varianza della key velocity media, ipotesi che andrà verificata con le successive analisi statistiche.

Un'altro risultato notevole è rappresentato dalla relazione che sembra legare l'intensità dell'esecuzione (key velocity) all'intensità del movimento: entrambi i parametri, nell'esecuzione affetta da DAF, aumentano drasticamente per gli stessi soggetti (Bart, Lukas, Clara). Questo risultato può essere spiegato con il fatto che la pressione sui tasti è strettamente collegata al movimento delle dita, e più in generale del corpo, per cui può essere vista anch'essa come una forma di misura del movimento corporeo. Il test ANOVA riguardante i dati di movimento indica come i fattori responsabili di questi drastici incrementi devono essere ricercati nelle caratteristiche soggettive (attitudini personale, educazione al movimento) dei musicisti. I risultati dell'ANOVA suggeriscono infatti che i soggetti che si muovono poco in condizioni normali tendano a muoversi poco anche in presenza di DAF, mentre i soggetti che si muovono molto abitualmente tendano a muoversi ancora di più in presenza di DAF.

Infine, il fatto che l'organista riesca a mantenere costanti tempo e key velocity media in tutte e tre le condizioni sembra confermare l'ipotesi che gli organisti sviluppino con l'esperienza strategie opportune per contrastare gli effetti negativi del DAF. In particolare, ed in relazione con quanto appena detto riguardo l'incremento della key velocity, si può notare come l'organista mantenga in tutte e tre le esecuzioni valori di questo parametro molto alti se confrontati con quelli dei pianisti. Allo stesso modo la varianza del parametro risulta essere molto ridotta rispetto a quelle dei pianisti, ad indicare come le variazioni di espressività legate alla dinamica siano pressoché assenti. Ciò si può spiegare facilmente con il fatto che l'intensità del suono dell'organo non è sensibile alle differenze di pressione esercitata sul tasto, per cui gli organisti si abituerebbero a suonare a intensità costante. Inoltre si può ipotizzare, come suggerito anche dalle risposte dell'organista al questionario, che un movimento delle dita forte e sicuro, e quindi una dinamica forte, aiuti l'esecutore a dissociarsi dal feedback. Ulteriori informazioni al riguardo verranno dall'analisi statistica delle esecuzioni degli altri organisti che hanno partecipato all'esperimento.

## SVILUPPI DELLA RICERCA

Innanzitutto la ricerca verrà completata con l'analisi statistica dei dati delle esecuzioni dei rimanenti sei organisti. Ciò permetterà di consolidare i risultati generali e di fare confronti più significativi tra i risultati dei pianisti e quelli degli organisti. Analisi più approfondite verranno poi fatte sull'espressività, ed in particolare sulle variazioni di varianza e sulla correlazione dei parametri tra le diverse esecuzioni. Verrà approfondita anche l'analisi dell'esecuzione in assenza di feedback, tralasciata per il momento in quanto secondaria ai fini della nostra ricerca. Per quanto riguarda l'analisi del movimento, verranno normalizzate e quantificate le variazioni nella periodicità del beccheggio della testa. Infine, ad appendice della ricerca, si concluderanno gli studi sulle misurazioni (rappresentate in Figura 10) del ritardo del feedback dell'organo pneumatico della chiesa di Sant'Anna a Gand. I risultati preliminari di questo studio mostrano come, per le basse frequenze, il ritardo del suono che arriva all'organista è dell'ordine dei 150 ms.

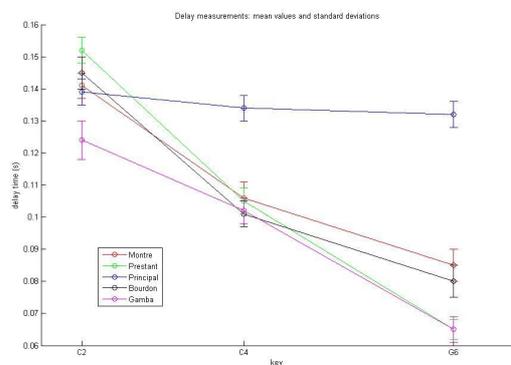


Figura 10. Misurazione dei ritardi del suono di alcune note (C2, C4, G6) per vari registri dell'organo di Sant'Anna a Gand.

## RIFERIMENTI

1. Marc Leman. *Embodied music cognition and mediation technology*, The MIT Press, Cambridge, Massachusetts, 2008.
2. James J. Gibson. *The ecological approach to visual perception*, Houghton Mifflin, Boston, 1979.
3. Steven A. Finney. *Auditory feedback and musical keyboard performance*, *Music Perception*, 15, 153-174, 1997.
4. Steven A. Finney. *Disruptive effects of delayed auditory feedback on motor sequencing*, dissertazione di dottorato non pubblicata, Brown University, Providence, RI, 1999.
5. Donald G. MacKay. *The organization of perception and action*, Springer, New York, 1987.
6. Karl U. Smith. *Delayed sensory feedback and behaviour*, Saunders, Philadelphia, 1962.
7. Anne Gates, John L. Bradshaw. *Effects of auditory feedback on a musical performance task*, *Perception & Psychophysics*, 16, 105-109, 1974.
8. Louise J. Banton. *The role of visual and auditory feedback during the sight-reading of music*, *Psychology of Music*, 23, 3-16, 1995.
9. Bruno H. Repp. *Effects of auditory feedback deprivation on parameters of expressive piano performance*, *Music Perception*, 16, 409-438, 1999.
10. Steven A. Finney, Caroline Palmer. *Auditory feedback and memory for music performance: Sound evidence for an encoding effect*, *Memory & Cognition*, 31, 51-64, 2003.
11. Peter Q. Pfordresher. *Auditory feedback in music performance: The role of melodic structure and musical skill*, *Journal of Experimental Psychology: Human Perception and Performance*, 31, 1331-1345, 2005.
12. Zebulon Highben, Caroline Palmer. *Effects of Auditory and Motor Mental Practice in Memorized Piano Performance*, *Bulletin of the Council for Research in Music Education*, 158, 58-67, 2004.
13. Eckart O. Altenmüller. *Music In Your Head*. *Scientific American Mind*, 14(2), 2004.

# Dalla Computer Music al Sound and Music Computing: Contesti culturali e tecnologici\*

**Nicola Bernardini**  
Conservatorio “Cesare Pollini”  
Padova  
nicb@sme-ccppd.org

**Alvise Vidolin**  
Conservatorio “Benedetto Marcello”  
Venezia  
vidolin@dei.unipd.it

## SOMMARIO

La musica informatica ha superato mezzo secolo di vita e riteniamo sia utile fare un bilancio del rapporto che lega la musica alla tecnologia nel passaggio dalla computer music al sound and music computing, allargando l’analisi all’evoluzione del contesto sociale e culturale in cui tale disciplina si è sviluppata. La computer music nasce in una fase storica dominata da un sistema politico di “blocchi contrapposti” in cui la cultura aveva un ruolo molto importante nei paesi occidentali anche per dimostrare al blocco opposto ed al resto del mondo i pregi e la ricchezza di una cultura libera, progressista ed innovatrice.

La fine della contrapposizione tra “blocchi” ha trasformato in maniera sostanziale la funzione delle politiche culturali dei paesi occidentali. Alla classica contrapposizione tra politiche “conservatrici” e politiche “progressiste” si è andata sovrapponendo una contrapposizione trasversale tra politiche “semplicitistiche” e politiche “consapevoli” con pesanti risultati sul piano culturale. Avvicinandoci alla fine del XX secolo, molti altri elementi si aggiungono a modificare il contesto culturale e tecnologico, favorendo il passaggio alla *sound and music computing*:

1. L’avvento del *internet*
2. La riproducibilità del contenitore
3. L’ampia affermazione del *Software Libero*
4. Lo sviluppo delle tecnologie *P2P*

Oggi, abbiamo tutti accesso a tecnologie di potenza insperata ed inattesa, coniugata con mobilità, reti, nanotecnologie e quant’altro. Ma sarà possibile generare cultura senza conoscere la storia, le radici che costituiscono le nostre basi del pensiero? Sarà possibile che una nuova cultura emerga dal rumore creato dalla miriade di sforzi individuali di una generazione abbandonata a se stessa?

Questo intervento vuol essere un invito a riflettere sui rapporti fra musica, tecnologia e cultura per rivedere “dall’interno le ragioni che stimolano l’evoluzione estetica e il processo creativo musicale.

## Parole Chiave

Musica Contemporanea; Informatica Musicale; Sound and Music Computing.

## INTRODUZIONE

Nel 1994, gli autori di questo saggio scrissero una breve panoramica delle risorse economiche alla base della ricerca e della produzione nel campo della musica elettronica dagli esordi sino a quel momento [5]. Scopo dichiarato di quella panoramica era quello di analizzare il passato e il presente (di allora) tentando di formulare delle ipotesi di sviluppo futuro.

Pur lasciando essenzialmente inalterata quella breve analisi, può essere interessante considerare i cambiamenti epocali intervenuti in seguito, i quali hanno introdotto modifiche rilevanti nel modo di pensare e di concepire la musica prodotta con l’ausilio di tecnologie informatiche e hanno reso urgenti le riflessioni sulle problematiche della conservazione e del recupero del patrimonio musicale. Alla base di tutti i cambiamenti di cui parleremo in questo saggio ce ne sono fondamentalmente tre, strettamente collegati tra loro: a) la fine della contrapposizione tra “blocchi” e la conseguente affermazione della globalizzazione dei mercati e delle società, b) l’affermazione delle *società dell’informazione* [7], e c) l’affermazione delle tecnologie legate alla mobilità e alla disponibilità di risorse computazionali e d’informazione *per tutti, dappertutto, in qualsiasi momento*. Naturalmente, tutti questi cambiamenti erano già in corso negli anni ’90 – e si può facilmente risalire ai loro prodromi negli anni ’80. Tuttavia, la diffusione estesa a livello sociale delle implicazioni che essi hanno comportato è avvenuta nel passaggio del millennio – ed è tuttora in corso. E, altrettanto naturalmente, le problematiche introdotte da questi cambiamenti sono lungi dall’aver trovato soluzione. Eppure non si tratta di problematiche di poco conto: fondamentalismi religiosi, scarsità di risorse energetiche, flussi migratori, terrorismo, *digital divide* appaiono – oggi – come problemi di complessità insormontabile.

## LE POLITICHE DELLA CULTURA E LA MUSICA, OGGI

La fine della contrapposizione tra blocchi di nazioni ad economia capitalista e blocchi di nazioni ad economia socialista ha trasformato in maniera sostanziale la funzione delle politiche culturali dei paesi occidentali. La svanita necessità di dimostrare al blocco opposto ed al resto del mondo i pregi e la ricchezza di una cultura libera, progressista ed innovatrice si è coniugata nel peggiore dei modi con l’emergenza

di fenomeni estremamente complessi da capire, analizzare ed affrontare – quali ad esempio il terrorismo planetario, la scarsità di risorse in un mondo che ha ripreso a crescere vertiginosamente, il divario sempre crescente tra nord e sud del mondo, ecc.

In questo contesto, alla classica contrapposizione tra politiche “conservatrici” (tipicamente collegate alla promozione di ideologie individualistiche), e politiche “progressiste” (collegate invece ad una attenzione nei confronti del sociale) si è andata sovrapponendo una contrapposizione trasversale tra politiche “semplicistiche” (p.es.: cacciare gli immigrati, abolire le tasse, liberare gli affari dalle leggi che li regolano, ecc.), e politiche “consapevoli” (p.es.: affrontare i problemi a livello globale, promuovere soluzioni sostenibili, ecc.). Entrambe queste politiche funzionano su un consenso di massa difficilmente raggiungibile senza precisi interventi a livello culturale. Tuttavia, per raggiungere un consenso culturale su politiche “semplicistiche” è sufficiente “giocare al ribasso”: è importante giocare su emozioni primarie e primordiali (paura, desiderio, rabbia, piccole felicità a basso costo, ecc.), paradossalmente sostituendo alla capacità individuale di riflessione una capacità di convincimento legata al bombardamento ripetuto effettuato attraverso mezzi di comunicazione di massa. È invece molto più difficile raccogliere consenso attorno alle politiche “consapevoli”: i processi di acquisizione della consapevolezza sono processi molto lenti, faticosi e quindi, intrinsecamente, rischiosi – essi richiedono tempi lunghi (ben al di là dei cicli elettorali), il dispiegamento di strategie educative di alto profilo, e lo sviluppo di una cultura che non viva sui propri allori (i cosiddetti *beni culturali*) ma che si sappia costruire e ri-costruire giorno per giorno in un incessante susseguirsi di innovazioni guidate dal filo rosso della memoria.

È proprio su questa difficoltà che si è arenata la macchina ideologica “progressista”: le perdite di consenso che sarebbero dovute essere considerate come fisiologiche sono state interpretate come inequivocabili segnali di precise richieste politiche (di semplificazione concettuale) da parte della base elettorale. A questi segnali si è risposto e si sta rispondendo nel peggiore dei modi: inseguendo il conservatorismo “semplicistico” sul suo terreno – cercando di far diventare la complessità della società odierna un problema di nicchia, relegato alle *élites* degli studiosi e degli accademici. Il problema è che, invece, questa complessità viene vissuta quotidianamente, in maniera più o meno tangibile, più o meno tragica, dalla stragrande maggioranza dei popoli. Ma per questi rimangono, appunto, solo soluzioni semplicistiche che non fanno altro che rimandare i problemi ad altri momenti ed in altri luoghi senza mai risolvere alcunché. La politica della cultura, così come l’abbiamo conosciuta dal secondo dopoguerra sino agli anni ’90, non esiste più.

In questo contesto, la produzione musicale non fa eccezione. L’attenzione delle “politiche culturali” di schieramenti di qualsiasi colore è esclusivamente rivolta ormai a fenomeni di massa – e quindi alla musica di consumo. Quest’ultima segue rigorose logiche di mercato – che poco hanno a che fare con lo sviluppo delle capacità intellettuali, mentre gli

spazi riservati alla creazione musicale con funzioni speculative (ad es. quelle di proporre nuove visioni, nuovi modi di pensare, ecc.) si stanno rapidamente estinguendo – per mancanza di fondi, di supporto politico, di attenzione sociale. La produzione di musica contemporanea, mortificata e vilipesa dalle necessità del consenso “semplicistico”, si ritira per forza di cose in se stessa, in ristretti gruppi di persone che sono ancora in grado di *pensare* la musica con l’acutezza e la profondità che ci è stata tramandata dalle generazioni precedenti. Naturalmente, questo “ritiro” viene generalmente tacciato di *élitismo* (cf., ad es., [9]) – come se i compositori contemporanei non avessero a cuore il problema del pubblico al quale rivolgersi. Anche questo fa parte delle ideologie “semplicistiche”: confondere l’effetto con la causa, in modo da trasformare la vittima in colpevole.

Purtroppo, la stragrande maggioranza dei compositori – anche quando possiede doti notevoli sul piano musicale – è totalmente sprovvista degli strumenti intellettuali per cogliere ed affrontare una situazione così degenerata. Nei casi migliori, essa si racchiude in furori e rancori personalizzati che non permettono ulteriori elaborazioni del problema. Nei casi peggiori, essa trasforma la propria produzione musicale in prodotti ibridi (fintamente *colti* e realmente *commerciabili*) che si sostituiscono, con la complicità di imprenditori e direttori artistici in cerca di facili successi, alla produzione musicale contemporanea propriamente detta. Così, il mondo della musica si riempie di improvvisatori di piano bar che si spacciano per compositori “di musica contemporanea”, di giovani che invocano “contaminazioni” artistiche improbabili, di “artisti sonori” quanto mai estranei alla musica e così via. Tutto questo marasma di produzioni musicali senza senso, senza storia e senza futuro rientra comunque perfettamente nella cultura promossa dalle politiche “semplicistiche”: il messaggio sotterraneo che filtra attraverso di esso indica che esiste una soluzione semplice alle complessità proposte dalla musica contemporanea (che erano e rimangono, ovviamente, un possibile specchio della realtà esistente) – quella di abbandonare qualsiasi possibilità di riflessione e di speculazione per lasciarsi andare a quell’edonismo di facciata tanto caro alla politica odierna.

La musica elettronica non versa in condizioni certo migliori. Già le conclusioni di [5] lamentavano la poca disponibilità degli enti lirici e delle istituzioni concertistiche a impegnare fondi *extra* (pur contenuti se paragonati, per esempio, ai costi di allestimento di un’opera lirica) per l’impiego di nuove tecnologie nelle compagnie strumentali dei (già rari) lavori commissionati. E pensare che istituzioni ed enti lirici avrebbero avuto e avrebbero tuttora lo strumento legislativo adeguato (la “legge Corona” della seconda metà degli anni ’60, non contraddetta sostanzialmente dalla normativa successiva – cf. [11]) per finanziare nuove creazioni e produzioni. Ma è significativo che queste leggi vengano generalmente vissute dagli impresari e dai direttori artistici non come validi sostegni economico-finanziari a supporto della ricerca e della sperimentazione musicale, ma piuttosto come fastidiose gabelle che sottraggono fondi alla produzione più tradizionale. L’avvento delle ideologie “semplicistiche” ha dato man forte a questa ritrosia – perché complicarsi la vita con la tecnolo-

gia quando tanta buona musica è stata scritta per gli organici tradizionali? – e la creazione di nuovi lavori che utilizzino tecnologie avanzate è diventato ormai un fatto più unico che raro. E di nuovo, spesso sono i compositori stessi a rifiutare la tecnologia: terrorizzati dal degrado descritto poc'anzi e dalla scarsità di lavoro sono pronti a compiacere i propri committenti in tutto pur di veder eseguita una propria opera.

A questa problematica se ne aggiunge un'altra, più sottile. Gli anni di fagocitazione strumentale da parte della musica di consumo (anch'essi descritti in [5]) hanno portato a gravi distorsioni sulla capacità di utilizzo "consapevole" delle tecnologie informatiche da parte dei compositori contemporanei. Paradossalmente, alla generazione di compositori che era cresciuta nella scoperta delle nuove tecnologie (e delle loro grandi limitazioni, allora) è seguita una generazione di compositori incapace di distinguersi da un uso puramente utilitaristico di queste ultime. In sintesi, le tecnologie informatiche hanno perso le loro funzioni di analisi, di scoperta e di invenzione di nuovi mondi sonori per ricadere nel loro ruolo più semplice – quello di strumenti ai quali non si chiede di più di quanto sia specificato nel loro libretto di istruzioni. In questo ennesimo "semplicismo", il senso dell'uso delle tecnologie avanzate di cui oggi disponiamo evapora lasciando spazio ad un'utilizzazione alienata e abbruttita di *software* e *hardware* pre-confezionati per altri scopi.

#### LA MUSICA NELLA SOCIETÀ DELL'INFORMAZIONE

Eppure, la società dell'informazione (così come viene descritta ad es. in [7]) introduce alcune opportunità che andrebbero colte in tutta la loro portata. L'avvento del *internet* commerciale e la massificazione delle connessioni tra le tecnologie informatiche rivoluziona il concetto di "proprietà intellettuale" – concetto quanto mai evanescente che, appunto, la rete mette in crisi proponendo invece un modello distribuito della conoscenza. Questa società segna la fine dei "contenitori", dei "supporti" usati dall'editoria convenzionale per ricavare un profitto variabile (spesso molto ingente) da un *contenuto* dai costi fissi. È anche la fine del "diritto d'autore" inteso come la possibilità da parte di ristretti oligopoli di lucrare indefinitamente sulla diffusione di cultura e conoscenza. È l'inizio di una diversa diffusione dei contenuti, basata sul riconoscimento della paternità (*authorship*) e su pratiche di diffusione più attente alle possibilità e modalità di accesso che all'eventuale profitto da esso derivabile nell'immediato<sup>1</sup>. La riproducibilità del *contenitore* – divenuta perfetta e a costo zero – perde qualsiasi possibilità di rappresentare un valore in sé, restituendo tale valore al *contenuto*. Alla svalutazione della riproducibilità si contrappone la rivalutazione del lavoro di produzione dei contenuti [3]. Naturalmente, con la sparizione degli editori sparirà anche la capacità di elaborare una politica culturale per la musica contemporanea. Questa capacità, benché spesso utilizzata in forma privatistica per difendere oligopoli chiusi ed arroccati in posizioni dominanti, ha avuto comunque il merito di cogliere e selezionare la migliore produzione musicale del

<sup>1</sup>Si vedano ad esempio le licenze d'uso *Creative Commons*, che vanno diffondendosi sempre più rapidamente nelle comunità scientifiche, umanistiche ed artistiche (cf. <http://creativecommons.org>)

dopoguerra. Sparita questa capacità, rimarrà il problema di elaborare valutazioni critiche della produzione contemporanea – lavoro gravoso e, ormai, di scarsa gratificazione; tuttavia, senza questa elaborazione la produzione culturale di qualità affogherà nel brusio indistinto della "musica inutile" prodotta ogni giorno. Nella società dell'informazione questo è però un problema comune: l'informazione diventa la risorsa fondamentale (al pari dell'energia nella società industriale) e la capacità di selezionare l'informazione "più pulita" diventa uno strumento essenziale di sopravvivenza. La ricerca in questo ambito è quindi molto attiva, ed è prevedibile che la rete stessa fornisca in tempi brevi una risposta efficace a questo problema<sup>2</sup>.

Per capire la portata rivoluzionaria della rete, basterà considerare la fragilità della memorizzazione digitale dei dati. Come si può combattere la volatilità della memoria digitale? Un sistema può essere quello della ridondanza. Più copie esisteranno dello stesso contenuto, più alta sarà la probabilità che almeno una copia (dalla quale sono ricavabili ulteriori copie identiche) rimanga leggibile. Incidentalmente, più queste copie sono distribuite su macchine diverse, piattaforme diverse, luoghi diversi, ecc. e più alta sarà la probabilità di trovare sempre almeno una copia del contenuto desiderato. Ed ecco che si profila all'orizzonte una prospettiva interessante: la copia digitale distribuita in rete e gli strumenti tecnologici che la facilitano (le cosiddette *tecnologie P2P*) – tanto vituperate da editori e produttori di contenuti – sono in realtà il più potente antidoto contro la perdita della memoria, poiché si tratta di strumenti che rendono triviale la ridondanza distribuita di contenuti di qualsiasi genere. La rivoluzione della società dell'informazione porta quindi con sé novità estremamente dirompendi (quali quelle, per esempio, delle sopraccitate *tecnologie P2P*). Se invece di essere ostacolate da ideologie aprioristiche esse venissero analizzate ed utilizzate nella maniera più utile, esse potrebbero fornire la soluzione a problemi estremamente urgenti e complessi che la conservazione si trova oggi ad affrontare.

Un'altra innovazione che si è definitivamente affermata con l'avvento dell'*internet* commerciale è stata quella concernente il *Software Libero*. Il *Software Libero* è software prodotto dal contributo complessivo di milioni di volontari attraverso l'auto-organizzazione distribuita della rete: come tale, esso è accessibile, utilizzabile, distribuibile e modificabile da chiunque [12]. Negli ultimi dieci anni, il *Software Libero* si è affermato al di là di ogni più rosea previsione – e l'adozione da parte di tutte le categorie sociali e professionali sta progredendo in forma geometrica anno per anno [14]. Il *Software Libero* è particolarmente interessante (e diffuso) in ambito musicale: le sue caratteristiche intrinseche di versatilità e di adattabilità lo rendono infatti particolarmente consono agli ambiti professionali della ricerca musicale ed affini e il suo sviluppo in quest'area è stato particolarmente intenso [4].

Anche le tecnologie dedicate al suono ed alla musica hanno

<sup>2</sup>Si vedano ad esempio le ricerche riguardanti i sistemi di *raccomandazione* (ad. es. [2]) e quelli di *reputazione* (ad. es. [8]).

	Musica Elettronica	Sound & Music Computing
Discipline	Composizione Elaborazione del Suono Psicoacustica	Composizione Elaborazione del Suono Psicoacustica Interazione cross-modale Design del Suono <i>Music Information Retrieval (MIR)</i>
Applicazioni	Musica Contemporanea	Musica Contemporanea Restauro di documenti sonori Installazioni Sonore Giochi Oggetti Quotidiani
Domini	Strumenti Musicali	Strumenti Musicali Intrattenimento Sicurezza Salute

Table 1. L'evoluzione della *Musica Elettronica* (Fonte: [13])

subito un notevole impulso nella nuova società dell'informazione. Alla consueta accezione "musica elettronica" (e di seguito *computer music*) si è andato sostituendo il più generale *Sound and Music Computing*, denominazione sotto la quale confluiscono pariteticamente tutte le discipline collegate al suono, alla musica e all'informatica (si veda la Tav.1).

## CONCLUSIONI

Il volume *Il Complesso di Elettra – mappa ragionata dei centri di ricerca e produzione musicale in Italia* [1] segnava, quindici anni orsono, l'inizio di un nobile e visionario tentativo: quello di coordinare i numerosi centri di produzione di musica elettronica italiani attraverso il CEMAT (CEntri Musicali ATtrezzati), una mappa delle strutture di produzione musicale equipaggiate con tecnologie informatiche. Quel volume segnava il riconoscimento ufficiale, in Italia, della nuova informatica (quella dei personal computers, della mobilità, delle reti) nel novero dello strumentario tecnologico della musica contemporanea: in effetti, a fianco degli studi sperimentali e pionieristici (che avevano avuto accesso per primi a sistemi di calcolo sufficientemente potenti per l'informatica musicale) erano sorti numerosi altri centri di produzione collegati alla ricerca ed alla realizzazione di nuovi lavori musicali.

A quindici anni di distanza, il CEMAT, diventato federazione, ha pubblicato un secondo volume che "fotografa", in qualche modo, la realtà italiana della musica contemporanea: il *Libro bianco sulla diffusione della musica contemporanea in Italia* [10]. Tra le altre cose, questo volume riporta un'interessante collezione di dati statistici e quantitativi che sottolineano quel degrado del supporto istituzionale della musica contemporanea che abbiamo con altri mezzi espresso in questo saggio (ad es. il crollo dei contributi alla musica contemporanea da parte del Fondo Unico dello Spettacolo – 50% dal 2002 al 2006, la crescita dei contributi locali e degli *sponsors* privati, la frammentazione della produzione, ecc.), alla quale rimandiamo volentieri il lettore per ulteriori approfondimenti.

In questa sede, riteniamo opportune alcune osservazioni nel

	Complesso di Elettra	Libro Bianco	%
anno di pubblicazione	1995	2008	
n.pagine	147	435	+196%
n.saggi	11	7	-36%
n.schede	28	93	+232%
n.schede dei centri musicali attrezzati	25	9	-64%

Table 2. Confronto tra *Complesso di Elettra* e *Libro Bianco*

confronto tra i due "contenitori". Queste osservazioni sono riassunte schematicamente in Tav.2. Innanzitutto, la consistenza è assai diversa (quasi il quadruplo delle pagine e ben più del quadruplo delle schede per il *Libro bianco*), il che è solo parzialmente giustificato dal fatto che il primo volume intendesse essere una mappa dei soli centri musicali di produzione con mezzi tecnologici, mentre il secondo riguarda tutte le realtà che si occupino, a diverso titolo, di musica contemporanea. Affinando lo sguardo non si può non notare invece che più della metà dei centri citati nel *Complesso di Elettra* siano letteralmente scomparsi (16 su 25). Ecco forse perché il CEMAT oggi non pubblica più un nuovo volume sulla sua base federativa (i centri musicali attrezzati, per l'appunto) ma accoglie e fa proprio un volume del più ampio coordinamento R.It.M.O (Rete Italiana dei Musicisti Organizzati): per la ragione più semplice, più cruda e più triste – perché questi centri sono in via di sparizione. Resta da capire se questa sparizione è legata ad una integrazione silenziosa e capillare in strutture ed associazioni musicali genericamente dedicate alla musica contemporanea (quelle riportate, appunto, nel *Libro bianco*), o se si tratta invece di una sparizione senza mezzi termini, anche legata agli argomenti esposti nelle sezioni precedenti. Un'ulteriore analisi qualitativa mostra che la maggioranza dei centri rimasti soddisfa essenzialmente le esigenze dei compositori che li hanno fondati. In pratica, i centri non servono più a promuovere linee o tendenze musicali, non sono più i catalizzatori di ambienti intellettuali e di politiche culturali: essi sono semplicemente gli strumenti e gli ambienti di lavoro di coloro che li abitano.

Si potrebbe ritenere che anche questo sia un segno dei tempi: l'informatica è ormai pervasiva e i compositori non hanno più bisogno di rivolgersi a dei "centri di produzione" per produrre la propria musica – e i pochi che rimangono utilizzano i propri centri come ambienti di lavoro e di promozione dei propri prodotti musicali. Ma la verità potrebbe anche essere un'altra – già enunciata in questo saggio: sta venendo a mancare la volontà di affrontare la complessità del comporre per e con gli strumenti informatici – complessità che ha richiesto e richiede tuttora la partecipazione attiva di ricercatori, scienziati e tecnici nel processo compositivo. Senza questa volontà, i compositori si limiteranno ad utilizzare *software* preconfezionati dando così luogo a risultati preconfezionati.

Siamo quindi di fronte ad un paradosso: nel momento in cui si va affermando il ruolo strutturale ed imprescindibile dell'informatica in tutti gli aspetti del nostro quotidiano, spariscono i centri (intesi non solo come strutture fisiche ma anche come aggregazioni di intelletti, di culture e di metodo-

logie) che pure questo ruolo (per quanto riguarda il suono e la musica) hanno contribuito a creare. Contemporaneamente (e trasversalmente), la musica elettronica, che si è trasformata ed arricchita nel più generale *sound and music computing* per far fronte alla molteplicità di ruoli assunta dagli strumenti informatici dedicati alla musica, viene abbandonata dai compositori attratti da soluzioni musicali “semplicistiche” o altrimenti terrorizzati dalla crescente difficoltà a far accettare le nuove tecnologie ai propri committenti. In pratica, nel momento in cui si potrebbe cogliere in pieno i frutti di quasi mezzo secolo di febbrile ricerca di intelletti brillanti e profondamente musicali, la sparizione di politiche culturali degne di questo nome e la profonda apatia che ha conseguentemente investito i compositori rischiano di lasciare ad altri (altre nazioni, altre musiche, altri contesti) i risultati migliori.

Non ci sono scorciatoie, non ci sono ricette semplici. Per evitare questa disfatta è necessario che tutti gli intelletti “consapevoli” si impegnino nel contrastare le tendenze “semplificatorie” con tutti i mezzi a loro disposizione: bisogna capire per esempio che l’informatica di oggi, ubiqua, pervasiva, mobile, condivisa e perfino democratica, è pronta a fornire mezzi potentissimi per realizzare questa azione di contrasto. È un passo piccolo, paragonato ai risultati che si potrebbero ottenere, ma bisogna farlo.

#### RIFERIMENTI

1. AA.VV. *Il Complesso di Elettra*. CIDIM, 1995.
2. G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):737–749, 2005.
3. P. Aigrain. *Causa Comune. L’informazione tra bene comune e proprietà*. Eretica Speciale. Nuovi Equilibri, 2007.
4. N. Bernardini, D. Cirotteau, F. Ekanayaka, and A. Glorioso. The AGNULA/DeMuDi distribution: GNU/Linux and Free Software for the pro audio and sound research domain. In *Proceeding of the Sound and Music Computing Conference ’04*, Paris, 2004.
5. N. Bernardini and A. Vidolin. Piccola Economia della Musica Elettronica. In *Il Complesso di Elettra*, pages 23–26. CIDIM, 1995.
6. N. Bernardini and A. Vidolin. La musica elettronica ed il restauro dei documenti sonori. *Economia della Cultura*, XVIII(2), 2008. in press.
7. M. Castells. *The Information Age: the rise of the Network Society*, volume 1. Blackwell Publishing, Inc., 2000.
8. A. Cheng and E. Friedman. Sybilproof reputation mechanisms. In *Proceedings of SIGCOMM’05 Workshops*. ACM, 2005.
9. A. Frova. *Armonia celeste e dodecafonìa. Musica e scienza attraverso i secoli*. Scienza. BUR Biblioteca Univ. Rizzoli, 2006.
10. R.It.M.O. *Libro bianco sulla diffusione della musica contemporanea in Italia*. CEMAT, 2008.
11. M. Ruggieri. L’Intervento Statale a Favore della Creazione Musicale. *Economia della Cultura*, X(3):311–324, 2000.
12. R. Stallman and L. Lessig. *Free Software, Free Society: Selected Essays of Richard M. Stallman*. Free Software Foundation, 2002.
13. The S2S<sup>2</sup> Consortium. *A Roadmap for Sound and Music Computing, Version 1.0*. The S2S<sup>2</sup> Consortium, 2007.
14. D. Wheeler. Why open source software / Free Software (oss/fs, floss, or foss)? look at the numbers! [http://www.dwheeler.com/oss\\_fs\\_why.html](http://www.dwheeler.com/oss_fs_why.html), 2007.



# Voice Painter: un'interfaccia multimodale per dipingere con la voce

Amalia de Götzen, Riccardo Marogna, Federico Avanzini  
Dipartimento di Ingegneria dell'Informazione  
Università di Padova  
{degotzen, marognar, avanzini}@dei.unipd.it

## SOMMARIO

Il presente lavoro descrive un sistema - denominato *voice painter* - che permette di dipingere su una tela virtuale attraverso l'uso della propria voce e del movimento del corpo. Tale sistema, trasforma lo spettatore della performance multimodale in un performer/en-attore, proponendo così una nuova metafora artistica.

## Parole Chiave

Enazione; Informatica Musicale; Sound and Music Computing; Performing Arts.

## INTRODUZIONE

L'argomento di questo lavoro è inerentemente multidisciplinare: per scoprire ed enfatizzare l'approccio enattivo nelle produzioni artistiche (danza, musica, pittura, scultura, ecc.) usando le nuove tecnologie, è necessario colmare il divario tra tecnologia e arte prendendo in considerazione i suggerimenti e le richieste degli artisti e le limitazioni e possibilità proposte dai tecnici.

Le arti e l'enattività possono essere considerati strettamente correlati anche nei lavori del passato, dove le tecnologie possono essere, semplicemente, un pennello e qualche colore, o il materiale di una scultura e gli strumenti utilizzati dall'artista per creare il proprio lavoro. In questi casi il ruolo dell'enazione, per coloro che esploreranno l'oggetto artistico, è limitato al muoversi intorno a quest'ultimo per coglierne i vari aspetti del suo significato. Due esempi molto noti ed appropriati sono i lavori di Close nella pittura [7] e quelli di Serra nella scultura [14].

Tuttavia, le nuove tecnologie di espressione artistica che contemplano l'interazione multimodale danno all'artista nuovi strumenti e nuovi modi di pensare al proprio lavoro, coinvolgendo gli spettatori in un'interazione enattiva nella fruizione di una specifica opera d'arte.

La realizzazione del sistema *voice painter* prende spunto dall'idea di esplorare ed enfatizzare l'approccio enattivo in ambito artistico attraverso l'utilizzo di tecnologie appropriate. Nelle forme artistiche classiche molto spesso lo spettatore è forzato dall'opera stessa ad esplorare, a muoversi e ad usare il proprio corpo per mettersi in relazione con l'opera d'arte. La mediazione della tecnologia, può enfatizzare questo tipo di relazione ed il sistema che viene descritto nei paragrafi che seguono ne è un esempio. Noi pensiamo che questo

approccio sia genuinamente enattivo e che suggerisca una "terza via" tra arti allografiche ed arti autografiche.

Questo saggio è organizzato in tre sezioni. La prima sezione riassume lo stato dell'arte e i principali punti di vista riguardanti il come le esperienze enattive informano le rappresentazioni artistiche. In questo contesto, le differenze tra arti autografiche e arti allografiche verranno analizzate nei termini delle loro proprietà enattive. La seconda sezione presenterà e descriverà i dettagli tecnici del sistema *voice painter*: verranno delineate le componenti *hardware* e *software* e verranno motivate le strategie principali di design dell'interazione. La terza sezione è dedicata alle possibili applicazioni del sistema, dalla discussione di uno scenario reale nel quale il sistema è già stato sperimentato da numerosi utenti sino ai miglioramenti futuri e all'analisi di nuovi scenari applicativi.

## ARTI AUTOGRAFICHE E ARTI ALLOGRAFICHE

Una delle principali categorizzazioni tra le diverse forme d'arte è quella di Goodman [10] il quale definisce le arti 'autografiche' e le arti 'allografiche':

the former cannot be noted and do not contemplate performance, while the latter can be translated into conventional notation, and the resulting 'score' can be performed with a certain freedom of variation.

(le prime non possono essere trascritte e non contemplano l'esecuzione, mentre le seconde possono essere trascritte in notazione convenzionale, e la 'partitura' risultante può essere eseguita con una certa libertà di variazione). La pittura e la musica sono le due espressioni artistiche generalmente utilizzate come esempi di questa distinzione. È difficile determinare le regole che abbiano generato un dato dipinto, e non c'è notazione o scrittura che possa aiutare qualcun altro a produrre una replica esatta dell'opera d'arte originale: è persino possibile definire ogni copia un 'falso'. In musica il punto di vista è completamente diverso: ciascuna copia/esecuzione di un brano è una possibile interpretazione. La scrittura permette a tanti musicisti diversi di suonare un dato lavoro musicale: i segnali musicali 'discreti' vengono dapprima scritti dal compositore e poi interpretati dai musicisti. Si può dire che mentre le arti autografiche sono arti ad uno stadio, le arti allografiche sono arti a due stadi. La distanza tra queste due forme d'arte può ridursi drasticamente nelle esecuzioni moderne nelle quali, per esempio, un dipinto può essere il risultato di una performance dal vivo: un/a ballerino/a che dipinge con il proprio corpo o un musicista

che controlli una tecnologia multimodale che produce una parte video mentre suona.

### L'Enattività nelle Arti

Uno dei principali risultati del progetto ENACTIVE<sup>1</sup> è stata la profonda e fruttuosa riflessione sul ruolo dell'enattività nel processo creativo artistico [3]. Questo argomento è particolarmente difficile da affrontare perché collega insieme concetti astratti che sfuggono alle definizioni (enazione, creazione): la conferenza Enactive/07, tenutasi a Grenoble, ha messo insieme numerosi contributi che possono essere analizzati per tracciare una sorta di "filo rosso" che colleghi diverse espressioni artistiche con l'intento comune di esplorare il processo creativo enattivo.

La teoria della percezione enattiva dichiara che non è possibile disaccoppiare schematicamente la percezione e l'azione, dal momento che ogni tipo di percezione è intrinsecamente attiva e premeditata. La "concezione enattiva" dell'esperienza è qualcosa che un animale *enagisce* nell'esplorazione del proprio ambiente [15, 20]. Da questo punto di vista, il soggetto degli stati mentali è tutto il corpo dell'animale (*embodied*) situato nel suo ambiente. La conoscenza enattiva viene poi immagazzinata in forma di risposte motorie acquisite attraverso l'atto del "fare" [12].

Il sapere enattivo può essere acquisito anche esplorando un dipinto o una scultura se il fruitore è immerso in questo anello di azione-percezione. Il tipico esempio di un'arte enattiva è la musica: un violinista ha bisogno di sentire e di ascoltare il suono per correggere ed affinare l'esecuzione. In questa prospettiva molte applicazioni artistiche enattive create con il supporto di tecnologie diverse esplorano strumenti virtuali attraverso diversi tipi di gesti e di posture. Nello stesso tempo, queste applicazioni devono prendere in considerazione il feedback specifico ricevuto, ad es., esplorando una superficie o utilizzando un arco su una corda [5]: noi percepiamo attraverso le nostre mani e le nostre dita specifiche sensazioni aptiche che stimolano il fruitore/esecutore a reagire per capire. Gli strumenti musicali virtuali devono quindi essere corredati da tecnologie aptiche che possano restituire le superfici e le forze coinvolte nel suonare uno strumento reale. Allo stesso modo, molte creazioni artistiche sono concentrate sull'uso della retroazione aptica per esplorare l'ambiente e, per esempio, un oggetto/scultura sul quale potremmo volgere il nostro sguardo.

### Dipingere con la voce

L'uso della voce come strumento enattivo è un argomento piuttosto inesplorato, in particolare nel contesto di possibili applicazioni artistiche. La voce è un mezzo di comunicazione umana universale, e viene utilizzato anche per trasmettere elementi non-verbali e paralinguistici: emozioni, prosodia, stress. In più, la voce e la parola sono sempre accompagnate da altri canali di comunicazione non-verbale, tra i quali l'espressione facciale, il gesto e il linguaggio del corpo. Una visione influente nella psicologia cognitiva e nella linguistica stabilisce che il gesto e il parlato formano un

sistema di comunicazione unico, e che il gesto è una componente integrale della comunicazione anziché un semplice accompagnamento o ornamento [13].

Queste osservazioni forniscono la motivazione di base per lo sviluppo di un'interfaccia che utilizzi l'espressione vocale come mezzo per creare segni visivi. L'idea centrale è di esplorare le caratteristiche più rilevanti dell'espressione vocale (includendo anche caratteristiche prosodiche come l'altezza, l'intensità, ecc.) e di correlarle a caratteristiche grafiche, creando così uno strumento semplice e versatile che possa essere utilizzato da un esecutore esperto ma anche da un utilizzatore alle prime armi.

L'esecutore potrà dipingere su uno schermo nero utilizzando la propria voce, e potrà scoprire, nel contempo, le correlazioni tra il suono prodotto ed il dipinto che apparirà in tempo reale. La voce potrà essere considerata come un pennello: per disegnare sull'intera superficie dello schermo l'esecutore dovrà muoversi, coinvolgendo così tutto il corpo e non la sola voce. L'anello chiuso azione-percezione verrà così ricreato con l'aiuto di un sistema che verrà descritto nella sezione che segue.

### DESCRIZIONE DEL SISTEMA

Il sistema è costituito da un array di 8 microfoni i cui segnali sono processati in tempo reale per la localizzazione ed il tracking della posizione dell'utente. Si suppone che l'utente si muova in un'area ad una determinata distanza dallo schermo, definendo un piano di interesse. Il rendering grafico è retro-proiettato sullo schermo al fine di ottenere una perfetta corrispondenza tra la posizione della sorgente vocale e la posizione del rendering grafico da questa prodotto. In questo modo il feedback visivo è temporalmente e spazialmente correlato con l'evento sonoro. Dal segnale processato vengono inoltre estratti una serie di parametri indicativi di alcune caratteristiche notevoli del segnale vocale, che vengono poi mappati in altrettante caratteristiche del tratto grafico.

### Localizzazione e tracking in tempo reale della sorgente vocale

Il sistema audio realizza la localizzazione ed il tracking del performer attraverso un algoritmo organizzato in 2 fasi. Dapprima, per ciascuna coppia di sensori viene stimato il ritardo di arrivo (TDOAE, *Time Delay Of Arrival Estimation*). Il vettore di ritardi così stimati viene processato attraverso un algoritmo di stima ai minimi quadrati al fine di ottenere la coppia di coordinate (x,y) nel piano di interesse. Una stima del rapporto tra le ATF (*Acoustical Transfer Functions*) [9]. Il vettore di ritardi risultante da questa prima fase viene confrontato con gli elementi di una matrice pre-calcolata contenente un set di vettori ottenuti tramite discretizzazione del piano di interesse ed il successivo calcolo, per ciascuna posizione così ottenuta, del corrispondente vettore di ritardi. La procedura di *fitting* ai minimi quadrati del vettore stimato con gli elementi della matrice delle posizioni realizzabili fornisce quindi una stima della posizione (x,y) della sorgente vocale. Questa fase della procedura di localizzazione può essere molto pesante dal punto di vista computazionale se si desidera una certa accuratezza nella stima; essa può es-

<sup>1</sup><http://www.enactive.org>

sere altresì alleggerita con l'introduzione di un algoritmo di tracking della sorgente. Tale algoritmo di stima della posizione è stato implementato in linguaggio C ed integrato nella piattaforma per l'elaborazione audio in tempo reale Pure Data [17]. I segnali audio vengono acquisiti ad una frequenza di campionamento pari a 32 kHz (un compromesso tra accuratezza e velocità).

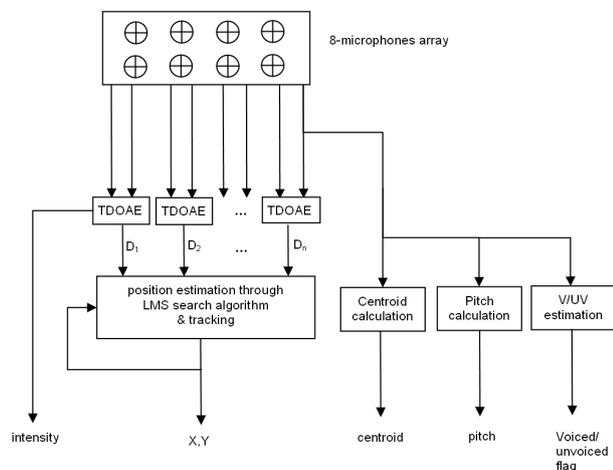


Figura 1. Schema a blocchi del sistema di localizzazione ed estrazione delle features.

### Mappatura dei parametri vocali e rendering grafico

La voce può essere definita da numerose caratteristiche che gli esseri umani sono in grado di controllare tramite la loro emissione vocale.

L'interfaccia multimodale è stata progettata al fine di realizzare una mappatura delle caratteristiche rilevanti della voce in qualità grafiche ben riconoscibili (forma, dimensione, colore) seguendo un approccio il più possibile naturale.

Tale mappatura è illustrata in Fig.2. Tra i diversi parametri che possono caratterizzare la voce umana ne sono stati identificati alcuni notevoli che potessero essere controllati facilmente dal performer e stimati in tempo reale con un certo grado di robustezza:

- *Intensità*, calcolata come RMS il quadrato della pressione del segnale;
- *Centroide*, il centro di gravità della ampiezza dello spettro calcolata su un frame audio;
- *Suono vocalizzato/non vocalizzato*, caratteristica che dipende dalla presenza o meno di vibrazioni pseudo periodiche
- *Altezza*, l'attributo soggettivo della frequenza di un suono;

Sebbene per le prime due caratteristiche sia stata utilizzata la definizione comune e siano facili da estrarre, le altre due non sono altrettanto banalmente stimabili. L'attributo relativo al suono vocalizzato indica la presenza di periodicità dovuta alle vibrazioni delle corde vocali. Questo tipo di segnali possono essere ricavati tramite diversi approcci. Qui è stata im-

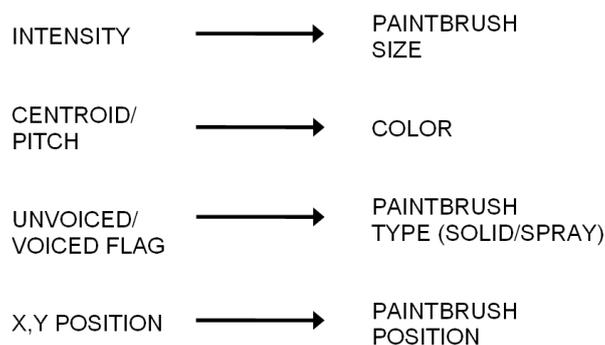


Figura 2. Mappatura delle caratteristiche vocali nelle caratteristiche grafiche.

plementata la tecnica che combina il rilevamento dello zero-crossing e l'estrazione mediante cepstrum [1]. L'estrazione dell'altezza è in questo caso facilitata dalla struttura relativamente semplice dei segnali vocalizzati, così che il problema si riduce alla stima della frequenza fondamentale. Questa è stimata usando un algoritmo che estrae e correla le componenti spettrali armoniche su frames successive del segnale vocale [16].

Il rendering grafico è stato realizzato utilizzando la libreria GEM (Graphical environment for Multimedia [21]), che ben si integra nella piattaforma Pure Data. Sono state utilizzate caratteristiche grafiche ben riconoscibili (dimensione, colore, geometria) con l'intento di seguire un approccio il più possibile naturale. Questo è un problema non triviale dato che il termine "naturale" non può essere definito a priori, ma può essere solo definito sulla base delle proprie esperienze personali o su scelte arbitrarie (questo problema può essere confrontato con il suo duale, la sonificazione). La naturalezza di un dato mapping dovrebbe quindi essere testata su diversi soggetti in modo da essere giustificata ed eventualmente modificata. Questo risulta particolarmente vero in alcune delle scelte fatte per il voice painter: la corrispondenza tra la frequenza di un suono con la frequenza della luce, anche se piuttosto logica scientificamente, potrebbe non necessariamente essere di facile (e.g. naturale) comprensione per l'utente. Un'altra possibile mappatura potrebbe essere quella sfruttando l'insieme RGB dei colori, associando suoni gravi a colori "caldi" e suoni acuti a colori "freddi". L'effetto visivo risultante è una sorta di schizzo astratto che può contenere elementi geometrici ben definiti (corrispondenti a brevi tratti vocalizzati) e tratti nebulari dovuti a segnali non vocalizzati, entrambi scalati a diverse dimensioni e con diversi colori (Fig.3).

### DISCUSSIONE

L'interfaccia è stata sottoposta ad un test preliminare in occasione della Conferenza ENACTIVE 07 svoltasi a Grenoble. Estratti video di alcune sessioni di test possono essere consultati al seguente link: [www.smc.dei.unipd.it/voicePainter.html](http://www.smc.dei.unipd.it/voicePainter.html). Ciascun utente poteva testare il sistema interagendo con la tela virtuale senza alcuna spie-

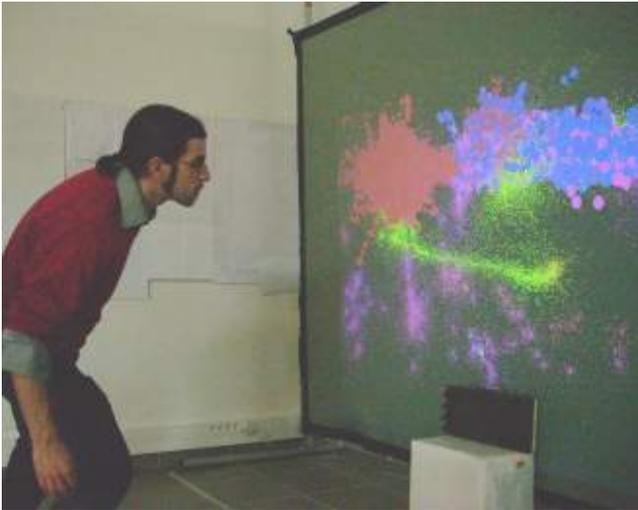


Figura 3. Il voice painter in azione.

gazione preliminare relativa al mapping e senza alcuno specifico task da compiere. L'unica indicazione data era "Usa la tua voce per dipingere". L'obiettivo del test era di valutare

- il modo in cui l'utente si relazionava allo schermo e come esplorava le diverse caratteristiche vocali;
- quali caratteristiche vocali l'utente identificava nel mapping proposto;
- se l'utente era in grado di controllare la/le caratteristica/e identificate in modo da ottenere l'effetto grafico desiderato.

Alla fine del test all'utente era richiesto di giudicare la naturalezza dell'esperienza vissuta. Questo test preliminare ha rivelato che la maggior parte delle caratteristiche vocali utilizzate erano state individuate, con l'unica eccezione del mapping altezza/centroide-colore. Sebbene il test sia stato realizzato in maniera piuttosto informale, la presenza di molte persone che assistevano alle "performances" molto spesso inibiva l'utente, lasciandolo meno libero e spontaneo nell'esplorazione del sistema. Al tempo stesso il feedback visivo è risultato utile per incoraggiare l'utente ad usare la voce nei modi più disparati per ottenere effetti grafici interessanti e particolari. Il prossimo passo di questo lavoro riguarderà la realizzazione di test quantitativi con sistematica raccolta dei commenti degli utenti.

### Localizzazione e tracking del parlatore

Uno degli aspetti che si sta attualmente sviluppando riguarda il miglioramento della localizzazione e del sistema di tracking, con particolare riguardo alla robustezza degli algoritmi in contesti rumorosi e quindi più vicini a scenari reali. La stima tramite TDOA e l'algoritmo LMS descritti nella seconda parte di questo articolo, considerano solo sottofondi con rumori impulsivi o stazionari e la performance peggiora in presenza di bassi valori di SNR, specialmente quando il rumore è non stazionario. Per aumentare la robustezza di tale sistema, il primo miglioramento nella procedura di

tracking riguarda l'inclusione di un algoritmo più generale di rilevamento di attività vocale (VAD), che impiega statistiche più accurate che permettono il rilevamento della presenza o meno del parlato in qualsiasi regione dello spettro audio [19]. È stato sviluppato anche un sistema bimodale nel quale la localizzazione ed il tracking sono realizzati non solo sfruttando il segnale vocale ma anche tramite quello video. Una implementazione preliminare è stata sviluppata usando la piattaforma software EyesWeb per l'analisi gestuale [6].

Un ulteriore futuro sviluppo riguarda il problema della localizzazione simultanea di più sorgenti acustiche, che richiede l'utilizzo di più avanzati metodi statistici, come il *particle filtering* [2].

### Riabilitazione vocale

Oltre alle applicazioni in ambiti artistici e di intrattenimento, il sistema che abbiamo proposto apre possibili strade di ricerca nella riabilitazione vocale. Le tecniche derivate dall'HCI sono sempre più utilizzate come metodi per insegnare e rafforzare le capacità vocali in diversi contesti, con particolare riferimento al processo educativo di bambini che abbiano alcuni deficit di comunicazione [4, 8, 11, 18]. Possibili aree di applicazione includono sia disfunzioni della parola dovuti a disabilità fisiche o a problemi nel coordinamento (dysarthria, dyspraxia) per i quali gli esercizi di articolazione, effettuati molto spesso durante attività ludiche, sono usati per mostrare ai bambini come pronunciare correttamente i suoni e le sillabe. Precedenti lavori hanno mostrato che sistemi di training del parlato basati su tecniche informatiche possono essere molto utili per suggerire il miglioramento dell'articolazione [4] e per aiutare l'apprendimento della corretta pronuncia di suoni specifici [18].

Una seconda possibile area di applicazione riguarda i problemi di ritardo del linguaggio (i.e. situazioni in cui lo sviluppo del linguaggio non segue il pattern usuale, ma ha un andamento più lento del normale) [8] ed infine la terza area è relativa a problemi di comunicazione, legati allo spettro dei disordini di tipo autistico (ASD) [11]. Il voice painter presentato in questo lavoro incoraggia il comportamento ludico dell'utente, maggiormente a suo agio nell'interazione con una macchina che non lo intimidisce, ma che anzi riduce o addirittura elimina il disagio causato dall'interazione con altre persone.

### CONCLUSIONI

Questo articolo presenta un'interfaccia che apre numerose vie di esplorazione. Il voice painter è nato pensando ad applicazioni artistiche: creare uno strumento che potesse mediare tra le arti autografiche ed allografiche, dando vita a performances sofisticate basate su una notazione musicale (usata poi per dipingere), ad improvvisazioni o a semplici occasioni di intrattenimento. Lo strumento è stato mostrato alla Conferenza Enactive/07 ed ha suscitato particolare interesse tra artisti visivi. Diversi scenari applicativi sono quindi stati proposti, suggerendo futuri sviluppi: oltre alle applicazioni artistiche, risultano particolarmente promettenti quelle relative a terapie vocali e a problemi di comunicazione.

## RINGRAZIAMENTI

Si ringraziano Barbara Mazzarino e Giovanna Varni del DIST di Genova per i loro preziosi suggerimenti e per il loro aiuto nello sviluppo della prima implementazione del sistema, presentato alla Conferenza Enactive/07 a Grenoble..

## RIFERIMENTI

1. S. Ahmadi and A. S. Spanias. Cepstrum-based pitch detection using a new statistical v/uv classification algorithm. *IEEE Trans. on Speech and Audio Processing*, vol. 7, no. 3, May 1999.
2. F. Antonacci, D. Riva, M. Tagliasacchi, and A. Sarti. Efficient localization and tracking of two acoustic sources using particle filters with swarm intelligence. In *Proc. EURASIP European Sig. Process. Conf.*, Poznan, 2007.
3. Association ACROE, editor. *Proc. 4th Int. Conf. on Enactive Interfaces (ENACTIVE/07... Enaction in Arts)*, Grenoble, 2007.
4. O. Bälter, O. Engwall, A.-M. Öster, and H. Kjellström. Wizard-of-oz test of artur: a computer-based speech training system with articulation correction. In *Proc. 7th Int. ACM SIGACCESS Conf. on Computers and accessibility (ASSETS'05)*, pages 36–43, Baltimore, 2005.
5. C. Cadoz. Musical creation process and digital technology. the supra-instrumental gesture. In *4th International Conference on Enactive Interfaces (Enactive'07)*, Grenoble, France, November 2007.
6. A. Camurri, B. Mazzarino, and G. Volpe. Analysis of expressive gesture: The EyesWeb expressive gesture processing library. In A. Camurri and G. Volpe, editors, *Gesture-based Communication in Human-Computer Interaction*. LNAI 2915, Springer Verlag, 2004.
7. C. Close, R. Storr, K. Varnedoe, D. Wye, and G. D. Lowry. *Chuck Close*. The Museum of Modern Art, New York, 2002.
8. H. Fell, C. Cress, J. MacAuslan, and L. Ferrier. visiBabble for reinforcement of early vocalization. In *Proc. 6th Int. ACM SIGACCESS Conf. on Computers and accessibility (ASSETS'04)*, pages 161–168, Atlanta, 2004.
9. S. Gannot, D. Burshtein, and E. Weinstein. Signal enhancement using beamforming and non-stationarity with application to speech. *IEEE Trans. Signal Processing*, 49 (8), 2001.
10. N. Goodman. *Languages of Art*. Hackett, 1985.
11. J. Hailpern, K. Karahalios, J. Halle, L. DeThorne, and M.-K. Coletto. Visualizations: speech, language & autistic spectrum disorder. In *Proc. ACM Computer-Human Interaction Conf. (CHI'08)*, pages 3591–3596, Firenze, 2008.
12. A. Luciani, J. Florens, and N. Castagne. From action to sound: a challenging perspective for haptics. In *Workshop on Enactive Interfaces (Enactive'05)*, Pisa, Italy, January 2005.
13. D. McNeill. *Gesture and Thought*. University of Chicago Press, Chicago, 2005.
14. K. McShine, L. Cooke, J. Rajchman, B. Buchloh, and R. Serra. *Richard Serra Sculpture: Forty Years*. The Museum of Modern Art, New York, 2007.
15. A. Noe. *Action in perception*. MIT press, Cambridge, Mass., 2005.
16. M. Puckette and T. Apel. Real-time audio analysis tools for pd and msp. In *Proc Int. Computer Music Conf.*, pages 109–112, San Francisco, 1998.
17. M. S. Puckette. Pure data: another integrated computer music environment. *Proc. of Int. Computer Music Conf.*, pages 37–41, 1996.
18. L. I. Shuster, D. M. Ruscello, and A. R. Toth. The use of visual feedback to elicit correct /r/. *Am. J. Speech-Language Pathology*, 4:37–44, 1995.
19. S. G. Tanyer and H. Ozer. Voice activity detection in nonstationary noise. *IEEE Trans. Speech Audio Process.*, 8(4):478–482, 2000.
20. F. Varela, E. Thompson, and E. Rosch. *The Embodied Mind*. MIT Press, Cambridge, MA, 1991.
21. J. M. Zmölnig. Gem, <http://gem.iem.at>.



# Un modello per la generazione dinamica di paesaggi sonori

**Mattia Schirosa**

mattiaschirosa@yahoo.it  
Politecnico di Torino, III Facoltà

**Andrea Valle, Vincenzo Lombardo**

andrea.valle@unito.it, vincenzo@di.unito.it  
CIRMA (Centro Interdipartimentale di  
Ricerca sulla Multimedialità e  
l'Audiovisivo)  
Università di Torino

## ABSTRACT

Il progetto si propone di definire un modello per la generazione dinamica di paesaggi sonori e di implementarlo in un'applicazione che permetta l'esplorazione interattiva e in tempo reale degli stessi. Oltre ad essere utilizzabile autonomamente, l'applicazione risultante è pensata per essere integrata in un motore di realtà virtuale. In questo modo il paesaggio sonoro può completare acusticamente l'esplorazione di un paesaggio architettonico-urbanistico.

## PAROLE CHIAVE

Paesaggio sonoro; SoundScape Studies; Sound Design; Oggetto sonoro.

## INTRODUZIONE

Una definizione rigorosa del concetto di “paesaggio sonoro” (*soundscape*) richiede un'analisi dei molti studi che utilizzano il termine e che delimitano un campo di interessi in piena fase evolutiva, in cui la terminologia presenta forti ambiguità. Si considera come punto di partenza la ricerca sull'estetica ecologica ed acustica, iniziato dal lavoro del World Forum For Acoustic Ecology a partire dalle esperienze di R. Murray Schafer ([1]) che ha per la prima volta introdotto il termine “paesaggio sonoro”. In seguito il concetto si è sviluppato disordinatamente in differenti ambiti di ricerca: dal multimedia ([2]) alla psicoacustica ([3]), dagli studi sull'ambiente lavorativo ([4]) alla pianificazione urbanistica ([5]), dal game design ([6], [7]) alla realtà virtuale ([8]), dalla sonificazione dei dati ([9]) all'ubiquitous computing ([10], [11]). In particolare, è una nozione fondamentale per il design acustico ([12], [13]), la composizione elettroacustica ([14]), gli studi sui display uditivi ([15]). Una simile complessità interdisciplinare (caratteristica degli studi: [16], [17]) si riverbera sulla definizione stessa del termine, che tende ad essere fortemente polisemico. D'altra parte, già nella concisa definizione di Murray Schafer, un “paesaggio sonoro” è semplicemente “l'ambiente dei suoni”.

In sintesi, è possibile individuare tre diverse accezioni del termine:

- Musicale/relata al sound design ([19]). Quello musicale è certamente l'ambito di ricerca in cui è nata la sensibilizzazione all'ascolto attento del soundscape come “composizione musicale naturale”. L'etnomusicologia, il bruitismo, la *musique d'ameublement* e la *musica aneddotica* hanno fertilizzato la riflessione sul concetto di paesaggio sonoro come scenografia acustica, o come atmosfera scenica. Parallelamente, la *musica concreta* ha stimolato la riflessione compositiva sull'oggetto sonoro. Inoltre, i sound designer per il cinema e la tv contribuiscono alla diffusione del termine “paesaggio sonoro” creando scenografie rumoristiche dei film e perfezionando l'intersezione audiovisiva che è alla

base dell'illusione cinematografica ([34], in generale cfr. [30]).

Il concetto di paesaggio sonoro è presente in molte forme musicali contemporanee ([16]): ad esempio, l'*environmental music* compone lo spazio stesso lavorando sull'interattività di installazioni sonore con l'ascoltatore, la *soundscape composition*, praticata prevalentemente in forma elettroacustica, opera sulla giustapposizione di parti paesaggistiche a volte unite a ridotte partiture musicali, l'*ambient music* simile alla precedente, si differenzia per una maggior presenza di partiture musicali e per l'utilizzo di strumenti acustici

- Eco-antropologica ([18]). In una simile accezione, attraverso un'attenta analisi sociale ed antropologica del design acustico, l'accento è posto sulla dimensione ambientale e culturale del suono. Una prospettiva eco-antropologica mira altresì alla creazione di un'estetica acustica capace di guidare la qualità dei progetti applicativi. Infine, impone una riflessione sull'importanza di organizzare l'archiviazione delle tracce sonore che costituiscono il retaggio culturale di uno specifico territorio, al fine di tramandarne la testimonianza alle generazioni future.
- Architettonico-urbanistica ([20]). La tecnologia elettroacustica in combinazione con l'acustica architettonica permette di pensare all'emergere del suono nello spazio in modo nuovo, per sensibilizzare il cittadino all'ascolto cosciente dei propri paesaggi sonori quotidiani, fino alla possibilità di partecipare attivamente alla composizione e progettazione delle sonorità e delle atmosfere che riempiono le città. In tal senso in questi ultimi anni, diversi progetti di architettura stanno cercando di dare il loro contributo sperimentale ([35], [36]). Un concetto nuovo ed interessante prodotto da questo ambito è la “liuteria urbana”, intesa come progettazione congiunta -architettonica e dei materiali- per la fabbricazione di componenti monumentali all'interno degli spazi pubblici, capaci di agire come casse di risonanza per l'ambiente sonoro circostante [36]

Va notato che la ricchezza di determinazioni che riceve il concetto di paesaggio sonoro è estremamente importante perché mette in luce come il problema dello spazio sonoro non possa essere risolto soltanto in termini acustici o psico-acustici. Il modello tradizionale dei processi acustici tratta il suono come trasferimento di energia attraverso la catena “fonte - canale di trasmissione - ricevente - effetti” ampliato dal modello di trasmissione dei messaggi con i concetti di codifica e decodifica ([32]) intesi come stati significativi del trasferimento dei segnali. Questo approccio si concentra sul calcolo della qualità del suono esclusivamente attraverso un parametro misurabile, la “fedeltà” del segnale alla fine della

catena di trasmissione. È necessario tuttavia estendere l'analisi ad una riflessione fenomenologica e semiotica, non lontana ad esempio dagli studi sull'audiovisione ([21]). È evidente, poi, l'importanza di differenti processi d'ascolto nella percezione del paesaggio sonoro: ascolto indicale, inteso come ricerca della fonte o della causa del suono, individuazione figurativa di senso; ascolto simbolico, che si manifesta nella ricerca di un senso prestabilito e culturalmente codificato; ascolto iconico, come generazione di nuovi significati a partire dal materiale sonoro ([23]).

Ancora, una definizione cognitiva e fenomenologica di suono è alla base del concetto di oggetto sonoro ([22], [23]), che può essere utilmente integrato in termini di plausibilità ecologica e cognitiva attraverso la teoria dell'analisi della scena uditiva ([24]). Il concetto di oggetto sonoro permette di scomporre il paesaggio sonoro in unità elementari utilizzabili nella modellazione della struttura complessa. Un oggetto sonoro è un'unità di analisi che pertiene all'ascolto e non coincide necessariamente con un segnale audio. Non è sempre agevole definire una correlazione tra livello acustico e livello fenomenologico, cioè individuare un segnale che corrisponda ad un oggetto sonoro ([28]). Il problema è particolarmente evidente nel momento in cui gli oggetti sonori individuati si presentano in sovrapposizione. Maggiore è la complessità del paesaggio sonoro, minore è la possibilità di ricondurre ogni oggetto sonoro ad un materiale sonoro. Nell'analisi di un paesaggio ampio, alla difficoltà nella percezione e nel riconoscimento degli oggetti sonori si aggiunge allora la difficoltà di prelevamento dei segnali corrispondenti: ad esempio, nel paesaggio sonoro di una cucina la scomposizione della struttura in elementi unitari è molto vaga, alcuni elementi sono di facile riconducibilità, il fuoco, l'acqua, lo sbattere degli sportelli, le voci dei cuochi, ecc., ma esiste un tappeto

composto da innumerevoli eventi sonori che si sovrappongono, rendendone difficile la riconducibilità a fonti specifiche.

Si può parlare di "materiale sonoro" per descrivere la realizzazione effettiva di un oggetto sonoro in termini di segnale audio, a seguito della registrazione e dell'elaborazione digitale.

### UNA DEFINIZIONE OPERATIVA DI PAESAGGIO SONORO

La complessità semantica del termine richiede di formulare una definizione esplicita di "paesaggio sonoro", che possa condurre operativamente ad una modellizzazione. Qui di seguito con "paesaggio sonoro" si intenderà dunque una struttura complessa di oggetti sonori (realizzati in materiali sonori semplificati) relativi a un certo contesto geo-culturale, rispetto alla quale è definita una trasformazione spaziale e semiotica che dipende dall'ascoltatore.

La trasformazione è spaziale perché il modello generativo del paesaggio sonoro deve tenere conto della posizione delle fonti sonore e dell'ascoltatore, selezionando i soli oggetti sonori significativi rispetto all'area percettiva dell'ascoltatore e applicando opportune trasformazioni acustiche relative alle condizioni di ascolto.

La trasformazione è anche semiotica perché la costruzione di senso passa attraverso l'ascolto figurativo/causale, iconico e simbolico, che influenzano la scelta degli oggetti sonori costituenti la struttura. Tale dimensione semiotica implica un attento processo di selezione dei materiali sonori di partenza. L'intenzione è quella di ricreare una particolare "atmosfera scenica" che include esplicitamente una dimensione emotiva e culturale ([31]). Dunque, ad un'analisi acustica e fenomenologica del paesaggio sonoro deve aggiungersi un'analisi geo-culturale dello stesso.

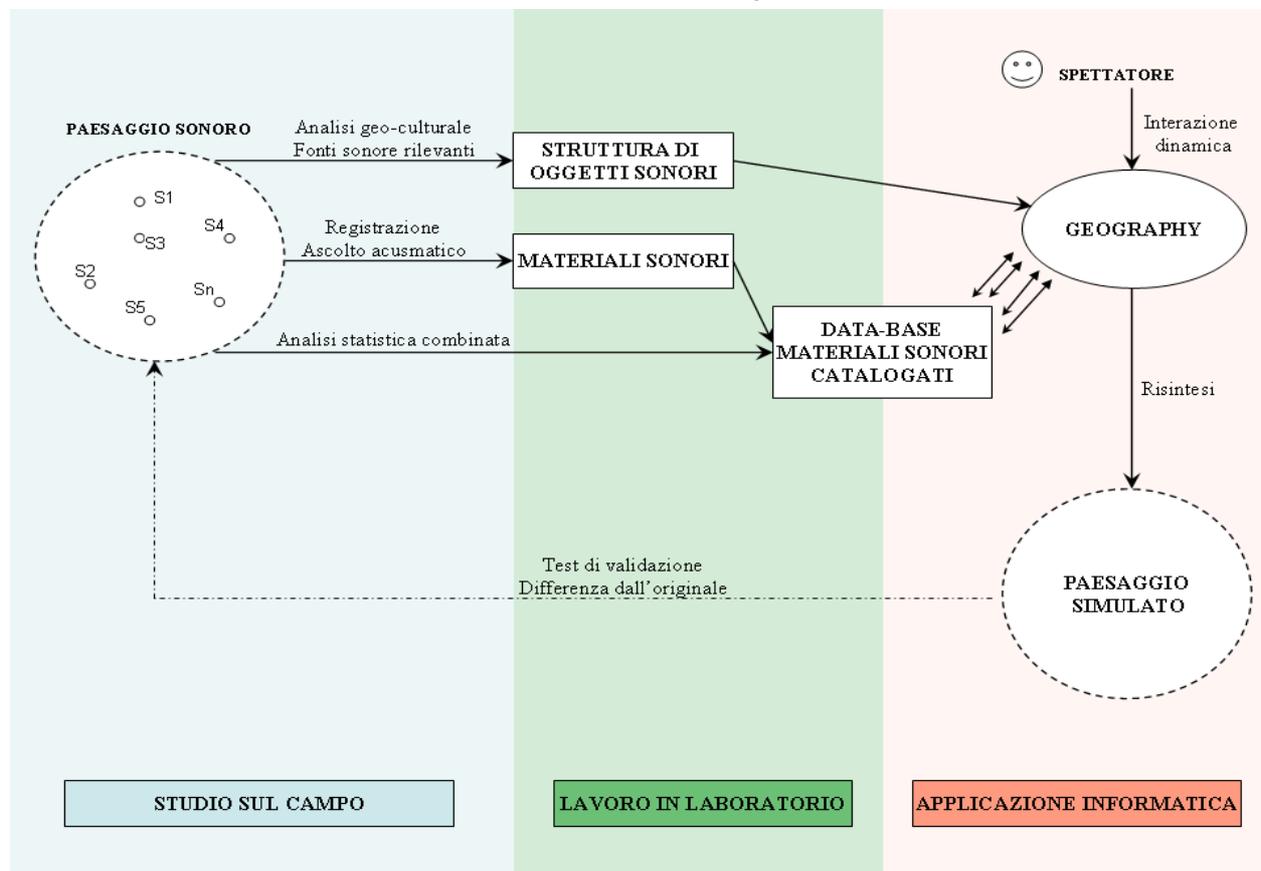


Figura 1. Schema complessivo del modello.

## IL MODELLO

Nonostante l'abbondanza degli usi del termine "soundscape", non risultano modelli e relative applicazioni il cui obiettivo sia la simulazione di un paesaggio sonoro a partire da una analisi/archiviazione dell'esistente. *Listen* ([25]) permette la generazione e il controllo di soundscapes interattivi per l'augmented reality, ma non prevede un modello del paesaggio sonoro. *Tapestrea* ([19]) consente la generazione in tempo reale di "suono ambientale", ma non definisce alcuna relazione tra suono e spazio.

Il modello proposto, che assume la definizione di "paesaggio sonoro" precedentemente discussa, prevede quattro componenti:

- Rappresentazione dello spazio il cui paesaggio sonoro è oggetto di studio. Essa deve includere una mappatura delle sorgenti e dei percorsi possibili per l'esplorazione
- Definizione e standardizzazione di una base dati audio di catalogazione dei componenti del paesaggio sonoro. Essa richiede:
  - registrazione ed eventuale elaborazione digitale dei campioni
  - standardizzazione della procedura di selezione oggetti sonori rilevanti
  - standardizzazione della procedura di generazione dei materiali sonori utilizzabili
- Rappresentazione del paesaggio sonoro attraverso i dati raccolti e definizione di un algoritmo generativo per la sintesi del paesaggio sonoro, a partire dai materiali sonori, dallo spazio e dal percorso dell'ascoltatore
- Definizione di criteri per la validazione qualitativa del risultato.

Innanzitutto il modello necessita di una rappresentazione dello spazio sonoro capace di simulare la percezione acustica del movimento di un ascoltatore rispetto alle fonti sonore. Nel modello le fonti sono semplificate: sono considerate immobili e localizzate puntualmente. In secondo luogo è necessario individuare parametri descrittivi degli oggetti sonori che, modificati dinamicamente, permettano di simulare la percezione della spazialità.

In sintesi (vedi Fig. 1), dato un paesaggio sonoro complesso, al fine di simulare un percorso interattivo esplorativo dello spazio, il modello proposto prevede: (A) una procedura di lavoro per la creazione del database di campioni audio che prevede due fasi: un'analisi geo-culturale del paesaggio sonoro al fine di ricavare la struttura modello degli oggetti sonori ed una fase di creazione e catalogazione dei materiali sonori registrati a partire dalle fonti reali esistenti (la procedura

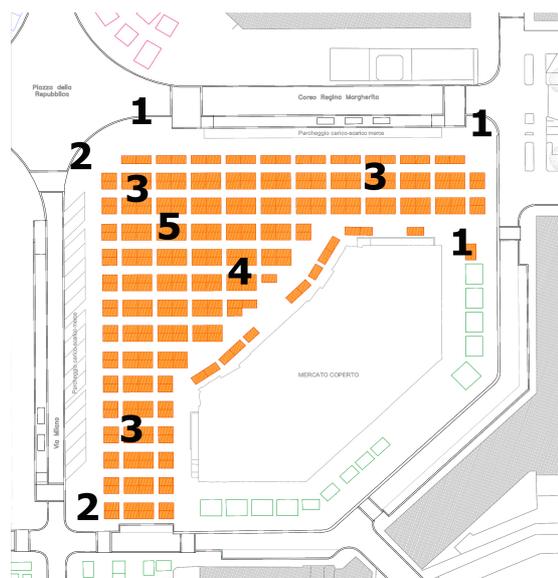


Figura 2. Pianta del mercato di Porta Palazzo e suddivisione in aree costituite da oggetti sonori omogenei

comprende i punti a. e b. precedenti); (B) un algoritmo capace di organizzare i materiali nello spazio e di sintetizzarli dinamicamente in relazione ad una trasformazione spaziale dovuta al percorso di un ascoltatore; (C) la definizione di test di validazione della simulazione del paesaggio sonoro.

I tre componenti del modello sono descritti più in dettaglio qui di seguito.

### (A) LA PROCEDURA DI CREAZIONE DELLA BASE DATI DI CAMPIONI AUDIO

La procedura si sviluppa in quattro fasi tra loro collegate e interdipendenti (Fig. 3).

La prima fase prevede l'esplorazione "distratta" del paesaggio: si richiede di concentrarsi percettivamente sulla dimensione passiva dell'"udire"([30]). In questa fase si individuano gli oggetti sonori più importanti del paesaggio nel suo complesso, quelli che si impongono all'attenzione anche dei frequentatori meno consapevoli. Vengono inoltre effettuate brevi interviste con diverse tipologie di fruitori del paesaggio in analisi, a proposito delle loro esperienze sonore.

Nella seconda fase, attraverso un ascolto attivo, si scompone lo spazio in aree per mappare i luoghi dove gli oggetti sonori vengono rilevati. La scomposizione viene effettuata al fine di individuare aree geografiche costituite da oggetti sonori

I FASE	<i>Esplorazione distratta</i>	Passeggiata esplorativa	► Individuazione oggetti sonori maggiormente rilevanti
II FASE	<i>Ascolto attivo, mappatura</i>	Registrazione sul campo	► Suddivisione in aree geografiche secondo la localizzazione degli oggetti sonori. Individuazione oggetti meno percettibili
III FASE	<i>Ascolto acusmatico</i>	In laboratorio	► Creazione materiali sonori scelti per rappresentazione indicale. Semplificazione degli oggetti raggruppati per similitudine fenomenologica nei materiali. Inizio rilevazione statistica
IV FASE	<i>Ascolti mirati</i>	Sul campo e in laboratorio	► Raffinamento materiali mancanti. Mappatura dei materiali nello spazio. Risultati dello studio statistico

Figura 3. Le fasi della procedura di creazione della base dati

omogenei (nel caso di un mercato, le aree della struttura geografica possono dipendere dalla tipologia delle bancarelle, dall'individuazione di aree di transito pedonale, di aree per automezzi, di aree di carico/scarico). Si inizia a costruire una mappa sonora (Fig. 2) con la divisione in aree e l'assegnazione degli oggetti sonori di riferimento, utile per l'implementazione spaziale dell'algoritmo.

Inoltre si effettua un primo studio approfondito del paesaggio sonoro al fine di individuare gli oggetti sonori meno percettibili e consueti. Tale fase di esplorazione viene realizzata durante le registrazioni sul campo, dove l'occhio ancora aiuta l'orecchio nella codifica e catalogazione degli oggetti sonori rilevanti, delle impronte sonore e dei segnali che descrivono ogni parte del paesaggio.

Dopo una prima esplorazione più osservativa, le informazioni vengono continuamente raffinate attraverso ulteriori e differenti stadi di analisi.

Nella terza fase si effettua l'analisi acustica delle registrazioni in laboratorio, e si inizia la creazione dei materiali sonori desunti dagli oggetti sonori modello. Le porzioni di segnale registrato che rappresentano all'ascolto indicale uno specifico oggetto sonoro creano un nuovo materiale sonoro. Oggetti sonori che presentano caratteristiche fenomenologiche molto simili vengono raggruppati in uno stesso materiale sonoro (se i suoni del cucchiaino, della forchetta, del coltello avevano permesso di individuare tre oggetti sonori specifici, l'ascolto ridotto può ricondurli ad un unico materiale per similarità delle caratteristiche sonore). Viene altresì effettuato uno studio del comportamento statistico dell'oggetto sonoro (quantità e modo di ripetizione). Lo studio statistico servirà poi in fase di implementazione dell'algoritmo.

La quarta fase è una fase di consolidamento delle informazioni: conduce alla creazione definitiva della mappa sonora.

L'annotazione del paesaggio sonoro avviene attraverso i seguenti parametri (vedi Allegato 1):

*Struttura geografica*: indica l'area che localizza la presenza di materiali sonori specifici. Permette la mappatura dei suoni nello spazio. Durante le registrazioni sul campo viene riferita agli oggetti sonori, a cui vengono fatti corrispondere durante le rilevazioni in laboratorio i materiali definitivi. Esempi di aree geografiche costituenti la struttura: bancarella frutta, bancarella formaggi, bancarella verdura, vie di transito...

*Struttura indicale*: descrive la dimensione figurativa evocata dall'oggetto sonoro, la scenografia sonora della narrazione, a quale evento la sensazione sonora rimanda. Vengono annotati suggerimenti durante la rilevazione sul campo, poi mediati e corretti in fase di analisi in laboratorio.

*Intensità*: indica l'intensità del suono espressa in termini di dinamica musicale, attraverso una scala di 5 unità (1: pianissimo, 5: fortissimo)

*Statistica*: indica il numero di accadimenti del materiale sonoro nell'unità di tempo

*Motivo rilevazione*: descrizione semantica del materiale sonoro. I valori possibili sono: tonica, segnale, impronta sonora, ascolto iconico, causale, simbolico. È prevista una descrizione aperta.

*Durata*: durata temporale del materiale sonoro

*Distanza registrazione*: parametro qualitativo che indica la distanza dalla quale è avvenuta la registrazione dell'oggetto sonoro poi catalogato in materiale di interesse.

*Ritmo*: descrizione della ripetitività dell'evento sonoro, inteso come susseguirsi di unità sonore costituenti un solo materiale sonoro (ad esempio, il battere il coltello per tagliare la carne è un materiale sonoro costituito da suoni analoghi ripetuti velocemente con un ritmo particolare), o connessioni con altri materiali (in prossimità di un acquisto il ritmo tra fruscio della sonora).

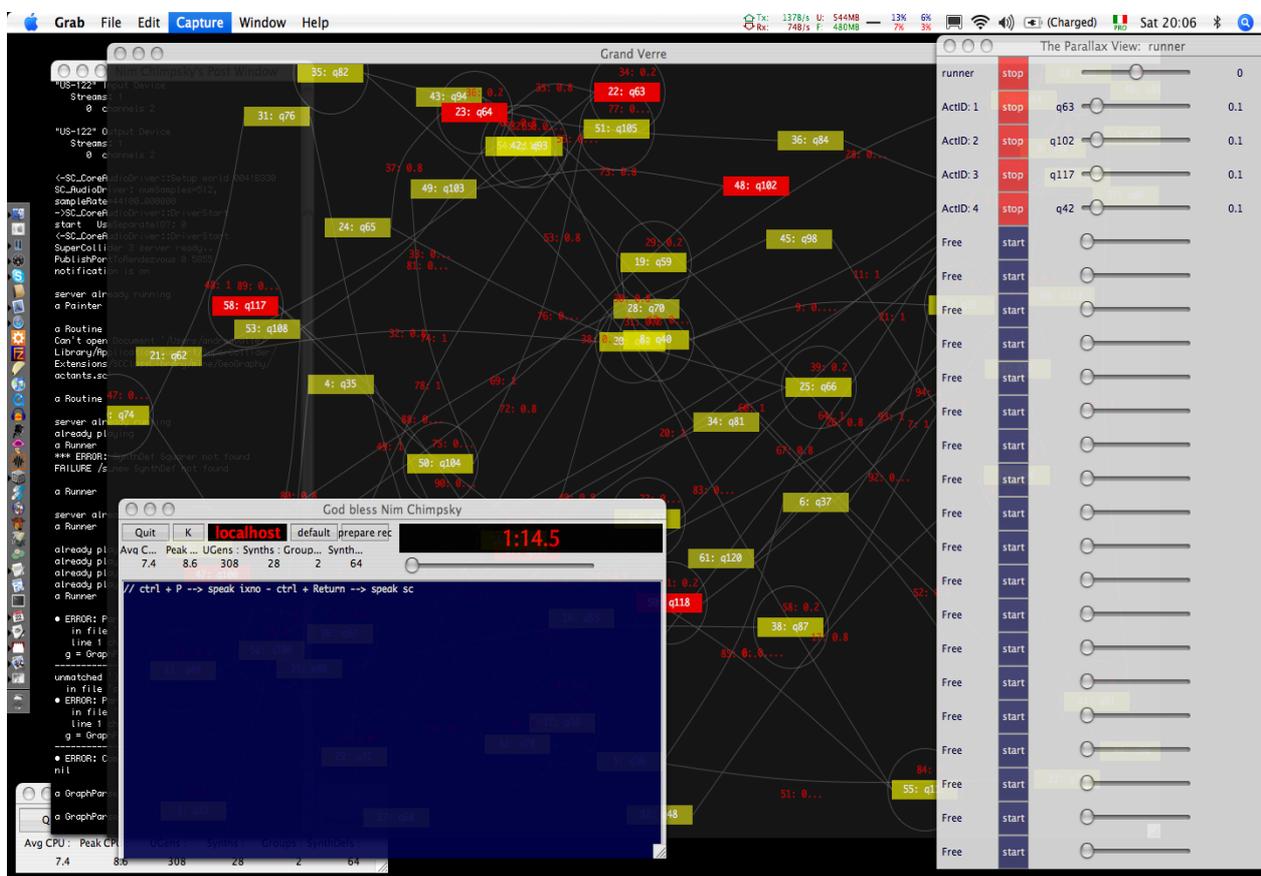


Figura 4. Esempio di struttura di grafi. Sulla destra sono elencati nell'interfaccia di controllo dell'applicazione 5 possibili attori di questa struttura

carta, della plastica e il tintinnio delle monete acquista una sua rilevanza informativa).

### (B) L'ALGORITMO GENERATIVO

Il modello proposto per generare un paesaggio sonoro di sintesi a partire dalle informazioni risultanti dall'annotazione di un paesaggio reale espande il sistema GeoGraphy ([26], [27]). GeoGraphy prevede due livelli.

Il primo organizza insiemi di materiali sonori attraverso una rappresentazione in grafi: i grafi descrivono possibili strutture di generazione di sequenze di materiali sonori. Tale struttura è descritta attraverso le connessioni e le relazioni tra vertici e archi, come illustrato in Figura 5. Ogni etichetta (1 e 2) rappresenta un vertice a cui è associato un materiale sonoro, gli archi (1, 2, 3, 4) connettono i vertici e illustrano l'intervallo temporale che separa l'attacco di un materiale sonoro dall'attacco del successivo. Il grafo è orientato e multiplo: gli archi rappresentano un possibile percorso da un vertice di partenza ad un vertice di arrivo, ogni vertice può avere più archi in entrata e in uscita. Nel grafo, un coppia (arco 2) indica che il materiale sonoro può essere ripetuto. Gli attanti (Fig. 4)

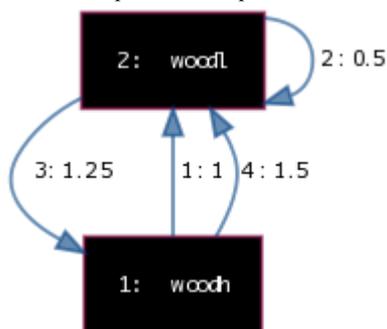


Figura 5. Vertici e Archi dei grafi

sono elementi dinamici dell' algoritmo che percorrono la struttura del grafo: ogni volta che un vertice è attivato da un attante, si ha generazione di un 'evento sonoro per la durata prevista dell'etichetta del vertice. Quindi l'attante seleziona un arco, attende un tempo pari all'etichetta dell'arco prescelto, attiva il vertice di arrivo, e via di seguito. Supponendo di inserire un attante a partire dal vertice 1, sono possibili diversi percorsi che equivalgono a diverse sequenze di materiali sonori. Un esempio potrebbe essere:

woodh - 1s di attesa > woodl - 0.5s > woodl - 1.25s > woodh - 1.5s > woodl etc...

La probabilità di scelta di ogni arco è uniforme.

Lo studio della statistica di generazione dei materiali sonori, avvenuta in fase di costruzione del data base del paesaggio in analisi, fornisce le informazioni necessarie alla definizione della struttura dei grafi. Ogni grafo rappresenta cioè una certa struttura di oggetti sonori e il suo comportamento. Le diverse topologie dei grafi permettono di descrivere comportamenti a differenti gradi di complessità delle sorgenti sonore (ad esempio, una specifica sintassi sonora del carico/scarico di un automezzo).

Nel secondo livello i vertici dei grafi ricevono una posizione in uno spazio. All'interno di tale spazio si muove un secondo elemento dinamico, un "Ascoltatore". L'Ascoltare può essere pensato come una funzione di filtraggio ed elaborazione delle sequenze generate dagli attanti. Ad ogni istante di tempo, l'algoritmo tiene conto dei suoni generati, dei suoni ancora non estinti, e delle loro aree di diffusione energetica. Calcola poi la posizione e l'orientamento dell'ascoltatore ed elimina tutte le componenti del paesaggio che risultano fuori dal raggio

percettivo. Seleziona, dunque, i materiali di interesse del dato istante (in Fig. 6 il vertice che risulta dentro l'area di percezione), calcola la traiettoria dell'ascoltatore tenendo traccia delle posizioni rilevate agli istanti di tempo t-1 e t+1 (Fig. 6), applica le trasformazioni previste dal modello per simulare la traiettoria del percorso. L'orientamento dell'Ascoltatore permette di calcolare la dislocazione di ogni vertice rispetto ad un asse definito, un parametro utile per la spazializzazione del suono.

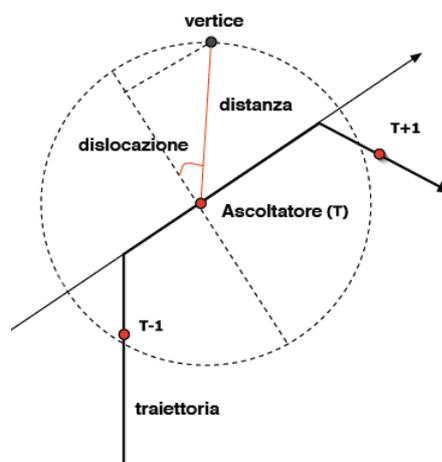


Figura 6. Le trasformazioni dovute al percorso

Dal secondo livello, infine, le sequenze generate dagli attanti e filtrate dalla funzione ascoltatore vengono spedite assieme alla lista dei parametri descrittivi ad un modulo esterno dell'algoritmo, dedicato all'implementazione audio, che si occupa della risintesi dei campioni della base dati. Tale modulo dedicato permette un'astrazione rispetto all'implementazione audio, lasciando la possibilità di modificare facilmente il modello attraverso il raffinamento degli algoritmi scelti per la sintesi sonora.

Il modello è implementato nel linguaggio di programmazione per l'audio SuperCollider ([29], cfr. Fig. 4). Estremamente espressivo ed efficiente, SuperCollider permette all'applicazione di operare interattivamente in tempo reale e di essere agevolmente inserita come componente audio in un'architettura di rete che comprende altre applicazioni. Un caso tipico è quello della sincronizzazione con un motore di realtà virtuale ai fini dell'esplorazione audio-visiva di un spazio urbanistico-architettonico.

### (C) I TEST DI VALIDAZIONE

La simulazione è validata attraverso test di ascolto e di orientamento all'interno dello spazio dinamico generato, ponendo quattro differenti tipologie di ascoltatori nella possibilità di esplorare il paesaggio sonoro. Si prevede l'installazione in uno spazio chiuso semibuio, dove l'ascoltatore dispone di un'interfaccia (tipicamente con joystick) per controllare il movimento a velocità costante all'interno dello spazio simulato. Una carta architettonica semplificata del paesaggio aiuta lo spettatore a individuarne i confini. Durante il test, il soggetto tenta di orientarsi all'interno del paesaggio attraverso l'ascolto, con una navigazione libera ma anche attraverso la realizzazione di tasks specifici ("raggiungi un punto preciso del paesaggio", "dove pensi di trovarti?", "cosa è avvenuto?", "quali differenze noti rispetto alla posizione precedente?"). Ai soggetti del test viene chiesto di commentare le loro esperienze sia durante la fruizione della

simulazione (archiviato tramite registrazione), sia in seconda analisi, tramite interviste.

Le tipologie di ascoltatori scelte sono:

- Frequentatori assidui (A)
- Frequentatori occasionali (O)
- Ascoltatori con orecchio educato (musicisti, sound designer) (E)
- Ascoltatori non sensibilizzati all'ascolto musicale o paesaggistico (N)

Inoltre, vengono effettuate interviste al fine di ottenere descrizioni qualitative/emozionali dell'esperienza esplorativa.

### CASE-STUDY: IL MERCATO DEL BALÒN A TORINO

Il modello presentato per la creazione di paesaggi sonori è costruito e studiato attraverso la simulazione del paesaggio sonoro del mercato storico del Balòn di Torino.



Figura 7. Panoramica del mercato del Balòn

Il mercato è un caso tipico di paesaggio sonoro socio-culturalmente importante. In particolare il Balòn ha una tradizione ultracentenaria, è il mercato all'aperto più grande in Europa, ed è l'espressione commerciale del retaggio culturale della città di Torino. Nei secoli ha conservato la sua tenace identità contraddistinta dalla storica ostinazione dei lavoratori della zona ad entrare nei processi decisionali di gestione dello spazio, una connotazione popolare autentica che la rende una delle zone della città di Torino dove sopravvive più che altrove la coesistenza di realtà sociali differenti e di culture diverse; ciò si manifesta anche nella sua disponibilità ad accogliere le professionalità straniere che confluiscono nel capoluogo Piemontese ([33]). La convivenza di mercanti provenienti da tutt'Italia e da gran parte del mediterraneo, dal Sud America, dall'Oriente e dall'Est Europa rende il Balòn una grande ricchezza acustica di voci, lingue, dialetti e idiomi. Presenta, inoltre, una forte complessità: molte sono le sorgenti sonore qualitativamente differenti, basti pensare che ogni giorno il mercato serve 20.000 persone, mentre il sabato raggiunge addirittura le 80.000, e necessita di 5.000 addetti per svolgere le mansioni commerciali, distributive, di pulizia e manutenzione. La sua frenetica attività si può desumere dall'estrema varietà del paesaggio sonoro ([33]).

È stato scelto il mercato, e in particolare la ricchezza compositiva del paesaggio sonoro del Balòn, perché è una caratteristica della società italiana, della sua cultura commerciale ed alimentare, e presenta sonorità assolutamente uniche, un patrimonio archeologico d'interesse culturale.

L'analisi del case-study (I fase della procedura di creazione della base di dati) è iniziata approfondendo la dimensione socio-culturale del luogo attraverso gli studi condotti dal comune di Torino per la riqualificazione del territorio, e con brevi interviste colloquiali ai lavoratori locali, ai clienti e alle figure di riferimento del mercato (i rappresentanti amministrativi delle diverse aree del mercato, venditori eletti autonomamente dai propri colleghi). Tali interviste venivano svolte durante le prime perlustrazioni "distratte" in cui si annotavano gli oggetti sonori più comuni: il suono della plastica, subito annotato come una tonica (inteso come segnale costante è onnipresente) del paesaggio sonoro, le voci e le grida dei mercanti, altra tonica di riferimento di intensità predominante in alcune aree del paesaggio, i rumori degli automezzi e dei trasporti pubblici che invadevano lo spazio dalle aree confinanti. Ad esempio, dalle testimonianze dei clienti è emerso come l'arrivo del Tram numero 4 facesse automaticamente pensare al mercato del Balòn, ed effettivamente tale fonte sonora permea tutto lo spazio con caratteristiche dinamiche e timbriche differenti a seconda della posizione dell'ascoltatore. Ci si è poi focalizzati (II fase della procedura) sui segnali sonori che permettevano di individuare i tipi di bancarelle. Si è così presentata la necessità di costruire una mappa sonora per localizzare ogni tipo di bancarella nello spazio mappandone gli oggetti sonori di riferimento. È stato possibile individuare le seguenti aree geografiche omogenee: bancarella della frutta (oggetti sonori: frutta dura mossa, frutta gettata sul banco, cocomeri sbattuti, bilancia elettronica semi-silenziosa, sequenza di acquisto: fruscio plastica, fruscio carta, tintinnio monete, apertura e chiusura cassa, voci...), bancarella della verdura (oggetti sonori: verdura lavata, spruzzino, bilancia in metallo rumorosa, ortaggi pesanti gettati sul banco, sequenza di acquisto, voci...), bancarella frutta e verdura (analoga alle precedenti), bancarella formaggi e salumi (più silenziosa e meno caratteristica, individuati particolari frusci, coltello che taglia, sbattito di prodotti pesanti, principalmente prosciutti), acciugai (suono della latta, acciughe sbattute ripetutamente su legno, olive buttate in olio...); sono poi state definite delle macro-aree che presentavano differenze di intensità delle voci e di attività sonora all'interno del mercato (Vedi Fig 2: 3 Attività di confine, 4 Attività combinata frutta verdura – formaggi acciugai, 5 Attività ad alta intensità) e delle aree di invadenza dei suoni meccanici e di motore (2 Area carico scarico camion, prossimità di strada con passaggio tram, 1 Area carico scarico piccoli automezzi e carrelli a mano, prossimità di strada con passaggio di qualsiasi tipo di automezzo). Infine sono stati raggruppati insieme gli oggetti sonori che descrivono aree di difficile mappatura, interni alle aree 3, 4, 5 e presenti in punti differenti, relativi cioè al transito pedonale, al transito dei carretti a mano, alla raccolta delle casse vuote (in legno, plastica e compensato), allo smontaggio/montaggio delle strutture delle bancarelle.

Durante la III fase della procedura sono state scelte le migliori registrazioni sul campo (i materiali sonori) descrittive degli oggetti sonori individuati. Tipicamente, si è preferito disporre di diversi materiali per ogni oggetto, in modo da non disperdere la complessità delle fonti reali. La fase ha incluso una rilevazione statistica del comportamento dei materiali creati e la mappatura degli stessi nella mappa sonora implementata poi nell'algoritmo di sintesi. Durante la IV fase è stata approfondita l'analisi delle aree geografiche: l'ascolto acusmatico delle registrazioni ha permesso l'individuazione di materiali sonori inaspettati e nuovi; ad essa si è aggiunta un'indagine sul campo per la loro catalogazione semantica. Ad esempio, la differenza costruttiva delle casse di distribuzione non era stata notata in prima analisi, ma l'ascolto acusmatico ha rilevato materiali sonori sconosciuti, che l'analisi

approfondita della IV fase ha permesso di attribuire alla fonte sonora originaria.

### STATO DI AVANZAMENTO DEL PROGETTO

Al momento presente, il progetto di ricerca è corso di realizzazione: la prima fase di definizione e standardizzazione delle procedure che costituiscono il modello è completa, i materiali audio sono stati raccolti e organizzati, l'applicazione è in stato avanzato di implementazione, i test qualitativi sono definiti ma non sono stati ancora realizzati

### RIFERIMENTI BIBLIOGRAFICI

- [1] Murray Schafer R., *The Tuning of the World*. McClelland & Steward and Knopf, Toronto and New York, 1977.
- [2] Matthew Burtner, *Ecoacoustic and shamanic technologies for multimedia composition and performance*, Organised Sound, 10(1):3-19, 2005.
- [3] Fontana F., Rocchesso D. and Ottaviani L., *A structural approach to distance rendering in personal auditory displays*, In Proceedings of the International Conference on Multimodal Interfaces (ICMI 2002), 2002.
- [4] McGregor I., Crerar A., Benyon D. and Macaulay C., *Sounfields and soundscapes: Reifying auditory communities*, In Proceedings of the 2002 International Conference on Auditory Display, 2002.
- [5] Rubin B. U., *Audible information design in the new york city subway system: A case study*, In Proceedings of the International Conference on Auditory Display '99, 1998.
- [6] Droumeva M. and Wakkary R., *The role of participatory workshops in investigating narrative and sound ecologies in the design of an ambient intelligence audio display*, In Proceedings of the 12 th International Conference on Auditory Display, 2006.
- [7] Friberg J. and Gärdenfors D., *Audio games: New perspectives on game audio*, In Proceedings of the 2004 ACM SIGCHI International Conference on Advances in computer entertainment technology, pages 148-154, New York, 2004, ACM Press.
- [8] Serafin S., *Sound design to enhance presence in photorealistic virtual reality*, In Proceedings of the 2004 International Conference on Auditory Display, 2004.
- [9] Hermann T., Meinicke P. and Ritter H., *Principal curve sonification*, In Proceedings of International Conference on Auditory Display, 2000.
- [10] Butz A. and Jung R., *Seamless user notification in ambient soundscapes*, In IUI '05: Proceedings of the 10th international conference on Intelligent user interfaces, 2005, ACM Press.
- [11] Kilander F. and Lönnqvist P., *A whisper in the woods - an ambient soundscape for peripheral awareness of remote processes*, In Proceedings of the 2002 International Conference on Auditory Display, 2002.
- [12] VV.AA., *The tech issue ...to be continued*. Soundscape, 3(1), 2002.
- [13] VV.AA., *Acoustic design*, Soundscape, 5(1), 2004.
- [14] Hildegard Westerkamp, *Linking soundscape composition and acoustic ecology*, Organised Sound, 7(1), 2002.
- [15] Mauney B. S. and Walker B. N., *Designing systems for the creation and evaluation of dynamic peripheral soundscapes: a usability study*. In Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting-2004, 2004.
- [16] Mayr A., editor, *Musica e suoni dell'ambiente*, CLUEB, Bologna, 2001.
- [17] Colimberti A., editor, *Ecologia della musica: Saggi sul paesaggio sonoro*, Donzelli, Roma, 2004.
- [18] Truax B., *Acoustic Communication*, Greenwood, Westport, CT, 1984.
- [19] Misra A., Cook P. R. and Wang G., *Musical Tapestry: Re-composing Natural Sounds*, In Proceedings of the International Computer Music Conference (ICMC), 2006.
- [20] Augoyard J.-F. and Torgue H., *Repertorio degli effetti sonori*, Lim, Lucca, 2003.
- [21] Chion M., *L'audiovisione*, Lindau, Torino, 1999.
- [22] Schaeffer P., *Traité des objets musicaux*, Seuil, Paris, 1st edition, 1966.
- [23] Lombardo V. and Valle A., *Audio e multimedia*, Apogeo, Milano, 3a edition, 2008.
- [24] Bregman A., *Auditory Scene Analysis*, The MIT Press, Cambridge, Mass. and London, 1990.
- [25] O. Warusfel and G. Eckel, *LISTEN-Augmenting everyday environments through interactive soundscapes*, Virtual Reality for Public Consumption, IEEE Virtual Reality 2004 Workshop, Chicago IL, 27, 2004.
- [26] Valle A. and Lombardo V., *A two-level method to control granular synthesis*. In XIV CIM 2003 Computer Music: Past and Future, Proceedings, pages 136-140, Firenze, 2003.
- [27] Valle A., *Geography: a real-time, graph-based composition environment*. In NIME 2008: Proceedings, pages 253-256, 2008.
- [28] Rocchesso D. and Fontana F., editors, *The Sounding Object*, Edizioni di Mondo Estremo, Firenze, 2003.
- [29] Wilson S., Cottle D. and Collins N., editors, *The SuperCollider Book*, The MIT Press, Cambridge, Mass., 2008.
- [30] Agostini L., *Creare Paesaggi Sonori*, Edizioni Lulu.com, 2007
- [31] G. Böhme, "Atmosfere acustiche. Un contributo all'estetica ecologica", in [17].
- [32] B. Truax, "Model and strategies for acoustic design", in: H. Karlsson, editor, "Hör upp! Stockholm, Hey Listen! - Papers presented at the conference", The Royal Swedish Academy of Music, Stockholm, 1998
- [33] Studio di Ingegneria ed Urbanistica Vittorio Cappato, editor, *50 Centesimi al Kilo: La riqualificazione del mercato di Porta Palazzo dal progetto al cantiere*, Pubblicazione città di Torino, Torino, 2006
- [34] Walter Murch, *In un batter d'occhi: prospettiva sul montaggio cinematografico nell'era digitale*, Trad., Edizioni Lindau, Torino, 2001
- [35] P. Amphoux, *L'identità sonora des villes européennes, Guide méthodologique à l'usage des gestionnaires de la ville, des techniciens du son et des chercheurs en sciences sociales*, publication IREC, EPFL, Lausanne, e CRESSON, Ecole d'Architecture de Grenoble, 1993.
- [36] Acirene atelier de traitement culturel et estétique de l'environnement sonore, *Résumé de l'étude de conception et d'aménagement du paysage sonore du secteur de la sucrerie - st. cosme*, documento fornito sul sito dell'atelier per una chiamata a progetto di liuteria urbana e pianificazione territoriale, www.acirene.com, 2007.

## Scheda rilevazione materiali sonori Mercato Frutta e Verdura

Nome	Sotto-Strut Geo	Strut indicale	Area Temp	Intensità	Stocastica (volte/ora)	Motivo Rilevazione	Time code start	Durata (s) o stop	Dist ripresa	Ritmo	Note
CamioncinDis tribuzione1	Area confine 1	La meccanizzazione del mercato,distribuzione	Piena attività	F nell'area di appartenenza P nelle aree vicine	x	Segnale Presenza di pericolo, necessità di fare attenzione, arrivo rifornimenti		17,22	distante		Registrazione p per le aree vicine
FrenataBus1	Rilevata Area confine 1 (Da Strada)	Il passaggio, transito e mobilità della città	Piena attività	FF nell'area di appartenenza F nelle aree vicine	x	Segnale, icona della mobilità cittadina e del traffico del mercato		4,73	distante		Registrazione p per le aree vicine
CaricoCassette Fruttal,egno	Rilevata area mercato 1 vicino confine	distribuzione	Piena attività	F nell'area M in vicinanza	xxx	Segnale cass piena in camioncino attività del camioncino di distribuzione		1,7	Media distanza (5 m circa)		Un po troppe voci di fondo meglio trovare un'altro campione
Cassal,egnoget ataAterra	Area mercato angolo 1	Attività bancarella cassa vuota per terra	Piena attività	F nell'area M in vicinanza	xx	Segnale, attività bancarella, allattamento continuo		1,76	Media distanza		
GridaMercanti2	Area mercato angolo 1	Richiamo aquitenti, mercantaggio	Piena attività	FF nell'area Tonica in vicinanza	Tonica	Tonica		3	Vicino		Voce singola
GridaMercanti4	Area mercato angolo 1	A pproccio mercantaggio	Piena attività	F nell'area, pp vicino				1,2			Voce di accoglienza nella bancarella
MoneteInCassa 1	Area mercato angolo 1	Incasso, spesa, mercantaggio	Piena attività	M nell'area pp vicino	xxxxxx	Segnale, icona del commercio, simbolo dello scambio avvenuto		1,58	Vicino		
FruscioPlastica 1	Tutto il mercato	Compravendita, mercantaggio, prodotti messi in busta	Piena attività	P nell'aera Niente Vicino	xxxxxxxx	Imponta sonora della spesa moderna, icona del trasporto di alimenti		2,23	Vicino	C - c - p - c - c - p - p - c - c - p	In prossimità di uno acquisto, il ritmo tra fruscio carta e plastica è info rilevante

## Allegato 1: Scheda rilevazione dei materiali sonori

# Del suonare l'elettronica

## Gabriel Bacalov

Master MASP, SDA Luigi Bocconi  
Milano

## Massimo Marchi

AGON acustica informatica musica  
Milano

## Francesco Vitale

AGON acustica informatica musica  
Milano

### SOMMARIO

In questo articolo affronteremo le tematiche relative alla sostenibilità delle opere di *live electronics*, proponendo una visione unitaria di aspetti tradizionalmente affrontati separatamente e in ambiti autonomi, ovvero quello delle tecnologie applicate, quello della musica e quello dell'economia dello spettacolo.

### Parole Chiave

*Live electronics*; economia dello spettacolo.

### INTRODUZIONE

Negli ultimi anni si è sempre più diffusa nell'ambito della musica elettroacustica la consapevolezza – che AGON, in virtù della sua quasi ventennale esperienza nella produzione, nella ricerca e nell'organizzazione nel campo delle *performing arts* e delle nuove tecnologie, promuove attivamente dall'inizio di questo decennio – della necessità di affrontare il problema della conservazione e della riproducibilità delle opere musicali che impiegano il mezzo elettronico, le quali peraltro costituiscono una parte assai consistente del repertorio musicale del secondo Novecento.

Il vertiginoso progresso tecnologico degli ultimi decenni ha infatti rapidamente relegato nell'obsolescenza gran parte delle macchine impiegate nelle produzioni musicali contemporanee, le quali a breve termine potrebbero trovarsi a non essere più eseguibili per due principali motivi:

- irreperibilità degli strumenti elettronici originali;
- impossibilità di sopperire alla mancanza di questi, dovuta all'impiego di tecnologie software e hardware proprietarie che non consentono la migrazione ad altri sistemi.

Già oggi, del resto, uno dei maggiori ostacoli all'effettiva riproposizione di opere che prevedono l'utilizzo di strumenti elettronici è costituito dalla difficile reperibilità di questi, ormai divenuti rari se non introvabili.

Alle difficoltà già evidenziate si aggiungono ostacoli di natura più eminentemente economica, che spesso vengono trascurati nell'analisi delle problematiche che rendono tanto difficoltosa l'esecuzione di opere elettroniche.

### ASPETTI METODOLOGICI

Il dibattito su come affrontare questa drammatica situazione è più che mai vivo e attivo in tutta la comunità mondiale dei musicisti informatici e, al di là delle sottili differenze di impostazione, essa sembra riconoscere pressoché unanimemente come unica soluzione al problema della sostenibilità la riproposizione in un formato aperto [5] (anche se le caratteristiche specifiche di tale formato sono ancora oggetto di discussioni e proposte) delle opere concepite in formato proprietario.

Come affermato da Bernardini e Vidolin [2], è evidente che questa rappresentazione dell'opera "*dovrà essere in grado di sopravvivere all'invecchiamento e alle rivoluzioni tecnologiche*", per evitare che il problema sia semplicemente

differito e torni a proporsi in un secondo momento. La loro proposta per un "*glossario multimediale che copra ogni processo elettroacustico dal vivo*" ci appare la più completa e ragionevole e richiede, per ogni elemento che intervenga nel brano:

- una descrizione algoritmica del processo di generazione/elaborazione del suono;
- una risposta all'impulso del sistema;
- un esempio audio.

Condividiamo inoltre la scelta del paradigma *orchestra & score* da loro proposto (senza dilungarci in questa sede sugli aspetti peculiari dell'auspicabile formato), in quanto, oltre ai vantaggi già esposti nell'articolo citato (separazione tra dati sul sistema e dati sulla performance e riduzione della complessità durante l'esecuzione), esso limita gli interventi necessari sulla partitura originale.

Non prendiamo in considerazione qui il caso di opere future, fermo restando che queste dovrebbero già in partenza adottare questo paradigma, concentrandoci invece sull'attività di ricostruzione di opere già esistenti che adottino uno o più formati software o hardware proprietari.

Per quanto riguarda le metodologie e le procedure da seguire in questo genere di operazione proviamo a tracciare qui uno schema, il quale mostra notevoli affinità con quello a tre stadi proposto da Brooke Wetzel [3]; tale schema tuttavia non ci pare del tutto esaustivo, poiché lascia aperti alcuni dubbi, che noi cercheremo di chiarire più avanti, sulle competenze necessarie alla buona riuscita del lavoro:

- la prima fase prevede un lavoro analitico, per l'estrazione dei processi algoritmici e dei parametri oggettivi;
- la seconda fase prevede la ricostruzione del sistema, a partire dalla precedente, con le tecnologie disponibili, allo scopo di testare il modello estrapolato;
- la terza, infine, richiede la sperimentazione pratica del modello e della ricostruzione, nel corso di prove e di performance vere e proprie.

Ovviamente è previsto un feedback tra le varie fasi del lavoro, allo scopo di affinare sempre maggiormente il modello. L'elaborazione dei dati raccolti sarà infine volta ad una proposizione di questi che raggiunga il massimo livello di astrazione, vale a dire il modello algoritmico di generazione/elaborazione del suono e le caratteristiche fisiche di questo, espresse in unità di misura standard. Da tale tipo di descrizione sarà quindi possibile enucleare di caso in caso le specifiche di alto livello per l'implementazione del sistema in qualunque ambiente di programmazione per la musica esistente o futuro.

Forse il principale ostacolo alla realizzazione di questo lavoro si trova proprio al primo stadio di lavorazione, e consiste nel reperimento del materiale originale, innanzitutto le macchine sulle quali eseguire i test per la modellazione che dovrebbero essere, per garantire una buona riuscita, in perfette condizioni. Non è da trascurare, in questa delicatissima fase, un'attività di documentazione quanto più rigorosa possibile che porti a

raccogliere tutti i materiali disponibili sull'opera, che Brooke Wetzel nel suo articolo suddivide in:

- partitura e annotazioni tecniche;
- ambiente esecutivo originale;
- commenti posteriori alla composizione;
- registrazioni;
- modifiche apportate dal compositore.

### CASE STUDY: *LE TEMPS ET L'ÉCUME* DI GÉRARD GRISEY

Lo scorso anno, nel prepararci al decennale della scomparsa di Gérard Grisey, abbiamo deciso di investire delle risorse nello studio dei pochi brani in cui il compositore francese aveva deciso di utilizzare il mezzo elettronico, nell'ottica di sensibilizzare gli enti di programmazione per inserire questi pezzi nelle loro stagioni. Ciò s'immette in un progetto più ampio che AGON da tempo persegue: lo studio sulla reale possibilità di eseguire i pezzi di repertorio con elettronica. Fatto che sembrerebbe ovvio ma che molto spesso è fonte di insospettabili e, talvolta, insormontabili problemi; non solo legati all'esecuzione musicale in sé e in particolare alla peculiarità e alla scarsa reperibilità degli strumenti, ma anche, come si vedrà in seguito, ad aspetti tradizionalmente trascurati dagli addetti ai lavori e delegati ad altri soggetti, come l'organizzazione e l'economia dello spettacolo.

I cinque pezzi con elettronica di Grisey che abbiamo preso in esame sono:

- *Prologue (1976/85)*, per viola e live electronics;
- *Jour, contre-jour (1978/79)*, per organo elettronico, 13 musicisti e nastro magnetico;
- *Les Chants de l'Amour (1982/84)*, per 12 voci miste e voci sintetizzate al computer;
- *Le temps et l'écume (1988/89)*, per 4 percussionisti, 2 sintetizzatori e orchestra da camera;
- *Le Noir de l'étoile (1989/90)*, per 6 percussionisti, nastro magnetico e ritrasmissione in situ di segnali astronomici;

Per alcuni di questi il problema è relativo:

- per *Les Chants de l'Amour* esiste un nastro ADAT quadrifonico;
- per *Le Noir de l'étoile* abbiamo un nastro ADAT di sincronizzazione per i percussionisti ed un DAT per le pulsar;
- per *Prologue* è stato eseguito un lavoro di rifacimento da parte di Eric Daubresse.

Le maggiori criticità, dal punto di vista della reperibilità degli strumenti elettronici, si sono riscontrate in *Jour, contre-jour* e *Le temps et l'écume* [4]. Ed è su questo secondo pezzo, anche grazie ad una richiesta pervenuta da Casa Ricordi, che si è concentrata la nostra attenzione: il brano, per 4 percussionisti, 2 sintetizzatori e orchestra da camera, prevede l'impiego di due Yamaha DX-7, con due assistenti per la gestione dei parametri durante l'esecuzione.

L'archivio Ricordi, che disponeva soltanto delle cartucce master su cui sono salvati i preset dei timbri per i due sintetizzatori, ci ha richiesto una duplicazione di queste per non dover noleggiare gli originali; il che non è stato possibile – ci siamo dovuti limitare ad eseguire un backup dei dati di sistema esclusivo su CD-ROM – essendo ormai i supporti del DX-7 fuori produzione e quindi di assai difficile reperibilità.

L'occasione, colta al balzo anche in virtù dell'enorme popolarità del modello di sintetizzatore in questione, ci ha portati a una prima riflessione sulla possibilità di rendere l'esecuzione di *Le temps et l'écume* – e, potenzialmente, quella di ogni pezzo che richieda l'impiego di strumenti elettronici o informatici – svincolata dalla necessità di utilizzare un particolare hardware, di modo che in futuro (o anche adesso) il brano possa essere eseguito al di là della disponibilità della macchina sulla quale e per la quale è stato scritto.

Riassumendo il lavoro da noi svolto, possiamo articolarlo nelle

seguenti fasi:

- *reperimento e lettura dei dati sui supporti originali*: il materiale fornitoci da Casa Ricordi consisteva in tre cartucce per DX-7 e un floppy disk per SY-77, synth che l'autore aveva previsto come alternativa in assenza del primo;
- *confronto dei dati pervenuti su diverse tipologie di supporto*: nel nostro caso si trattava di poche cartucce dal contenuto pressoché identico;
- *analisi comparativa dei dati e della partitura*: nel nostro caso disponiamo dell'edizione originale con annotazioni dell'autore. Tale analisi è appunto finalizzata a verificare il contributo delle macchine al brano;
- *documentazione e analisi della macchina originale*, allo scopo di ottenere un modello algoritmico quanto più dettagliato del funzionamento interno del DX-7;
- *test dei valori dei singoli parametri sulla macchina originale*: abbiamo utilizzato un TX-816 in nostro possesso, modulo expander che implementa la medesima tecnologia del DX-7;
- *ricostruzione delle specifiche*: operazione che si compie sostituendo nel modello algoritmico i parametri fisici a quelli proprietari;
- *implementazione del modello a scopo di verifica*: con Max/MSP abbiamo realizzato l'emulazione di alcune voci previste nel brano, basandoci sui dati da noi raccolti, con risultati piuttosto soddisfacenti dal punto di vista della resa acustica;
- *realizzazione finale di un'edizione critica della partitura*, con le annotazioni per la parte elettronica, comprendente il modello per la sintesi e tutti i parametri in unità di misura standard.

Attendiamo ora l'occasione per testare "sul campo", ovvero con un'esecuzione, il nostro lavoro.

### COMPETENZE

Un lavoro di questo genere, che a prima vista potrebbe essere assimilabile alla progettazione di emulazioni software di strumenti elettronici, necessita invece di competenze che vanno ben al di là di quelle unicamente ingegneristiche e informatiche richieste a un progettista di software audio commerciali.

A queste, sicuramente necessarie, è auspicabile che siano affiancate conoscenze non trascurabili in almeno altri due campi applicativi:

- *musicale*, per affrontare consapevolmente il brano non solo dal punto di vista dell'elettronica, ma con una più completa visione d'insieme, senza trascurare aspetti che sfuggirebbero ad un tecnico;
- *musicologico/filologico*, per strutturare in maniera critica sia le fasi di ricerca preliminare dei materiali, sia la successiva elaborazione e proposizione dei dati raccolti; la ricerca e il confronto delle fonti (partiture, note tecniche, equipaggiamento, commenti, registrazioni ecc...) presenta notevoli analogie con i compiti svolti dal tipico filologo musicale.

I materiali sui quali si opera sono infatti fondamentalmente classificabili in tre generi:

- *la partitura e documentazione*, nella loro accezione più estesa, ovvero l'insieme di tutte le prescrizioni operative per la corretta esecuzione del brano, annotazioni tecniche, note di esecuzione, commenti, testimonianze degli interpreti originali e del compositore stesso;
- *l'ambiente esecutivo*, ovvero tutto l'insieme di caratteristiche fisiche e ingegneristiche delle macchine impiegate, unite a tutte le modalità di interazione con queste;
- *i materiali sonori e i dati di generazione/elaborazione*, ovvero i campioni, i nastri, i parametri di elaborazione o

sintesi del suono che determinano nello specifico il contributo dello strumento elettronico al brano.

Solo con un'analisi comparativa di queste tre tipologie di materiali di partenza si potrà ottenere, con la massima efficienza, un risultato di buona qualità; e ci sembra che solo una figura che possieda l'insieme di competenze tecniche, musicali e musicologiche sopra citate possa fornire le dovute garanzie di serietà e professionalità.

### ASPETTI ORGANIZZATIVI ED ECONOMICI

Oltre all'analisi di tipo informatico-musicologico, la realizzazione di questo tipo di lavoro richiede una preparazione manageriale completa [6], fondata su competenze economiche, tecniche e culturali. Inoltre, da parte degli organizzatori, sembrerebbe essere di notevole rilevanza il possesso di una consapevolezza, di una sensibilità relativa al ruolo storico, sociale ed ai vantaggi economici correlati al mezzo tecnologico applicato alla musica ed in generale all'arte.

L'esigenza di acquisire un bagaglio di competenze miste – da un lato elementi umanistici, dall'altro economico-finanziari, organizzativi e tecnici – sembra essere ormai un fatto accettato praticamente da tutti gli operatori del settore della musica contemporanea, mentre si presenta ancora poco sviluppata la sensibilità nei confronti del secondo punto.

Emerge una quadro globale in cui i manager culturali sembrano non volere accettare l'enorme valore della musica elettronica nell'ambito della musica contemporanea con organico tradizionale. Infatti, nel contesto del secondo novecento, la programmazione di brani di musica elettronica appare sempre e comunque minoritaria rispetto a brani scritti esclusivamente per strumenti acustici. I risultati della nostra analisi sulla composizione della programmazione di vari Festival di musica contemporanea in Europa confermano quanto detto: nel Vecchio Continente<sup>1</sup> lo spazio dedicato alla musica elettronica è solo del 17,8% (vedi tabella).

Festival	Brani con elettronica	Brani totali	Percentuale di brani con elettronica
Manca Nizza, 2007 e 2008	22	60	36,67%
Milano Musica Milano, 2007 e 2008	18	117	15,38%
Biennale di Venezia Venezia, 2008	10	113	8,85%
New Music Festival Stuttgart, 2007 e 2008	6	51	11,76%
Musikfest Berlin, 2008	0	48	0,00%
New Music Festival "Arena" Riga, 2008	11	47	23,40%
Huddersfield Contemporary Music Festival Huddersfield, 2007	16	31	51,61%
<b>Totale (media)</b>	<b>83</b>	<b>467</b>	<b>17,77%</b>

Tabella 1

Questo squilibrio nella programmazione appare perlomeno paradossale se analizzato prendendo in considerazione il peso storico, estetico e talvolta quantitativo dell'elettronica negli ultimi 50 anni della storia della musica.

<sup>1</sup> Un'eventuale analisi del medesimo tipo per gli Stati Uniti non sarebbe stata rappresentativa in quanto in Nord America viene generalmente prediletta una formula che prevede la separazione della musica contemporanea strumentale da quella elettronica in Festival di distinti.

Oltretutto se si pensa a quanto la tecnologia abbia pervaso praticamente tutti gli aspetti della vita dell'uomo, diventandone – anche negativamente – elemento essenziale, appare quantomeno bizzarro che si investa così poco in programmazioni che tengano conto di un mezzo che è diventato simbolo e ritratto della civiltà occidentale del Ventesimo Secolo.

Tralasciando il discorso relativo alla mancanza di consapevolezza su ciò che la musica elettronica rappresenta per la nostra civiltà ed escludendo motivazioni estetiche prettamente legate al gusto personale di chi organizza manifestazioni musicali, possiamo ragionevolmente ipotizzare che vi sia un'ulteriore spiegazione alla base di questa carenza di programmazione: presumiamo che esista una tendenza generale che racchiude un timore, un'apprensione nei confronti dell'incremento dei costi nelle produzioni che utilizzano le nuove tecnologie.

Spesso gli organizzatori dimostrano una preoccupazione nei confronti dei costi relativi all'allungamento dei tempi delle prove, del sound-check, del montaggio e dello smontaggio ed al reperimento del materiale tecnico-musicale.

Oltretutto, analizzando il quadro da una prospettiva globale, questa riluttanza organizzativa si manifesta sia nel sistema culturale anglosassone basato sull'intervento privato – dalle grandi aziende ai singoli spettatori – che nel sistema europeo in cui è lo Stato a sostenere l'arte, tuttavia con risultati anche molto diversi a seconda dei Paesi – si pensi al sostegno che ricevono centri come l'IRCAM in Francia (15,26 milioni di Euro), STEIM nei Paesi Bassi (580.000 Euro) o il Centro Tempo Reale in Italia (circa 320.000 Euro)<sup>2</sup>.

In entrambi i sistemi vi è la necessità di un intervento esterno per mantenere positivo il rapporto tra le varie voci di costo ed i ricavi. Il sostegno, privato o pubblico, è indispensabile in quanto il settore artistico, definito a produttività stagnante – suddiviso in beni, attività ed in parte industrie culturali – è soggetto alla cosiddetta malattia dei costi di Baumol e Bowen [1], ovvero ad un incontrollato aumento del costo del lavoro in funzione della produttività.

Questo è senza dubbio uno dei motivi principali che rende apparentemente svantaggioso il settore culturale da un punto di vista meramente economico.

Anche in questo tipo di ottica il caso del brano da noi esaminato si presenta paradossale. Questa volta però, in senso positivo.

Se analizzato dal punto di vista dell'efficienza, – il raggiungimento di determinati risultati utilizzando il minor numero possibile di risorse – il timore nei confronti dei costi potrebbe essere soppresso o almeno eluso. In primis, il lavoro svolto su *Le temps et l'écume* di Gérard Grisey dimostra che esiste una possibilità concreta per abbattere i costi: in questo caso, il lavoro di ricerca tecnica, musicologica e filologica permetterà a chi volesse eseguire il brano un risparmio relativo alle risorse umane (2 esecutori e 1 addetto alla regia del suono e alla gestione dei parametri di esecuzione invece di 2 esecutori, 2 assistenti e 1 addetto alla regia del suono), oltre al già citato abbattimento dei costi relativi al reperimento e al noleggio della strumentazione elettronica.

Le nuove tecnologie informatiche rappresentano quindi un concreto mezzo per riportare alla luce numerosi brani musicali ormai abbandonati perché poco economici e tecnicamente complessi.

Allargando la dissertazione a tutte le *performing arts*, ci sembra opportuno proporre una riflessione concernente i vantaggi offerti dall'informatica: senza volere assolutamente mettere in discussione la bellezza, la grandiosità e l'importanza delle arti del passato, possiamo – sempre e

<sup>2</sup> Ultimi dati disponibili (IRCAM e STEIM 2006, Tempo Reale 2003).

comunque nell'ottica del risparmio - affermare che le potenzialità dell'informatica sono innumerevoli e offrono una concreta occasione per dar vita ad opere artistiche ampiamente sostenibili dal punto di vista economico.

## CONCLUSIONI

Per quanto riguarda le problematiche di carattere tecnico-musicale, specie quelle relative alla sostenibilità, riteniamo che il dibattito in corso vada nella giusta direzione, che sia necessario proseguire in tal senso in maniera sistematica e con interventi consistenti sul repertorio.

Per concludere, desideriamo rilevare un ulteriore presunto problema del "sistema musica elettronica" e in generale del "sistema spettacolo".

Si è detto che i Festival di contemporanea privilegiano la musica strumentale acustica e che le possibili cause potrebbero annidarsi in lacune socio-culturali e preoccupazioni di tipo economico da parte degli addetti ai lavori.

Oltre a questi due ipotesi, vogliamo proporre un ulteriore spunto di riflessione:

- è sensato pensare che esista un profondo divario culturale, estetico e linguistico tra il grande pubblico e le opere elettroniche oppure tra il grande pubblico e la musica contemporanea in generale?
- se le istituzioni teoricamente responsabili dell'educazione dei cittadini non prevedono nemmeno un insegnamento musicale sufficiente, è possibile che siano gli stessi operatori culturali a sopperire, nel proprio interesse, a questa carenza?
- di conseguenza, è possibile che un maggiore impegno degli organizzatori musicali nell'ambito della promozione del pubblico e in un utilizzo più convinto degli strumenti del marketing culturale possa portare, in una sorta di circolo virtuoso, ad una crescita di interesse e ad un conseguente aumento della programmazione di brani di musica elettronica?

## RIFERIMENTI

1. W.J. Baumol, W.G. Bowen, *Performing arts - The economic dilemma*, New York, Twentieth Century Fund, 1966
2. N. Bernardini, A. Vidolin, *Sustainable live electro-acoustic music*, Proceedings SMC '05, 2nd Sound and Music Computing Conference, November 24-26 2005, Salerno, Italy
3. D. Brooke Wetzel, *A model for the conservation of interactive electroacoustic repertoire: analysis, reconstruction, and performance in the face of technological obsolescence*, Organized Sound vol. 11/3, December 2006, Cambridge University Press
4. G. Grisey, *Le temps et l'écume - partition*, Paris, Ricordi 1989
5. M.S Puckette, *New Public-Domain Realizations of Standard Pieces for Instruments and Live Electronics*, Proceedings of the International Computer Music Conference, 2001
6. S. Salvemini, *Quando "carmina dant panem": la cultura come risorsa economica*, Economia & Management, 2005/3

# Principi elementari per una sonificazione del gesto

**Maurizio Goina**

Scuola di Musica e Nuove Tecnologie  
Conservatorio “Giuseppe Tartini” di Trieste

**Pietro Polotti**

Dipartimento di Informatica  
Università di Verona  
Scuola di Musica e Nuove Tecnologie  
Conservatorio “Giuseppe Tartini” di Trieste

## SOMMARIO

In questo lavoro viene condotta un'indagine sulle relazioni tra gesto e suono per mezzo di una sonificazione elementare del gesto. L'ispirazione per questo lavoro viene dai principi del Bauhaus ed in particolare dalla ricerca di Paul Klee nel campo delle forme e della rappresentazione pittorica. In linea con queste idee, l'obiettivo principale di questo lavoro è quello di ridurre il gesto alla combinazione di un numero essenziale di componenti elementari (gestalt) usati per controllare un insieme di suoni corrispondentemente piccolo. Sulla base di una realizzazione software dimostrativa introduciamo qui una linea di ricerca che è ancora in corso di sviluppo. L'obiettivo futuro è quello di realizzare un nuovo sistema che sia uno strumento atto alla composizione e/o all'improvvisazione, come pure un'interfaccia per la danza interattiva.

## Parole Chiave

Bauhaus; Klee; analisi del gesto; sonificazione.

## INTRODUZIONE

Questo lavoro si ispira alle affinità tra pensiero del Bauhaus [1] e musica, particolarmente evidenti, per esempio, nell'attività di Paul Klee. Andrew Kagan, nel suo saggio 'Paul Klee, Art and Music', afferma: “fra tutti quelli interessati al rapporto tra musica e pittura nessuno più di Paul Klee dedicò così tanto tempo ed energie, e nessuno ottenne risposte, soluzioni e intuizioni così interessanti” in merito [2]. Klee fu violinista e per un periodo suonò nell'orchestra municipale di Berna, e con altre istituzioni musicali svizzere, come musicista semi-professionista. Ci piace immaginare che la pratica del violino sia stata per Klee una fonte di ispirazione per lo sviluppo dei principi formali della sua pittura, e che le linee e le curve che popolano i suoi quadri fossero in qualche modo collegati alla pratica della condotta dell'arco sulle corde del violino (vedi Figura 1).

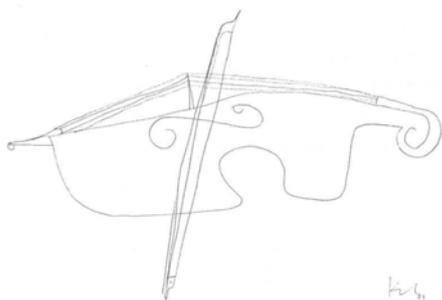


Figura 1. 'Violino ed archetto'. Disegno a matita di Paul Klee [3].

Seguendo gli insegnamenti di Klee, partiamo dal punto (Punkt), che egli chiama l'agente del movimento. Muovendosi, il punto genera delle linee (Linie). Questo concetto è mostrato da Klee nel primo esempio del suo "Pedagogisches Skizzenbuch" (Album pedagogico) [4], un libro concepito come manuale di base per il corso di Teoria del Design al Bauhaus. Il punto è concepito come elemento originante che genera linee e superfici (Fläche). In modo analogo consideriamo un gesto come il prodotto di sequenze di punti, le quali formano strutture che diventano complesse a differenti livelli. L'approccio è sostanzialmente astratto. Anche se la nostra indagine si estende anche all'ambito del visivo, non cerchiamo alcun riferimento a modelli o metafore fisici, come invece è stato fatto in altri lavori (vedi ad esempio [5] e [6]).

Questa ricerca vuole essere il punto di partenza di un lavoro che sembra promettente e che sarà sviluppato ulteriormente in futuro. Come spiegheremo in dettaglio successivamente, la nostra indagine si estende anche all'ambito del visivo. L'obiettivo complessivo è quello di identificare delle forme (gestalt) elementari, intese come strutture percettive/espressive basilari sia visuali che uditive, generate e controllate dal gesto, al fine di definire un mapping efficace (ed indiretto) tra forme e suoni. Qui presentiamo la prima realizzazione di uno strumento dimostrativo, realizzato con Max/MSP/Jitter, che implementa una sonificazione dei gesti generati dai movimenti di una sfera colorata. Nella seconda sezione viene innanzitutto proposta una panoramica dei principi estetici del Bauhaus ed in particolare di Paul Klee, in relazione al nostro lavoro. Nella terza sezione vengono poi definiti dei suoni elementari usati per la sonificazione di due categorie principali di gesti: movimenti in linea retta e movimenti curvi. Nella quarta sezione viene quindi sviluppato il tema delle Forme Elementari per la Sonificazione del Gesto (Elementary Gestalts for Gesture Sonification, EGGS) e viene infine illustrato il prototipo di un sistema basato su tali principi. L'articolo si conclude con una discussione dei possibili sviluppi futuri, alcuni dei quali già in corso d'opera.

## KLEE, IL BAUHAUS E I PRINCIPI DEL DESIGN E DELLA PRODUZIONE ARTISTICA

Nel 1921, Klee entrò a far parte della Scuola di Arte ed Architettura Bauhaus, la celebre scuola tedesca degli anni venti, dove insegnò fino al 1931, assieme ad altri importanti artisti come il pittore russo Wassily Kandinsky, l'architetto e designer tedesco Walter Gropius, il pittore tedesco Joseph Albers, il pittore e fotografo ungherese László Moholy Nagy, l'architetto tedesco Mies van der Rohe, il coreografo tedesco Oskar Schlemmer, ed altri. Fondamentalmente, l'insegnamento del Bauhaus promuoveva l'interazione interdisciplinare, e dalla fusione delle competenze delle facoltà fu creato un nuovo campo di studi: il design industriale. Il Bauhaus rappresentò la più importante rivoluzione nell'educazione estetica del

ventesimo secolo [7]: vi si promuoveva l'idea che il linguaggio visivo facesse parte di un linguaggio universale, non figurativo, e spesso venivano fatti dei paralleli con il linguaggio 'universale' della musica (vedi [2], [8], [9] e [10]). In particolare, l'interesse di Klee negli aspetti musicali della pittura ha a che vedere con il ritmo. Nel suo saggio [2] Klee afferma che fu il Cubismo a fornire a Paul Klee le prime solide basi per il suo pensiero musicale-pittorico: "Nell'alternanza tra superfici chiare e scure del Cubismo egli riconobbe un collegamento con il ritmo musicale .../... Durante la sua carriera, Klee continuò a elaborare e a raffinare il concetto di ritmo pittorico". D'altra parte "Klee stesso, particolarmente durante i primi anni della sua carriera, era estremamente cauto nel tracciare analogie tra le arti" e "l'effettiva applicazione da parte di Klee di modelli musicali alla sua arte avvenne solamente attraverso un lentissimo e lunghissimo processo di evoluzione". Inoltre Klee "credeva nell'asserzione di Goethe che colore e suono non possono essere confrontati direttamente .../... ma entrambi sono riferibili ad una formula universale." Questi concetti sembrano perfettamente in linea con le odierne ricerche sulla cross-modalità [11]. Quello che viene perseguito in questa ricerca è l'esplorazione delle caratteristiche della cross-modalità (per esempio la formula universale di Goethe) studiando una versione astratta della terna formata da gesto, suono, immagine. Nel suo Quaderno di schizzi pedagogici, Klee delinea un percorso didattico per i suoi studenti al Bauhaus, ma, allo stesso tempo, presenta i principi generali della sua ricerca artistica. Nella prima parte del libro, Klee introduce la trasformazione del punto statico in dinamiche lineari. Usando le parole della prefazione al quaderno, scritta da Sybil Moholy Nagy, la linea, essendo una sequenza di punti, "si sposta, circonda e crea dei contorni di figure piane passive e delle figure piane attive" (vedi le Figure 2, 3, 4, 5 e 6).

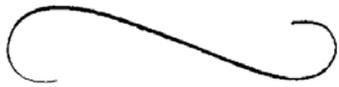


Figura 2. Linea attiva. L'agente è un punto in movimento. Movimento curvilineo. Disegno di Paul Klee [4].



Figura 3. Linea attiva. L'agente è un punto in movimento. Movimento curvilineo. Disegno di Paul Klee [4].

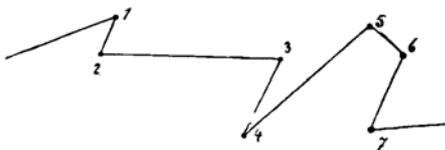


Figura 4. Linea attiva che si muove tra punti definiti. Movimento rettilineo. Disegno di Paul Klee [4].

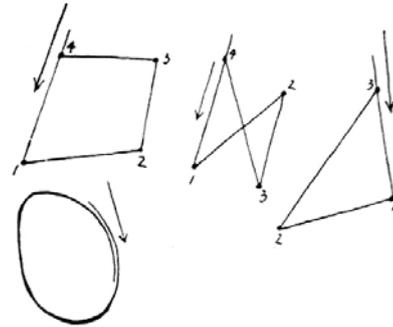


Figura 5. Linea mediale, a metà tra il movimento di un punto e l'effetto di superficie. Il carattere lineare lascia il posto all'idea di superficie. Disegno di Paul Klee [4].

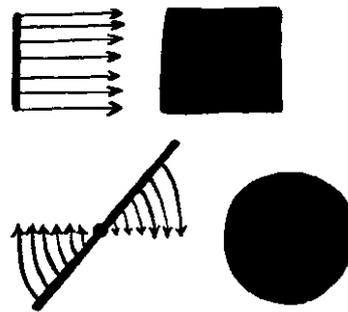


Figura 6. Linee passive che risultano da una superficie attiva. Disegno di Paul Klee [4].

Questo approccio astratto alla rappresentazione visiva è in qualche modo equivalente a quello che si vuole realizzare nel dominio uditivo: liberare il suono dal compito di esprimere e simbolizzare qualche cosa di per sé e dare ad esso una vita autonoma in relazione alla linearità o circolarità del gesto come elementi primi costruttivi.

Klee considerava la musica come un modello maturo per quello che egli voleva realizzare nel campo visivo: considerava "...la grandezza suprema di Mozart, Bach e Beethoven..." come guida per raggiungere "...una egualmente monumentale ed universale arte visiva del futuro..." [2]. Il fine ultimo era "...scoprire le proprietà estetiche generali che potessero essere isolate dai traguardi raggiunti dai titani della musica e quindi tradurre tali scoperte in termini visivi pratici, concreti ed efficaci" [2].

Quello che stiamo cercando di realizzare qui è una sorta di processo inverso, dal gesto al suono, considerando la lezione di Klee sui punti e le linee per definire un nuovo modo di comporre o, potremmo dire, "disegnare" il suono per mezzo del gesto.

#### VERSO UNA SONIFICAZIONE DEL GESTO PER MEZZO DI SUONI ELEMENTARI

A questo livello della ricerca lo scopo è quello di creare uno strumento virtuale che produca suoni "astratti" passando attraverso un'analisi e classificazione del gesto, dove il gesto è a sua volta concepito in termini astratti. L'obiettivo è quello di cercare relazioni originali tra gesto e suono tramite la ricombinazione di categorie elementari, partendo dal presupposto che non vi sia alcuna relazione necessaria tra gesto e suono. Al contrario, l'obiettivo è di mostrare come sia possibile costruire nuove ed efficaci relazioni tra gesto e suono definendo relazioni astratte secondo un mapping appropriato.

L'idea principale è di definire un numero di componenti elementari di traiettorie riferite al gesto, e di associare a ciascuno di essi una determinata categoria di suoni. In questa sezione vengono esposti i principi adottati ed i risultati preliminari ottenuti.

### Gesti elementari

Il nostro agente del movimento, l'equivalente del punto di Klee, è una sfera colorata. Il suo movimento produce linee diritte e curve in uno spazio bidimensionale, le quali controllano la produzione del suono. Klee comincia con un punto su un foglio, l'elemento base, a generare linee e, ad un livello più alto, superfici. In modo simile, qui è la posizione della sfera colorata che genera gesti e quindi suoni. La produzione del suono parte in questo modo dal gesto di chi tiene in mano la sfera. La ricerca sull'analisi e l'interpretazione del gesto è un argomento molto vasto [12]. Un gesto è in relazione diretta con il movimento ed è caricato di significati che hanno a che vedere, oltre che con la dinamica, la forza, l'inerzia, in quanto caratteristiche del movimento, anche con l'espressività. In questo lavoro tutto ciò è messo in secondo piano a favore di un approccio geometrico astratto all'analisi del gesto. A questo stadio, abbiamo deciso di ridurre l'insieme delle componenti del gesto a quello dei segmenti rettilinei e dei segmenti curvilinei. Per mezzo di queste due categorie semplici e molto generiche, scomponiamo il gesto in una sequenza di linee diritte e linee curve. Coerentemente, la sonificazione del gesto sarà una sequenza di suoni corrispondenti alle due categorie gestuali. Questo è il principio geometrico di base che controlla la selezione dell'una o dell'altra famiglia di suoni. Accanto a questa parte astratta, vengono considerati alcuni parametri secondari allo scopo di rendere la risposta sonora più percettivamente coerente con l'evoluzione dinamica del gesto. Ciò è descritto nei paragrafi successivi in modo dettagliato.

### Suoni elementari

Come già detto, si vogliono definire due categorie principali di suoni, corrispondenti a movimenti dritti e a movimenti circolari. Al momento vengono usati suoni generati da una semplice sintesi additiva: due armoniche forniscono il "suono lineare" e otto parziali inarmoniche formano il "suono curvilineo". In più, il suono curvilineo produce un glissando veloce e continuo, più precisamente un glissando infinito di Shepard (vedi [13] e [14]). Qui la scelta di un glissando infinito è stata fatta per applicare al suono il concetto di rotazione. Ci sono dei precedenti in questa scelta, per esempio Verfaillie, Quek e Wanderley usano il glissando di Risset per la sonificazione dei movimenti circolari della campana del clarinetto nel loro studio sui movimenti ancillari dell'esecuzione musicale [15].

Nella categoria del suono "circolare" viene poi fatta una ulteriore distinzione, tra rotazione in senso orario e rotazione in senso anti-orario (vedi Tabella 1). Inoltre, delle variazioni dinamiche di parametri di tali suoni astratti, come il pitch e l'intensità, vengono fatti corrispondere alla posizione assoluta, alla velocità ed all'accelerazione del gesto. Questi attributi ci riportano da un mondo totalmente astratto ad un mondo fisico. Ciò è necessario per evitare il rischio di monotonia. L'altezza del suono dipende dalla posizione spaziale assoluta (una posizione più alta nello spazio corrisponde ad un'altezza più alta del suono), e il volume sonoro è in relazione diretta con un valore mediato tra velocità ed accelerazione.

### SISTEMA PER UNA SONIFICAZIONE ELEMENTARE

Come già detto, ci siamo soffermati su due tipi di movimento, dritto e curvilineo (vedi Figure 2, 3 e 4), limitando la presente analisi al solo caso bi-dimensionale.

Anche solo con un tale semplice mapping, il suono risultante è tuttavia sorprendentemente ricco ed efficace, ed il sistema rivela inaspettate potenzialità in termini di esplorazione della relazione tra suono e movimento.

quiete	silenzio
movimento curvilineo in senso orario	glissando di Shepard ascendente + glissando semplice, dipendente dall'altezza del punto
movimento curvilineo in senso anti-orario	glissando di Shepard discendente + glissando semplice, dipendente dall'altezza del punto
movimento rettilineo	glissando semplice, dipendente dall'altezza del punto

Tabella 1. Relazione tra movimento e suono (mapping).

Per i nostri scopi abbiamo considerato una serie di sistemi per il riconoscimento del gesto. Per esempio, MnM [16] è un pacchetto incluso nella libreria esterna per Max/MSP FTM [17], e con esso viene fornito un Gesture-Follower (sistema di riconoscimento del gesto) (vedi anche [18] e [19]). Sfortunatamente questo strumento non è adatto per i nostri scopi poiché è concepito per riconoscere bene un singolo oggetto tra molti, mentre quello di cui abbiamo bisogno è di individuare delle caratteristiche astratte inerenti diversi tipi di traiettorie. Il nostro scopo è cioè quello di identificare una caratteristica comune ad infiniti oggetti. MnM necessita di essere istruito con molti singoli oggetti di una stessa famiglia se vogliamo che ne riconosca di simili. Noi invece, focalizzandoci sul gesto, abbiamo voluto trovare un algoritmo comune adatto a tutte le categorie gestuali definite: ad esempio consideriamo cerchi e spirali come oggetti appartenenti alla stessa categoria dei movimenti curvilinei.

Per implementare il sistema di tracking del movimento abbiamo iniziato a provare anche l'utilizzo di cv.jit [20], un pacchetto di oggetti esterni per Max/Msp/Jitter orientato alla computer vision. Sicuramente sfrutteremo le numerose e interessanti funzioni di tali oggetti per gli sviluppi futuri del nostro sistema.

Per l'implementazione attuale abbiamo impiegato Max/MSP/Jitter assieme a Jamoma [21][22]. Jamoma è un tool modulare basato su Max/MSP/Jitter, e serve a costruire velocemente delle patch in Max/MSP/Jitter semplificandone la struttura e abbreviando i tempi di programmazione. Ogni modulo infatti svolge una determinata funzione, dispone di interfaccia grafica che permette una rapida configurazione, ed è facilmente combinabile con altri moduli, oltre ad essere perfettamente inseribile in una normale patch. I due moduli qui impiegati sono jmod.input%, che gestisce l'input della telecamera, e jmod.motion% che calcola i parametri del movimento (vedi figura 7).

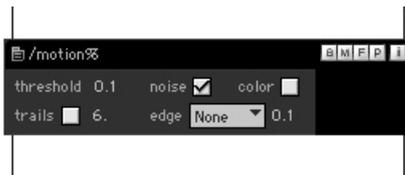


Figura 7. Uno dei moduli di Jamoma, il jmod.motion%.

Il nostro sistema realizza il tracking del colore mediante la combinazione di due tecniche: dapprima viene applicato un chromakey all'immagine catturata dalla camera, in maniera da selezionare il colore della sfera che usiamo per generare il gesto. Poi l'immagine viene trasformata in forma binaria, tenendo solo l'informazione di quali sono i pixel del colore selezionato. Infine vengono selezionate e tracciate solamente le parti in movimento di quel colore (tramite jmod.motion%). Si ottiene così che anche se sullo sfondo c'è qualche altro oggetto statico dello stesso colore, oppure anche qualche oggetto in movimento ma di un colore diverso, questi non vanno ad influire sul tracking della nostra sfera.

I dati spaziali così raccolti, concernenti il gesto, vengono elaborati in tempo reale in maniera da restituire cinque valori. Il primo va da 0 a 3 a seconda che ci sia quiete (nessun movimento), movimento curvilineo in senso anti-orario (Figura 9), movimento rettilineo (Figura 8), movimento curvilineo in senso orario (vedi Tabella 1). Il secondo valore è la velocità scalare del gesto. Il terzo è l'angolo, in radianti, del vettore velocità, calcolato rispetto all'origine. Il quarto valore è l'angolo totale, in radianti, calcolato dall'inizio di ogni sessione; quest'ultimo parametro è utile per avere un valore angolare che vari in maniera continua, senza il salto del valore di  $2\pi$  che altrimenti si avrebbe tra la fine di un angolo giro e l'inizio del seguente.

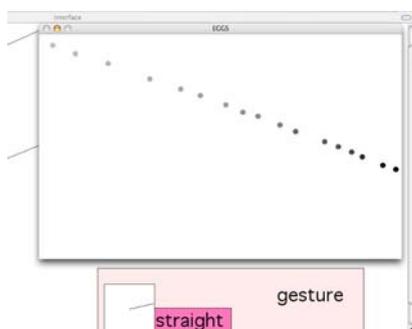


Figura 8. Riconoscimento e classificazione della traiettoria: movimento rettilineo.

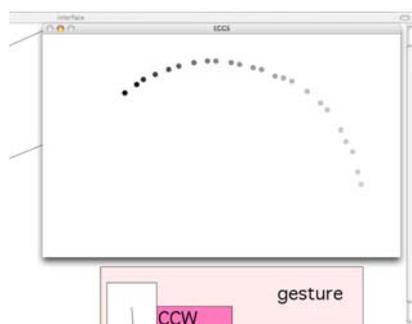


Figura 9. Riconoscimento e classificazione della traiettoria: movimento curvilineo anti-orario.

Da un punto di vista tecnico, la distinzione tra movimenti rettilinei e curvilinei è ottenuta misurando la variazione dell'angolo dei due segmenti definiti da 3 punti consecutivi (Figure 10 e 11), cioè l'accelerazione centripeta del moto.

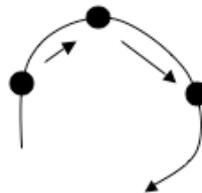


Figura 10. Tre punti consecutivi vengono impiegati per calcolare il tipo di moto.

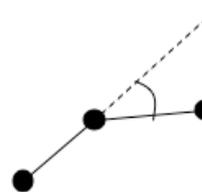


Figura 11. Deviazione angolare (accelerazione centripeta del moto).

Una variazione (in valore assoluto) prossima allo zero (o vicina al valore di mezzo angolo giro  $\pi$ ) è classificata come traiettoria rettilinea, altrimenti viene classificata come traiettoria curvilinea.

### POTENZIALITÀ PERFORMATIVE E SVILUPPI FUTURI

Ispirandoci agli esempi pittorici di Klee, dove le immagini vengono formate a partire da elementi semplici (come ad esempio nell'esempio pittorico delle Figure 12 e 13), abbiamo iniziato a realizzare una polifonia a due voci, tracciando il movimento di due sfere colorate diversamente. In futuro le voci potrebbero essere più di due, e si potrebbe ad esempio fare il tracking di più punti del corpo di un performer, il quale danzando potrebbe creare da solo delle situazioni sonore polifoniche.



Figura 12. "Häuser in der Landschaft" (Case nel paesaggio). Dipinto di Paul Klee [3].

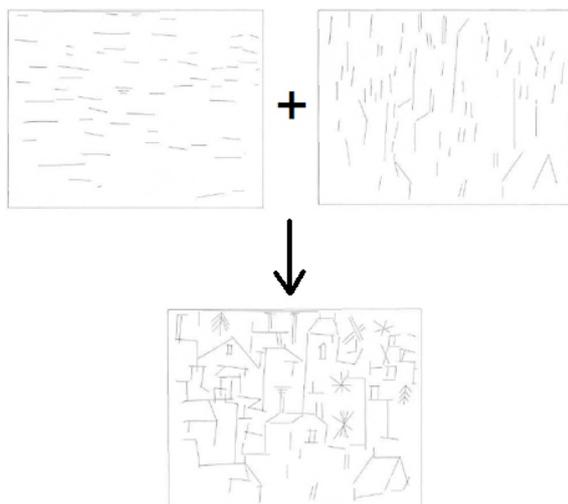


Figura 13. Analisi di “Häuser in der Landschaft” [3]. Poliphonia.

Ancora una volta ci siamo ispirati a Klee ed abbiamo iniziato ad implementare i suoi principi di formazione delle figure piane a partire da elementi semplici (Figure 5 e 6). In questo caso abbiamo usato i nostri “mattoni” elementari, i segmenti rettilinei ed i segmenti curvilinei, per definire dei gesti composti che vanno a formare figure piane, per ora limitate a cerchi e triangoli.

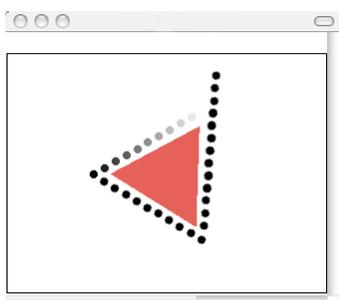


Figura 14. Il sistema riconosce un triangolo.

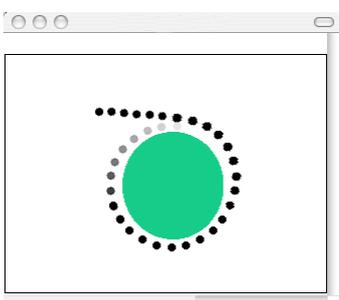


Figura 15. Il sistema riconosce un cerchio.

Quando una spezzata di retta si chiude a formare un triangolo il sistema lo riconosce come tale, lo sonifica e lo visualizza (vedi Figura 14). Lo stesso vale per un tratto di traiettoria circolare, che quando compie un giro completo viene riconosciuto come cerchio (vedi Figura 15). In futuro vogliamo estendere questo insieme di sub-categorie anche ad altre figure piane derivate dai segmenti elementari rettilinei e curvilinei, e cioè ad altre poligonali, ellissi, figure a otto,

spirali, ecc., da essere associate a suoni appropriati. Questo costituisce un primo passo in questa ricerca di gestalt elementari che abbiano una valenza cross-modale, ovvero a cavallo tra il visuale e l’uditivo.

Quello presentato è anche un primo e semplice sistema performativo. Sono concepibili molte altre possibilità in termini di articolazione delle combinazioni e di mapping elementare. Ad esempio, stiamo sperimentando un semplice processo di accumulazione, gestito da momenti di movimento e momenti di quiete: l’inizio della quiete è il segnale di partenza di un processo di ricomposizione che utilizza quanto generato e registrato nei momenti di moto. Il serrato alternarsi di movimenti e di momenti di quiete crea così delle situazioni polifoniche, anche utilizzando un unico punto generatore, in cui ogni ripresa viene disattivata automaticamente a tempo.

Inoltre, come in ogni attività musicale, il tema dell’apprendimento è fondamentale. L’esercizio è importante per capire le possibilità dello strumento ed ottenere risultati rilevanti. Allo stesso tempo non è richiesta all’esecutore un’abilità tecnica avanzata in quanto anche gesti semplici producono una sonificazione significativa. Inoltre, come già anticipato, seguendo ancora una volta gli insegnamenti di Klee e del Bauhaus, ed il paradigma “Punkt, Linie, Fläche” (punto, linea, piano), stiamo lavorando ad un’estensione del sistema per definire anche la sonificazione del piano. Da un punto di vista del suono, questo corrisponderà ad una texture sonora. Più in generale, nel futuro vorremmo anche sviluppare l’idea di usare il gesto come controllo sia per la generazione di suono che di immagini, come già in parte fatto con i triangoli e i cerchi. Si possono immaginare tre modi di mappare suoni con immagini: dal suono all’immagine, dall’immagine al suono, e la contemporanea generazione di suono e immagini. Con il sistema qui proposto, l’obiettivo finale sarebbe di cercare una nuova relazione tra suono ed immagine per mezzo della ricombinazione di categorie astratte controllate dal gesto. L’intenzione è di scoprire se la definizione di categorie astratte del gesto e la definizione di un mapping efficace ed indipendente sia per la generazione del suono che per quella dell’immagine rivelerà inaspettate relazioni tra immagini e suoni. Un tale sistema potrebbe essere usato per una ricerca di tipo artistico sui domini cross-modali e multi-modali, in particolare nell’ambito audiovisivo. Si può immaginare la possibilità di usare il sistema nelle arti performative e nella danza interattiva. In questo modo un danzatore, diventando un compositore audio-visivo, sarebbe l’esecutore ideale di uno strumento capace di rivelare nuove relazioni tra ascolto e visione.

## RIFERIMENTI

1. Andrew Kennedy. *Bauhaus*. Flame Tree Publishing, London, 2006.
2. Andrew Kagan. *Paul Klee, Art & Music*. Cornell University Press, Ithaca and London, 1987.
3. Paul Klee. *Teoria della Forma e della Figurazione*, trad. Mario Spagnol e Francesco Saba Sardi. Feltrinelli, Milano, 1984.
4. Paul Klee. *Pedagogical Sketchbook*, trans. Sibyl Moholy-Nagy. Frederick A. Praeger, New York, 1965.
5. Claude Cadoz, Annie Luciani, and Jean-Loup Florens. ‘Artistic creation and computer interactive multisensory simulation force feedback gesture transducers’. In *Proc. Conf. on New Interfaces for Musical Expression (NIME)*, pages 235–246, Montreal, Canada, May 2003.
6. Davide Rocchesso and Federico Fontana, editors. *The Sounding Object*. Mondo Estremo, Firenze, 2003.

7. Murray Schafer. *The Soundscape: Our Sonic Environment and the Tuning of the World*. Destiny Books, Rochester, Vermont, 1977.
8. Pierre Boulez. *Il paese fertile: Paul Klee e la musica*. Abscondita, Milano, 2004.
9. Wassily Kandinsky. *Point and Line to Plane*. Dover, New York, 1979.
10. Isabella Simonini. 'Storia del Basic Design', in Giovanni Anceschi, Massimo Botta e Maria Amata Garito, *L'ambiente dell'apprendimento - Web design e processi cognitivi*. McGraw Hill, Milano, Italia, 2006.
11. Antonio Camurri, Carlo Drioli, Barbara Mazzarino, and Gualtiero Volpe, 'Controlling Sound with Senses: multimodal and crossmodal approaches to control of interactive systems'. In Pietro Polotti and Davide Rocchesso, eds. *Sound to Sense, Sense to Sound. A State of the Art in Sound and Music Computing*. Logos Verlag, Berlin, 2008.
12. Antonio Camurri. and Gualtiero Volpe, eds., *Gesture-based Communication in Human-Computer Interaction, LNAI 2915*, Springer Verlag, February 2004.
13. [http://en.wikipedia.org/wiki/Shepard\\_tone](http://en.wikipedia.org/wiki/Shepard_tone)
14. Curtis Roads. *Computer Music Tutorial*. The MIT Press, Massachusetts, 1996.
15. Vincent Verfaillie, Oswald Quek, Marcelo M. Wanderley. 'Sonification of musicians' ancillary gestures'. In *Proceedings of the 12th International Conference on Auditory Display, London, UK June 20 - 23, 2006*.
16. Frederic Bevilacqua, Remy Mueller and Norbert Schnell. 'MnM: a Max/MSP mapping toolbox', in *Proceedings of the New Interfaces for Musical Expression Conference, NIME, Vancouver, Canada, 2005*.
17. Norbert Schnell, Riccardo Borghesi, Diemo Schwarz, Frederic Bevilacqua, Remy Muller. 2005. 'FTM – Complex Data Structures for Max.' *Proc. of ICMC 2005*. International Computer Music Association. Barcelona, Spain.
18. [http://ftm.ircam.fr/index.php/Gesture\\_Follower](http://ftm.ircam.fr/index.php/Gesture_Follower)
19. Frederic Bevilacqua, Fabrice Guédy, Norbert Schnell, Emmanuel Fléty, Nicolas Leroy. 'Wireless sensor interface and gesture-follower for music pedagogy.' In *Proceedings of the 2007 Conference on New Interfaces for Musical Expression (NIME07), New York, NY, USA*.
20. <http://www.iamas.ac.jp/~jovan02/cv/>
21. <http://jamoma.org/>
22. Tim Place and Trond Lossius. 'Jamoma: a modular standard for structuring patches in Max', in *Proceedings of the International Computer Music Conference 2006*.

# **Meta-Morfologie Musicali: un ambiente di composizione elettroacustica**

**Riccardo Dapelo, Roberto Doati,  
Emilio Pozzolini, Alessandro Quaranta,  
Martino Sarolli**

Scuola di Musica e Nuove Tecnologie,  
Conservatorio di Musica “Niccolò Paganini”  
di Genova

**Corrado Canepa, Barbara Mazzarino**  
InfoMus Lab – Casa Paganini, DIST –  
Università di Genova

## **SOMMARIO**

*Meta-Morfologie Musicali* è un progetto di diffusione della cultura musicale elettroacustica italiana realizzato da allievi e docenti del Conservatorio e dell’Università di Genova. Un ambiente interattivo con feedback visivo e sonoro consente anche a un pubblico di non addetti ai lavori di trasformare, generare e mescolare eventi sonori tratti da quattro opere fondamentali del repertorio elettroacustico: *Incontri di fasce sonore* di Franco Evangelisti, *Continuo* di Bruno Maderna, *Thema (Omaggio a Joyce)* di Luciano Berio, *Post-Prae-Ludium per Donau* di Luigi Nono. Oltre a una descrizione dei principi ispiratori e del funzionamento dell’ambiente, si affronteranno le problematiche di costruzione dell’interfaccia e di *interaction design*.

## **Parole Chiave**

Musica Elettroacustica, interazione multimodale, interfacce interattive, *interaction design*, EyesWeb.

## **PRINCIPI ISPIRATORI**

*Meta-Morfologie Musicali* (d’ora in avanti *M-MM*) è un progetto di diffusione della cultura tecnico-scientifica che si occupa della valorizzazione del patrimonio musicale elettroacustico del nostro Paese, nonché di un’attività di formazione e di divulgazione al fine di stimolare l’interesse alle tematiche della ricerca e della sperimentazione musicale attraverso l’impiego delle nuove tecnologie multimediali.

La cultura musicale si è fortemente contaminata con quella scientifica a seguito della nascita della Musica Elettroacustica negli anni ’50 del secolo scorso. La realizzazione di opere pensate e realizzate con tecnologie elettroniche ha portato a nuovi sviluppi del linguaggio, delle forme e dei contenuti musicali, con il risultato che la musica d’oggi è il frutto di uno stretto e complesso rapporto fra compositori, scienziati, esecutori e pubblico. In questo senso l’Italia ha avuto un ruolo di estrema importanza, a cominciare dallo Studio di Fonologia Musicale della RAI di Milano e anche grazie ai corsi di Musica Elettronica istituiti nei Conservatori già dalla metà degli anni ’60. Con l’avvento delle tecnologie informatiche, meno costose, più semplici ed efficienti da usare, l’impatto sociale è oggi divenuto ancora più forte, ma a questo non corrisponde una reale conoscenza delle qualità estetiche ed espressive della musica elettroacustica.

Per questo motivo si è pensato a un progetto che avesse come obiettivo principale la diffusione della cultura elettroacustica, realizzando un ambiente multimediale che consente anche a un pubblico di non addetti ai lavori di trasformare, generare e mescolare eventi sonori tratti da quattro opere fondamentali del repertorio elettroacustico: *Incontri di fasce sonore* (1956-57) di Franco Evangelisti, *Continuo* (1958) di Bruno Maderna,

*Thema (Omaggio a Joyce)* (1958) di Luciano Berio, *Post-Prae-Ludium per Donau* (1987) di Luigi Nono.

Anche se può apparire tale, sarebbe errato definire *M-MM* un’installazione. Non si tratta infatti di un assemblaggio artistico in cui il nesso di causalità tra azione dell’utente e risultato può non essere evidente”, peculiarità, questa, del pensiero creativo. Per sfruttare pienamente le potenzialità di *M-MM* è necessaria una guida, una sorta di assistente, come un tempo era indispensabile negli studi di musica elettronica il cosiddetto “tecnico del suono” che istruiva i compositori all’uso di apparecchiature che allora non facevano certo parte dell’organologia musicale. Più corretto sarebbe quindi definirlo una sorta di laboratorio sperimentale in cui è possibile ricreare virtualmente le operazioni di generazione, trasformazione e montaggio di materiali sonori tipiche della musica elettroacustica.

Sono tre i principi su cui si fonda *M-MM*:

- il rifiuto di una mera simulazione di singole apparecchiature;
- la condivisione di alcuni aspetti dello strutturalismo di Piaget, in particolare quando afferma che per conoscere un oggetto il bambino agisce su di esso, lo modifica, lo trasforma. E con la comprensione del processo di tali cambiamenti, raggiunge anche la comprensione del modo in cui l’oggetto viene costruito [1];
- la scelta di una forma di interazione non limitata al concetto di bottone o di tasto.

La nostra intenzione infatti non era quella di costruire l’ennesimo sintetizzatore, ma di fare agire il fruitore su strutture musicali complesse e non in modo generico, ma attraverso l’ottica e il pensiero di alcuni dei grandi protagonisti della musica del XX secolo. In alcuni casi re-interpretando l’opera (tradizionalmente per la musica scritta si parla di interpretazione, ma essendo la musica su nastro una consegna definitiva dell’interpretazione del compositore alla Storia, in mancanza di una partitura ci troviamo di fronte a una vera e propria ri-creazione), in altri simulando parte del percorso compositivo dell’autore, in altri ancora “giocando” (nel doppio significato, assente nei lemmi della lingua italiana, di “suonare” e “giocare”) con gli strumenti del comporre elettronico.

Da almeno 20 anni [2] è pratica diffusa presso alcune Scuole di Musica Elettronica la ricostruzione, più o meno filologica [3], di opere del repertorio elettroacustico. È convinzione di chi scrive che tale pratica sia ausilio fondamentale, non solo per la disciplina analitica, ma anche per l’approfondimento delle tecniche di elaborazione del suono, per l’approfondimento della Storia e delle tecniche compositive .

La possibilità di agire su parametri globali di eventi sonori *all’origine di o creati per* le opere musicali scelte, viene

consentita dalla predisposizione di un ambiente esecutivo che simula o formalizza quello originalmente costituito da apparecchiature analogiche collegate fra loro, consultazioni di tabelle, interventi manuali di compositori e assistenti, ecc. Più semplicemente, in *M-MM* la comunicazione avviene attraverso una serie di gesti “puri” su uno schermo che comprende: toccare, trascinare, traslare, espandere, comprimere, spostare, lanciare, delimitare.

Anche quando si tratti di percezione uditiva e di arte sonora sembra imprescindibile, per l’essere umano, la necessità di ricorrere a una rappresentazione visiva, specie in connessione a esperienze percettive profonde. Per questa ragione i movimenti sono compiuti su immagini che presentano una certa analogia o affinità formale con l’evento sonoro denotato, oppure ne sono una vera e propria rappresentazione grafica. La presa di contatto con alcuni processi compositivi caratteristici del repertorio elettroacustico avviene quindi modellando con le mani delle immagini-suono.

## L'INTERFACCIA

La progettazione dell’interfaccia in tutti i suoi aspetti complessi e intrecciati (la progettazione hardware, il tipo di interazione, i risultati ottenibili attraverso l’analisi del gesto, le primitive più efficaci per la rappresentazione e la manipolazione delle opere, le ipotesi di reazione di un fruitore di cultura media), ha richiesto un lungo lavoro di discussione, studio e test che ha coinvolto studenti e docenti del Conservatorio e dell’Università. Durante la fase di progettazione sono stati realizzati diversi prototipi da sottoporre a studenti e a soggetti scelti casualmente per valutarne le reazioni.

L’idea emersa in fase di progetto si è focalizzata sulla possibilità di utilizzare il paradigma “apprendere mediante il fare” [4], sfruttando sensi differenti quali tatto e vista. Da subito è comparsa la necessità di rendere l’azione, alla base dell’interazione, il più naturale e semplice possibile, ma sufficiente per stimolare l’esplorazione. Per questo si è cercato di utilizzare una struttura che fornisca anche un feedback tattile congruente sia con il gesto che con il feedback sonoro. In pratica si è sviluppata un’interfaccia interattiva che permette la manipolazione di forme alle quali sono associati importanti aspetti delle opere musicali scelte.

Ogni gesto e comportamento del fruitore viene rilevato e analizzato tramite piattaforma EyesWeb, quindi inviato tramite protocollo OSC (OpenSoundControl) a due computer portatili con piattaforma Max/MSP - Jitter che gestiscono rispettivamente la generazione dell’audio e la generazione dei contenuti visivi con il controllo delle varie scene. Non è da escludersi, dopo un congruo periodo di test e un accurato lavoro di ottimizzazione, la possibilità di ridurre a una sola macchina la generazione dei contenuti audiovisivi.

### Hardware

Per mantenere la manipolazione semplice e naturale introducendo un feedback tattile, si è realizzata una superficie “sensorizzata” in tessuto. Più in dettaglio le forme visive vengono retro-proiettate sopra uno schermo in tessuto, su cui il fruitore le manipola toccandole, trascinandole e così via. Il risultato ottenuto ricorda, nel comportamento, un *touch-screen*, ma con la differenza che nel nostro caso è presente una deformazione della superficie (dovuta all’impiego del tessuto) e le dimensioni sono decisamente più rilevanti (cm 90x120 di luce interna).

Il tracking delle mani è stato effettuato utilizzando una illuminazione diffusa e radente la superficie, e con l’impiego di una telecamera a infrarossi posizionata dietro lo schermo. In particolare, sopra la superficie di tela, è stata creata una

“lama”, di luce a infrarosso, generata da 150 Led IR 850 nm 26 gradi opportunamente fissati a una cornice con la telecamera (in Bianco e Nero) posizionata dalla parte opposta e utilizzata con una risoluzione 720x576 a 25 fps. Il tipo di illuminazione scelto permette di identificare un oggetto quando questo si trova a contatto con la tela o nelle sue immediate vicinanze, a discapito di una maggior precisione nel riconoscimento della sua forma.

### Software di tracking

Il riconoscimento dei gesti di manipolazione, ad esempio tocco e trascinamento, sono stati realizzati sfruttando la piattaforma software EyesWeb ([www.eyesweb.org](http://www.eyesweb.org)), sviluppata dal laboratorio InfoMus del DIST- Università di Genova, e le sue librerie [5].

Nello specifico, dal segnale video vengono identificati i blob che rappresentano gli oggetti (ad esempio mani e dita) che “toccano” la superficie. Una volta distinti tra loro i blob, la piattaforma mette a disposizione diversi moduli software per estrarne caratteristiche cinematiche, come posizione, direzione, velocità del movimento, o caratteristiche legate al tipo di movimento eseguito (diretto, morbido, energico, ecc.). Tutte queste caratteristiche vengono generate in tempo reale, ed essendo grandezze normalizzate, possono essere utilizzate per controllare differenti tipologie di feedback audio e/o video [6].

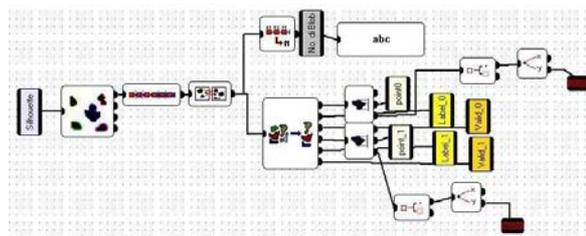


Figura 1. Estratto della patch EyesWeb che permette di separare i singoli blob, seguirli nel tempo, ed estrarne alcune caratteristiche.

Per la parte audio sono stati implementati in Max/MSP algoritmi di risintesi e/o di generazione delle opere in tempo reale per permettere la massima rapidità di feedback durante la manipolazione da parte del fruitore.

### Mapping

Tutte le manipolazioni sono possibili, non vi sono costrizioni a priori, ma per mantenere il concetto di semplicità e soprattutto di congruenza tra gesto, parte visiva e parte audio, solo alcune manipolazioni modificano gli oggetti associati a ciascuna opera. La scelta di queste manipolazioni è stata la parte fondamentale dell’*interaction design* in quanto, se venisse a mancare una oggettiva congruenza tra movimento visivo, fisico e movimento acustico, verrebbe automaticamente a crearsi un paradosso percettivo. Per questo motivo solo alcuni gesti generano modificazioni alle sorgenti sonore e, per questo stesso motivo, ogni singolo movimento attivo è stato pensato per essere semplice ma evocativo rispetto all’unità sonora estratta da ciascuna opera.

### Generazione dei contenuti audiovisivi

Anche questo aspetto è stato oggetto di lunghe e intense discussioni da parte del team di sviluppo, durante le quali sono state anche analizzate ed esaminate alcune considerazioni di Paul Klee sul concetto di figura e sulle sue implicazioni espressive [7].

Per la parte audio sono stati implementati in Max/MSP 4.6.3 e Csound5 (per alcune risintesi in tempo differito), algoritmi di risintesi e/o di generazione in tempo reale delle opere scelte, al fine di consentire la massima rapidità di feedback durante la



Figura 2. Caricamento delle icone come *texture*

manipolazione da parte del fruitore. Gli algoritmi audio sono ottimizzati (con controllo dei rispettivi *mute enable* per essere inseriti nella catena Dsp solo quando attivati dal fruitore) e predisposti per ricevere i messaggi OSC provenienti da EyesWeb. Dall'analisi del gesto vengono estratti in ogni scena i parametri ritenuti più significativi che generano (con opportuni riscaldamenti) trasformazioni degli oggetti audio e video.

In considerazione della sua rapidità di esecuzione e del basso carico computazionale, per la parte video si è optato per la generazione in linguaggio OpenGL (in ambiente Jitter 1.6.3) con un uso intensivo delle tecniche di *layering* (stratificazioni di più livelli video sovrapposti) e di *texture rendering* (stratificazione di *texture* multiple) per preservare l'indipendenza spaziale, di colore e dimensione degli oggetti manipolabili rispetto a quelli di sfondo (si vedano figure 2 e 3)

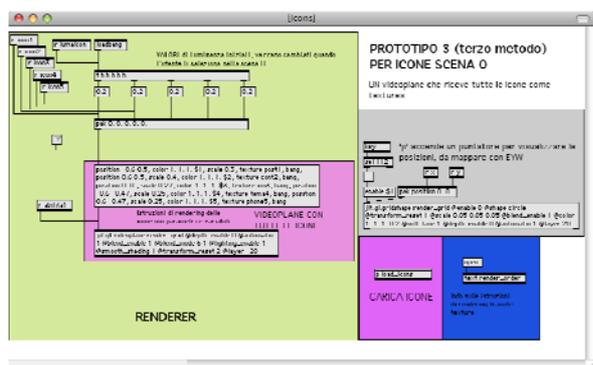


Figura 3. Chiamata sequenziale delle *texture* (multi texturing)

Il grande impegno nella programmazione ha riguardato:

- la fluidità delle risposte dei contenuti audiovisivi;
- il ripristino degli stati iniziali di ogni patch dopo una serie di azioni da parte del fruitore;
- la valutazione dell'assenza di interazione come segnale per ritornare alle schermate precedenti con o senza cambiamenti in base ai comportamenti del fruitore, nel modo più comprensibile e intuitivo;
- l'ottimizzazione del peso di ogni patch (in particolare l'eliminazione di tutte le *number boxes* e degli elementi di GUI che possono rallentare il frame rate video o pesare sull'audio) (figura 4)

## DESCRIZIONE E IMPLEMENTAZIONE

Cosa vede il fruitore quando si avvicina a *M-MM*?

1. La schermata iniziale consiste nella simulazione di una superficie fluida attraverso la quale appaiono 4 icone che rappresentano le opere musicali come se fossero immerse nel fluido. La superficie reagisce al contatto con la mano: il fluido si increspa nel punto di contatto e segue il percorso del fruitore con ondulazioni.

2. Se durante il percorso viene sfiorata un'icona essa si illumina (a simulare quasi un'emersione dal fondo) e si ascolta un estratto sonoro di circa un minuto.
3. Quando sull'icona viene esercitata una pressione oltre un determinato intervallo di tempo, si entra nell'ambiente di sperimentazione per quell'opera.
4. Compaiono altre icone che di volta in volta indicheranno un sistema di notazione, un'analisi fisica del suono, una sorgente sonora, un percorso, una rappresentazione visiva di un parametro sonoro.
5. A questo punto il fruitore può scegliere un'icona e nell'ambito dei gesti consentiti, sperimentare direttamente la manipolazione dell'evento sonoro.

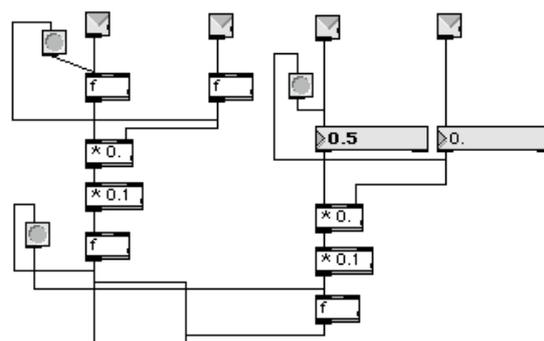


Figura 4. Gli oggetti *<f>* (a sinistra) a sostituzione dei floating number boxes (a destra)

## Incontri di fasce sonore (1956-57)

Questa opera di Franco Evangelisti gode della stampa di una partitura redatta dall'autore [8]. Si tratta di una partitura realizzativa, anche se contrariamente a quanto affermato da Evangelisti, è stata redatta dopo la composizione: «Il metodo qui raffigurato è stato quello che ha permesso al tecnico la realizzazione». Va perciò collocata in quel clima di confluenza di credo serialista in opposizione alla tradizione e di malcelato desiderio di subentrarle seguendola, tipico di quegli anni: «La notazione [in questa stampa] permette anche la visione della partitura in senso tradizionale».

Dal momento che sono riportate indicazioni tecniche, anche se relative ad apparecchiature che solo in parte sono emulabili nel dominio digitale, in *M-MM* il fruitore può a sua volta "interpretare" alcune della fasce sonore che compongono l'opera (ciascuna somma di 7 sinusoidi) sintetizzate in tempo reale e di cui si possono mutare l'ambito frequenziale, la trasposizione, la durata, sempre nel rispetto dei valori posti dall'autore nella tabella di frequenze per i "Gruppi e loro Elementi" come l'autore definisce le sue fasce sonore.

Come sfondo sullo schermo è presente la riproduzione della partitura; le fasce modificabili dal fruitore appaiono in colorazione differente tra di loro e leggermente in rilievo rispetto allo sfondo. Il brano viene riprodotto automaticamente all'apertura della patch mentre un cursore scorre lungo la sequenza di fasce. Arrivato al termine della pagina si passa alla successiva (per un totale di 1'30", le prime quattro pagine circa). Se il fruitore non compie nessuna azione, la scena si ripete, dopodiché si ritorna alla schermata iniziale.

Selezionando una fascia (la modalità di selezione è comune a tutte le scene: toccare un oggetto per un certo intervallo di tempo) la sequenza audio sfuma, lo sfondo inverte i colori e la fascia selezionata viene zoomata sull'asse *z*. Visualizzata al centro dello schermo, sarà possibile ascoltarla in loop. Modalità di editing sulla singola fascia:

- muovendo l'oggetto fascia con una mano sull'asse y avverrà una traslazione da un Elemento a un altro all'interno del Gruppo a cui appartiene;

- dilatando o comprimendo l'oggetto sull'asse y (con due mani) si otterrà una modifica dello spettro mediante interpolazione fra gruppi diversi (ogni Gruppo è caratterizzato da diverse larghezze di banda critica).

Una volta terminato l'editing (assenza di ulteriori azioni da parte del fruitore) la fascia ritorna automaticamente al suo posto nella sequenza iniziale, che riacquista il volume originale. La partitura sullo sfondo riprende il colore originale, ma la fascia o le fasce modificate suoneranno differenti nella partitura.



Figura 5. Interfaccia gestuale per *Incontri di fasce sonore*

#### *Continuo* (1958)

Nel caso di *Continuo* il fruitore non re-interpreta l'opera, ma viene messo in condizione di sperimentare la prassi compositiva di Bruno Maderna fondata su un'attenta investigazione, anche a forte componente ludica, sia dei materiali sonori che dei processi di trasformazione. «Progetti scritti, schemi non ne faceva mai. Maderna raccoglieva del materiale, ricercava, catalogava i suoni, li chiamava in qualche modo. Questi materiali non sono rimasti perché Maderna li considerava elementi primari, da costruzione. Costruiva questi colori, questi impasti e se li metteva da parte. A questo punto iniziava la costruzione. Anche la composizione era in un certo senso improvvisata». [9]

Con *M-MM* si possono esplorare i 2 materiali musicali con cui è stata composta l'opera: fasce di rumore bianco filtrate e impulsi filtrati).

Durante l'ascolto di una parte del brano (risintesi di 1' circa) che contiene le due tipologie sonore su cui il fruitore agirà, viene generato sullo schermo il corrispondente sonogramma ( $y$ =frequenza,  $x$ =tempo, colore=intensità). Sull'asse  $y$  è rappresentata una tastiera di pianoforte (indicazione ritenuta necessaria per far meglio comprendere al fruitore che la rappresentazione è nel dominio del tempo, mentre alcune delle modifiche successive si svolgeranno nel dominio della frequenza con conseguente rotazione dell'immagine).

Si può ora scegliere (con la consueta modalità della pressione prolungata) una delle due tipologie che troverà marcate separatamente. Al termine dell'ascolto si creano due riquadri che evidenziano le due tipologie.

Modalità di editing sugli impulsi: compare in trasparenza una specie di scatola rettangolare poggiata sul lato corto in cui una pallina rimbalza producendo un treno di impulsi

- il fruitore preme lo schermo e crea un impulso secco;
- il fruitore lancia la pallina e la forza di lancio aumenta la velocità (frequenza) degli impulsi;
- rimbalzando sull'asse  $y$  si varia la frequenza centrale del filtro passa-banda usato da Maderna;
- se si usa 1 dito si genera 1 impulso, se si usa la mano uno sciame di impulsi.

Modalità di editing sul rumore bianco: in ascolto abbiamo un rumore bianco non filtrato. La visualizzazione del sonogramma comincia a ruotare lentamente in senso antiorario, per diventare: asse  $y$ =intensità e asse  $x$ =frequenza. Ora la figura rappresenta un banco di filtri controllabili in frequenza e ampiezza, con involuppo spettrale delineabile dal fruitore. Trascinando un dito verso l'alto si alza il livello del filtro alla frequenza corrispondente.

#### *Thema (Omaggio a Joyce)* (1958)

È nota a tutti la capacità, virtù rara nei compositori, con cui Luciano Berio riusciva a scrivere delle proprie opere senza farne un surrogato verbale, bensì aggiungendo significato a quello già presente nella sua musica. Abbiamo pertanto deciso di definire l'ambito di azione su *Thema (Omaggio a Joyce)* partendo dal testo del 1959 [10] in cui viene descritto passo a passo il percorso dell'autore verso la composizione.

In particolare, grazie alla recente pubblicazione in cd [11] della trasmissione radiofonica che è all'origine dell'opera in questione (nella prefazione alla III edizione del suo *Opera aperta*, Eco la definisce «una dimensione puramente musicale» del documentario sonoro), siamo stati in grado di ricostruire uno dei primi esperimenti, poi abbandonato da Berio nella realizzazione definitiva perché ritenuto di eccessiva complessità polifonica, quello cioè in cui si porta a compimento «l'incontro organizzato di tre lingue diverse». In *M-MM* parole prese dalla lettura originale del testo joyciano in inglese, italiano e francese (a cura di Cathy Berberian, Ruggero de Daninos e Umberto Eco), possono comporre una trama polifonica a tre voci variando rapporti di tempo, altezza e dinamiche.

Sullo sfondo è presente una immagine tratta da uno schema di Joyce che illustra alcuni aspetti di *Ulysses* (il cosiddetto schema Linati, realizzato dall'autore nel 1920 circa per spiegare a un amico alcuni passaggi dell'opera [12]). Nella parte superiore dello schermo sono presenti 3 colonne di 4 parole, tratte dalle letture originali nelle 3 lingue. A ciascuna parola corrisponde la sua registrazione audio che si può ascoltare secondo le modalità consuete di selezione, sfiorando un'icona che si illumina. Con una pressione prolungata viene selezionata, cambia colore e compare nella parte sottostante dello schermo. Una volta selezionate le parole da manipolare (al massimo 5) e fatta la scelta, la sequenza viene riportata nella parte sottostante dello schermo in una traccia. La traccia è a sua volta duplicabile per ottenere un totale di 3 tracce traslando con la mano verso il basso la traccia originale.

Modalità di editing sulle singole tracce:

- ritardo temporale tramite spostamento sull'asse  $x$  della seconda o terza traccia;
- espansione/compressione temporale senza trasposizione dell'altezza agendo sulle singole parole nella colonna con due mani;

- variazioni dinamiche ottenute pescando da un menu grafico che compare a destra con 4 diversi involuppi: *crescendo*, *diminuendo*, *crescendo-diminuendo*, *diminuendo-crescendo*.

L'implementazione audio prevede il playback di una serie di buffer predisposti per caricare e disporre in sequenza le parole scelte dal fruitore e fa uso dell'oggetto MaxMSP gizmo~ per effettuare le espansioni/compressioni temporali nel dominio spettrale.

#### *Post-Prae-Ludium per Donau* (1987)

Fra i diversi insegnamenti dell'amico-maestro Bruno Maderna, Luigi Nono annoverava quello di un uso dell'elettronica che non fosse la mera realizzazione di un progetto scritto: «...Bruno Maderna, geniale maestro anche nell'iniziarmi a questo studio... Per qualche ora si lavorò alla realizzazione di un grafico acustico, preparato "per bene" da me a Venezia... Il risultato fu una "cosa" pressoché insignificante e una

“sganassada” formidabile di Marino e Bruno. Primo insegnamento immediato: lo studio elettronico esige non “preprogettati” al tavolino, bensì studio-sperimentazione-ascolto sempre in *tempo reale* in ogni suo attimo...» [13]

Ecco quindi perché in *M-MM*, in modo analogo a *Continuo*, si potrà agire non sulla partitura di *Post-Prae-Ludium per Donau* [14] (che come per tutte le opere degli anni '80 aveva funzione di nota sommaria per ricordare quanto stabilito di comune accordo con gli interpreti), ma direttamente sul suono della parte per tuba, registrata per l'occasione da Michele Odetti. Il fruitore può quindi provare per esperienza diretta una fase importante della pratica compositiva noniana. Una volta compiuta la scelta fra le articolazioni strumentali disponibili, potrà “lasciarsi andare” nella creazione di eventi sonori ottenuti con la proiezione del suono nello spazio, la traslazione nel tempo e la trasposizione di altezza.

Sullo sfondo compare la partitura con i materiali della sola tuba. Il fruitore, con un gesto di sfioramento dell'icona di una tuba che cambia stato (passa a una luminosità inferiore) può attivare il live electronics sentendo la differenza con la sola tuba. Selezionando il live electronics compare in trasparenza anche un'animazione che simula il comportamento dell'esecutore che apre e chiude i fader dei diversi circuiti.

Al termine dell'ascolto compaiono 6 icone corrispondenti a: 3 diverse articolazioni di tuba (lunghezza media 10") prese dalla prima parte dell'opera (fra 0'00" e 5'20") e 3 note lunghe (media di 5") e acute prese dall'inizio della seconda parte.

Modalità di editing:

- prima parte (riverbero e ritardi). Trascinando le icone delle articolazioni con la mano si crea una scia visiva a cui corrisponde l'ascolto di 4 ritardi temporali (uno diverso per ogni canale). La lunghezza della scia (e il feedback delle linee di ritardo) sarà determinata dalla durata del trascinamento. Dopo aver creato le scie, ogni successivo sfioramento delle icone da parte del fruitore farà ripartire la linea di ritardo, consentendo di creare contrappunti temporali. Il materiale preregistrato della tuba è già passato nel riverbero di 5" indicato da Nono in partitura;

- seconda parte (4 trasposizioni e spazializzazione). Toccando un'icona delle note lunghe e acute, compaiono due cerchi che indicano la spazializzazione incrociata (2 trasposizioni vanno in senso orario, 2 in senso antiorario). Il fruitore può trascinare l'icona in uno dei due cerchi (o in entrambi) e sente il suono ruotare con le trasposizioni. Terminata l'esecuzione del ciclo l'icona torna alla sua posizione originale.

Combinando le interazioni tra note lunghe e articolazioni, sarà possibile sperimentare sovrapposizioni tra le due tipologie di materiali musicali e il loro trattamento.

## SVILUPPI FUTURI

*Meta-Morfologie Musicali* mette in atto più di un'interazione; non solo quella fra il singolo fruitore e l'ambiente audiovisivo, ma anche quella fra lui e il pubblico che assiste alla fruizione. Attraverso gli sbagli, le scoperte e in generale le azioni di un individuo, si aumenta la comprensione del funzionamento e delle potenzialità dell'ambiente. Per questa ragione sarebbe nostro interesse pensare a un utilizzo collettivo e quindi alla realizzazione di un'interfaccia multiutente.

Un altro ampliamento potrebbe riguardare la disponibilità di un maggior numero di materiali sonori su cui agire. Il risultato della manipolazione potrebbe essere salvato in un archivio per poi passare a un II livello per la composizione nel tempo dando così vita a una nuova, originale opera elettroacustica.

*Meta-Morfologie Musicali* sarà aperto al pubblico durante il Festival della Scienza di Genova 2008 (23 ottobre - 4 novembre) negli spazi di Casa Paganini. Una registrazione audiovisiva del suo utilizzo è disponibile al seguente indirizzo: <http://www.conservatoriopaganini.org/video/video.html>.

## RINGRAZIAMENTI

*Meta-Morfologie Musicali* è un progetto finanziato dal Ministero dell'Università e della Ricerca, Decreto Direttoriale 21.03.2006, per il periodo aprile 2007 - aprile 2008. Desideriamo ringraziare quanti hanno contribuito alla sua realizzazione, con il loro lavoro e con suggerimenti, osservazioni, soluzioni tecniche, critiche. In particolare Franco Beretta, Talia Pecker Berio, Antonio Camurri, Paolo Coletta, Alessandro Fausto, Nicola Ferrari, Maddalena Novati, Michele Odetti, Emanuela Quaranta, Matteo Ricchetti, Alvisè Vidolin.

## RIFERIMENTI

1. Jean Piaget. *Psicologia dell'intelligenza*, (1947), Giunti-Barbera, Firenze, 1952.
2. Paolo Furlani e Paolo Zavagna. *Uno studio musicale all'elaboratore. Analisi e sintesi di Studie I di K. Stockhausen*, Campanotto editore, Udine, 1988.
3. Alvisè Vidolin. "Re-Synthesis of Analogue Electronic Music Compositions by Computer: a Teaching Experience", in *Science, Philosophy and Music, Proceedings of the XXth International Congress of History of Science ( Liège, 20-26 July 1997)*, a cura di Erwin Neuwander e Laurence Bouquiaux, Brepols Publishers, Turnhout (B), 2002.
4. *Enaction and Enactive Interfaces: A handbook of terms*, Ed. Luciani-Cadoz, 2007.
5. A. Camurri, B. Mazzarino, G. Volpe, "Analysis of Expressive Gesture: The EyesWeb Expressive Gesture Processing Library", in A. Camurri, G. Volpe (Eds.), "Gesture-based Communication in Human-Computer Interaction", LNAI 2915, pp. 460-467, 2004.
6. A. Camurri, B. Mazzarino, G. Volpe, "Expressive interfaces", *Cognition, Technology & Work*, 6(1): 15-22, Springer-Verlag, February 2004.
7. Paul Klee, *Das bildliche Denken*, Benno Schwabe, Basilea, 1956 (trad. it. *Teoria della forma e della figurazione*, Feltrinelli, Milano, 1959) e *Unendliche Naturgeschichte*, Benno Schwabe, Basilea, 1970 (trad. it. *Teoria della forma e della figurazione volume II. Storia naturale infinita*, Feltrinelli, Milano, 1970).
8. Franco Evangelisti. *Studio elettronico. Incontri di fasce sonore*, Universal Edition UE 12863, Vienna, 1958.
9. Marino Zuccheri, Comunicazione in occasione della intitolazione del Conservatorio di Musica di Cesena a "Bruno Maderna", Cesena, 19-20 novembre 1990.
10. Luciano Berio, "Poesia e musica - un'esperienza", in *Incontri Musicali* n. 3, Milano, 1959; ripubblicato in *La musica elettronica. Testi scelti e commentati da Henri Pousseur*, Feltrinelli, Milano, 1976; e in *Nuova musica alla radio. Esperienze allo Studio di Fonologia della Rai di Milano 1954-1959*, a cura di Veniero Rizzardi e Angela Ida De Benedictis, Rai-Eri/Cidim/Amic, 2000.
11. Luciano Berio e Umberto Eco, "Omaggio a Joyce. Documenti sulla qualità onomatopeica del linguaggio poetico" (1958), in *Nuova musica alla radio. Esperienze allo Studio di Fonologia della Rai di Milano 1954-1959*, a cura di Veniero Rizzardi e Angela Ida De Benedictis, Rai-Eri/Cidim/Amic, 2000, con CD audio allegato.
12. Peter Spielberg, *James Joyce's Manuscripts and Letters at the University of Buffalo*, University of Buffalo, 1962.
13. Luigi Nono, "Per Marino Zuccheri", in Roberto Doati e Alvisè Vidolin (a cura di), *Nuova Atlantide - Il continente della musica elettronica 1900-1986*, Edizioni La Biennale di Venezia - ERI, Venezia, 1986
14. Luigi Nono, *Post-Prae-Ludium per Donau*, Casa Ricordi 134668, Milano, 1987



# Sostituzione Incrociata della Fase nel Dominio di Scala

Antonio De Sena  
Dipartimento di Informatica  
Università degli Studi di Verona

## SOMMARIO

La trasformata di scala è uno strumento matematico che viene utilizzato in diversi ambiti. La si sfrutta quasi esclusivamente per scopi legati alla normalizzazione della dimensione di un fenomeno o per il riconoscimento di forme, grazie alla sua principale proprietà, ovvero l'invarianza alla scala. In questo articolo verranno presentati esperimenti in un altro ambito: l'elaborazione del suono (sound processing) ed in particolare nel campo degli effetti audio digitali. Scopo principale di questo lavoro è di investigare i comportamenti di tale trasformata in questo campo. Nell'articolo verrà descritto un particolare esperimento: la sostituzione incrociata della fase di scala. Dati due campioni sonori di strumenti musicali riproducenti la stessa nota ne verranno creati altri due ottenuti scambiando la fase degli originali. Questa operazione può essere interpretata come lo scambio delle informazioni legati alla dimensione (o alla forma) di un evento sonoro con quelle di un altro evento sonoro.

## Parole Chiave

Scale Domain; Scale Transform; Mellin Domain; Mellin Transform; Fast Mellin Transform; Mellin Pizzicato; Shape-Dimension Decoupling; Digital Audio Effects;

## INTRODUZIONE

In questo articolo verrà presentato un esperimento nel dominio di scala applicato a campioni sonori di strumenti musicali diversi. L'obiettivo di tale esperimento è l'esplorazione del dominio di scala. Con questi esperimenti si sonda la capacità di questo strumento nel campo degli effetti audio digitali e, più in generale, nel campo della teoria dei segnali ed elaborazione dei segnali audio.

Più in dettaglio, verrà sfruttata la capacità di disaccoppiamento dimensione-forma della trasformata di scala in modo tale da sostituire le informazioni legate alla dimensione (o alla forma) di un campione sonoro con quelle di un altro campione. Dal punto di vista pratico questo consiste nella sostituzione della fase di scala di un campione con quella di un altro campione. Questi esperimenti sono stati eseguiti avvalendosi dell'algoritmo veloce per il calcolo della trasformata di scala [5]. La versatilità e la velocità di questo algoritmo hanno permesso l'esecuzione di numerosi test. In questo articolo riporteremo i più interessanti.

I risultati di questo esperimento ci suggeriscono come si possano ottenere diverse situazioni. Ad esempio si possono ottenere campioni sonori che sono il "morphing" tra lo stru-

mento del primo campione e lo strumento del secondo, o campioni in cui "emergono" delle melodie, o, ancora, campioni che mostrano come certe proprietà sono legate solo alla fase come l'effetto vibrato.

Nella prima sezione si introdurrà la trasformata di scala e il relativo dominio. Verrà data la sua definizione e verranno illustrati i concetti e le proprietà base come l'invarianza alla scala. Nella sottosezione seguente verranno introdotti i concetti di periodicità in scala e la definizione di funzioni dilatocicliche. Infine verrà introdotto lo sviluppo in serie di scala con relativa interpretazione.

La seconda sezione introdurrà i primi esperimenti fatti in questo dominio in lavori precedenti, e le successive due sezioni spiegheranno che tipo di esperimento è stato condotto in questo lavoro e i relativi risultati ottenuti.

## LA TRASFORMATA DI SCALA E IL DOMINIO DI SCALA

La trasformata di scala può essere definita in diversi modi. Ad esempio, si può partire dalla trasformata di Mellin, vincolando il parametro complesso [5, 7], oppure si può utilizzare la Teoria Unificata dei Segnali, creando una trasformata invariante alla scala partendo dal gruppo abeliano  $(\mathbb{R}^+, \cdot)$  [1, 7] o, ancora, si può utilizzare il Metodo degli Operatori creando una trasformata partendo dall'operatore di scala [2, 7].

In ogni caso, la definizione della trasformata è sempre la stessa ed è esprimibile con la seguente equazione:

$$D_f(c) = \frac{1}{\sqrt{2\pi}} \int_0^\infty f(t) e^{(-jc - \frac{1}{2}) \ln t} dt. \quad (1)$$

La trasformata di scala inversa può essere così definita:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty D_f(c) e^{(jc - \frac{1}{2}) \ln t} dc. \quad (2)$$

La proprietà principale della trasformata di scala è l'invarianza di scala. Questo significa che se  $f$  è una funzione e  $g$  è una sua versione in scala, il modulo della trasformata delle due funzioni è identico. Una modifica in scala è una compressione o espansione dell'asse temporale del segnale originale che ne preserva l'energia. Quindi, una funzione  $g(t)$  può essere ottenuta con una modifica in scala da una funzione  $f(t)$  se  $g(t) = \sqrt{\alpha} f(\alpha t)$ , con  $\alpha \in \mathbb{R}^+$ . Quando  $\alpha < 1$  la modifica risulta essere una espansione in scala, quando  $\alpha > 1$  la modifica risulta essere una compressione in scala. Data una modifica in scala con parametro  $\alpha$ , la tras-

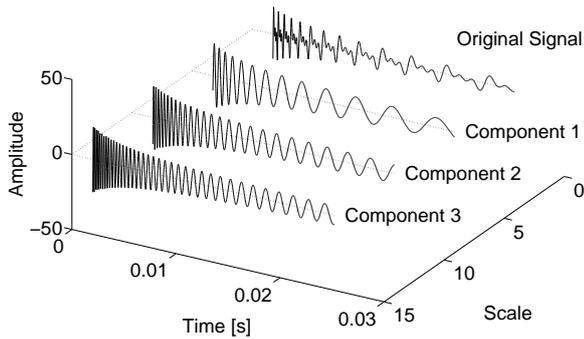


Figura 1. Esempio di decomposizione in scala (per semplicità, fase, energia e altre grandezze non sono state prese in considerazione).

formata del segnale originale e quella del segnale modificato sono in relazione. Questa relazione può essere espressa con la seguente equazione:

$$D_g(c) = \alpha^{jc} D_f(c). \quad (3)$$

In (3) la modifica in scala è presente sotto forma di fattore moltiplicativo ( $\alpha^{jc}$ ) e questo rappresenta una differenza solamente nella fase. Pertanto, il modulo del segnale originale è identico a quello del segnale modificato:

$$|D_g(c)| = |D_f(c)|. \quad (4)$$

### Interpretazione della Scala

Un interessante parallelo è quello tra la trasformata di Fourier e la trasformata di scala. In particolare, possiamo ridefinire i concetti presenti nella teoria legata alla trasformata di Fourier anche nella teoria legata alla trasformata di scala. Ad esempio, possiamo definire il concetto di peridicità di scala (*scale periodicity*, [5–7]) come segue: una funzione  $f(t)$  è periodica in scala con periodo  $\tau$  se soddisfa l'equazione  $f(t) = \sqrt{\tau} f(t\tau)$ , dove  $\tau = b/a$ , con  $a$  e  $b$  l'istante temporale iniziale e finale del periodo di scala.  $C_0 = 2\pi / \ln \tau$  è la scala fondamentale (“fundamental scale”) associata alla funzione periodica (in scala). In maniera analoga possiamo definire la serie in scala (“scale series”, [5–7]) e il teorema di Parseval per la scala [6]. Nella teoria legata alla trasformata di Fourier possiamo interpretare un segnale come somma infinita di seni e coseni. Analogamente, nella teoria legata alla trasformata di scala possiamo interpretare un segnale come somma infinita di funzioni periodiche smorzate in tempo e in frequenza (funzioni dilatocicliche, figura 1).

Quindi, invece di seni e coseni abbiamo le seguenti funzioni:

$$dd\sin(t) = \frac{\sin(c \ln t)}{\sqrt{t}} \quad (5)$$

$$dd\cos(t) = \frac{\cos(c \ln t)}{\sqrt{t}}, \quad (6)$$

dove  $c$  è la scala.

### ESPERIMENTI NEL DOMINIO DI SCALA

Il motivo di questi esperimenti è legato all'esplorazione del dominio di scala. In letteratura la trasformata di scala viene per lo più utilizzata allo scopo di ottenere normalizzazioni delle dimensioni di vari fenomeni. In questo lavoro (e precedenti, cfr. [3,4]) si prova ad esaminare la trasformata di scala, ed il relativo dominio, anche da altri punti di vista. In particolare, eseguendo dei test per capire come avviene la separazione dimensione-forma e come può essere sfruttata per ottenere effetti audio o estrarre informazioni. Inoltre, questo tipo di sperimentazioni portano a scoprire nuove idee o proprietà precedentemente ignorate o sconosciute (ad esempio il fatto che l'effetto vibrato viene preservato, come si vedrà nelle successive sezioni).

Le prime sperimentazioni eseguite da questi punti di vista sono state presentate alla conferenza DAFx nel 2004 [3]. Le sperimentazioni erano basate su filtri e modifiche al modulo e alla fase nel dominio di scala. I risultati ottenuti ci hanno permesso di capire meglio il funzionamento pratico da un punto di vista tempo-frequenza (oltre che scala) di queste elaborazioni. Dal comportamento di un particolare filtro sperimentato in quel lavoro è stato sviluppato un interessante effetto audio chiamato “Mellin pizzicator” [4]. Questo effetto simula l'effetto pizzicato su un qualsiasi campione audio, sfruttando appunto un filtro nel dominio di scala.

Da esperimenti condotti per esplorare la fase nel dominio di scala abbiamo ottenuto indicazioni su come modificare la dimensione percepita di un semplice evento sonoro.

Questi risultati sono stati utilizzati nel concorso Freesound (Freesound Competition<sup>1</sup>) per cambiare la dimensione percepita di campioni sonori di gocce d'acqua.

### SOSTITUZIONE INCROCIATA DELLA FASE

Questo lavoro è incentrato su una singola tipologia di esperimenti: la sostituzione incrociata di fase tra coppie di campioni sonori.

Gli obiettivi di tale esperimento sono molteplici. Innanzitutto si è cercato di approfondire le conoscenze in questo dominio. Inoltre si è tentato di sfruttare le capacità di disaccoppiamento dimensione-forma della rappresentazione modulo-fase nel dominio di scala per creare effetti audio digitali. Così facendo si potrà anche capire meglio il funzionamento di questo disaccoppiamento.

Nell'esperimento sono stati utilizzati diversi campioni sonori. Questi sono tutti campioni della stessa nota riprodotta con diversi strumenti musicali.

A partire da 10 campioni, l'elaborazione ha portato alla creazione di 90 nuovi campioni. Gli strumenti presi in esame sono stati il flauto alto, il sassofono alto, il clarinetto basso, il flauto basso, il clarinetto contralto in Mi $\flat$  (Eb clarinet), il corno francese, il sassofono soprano, il trombone tenore, la tromba e la viola suonata con archetto. I risultati sono stati analizzati ed interpretati tenendo presente i campioni originali e il funzionamento e le varie interpretazioni del dominio di scala.

<sup>1</sup><http://www.freesound.org/competition.php>

Tutti i campioni sono stati pre-elaborati (normalizzazione, taglio del silenzio, ecc.) per rendere più semplice l'analisi dei risultati ed evitare problemi di sincronizzazione (la trasformata di scala non è invariante allo spostamento temporale).

La trasformata di scala può essere calcolata al computer sfruttando l'algoritmo FMT (e l'algoritmo inverso IFMT) [5].

Questo algoritmo per il calcolo veloce delle trasformate della famiglia  $\beta$ -Mellin può essere utilizzato per calcolare la trasformata di scala dato che quest'ultima appartiene alla famiglia citata. La sua efficienza computazionale ( $O(n \ln^2 n)$ ) la rende la trasformata più adatta a questo tipo di calcoli ed esperimenti e la sostituzione incrociata della fase diventa un'operazione relativamente semplice e veloce.

## ESPERIMENTI, RISULTATI E COMMENTI

Con la tecnica descritta nella sezione precedente sono stati condotti numerosi esperimenti e sono stati ottenuti diversi risultati.

In generale i campioni sintetizzati mantengono molte caratteristiche dei campioni utilizzati per generali, in particolare modo vengono preservate molte caratteristiche del campione la cui fase viene conservata.

Inoltre, spesso si percepiscono componenti o modulanti smorzate in ampiezza e in frequenza (rispetto al tempo, cfr. eq. 5). Questo comportamento è dovuto al funzionamento della trasformata di scala e delle sue autofunzioni in un'interpretazione di sviluppo in serie di scala (cfr. sezione "Interpretazione della Scala"). Negli spettrogrammi queste componenti sono visibili sotto forma di curve iperboliche nel piano tempo-frequenza.

Alcuni campioni sintetizzati presentano un suono non melodioso, quasi rumoroso, ambientale. Questo tipo di campioni hanno tutti la fase di una tromba o di un trombone tenore. È difficile dare una spiegazione di tale comportamento. In ogni caso, da un punto di vista musicale, la loro fase mal si adatta a qualsiasi altro campione.

Quando si utilizza la fase di tromba, trombone tenore, clarinetto contralto o sassofono alto c'è una tendenza di quest'ultimi a dominare e a prevalere in maniera netta. Il timbro o comunque l'indistinguibile "firma" di tali strumenti è presente nella fase ed è molto più forte che negli altri strumenti.

Molto interessante è l'esperimento condotto con campioni sonori con effetto vibrato. Infatti, quando si utilizza la fase di un campione con effetto vibrato tale effetto verrà conservato intatto nel campione sintetizzato. Un esempio pratico è quello tra un campione di sax alto che non presenta il vibrato (figura 2) e un campione di flauto basso suonato con effetto vibrato (figura 3). Nello spettrogramma del campione sintetizzato (figura 4) si possono ancora notare le componenti legate all'effetto vibrato.

In alcuni casi i campioni sintetizzati conservano le caratteristiche di entrambi i campioni originali, oppure si può per-

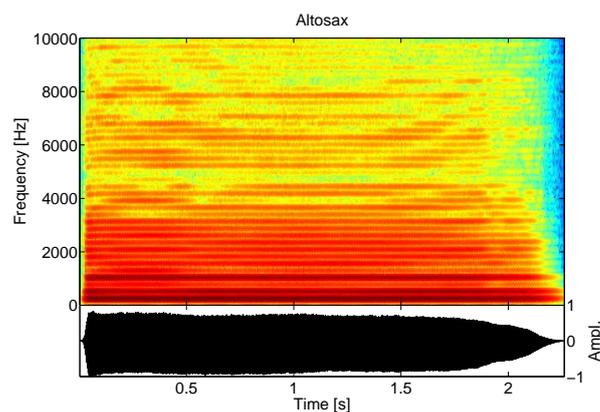


Figura 2. Spettrogramma di un campione sonoro prodotto con un sassofono alto.

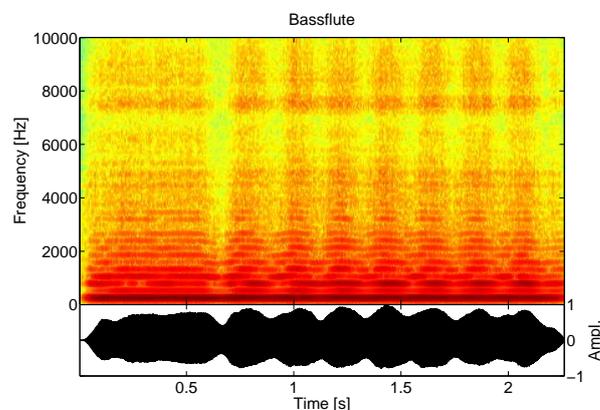


Figura 3. Spettrogramma di un campione sonoro prodotto con un flauto basso (con vibrato).

cepire il cambiamento (morphing) tra uno e l'altro al passare del tempo. Viola, sax alto, flauto alto, clarinetto basso, corno francese sono degli esempi. In particolare modo quando si lavora con la coppia flauto alto e sax alto o sax alto e viola o, ancora, corno francese e viola e altri. Questo comportamento può essere dovuto a similarità presenti nelle fasi dei campioni nella coppia da elaborare.

Qualche esperimento ha portato anche ad ottenere uno strano effetto di generazione di componenti in frequenza, simile all'effetto sonoro di un gong in cui le alte frequenze "emergono" al passare del tempo. Difficile dire qual'è la causa scatenante. Ad esempio, utilizzando il modulo del campione del clarinetto basso con la fase del sax alto o del sax soprano o della viola. Anche il clarinetto contralto quando si sostituisce la sua fase con quella del clarinetto basso e della viola produce un tal risultato.

Infine, un altro interessantissimo effetto è la generazione di

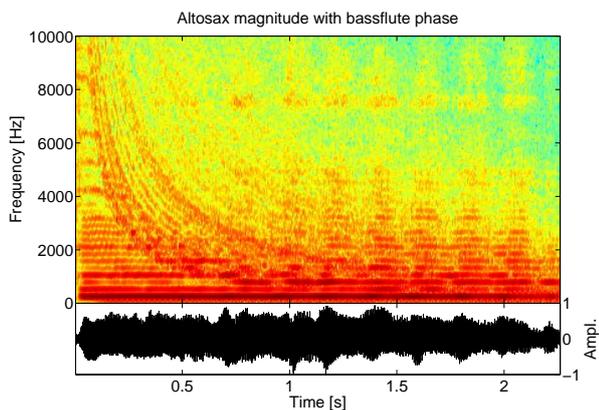


Figura 4. Spettrogramma di un campione sonoro sintetizzato utilizzando il modulo di scala di un sax alto e la fase di un flauto basso. Il vibrato del flauto è preservato.

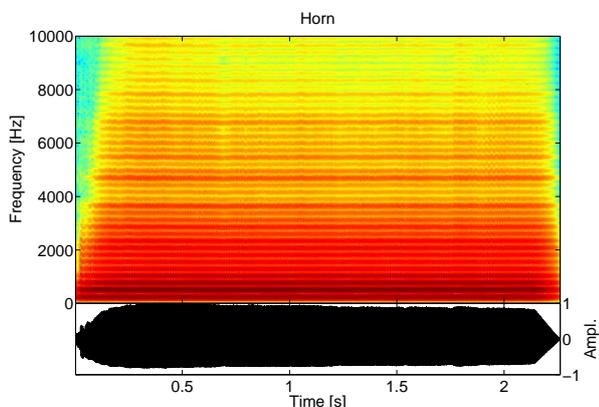


Figura 5. Spettrogramma di un campione sonoro prodotto con un corno francese.

una semplice melodia. In particolare utilizzando la fase del campione del clarinetto contralto. È possibile che tale effetto sia dovuto al fatto che la fase del campione in questione elimini delle frequenze (fondamentali e armoniche) in istanti temporali diversi. Ciò crea una semplice melodia. Un esempio (figura 7) può essere ottenuto con la coppia corno francese (figura 5) e clarinetto (figura 6).

## CONCLUSIONI

In questo articolo sono stati descritti i risultati ottenuti con diversi esperimenti di sostituzione incrociata della fase di scala di coppie di campioni sonori di strumenti musicali. L'obiettivo di tali esperimenti è di investigare sulla possibilità di utilizzare la trasformata di scala nel campo degli effetti audio digitali, e di studiare il comportamento della capacità di disaccoppiamento dimensione-forma di tale trasformata. I risultati suggeriscono come certe informazioni legate alla fase siano predominanti (vibrato, timbro o co-

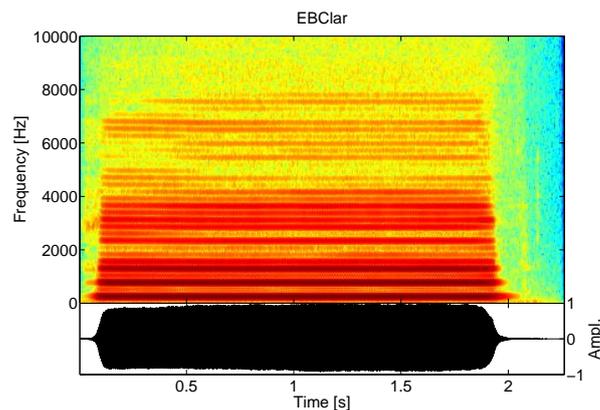


Figura 6. Spettrogramma di un campione sonoro prodotto con un clarinetto contralto.

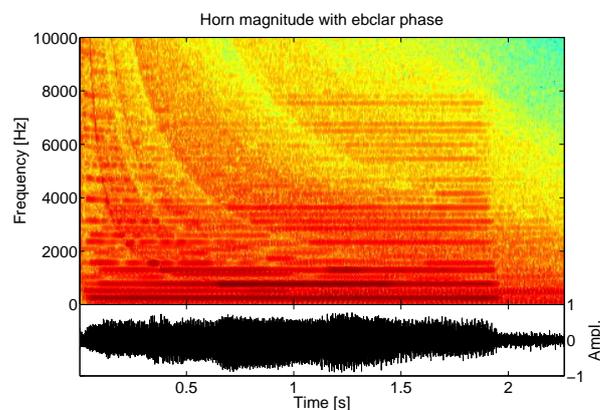


Figura 7. Spettrogramma di un campione sonoro sintetizzato utilizzando il modulo di scala di un corno francese e la fase di un clarinetto contralto. In questo caso, la fusione produce tagli delle armoniche e “nascono” delle melodie.

munque le caratteristiche dello strumento musicale). Inoltre alcune elaborazioni mostrano come la fase potrebbe essere utilizzata o manipolata per generare effetti sonori molto particolari, come la creazione di semplici melodie “estrate” dai campioni di partenza, o “esplosioni” delle frequenze tipo effetto “gong”.

## RIFERIMENTI

1. G. Cariolaro. La teoria unificata dei segnali. In G. Cariolaro, editor, *La teoria unificata dei segnali*. UTET Università, 1996.
2. L. Cohen. The scale representation. *IEEE Trans. on signal processing*, 41(12):3275–3291, December 1993.
3. A. De Sena and D. Rocchesso. A fast Mellin transform with applications in dafx. In *Proc. of the 7th Int.*

*Conference on Digital Audio Effects (DAFx'04)*, pages 65–69, October 2004. Naples, Italy, October 5-8.

4. A. De Sena and D. Rocchesso. The Mellin pizzicator. In *Proc. of the 9th Int. Conference on Digital Audio Effects (DAFx'06)*, pages 95–99, September 2006. Montreal, Quebec, Canada, September 18-20.
5. A. De Sena and D. Rocchesso. A fast Mellin and scale transform. *EURASIP Journal on Advances in Signal Processing*, 2007:Article ID 89170, 9 pages, 2007. doi:10.1155/2007/89170.
6. H. Sundaram, S. Joshi, and R. Bhatt. Scale periodicity and its sampling theorem. *IEEE Trans. on signal processing*, 45(7):1862–1864, July 1997.
7. A. De Sena. A Computational Framework for Sound Analysis with the Mellin and Scale Transform. *Ph.D. Thesis*, <http://mordente.sci.univr.it/~antonio/univr/phdthesis.pdf>, February 2008.



# La Computer Vision per l'annotazione automatica di documenti audio

**Lauro Snidaro**

Laboratorio AVIRES

Dipartimento di Matematica e Informatica

Università degli Studi di Udine

**Sergio Canazza**

Laboratorio AVIRES

Dipartimento di Scienze Storiche e Documentarie

Università degli Studi di Udine

## SOMMARIO

Durante il processo di conservazione attiva dei documenti audio è importante non trascurare l'informazione contestuale presente nel supporto sonoro. In particolare, le riprese video del trasferimento A/D del segnale audio (disco fonografico in rotazione, nastro magnetico in scorrimento, ecc.) possono essere un mezzo per memorizzare le informazioni sullo stato del supporto. Nel caso di archivi audio di rilevanti dimensioni, è economicamente improponibile estrarre manualmente le informazioni di interesse dai filmati. Viene presentato, in questa sede, un sistema in grado di individuare automaticamente eventi di interesse presenti nelle registrazioni video di documenti audio.

## Parole Chiave

Annotazione automatica; Visione Artificiale; Conservazione attiva; Documenti audio.

## INTRODUZIONE

Dalla carta usata nel 1860 (a cui risale *Au Clair de la Lune*, la più antica registrazione audio di cui ci è stato tramandato il supporto, effettuata da Édouard-Léon Scott de Martinville mediante fonografo<sup>1</sup>), sino al moderno Blu-ray Disc, il campo delle memorie audio costituisce una vera Torre di Babele: in 150 anni sono stati prodotti un numero enorme di supporti (analogici e digitali) incompatibili tra loro. Sono chiare l'urgenza, l'importanza e la complessità della loro conservazione, che può essere articolata in: a) passiva<sup>2</sup> (difesa del supporto dagli agenti ambientali, senza alterarne la struttura) e b) attiva (trasferimento dei dati nei nuovi media). Poiché le memorie audio sono caratterizzate da un'aspettativa di vita relativamente bassa – se confrontata con quella di altri monumenti – la conservazione passiva risulta insufficiente. Inoltre, i beni culturali musicali non hanno la possibilità di

<sup>1</sup> Brevettato il 25 marzo 1857. Utilizzato come strumento di laboratorio per studi di acustica, con funzione simile all'oscilloscopio, era in grado di trascrivere graficamente le onde sonore su un mezzo visibile (vetro annerito o rotolo di carta), ma non c'era modo di riprodurre il suono registrato.

<sup>2</sup> La conservazione passiva si suddivide a sua volta in indiretta – che non comporta il coinvolgimento fisico del supporto sonoro – e diretta, nella quale il supporto viene trattato, senza comunque alterarne struttura e composizione. Nella conservazione passiva indiretta rientrano: la prevenzione ambientale (che si esplica attraverso il controllo dei parametri ambientali che sono, in ordine decrescente di pericolosità per i nastri magnetici: umidità relativa, temperatura, inquinamento, luce), la formazione del personale addetto alla conservazione, l'educazione dell'utente. La conservazione passiva diretta comprende gli interventi di: realizzazione di custodie di protezione; spolveratura delle raccolte; disinfestazione degli archivi con gas inerti; periodico svolgimento e riavvolgimento dei nastri magnetici.

vicariare o rigenerare – partendo dal segnale audio – l'informazione perduta: è quindi di fondamentale importanza trasferire nel dominio digitale (oltre al segnale audio) tutta l'informazione contestuale presente nel supporto. In questo senso, molti archivi inseriscono nella copia conservativa le riprese video effettuate durante il trasferimento A/D del segnale audio (scorrimento del nastro, rotazione del disco o del cilindro di cera). Il video, infatti, offre informazioni sulle alterazioni (intenzionali o meno) e sulle corrotture del supporto. Questi dati sono necessari a fini archivistici, per studi musicologici e nelle operazioni di restauro audio [1], [2], [3].

Nel caso di archivi di medie/grosse dimensioni, è economicamente improponibile estrarre manualmente le informazioni di interesse dai filmati. Diventa quindi prezioso poter contare su di un sistema in grado di individuare automaticamente eventi di interesse presenti nelle registrazioni video di documenti audio.

Gli autori hanno utilizzato strumenti sviluppati nell'ambito della visione artificiale per rilevare automaticamente discontinuità presenti nel nastro magnetico. In particolare, sono state impiegate tecniche di *background subtraction* con impostazione automatica della soglia [4], ottimizzata al fine di individuare la presenza di nastro *leader*. In questo modo sono annotati automaticamente gli istanti temporali d'inizio e di fine di ogni tratto di nastro magnetico (registrabile), rispetto al nastro di plastica utilizzato durante la fase di montaggio dai compositori e dai tecnici. Nel caso il nastro coprisse solo una percentuale dell'immagine (caso molto comune), è possibile impostare una *regione di interesse* (*Region Of Interest, ROI*), in modo da scartare durante l'elaborazione i dettagli non rilevanti (testina di lettura, sfondo, ecc.). Questo approccio è mutuato dalle tecniche utilizzate per il rilevamento dei cambi di scena nel campo dell'annotazione automatica di sequenze video [5].

A questa tecnica di base sono poi stati accoppiati altri algoritmi al fine di rilevare specifiche alterazioni sul nastro (intenzionali o meno) nelle sotto-regioni selezionate (ossia: in cui il sistema ha rilevato una discontinuità). In questo modo vengono automaticamente annotati gli istanti temporali in cui nel nastro compaiono: a) giunte; b) segni; c) perdite di pasta magnetica. Nel caso in cui al video fosse sincronizzato il segnale audio, ognuna di queste discontinuità può essere facilmente allineata al corrispondente evento sonoro.

Nel caso di videoregistrazioni di dischi fonografici in rotazione, il sistema rileva l'evoluzione temporale della posizione del braccio. Da questa funzione, può essere automaticamente calcolata (conoscendo la velocità di rotazione usata per leggere il disco) la variazione del *pitch* nel segnale audio dovuta alle deformazioni presenti nel supporto

fonografico rispetto al suo piano (dischi cosiddetti *imbarcati*), oppure a difetti di bilanciamento del piatto. Questo dato è di estrema utilità nel caso si voglia procedere, in fase di restauro audio, a riconoscere le oscillazioni del *pitch* dovute a: a) difetti del sistema di registrazione; b) deterioramento del supporto; c) imperfezioni del sistema di lettura.

L'approccio sperimentato si basa sull'analisi dello spostamento delle *features* di Lucas-Kanade rilevate sul braccio del giradischi. La tecnica sviluppata consiste in primo luogo nel *clustering* delle *features* non statiche all'interno di una ROI prefissata; viene quindi valutata la variazione di inclinazione del *cluster* nel tempo determinando quindi la variazione di *pitch*.

## DISCONTINUITA' PRESENTI SU NASTRI MAGNETICI

La ripresa dello svolgimento di un nastro, durante il trasferimento dei dati audio in esso contenuti su supporto digitale, permette di raccogliere importanti informazioni utili allo studio del segnale registrato.

Una sequenza video costituisce una ricca fonte di informazioni sullo stato del nastro. In particolare, è possibile evincere alterazioni quali corrottele del supporto o interventi di editing (giunte, segni). Queste informazioni rappresentano preziosi metadati con cui annotare la copia riversata su supporto digitale. Essi possono essere infatti associati all'informazione acustica fornendo importanti elementi per un'analisi della storia del documento audio originale e per il suo restauro.

L'estrazione manuale di metadati mediante ispezione di sequenze video da parte di un operatore (annotazione manuale) non è concepibile per archivi di medie/grosse dimensioni [6]. In questo lavoro vengono quindi proposte delle tecniche di visione artificiale in grado di elaborare automaticamente le sequenze video ottenute dalle riprese di documenti audio su nastro durante la loro riproduzione. Queste tecniche sono in grado di rilevare le discontinuità presenti sul nastro e di riconoscere elementi notevoli quali giunte.

Negli esperimenti effettuati, la tecnica di base impiegata per rilevare discontinuità in una sequenza video è la nota *background subtraction*: il fotogramma corrente  $I_t$  all'istante  $t$  viene confrontato con uno di riferimento  $I_{bck}$  (background), acquisito in un istante temporale precedente. Il confronto, effettuato mediante differenza in valore assoluto fra matrici, permette di ottenere l'immagine delle differenze  $D_t = |I_t - I_{bck}|$ . All'immagine delle differenze  $D_t$  viene applicata una soglia  $th$  ottenendo quindi un'immagine binaria  $B_t$ . Un'immagine binaria può rappresentare solo due colori, convenzionalmente il bianco e il nero. Con il nero vengono in genere indicate le regioni dell'immagine che non hanno subito cambiamenti, mentre in bianco vengono indicate le differenze sostanziali fra le due immagini oggetto del confronto, quelle cioè che hanno superato la soglia  $th$ . Negli esperimenti qui proposti è stato impiegato l'algoritmo di *sogliatura automatica* basato sui numeri di Eulero i cui dettagli sono riportati in [4]. Questo algoritmo calcola automaticamente per ogni frame la soglia  $th$  ottimale in base ad un criterio legato alla connettività delle componenti connesse presenti in  $B_t$ . E' stato inoltre applicato un passo di filtraggio come ulteriore elaborazione. In particolare, è stato impiegato un semplice *voting* per rimuovere piccole componenti connesse spurie, dovute a rumore o a improvvise variazioni di luminosità locale, nella matrice binaria  $B_t$  al fine di migliorarne la qualità.

In Figura 1e in Figura 2 sono visibili alcuni fotogrammi estratti da sequenze video diverse su cui sono stati condotti gli esperimenti. Nella prima colonna sono stati riportati i frame

sorgente, mentre nella seconda sono state riportate le immagini binarie risultato delle procedure di *change detection* e di *sogliatura*. Per entrambe le sequenze, il fotogramma di background era stato acquisito precedentemente ai fotogrammi illustrati.

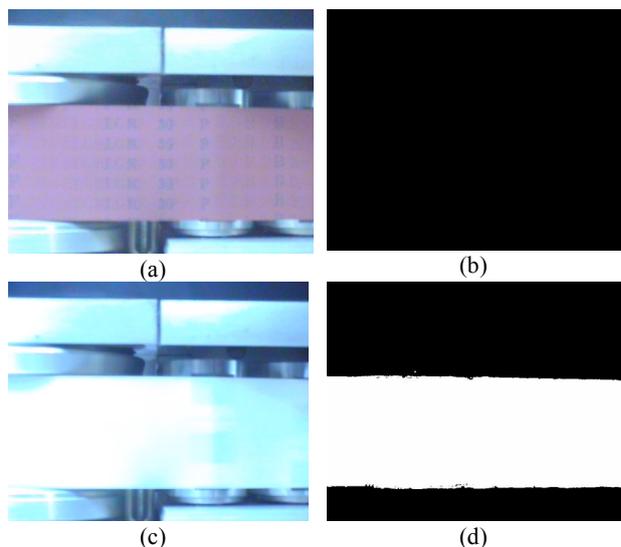


Figura 1. Nella prima colonna sono presenti due immagini estratte dal video di un nastro in fase di riproduzione. Nella seconda colonna sono visibili le corrispondenti elaborazioni. (a) Dorso del nastro magnetico e (b) nessuna anomalia rilevata. Il nastro *header* in (c) viene rilevato come discontinuità (d).

La Figura 1 evidenzia come la tecnica della *change detection* permetta di rilevare discontinuità nella composizione del nastro. La Figura 1(b) è completamente nera poiché non è stata rilevata nessuna differenza significativa tra il fotogramma corrente (a) e quello di background. In (d) viene invece rilevata una differenza consistente (regione bianca) fra l'immagine di riferimento e quella corrente (c). In questo caso, è evidente come la sezione *header* del nastro viene riconosciuta con precisione.

Le procedure sopra descritte (si consulti [4] per ulteriori dettagli realizzativi) forniscono quindi un mezzo efficace per rilevare variazioni nella composizione del nastro. Il conteggio dei pixel bianchi nelle immagini binarie e una soglia fissata a priori sulla percentuale di pixel cambiati rispetto alla *regione di interesse* (*Region Of Interest*, ROI) permettono di decidere se si è in presenza di una discontinuità o meno. La ROI può essere impostata per focalizzare l'attenzione dei vari algoritmi solo su una sottoregione dell'immagine. Come si può vedere nei frame sorgente della Figura 1, il nastro occupa circa il 50% dell'immagine; mentre altri dettagli come le testine del lettore non sono rilevanti ai fini dell'elaborazione e dovrebbero essere eliminati impostando la ROI sulla regione corrispondente al solo nastro.

L'approccio appena descritto è molto simile alle tecniche di *scene cut detection* per l'annotazione automatica di filmati televisivi o cinematografici [5], [6].

La Figura 2 illustra invece come altri tipi di informazioni possano essere estratte mediante l'analisi delle riprese di nastri in svolgimento. I passi di elaborazione di base sono gli stessi dell'esperimento precedente, ma in questo caso sono richiesti dei procedimenti aggiuntivi per rilevare giunte o discontinuità specifiche.

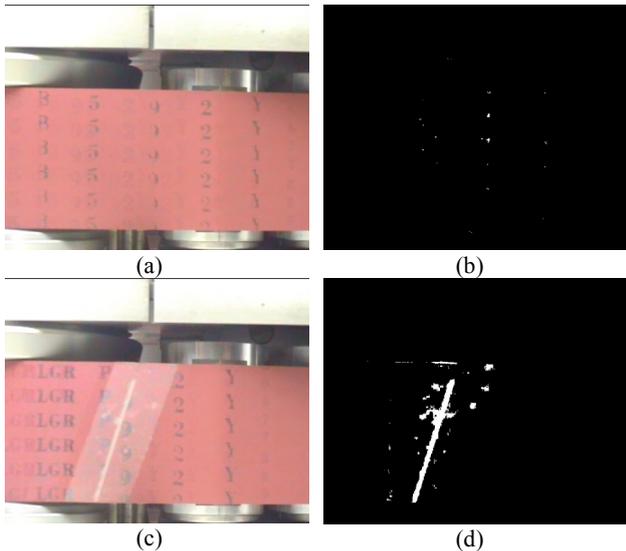


Figura 2. Discontinuità rilevate su un nastro in fase di riproduzione. (a) Dorso del nastro magnetico e (b) nessuna anomalia rilevata. La giunta visibile in (c) viene rilevata come discontinuità dal sistema (d).

Il frame (b) è quasi completamente nero ad indicare che non vi sono cambiamenti significativi fra il frame corrente (a) e l'immagine di riferimento. Le piccole componenti connesse osservabili in (b) rappresentano delle piccole variazioni dovute allo scorrimento dei caratteri alfanumerici visibili in (a). Questi piccoli cambiamenti (numero di pixel bianchi) non superano la soglia di attenzione impostata a priori dall'operatore e quindi non rappresentano un'anomalia del nastro. Il fotogramma (d) evidenzia come la giunta venga rilevata dalla change detection. In questo caso, il numero di pixel cambiati passa la soglia di attenzione. Tuttavia, al fine di discriminare il tipo di anomalia è necessario un passo ulteriore di elaborazione. L'individuazione del segmento in (d) corrispondente alla giunta può essere effettuata mediante l'applicazione della trasformata di Hough all'immagine binaria [7]. Queste informazioni – opportunamente allineate col segnale audio – possono essere annotate automaticamente dal sistema direttamente nella copia conservativa.

E' utile notare che le tecniche impiegate, in particolare la *sogliatura automatica* basata sui numeri di Eulero, hanno consentito di analizzare sequenze video acquisite senza particolari accorgimenti volti a creare delle condizioni normalizzate. In particolare, si può osservare come i fotogrammi in Figura 1 e in Figura 2 siano stati acquisiti in condizioni di illuminazione sensibilmente differenti. Questo facilita di molto le videoregistrazioni stesse che possono essere effettuate dagli archivi senza vincoli particolari sulle condizioni di illuminazione, mediante hardware *entry level* e senza l'utilizzo di personale addestrato a riprese video professionali.

## DEFORMAZIONI DEI SUPPORTI FONOGRAFICI

Le deformazioni dei supporti fonografici (dischi *imbarcati* o sbilanciati) causano una sensibile variazione del *pitch* del segnale audio (la Figura 3 mostra un esempio di disco afflitto da un'importante deformazione). Questo è dovuto al movimento ondulatorio a cui è soggetto il braccio del giradischi (e quindi la puntina).



Figura 3. Disco fonografico afflitto da un'importante deformazione rispetto al suo piano. La corruttela è stata causata da un'escursione della temperatura (il disco è stato esposto al sole per 4 ore, all'interno di un'autovettura in sosta, a una temperatura di oltre 50°C).

Poiché le caratteristiche del moto oscillatorio possono essere messe in relazione con la variazione del *pitch*, esse costituiscono importanti metadati utili al restauro del segnale audio. Si propone in questa sede di utilizzare tecniche di visione artificiale per analizzare e annotare automaticamente le videoregistrazioni di dischi fonografici in rotazione.

In questo esperimento è stato impiegato un approccio diverso rispetto a quello basato sulla *change detection* visto precedentemente. In questo caso è stato utilizzato l'algoritmo di *features tracking* comunemente noto come Lucas-Kanade *tracker* [8]. L'algoritmo individua dei punti notevoli nell'immagine (*features*) che possono essere impiegati per confrontare due frame successivi e valutare gli spostamenti fra uno e l'altro. La tecnica, inizialmente concepita per l'allineamento di immagini, viene qui utilizzata nella sua implementazione come *features tracker*, che è in grado di tenere traccia degli spostamenti delle *features* da un frame al successivo.

Nella Figura 4 sono stati riportati alcuni fotogrammi tratti da una delle sequenze impiegate negli esperimenti. La prima colonna riporta (a) il punto più elevato e (d) quello più basso dell'oscillazione del braccio. Nella seconda colonna si possono osservare le rispettive elaborazioni in cui sono visibili le feature individuate sul braccio del giradischi.

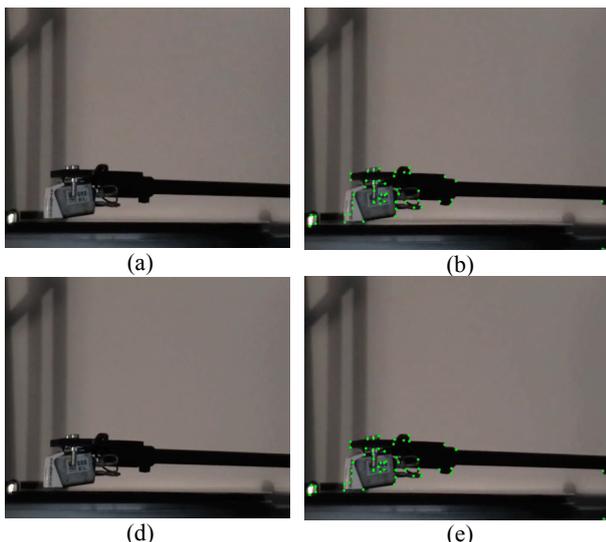


Figura 4. Immagini sorgente e relative elaborazioni tratte da una ripresa video del braccio di un giradischi durante la riproduzione. Il supporto sul piatto è deformato e ciò causa un'oscillazione del braccio. Nelle immagini sono state evidenziate: (a) posizione più bassa del braccio durante l'oscillazione e (d) posizione più alta. Nelle immagini (b),(e) si possono notare le *features* di Lukas-Kanade rilevate sulla testa del braccio.

Durante gli esperimenti il *tracker* di Lucas-Kanade ha correttamente mantenuto traccia delle *features* rilevate nel primo frame delle sequenze video utilizzate. Il *tracker* ha quindi permesso di registrare gli spostamenti delle *features* durante le oscillazioni del braccio del giradischi dovute alla riproduzione di supporti deformati.

Nella Figura 5 è stata riportata l'evoluzione temporale della coordinata *y* di una *feature* localizzata sul braccio. Sull'asse delle ordinate è riportato il numero di *frames*, mentre le ascisse rappresentano la posizione in pixel sul piano immagine. E' chiaramente visibile l'andamento oscillatorio. I fotogrammi (a) e (d) nella Figura 4 distano 29 frame, dato riscontrabile nelle oscillazioni del grafico di Figura 5.

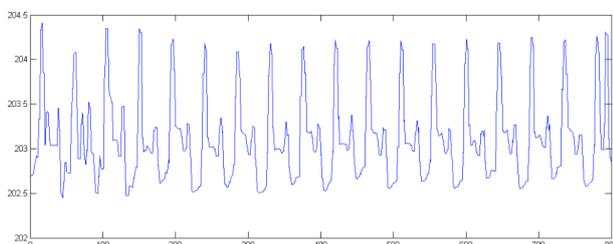


Figura 5. Evoluzione temporale della coordinata *y* di una *feature* di Lucas-Kanade posizionata sul braccio. Le evidenti oscillazioni indicano la presenza di un supporto imperfetto.

## CONCLUSIONI

Nel caso di archivi sonori di medie/grosse dimensioni, è economicamente improponibile estrarre manualmente le informazioni di interesse dai filmati effettuati per conservare l'informazione contestuale dai supporti fonografici. Diventa quindi prezioso poter contare su di un sistema in grado di annotare automaticamente eventi di interesse presenti nelle registrazioni video dei dischi in rotazione. Questi metadati possono essere vantaggiosamente utilizzati in fase di restauro del segnale audio per riconoscere le alterazioni inserite durante la fase di registrazione/produzione del supporto, quelle dovute

al deterioramento del supporto o alle imperfezioni del sistema di lettura.

Le applicazioni sono state sviluppate in C++ e sono in grado di operare in tempo reale su segnali video a risoluzione di 320x240 *pixels*, e utilizzando un elaboratore dotato di (singolo) processore a 3 GHz. Le sequenze video possono essere acquisite mediante un *camcorder* a risoluzione PAL e successivamente scalate e compresse in DivX a qualità medio-alta.

Grazie alle basse richieste computazionali e al fatto che non viene richiesto hardware specifico, si ritiene che il sistema possa essere un pratico ausilio per gli archivi di documenti sonori e le audio Digital Libraries, che possono in questo modo annotare automaticamente l'informazione contestuale dei supporti fonografici utilizzando personale con modesto addestramento.

## RIFERIMENTI

- [1] Orcalli, A. (2006). Orientamenti ai documenti sonori. In Canazza, S. e Casadei Turronei Monti, M. (a cura di), *Rimediazione dei documenti sonori*, pp. 15-94, Udine: Forum.
- [2] Canazza, S. (2007). Note sulla conservazione attiva dei documenti sonori su disco. In Atti del Convegno annuale del Laboratorio per la Divulgazione Musicale (Ladimus). *Il suono riprodotto: storia, tecnica e cultura di una rivoluzione del Novecento*, Torino: EDT.
- [3] Canazza, S. (2006). Conservazione attiva e restauro audio dei 78 giri. Un caso di studio: Eternamente, In Canazza, S. e Casadei Turronei Monti, M. (a cura di), *Rimediazione dei documenti sonori*, pp. 695-715, Udine: Forum.
- [4] L. Snidaro and G.L. Foresti, "Real-time thresholding with Euler numbers", *Pattern Recognition Letters*, Vol. 24, n. 9-10, pp. 1533-1544, June 2003.
- [5] Y. Liu, D. Zhang, G. Lu, and W.Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262-282, 2007.
- [6] A. Del Bimbo, Visual information retrieval, Morgan Kaufmann, San Francisco, CA, 1999.
- [7] N. Aggarwal and W.C. Karl, "Line detection in images through regularized hough transform" *IEEE Transactions on Image Processing*, Vol.15, n°3, March 2006.
- [8] J. Shi and C. Tomasi, "Good Features to Track", IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WS, June 1994, pages 593-600.

# Analisi del comportamento di fruitori di sistemi di ascolto attivo: il caso dell'*Esploratore d'Orchestra*

Francesca Cavallero, Antonio Camurri, Corrado Canepa, Gualtiero Volpe

Casa Paganini - InfoMus Lab  
DIST – Università degli Studi di Genova

## SOMMARIO

Il presente lavoro si inserisce nell'ambito della ricerca su modelli di ascolto attivo in ambienti virtuali interattivi. Obiettivi di questo elaborato sono l'analisi e la valutazione di un nuovo paradigma di ascolto attivo, alla base dell'installazione "Esploratore d'Orchestra" realizzata a Casa Paganini da InfoMus Lab in occasione della mostra "Cimenti di Invenzione e Armonia – Metamorfosi del Senso", durante il Festival della Scienza 2007 (Camurri, Canepa, Volpe 2007b). A partire da strategie di monitoraggio, rilevamento ed elaborazione dei dati relativi a soggetti/utilizzatori dell'installazione, viene indagato, da un punto di vista multidisciplinare, coinvolgendo anche aspetti di *interaction design*, il concetto di fruizione attiva di tale ambiente virtuale. La ricerca è stata svolta nell'ambito della tesi di laurea specialistica in Scienze dello Spettacolo (Cavallero 2008), Facoltà di Lettere e Filosofia, Università degli Studi di Genova.

Attualmente, lo studio è focalizzato sulla interazione tra soggetti/fruitori con particolare riferimento a modelli di interazione sociale e di *entrainment* (Camurri et al 2008). La ricerca è parzialmente supportata dal progetto UE 7° Programma Quadro ICT SAME ([www.sameproject.eu](http://www.sameproject.eu)).

## Parole Chiave

Informatica Musicale; Sound and Music Computing; Spazio Attivo; Ascolto Attivo; Interaction Design; Ambienti Multimodali Interattivi.

## INTRODUZIONE

Nello sviluppo di sistemi multimediali interattivi è opportuno tenere presenti due aspetti: il ruolo e l'utilizzo dell'ambiente in quanto superficie fisica che muta in spazio attivo, in grado di osservare il fruitore e "dialogare" con lui, ad esempio attraverso la produzione di contenuti visuali e sonori; in secondo luogo, l'analisi dei gesti compiuti dagli utenti, ponendo l'accento sulle modalità di esecuzione del movimento e sul rilevamento di particolari caratteristiche del medesimo. Un Ambiente Multimodale Interattivo, pertanto, conferisce all'utente la possibilità di poter variare dinamicamente e gradualmente le proprietà dell'ambiente stesso, come conseguenza del suo comportamento, in un certo senso "plasmandolo". Oltre alla capacità di tali dispositivi di "osservare" caratteristiche del movimento generali (con riferimento al gesto e all'espressività), risultano pertanto evidenti l'importanza e la poliedricità di ruolo assunta dal progettista di un AMI come l'*Esploratore d'Orchestra*, in equilibrio fra coreografo, regista, compositore.

L'*Esploratore d'Orchestra* (Camurri et al. 2007a) è un ambiente multimodale interattivo in grado di rilevare, (attualmente attraverso una telecamera, in futuro tramite

sistemi mobili), caratteristiche a più livelli di dettaglio e di sensibilità nel comportamento degli utenti, ed in particolare nei loro gesti e nei loro movimenti. Tali informazioni sono successivamente utilizzate nella gestione in tempo reale dei processi di manipolazione e di *rendering* di musica preregistrata, grazie ad un'applicazione basata sulla piattaforma software EyesWeb XMI per l'analisi multimodale. Con "esperienza di ascolto attivo" s'intende la possibilità, per lo spettatore, di operare interattivamente con un contenuto musicale, modificandolo e modulandolo in tempo reale durante

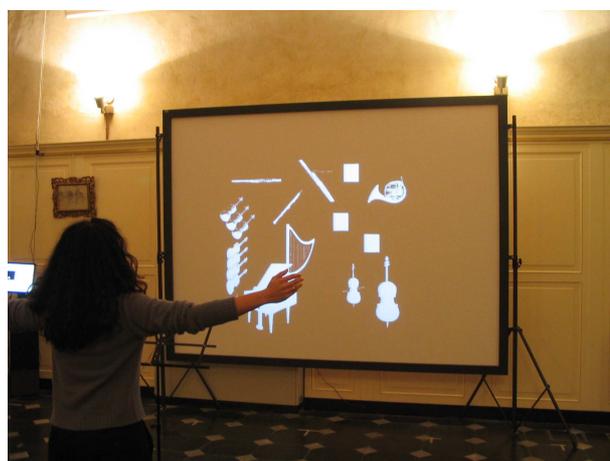
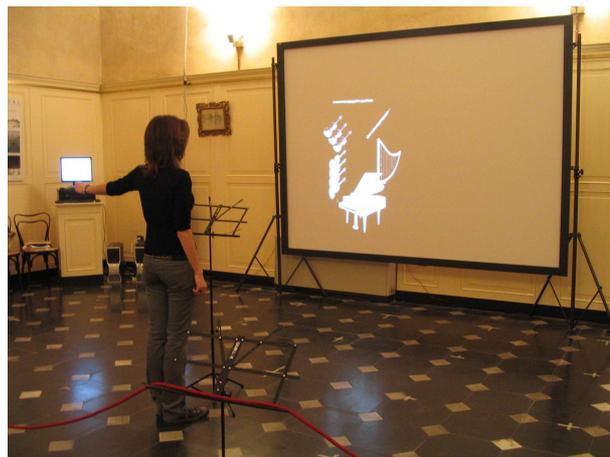


Figura 1-2. Il fruitore si trova nella posizione del "Direttore d'Orchestra": grazie al movimento delle braccia, attiva una o l'altra parte dell'orchestra virtuale.

l'ascolto. Le singole sezioni dell'orchestra virtuale sono disposte in diverse regioni dello spazio: esplorandolo, si attivano i "domini" dei diversi strumenti (o di sezioni strumentali) man mano che questi sono raggiunti dall'utente. L'esplorazione di un brano musicale all'interno di una dimensione spaziale (portando quindi la dimensione del tempo a convivere strettamente con essa), altera in modo radicale il brano stesso, perché cambia il modo in cui esso è percepito nella dimensione temporale.

Il lavoro include analisi, anche di tipo statistico, del comportamento e delle diverse modalità d'interazione concretizzate dagli utenti all'interno dell'installazione, alla luce di studi antropologici ed etnografici (ad es., Veron e Levasseur 1983): coerentemente, sono stati scelti fattori determinanti (percorso seguito all'interno dello spazio attivo, riferimento principale – visivo o sonoro – assunto a guida dell'esplorazione, tempo di permanenza, correttezza d'interazione rispetto ad un modello "standard" di fruizione) posti poi in relazione a parametri quali ad esempio il sesso e l'età dei visitatori. L'analisi ha fornito informazioni circa il "potenziale interattivo" dei soggetti coinvolti, e rivelando, nello specifico, come alcune fasce d'età presentino una maggiore attitudine a relazionarsi con contesti virtuali/interattivi. Inoltre, è stato possibile rilevare i diversi atteggiamenti assunti dai visitatori osservati rispetto alle due tipologie di feedback (sonoro e visivo) proposte dall'installazione: in particolare, anche con riferimento specifico a distinte fasce d'età, è risultato evidente come lo stimolo visivo abbia una certa preponderanza causale nell'influenzare l'interazione dei soggetti.

A partire da un approccio "ecologico" (Gibson 1979), secondo il quale sarebbe l'ambiente, sulla base di caratteristiche proprie (*affordances*), ad emettere gli stimoli necessari alla sua stessa fruizione, sono inclusi nella ricerca anche aspetti di *interaction design* (Preece, Rogers, Sharp 2002), il cui fine è proprio creare la migliore interazione possibile fra ambiente e fruitore.

## ELEMENTI DI INTERACTION DESIGN

In un AMI risulta di fondamentale importanza l'interazione sussistente fra l'utente/fruitore e l'ambiente virtuale di riferimento: questo è vero anche in relazione alla funzionalità del dispositivo, delineando una prospettiva di studio che si inserisce nell'ambito disciplinare dell'Interaction Design.

All'interno di installazioni come l'*Esploratore d'Orchestra*, ad esempio, la questione assume grande rilevanza, perché sia a livello dell'ambiente virtuale in sé, sia a livello di interazione concreta col dispositivo tecnologico, la figura dell'utente si rivela centrale, dando per scontato che in quanto interfaccia è necessariamente sottostato al processo di *ideazione – progettazione – sviluppo – implementazione* all'interno del quale si inseriscono le ordinarie attività di interaction design.

A questo proposito, esistono tre caratteristiche "chiave" del "processo" di interaction design che aiuteranno a meglio comprenderne le implicazioni sul piano concreto:

*Focalizzazione sull'utente*: consistente in spunti per valutazioni e feedback da parte dell'utente stesso;

*Obiettivi di usabilità specifica e di esperienza dell'utente*: dovrebbero essere identificati, chiaramente documentati e accettati prima dell'avvio del progetto. Aiutano gli sviluppatori a scegliere fra diverse alternative di design e ad eseguire verifiche sui progressi di sviluppo del prodotto;

*Iterazione*: permette modifiche e rifiniture basate su continui feedback. Nel momento in cui utenti e designers cominciano a discutere requisiti, bisogni, speranze ed aspirazioni allora emergeranno differenti idee su ciò che è necessario, su ciò che sarà d'aiuto, su ciò che potrà essere realizzato. È fondamentale, per l'iterazione, che sussista un flusso informativo reciproco e continuo che permetta "modifiche in corso d'opera". Per quanto designers ed utenti possano

ritenere, rispettivamente, buona o chiara la loro visione del prodotto, spesso si rivelerà necessario rivedere le idee alla luce dei feedback ottenuti. Uno studio come quello in oggetto nasce proprio con l'intenzione di validare e verificare attraverso l'osservazione aspetti propri dell'installazione nelle condizioni più "ecologiche" possibili, ossia quando concretamente essa viene fruita dagli utenti, durante una mostra: una situazione evidentemente molto diversa da quella "asettica" creata in laboratorio, ad esempio.

Ciò che desta qui particolare interesse è proprio l'attenzione focalizzata sull'utente: chi è in realtà l'utente che fruirà, poniamo, dell'installazione ancora in fase progettuale, e quali sono le sue caratteristiche?

Nello specifico dell'ambiente virtuale, è evidente come non sia possibile prescindere da caratteristiche del soggetto che riguardano percezione e scienze cognitive; tuttavia, più in generale, ci sono alcuni aspetti concernenti l'utente che, proprio a livello cognitivo, hanno una grande influenza sui processi di interaction design: **attenzione** (si tratta di una "procedura" mentale di selezione a partire da una gamma di opzioni disponibili: l'attenzione coinvolge il nostro sistema visivo e/o uditivo); **percezione e recognition** (le combinazioni di media diversi devono anche essere progettate per permettere agli utenti di riconoscere l'informazione composita rappresentata: l'uso integrato di suono ed animazioni ha bisogno di essere coordinato così come accade in una sequenza logica e tenendo conto della multimodalità); **memoria** (coinvolge i richiami ai vari domini della conoscenza che ci permettono di agire in modo appropriato); **apprendimento** (si può considerare o nei termini di come usare un'applicazione *computer-based*, oppure usare un'applicazione *computer-based* per apprendere un determinato argomento: uno dei principali vantaggi, nelle tecnologie interattive, è che si possono prevedere modi alternativi di rappresentazione ed interazione con l'informazione che non sono possibili con i mezzi tradizionali); **"reading speaking and listening"**; **risolvere problemi, pianificare, ragionare, prendere decisioni**.

Per quanto concerne l'*Esploratore*, è l'attenzione dell'utente ad essere sollecitata nel momento in cui esplora lo spazio, all'interno del quale particolari riceve stimoli di tipo sonoro e visivo, che ad esempio consentono di riconoscere un particolare strumento o sezione; sulla base dell'esperienza che matura all'interno dell'installazione, l'utente impara a correlare suono e strumento, scopre relazioni tra le voci e più in generale elementi di orchestrazione e ritmici, può decidere se fruire dell'ambiente in modalità d'esplorazione o di direzione. Anche il semplice fatto che le icone degli strumenti disegnate sullo schermo siano costituite da una sagoma bianca su sfondo scuro, ad esempio, pone a favore di un'intuitività più lineare all'interno dello spazio attivo.

La ricerca in oggetto, sulla valutazione del comportamento di utenti, è stata condotta a diretto contatto con gli stessi nell'ambito di una mostra, e si riconduce all'alveo delle attività di progettazione e verifica che pongono l'utente al centro dello sviluppo di dispositivi interattivi. Si tratta in effetti di uno studio sul campo, eseguito in condizioni "naturali" (ecologiche), con lo scopo di incrementare la conoscenza riguardo ciò che abitualmente fanno gli utenti. Tramite l'osservazione e l'analisi dei dati ottenuti si vogliono identificare le necessità che conducono a nuove tipologie di dispositivi e sono di aiuto nella valutazione dei prototipi.

Dunque, se da una parte l'utente si trova coinvolto a livello quasi ontologico nel processo di interaction design, dall'altra, per quanto concerne la progettazione di installazioni interattive e di ambienti virtuali in genere, risulta evidente come entrambi gli estremi della curva d'interazione (che coinvolge *macchina – uomo – ambiente*) necessitino di un

monitoraggio continuo e ragionato, e come la fruibilità sia, in ultima analisi, un concetto tutt'altro che astratto.

## L'ESPLORATORE D'ORCHESTRA

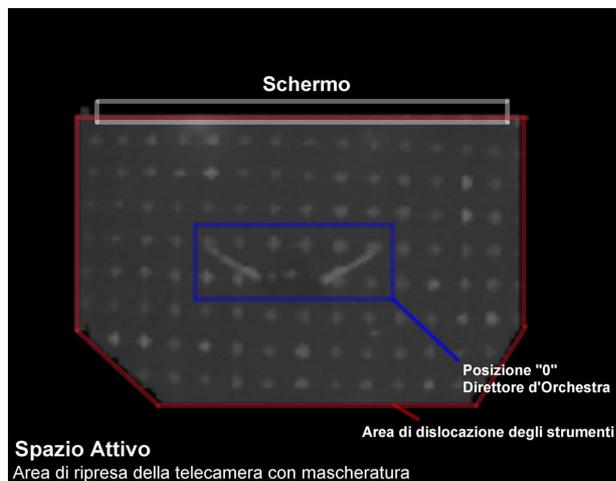


Figura 3. La telecamera, perpendicolare al terreno, riprende lo spazio fruito dall'utente inviando al software informazioni sul suo movimento.

Il movimento del visitatore nello spazio attivo viene rilevato da una telecamera ed analizzato da un'applicazione software, basata sulla piattaforma EyesWeb XMI per l'interazione multimodale. Risultato dell'analisi è l'estrazione di caratteristiche del movimento a vari livelli di complessità, dalla posizione del soggetto nello spazio a caratteristiche dinamiche del gesto (ad esempio, contrazione ed espansione, curve di energia, simmetrie; e più ad alto livello "esitazione", "incertezza", "slanci impulsivi", "dirittezza"), calcolate su singole parti del corpo, e/o sull'intero corpo.

Le singole sezioni dell'orchestra virtuale sono disposte in diverse regioni dello spazio. Esplorando lo spazio si attivano le regioni man mano che sono raggiunte: inoltre, alcuni gesti particolari o caratteristiche del movimento possono, in determinati contesti, attivare ulteriori sezioni o modularne il contenuto audio.

Due modalità di fruizione attiva dell'ambiente virtuale sono state definite nell'installazione: nel primo caso, corrispondente alla zona dello spazio attribuita al "Direttore" dell'orchestra virtuale, è possibile attivare, in posizione statica, l'intera orchestra, o parti di essa, tramite l'allargamento delle braccia, ove poter ascoltare il brano nella sua interezza; nel secondo caso, corrispondente alla modalità di fruizione che definiamo di "Esplorazione", è possibile muoversi (in questo senso, si ha una vera e propria interazione dinamica) all'interno dello spazio e scoprire le diverse zone dedicate a particolari strumenti o sezioni, estrapolandone la singola traccia dal brano *in loop* sulla macchina. Inoltre, il modo con cui un soggetto esplora lo spazio viene analizzato affinché, ad esempio, ad un approccio "irruento" verso una sezione corrisponda una modulazione ("spavento") della sezione raggiunta. Questo consente di avere un sistema più complesso di una semplice regolazione di volume delle differenti sezioni, in cui il brano viene eseguito con una interpretazione tutt'uno con le qualità espressive naturali del movimento del soggetto.

Il brano, *Borderline* (durata 3'30"), di M. Canepa, L. Cresta, A. Sacco, è un pezzo originale di musica da film che non è mai stato eseguito in pubblico. Non possiede una caratterizzazione eccessivamente forte, ed è molto adatta come colonna sonora per l'esplorazione. Dal momento che gli utenti non la conoscono, *Borderline* aiuta i soggetti fruitori a prestare attenzione a ciò che stanno ascoltando. Il brano consiste di 13

tracce stereo ed include i seguenti strumenti musicali e sezioni: arpa, violoncello, corno, flauto, contrabbasso, oboe, fagotto, percussioni, piano, violini e pizzicati. Gli strumenti sono dislocati sul palco e sono associati a 13 funzioni potenziali Gaussiane/algoritmiche.

## ARCHITETTURA DEL SISTEMA

Nella versione "Festival della Scienza 2007" sono stati impiegati due computer collegati in rete locale, in modo da gestire separatamente le componenti grafica e audio coinvolte nell'installazione.

La componente audio si basa sulla gestione attraverso EyesWeb XMI di un contenuto audio multicanale. Nella configurazione tipica dell'*Esploratore*, due canali audio (registrazione stereofonica) sono associati ad ogni strumento dell'orchestra virtuale o anche ad intere sezioni orchestrali. In principio, ogni canale audio è processato separatamente, mantenendo comunque la sincronizzazione tra le tracce. Uno o più moduli di elaborazione del suono al computer operano in tempo reale su ogni singolo canale audio. I moduli software di elaborazione del suono ricevono valori dei parametri dalla componente del sistema che gestisce il controllo della performance: le tracce così elaborate vengono quindi mixate in tempo reale, e anche il mixaggio è sottoposto alla componente di controllo. Infine, ulteriori effetti e tecniche di *sound processing* possono essere applicati all'uscita mixata, sempre controllati in tempo reale dalla componente di controllo. L'uscita tipica è un audio stereo o quadrifonico, diffuso nell'area nella quale il fruitore si sta muovendo, riproducendo così nello spazio fisico le posizioni delle diverse sezioni orchestrali.

Gli input della componente di controllo sono costituiti dai dati sul movimento estratti attraverso una videocamera ed eventualmente ulteriori sensori (per es. accelerometri<sup>1</sup>).

Le immagini vengono analizzate da EyesWeb XMI, in modo da ritagliare il corpo del fruitore dallo sfondo e tracciarne in tempo reale la posizione (x,y) e la traiettoria (*body segmentation, localization, and tracking module*).

Il modulo di interazione con lo spazio riceve in ingresso questa posizione (x,y), mentre il modulo di analisi del gesto espressivo riceve la posizione del fruitore, immagini elaborate (per es. la silhouette del corpo), informazioni sulla dinamica del movimento, ed eventualmente, dati filtrati dai sensori: a partire da questi elementi, il modulo di analisi del gesto espressivo estrae caratteristiche espressive (*features*). Le *features espressive* sono indizi che contribuiscono a valutare il contenuto espressivo del gesto e offrono una descrizione qualitativa di alto livello del gesto che il fruitore sta compiendo<sup>2</sup>. Il movimento di un ascoltatore attivo nello spazio può dunque essere classificato come Diretto o Flessibile, Impulsivo o Sostenuto. A partire da questo, saranno possibili ulteriori caratterizzazioni ad alto livello del comportamento del fruitore. Per esempio, un movimento flessibile, non fluido e sostenuto, può indicare esitazione; invece un movimento "diritto", veloce e fluido, può indicare decisione e determinazione<sup>3</sup>.

<sup>1</sup> Non utilizzati nella versione "Festival della Scienza 2007".

<sup>2</sup> Per esempio, in un recente studio [Camurri e Volpe 2006], è stato utilizzato un set di *features espressive* di traiettorie gestuali per classificare il gesto espressivo in relazione alle dimensioni *Space* e *Time* definite nella *Effort Theory* del grande coreografo e teorico Rudolf Laban.

<sup>3</sup> Il modulo di analisi del gesto espressivo è definito dal *multi-layered framework for expressive gesture processing* [Camurri et al. 2005], sviluppato nel progetto europeo **IST MEGA** (*Multisensory Expressive Gesture Applications*) e ulteriormente raffinato nel progetto europeo **IST Tai-Chi** (*Tangible Acoustic Interfaces for Computer-Human Interaction*).

Il modulo che modella l'interazione spaziale computa e sovrappone allo spazio fisico un insieme di funzioni potenziali bidimensionali.

Almeno una funzione potenziale è associata a ogni canale audio d'ingresso, per controllare il mixaggio in tempo reale. Altre funzioni potenziali possono essere create per controllare i moduli degli effetti e della elaborazione del suono (presenza, riverbero, *sound morphing*, ecc.). Mentre il fruitore esplora lo spazio fisico, vengono calcolati i valori di queste funzioni potenziali in relazione alla sua posizione (x,y). I valori dei potenziali associati al mixaggio in tempo reale sono usati come pesi per i livelli del suono dei canali corrispondenti.

I valori dei potenziali associati a parametri o effetti dei moduli di *sound processing*, operano sui moduli corrispondenti.

Il modulo di analisi del gesto espressivo può modificare l'attuale valore dei parametri delle funzioni potenziali, così che il loro profilo può essere modellato dal gesto espressivo che il fruitore compie esplorando l'orchestra virtuale.

La componente di controllo può perciò operare sulla componente audio in due modi differenti:

- Attraverso una corrispondenza diretta delle caratteristiche del gesto espressivo con i parametri di elaborazione del suono: per esempio, un più alto valore nella Quantità di Moto può essere associato a una maggiore riverberazione del suono.
- Attraverso una corrispondenza del valore delle funzioni potenziali con i parametri di elaborazione del suono: per esempio, un movimento esitante può determinare un maggiore sparpagliamento degli strumenti nello spazio dell'orchestra.

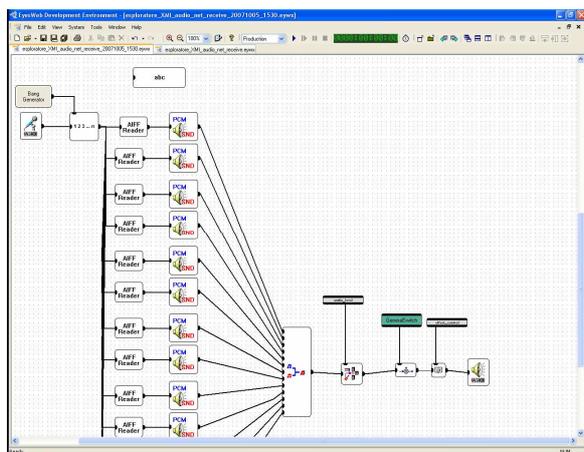


Figura 4. Snapshot di EyesWeb in cui è evidente l'input audio multicanale (2 canali per ogni strumento), per le cui tracce si procede al mixaggio in tempo reale.

## L'ESPERIMENTO: PARAMETRI E METODOLOGIA DI RILEVAMENTO DATI

**Periodo monitorato:** 1-6 novembre 2007, dalle ore 10.00 alle 17.00 (tranne il 3 novembre, dalle ore 14.00 alle 17.00, e il 5 novembre, dalle ore 10.00 alle 13.00).

**Parametri seguiti nel rilevamento dati:** precisando che ogni valore viene riferito al "Soggetto-Fruitore", ossia a chi concretamente interagisce con l'installazione, teoricamente un soggetto singolo, in alcuni casi (bambini molto piccoli), l'installazione è stata fruita da due soggetti contemporaneamente, in modo tale (posizione molto ravvicinata, adulto che reca in braccio un bambino) che la telecamera li rilevasse come un *unicum*. In questi casi, si è considerato chi, fra i due soggetti, tenesse l'atteggiamento

dominante in senso direzionale e comportamentale. Nello specifico, si riportano di seguito gli elementi considerati:

**Gruppo/Singolo:** appartenenza o meno del soggetto ad un gruppo di visitatori e, nel caso, quantificato;

**Ora Accesso:** ora di accesso all'installazione, assunta come coincidente con l'inizio della concreta interazione;

**Età:** si è fatto ricorso ad un criterio di suddivisione in fasce d'età consistente in un intervallo di anni 10.

**Sesso;**

**Direzione:** si è considerata la direzione presa dal soggetto a partire dalla "Posizione 0", Direttore d'Orchestra, nel momento in cui egli l'abbandona per andare in esplorazione dello spazio circostante;

**Riferimento principale:** il soggetto può scegliere come guida la mappa degli strumenti proiettata sullo schermo, oppure prediligere una più autentica esplorazione dell'ambiente virtuale, facendo a meno della mappa stessa;

**Tempo di permanenza:** i valori si riferiscono al tempo di permanenza del soggetto nella posizione del *Direttore* o in *Esplorazione*.

**Tipologia d'interazione:** l'interazione può rivelarsi *corretta* oppure presentare *anomalie*.

*Interazione corretta:* il soggetto riesce a relazionarsi senza difficoltà con l'ambiente virtuale gestendo la mappa in maniera piana e coerente, nel caso essa venga scelta come riferimento principale; il soggetto decide di esplorare l'ambiente senza avvalersi della mappa, affidandosi solo ai suoi sensi; il soggetto esplora autonomamente l'ambiente, utilizzando la mappa come verifica della sua posizione dello spazio; il soggetto presenta interesse per specifici strumenti e, sulla base della mappa, raggiunge senza difficoltà la posizione corrispondente nello spazio.

*Interazione "affetta" da anomalie:* nella posizione del Direttore, il soggetto assume una gestualità coerente al personaggio ma non adeguata allo spirito dell'installazione, poiché un tale comportamento presuppone che col semplice gesto vengano azionati gli strumenti proiettati sullo schermo; in esplorazione, il soggetto allarga le braccia nelle posizioni corrispondenti ai singoli strumenti, nella speranza di modificare il volume ed in qualche caso le modalità d'esecuzione; in esplorazione, il soggetto ha difficoltà nel relazionare la sua posizione nello spazio alla mappa che pure ha scelto come riferimento principale; in esplorazione, quando il soggetto si trova nella posizione corrispondente ad un particolare strumento, ne simula le modalità d'esecuzione; in esplorazione, il soggetto fraintende la funzione dello schermo o del pavimento e li scambia per superfici sensibili.

**Proposta d'interazione fra più soggetti:** il soggetto, influenzato dalle possibilità offerte dall'installazione visitata immediatamente prima dell'*Esplorazione*, *Mappe per Affetti Erranti* (Camurri et al. 2008), s'informa sull'eventualità di poterne replicare, anche in questo caso, l'aspetto collaborativo fra più soggetti.

**Procedura:** sulla base dell'osservazione dei comportamenti assunti dai visitatori, si è provveduto ad eseguire una sistematizzazione dei dati raccolti su base giornaliera e poi sull'intero periodo monitorato, ove poterne gestire l'analisi a livello parziale e totale, ed avere un pattern d'informazioni discreto; i dati, riportati su fogli di calcolo, hanno subito una prima elaborazione al fine di convertire numericamente alcuni di essi e renderli adatti all'analisi, condotta suddividendo i soggetti per sesso (M; F) e per fasce d'età dell'ampiezza di 10 anni ciascuna, da 0 a 70. Per ognuna di queste classi sono stati analizzati i valori relativi a quattro dei parametri rilevati: **direzione intrapresa** a partire dalla "posizione 0" (ovvero se si preferisce intraprendere l'esplorazione dalla parte sinistra o destra dell'orchestra, a partire dal leggio); **tempo trascorso** nelle diverse modalità di fruizione; **qualità/tipologia della fruizione;** **riferimento principale** assunto come guida della

fruizione stessa (Mappa; Ambiente). Perché sono stati selezionati questi, e non altri parametri che pure si sono rilevati? In effetti, dopo un attento esame, si è notato come le categorie elette ad elaborazione successiva presentano il più alto livello di obiettività e meglio si prestano ad una quantificazione ed ad un'analisi di tipo statistico. Ad esempio, parametri come l'appartenenza o meno dell'utente ad un gruppo di visitatori, e se sì, di quanti (nel caso in cui la fruizione potesse esserne influenzata), oppure l'ora in cui esso mette piede nell'installazione, non risultano decisivi per il comportamento tenuto, in quanto contaminati da fattori eteronomi come, ad esempio, la successiva incorporazione di ritardatori o l'incalzare di un altro gruppo, circostanza, questa, che tuttavia non ha influenzato i tempi di fruizione: ogni utente è stato libero di interagire con l'ambiente virtuale per quanto i tempi della visita (all'incirca uguali per tutti) lo concedessero.

Infine, con riguardo al parametro denominato **“Riferimento principale”**, occorre fare una breve digressione: nella concezione originale dell'installazione si è posto l'accento sull'aspetto proprio dell'esplorazione che l'ambiente virtuale offre, in particolare, esplorazione basata su una forma “attiva” di ascolto. L'utente, immerso nello spazio, avrebbe dovuto lasciarsi guidare dall'udito alla scoperta dell'orchestra invisibile. Se questo era del tutto vero ed attuabile nella “versione 2006” dell'*Esploratore*, che quanto a feedback visivo prevedeva soltanto la proiezione sullo schermo di una sorta di ombra colorata dell'utente, nell'allestimento del 2007, a seguito delle variazioni apportate al feedback visuale (proiezione della mappa degli strumenti dislocati nello spazio), l'intenzionalità originale ha subito, a livello dei fatti, una sorta di “deviazione d'utilizzo”. In effetti, la stragrande maggioranza dei visitatori, invece di lasciarsi guidare dai suoni, ed eventualmente verificare sulla mappa la propria posizione nello spazio, tendeva a seguirla in modo pedissequo, in alcuni casi scegliendo preventivamente quale strumento recarsi a sentire: si tratta di un indizio considerevole della preponderanza dell'organo della vista anche in un'installazione nata in virtù di un potenziamento delle qualità d'ascolto. Sulla base di questa situazione, si è ritenuto utile effettuare l'analisi solo sul periodo complessivo, vista la stragrande maggioranza di soggetti che hanno assunto la mappa come guida della propria fruizione.

**Risultati:** si riportano, di seguito, alcuni dei risultati complessivi più significativi (ottenuti dall'analisi dei dati dell'intero periodo di monitoraggio, fra l 1 e 6 novembre 2007) derivanti dall'attività di rilevamento.

Totale visitatori dell'installazione nel periodo monitorato: 481  
 Totale fruitori dell'installazione monitorati: 154  
 Totale soggetti femmine: 79  
 Totale soggetti maschi: 75  
 Totale soggetti divisi secondo fasce d'età:

Fasce d'età	Numero totale
0 < 10	42
10 < 20	30
20 < 30	26
30 < 40	17
40 < 50	20
50 < 60	12
60 < 70	7

## RISULTATI

Mentre il numero di soggetti femmine e soggetti maschi che, nel corso dei 6 giorni di monitoraggio, hanno fruito dell'installazione, è equiparabile (79 visitatrici e 75 visitatori), scorrendo lo sguardo sulle fasce d'età è possibile notare come bambini (di età inferiore ai 10 anni) e ragazzi (fra gli 11 e i 20

anni), siano stati gli ospiti più frequenti, anche a causa dell'elevato numero di classi scolastiche in visita.

*Rapporto fra sesso e direzione intrapresa dall'utente all'uscita della “posizione 0” (Direttore d'Orchestra):* non è difficile rilevare che, con riguardo a questo primo pattern di dati, la preponderanza sia da attribuire al lato sinistro, per entrambi i sessi, sebbene le donne abbiano, in questo senso, un maggiore incremento (65,82% contro 58,67%), che invece gli uomini detengono sul lato destro (41,33% contro 34,18%). Tale preferenza per il lato sinistro potrebbe essere imputabile alla presenza, in quella zona, degli strumenti più “popolari” (archi, pianoforte, arpa) e, secondariamente, dotati del timbro più squillante; sul lato destro, invece, sono collocati strumenti dal timbro più profondo, e va inoltre segnalata la presenza nella mappa di tre “icone fantasma”, cioè zone mancanti della sagoma dello strumento corrispondente, che nascondevano strumenti percussivi. Ancora è da sottolineare, in questo caso, la supremazia dell'organo della vista, che condiziona il fruitore fin dalla “posizione 0”.

*Rapporto fra sesso e qualità/tipologia di fruizione:* la percentuale di interazioni che risultano corrette è, in totale, superiore (56,86% contro 43,14% d'interazioni che presentano anomalie e problemi), segno che l'“immersione” nell'ambiente virtuale ha trovato un buon riscontro a livello di adattamento del soggetto coinvolto. Delle motivazioni sottese al predominio della mappa utilizzata come riferimento principale in esplorazione, si è già detto.

*Rapporto fra sesso e tempo di permanenza nelle due diverse modalità di fruizione (Direzione ed Esplorazione):* le percentuali, in questo caso, mostrano, da parte di entrambe le categorie di soggetti, una netta prevalenza del tempo trascorso in esplorazione; gli uomini, tuttavia, dimostrano una certa preferenza per la modalità di direzione (36,88% contro il 32,30% delle donne, rispetto al totale del tempo trascorso nell'installazione).

*Rapporto fra fasce d'età e direzione intrapresa dall'utente all'uscita della “posizione 0” (Direttore d'Orchestra):* la preferenza per il lato sinistro dell'orchestra virtuale è pressoché evidente nella maggior parte delle fasce prese in considerazione, sebbene in quelle comprese fra i 20 e i 30 anni, fra i 50 e i 60, i 60 e i 70, la destra sia in netto recupero (rispettivamente 53,85% contro 46,15%, 50%, 57,14% contro 42,86% del lato sinistro). E' possibile che in queste fasce d'età sia maggiore l'istinto ad esplorare strumenti meno conosciuti, meno “inflazionati”, e, nel caso delle “icone fantasma”, assolutamente ignoti, comunque dal timbro più scuro. Si potrebbe ravvisare, in questo, un atteggiamento più sicuro di sé e più orientato a scopo conoscitivo.

*Rapporto fra fasce d'età e qualità/tipologia di fruizione:* la percentuale più alta d'interazioni corrette si ha nella fascia di età compresa fra i 10 e i 20 anni (il 70% dei soggetti appartenenti alla fascia), probabilmente a causa di una maggiore abitudine all'interazione con interfacce grafiche/sonore (*videogames* ecc.) e la minor presenza ed incidenza di sovrastrutture mentali e preconcetti nei confronti di nuove modalità di fruizione spaziale mediate da quella che può definirsi tecnologia impiegata in modo “non convenzionale”. La fascia d'età che invece ha presentato il maggior numero di problemi è quella tra i 40 e i 50 anni (il 60% dei soggetti appartenenti alla fascia).

*Rapporto fra fasce d'età e tempo di permanenza nelle due diverse modalità di fruizione (Direzione ed Esplorazione):* in tutte la fasce d'età è evidente la prevalenza temporale in modalità esplorativa, sebbene il maggior numero di soggetti che hanno trascorso più tempo in direzione siano quelli compresi tra 0 e 10 anni e tra i 60 e i 70 anni (il 39,79% e il 40% contro percentuali nettamente inferiori relative alle altre fasce d'età): è probabile che l'ensemble dell'orchestra,

ubbidiente alla gestualità di chi si trova in direzione, abbia esercitato una maggiore impressione sui fruitori più giovani e su quelli più anziani, ed anche più entusiasti nell'agire e reagire sulla base del feedback sonoro.

**Rapporto fra sesso e Riferimento principale:** solo 13 visitatori su 154 fruitori (l'8,44% del totale) nell'arco dei 6 giorni hanno scelto l'ambiente come fonte primaria dell'interazione: 4 soggetti femmine e 9 soggetti maschi, il 5,06% ed il 12% delle rispettive categorie; inoltre, si tratta del 30,77% dei soggetti femmine e del 69,23% dei soggetti maschi fra coloro che hanno eletto l'ambiente a riferimento principale, con una netta prevalenza di questi ultimi.

**Rapporto fra fasce d'età e Riferimento principale:** il 46,15% (ed il 14,29% della categoria) dei visitatori che hanno scelto come riferimento principale l'ambiente (6 visitatori su 13) è costituito da soggetti di fascia d'età inferiore ai 10 anni: si tratta di un dato estremamente significativo perché indica come, in effetti, siano i soggetti più giovani ad interagire con maggior naturalezza e a muoversi meglio all'interno di un ambiente virtuale. Mentre gli adulti tendono ad assumere un atteggiamento più orientato in senso cognitivo, lasciandosi guidare dalla vista per esperire l'esplorazione, i bambini preferiscono un coinvolgimento più "sensazionale", che li

porta a scoprire un livello meno immediato di conoscenza dell'ambiente circostante.

## SVILUPPI FUTURI

Alla luce di quanto emerso, molti sono gli aspetti che potranno essere coinvolti nella ricerca, con riguardo, per esempio, ad una descrizione qualitativa di alto livello di caratteristiche del comportamento del fruitore: in particolare, sono possibili forme di *social interaction* fra più utenti che si muovono all'interno dello spazio attivo, e il cui movimento è indagato specificamente dal punto di vista dell'espressività. Tali elementi sono stati introdotti per la "forma evoluta" dell'*Esploratore d'Orchestra*: l'installazione *Mappe per Affetti Erranti*, presentata nell'edizione 2007 del Festival della Scienza, in cui vengono definite mappe espressive che associano qualità del gesto ad interpretazione musicale, ma dove è anche possibile sperimentare dinamiche di gruppo (*social network*), poiché solo la fruizione multipla offre la possibilità di ascolto di un brano nella sua interezza.

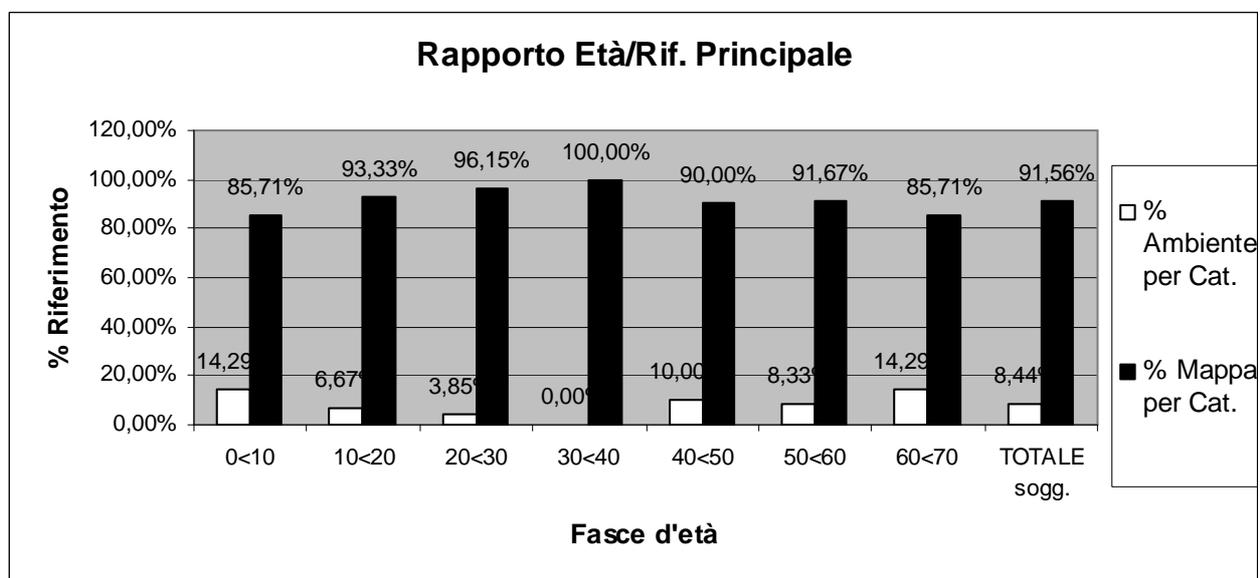


Figura 5. Totale soggetti e riferimento principale assunto, secondo le fasce d'età.

## RIFERIMENTI

1. Antonio Camurri, Corrado Canepa, Gualtiero Volpe. *Active listening to a virtual orchestra through an expressive gestural interface: The Orchestra Explorer*, Proc. Intl. Conf. NIME-2007 New Interfaces for Music Expression, New York University, 2007a.

2. Antonio Camurri, Corrado Canepa, Gualtiero Volpe. *Ascolto attivo di un'orchestra virtuale: un'interfaccia basata sul gesto espressivo*, *Metamorfosi del Senso* – Catalogo della mostra Cimenti di Invenzione e Armonia (a cura di N. Ferrari), Quaderni di Casa Paganini, 2, Genova, 2007b.

3. Antonio Camurri, Corrado Canepa, Paolo Coletta, Barbara Mazarino, Gualtiero Volpe. *Mappe per Affetti Erranti: a Multimodal System for Social Active Listening and Expressive Performance*, Proc. Intl. Conf. NIME-2008 New Interfaces for Music Expression, Genova, 2008.

4. Antonio Camurri, Giovanni De Poli, Marc Leman, Gualtiero Volpe. *Toward Communicating Expressiveness and Affect in Multimodal Interactive Systems for Performing Art and Cultural Applications*, IEEE Multimedia Magazine, Vol. 12, no. 1, 2005.

5. Antonio Camurri, Giovanna Varni, Gualtiero Volpe. *Emotional Entrainment in Music Performance*, Proc. 8<sup>th</sup> IEEE Intl Conf on Automatic Face and Gesture Recognition. Sept. 17-19, Amsterdam, 2008.

6. Antonio Camurri, Gualtiero Volpe. *Multimodal and cross-modal analysis of expressive gesture in tangible acoustic interfaces*, Proc. RO-MAN 2006 – The 15<sup>th</sup> IEEE Intl. Symposium on Robot and Human Interactive Communication, IEEE CS Press, 2006.

7. Francesca Cavallero. *Studi ed esperimenti su modelli di ascolto attivo in ambienti virtuali interattivi*, Tesi di Laurea Specialistica in « Scienze dello Spettacolo », Università degli Studi di Genova, Marzo 2008.

8. Mihály Csikszentmihalyi. *Flow: The psychology of optimal experience*, Harper & Row, New York, 1990.

9. James Jerome Gibson. *The ecological approach to visual perception*, Houghton Mifflin, Boston, 1979.

10. J. Preece, Yvonne Rogers, Helen Sharp, *Interaction Design beyond human-computer interaction*, John Wiley & Sons, Inc., 2002.

11. Eliséo Veron and Martine Levasseur. *Ethnographie de l'exposition*, Bibliothèque publique d'Information, Centre Georges Pompidou, Paris, 1983.



# Modelli VCF discreti a retroazione istantanea

Marco Civolani    Federico Fontana  
Dipartimento di Informatica  
Università degli studi di Verona

## SOMMARIO

Il progetto riguarda la modellazione numerica in tempo reale di filtri parametrici analogici per l'elaborazione di segnali audio. In particolare, vengono presentati i modelli di due Voltage Controlled Filter (VCF) largamente presenti nelle produzioni musicali degli ultimi quarant'anni: il primo è quello progettato da Robert Moog nel 1967 e integrato negli omonimi sintetizzatori; il secondo è quello ideato da David Cockerell, installato a bordo del sintetizzatore EMS VCS3 MkII. I due modelli sono stati derivati seguendo strategie differenti, aventi un punto in comune: in entrambi i casi viene infatti preservata l'istantaneità dell'anello di retroazione a guadagno variabile, che conseguentemente è stato modellato senza l'ausilio di elementi di ritardo.

## IL VCF MOOG

Il filtro in questione è descritto dallo schema a blocchi riportato in Figura 1: la struttura è sostanzialmente costituita da un anello contenente una serie di quattro filtri del primo ordine identici. Ciascuno di essi possiede una caratteristica passabasso che in regime lineare è espressa con buona approssimazione dalla seguente funzione di trasferimento

$$G(s) = \frac{\omega_c}{\omega_c + s} \quad (1)$$

Nel VCF Moog i parametri  $\omega_c$  e  $k$  possono essere variati dinamicamente: il primo determina la frequenza di taglio, il secondo il guadagno di retroazione. Quest'ultimo può assumere valori compresi tra 0 e 4. Aumentando il valore del guadagno di retroazione, si invia una maggior quantità di segnale  $y(t)$  in ingresso: in questo modo, il comportamento di puro filtro passabasso del sistema (taglio ai 3 dB in corrispondenza di  $\omega_c$ , con  $k = 0$ ) diventa sempre più simile a quello di un filtro risonante. Il limite di stabilità si ha per  $k = 4$ : in questa condizione il VCF oscilla ad una frequenza pari a  $\omega_c$  rad/s.

L'approccio seguito per la modellazione del VCF è stato indirizzato fin dall'inizio a preservare nel dominio discreto la struttura del dispositivo stesso: in pratica, ciascun filtro analogico del primo ordine costituente la serie di cui sopra è stato mappato in un suo corrispondente digitale. Una tale strategia di progettazione si è resa possibile grazie all'esistenza di un versatile strumento formale in grado di trattare gli anelli di retroazione privi di ritardo (cosa impossibile sfruttando unicamente le tradizionali procedure sequenziali): grazie a questo strumento, la versione digitale del VCF Moog otte-

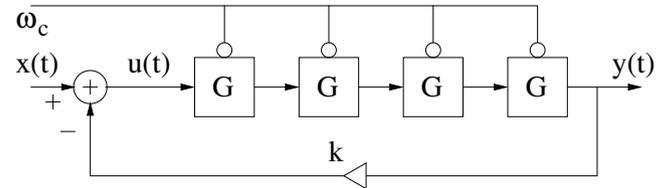


Figura 1. Struttura del VCF Moog.

nuta mediante trasformazione bilineare possiede una risposta in frequenza molto simile a quella della controparte analogica. Il modello ottenuto inoltre consente l'accesso diretto ai parametri e si è rivelato robusto alle loro variazioni.

## CALCOLO DELL'ANELLO SENZA RITARDO

Il problema maggiore nella discretizzazione del VCF risiede nel fatto che digitalizzando  $G(s)$  non si ottiene alcun blocco di ritardo puro: essa infatti risponde a ciascun input senza introdurre alcun ritardo temporale. Ciò implica che il filtro non può essere simulato nel dominio digitale alle usuali frequenze di campionamento mediante l'ausilio delle tradizionali procedure sequenziali. Esistono, a tal proposito, diversi metodi basati sulla topologia dei grafi in grado di rilevare (anche in strutture filtranti di ordine elevato) e conseguentemente aggregare questo tipo di anelli di retroazione [5]; nel caso del VCF Moog, la rilevazione dell'anello risulta immediata, così come la determinazione della relativa funzione di trasferimento "aggregata" [8]:

$$H(s) = \frac{\{G(s)\}^4}{1 + k\{G(s)\}^4} = \frac{1}{k + \{1 + s/\omega_c\}^4}. \quad (2)$$

Se a questo punto si implementasse un filtro del quarto ordine che realizzasse  $H(s)$  si perderebbero le proprietà strutturali del VCF, e quindi l'accesso diretto ai parametri di regolazione risulterebbe impossibile. Nel 2000 Härmä propose un algoritmo in grado di calcolare nel discreto un anello privo di ritardo come quello riportato in Figura 1 [4]: esso è stato poi generalizzato a reti filtranti lineari [1] e non lineari [2,3] contenenti un qualunque tipo di anello di retroazione senza ritardo. Tale algoritmo prevede che in prima istanza venga calcolato l'output del filtro  $y[n]$  e che, successivamente, si aggiorni lo stato dei blocchi filtranti costituenti l'anello. L'algoritmo in questione risulta particolarmente utile nel caso del VCF Moog dal momento che consente di preservare la struttura di Figura 1.

La serie composta dalle quattro celle filtranti del primo or-

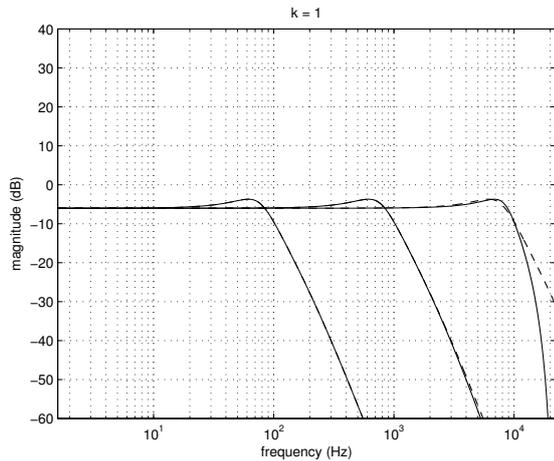


Figura 2.  $k = 1$ : modulo della risposta in frequenza del VCF Moog digitale (linea continua); modulo della risposta in frequenza del VCF Moog analogico (linea tratteggiata). Frequenze di taglio: 100, 1000 e 10000 Hz.

dine costituisce un sistema lineare la cui uscita  $y[n]$  è data dalla sovrapposizione della sua evoluzione libera con quella forzata; la prima è una funzione lineare  $s$  dello stato dei quattro filtri passabasso, mentre la seconda dipende linearmente dal segnale  $u[n]$ :

$$y[n] = bu[n] + s[n]. \quad (3)$$

Inoltre (si veda Figura 1), vale

$$u[n] = x[n] - ky[n]. \quad (4)$$

Sostituendo (4) in (3) si ottiene

$$y[n] = \frac{bx[n] + s[n]}{1 + bk}. \quad (5)$$

A ogni passo, il contributo istantaneo dell'evoluzione libera  $s[n]$  all'uscita  $y[n]$  può essere determinata forzando  $u[n] = 0$ . In questo modo, la seguente procedura può essere implementata al fine di determinare  $y[n]$ : i) si apra l'anello tagliando il ramo in cui scorre  $u$ ; ii) si determini l'uscita (cioè  $s[n]$ ); iii) si calcoli  $y[n]$  mediante (5). A questo punto,  $u[n]$  è immediatamente determinato tramite (4) e quindi lo stato delle quattro celle filtranti può essere aggiornato di conseguenza.

### TRASFORMAZIONE AL DISCRETO

In letteratura esistono diverse tecniche utili a mappare un filtro passabasso analogico del primo ordine in uno digitale di ordine basso [6]: per la discretizzazione del VCF Moog è stata adottata la ben nota trasformazione bilineare: è stata inoltre sfruttata la tecnica del warping sulla funzione di trasferimento discreta  $G(z)$  ottenuta da  $G(s)$ . In particolare, è stata effettuata una pre-compensazione della frequenza di taglio al valore  $\tilde{\omega}_c$ , che è la trasformazione bilineare esatta di  $\omega_c$ . Sostituendo  $s$  con  $(2/T)(z-1)/(z+1)$  in  $G(s)$  si ottiene la corrispondente cella passabasso pre-compensata:

$$G(z) = b_0 \frac{1 + z^{-1}}{1 + a_1 z^{-1}}, \quad (6)$$

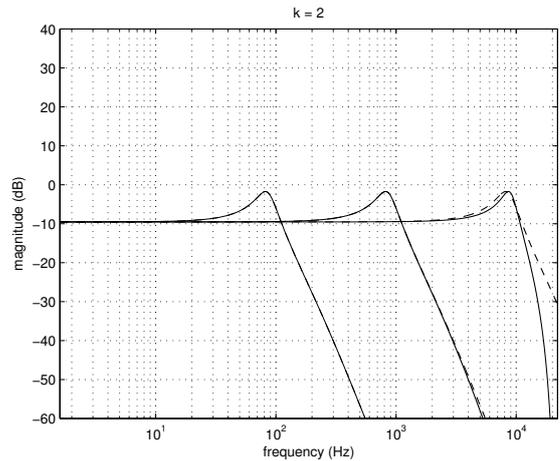


Figura 3.  $k = 2$ : modulo della risposta in frequenza del VCF Moog digitale (linea continua); modulo della risposta in frequenza del VCF Moog analogico (linea tratteggiata). Frequenze di taglio: 100, 1000 e 10000 Hz.

dove:  $b_0 = T\tilde{\omega}_c/(T\tilde{\omega}_c + 2)$  e  $a_1 = (T\tilde{\omega}_c - 2)/(T\tilde{\omega}_c + 2)$ . Dalla definizione di trasformazione bilineare, la pulsazione antitrasmformata  $\tilde{\omega}_c$  del filtro digitale è così calcolata:

$$\tilde{\omega}_c = \frac{2}{T} \tan\left(\frac{\omega_c T}{2}\right). \quad (7)$$

Dal momento che l'anello contiene quattro blocchi passabasso in serie, allora in Eq. (3) vale  $b = b_0^4$ . Riscrivendo ora  $G(z)$  in forma canonica,  $s[n]$  può essere ottenuto in questo modo:

$$s[n] = b_0^3 s_1[n] + b_0^2 s_2[n] + b_0 s_3[n] + s_4[n]. \quad (8)$$

Le equazioni (8), (5), e (4) possono essere applicate in questo ordine per realizzare la procedura precedentemente illustrata. Infine, gli stati  $s_1, s_2, s_3, s_4$  dei blocchi passabasso vengono aggiornati in successione.

### RISULTATI (VCF MOOG)

Le Figure 2, 3, 4, 5 mostrano: in linea tratteggiata, il modulo della risposta in frequenza del VCF Moog analogico ottenuta tracciando  $|H(j2\pi f)|$  come in (2) e, in linea continua, lo spettro delle risposte all'impulso ottenute eseguendo in Matlab la procedura di cui sopra, rispettivamente per valori del guadagno  $k$  pari a: 1, 2, 3, 4. Tutte le risposte sono state tracciate per valori della frequenza di taglio pari a 100, 1000 e 10000 Hz. Per frequenze di taglio pari a 100 e 1000 Hz si ottengono risultati molto accurati, a patto che  $\omega_c$  sia pre-compensata usando (7). A 10 kHz l'effetto del warping in frequenza diviene visibile: il modulo della risposta in frequenza del sistema a tempo discreto cala infatti molto rapidamente oltre tale valore di frequenza. La risposta del sistema analogico e quella del sistema digitale divergono tuttavia sensibilmente differenti solo quando il filtro è portato al limite di stabilità, cioè per  $k = 4$  (per tale valore di guadagno di retroazione il filtro diviene instabile!). Ad ogni modo, tali discrepanze appaiono ben al di sotto dell'ampiezza registrata alla frequenza di picco e la loro effettiva udibilità andrebbe

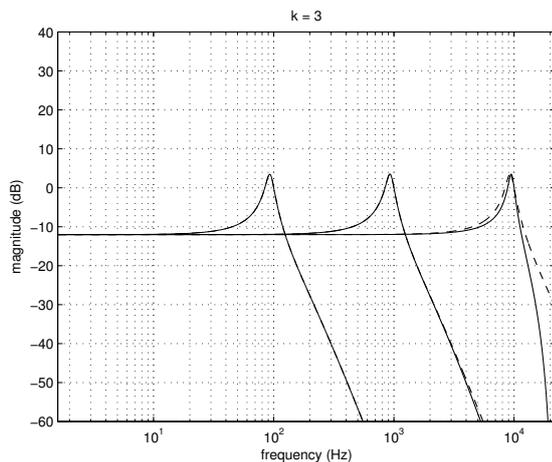


Figura 4.  $k = 3$ : modulo della risposta in frequenza del VCF Moog digitale (linea continua); modulo della risposta in frequenza del VCF Moog analogico (linea tratteggiata). Frequenze di taglio: 100, 1000 e 10000 Hz.

valutata mediante prove di ascolto che mettano a confronto il modello numerico del VCF con la sua controparte analogica.

#### IL VCF EMS

Il VCF integrato nel sintetizzatore analogico modulare EMS VCS3 sfrutta i diodi a giunzione come resistenze variabili per il controllo della frequenza di taglio: tale circuito si basa su una struttura *ladder*, molto simile a quella implementata nel VCF Moog [7–11]; la principale differenza tra i due circuiti risiede nel fatto che in quest'ultimo sono sfruttati i transistor BJT per il controllo della frequenza di taglio. In Figura 6 compare lo schema circuitale del VCF. Si tratta di un amplificatore differenziale caricato con una rete diodi-capacità, in cui i diodi fungono appunto da resistori controllati in corrente. Il valore di resistenza da essi assunto determina la frequenza di taglio del filtro e dipende dalla corrente di controllo  $I_0$ , imposta per mezzo del transistor Q4 che funge da generatore di corrente controllato dalla tensione  $V_{CV}$ .

Il transistor Q3 costituisce invece una fonte attiva di tensione costante e indipendente dalla corrente che fruisce nei due rami paralleli del differenziale: esso consente di mantenere una tensione fissa pari a 9V sugli anodi di D13 e D14, garantendo la polarizzazione dei diodi necessaria al transito di corrente nei rami stessi con il verso indicato in Figura 6.

Il VCF è altresì dotato di un circuito di retroazione a guadagno variabile (da 0 a 10 circa), costituito fondamentalmente dall'amplificatore differenziale basato sull'operazionale U1, la cui uscita viene mandata sia alla base di Q2 (chiudendo l'anello di feedback), sia al vero e proprio output del filtro. Per ottenere la retroazione, l'uscita differenziale del filtro ad anello aperto (cioè la tensione ai capi di C4) viene posta in ingresso al suddetto amplificatore passando attraverso una cella R-C (costituita da C7 e R7 su un ramo e da C8 e R8 sull'altro). Questa funge da filtro passa-alto del primo ordine e ha lo scopo di eliminare la componente continua dal

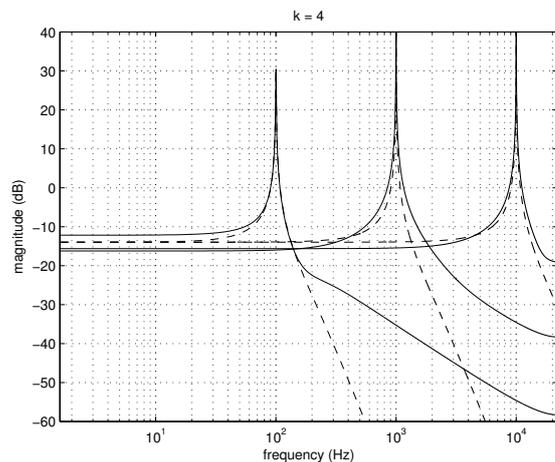


Figura 5.  $k = 4$ : modulo della risposta in frequenza del VCF Moog digitale (linea continua); modulo della risposta in frequenza del VCF Moog analogico (linea tratteggiata). Frequenze di taglio: 100, 1000 e 10000 Hz.

segnale in uscita. Lo stesso compito è assolto da C5 e C6, che vengono sfruttati per bloccare la componente continua della tensione di base di Q1 e Q2 rispettivamente.

Si considerino infine i tre diodi D9, D11 e D13 e i corrispondenti D10, D12 e D14: essi possono essere visti come due resistenze di *pull-up* anch'esse controllate in corrente e necessarie a fare in modo che gli ingressi dell'amplificatore operazionale U1 non siano entrambi posti a 9V (rendendo di fatto, il filtro inutilizzabile). I tre diodi potrebbero essere sostituiti da un resistore vero e proprio, ma ciò limiterebbe l'efficacia della corrente di polarizzazione  $I_0$  sulla variazione della frequenza di taglio del filtro e introdurrebbe una linearizzazione nel comportamento dello stesso: ciò sarebbe da molti avvertito come un effetto indesiderato, dal momento che il carattere e le sonorità tipiche di circuiti come questo sono dovute principalmente alle loro caratteristiche non lineari. Il filtro in questione è stato analizzato dal punto di vista circuitale al fine di estrapolare le equazioni differenziali non lineari in grado di descriverne il comportamento nel dominio del tempo in regime di grandi segnali: a tal scopo, i diodi sono stati modellati con una caratteristica tensione-corrente esponenziale, mentre per i transistor è stata adottata una versione semplificata del modello di *Ebers-Moll*. Si è inoltre proceduto a semplificare altre parti del circuito: ci si è concentrati in particolare sul ruolo dei transistor Q4 e Q3. Il primo è stato sostituito da un *generatore ideale di corrente* di intensità variabile, in modo da consentire la regolazione della frequenza di taglio; il secondo è stato invece trattato come *generatore ideale di tensione*, il cui valore è stato posto a 9V. Infine è stato introdotto il circuito di retroazione a guadagno variabile: l'amplificatore operazionale U1 è stato modellato con le usuali assunzioni di linearità e idealità nella caratteristica di input-output.

Lo sviluppo del modello è stato suddiviso in due fasi: dapprima si è considerato il filtro ad anello aperto (senza il modulo deputato alla retroazione). Successivamente si è tenuto

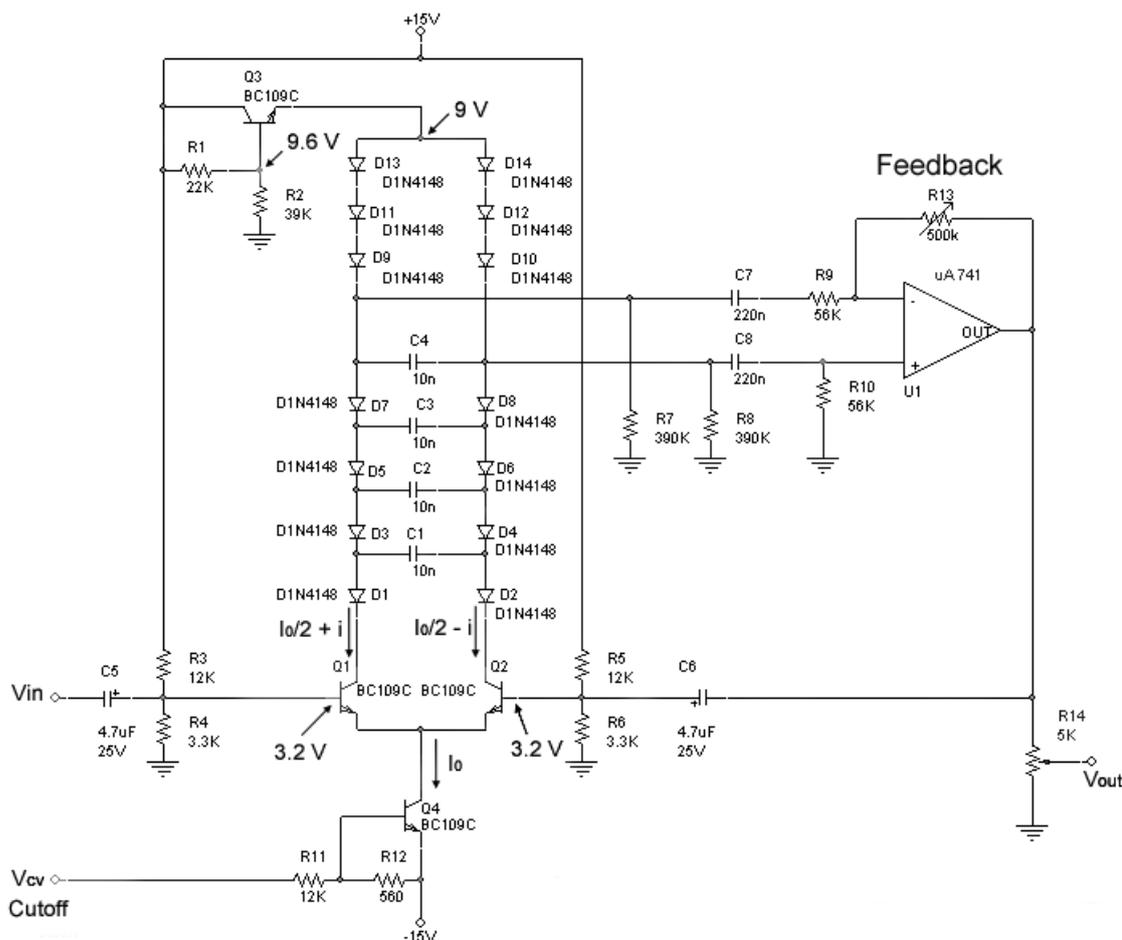


Figura 6. Schema circuitale del VCF dell'EMS VCS3 MkII.

conto della retroazione: al circuito ad anello aperto è stato cioè aggiunto l'amplificatore differenziale basato sull'amplificatore operazionale U1.

### MODELLO NON LINEARE AD ANELLO APERTO

Lo schema circuitale del filtro ad anello aperto è riportato in Figura 7. Dallo studio del circuito del VCF emerge che è possibile assumere che i diodi nei due rami del ladder siano sempre polarizzati direttamente. Il loro comportamento in regime di grandi segnali è quindi descritto con buona approssimazione da Eq. (9):

$$i = I_{inv} \left( e^{\frac{v}{\eta V_T}} - 1 \right), \quad (9)$$

Dove:  $I_{inv}$  è la corrente di polarizzazione inversa del diodo,  $\eta$  è il coefficiente di emissione, parametro il cui valore varia da diodo a diodo: per il modello 1N4148 (adottato nel VCF),  $\eta = 1.836$ .  $V_T$  è il cosiddetto thermal voltage:  $V_T = kT/q \approx 0.026$  mV a  $25^\circ\text{C}$ , dove  $q = 1,60206 \cdot 10^{-19}$  C è la carica dell'elettrone,  $k$  è la costante di Boltzmann e  $T$  è la temperatura espressa in Kelvin. La seguente equazione può

essere direttamente derivata da Eq. (9):

$$v = \eta V_T \ln \left( \frac{i}{I_{inv}} \right). \quad (10)$$

I transistor Q1 e Q2 sono stati trattati come prescritto dal modello di Ebers-Moll, in grado di descriverne con buona approssimazione il comportamento non lineare in zona attiva. In Figura 8 è riportato il circuito equivalente al BJT NPN secondo il modello di Ebers-Moll. I due diodi nel modello rappresentano le giunzioni base-emettitore e base-collettore. Le equazioni che caratterizzano il modello sono le seguenti:

$$I_E = I_{ES} \left( e^{\frac{V_{BE}}{V_T}} - 1 \right) \quad (11)$$

$$I_C = \alpha_0 I_{ES} \left( e^{\frac{V_{BE}}{V_T}} - 1 \right), \quad (12)$$

dove  $I_E$  è la corrente all'emettitore,  $I_C$  è la corrente al collettore,  $\alpha_0$  è il guadagno in corrente in corto circuito, in configurazione a base comune ( $\alpha_0 \approx 1$ ),  $I_{ES}$  è la corrente inversa del diodo tra base ed emettitore e  $V_{BE}$  è la tensione tra

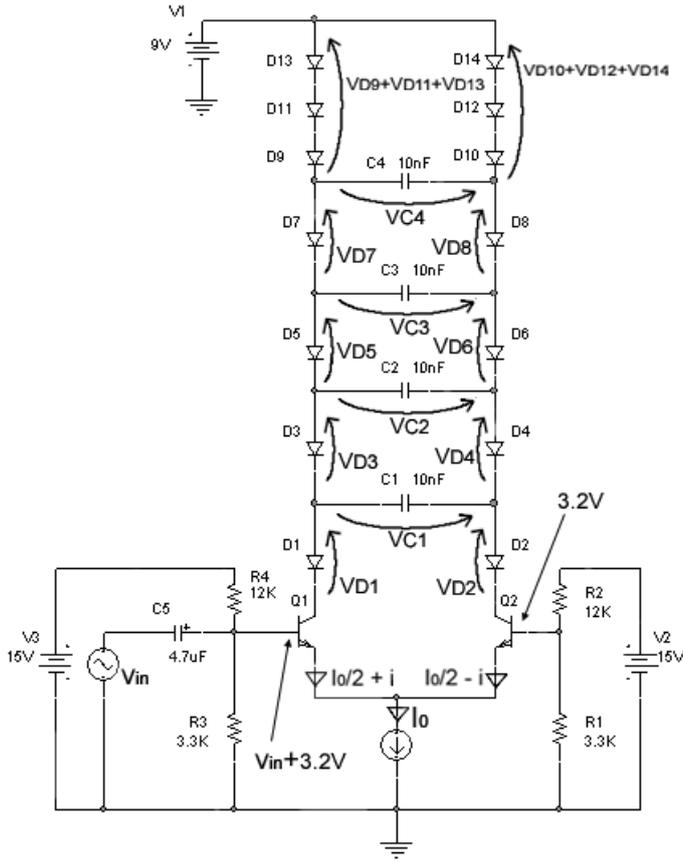


Figura 7. Schema circuitale del VCF EMS non lineare ad anello aperto.

base ed emettitore. Vale inoltre che:

$$\alpha_0 = I_C / I_E \quad (13)$$

considerando quindi  $\alpha_0 \approx 1$ , si evince che l'intensità della corrente di collettore può essere considerata pari a quella della corrente di emettitore; di conseguenza, la corrente di base può essere considerata nulla. Pertanto l'equazione descrittiva fondamentale per il modello risulta la (11). Volendo a questo punto applicare il modello di Ebers-Moll al VCF ad anello aperto, bisogna innanzitutto capire come esso consenta di descrivere tensioni e correnti in gioco nel circuito che sta alla base del VCF stesso, cioè l'amplificatore differenziale a coppia long-tailed basato sui transistor Q1 e Q2. Si veda a tal proposito Figura 9. Applicando la relazione (11) rispettivamente a Q1 e a Q2, si ottiene:

$$\begin{aligned} i_{C1} &= I_{ES1} \left( e^{\frac{v_{BE1}}{V_T}} - 1 \right) \approx I_{ES1} \left( e^{\frac{v_{BE1}}{V_T}} \right) \\ i_{C2} &= I_{ES2} \left( e^{\frac{v_{BE2}}{V_T}} - 1 \right) \approx I_{ES2} \left( e^{\frac{v_{BE2}}{V_T}} \right) \end{aligned} \quad (14)$$

Quindi:

$$v_{BE1} = V_T \ln \left( \frac{i_{C1}}{I_{ES1}} \right) \quad \text{e} \quad v_{BE2} = V_T \ln \left( \frac{i_{C2}}{I_{ES2}} \right) \quad (15)$$

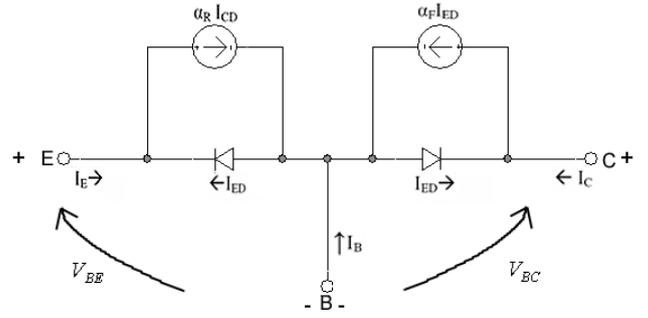


Figura 8. Circuito equivalente al BJT NPN proposto dal modello di Ebers-Moll.

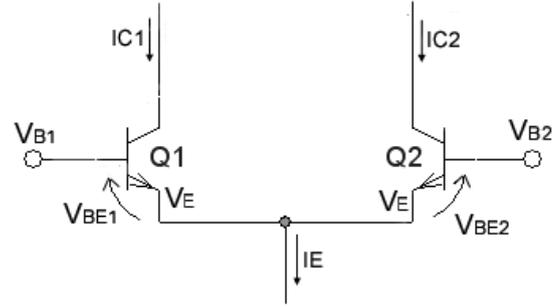


Figura 9. Il modello di Ebers-Moll applicato all'amplificatore a coppia long-tailed.

Risulta ora interessante studiare il comportamento dell'amplificatore differenziale prendendo come ingresso la differenza delle due tensioni base-emettitore di Q1 e Q2. Pertanto si procede calcolando:

$$v_{BE2} - v_{BE1} = V_T \ln \left( \frac{i_{C2} \cdot I_{ES1}}{i_{C1} \cdot I_{ES2}} \right) \quad (16)$$

inoltre, assumendo che la corrente inversa della giunzione tra base ed emettitore di Q1 sia uguale a quella di Q2 (questo è possibile adottando una coppia di transistor selezionati, *matched-pair* in inglese), ovvero:  $I_{ES1} = I_{ES2}$ , si ottiene:

$$v_{BE2} - v_{BE1} = V_T \ln \left( \frac{i_{C2}}{i_{C1}} \right) \quad (17)$$

Si consideri ora quanto segue:

$$v_{BE2} - v_{BE1} = (v_{B2} - v_E) - (v_{B1} - v_E),$$

dal momento che i due emettitori sono direttamente collegati tra loro, si verifica che:

$$v_{BE2} - v_{BE1} = v_{B2} - v_{B1} \quad (18)$$

e, calcolando l'esponenziale di entrambi i membri Eq. (17):

$$e^{\frac{v_{B2} - v_{B1}}{V_T}} = \frac{i_{C2}}{i_{C1}} \quad (19)$$

Quindi, come suggerito anche in [10]:

$$i_E = i_{C1} + i_{C2} = i_{C1} + i_{C1} e^{\frac{v_{B2} - v_{B1}}{V_T}}, \quad (20)$$

da cui:

$$i_{C1} = \frac{i_E}{1 + e^{\frac{v_{B2} - v_{B1}}{V_T}}}$$

che si può riscrivere come:

$$\begin{aligned} i_{C1} &= \frac{2i_E}{2 \left( 1 + e^{\frac{v_{B2} - v_{B1}}{V_T}} \right)} \\ &= \frac{i_E}{2} \left( \frac{1 + 1 + e^{\frac{v_{B2} - v_{B1}}{V_T}} - e^{\frac{v_{B2} - v_{B1}}{V_T}}}{1 + e^{\frac{v_{B2} - v_{B1}}{V_T}}} \right) \\ &= \frac{i_E}{2} \left( 1 + \frac{1 - e^{\frac{v_{B2} - v_{B1}}{V_T}}}{1 + e^{\frac{v_{B2} - v_{B1}}{V_T}}} \right) \end{aligned}$$

dal momento che:

$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1}$$

operando le dovute sostituzioni, si ottiene:

$$\begin{aligned} i_{C1} &= \frac{i_E}{2} \left[ 1 - \tanh \left( \frac{v_{B2} - v_{B1}}{2V_T} \right) \right] \\ &= \frac{i_E}{2} \left[ 1 + \tanh \left( \frac{v_{B1} - v_{B2}}{2V_T} \right) \right] \end{aligned} \quad (21)$$

e

$$\begin{aligned} i_{C2} = i_E - i_{C1} &= i_E - \frac{i_E}{2} \left[ 1 + \tanh \left( \frac{v_{B1} - v_{B2}}{2V_T} \right) \right] \\ &= \frac{i_E}{2} \left[ 1 - \tanh \left( \frac{v_{B1} - v_{B2}}{2V_T} \right) \right] \end{aligned} \quad (22)$$

Le ultime due relazioni sono state sfruttate per ottenere il sistema di equazioni differenziali non lineari che costituisce il modello del VCF in regime di grandi segnali. A tal scopo è stato applicato il principio di Kirchoff alle quattro maglie costituite dalle celle diodi-capacità, ottenendo il seguente sistema di equazioni algebriche:

$$\begin{cases} v_{D3} + v_{C2} - v_{D4} - v_{C1} = 0 \\ v_{D5} + v_{C3} - v_{D6} - v_{C2} = 0 \\ v_{D7} + v_{C4} - v_{D8} - v_{C3} = 0 \\ 3v_{D9} - 3v_{D10} - v_{C4} = 0 \end{cases}, \quad (23)$$

si veda Figura 8. Applicando le proprietà dei logaritmi alle equazioni costituenti i modelli dei componenti precedentemente introdotti e considerando che  $C_1 = C_2 = C_3 = C_4 = C$ , dopo alcune manipolazioni algebriche il sistema

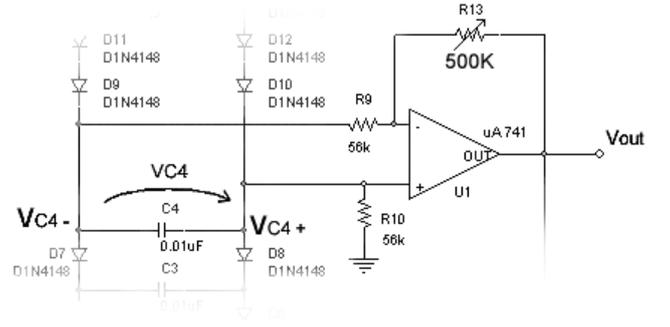


Figura 10. L'amplificatore differenziale usato nel VCF come circuito di retroazione a guadagno variabile.

(23) può essere riscritto come segue:

$$\begin{cases} \dot{v}_{C1} = \frac{I_0}{2C} \frac{1 - e^{\frac{v_{C1} - v_{C2}}{\gamma}}}{1 + e^{\frac{v_{C1} - v_{C2}}{\gamma}}} + \frac{I_0}{2C} \left( 1 + \tanh \left[ \frac{v_{in}}{2V_T} \right] \right) \\ \dot{v}_{C2} = \frac{I_0}{2C} \left( \frac{1 - e^{\frac{v_{C1} - v_{C2}}{\gamma}}}{1 + e^{\frac{v_{C1} - v_{C2}}{\gamma}}} + \frac{1 - e^{\frac{v_{C2} - v_{C3}}{\gamma}}}{1 + e^{\frac{v_{C2} - v_{C3}}{\gamma}}} \right) \\ \dot{v}_{C3} = \frac{I_0}{2C} \left( \frac{1 - e^{\frac{v_{C2} - v_{C3}}{\gamma}}}{1 + e^{\frac{v_{C2} - v_{C3}}{\gamma}}} + \frac{1 - e^{\frac{v_{C3} - v_{C4}}{\gamma}}}{1 + e^{\frac{v_{C3} - v_{C4}}{\gamma}}} \right) \\ \dot{v}_{C4} = \frac{I_0}{2C} \left( \frac{1 - e^{\frac{v_{C3} - v_{C4}}{\gamma}}}{1 + e^{\frac{v_{C3} - v_{C4}}{\gamma}}} + \frac{1 - e^{\frac{v_{C4}}{3\gamma}}}{1 + e^{\frac{v_{C4}}{3\gamma}}} \right) \end{cases} \quad (24)$$

dove  $\gamma$  è una quantità ausiliaria definita per mantenere la notazione più leggibile:

$$\gamma \triangleq \eta \cdot V_T \approx 1.836 \cdot 0.026 \approx 0.048, \quad (25)$$

Il sistema (24) rappresenta il modello del filtro non lineare ad anello aperto e ne descrive il comportamento in regime di grandi segnali.

### MODELLO NON LINEARE AD ANELLO CHIUSO

Come precedentemente anticipato, il VCF è dotato di un circuito di retroazione a guadagno variabile basato sull'amplificatore operazionale U1, che è stato modellato con le usuali assunzioni di linearità. In particolare, questa parte del circuito (riportata in Figura 10) può essere descritta applicando il principio di sovrapposizione degli effetti. In questo modo, si ottiene la seguente relazione di ingresso-uscita:

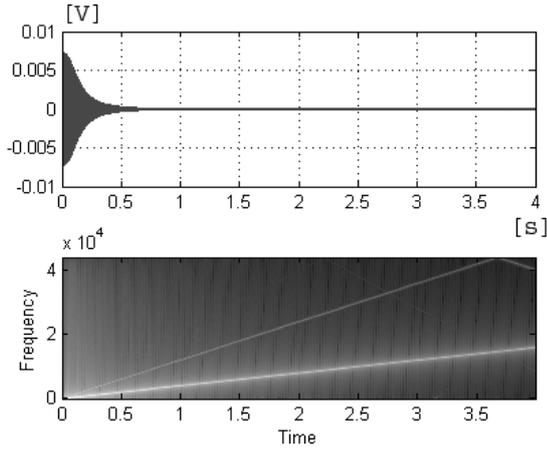
$$v_{out} = v_{C4+} \frac{R_9 + R_{13}}{R_9} - v_{C4-} \frac{R_{13}}{R_9}. \quad (26)$$

Considerando ora come si comportano correnti e tensioni in un amplificatore differenziale a coppia long-tailed e applicando tali considerazioni ai due rami del ladder che poggiano su Q1 e Q2, si determina quanto segue:

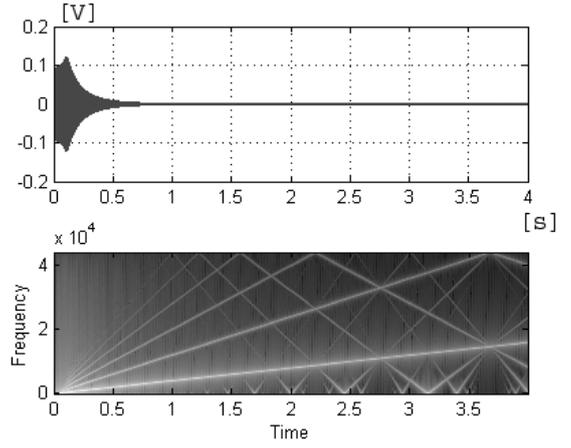
$$v_{C4+} = -v_{C4-}.$$

da cui:

$$v_{C4} = v_{C4+} - v_{C4-} = 2v_{C4+}.$$



**Figura 11.** Sweep sinusoidale lineare da 20 Hz a 8 kHz, ampiezza di picco pari a 10 mV.  $I_0 = 30 \mu\text{A}$ , senza retroazione ( $R_{13} = 0$ ).



**Figura 12.** Sweep sinusoidale lineare da 20 Hz to 8 kHz, ampiezza di picco pari a 100 mV.  $I_0 = 30 \mu\text{A}$ , guadagno di retroazione  $R_{13}/R_9 = 2$ .

Quindi, Eq. (26) può essere così riscritta:

$$\begin{aligned} v_{out} &= v_{C_{4+}} \frac{R_9 + R_{13}}{R_9} + v_{C_{4+}} \frac{R_{13}}{R_9} \\ &= v_{C_{4+}} + 2v_{C_{4+}} \frac{R_{13}}{R_9} \\ &= \frac{v_{C_4}}{2} + v_{C_4} \frac{R_{13}}{R_9}. \end{aligned} \quad (27)$$

Dove il termine  $R_{13}/R_9$  rappresenta il guadagno del circuito di retroazione. Ora, dal momento che  $R_{13}$  varia da 0  $\Omega$  a 500 k $\Omega$  e  $R_9 = 56$  k $\Omega$ , tale guadagno varia da 0 a circa 10. Considerando Eq. (21), la relazione (27) può essere sostituita direttamente nel sistema (24) ottenendo:

$$\begin{cases} \dot{v}_{C_1} = \frac{I_0}{2C} \left( \frac{1-e^{-\frac{v_{C_1}-v_{C_2}}{\gamma}}}{1+e^{-\frac{v_{C_1}-v_{C_2}}{\gamma}}} + \tanh \left[ \frac{v_{in} - \frac{v_{C_4}}{2} - v_{C_4} \frac{R_{13}}{R_9}}{2V_T} \right] + 1 \right) \\ \dot{v}_{C_2} = \frac{I_0}{2C} \left( \frac{1-e^{-\frac{v_{C_1}-v_{C_2}}{\gamma}}}{1+e^{-\frac{v_{C_1}-v_{C_2}}{\gamma}}} + \frac{1-e^{-\frac{v_{C_2}-v_{C_3}}{\gamma}}}{1+e^{-\frac{v_{C_2}-v_{C_3}}{\gamma}}} \right) \\ \dot{v}_{C_3} = \frac{I_0}{2C} \left( \frac{1-e^{-\frac{v_{C_2}-v_{C_3}}{\gamma}}}{1+e^{-\frac{v_{C_2}-v_{C_3}}{\gamma}}} + \frac{1-e^{-\frac{v_{C_3}-v_{C_4}}{\gamma}}}{1+e^{-\frac{v_{C_3}-v_{C_4}}{\gamma}}} \right) \\ \dot{v}_{C_4} = \frac{I_0}{2C} \left( \frac{1-e^{-\frac{v_{C_3}-v_{C_4}}{\gamma}}}{1+e^{-\frac{v_{C_3}-v_{C_4}}{\gamma}}} + \frac{1-e^{-\frac{v_{C_4}}{3\gamma}}}{1+e^{-\frac{v_{C_4}}{3\gamma}}} \right) \end{cases} \quad (28)$$

Il sistema (28) descrive il VCF EMS ad anello chiuso, in regime di grandi segnali.

### SIMULAZIONI NUMERICHE E RISULTATI (VCF EMS)

Una volta ottenuto il modello, si è proceduto alla sua simulazione in ambiente Matlab: per quanto riguarda l'integrazione numerica, ci si è avvalsi del noto metodo di Runge-Kutta esplicito del quarto ordine. In particolare, riscrivendo il sistema (28) in questo modo:

$$\dot{y} = F(t, v_C)$$

si ottiene la classica rappresentazione dello schema del metodo:

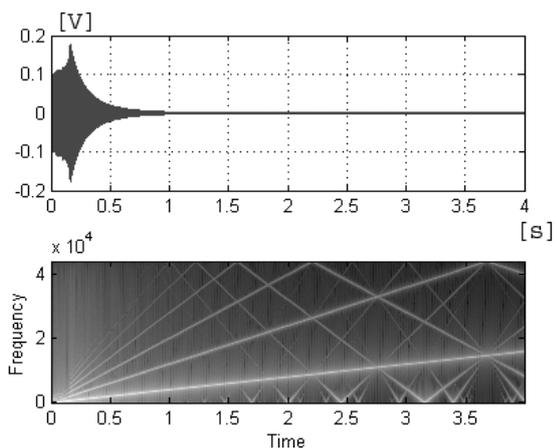
$$\begin{aligned} K_0 &= hF(t_i, v_{C_i}) \\ K_1 &= hF\left(t_i + \frac{h}{2}, v_{C_i} + \frac{K_0}{2}\right) \\ K_2 &= hF\left(t_i + \frac{h}{2}, v_{C_i} + \frac{K_1}{2}\right) \\ K_3 &= hF(t_i + h, v_{C_i} + K_2) \end{aligned} \quad (29)$$

dove  $t_i$  è l' $i$ -esimo passo di integrazione e  $v_{C_i}$  è un vettore  $4 \times 1$  rappresentante le tensioni ai capi dei quattro condensatori all' $i$ -esimo passo di integrazione. Tali tensioni possono quindi essere aggiornate nel seguente modo:

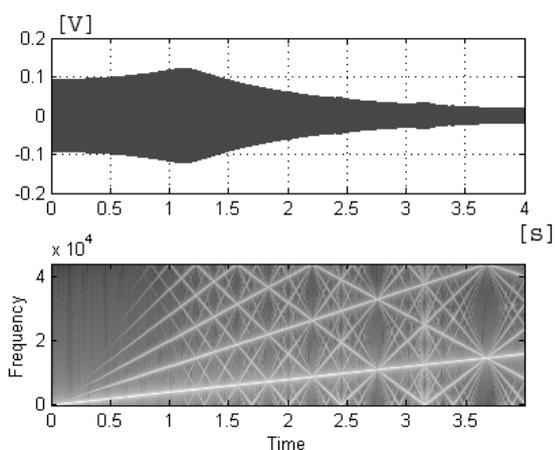
$$v_{C_{i+1}} = v_{C_i} + \frac{1}{6} (K_0 + 2K_1 + 2K_2 + K_3) \quad (30)$$

Osservando il sistema (28) si evince la presenza della retroazione nel modello: ad ogni passo di integrazione è infatti necessario conoscere a priori il valore della caduta di tensione ai capi di  $C_4$ . Al fine di determinare tale valore a meno di una soglia prefissata, è stato implementato uno schema di iterazione funzionale basato sulla tecnica del punto fisso: tale schema viene lanciato ad ogni passo di integrazione.

Le Figure da 11 a 14 mostrano i risultati di quattro simulazioni ottenute ponendo in ingresso al modello sweep sinusoidali da 20 Hz a 8kHz aventi ampiezza fissa e durata pari a 4 s; la frequenza di campionamento è  $F_s = 88.2$  kHz. In ciascuna immagine sono visibili: in alto la forma d'onda del segnale in uscita, ed in basso lo spettrogramma dello stesso. La caratteristica passabasso del sistema è ben visibile in tutte le figure riportate: l'ampiezza dell'uscita decresce infatti col passare del tempo, mentre la frequenza del segnale aumenta. Si rileva inoltre come all'aumentare dell'ampiezza del segnale in ingresso corrisponda un aumento della distorsione armonica da parte del modello. Le ultime tre figure mostrano invece l'effetto della variazione del guadagno di retroazione, determinato dal valore della resistenza variabile  $R_{13}$  e l'effetto della variazione dell'intensità della corrente di controllo  $I_0$ : come previsto, ad un aumento di quest'ul-



**Figura 13.** Sweep sinusoidale lineare da 20 Hz to 8 kHz, ampiezza di picco pari a 100 mV.  $I_0 = 30 \mu\text{A}$ , guadagno di retroazione  $R_{13}/R_9 = 5$ .



**Figura 14.** Sweep sinusoidale lineare da 20 Hz to 8 kHz, ampiezza di picco pari a 100 mV.  $I_0 = 300 \mu\text{A}$ , guadagno di retroazione  $R_{13}/R_9 = 2$ .

tima corrisponde un aumento della frequenza di taglio del filtro. Le figure appena commentate mostrano inoltre come la versione numerica del filtro sia affetta da aliasing: questo problema verrà affrontato nelle successive implementazioni del modello.

Il codice Matlab del VCF è stato infine tradotto in linguaggio C, sotto forma di modulo external per l'ambiente di simulazione e sintesi in tempo reale Pure Data: in questo modo è stato possibile appurare che il numero di iterazioni di punto fisso necessarie per raggiungere la convergenza non è mai abbastanza alto da generare artefatti udibili nel segnale in uscita. La piattaforma di riferimento è un Apple MacBook con processore Intel Core 2 Duo a 2.0 GHz e sistema operativo MacOS X v. 10.4.

## RIFERIMENTI

1. F. Fontana. Computation of linear filter networks containing delay-free loops, with an application to the waveguide mesh. *IEEE Trans. on Speech and Audio Processing*, 11(6):774–782, Nov. 2003.
2. F. Fontana, F. Avanzini, and D. Rocchesso. Computation of nonlinear filter networks containing delay-free paths. In *Proc. Conf. on Digital Audio Effects (DAFx-04)*, pages 113–118, Naples, Italy, Oct. 2004.
3. F. Avanzini, F. Fontana, and D. Rocchesso. Efficient computation of nonlinear filter networks with delay-free loops and applications to physically-based sound models. In *Proc. of The Fourth International Workshop on Multidimensional Systems, 2005 (NDS 2005)*, pages 110–115, Wuppertal, Germany, July 2005. IEEE.
4. A. Härmä. Implementation of frequency-warped recursive filters. *EURASIP Signal Processing*, 80(3):543–548, Mar. 2000.
5. J. Szczupak and S. K. Mitra. Detection, location, and removal of delay-free loops in digital filter configurations. *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-23(6):558–562, 1975.
6. C. Wan and A. M. Schneider. Further improvements in digitizing continuous-time filters. *IEEE Trans. on Signal Processing*, 45(3):533–542, Mar. 1997.
7. R. A. Moog, “A voltage-controlled low-pass high-pass filter for audio signal processing,” in *Audio Engineering Society Convention*. AES, Oct. 1965, Preprint 413.
8. T. Stilson and J. O. Smith, “Analyzing the Moog VCF with considerations for digital implementation,” in *Proc. Int. Computer Music Conf.*, Hong Kong, 1996.
9. A. Huovilainen, “Nonlinear digital implementation of the Moog ladder filter,” in *Proc. Conf. on Digital Audio Effects (DAFx-04)*, Naples, Italy, Oct. 2004, pp. 61–64.
10. T.E. Stinchcombe, “Derivation of the Transfer Function of the Moog Ladder Filter,” technical report, 2005
11. F. Fontana, “Preserving the structure of the moog VCF in the digital domain,” in *Proc. Int. Computer Music Conf.*, Copenhagen, Denmark, 27–31 Aug. 2007.
12. M. Civolani and F. Fontana, “A nonlinear digital model of the EMS VCS3 voltage-controlled filter,” in *Proc. Conf. on Digital Audio Effects (DAFx-08)*, Espoo, Finland, Sep. 2008.
13. R. C. Jaeger and T. N. Blalock, *Microelectronic Circuit Design*, McGraw-Hill, New York, 2nd edition, 2004.
14. F. Fontana and F. Avanzini, “Computation of delay-free nonlinear digital filter networks. Application to chaotic circuits and intracellular signal transduction,” *IEEE Trans. on Signal Processing*, 56(10), Oct. 2008.

# La ricostruzione dello Studio di Fonologia della Rai di Milano

Maddalena Novati  
Consulente musicale RAI  
Produzione Radiofonia

## SOMMARIO

Il 17 settembre 2008 alle ore 18 si è ufficialmente aperta la sala XXXVI del Museo degli Strumenti Musicali presso il Castello Sforzesco di Milano dedicata allo Studio di Fonologia Musicale di Milano della RAI.

In essa sono esposti gli apparati originali nell'assetto del 1968 ed è possibile visionare la storia dello Studio attraverso un maxischermo e tre monitor touchscreen che consentono la consultazione di un ricco database contenente illustrazioni, filmati, quasi mille fotografie, esempi audio e partiture. nei prossimi anni il progetto prevede di trasportare presso il Museo anche l'intero archivio della fonologia comprendente corrispondenza, foto, partiture, manoscritti, i circa 400 nastri audio che contengono oltre 200 ore di musica, nonché l'attuazione del secondo riversamento audio in digitale a 24 bit e 96 kHz che sarà curato dal Laboratorio Audio del Centro di Produzione TV di Milano della RAI e dalla Produzione Radio.

## Parole Chiave

Musica Elettronica; Conservazione beni musicali.

## LO STUDIO DI FONOLOGIA MUSICALE DI MILANO DELLA RAI

Lo Studio di Fonologia Musicale di Milano della RAI, progettato dal fisico Alfredo Lietti, nasce ufficialmente nel giugno del 1955 ad opera dei musicisti Luciano Berio e Bruno Maderna e resta in attività fino al 28 febbraio 1983, quando il tecnico Marino Zuccheri va in pensione.

La prima composizione realizzata per convincere i dirigenti RAI ad aprire lo Studio fu *Ritratto di città* su testo di Roberto Leydi, musica di Luciano Berio e Bruno Maderna[1].

Lo Studio nasce principalmente per due scopi: la produzione sperimentale di musica elettronica e la realizzazione di commenti e colonne sonore per la radio e la televisione.

Berio attingeva spunti e informazioni guardando all'America (incuriosito dalla *tape music* di Ussachevsky e Luening) e alla Francia (attraverso la sua amicizia con Schaeffer e il Club d'Essai); Maderna portava i contributi dallo studio di Colonia attraverso la sua amicizia con Stockhausen e Meyer-Eppler, e dai corsi estivi di Darmstadt.

*"Luciano Berio ed io abbiamo avuto la possibilità di fondare a Milano uno Studio di Musica Elettronica. L'esperienza più palesemente positiva per noi, fino ad ora, è stato l'incontro fra tecnici e noi musicisti. I tecnici, infatti, ci sono venuti incontro con tale interesse e comprensione da fare loro i nostri stessi desideri."* In queste lusinghiere parole, scritte da Maderna nel 1956, è riassunto tutto lo spirito non solo di collaborazione ma di assoluta amicizia che fin dall'inizio aveva impregnato di sé la sperimentazione presso lo Studio di Fonologia Musicale della RAI di Milano. Nella *repubblica veneta del quinto piano di Corso Sempione* la lingua ufficiale dello studio è il veneto, come ricordava spesso Marino Zuccheri riferendosi sia ai colleghi tecnici che ai compositori, con l'eccezione di Berio, la *contaminazione ligure*.

All'inizio le apparecchiature a disposizione sono semplicemente alcuni magnetofoni, giradischi con i quali poter cambiare velocità, qualche filtro, un oscillatore e le Onde Martenot. Il salto di qualità avviene l'anno successivo con l'acquisto dei mitici nove oscillatori e con la voce di Cathy Berberian considerato il *decimo oscillatore*: con lei Berio realizza fra l'altro *Thema – (Omaggio a Joyce)* e *Visage*.

Fra i musicisti che lavorarono o frequentarono lo Studio ricordiamo, fra gli italiani, Berio, Maderna, Nono, Castiglioni, Clementi, Donatoni, Gentilucci, Manzoni, Marinuzzi Jr., Paccagnini, Sciarrino, Sinopoli, Togni, Vlad. e fra gli ospiti stranieri soprattutto Cage e Pousseur.

Tra le prime composizioni di musica elettronica spiccano *Mutazioni* di Luciano Berio, *Continuo* e *Notturmo* di Bruno Maderna [2]. A questi due autori si devono inoltre le successive *Thema (Omaggio a Joyce)* e *Visage* (Berio); *Syntaxis* e *Invenzione su una voce* (Maderna).

Le più singolari colonne sonore e sonorizzazioni per produzioni radiofoniche e televisive sono *La Loira*, *Peter Pan*, *La fanciulla di neve*, *L'uccellino azzurro* (con musiche di L. Berio); *Macbeth*, *L'augellin Belverde*, *il Cavallo di Troia*, *L'altro mondo ovvero Gli stati e imperi del Sole e della Luna* (con musiche di B. Maderna)[3].

Inoltre sono stati registrati e/o elaborati programmi partecipanti al Prix Italia (il concorso internazionale che premia i migliori prodotti radiotelevisivi) ed ideate tecniche nuove ed effetti

particolari per brani di musica leggera ad opera soprattutto di Mario Migliardi.

Il Prix Italia fu vinto per ben sette volte dalla Fonologia con opere (interamente realizzate o semplicemente transitate per essere elaborate elettronicamente) come *La notte di un nevrastenico* (Rota, 1959), *Attraverso lo specchio* (Castiglioni, 1961), *Il dio di oro* (Paccagnini, 1964), *Johann Sebastian* (Negri, 1967), *La fine del mondo* (Negri, 1969), *Giochi di fanciulli* (Liberovici, 1970), *Perelà* (Liberovici, 1971), *Ages* (Maderna, 1972), *Diario Imaginario* (Berio, 1975), *Love's Body* (Renosto, 1974), *La voce dell'inferno* (Sciarrino, 1981).

A Fonologia hanno collaborato scrittori prestigiosi come Sermonti, Quasimodo, Eco e registi come Bachelletti, Pressburger, Puecher. Non va dimenticata la collaborazione con il Piccolo Teatro appena fondato da Giorgio Strehler e Paolo Grassi per il quale la Fonologia, attraverso le composizioni di Fiorenzo Carpi, ha dato il suo contributo nella elaborazione delle musiche di scena per gli spettacoli più innovativi.

#### LA LIUTERIA DEL XX SECOLO OSPITATA NELLA SALA XXXVI DEL CASTELLO SFORZESCO

Il 17 settembre 2008 alle ore 18 si è ufficialmente aperta la sala XXXVI del Museo degli Strumenti Musicali presso il Castello Sforzesco dedicata allo Studio di Fonologia Musicale di Milano della RAI. L'iniziativa si è potuta concretizzare grazie al contributo di MITO Associazione Festival Internazionale Milano Torino - Festival Settembre Musica.

Grazie a MITO si sono potute trasferire tutte le apparecchiature dell'ex Studio di Fonologia Musicale di Milano dalla sede RAI al Museo degli Strumenti Musicali del Castello Sforzesco e rendere disponibile tale patrimonio al pubblico degli appassionati e degli studiosi (fig. 1).



Fig. 1. Trasloco delle apparecchiature al Museo degli Strumenti Musicali

La RAI (Centro di Produzione TV di Milano e Direzione Radio) e le Civiche Raccolte del Castello Sforzesco hanno potuto così ricreare nella sala XXXVI del Museo degli Strumenti Musicali l'intero Studio di Fonologia con gli arredi originali Giò Ponti e quasi tutto l'originale ambiente di lavoro in modo tale da accogliere per la prima volta in un Museo strumenti e opere storicamente correlati<sup>1</sup>.

La storia della Fonologia è illustrata attraverso tre monitor touchscreen e un maxi-schermo: il visitatore toccando i monitor potrà fare ricerca e consultare l'intero archivio digitale delle opere e delle macchine, ascoltare in cuffia esempi musicali e vedere filmati in cui Berio e Maderna insieme a Marino Zuccheri lavorano e sviluppano le loro composizioni<sup>2</sup>.



Fig. 2. Schermata indice del database

La consultazione è possibile attraverso un potente database<sup>3</sup> (fig. 2) riccamente illustrato e corredato

<sup>1</sup> Il progetto di allestimento è stato curato dallo Studio De Lucchi, architetti Laura Parolin e Angelo Micheli. In RAI sono stati conservati e ritrovati sia i tavolini originali, sia la rastrelliera dove alloggiavano le bobine in attesa di lavorazione, sia il portacavi, la lavagna, il portacenere, il cestino, l'attaccapanni, la lampada e quasi tutti questi arredi sono stati inseriti nella loro identica collocazione. Si è anche trasferita una intera porta Giò Ponti che chiudeva gli studi di registrazione con il suo telaio.

<sup>2</sup> Il database contiene 115 esempi audio, 43 esempi video, 144 schede riguardanti i compositori, 527 schede riguardanti i brani, 592 riguardanti le bobine, 57 gli strumenti, 441 nomi in anagrafica; inoltre vi sono 777 foto suddivise in 109 riguardanti gli strumenti, 350 riguardanti le persone, 267 riguardanti le bobine, 51 i brani, fra cui 42 partiture e 9 foto tratte da articoli del Radiocorriere.

<sup>3</sup> i cui contenuti sono stati da me curati per la parte delle opere musicali, della ricerca iconografica e filmica, da Giovanni Caprioli per la maggior parte delle biografie e per la parte organologica dall'ingegnere Antonio Rodà

di filmati, con quasi mille fotografie, esempi audio e partiture: si tratta di un'applicazione web-based che visualizza informazioni aggregate e oggetti multimediali relativi ai singoli compositori, alle singole opere e agli apparati, elaborata dal Laboratorio LIM dell'Università Statale di Milano<sup>4</sup>. Il laboratorio LIM ha curato anche la nuova applicazione multimediale, che ospiterà la ricerca su *Serenata IV* di Maderna della dottoressa Angela Ida De Benedictis dove saranno per la prima volta correlate in tempo reale partiture diverse a differenti audio, con una sincronizzazione evoluta (realizzata con MX e un linguaggio di programmazione) particolarmente adatta a situazioni in cui sia disponibile una grande quantità di materiali alternativi (ad esempio, più partiture, più audio, più schemi relativi allo stesso brano).

Per quanto riguarda gli apparati, poiché il posizionamento della macchine è mutato svariate volte nel corso degli anni, la ricostruzione attuale, basandosi su foto e filmati, ricrea l'assetto dopo il 1968.

Nel database degli strumenti è possibile ripercorrere l'evoluzione tecnologia di ogni singolo apparato e la sua collocazione nelle versioni originarie del 1955/56 e nel successivo riassetto del 1968.

Nella sala sono esposti i seguenti dispositivi (fig. 3).

Le apparecchiature per la generazione di suoni elettronici: generatori di onde sinusoidali, generatori di onde quadre, generatore di impulsi, generatore di rumore bianco.



Fig. 3 Allestimento della Sala XXXVI

Le apparecchiature per la trasformazione e la combinazione dei suoni: modulatori, variatori di

del CSC Centro di Sonologia Computazionale dell'Università di Padova

<sup>4</sup> Il database è stato sviluppato dal LIM, dott. Goffredo Haus, dott. Luca Ludovico, dott. Adriano Baratè, Antonello D'Aguanno.

velocità e di durata, traslatori di frequenza, differenti tipi filtri e di eco, miscelatori a più canali. Le apparecchiature per la registrazione e la riproduzione del suono: due magnetofoni a quattro piste (gli unici per il momento rimasti poiché non sono stati ancora recuperati i magnetofoni a una e due piste).

Le apparecchiature per l'ascolto: cinque altoparlanti originali Tannoy che permettono di ascoltare il suono registrato e riprodotto dai magnetofoni o prodotto dai generatori di suono.

Il progetto da me curato si propone nei prossimi anni di trasportare presso il Museo degli Strumenti Musicali anche l'intero archivio della FONOLOGIA comprendente corrispondenza, foto, partiture e manoscritti ma soprattutto i circa 400 nastri audio da un quarto di pollice mono o stereo e nastri da un pollice a quattro tracce, registrati con tecnica analogica, per un totale di oltre 200 ore di musica. Il progetto prevede anche l'attuazione del secondo riversamento audio in digitale a 24 bit e 96 kHz che sarà curato dal Laboratorio Audio del Centro di Produzione TV di Milano della RAI e dalla Produzione Radio.

Avevo fatto una promessa a Luciano Berio nel 1996 durante le prove di *Outis* al Teatro alla Scala di Milano: ricostruire un'immagine il più possibile completa dello Studio di Fonologia Musicale di Milano della RAI recuperando da una parte i materiali musicali (nastri, partiture, progetti) e dall'altra le apparecchiature<sup>5</sup>.

Ci sono riuscita: le macchine sono state recuperate dalla sede RAI di Torino (dove erano state alloggiate in attesa dell'apertura di un mai attuato Museo della Rai)<sup>6</sup>, i nastri mancanti in buona parte sono stati individuati presso altri Archivi pubblici o privati e confluiti sotto forma di copie digitali nell'Archivio RAI.

L'archivio si compone dunque di un fondo storico e di un fondo acquisito.

Nel fondo storico sono depositati sia i nastri contenenti le musiche composte a Fonologia sia i nastri provenienti da altri studi di musica elettronica come Colonia, Parigi o Varsavia che Berio e Maderna richiedevano per ascolto e/o

<sup>5</sup> Il mio ruolo di responsabile dell'Archivio di Fonologia iniziò nel 1995 quando il mio diretto superiore Maestro Luigi Galvani andò in pensione e mi affidò le chiavi del mitico armadio contenente i nastri della Fonologia. Da allora ho cercato di recuperare tutte le informazioni possibili sulla storia e sulla produzione dello Studio, iniziando la catalogazione completa dei nastri presenti e la loro informatizzazione e digitalizzazione.

<sup>6</sup> Il progetto per il recupero delle macchine che giacevano nello scantinato della sede RAI di Torino, sono riuscita a realizzarlo grazie al dott. Massimo Ferrario, dirigente responsabile del Centro di Produzione TV di Milano che nel 2003 le fece trasferire a Milano e grazie alla collaborazione di Giovanni Belletti del Laboratorio Audio.

trasmissione radiofonica, sia copie di registrazioni di concerti diretti da loro stessi<sup>7</sup>.

Il fondo acquisito (contraddistinto dalla sigla di catalogazione Z) comprende invece le copie digitali di nastri recuperati presso i seguenti archivi: Archivio Ricordi, Archivio Paccagnini, Archivio di Tempo Reale, Nastroteca centrale della RAI.

La ricerca ovviamente continua e il lavoro di recupero e di conservazione dei nastri sarà uno degli obiettivi primari dei prossimi anni da conseguire insieme alla catalogazione della corrispondenza e dei materiali iconografici.

Per quanto riguarda il primo riversamento di tutti i nastri analogici su supporti digitali a 16 bit e 44.1 e 48 KHz, l'ho terminato nel 1998 con i tecnici della Radiofonia<sup>8</sup>.

Nel 2005 ho iniziato i lavori di un ulteriore secondo riversamento in collaborazione con il Laboratorio Audio del Centro di Produzione di Milano della RAI ed il Laboratorio Mirage dell'Università di Udine sede di Gorizia<sup>9</sup>.

---

<sup>7</sup> Il mio censimento al 23-9-1996 contava 29 bobine con sigla A (materiali acustici di ascolto e di studio e musica strumentale), 40 con sigla E (materiali di sola musica elettronica), 65 con sigla Q (nastri a quattro piste), 42 con sigla R (musiche di sonorizzazione per la prosa radiotelevisiva e l'effettistica), e 211 con la nuova sigla FON: da me aggiunte per i nastri senza classificazione originale. Attualmente i nastri presenti nell'archivio sono oltre 600 poiché si sono aggiunte le 210 copie su dat (contrassegnati con la mia nuova schedatura copie Z) per i riversamenti da bobine originali presenti nella nastroteca centrale RAI di Roma o da ritrovamenti presso altri archivi.

<sup>8</sup> Nel 1991 Claudio Guglielmotto (allora capo delle risorse tecniche della radiofonia di Milano) affidò un primo incarico al tecnico Lucio Cavallarin di riversare in digitale una parte dei nastri: quelli siglati E su Dat 16 bit, 48 KHz e una parte di quelli siglati Q su supporto video D2, l'unico supporto digitale allora disponibile in RAI, in grado di registrare quattro piste sincronizzate – riversamento effettuato dal 1991 al 1993. Questo fu il primo ma parziale tentativo di restauro conservativo, il successivo, cioè il mio, invece portò a termine tutto il trasferimento in digitale dell'intero fondo analogico soprattutto con la collaborazione del tecnico Gianfranco Bendin.

<sup>9</sup> Laboratorio Audio è formato da Giovanni Belletti, Alberto Zanon e Fabio Ferrarini. Giovanni Belletti, in Rai dal 1972, cugino di Marino Zuccheri, è responsabile del Laboratorio Audio di Milano della RAI dal 1997; ebbe modo di frequentare lo studio di Fonologia approntando per L. Nono un congegno tecnico servito all'esecuzione delle musiche *Per Luigi Dallapiccola* presso il Teatro alla Scala di Milano. Per il Laboratorio Mirage invece hanno partecipato ai lavori il musicologo Giovanni De Mezzo, l'ing. Sergio Canazza e l'ing. Antonio Rodà.

Attualmente sono stati riversati tutti i nastri a quattro piste a 24 bit, 96 KHz.

## CONCLUSIONI

Sono passati 13 anni dal primo giorno in cui ho aperto il mitico armadio, molto lavoro mi attende ancora ma il bilancio può considerarsi positivo poiché ora le macchine e le opere saranno al sicuro da possibili dimenticanze o addirittura da incaute dismissioni. Tutti potranno vederle, studiarle, ascoltarle e mi piace pensare che Berio, Maderna, Nono, Lietti e Zuccheri facciano delle belle passeggiate sugli spalti del Castello in compagnia di Leonardo da Vinci magari progettando insieme congegni e musiche per gli ospiti illustri e godendo delle feste paradisiache ordinate da qualche rampollo discendente dagli Sforza.

## RIFERIMENTI

1. Veniero Rizzardi e Angela Ida De Benedictis (a cura di), "NUOVA MUSICA ALLA RADIO" Esperienze allo studio di Fonologia della RAI di Milano, 1954 - 1958. *Archivi della Musica Italiana Contemporanea*, n°. 1. Quaderni CIDIM, Roma, 2000.
2. Nicola Scaldaferrì, *Musica nel laboratorio elettroacustico. Lo Studio di Fonologia di Milano e la ricerca musicale negli anni Cinquanta*. Quaderni di Musica/Realtà - LIM, Lucca, 1997.
3. Angela Ida De Benedictis, *Radiodramma e arte radiofonica. Storia e funzioni della musica per radio in Italia*, De Sono/EDT, 2004.

# Un approccio alle interfacce acustiche tangibili basato sulla trasformata wavelet di onde dispersive

Antonio Camurri<sup>1</sup>, Corrado Canepa<sup>1</sup>, Chiara Erra<sup>1,2</sup>  
Alberto Massari<sup>1</sup>, Gualtiero Volpe<sup>1</sup>

<sup>1</sup>Casa Paganini - InfoMus Lab  
DIST  
Università degli Studi di Genova

Luciano Fava<sup>2</sup>

<sup>2</sup>Dipartimento di Scienze e Tecnologie Avanzate  
Università del Piemonte Orientale “Amedeo Avogadro”  
Alessandria

## SOMMARIO

La ricerca su interfacce acustiche tangibili ha come obiettivo la trasformazione di oggetti di uso quotidiano, come ad esempio un tavolo o una sedia, in superfici “sensibili” capaci di controllare applicazioni software attraverso il tatto. Superfici sensibili all’interazione tattile possono trasformarsi, così, in strumenti dove l’espressione musicale si realizza nel contatto e nella manipolazione dell’oggetto stesso. In questa direzione, la ricerca presentata in questo articolo espone un approccio basato sull’analisi modale delle onde di Lamb nelle lastre sottili, combinato con l’analisi wavelet del segnale dispersivo nel dominio tempo-frequenza allo scopo di migliorare l’algoritmo TDOA (Time Delay Of Arrival) per la localizzazione del tocco su un’interfaccia tangibile.

## Parole Chiave

nuove interfacce per l’espressione musicale; interfacce multimodali; interfacce acustiche tangibili.

## INTRODUZIONE

Negli ultimi anni, uno degli obiettivi della ricerca nel settore della interazione uomo-macchina è stata la realizzazione di nuove interfacce multimodali in grado di rendere più diretta e naturale l’interazione. Nel nostro lavoro, ci si propone di trasformare oggetti di uso quotidiano (un tavolo, una sedia), in superfici sensibili capaci di controllare applicazioni software solo attraverso il tatto. Riuscire a rendere una superficie “sensibile” all’interazione tattile da parte dell’utente vuol dire disporre di un nuovo oggetto che può divenire, ad esempio, uno strumento dove l’espressione musicale si realizza proprio attraverso la manipolazione e il contatto con lo strumento stesso [1, 2, 3]. Alcuni studi hanno affrontato il problema di rendere interattive superfici ampie, tali da non essere realizzabili con tecnologie touch-screen convenzionali (ad esempio [4]); altre direzioni della ricerca si sono focalizzate su nuove interfacce musicali (si veda [5, 6]). In [6], ad esempio, la ricerca mira al riconoscimento del modo espressivo in cui ha luogo l’interazione tattile, cercando una corrispondenza tra l’interazione fisica e la produzione del suono.

In questa prospettiva, “la tecnologia è tesa a interrogare la superficie e trasformarla in uno strumento creativo, come in una metamorfosi della sua funzione originaria” [7]. Migliore sarà la comprensione della fenomenologia delle vibrazioni causate dal tocco, maggiori saranno le informazioni che potranno permettere un utilizzo “espressivo” del segnale raccolto, donando il “senso del tatto” agli oggetti.

Più in dettaglio, obiettivi della ricerca sono il raggiungimento di un’accuratezza della localizzazione del tocco sulla superficie e l’estrazione di caratteristiche qualitative del segnale.

Un approccio alle *interfacce acustiche tangibili* è stato realizzato nel corso del progetto europeo IST TAI-CHI (Tangibile Acoustic Interfaces for Computer Human Interaction) [8].

Per la localizzazione della sorgente in un piano, pochi sensori di posizione nota sono sufficienti per calcolare il punto di contatto per mezzo di differenze nei tempi di arrivo (TDOA, Time Delay Of Arrival) del segnale dispersivo.

Usando la tecnica convenzionale di attraversamento di una soglia da parte degli impulsi rilevati ai sensori, possono generarsi degli errori come risultato del comportamento dispersivo del segnale [9]. Anche la cross-correlazione di due segnali causati dallo stesso evento acustico e rivelati in punti diversi della superficie possono condurre, in mezzi dispersivi, a risultati non utilizzabili. Se tuttavia dal segnale viene estratta una singola frequenza, le differenze di tempo ottenute da quella componente in frequenza possono fornire un miglioramento della qualità del risultato [10].

La trasformata wavelet è stata identificata come il modo più efficace di isolare nell’onda la componente in frequenza scelta ed è stata già applicata con successo nel campo dei Controlli Non Distruttivi (Non Destructive Testing, NDT). Le wavelet sono funzioni localizzate nel tempo e quindi sono adatte a ricostruire o a estrarre caratteristiche di segnali impulsivi con *sharp dynamics* [11]. Esse permettono inoltre di analizzare il segnale con differenti risoluzioni di scala, una proprietà particolarmente importante in applicazioni in presenza di rumore di fondo.

Basandosi su tali premesse, nei paragrafi successivi viene proposta una tecnica per interfacce acustiche tangibili.

Si noti che nei termini “analisi modale” e “interfacce multimodali”, l’aggettivo “modale” si riferisce a contesti e teorie differenti.

## TEORIA

### Onde nei solidi

Un piccolo impatto su di una superficie, ad esempio effettuato con un dito, una penna, una chiave, provoca un’onda acustica che si propaga all’interno del materiale.

Le onde che si propagano in una lastra isotropa sono costituite da una componente trasversale polarizzata verticalmente e da una componente longitudinale. Quando la lunghezza d’onda è molto più grande dello spessore della lastra, le due componenti sono accoppiate dalla riflessione delle due superfici libere, il

materiale si comporta come una guida d'onda e le onde che si generano vengono definite come onde di Lamb [12].

Il rapporto tra le ampiezze di queste onde parziali trasversali e longitudinali e la loro velocità di propagazione variano con la frequenza, cioè la propagazione è dispersiva. La teoria individua due modi: uno estensionale o simmetrico (S) e uno flessionale o antisimmetrico (A), le cui velocità di propagazione in funzione della frequenza sono definite dalle rispettive relazioni di dispersione:

$$c_e = (E / \rho(1 - \nu^2))^{1/2}$$

$$c_f = (Eh^2 / 12\rho(1 - \nu^2))^{1/4} \omega^{1/2}$$

dove  $E$  è il modulo di elasticità del materiale,  $\rho$  la densità,  $\nu$  il coefficiente di Poisson,  $h$  lo spessore della lastra e  $\omega$  la frequenza [13].

Questi due modi hanno caratteristiche fisiche e comportamento diverso, avendo componenti in frequenza e velocità di propagazione differenti. In particolare, il modo estensionale è non dispersivo: tutte le frequenze si propagano con la stessa velocità e la forma del fronte d'onda non cambia durante la propagazione; al contrario la componente flessionale, dispersiva, subirà un cambiamento nella forma del segnale propagato all'aumentare della distanza dal punto di impatto, con conseguente difficoltà nel riconoscimento e trattamento del segnale.

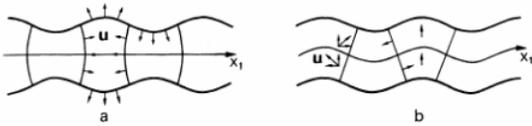


Fig 1. Onde di Lamb. (a) Simmetrica – da entrambi i lati del piano mediano, le componenti longitudinali sono uguali e le componenti trasversali sono opposte. (b) Antisimmetrica – da entrambi i lati del piano mediano, le componenti trasversali sono uguali e le componenti longitudinali sono opposte.

I vantaggi nell'utilizzo delle onde di Lamb sono (i) la riduzione dell'attenuazione dovuta alla diffrazione, poiché per le onde guidate l'energia decresce come  $r^{-1}$  e non come  $r^{-2}$ , e (ii) l'accessibilità su tutta la superficie al rilevamento degli spostamenti meccanici.

#### Algoritmo TDOA

La tecnica scelta per individuare la posizione di tocchi su di una lastra si basa sul tempo di ritardo (Time Delay Of Arrival TDOA) con cui il segnale perviene in posizioni note: dalla misura dei tempi di arrivo dell'onda prodotta dall'impatto sulla lastra rilevati da una serie di sensori posti a contatto della superficie in posizioni diverse e nota la velocità del suono nel materiale, si ricava la distanza percorsa dall'onda e infine la posizione.

Diversi metodi sono stati esaminati per estrarre i tempi di ritardo da una serie di segnali accelerometrici:

1. tempo di attraversamento di una soglia da parte del segnale
2. tempo di attraversamento di una soglia da parte dell'involuppo dell'onda
3. tempo di attraversamento di una soglia da parte della wavelet (reale) dell'onda
4. tempo di attraversamento di una soglia da parte della wavelet (complessa)

5. tempo a cui si registra il massimo della wavelet (complessa).

In tutti i casi il segnale è stato normalizzato rispetto al suo massimo, per tener conto dell'attenuazione dell'onda all'aumentare della distanza percorsa nel materiale. Nel caso delle wavelet complesse si è considerato il modulo delle ampiezze.

Gli algoritmi basati sul tempo di attraversamento di soglia del segnale o del suo involuppo sono ormai consolidati; per algoritmi basati su trasformazioni wavelet sono stati confrontati algoritmi diversi, alla ricerca di quello più adatto al caso di studio.

Il metodo di estrazione del time delay basato sul massimo della wavelet deriva dall'assunzione che la wavelet presenti un massimo in corrispondenza del tempo di arrivo di una determinata frequenza. Tale ipotesi è stata verificata con successo nel campo dei controlli non distruttivi, ma non ha - al momento - dato risultati positivi per l'applicazione ad una superficie tangibile.

L'algoritmo basato sul tempo di attraversamento della wavelet si fonda sull'assunzione che al tempo in corrispondenza del quale la wavelet centrata su una frequenza predefinita del segnale passa una soglia fissata si ha l'arrivo del segnale. Rispetto al tempo di attraversamento del segnale, questa tecnica presenta il vantaggio di potersi focalizzare su una determinata frequenza, e quindi su una velocità di propagazione unica e determinata.

#### Trasformata wavelet

Il termine wavelet indica un'onda  $\psi(t)$  limitata nel tempo che può essere reale o complessa e che soddisfa a due condizioni: (i) deve avere valore medio zero e (ii) avere un'energia finita [14]. A partire dalla funzione base  $\psi(t)$ , detta anche *mother wavelet*, attraverso una serie di operazioni di trasformazione di scala e traslazione, viene ricavata una base ortonormale di funzioni

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right)$$

chiamate *daughter wavelets*, caratteristiche per il particolare segnale analizzato. Esistono diverse famiglie di mother wavelets tra cui scegliere la più adatta per l'applicazione specifica.

La trasformata wavelet di un segnale  $f(t)$  alla scala  $s$  e posizione  $u$  è calcolata dal prodotto scalare di  $f(t)$  con una wavelet

$$Wf(u,s) = \int f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt = \langle f(t), \psi_{u,s}(t) \rangle$$

dove  $*$  indica il complesso coniugato [13]. Se le funzioni di base vengono dilatate di un fattore di scala  $2^j$ , anche il segnale verrà scalato dello stesso fattore. Quindi, sulla base:

$$\psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t-2^j n}{2^j}\right)$$

per  $n = 1, 2, \dots, N$  e per  $j = 1, 2, \dots, N$  con  $N = 2^j$ ,

il segnale sarà rappresentato come

$$f(t) = \sum_{j=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \langle f(t), \psi_{j,n}(t) \rangle \psi_{j,n}(t)$$

dove la sommatoria dei coefficienti della wavelet

$$d_j(t) = \sum_{n=-\infty}^{+\infty} \langle f(t), \psi_{j,n}(t) \rangle \psi_{j,n}(t)$$

fornisce l'approssimazione del segnale alla scala  $2^j$ . Quindi, ogni coefficiente  $d_j(t)$  cui corrisponde una specifica frequenza contiene inoltre un'informazione di localizzazione temporale all'interno del segnale. Tale tecnica risulta pertanto particolarmente adatta per la rappresentazione di fenomeni impulsivi.

## ESPERIMENTO

### Setup di misura

Per selezionare l'algoritmo TDOA più promettente e verificare l'accuratezza dei tempi di ritardo, è sufficiente concentrarsi inizialmente su un problema di localizzazione monodimensionale in cui il tocco e i sensori si trovano tutti lungo la stessa linea.

L'estensione al caso di localizzazione bidimensionale è solo un problema geometrico, che richiede di collocare un ulteriore sensore e che introduce una formula algebrica che conterrà una variabile angolo oltre alla variabile distanza [15].

In questo lavoro, è stata utilizzata una lastra in plexiglass di dimensioni 210x840 mm e di spessore 3 mm, colpita, in questa prima fase con una punta rigida, in modo da ottenere una serie di misure più facilmente ripetibili e confrontabili. Il segnale rilevato è l'accelerazione lungo l'asse normale al piano della lastra, per mezzo di accelerometri BU-21771 Knowles e acquisito su elaboratori per mezzo di una scheda National Instrument NI-USB 6210.

Per questa misura, gli accelerometri sono posti ad una distanza di 200 mm, distanza scelta in modo da adattarsi alla dimensione del campione di plexiglass, e da rendere agevole il confronto con i risultati sperimentali riportati in bibliografia.

### Separazione dei modi

Come si è visto nelle premesse, la vibrazione prodotta da un impatto sulla superficie ha un contenuto complesso e, in particolare, nel segnale sono contenuti il modo simmetrico e antisimmetrico di ordine più basso ( $S_0$ ,  $A_0$ ). Questo comportamento ha effetto sulla definizione della modalità di misura.

Il metodo proposto in questo lavoro, si fonda su due considerazioni fisiche. La prima è che, separando i diversi modi dell'onda acustica ai sensori, si possono ottenere maggiori informazioni sulla natura dell'evento che l'ha provocata [16]. In secondo luogo, è necessario effettuare una misura della velocità di gruppo su un segnale che è a banda larga e dispersivo [17]. Questo rende utile implementare l'algoritmo nel dominio tempo-frequenza, impiegando la trasformata wavelet come metodo di analisi del segnale.

Il tipo di set up scelto per la misura di velocità è con quattro sensori posizionati in due coppie affacciate sulle due superfici libere della lastra (fig. 2). Con questo metodo l'impulso può essere scomposto facendo la somma o la differenza del segnale raccolto nella stessa posizione, isolando così il modo simmetrico da quello antisimmetrico.



Fig 2. Posizionamento dei sensori contrapposti sulla lastra

Per il modo simmetrico, chiamate  $S_0(+)$  e  $S_0(-)$  le deformazioni delle superfici superiore (+) e inferiore (-) della lastra, si ottiene:

$$\begin{aligned} S_0^+ + S_0^- &= 0 \\ S_0^+ - S_0^- &= 2|S_0| \end{aligned}$$

Per il modo antisimmetrico, analogamente, date le deformazioni raccolte da ogni accelerometro sulla faccia superiore (+) e inferiore (-):

$$\begin{aligned} A_0^+ + A_0^- &= 2|A_0| \\ A_0^+ - A_0^- &= 0 \end{aligned}$$

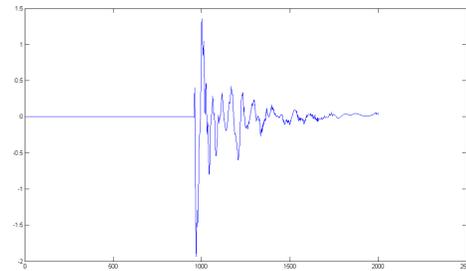


Fig 3. Modo  $A_0$  estratto da una coppia di accelerometri

In figura 3 è riportato un esempio di *time-history* della componente antisimmetrica del segnale, misurata durante uno degli esperimenti.

### Misura della velocità di propagazione

Con il setup sperimentale proposto è stato possibile procedere innanzitutto alla misura della velocità di propagazione dell'onda nel materiale utilizzato. Sono stati effettuati una serie di tocchi in posizioni esterne rispetto alle coppie di sensori, in modo da poter misurare dei tempi di ritardo tra posizioni note e risalire quindi alla velocità di propagazione.

I tempi di ritardo sono stati misurati con tutti e cinque gli algoritmi TDOA descritti e i risultati sono stati confrontati.

Si è potuto osservare come il metodo basato sull'attraversamento di soglia da parte della wavelet complessa abbia dato i risultati più ripetibili e con dispersione minore. Il segnale utilizzato in questo caso è stato il modo antisimmetrico flessionale. La figura 4 riporta i tempi di ritardo misurati con tale algoritmo per diverse distanze tra le coppie di sensori. La velocità del modo  $A_0$  misurata con questa tecnica per il materiale e la lastra utilizzata è stata di 665 m/s.

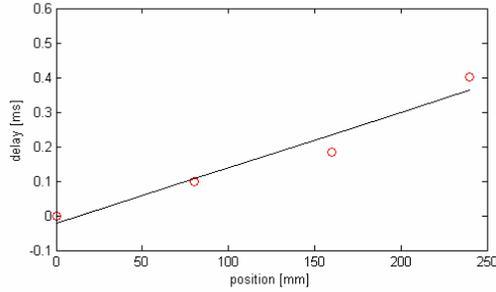


Fig. 4. Velocità di propagazione estrapolata con l'algoritmo Complex Wavelet

La trasformata wavelet del segnale è stata calcolata utilizzando il software Wavelet Toolbox™ di MATLAB. La mother wavelet scelta è la Morlet complessa (figura 5), poiché il pacchetto d'onda di Lamb può essere trattato come un involuppo Gaussiano che modula la frequenza portante eccitata selezionata [18]

$$\psi(x) = \pi^{-1/4} \exp(i\omega_0 x) \exp(-x^2/2)$$

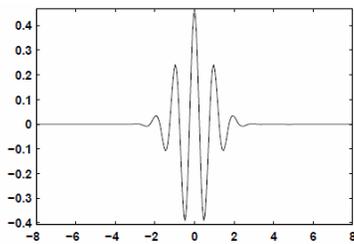


Fig. 5. Funzione base per la wavelet Morlet

Si è fatta la scelta di utilizzare la wavelet ad una frequenza che una preliminare analisi di Fourier standard aveva rivelato essere una frequenza caratteristica del segnale [19].

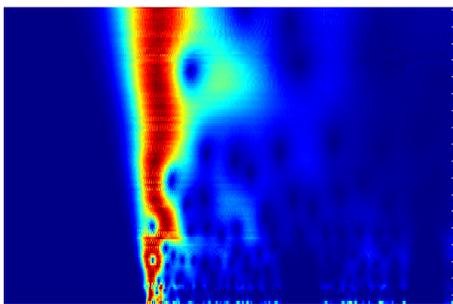


Fig. 6. Esempio di wavelet del segnale da uno degli esperimenti

La figura 6 mostra un esempio di trasformata wavelet di uno dei segnali raccolti durante gli esperimenti. Sull'asse delle ascisse si trova l'asse dei tempi, l'asse delle ordinate rappresenta il parametro di scala, proporzionale all'inverso di una frequenza. Sul terzo asse, rappresentato dalla scala di colore, viene riportata l'ampiezza della wavelet per la scala ed il tempo corrispondenti. Un taglio ad una scala predeterminata corrisponde ad una time-history, di lunghezza pari alla lunghezza del segnale di partenza, dell'ampiezza del segnale ad una particolare frequenza.

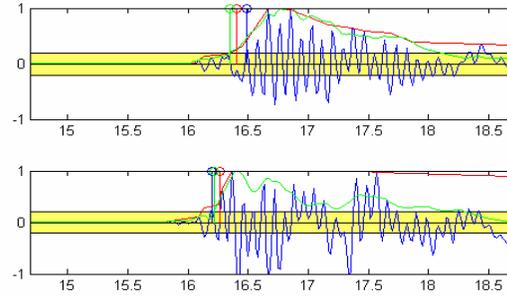


Fig. 7. Confronto dell'attraversamento della soglia tra segnale, involuppo e wavelet complessa

In figura 7 vengono confrontati tra loro il segnale, l'involuppo e la wavelet alla scala selezionata. Nel grafico è anche riportata la soglia utilizzata, e si sono evidenziati i tempi di attraversamento.

### Localizzazione dei tocchi

Una volta nota la velocità di propagazione dell'onda acustica nel materiale è possibile procedere alla validazione degli algoritmi per la localizzazione dei tocchi.

I tocchi sono stati ripetuti in posizioni note sulla linea che congiunge le due coppie di accelerometri, con i punti fissati a distanza di 2 cm uno dall'altro.

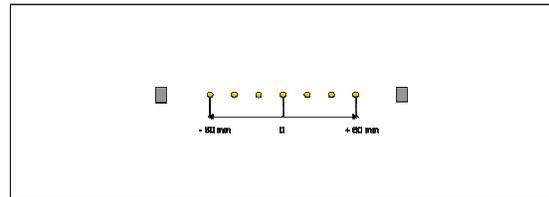
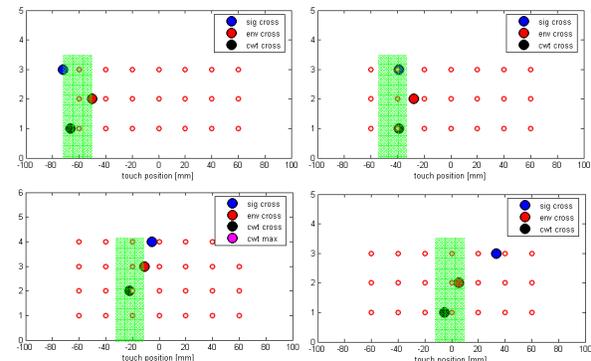


Fig. 8. Posizioni dei tocchi rispetto alla lastra ed agli accelerometri

Per ogni posizione, dalle misure raccolte dagli accelerometri si è ricostruita la posizione dell'impatto, cioè la posizione teorica calcolata con tutti gli algoritmi proposti. Vengono di seguito riportati i risultati relativi agli algoritmi che hanno mostrato le performance migliori, ossia il superamento di soglia da parte del segnale, del suo involuppo e della wavelet complessa determinata in precedenza.

In figura 9 sono rappresentati i risultati forniti dai tre algoritmi, in termini di posizioni calcolate, visualizzati con un intervallo di incertezza di  $\pm 5$  mm rispetto alla posizione vera dell'impatto.



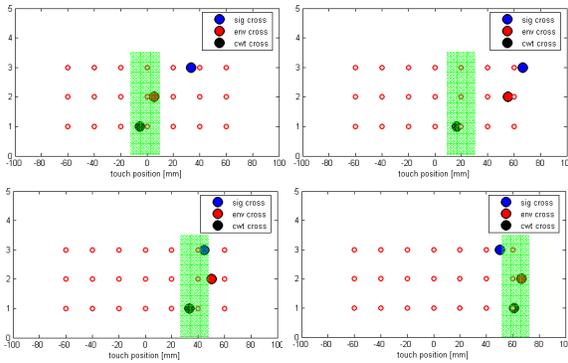


Fig. 9. Confronto tra le posizioni dei tocchi individuate dai diversi algoritmi rispetto alle posizioni reali

Si può rilevare come l'algoritmo basato sull'attraversamento di soglia della wavelet complessa sia stato costantemente in grado di rilevare la posizione del tocco con un errore inferiore a  $\pm 5$  mm.

Gli altri algoritmi, invece, pur se alcune volte si sono dimostrati accurati, in alcuni casi hanno fornito un calcolo di posizione con un errore di oltre 20 mm rendendo impossibile individuare il punto del tocco.

## CONCLUSIONI

Sono stati confrontati diversi algoritmi per la localizzazione di un tocco su una lastra piana. L'algoritmo basato sul tempo di attraversamento di soglia della wavelet si è rivelato più accurato e consistente rispetto ad altri algoritmi basati sull'attraversamento di soglia del segnale.

Si ritiene che ciò sia da attribuirsi al fatto che il segnale in arrivo è stato scomposto nella sua componente simmetrica ed antisimmetrica, e che l'algoritmo della wavelet abbia permesso di calcolare il tempo di arrivo di una particolare frequenza del modo antisimmetrico. Il concentrarsi su una particolare frequenza è stato necessario data la particolare natura dispersiva di tale modo, caratterizzato da una diversa velocità di propagazione per frequenze diverse.

Si è rivelato altresì importante calcolare la velocità di propagazione di tale frequenza nel materiale scelto utilizzando lo stesso metodo già implementato per la localizzazione dei tocchi.

Tale accresciuta accuratezza e attendibilità dei risultati rende possibile la trasformazione di una superficie qualsiasi in una interfaccia tangibile con l'utilizzo di un limitato numero di sensori. Esperimenti sono in corso nella progettazione di interfacce per l'espressione musicale e in particolare per ascolto attivo di musica preregistrata.

## RINGRAZIAMENTI

La ricerca è parzialmente supportata da Programma EU-ICT 2007.1.5 SAME (Sound and Music for Everyone, Everyday, Everywhere, Every way, [www.sameproject.eu](http://www.sameproject.eu)).

## RIFERIMENTI

1. C. Bornand, A. Camurri, G. Castellano, S. Catheline, A. Crevoisier, E. Roesch, K. Scherer, G. Volpe, *Usability evaluation and comparison of prototypes of tangible acoustic interfaces*. Proceedings ENACTIVE 2005, 2nd Enactive International Conference, Genova, November 2005.
2. A. Crevoisier and P. Polotti, *Tangible Acoustic Interfaces and their Application for the Design of New Musical Instruments*, Proceedings of the 2005 International Conference on New Interfaces for Musical Expression, Vancouver, BC, Canada, 2005.
3. D. Rocchesso, F. Fontana, *The Sounding Object*, Edizioni di Mondo Estremo, 2003.
4. J. A. Paradiso, C. K. Leo, N. Checka, K. Hsiao, *Passive Acoustic Sensitive for Tacking Knocks Atop Large Interactive Displays*, Proceedings of the 2002 IEEE International Conference on Sensors, Volume 1, Orlando, Florida, June 11-14, 2002, pp. 521-527.
5. A. Camurri, G. Volpe *Multimodal and crossmodal analysis of expressive gesture in tangible acoustic interfaces*, Proc. ROMAN 2006 – The 15th IEEE Intl. Symposium on Robot and Human Interactive Communication, 6-8 Sept 2006, University of Hertfordshire, pp.781-786, IEEE CS Press.
6. A. Camurri A., C. Canepa, S. Ghisio, G. Volpe, *Automatic Classification of Expressive Hand Gestures on Tangible Acoustic Interfaces According to Laban's Theory of Effort*, Proc. Intl Gesture Workshop 2007, Springer.
7. *Metamorfosi del Senso, Catalogo della mostra Cimenti di Invenzione e Armonia a Casa Paganini*, a cura di N. Ferrari, Quaderni di Casa Paganini, 2007b.
8. <http://www.taichi.cf.ac.uk/>
9. H. Jeong, Y.S. Jang, *Fracture Source Location in Thin Plates using the Wavelet Transform of Dispersive Waves*, IEEE Transactions on ultrasonics, ferroelectrics and frequency control, Vol. 47, N. 3, May 2000.
10. Kishimoto K., Inoue H., Hamada M., Shibuya T., *Time frequency analysis of dispersive waves by means of wavelet transform*, J. Applic. Mech. 1995; 62:841 – 6.
11. Y. Y. Kim, E. H. Kim, *Effectiveness of the Continuous Wavelet Transform in the Analysis of some Dispersive Elastic Waves*, J. Acoust. Soc. Am. 110 (1), July 2001.
12. D. Royer, E. Dieulesaint, *Elastic Waves in Solids*, vol. I-II, Springer-Verlag, 1999.
13. S. M. Ziola, M. R. Gorman, *Source Location in Thin Plates using Cross-correlation*, J. Acoust. Soc. Am. 90 (5), November 1991.
14. S. Legendre, D. Massicotte, J. Goyette, T. K. Bose, *Wavelet-Transform-Based Method of Analysis for Lamb Wave Ultrasonic NDE Signals*, IEEE Transactions on Instrumentation and Measurement, Vol. 49, N. 3, June 2000.
15. A. Tobias, *Acoustic-emission source location in two dimensions by an array of three sensors*, Non-Destructive Testing, vol. 9, Feb. 1976, p. 9-12, 1976.
16. M. Surgeon, M. Wevers, *Modal Analysis of Acoustic Emission Signals from CFRP Laminates*, NDT&E International 32, 1999.
17. R. Chona, C. S. Suh, G. A. Rabroker, *Characterizing defects in multilayer materials using guided ultrasonic waves*, Optics and Lasers in engineering 40 (2003) 371-378.
18. A. Z. Abid, M. A. Gdeisat, D. R. Burton, M. J. Lalor, *Ridge Extraction Algorithms for One Dimensional continuous Wavelet Transform: a Comparison*, Sensors and their Applications XIV, Journal of Physics: Conference Series 76 (2007).
19. E. Ribaldone, *Digital Data Analysis*, CRF Report, Torino, 2007.



# Un Modello per la Simulazione in Tempo reale di un Riverbero a Piastra

Stefano Zambon  
Dipartimento di Informatica  
Università degli Studi di Verona

## SOMMARIO

Questo lavoro è basato su una recente tecnica per la discretizzazione dell'equazione delle onde in due dimensioni, basata su una trasformazione geometrica che permette di ridurre lo studio a sistemi monodimensionali. Il metodo, precedentemente sviluppato per il caso semplice di un dominio quadrato, è stato esteso a forme rettangolari di dimensioni generiche e diverse condizioni al contorno sono state considerate. In seguito, viene fornita un'applicazione alla simulazione del riverbero a piastra, riconducendo l'equazione delle lamine metalliche a quella della corda rigida e con perdite, largamente studiata in letteratura per la sintesi di suoni di pianoforte. Viene mostrato come la nuova tecnica sia largamente più efficiente dei metodi tradizionali, e pertanto adatta ad un'implementazione in tempo reale.

## Parole Chiave

Sintesi per Modelli Fisici; Riverbero a Piastra; Equazione delle Onde.

## INTRODUZIONE

I benefici della sintesi per modelli fisici sono ben noti per la simulazione di strumenti acustici. Tramite queste tecniche, è possibile ottenere un'alta qualità sonora e, inoltre, viene offerta la possibilità di un controllo intuitivo sia nell'interazione con il musicista che nella parametrizzazione dello strumento virtuale. Tra le tecniche usate con più successo si citano differenze finite [24], sintesi modale [1] e digital waveguides [22, 23]. Recentemente, questi metodi sono stati applicati anche alla simulazioni di effetti audio basati su unità meccaniche. Un esempio è dato dal riverbero a piastra, un effetto che deve la sua popolarità grazie all'impiego massiccio nella musica *pop* negli anni sessanta; l'unità è composta da una sottile lamina di metallo, di forma rettangolare, a cui vengono applicati due o più trasduttori elettrici [10]. La piastra si comporta, a tutti gli effetti, come un risonatore distribuito spazialmente, per cui è possibile sfruttare le riflessioni che si creano al suo interno per la riverberazione artificiale del suono.

Esistono da diverso tempo modelli matematici accurati per la vibrazione di piastre metalliche [11]; di recente, si è provato a discretizzare le equazioni del moto ai fini di simulare il comportamento di un riverbero a piastra [8]. Con questo approccio si ottiene una qualità audio migliore rispetto a tradizionali riverberi digitali basati su filtri allpass e linee di ritardo [12]; inoltre, è possibile specificare la posizione di sorgenti e punti di ascolto, ottenendo così un valido spazial-

izzatore, utilizzabile anche in contesti multimediali. D'altro canto, la complessità computazionale di tali modelli è estremamente elevata, per cui ad oggi non sono presenti implementazioni accurate in tempo reale; l'unica realizzazione presentata finora, pur scendendo a diversi compromessi sulla qualità, necessita di hardware particolare per il suo funzionamento [16].

Questo lavoro propone una tecnica particolarmente efficiente per la discretizzazione dell'equazione della piastra, adatta ad un'implementazione in tempo reale. L'idea di base prevede di ricondurre il moto a strutture ad una dimensione spaziale, per le quali le digital waveguides [22] sono un metodo di simulazione estremamente efficiente. La tecnica è particolarmente interessante dal punto di vista teorico, in quanto è uno dei pochi esempi in cui le digital waveguides sono usate per rappresentare fronti d'onda multidimensionali. In letteratura sono stati trattati solamente pochi esempi di questa tecnica a situazioni estremamente idealizzate, quali la membrana circolare [26] e quella quadrata con condizioni libere al contorno [7]. Tuttavia, alcune idee che sono alla base di questi lavori sono state trattate in algoritmi di sintesi *physically inspired*, ad esempio da Rocchesso per la simulazione di stanze [18] e da Essl per risonatori generici [13]. In entrambi i casi, viene data un'interpretazione geometrica all'uso di linee di ritardo usate per risonatori multidimensionali, sebbene le tecniche non siano strettamente correlate all'equazione matematica che descrive il moto in tali sistemi.

## EQUAZIONE DELLE ONDE MONODIMENSIONALE

In questa sezione vengono illustrati brevemente gli strumenti matematici che sono alla base dei più diffusi metodi di sintesi per risonatori monodimensionali, quali ad esempio corde e tubi acustici. In particolare, viene illustrato il legame tra le due soluzioni all'equazione delle onde, che sarà di fondamentale importanza per capire meglio il funzionamento della tecnica *digital waveguide extraction* presentata nella prossima sezione.

L'equazione delle onde è un'equazione alle derivate parziali (PDE) che descrive, seppur con diverse approssimazioni, il moto di diversi sistemi di interesse in campo musicale. La derivazione è presente in molti testi di acustica (ad es. [15]), e conduce alla seguente equazione iperbolica:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2} \quad (1)$$

dove, nell'esempio di vibrazioni trasversali lungo una corda,  $u$  è lo spostamento trasversale e  $c$  è la velocità di propagazione

nelle onde. È da notare come questa equazione sia valida nel caso ideale, dove non sono presenti fenomeni di perdita e dispersione, e solamente per lievi valori della variabile  $u$ . Tuttavia, lo studio dell'equazione è di grande interesse dal momento che i modelli più realistici sono ottenuti aggiungendo termini alla (1); ad esempio, è possibile includere termini per dispersione e dissipazione [6] e, con qualche accorgimento, è possibile l'estensione alle vibrazioni di elevata ampiezza e all'accoppiamento con le vibrazioni longitudinali [3].

### Soluzione analitica per espansione modale

Per completare la descrizione del problema, bisogna imporre le condizioni iniziali e al contorno. Un modello di interesse per gli strumenti acustici prevede in genere condizioni al bordo ad estremi fissati, per cui  $u(0, t) = u(L, t) = 0$  in una corda di lunghezza  $L$ . Le condizioni iniziali vengono specificate indicando lo spostamento trasversale lungo l'intera corda all'istante  $t = 0$ :

$$u_0(x) = u(x, 0) = \sum_{k=1}^{\infty} A_k \cos\left(\frac{k\pi x}{L}\right) \quad (2)$$

Separando le variabili e tenendo in considerazione le condizioni al contorno, è immediato ricavare la seguente soluzione in termini di espansione di Fourier [15], proposta inizialmente da Bernoulli nel 1738:

$$u(x, t) = \sum_{k=1}^{\infty} A_k \sin\left(\frac{k\pi x}{L}\right) \cos(2\pi f_k t), \quad (3)$$

dove  $f_k = (kc)/(2L)$  sono le frequenze naturali di vibrazione della corda. Da questa soluzione si può ricavare facilmente un'implementazione per sintesi modale [1,25] nella quale ad ogni termine della sommatoria (3) viene associato un risonatore digitale, tipicamente realizzato tramite un filtro passa-banda del secondo ordine. Ovviamente, bisogna limitarsi ad un numero finito  $N = F_s/(2f_1)$  di risonatori, che in genere dipende dalla frequenza di campionamento.

### Soluzione analitica per onde sovrapposte

Nel caso la frequenza fondamentale  $f_1$  sia bassa, è necessario usare un elevato numero di risonatori del secondo ordine nell'approccio a sintesi modale. Ad esempio, un Do della prima ottava del pianoforte (C1) richiede più di 1500 risonatori ad una frequenza di campionamento  $F_s = 44.1 \text{ KHz}$ .

Le digital waveguides [22,23], al contrario, permettono un'implementazione di gran lunga più efficiente, avente costo computazionale indipendente dalla frequenza di campionamento. L'idea è in qualche modo simile a quella della sintesi modale, in quanto anche in questo caso si sceglie di discretizzare una soluzione analitica all'equazione, anziché discretizzare direttamente l'equazione stessa. La soluzione scelta in questo caso è quella per onde viaggianti (*traveling waves*), proposta da D'Alembert nel 1770:

$$u(x, t) = \sum_{k=1}^{\infty} \frac{A_k}{2} \left( \sin\left(\frac{k\pi}{L}(x - ct)\right) + \sin\left(\frac{k\pi}{L}(x + ct)\right) \right) = \frac{1}{2} (u_0(x - ct) + u_0(x + ct)) \quad (4)$$

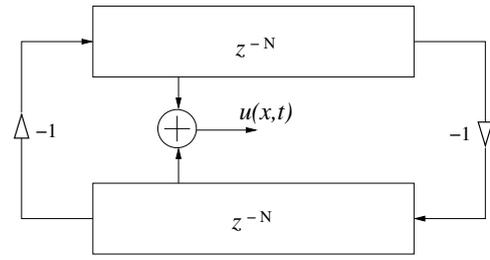


Figura 1. Schema computazionale di una digital waveguide usando due linee di ritardo di uguale lunghezza.

Se l'Eq. (4) viene discretizzata usando come passo di campionamento temporale  $dX = c/F_s$ , possiamo realizzare il sistema tramite una linea di ritardo bidirezionale, come mostrato in Fig. 1. La struttura può essere realizzata molto economicamente implementando le linee di ritardo con dei buffer circolari, per cui ad ogni istante le uniche operazioni necessarie sono l'aggiornamento dei puntatori e la somma dei valori in corrispondenza dell'output, indipendentemente dalla lunghezza delle linee di ritardo e dalla frequenza di campionamento.

Dall' Eq.(4) è possibile inoltre notare come le due soluzioni siano collegate tra di loro: se la condizione iniziale ammette sviluppo in serie di Fourier, come solitamente avviene nei casi di interesse pratico, è possibile passare facilmente dalla soluzione modale a quella per onde viaggianti tramite l'uso delle note formule trigonometriche di prostaferesi.

### EQUAZIONE DELLE ONDE IN DUE DIMENSIONI

Come già notato per i sistemi monodimensionali, anche nel caso bidimensionale i modelli matematici realistici sono costruiti aggiungendo termini all'equazione ideale. Inoltre, nella situazione di nostro interesse la parte spaziale della soluzione rimane sostanzialmente invariata nei due casi, per cui ha senso studiare approfonditamente l'equazione ideale e derivare da questa la trasformazione spaziale che ci sarà utile in seguito.

L'equazione delle onde in due dimensioni [15] è usata per descrivere le vibrazioni trasversali di una membrana:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \nabla^2 u. \quad (5)$$

Nella (5)  $\nabla^2$  è l'operatore laplaciano e generalizza in più dimensioni l'operatore di derivata seconda presente nella (1). Le condizioni iniziali sono date similmente al caso monodimensionale, mentre per le condizioni al contorno consideriamo ancora una volta di fissare lo spostamento agli estremi, ponendo  $u = 0$  ai bordi del rettangolo  $[0, L_x] \times [0, L_y]$ .

### Soluzione analitica per espansione modale

Sfruttando la forma particolare del dominio geometrico, è possibile anche in questo caso ottenere una soluzione modale separando le variabili e imponendo le condizioni al contorno separatamente per ciascuna dimensione spaziale [15]. Come risultato generale, la soluzione  $u(x, y, t)$  è ancora esprimibile come una somma di modi normali, ciascuno dei quali è

formato dal prodotto di due sinusodi per la parte spaziale e una cosinusode per la parte temporale:

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} A_{m,n} \sin\left(\frac{m\pi x}{L_x}\right) \sin\left(\frac{n\pi y}{L_y}\right) \cos(2\pi f_{m,n} t), \quad (6)$$

dove i coefficienti  $A_{m,n}$  sono definiti dalle condizioni iniziali e le frequenze di vibrazione  $f_{m,n}$  sono calcolate come:

$$f_{m,n} = \frac{c}{2} \sqrt{\left(\frac{m}{L_x}\right)^2 + \left(\frac{n}{L_y}\right)^2}. \quad (7)$$

A differenza del caso monodimensionale, le frequenze dei modi non sono riconducibili ad un'unica serie armonica, per cui non è possibile applicare direttamente la sintesi per digital waveguide. Di fatto, la grande maggioranza dei metodi proposti in letteratura quali differenze finite [24] e digital waveguide mesh [9] si basa sulla discretizzazione diretta dell'equazione; alternativamente, è possibile discretizzare la soluzione (6) tramite sintesi modale [1,25], sebbene il numero di risonatori necessari sia spesso proibitivo per un'implementazione in tempo reale.

### Digital waveguide extraction

In questa sezione, viene proposta una soluzione equivalente alla (6), avente il vantaggio di essere espressa come sovrapposizione di onde viaggianti e, pertanto, implementabile efficientemente in termini di digital waveguides. Soluzioni simili sono già state pubblicate nel caso di dominio circolare [26] e quadrato [7]. La trattazione che segue può essere vista come un'estensione al caso rettangolare e con condizioni al contorno più generali di quanto presente in [7]; pertanto, laddove possibile, si tenterà di far uso di una notazione simile nella derivazione.

L'osservazione di partenza è che l'insieme dei modi di vibrazione può essere ripartito in sottoinsiemi formati da serie armoniche, ovvero aventi frequenze multiple di un'unica frequenza fondamentale. Quindi, la soluzione (6) può essere riscritta come:

$$u(x, y, t) = \sum_{\mathbf{q}} u_{\mathbf{q}}(x, y, t), \quad (8)$$

dove ciascuna componente  $u_{\mathbf{q}}(x, y, t)$  è data dalla seguente serie armonica:

$$u_{\mathbf{q}}(x, y, t) = \sum_{k=1}^{\infty} A_{k\mathbf{q}} \sin\left(\frac{kq_x \pi x}{L_x}\right) \sin\left(\frac{kq_y \pi y}{L_y}\right) \cos(2\pi f_{k\mathbf{q}} t). \quad (9)$$

È facile notare le similarità dell'equazione (9) con la soluzione modale dell'equazione delle onde monodimensionale (3); in questo caso, tuttavia, non è immediata l'applicazione delle formule di prostaferesi per ricondurci ad una forma di onde viaggianti.

Come già fatto notare in un precedente lavoro [18], tutti i modi nella serie di Eq. (9) corrispondono ad un'onda piana viaggiante lungo la direzione data dal vettore  $[q_x/L_x, q_y/L_y]$ . È facile visualizzare questo fatto analizzando i fronti d'onda chiusi corrispondenti a tali modi, illustrati in Fig. 2. L'idea di

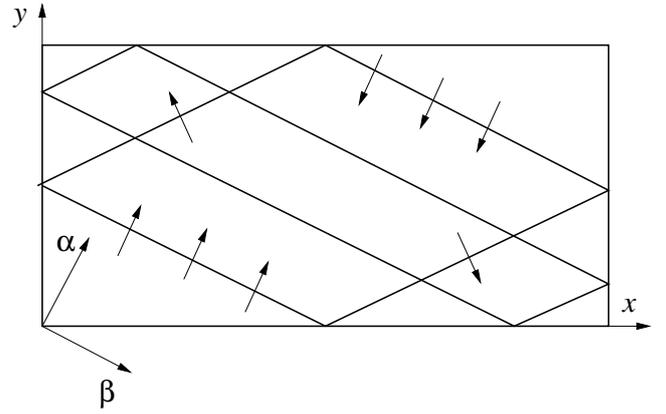


Figura 2. Propagazione dei fronti d'onda corrispondenti alla digital waveguide estratta avente fondamentale  $q_x = 1, q_y = 1$ .

base delle digital waveguide extraction [7] consiste nell'applicare un cambio di coordinate all'espressione (9), portandoci in un sistema di riferimento parallelo alle direzioni di propagazione del fronte d'onda:

$$\alpha_q = \frac{q_x L_y x + q_y L_x y}{N_q}, \beta_q = \frac{q_x L_y x - q_y L_x y}{N_q} \quad (10)$$

dove  $N_q = \sqrt{q_x^2 L_y^2 + q_y^2 L_x^2}$ . Applicando ripetutamente le formule di prostaferesi all'espressione ottenuta, con alcune semplificazioni si arriva alla formulazione equivalente in termini di onde viaggianti:

$$u_{\mathbf{q}}(x, y, t) = \frac{1}{4} [(u_{\mathbf{q},0}(\alpha_q - ct) + u_{\mathbf{q},0}(\alpha_q + ct)) - (u_{\mathbf{q},0}(\beta_q - ct) + u_{\mathbf{q},0}(\beta_q + ct))] \quad (11)$$

dove si è introdotta una lunghezza effettiva  $L_q = L_x L_y / N_q$  e l'eccitazione iniziale del sistema monodimensionale  $u_{\mathbf{q},0}$  è calcolata come:

$$u_{\mathbf{q},0}(\gamma) = \sum_{k=1}^{\infty} A_{k\mathbf{q}} \cos\left(\frac{k\pi\gamma}{L_q}\right). \quad (12)$$

L'equazione (11) corrisponde, a livello formale, ad un sistema monodimensionale, in cui sono espresse le relazioni tra le onde viaggianti. È pertanto possibile un'implementazione efficiente tramite una digital waveguide in cui l'ampiezza viene misurata in due punti  $\alpha_q, \beta_q$ ; di seguito, sottraendo tra di loro questi valori si ottiene la soluzione nel punto  $x, y$  della membrana. Rispetto allo schema di Fig. 1, non c'è alcuna riflessione al bordo in questo caso; questo dipende dal fatto di avere una funzione cosinusoidale nell'Eq. (12), corrispondente a condizioni al contorno di tipo libero.

Può capitare che i valori di  $\alpha_q, \beta_q$  calcolati con la trasformazione (10) non cadano all'interno dell'intervallo  $[0, L_q]$  che descrive il dominio della digital waveguide implementata. Il rimedio, simile a quello proposto in [7], consiste nel riflettere i valori nell'intervallo desiderato sfruttando il fatto che la soluzione è spazialmente periodica con periodo  $2L_q$ . Inoltre, le posizioni di ascolto  $\alpha_q, \beta_q$  non sono generalmente un multiplo del passo di campionamento spaziale  $dX$ ;

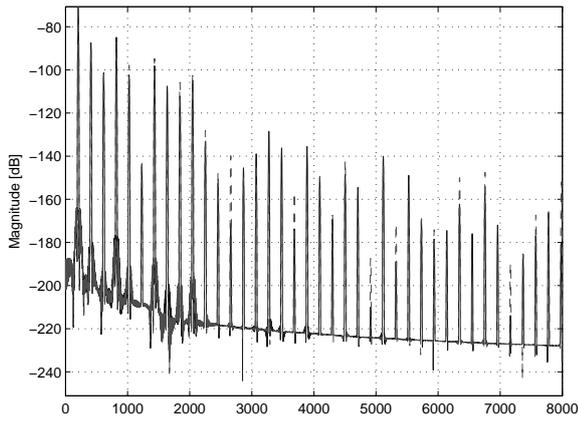


Figura 3. Spettro di una simulazione dei modi di una membrana rettangolare corrispondenti alla waveguide di indice  $q_x = 2, q_y = 1$ . La soluzione modale (linea continua) è confrontata con quella per digital waveguide extraction (linea tratteggiata).

sarebbe pertanto necessario usare anche in questo caso dei filtri di interpolazione per realizzare i ritardi frazionari. Tuttavia, l'errore che si commette trascurando di implementare questi filtri è generalmente trascurabile, in quanto corrisponde a modificare leggermente l'ampiezza dei modi; infatti, per la qualità della riverberazione è sufficiente approssimare l'andamento generico delle ampiezze [2]. In Fig. 3 vengono confrontati gli spettri di una simulazione per una singola waveguide estratta, usando l'implementazione per digital waveguide e per sintesi modale. È possibile osservare come l'errore nelle ampiezze sia limitato a pochi dB e sia praticamente assente per i primi modi in bassa frequenza.

Nelle Fig. 4,5 viene mostrata la differenza nell'approssimazione per la simulazione di una membrana avente  $L_x = 2, L_y = 1, c = 200$ . L'output della simulazione completa con sintesi modale con 41607 risonatori viene confrontato con quello ottenuto usando 750 waveguide estratte. È possibile notare alcune distorsioni nella forma d'onda, dovute in parte all'approssimazione in frequenza e in parte alla risposta transitoria dei filtri allpass usati per i ritardi frazionari. L'influenza di questi e altri artefatti nella qualità percettiva del riverbero restano un problema aperto, che vale probabilmente la pena investigare tramite opportuni esperimenti percettivi. Al momento, è possibile analizzare quantitativamente le risposte ottenute, confrontandole con il modello di riferimento dato dall'espansione modale completa; un esempio è mostrato in Fig. XX.

### UN MODELLO PER IL RIVERBERO A PIASTA

Ci concentriamo ora su un modello matematico per descrivere le vibrazioni in una sottile lamina metallica, con caratteristiche simili a quelle presenti in un'unità di riverbero a piastra. Tralasciamo in questo lavoro tutta la modellazione delle componenti elettriche del sistema originale, che comunque contribuiscono solo a colorare differentemente il suono ma non influenzano le caratteristiche di evoluzione temporale del sistema di riverberazione. Un'equazione che

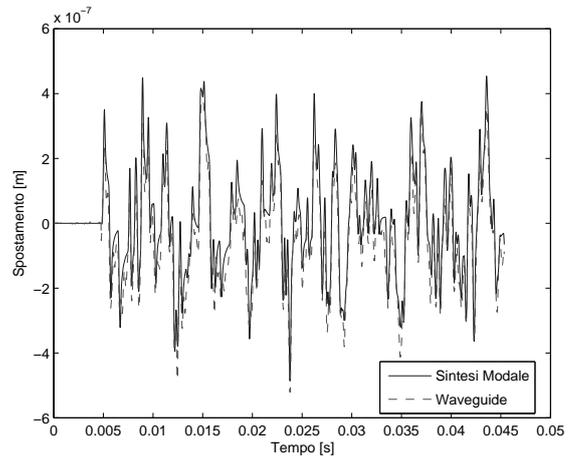


Figura 4. Output della simulazione di una membrana rettangolare, usando sintesi modale con 41607 risonatori (linea continua) e digital waveguide extraction con 750 waveguide (linea tratteggiata).

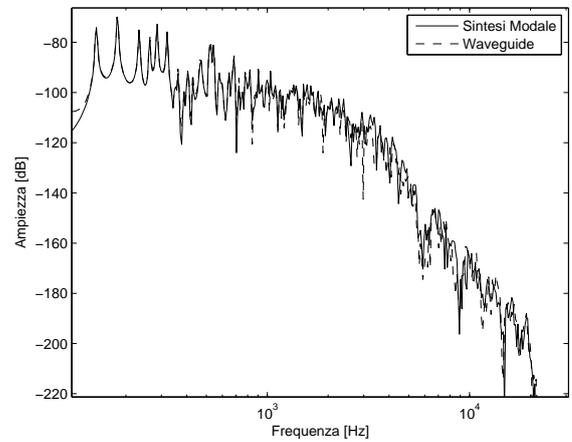


Figura 5. Spettro della risposta ottenuta eccitando la membrana con una forza avente caratteristiche passabasso. Le ampiezze sono mediate in intervalli di frequenze spaziate logicamente.

descrive piuttosto bene il comportamento di una piastra sottile e rigida, sottoposta a lievi eccitazioni, può essere ricavata a partire dalle note equazioni per le piastre senza perdite presenti in testi di acustica [11, 15]. I modelli proposti in letteratura [8, 25] sono entrambi riconducibili alla seguente equazione alle derivate parziali:

$$\frac{\partial^2 u}{\partial t^2} = -\kappa^2 \nabla^4 u + c^2 \nabla^2 u - 2\sigma \frac{\partial u}{\partial t} + b \frac{\partial}{\partial t} \nabla^2 u + f(x_{in}, y_{in}, t) \quad (13)$$

dove i termini di smorzamento, derivanti da perdite termoe-lastiche e di radiazione [11], sono sintetizzati in un coefficiente per le perdite indipendenti dalla frequenza  $\sigma$  e un coefficiente per le perdite dipendenti dalla frequenza  $b$ . È presente un termine forzante  $f(x_{in}, y_{in}, t)$ , corrispondente nel nostro caso al segnale "asciutto", e il termine  $\kappa^2$  rappresenta invece la durezza del materiale ed è dato da

$$\kappa^2 = \frac{Eh^2}{12\rho(1-\nu^2)} \quad (14)$$

dove  $E$  è il modulo di Young,  $h$  lo spessore della piastra e  $\rho, \nu$  sono rispettivamente la densità e il coefficiente di Poisson del materiale.

L'equazione (13) può essere pensata come una generalizzazione al caso multidimensionale alla ben nota equazione per la corda con rigidità e perdite, che costituisce ad esempio un buon modello per le corde del pianoforte [6]. Vedremo come sfruttare questa similarità per poter utilizzare, tramite la digital waveguide extraction, tecniche consolidate per la simulazione delle caratteristiche di dispersione e perdita, sviluppate nell'ambito della sintesi di suoni di pianoforte.

### Soluzione per espansione modale

Per l'Eq. (13) consideriamo condizioni iniziali a riposo, dal momento che l'eccitazione è condensata nel termine forzante esterno  $f$ . Come condizioni al contorno, imponiamo ancora una volta quelle ad estremi fissati, come nel caso ideale; tuttavia, avendo in questo caso una derivata spaziale del quarto ordine, è necessario imporre un'altra condizione. Scegliendo di annullare la derivata seconda di  $u$  al bordo, condizione equivalente ad estremi morsettati, e procedendo ancora una volta per separazione di variabili, si può arrivare a scrivere la soluzione nella seguente forma:

$$u(x, y, t) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \sin\left(\frac{m\pi x}{L_x}\right) \sin\left(\frac{n\pi y}{L_y}\right) u_{m,n}(t) \quad (15)$$

In questo caso, non è possibile risolvere immediatamente la parte temporale, in quanto questa dipende ovviamente dall'eccitazione del sistema. Tuttavia, è possibile ottenere un insieme di equazioni differenziali ordinarie (ODE), disaccoppiate tra di loro, per descrivere l'evoluzione temporale  $u_{m,n}(t)$  del modo  $(m, n)$ . Per farlo, sostituiamo la (15) in (13) ed integriamo in entrambe le direzioni spaziali lungo il dominio; grazie all'ortogonalità dei modi, rimane un solo termine nella serie (15), per cui si arriva alla seguente equazione:

$$\frac{d^2 u_{m,n}}{dt^2} + a_{1,m,n} \frac{du_{m,n}}{dt} + a_{0,m,n} u_{m,n} = b_{m,n} F_{m,n}(t) \quad (16)$$

I coefficienti sono collegati ai parametri dell'Eq. (13) tramite:

$$a_{1,m,n} = 2\sigma + b(C_m^2 + C_n^2) \quad (17)$$

$$a_{0,m,n} = \kappa^2(C_m^4 + C_n^4 + 2C_m^2 C_n^2) + c^2(C_m^2 + C_n^2) \quad (18)$$

$$b_{m,n} = \frac{4}{L_x L_y} \quad (19)$$

$$F_{m,n} = \int_0^{L_x} \int_0^{L_y} f(x, y, t) \sin(C_m x) \sin(C_n y) dx dy \quad (20)$$

dove  $C_m = m\pi/L_x$  e  $C_n = n\pi/L_y$ . La soluzione dell'Eq. (16), nel caso di eccitazione impulsiva, corrisponde ad una sinusoide smorzata esponenzialmente:

$$u_{\delta,m,n} = A_{m,n} e^{-\frac{t}{\tau_{m,n}}} \sin(\omega_{m,n} t) \quad (21)$$

$$\omega_{m,n} = \sqrt{a_{0,m,n} - a_{1,m,n}^2} \quad (22)$$

$$\tau_{m,n} = \frac{2}{a_{1,m,n}} \quad (23)$$

$$A_{m,n} = \frac{b_{m,n}}{\omega_{m,n}} \quad (24)$$

$$(25)$$

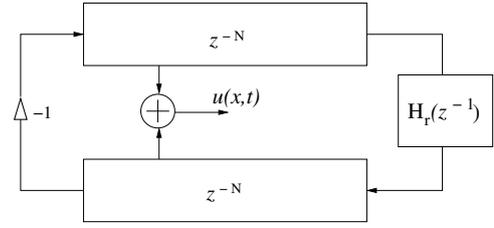


Figura 6. Schema a blocchi per la simulazione di corde dispersive e perdite includendo un filtro  $H_r$  nell'anello della digital waveguide.

Come ben noto dalla teoria delle equazioni differenziali, è possibile ottenere la soluzione per un generico termine forzante  $F_{m,n}$  semplicemente per convoluzione con la risposta all'impulso  $u_{\delta,m,n}$ .

La soluzione analitica qui riportata è piuttosto simile alla soluzione nel caso ideale (6); le differenze sono che, in questo caso, non è possibile raggruppare le frequenze dei modi in serie armoniche e sono inoltre presenti termini di smorzamento. Tuttavia, è sempre possibile una discretizzazione diretta della soluzione tramite sintesi modale; particolarmente interessante è la discretizzazione tramite il metodo dell'invarianza all'impulso, usata in [3] nel caso molto simile dei modi naturali della corda del pianoforte.

### Simulazione per waveguide dispersive e con perdite

Grazie alle forma delle condizioni al contorno, la parte spaziale della soluzione rimane invariata quando si aggiungono i termini di perdita e dispersione all'equazione ideale della membrana rettangolare. Per questo motivo, è possibile applicare allo stesso modo la tecnica di digital waveguide extraction alla soluzione (15). Tuttavia, in questo caso non è possibile la simulazione diretta con una digital waveguide, in quanto le frequenze dei modi sono distribuite inarmonicamente ed inoltre ciascun modo decade ad una velocità diversa.

Per questo, è necessario modificare la struttura della digital waveguide di Fig. 1 aggiungendo nell'anello di ritardo un filtro  $H_r(z)$  per ottenere una risposta inarmonica e con tempi di decadimento dipendente dalla frequenza, v. Fig. 6. L'approccio è stato abbondantemente usato in letteratura per la simulazione delle corde del pianoforte tramite digital waveguides [3, 4, 6]. La tecnica più usata prevede di decomporre il filtro in tre parti,  $H_r(z) = H_{lp}(z)H_d(z)H_{fd}(z)$  dove  $H_{lp}$  è un filtro passa basso per le perdite dipendenti dalla frequenza,  $H_d$  è un filtro allpass per approssimare le caratteristiche di dispersione e  $H_{fd}$  è un'implementazione di un ritardo frazionario necessario per l'accordatura fine del sistema. In aggiunta, è necessario inserire l'eccitazione esterna nelle linee di ritardo; il problema è facilmente risolvibile in questo caso dal momento che stiamo assumendo che la forza di eccitazione  $f$  non dipenda dal moto della piastra. È tuttavia necessario inserire l'eccitazione lungo due posizioni diverse lungo la waveguide, calcolate ancora una volta grazie alla trasformazione spaziale di Eq. (10) del punto di eccitazione  $(x_{in}, y_{in})$ .

In questo lavoro, si è scelto di usare un filtro ad un polo per

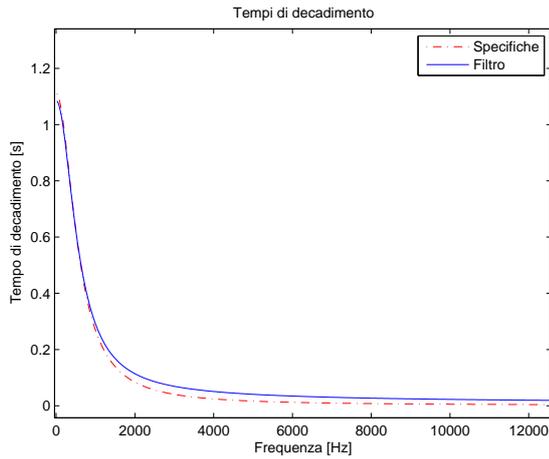


Figura 7. Specifiche e risultati per il filtro di perdita proposto, nel caso dei tempi di decadimento per la waveguide di indice  $q_x = 3, q_y = 1$ .

implementare le perdite:

$$H_{lp}(z) = g \frac{1 + a_1}{1 + a_1 z^{-1}}. \quad (26)$$

Si è usata la tecnica di progettazione basata sulla minimizzazione dell'errore nei tempi di decadimento proposta da Bank in [5]. Similarmente a quanto avviene per le corde di pianoforte, il metodo è particolarmente efficace in quanto i tempi di decadimento ottenuti sono pressoché identici a quelli predetti dalla (25). In Fig. 7 vengono mostrati i risultati ottenuti partendo dalle specifiche per una particolare digital waveguide estratta, ma i risultati sono ugualmente validi per tutti i valori dell'indice ( $q_x, q_y$ ) della waveguide. I parametri del riverbero a piastra sono identici a quelli usati in [8], e corrispondono ad una piastra di acciaio di dimensione  $L_x = 2, L_y = 1$ .

Anche per la simulazione delle caratteristiche dispersive si è fatto riferimento ad una tecnica consolidata per le corde del pianoforte, proposta da Rocchesso e Scalcon in ([20]), che è a sua volta basata su un metodo iterativo per il design di filtri allpass [14]. Le specifiche sul ritardo di fase per il filtro di dispersione possono essere espresse sfruttando le proprietà di un filtro inserito in un anello di ritardo:

$$\angle H_d(e^{j\theta_k}) = N\theta_k - 2k\pi - \angle H_{lp}(e^{j\theta_k}) \quad (27)$$

dove le frequenze  $\theta_k = 2\pi f_k/F_s$  sono ottenute a partire dalla soluzione modale (25). È da notare come il ritardo di fase del filtro di perdita  $H_{lp}$  venga sottratto dalle specifiche. Partendo da queste, un numero  $K$  di frequenze è scelto per la progettazione del filtro; valori tipici di  $K$  cadono nell'intervallo 20-100. Se l'ordine desiderato del filtro è  $N_d$ , il metodo calcola per ogni parziale  $k$  le quantità  $\alpha_k = -0.5 (\angle H_d(e^{j\theta_k}) + N_d\theta_k)$ , ed i coefficienti del filtro  $a_n$  sono ottenuti risolvendo il sistema:

$$\sum_{n=1}^{N_d} a_n \sin(\alpha_k + n\theta_k) = -\sin(\alpha_k), \quad k = 1 \dots K. \quad (28)$$

Dal momento che il sistema è sovradeterminato, viene risolto tramite minimi quadrati. Il metodo procede poi iterativamente compensando l'errore per un fattore pari all'inverso

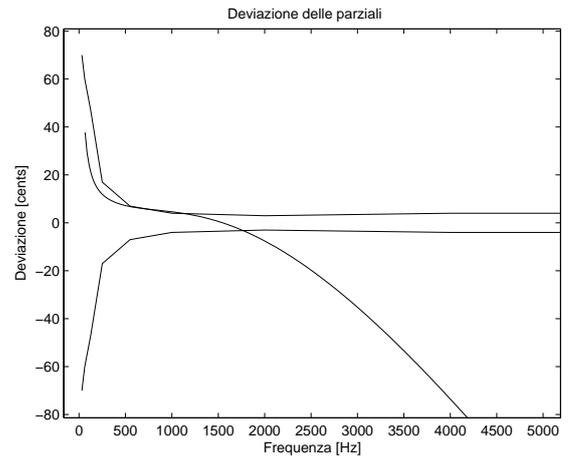


Figura 8. Deviazioni delle frequenze parziali rispetto ai valori nominali, usando come approssimazione un filtro di dispersione un allpass di ordine  $N_d = 2$  progettato con il metodo citato nel testo. Le bande orizzontali indicano i margini entro i quali la differenza non è percepibile.

del modulo del denominatore del filtro  $H_d(z)$ ; eventualmente, può essere applicata una pesatura aggiuntiva basata su criteri percettivi, in modo da dare maggior enfasi alle bande di frequenza in cui l'orecchio è più sensibile all'inarmonicità. Rocchesso e Scalcon hanno anche svolto una serie di esperimenti psicoacustici per studiare la banda di frequenza entro la quale è necessario approssimare con precisione l'inarmonicità [21]; con un ordine del filtro moderato (2-18), è possibile simulare accuratamente le caratteristiche di dispersione delle corde del pianoforte.

Nel caso del riverbero a piastra, purtroppo, è difficile ottenere risultati ugualmente accurati. Questo dipende dalla maggior inarmonicità presente nelle digital waveguide estratte, derivante dal fatto che il materiale è sensibilmente più rigido rispetto al caso delle corde del pianoforte. Inoltre, è opportuno usare filtri di dispersione con ordine ridotto, in quanto essi devono essere replicati per ogni waveguide, e quindi anche diverse centinaia di volte. Tuttavia, non è probabilmente necessario avere lo stesso livello di precisione nel nostro caso, in quanto la risposta è composta da diverse serie armoniche sovrapposte; anche nei tradizionali riverberi implementati tramite linee di ritardo interconnesse (Feedback Delay Networks), sono solitamente impiegati allpass di ordine limitato per ovviare agli artefatti udibili nel caso armonico [2]. Pertanto, in questo lavoro si è scelto di usare  $N_d = 2$  come ordine del filtro allpass; l'approssimazione delle frequenze ottenuta rimane buona approssimativamente per le prime 15-20 parziali, corrispondenti ad una frequenza di circa 2-3000 Hz nel caso delle waveguide più gravi. I risultati ottenuti sono riassunti in Fig. 8.

## CONCLUSIONI

Una recente tecnica per la simulazione di risonatori multidimensionali tramite digital waveguides è stata rivista ed ampliata, arrivando per la prima volta ad un'applicazione per un modello di un sistema acustico reale. Il metodo è applicato all'equazione alle derivate parziali che descrive il moto di un

riverbero a piastra; partendo da una soluzione modale, ci si è ricondotti ad un sistema monodimensionale formalmente simile a quello usato per modellare le corde dispersive e con perdite, problema per cui esistono soluzioni consolidate in letteratura.

Da un punto di vista computazionale, il risparmio è notevole rispetto alle tecniche basate su griglie, quali differenze finite e digital waveguide mesh. Nell'esempio usato, il costo totale è approssimativamente di 8 MPOS (moltiplicazioni per istante di campionamento), per ciascuna waveguide estratta; nel nostro caso, usando 750 waveguides, si arriva ad un costo totale di 6000 MPOS, di gran lunga inferiore rispetto alle  $10^{10}$  operazioni richieste da uno schema alle differenze finite applicato allo stesso sistema [8]. Tuttavia, può darsi che l'applicazione diretta della sintesi modale risulti più efficiente in alcuni casi, ad esempio in presenza di alta dispersione nel materiale; di fatto, l'uso della memoria e la possibilità di parallelizzare il più possibile il codice sono fattori rilevanti per sfruttare l'hardware moderno. Sotto questo aspetto, la sintesi modale offre uno schema di calcolo più semplice e adatto alle caratteristiche architettoniche menzionate.

Esistono diverse direzioni verso cui ampliare questo lavoro; tra tutte, la più interessante è l'estensione al caso tridimensionale, ovvero alla simulazione di stanze rettangolari. Data la similarità del problema, è lecito ritenere che un'estensione sia possibile senza troppi sforzi; inoltre, in questo caso anche la soluzione modale è impraticabile per un'implementazione efficiente in tempo reale, e in genere la dispersione è di gran lunga più ridotta rispetto alle piastre. Un altro dominio che può avere una rilevanza applicativa è quello sferico; una simulazione approssimata tramite reti di ritardo interconnesse è già stata proposta in [19]. È difficile invece pensare di estendere la tecnica a domini non regolari, ovvero a forma rettangolare o circolare, in quanto in questo caso è anche arduo esprimere una soluzione in forma modale. Altre caratteristiche che potrebbero essere incluse sono, ad esempio, la modellazione delle nonlinearità dovute all'eccitazione ad elevate intensità; in letteratura esistono diverse soluzioni per la modulazione della tensione, usando sintesi modale [17], o per la simulazione dell'accoppiamento nonlineare tra vibrazioni trasversali e longitudinali [3].

## RIFERIMENTI

1. J. Adrien. *Representation of Musical Signals*, chapter The missing link: modal synthesis. MIT Press Cambridge, MA, USA, 1991.
2. D. Arfib, F. Keiler, and U. Zölzer. *DAFX-Digital Audio Effects*. John Wiley & Sons, New York, NY, USA, 2002.
3. B. Bank. *Physics-based Sound Synthesis of String Instruments Including Geometric Nonlinearities*. PhD thesis, Budapest University of Technology and Economics, Hungary, Feb. 2006. URL: <http://www.mit.bme.hu/~bank/phd>.
4. B. Bank, F. Avanzini, G. Borin, G. D. Poli, F. Fontana, and D. Rocchesso. Physically Informed Signal Processing Methods for Piano Sound Synthesis: A Research Overview. *EURASIP Journal on Applied Signal Processing*, 2003(10):941–952, 2003.
5. B. Bank and V. Valimaki. Robust loss filter design for digital waveguide synthesis of string tones. *Signal Processing Letters, IEEE*, 10(1):18–20, 2003.
6. J. Bensa, S. Bilbao, R. Kronland-Martinet, and J. Smith III. The simulation of piano string vibration: From physical models to finite difference schemes and digital waveguides. *The Journal of the Acoustical Society of America*, 114:1095, 2003.
7. S. Bilbao. Fast modal synthesis by digital waveguide extraction. *Signal Processing Letters, IEEE*, 13(1):1–4, 2006.
8. S. Bilbao. A Digital Plate Reverberation Algorithm. *Journal of the Audio Engineering Society*, 55(3):135, 2007.
9. S. Bilbao, S. University, and D. of Electrical Engineering. *Wave and Scattering Methods for the Numerical Integration of Partial Differential Equations*. PhD thesis, Stanford University, 2001.
10. B. Buontempo. "building your own plate reverb". Available online at <http://www.prosoundweb.com/recording/tapeop/2002>.
11. A. Chaigne and C. Lambourg. Time-domain simulation of damped impacted plates. I. Theory and experiments. *The Journal of the Acoustical Society of America*, 109:1422, 2001.
12. J. Dattorro. Effect Design: Part I Reverberator and Other Filters. *Journal of the Audio Engineering Society*, 45:660–684, 1997.
13. G. Essl, S. Serafin, P. Cook, and J. Smith. Theory of Banded Waveguides. *Computer Music Journal*, 28(1):37–50, 2004.
14. M. Lang and T. Laakso. Simple and robust method for the design of allpass filters using least-squares phase error criterion. *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on [see also Circuits and Systems II: Express Briefs, IEEE Transactions on]*, 41(1):40–48, 1994.
15. P. Morse and K. Ingard. *Theoretical Acoustics*. Princeton University Press, 1986.
16. E. Motuk, R. Woods, and S. Bilbao. Implementation of finite difference schemes for the wave equation on fpga. In *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005 (ICASSP'05)*, 2005.
17. S. Petrausch and R. Rabenstein. Tension modulated nonlinear 2d models for digital sound synthesis with the functional transformation method. In *13th European Signal Processing Conference (EUSIPCO 2005)*, Antalya, Turkey, 2005.

18. D. Rocchesso. The Ball within the Box: A Sound-processing Metaphor. *COMPUTER MUSIC JOURNAL*, 19:47–57, 1995.
19. D. Rocchesso and P. Dutilleux. Generalization of a 3-D Acoustic Resonator Model for the Simulation of Spherical Enclosures. *EURASIP JOURNAL ON APPLIED SIGNAL PROCESSING*, 1:15–26, 2001.
20. D. Rocchesso and F. Scalcon. Accurate dispersion simulation for piano strings. In *Proc. Nordic Acoustical Meeting*, 1996.
21. D. Rocchesso and F. Scalcon. Bandwidth of perceived inharmonicity for physical modeling of dispersive strings. *Speech and Audio Processing, IEEE Transactions on*, 7(5):597–601, 1999.
22. J. Smith. Physical Modeling Using Digital Waveguides. *Computer Music Journal*, 16:74–74, 1992.
23. J. Smith. Physical Audio Signal Processing. Available online at <http://ccrma.stanford.edu/~jos/pasp/pasp.html>, 2008.
24. J. Strickwerda. Finite Difference Schemes and Partial Differential Equations, 1989.
25. L. Trautmann and R. Rabenstein. *Digital Sound Synthesis by Physical Modeling Using the Functional Transformation Method*. Kluwer Academic/Plenum Publishers, 2003.
26. S. Zambon and D. Rocchesso. Space-time simulations of circular membranes by parallel comb filters. In *Proc. 4th Conference "Understanding and Creating Music"*, Caserta, Italy, Nov. 2004.

# Un campionatore per “Musica su due dimensioni” di Bruno Maderna

Marco Gasperini

Conservatorio

“Benedetto Marcello”

di Venezia

## SOMMARIO

In occasione dell'esecuzione di *Musica su due dimensioni* (1958), per flauto e registrazione stereofonica, di Bruno Maderna, nel corso del “Laboratorio di Musica e Tecnologie Digitali” del Conservatorio “B. Marcello” di Venezia nel febbraio 2008 (flautista, Elena Gabbrielli), è stato approntato un ambiente esecutivo in Max/MSP che affronta in maniera radicale i problemi interpretativi posti da questo brano. A tal fine sono state studiate delle possibilità per rendere più versatile la diffusione/esecuzione della registrazione stereofonica nell'intento di realizzare effettivamente quella “interpretazione bilaterale del solista e del tecnico”[1] auspicata da Maderna nelle istruzioni pubblicate con la partitura del brano.

## Parole Chiave

Bruno Maderna, musica elettroacustica, opera aperta, .

## INTRODUZIONE

*Musica su due dimensioni*, composto da Maderna nello Studio di Fonologia della RAI di Milano, è uno dei primi lavori rilevanti ad utilizzare contemporaneamente suoni acustici strumentali *dal vivo* con altri suoni (di origine elettronica e strumentale) diffusi da altoparlanti. Tuttavia un primo studio sulle possibilità di questa combinazione da parte del compositore risale ad un'opera del 1952 che porta lo stesso titolo, composta presso l'Università di Bonn, il che dimostra come fin dai primi esperimenti con gli strumenti elettronici egli pensasse all'integrazione tra la dimensione acustica e quella elettroacustica, suscitando per altro la perplessità di Werner Meyer-Eppler riguardo l'ammissibilità di tale *organico*[2][3]. Nel lavoro del 1958 Maderna affronta in modo originale il problema della convivenza tra la struttura cronologicamente definita della registrazione e l'esecuzione strumentale del solista, recependo alcune delle istanze estetiche che andavano diffondendosi sulla scorta dell'operato di John Cage: il compositore non definisce puntualmente delle sincronie tra il flauto ed il nastro magnetico (com'è il caso della forma definitiva di un'altra composizione per strumenti e nastro magnetico che veniva composta in quegli anni, *Kontakte* (1959-60), di Karlheinz Stockhausen [4]), ma individua dei campi cronologici in cui il flautista può interagire, sulla base della partitura, con gli accadimenti *proposti* dal nastro.

La composizione in esame nasce in un periodo di grande interesse nei confronti delle cosiddette forme *aperte*: il superamento dello strutturalismo integrale avveniva attraverso il ricorso a metodi di organizzazione statistici, casuali ed informali. Solo tra i frequentatori dello Studio di Fonologia, tra il 1956 ed il 1960, della questione se ne occuparono, oltre a Maderna, anche Luciano Berio, Henry Pousseur e John Cage, ospitati su invito di Berio, nonché Umberto Eco [5]. Quest'ultimo dà testimonianza di tale interesse nell'esordio del suo *Opera aperta*, prendendo lo spunto per il suo discorso

sull'indeterminazione nelle poetiche delle avanguardie proprio dall'attività dei summenzionati compositori, con cui era stato in rapporto diretto. In questo lavoro Eco sottolinea la problematicità nuova posta dalle opere musicali composte secondo i principi dell'*apertura* formale: “(...)queste nuove opere musicali consistono (...) non in un messaggio conchiuso e definito, non in una forma organizzata univocamente, ma in una possibilità di varie organizzazioni affidate all'iniziativa dell'interprete, e si presentano quindi non come opere finite che chiedono di essere rivissute e comprese in una direzione strutturale data, ma come opere 'aperte', che vengono portate a termine dall'interprete nello stesso momento in cui le fruisce esteticamente.”[6]

Se espressioni quali *opera aperta*, *apertura formale*, etc., si riferiscono a delle categorie storizzate è anche vero che gli ambiti estetici che sono stati via via compresi da questi significanti si sono allargati esponenzialmente. Tuttavia proprio questa diffusione dimostra l'esorbitare da quella concezione originale che metteva in evidenza un'opposizione dialettica con un'ipotetica opera o forma in qualche modo *conclusa*. Dovremmo insomma considerare quelle espressioni come provvisorie ed indicative di una fase di definizione del problema. Esistono, è vero, nel campo musicale, una varietà di termini che specificano ognuno una particolare tecnica (musica aleatoria, indeterminata, stocastica, algoritmica, etc.) ma si tratta appunto di definizioni *parziali* che si riferiscono semplicemente ad una piuttosto che ad un'altra metodologia operativa, ma che non li contemplano da un livello superiore. Tali espressioni relative all'*apertura* saranno dunque utilizzate con le suddette riserve, non disponendo attualmente di una definizione più opportuna e condivisa, sulla quale ci si auspica che sia possibile sforzo teorico collettivo. Con questo termine si vorranno dunque indicare quelle attitudini formali indirizzate non già alla definizione di precisi svolgimenti cronologici ma alla delineatura di campi, più o meno ampi, di possibilità formative.

Nella seconda metà degli anni Cinquanta vedono la luce vari lavori che lasciano agli interpreti scelte di ordine formale. Sono perlopiù opere destinate ad un esecutore: per citarne alcuni, *Klavierstück XI* (1956) per pianoforte e *Zyklus* (1959) per un percussionista di Stockhausen; *Mobile* (1958) per due pianoforti di Pousseur; *Troisième sonate* per pianoforte di Boulez; *Sequenza I* (1958) per flauto solo di Berio; *Proporzioni* (1958) per flauto solo e *Aleatorio* (1959) per quartetto d'archi, tra gli altri, di Franco Evangelisti.

La natura dei mezzi tecnologici utilizzati nella composizione elettroacustica inibivano la possibilità di simili operazioni anche in quell'ambito: certamente in fase compositiva venivano utilizzate metodologie empiriche ed aleatorie (si pensi, per esempio, ad *Artikulation* di Gyorgy Ligeti) ma operazioni simili a quelle tentate nella musica strumentale erano impedito dal mezzo che convogliava l'opera elettroacustica: il nastro era un supporto fisso che non poteva essere maneggevolmente *ricomposto* fuori dallo studio. Non che il problema non fosse sentito in tutta la sua urgenza: il caso più emblematico in questo senso è quello di *Kontakte* (1958-

1960), per suoni elettronici, pianoforte e percussioni di Stockhausen, brano che, come è noto, inizialmente era stato concepito in tutt'altra forma da quella che assunse. Le varie strutture che formavano il nastro dovevano essere diffuse secondo ordini variabili e gestibili dal regista del suono e a queste strutture dovevano rispondere estemporaneamente i due strumentisti. L'insoddisfazione nei confronti dei risultati ottenuti durante le prove spinsero il compositore a scrivere una partitura definita nei minimi dettagli. Un tentativo riuscito di estendere alla produzione elettroacustica i risultati ottenuti nella pratica strumentale fu *Scambi* (1957) di Pousseur, composto e realizzato nello Studio di Milano, opera consistente in una serie scollegata di strutture che possono essere combinate e/o sovrapposte secondo alcuni principi dati di concatenamento.

*Musica su due dimensioni* si inserisce in questo ambito di ricerca e qui Maderna definisce una soluzione che informerà tutto il suo operato successivo. Egli deriva dalla tecnica empirica di composizione in studio la concezione della forma *aperta*: è a contatto con questo nuovo rapporto con la materia sonora che il suo pensiero musicale si indirizza verso l'utilizzo di componenti più o meno aleatorie che tanta rilevanza assumeranno nel seguito del suo operare e *Musica su due dimensioni* è il primo importante raggiungimento in questa direzione. "L'interprete deve intervenire nella struttura stessa dell'opera, rivelare secondo il proprio giudizio questo o quel possibile aspetto dell'opera...Le opere aperte, mobili, sono un'avventura necessaria del pensiero creativo del nostro tempo, al quale bisognava logicamente arrivare. E' un'acquisizione importante, perfino pericolosa, perché questa manipolazione dell'imprevedibile deve condurre al dischiudersi, al fiorire di bellezze che il compositore ha voluto multiple e continuamente nuove: ad una glorificazione della forma dunque e non alla sua negazione." [7] Le tecnologie dell'epoca, come si è detto, non permettevano una simile libertà anche per l'esecuzione della dimensione elettroacustica, imponendo al compositore la definizione di una struttura data una volta per sempre. *Musica su due dimensioni* è perciò luogo di una contrapposizione irrisolvibile: da un lato una partitura (quella del flauto), prevede delle zone di libertà formativa; d'altra parte a questa componente di alea, nel compiersi temporale di una delle due *dimensioni*, si contrappone l'irreversibilità nello svolgersi della registrazione, a discapito delle indicazioni di partitura in cui Maderna auspica una libertà bilaterale: "Tutta l'esecuzione di questa composizione deve avvenire su una specie di interpretazione bilaterale del solista e del tecnico, interpretazione che si può inventare di volta in volta." [1] L'idea di una lettura più vicina allo spirito originale di quest'opera rispetto alla tradizione interpretativa storica, ci ha spinti ad elaborare un sistema che permetta l'esecuzione, da parte del regista del suono, di alcuni frammenti tratti dalla registrazione, così che questi possa veramente partecipare in modo attivo all'invenzione, ogni volta rinnovata.

### TRADIZIONE INTERPRETATIVA

L'analisi di alcune interpretazioni del brano sono state fondamentali per definire l'ambito di interventi ammissibili. Il primo esempio preso in considerazione è fornito da un interprete d'eccezione: Luciano Berio. Si è avuta infatti la possibilità di esaminare la versione del nastro da lui utilizzata per un concerto tenuto a New York nel novembre 1995 con il flautista Michele Marasco. Si sono potuti notare degli interventi consistenti sulla registrazione, sotto forma di tagli, in particolare:

- 1) 12" iniziali del primo intervento;
- 2) tre tagli nel terzo intervento rispettivamente di 28", 36" e 54".

Confrontandola con il nastro originale la versione di Berio evidenzia un chiaro intento interpretativo: eliminazione dei silenzi e *condensazione* delle parti più diluite. Si prenda, per esempio il primo taglio: l'attacco originale prevede sei secondi di suono e sei di silenzio, un attacco che potrebbe essere definito *a singhiozzo*, cui segue un minuto praticamente ininterrotto di suoni. Berio taglia tutto l'attacco facendo cominciare l'intervento del nastro dopo i primi dodici secondi, scelta che segnala l'intenzione di marcare l'esordio del nastro affermativamente, al contrario dell'esordio originale carico di tensione e ambiguità.

Anche nella lettura di Renato Rivolta [8], sebbene non vi siano tagli di materiale sonoro, vengono eliminate le pause che separano i vari frammenti della terza parte, riducendone la durata dai 5'27" originali a 4'51".

Nella versione di Roberto Fabbri e Alvisi Vidolin [9] il nastro non viene ritoccato, ma accompagna tutto il brano, praticamente senza soluzione di continuità, a partire dalla fine dell'introduzione flautistica, comprendendo anche quello che, da partitura, dovrebbe essere un *a solo* di flauto (parte **IV**).

Un'altra lettura venuta all'attenzione è quella del flautista Giorgio Cagnoni, in una registrazione del 1974 al Teatro Valli di Reggio Emilia [10] che, se pur discutibile, presenta qualche interesse per il fatto che segmenta diversamente il materiale del nastro rispetto alla lettura di Rivolta ed elimina praticamente tutto il terzo intervento del nastro, mantenendo solamente il primo minuto (del resto Maderna non dà una segmentazione univoca del materiale sonoro, limitandosi a dire che "le parti successive *-alla seconda-* si intrecciano *ad libitum*, secondo l'interpretazione del solista (...)" [1]).

Questo rapido *excursus* sulla tradizione interpretativa ci ha dimostrato come gli interpreti del brano siano sempre intervenuti più o meno liberamente sulla forma, in relazione alla distribuzione e durata degli interventi della registrazione, ma abbiano sempre mantenuto una certa predilezione per quella che potremmo definire la *disposizione editoriale* (ciò vale parzialmente anche per la parte flautistica).

### GENERALITA' MACROSCOPICHE

*Musica su due dimensioni* è il primo nastro composto nello studio di Milano da Maderna in cui vengano utilizzati consistentemente, insieme a suoni elettronici, suoni strumentali e loro elaborazioni; tutti i lavori precedenti furono infatti realizzati unicamente con l'utilizzo di suoni elettronici, eventualmente sottoposti ad elaborazioni quali modulazione ad anello, riverberazione, filtraggio, etc. (*Notturmo*, *Syntaxis*, *Continuo*).

La forma globale del pezzo come viene illustrata da Maderna in partitura, è la seguente:

1. **sezione I**: flauto solo (indicata d'ora in poi come **I**);
2. **sezione II**: nastro solo per 30-120 secondi, poi flauto e nastro ed infine flauto solo (indicata come **II**);
3. **sezione III**: flauto e nastro (indicata come **III**);
4. **sezione IV**: flauto solo (indicata come **IV**);
5. **sezione V**: flauto e nastro (indicata come **V**).

Si può notare, a livello macroscopico, come gli interventi della registrazione in **II** e **III** siano informati da sonorità *strumentali* e da figure melodiche molto evidenti, con un progressivo addensamento del materiale ed incremento dinamico, culminante nella parte conclusiva di **III**. Lo scarto col materiale di **V** è brutale: esplosioni di sonorità molto penetranti (fasce di rumore modulate ad anello e riverberate), completamente slegate da modalità strumentali ed in netta opposizione con tutto il materiale precedente. La tensione a questo punto si distende progressivamente fino al ritorno a sonorità più morbide e *flautistiche*.

Non si intenda per sonorità *strumentali* che tutti i suoni siano ricavati da questo piuttosto che da quello strumento, ma che il criterio di scelta e di organizzazione dei vari materiali si

riferisce a modelli strumentali, in questo caso in maniera consistente al flauto. Le tecniche utilizzate nella composizione in studio non sono volte favorire una netta separazione percettiva tra i vari livelli del materiale (suoni strumentali, suoni strumentali elaborati, suoni sintetici) ma contribuiscono a creare una massa continuamente cangiante, portando in luce ora un elemento ora un altro, ora gli elementi di primo piano ora quelli sullo sfondo. Si vuole notare qui, per inciso, che Maderna pare recepire parzialmente l'osservazione postagli da Meyer-Eppler a proposito di *Musica su due dimensioni* (1952), sulla problematicità della convivenza tra suoni sintetici e strumenti *dal vivo*, suggerendogli, come soluzione, la possibilità di inserire i suoni strumentali nel processo di organizzazione elettronica[3].

Passando a considerare le possibilità di intervento formale degli esecutori è necessario sottolineare come i fattori aleatori non siano costanti nel corso del brano, ma distinti secondo tre livelli gradualità: il primo riguarda l'interazione tra flauto e registrazione in **II**, dove viene lasciata agli interpreti la scelta del momento di attacco del flauto, nei limiti indicati in partitura. Già questo tipo di scelta, tutto sommato limitata, richiede una serie decisioni interpretative: si potrebbe optare per la confusione iniziale tra le due fonti acustiche, oppure mantenere l'alternanza/opposizione dell'esordio facendo entrare il flauto sulla lunga pausa che si trova tra l'11" e l'59" del nastro; vanno successivamente definite anche le coincidenze che devono verificarsi tra le due parti, nonché il momento in cui il flautista si ritroverà da solo a concludere **II**. Agli interpreti quindi in questo caso viene affidato il compito di distribuire sul nastro la parte flautistica.

Il secondo livello di alea, è quello di **III** in cui non viene fornita, al solista, una disposizione delle funzioni musicali notate in partitura, ma gli si attribuisce la facoltà di liberamente interpolare e ripetere i vari frammenti dati, fatto salvo il vincolo di non ripetere alcuna delle figure poste tra parentesi quadre: sta agli esecutori definire l'ordine cronologico (o un modo di ordinare, nel caso si voglia una esecuzione estemporanea) dei vari frammenti del flauto. La stessa cosa vale per **V**, ma qui non vengono dati vincoli e possono essere utilizzati anche i frammenti di **III**, con l'esclusione di quelli posti tra parentesi (terzo livello di aleatorietà, massima entropia formale). "La ripetizione di questi frammenti non può che essere alternata, vale a dire che lo stesso frammento non deve venire ripetuto immediatamente, ma solo dopo una serie di frammenti diversi." [1] Si può notare come le *briglie* formali tendano progressivamente ad allentarsi nel corso del brano, mettendo a disposizione del solista una sempre maggiore gamma di materiali utilizzabili.

### ANALISI NASTRO E SUDDIVISIONE

Il primo intervento del nastro si pone come ponte tra l'esordio del flauto ed il suo secondo intervento. Come si è detto esiste un campo di possibilità cronologiche di azione. Nella nostra interpretazione il flauto entra dopo cinquantotto secondi e resta solo all'inizio del penultimo rigo (pedale di re), punto formalmente importante in quanto ripresa del materiale melodico iniziale della sezione. Può essere suddiviso in tre parti (0"-1'11"; 1'11"-1'58", 1'58"-3'14"), con una parte centrale caratterizzata da eventi sporadici dinamicamente tenui, inframezzati da lunghe pause. Questo intervento del nastro si caratterizza per l'utilizzo di evidenti figurazioni melodiche: i suoni sono sia di origine strumentale (frammenti melodici registrati), segnatamente flauto, che politimbriche (successioni di altezze formate con diversi materiali). Questo materiale melodico fa da contrappunto al flauto utilizzando figure estranee a quelle del solista (p.es. insistenza sul semitono, intervallo utilizzato con grande parsimonia nel flauto).

Una figura melodica importante è quella riportata in fig.1, presentata a 27.4", che in varie forme parziali e/o trasposte si ripresenta varie volte nel corso del brano.



Figura 1: tema *esposto* a 27.4"

In concomitanza con la fig.1 (dopo il re bemolle) si presenta regolarmente l'inciso di fig.2, ma lo troviamo pure isolato, p.es. a 2'47" (trasposto una sesta minore sopra, fa-sol).



Figura 2.

Le varie apparizioni e metamorfosi di queste cellule testimoniano della prassi compositiva nello Studio di Fonologia: punto di partenza erano alcuni incisi, più o meno articolati, che venivano ripetuti, variati, modificati, frammentati, giustapposti con altri, etc. (per uno studio analitico della prassi compositiva nello Studio di Milano si rimanda a [11]).

Si sottolinea l'importanza di un tipo di analisi anche intervallare sul nastro, vista la congruenza sintattica (e armonica) con le risultanze analitiche della parte di flauto (per cui si rimanda a [12]).

L'attenzione al contenuto melodico della dimensione elettroacustica risulta ancora più evidente analizzando il secondo intervento del nastro (**III**): l'attacco di questa sezione è costituito dalla registrazione dell'ultima parte di **I** (fig.3). Per tutta la durata di questo intervento, inoltre, è presente, nel canale destro, la registrazione integrale di **II**, quasi una sorta di lungo ritardo del flauto, che fa da ossatura al resto del materiale.



Figura 3: inciso melodico finale di **I**.

Se in **I** e **III** le sonorità della registrazione erano modellate su caratteristiche *strumentali* in **V** è il flauto ad essere portato verso sonorità esasperate per contrastare gli altoparlanti. L'intervento della registrazione in **V** è infatti molto diverso dai precedenti, segnalandosi per eventi sonori aggressivi, prevalentemente elettronici (rumore bianco filtrato, modulazione ad anello), separati da silenzi relativamente lunghi (5-6 secondi). Il flauto risponde a tono: *fff*, *ffff*, *il più ff possibile*, e via dicendo.. Sembra di intravedere in questo episodio l'opposizione, così densa di sviluppi, tra il singolo ed un fattore oppressivo, tema che verrà sviluppato da Maderna

nelle opere degli anni '60 (*Hyperion*, concerti per oboe, *Grande aulodia*). Dopo 4'30" dall'inizio di **V** compaiono nuovamente suoni flautistici, e si assiste ad un progressivo rilassarsi del magma sonoro precedente, fino alla conclusione.

### AMBIENTE ESECUTIVO

L'ambiente esecutivo per la diffusione della registrazione stereofonica è stato realizzato in ambiente **Max/MSP** e prevede l'utilizzo di un calcolatore, una tastiera MIDI ed un *controller* MIDI.

Per **II** è previsto semplicemente un lettore di file che legge il primo intervento; il regista del suono potrà controllare l'equilibrio dinamico con il flautista con un *fader* del *controller* MIDI.

Per quanto riguarda invece **III** e **V**, sulla base delle risultanze analitiche, sono state isolate le unità costitutive e, secondo questa suddivisione, il nastro è stato frammentato in parti più o meno lunghe (dai sei ai trenta secondi). Il fatto che nel nastro di **III** fosse presente integralmente e senza elaborazioni la registrazione di **II**, non poteva essere trascurato: si è proceduto perciò alla registrazione di questa parte, che viene riprodotta indipendentemente dal resto.

I frammenti così individuati vengono caricati in memoria singolarmente ad inizio esecuzione. Ogni frammento è associato ad un tasto di una tastiera MIDI, su due ottave contigue, distribuiti a seconda della loro provenienza (quelli di **III** su un'ottava e quelli di **V** sulla successiva). Per evitare la possibilità di *buchi* indesiderati tra un frammento e l'altro e per avere delle possibilità polifoniche, il tutto è stato duplicato su altre due ottave: il sistema permette quindi una polifonia a due voci. Il numero associato al tasto premuto (*key-number*) determina il frammento da leggere, mentre la velocità (*key-velocity*) determina il guadagno di due filtri passa-banda, uno per canale, con banda passante tra i 3000 ed i 4000 Hertz, e la frequenza di taglio di un passa-basso tra i 500 (velocità 0) ed i 15000 Hertz (velocità 127), nonché un traslatore di frequenza (Hilbert) che trasla il segnale, al massimo, di 7.5 Hertz. Questi controlli servono al controllo della dinamica della lettura dei vari frammenti. Oltre che tramite la velocità di pressione si è ritenuto opportuno, data la lunghezza di molti frammenti, poter modificare la dinamica anche dopo l'attacco con i potenziometri di un controller MIDI. Il regista del suono ha in questo modo un'ampio ventaglio di possibilità esecutive che vanno dalla scelta del frammento da riprodurre in un dato istante, al modo di attacco e alla sua evoluzione dinamica.

L'interfaccia grafica permette di visualizzare la forma d'onda dei frammenti corrispondenti agli ultimi due tasti premuti ed il nome dei tasti, nonché la velocità di pressione (che può variare dinamicamente). Un indicatore posto al di sotto della forma d'onda segnala il punto corrente di lettura(fig.4).



Figura 4: interfaccia ambiente esecutivo.

Il regista del suono dovrà avere una conoscenza perfetta dei vari frammenti e della loro disposizione sulla tastiera; dovrà inoltre sviluppare una sensibilità esecutiva in rapporto al flautista.

L'utilizzo di un sistema di diffusione ottonico ha posto il problema della messa nello spazio del brano. Innanzitutto si è deciso di non amplificare il flauto in quanto sarebbe stato completamente contrastante con il presupposto del brano, eliminando la relazione spaziale tra l'altrove del suono diffuso dagli altoparlanti e l'immanenza del suono del flauto: "(...)la prossimità del suono del flauto (...) è espressione di una vicinanza molto più vicina del più vicino diffusore sonoro, fosse anche il flauto durante l'esecuzione posto così distante da trovarsi alle soglie della percezione acustica dell'ascoltatore."[11]

Dato che la registrazione è a due canali ed il sistema elettroacustico a disposizione disponeva di otto altoparlanti, si è studiata una stereofonia estesa su otto canali. Sono state definite due modalità basilari di diffusione:

1. stereofonia frontale (figura 5a), per cui un canale viene assegnato, per esempio, ad un altoparlante a sinistra e l'altro a destra;
2. stereofonia laterale (figura 5b), per cui i canali sono distribuiti in senso ortogonale rispetto alla modalità precedente (un canale davanti e l'altro dietro).

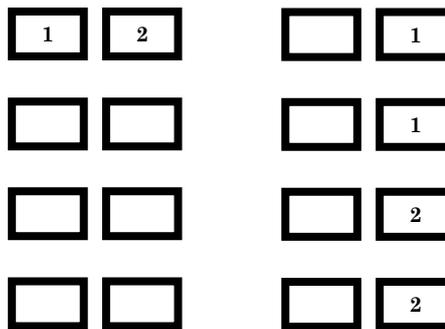


Figura 5a: stereofonia frontale

Figura 5b: stereofonia laterale

Estendendo queste due disposizioni agli altri lati si ottengono quattro possibilità di posizionamento dei segnali della registrazione. Oltre a ciò i diffusori in asse con una data sorgente tra quelle così individuate, ricevono lo stesso segnale, ritardato ed attenuato proporzionalmente alla distanza tra gli altoparlanti (fig.4). Ognuna delle due voci può essere assegnata indipendentemente a qualsiasi di queste possibili configurazioni.

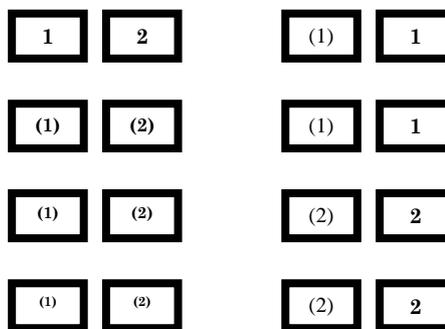


Figura 6

L'ambiente esecutivo permette di impostare questi valori di volta in volta diversamente, a seconda delle sale e del sistema utilizzato.

Nel corso dell'esecuzione queste varie disposizioni spaziali sono state distribuite nel seguente modo: in **II** posizione la stereofonia è frontale; in **III** le due voci vengono disposte a doppio coro (stereofonie laterali); in **V** una voce è sui due altoparlanti frontali e l'altra inizialmente sui due posteriori, per passare progressivamente alla stereofonia laterale sinistra ed infine sovrapporsi, nel finale, alla prima sui due diffusori frontali.

## CONCLUSIONI

Il lavoro analitico svolto su nastro e partitura di *Musica su due dimensioni*, oltre che sulla tradizione interpretativa dell'opera, ha portato alla definizione di un contesto esecutivo in cui i due solisti possano effettivamente ricomporre e concertare la forma del brano, permettendo inoltre nell'esecuzione una relazione *cameristica* tra i due, introducendo problemi di *tocco* e di *ritmo*.

## RIFERIMENTI

1. Maderna, Bruno. *Musica su due dimensioni*, Suvini Zerboni, Milano, 1960.
2. Maderna, Bruno. *Musica su due dimensioni* (1952), a cura di Nicola Scaldaferrì. Riedizione critica delle opere di Bruno Maderna diretta da Mario Baroni e Rossana Dal Monte, Suvini Zerboni, Milano.
3. Montecchi, Giordano. Continuo di Bruno Maderna, in *I quaderni della Civica Scuola di Milano*, Milano, 1992.
4. Maconie, Robin. *The works of Stockhausen*, Oxford University Press, 1976.
5. Luciano Berio, Umberto Eco, Roberto Leydi rievocano lo Studio di Fonologia a quarant'anni dalla fondazione, in AA.VV, *Nuova Musica alla Radio*, a cura di Veniero Rizzardi e Angela De Benedictis, CIDIM-ERI-AI-AMIC, Treviso, 2000.
6. Eco, Umberto. *Opera Aperta*, Bompiani, Milano, 1962.
7. Boucourechliev, André. *La musique sérielle, aujourd'hui*, in "Previe", XV. Citato in Sandro, Gorli. *Premessa* all'edizione critica di Don Perlimplino (1961), Suvini Zerboni, Milano, 2001.
8. Bruno Maderna, *Musica elettronica*, Stradivarius STR33349.
9. *Flute XXth century-vol.2*, ARTS47702.
10. [http://www.magazzinisonori.it/esplora\\_contenuti/autori\\_eseutori/maderna\\_bruno.aspx](http://www.magazzinisonori.it/esplora_contenuti/autori_eseutori/maderna_bruno.aspx)
11. Scaldaferrì, Nicola. *Lo Studio di Fonologia Musicale della Rai di Milano negli anni cinquanta*, tesi di laurea, Università degli studi di Bologna, 1994.
12. Del Prete, Michele. *Analisi di Musica su due dimensioni per flauto e registrazione stereofonica di Bruno Maderna*, scritto presentato per il Corso di Analisi della Musica Contemporanea I, Conservatorio "B. Marcello", Venezia, A.A 2005/2006, inedito.



# Sviluppo di un'interfaccia audio-aptica basata sulla profondità spaziale

**Stefano Delle Monache, Delphine Devallez  
e Pietro Polotti**  
Dipartimento di Informatica  
Università Degli Studi di Verona

**Davide Rocchesso**  
Dipartimento delle Arti e  
del Disegno Industriale  
Università IUAV di Venezia

## SOMMARIO

La spazializzazione in profondità di sorgenti sonore consente di effettuare una esposizione gerarchica di molteplici stream audio e potrebbe pertanto costituire uno strumento efficace nello sviluppo di nuove interfacce uditive. In questo lavoro viene presentata una interfaccia audio-aptica per una navigazione uditiva basata sulla resa acustica della distanza al fine di ordinare le sorgenti sonore in profondità. L'interfaccia aptica è composta da un nastro conduttivo, in grado di rilevare la posizione del tocco lungo di esso. La posizione sul nastro viene mappata sul punto di ascolto di uno spazio virtuale rettangolare, modellato secondo una Digital Waveguide Mesh, in modo da fornire delle misure di distanza di cinque sorgenti sonore equispaziate. In particolare, il modello enfatizza l'effetto di riverberazione, uno degli indizi acustici principali della distanza. In una cornice ecologica di interazione, la sintesi e riverberazione dei suoni utilizzano modelli fisici. I contenuti sonori sono organizzati per analogia con l'esercitazione di Basic Design dell'Antiprimadonna.

## Parole Chiave

Interfaccia Audio-aptica; Spazializzazione del Suono; Distanza Acustica; Sintesi per Modelli Fisici; Sonic Interaction Design; Basic Design.

## INTRODUZIONE

Nella vita quotidiana siamo abituati ad utilizzare il senso dell'udito per monitorare la realtà circostante. L'abilità di effettuare una discriminazione delle sorgenti sonore ci induce ad organizzare le informazioni presenti nella scena uditiva secondo modalità di attenzione di tipo gerarchico, nel tipico rapporto di figura-sfondo. Sin dagli albori della musica elettroacustica, le tecniche di spazializzazione del suono, sia quelle basate sulla direzionalità che quelle basate sulla profondità, sono state impiegate per disegnare spazi acustici e migliorare, anche in senso creativo-espressivo, la scena uditiva.

Il lavoro in oggetto consiste in una interfaccia audio-aptica per lo scorrimento di sorgenti sonore, basato sulla resa acustica della loro distanza relativa, al fine di ordinarle secondo un criterio di profondità spaziale.

Laddove la maggior parte della ricerca sul design di nuove interfacce sonore è focalizzata sulla spazializzazione direzionale di sorgenti sonore multiple, l'interfaccia ivi presentata si basa sulle qualità informative della distanza: un'informazione basata sulla profondità spaziale è in grado di fornire delle relazioni gerarchiche tra gli oggetti (sonanti) e pertanto

di richiamare l'attenzione dell'utente sulla sorgente sonora più prossima (principio di "acoustic zooming"), pur continuando l'ascolto delle altre sullo sfondo.

L'articolo ha la seguente struttura: nella sezione INTERFACCIA UTENTE viene descritta l'interfaccia utente, in particolare la progettazione e il mapping di controllo tra il nastro conduttivo e lo spazio acustico virtuale; la sezione PRODUZIONE DELLE SORGENTI SONORE descrive l'approccio utilizzato nel design delle sorgenti sonore, secondo una prospettiva ecologica, e introducendo gli strumenti per la sintesi per modelli fisici contenuti nel Sound Design Toolkit; nella sezione ORGANIZZAZIONE DELLE SORGENTI SONORE viene affrontato il tema dell'organizzazione delle sorgenti sonore sotto un profilo di Basic Design; nella sezione MODELLIZZAZIONE DELLO SPAZIO ACUSTICO viene data una descrizione del funzionamento e della modellizzazione dello spazio acustico virtuale.

## INTERFACCIA UTENTE

### Design e Mapping

In questa sezione viene presentata l'interfaccia per la navigazione tra diversi stream sonori ordinati in profondità in relazione all'input aptico dell'utente. L'informazione uditiva della distanza è resa mediante una Digital Waveguide Mesh (DWM) che simula una membrana 2D (v. Sezione MODELLIZZAZIONE DELLO SPAZIO ACUSTICO). L'interfaccia aptica consiste in un nastro conduttivo, il sensore *SlideLong* della Infusion Systems<sup>1</sup>, in grado di rilevare la posizione lineare del tocco lungo il nastro stesso, mediante variazione della resistenza offerta del circuito elettrico formato dal sensore e dalla scheda di acquisizione. La tipologia di sensore, e di interazione, è quella comunemente impiegata per la navigazione delle librerie musicali nei lettori portatili che montano sensori di posizione. Peraltro il sensore di posizione impiegato è dotato di un'area attiva di  $21 \times 2 \text{ cm}^2$ , così da rendere chiara la percezione dell'informazione relativa alla resa acustica della distanza. Inoltre la geometria rettangolare del nastro consente una intuitiva analogia con la geometria, anch'essa rettangolare, della DWM che simula l'ambiente uditivo. Il controller ritorna un valore dipendente dalla posizione rispetto alla quale si esercita la pressione: la posizione sul nastro del dito dell'utente è mappata sulla posizione di un microfono virtuale applicato alla membrana, così da consentire una interazione continua.

<sup>1</sup>[http://infusionsystems.com/catalog/product\\_info.php/products\\_id/52](http://infusionsystems.com/catalog/product_info.php/products_id/52)

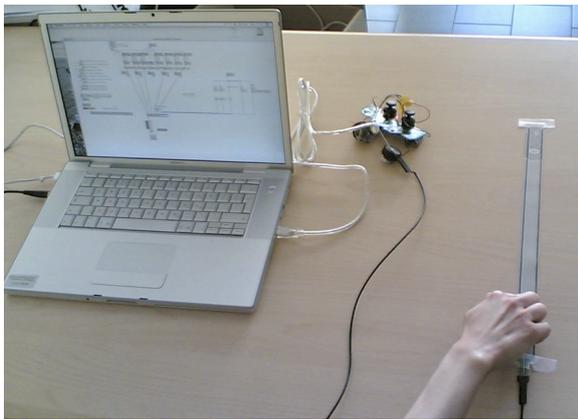


Figura 1. Immagine dell'implementazione.

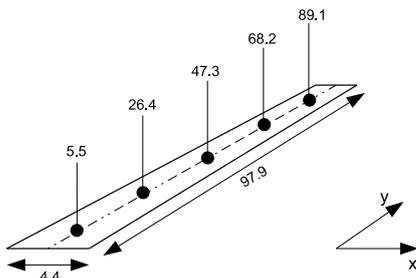


Figura 2. La membrana virtuale con le posizioni delle cinque sorgenti sonore. Tutte le misure sono in centimetri.

### Implementazione del prototipo

L'implementazione dell'interfaccia è mostrata nella figura 1. Il sensore ritorna un valore corrispondente alla posizione del tocco lungo il nastro. Un normale gamepad, collegato alla porta USB del computer, viene impiegato come scheda di acquisizione delle variazioni rilevate dal trasduttore: questi controller hanno il particolare vantaggio di essere piuttosto economici e di essere corredati di ingressi analogici compatibili con il segnale elettrico in uscita del sensore (0-5 volt) [8]. I valori in ingresso relativi alla posizione del tocco sono quindi letti in Max/MSP<sup>2</sup> mediante l'oggetto *hi* (*human interface*) e riscaldati in numeri decimali nel range 0-255. I valori in ingresso così riscaldati costituiscono la posizione di ascolto di entrata per il calcolo dei segnali acustici nella DWM. In questo modo, attraverso lo spostamento del dito sul nastro, l'utente può esplorare un ambiente virtuale composto di cinque stream audio collocati in posizioni diverse, come descritto nella figura 2. Inoltre da un punto di vista pratico, e coerentemente con l'obiettivo di rendere una informazione monodimensionale della distanza, viene stimata la sola posizione del tocco lungo il nastro. Viene trascurata la misura della larghezza del nastro, la cui grandezza, esigua, per lo più coincide con quella del dito. Per concludere, il sensore in esame è volto a fornire un'interfaccia intuitiva ed efficace, e dotata di una risposta immediata e coerente con i gesti esercitati dall'utente.

<sup>2</sup><http://www.cycling74.com/products/maxmsp>

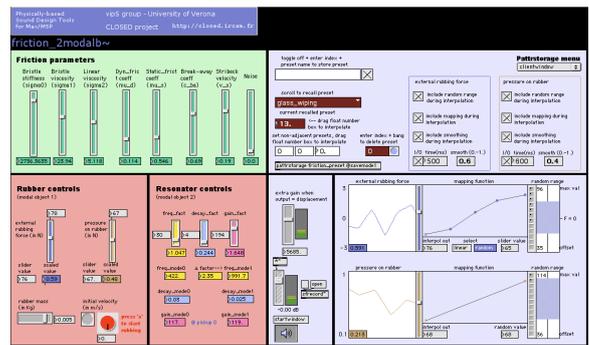


Figura 3. Interfaccia di controllo del modello di frizione.

### PRODUZIONE DELLE SORGENTI SONORE

It's not only what we hear that tell us what we know; what we know tell us what we hear.<sup>3</sup>

Un ulteriore aspetto riguarda il design e la generazione delle sorgenti sonore, in continuità con la prospettiva ecologica [6] sottesa all'interfaccia in esame. Per definizione il suono necessita di un evento produttore (sorgente), ossia il movimento di uno o più corpi. Le caratteristiche del suono risultante sono determinate dalle proprietà fisiche degli oggetti coinvolti, quali materiale, forma e dimensioni, e dalle tipologie delle interazioni coinvolte, in termini di forze applicate. Per contro gli esseri umani sono in grado di elaborare tutta una serie di informazioni complesse dalla percezione uditiva degli eventi sonori così prodotti: dal riconoscimento della sorgente sonora, dell'oggetto che ha generato il suono e delle sue proprietà, all'identificazione dell'interazione che l'ha generata, alla formazione di un giudizio di tipo semantico sull'evento sonoro percepito e, in definitiva, di una inferenza circa l'intenzionalità e l'espressività del gesto che l'ha prodotto [15].

Pertanto, nel design di interfacce uditive una sintesi sonora basata sulla modellazione delle sorgenti sonore, secondo il loro comportamento fisico, consente di sfruttare al meglio le qualità informative degli eventi sonori. In questo modo è possibile definire un'associazione naturale tra gestualità, espressività e controllo parametrico del modello di sintesi.

### Il Sound Design Toolkit (SDT)

Per la produzione del corpus sonoro si è fatto uso del Sound Design Toolkit (SDT)<sup>4</sup>. L'SDT consiste di una serie di strumenti di sintesi rivolti al design sonoro, coerenti sul piano percettivo e fisico. La figura 3 mostra uno screenshot dell'interfaccia di controllo del modello di frizione. Inizialmente

<sup>3</sup>Cit. R.A. Cole e J. Jakimik: Understanding speech: how words are heard, in G. Underwood (Ed.): Strategies of information processing, New York, Academic Press, 1978, p. 113.

<sup>4</sup>Disponibile per il download al seguente url: <http://closed.ircam.fr/deliverables.html>. Il pacchetto SDT eredita la ricerca iniziata con l'analisi fisica dei suoni di impatto, poi completata da altri modelli fisici di basso livello, come la frizione, e quindi seguita modelli di sintesi di alto livello che integrano i modelli di basso livello in paradigmi di sintesi più elaborati, come il rolling e il crumpling. Attualmente gli algoritmi disponibili sono compilati sia per Pure Data che per Max/MSP, nella forma di patch ed *external*.

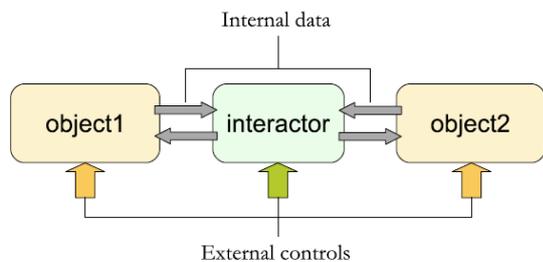


Figura 4. Struttura comune degli algoritmi di contatto tra solidi.

concepito nell'ambito del progetto europeo di ricerca SOB (Sounding Object) [12], l'attuale pacchetto software è mantenuto e sviluppato dall'Università di Verona nell'ambito del progetto europeo CLOSED<sup>5</sup> (Closing the Loop Of Sound Evaluation and Design) [14].

La maggior parte degli algoritmi sono caratterizzati da una struttura modulare del tipo risonatore-interattore-risonatore, che rappresenta l'interazione tra due oggetti risonanti, come si evidenzia nella figura 4.

Secondo questa struttura, ad ogni istante (campiono) entrambi gli oggetti inviano i rispettivi dati di stato (posizione e velocità al punto di interazione) all'interattore, il quale ritorna in cambio le nuove forze (opposte) così calcolate agli oggetti. Sulla base delle nuove forze applicate, gli oggetti sono in grado di calcolare il loro nuovo stato interno relativo all'istante successivo. In altre parole c'è una comunicazione bidirezionale fra i tre modelli.

I risonatori sono modellati mediante tecniche modale e digital waveguide. La prima descrive un oggetto in vibrazione attraverso i suoi modi di risonanza (frequenza, attacco e relativa energia). La seconda simula invece la propagazione delle onde attraverso un corpo elastico (ad esempio, nel caso di una sola dimensione, una corda ideale). Grazie alla modularità della configurazione adottata, è possibile connettere una qualsiasi coppia di risonatori attraverso complessi (non lineari) modelli di interazione. Nel corso della ricerca sono stati sviluppati diversi modelli per la sintesi sonora di eventi semplici, quali impatti, frizioni, bolle. In molti casi questi eventi sonori sono risultati essere delle vere e proprie primitive sonore di molti eventi complessi. Ad esempio, i suoni di rotolamento, di rimbalzo, di accartocciamento vengono implementati mediante complesse articolazioni temporali della generazione degli eventi elementari. Il modello di frizione è impiegato invece per simulare suoni di sfregamento, di scricchiolio e di frenata. Il modello di bolla costituisce invece la base dei suoni di gorgoglio, di gocciolamento e di versamento.

Data la coerenza fisica dei modelli, sono intuitibili l'associazione dei parametri di controllo con interazioni fisiche e continue, e la descrizione di risonatori e interattori mediante le loro proprietà fisiche e geometriche. Ad esempio, nel simulare una corda percossa, è possibile impostare la lunghezza, la massa e la tensione (risonatore waveguide), la massa del martello (risonatore modale), la forma e la rigidità del feltro del martello (interattore dell'impatto), la posizione della

percussione lungo la corda, la forza e la velocità applicate al martello, e così via.

Questo tipo di approccio al design del suono è comune alla disciplina del Sonic Interaction Design (SID) [13]. Di recente acquisizione, il SID sintetizza gli ambiti di ricerca e studio sull'acustica ecologica, sul paesaggio sonoro, sulle condotte di ascolto e sul design dell'interazione. Nel SID, l'aspetto funzionale caratteristico del suono, ossia la sua capacità di portare con sé delle informazioni, gioca un ruolo fondamentale, al fine di creare o rivelare nuove funzionalità (quando queste non sono direttamente apparenti), valorizzare l'identità sonora di oggetti o migliorare la performance e l'usabilità durante l'interazione. L'informazione sonora concorre con quella visiva, tattile e aptica alla formazione dell'esperienza del mondo che ci circonda.

## ORGANIZZAZIONE DELLE SORGENTI SONORE

La questione dell'organizzazione delle sorgenti sonore attiene alla forma ed espressività dei contenuti sonori e alla loro leggibilità, ossia alle strutture di controllo della sintesi sonora, modellazione degli eventi sonori e dei gesti, e loro distribuzione nel tempo.

Sotto questo aspetto, l'approccio seguito è quello del Basic Design [11] e del "problem solving", e cioè del descrivere, del configurare esperienze percettive anche complesse sulla base di grammatiche elementari condivise [9]. Anche se in continua evoluzione per la comunicazione grafica, la sintassi di base di queste grammatiche rimane quella elaborata dopo il periodo del Bauhaus e fa riferimento ad elementi visivi quali punto, linea, superficie, testura e colore, organizzati, tra gli altri, secondo principi di scala, contrasto, movimento, ritmo, equilibrio.

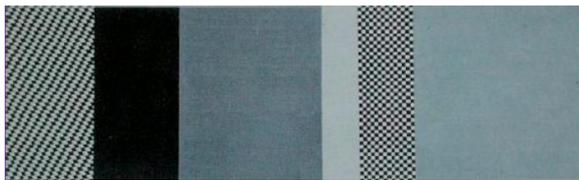
Sul piano del Sonic Interaction Design, tale metodo consente di affrontare la complessità del design dell'interazione, attraverso un'analisi delle relazioni e degli elementi primari che la compongono [4, 5]. Inoltre un vocabolario del Basic Design si rivela particolarmente fecondo nel gettare un ponte verso la costruzione di una grammatica condivisa tra i diversi ambiti del design: in particolare, dal punto di vista della creazione sonora, questo porta ad una distinzione anche nella pratica, del suono funzionale dal suono musicale, così superando la mera giustapposizione delle grammatiche e retorica musicali al design del suono.

## L'Antiprimadonna

Il corpus delle sorgenti sonore è organizzato per analogia secondo i principi dell'Antiprimadonna [1], un'esercitazione ideata da Tomàs Maldonado, esponente della scuola di ULM [7]: nella figura 5, sette bande verticali di larghezza variabile, cinque di diverso colore (tinta piatta) e due di una diversa trama di bianco e nero, vanno organizzate in modo che nessuna, e in generale nessuna parte della configurazione, risalti rispetto alle altre e quindi eserciti il ruolo della primadonna. Obiettivo dell'esercitazione è di produrre pattern non gerarchizzanti, e quindi sviluppare l'abilità a realizzare gerarchie percettive.

In questa realizzazione, l'Antiprimadonna sonora è composta da cinque pannelli sonori suddivisibili in due texture e tre gesti. L'intera composizione costituisce un paesaggio di

<sup>5</sup><http://closed.ircam.fr>



**Figura 5.** Docente Toms Maldonado, *Antiprimadonna*, Giovanni Anceschi 1962/1963.



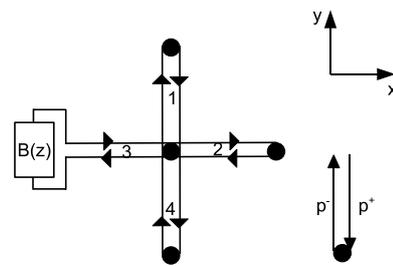
**Figura 6.** Rappresentazione dei 5 pannelli sonori nell'*Antiprimadonna*.

eventi sonori e interazioni elementari, realizzati con alcuni dei modelli di sintesi disponibili nel SDT e così strutturati, come da figura 6:

- Primo pannello: ospita una texture di suono di “frittura”. Per la sintesi sonora è stato utilizzato il modello fisico del crumpling (accartocciamento). La texture ricorda il tipico suono del “soffritto”.
- Secondo pannello: ospita un suono di cigolio piuttosto regolare, e precisamente del cardine di metallo di una altilena. Il suono è realizzato mediante modello di frizione.
- Terzo pannello: ospita suoni di impatto su legno. La distribuzione temporale degli impatti richiama l’immagine del tipico tamburellare con le nocche della mano su un oggetto cavo di piccole dimensioni, quasi ad esplorarne la forma. Lo spazio timbrico è ottenuto mediante interpolazioni dinamiche tra due principali configurazioni del modello di impatto.
- Quarto pannello: ospita un suono di cigolio del cardine di legno di un armadietto. Il suono è realizzato mediante modello di frizione. La configurazione temporale delle forze (pressione esercitata e velocità di sfregamento) esplicita in modo efficace l’immagine del movimento dello sportello e della gestualità associata.
- Quinto pannello: ospita una texture sonora di gocciolamento, realizzata con il modello del bubblestream.

Sul piano design del suono, questo approccio appare particolarmente efficace: dal punto di vista spettrale le cinque sorgenti occupano aree frequenziali diverse, senza particolari sovrapposizioni che possano causare fenomeni di masking. Sotto il profilo ritmico e temporale le sorgenti sono distribuite all’interno del quadro unitario secondo mutevoli relazioni di simmetria e asimmetria, ripetizione e cambiamento, elementi di dissonanza, in modo da conseguire stati di stabilità e sorpresa.

Nello specifico la ricerca dell’assenza di una gerarchia tra i cinque pannelli sonori concorre, con la tecnica della spazializzazione in profondità, alla maggior efficacia dello scorri-



**Figura 7.** Zoom su un nodo collegato agli altri nodi attraverso le guide d’onda 1, 2 e 4. La guida d’onda 3 è rivolta verso un margine parzialmente assorbente. I triangoli riempiti in neretto rappresentano le unità di ritardo secondo il loro orientamento.

mento audio-aptico delle sorgenti sonore, laddove la spazializzazione svolge la funzione esattamente opposta: creare ed enfatizzare delle gerarchie percettive chiare e precise tra le sorgenti.

### MODELLIZZAZIONE DELLO SPAZIO ACUSTICO

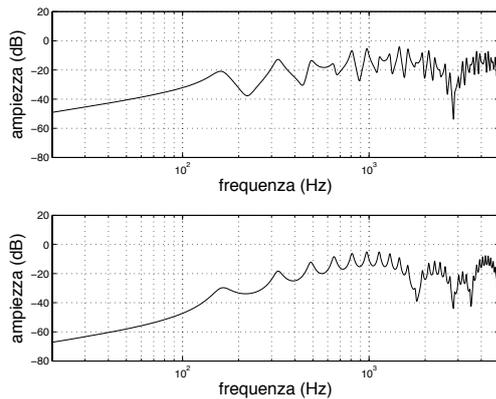
Nel campo della spazializzazione del suono, recenti ricerche si sono concentrate sulla modellazione fisica della propagazione acustica. In particolare, la scelta di un approccio fisico come la modellazione con tecnica digital waveguide appare convincente e sensato, dati l’accessibile formulazione del modello e il carattere intuitivo del relativo controllo mediante parametri che hanno un peso fisico. Ad esempio è stata testata l’efficacia della Digital Waveguide Mesh (DWM) nella modellazione di un parallelepipedo rettangolare, al fine di fornire degli indici di profondità acustica [3]. Esperimenti di ascolto effettuati con questo modello hanno evidenziato la capacità della DWM nel rendere la distanza apparente tra sorgenti sonore.

#### La Digital Waveguide Mesh

L’ambiente acustico virtuale in esame consiste in un reticolo rettilineo a due dimensioni, le cui guide d’onda digitali simulano la trasmissione dell’onda sonora tra i vari nodi interni che lo compongono. Ciascuna guida d’onda modella la scomposizione dell’onda di un segnale di pressione  $p$  nelle sue componenti  $p^+$  e  $p^-$ . Ciascun nodo propaga senza dispersione 4 segnali di input in arrivo dalle direzioni ortogonali  $p_1^+, \dots, p_4^+$  nei rispettivi segnali di output  $p_1^-, \dots, p_4^-$  (v. figura 7). Nella resa acustica di uno spazio 3D le proprietà delle pareti svolgono un contributo importante. Questo è vero anche nel caso di ambienti acustico bidimensionali dal momento che le onde orizzontali interagiscono con i margini della superficie. Le riflessioni dai margini sono modellate mediante Digital Waveguide Filter (DWF), i cui coefficienti sono stati impostati in modo da simulare specifiche proprietà riflesse delle superfici reali [2]. Infine, il numero di nodi viene convertito nelle dimensioni della relativa membrana una volta determinati la velocità del suono e la frequenza di campionamento della simulazione. Per via dell’alto consumo di risorse computazionali richieste, si è scelto di impostare le dimensioni del reticolo in  $90 \times 5$  nodi, che corrispondono ad una membrana di  $97.9 \times 4.4 \text{ cm}^2$  ad una frequenza di campionamento di 44.1 kHz.

## Proprietà acustiche della Membrana

Al fine di esaminare gli indici della distanza acustica all'interno dell'ambiente virtuale, le risposte all'impulso vengono calcolate sulla membrana a differenti distanze, le quali corrispondono alle posizioni dei cinque stream audio. Si presume la natura puntuale della sorgente sonora e la sua collocazione sull'asse  $y$  della mesh, a 16 cm dal margine. Le misurazioni delle risposte all'impulso sono effettuate alle 5 posizioni, sull'asse  $y$  della membrana (v. figura 2). La simulazione dell'ambiente di ascolto è stata effettuata in ambiente `Matlab`. La figura 8 mostra la risposta in frequenza fino a 5 kHz, misurata rispettivamente a 26.4 cm e 89.1 cm della membrana virtuale.



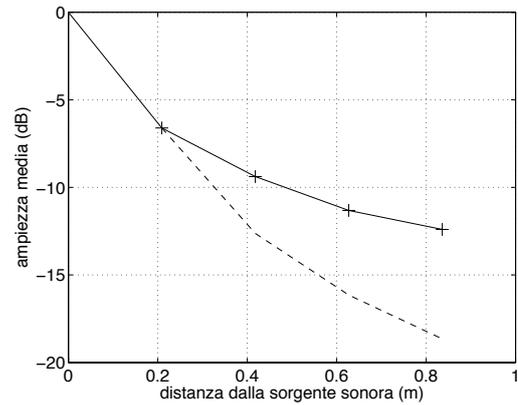
**Figura 8.** Risposte in frequenza fino a 5 kHz sulla membrana. Figura in alto: 26.4 cm. Figura in basso: 89.1 cm.

### Livello complessivo dell'intensità

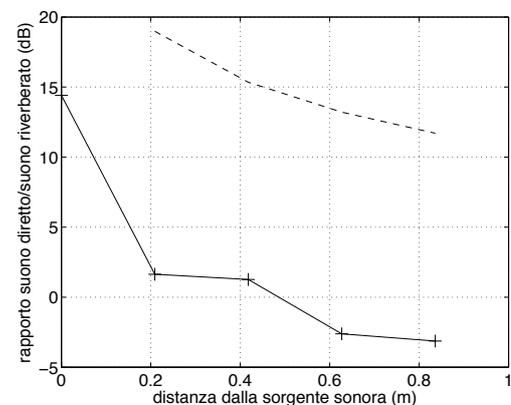
La figura 9 mostra la variazione dell'energia totale in relazione alla distanza: nel confronto con l'attenuazione dell'energia nello spazio aperto, caratterizzata dalla ben nota regola dei 6 dB di attenuazione, si può notare come l'intensità complessiva sulla membrana si attenui in maniera significativamente minore. Questo comportamento permette di ascoltare anche la sorgente sonora più lontana, a qualsiasi posizione sulla membrana, senza soffrire di eventuali manipolazioni del volume introdotte dall'utente.

### Rapporto tra le energie del suono diretto e riverberato

Per rendere l'informazione sulla distanza, lo spazio acustico virtuale così progettato si serve in particolar modo dell'indice del rapporto tra l'energia del segnale diretto e l'energia del segnale riverberato. Per ciascuna risposta all'impulso, il ritardo del suono diretto viene dedotto dalla distanza tra la sorgente sonora e il punto di ascolto, e quindi rimosso dalla risposta stessa. Quindi l'energia diretta viene integrata tra i primi 2.5 ms delle risposte all'impulso senza ritardo, il che approssima la durata delle Head Related Impulse Responses misurate in condizioni anecoiche per poi catturare la traiettoria diretta del segnale sonoro [10]. Infine l'energia della riverberazione è calcolata in base alla coda delle risposte all'impulso senza ritardo. La figura 10 mostra i valori dei rapporti dell'energia diretta e quella di riverbero per diverse distanze sulla mesh.



**Figura 9.** Ampiezza media della risposta all'impulso come funzione della distanza. Linea continua: mesh 2D. Linea tratteggiata: spazio aperto di riferimento.



**Figura 10.** Rapporto tra l'energia del segnale diretto e quella del segnale riverberato come funzione della distanza. Linea continua: mesh 2D. Linea tratteggiata: approssimazione di un ambiente naturale.

Per un raffronto il rapporto  $v$  tra l'energia diretta e quella di riverbero è stato calcolato per un ambiente naturale, modellato da Zahorik secondo la funzione  $v = -3.64 \log_2(r) + 10.76$  [16]. I valori dei rapporti risultano essere notevolmente inferiori nella mesh 2D rispetto allo spazio acustico naturale, come si può notare da un ulteriore confronto della linea solida con quella tratteggiata, nella figura 10. Da ciò si desume che la quantità di riverberazione viene enfatizzata nell'ambiente virtuale.

## CONCLUSIONE

Alla fase di realizzazione è seguita una parte di fruizione libera da parte di partecipanti a workshop nell'ambito dei quali è stata presentata l'interfaccia. In media le persone coinvolte hanno evidenziato l'aspetto intuitivo e divertente dell'interfaccia aptica.

Questo studio ha offerto diversi interessanti spunti di ricerca, sia in termini di interaction design che di auditory display e di percezione uditiva. Tra i vari punti emersi, la natura e il numero di sorgenti sonore concomitantemente presenti nell'archivio influiscono sensibilmente sulla qualità della per-

cezione. La questione potrebbe essere risolta mediante un giusto compromesso tra sovraccarico uditivo e obiettivi: un design del suono ecologicamente orientato e attento ai profili spettrali e temporali dell'intero corpus sonoro spiega la sua efficacia quanto più riesce a mettere a fuoco gli aspetti funzionali. Sotto questo aspetto un approccio di Basic Sonic Interaction Design si rivela essere molto fecondo. Sono in programma ulteriori indagini relative all'impiego dell'interfaccia in oggetto: in particolare la configurazione di menù uditivi, mediante sonificazione delle voci che li compongono, spazializzata in profondità tramite DWM e accessibile attraverso interfaccia tattile. Un esempio è la sonificazione mediante diversi livelli di profondità della barra delle applicazioni dei sistemi operativi.

## RIFERIMENTI

1. G. Anceschi. Basic Design, fundamenta del design. In G. Anceschi, M. Botta, and M. A. Garito, editors, *L'ambiente dell'apprendimento – Web design e processi cognitivi*, pages 57–67. McGraw Hill, Milan, Italy, 2006.
2. F. Fontana. *Physics-based models for the acoustic representation of space in virtual environments*. PhD thesis, University of Verona, Italy, 2003.
3. F. Fontana and D. Rocchesso. A physics-based approach to the presentation of acoustic depth. In *Proc. of the 2003 International Conference on Auditory Display*, pages 79–82, July 2003.
4. K. Franinovic. Basic interaction design for sonic artefacts in everyday contexts. In *Focused - Current Design Research Projects and Methods*, Bern, Switzerland, 2008. Swiss Design Network.
5. K. Franinovic and Y. Visell. Strategies for sonic interaction design: from context to basic design. In *Proc. of the 14th International Conference on Auditory Display, Paris, France*, June 2008.
6. W. W. Gaver. What in the world do we hear? an ecological approach to auditory source perception. *Ecological Psychology*, 5(1):1–29, 1993.
7. E. H. Lindinger. *La Scuola di Ulm*. Costa & Nolan, Genova, 1988.
8. A. Jensenius, R. Koehly, and M. Wanderley. Building low-cost music controllers. *LNCS*, 3902:123–129, 2006.
9. E. Lupton and J. C. Phillips. *Graphic Design: The New Basics*. Princeton Architectural Press, 2008. <http://gdbasics.com/>.
10. H. Møller, M. Sørensen, D. Hammershøi, and C. Jensen. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.*, 43:300–321, 1995.
11. D. Riccò. Basic design sinestesico. fondamenti teorici. In G. Anceschi, M. Botta, and M. A. Garito, editors, *L'ambiente dell'apprendimento – Web design e processi cognitivi*, pages 97–110. McGraw Hill, Milan, Italy, 2006.
12. D. Rocchesso and F. Fontana, editors. *The Sounding Object*. Mondo Estremo, 2003. <http://www.soundobject.org/>.
13. D. Rocchesso, S. Serafin, F. Behrendt, N. Bernardini, R. Bresin, G. Eckel, K. Franinovic, T. Hermann, S. Pauletto, P. Susini, and Y. Visell. Sonic interaction design: Sound, information and experience. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, April 2008.
14. P. Susini, N. Misdariis, G. Lemaitre, D. Rocchesso, P. Polotti, K. Franinovic, Y. Visell, and K. Obermayer. Closing the loop of sound evaluation and design. In *Proc. of the 2nd ISCA/DEGA Tutorial and Research Workshop on Perceptual Quality of Systems, Berlin, Germany*, September 2006.
15. N. J. Vanderveer. *Ecological acoustics: human perception of environmental sounds*. PhD thesis, Cornell University, USA, 1979.
16. P. Zahorik. Direct-to-reverberant energy ratio sensitivity. *J. Acoust. Soc. Am.*, 112(5):2110–2117, November 2002.

# OMAGGIO A STOCKHAUSEN

## technical set-up digitale per una performance di SOLO [Nr.19] *für melodieninstrument mit rückkopplung<sup>1</sup>*

a cura di

**Enrico Francioni**

*via G.Verdi, 15*

*I-61021 CARPEGNA PU - Italy*

*francioni61021@libero.it*

*http://www.myspace.com/enricofrancioni*

in collaborazione con il:

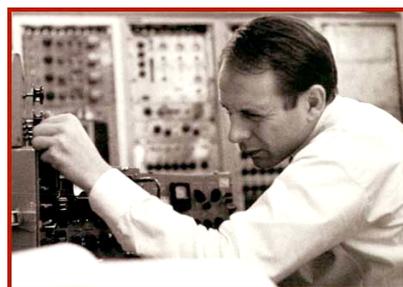
**L.E.M.S.**

*Laboratorio Elettronico per la Musica Sperimentale  
del Conservatorio Statale di Musica "G.Rossini" di Pesaro*

*P.tta Olivieri*

*I-61100 PESARO PU - Italy*

*http://www.conservatoriorossini.it/*



### INTRODUZIONE

**OMAGGIO A STOCKHAUSEN** è un progetto di ricerca nel settore della musica elettroacustica finalizzato alla trasposizione in ambito digitale dell'allestimento tecnico analogico originale (*technical set-up*) quale strumento irrinunciabile all'esecuzione di *Solo für melodieninstrument mit rückkopplung<sup>2</sup>* - lavoro del 1966 di **Karlheinz Stockhausen** (scomparso il 5 dicembre 2007).

L'idea di questo progetto nasce nell'a.a. 2005/2006 nel corso dello studio delle linee di ritardo e si sviluppa come un lavoro curato dallo scrivente (allora allievo) in collaborazione con il L.E.M.S. (*Laboratorio Elettronico per la Musica Sperimentale*) del Conservatorio Statale di Musica "G.Rossini" di Pesaro (Italia), istituzione alla quale va riconosciuto il merito di avere incoraggiato e sostenuto l'attività già dalla fase progettuale, sino alle *performances* pubbliche<sup>3</sup>.

Gli obiettivi più significativi raggiunti con la realizzazione di questo lavoro sono stati principalmente:

- l'aver voluto e l'aver poi concretizzato il desiderio di essere in condizione di eseguire *Solo* in una versione per solista<sup>4</sup> e Personal Computer (Mac) superando quelle difficoltà che l'allestimento originale (per forza di cose concepito con caratteristiche analogiche) pone all'interprete;
- l'aver sviluppato, attraverso l'impiego di un software di pubblico dominio, un codice linguistico informatico-musicale,

*in ambiente Macintosh, che costituisce un valido supporto per la performance in real-time, oltre che una base indispensabile per realizzare in ambiente digitale l'allestimento tecnico di partenza indicato dall'autore (technical set-up);*

- l'aver arricchito timbricamente il brano intervenendo sia sulla parte tecnica (l'algoritmo), sia nella parte tecnico-strumentale (la partitura);
- l'aver prodotto una versione scritta della parte (sei pagine destinate alla Version I), sulla base delle indicazioni date dall'autore e facendo uso di un editor di notazione musicale.

Tra i punti di forza di questo progetto si possono individuare:

- l'importanza di aver impiegato, in ambiente Macintosh, un frontend per il real-time (come MacCsound<sup>5</sup>) ed un linguaggio (Csound<sup>6</sup>) diffusi e distribuiti sul web in maniera libera (free) – open source – in continua espansione e accessibili a tutti;
- l'essere riusciti a rendere completamente autonomo l'esecutore nella parte relativa alla performance (One-man performance), ma anche libero di muoversi all'interno dell'azione di controllo dell'algoritmo;
- l'aver contribuito a dare una maggiore chiarezza di linguaggio, una funzionalità ed un aspetto estetico più positivo (rispetto alle 'storiche' precedenti versioni), dell'interfaccia grafica utente utile alla performance.

# 1. DALL'ANALOGICO AL DIGITALE

## 1.1 LA COSTRUZIONE DELL'ALGORITMO

Partendo dallo *Schematic Diagram* che accompagna la partitura e commentato in fig.1:

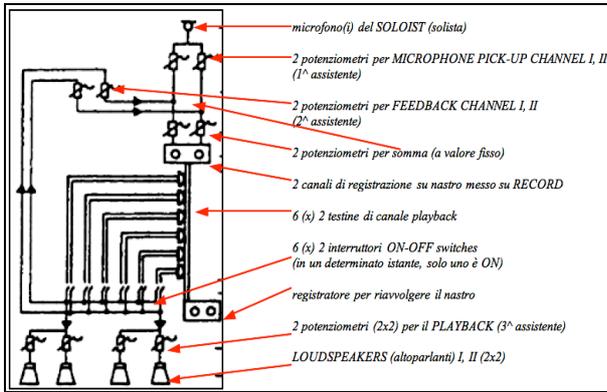


fig.1 - Schematic diagram originale

per prima cosa si trattava di capire con precisione il percorso del segnale e creare un altro diagramma di flusso più utile alla stesura del codice che mettesse in evidenza con chiarezza:

- le varie parti del meccanismo di controllo (gestito per via analogica dagli Assistenti<sup>7</sup>), che in sostanza costituiscono la parte principale del technical set-up del pezzo;
- il percorso delle due linee (L/R) della stereofonia e
- i riferimenti principali alle variabili che poi entrano in gioco nel codice dell'algoritmo scritto con MacCsound (compilato sulla base delle informazioni del nuovo diagramma di flusso).

Si decide di prendere in considerazione il *Form-Schema* della Versione I che ha una durata complessiva di 10':39":8" (fig.7).

Al *Form-Schema* scelto sono state associate le sei pagine, poi interpretate secondo l'Interpretations-Schema indicato dall'autore.

Il diagramma di flusso risultante scaturito dalle informazioni acquisite, anche tenendo conto delle Note introduttive presenti in partitura relative alla parte dell'allestimento tecnico, in maniera sintetica è quello che appare in fig.2.

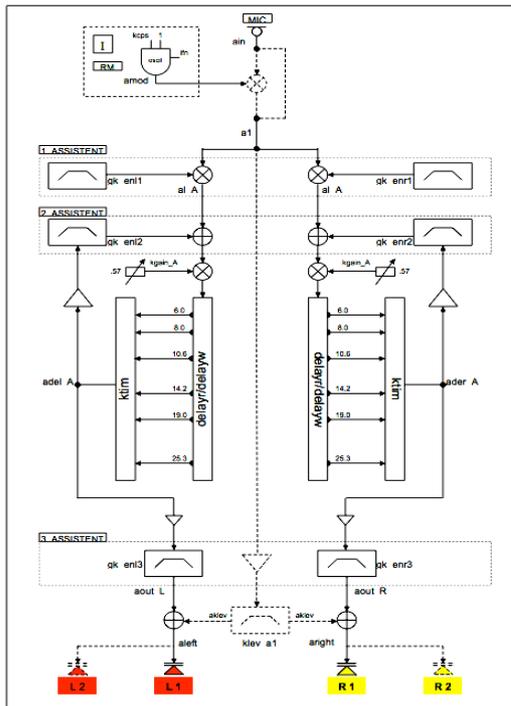


fig.2 - Nuovo diagramma di flusso

Su questa base è stato elaborato il codice dell'algoritmo dell'intero allestimento tecnico comprendente i files orchestra (*orc*) e score (*sco*), ma che in MacCsound sono compresi in un unico file *csound* (*csd*).

Cercheremo ora di analizzare da vicino il codice del file *.csd* nella parte relativa all'orchestra e, in maniera conseguente, allo score.

Il problema da affrontare immediatamente era l'automazione relativa agli involuppi (perché non altro che di involuppi si trattava) disegnati dall'azione di apertura o di chiusura degli *sliders* dei tre Assistenti: due *sliders* per ogni Assistente, per un totale di sei *sliders*, che si muovono su due canali (L/R), secondo le indicazioni del *Form-Schema*, nel nostro caso il I.

Attraverso sei variabili globali (variabili di tipo *gk*) invocate a monte di tutta l'orchestra, abbiamo cercato di rendere indipendenti le sei condotte di involuppo degli Assistenti.

Le sei variabili globali infatti vanno a far riferimento ad altrettanti strumenti (*instr*) che mettono in azione due indici (riferiti a due tabelle e quindi alle rispettive funzioni che fanno uso della GEN2 all'interno dello score) i quali scandiscono in maniera sincronica il tempo di ciascun evento e contestualmente la durata dell'involuppo all'interno dello stesso evento.

Ad esempio nella seguente apertura del canale Left del I Assistente all'interno del ciclo A, è stato preso in considerazione un evento della durata complessiva di 6 x 3 secondi all'interno del quale lo *slider* rimane aperto per i primi 6 secondi (fig.3):

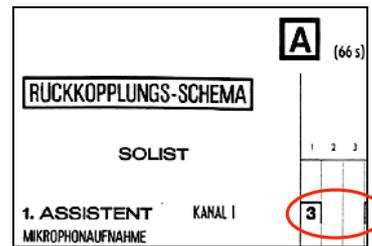


fig.3 - Apertura di sei secondi del Kanal I (Assistent 1)

Relativamente a questa parte, il modello di orchestra generato e poi replicato per tutti e sei gli *sliders* degli Assistenti, è stato il seguente (per comodità di illustrazione viene riportato di seguito solo quello del canale sinistro del primo Assistente).

;generazione della variabile globale - Assistente 1, Left

```

ia = .0001
ib = .0001
ifn1 = 1 ;riferimento alla funzione 1
ifn2 = 2 ;riferimento alla funzione 2
indx = init - 1
ilen = fflen(ifn1)
reset:
indx = indx + 1
if indx <= ilen - 1 goto again
turnoff

again:
idur = table indx, ifn1
idur_env = table indx, ifn2
timeout 0, idur, contin
reinit reset

contin:
gk_enr1 = linseg 0, ia, 1, idur_env - (ia + ib), 1, ib, 0
    
```

Lo stesso principio per la generazione degli involuppi (che sono poi andati a moltiplicarsi al segnale) è stato applicato anche alla parte relativa all'attivazione degli effetti timbrici come la RM (*Ring Modulation*).

In fig.4a il profilo d'involuppo generato dal codice compilato, relativo ai primi 66 secondi (ciclo A) del Canale Left

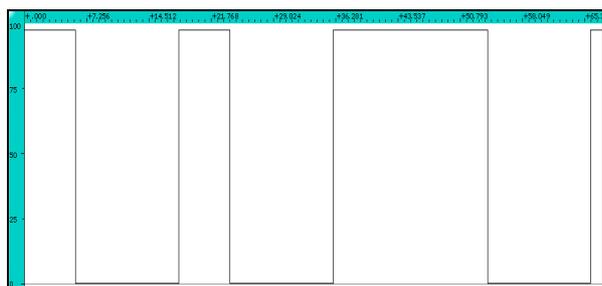


fig.4a – Profilo di involuppo generato dal codice

dell'Assistente I e, per un confronto, in fig.4b il frammento corrispondente tratto direttamente dal *Form-Schema*:

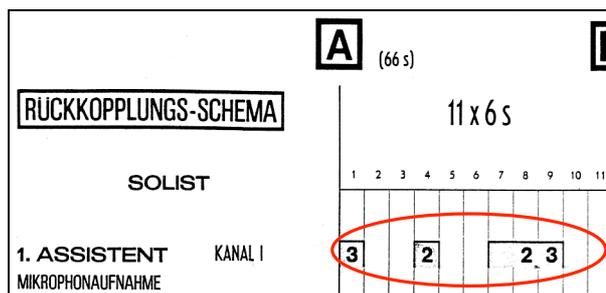


fig.4b – Profilo di involuppo come da *Form-Schema*

Per quanto riguarda invece la virtualizzazione della scatola dei pulsanti (*switches*), per la commutazione alle prese (*taps*) della

*delay-line*, si è proceduto in maniera più semplice come segue:

*;delay-line e switches, con valori di deltapi variabili nel tempo*

```

ktim          linseg 6,66,6,p4,14.2,113.6-p4,14.2,p4,19,133-p4,19,p4,25.3,151.8-p4,25.3,p4,10.6,95.4-p4,10.6,p4,8,80-p4,8
               ;gestione switches delle linee di ritardo
adummyl      delayr 26          ;valore massimo della delay-line
adel_A       deltapi ktim
               delayw al_A + (adel_A * gk_enl2) * kgain_A
    
```

Quest'ultima strategia ha dato immediatamente il risultato atteso ed è stata applicata ad entrambi i canali. Siamo in presenza di un vero e proprio commutatore, uno *switch*.

In questo caso il tempo p4 (che rappresenta la rampa *salita e discesa* dello *slider*) è stato prudentemente scelto con un valore minimo (0.0001 sec.) per evitare effetti di disturbo (*spike*).

Come si può osservare contestualmente, per la realizzazione delle *delay-lines* (una per canale) è stata impiegata la tripletta di opcodes *delayr/deltapi/delayw*, questo per poter avere la massima libertà nella gestione dei parametri.

Per la realizzazione della parte grafica, che appare nella schermata *i/o* di MacCsound, relativamente all'orchestra si è fatto ancora ricorso all'opcode *timeout*, anche se questa volta coinvolto in maniera molto meno articolata.

I valori di controllo (*k*) o di inizializzazione (*i*) sono stati fatti passare dall'orchestra all'*i/o* grazie all'uso dell'opcode *outvalue* ed in maniera inversa, dall'*i/o* all'orchestra, attraverso l'opcode *invalue*.

E' stata scritta quindi la parte di orchestra relativa alla gestione grafico/temporale; parliamo ad esempio del metronomo di scansione di inizio periodo (M1) e del metronomo della scansione degli impulsi all'interno dei periodi (M2) - strumenti

indispensabili per la *performance* del solista - che hanno seguito la seguente sintassi:

```

;metronomo che scandisce l'inizio di ogni periodo
idur = p4
daqua:
kenv linseg 0, p5, 1, p6, 1, p7, 0, idur - (2 * p5 + p6), 0
      timeout 0, idur, continua
      reinit daqua
continua:
      rireturn
      outvalue "M1", kenv
;metronomo che scandisce tutti gli impulsi
idur = p4
daqua:
kenv linseg 0, p5, 1, p6, 1, p7, 0, idur - (2 * p5 + p6), 0
      timeout 0, idur, continua
      reinit daqua
continua:
      rireturn
      outvalue "M2", kenv
    
```

La sintassi relativa alla gestione grafica della progressione delle linee del tempo in *i/o* all'interno di ogni ciclo, (raffigurata con sei meter di diverso colore disposti in orizzontale in basso all'interfaccia grafica) risulta invece ben più complessa:

```

;interfaccia: gestione progressione dei periodi (in corrispondenza al rispettivo ciclo)
ktime1 linseg 0, p4, 1/6, 66 - (p4 * 2), 1/6, p4, 0, p3 - 66, 0
ktime2 linseg 0, 66, 0, p4, 1/14.2, 113.6 - (p4 * 2), 1/14.2, p4, 0, p3 - (66 + 113.6), 0
ktime3 linseg 0, 66 + 113.6, 0, p4, 1/19, 133 - (p4 * 2), 1/19, p4, 0, p3 - (66 + 113.6 + 133), 0
ktime4 linseg 0, 66 + 113.6 + 133, 0, p4, 1/25.3, 151.8 - (p4 * 2), 1/25.3, p4, 0, p3 - (66 + 113.6 + 133 + 151.8), 0
ktime5 linseg 0, 66 + 113.6 + 133 + 151.8, 0, p4, 1/10.6, 95.4 - (p4 * 2), 1/10.6, p4, 0, p3 - (66 + 113.6 + 133 + 151.8 + 95.4), 0
ktime6 linseg 0, 66 + 113.6 + 133 + 151.8 + 95.4, 0, p4, 1/8, 80 - (p4 * 2), 1/8, p4, 0
k1 phasor ktime1
k2 phasor ktime2
k3 phasor ktime3
k4 phasor ktime4
k5 phasor ktime5
k6 phasor ktime6
outvalue "PH1", k1 ;progressione della linea grafica dei periodi all'interno del primo ciclo (meter)
outvalue "PH2", k2 ;progressione della linea grafica dei periodi all'interno del secondo ciclo (meter)
outvalue "PH3", k3 ;progressione della linea grafica dei periodi all'interno del terzo ciclo (meter)
outvalue "PH4", k4 ;progressione della linea grafica dei periodi all'interno del quarto ciclo (meter)
outvalue "PH5", k5 ;progressione della linea grafica dei periodi all'interno del quinto ciclo (meter)
    
```

outvalue	"Ph6",	k6	;progressione della linea grafica dei periodi all'interno del sesto ciclo (meter)
outvalue	"Ph1",	k1 * 6	;progressione valori temporali numerici dei singoli periodi del primo ciclo (text)
outvalue	"Ph2",	k2 * 14.2	;progressione valori temporali numerici dei singoli periodi del secondo ciclo (text)
outvalue	"Ph3",	k3 * 19	;progressione valori temporali numerici dei singoli periodi del terzo ciclo (text)
outvalue	"Ph4",	k4 * 25.3	;progressione valori temporali numerici dei singoli periodi del quarto ciclo (text)
outvalue	"Ph5",	k5 * 10.6	;progressione valori temporali numerici dei singoli periodi del quinto ciclo (text)
outvalue	"Ph6",	k6 * 8	;progressione valori temporali numerici dei singoli periodi del sesto ciclo (text)

Un altro problema tecnico da risolvere per l'esecuzione in *One-man performance* è stato quello di dare la possibilità all'esecutore di avere il tempo di potersi preparare dal momento in cui viene attivato lo *Start-Render* su MacCsound (un'attivazione consigliata in questo caso è senz'altro quella col comando da tastiera:  $\text{⌘R}$ ). Attraverso l'impiego dell'opcode *schedule* sono state gestite le chiamate degli strumenti proprio all'interno dell'orchestra ordinando a Csound di iniziare un pre-conteggio di sei secondi dopo il quale saranno avviate le azioni degli Assistenti e dell'intero algoritmo.

Nella progettazione iniziale dell'algoritmo si era deciso di non applicare l'automatismo delle *perforazioni*, anche se non sarebbe un problema attivarlo all'interno delle aperture degli *sliders* in maniera pseudo-casuale. In questo caso le *perforazioni*, segnalate nel *Form-Schema* con valori numerici all'interno dei *box* grafici, sarebbero effettuate dal I Assistente addetto al *Microphone* e dal II Assistente addetto al *Feedback*.

Allo stato attuale del progetto l'interprete ha la possibilità di provare il pezzo partendo: dal ciclo A, quindi dall'inizio, oppure dal ciclo C, o dal ciclo E, facendo uso di tre distinti file *.csd* dedicati allo studio.

Applicando un analogo procedimento per la progettazione di questo algoritmo (in file *csd.*), sono in preparazione altri cinque file per l'esecuzione delle altre *Versions* (dalla II alla VI).

## 1.2 L'INTERFACCIA GRAFICA

Sempre facendo uso di MacCsound, del californiano Matt Ingalls, l'interfaccia grafica (fig.5), concepita per questa versione di *Solo*, ha avuto come principale obiettivo quello di voler rendere ben visibili all'interprete, durante la *performance*, i seguenti elementi<sup>8</sup>:

- due **metronomi** (*outvalue*) realizzati con dei *meter* "pulsanti" che indicano:
  - il primo (*metronome\_1*), di colore rosso, la scansione dell'inizio dei singoli periodi (quindi l'inizio di ciascun pentagramma) ed
  - il secondo (*metronome\_2*), di colore bianco, la scansione di tutte le pulsazioni all'interno dei periodi (come vedremo il numero intero di pulsazioni da scandire all'interno di ogni periodo sarà deciso dall'interprete in maniera più conveniente possibile sulla base delle linee tratteggiate presenti su ogni pentagramma);
- sei *meter* orizzontali (*outvalue*), di vari colori, che visualizzano graficamente ed in maniera lineare, il **trascorrere fisico del tempo dei singoli periodi all'interno del rispettivo ciclo (A, B, C, D, E, F)** affiancati ognuno da due corrispondenti caselle di testo attive su *value display* per segnalare rispettivamente l'istante temporale del singolo periodo (*t*) e la successione numerica dei periodi (*n*) in relazione al ciclo.

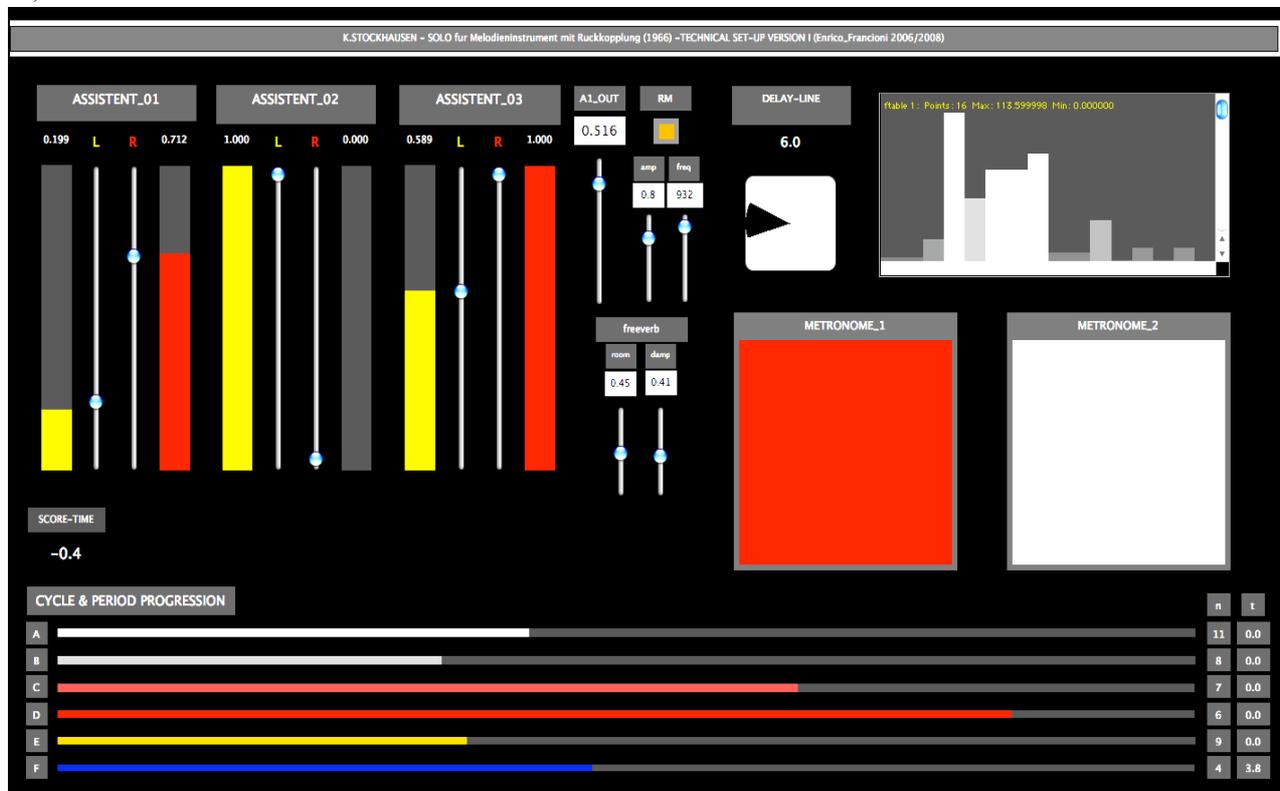


fig.5 – L'interfaccia grafica per Solo compilata su MacCsound 1.3b4

Nell'interfaccia (*i/o*) inoltre possiamo osservare:

- i movimenti degli **sliders** (*outvalue*) dei tre Assistenti (2+2+2) disposti in coppie stereofoniche – per un totale di sei sliders - , affiancati da sei coppie relative di *meter* (di

colore giallo per il L e di colore rosso per il R) che ne favoriscono la visualizzazione e dalle relative caselle di testo attive su *value display* dei rispettivi canali;

- uno **knob** (potenziometro) che indica (*outvalue*) il valore in

secondi della tap in uso dalla **delay-line** in un dato momento (nell'interfaccia abbiamo un unico commutatore in quanto le taps sono le stesse per entrambi i canali L/R). I valori in secondi di ritardo temporale delle taps, usati per la Version I, nell'ordine d'esecuzione sono: 6, 14.2, 19, 25.3, 10.6, 8;

- il livello del **segnale principale** (ain) in uscita gestito con uno slider (invalue) che agisce su entrambi i canali e che va a sommarsi al segnale ritardato; i valori in decremento, visibili nella casella di testo attiva su value-display, possono variare tra 0.0 e 0.6;
- uno slider (outvalue) che traccia il percorso programmato della frequenza della **RM** (Ring Modulation), affiancato da un altro slider (invalue) che ne regola la presenza e le relative caselle di testo attive su value display;
- due sliders relativi alla gestione del **riverbero**: per quanto riguarda la room lo slider (outvalue) traccia il profilo di valori programmato in orchestra in stretta relazione con la partitura musicale del solista, mentre per la damp i valori vengono passati in maniera attiva dall'interfaccia (invalue); anche qui i valori numerici sono indicati con relative caselle di testo attive su value display;
- un value-display (outvalue) che funge da **cronometro** (time), con pre-conteggio, dal momento dello Start Render;
- un **Inspector** riassuntivo di tutte le Funzioni presenti nello score (sco);
- un **Display** che ha la funzione di mostrare tutte le informazioni relative al setting, alla successione degli eventi e alla presenza di eventuali errori nella compilazione del codice.

## 2. PAGINE E TIMBRI

Seguendo le indicazioni dell'Interpretations-Schema, all'interno del Form-Schema Version I, ho quindi ri-composto e trascritto le sei pagine, con l'aiuto di un editor di notazione musicale<sup>9</sup> (si riporta come esempio la pagina relativa al Ciclo C - fig.6).

fig.6 – La pagina corrispondente al Ciclo C

La pagina è stata scritta tenendo presenti le seguenti istruzioni dell'Interpretations-Schema:

- **pagine:** confrontandola con la precedente (la B); usando più materiale da questa pagina (la C) che dalla precedente e confrontando le due pagine alternandole irregolarmente tra loro (per esempio, C B C C B C B C C, ecc.);
- **che cosa:** combinazione di parti e elementi (di B e di C);
- **in che modo:** parti ed elementi che sono più o meno gli stessi, o differenti l'uno dall'altro (combinazione); in base a: registro, pitch, dinamica, timbro, articolazione...;
- **uso di pause:** relativamente lunghe o relativamente corte (combinazione);
- **entrata di parti o elementi:** le entrate di parti ed elementi dovrebbero essere sincroni con le entrate del playback. Inoltre, pentagrammi e parti potrebbero essere scelti e sovrapposti in maniera tale che risultino più accordi possibili (Akkorde).

Le linee verticali tratteggiate indicano la scansione temporale e hanno oltre alla funzione di aiutare l'interprete nella fase di riscrittura del materiale, anche quella di dare un utile riferimento metronomico nella fase della performance.

Come indicato dall'autore ogni pagina corrisponde ad un ciclo (sei cicli per ogni Version) e ogni pentagramma corrisponde ad un periodo; all'interno del ciclo C, ad esempio, ogni periodo ha la durata di 19 secondi, che moltiplicati per i sette pentagrammi (o periodi) ci dà una durata totale del ciclo di 133 secondi.

Nel nostro caso l'effetto reale del contrabbasso è da considerarsi all'ottava inferiore trattandosi di uno strumento traspositore. L'accordatura dello strumento è *orchestra* (sol, re, la, mi), ma personalmente penso che un'accordatura *solista* (la, mi, si, fa#) potrebbe rendere molto più efficace e brillante l'esecuzione. Per poter eseguire con accordatura *solista* e rimanere fedeli alle indicazioni di altezza date dell'autore, potrebbe essere giusto allora trasporre un tono sotto tutta la notazione dell'intavolatura.

Frequentemente in questa versione sono stati utilizzati suoni armonici naturali o artificiali<sup>10</sup>: una pratica strumentale che a ragione rientra sia nello stile esecutivo di questo pezzo in particolare, ma anche, da un punto di vista filologico, nell'ambito della ricerca timbrica che si stava avviando in questo periodo storico. I suoni armonici inoltre sono stati introdotti per soddisfare le richieste di registro, per ragioni tecniche, oltre che per ottenere un maggiore interesse sulla materia suono.

Nel corso della performance incontriamo in partitura i simboli **N, I, II e III**: essi richiamano l'esecutore all'organizzazione ed all'uso di quattro diversi timbri. Quando si incontrano questi simboli all'inizio, o all'interno del pentagramma, il suono sarà sottoposto ad un'elaborazione che potrà essere di natura elettroacustica, o di tipo tecnico-strumentale.

In sintesi nel nostro caso specifico abbiamo inteso:

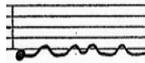
- con **N arco ordinario** condotto nella maniera convenzionale: col crine tra il tasto e il ponte;
- con **I** l'intervento di una **Rimodulazione** (Ring Modulation-RM) 'attiva', in quanto gestita in base ad alcuni centri frequenziali ricavati dall'osservazione della successione delle altezze (in corrispondenza alle zone sottoposte al timbro I); qui l'interprete continua ad usare la tecnica dell'arco ordinario,
- con **II arco condotto all'estremo ponte**;
- con **III arco col legno**; quindi arco condotto dalla parte della bacchetta tra il tasto e il ponte.

Per quanto riguarda invece le altre indicazioni timbriche si sono applicate le seguenti tecniche:

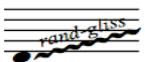
- **ETWAS GERAUSCHHAFT** (più o meno rumoroso): **arco multifonico** (verso il ponte cercando di sollecitare il 2°, il 3°, o il 4° armonico rispetto alla fondamentale);
- **GERAUSCHHAFT** (rumoroso): tecnica dell'**arco jeté e/o ricochet**;
- **SEHR GERAUSCHHAFT** (molto rumoroso): tecnica dell'**arco tremolo** insieme ad una **RM**.

E' evidente che la presenza simultanea di più indicazioni timbriche corrisponde automaticamente ad un più alto livello di elaborazione del suono.

La notazione del glissando randomico, indicato genericamente dall'autore con una linea del tipo:



è stata ridotta graficamente al simbolo:



### 3. VERSIONI STORICHE DI SOLO

Di seguito vengono riportati dei commenti ed alcune strategie adottate nella realizzazione del *technical set-up* di **Solo**.

Le versioni storiche prese in considerazione sono quelle per flauto, per sassofono, per trombone, per vibrafono, per contrabbasso e per sintetizzatore, anche se fino ad oggi se ne sono sperimentate e registrate numerose per svariati strumenti musicali.

#### 3.1 VERSIONE PER FLAUTO

Riporto ora in sintesi alcuni commenti tratti dalle note curate dallo stesso autore e allegate al CD della versione per flauto del 1995 – al flauto Dietmar Wiesner.

*“Da tempo degli strumentisti mi chiedevano un brano per solista e nastro (clarinetto, clarinetto basso, arpa, contrabbasso, trombone...). I solisti mi suggerivano un inventario di tecniche esecutive e le possibilità del loro strumento, ma io in realtà pensavo di comporre un pezzo polifonico in modo simile a *Gesang der Jünglinge* del 1955, anche se lì avevo usato la voce.*

*Il brano che cercavo di scrivere doveva prendere forma dal solista stesso, solista che sarebbe stato aiutato da alcuni Assistenti.*

*In questo tipo di esecuzione il solista deve essere cosciente dei cicli, dei periodi e degli elementi, cioè deve confrontarsi musicalmente con ciò che viene replicato attraverso la linea di ritardo, così come nella vita le cose ritornano simultaneamente.*

*Suonare in maniera critica sopra ciò che è presente, suonare in maniera complementare.*

*E' in questo senso che nell'esecuzione l'interprete è il testimone di un processo di multireazione, a più livelli.*

*Solamente quando la musica ci rende sensibili ad un pensiero a più dimensioni, nonché ad un processo di sovrapposizione della struttura, allora si potrà dire di essere giunti ad un livello superiore di composizione per il solista.*

*Quando il solista pensa ad una sua interpretazione si dovrà avvertire una relazione tra ciò che è e ciò che sarà, un'interazione a molti strati.*

*Nella musica d'oggi è molto forte il concetto di **giustapposizione**, ma io ho in mente un concetto **spazio/temporale**, all'interno del quale è possibile muoversi in tutte le direzioni, proprio come quando osserviamo una scultura (spazio/tempo).*

*L'accumulazione acustica e la verticalizzazione rende possibile questa esperienza.*

***Solo** può essere eseguito anche con più strumenti; l'esecutore può impiegare molte variazioni e intervenire in molti modi esecutivi. Quando l'esecutore deve dar forma a più blocchi, o strutture accordali, o concentrarsi su parti, o elementi, deve concentrarsi su ciò che si sente realmente.*

*Già nel 1966 a Tokyo ci si rese conto come fosse difficile realizzare questo brano; il brano dovrebbe essere elaborato, ...inventato lì per lì, anche se convenzionalmente tutti gli strumentisti si organizzano per tempo a stendere una versione con una sequenza di pagine e con una struttura del materiale. La sincronizzazione tra playback e registratore è controllata da un click nella cuffia del solista”<sup>11</sup>.*

#### 3.2 VERSIONE PER TROMBONE

Questa versione è stata elaborata da Benny Sluchin ed eseguita al trombone da Barry Anderson.

*“Alla fine del 1992, usando una versione di Max sul computer NEXT dell'IRCAM-SIM (con un processore per il suono) è stata preparata una versione, con la collaborazione di Cort Lippe, che è stata poi eseguita con successo in concerti pubblici nel febbraio e nel marzo del 1993.*

*Il programma in uso è stato scritto in MSP su un Macintosh Apple G3 ed è stato preparato all'IRCAM in collaborazione con Carl Harrison-Faia.*

*Nella nostra realizzazione, il computer produce i ritardi, esegue e coordina gli incarichi dei quattro Assistenti (come previsto dal compositore) e dà un aiuto all'interprete attraverso l'uso di un'interfaccia.*

*La finestra principale fornisce tutte le informazioni necessarie all'esecuzione del pezzo. C'è anche la possibilità di poter provare il pezzo iniziando da vari punti, ed anche a velocità diverse. Attraverso l'interfaccia è possibile un controllo dei diversi eventi registrati, la retroazione ed il playback, come pure la lunghezza del ritardo.*

*Un modulo speciale permette un meccanismo di perforazioni nell'ingresso, nella retroazione e nel playback.*

*Nella versione per trombone (Form-Schema Version V) il timbro I è stato realizzato con una sordina wa-wa, il II con una sordina ordinaria e il III attraverso un cappello. I cambiamenti veloci di timbri sono stati effettuati con l'aiuto di una sordina già in posizione.*

*Sono state scritte delle patch in MSP per soddisfare le indicazioni del compositore in modo da poter avere la maggior parte delle informazioni metronomiche [...]. Le trasformazioni in tempo reale sono effettuate liberamente, senza dover sottrarre così potenza di memoria e di elaborazione al computer.*

*L'interfaccia è stata disegnata per dare le informazioni necessarie allo strumentista: il tempo trascorso nel ciclo attuale, il numero del periodo, il numero del ciclo e l'azione di un metronomo quando l'interprete ne ha avuto bisogno. In questo modo l'interprete può essere preciso e suonare nel tempo collocato in ogni periodo.*

*La fedeltà e la precisione di questa versione al computer permette una prestazione che è molto vicina alle intenzioni originali del compositore.*

*E' anche possibile la realizzazione delle altre versioni (Form-Schema) con l'uso dello stesso programma”<sup>12</sup>.*

#### 3.3 VERSIONE PER SASSOFONO

La versione per sassofono (Form-Schema, Version I) era stata curata dal sassofonista Gianpaolo Antongirolami (per quanto concerne la parte relativa all'esecuzione) e da Eugenio Giordani (per la parte dell'installazione digitale di elaborazione in tempo reale); il lavoro è stato allestito in occasione del settantesimo compleanno dell'autore (1998).

La versione, realizzata avvalendosi di un sistema di sintesi digitale Kyma/Capibara (con processore Motorola 56002), cercava di ovviare ai problemi legati al sistema analogico originale come l'eccessivo accumulo di rumore di fondo del nastro, l'oscillazione del nastro che provoca *wow & flutter* e l'alto numero di Assistenti: ben quattro nella versione del 1966.

Nella versione ottenuta, per vari motivi legati all'utilizzo delle risorse del sistema impiegato (Kyma/Capybara), i realizzatori si sono occupati della messa in opera dell'automazione della parte relativa ai feedback multipli e agli involucri del segnale dei tre Assistenti (riducendo peraltro i quattro Assistenti alla sola regia del suono).

Antongirolami e Giordani con questa soluzione rinviano all'esecutore il compito di dover effettuare le variazioni timbriche previste dallo score (I, II, III), tuttavia gli stessi hanno anche realizzato delle *performance* dal vivo in cui avevano previsto delle variazioni timbriche attraverso l'uso di un Phase-Vocoder<sup>13</sup>.

### 3.4 VERSIONE PER VIBRAFONO

Questa versione è stata progettata ed eseguita al vibrafono da Robert Esler<sup>14</sup> nel 2006.

*“Tutte le modulazioni di timbro e di rumore erano basate intorno alle qualità armoniche del vibrafono programmato in Pd (Pure Data)<sup>15</sup>. Il vibrafono non ha un timbro intrinsecamente complicato ed è decisamente uno degli strumenti dal timbro più ‘puro’ a causa della sua struttura armonica. I timbri per questa versione di Solo erano progettati per aggiungere complessità al suono: con I si attiva una modulazione, con II e III invece si attivano due diversi Vocoder. Non era stato fatto alcun tentativo per preparare, o alterare fisicamente lo strumento (vibrafono), troppo scomodo nel contesto del pezzo e non opportuno all'interno del concetto di autonomia digitale; l'elaborazione era stata realizzata per via digitale. Tuttavia, era importante la parte del procedimento del progetto pur di non perdere l'identità dello strumento all'interno delle combinazioni timbriche. Ogni alterazione ha mirato ad ottenere una complessità che ci piacesse esteticamente, conservando questa neutralità del vibrafono. Un effetto in più che abbiamo incluso in questa versione è stato il glissando. Il vibrafono infatti non è in grado di eseguire i glissando, quindi il computer simula questa condotta. L'effetto è controllato dall'interprete impiegando una ‘modulation wheel’ su una tastiera MIDI.*

#### TRE DISTINZIONI TIMBRICHE

I. Ring Modulation (Modulazione ad anello) - modulando la frequenza in misura di 3.3 volte il segnale di entrata. Questo contribuisce a dare un colore sulle fasce di frequenza alte e basse delle componenti enarmoniche.

II. Vocoder A – usa un campione di clarinetto come impronta sul timbro. Questo campione è un chorus di clarinetti, e non uno solo per evitare la presenza di una nota di pedale quando si usa questo effetto. Quando si elabora il vibrafono, questo vocoder schiarisce con una miscela di parziali in aggiunta al suono dello strumento, oscurando leggermente l'identità dello stesso rispetto alla Ring Modulation.

III. Vocoder B – usa un gong come impronta sul timbro. Il gong satura lo spettro armonico e smorza in misura maggiore il complesso miscuglio.

#### GRADAZIONI RUMOROSE

**Rumoroso** (*Geräuschhaft*) – il rumore bianco filtrato viene moltiplicato al segnale. Questo effetto era stato scelto per creare una simulazione di un sottofondo di forte rumore. L'effetto assume significato se è impercettibile, ma non trasparente.

**Poco Rumoroso** (*Etwas geräuschhaft*) – sono state moltiplicate sul segnale una serie di onde 'sawtooth' (a dente di sega). L'intenzione era quella di creare un brusio che fosse più a fuoco del rumore bianco, come se il vibrafono in qualche modo fosse 'preparato'.

**Molto Rumoroso** (*Sehr geräuschhaft*) – un vocoder parziale. Essenzialmente, questo modulatore 'randomico' di rumore risintetizza il segnale attraverso un errore di calcolo. I risultati sono delle spesse bande di parziali enarmoniche che sostengono gli armonici del vibrafono.

*Questa tavolozza di rumori e di timbri funziona molto bene in combinazione al vibrafono. Siamo molto soddisfatti del fatto di aver creato textures diverse durante le stratificazioni (layers) del feedback-loop, tuttavia lo strumento conserva sempre un'identità. Dal punto di vista estetico si preferiva l'uso di un computer per creare alterazioni di timbro/rumore, c'era un maggiore controllo nel progetto, rispetto alle limitazioni che presentava invece l'elaborazione del suono per via acustica. Il procedimento si è trasformato in un esercizio di 'sound-design' per l'interprete, mettendo in gioco una grande libertà e un alto livello di creatività nel momento in cui vengono prodotti in tempo reale i moduli di DSP.*

*Gli Assistenti erano sincronizzati attraverso una 'master time' che esegue le loro azioni attraverso un file di testo (q-list). Il sistema era progettato per importare un messaggio di on/off da ciascun fader controllandone poi il feedback-loop. La parte del procedimento di interpretazione è costituita dalla programmazione degli Assistenti all'interno del Form-Schema. Poiché l'interprete presumibilmente non eseguirà il pezzo esattamente sempre in maniera identica, tra le performances c'è una certa quantità di imprevedibilità. Questo obbedisce alle istruzioni di Stockhausen in merito alla spontaneità delle azioni degli Assistenti: osserviamo infatti che molti dei periodi sono troppo brevi per chiunque voglia produrre un risultato altamente convincente. Il feedback-loop era progettato per simulare la configurazione originale dal 1966 che consisteva in un nastro molto lungo con sei testine (tap), una per ogni ciclo. Ogni testina era collocata alla distanza corrispondente al valore di ritardo di tempo di ogni periodo. La versione digitale di questo processo impiega sei delay-lines variabili [...]”<sup>16</sup>.*

### 3.5 ALTRE VERSIONI

Sempre nel booklet del CD di Stockhausen del 1995 leggiamo di altre versioni.

*“Nel 1994 mio figlio Simon elaborò una versione, la V, con un sequencer e un campionatore, quindi registrò il risultato su un sistema a sette tracce (multitraccia). Usò campioni di suono di flauto che vennero rielaborati attraverso le tecniche dello stretching, della trasposizione, del filtraggio e del reverse-playback...”*

*“In un'altra versione per sintetizzatore (Kürten, Stoccarda) mio figlio Simon caratterizzò i vari timbri in questo altro modo: N, con vocal-like sound, I con glasses, II con metallic e III con suoni di ottoni”<sup>17</sup>.*

\*

Il Maestro Eugenio Giordani mi riferisce dell'esistenza di una versione *analogica* in cui si faceva uso del nastro magnetico, eseguita negli anni ottanta al flauto, o anche al contrabbasso.

Mi ha anche confermato che la versione di *Solo* (per contrabbasso e nastro) che io stesso ascoltai a Perugia presso il Teatro Comunale nel 1987, era stata curata proprio dal gruppo del quale anch'egli faceva parte: l'Ensemble ElectraVox.

Relativamente alla parte tecnica, le performances erano gestite da Assistenti d'eccezione quali: Ceccarelli, Baggiani, Branchi, Mollia e lo stesso Giordani.

Le parti meccaniche dell'allestimento tecnico erano state progettate e realizzate ex-novo in modo pressoché artigianale. Il Maestro Giorgio Klauer scrive:

*“Per quanto riguarda il mio lavoro su Solo, a suo tempo ne avevo seguito l'interpretazione a cura del CRFMW di Liegi con un violoncellista, Arne Deforce. L'interpretazione era stata realizzata senza effetti. In seguito ho formulato uno schema relativo alla prima versione, per il flauto, aggiungendo alcune elaborazioni timbriche che sopperissero alla ridotta versatilità timbrica dello strumento: faceva parte del mio lavoro di tesi per il conservatorio; si trattava anche nel mio caso di una One-man performance...”*

*[...] ho prodotto il patch in Max/MSP 4.6. Nelle esecuzioni non ho incontrato difficoltà particolari, perché il pezzo è efficace, se lo si riesce a interpretare - penso in particolar modo alla valorizzazione delle risorse di prassi esecutiva che entrano in gioco assieme allo strumento con il quale lo si esegue.*

*[Nel 2008]...l'ho ripresa e con alcuni adattamenti è stata eseguita da una giovane violoncellista del conservatorio di*

*Como, con la regia di alcuni allievi della scuola di musica e tecnologie del suono, con esiti abbastanza soddisfacenti (...si tratta di un progetto didattico).*

*[...] la seconda parte di 'Wie di Zeit vergeht' e alcuni brani dall'articolo sulla prassi esecutiva della musica elettroacustica pubblicato in 'Perspectives of new music' [...] mi hanno aiutato abbastanza a chiarirmi le idee sull'interpretazione da dare (certo non sulle sonorità da ricercare)”<sup>18</sup>.*

\*

In queste ultime ore, mentre mi accingo a terminare il presente articolo, vengo a conoscenza di una nuova versione (2008) di Solo realizzata in multitraccia (ProTools) ed interpretata da Federico Bagnasco al contrabbasso. La Version utilizzata in questo caso è la III e sarà probabilmente destinata alla pubblicazione in CD per la rivista *Suonosonda*<sup>19</sup>. Ad un primo ascolto appare evidente la cura che l'interprete ed i tecnici hanno dedicato allo studio dei timbri e alla dimensione verticale del pezzo.

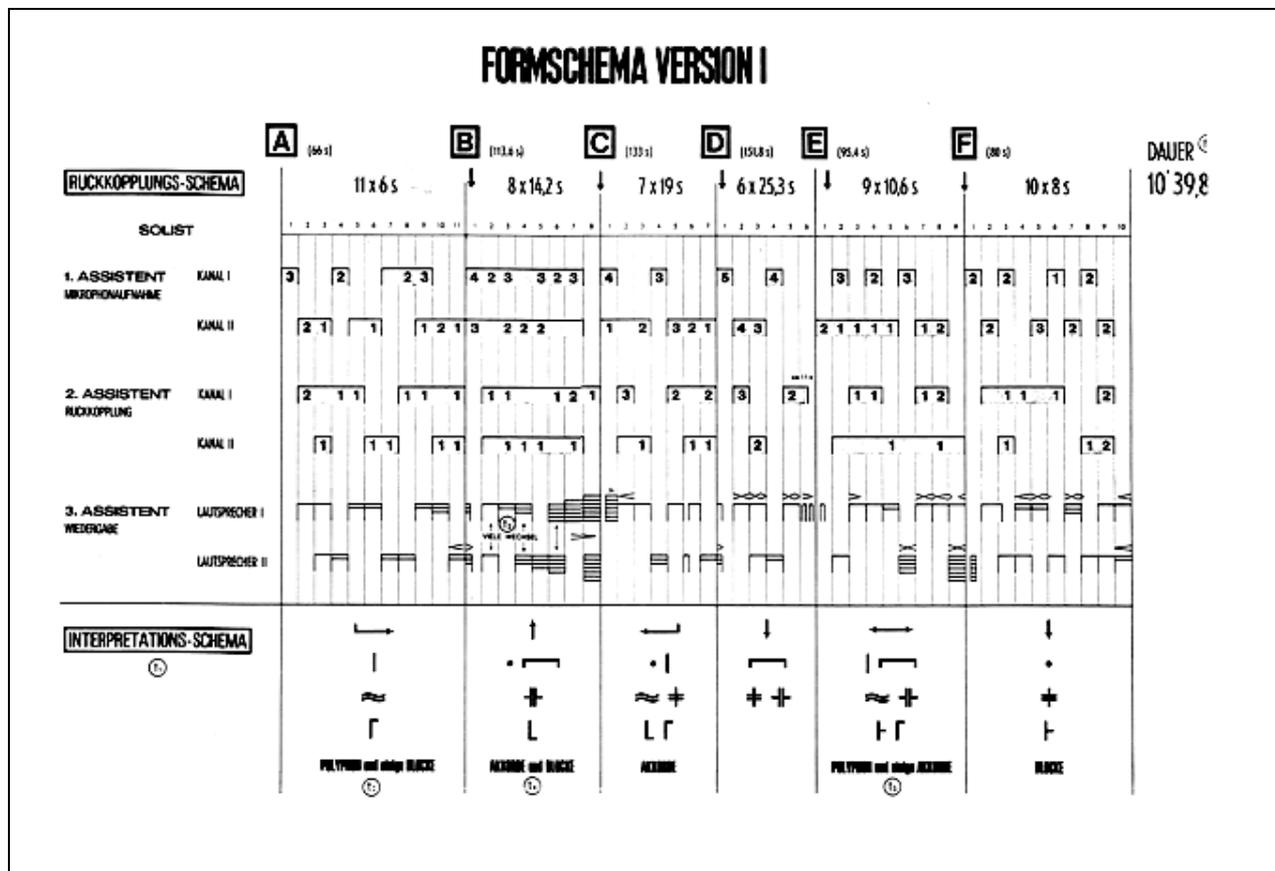


fig.7 – Il Form-Schema Version I

## NOTE E RIFERIMENTI

<sup>1</sup> Stockhausen, K. – “SOLO NR.19, FÜR MELODIEINSTRUMENT MIT RÜCKKOPPLUNG”, (Universal Edition-Wien) UE 14789 (1969).

<sup>2</sup> “Solo è un pezzo per strumento melodico e feedback che permette ad un singolo esecutore di fare musica polifonica. Scritto nel 1966 ed eseguito in prima assoluta a Tokyo nello stesso anno, la sua origine risale al periodo di Kontakte, in particolare all’invenzione della testina di copia, un riarrangiamento delle testine di riproduzione, registrazione e cancellazione del nastro, che permette al compositore di creare suoni complessi attraverso l’accumulazione. Solo è un’espressione unica dell’interpretazione tra influenze personali e meccaniche nella sintesi della musica elettronica. Solo è un pezzo di teatro intimista, l’esecutore rappresenta il compositore, il suo strumento, la voce dell’autore.

*L’artista non è più visto come un creatore, ma come un intermediario [...] e trasforma ciò che sente in qualcosa di nuovo”.*

Da Maconie, R. – “THE WORKS OF KARLHEINZ STOCKHAUSEN”, Marion Boyars, London 1976.

<sup>3</sup> Solo in questo allestimento tecnico è stato per ora eseguito nell’ambito de “Il Suono aperto” *Computer Music, Environmental Sound Art*, a cura del L.E.M.S. in collaborazione col Conservatorio Statale di Musica “G.Rossini” di Pesaro (Auditorium Pedrotti del Conservatorio Rossini - Pesaro, 18 giugno 2008) e del Festival “Villa e Castella 2008” (Villa Fiorini – Barchi PU, 6 luglio 2008).

Per un’ascolto di alcuni estratti dell’interpretazione: <http://www.myspace.com/enricofrancioni>.

<sup>4</sup> Nel nostro caso lo strumento solista è il contrabbasso.

<sup>5</sup> Il frontend impiegato per Solo è MacCsound v1.3b4 del 2006.08.07 - beta version (che è anche l’unica versione ufficialmente disponibile in rete fino ad oggi). Alla pagina internet destinata al download <http://www.csounds.com/matt/d+d.html> l’autore del programma Matt Ingalls dichiara spiritosamente: *but only minor changes still left to do - should be reliable!*

Questo frontend attualmente gira sul mio PowerBook G4 (processore: 1.67 Ghz PowerPC G4, memoria: 1.5 GB DDR2 SDRAM, Mac OS X 10.4.11) poggiando sull’ultima versione rilasciata di Csound [5.08].

La precedente versione di MacCsound, la 1.2a5 sembra manifestare, ma solamente per alcuni impieghi, una maggiore stabilità anche se gira con Csound Version 4.23f12 (9 Mar 2005) con tutti i limiti del caso.

<sup>6</sup> Per un download dei files si faccia sicuro riferimento alla pagina <http://www.csounds.com/downloads>.

<sup>7</sup> Normalmente si indicano tre Assistenti principali anche se nelle stesse note introduttive dell’autore, si parla dell’impiego di ben quattro Assistenti all’esecuzione.

<sup>8</sup> Personalmente ritengo che non sarebbe da scartare la possibilità di estendere anche all’ascoltatore la visione dell’interfaccia grafica durante la *performance* per comprendere meglio il senso dell’opera.

<sup>9</sup> Nel caso specifico si è fatto uso di Finale 2005 della “MakeMusic!”.

<sup>10</sup> Ciò si può osservare in un’altra partitura utile all’esecuzione in cui compare, oltre al pentagramma dove è notato l’effetto reale, anche un secondo pentagramma che riporta l’intavolatura sullo strumento.

<sup>11</sup> Note liberamente scelte e tradotte dall’inglese dal CD: “STOCKHAUSEN 45”, Solo – Spiral (Stockhausen-Verlag, Kurten-Germany, 1995).

<sup>12</sup> Note liberamente scelte e tradotte dall’inglese da: Sluchin, B. – “A COMPUTER-ASSISTED VERSION OF STOCKHAUSEN’S SOLO FOR A MELODY INSTRUMENT WITH FEEDBACK”, Institut de Recherche et Coordination Acoustique/Musique (IRCAM) 1 Place Igor Stravinsky Paris F-75004, France [sluchin@ircam.fr](mailto:sluchin@ircam.fr), in *Computer Music Journal*, 24:2, pp. 39–46, Summer 2000 © 2000 Massachusetts Institute of Technology.

<sup>13</sup> Note liberamente scelte da: Antongirolami, G./Giordani, E. – “K. STOCKHAUSEN – «SOLO» (PER SASSOFONO E

WORKSTATION KYMA/CAPYBARA – 66)”, in Atti del Convegno di “La terra fertile” - 1988, L’Aquila, Istituto Gramma pp. 14-17.

<sup>14</sup> Ringrazio Robert Esler per le informazioni gentilmente fornite sulla sua versione scritta in Pd. Per un ulteriore approfondimento tecnico sul programma, sull’algoritmo e per una verifica d’ascolto si faccia riferimento al sito internet: <http://www.robertesler.com/>

<sup>15</sup> Pd (Pure Data), molto simile per concezione a Max/MSP, è *free* ed è disponibile alla pagina: <http://circa.ucsd.edu/~msp/software.html>

<sup>16</sup> Note liberamente scelte e tradotte dall’inglese da: Esler, R. - DIGITAL AUTONOMY IN ELECTROACOUSTIC MUSIC PERFORMANCE: RE-FORGING STOCKHAUSEN University of California, San Diego; articolo presentato dallo stesso Esler in occasione della ICMC del 2006.

<sup>17</sup> Cfr. “STOCKHAUSEN 45”, Solo – Spiral (Stockhausen-Verlag, Kurten-Germany, 1995).

<sup>18</sup> Da un recente scambio di messaggi di posta elettronica col M° Giorgio Klauer che ringrazio pubblicamente per la disponibilità dimostrata (26 luglio e 5 agosto 2008).

<sup>19</sup> Sito internet: [www.suonosonda.org](http://www.suonosonda.org).



# Comporre (con) oggetti audio-visivi: un ambiente di lavoro in via di sviluppo

Alessio Santini  
Université Paris 8  
CICM - MSH Paris Nord  
alessio.santini@gmail.com

## SOMMARIO

In questo articolo mi propongo di illustrare la mia pratica di composizione audio-visiva, sviluppata in seno a una ricerca di dottorato ancora in corso, che mi trovo a svolgere presso il Centre de Recherche Informatique et Création Musicale, Université Paris8.

Dopo aver presentato alcuni tratti fondamentali dell'interazione audio/video, quali la visualizzazione del suono, la sonificazione di immagini e l'oggetto audio-visivo (inteso come oggetto informatico), proporrò alcuni esempi di convergenza compositiva : gesti creativi dove i due media vengono concepiti/creati/manipolati nello stesso momento.

Nella seconda parte illustrerò il mio metodo di composizione, gli strumenti software che utilizzo, e una pièce audio-visiva. Seguiranno alcune considerazioni conclusive.

## Parole Chiave

Max5 ; Jitter ; oggetto audio-visivo ; interazione ; modulare ; tempo reale

## 1. INTRODUZIONE

Lo sviluppo recente della capacità di calcolo dei computer ha introdotto la possibilità di trattare il video digitale in tempo reale. Questa capacità nuova è stata accompagnata dalla fioritura di applicazioni software che, con alcune differenze a seconda dei casi, permettono all'utente di lavorare in situazione di interazione inter-mediale : si pensi alla vasta gamma di programmi per *vjing*<sup>1</sup>, ed alle integrazioni a software già affermati come le librerie Gem per Pure Data<sup>2</sup>, o Jitter per Max/MSP, introdotta da Cycling<sup>74</sup><sup>3</sup> nel 2003. Da ricordare anche i software *Open Source* come Processing<sup>4</sup>, che stanno riscuotendo sempre maggior successo tra gli utenti, che dunque possono contare su comunità di supporto online molto attive.

Questi strumenti hanno dato origine, o per meglio dire nuovo vigore, all'interessamento di musicisti e creatori sonori verso il mezzo visivo, attraverso tecniche di interazione audio/video come la visualizzazione del suono e la sonificazione di immagini.

## 2. VISUALIZZAZIONE, SONIFICAZIONE, E OGGETTO AUDIO-VISIVO

Mi propongo di illustrare brevemente di seguito queste tre tecniche di interazione suono/immagine, citando il lavoro di alcuni ricercatori e artisti che più hanno influenzato la mia personale riflessione in proposito.

### 2.1. Visualizzazione del suono

La mia esperienza intorno alla questione è strettamente legata alle attività svolte dal gruppo di lavoro "Visualisation du Son"<sup>5</sup> dell'AFIM<sup>6</sup>. Dal 2004 al 2006, il gruppo ha condotto alcune ricerche in cui la tematica della visualizzazione del suono è stato il punto di partenza per approfondire, anche a livello teorico, le questioni della relazione suono-immagine, come vista dal compositore. Lo scopo era di integrare il visuale al musicale, per estendere l'espressività sonora fino al dominio del visivo [1].

È possibile individuare diversi tipi di visualizzazione : considerando le finalità dell'operazione si può parlare di visualizzazione « scientifica » (ad esempio un sonogramma), o « artistica », che è poi quella che osserveremo più da vicino in questo articolo.

Generalmente, dal punto di vista pratico, si considera visualizzazione del suono la generazione/manipolazione di immagini realizzata utilizzando valori estratti dall'analisi di un flusso audio ; la definizione è corretta, ma non esaustiva. Si possono citare ancora due casi :

- L'immagine viene controllata non dai valori estratti dall'analisi, bensì dagli stessi parametri che gestiscono il suono [2]. Per esempio : calcolare tramite analisi di un flusso audio, il valore del parametro *time* di un effetto di *delay* con un elevato valore di *feedback*, può risultare molto complicato ; una soluzione più semplice può essere quella di utilizzare direttamente il valore di detto parametro, magari dopo opportune operazioni di regolazione.

- L'immagine non ha relazioni dirette col materiale sonoro. Un esempio eccellente sono le visualizzazioni, opera di Brian O'Reilly, di alcune composizioni di Curtis Roads raccolte nel CD/DVD "Point Line Cloud"<sup>7</sup>, dove la visualizzazione non è necessariamente prodotta dal suono quanto, « ispirata » : l'artista si serve di video di repertorio e immagini che, filtrati da effetti e montati su molteplici livelli tramite operazioni di *compositing*<sup>8</sup>, integrano l'esperienza sonora, evidenziandola, sostenendola.

### 2.2. Sonificazione di immagini

Le operazioni di sonificazione dell'immagine, consistono in genere nella messa in atto del processo inverso rispetto alla visualizzazione del suono : la tecnica si basa sostanzialmente sull'analisi di matrici video.

Questo tipo di relazione è stata trattata in modo approfondito, tra gli altri, da Woon Seung Yeo e Jonathan Berger, delle cui

<sup>1</sup> si vedano ad esempio Arkaos vj (<http://www.arkaos.net>), Modul8 (<http://www.garagecube.com/modul8>), e VDMX (<http://www.vidvox.com>)

<sup>2</sup> <http://www.puredata.org>

<sup>3</sup> <http://www.cycling74.com>

<sup>4</sup> <http://www.processing.org>

<sup>5</sup> <http://gtv.mshparisnord.org>

<sup>6</sup> Association Française d'Informatique Musicale, <http://www.afim-asso.org>

<sup>7</sup> pubblicato da Asphodel, 2004, ASP3000

<sup>8</sup> <http://en.wikipedia.org/wiki/Compositing>

ricerche è disponibile ampia documentazione, specialmente riguardo le tecniche di sonificazione dette di *scanning e probing* [3,4].

In generale, differenti tipi di analisi di un flusso di immagini, offrono molteplici possibilità di utilizzare i dati estratti per produrre/manipolare il suono.

### 2.3. Il caso dell'oggetto audio-visivo

Una terza possibilità nell'interazione di suono e immagine, si riscontra nell'« oggetto audio-visivo ».

Questo termine, sempre più comunemente usato, implica dal mio punto di vista una riflessione da condurre su diversi livelli, ma in questa prima parte del testo mi limiterò a presentare un solo aspetto della questione, quello che successivamente definirò "produttore". Questa scelta è giustificata dal fatto che, anche nei due paragrafi precedenti, ho parlato dei tipi di interazione considerando esclusivamente le tecniche di produzione ; volendo per il momento mantenere questo angolo di osservazione, illustrerò qui l'oggetto audio-visivo inteso come strumento informatico : algoritmo atto a produrre suono e immagine sincronicamente, mediante comandi comuni.

Cyrille Henry, artista e ricercatore francese, considera l'oggetto audio-visivo proprio da questo punto di vista : un oggetto informatico che genera audio-visioni tramite un medesimo gesto performativo [5]. È ciò che Benoît Courribet, nel descrivere alcune tecniche di creazione audio-visiva, definisce "approccio parallelo" [6] : controllare l'audio e il video attraverso un unico set di parametri dinamici detti « intenzionali », in quanto pre-esistenti rispetto ai media che andranno a generare/manipolare.

I due media vengono gestiti da parametri comuni ; nel tipo di mappatura dei dati si esplicita la scelta, arbitraria, dell'artista.

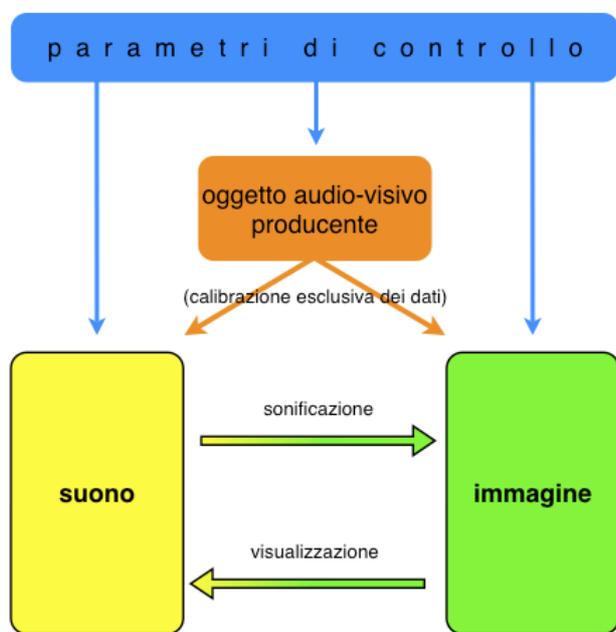


Figura 1. Classificazione dei rapporti interattivi

### 3. CASI DI CONVERGENZA NELLA COMPOSIZIONE DI SUONO E IMMAGINE

La possibilità di controllare un oggetto informatico audio-visivo in tempo reale, ha portato alla produzione di molti eventi e performance che impiegano i due media allo stesso tempo ; accenno di seguito a due casi diversi tra loro :

- CHDH<sup>9</sup> : Cyrille Henry e Nicolas Montgermont eseguono una performance audiovisiva, di natura minimale, a tratti industriale. Ogni algoritmo impiegato genera un suono sintetico e un'immagine tridimensionale, che evolvono in stretta connessione, controllati dallo stesso gesto. Nel corso della performance, i musicisti passano da una fase in cui gli oggetti sono utilizzati in solo (per permettere al pubblico di appropriarsi delle relazioni suono/immagine di ognuno), ad un momento di improvvisazione che coinvolge più oggetti alla volta. L'esibizione è caratterizzata da una forte componente ritmica, e la costruzione di poliritmie audio-visive realizzata nella seconda parte risulta di forte impatto.

- Klipp av<sup>10</sup> : il duo formato da Nicolas Collins e Fredrik Olofsson propone delle performance live in cui flussi audio e video (precedentemente archiviati o catturati in diretta con microfoni e telecamere) sono manipolati in maniera sincrona [7]. I due media sono tagliati e ripetuti secondo parametri comuni, che fanno capo ad una libreria di oggetti per la manipolazione ritmica del suono in stile *breakbeat*, sviluppata da Collins per Supercollider<sup>11</sup>.

### 4. UN APPROCCIO PERSONALE

Vengo adesso ad illustrare il mio percorso, gli strumenti informatici, e la metodologia che utilizzo nella composizione audio-visiva.

Lo scopo della mia ricerca è elaborare tecniche per la composizione di pièces audio-visive fisse, considerabili come cortometraggi. Per raggiungere lo scopo, è importante affrontare la questione dell'interazione suono/immagine da un duplice punto di vista :

- tecnico : programmazione di algoritmi di interazione e concepimento di un metodo di composizione soddisfacenti
- sintattico : elaborazione di una sintassi ibrida, che consideri i media impiegati secondo le loro proprietà specifiche, morfologiche e semantiche, e che le utilizzi come elemento compositivo

L'equilibrio sintattico è necessario perché, l'oggetto risultante, sia effettivamente audio-visivo, prodotto di un linguaggio ibrido. Si considerino ancora una volta le visualizzazioni di O'Reilly su Roads : pur essendo il risultato formalmente considerabile un cortometraggio, si tratta ancora di visualizzazione del suono, in quanto costruito su logiche esclusivamente musicali.

Sull'approfondimento dell'aspetto sintattico mi riprometto di tornare in futuro ; in questo articolo, sarà curata la descrizione del solo aspetto tecnico.

#### 4.1. Dalla granulazione del suono alla micro-frammentazione temporale audio-visiva

Il mio interesse attivo verso l'interazione suono-immagine, è nato in seguito allo sviluppo di alcune patch per la granulazione del suono, che ho scritto in Max/MSP negli ultimi due anni<sup>12</sup>. L'enorme potenziale espressivo di questa tecnica, mi ha stimolato a tentare di applicare un procedimento simile anche al dominio del video. Ho dunque realizzato una patch che operasse, sincronicamente, una micro-frammentazione

<sup>9</sup> sono disponibili alcuni esempi all'indirizzo <http://www.chdh.free.fr>

<sup>10</sup> sono disponibili alcuni esempi all'indirizzo <http://www.klippav.org>

<sup>11</sup> la libreria è disponibile in download all'indirizzo <http://www.cus.cam.ac.uk/~nc272/bbcut2.html>

<sup>12</sup> una patch per la granulazione del suono è disponibile in download all'indirizzo <http://www.cycling74.com/wiki/bin/view/Share/AlessioSantini>

temporale sui due file, un suono e un filmato, in lettura. All'inizio, il motore sonoro della patch produceva un solo grano alla volta, e il punto di lettura del file audio era sincronizzato col punto di lettura del file video : ad ogni grano sonoro corrispondeva un *frame* del video<sup>13</sup>.

I limiti dell'approccio sincronico sono subito risultati evidenti : pretendere di sincronizzare due flussi di dati così diversi tra loro evidenziava dei problemi di corrispondenza, specialmente in occasione di sequenze di grani di densità particolarmente alta<sup>14</sup>. Impostare dei valori di soglia che una volta superati interrompessero il rapporto di sincronia, e inviassero al video degli effetti per sostenere visivamente la percezione auditiva di un flusso di grani molto brevi, è stata una prima, efficace, soluzione [8].

#### 4.2. Triplice classificazione dell'oggetto audio-visivo

Prima di proseguire è necessario illustrare alcune considerazioni aggiuntive sul concetto di oggetto, poiché la sola definizione data fino a questo punto non è più sufficiente. Se da un lato, la definizione di oggetto informatico atto a produrre audio e video, trova riscontro nello stato delle cose, ed è condivisibile, è pur vero che il concetto di oggetto, storicamente, è denso di sfumature e significati diversi. Nella mia attività compositiva ne ho individuati tre : sorgente, produttore, risultante.

##### 4.2.1. Oggetto come sorgente

Seppur essendo la sintesi contemplata teoricamente nell'esposizione che segue, prediligo manipolare eventi già esistenti, audio e video, sia che questi siano sincronizzati tra loro in origine, o che siano inizialmente indipendenti l'uno dall'altro : suoni campionati e immagini, fisse e in movimento. La raccolta di oggetti-sorgente, è il primo passo da compiere, nel metodo compositivo che verrà illustrato. Questa operazione trova i suoi fondamenti teorici nel pensiero schaefferiano ; si pensi al concetto di oggetto pre-esistente, e di ascolto ridotto [9].

##### 4.2.2. Oggetto come produttore

Come precedentemente illustrato, l'oggetto produttore è lo strumento informatico : l'applicazione, il software che permette di elaborare in tempo reale i due media allo stesso tempo.

##### 4.2.3. Oggetto come risultante

Il risultato del gesto creativo, scritto, o eseguito dal vivo, è un prodotto, espressione di una scrittura ibrida, composto integrando elementi sonori e visivi : è la forma in cui vanno a fissarsi in maniera definitiva i risultati di tutte le relazioni interattive create dal compositore.

L'oggetto risultante, nella sua fissità (trattandosi di una forma assimilabile all'idea di cortometraggio), non è solo un punto di arrivo di un processo di composizione che si è appena concluso, ma pure il punto di partenza per tutti i processi che seguiranno. L'analisi di questo oggetto è utile per poter considerare l'efficacia delle relazioni audio/video che sono state usate durante la composizione, e per valutare la componente di « equilibrio sintattico » presente.

#### 4.3. Metodologia

Come emergerà, la pratica di composizione è largamente influenzata dalle tecniche di composizione a oggetti [10].

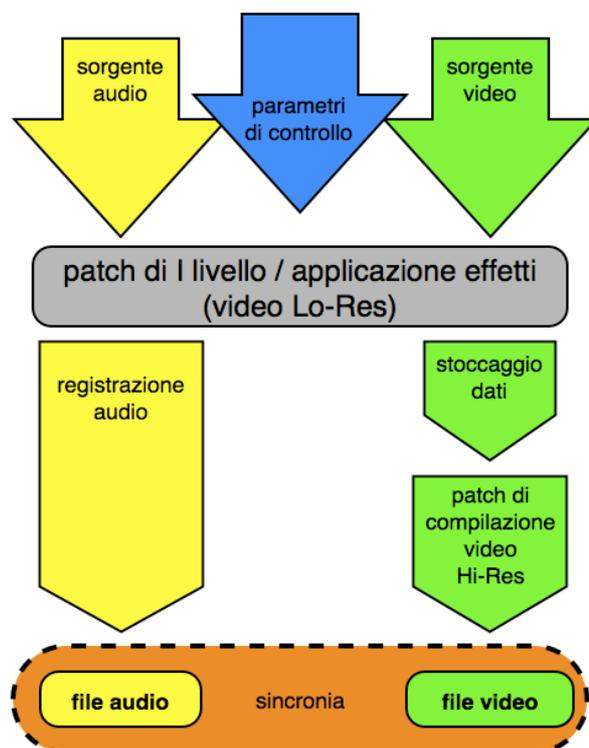


Figura 2. Dall'oggetto-sorgente all'unità audio-visiva

##### 4.3.1. Archivio

Il primo passaggio consiste nel costituire un archivio di oggetti-sorgente. Essendo in questa sede mia intenzione esporre solo l'aspetto tecnico della composizione, non mi soffermerò su considerazioni concernenti il valore semantico dell'oggetto pre-esistente, che pure è parte importante della mia riflessione intorno a queste questioni [8].

##### 4.3.2. Unità audio-visive

Tutte le operazioni che seguono sono svolte da patch scritte nell'ambiente di programmazione Max5.

Una prima patch opera un'azione di « delinearizzazione » sull'asse temporale delle due componenti (audio e video) secondo i valori di parametri comuni, e rende un nuovo file audio-visivo, sincronizzato, che definisco unità audio-visiva di I livello. Attualmente la generazione di questo file avviene attraverso pratiche di micro-frammentazione temporale di cui sopra, ma non è escluso lo sviluppo di altre tecniche. Ad esempio, una patch di sintesi sonora con visualizzazione del suono porterebbe allo stesso risultato : una componente audio e una video, sincronizzate nel loro sviluppo temporale.

Le modalità di scrittura delle unità di I livello sono diverse : spesso i valori sono inviati ai parametri della patch da un sequencer, ancora sviluppato in Max5. Una seconda opzione è data dall'improvvisazione, realizzata con un guanto su cui sono applicati alcuni sensori, o, più recentemente, con delle interfacce che sto sperimentando su un controller multitouch<sup>15</sup>. Qualunque sia la metodologia scelta, i dati vengono stoccati in un oggetto *coll*<sup>16</sup> ; è infatti disponibile una patch di scrittura e una di lettura, per poter intervenire sui risultati delle improvvisazioni.

<sup>13</sup> dei piccoli clip realizzati con questo primo software sono disponibili, a titolo d'esempio, all'indirizzo <http://www.alessiosantini.eu/pags/media.htm>

<sup>14</sup> considerando ad esempio un framerate di 25fps la sincronia viene neutralizzata in occasione di flussi di grani più brevi di 40 millisecondi

<sup>15</sup> [http://www.jazzmutant.com/lemur\\_overview.php](http://www.jazzmutant.com/lemur_overview.php)

<sup>16</sup> <http://www.cycling74.com/docs/max5/refpages/max-ref/coll.html>

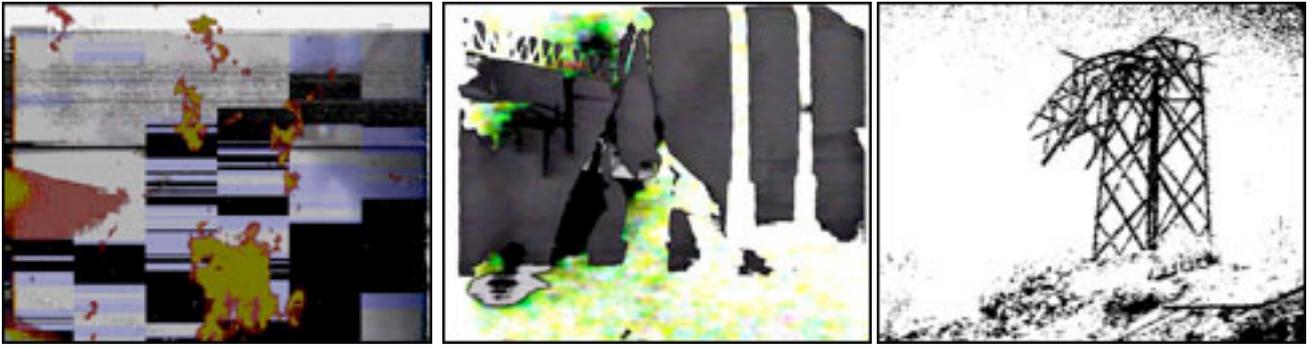


Figura 3. Montaggio di unità audio-visive attraverso tecniche di compositing : immagini estratte da Persistenza del Ferro

Una volta ottenuto il primo file, è possibile processare di nuovo le componenti audio e video attraverso delle patch che applicano effetti ai due media impiegando tecniche di visualizzazione, sonorizzazione, e ancora, variando parametri comuni.

Ogni volta che una unità passa attraverso uno di questi trattamenti audio-visivi, è definita per aumento di livello ; questa terminologia è impiegata solo per comodità.

Le patch per la produzione di unità audio-visive di livello superiore, nella maggior parte dei casi prevedono la lettura sincronizzata dei due files ; questa avviene attraverso gli oggetti Max5 *groove~*<sup>17</sup> e *jit.qt.movie*<sup>18</sup>.

Questo processo di creazione, per la ripetizione dei passaggi, risulta indubbiamente macchinoso, ma permette di seguire lo sviluppo dell'unità effetto dopo effetto, e di agire e correggere il risultato ottenuto ad ogni passaggio.

#### 4.3.3. Compilazione della componente video

In ognuna di queste fasi, è compreso un passaggio aggiunto per la compilazione in tempo differito del file video. Questo accorgimento si è reso necessario per ottenere un risultato di una qualità che, in regime di tempo reale, sarebbe compromessa ; soprattutto mi riferisco al *framerate*, che si abbassa sensibilmente se si domanda alla macchina di eseguire le operazioni di trattamento e contemporaneamente di renderizzare il risultato in un nuovo file.

Durante la fase di improvvisazione/scrittura della frase, il video viene trattato in bassa risoluzione, e i dati vengono archiviati in una lista tramite *coll*. In un secondo momento la lista viene poi riaperta in un'altra patch, che la rilegge applicando i medesimi parametri al video, ma in alta risoluzione, *frame by frame*.

La componente audio, che non comporta problemi di questo tipo, viene registrata già durante la prima fase. Le due componenti vengono in seguito riunite in un unico file tramite un qualsiasi programma di editing video.

#### 4.3.4. Montaggio

Dopo aver compilato una quantità soddisfacente di unità audiovisive, può iniziare la fase di montaggio. Questa è realizzata con un comune software per il montaggio video, e segue delle logiche di costruzione, ancora una volta, riconducibili a quelle della composizione ad oggetti : le unità audio-visive sono disposte sulla *timeline*, tagliate, ripetute, spostate, etc.

Come è evidente, il montaggio deve essere condotto con tecniche strettamente connesse alla natura di ogni medium : per le componenti sonore si impiegherà il missaggio, per quelle visive il *compositing*.

In questa fase, è possibile sciogliere alcune relazioni generate dai rapporti di interazione intermediale, istituiti durante la composizione delle singole unità audiovisive. Questo non deve

stupire : l'autore si trova a miscelare una grande quantità di segni, molto diversi tra loro. Bisogna considerare, ancora una volta, le differenze, anche di tipo morfologico, con cui si sta lavorando [2,6] : se da un lato, trovarsi a mixare otto tracce audio fa parte della comune pratica compositiva musicale, inserire altrettanti flussi video sullo spazio bidimensionale di destinazione, filtrando i livelli per colore, o luminosità, o ancora distribuendoli in diversi punti di detto spazio, può risultare molto arduo, e spesso anche controproducente. Spesso ci si trova a sopprimere una traccia video e, all'occorrenza, a riprocessare il reso visuale su cui sta lavorando (il « master video ») con un effetto globale i cui parametri sono in rapporto di interazione con la componente audio rimasta « orfana » del suo video.

Oltre a questioni di carattere tecnico, si può considerare che la scelta, arbitraria, è un tratto saliente della composizione e della produzione artistica in genere : giunto a questo passaggio, l'autore ha il diritto-dovere, in base alla sua sensibilità e al suo linguaggio, di rompere quelle relazioni che lui stesso aveva istituito, e tradire il rigore formale del particolare, per perseguire il suo obiettivo artistico, e sviluppare al meglio la sua poetica [11].

#### 4.3.5. Esempio di un oggetto audio-visivo risultante : Persistenza del Ferro

"Persistenza del Ferro" è una pièce audio-visiva di 4 minuti composta utilizzando gli strumenti e le tecniche descritti in questo articolo<sup>19</sup>. L'archivio di oggetti sorgente è costituito da dei campioni sonori registrati in diverse occasioni, e da da video e foto che testimoniano lo stato, presente e passato, delle miniere di ferro dell'Isola d'Elba. L'obiettivo era di trattare la tematica delle difficili condizioni di lavoro dei minatori nella prima metà del '900.

Si può considerare questo risultato come una sorta di cortometraggio, non narrativo, dove elementi intelligibili emergono sporadicamente da un flusso audio-visivo in gran parte astratto.

La composizione è stata molto influenzata dalle tecniche di trattamento e montaggio della musica granulare [12,13]; questo emerge, oltre che dal risultato sonoro, dalla struttura generale della pièce. Come oggetto risultante, presenta ancora molti limiti, soprattutto dal punto di vista sintattico, come ho evidenziato in altra sede [8].

## 5. UN AMBIENTE DI LAVORO IN VIA DI SVILUPPO

Vediamo ora più nel dettaglio la strumentazione software utilizzata per la produzione di unità audio-visive.

Il sistema che vado ad illustrare è stato scritto su Max/MSP/Jitter, e portato su Max5 subito dopo la sua uscita.

In un primo momento, ogni oggetto audio-visivo produttore era una patch chiusa e indipendente, con possibilità interattive

<sup>17</sup> <http://www.cycling74.com/docs/max5/refpages/msp-ref/groove~.html>

<sup>18</sup> <http://www.cycling74.com/docs/max5/refpages/jit-ref/jit.qt.movie.html>

<sup>19</sup> la pièce è visibile online all'indirizzo <http://www.alessiosantini.eu/pags/media.htm>

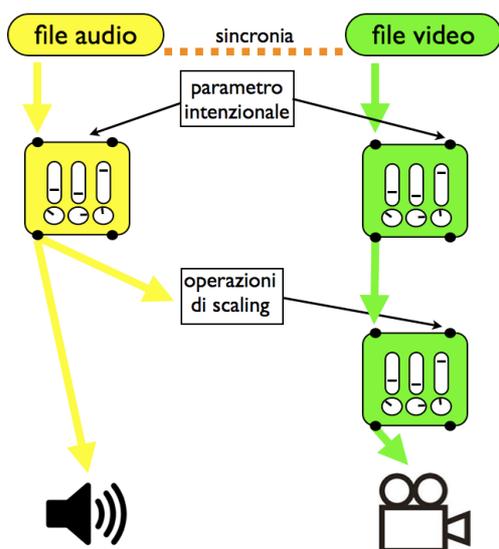


Figura 4. Esempio di relazioni audio/video nel sistema modulare

predefinite ; col passare del tempo, si è optato per lo sviluppo di un sistema modulare, più dinamico e flessibile, che permettesse di realizzare e testare velocemente diverse possibilità di interazione, secondo le esigenze del singolo caso.

Recentemente, tutta la parte del sistema che riguarda il trattamento video, è stata riadattata per operare in Open GL<sup>20</sup>. Oltre che migliorare l'efficienza della patch, alleggerendo il lavoro della CPU, questo passaggio introduce anche possibilità aggiunte per la disposizione nello spazio delle unità audio-visive, visto che in questo modo risultano « proiettate » su dei livelli inseriti in un ambiente 3D. Va aggiunto comunque, che per ragioni di ordine estetico, essendo ancora molto legato alla bidimensionalità degli oggetti sorgente che sono solito utilizzare, attualmente non sento la necessità di avvalermi di questa possibilità durante la composizione, o almeno non in modo evidente.

### 5.1. Un sistema modulare

Per la creazione di un ambiente di lavoro modulare, le patch per l'applicazione di effetti audio-visivi sopracitate sono state convertite in moduli. Ogni modulo viene caricato nella patch dentro un *bpatcher*<sup>21</sup>, e comunica con gli altri moduli tramite ingressi e uscite che veicolano tutti i tipi di dati coinvolti : il flusso audio o video, e i valori numerici di tutti i parametri del modulo. I dati coinvolti nell'interazione sono all'occorrenza riadattati per scale di valori diverse, a discrezione dell'utente.

Si possono individuare diverse classi di moduli, in base alla funzione svolta :

- moduli di I livello : impiegati per la creazione di un'unità audio-visiva di I livello, agiscono in modo sincrono sugli oggetti sorgente, oppure generando i media. Svolgono funzioni di generazione/sintesi, lettura files, manipolazione sincrona sull'asse temporale
- moduli effetto : applicano effetti, in maniera esclusiva, su un medium. La rete interattiva tra audio e video si compone di moduli di questa classe connessi tra loro. In genere si tratta di effetti molto semplici : la connessione di più funzioni permette all'occorrenza di creare reti complesse e di lavorare in modo più preciso sul trattamento dei media
- moduli di sistema : sono i controlli generali della patch (motore on/off, DSP on/off, stoccaggio dati per il richiamo dei

presets, etc)

- moduli di uscita : audio out, video out, finestra video, registrazione, etc

Se consideriamo ora l'interfaccia utente, risulta evidente che il sistema modulare di cui mi servo non presenta particolari qualità : ogni modulo ha tutti i parametri bene in vista, controllabili all'occorrenza con il mouse, ma non si può dire che dietro la costruzione della patch vi sia un pensiero ergonomico. Allo stato attuale della ricerca non rientra tra le priorità : si tratta di uno strumento pensato per l'attività in studio, il metodo compositivo è talmente frazionato, e il lavoro avanza così lentamente, che già disporre di tutti i parametri di cui si ha bisogno, bene in evidenza, risulta più che sufficiente.

La necessità di pensare un'interfaccia ergonomica diviene prioritaria se invece si vuole utilizzare questo sistema in occasione di performance dal vivo. Le combinazioni modulari dovranno allora essere cristallizzate in patch ben definite, con una accurata selezione di quei controlli che dovranno essere visibili ; non affronterò in questa sede la questione, ma sicuramente una delle novità di Max5 rispetto alla versione precedente, ovvero la « modalità presentazione<sup>22</sup> », offre diversi vantaggi in questo senso : è adesso possibile pensare la patch, già al momento della sua programmazione, sia dal punto di vista della funzionalità, che da quello dell'interfaccia utente. Questo porterà ad interessanti risultati, oltre che rendere le cose più facili durante la programmazione della patch.

## 6. CONCLUSIONI

In questo articolo, mi sono curato di esporre la parte tecnica della ricerca che sto conducendo : la metodologia, e lo strumento di composizione.

L'utilizzo di uno strumento informatico organizzato secondo dinamiche modulari, permette di poter integrare continuamente il proprio linguaggio senza per questo dover intervenire sulla struttura generale dello strumento stesso, né sul metodo ; è infatti agevole poter inserire un nuovo modulo nella libreria tutte le volte che si desidera, o modificarlo per rendere più efficiente l'algoritmo di cui si compone.

La struttura modulare permette inoltre di elaborare soluzioni specifiche, strategie locali, adattabili a esigenze sempre diverse ; questo è importante, in quanto ogni composizione è un episodio unico, che richiede attenzioni particolari [2]. La, seppur parziale, fissità di alcuni software che propongono molte funzionalità in partenza, rende più ardua questa ricerca di specificità.

Le prossime tappe di questa ricerca saranno :

- lo sviluppo di nuovi moduli da aggiungere alla libreria già esistente
- la composizione di nuove pièces e studi audio-visivi, da analizzare, e proporre in pubblico
- l'adattamento di tutto il sistema per operare in ambito di performance.

## RIFERIMENTI

1. Anne Sedes "Groupe de Travail Visualisation du Son", Journées d'Informatique Musicale, Paris, 2005
2. Benoît Courribet, "Réflexions sur les Relations Musique/ Vidéo et Stratégies de Mapping pour Max/MSP/Jitter", Journées d'Informatique Musicale, Paris, 2005
3. Woon Seung Yeo and Jonathan Berger. "Application of Image Sonification Methods to Music", ICMC, Barcelona, 2005

<sup>20</sup> <http://en.wikipedia.org/wiki/OpenGL>

<sup>21</sup> <http://www.cycling74.com/docs/max5/refpages/max-ref/bpatcher.html>

<sup>22</sup> <http://www.cycling74.com/products/maxuppatching>

4. Woon Seung Yeo and Jonathan Berger. "Raster Scanning A New Approach to Image Sonification, Sound Visualization, Sound Analysis And Synthesis ", ICMC, 2006

5. Ali Momeni and Cyrille Henry. "Dynamic Independent Mapping Layers for Concurrent Control of Audio and Video Synthesis", Computer Music Journal, Vol 30 N 1, The MIT Press, 2006

6. Benoît Courribet. "Compositing audio-visual art : the issue of time and the concept of transduction", actes des SMC 2006, Marseille

7. Nicolas Collins and Fredrik Oloffson, "Klippav : Live Algorithmic Splicing and Audiovisual Event Capture", Computer Music Journal, Vol 30 N 2, The MIT Press, 2006

8. Alessio Santini. "Prémises pour une étude sur l'objet logiciel audiovisuel, considérations et premiers résultats", Journées d'Informatique Musicale, Albi, 2008

9. Pierre Schaeffer, *Traité des objets musicaux*, Seuil, Paris, 1966

10. Osvaldo Budon. "Composing with Objects, Networks, and Time Scales. An Interview with Horacio Vaggione", Computer Music Journal, Vol 24 N 3, The MIT Press, 2000

11. Horacio Vaggione. "Some Ontological Remarks about Music Composition Processes", Computer Music Journal, Vol 25 N 1, The MIT Press, 2001

12. Curtis Roads, *Microsound*, MIT Press, Camb, MA, 2001

13. Caires, C., "Vers une écriture des processus de micromontage", Journées d'Informatique Musicale, 2003

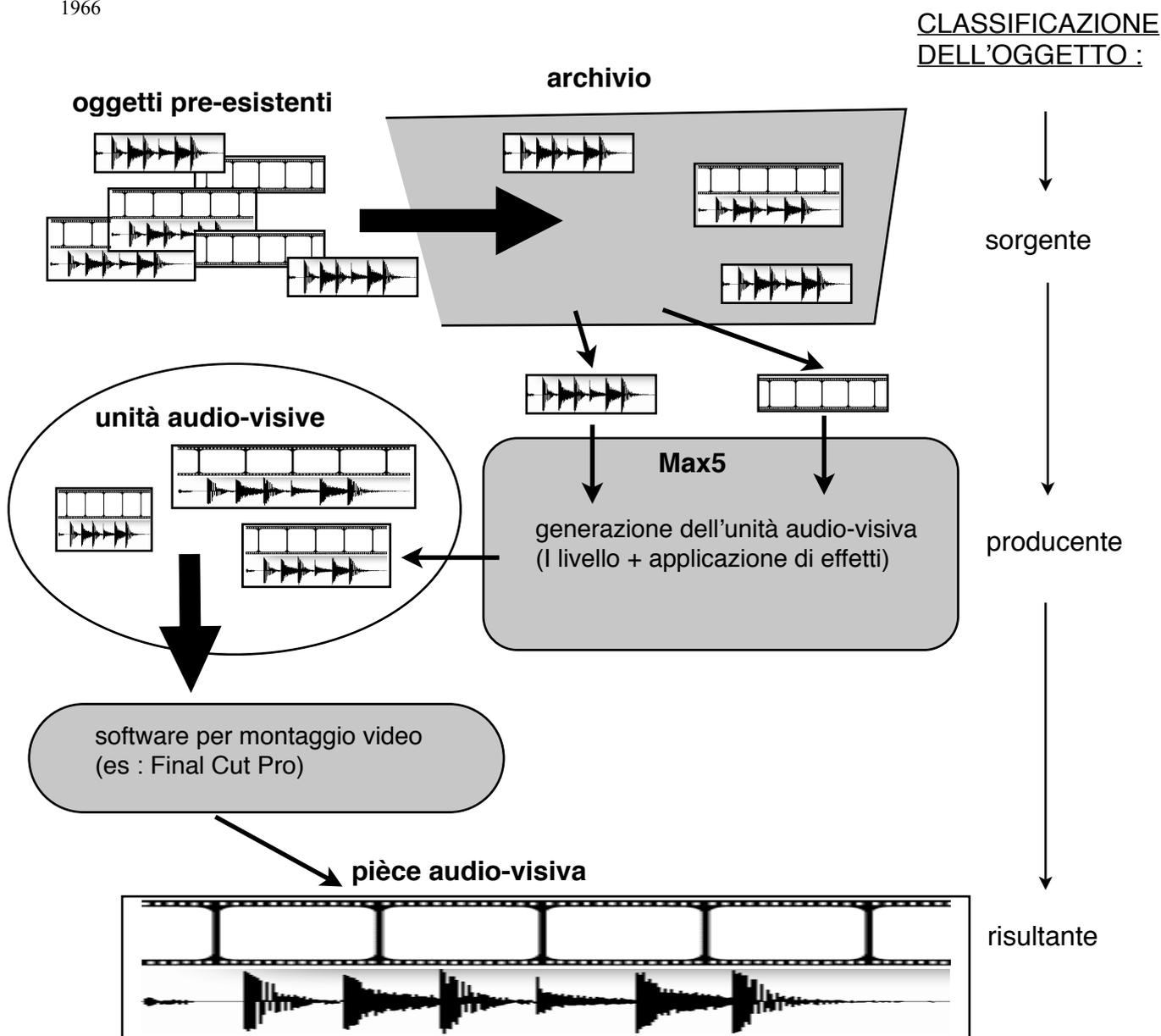


Figura 5. Schema riassuntivo del metodo di composizione

# Generazione di suoni e musica mediante sistemi caotici

Costantino Rizzuti  
Dipartimento di Linguistica  
Università della Calabria

## SOMMARIO

Negli ultimi anni si è verificato un crescente interesse a impiegare in campo musicale conoscenze provenienti da diversi settori della ricerca scientifica. Sono infatti ormai numerose le applicazioni di sistemi provenienti dalle ricerche in settori quali la complessità, il caos, l'intelligenza e la vita artificiale per la generazione di suoni e musiche.

Il presente lavoro illustra alcuni dei risultati ottenuti nella sperimentazione volta a impiegare sistemi dinamici complessi e caotici per la generazione di suoni e musiche.

L'attività di ricerca è spinta principalmente da un duplice intento: il primo, prevalentemente scientifico, consiste nel definire nuovi strumenti per lo studio e la comprensione dei sistemi complessi attraverso metodi non convenzionali; il secondo, più artistico, è finalizzato alla creazione di nuovi strumenti musicali.

## Parole Chiave

Complessità; Caos; Sintesi del Suono; Musica Generativa; Sound and Music Computing.

## INTRODUZIONE

A partire dagli anni '80 lo sviluppo di nuove discipline scientifiche volte all'investigazione dei fenomeni caotici e allo studio della geometria frattale hanno fortemente influenzato anche la ricerca in campo musicale. Nel corso degli anni sono stati numerosi i compositori che si sono dedicati a investigare le possibilità di impiego di sistemi dinamici non lineari per la generazione di suoni e musiche [1-4].

Da un differente punto di vista, diversi ricercatori si sono dedicati alla creazione di rappresentazioni uditive di insiemi di dati vasti e caratterizzati da un alto grado di complessità. Witten [5], ad esempio, propone l'impiego di metodi di rappresentazione uditiva poiché sostiene che molto spesso le serie di dati impiegate in campo scientifico sono vaste ed estremamente complesse e ciò ne rende difficile la comprensione tramite la semplice visualizzazione.

D'altra parte Pressing [1] avanza l'ipotesi che sfruttando le proprietà caratteristiche dell'udito probabilmente le rappresentazioni uditive rendono possibile seguire e distinguere l'evoluzione di dati caratterizzati da più di tre dimensioni in maniera più facile che tramite l'impiego della visualizzazione tridimensionale.

Il presente lavoro illustra le attività di ricerca svolte finora nell'ambito del corso di Dottorato di Ricerca in "Psicologia della Programmazione e Intelligenza Artificiale" presso l'Università della Calabria. Tale attività è principalmente volta a sperimentare l'impiego di sistemi dinamici complessi

e caotici per la generazione di suoni e musiche. La ricerca è spinta da un duplice obiettivo:

- definire nuovi strumenti per lo studio e la comprensione dei sistemi complessi attraverso metodi non convenzionali;
- sviluppare nuovi strumenti e nuove tecniche che possano contribuire nell'opera di musicisti e compositori volti a esplorare nuovi territori sonori e musicali.

Circa il primo obiettivo, si cerca di definire nuovi strumenti d'analisi che possano essere impiegati in maniera complementare alla visualizzazione 3D tramite la definizione di metodi di rappresentazione uditiva. La rappresentazione acustica di serie di dati sembra essere particolarmente adatta per il riconoscimento di strutture e nell'analisi della ricorrenza di andamenti caratteristici. Lo sviluppo di tali tecniche può costituire un importante strumento di supporto allo studio e all'interpretazione dei sistemi complessi, poiché, tramite l'ascolto, probabilmente, si possono cogliere informazioni qualitative sull'evoluzione temporale di serie complesse più di quanto sia possibile fare tramite la visualizzazione.

Circa il secondo obiettivo l'attività è volta a sperimentare l'impiego artistico dei sistemi dinamici provenienti dalle ricerche sulla complessità e il caos. A tal fine la ricerca è volta a definire delle metodologie basate su tali sistemi per la generazione di suoni, di strutture melodiche e di interi brani musicali. Inoltre, si cerca di creare "strumenti musicali" innovativi che possano essere utilizzati per la produzione di brani musicali, per la realizzazione di performance live o di installazioni audio/video. Tra le problematiche principali in questa direzione spicca la necessità di definire dei sistemi di interazione a più livelli fra musicista e strumento musicale. I sistemi d'interazione devono essere da una parte sufficientemente semplici da poter essere gestiti in tempo reale, mentre dall'altra devono essere capaci di fornire un controllo adeguato dell'espressività del materiale musicale prodotto. In questa direzione è importante, quindi, riuscire a dotare gli strumenti musicali di opportuni sistemi di analisi e sintesi che possano renderli realmente interattivi, ossia capaci di rispondere alle informazioni provenienti dai musicisti o da altre fonti esterne di informazione.

La definizione di codici capaci di trasformare le serie temporali, ossia le sequenze di dati relative all'evoluzione nel tempo delle variabili che caratterizzato univocamente lo stato di un dato sistema dinamico, nella variazione di opportuni parametri sonori/musicali è un problema che determina il conseguimento di entrambi gli obiettivi. Nel seguito si il-

lustreranno alcuni dei sistemi di codifica e si presenteranno le applicazioni finora sviluppate.

Nella prossima sezione si introdurranno le tecniche di generazione di suoni e musica mediante sistemi dinamici. Nella terza sezione si presenteranno alcuni sistemi di codifica per la generazione di melodie. Nella quarta sezione si presenteranno due strumenti musicali che impiegano dei sistemi dinamici per la generazione di suoni di sintesi. Nella quinta sezione si presenterà l'uso di ambienti immersivi 3D audio/video per l'esplorazione dei sistemi complessi. Con la sesta sezione si conclude il lavoro fornendo delle considerazioni finali sullo stato della ricerca e sulle direzioni che si intende seguire in futuro.

## GENERAZIONE DI SUONI E MUSICA MEDIANTE SISTEMI DINAMICI

La generazione di suoni e musica mediante sistemi dinamici può essere compiuta seguendo diversi approcci. La scelta del metodo da impiegare dipende sia dal tipo di sistema dinamico che si intende utilizzare sia dal tipo di evento da produrre (note, parametri di controllo, campioni audio).

Per quanto riguarda la tipologia del sistema dinamico è possibile far ricorso a sistemi provenienti dalle ricerche sulla complessità e il caos [1–6], oppure sistemi utilizzati nelle ricerche sulla Vita Artificiale (Artificial Life), ad esempio gli Automi Cellulari (CA) [7–9], inoltre, possono essere impiegati anche metodi evolutivi tramite l'impiego degli Algoritmi Genetici (GA) [10, 11].

Diversi metodi sono stati proposti per trasformare serie temporali in sequenze di note musicali [1, 2, 5, 6]. In letteratura, la generazione di melodie a partire dall'evoluzione di sistemi dinamici sembra essere l'approccio maggiormente utilizzato; da un punto di vista musicale questo può essere pensato come una operazione di alto livello poiché trasforma le serie temporali in note che, rispetto ai campioni audio o ai parametri di sintesi, sono eventi di alto livello. Generalmente questi sistemi di codifica producono dei file MIDI che possono essere trasformati in suoni tramite un qualsiasi dispositivo General MIDI.

Un altro approccio consiste nel realizzare la codifica delle serie temporali a un livello inferiore rispetto al precedente [12]. In questo caso infatti le serie temporali vengono trasformate nell'evoluzione dei parametri di controllo di un sistema di sintesi sonora. In tal modo è possibile trasformare l'evoluzione di un sistema dinamico nella variazione di alcune caratteristiche di un suono di sintesi. Il vantaggio offerto da questo approccio consiste nel lasciare al musicista la libertà di definire i parametri musicali di alto livello (altezza, intensità e durata della nota da sintetizzare).

In letteratura sono riportati anche casi in cui la trasformazione delle serie temporali viene realizzata ad un livello inferiore rispetto al precedente producendo gli elementi fondamentali del suono: i campioni audio [3, 4].

Generalmente un sistema per la generazione di suoni e musica a partire da sistemi dinamici può essere suddiviso in tre parti:

- un sistema dinamico che genera delle serie temporali,
- un sistema di codifica che trasforma le serie temporali in

parametri musicali,

- un motore di sintesi che genera dei suoni in funzione dei dati forniti dal processo di codifica.

La figura 1 mostra una rappresentazione schematica di un sistema per la generazione di suoni e musica tramite sistemi dinamici.

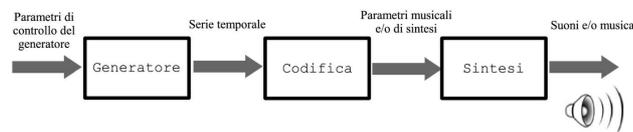


Figura 1. Schema di un sistema per la generazione di suoni e musica tramite sistemi dinamici.

L'aspetto più importante nella rappresentazione uditiva di sequenze di dati consiste nella codifica delle serie temporali. In letteratura spesso il sistema di codifica viene anche chiamato "mapping" sottolineando che nella maggior parte dei casi questa operazione consiste nella definizione di una opportuna funzione che instaura una relazione univoca tra gli elementi dello spazio delle fasi del sistema dinamico e gli elementi nello spazio dei parametri musicali. La figura 2 mostra schematicamente la relazione che è possibile instaurare tramite il sistema di codifica fra una serie di dati e la sua rappresentazione uditiva.<sup>1</sup>

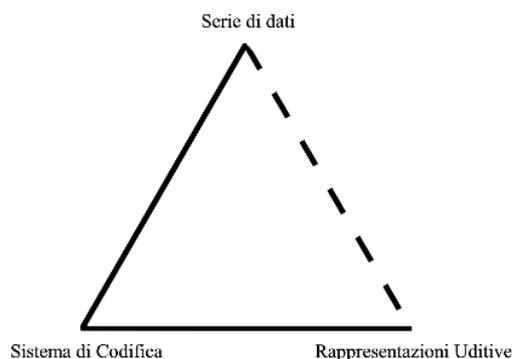


Figura 2. Triangolo semantico relativo alla rappresentazione uditiva di serie di dati.

## SISTEMI DI GENERAZIONE DI MELODIE

Una semplice codifica, da impiegare in un sistema di musica generativa per produrre sequenze di note, può essere definito tramite una relazione tra il campo di valori in cui è definita la serie temporale prodotta dal sistema dinamico e un determinato intervallo di note MIDI.

Questa trasformazione può essere realizzata tramite un mapping lineare che trasforma la variabile  $x$ , definita in  $[a, b]$ , nella variabile  $y$ , definita in  $[c, d]$  [13]:

$$y = \frac{(d - c)}{(b - a)}(x - a) + c \quad (1)$$

<sup>1</sup>La figura utilizzata è spesso impiegata negli studi sul linguaggio e la comunicazione e viene chiamata "triangolo semantico".

Nel caso in cui la variabile  $x$  é continua, allora lo sarà anche la  $y$ ; altrimenti se  $x$  é discreta anche la  $y$  sarà discreta. Questo semplice schema di codifica soffre di un grosso limite: l'impossibilità di definire un preciso vocabolario di altezze da utilizzare nella codifica delle serie temporali in sequenze di note. Le sequenze musicali ottenute attraverso questo approccio spesso sono musicalmente poco interessanti poiché la definizione di un intervallo di note MIDI comporta l'impiego di una collezione di note di tipo cromatico<sup>2</sup>. Inoltre poiché la maggior parte dei sistemi dinamici sono caratterizzati da un andamento pressoché continuo questo tipo di codifica porta alla generazione di scale cromatiche ascendenti e discendenti che risultano estremamente noiose all'ascolto e scarsamente interessanti dal punto di vista musicale.

Questo problema può essere superato definendo dei sistemi di codifica differenti che diano la possibilità di scegliere liberamente il vocabolario di altezze da utilizzare. In tal modo é possibile consentire all'utente di determinare il tipo di materiale musicale prodotto dal sistema di musica generativa a seconda della collezione di altezze adottata.

### Codifica Geometrica

La *codifica geometrica* é un sistema che consente all'utente di definire la collezione di altezze da utilizzare nella codifica delle serie temporali. In tal modo l'utente può definire il tipo di scala musicale e la tonalità da utilizzare. Inizialmente é necessario scegliere l'ampiezza dell'intervallo di altezze che s'intende utilizzare nel processo di codifica tramite la definizione di un estremo superiore e di un estremo inferiore (ad esempio si può considerare un intervallo di due ottave centrato attorno al Do centrale del pianoforte). La scelta delle due note che limitano l'intervallo di altezze può essere determinante per la definizione della tonalità o del modo che caratterizza la collezione di altezze.<sup>3</sup> Successivamente é necessario scegliere il tipo di scala musicale da utilizzare. L'elenco di seguito riporta alcune delle principali scale musicali, tra le tante possibili, da utilizzare per la definizione della collezione di altezze:

- diatonica - sette note per ottava
- pentatonica - cinque note per ottava
- scala basata su triadi - primo, terzo e quinto grado di una scala diatonica.

La scelta di una scala musicale consente di definire univocamente la collezione di altezze musicali da utilizzare, che può essere rappresentata tramite l'insieme ordinato  $S_h$  che

<sup>2</sup>Una collezione di note basata sull'impiego di una *scala cromatica*, ossia una successione ordinata per altezze ascendente o discendente dei dodici semitoni della scala temperata costituenti l'ottava.

<sup>3</sup>La tonalità é, in senso lato, il rapporto di gerarchia che si stabilisce, all'interno di una successione determinata di suoni, nei confronti di una tonica, ossia di un suono prescelto attorno al quale, e verso il quale, gravitano tutti gli altri.

Il modo é la scala che inizia a partire da un determinato grado di una scala diatonica (ad esempio il modo Ionico é costruito a partire dalla tonica, ossia dal primo grado della scala di riferimento, mentre il modo Eolio é costruito sul sesto grado).

comprende tutte le  $h$  note generabili dal processo di codifica:

$$S_h = \{nota_1, nota_2, \dots, nota_h\} \quad (2)$$

Prendendo in considerazione un sistema dinamico caratterizzato da uno spazio delle fasi  $n$ -dimensionale, per realizzare questo tipo di codifica é necessario calcolare il valore minimo  $m$  e il valore massimo  $M$  per ognuna delle  $n$  serie temporali che caratterizzano il sistema. In tal modo si otterranno  $n$  coppie di valori del tipo  $[m_i, M_i]$  che definiscono l'intervallo di variazione relativo alla  $i$ -esima serie temporale del sistema dinamico.

L'insieme di tutte le coppie di valori può essere pensato come l'insieme dei vertici che definiscono un parallelepipedo in uno spazio a  $n$  dimensioni all'interno del quale é vincolata l'evoluzione del sistema dinamico.

Se si prende in considerazione solo l' $i$ -esima coordinata dello spazio delle fasi la lunghezza dello spigolo può essere calcolata come:  $(M_i - m_i)$ . Ogni spigolo può essere suddiviso in  $h$  segmenti di egual lunghezza; la lunghezza di ogni segmento  $l_i$  può essere calcolata tramite la relazione:

$$l_i = \frac{M_i - m_i}{h} \quad (3)$$

Il processo di codifica può essere realizzato associando ad ognuna delle  $n$  serie temporali  $x_i(t)$  ( $i = 1, \dots, n$ ), che caratterizzano l'evoluzione del sistema dinamico, una sequenza di numeri interi  $p_i(t)$  ( $i = 1, \dots, n$ ) definiti nell'intervallo  $[1, h]$ . Questo processo di codifica determina una quantizzazione delle serie temporali secondo il numero di note dell'insieme  $S_h$ . Considerando il caso dell' $i$ -esima serie temporale l'indice del segmento in cui cade il valore corrente  $x_i(t)$  della serie temporale può essere calcolato tramite la relazione:

$$p_i(t) = 1 + \frac{x_i(t) - m_i}{l_i} \quad (4)$$

Il valore ottenuto per l'indice  $p_i(t)$  definirà quale elemento dell'insieme  $S_h$  si sceglierà come altezza della prossima nota da suonare. Le note possono essere generate secondo due differenti approcci:

- asincrono - si genera una nota ogni qualvolta si determina una variazione rispetto al valore precedente di uno degli indici  $p_i(t)$ ,
- sincrono - prima di effettuare la codifica si campionano le serie temporali secondo un determinato periodo di campionamento.

### Chaotic Performer

Il *Chaotic Performer* é una applicazione sviluppata mediante l'ambiente di programmazione visuale Pure Data (PD) e finalizzata alla creazione di un sistema d'interazione fra musicista e un sistema di musica. Il sistema di interazione deve consentire di instaurare un flusso d'informazione bidirezionale fra musicista e sistema generativo, cosí da rendere possibile la realizzazione di una mutua interdipendenza fra i gesti musicali dell'utente e il materiale sonoro prodotto dal sistema generativo. L'obiettivo principale consiste nel realizzare un sistema di musica generativa capace di avere in-

sieme ad uno o piú musicisti un controllo simultaneo e condiviso sul processo di produzione del materiale musicale. Come afferma Winkler: “*Interaction is a two way street*” [14], per tale motivo si é provveduto a dotare il Chaotic Performer di un “*ascoltatore*”, ossia di un sistema capace di estrarre e organizzare le informazioni relative al materiale musicale prodotto da un musicista. In seguito i dati raccolti dall’*ascoltatore* vengono utilizzati da una seconda parte del sistema, chiamato “*compositore*”, per la generazione di melodie tramite l’impiego di sistemi dinamici.

La figura 3 mostra l’interfaccia grafica del Chaotic Performer; in alto é posizionata la zona relativa all’*ascoltatore* in cui vengono visualizzate, tramite tabelle, le informazioni raccolte sulla performance del musicista. L’*ascoltatore* riceve un flusso di dati MIDI ed estrae le seguenti informazioni:

- stima del tempo
- classificazione delle note in classi di altezza (es. DO, DO#, RE, Mib, ....., SI)
- durata delle note
- intervallo di tempo fra l’attacco di due note successive (Delta-Time o IOI-Inter Onset Interval)
- intensità dell’note.

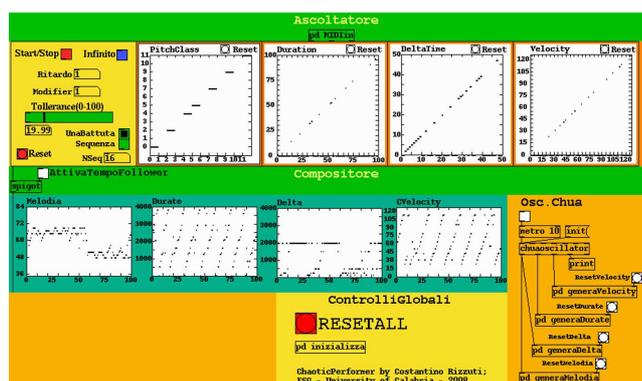


Figura 3. Interfaccia grafica del Chaotic Performer.

La parte sottostante all’*ascoltatore* é dedicata al *compositore*. Sulla destra é collocato il sistema di musica generativa in cui si utilizza l’oscillatore di Chua (descritto in dettaglio in una prossima sezione) come sistema dinamico per la generazione delle sequenze di note. Le tabelle che si trovano in questa parte dell’interfaccia grafica mostrano l’andamento dei seguenti parametri musicali: altezza, durata, delta-time, intensità. Il sistema di musica generativa impiega la codifica geometrica per trasformare le serie temporali prodotte dall’oscillatore di Chua nelle sequenze dei quattro parametri musicali. Nel processo di codifica si impiegano le informazioni raccolte dall’*ascoltatore* come collezioni di dati ammissibili per i diversi parametri musicali.

L’impiego del Chaotic Performer é ancora in fase di sperimentazione e di sviluppo, ma i risultati forniti dai primi esperimenti sono estremamente incoraggianti. Giá allo stato attuale il sistema fornisce un buon livello d’interazione e produce del materiale musicale interessante e dotato di un

certo grado di complessità. Il Chaotic Performer può essere impiegato per lo sviluppo di sistemi di accompagnamento automatici e per la generazione di linee melodiche secondo un preciso stile musicale. Quest’ultima possibilità potrebbe risultare interessante per la composizione poiché pur non utilizzando il sistema di analisi in tempo reale é possibile inserire nelle tabelle i dati desiderati per la generazione delle sequenze melodiche.

In fine uno sviluppo per migliorare l’impiego di questa applicazione nelle performance dal vivo consiste nel dotare il sistema di composizione di un “*ascoltatore interno*” che consenta di compiere un confronto in tempo reale fra il materiale prodotto dal sistema di musica generativa e i dati esterni provenienti dai musicisti.

## SISTEMI PER LA SINTESI DEL SUONO

I sistemi dinamici caotici possono essere utilizzati anche per realizzare algoritmi di sintesi del suono ed elaborazione di segnali. In questo ambito sono stati creati due “strumenti musicali”, attualmente ancora in fase prototipale: il “*Timbralizer*” basato sulla codifica dell’evoluzione di serie temporali sotto forma di cambiamenti del timbro e il “*Chaotic Modulator*” basato su un sistema di sintesi del suono in cui si impiega l’oscillatore di Chua.

### Timbralizer

Il *Timbralizer* [12] é uno strumento in cui si controllano i parametri di un sistema di sintesi sonora mediante l’impiego di sistemi caotici. Questo tipo di approccio sembra essere molto promettente, specialmente dal punto di vista delle applicazioni musicali, poiché lascia la possibilità all’utente di controllare, tramite un qualsiasi dispositivo MIDI, i parametri musicali delle note (altezza, intensità, durata).

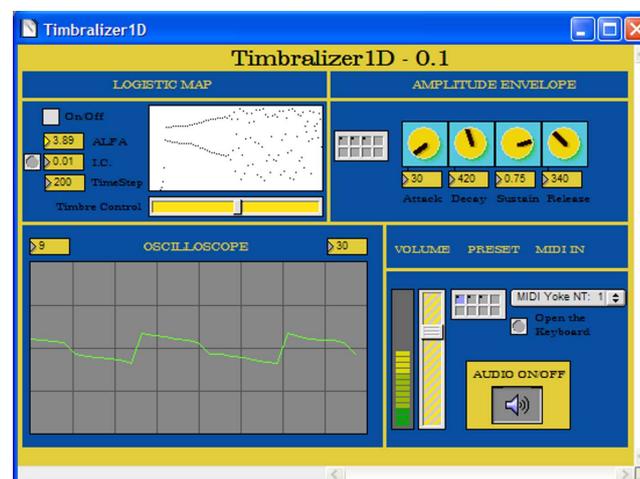


Figura 4. Interfaccia grafica del Timbralizer-1D.

L’immagine in figura 4 mostra la finestra del prototipo del *Timbralizer-1D*<sup>4</sup>. Sono state sviluppate tre differenti versioni da impiegare con sistemi a una dimensione, a due di-

<sup>4</sup>Il *Timbralizer-1D* può essere scaricato dal sito: <http://www.musichenza.it/timbralizer/>

mensioni e tridimensionali. Il prototipo della versione monodimensionale è completamente funzionante, mentre le altre due versioni sono ancora in fase di sviluppo. Nell'immagine è possibile notare in alto a sinistra la finestra che visualizza l'evoluzione del sistema dinamico e in basso la finestra che visualizza la forma d'onda generata dal sintetizzatore. Nel Timbralizer-1D la serie temporale prodotta da un sistema dinamico mono-dimensionale<sup>5</sup> viene trasformata nell'evoluzione del timbro del suono generato dallo strumento. Il timbro, definito come la qualità del suono che consente di distinguere tra due suoni della stessa altezza e intensità, è in genere associato al "colore" del suono o più precisamente alla distribuzione e alla variazione nel tempo del suo contenuto spettrale. La caratteristica che differenzia il timbro dalle altre qualità di un suono consiste nell'impossibilità di discretizzare e ordinare il timbro dei suoni lungo una scala unidimensionale. Infatti, frequentemente due suoni differiscono simultaneamente per diversi aspetti e ciò spinge a pensare al timbro come ad una grandezza multidimensionale. Numerosi ricercatori hanno proposto differenti tipi di rappresentazioni spaziali del timbro per l'analisi, la classificazione o per la sintesi dei suoni. L'idea di base consiste nella possibilità di definire un insieme di coordinate all'interno di uno spazio timbrico al fine di definire un timbro tramite le coordinate date. Wessel [15] ha creato uno spazio timbrico a due dimensioni tramite il quale era possibile controllare un sistema di sintesi additiva. Una dimensione era utilizzata per controllare la distribuzione del contenuto spettrale, mentre la seconda dimensione era utilizzata per controllare il tempo di attacco dell'involuppo d'ampiezza del suono e il grado di sincronia fra l'attacco delle varie componenti spettrali.

Questi spazi timbrici possono essere utilizzati anche per la creazione di "sound morphing", ossia delle trasformazioni, definite da opportune funzioni di morphing, che consentono di interpolare due timbri differenti in maniera da ottenere i due suoni originari agli estremi della funzione di morphing e dei suoni ibridi che posseggono, con gradi diversi, le caratteristiche di entrambi i suoni per tutti i valori intermedi.

Il Timbralizer impiega una codifica basata su un sound morphing tra due ben definite forme d'onda per trasformare una serie temporale nell'evoluzione del timbro. Il sistema di sintesi impiegato è di tipo additivo ed è caratterizzato dalla somma di dieci parziali sinusoidali. Un'opportuna funzione di mapping consente di controllare tramite un unico parametro l'ampiezza di tutte le parziali. Questa funzione è stata definita in maniera da ottenere al variare del valore del parametro un sound morphing tra una forma d'onda a dente di sega e un'onda quadra.

Il sintetizzatore è stato dotato di una interfaccia MIDI che consente il controllo mediante una tastiera per la definizione dei parametri necessari alla generazione delle note, mentre il contenuto spettrale del suono prodotto dallo strumento evolve seguendo la serie temporale prodotta dal sistema dinamico preso in considerazione.

Le prove effettuate hanno mostrato che questo strumento consente di realizzare una buona rappresentazione uditiva dell'evoluzione del sistema dinamico. È possibile infatti riuscire a seguire in maniera qualitativa e riconoscere tramite

<sup>5</sup>Nell'esempio riportato è stata impiegata la mappa logistica definita tramite la relazione:  $x_{n+1} = \alpha x_n(1 - x_n)$

l'ascolto i comportamenti topologicamente diversi caratteristici dei sistemi dinamici caotici. Nel caso della mappa logistica, ad esempio, è possibile riuscire a distinguere le tre diverse zone che caratterizzano il comportamento di tale sistema:

- convergenza ad un punto fisso ( $\alpha < 2$ )
- comportamento periodico ( $2 \leq \alpha < 3.5$ )
- comportamento caotico ( $3.5 \leq \alpha < 4$ ).

Da un punto di vista artistico-musicale questo strumento può essere utilizzato per generare dei suoni caratterizzati da una evoluzione del timbro estremamente ricca e interessante che si avvicina molto alla complessità dei suoni prodotti dagli strumenti naturali. È possibile controllare l'evoluzione del timbro operando su due distinti fattori:

- modifica dei parametri del sistema dinamico - generalmente le zone caotiche e le zone di transizione tra comportamento periodico e caotico sono le zone in cui si ottengono sonorità più ricche e interessanti
- modifica del periodo d'iterazione del sistema dinamico - in tal modo è possibile ottenere degli effetti di modulazione (tipo phaser, flanger) per valori inferiori a 100ms, oppure complesse evoluzioni del timbro per valori superiori.

#### Chaotic Modulator

Il *Chaotic Modulator* è uno strumento musicale in cui si impiega l'oscillatore di Chua [16, 17] all'interno di un processo di sintesi del suono per modulazione di frequenza. Il prototipo del Chaotic Modulator è stato creato impiegando l'ambiente di programmazione visuale Pure Data. La simulazione dell'oscillatore di Chua è stata realizzata creando un apposito plug-in per Pure Data che compie l'integrazione numerica<sup>6</sup> del sistema di equazioni differenziali dello oscillatore di Chua per generare un segnale di tipo audio.

Di seguito è riportato il sistema di equazioni differenziali relativo alla formulazione adimensionale dell'oscillatore di Chua:

$$\begin{aligned} \dot{x} &= k\alpha[y - x - f(x)] \\ \dot{y} &= k(x - y + z) \\ \dot{z} &= k(-\beta y - z) \end{aligned} \quad (5)$$

dove  $f(x)$  è definito come:

$$f(x) = bx + \frac{1}{2}(a - b)|x + 1| - |x - 1|$$

L'oscillatore di Chua genera tre distinte serie temporali ciascuna relativa a una delle tre variabili ( $x, y, z$ ) che caratterizzano lo stato del sistema. Le tre serie temporali prodotte mediante l'integrazione numerica del sistema di equazioni differenziali sono state utilizzate per modulare in frequenza

<sup>6</sup>L'integrazione numerica del sistema d'equazioni si realizza impiegando il metodo di Runge-Kutta del quarto ordine.

tre oscillatori sinusoidali. Le frequenze portanti dei tre oscillatori sono state definite tramite dei rapporti armonici in modo che in assenza di modulazione forniscano un suono contenente la prima, la seconda e la quarta armonica. La frequenza fondamentale del sintetizzatore può essere controllata attraverso una tastiera o qualsiasi altro dispositivo MIDI. La figura 5 mostra lo schema del sistema di sintesi utilizzato.

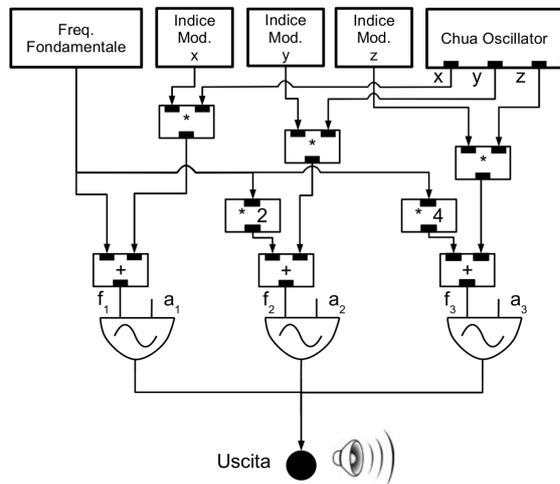


Figura 5. Schema del sistema di sintesi impiegato nel Chaotic Modulator.

Attraverso il controllo dei tre indici di modulazione (in alto nella figura 5) è possibile modificare il contenuto spettrale del suono passando da un suono con quasi esclusivamente parziali armoniche, nel caso di valori bassi dei tre indici, a un suono dal timbro articolato e con una forte presenza di componenti inarmoniche e di rumore, ottenuto per valori alti dei tre indici.

### AMBIENTI IMMERSIVI 3D AUDIO/VIDEO

Lo studio e l'esplorazione dei sistemi complessi può essere condotto anche tramite degli ambienti immersivi 3D audio e video. L'impiego di questi ambienti consente di combinare simultaneamente la visualizzazione e l'ascolto dei suoni e delle musiche generate dai diversi sistemi dinamici.

La creazione di questi ambienti viene realizzata impiegando Audioscape<sup>7</sup>, un framework per la creazione di ambienti virtuali immersivi audio/video che consente di simulare la propagazione del suono in un ambiente tridimensionale [18, 19]. Questa architettura, basata sulla modellizzazione dei fenomeni fisici che caratterizzano l'acustica, consente di ricreare un ambiente virtuale 3-D sia visivo che acustico. Tramite Audioscape sono stati creati e sperimentati diverse applicazioni finalizzate sia alla generazione di suoni e musiche sia per la produzioni di accompagnamenti musicali mediante sistemi caotici [20].

Il risultato di questa attività di ricerca è stato presentato sotto forma di installazione multimediale nell'ambito della mostra "Forme Caotiche, un viaggio fra arte e scienza" svolta dal 19 al 24 Gennaio 2008 presso il Complesso Monumentale del

<sup>7</sup><http://www.audioscape.org>

Valentianum di Vibo Valentia. Per questa occasione è stato creato un mondo virtuale tridimensionale in cui sono stati collocati gli attrattori generati dall'oscillatore di Chua. Lo scopo dell'installazione è stato consentire ai visitatori di interagire con gli attrattori caotici all'interno di un ambiente 3D e presentare i sistemi di generazione di suoni e musica in cui s'impiegano questi attrattori.

La figura 6 mostra alcune immagini dell'ambiente virtuale e dell'installazione svolta presso la mostra Forme Caotiche.

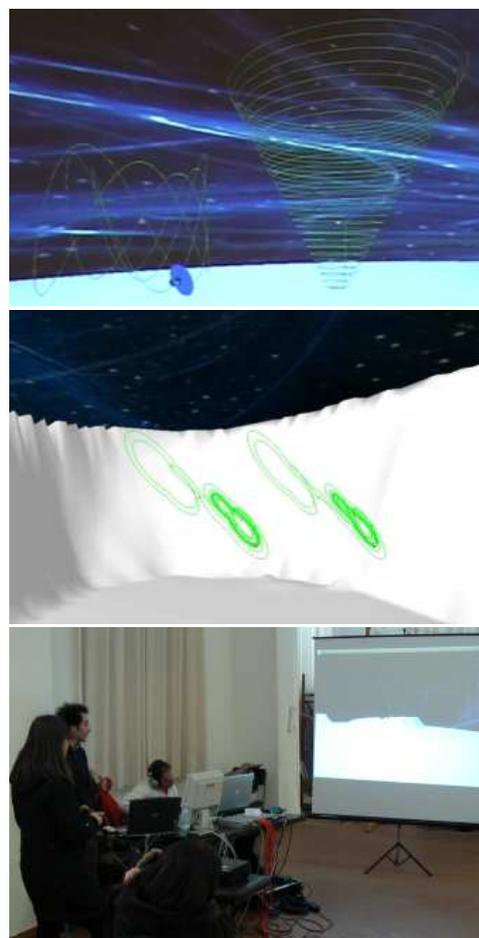


Figura 6. Immagini dell'installazione presentata alla mostra Forme Caotiche.

L'ambiente virtuale è stato suddiviso in quattro diverse aree:

1. Presentazione dei suoni e delle musiche prodotte dagli attrattori dell'oscillatore di Chua.
2. Spazializzazione tridimensionale riprodotta tramite un impianto dotato di otto altoparlanti. Gli spettatori hanno assistito all'evoluzione di alcune traiettorie tridimensionali (prima immagine della figura 6) nell'ambiente grafico e hanno avuto la possibilità di seguire il movimento della sorgente sonora tramite la spazializzazione tridimensionale.
3. Sperimentazione della sensibilità alle condizioni iniziali (seconda immagine della figura 6). È stata mostrata

l'evoluzione simultanea di due attrattori caratterizzati da condizioni iniziali fra loro molto simili. Per un certo periodo di tempo i due attrattori seguono la stessa traiettoria, in seguito le traiettorie evolvono in maniera autonoma. L'evoluzione di ogni attrattore è stata trasformata in una sequenza di note con un dato suono. Inizialmente è possibile ascoltare un'unica melodia poiché gli attrattori seguono la stessa evoluzione, in seguito, dopo la separazione, è possibile distinguere due differenti melodie.

4. Il giardino degli attrattori, presentazione di una galleria di sei diversi attrattori che evolvono in tempo reale.

## CONCLUSIONI

Nel presente lavoro si è provveduto a illustrare alcune applicazioni che impiegano sistemi dinamici caotici per la generazione di suoni e musica. Allo stesso tempo sono state affrontate, da un punto di vista teorico, alcune delle problematiche che stanno alla base della rappresentazione uditiva di sequenze di dati. Infine è stato illustrato l'impiego di ambienti 3D audio/video per lo studio e la rappresentazione di sistemi complessi.

Le prossime attività di ricerca saranno particolarmente concentrate sulla realizzazione di esperimenti percettivi per valutare la qualità delle rappresentazioni uditive fornite dai sistemi qui presentati. Inoltre si cercherà, tramite la collaborazione con musicisti e compositori, di utilizzare tali sistemi per la creazione di performance musicali e installazioni.

## RIFERIMENTI

1. J. Pressing, *Nonlinear Maps as Generators of Musical Design*, Computer Music Journal, Vol. 12, No. 2 Summer, MIT, pp. 35-45, 1988.
2. R. Bidlack, *Chaotic systems as Simple (but Complex) Compositional Algorithms*, Computer Music Journal, Vol. 16, No. 3, MIT, pp. 33-47, 1992.
3. A. Di Scipio, I. Prignano, *Synthesis by Functional Iterations. A Revitalization of Nonstandard Synthesis*, Journal of New Music Research, Vol 25, pp.31-46, 1996.
4. A. Di Scipio, *The Synthesis of Environmental Sound Texture by Iterated Nonlinear Functions, and its Ecological Relevance to perceptual Modeling*, Journal of New Music Research, Vol 31, No.2, pp. 109-117, 2002.
5. M. Witten, *The Sounds of science: II. Listening to Dynamical systems - Towards a Musical Exploration of Complexity*, Computers Mathematic with Applications, Vol.32, No.1, Elsevier Science, pp.145-173, 1996.
6. E. Bilotta, P. Pantano, S. Gervasi, *Readings Complexity in CHUA's Oscillator Through Music. Part I: A New Way Of Understanding Chaos*, International Journal of Bifurcation and Chaos, Vol.15, No. 2, pp. 253-382, 2005.
7. E. Bilotta, P. Pantano, V. Talarico, *Music Generation Through Cellular Automata: How to Give Life to Strange Creatures*, Proceedings of Generative Art GA 2000, Milano, 2000.
8. E. Bilotta, P. Pantano, *Synthetic Harmonies: an approach to musical semiosis by means of cellular automata*, Leonardo, Vol. 35, No. 2, pp. 153-159, 2002.
9. K. Mc Alpine, E.R. Miranda, S. Hoggar, *Making Music with Algorithms: A Case study*, Computer Music Journal, Vol. 23, No. 2, MIT, pp. 19-30, 1999.
10. Y. Khalifa, R. Foster, *A two-state autonomous evolutionary music composer*, F. Rothlauf et al.(Eds): EvoWorkshops 2006, LNCS 3907, pp. 717-721, Springer-Verlag, Berlin Heidelberg, 2006.
11. E. Bilotta, P. Pantano, E. Cupellini, C. Rizzuti, *Evolutionary Methods for Melodic Sequences Generation from Non-linear Dynamics Systems*, in M. Giacobini et al. (Eds.): Evoworkshops 2007, LNCS 4448, pp.585-592, Springer-Verlag Berlin Heidelberg, 2007.
12. C. Rizzuti, *Mapping Chaotic Dynamical Systems Into Timbre Evolution*. Proc. of "SMC 07" Conference, Lefkada, Greece, 11-13 July, 2007, ISBN 978-960-6608-75-9, pp. 22-29, National and Kapodistrian University of Athens, Athens, 2007.
13. C. Dodge, *Computer Music*, Schirmer, New York, 1997.
14. T. Winkler, *Composing Interactive Music - Techniques and Ideas Using Max*, The MIT Press, Cambridge Massachusetts, 2001.
15. D.L. Wessel, *Timbre Space As A Musical Control Structure*, Rapport IRCAM 12/78, 1978.
16. L.O. Chua, C.W. Wu, A. Huang, G.Q. Zhong, *A universal circuit for studying and generating chaos. I. routes to chaos*, IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 40:732744, 1993.
17. L.O. Chua, C.W. Wu, A. Huang, G.Q. Zhong., *A universal circuit for studying and generating chaos. II. strange attractors*, IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 40:745761, 1993.
18. M. Wozniowski, Z. Settel, J.R. Cooperstock, *A paradigm for physical interaction with sound in 3-D audio space*, In Proceedings of International Computer Music Conference (ICMC), 2006.
19. M. Wozniowski, Z. Settel, J.R. Cooperstock, *User-specific audio rendering and steerable sound for distributed virtual environments*, In Proceedings of International conference on auditory displays (ICAD), 2007.
20. E. Bilotta, J.R. Cooperstock, E. Cupellini, P. Pantano, C. Rizzuti, M. Wozniowski, *Exploring Musical Mappings and Generating Accompaniment with Chaotic Systems*, In Proceedings of International Computer Music Conference (ICMC), 2008.



# DepThrow: uno strumento di indagine sulla percezione uditiva della distanza in forma di gioco audio

**Stefano Papetti**  
Dipartimento di Informatica  
Università di Verona

**Delphine Devallez**  
Dipartimento di Informatica  
Università di Verona

**Federico Fontana**  
Dipartimento di Informatica  
Università di Verona

## SOMMARIO

*DepThrow* è una applicazione software per Max/MSP<sup>1</sup> che sfrutta la percezione uditiva della distanza, e fa uso di modelli fisici per la simulazione della dinamica, della sorgente sonora e dell'ambiente acustico. L'applicazione è in forma di "gioco audio": durante l'esperienza di gioco l'utente non dispone di alcuna rappresentazione visiva e deve fare esclusivo affidamento sull'informazione uditiva disponibile. L'azione di gioco consiste nel lanciare un oggetto sonoro in un tubo aperto alle estremità e inclinato, e l'obiettivo è quello di fare arrivare l'oggetto il più lontano possibile, evitando che fuoriesca dal tubo. L'informazione sulla posizione corrente dell'oggetto all'interno del tubo viene fornita da un *feedback* audio continuo. La prestazione del giocatore dipende fortemente dalla sua capacità di percepire la distanza, dinamicamente variabile, dell'oggetto nel tubo. Per tale ragione l'applicazione rappresenta uno strumento per l'analisi della percezione e dell'usabilità dell'informazione uditiva di distanza nel contesto della *Human-Computer Interaction* (HCI). L'algoritmo di spazializzazione qui descritto, così come i risultati della suddetta analisi, sono inoltre di potenziale interesse per l'esecuzione e la composizione di musica elettroacustica.

## Parole Chiave

Spazializzazione dinamica; Digital Waveguide Mesh; Suono di rotolamento; Mappatura del gesto; Modelli fisici.

## INTRODUZIONE

La possibilità di spazializzare dinamicamente ed in tempo reale il segnale di una sorgente sonora risulta di grande interesse sia in campo musicale che in ambito HCI. Per quanto riguarda l'ambito dell'interazione uomo-macchina, la resa acustica della distanza è stata finora tenuta in scarsa considerazione, ciononostante essa promette di offrire nuove e potenti modalità per la comunicazione di informazioni. Alcuni esempi di applicazioni HCI che potrebbero trarre vantaggio da tale approccio sono la sonificazione di interfacce (ad esempio, l'interfaccia di un sistema operativo o di apparecchiature domestiche) e la simulazione di realtà virtuale.

Il modello di spazializzazione dinamica implementato in *DepThrow* consente di variare in tempo reale la posizione della sorgente sonora all'interno di un ambiente virtuale. Tale modello simula la propagazione delle onde in un ambiente acustico e risulta perciò essere fisicamente coerente. Grazie a tale coerenza fisica, i parametri di controllo del modello di

<sup>1</sup><http://www.cycling74.com/products/>

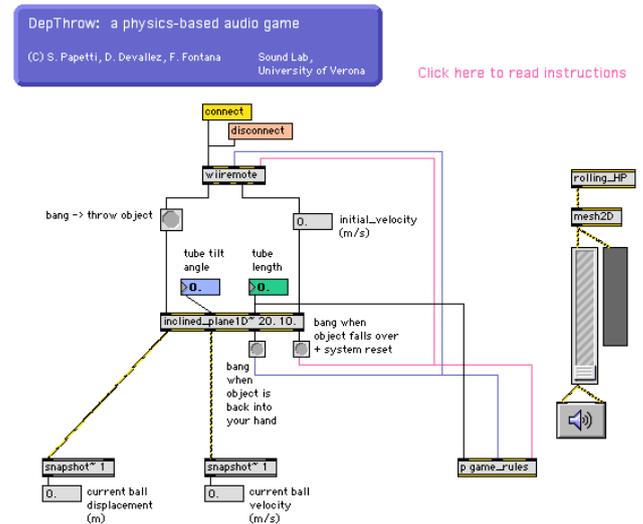


Figura 1. Il patch Max/MSP dell'applicazione.

spazializzazione corrispondono direttamente o sono riconducibili a grandezze fisiche quali, ad esempio, le dimensioni e le proprietà riflettenti dell'ambiente, la distanza della sorgente sonora, etc.

Per quanto riguarda la dinamica di gioco, la simulazione di uno spazio acustico consente di fornire all'utente informazioni sulla posizione dinamica di un oggetto sonoro in un ambiente virtuale.

## STRUTTURA E FUNZIONAMENTO

In figura 1 è rappresentato il patch Max/MSP dell'applicazione. In esso, tre modelli fisici interagiscono e permettono di simulare rispettivamente: la dinamica (un piano inclinato), la sorgente sonora (una sfera rotolante su di una superficie piana), e le proprietà acustiche dell'ambiente (un tubo aperto alle estremità).

L'azione di lancio avviene mediante un piccolo carrello provvisto di ruote sul quale è fissato un controller *Nintendo Wii Remote*<sup>2</sup>. Tale controller dispone, tra le altre cose, di un accelerometro tridimensionale, un attuatore di vibrazione, diversi bottoni e connettività Bluetooth.

<sup>2</sup>[http://en.wikipedia.org/wiki/Wii\\_Remote](http://en.wikipedia.org/wiki/Wii_Remote)

I tre modelli fisici sono interconnessi in una struttura di tipo *top-down* secondo lo schema seguente:

1. **mappatura del gesto:** grazie all'accelerometro, la forza applicata al carrello viene campionata e utilizzata per il calcolo della velocità con cui la sfera virtuale viene lanciata all'interno del tubo inclinato;
2. **dinamica:** la velocità iniziale calcolata al passo precedente viene utilizzata come parametro di ingresso per la simulazione della fisica di un piano inclinato;
3. **sonificazione:** la velocità risultante dalla simulazione precedente controlla la sintesi in tempo reale del suono di rotolamento di una sfera su di una superficie piana;
4. **spazializzazione:** il segnale sonoro sintetizzato dal modello di rotolamento viene dato in ingresso ad un processo di spazializzazione che simula l'acustica di un tubo. La posizione corrente dell'oggetto lungo il piano inclinato controlla il parametro di distanza del modello di spazializzazione. Il modello permette infatti di cambiare la posizione della sorgente sonora in tempo reale, e quindi di simulare l'avvicinamento o l'allontanamento dall'ascoltatore.

L'attuatore di vibrazione fornito dal controller *Wii Remote* viene utilizzato per fornire un *feedback* tattile quando la sfera virtuale torna alle mani dell'utente.

#### MAPPATURA DEL GESTO

L'interfaccia scelta per effettuare il lancio virtuale della sfera all'interno del tubo consiste in un piccolo carrello con ruote sul quale è montato un controller *Wii Remote* (v. Fig. 2).

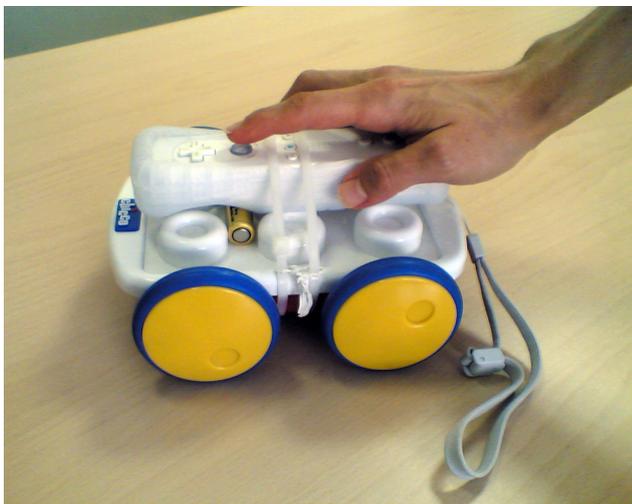


Figura 2. L'interfaccia di controllo del gioco consiste in un carrello sul quale è fissato un controller *Wii Remote*.

Una volta impugnato il controller *Wii Remote*, l'azione di gioco consiste nel fare scorrere il carrello fino a raggiungere la velocità di lancio desiderata, e premere il tasto **A** del controller quando si vuole rilasciare la sfera virtuale. La velocità del lancio viene calcolata a partire dall'informazione

di accelerazione lungo l'asse longitudinale (cioè coassiale alla direzione di lancio) fornita dall'accelerometro integrato. In particolare, il calcolo della velocità avviene mediante una semplice integrazione numerica dei dati di accelerazione forniti durante la pressione del tasto **A**. La formula di integrazione numerica adottata è la seguente:

$$v(n) = \sum_{i=0}^n a(i)dT \quad (1)$$

dove  $v(n)$  è la velocità calcolata al campione  $n$ -esimo,  $a(i)$  è l'accelerazione al campione  $i$ -esimo, il cui contributo viene considerato solo se  $a(i) \geq 0$ , e  $dT$  è il passo di integrazione, pari all'inverso della frequenza di acquisizione di dati dal controller (max 100 Hz).

Durante la fase di sviluppo dell'applicazione, la parte di mappatura del gesto è risultata piuttosto difficoltosa, e ciò principalmente a causa della scarsa precisione numerica fornita dall'accelerometro, e della impossibilità di mantenere il controller in posizione perfettamente orizzontale. Per questo stesso motivo è risultato più efficace integrare l'accelerazione in modo controllato, e solamente durante la pressione del tasto **A** invece che per l'intera durata del gesto. Inoltre è noto che lo scambio di dati via Bluetooth introduce una latenza nell'ordine della decina di millisecondi, e ciò concorre alla produzione di un percepibile ritardo tra l'azione di lancio e l'effettiva partenza della sfera virtuale. I suddetti problemi sarebbero probabilmente risolvibili adottando una politica più raffinata per il calcolo della velocità lungo l'asse longitudinale, e utilizzando un accelerometro di qualità più elevata. Tuttavia, allo stato attuale, la mappatura realizzata è risultata essere sufficientemente adeguata agli scopi dell'applicazione.

#### DINAMICA DEL PIANO INCLINATO

Il primo modello della catena è una simulazione semplificata della fisica del piano inclinato, espressa dalla seguente equazione:

$$v(t) = v_0 + a \cdot t \quad (2)$$

dove  $v_0$  è la velocità iniziale dell'oggetto e l'accelerazione  $a$  è costante e pari a:

$$a = g \cdot \sin \omega \quad (3)$$

dove  $g$  è l'accelerazione di gravità, e  $\omega$  è l'angolo di inclinazione espresso in gradi radianti.

In questa sede non è stata considerata l'eventuale presenza di attrito dovuto al contatto tra la superficie dell'oggetto e quella del piano, in quanto si suppone che il primo possa solo rotolare (caso in cui l'attrito è trascurabile) ma non scivolare lungo il piano. I parametri che descrivono il modello risultano quindi essere solamente la lunghezza e l'angolo di inclinazione del piano.

Il modello accetta in ingresso la velocità iniziale di un oggetto che, partendo dall'estremità inferiore, risalga lungo il piano, e ne restituisce posizione e velocità correnti.

Per quanto riguarda l'implementazione del modello come oggetto per Max/MSP, si è fatto in modo di fornire valori in uscita con frequenza pari alla frequenza di campionamento audio<sup>3</sup>, permettendo così di controllare in modo quasi continuo gli oggetti connessi in uscita, e riducendo al minimo eventuali artefatti dovuti alla discontinuità dei valori. Non essendo attualmente disponibili librerie Max/MSP per la simulazione della dinamica<sup>4</sup> che soddisfino il requisito di tale frequenza di aggiornamento, si è deciso di sviluppare un oggetto *ad hoc*.

## SUONO DI ROTOLAMENTO

La sorgente sonora viene generata utilizzando un modello di rotolamento [6] originariamente sviluppato e implementato nel corso del progetto europeo *SOB* [8] (the Sounding Object), e successivamente ripreso nell'ambito dell'attuale progetto europeo *CLOSED*<sup>5</sup> (Closing the Loop Of Sound Evaluation and Design).

Il modello di rotolamento combina in sé diverse componenti: il suo nucleo è costituito da un modello fisico di impatto [4] il quale viene pilotato da altri due modelli pseudo-fisici. Questi simulano rispettivamente: 1) la geometria delle superfici in contatto, la quale ha effetto sulla successione temporale di micro-impatti che generano il suono di rotolamento; 2) l'asimmetria della sfera, che genera una modulazione ciclica della forza di impatto. Si vedano a tale proposito le figure 3 e 4.

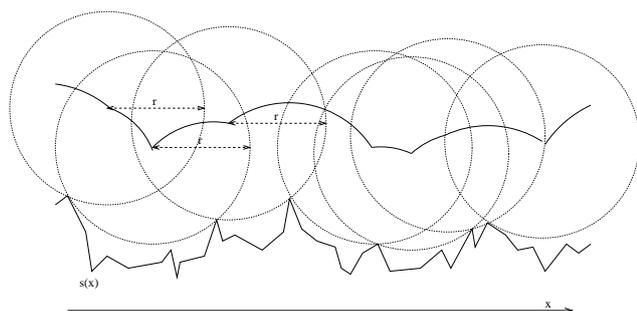


Figura 3. Schematizzazione della traiettoria di una sfera su di una superficie irregolare. Immagine tratta da [8]

Per quanto riguarda l'interazione di contatto tra la sfera e la superficie di rotolamento, vengono considerate solo le componenti normali di impatto e vengono invece trascurate le componenti di frizione (che nel caso del rotolamento sono comparativamente piccole). Il modello non-lineare di impatto adottato è conforme alla struttura modulare "risonatore-interattore-risonatore", rappresentando così l'interazione tra due oggetti risonatori. I risonatori sono modellati secondo la tecnica di *sintesi modale* [1], la quale descrive un oggetto in vibrazione mediante i suoi modi di risonanza, e in particolare descrivendo frequenza, tempo di decadimento e peso

<sup>3</sup>Ovvero pari al *signal rate*, in contrasto alla frequenza di aggiornamento dei controlli detta *control rate*, la quale risulta generalmente inferiore di diversi ordini di grandezza alla prima.

<sup>4</sup>Si consideri ad esempio *pmpd*, disponibile all'URL: <http://drpichon.free.fr/pmpd/>, i cui oggetti sono in grado di gestire solamente segnali a *control rate*.

<sup>5</sup><http://closed.ircam.fr>

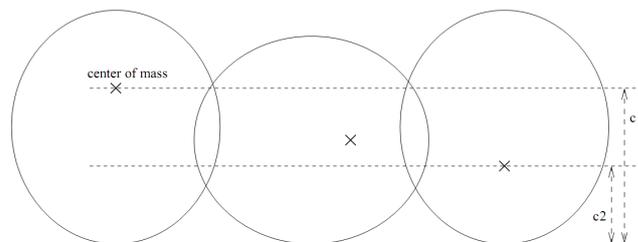


Figura 4. Esempificazione dello spostamento del centro di massa di una sfera asimmetrica. Immagine tratta da [6]

modale di ciascun modo. Nell'implementazione considerata l'interazione di impatto avviene tra una sfera e un tubo ma, per semplicità, solo quest'ultimo viene considerato come un vero e proprio risonatore, mentre la sfera viene modellata come una massa puntuale priva di modi risonanti.

Il profilo della superficie piana viene simulato mediante un campione audio pre-registrato, contenente rumore bianco processato attraverso un filtro passa-banda. Questo segnale viene dato in ingresso ad uno speciale filtro, il cui compito è quello di calcolare la traiettoria della sfera (o meglio, del suo centro di massa) rotolante su tale superficie (v. Fig. 3). Ovviamente la traiettoria dipende anche dal diametro della sfera. Il segnale così filtrato viene utilizzato per modulare su scala microscopica la distanza tra gli oggetti risonatori del modello di impatto.

Per semplicità l'eventuale rimbalzo della sfera non viene preso in considerazione.

## SPAZIALIZZAZIONE

Alcune recenti ricerche nell'ambito della spazializzazione del suono hanno avuto come obiettivo la modellazione fisica della propagazione di onde acustiche. In particolare, un approccio basato sulla fisica quale la modellazione mediante *digital waveguide* trova motivazione nella semplicità della formulazione del modello, e nell'intuitività dei parametri di controllo dovuta alla loro corrispondenza con grandezze fisiche. Fontana e Rocchesso [3] hanno sottolineato l'efficacia di una struttura *digital waveguide mesh* (DWM) che modelli un parallelepipedo rettangolare al fine di fornire indizi uditivi di profondità e distanza. Grazie agli esperimenti svolti, gli autori hanno dimostrato che tale modello è in grado di simulare la distanza apparente di una sorgente sonora.

### Digital waveguide mesh

In *DepThrow* l'ambiente acustico è simulato grazie ad una struttura *digital waveguide mesh* bidimensionale rettilinea, grazie alla quale viene modellata la trasmissione di onde acustiche tra le giunzioni interne della *mesh*. In particolare, ogni *waveguide* modella la decomposizione di un segnale di pressione  $p$  nelle sue componenti d'onda  $p^+$  e  $p^-$ , ed ogni nodo della *mesh* separa (fenomeno di *scattering*) senza perdite i quattro segnali in ingresso  $p_1^+$ , ...,  $p_4^+$ , provenienti da direzioni ortogonali, nei corrispondenti segnali di uscita  $p_1^-$ , ...,  $p_4^-$  (v. Fig. 5).

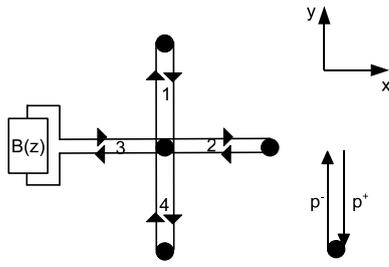


Figura 5. Particolare di un nodo della mesh connesso ad altri nodi mediante le waveguides 1, 2 e 4. Nell'esempio raffigurato la waveguide 3 non è connessa ad altri nodi, ma ad un bordo della mesh parzialmente assorbente. I triangoli lungo le waveguides rappresentano ritardi unitari.

Per uno spazio reale, le proprietà dei materiali delle pareti contribuiscono alla sua acustica. Lo stesso deve valere anche nel caso di un ambiente acustico virtuale e semplificato al caso bidimensionale, in quanto le onde interagiscono con i bordi della superficie. Le riflessioni ai bordi sono modellate mediante *digital waveguide filters* (DWF), i cui coefficienti sono stati regolati in modo da modellare delle specifiche proprietà di riflessione di superfici reali [2].

Una volta fissate la velocità di propagazione del suono e la frequenza di campionamento, il numero di nodi implementati nella mesh risulta corrispondere a specifiche dimensioni della superficie. Nel caso qui implementato, onde ridurre il carico computazionale il segnale in ingresso alla DWM viene sottocampionato, e le dimensioni della mesh sono state fissate pari a  $120 \times 7$  nodi. Considerando una frequenza di campionamento pari a 22050 Hz, tale configurazione corrisponde ad una membrana avente dimensioni  $2.62 \times 0.14 \text{ m}^2$ .

### Proprietà acustiche della membrana

Al fine di analizzare gli indizi uditivi di distanza forniti dall'ambiente virtuale, sono state calcolate le risposte all'impulso di differenti punti della membrana. In particolare, le misurazioni delle risposte all'impulso sono state effettuate in sei diverse posizioni lungo l'asse longitudinale della membrana, e la sorgente sonora, considerata puntiforme, è stata collocata sull'asse longitudinale, ad una distanza pari a 0.11 m da una delle estremità della membrana.

Durante la fase di sviluppo, la simulazione dell'ambiente acustico è stata effettuata in MATLAB. La figura 7 mostra due risposte in frequenza misurate in punti differenti della membrana virtuale.

### Intensità acustica

In figura 8, al variare della distanza, vengono confrontati gli andamenti dell'energia totale del segnale nella mesh e in uno spazio aperto di riferimento. Dal confronto appare evidente che, in generale, l'intensità diminuisce più lentamente nel caso della mesh, e ciò significa che lungo la membrana è possibile udire sorgenti sonore a distanze decisamente superiori.

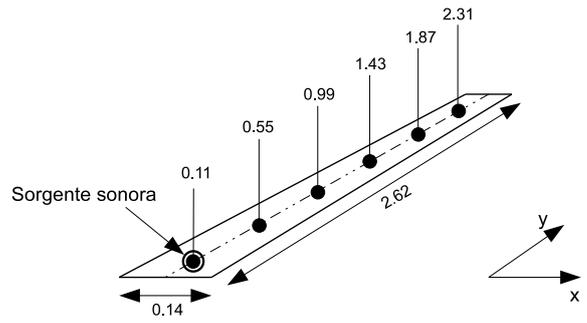


Figura 6. Raffigurazione della membrana virtuale su cui vengono evidenziate la posizione della sorgente sonora e le posizioni di misurazione delle risposte all'impulso. Tutti i valori sono espressi in metri.

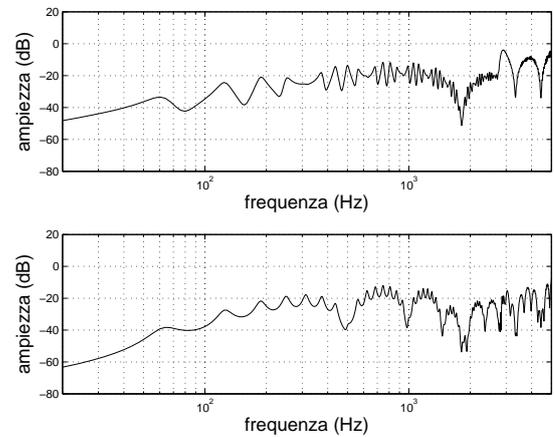


Figura 7. Risposte in frequenza della membrana fino a 5 kHz. La misurazione è stata effettuata rispettivamente alle distanze di 0.55 m (sopra) e 2.31 m (sotto) lungo l'asse longitudinale.

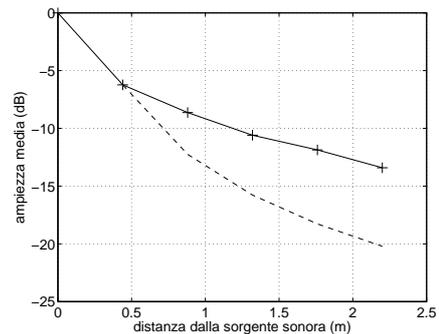
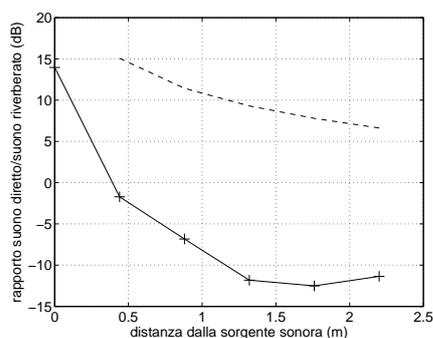


Figura 8. Ampiezza media della risposta all'impulso come funzione della distanza lungo la membrana. Linea continua: mesh bidimensionale. Linea tratteggiata: spazio aperto di riferimento.



**Figura 9.** Rapporto tra l'energia del segnale diretto e quella del segnale riverberato come funzione della distanza. Linea continua: *mesh* bidimensionale. Linea tratteggiata: approssimazione di un ambiente naturale.

#### Rapporto tra le energie del suono diretto e riverberato

L'ambiente acustico virtuale qui adottato si propone di fornire l'informazione di distanza facendo principalmente affidamento all'indizio percettivo dato dal rapporto tra l'energia del suono diretto e quella del suono riverberato (*direct-to-reverberant energy ratio*, abbreviato con *d-to-r* da qui in avanti).

Per ciascuna delle risposte all'impulso estratte, il ritardo del suono diretto è stato dedotto dalla distanza tra la sorgente sonora e il punto di ascolto, ed è stato rimosso. In seguito, al fine di isolare il percorso diretto del segnale sonoro [5], l'energia diretta è stata integrata nei primi 2.5 ms delle risposte all'impulso senza ritardo, approssimando così la durata delle *Head Related Impulse Responses* misurate in condizioni anecoiche. Infine, l'energia del segnale riverberato è stata calcolata dalla coda delle risposte all'impulso senza ritardo.

In figura 9 sono mostrati i valori del rapporto *d-to-r* al variare della distanza lungo la *mesh*. Si noti che gli artefatti numerici della simulazione DWM, particolarmente evidenti in prossimità dei bordi della *mesh*, potrebbero essere responsabili del leggero incremento del rapporto tra le energie intorno a 2 m.

Al fine di effettuare un confronto, il rapporto *d-to-r* è stato calcolato per un ambiente naturale mediante la seguente funzione proposta da Zahorik in [11]:

$$v = -3.64 \log_2(r) + 10.76 \quad (4)$$

Dal confronto tra le linee continua e tratteggiata in figura 9, emerge che i valori del rapporto *d-to-r* sono di gran lunga inferiori nella *mesh* bidimensionale rispetto all'ambiente naturale, e ciò significa che la riverberazione viene esagerata nell'ambiente virtuale.

Nel modello proposto non vengono considerati altri indizi uditivi di distanza, quali ad esempio la *Interaural Level Difference* (ILD) e lo spettro del segnale. Ciò trova motivazione nel relativamente scarso contributo alla percezione della distanza dato da tali indizi. Innanzitutto le ILDs, che sono

dovute alla differenza di intensità del segnale ricevuto dalle due orecchie, non sono di interesse per il caso qui considerato in quanto, sul piano mediano, risultano essere nulle e non forniscono quindi alcuna informazione di distanza per sorgenti sonore in tale posizione [9]. Per quanto riguarda l'attenuazione delle alte frequenze dovuta alla propagazione delle onde nell'aria, esse occorrono per distanze molto elevate (superiori a 15 m), tali da essere attualmente al di fuori dalla portata<sup>6</sup> della *mesh* qui considerata. In definitiva, nel caso presente risulta che le uniche informazioni di interesse per la valutazione della distanza di una sorgente sonora nella *mesh* sono l'intensità e, ancor di più, il rapporto *d-to-r*.

## CONCLUSIONE

L'applicazione è attualmente in fase prototipale e sono stati programmati alcuni esperimenti al fine di validare il modello di spazializzazione e fornire nuove informazioni per l'analisi della percezione uditiva della distanza.

Le prestazioni dell'utente durante l'azione di gioco permettono di valutare sia il coinvolgimento che gli aspetti di "giocabilità", i quali dipendono fortemente dalla qualità della mappatura del gesto (nel nostro caso, l'atto di lanciare una sfera virtuale). In altre parole, la mappatura del segnale di accelerazione (acquisito dal controller *Wii Remote*) sui parametri di controllo dell'ambiente dinamico (piano inclinato) e del modello di rotolamento deve risultare naturale e realistica per l'utente. Di più, l'utente deve avere la sensazione di interagire in modo efficace e coerente con l'ambiente virtuale fornito dal gioco. L'attuale mappatura è il risultato di uno sviluppo iterativo e della valutazione informale del prototipo. Al fine di effettuare una valutazione più precisa e eventualmente migliorare la mappatura verranno svolti ulteriori esperimenti.

Il principale problema posto dall'applicazione consiste nella possibilità per l'utente di percepire la posizione dell'estremità opposta del tubo virtuale, la qual cosa è fondamentale per evitare di lanciare la sfera troppo lontano. La capacità umana di percepire uditivamente la distanza è relativamente scarsa se confrontata con la percezione della direzione [12]. Una delle questioni problematiche relative agli esperimenti sulla percezione della distanza riguarda la difficoltà con cui i soggetti sono in grado di distinguere gli indizi uditivi di distanza dalle caratteristiche spettrali del suono stesso. Ciò fa credere che gli utenti del gioco potrebbero avere bisogno di alcune informazioni a priori riguardo alla sorgente sonora, o di una fase preliminare di training. Diversi studi sono concordi nel riscontrare una sistematica deviazione tra la distanza percepita e la distanza fisica della sorgente sonora, così come nel sottolineare una alta variabilità nelle risposte dei soggetti. Inoltre, la percezione della distanza sembra essere estremamente scarsa in mancanza di riverberazione. Uno studio recente [9] ha infatti mostrato che in ambienti riverberanti la percezione della distanza risulta essere più accurata. *DepThrow* offre quindi un potenziale strumento per la validazione di questa affermazione.

<sup>6</sup>A causa dei limiti computazionali dell'hardware moderno.

Come nota finale, si segnalano di seguito alcuni articoli di interesse per il presente lavoro. In [10] si illustra uno studio sugli indizi percettivi sfruttati dagli utenti per localizzare la posizione di una sfera virtuale che rotola o scorre all'interno di un tubo. Gli autori hanno riscontrato la capacità relativamente buona degli utenti di stimare la lunghezza del tubo semplicemente inclinandolo e percependo (attraverso indizi uditivi e/o tattili) la dinamica dell'oggetto virtuale. Lo stesso modello fisico di rotolamento qui adottato è stato precedentemente utilizzato in [7] per generare un *feedback* uditivo continuo che permettesse di mantenere in equilibrio una sfera virtuale rotolante lungo una barra inclinabile. Nella dinamica degli esperimenti effettuati, i soggetti dovevano far scorrere la sfera virtuale fino ad una posizione specifica lungo la barra. Ne è emerso che, anche nel caso in cui non fosse stato fornito alcun *feedback* visivo, tutti i soggetti sono stati in grado di svolgere il compito loro assegnato.

## RIFERIMENTI

1. J. M. Adrien. The missing link: Modal synthesis. In G. De Poli, A. Piccialli, and C. Roads, editors, *Representations of Musical Signals*, pages 269–297. MIT Press, Cambridge, MA, 1991.
2. F. Fontana. *Physics-based models for the acoustic representation of space in virtual environments*. PhD thesis, University of Verona, Italy, 2003.
3. F. Fontana and D. Rocchesso. A physics-based approach to the presentation of acoustic depth. In *Proc. of the 2003 International Conference on Auditory Display*, pages 79–82, July 2003.
4. D. W. Marhefka and D. E. Orin. A compliant contact model with nonlinear damping for simulation of robotic systems. *IEEE trans. on systems, man, and cybernetics - Part A: systems and humans*, 29(6):566–572, November 1999.
5. H. Møller, M. Sørensen, D. Hammershøi, and C. Jensen. Head-related transfer functions of human subjects. *J. Audio Eng. Soc.*, 43:300–321, 1995.
6. M. Rath. An expressive real-time sound model of rolling. In *Proc. of the Int. Conference on Digital Audio Effects (DAFx-03), London, UK, September 2003*.
7. M. Rath and D. Rocchesso. Continuous sonic feedback from a rolling ball. *IEEE MultiMedia*, 12(2):60–69, 2005.
8. D. Rocchesso and F. Fontana, editors. *The Sounding Object*. Mondo Estremo, 2003. Available at <http://www.soundobject.org/>.
9. B. Shinn-Cunningham. Learning reverberation: Considerations for spatial auditory displays. In *Proc. of the 2000 International Conference on Auditory Displays*, 2000.
10. H.-Y. Yao and V. Hayward. An experiment on length perception with a virtual rolling stone. In *Proc. of the EuroHaptics Int. Conf., Paris, France, pages 275–278, 2006*.
11. P. Zahorik. Direct-to-reverberant energy ratio sensitivity. *J. Acoust. Soc. Am.*, 112(5):2110–2117, November 2002.
12. P. Zahorik, D. Brungart, and A. Bronkhorst. Auditory distance perception in humans: A summary of past and present research. *Acta Acustica united with Acustica*, 91:409–420, 2005.

# CHORALE

## Compositional Hyper ORganized ALgorithmic Environment

Gianni Della Vittoria  
 Laboratorio Elettronico per la  
 Musica Sperimentale (L.E.M.S.)  
 Conservatorio di Pesaro

### SOMMARIO

Chorale è un ambiente di lavoro in linguaggio python che trasforma algoritmi compositivi strutturati su più livelli in un file audio attraverso csound. L'utente può costruire algoritmi del tutto personalizzati per gestire qualsiasi strumento csound, anche preesistente, pilotando sia algoritmicamente che manualmente ogni parametro. L'organizzazione algoritmica su quattro livelli consente di controllare la complessità della composizione musicale sia nel dettaglio che nel disegno generale allo stesso tempo.

### Parole Chiave

Composizione algoritmica; python; csound

### COSA FA

Partiamo da un esempio.

Supponiamo di voler creare un pezzo di musica con un solo semplice strumento, un oscillatore sinusoidale, che funzioni attraverso la definizione di due parametri: ampiezza e frequenza (chiameremo questo strumento "player").

Supponiamo ora di voler creare un suono attraverso la consueta tecnica additiva, cioè attivando simultaneamente più istanze dello stesso oscillatore con i giusti valori di ampiezza e frequenza. Chiamiamo "vox" ogni istanza del player oscillatore. Così, per avere un suono di 5 secondi, poniamo, con certe caratteristiche, dovremo decidere quante voci sovrapposte utilizzare e quali rapporti di ampiezza e frequenza ci devono essere tra fondamentale e parziali. Chiameremo questo complesso col nome di "grain" (fig. 2). Attraverso csound è possibile fare tutto ciò in poche righe: si definisce l'oscillatore nell'orchestra e si danno i valori nello score.

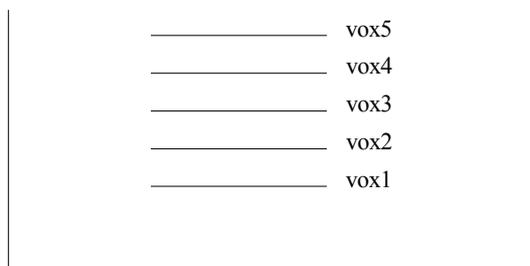


Figura 2. Grain

A questo punto vogliamo aggiungere un secondo suono, che parta di seguito al primo. Procederemo nel modo che abbiamo appena descritto, e questa volta, oltre ai rapporti interni di frequenza e ampiezza, occorre decidere anche i rapporti tra le fondamentali dei due grain. In csound si tratterà

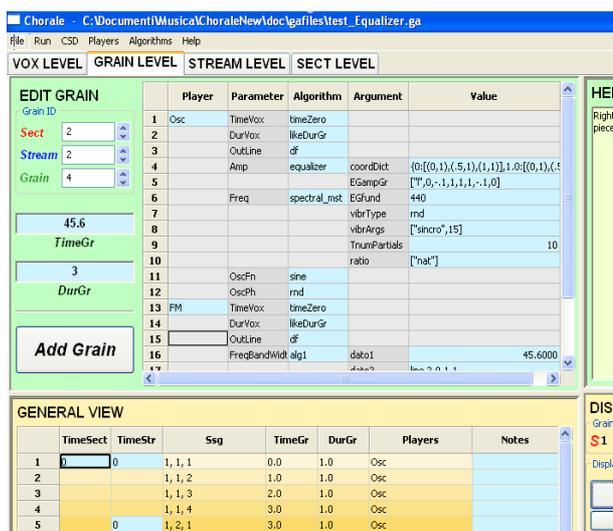


Figura 1. Dettaglio di Chorale: finestra Edit Grain

semplicemente di aggiungere i nuovi valori sincronizzando le entrate delle voci del secondo grain sul giusto tempo di attivazione.

Potremmo proseguire così per tanti altri grain: l'unica difficoltà sarà che lo score di csound presto giungerà a dimensioni notevoli, talvolta difficili da controllare.

Pensiamo infatti a che cosa succederebbe se, oltre a modificare la fondamentale di ogni grain, volessimo anche intervenire sulla ratio dei rapporti tra le frequenze interne al grain in maniera tale che essa cambi secondo una logica generale e che anche la relazione interna dei singoli pesi d'ampiezza di ogni parziale sia soggetta ad un piano globale che attraverso l'intera serie di grain. Molto probabilmente perderemo presto la pazienza e magari rinunceremo al disegno compositivo che ci eravamo prefissati per accontentarci di un calcolo con meno variabili.

C'è però una possibilità. Se è vero che abbiamo in mente un piano compositivo generale, la cui complessità risiede più nella realizzazione che non nella sua ideazione, possiamo contare sul fatto che un programma può fare il "lavoro sporco" al posto nostro. Esistono già tanti programmi che fanno questo, magari proprio creando partiture per csound, ma molti di essi costringono ad applicare una delle vie che sono già state previste dall'autore del programma, lasciando poche speranze a che l'algoritmo ricalchi ciò che volevamo.

Chorale, invece, fa proprio questo: consente all'utente di definire esattamente la propria logica compositiva attraverso cui pilotare ogni parametro.

L'utente ha così la possibilità di crearsi non solo il proprio suono, ma anche il proprio personale percorso compositivo,

attraverso la creazione degli “algoritmi”. Naturalmente Chorale contiene dentro di sé già diversi algoritmi, per dare la possibilità di lavorare subito senza troppe conoscenze, ma anche per mostrare come produrne altri per imitazione in linguaggio python.

Riprendiamo il nostro esempio. Abbiamo detto che con Chorale possiamo ottenere una serie di grain (che definiremo “stream”, fig. 3) impostando solamente alcuni algoritmi che si incarichino di definire i parametri. Avremo quindi un algoritmo per decidere la serie delle fondamentali, uno per definire come varia la ratio tra fondamentale e parziali di ciascun grain, uno per definire i tempi d'entrata del grain, durate, e così via. L'interfaccia grafica carica tutti gli algoritmi di default con i rispettivi dati, cosicché l'utente modifica all'istante solo ciò che gli interessa.

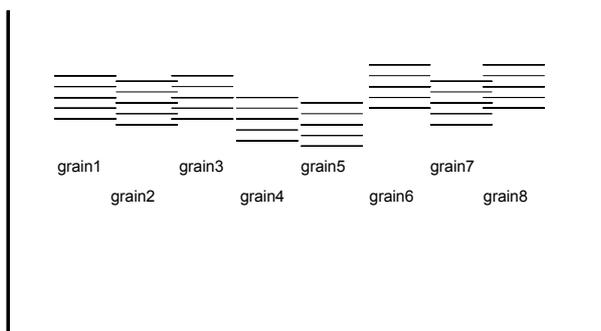


Figura 3. Stream

Ora vogliamo aggiungere un secondo stream, simile a quello che abbiamo appena terminato, ma con la particolarità di avere suoni in relazione col primo stream già calcolato. Chorale tiene in memoria i dati già calcolati proprio per servirsene per eventuali interdipendenze.

Tanti stream si possono riunire in una “sect” (fig. 4), il livello più alto di Chorale, e tante sect compongono il pezzo di musica.

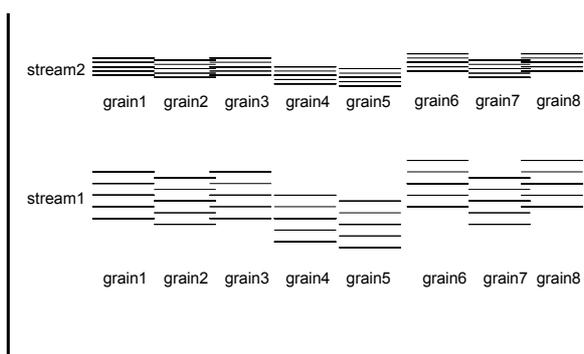


Figura 4. Sect

Anche la sect consente la definizione di algoritmi. Questi hanno la funzione di scegliere gli algoritmi del livello stream e di fornire loro i rispettivi dati.

Ecco la particolarità di Chorale: “creare” algoritmi di algoritmi. Questa piramide gerarchica di algoritmi consente quindi di scrivere un pezzo articolatissimo, dove ogni algoritmo sceglie l'algoritmo (e gli fornisce i dati) del livello immediatamente inferiore.

## COME LAVORA CHORALE

Chorale è suddiviso nei 4 livelli vox, grain, stream, sect. L'utente può lavorare a qualunque livello per l'inserimento dei dati e la scelta dei processi algoritmici.

Analizziamo la struttura algoritmica del player Oscillator. Esso prevede i seguenti sette parametri:

- |            |                                  |
|------------|----------------------------------|
| 1. Time    | (tempo d'attacco)                |
| 2. Dur     | (durata)                         |
| 3. OutLine | (pista su cui versare la vox)    |
| 4. Amp     | (ampiezza)                       |
| 5. Freq    | (frequenza)                      |
| 6. OscFn   | (tipo di onda)                   |
| 7. OscPh   | (fase iniziale dell'oscillatore) |

I primi tre parametri sono necessari a quasi tutti i players, gli altri quattro sono tipici dell'oscillatore. La pista di OutLine raccoglie tutte le voci che dovranno poi essere processate da altri player. Se, per esempio, alcune voci dovranno subire un riverbero e le altre no, sarà sufficiente versare le voci su due OutLine diverse, così il player Reverb processerà le prime attraverso la sua InLine escludendo le seconde.

Dovendo costruire una sola vox possiamo decidere di inserire direttamente i valori per ogni parametro a livello vox, ma per fare un insieme di voci possiamo ricorrere agli algoritmi del grain level. Per ogni parametro un algoritmo del grain level si incaricherà di calcolare i rispettivi valori del vox level. Allo stesso modo accadrà che gli algoritmi del livello stream calcoleranno i dati per il livello grain, e quelli del livello sect per il livello stream.

Nel diagramma di figura 5 viene rappresentata la struttura algoritmica applicata al player Oscillator. La freccia continua indica il lavoro degli algoritmi: l'output del livello superiore viene ad essere l'input del livello immediatamente inferiore, mentre la freccia tratteggiata rappresenta la conversione in dati csound.

Per quanto riguarda il tempo, la timeline viene elaborata come qualsiasi altro parametro attraverso specifici algoritmi. Timevox si riferisce al tempo di attivazione del player relativamente al proprio grain: timevox = 1 significa che partirà 1 secondo dopo l'inizio del grain cui appartiene. A questo punto l'algoritmo TimeVox a livello grain può servire proprio per sfasare in modo più o meno ordinato l'ingresso di ogni vox nel grain. L'algoritmo TimeVox a livello stream sarà utile per impostare il suddetto algoritmo TimeVox a livello grain ogni volta diverso per ciascun grain. TimeVox a livello sect potrà fare lo stesso nei confronti del livello stream.

La medesima cosa vale per TimeGr (relativo a inizio stream) e TimeStr (relativo a inizio sect).

## COME SONO ORGANIZZATI GLI ALGORITMI

L'ordine di esecuzione degli algoritmi avviene rispettando i seguenti criteri di precedenza:

1. Prima tutti gli algoritmi di livello superiore, poi quelli di livello immediatamente inferiore e così via fino al livello vox, ricordando che l'utente può cominciare, o intervenire, a qualsiasi livello.
2. All'interno di uno stesso livello si terrà conto delle reciproche necessità di precedenza, o dependencies.

Per quanto riguarda il secondo punto, la flessibilità di Chorale si manifesta anche in questo importante aspetto: non esiste un ordine predeterminato di esecuzione tra i diversi parametri, esso viene di volta in volta fissato secondo le particolari

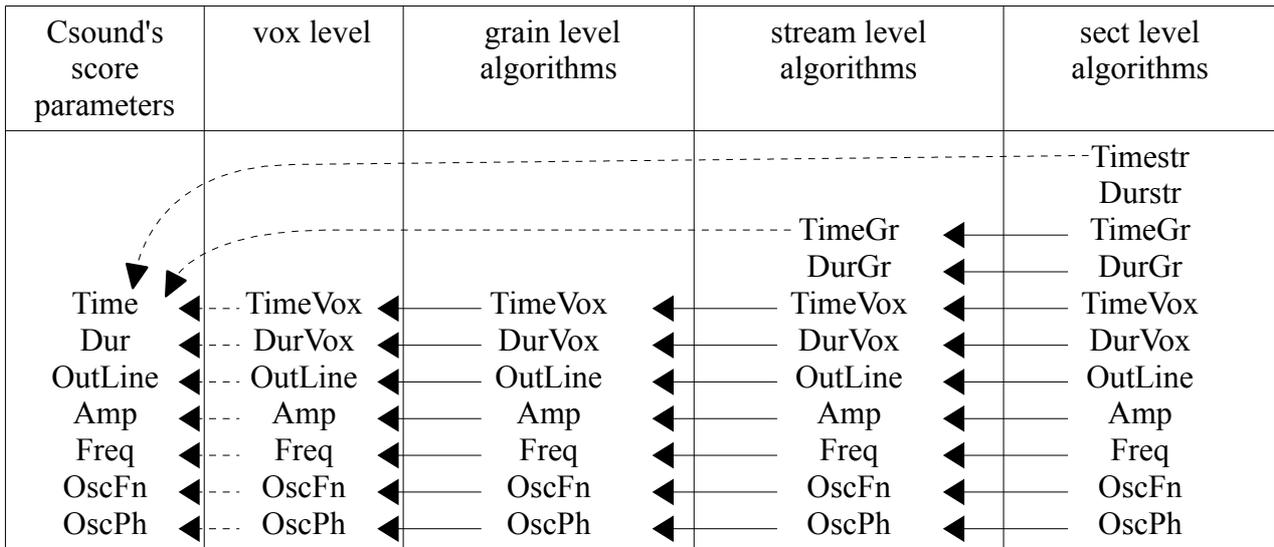


Figura 5. Struttura algoritmica applicata al player Oscillator

esigenze del momento. Per esempio, se in uno stream i grain dispari dovessero calcolare la frequenza in base alla durata, e i grain pari, per particolari esigenze, la durata in base alla frequenza, basterà semplicemente creare questi due algoritmi grain level: uno per la frequenza con dipendenza dalla durata, uno per la durata con dipendenza dalla frequenza. A livello stream ci saranno i criteri per la scelta alternata tra i due. Ogni algoritmo è pertanto caratterizzato dalle dependencies, ovvero quei parametri i cui dati devono essere già stati calcolati perché necessari all'esecuzione dell'algoritmo stesso. In base alle dependencies Chorale calcolerà automaticamente ogni volta l'ordine corretto di esecuzione degli algoritmi, senza che l'utente debba preoccuparsi di nulla.

### ENVELOPE GENERATORS

Uno dei limiti di molti software che creano partiture per csound è che, una volta definito lo strumento dell'orchestra, i parametri p-fields rappresentano delle costanti. I generatori d'involuppo di Chorale, invece, consentono di definire ciascun parametro del player attraverso una funzione, anziché un valore costante.

Consideriamo il solito oscillatore a livello vox. Per parametri quali l'ampiezza, la frequenza, o altro, è spesso utile poter manovrare gli strumenti modificandone il comportamento nel tempo. Se volessimo fare un glissato esponenziale da 500Hz a 1000Hz nella prima metà della nota, e da 1000Hz a 100Hz nella restante parte, sarà sufficiente scrivere:

```
[ 'e', 500, .5, 1000, .5, 100 ]
```

Per ottenere la stessa cosa in csound dovremmo scrivere:

```
kenv linseg 500, p3/2, 1000, p3/2, 0
```

Come si vede, il valore .5 indica automaticamente la metà di p3, la durata della nota; se si desiderasse esprimere un valore assoluto in secondi basterà utilizzare valori negativi. E' anche possibile servirsi contemporaneamente di valori assoluti e relativi nello stesso opcode: i valori relativi verranno calcolati non su p3, ma su p3-(somma di tutti i valori assoluti).

Un evidente vantaggio degli EG sta nel fatto che non si è costretti ad usare un numero prestabilito di elementi. Pensiamo infatti a come è costruito un oscillatore in csound:

```
a1 oscili kamp,kfreq,ifn,iph
```

Ora, volendo variare kamp con spezzate lineari dovremo far precedere una riga come questa:

```
kamp linseg p4,p5,p6,p7,p8
```

per ottenere una spezzata a tre vertici. Ma se ci servissero più vertici, 100 ad esempio? Evitando di scrivere un nuovo strumento ogni volta che cambiamo, potremmo aggiungere una lunga sfilza di p-fields, rendendo però scomodo azzerarli tutti quando non servono. Una buona alternativa sarebbe ricorrere a un controllo da tabella in partitura, con l'incomodo di creare uno spaventoso numero di tabelle e la poca flessibilità della sintassi.

Con gli EG questo problema è scavalcato. Un player di Chorale è costruito in puro linguaggio csound, ma, grazie ad una speciale orchestra che fornisce appunto l'immediata utilizzazione degli EG, possiamo tranquillamente servirci dello stesso player con piena flessibilità di numero di elementi e di tipo di involuppo (è possibile alternare spezzate lineari, random controllati, operazioni tra gli stessi EG con la massima libertà). Ecco quelli più utilizzati:

- float = costante
- 'l' = spezzate lineari
- 'e' = segmenti esponenziali
- 'u' = interpolazione cubica
- 't' = transeg di csound
- 'o' = oscillatore
- 'f' = lettura di tabella e distorsione non lineare
- 's' = somma di involuppi
- 'p' = prodotto di involuppi
- 'r' = random controllato

Non c'è limite al numero di segmenti nell'involuppo e con le operazioni tra involuppi si possono deformare quelli esistenti nei modi più disparati. Grazie alla sintassi degli EG gli stessi algoritmi sono in grado di produrre forme di involuppo molto elaborate.

## SUL CONCETTO DI GRAIN

Abbiamo inizialmente affermato che il grain è un insieme di voci e che la vox è l'istanza di un player. Il grain dell'esempio iniziale realizza una normalissima sintesi additiva di voci dello stesso tipo, cioè del player Oscillator, ma un grain può essere composto anche da voci istanziate da player diversi. Tanto per fare un esempio, si potrebbe creare un grain con un timbro tempovariante grazie al contributo studiatamente sfasato di diversi tipi di player. Questi player possono essere dei generatori di suono, ma anche processori di suono applicati esclusivamente al grain cui appartengono (fig.6).

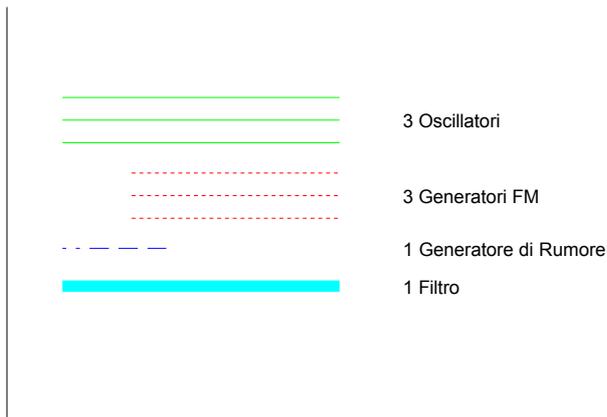


Figura 6. Grain composto da 8 voci di 4 diversi player

Il fatto che il grain dell'esempio possieda, accanto ai diversi generatori indipendenti, anche un filtro, cioè un processore della sintesi degli altri player, fa del grain un'unità complessa che va al di là del semplice agglomerato additivo. In questo modo infatti un secondo grain simile e sovrapposto al precedente potrebbe avere valori di filtraggio diversi dal primo, senza timore di interferenze. Gestendo la cosa attraverso gli algoritmi del livello stream non sarà difficile ottenere una estesa serie di grain con sottili variazioni.

La stessa durata del grain non ha limiti di definizione. Potrebbe essere ridottissima, come per un grano nella sintesi granulare, o lunga come un'intera sezione: dipende da come l'utente preferisce interpretare il grain, e, di conseguenza, lo stream e la sect.

## CONCLUSIONI

Lo scopo di Chorale è quello di consentire una più approfondita riflessione sui processi compositivi e non unicamente sulla sintesi del suono o sul missaggio, in modo che l'ideazione compositiva si articoli con nuovi mezzi. L'utente può creare gli algoritmi che desidera a tutti i livelli o sfruttare quelli già presenti. Può anche creare tutti i players che vuole o sfruttare i propri vecchi strumenti csound, ora con la potenza e la flessibilità degli algoritmi di Chorale.

La possibilità di coordinare l'attenzione ai particolari simultaneamente al piano complessivo dell'opera è al centro dell'approccio compositivo a cui aspira Chorale.

## RIFERIMENTI

1. D. Cope, *The algorithmic composer*, A-R Editions, Madison, Wis., 2000
2. D.R. Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid*, Basic Books, New York, 1979
3. R. Bianchini, A. Cipriani, *Il suono virtuale*, ConTempo, Roma, 2001

GENERAL VIEW							
	TimeSect	TimeStr	Ssg	TimeGr	DurGr	Players	Notes
1	0		1, 1, 1	0.0	1.0	Osc	
2			1, 1, 2	1.0	1.0	Osc	
3			1, 1, 3	2.0	1.0	Osc	
4			1, 1, 4	3.0	1.0	Osc	
5	0		1, 2, 1	3.0	1.0	Osc	
6			1, 2, 2	3.0	1.0	Osc	
7			1, 2, 3	3.0	1.0	Osc	
8			1, 2, 4	3.0	1.0	Osc	
9			1, 2, 5	3.0	1.0	Osc	
10	0		1, 3, 1	3.0	1.0	Osc	
11			1, 3, 2	3.0	1.0	Osc	
12	0	0	2, 1, 1	3.0	1.0	Osc	
13			2, 1, 2	3.0	1.0	Osc	
14			2, 1, 3	3.0	1.0	Osc	
15	0		2, 2, 1	3.0	1.0	Osc	

Figura 7. Finestra General View in cui sono riportati tutti gli eventi creati

DISPLAY		TimeSect 0	TimeGr 0.0	
Grain ID		TimeStr 0	DurGr 1.0	
S1	s1	g1		
Player	Parameter	Algorithm	Argument	Value
1	Osc	TimeVox	bimeZero	
2		DurVox	likeDurGr	
3		OutLine	df	
4		Amp	equalizer	coordDict {0:{{(0,0),(5,1),(1,0)},1,C
5			EGampGr	[T,0,-1,1,1,-1,0]
6		Freq	spectral_met	EGfund 100
7			vibrType	no
8			vibrArgs	
9			TrunPartial	15
10			rabo	["nat"]
11	OscFn	sine		
12	OscPh	rmd		

Figura 8. Finestra Display in cui si descrivono tutti i parametri del grain selezionato

# SUONI SCOLPITI E SCULTURE SONORE: alcuni esempi di installazioni d'arte elettroacustica

Silvia Lanzalone

CRM - Centro Ricerche Musicali, Roma  
silvia.lanzalone@crm-music.it

## SOMMARIO

Il presente scritto descrive alcune caratteristiche tecniche, linguistiche ed espressive di tre installazioni d'arte elettroacustica da me realizzate, presentando così i risultati di una ricerca artistica "in itinere" il cui principio estetico comune fa proprio il concetto di intermedialità.

*Il suono incausato*, improvvisa-azione per clarinetto sospeso clarinetista ed elettronica (2005), *Voce*, per piccolo coro di vetro, voce femminile ed elettronica (2006) e *Risonanze dalla Terra*, sculture sonore in terracotta, canapa e suoni elettronici (2007), sono lavori di arte elettroacustica in cui il mezzo tecnologico, anziché sovrapporsi al suono naturale per trasformarlo con artifici algoritmici, prende parte all'intera struttura fisico-acustica di un nuovo "strumento", concepito e progettato insieme alla composizione musicale in modo da dare vita ogni volta ad uno specifico ed autonomo oggetto d'arte, sonora e plastica insieme.

## Parole Chiave

Installazioni d'arte elettroacustica; Arte intermediale; Estemporaneità e teatralità del gesto esecutivo.

## IL SUONO INCAUSATO (2005)

*Il suono incausato* propone una singolare installazione intorno al clarinetto, in cui il rapporto con lo strumento non è più perseguito nella tipica concatenazione logico-temporale per cui il suono (effetto) è ottenuto tramite eccitazione della colonna d'aria presente nel tubo attraverso l'ancia da parte dell'esecutore (causa). Il nuovo strumento, denominato <<clarinetto sospeso>>, è acusticamente autonomo poiché, anziché essere "suonato" dall'interprete, viene da lui "esplorato" attraverso una serie di interventi atti soprattutto a modificare le caratteristiche fisiche del tubo e le combinazioni di fori aperti o chiusi, creando così nuove condizioni del sistema vibrazionale. In questo brano il clarinetto è posto su di un sostegno ed appare come sospeso, distante dallo strumentista, nonché fisicamente modificato. Al clarinetto sospeso sono infatti aggiunti: un altoparlante, tappi o leve di sughero da applicare rispettivamente nei fori o sotto le chiavi, tronchi di cono di diverse grandezze e materiali da inserire nei fori, nella campana o nel barilotto. Le manipolazioni effettuate sul clarinetto sospeso, da me definite "esplorazioni", possono amplificare o attenuare alcune specifiche risonanze, possono introdurre dei ritmi timbrici in sovrapposizione al suono di base, possono agire per imitazione o per contrasto rispetto ad una particolare modalità di articolazione del suono preesistente, possono non produrre alcun effetto, coerentemente con le molteplici tipologie di connessione tra azione ed evento. Le esplorazioni sono effettuate secondo un criterio di estemporaneità, attraverso azioni che mettono in rilievo una componente gestuale e teatrale dell'esecuzione.

*Il suono incausato* è dedicato a Massimo Munari, interprete del brano, ed è stato eseguito in diversi festivals di musica contemporanea. L'opera ha vinto il Premio Internazionale di Composizione Franco Evangelisti 2006 ed è pubblicata dalla casa editrice Suvini Zerboni.



Figura 1. Silvia Lanzalone, *Il suono incausato*: clarinetto sospeso.

Nessun suono può essere realmente "incausato" dal punto di vista fisico; tuttavia nel clarinetto sospeso i termini della relazione strumento/esecutore/suono trovano una diversa relazione consequenziale. L'altoparlante non costituisce parte visibile dello strumento: scenicamente il clarinetto è "solo" sul palco e l'interprete vi si pone in relazione solamente dopo che lo strumento emette dei suoni. L'installazione dà luogo ad un meta-strumento, risultato di un orientamento estetico che considera la tecnologia elemento da integrare allo strumento classico, che viene così tecnicamente modificato, in modo che ne risultino ampliate le possibilità artistiche ed espressive.

Il clarinetto sospeso prevede in totale: sette tappi di sughero che hanno la funzione di chiudere i fori e tre leve dello stesso materiale che servono a tenere aperte o chiuse le chiavi, da disporre secondo la numerazione fornita in partitura; quattro tronchi di cono, di cui tre in cartone e uno in alluminio, che hanno la funzione di amplificare il suono e di esaltarne alcuni armonici in determinate zone di irradiazione dello strumento, quali la campana oppure alcuni fori. La caratteristica fondamentale dello strumento così configurato è l'assenza di imboccatura e quindi di ancia. Al tubo, all'altezza del barilotto, è applicato un sistema composto da un altoparlante inserito in un tronco di cono di piombo, materiale scelto perché acusticamente isolante, ma non fonoassorbente. Il cono, chiuso all'estremità più larga ed inserito con l'altra estremità aperta nel tubo del clarinetto come una sorta di imbuto, svolge la funzione di convogliare l'aria ed il suono all'interno dello strumento. L'altoparlante infatti, durante la riproduzione di suoni ad elevata energia sonora, trasmette tramite il movimento meccanico della membrana una significativa vibrazione alla colonna d'aria presente nel tubo del clarinetto, provocando una risonanza aggiuntiva. La combinazione di ancia battente e tubo di forma cilindrica fa sì che il corpo del clarinetto sia paragonabile ad un tubo chiuso, ma la presenza dell'ancia sposta il primo punto nodale ad 1/14 della lunghezza del tubo. Nel clarinetto sospeso la sostituzione dell'ancia con il cono di piombo introduce una deformata che agisce sulla lunghezza del tubo e, probabilmente, sulla posizione del nodo. La qualità timbrica dei suoni che provengono dal clarinetto sospeso non sempre rispecchia ciò che normalmente si è abituati a sentire dallo strumento tradizionale. All'inizio del brano, per esempio, un tipico suono di clarinetto si evolve nella direzione di un suono/rumore simile ad un timbro elettronico, seppure con "l'impronta spettrale" propria della caratteristica struttura ad armoniche dispari dovuta alle risonanze che lo strumento impone al suono in esso immerso. La trasformazione del timbro nel clarinetto sospeso è frutto dell'integrazione tra parte elettronica e parte meccanica: il clarinetto impone al suono diffuso dall'altoparlante un contributo conforme alle proprie caratteristiche fisico-acustiche, che sono modificabili attraverso la chiusura o

l'apertura dei fori, o tramite l'aggiunta dei coni. Un possibile intervento è infatti costituito dall'inserimento di un cono in uno dei fori dello strumento: la maggiore efficacia si ottiene inserendo il cono nel foro dove l'aria, in base al suono emesso, esce con più energia e da cui viene irradiata una parte significativa dello spettro acustico. E' noto che i fori del clarinetto sono caratterizzati da diversa dimensione e spaziatura e che il tubo risponde in maniera differente alle diverse frequenze di eccitazione. Il suono è irradiato dai fori, la componente più grave attraverso il primo foro aperto, i successivi cinque o sei armonici dal secondo e terzo foro aperto, la restante parte attraverso i fori successivi. Fori di grande diametro sono comparabili all'apertura finale del tubo dello strumento, che quindi risulta accorciato quando sono aperti; fori di diametro medio riducono il tubo in modo parziale e fori molto piccoli hanno la funzione di esaltare i più alti modi di vibrazione. Come regola generale per la definizione del timbro, la parte di tubo inferiore al primo foro aperto agisce come filtro passa-alto, poiché l'energia delle frequenze acute si disperde dai fori successivi, mentre il primo foro aperto irradia la frequenza di taglio; a frequenze inferiori ad essa vi è una forte attenuazione (è noto che la frequenza di taglio può essere modificata dai costruttori di strumenti, ma secondo schemi complessi e tenendo conto anche delle distanze dei fori oltre che del loro diametro: generalizzando si può dire che, per un determinato spettro acustico, aumentando il diametro dei fori si ottiene un timbro "chiaro" in quanto si alza la frequenza di taglio, mentre restringendo il diametro si ottiene una frequenza di taglio più grave e quindi un timbro più "scuro"). Relativamente al cono di cartone (è evidente che il cono deve essere aperto ad entrambe le estremità e che ad essere inserita nel foro è l'estremità di diametro più piccolo, la quale deve avere una misura adeguata al diametro del foro in questione): se, ad esempio, un foro irradia una frequenza grave dello spettro di un dato timbro (non necessariamente corrispondente al "pitch" percepito), l'inserimento del cono di cartone all'interno di tale foro riuscirà molto probabilmente ad esaltarla, agendo a sua volta come un filtro passa-banda, mentre l'inserimento del cono all'interno di un foro diverso probabilmente tenderà ad amplificare una o più frequenze acute dello stesso spettro, mettendole in evidenza. Il cono è infatti un componente della famiglia delle cosiddette "trombe di Bessel" ed insieme al cilindro è utilizzato normalmente come forma di anima per la costruzione di strumenti a fiato; presenta tutte le frequenze di risonanza a rapporto intero rispetto al fondamentale e quindi, nel caso sperimentato sul clarinetto sospeso, agisce come un ulteriore risonatore applicato allo strumento. Inoltre, il timbro di uno strumento a fiato è influenzato fortemente dalla forma dell'anima interna del tubo, nonché dalla presenza di aperture o chiusure alle estremità di esso, più che dal materiale di cui è composto; per cui, se si considera che le cavità interne conseguenti alla creazione dei fori nel tubo introducono già delle complicazioni alla struttura di per sé semplice della forma di Bessel utilizzata e che sicuramente tali variazioni contribuiscono alla configurazione del suono, applicare un cono ad un foro del clarinetto complica ulteriormente la forma dell'anima contribuendo a significative variazioni sulla conformazione del timbro. L'inserimento di un cono di cartone di lunghezza di circa 30 cm nella estremità aperta del tubo secondo il criterio accennato prima, provoca invece un consistente allungamento del tubo stesso (da 67 cm a circa 97 cm), nonché una modifica sostanziale della forma conica data dalla tradizionale campana del clarinetto che normalmente agisce come filtro passa-alto per le note gravi del registro sostituendo i fori mancanti e garantendo una buona omogeneità timbrica rispetto alle note dei registri più acuti. Il cono inserito nella campana, sostituendosi ad essa, ne modifica quindi la forma, agendo sul timbro naturale del clarinetto. Per rendere efficace il cono

aggiunto è necessario chiudere tutti i fori del tubo e agire sull'apertura/chiusura degli ultimi fori vicini alla campana, ottenendo così delle variazioni timbriche significative. Un altro caso di intervento effettuabile sulla campana (ve ne sono molti su tutto lo strumento, ma ne elenco solo alcuni) è costituito dalla sua chiusura dinamica e variata ottenibile con il palmo della mano, in modo da realizzare un effetto tipo sordina "wha-wha" sui suoni che hanno una forte irradiazione proprio in quella zona dello strumento (la sordina, effettuando delle variazioni alla forma del padiglione, agisce sul timbro dello strumento senza modificarne l'accordatura, che sostanzialmente dipende solo dai primi tre quarti della lunghezza complessiva del tubo). Alcuni di questi interventi sono più efficaci quando il suono e l'aria convogliati nel clarinetto sospeso hanno una forte dinamica ed una certa quantità di energia, in quanto viene messo in vibrazione un maggior numero di frequenze e le componenti armoniche acute dello spettro risultano più significative. Anche per questa ragione è stata necessaria una lunga fase di ottimizzazione del sistema fino a raggiungere il giusto equilibrio tra l'intensità sonora, l'energia dell'aria immessa dall'altoparlante ed una determinata configurazione fisico-acustica dello strumento.

L'opera *Il suono incausato* è stata da me definita un'improvvisazione, nell'intento di sottolineare il carattere estemporaneo e contemporaneamente l'impostazione teatrale dell'esecuzione. Il brano, senza la componente visiva e teatrale, non può essere compreso nella sua interezza ed il clarinetista svolge un ruolo fondamentale ai fini della gestione del materiale sonoro e del flusso temporale degli eventi. L'interprete interagisce con lo strumento secondo una modalità improvvisativa circoscritta a tre livelli di intervento: le variazioni microstrutturali del materiale sonoro all'interno delle unità formali; la temporizzazione, ossia la scelta della durata di ogni unità formale; la gestione delle azioni di carattere teatrale. L'estemporaneità del brano facilita una più viva interazione tra il clarinetto sospeso ed il clarinetista, poiché il suono incausato del clarinetto sospeso suggerisce modi e tempi di azione sulle configurazioni dello strumento, e le libere azioni del clarinetista possono essere gestite con lo spirito dell'esplorazione e della scoperta.

Tutti i livelli di intervento sul clarinetto sono guidati da indicazioni scritte in partitura che stabiliscono gli ambiti entro i quali l'interprete deve agire; si può così parlare di improvvisazione all'interno di una griglia formale prestabilita. La partitura assume dunque il carattere di schema che fissa gli accadimenti più significativi ai fini della costruzione della forma e della connotazione del materiale; è suddivisa in 23 riquadri, ciascuno dei quali descrive una configurazione del clarinetto, la tipologia di comportamenti del clarinetista su quella configurazione, gli ambiti approssimativi di durata. La configurazione del clarinetto è rappresentata in partitura in modo puntuale; l'azione del clarinetista sullo strumento è descritta in modo da consentire margini di libero intervento dell'esecutore all'interno di ciascun riquadro.

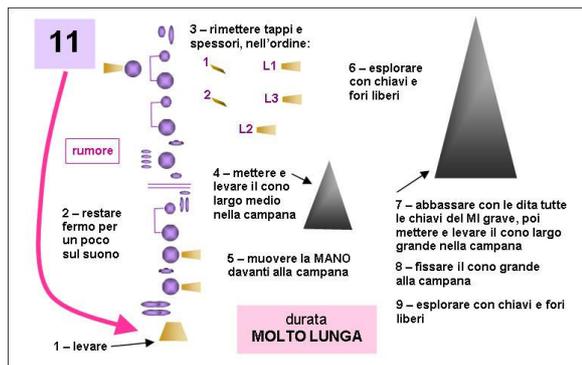


Figura 2. Silvia Lanzalone, *Il suono incausato*: partitura, riquadro n° 11.

## VOCE (2006)

La composizione *Voce* è un'opera caratterizzata da una componente sonora, visiva e teatrale, in cui il timbro della voce femminile è elaborato e poi simulato da uno strumento totalmente inventato – il «coro di vetro» – con cui l'interprete vocale entra in un rapporto di dialogo e di interazione.



Figura 3. Silvia Lanzalone, *Voce* (2006): coro di vetro.

Il coro di vetro è una struttura costituita da cavità risonanti di vetro, le cui diverse forme e dimensioni sono state scelte e raggruppate in modo da ottenere alcune risonanze tipiche delle formanti vocaliche. Le cinque vocali classiche A - E - I - O - U, insieme ad alcune delle rispettive varianti fonetiche, sono evocate da cinque elementi costituiti da risonatori di vetro bianco sostenuti da altrettanti piedistalli ad altezza d'uomo. L'insieme che ne deriva assume la consistenza acustica e visiva di un piccolo coro.

Nel corso del brano la voce femminile improvvisa su materiali costituiti da parole immaginarie, pronunciate e/o cantate in modo da mettere in evidenza tipologie di fonemi, suoni e modalità espressive secondo criteri di volta in volta diversi. Le parole, i fonemi ed i suoni emessi dalla cantante forniscono all'algoritmo di elaborazione del suono in tempo reale, materiale acustico per la generazione dei suoni del coro, la cui autonomia espressiva diviene progressivamente sempre più evidente fino ad assumere, verso la conclusione del brano, la forma di una conversazione caotica tra voci all'interno di un ipotetico sistema linguistico di totale invenzione. La modalità di scrittura dell'intero lavoro, relativamente ai livelli interni della struttura, è caratterizzata dall'inserimento di sezioni libere che sia l'interprete vocale, sia l'interprete informatico, devono eseguire con criteri improvvisativi. Ad un livello più alto invece il brano è definito con precisione nel suo percorso formale.

La composizione *Voce* è stata realizzata durante il biennio di II livello in Musica Elettronica del Conservatorio "Santa Cecilia" di Roma, docente Giorgio Nottoli, ed è stata eseguita in diversi festivals di musica contemporanea dalla voce di Angelina Yershova. Opera finalista al Concorso Internazionale di Composizione Franco Evangelisti 2006, *Voce* è stata eseguita presso la Sala Accademica del Conservatorio Statale di Musica di Santa Cecilia durante il 43° Festival di Nuova Consonanza, dalla cantante Silvia Schiavoni.

I risonatori sono cavità che, come è noto, hanno la funzione di trasformare il segnale acustico enfatizzando determinate regioni frequenziali, che acquistano così maggiore rilevanza di ampiezza rispetto ad altre componenti. Le frequenze centrali di tali regioni, dette frequenze formanti, sono totalmente dipendenti dalla forma e dal volume della cavità. Un risonatore di forma semplice, tipicamente una sfera con un collo che mette l'aria in comunicazione con l'esterno, genera una sola risonanza la cui frequenza è inversamente proporzionale al volume della cavità e alla lunghezza del collo, mentre è direttamente proporzionale alla grandezza dell'apertura. Il tratto vocale è un risonatore di notevole complessità la cui forma base è un tubo chiuso; da esso si genera una struttura spettrale ad armoniche dispari che viene poi modificata dalle varie cavità di risonanza dell'apparato vocale.

I risonatori del coro di vetro presentano un corpo circolare costituito da "colli" e rigonfiamenti di diametro variabile da 3 a 29.5 cm, simmetrici rispetto ad un'asse centrale la cui lunghezza varia da 15 a 40 cm. Alle estremità sono presenti due fori, uno dei quali, di grandezza differenziata tra i risonatori, costituisce la zona di principale irradiazione del suono; l'altro, di diametro di circa 10 cm, consente l'introduzione di un piccolo altoparlante che costituisce la fonte di emissione del segnale di eccitazione. I risonatori, in numero di 15, sono stati raggruppati in base alle relative frequenze di risonanza, in modo da costituire 5 gruppi di 3 elementi diversi per ciascun tipo di vocale: ognuno dei 3 elementi apporta il proprio contributo specifico alla definizione acustica della singola vocale. Tutti i risonatori hanno frequenze di risonanza che variano tra circa 250 Hz e circa 1500 Hz<sup>1</sup>, come indicato nella tabella riportata in figura 4: Ris indica il numero del risonatore<sup>2</sup> e la vocale associata, Ft indica la frequenza di centro banda e BW la larghezza di banda, in Hertz. In un timbro di voce sono facilmente riscontrabili circa sei formanti, ma sono sufficienti le prime due o tre formanti per rendere riconoscibili i timbri delle diverse vocali. Come si può notare osservando la tabella, le frequenze di risonanza dei risonatori assegnati alle vocali A - O - U, corrispondono con buona approssimazione alle prime due formanti tipiche delle rispettive vocali prodotte dalla voce umana. Nel caso dei risonatori delle vocali E - I, la seconda formante risulta collocarsi intorno a 1000 Hz, registro più grave delle risonanze necessarie alla simulazione di tali vocali. Le formanti necessarie, circa 1800 Hz per la produzione della E, circa 2300 Hz per la I, collocate nella regione spettrale comune alla terza formante (1500 - 3000 Hz), sono state ottenute tramite l'apporto di un'equalizzazione specifica del segnale inviato ai risonatori. L'altoparlante, che presenta caratteristiche tecniche idonee ad assecondare l'enfasi della zona spettrale suddetta, partecipa dunque alla configurazione dello spettro sonoro del coro di vetro, contribuendo a costruire un timbro definito fino alla terza formante. Le formanti del timbro vocale presentano differenze sostanziali a seconda del tipo di voce analizzata ed una grandissima varietà di sfumature al variare di molteplici fattori: esse dipendono dalla diversa conformazione degli organi della fonazione in ogni individuo, le cui differenze sono accentuate con il variare dell'età e del sesso, dalle caratteristiche della lingua parlata, ma anche probabilmente da meccanismi percettivi legati più a fenomeni cognitivi o semantici che non a fenomeni strettamente acustici. Altra caratteristica specifica delle risonanze della voce è la tempo-varianza, consentita dalla duttilità dell'organo della fonazione e necessaria alla produzione del parlato: una piccola variazione dinamica della forma delle labbra, ad esempio, può modificare leggermente anche una sola delle prime due formanti di un suono e causare così il passaggio da una vocale ad un'altra. La complessità del timbro vocale data dalla differenziazione fisica, linguistica e dalla tempo-varianza, non è un'area di ricerca che riguarda il coro di vetro, in quanto esso non è stato progettato con l'obiettivo di simulare in modo fedele la voce umana e la precisa pronuncia di tutti i fonemi.

<sup>1</sup> Ho effettuato le misure ponendo l'altoparlante ad un'estremità di ciascun risonatore e misurando le variazioni d'ampiezza del segnale in uscita all'altra estremità di esso. Il segnale mandato all'altoparlante era un suono sinusoidale a frequenza variabile (glissando da 50 a 15000 Hz): la misura della variazione d'ampiezza della risonanza ottenuta rispetto al segnale in ingresso è stata considerata significativa per l'individuazione della banda passante se superiore o uguale a 3dB e per la frequenza centrale se superiore o uguale a 6dB. La misura del tempo di decadimento è invece stata misurata ponendo un impulso come segnale d'ingresso.

<sup>2</sup> I numeri mancanti sono relativi ad altri risonatori su cui sono state effettuate le misure, ma che non sono stati scelti per il coro di vetro.

Ris A	20	2	10
Ft1	700	1000	650
BW1	(300-1000)	(700-1200)	(550-750)
Ft2	---	---	---
BW2	---	---	---
Ris E	5	8	13
Ft1	550	425	500
BW1	(550-750)	(400-450)	(500-600)
Ft2	---	---	950
BW2	---	---	(800 - 1000)
Ris I	7	15	1
Ft1	300	290/320	400
BW1	(250-400)	(280-340)	
Ft2	---	995	1200
BW2	---	(980-1025)	
Ris O	3	14	9
Ft1	650	400	650
BW1	(550-750)	(400-600)	(600-700)
Ft2	---	800	1300
BW2	---	(800 - 1000)	(1200-1400)
Ris U	11	12	---
Ft1	350	300	---
BW1	(250-400)	(200-300)	---
Ft2	1060	900	---
BW2		(850 - 950)	---

Figura 4. Voce: frequenze di risonanza del coro di vetro.

I risonatori del coro di vetro conferiscono al timbro di voce una caratterizzazione originale dovuta non solo alle forme, ma anche al materiale di cui sono composti, responsabile, ad esempio, di un tempo di decadimento del suono sufficientemente più lungo di quello dovuto alle risonanze naturali dell'apparato vocale<sup>3</sup>.

Il coro di vetro, installazione d'arte elettroacustica, prende spunto da uno strumento sonoro reale, il timbro della voce umana, più precisamente un piccolo coro: lo evoca, rappresentandone alcuni tratti significativi, ma soprattutto ne propone possibili deviazioni e rappresentazioni immaginifiche, nell'intento di offrire nuove suggestioni, ipotesi di mezzi espressivi rinnovati.

La conformazione timbrica di un suono di voce è tuttavia legata anche allo spettro del suono originariamente prodotto dalle corde vocali, che è simile ad un dente di sega con armoniche di ampiezze progressivamente decrescenti (il comportamento delle corde vocali può essere paragonato a quello di un'ancia), nonché ai modi con cui detto suono viene articolato (attacco e decadimento). Anche il risultato acustico del coro di vetro deriva dalla cooperazione di diversi fattori<sup>4</sup>:

1. risonatori:  
volume – forma – materiale;
2. segnale di eccitazione:  
spettro – ampiezza – involucro d'ampiezza;
3. riproduzione del segnale di eccitazione:  
equalizzazione digitale – amplificazione – diffusione.

Un modello meccanico esposto al Deutsches Museum di Monaco in Germania e al Miami Science Museum © 2002 Exploratorium<sup>5</sup>, simula le cinque vocali attraverso risonatori

<sup>3</sup> Il tempo di decadimento dei risonatori di vetro è stato osservato solo in via empirica, allo scopo di estrarre le informazioni, di tipo qualitativo, per me necessarie allo svolgimento del lavoro.

<sup>4</sup> La scelta combinata altoparlante-risonatore è stata effettuata attraverso test svolti con rumore bianco, impulsi, ma soprattutto con denti di sega e simili.

<sup>5</sup> [www.exploratorium.edu/exhibits/vocal\\_vowels/vocal\\_vowels.html](http://www.exploratorium.edu/exhibits/vocal_vowels/vocal_vowels.html).

costruiti sulla base delle misure delle aree che l'apparato vocale assume durante la loro emissione. Il modello, per me molto suggestivo, è stato un'importante fonte di riferimento per la scelta delle forme dei risonatori di vetro.

Il contributo estetico apportato dai risonatori del coro di vetro, oltre che sonoro, è infatti anche visivo, in misura coerente alle esigenze espressive dell'opera, che richiede all'installazione l'apporto di una "personalità" distintiva. Le forme dei risonatori sono state scelte in modo da evocare visivamente alcune caratteristiche significative delle forme che vengono modellate con la bocca durante la produzione delle vocali nel canto e nel parlato.

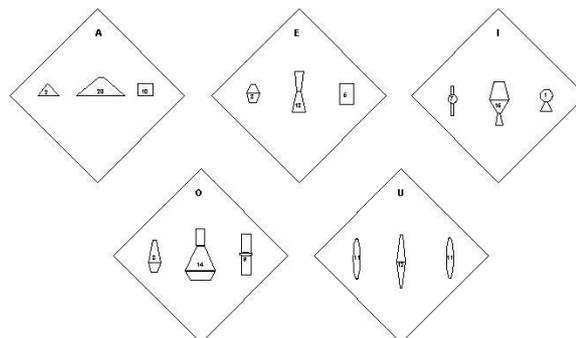


Figura 5. Voce: forme dei risonatori del coro di vetro.

Il disegno riportato in figura 5 indica in modo schematico le forme dei 15 risonatori e la vocale associata ad ogni elemento del coro, con la stessa numerazione utilizzata nella tabella delle frequenze formanti. Nella figura 6 sono mostrate le foto dei risonatori.

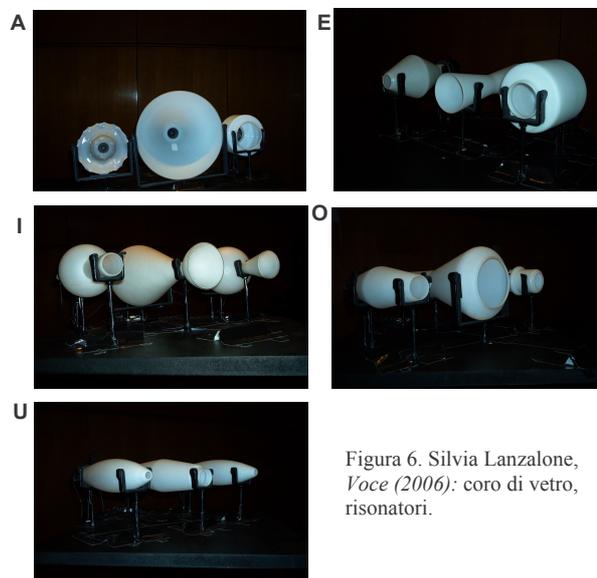


Figura 6. Silvia Lanzalone, Voce (2006): coro di vetro, risonatori.

I risonatori e gli altoparlanti ad essi associati sono stati collocati su supporti in ferro inseriti su basi quadrate di legno, in modo da simulare le "teste" dei personaggi del coro. Le basi quadrate sono state poggiate su 5 piedistalli, sempre in ferro, di altezza variabile tra 1.40 m e 1.70 m; ogni piedistallo sostiene quindi tre risonatori a cui è inviato lo stesso segnale, in modo che costituisca un corpo unitario e che somigli, acusticamente e visivamente, ad una sorta di "personaggio a tre teste".

La distribuzione dei cinque elementi del coro sul palco è libera, sia relativamente all'ordine delle vocali tra loro, sia rispetto alla forma della disposizione, se ad arco o lineare o su

gradini ad altezze differenziate, compatibilmente con un principio di efficace distribuzione sia delle grandezze o delle altezze (ad esempio la A presenta risonatori più grandi e vistosi rispetto alla U, che però è più alta), che delle sonorità. La voce femminile, per favorire il dialogo con gli elementi del coro, può spostarsi durante l'esecuzione collocandosi davanti, dietro, a destra o a sinistra dell'installazione.

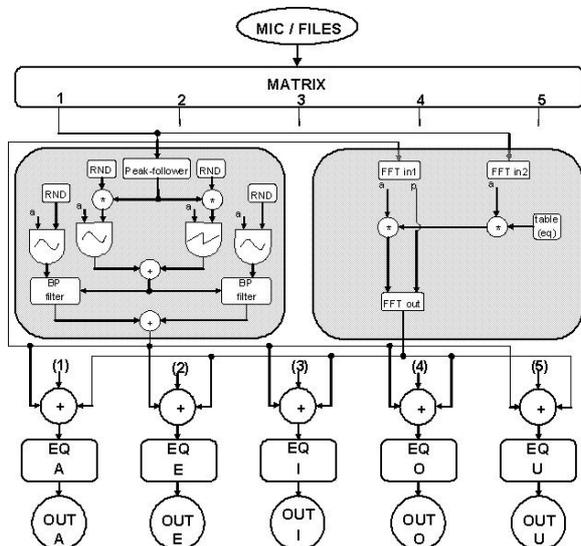


Figure 7. *Voce*: algoritmo di elaborazione del suono in tempo reale.

L'algoritmo di elaborazione del suono in tempo reale del brano è stato da me realizzato con il software Max/MSP<sup>6</sup>. La gestione dell'algoritmo è affidata ad un'interprete informatico; i suoni generati dal computer sono inviati al coro di vetro e da esso elaborati acusticamente. Dallo schema di figura 7 si vedono due blocchi di calcolo connessi in parallelo: a sinistra i segnali di due oscillatori a frequenza variabile in modo casuale sono elaborati da due filtri passa-banda con frequenza centrale modulata in modo random (sintesi); nell'algoritmo di destra si effettua la convoluzione tra il segnale di sintesi e il segnale di voce (elaborazione).

I materiali sonori presenti nel brano sono esclusivamente suoni di voce femminile, naturale o elaborata. All'interprete vocale è richiesto di improvvisare su diversi tipi di emissione, raggruppabili in tre categorie:

1. Voce sussurrata con respiri (sezioni 1-6):
  - vocali<sup>6</sup>;
  - consonanti fricative prolungate<sup>7</sup>;
  - plosive brevi con accenti e sforzando<sup>8</sup>.
2. Voce parlata (sezioni 3-5):
  - sillabe di durata differenziata<sup>9</sup>.
3. Suoni di gola (sezioni 7-8):
  - transizioni tra le vocali<sup>10</sup>.

Le elaborazioni digitali sono controllate in tempo reale dall'interprete informatico che regola, tra i vari parametri, anche le ampiezze dei suoni inviati al coro di vetro e la loro distribuzione spaziale tra i singoli elementi. Le elaborazioni finali dei suoni avvengono tramite il coro di vetro, che impone al suono le risonanze acustiche derivanti dalle caratteristiche fisiche dei diversi risonatori di cui è composto.

Il brano si configura come "un'opera aperta" o, più precisamente, come un'opera la cui forma è strettamente interconnessa con il carattere estemporaneo dell'esecuzione. La partitura è suddivisa in 10 sezioni, racchiuse da altrettanti riquadri contrassegnati da un numero d'ordine, in cui si descrivono: la durata approssimativa della sezione, le modalità esecutive della voce femminile, il numero di configurazione dell'algoritmo, l'andamento ritmico/dinamico del coro di vetro. Fino alla sezione 8 l'interprete vocale svolge un ruolo guida nella gestione dei tempi e delle microvariazioni formali. Il passaggio da una sezione all'altra del brano ed i relativi sincronismi, dipendono dai riferimenti biunivoci che l'interprete informatico e la voce femminile hanno stabilito in fase di studio della partitura. A partire dalla sezione 8 l'interprete esce gradualmente dalla scena e non emette più suoni; in corrispondenza di tale momento è introdotta in partitura una simbologia grafica specifica per il coro di vetro, relativa al ruolo solista che questo assume da qui sino alla fine del brano (sezioni 9 e 10). I simboli scelti sono indicativi di dinamica (grandezza del simbolo) e di durata (posizione e lunghezza del simbolo), ma evidenziano anche una peculiarità acustica del nuovo timbro assunto dal coro. Il coro infatti eredita dalla voce femminile una forma di articolazione assimilabile ad una sorta di parlato, all'interno del quale sono percepibili elementi simili ai fonemi, raggruppati in unità formali più grandi, simili a parole e frasi.

Figure 8. Silvia Lanzalone, *Voce*: prima e ultima pagina della partitura.

Il percorso formale è prestabilito da un punto di vista deterministico, in quanto è evidenziata in partitura la necessità di connessione temporale tra le diverse sezioni strutturali del brano. La definizione degli accadimenti sonori nelle sezioni è affidata però ad una modalità di creazione estemporanea del gesto, da cui deriva flessibilità formale e libertà espressiva.

<sup>6</sup> A - E - I - O - U, comprese le loro possibili varianti fonetiche.

<sup>7</sup> Z - S - SC - C - V - W - F.

<sup>8</sup> G - K - D - T - B - P.

<sup>9</sup> TE, NE, SE, VI, CI, QU, ZU, DU, ecc.

<sup>10</sup> AAAIIIIIOOOUEEEEE, ecc.

La personalità artistica dell'interprete vocale è essenziale alla determinazione musicale e teatrale dell'opera. Gli ambiti di libertà concessi all'interprete rispetto alla definizione dei micro-livelli di articolazione fonetica e musicale, rendono inoltre anche i materiali e le relazioni dialogiche con il coro estremamente duttili e suscettibili di modificazione sostanziale tra un'esecuzione e l'altra. Una determinata interprete vocale necessariamente conferisce all'opera una fisionomia molto variata rispetto ad un'altra. Gli ancoraggi formali costituiti dalle caratterizzazioni descritte o rappresentate in partitura garantiscono la definizione di un'identità strutturale, stilistica, nonché di un percorso-guida a cui legare la realizzazione delle linee espressive e dello sviluppo narrativo della composizione. Il brano dunque, nel suo essere strutturato tramite un criterio deterministico e, contemporaneamente, nel suo rigenerarsi attraverso un principio di flessibilità legata al gesto, vive simbolicamente, ad ogni sua rappresentazione, un processo di trasfigurazione: l'emancipazione del gesto creativo dalla <<pesantezza del cristallizzarsi in un auto-perfezionamento>> (Berdjaev) consente di rinnovare l'opera ogni volta che viene rappresentata, mettendone in evidenza un principio di libertà.

### RISONANZE DALLA TERRA (2007)

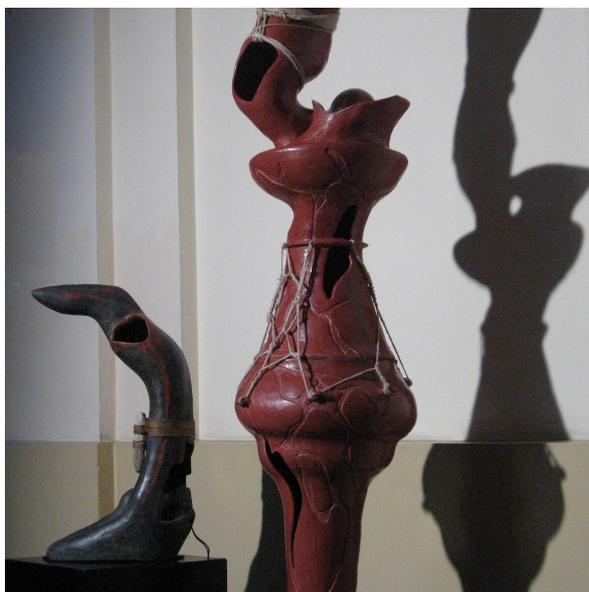


Figura 9. Silvia Lanzalone / Debora Mondovì, *Risonanze dalla Terra*: sculture in terracotta e canapa.

*Risonanze dalla Terra* è un'opera musicale e scultorea nata dall'integrazione di suono e forma nell'ambito di un progetto di ricerca e sperimentazione sulle sonorità e le risonanze producibili con la terracotta. La musica è stata composta con i suoni provenienti dalle esplorazioni acustiche effettuate su sculture e lastre di terracotta realizzate appositamente dall'artista Debora Mondovì. L'opera scultorea è stata progettata per risuonare a determinate frequenze dello spettro acustico e consiste di elementi cavi di diverse dimensioni entro cui i suoni vengono amplificati o attenuati coerentemente alla conformazione e grandezza della cavità. I materiali sonori sono stati prima registrati, poi selezionati, successivamente organizzati e sottoposti a processi di elaborazione digitale del segnale. Formalmente l'opera musicale si configura come un ciclo di quattro composizioni in quanto è suddivisa in altrettanti episodi detti *risonanze*, diversamente contraddistinti sia per natura dei materiali, sia per carattere ed articolazione interna: *risonanza I – dialogo*; *risonanza II – dalla terra*; *risonanza III – lamenti*; *risonanza IV – contrasti*.

La durata dell'opera è variabile in quanto ciascuna o più risonanze e/o l'intero brano, possono essere eseguiti anche in loop, ad libitum.

La prima esecuzione assoluta del brano si è svolta presso l'Accademia di Romania di Roma il 29 giugno 2007, in occasione della manifestazione SPAZIAPERTI5, durante l'evento-concerto *voce solitaria*, titolo di un'opera scultorea di Debora Mondovì prodotta per il brano *Solo* di Fausto Sebastiani, anch'esso composto in modo da sfruttare il principio delle risonanze delle sculture di terracotta. La collaborazione artistica instauratasi in occasione di questo evento è finalizzata alla produzione di opere che siano il risultato di una ricerca estetica rivolta all'integrazione dei linguaggi espressivi.

Il progetto d'installazione elettroacustica dell'opera *Risonanze dalla Terra* è stato sviluppato integrando le idee plastiche alle idee musicali, in modo da ottenere sculture formate da cavità risonanti con specifiche caratteristiche acustiche.

L'esplorazione dei materiali sonori di *Risonanze dalla Terra* si è svolta in più fasi, secondo criteri di produzione del suono differenziati e progressivi:

- percussione e fregamento con oggetti di diverse forme e materiali su sculture cave di terracotta dell'artista;
- percussione e fregamento con oggetti di diverse forme e materiali su lastre di terracotta appositamente realizzate dall'artista<sup>11</sup>;
- eccitazione delle lastre di terracotta tramite i suoni percussivi registrati durante le sperimentazioni delle fasi precedenti ed inviati elettricamente attraverso attuatori piezoelettrici;
- le lastre di terracotta sono inserite in appositi risonatori dello stesso materiale realizzati su misura dall'artista e vengono sottoposte ad eccitazione come descritto al punto precedente.

I materiali registrati durante le fasi di esplorazione sonora sulla terracotta sono stati elaborati digitalmente con algoritmi di filtraggio, time-stretching, distorsioni, trasposizioni di altezza, ecc., secondo criteri di progressiva trasformazione, spesso raggiunta assecondando la natura acustica dei suoni, altre volte ottenuta modificandola radicalmente. Le elaborazioni hanno prodotto nuovi elementi sonori dotati di senso formale autonomo a cui ho successivamente attinto per la composizione delle singole *risonanze*.

In genere, nei risonatori di forma semplice le frequenze di risonanza possono essere multiple della frequenza di risonanza fondamentale; nei risonatori di forma complessa tali frequenze possono avere rapporti non armonici e livelli di ampiezza non regolarmente decrescenti; inoltre si possono formare delle zone di attenuazione del suono, dette antirisonanze. In alcuni casi determinate frequenze di risonanza possono essere associate con facilità a determinate parti di una cavità complessa, ma per l'analisi di questa installazione, data la difficoltà di scomporre l'intera cavità in più risonatori semplici, mi è sembrato maggiormente conveniente analizzare tutte le frequenze di risonanza nel loro insieme come risultato di un contributo complessivo.

Il progetto delle sculture di *Risonanze dalla Terra* è stato ottimizzato per via empirica attraverso la creazione di alcuni prototipi. Le sculture, alla fine della fase sperimentale, hanno raggiunto una forma definitiva che consiste di due unità separate: la prima, alta 1.67 m, con due cavità risonanti; la seconda, alta 63 cm, che presenta una terza cavità<sup>12</sup>.

<sup>11</sup> Le lastre di terracotta hanno diverse dimensioni: 15\*40 cm circa la lastra più piccola, 40\*60 cm la più grande. Gli spessori variano tra 0.3 e 0.7 mm.

<sup>12</sup> Le dimensioni sono fornite senza includere la misura delle basi su cui le sculture sono poggiate.

All'interno delle sculture sono stati inseriti complessivamente quattro altoparlanti con caratteristiche tecniche adeguate a consentire una riproduzione ottimale dell'intero spettro sonoro. In figura 10 sono mostrate le posizioni in cui sono stati collocati gli altoparlanti<sup>13</sup>, che tengono conto anche dei principali punti di irradiazione del suono costituiti da fori di varia grandezza e forma posti in diverse zone alla scultura, alcuni dei quali sono visibili in figura 9.

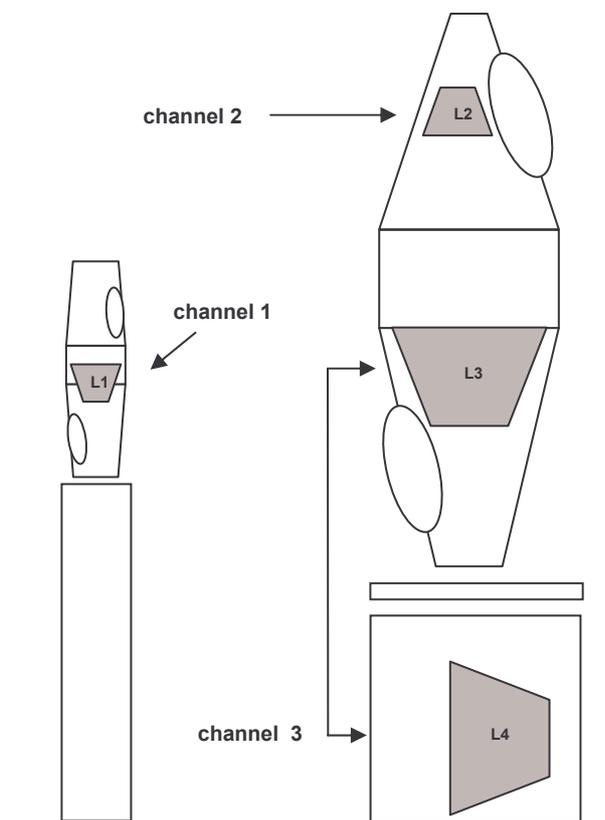


Figura 10. *Risonanze dalla Terra*: schema dell'installazione.

L'ascolto diretto, durante la composizione dell'opera, dei suoni plasmati dalle sculture, è stato pressoché essenziale alla definizione del brano relativamente agli aspetti sia formali che espressivi. Al fine di poter verificare il reale apporto, in termini di elaborazione sonora, delle risonanze ed antirisonanze prodotte dalle sculture, riporto alla fine del presente articolo, in figura 11, i grafici delle forme d'onda, dei sonogrammi e dell'analisi FFT effettuata su un suono percussivo estratto da *risonanza I - dialogo*. Nei grafici viene messo a confronto il suono originale con il suono diffuso dall'installazione: oltre ad un diverso decadimento del suono, più lungo per il suono in uscita dalle sculture, si riscontra che la regione media (300-800 Hz) presenta una più accentuata concentrazione di energia nel suono elaborato, mentre le frequenze più acute di detti suoni risultano attenuate. *Risonanze dalla Terra* è disponibile anche nella versione per supporto digitale, realizzata da una registrazione dell'esecuzione del brano diffuso tramite le sculture.

<sup>13</sup> Il quarto altoparlante è stato inserito nella base di legno ed ha la funzione di integrare il risultato acustico complessivo con un rinforzo dei suoni gravi.

## CONCLUSIONI

Algoritmi digitali di simulazione o di elaborazione dei suoni strumentali sono oggi adoperati abitualmente dai compositori per la realizzazione di opere musicali definite come "elettroacustiche" o "acusmatiche" o con "live electronics". In questi casi la macchina digitale o l'algoritmo di calcolo non si caratterizzano come emettitori diretti del segnale sonoro poiché questo, per essere udibile, necessita del transito verso la conversione D/A e poi verso gli altoparlanti. La catena elettroacustica, nonostante il notevole perfezionamento tecnologico degli ultimi decenni verso la massima linearità della risposta in frequenza, conferisce ancora al suono una trasformazione finale dovuta alla natura elettro-meccanica dei suoi componenti, impone dunque al suono una "personalità acustica", anche se ridotta. Il suono elettronico, seppure in misura diversa a seconda dei casi, resta decorrelato dalla fonte reale – lo strumento – da un punto di vista acustico, visivo e corporeo, e rimane dunque nell'ambito del "virtuale".

Le opere musicali fin qui descritte sono state ottimizzate per essere eseguite soltanto in forma di installazione; le relative versioni per supporto digitale sono costituite dalle registrazioni dei brani così come sono stati eseguiti in concerto. La struttura spettrale dei suoni utilizzati è strettamente dipendente dalle trasformazioni sonore apportate dal clarinetto, dalle cavità di vetro o dalle sculture di terracotta, installazioni a cui sono legati anche i risultati espressivi di ogni singola opera.

I brani presentati sono esempi di arte intermediale, caratterizzata da componenti sonore, visive e teatrali, in cui il suono è sia elettronico che acustico, in quanto prodotto da installazioni di arte elettroacustica in rapporto di interazione con l'azione dello strumentista, quando questi è previsto. In queste opere la musica ed il gesto esecutivo, lo strumento elettromeccanico ed il sistema digitale di elaborazione del suono, sono perfettamente integrati in un organismo unitario.

## RIFERIMENTI

1. AA.VV. *Manual of Phonetics*, a cura di Bertil Malmberg, North Holland Publishing Company, 1968.
2. Arthur H. Benade. *Fundamentals of musical acoustics*. Mineola, New York, 1990.
3. Arthur H. Benade. *I tubi sonori*. Zanichelli, Bologna, 1960.
4. Murray Campbell e Clive Greated. *The musician's guide to acoustic*. Dent, London, 1987.
5. Lindoro Massimo Del Duca. *Musica digitale*, Franco Muzzio Editore. Padova, 1987.
6. Gary Davis & Ralph Jones. *Yamaha. The sound reinforcement handbook*, Hal Leonard Corporation, Milwaukee, WI, U.S.A., 1990.
7. Charles Dogde, & Thomas A. Jerse. *Computer music. Synthesis, composition, and performance*. New York, 1985.
8. *Max, Reference Manual*, © 2000-2005 Cycling '74.
9. *MSP, Reference Manual*, © 2000-2005 Cycling '74.
10. Miller Puckette. *Theory and techniques of electronic music*, University of California, San Diego, 2003.
11. Phillip Rehfeldt. *New directions for clarinet*. University of California Press, Berkeley, Los Angeles, 1977.
12. Pietro Righini, Giuseppe Ugo Righini. *Il suono. Dalla fisica - all'uomo. Dalla musica - alla macchina*. Tamburini Editore. Milano, 1974.
13. Curtis Roads. *The computer music tutorial*, MIT Press. Cambridge Massachusetts, 1996.
14. Manfred R. Schroeder. *Computer Speech. Recognition, Compression, Synthesis*. Springer, 1999.
15. Johan Sundberg. *The science of the singing voice*. Northern Illinois University Press, 1987.

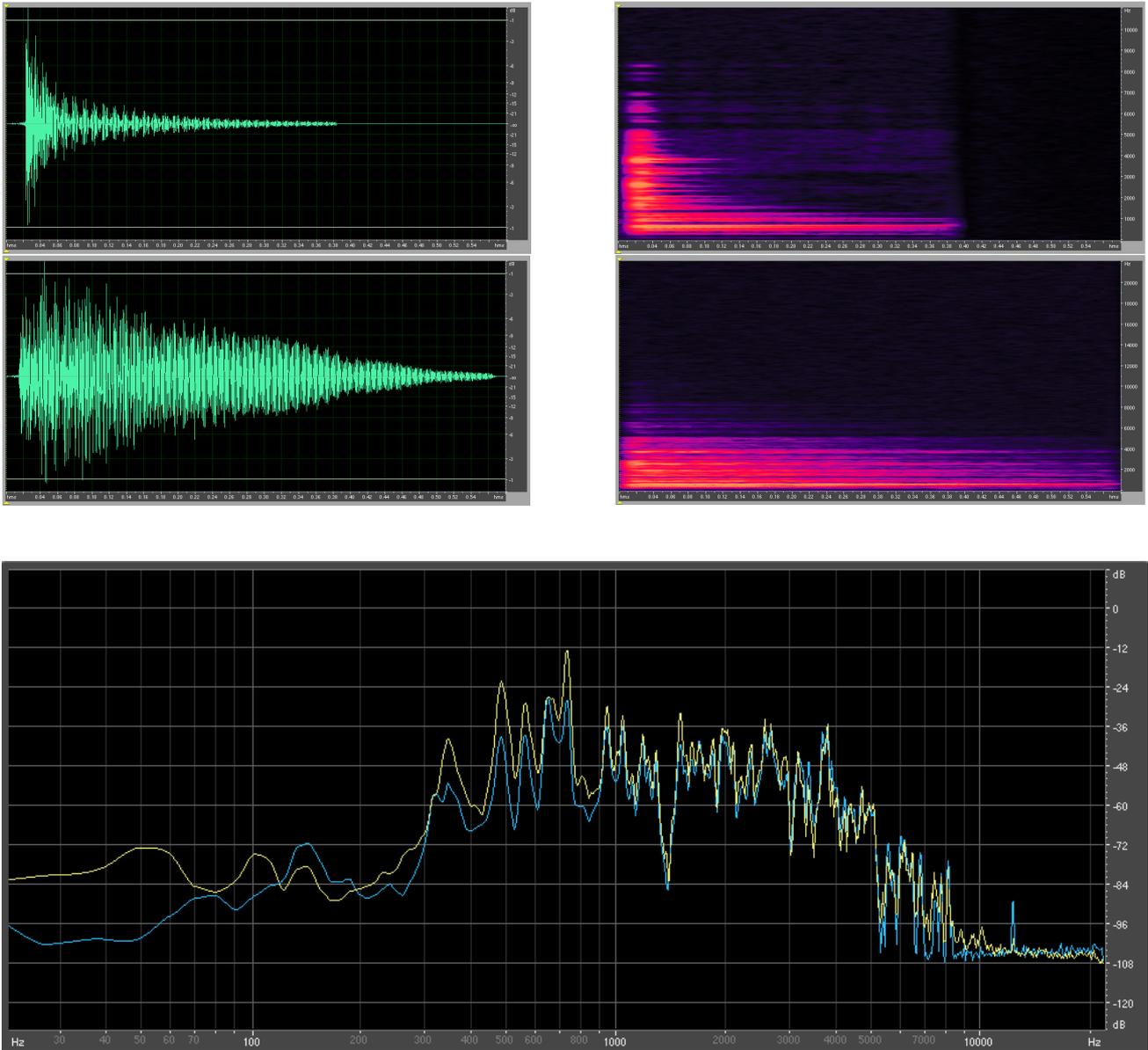


Figura 11. *Risonanze dalla Terra*: grafici per l'analisi di un suono percussivo da *risonanza I – dialogo*. In alto a sinistra: forme d'onda del suono originale (sopra) e del suono elaborato dalle sculture (sotto). In alto a destra: spettrogramma del suono originale (sopra) e del suono elaborato dalle sculture (sotto); In basso: FFT del suono originale in blu, FFT del suono elaborato in giallo.

# Oggetti Sonori nella Rete

Dante Tanzi

Laboratorio di Informatica Musicale  
Dipartimento di Informatica e Comunicazione  
Università degli Studi di Milano

## SOMMARIO

Nell'ambiente della rete le modalità della ricezione, la latenza temporale, la tipologia dei gesti interattivi e le interfacce per la navigazione influiscono sulla qualità e sulla forma degli eventi sonori.

## Parole Chiave

Rete, ricezione, produzione, latenza temporale, spazi risonanti, cybertempo, interattività, ambienti sonori navigabili.

## INTRODUZIONE

A partire dalla fine degli anni '90 l'ipertesto e lo sviluppo del web hanno avviato un nuovo processo di integrazione fra i media, basato sulle idee di connettività e di non linearità. Nel 1996 Brain Opera, uno dei più complessi sistemi interattivi per la produzione di materiale multimediale, ha mostrato la possibilità di intervenire sulle relazioni suono-immagine e gesto-suono attraverso l'uso di varie stazioni musicali e dell'interattività della rete [1]. Da allora le manifestazioni della presenza del suono nella rete si sono moltiplicate in quantità e qualità: pensiamo agli oltre 900 mila brani musicali in formato MIDI e MP3 scaricati ogni giorno, alle ambientazioni sonore dei siti Web e dei mondi VRML, ai concerti a distanza, alle sessioni di improvvisazione on-line e ai siti dedicati alla composizione collettiva. L'utilizzo della banda larga, il downloading e la diffusione dello scambio di files musicali hanno aperto la strada a nuove modalità di ricezione e di produzione dei contenuti musicali. Le modalità di ricezione si differenziano in base al numero delle sorgenti, alle funzioni implementate nelle interfacce di navigazione e ai metodi usati per la descrizione degli eventi musicali; ma esse dipendono anche dalle caratteristiche degli ambienti mediali e dal grado di interattività offerto dagli ambienti sonori navigabili. In alcuni casi la ricezione on-line è in grado di modificare il modello lineare di ascolto e di avvicinarlo ad una situazione di tipo acusmatico. Anche le modalità di produzione dipendono dal grado di interattività e dalle funzioni di tipo generativo, esplorativo e manipolativo implementate nelle applicazioni musicali on-line. Nei controlli interattivi l'utilizzo di simboli musicali e di altri elementi grafici segue criteri di importanza che vengono attribuiti dai progettisti alle componenti audio, simboliche, strutturali e grafiche dell'informazione musicale. Le più conosciute tra le applicazioni musicali on-line (da 'Koan Pro' fino a 'Fmol', da 'Cathedral, fino alle 'Reactable') richiedono agli audio-navigatori un approccio attivo, li sollecitano ad intraprendere percorsi di ideazione e di produzione, a diffondere le proprie esperienze e a partecipare a sessioni di produzione collettiva.

## LA CREATIVITA' NELLA RETE

Analogamente ad altri media, la rete è stata in grado di introdurre nuove proporzioni nella percezione del mondo e nei principi di differenziazione della realtà. Da ciò è conseguito

un mutamento di paradigma che ha influenzato tanto i processi comunicativi che le espressioni della creatività. Infatti la mappa topologica e informativa della rete propone un modello di interattività che sollecita una particolare attenzione nei confronti dei processi di de-localizzazione, mentre i processi di conoscenza distribuita tendono a globalizzare le proprietà emergenti. Le relazioni fra questi processi sono state illustrate da Geri Wittig nell'articolo *Situated and Distributed Knowledge - Production in Network Space*, pubblicato nel 2001: "Mentre l'enfasi fenomenologica sulla fisicità e il contesto viene correlata alla produzione locale di conoscenza, un sistema sociale basato sulla rete crea le capacità e le risorse per arricchire e amplificare la produzione di conoscenza distribuita. L'interazione e l'intreccio fra i contenuti espressi dai contesti locali e le dinamiche della conoscenza distribuita sono in grado di attivare i fenomeni emergenti di risonanza tra locale globale. Esiste una relazione dialettica tra le interazioni locali dei componenti di un sistema sociale in rete e le proprietà globali emergenti dell'intero, ed esistono relazioni di causalità che coordinano le due realtà. Nella realizzazione di un sistema di auto-organizzazione reticolare, l'interazione dei componenti sistemici da' origine ad un intero, o unità, che mantiene continuamente i propri limiti di identificazione. La caratteristica globale a sua volta obbliga i propri componenti a mantenere una specifica coerenza operativa, in caso contrario l'unità non sarebbe mantenuta. La chiusura organizzativa del sistema sociale di auto-organizzazione reticolare non indica una chiusura interattiva. Allo scopo di non interrompere la produzione di conoscenza, il sistema dipende da nuovi input, ma il sistema specifica che cosa - all'interno dell'ambiente globale della rete - è importante per se stesso. Un sistema sociale reticolare delimita le possibili interazioni con la rete, in funzione del mantenimento dell'autoproduzione, e per la realizzazione di una particolare prospettiva ed identità" [2]. Oltre ad illustrare le dinamiche fra locale e globale, il contributo di Geri Wittig mette in chiaro che una volta trasferiti da un piano mentale-soggettivo ad un piano sistemico, alcuni dei connotati della creatività individuale necessitano di essere descritti in modo appropriato. In effetti le proprietà sistemiche di un sistema sociale basato sulle reti assicurano la sopravvivenza della creatività, ma non necessariamente nella sua espressione individuale. Anche la creatività musicale non fa eccezioni, e si manifesta attraverso forme relazionali che innervano processi produttivi distribuiti, che non di rado giungono a sfidare molte convenzioni legate al linguaggio. Per questo le manifestazioni della creatività musicale on-line dovrebbero essere interpretate considerando con attenzione le caratteristiche dell'ambiente entro cui esse si sviluppano.

## SPAZI RISONANTI

Secondo Randall Packer nell'ambiente della rete esiste una dimensione sonora primaria, uno sfondo sonoro costituito dallo scorrere dei flussi di dati che si muovono in ogni direzione.

Nell'introdurre "Telemusic #2", installazione presentata nel 2003 all'Università De Montfort di Leicester nell'ambito del convegno 'Music without Walls? Music without Instruments' Packer ha spiegato come i flussi di dati vengono rivelati, interpretati statisticamente e poi trasformati in tessiture sonore le cui densità possono mutare in funzione delle velocità di trasferimento dei dati [3]. I profili formali delle tessiture sonore ottenute dipendono dal numero delle sorgenti, dall'ordine di accesso ai flussi di dati e dai modi in cui questi ultimi possono essere interpretati. Le transizioni da un flusso di dati all'altro possono attivare associazioni e di interferenze fra strutture, volumi, densità e timbri diversi. D'altra parte, secondo scrittore Erik Davis è l'interdipendenza tra diversi piani comunicativi a caratterizzare lo spazio della rete come *spazio risonante*. Davis è stato il primo ad utilizzare la *metafora* dello spazio acustico al fine di illustrare la complessità delle relazioni fra gli spazi mediali elettronici, e il primo a tentare di spiegare il perché del manifestarsi di alcune differenze temporali locali. Nell'articolo "Roots and wires. Polyhythmic Cyberspace and the Black Electronic" pubblicato nel 1997 Davis si esprimeva in questi termini: "Uno spazio acustico è capace di simultaneità, sovrapposizioni e non linearità, ma soprattutto, esso *risuona* [...]; la risonanza può essere assimilata ad una forma di causalità, ma si tratta di un tipo di causalità diversa da quella associata allo spazio visuale, poiché permette alle cose di risponderne vicendevolmente e in modo non lineare [...]. Proprio perché provenienti dalla dimensione acustica, le idee di risonanza e di non-linearità ci mettono in condizione di modulare le *differenze temporali locali* e di progettare l'ambiente dei media elettronici in un modo che un approccio visuale, grafico e testuale non potrebbe consentire" [4]. Dunque, se grazie agli spazi mediali elettronici il tempo cronologico riesce ad esprimere differenze temporali locali è lecito chiedersi cosa accade al tempo musicale in condizioni analoghe.

## TEMPO MUSICALE E CYBERTEMPO

In quanto universo comunicativo, la rete manifesta qualità che scaturiscono dall'interferenza fra reale e virtuale e da processi di dissociazione che investono i parametri tempo, distanza e velocità. Nell'articolo "Experiencing cybertime: computing as activity and event" del 1995, Lance Strate ha descritto il tempo della rete (denominato cybertempo) come una rete di segmenti temporali caratterizzati da differenti capacità evolutive: "Il cybertempo rappresenta l'intersezione di parecchi fenomeni temporali differenti. In primo luogo c'è il tempo-funzione degli elaboratori, il modo in cui esso costruisce un senso interno di tempo e il modo in cui misura il passaggio di tempo nel mondo esterno al fine di coordinare azioni ed eventi. In secondo luogo c'è la funzione di rappresentazione degli elaboratori [relativa a qualità astratte di tempo] e come, in quanto medium, essa possa convogliare un sensazione drammatica, fittizia o simbolica dello scorrere del tempo, o convogliare il senso di passato, presente o futuro. In terzo luogo c'è la nostra esperienza soggettiva di tempo, che riguarda il modo in cui noi interagiamo con e attraverso i media computerizzati, ma anche il modo in cui l'elaborazione dei dati costituisce una attività umana: un evento che influenza la nostra percezione del tempo, il nostro senso del sé e della comunità" [5]. Ma gli accadimenti che dipendono dagli elaboratori e dalle reti e i ritardi temporali possono anche ostacolare le aspettative di tipo cronologico: secondo Peter Anders, direttore di *Mindspace.net*, essi possono addirittura interrompere le relazioni causali fra i contenuti e le rispettive sorgenti [6]. Oltre ad influire sulle percezioni delle caratteristiche degli spazi navigabili, tali accadimenti possono sollecitare lo sviluppo di nuove sensibilità percettive. Nel

2004 Michel Bussière, compositore e direttore dell'istituto di Sonic Design dell'università di Carleton di Ottawa ha affrontato il tema dei mutamenti percettivi nell'ambiente della rete. Nell'articolo "Performance space meets cyberspace: seeking the creative idioms and technical model for live music on broadband" Bussière ha sostenuto che la banda larga consente l'emergere di proprietà comunicative del tutto nuove, che si manifestano in una dimensione che sta *al di là* delle usuali coordinate spaziali e temporali: "Le sensazioni che derivano dalle esperienze di presenza distribuita giungono ad alterare le qualità delle nostre risposte, ed in modo del tutto *distinto* da un processo intellettuale. Esse sono il prodotto della sensibilità umana "spalmata" lungo una realtà interoperazionale e multiluoogo; sono il paradosso di vivere uno spazio mentale condiviso mentre si risiede in spazi separati" [7]. Inoltre Bussière si è chiesto se e in quali termini la musica potrebbe accettare la sfida della latenza temporale; fino a che punto sarebbe opportuno attribuire ad essa una valenza estetica, e se sia già stato costituito un repertorio in grado di prefigurare i caratteri di un idioma musicale basato sulle interferenze fra tempo cronologico, tempo soggettivo e tempo della rete. In effetti sono molti gli artisti che nel loro lavoro utilizzano l'idea di latenza temporale. Ne citiamo solo uno, Atau Tanaka. Egli da tempo sostiene che solo accettando la latenza temporale come fatto naturale potrà essere possibile definire un linguaggio musicale che operi in accordo con la temporalità della rete [8]. Da 10 anni Tanaka applica alle sue creazioni la nozione di spazio acustico di Internet; assieme a Kasper Toeplitz ha realizzato nel 1998 una installazione multi-sito chiamata "Global String", strumento musicale che si serve della rete come *corpo risonante* e si avvale di un server dedicato alla sintesi del suono [9]. Allo scopo di investigare le relazioni tra spazio acustico reale e spazio risonante della rete, Tanaka ha realizzato nel 1999 "Constellations", una installazione di net-music che simula un sistema planetario in cui ciascun pianeta veicola il contributo di un compositore [10].

## NEGOZIARE I FLUSSI TEMPORALI

Ma l'utilizzo - sul piano estetico - degli effetti della latenza temporale non sembra poi essere del tutto condiviso, se c'è chi si è posto il problema di minimizzare tali effetti realizzando strumenti in grado di "negoziare" il tempo. In vista di questo obiettivo i ricercatori del M.T.G. (Music Technology Group) dell'Università Pompeu Fabra di Barcellona hanno sviluppato il progetto "Public Sound Objects" (PSOs), uno spazio web condiviso che ha la funzione di testare differenti soluzioni per la comunicazione musicale on-line. Lo scopo principale è quello di perfezionare le funzioni di uno strumento che *incorpori* la latenza della rete e *adatti* il proprio tempo interno ai ritardi nella comunicazione [11]. Se, per ipotesi, si potessero del tutto eliminare gli effetti provocati dalla latenza temporale rimarrebbe necessario confrontarsi col la ricezione multicanale di flussi musicali caratterizzati da differenti scansioni temporali in grado di mascherare o annullare la concatenazione fra gli eventi musicali, provocando continui aggiustamenti percettivi e reazioni che si ripercuotono nella sfera emotiva e motoria [12]. Ma se è vero che la sovrapposizione può generare interruzioni delle catene causali relative a singoli flussi musicali fino a determinare perdite di senso, è anche vero che la ricezione multicanale non impedisce di attribuire un senso musicale a posteriori. Mentre le modulazioni temporali, la risonanza fra piani comunicativi e i modi dell'interazione on-line (uno-uno, uno-molti, molti-molti) concorrono a ridefinire la percezione soggettiva delle proprietà degli spazi intermediali nell'ambiente della rete, esperienze analoghe - volte ad estendere e ad approfondire gli

ambiti della multimodalità - giungono a caratterizzare la produzione artistica di intermedia *fuori* dalla rete. Nel corso della visita all'installazione 'Pavillion' ideata da Randall Packer, la percezione del tempo e dello spazio dei visitatori viene volutamente dissociata allo scopo di separare ed elaborare diversi flussi temporali, contrastare le aspettative di continuità e sovvertire la differenza fra passato e futuro [13]. Jerome Joy, coordinatore con Peter Sinclair e Anne Roquigny del progetto di formazione e ricerca 'Locus Sonus' ha sottolineato che in seguito all'avvento della rete molte pratiche artistiche sono state indotte a dislocarsi in ambiti impropri o periferici, lasciandosi modulare dalle forme e dai ritmi che il nuovo universo comunicativo rende volta per volta attuali o disponibili [14].

## NAVIGAZIONE, ASCOLTO, COMUNICAZIONE

Si può affermare, con una metafora, che la musica viene *navigata* mediante l'ascolto, specie se ripetuto con intenzioni diverse. Diverse condotte di ascolto suggeriscono intenzioni che possono richiedere al tempo funzioni differenti: di unità metrica o cronologica, di unità appercettiva, di contenitore, di selettore, di puntatore, di materia-tempo, di durata. Di conseguenza la navigazione può essere orientata dall'esplorazione di ambiti distinti dell'informazione musicale: la durata, la materia-tempo, la densità, il linguaggio, lo sviluppo formale, l'articolazione discorsiva, le cadenze narrative. Esistono elementi di affinità tra la navigazione in rete e l'ascolto, che riguardano la collocazione di entrambi nel dominio temporale, il loro carattere ricettivo/attivo e le modalità di consultazione dei contenuti. Vi sono analogie nei modi in cui viene prefigurato o ripensato, nel tempo, un *ordine dello svolgimento degli eventi* che suggerisce a volte una pluralità di realizzazioni. Esistono somiglianze che riguardano le modalità di esplorazione dei contenuti effettuate con procedimenti di lettura che si sviluppano a diverse velocità, che utilizzano le modalità sinistra/destra e precedente/successivo, che prediligono un approccio ricorsivo e che coinvolgono più livelli informativi. Esistono similarità riguardo alla percezione soggettiva del cambiamento: spesso nel corso dell'analisi dei contenuti - musicali e non - l'inversione delle direzioni di lettura e il loro cambio di passo rivelano nuove caratteristiche e suggeriscono riflessioni che aprono la strada a mutamenti di prospettiva, dentro e fuori il dominio del tempo. Nell'ambiente della rete i contenuti musicali possono essere navigati in quanto eventi sonori esistenti; possono essere prodotti mediante l'interpretazione di flussi di dati, generati attraverso la sintesi oppure costruiti assemblando primitive musicali espresse da disegni, caratteri o grafismi. Ma ciò che caratterizza la navigazione dei contenuti musicali nella rete è la qualità dell'interazione e della ricezione, l'efficacia delle opzioni interlocutive ed il ruolo delle interfacce. Questi fattori sono in grado di separare o aggregare funzioni esplorative diverse, rivolte all'analisi delle strutture musicali, della materia sonora, e al monitoraggio dell'interazione. Le performances analitiche che accompagnano la navigazione e le modalità di produzione dei contenuti musicali dipendono dalla connettività, dal numero di interlocutori e dalle funzioni disponibili. Gil Weinberg ha distinto quattro differenti approcci che caratterizzano dal punto di vista comunicativo le applicazioni musicali on-line: l'approccio *Server*, che consente solo la connessione centralizzata ed esclude la comunicazione fra partecipanti; l'approccio *Bridge*, che permette alla rete di agire come un intermediario tra gli utenti, in modo da permettere scambi nei limiti di un medesimo spazio fisico interconnesso; l'approccio *Shaper*, che consente ai partecipanti un ruolo attivo: essi possono disegnare e modificare materiali generati

algoritmicamente da un server centrale, ascoltare e rispondere a sollecitazioni e messaggi collettivi, ma non comunicare singolarmente; infine l'approccio *Construction Kit*, che permette a utenti esperti di contribuire alla produzione musicale on-line, elaborando materiali musicali propri o modificando i contributi di altri utenti [15].

## PRODUZIONE MUSICALE ON-LINE

Negli ultimi 10 anni sono state sviluppate numerose applicazioni musicali on-line del tipo *Construction Kit*, che consentono ai navigatori percorsi di elaborazione musicale del tutto autonomi. Nel 2003 Alvaro Barbosa ha dato il nome di *Ambienti Sonori Condivisi* alle applicazioni progettate per esplorare la natura distribuita e collaborativa della rete e caratterizzate da un apprezzabile grado di interattività nella manipolazione e nella generazione di oggetti musicali [15]. Tali applicazioni potrebbero rientrare nella categoria denominata da Weinberg '*Shaper e Construction Kit*', ma se ne distinguono per essere più adatte all'improvvisazione sincrona, per essere accessibili a un largo pubblico e per non richiedere particolari conoscenze musicali. Secondo Barbosa l'applicazione "Cathedral", concepita e realizzata dal compositore William Duckworth nel 1996, è quella che più di altre incarna le caratteristiche proprie della produzione e della ricezione musicale on-line, vale a dire: a) la messa in questione dell'ordine gerarchico o sequenziale di accesso dei contenuti musicali; b) il superamento dell'isolamento dei singoli autori; c) la partecipazione a sessioni musicali interattive da luoghi diversi; d) la classificazione dei contenuti musicali per mezzo di marcatori multi-dominio. Nel 2003 lo stesso Duckworth ha affermato che le modalità di rappresentazione e di fruizione dei contenuti proposti da "Cathedral" possono differire radicalmente dalle forme di rappresentazione e di fruizione tradizionali come il concerto o il teatro musicale: "Per citare la ragione più ovvia, la forma dei contenuti musicali prodotta da "Cathedral" non sembra essere espressa secondo un tempo lineare. Ed in effetti non lo è, dal momento che un brano realizzato con "Cathedral" *non ha* un inizio, una sezione mediana, una fine. Due persone *non possono* fare esperienza dello stesso flusso di eventi né nel medesimo ordine di esecuzione né per la stessa durata; inoltre il tempo musicale cessa di essere un connotato che può essere stabilmente identificato, misurato e reso disponibile." Inoltre le modalità di produzione di "Cathedral" caratterizzano in senso *plurale e non-lineare* la presentazione e la ricezione dei flussi sonori: "C" è un ordine ovvio, ma non una mappa reale. Non c'è un modo "corretto" per andare avanti: lo sviluppo può avvenire in ogni direzione e nello stesso tempo, seguendo i percorsi di una *matrice di possibilità* interattive. Si tratta di una pluralità di forme, tutte quante individuali, che possono aver luogo; nessuna delle quali è simile, da qualsiasi punto di vista le si voglia osservare" [16]. L'orizzonte concettuale e formale delineato da Duckworth sembra costituire lo sviluppo coerente di alcune idee espresse da Marcos Novak. Da cyber-architetto e musicista Novak considera la navigazione in un senso assai diverso da quello accennato nel paragrafo precedente. Per questo egli è giunto ad elaborare il concetto di 'musica navigabile': cioè un ambiente da cui sia possibile *estrarre* una molteplicità di percorsi musicali, non importa se convenzionali oppure no, ma che sono in ultima istanza riferiti alle traiettorie tracciate al suo interno da chi interagisce con quel tipo di ambiente. Per Novak infatti "la musica deve cessare di essere rappresentata come una linea che va dal punto A al punto B": in altre parole occorrerebbe che la musica cessasse di essere intesa come organizzazione *del suono nel tempo* e iniziasse ad essere pensata come una attività di organizzazione applicata ad una *matrice* di potenzialità soniche, visuali e comportamentali [18].

## AMBIENTI SONORI NAVIGABILI

La musica nella rete può essere pensata come un processo potenziale che si declina nel tempo attraverso piani di azione diversi. A questo tipo di musica - progettata per ambienti sonori navigabili - è consentito giustificare l'assenza di relazioni tra progetto musicale e ascolto, mettendo in secondo piano l'identità dell'opera, l'unicità dell'autore e la tradizionale relazione fra i due. Negli ambienti sonori navigabili i contenuti estetici e i valori emozionali impliciti nella dimensione dell'ascolto entrano in risonanza con le pratiche comunicative della rete e ne sono condizionati, dando luogo alle seguenti possibilità: i processi musicali vengono *modificati dalle azioni* degli utenti della rete; i flussi sonori provenienti da più sorgenti vengono riconosciuti come un *unico processo*; la *successione dei gesti interattivi* contribuisce alla elaborazione della forma dei processi musicali in corso. L'interpretazione dei flussi sonori deve fare i conti con la presenza di eventi di tipo erratico, misurarsi con fattori che provocano scostamenti temporali e subire altre conseguenze come la perdita o il mascheramento di alcuni tratti sonori. In tali condizioni tutto ciò che un compositore può rivendicare è la progettazione di cornici di possibilità aperte al contributo dei navigatori, ben sapendo che i contenuti musicali da lui prodotti potranno essere selezionati ed estratti, miscelati e ripresentati sotto una nuova forma o in un nuovo contesto. Di conseguenza l'interpretazione dei processi musicali presenti nell'ambiente della rete richiede una metodologia di indagine che sappia valutare le *variazioni di status* (sorgenti, autori, contesto, relazioni intermediali) dei flussi e degli oggetti sonori e sia in grado di interrogarsi riguardo ai modelli di manipolazione e di interlocuzione che la rete propone. Notiamo tuttavia che, quale che sia lo status (originale, derivato, estratto, condensato) attribuito agli oggetti sonori presenti nella rete, non viene in alcun modo indebolita la ben nota definizione di *oggetto sonoro* proposta da Schaeffer: "Ogni fenomeno o evento sonoro percepito come un intero coerente, senza riguardo alla sua sorgente o significato" [19]. (Peralto, la frase potrebbe essere così estesa: .. 'al modo in cui viene prodotto e diffuso, e senza riguardo alle condizioni della sua ricezione'). E' chiaro che i controlli audiovisivi delle applicazioni musicali on-line e il grado di specificazione dell'informazione musicale sono determinanti nel facilitare o nell'ostacolare i processi di analisi e di elaborazione dei contenuti. Dunque solo un corretto feedback può aiutare i navigatori a contestualizzare i contenuti musicali nel modo più appropriato, a comprendere le soluzioni adottate per la realizzazione degli ambienti sonori navigabili e a valutare l'efficacia delle funzioni implementate nelle interfacce di navigazione.

## CONCLUSIONI

Nell'ambiente della rete le attività di produzione e di fruizione musicale si svolgono con modalità in parte diverse da quelle note. Ciò avviene per una serie di ragioni che possono essere così schematizzate. a) La natura cooperativa della rete induce i navigatori allo scambio di contenuti e alla costruzione di livelli di esperienza musicale condivisi. b) I profili di riconoscibilità degli oggetti sonori sono in relazione con le variabili dimensionali, temporali e socio-affettive degli universi acustici della rete. c) Il significato degli oggetti può essere modificato sia attraverso differenti modalità di ricezione (selettiva o cumulativa) sia attraverso diverse forme di lettura dei flussi di dati in arrivo. d) I processi interattivi della rete possono produrre risultati musicali che dipendono dal tipo e dalla successione degli interventi dei navigatori. e) Le difficoltà nella ricezione possono complicare il riconoscimento delle strutture musicali e causare un indebolimento degli indici emozionali dello spazio sonoro. f) Gli oggetti musicali della

rete tendono a perdere la natura di entità cui possono essere attribuite qualità fisse e misurabili. g) Nell'ambiente della rete la musica tende a perdere la natura di progetto espresso da un singolo autore e si trasforma in evento virtuale, le cui modalità di attuazione vengono affidate all'azione di più interlocutori.

## RIFERIMENTI

1. John Paradiso. The Brain Opera technology: New Instruments and Gestural Sensors for Musical interaction and Performance. *Journal of New Music Research*, Vol. 28, No. 2, pp. 130-149, 1999.
2. Geri Wittig. *Situated and Distributed Knowledge-production in Network Space*. [http://www.c5corp.com/research/situated\\_distributed.shtml](http://www.c5corp.com/research/situated_distributed.shtml)
3. Randall Packer. *Telemusic 2, inhabiting cyberspace*, 2000. <http://rhizome.org/discuss/view/29788/>
4. Erik Davis. *Roots and Wires. Polyrhythmic Cyberspace and the Black Electronic*, 1997. <http://www.techgnosis.com/cyberconf.html>
5. Lance Strate. *Experiencing cybertime: computing as activity and event*, Center for Teaching and Technology, Academic Computer Center, Georgetown University, Washington, DC. 1995. <http://www.helsinki.fi/science/optek/1995/n2/strate.txt>
6. Peter Anders. Antropic Cyberspace - Defining Electronic Space for First Principles, *Leonardo* 34:3, 2001 409-416, MIT Press, Cambridge MA.
7. Michel Bussi re, Performance space meets cyberspace: seeking the creative idiom and technical model for live music on broadband. *Leonardo Electronic Almanac*, 12(1), 2004. <http://mailman.mit.edu/pipermail/leaauthors/2004-January/000003.html>
8. Atau Tanaka. *Musical Implications of Media and Network Infrastructures: Perturbations of traditional artistic roles*. [www.csl.sony.fr/downloads/papers/2001/Atauh2ptm.pdf](http://www.csl.sony.fr/downloads/papers/2001/Atauh2ptm.pdf)
9. Atau Tanaka, Kasper Toeplitz. *Globalstring*, 1998. <http://www.sensorband.com/atau/globalstring/>
10. Atau Tanaka. *Constellations*, 1999. <http://www.sensorband.com/atau/constellations/>
11. Alvaro Barbosa. *Public Sound Objects*, 2005. <http://www.abarbosa.org/psa/index.html>
12. Dante Tanzi. Digital Time: Complexity and Perception, *Proceedings of International Computer Music Conference*, Sonic Arts Research Centre, Queens University Belfast, 24-29 August 2008.
13. Randall Packer. Composing with Media: zero in time and space. *Contemporary Music Review*, vol. 24 Part 6, 2005, 516-522. Editor Peter Nelson. Issue Editor Andrew Hugill.
14. Jerome Joy. *De la monstration aux t l chargements: pollinisations et  mancipation. Approches des dispositifs collectifs artistiques en r seau*. 2005. [http://joy.nujus.net/files/papers/2005\\_logs/2005\\_joy.html](http://joy.nujus.net/files/papers/2005_logs/2005_joy.html)
15. Gil Weinberg. The Aesthetics, History, and Future Challenges of Interconnected Music Networks", *Proceedings of International Computer Music Conference*, 2002 G teborg, Sweden, pp. 349-356.
16. Alvaro Barbosa. Displaced Soundscapes: a survey on network systems for music and sonic art creation, *Leonardo Musical Journal*, 15:53-59, MIT Press CA 2003.
17. William Duckworth. Perceptual and Structural Implications of "Virtual" music on the Web, *Ann. N.Y Acad. Sci.* 999:254-262. New York Academy of Sciences, 2003.
18. Markus Novak. Liquid architecture in Cyperspace, in *Cyberspace: First Steps*, pp 225, Ed Michael Benedikt, MIT Press, 1994, Cambridge, MA.
19. Pierre Schaeffer. *Trait  des objets musicaux*, Seuil, Paris, 1966.

# Rendering Acustico di liquidi in movimento rappresentati con modelli a particelle

**Carlo Drioli**  
Dipartimento di Informatica  
Università di Verona

**Davide Rocchesso**  
Dipartimento delle Arti e  
del Disegno Industriale  
Università IUAV di Venezia

## SOMMARIO

Il lavoro discute un approccio alla sonificazione di simulazioni numeriche rappresentanti liquidi in movimento. Vengono illustrate le caratteristiche dei modelli per la descrizione del moto del liquido e dell'emissione sonora, e vengono discussi i criteri per il controllo della sintesi audio mediante modelli fluidodinamici a particelle.

## Parole Chiave

Sonificazione di liquidi in movimento; Modelli fluidodinamici particellari; Emissione acustica per formazione di bolle.

## INTRODUZIONE

In numerose applicazioni, lo sfruttamento delle leggi fisiche che regolano un processo naturale si è rivelato una scelta efficace per capire, controllare o rappresentare il processo stesso (si vedano ad esempio gli studi sulla fonazione, il controllo predittivo di impianti industriali o l'animazione mediante computer graphics). La modellizzazione su base fisica dei processi ha dato prova di essere un mezzo efficace anche per la sintesi in tempo reale dei suoni in molti campi, inclusi la performance musicale, la realtà virtuale e aumentata, l'interazione uomo-macchina, la simulazione interattiva di processi fisici, l'animazione e la grafica al computer [7].

In questo lavoro si discutono alcuni aspetti riguardanti la riproduzione dei suoni generati da liquidi in movimento, nel particolare caso in cui la simulazione sia basata su modelli numerici a particelle della fluidodinamica. Il tema è particolarmente interessante per applicazioni di grafica e animazione al computer, dove fluidi animati realisticamente possono migliorare sostanzialmente la qualità percepita del rendering e della interazione. Tuttavia, sebbene questi modelli si prestino molto bene al calcolo delle grandezze necessarie a rappresentare graficamente un fluido in movimento, fornendo nuovi dati circa 30 volte al secondo, il loro peso computazionale a frequenze di campionamento audio diventa invece proibitivo. Si è dunque optato per una soluzione ibrida, in cui il moto delle particelle è usato per controllare opportuni algoritmi di sintesi audio che riproducono gli eventi sonori elementari, come la formazione di bolle o suoni di impatto in presenza di cavità risonanti, responsabili dell'emissione acustica nei liquidi. Nella progettazione di questi algoritmi si è seguito un approccio che tende a rappresentare gli eventi acustici in forma prototipica, con l'obiettivo di enfatizzarne anche in modo esagerato i tratti caratteristici ("cartonificazione"), piuttosto che ricercare il realismo tipico della tecnica del campionamento [5].

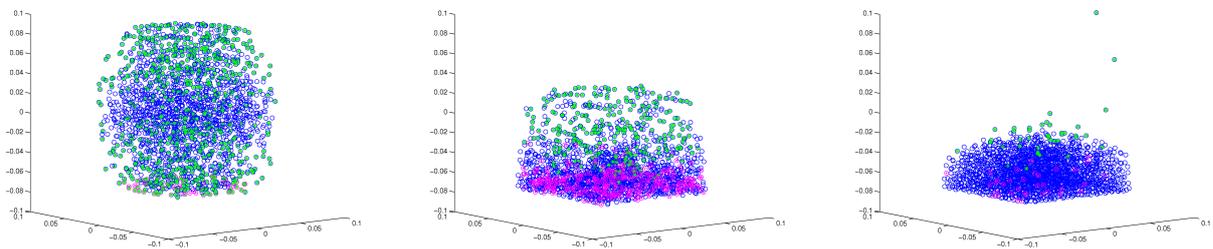
Nello studio viene presa in esame una configurazione sperimentale che rappresenta la caduta di un volume di liquido in un recipiente sottostante rigido e vuoto. Il tipo di setup sperimentale in esame appartiene ad una ampia famiglia di test di riferimento noti in letteratura con il nome di "collasso della diga" (breaking dam), rappresentante l'improvvisa rottura dei vincoli che mantengono in condizioni di riposo un volume di liquido, con il conseguente effetto di versamento in un bacino fino al raggiungimento del nuovo stato di riposo.

## METODO

Si fa riferimento qui ad un metodo di simulazione particellare dei fluidi, noto come smoothed particle hydrodynamics (SPH), che ha recentemente riscontrato un notevole interesse nell'ambito della computer graphics per la realistica delle animazioni di gas e liquidi [1,3]. In questo tipo di descrizioni numeriche, i fluidi sono modellati con particelle che rappresentano piccoli volumi del fluido. Ogni particella è descritta da un insieme di quantità fisiche (ad es., posizione, velocità, accelerazione, pressione, e densità) che sono aggiornate ad ogni passo di simulazione in base all'interazione con le particelle limitrofe. L'interazione è regolata dalle equazioni di Navier-Stokes per la conservazione della massa e del momento. Inoltre, è possibile includere nella simulazione forze esterne come quelle dovute alla gravità o a collisioni con oggetti solidi.

Per affrontare la sonificazione delle particelle in moto in modo rigoroso, sarebbe necessario calcolare la radiazione delle onde di pressione alla frontiera liquidi-aria a frequenze di campionamento audio. Questo porterebbe a un modello non utilizzabili a fini pratici a causa della elevata complessità computazionale, date le alte risoluzioni temporale e spaziale che questo implicherebbe. Il nostro approccio si propone di adottare risoluzioni temporali comunemente usate per l'aggiornamento dei parametri particellari (di solito da 10 a 30 frame per secondo) ed a identificare gli eventi per i quali vi sia probabilità di emissione acustica. A fronte della identificazione di un tale evento nel dominio del moto particellare, viene controllata l'attivazione di un evento acustico corrispondente. Un approccio simile si trova in letteratura per trattare la riproduzione dei rumori che accompagnano le vorticosità provocate da volumi di aria in movimento [8].

In una prima fase di studio, è stato individuato un insieme minimo di modelli di basso livello per la rappresentazione di eventi fondamentali responsabili di emissioni acustiche nei liquidi. In particolare, l'indagine si concentra su suoni che



**Figura 1. Evoluzione del sistema di 2000 particelle che modellano la caduta del liquido in un recipiente vuoto. I colori rappresentano alcune proprietà ed eventi di interesse, in particolare evidenziano le particelle con velocità nulla lungo la verticale (blu), le particelle in caduta (verde) e le particelle che subiscono variazioni repentine della velocità (probabili collisioni, viola).**

originano dalla formazione di cavità risonanti sotto la superficie del liquido (bolle), come si osserva nel gocciolamento o nel versare un liquido in un altro, e da impatti superficiali, come quelli che si osservano nella collisione tra due volumi liquidi, tra un solido ed una superficie liquida a riposo, o tra un liquido e una superficie solida [2, 6].

Nel seguito si illustreranno in primo luogo i criteri adottati per la sintesi degli eventi sonori elementari e si discuteranno poi i criteri per l'analisi, l'elaborazione e la sonificazione dell'evoluzione del liquido.

### Modelli di sintesi degli eventi sonori elementari

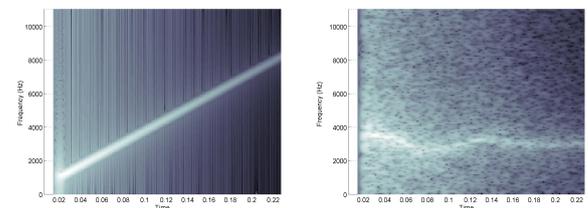
Il tipico suono emesso da una goccia che cade in un liquido a riposo ha origine nella formazione, al di sotto della superficie, di bolle d'aria le cui pareti oscillano a frequenze udibili ed il cui volume tende a ridursi rapidamente. La legge che descrive la risposta impulsiva di tale evento rappresenta la pressione acustica irradiata come una sinusoide smorzata e a frequenza crescente nel tempo:

$$p(t) = a \sin(2\pi f(t)t) e^{-dt} \quad (1)$$

in cui  $f(t)$  è la frequenza di risonanza tempo-variante,  $d$  fattore di smorzamento,  $a$  è l'ampiezza, e  $t$  il tempo. Per la simulazione di una singola bolla, sono state adottate le seguenti relazioni riportate in [6]: la frequenza iniziale è  $f_0 = 3/r$  (dove  $r$  è il raggio della bolla), il fattore di smorzamento è  $d = 0.043f_0 + 0.0014f_0^{3/2}$  e la frequenza istantanea è  $f(t) = f_0(1 + \sigma t)$  (in cui  $\sigma$  è il parametro che regola la velocità di incremento della frequenza). Il raggio iniziale della bolla è determinato da vari fattori, tra cui i principali sono la presenza di un livello minimo di liquido a riposo da permetterne la formazione e la massa del solido o liquido che vi cade dentro. La velocità di incremento della frequenza e, dunque, la velocità con cui la bolla collassa, è invece legata alla profondità alla quale si forma la bolla sotto la superficie del liquido.

Nell'istante di contatto della goccia o del solido in caduta con la superficie liquida, viene prodotto un rumore impulsivo dovuto a all'impatto tra le superfici e alla conseguente ecci-

tazione da parte di questo impulso di pressione delle cavità circostanti dovute ad aria intrappolata tra i volumi dei liquidi in gioco. Per la natura non rigida degli elementi che determinano le cavità, le risonanze hanno andamento variabile nel tempo e tipicamente è possibile osservare, in spettrogrammi di registrazioni di riferimento, delle "formanti" non costanti nel tempo. Si è dunque deciso di modellare il fenomeno con una sintesi sottrattiva in cui uno o più filtri IIR risonanti del secondo ordine sono eccitati da un breve impulso in corrispondenza dell'istante di contatto. Un modello simile, basato su uno schema di analisi/sintesi LPC, si trova descritto in letteratura per modellare il suono del battito di mani [4], dalle caratteristiche acustiche molto simili a quelle dell'evento di impatto in esame.



**Figura 2. Spettrogrammi degli eventi elementari impiegati: bolla (a sinistra) e impatto (a destra)**

### Analisi, elaborazione e sonificazione del moto particellare

Ad ogni frame, il processo di sonificazione si svolge in tre fasi consecutive:

1. lo stato delle particelle viene aggiornato attraverso le equazioni del modello SPH
2. la nuova configurazione particellare è analizzata e gli eventi che possono dar luogo ad emissioni acustiche sono individuati e analizzati

3. gli eventi individuati sono mappati in eventi audio di basso livello con opportuni parametri di sintesi

Per la descrizione dettagliata della risoluzione delle equazioni del moto particellare tramite modello SPH si rimanda alla letteratura citata (ad es. [1]). Si analizzeranno invece i due punti successivi.

L'analisi delle configurazioni istantanee del fluido ha un ruolo centrale nell'intero processo di sonificazione. A questo scopo, si è voluto identificare preliminarmente un insieme di configurazioni ed eventi notevoli osservati comunemente in situazioni sperimentali concernenti liquidi in moto. In Fig. 2 sono mostrati alcuni frame relativi all'evoluzione delle particelle del sistema. Dall'analisi dell'intera evoluzione sono immediatamente identificabili tre stati principali: inizialmente il volume di liquido precipita senza emettere suoni, fino al primo impatto della superficie inferiore con il fondo del recipiente. L'impatto del liquido con il contenitore vuoto costituisce il primo evento di interesse: in assenza di liquido a riposo non ci può essere formazione di bolle e gli unici eventi sonori in questa fase sono dovuti agli impatti liquido/solido (Figura 2,a)). Negli istanti successivi, il recipiente va via via riempiendosi e il livello crescente di liquido a riposo rende possibile, insieme con il suono da impatto liquido-liquido, anche l'emissione di suoni dovuti alla formazione di cavità sotto la superficie liquida (bolle) (Figura 2,b) e 2,c). Il processo continua in questa evoluzione che vede le particelle in caduta terminare la propria corsa nel liquido a riposo, eventualmente dopo essere rimbalzate in aria durante i primi impatti, fino al raggiungimento di una condizione finale di quiete.

Durante l'evoluzione del sistema, per ogni particella viene analizzata l'area circostante e viene redatta una lista delle particelle adiacenti (neighbors). Ciò permette di ottenere informazioni rilevanti quali la presenza di cluster di particelle che costituiscono volumi compatti, la presenza di particelle isolate, il verificarsi di configurazioni notevoli come le frontiere superficiali. Nel nostro caso, durante la fase di analisi vengono elaborate le seguenti informazioni usando i parametri di moto delle particelle (posizione, velocità, accelerazione) e le liste delle particelle adiacenti:

- presenza o meno di volume di liquido a riposo nel recipiente. In presenza di liquido a riposo viene stimato il livello  $h$  del liquido che si è depositato nel contenitore fino al frame presente
- presenza di particelle coinvolte in una collisione, sulla base delle informazioni di accelerazione
- identificazione di cluster di particelle e di particelle isolate in movimento
- formazione di configurazioni di frontiera, come superfici orizzontali di separazione liquido/aria, inferiori o superiori

Una volta terminata l'analisi della configurazione particellare, segue la fase di generazione dei segnali di controllo della sintesi audio. Nel design della mappatura fra stati del sis-

tema particellare ed eventi audio, vengono usati i parametri fisici delle particelle al fine di controllare la sintesi audio in modo coerente con l'evoluzione del sistema, secondo i seguenti criteri:

- la generazione di suoni da bolle d'aria è subordinata alla presenza di liquido a riposo, dunque di valori di  $h$  maggiori di zero
- per ogni collisione individuata viene generato un evento di tipo impatto e, se è  $h > 0$ , un evento di tipo bolla. Gli istanti di trigger degli eventi seguono una distribuzione uniforme all'interno del frame audio relativo ad un frame di visualizzazione.
- in presenza di cluster di collisioni, viene generato un evento di tipo impatto e bolla, con frequenze di risonanza inversamente proporzionali alle dimensioni del cluster
- valori di  $h$  crescenti (livelli maggiori di liquido nel recipiente) producono un incremento dei parametri che regolano le dimensioni possibili per le bolle e la velocità di incremento della frequenza (collasso della bolla).

In figura 3 si può osservare lo spettrogramma relativo al segnale di sintesi risultante dalla applicazione dei criteri illustrati. Si può notare come al momento del primo impatto siano presenti solo eventi di tipo impatto (che interessano la regione alta dello spettro) e non siano presenti eventi di tipo bolla. Dopo alcuni frame, essendo aumentato il livello nel recipiente, iniziano ad essere prodotti eventi di tipo bolla, dal caratteristico tratto crescente nello spettrogramma. Si può anche notare come la pendenza di queste curve che denotano l'incremento di frequenza sia minore all'inizio, per valori bassi di  $h$ , e vada aumentando con il riempirsi del recipiente.

## RISULTATI SPERIMENTALI E DISCUSSIONE

Sono state eseguite numerose simulazioni numeriche dell'esperimento descritto, e a partire da queste sono state valutate le varie fasi di analisi, mappatura e sonificazione. Per le simulazioni del sistema particellare è stato utilizzato, opportunamente modificato, il software di integrazione numerica SPH sviluppato da Takashi Amada nell'ambito di un progetto di animazione dei liquidi in tempo reale<sup>1</sup>

I parametri medi che regolano le distribuzioni statistiche dei parametri delle bolle (raggio e velocità di incremento della frequenza) e delle risonanze degli impatti sono stati aggiustati empiricamente, fino ad ottenere dei risultati acustici coerenti con il fenomeno allo studio. Sulla base di valutazioni informali, è possibile asserire che le sintesi audio, anche se isolate dal contesto visivo, risultano ben riconoscibili ed evocano efficacemente il versamento di un liquido in un recipiente, sebbene il livello di realistica sia sicuramente ancora distante da quello di un suono campionato. È interessante notare come ai fini di una efficace identificazione del fenomeno, più che la realistica timbrica, risulti essere importante la percezione del incremento del livello nel recipiente evocata dalle variazioni nel tempo delle caratteristiche degli eventi elementari, in particolare delle bolle. Il

<sup>1</sup>Disponibile all'indirizzo: <http://www.ss.ij4u.or.jp/amada/fluid/>

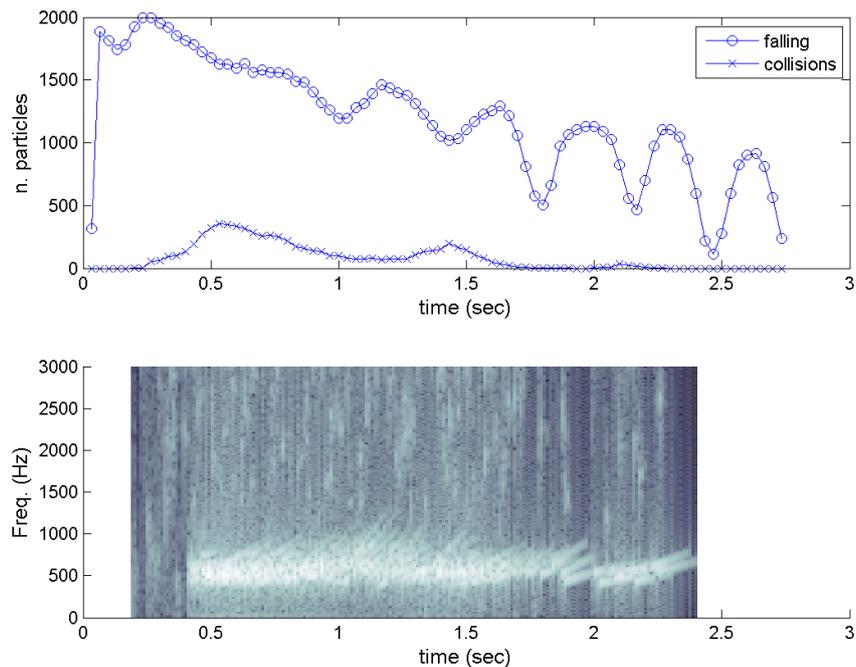


Figura 3. Risultato del controllo di eventi acustici a partire dall'analisi della configurazione particellare. La figura in alto mostra il numero di particelle in caduta e quelle coinvolte in collisioni. La figura in basso mostra lo spettrogramma del segnale risultante dalla combinazione di eventi di impatto e di bolle sotto la superficie.

risultato audiovisivo, evidenziando il sincronismo tra la visualizzazione del moto e le variazioni timbriche della sintesi sonora, è coerente con la percezione che si ha del fenomeno nell'esperienza quotidiana.<sup>2</sup>

## CONCLUSIONI

In questo lavoro si è affrontata la sintesi di suoni prodotti da liquidi in movimento mediante un approccio per modelli fisici. In particolare, si è adottato un modello numerico a particelle della fluido-dinamica per descrivere la caduta di un volume di liquido in un recipiente. I risultati sperimentali mostrano che anche con modelli di eventi acustici elementari relativamente semplici, il controllo basato su modelli fisici permette di evocare efficacemente l'esperienza di fenomeni complessi come il versamento di liquido in un recipiente. Ulteriori sviluppi sono auspicabili per quanto riguarda il raffinamento dei modelli elementari degli eventi acustici, per migliorare il grado di realismo, e nelle procedure di design delle mappe di controllo, attualmente progettate su base empirica.

## RIFERIMENTI

1. D. Enright, S. Marschner, and R. Fedkiw. Animation and rendering of complex water surfaces. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 736–744, New York, NY, USA, 2002. ACM.
2. G. J. Franz. Splashes as sources of sound in liquids. *The Journal of the Acoustical Society of America*, 31(8):1080–1096, 1959.
3. M. Muller, D. Charypar, and M. Gross. Particle-based fluid simulation for interactive applications. In *Proc. ACM SIGGRAPH/Eurographics 03*, pages 154–159, 2003.
4. L. Peltola, C. Erkut, P. R. Cook, and V. Välimäki. Synthesis of hand clapping sounds. *IEEE Trans. Audio, Speech and Language Proc.*, 15(3):1021–1029, 2007.
5. D. Rocchesso, R. Bresin, and M. Fernström. Sounding objects. *IEEE MultiMedia*, 10(2):42–52, 2003.
6. K. van den Doel. Physically-based models for liquid sounds. *ACM Transactions on Applied Perception*, 2(4):534–546, 2005.
7. K. van den Doel, P. Kry, and K. Pai. Foleyautomatic: Physically-based sound effects for interactive simulation and animation. In *Proc. ACM SIGGRAPH 01*, pages 537–544, 2001.
8. Y. Dobashi, T. Yamamoto, and T. Nishita. Real-time rendering of aerodynamic sound using sound textures based on computational fluid dynamics. *ACM Trans. on Graphics*, 22(3 (Proc. SIGGRAPH2003)):732–740, 2003.

<sup>2</sup>Alcuni file audio relativi alle simulazioni si trovano all'indirizzo: [http://mordente.sci.univr.it/~carlodrioli/liquids\\_modeling](http://mordente.sci.univr.it/~carlodrioli/liquids_modeling)

# *Echi tra le Volte*

## Un progetto di sound design rivolto agli edifici di culto

Andrea Taroppi

Conservatorio di Musica di Como

### SOMMARIO

Con il presente lavoro, derivato dalla mia Tesi Specialistica di Secondo Livello in Musica Elettronica e Tecnologie del Suono, discussa nel marzo 2008 presso in Conservatorio di Musica di Como, si intende indagare la relazione tra musica ed architettura, analizzando la possibilità di realizzare musica che possa considerarsi l'espressione in chiave acustica di uno specifico edificio.

Particolare interesse rivestirà il concetto di spazio, quale luogo di incontro e mutua influenza tra musica ed architettura.

Lo studio di questo tema, è stato la base teorica per lo sviluppo di *Echi tra le volte*, un'installazione sonora rivolta agli edifici di culto, dove il materiale sonoro deriva dai riverberi naturali del luogo, eccitati da treni di impulsi sinusoidali, a loro volta controllati in altezza da un algoritmo genetico.

### Parole Chiave

Musica e architettura, Informatica Musicale, Sound design, Algoritmo genetico, Composizione algoritmica.

### INTRODUZIONE: MUSICA E ARCHITETTURA

Lo scopo del progetto *Echi tra le volte* è quello di creare una musica che possa fortemente relazionarsi con uno specifico edificio. In questo senso, una musica siffatta dovrebbe poter essere considerata il naturale complemento, sul piano acustico, della fruizione, di norma esclusivamente visiva, di un edificio di pregio.

Ad esprimere assai efficacemente, l'idea che esista una stretta relazione tra musica ed architettura, viene spesso citata una famosa frase di Goethe secondo cui "L'architettura è musica pietrificata"; il che in effetti sembra anche lasciare intendere che musica e architettura siano due diverse facce della stessa medaglia, due modi diversi per esprimere qualcosa di comune ad entrambe.

Questo ponte che permetterebbe il legame fra la musica e l'architettura si trova nel comune impiego di misure e proporzioni da parte delle due discipline.

Sarebbe dunque nella matematica che possiamo trovare una sorta di linguaggio comune, grazie al quale prendono forma le espressioni musicali e quelle architettoniche.

All'origine di questo concetto possiamo indicare l'antica Grecia, ed in particolare Pitagora.

Infatti, fu questo grande pensatore che per primo definì una legge che lega le dimensioni di un oggetto produttore di un suono, con l'altezza del suono da esso generato; di fatto evidenziando un anello di congiunzione tra la percezione sonora e il mondo fisico.

Secondo Pitagora la coincidenza di musica, matematica e natura deriva dall'esistenza di tre musiche: quella strumentale, quella umana, suonata dall'organismo e quella mondana, suonata dal cosmo col movimento dei pianeti.

Ciò implicherebbe la possibilità di dedurre le leggi matematiche dell'universo da quelle musicali [1].

Questo concetto di armonia universale sta alla base della ricerca, costante nella storia del pensiero occidentale, di un principio unificatore che possa rivelarsi in grado di spiegare ogni aspetto dell'universo sensoriale, e viaggerà pressoché immutato attraverso i secoli da Platone agli Harmonices mundi di Keplero.

È possibile scorgere degli echi, financo nella moderna teoria delle stringhe, secondo la quale sarebbero i modi di vibrazione di stringhe nello spazio pluridimensionale a costituire le particelle elementari della materia.

Tornando al rapporto tra musica ed architettura, possiamo a questo punto parlare di una vera e propria estetica del numero che bene è descritta da Leon Battista Alberti (1404- 1472), il quale nel suo *de re Aedificatoria* afferma che: "*Quei numeri che hanno il potere di dare ai suoni la concinnitas, la quale riesce tanto gradevole all'orecchio, sono gli stessi che possono riempire di mirabile gioia gli occhi e l'animo nostro*" [cit. 2].

Pertanto, a garantire la bellezza di un edificio sta il fatto che le sue dimensioni rispettino precisi valori proporzionali, a loro volta corrispondenti a ben definiti rapporti armonici musicali.

Allo stesso modo, dovrebbe essere possibile realizzare una composizione musicale esteticamente valida, basandosi su rapporti numerici ricavati dalle misurazioni di un particolare edificio.

Secondo alcuni studiosi è esattamente questo che accade nel mottetto *Nuper rosarum flores* di Guillaume Dufay, che fu pressoché un contemporaneo dell'Alberti.

Le parti musicali del tenore e del controtenore del mottetto, sarebbero infatti costruite sulle dimensioni della cattedrale di Santa Maria del Fiore a Firenze, chiesa peraltro destinataria della composizione di Dufay [3].

Un esempio di architettura direttamente ispirata da una composizione musicale è invece la *Stretto House* di Steven Holl: una abitazione progettata come parallelo della *Musica per archi percussioni e celesta* di Bela Bartók [4].

Entrambe le opere inoltre impiegherebbero a livello strutturale un particolare rapporto numerico, il rapporto aureo<sup>1</sup>, che deve la sua fama proprio ad una sua forte ricorrenza nei luoghi più disparati, sia in natura, come in matematica che nell'arte [5].

Infine è possibile osservare il tentativo di ricavare elementi musicali da aspetti visivi legati ad un particolare luogo, anche in culture non occidentali. Un esempio è dato dagli Aborigeni Australiani, i quali sono soliti creare canzoni i cui profili melodici descrivono i profili paesaggistici di particolari aree territoriali [6].

Senza voler qui proporre un elenco di tutti i casi in cui si assiste ad una relazione fra musica ed architettura, impresa peraltro probabilmente impossibile data la quantità degli esempi a disposizione, quelli fin qui citati sembrano essere in linea con le affermazioni dell'Alberti, e cioè che il legame tra

<sup>1</sup> Il rapporto aureo si ottiene dividendo un segmento (AC) in due parti (AB, BC), tali che il rapporto tra l'intero segmento e la sezione maggiore sia uguale al rapporto fra la sezione maggiore e quella minore ( $AC/AB = AB/BC$ )

elemento visivo ed elemento acustico sarebbe da ricercarsi nelle caratteristiche quantificabili e misurabili di entrambi. In questo modo sarebbe addirittura possibile una sorta di traduzione dell'architettura in musica, e viceversa.

Bisogna però osservare che ciò porta inevitabilmente ad un percorso, per così dire a senso unico: dall'architettura alla musica o dalla musica all'architettura.

Ma se il nostro scopo è quello di trovare una relazione più profonda tra musica ed architettura, al di là di un processo di traslazione dovremmo guardare alla possibilità di dialogo, o di scambio informativo tra i due ambiti: dalla musica all'architettura e dall'architettura alla musica.

Un altro aspetto criticabile in una strategia di tipo analogico, basata cioè su corrispondenze numeriche, consiste nel fatto che questa non pare neppure una peculiarità delle sole due discipline in esame.

Infatti, anche la pittura fa uso di proporzioni e misurazioni, lo fa la scultura. In un certo senso anche la poesia deve fare i conti con ritmi misurati e proporzioni.

Tutto ciò suggerisce che forse non è negli aspetti quantificabili che possiamo trovare un più proficuo canale di comunicazione tra musica ed architettura. Piuttosto, potrebbe essere più promettente considerare con maggiore attenzione l'elemento spaziale, quale luogo di incontro ed interazione reciproca per le due discipline.

Lo spazio è infatti il luogo in cui forzatamente si trovano a coesistere musica e architettura, e nello stesso tempo è anche oggetto di trasformazione da parte di entrambe.

## DISEGNARE LO SPAZIO

L'influenza del luogo sugli eventi acustici che in esso avvengono è cosa nota. Lo *spazio acustico*, cioè "lo spazio tradizionale all'interno del quale l'esecuzione avviene" [7] agisce sempre sui suoni influenzando anche pesantemente il risultato acustico.

L'influenza dello spazio acustico può però essere impiegata a fini compositivi. Un esempio spesso citato in questo senso è l'impiego delle caratteristiche architettoniche della chiesa di San Marco a Venezia, da parte di compositori quali Andrea e Giovanni Gabrieli, al fine di creare particolari effetti di risonanza.

Vi sono inoltre studi che mostrano una relazione tra l'evoluzione degli organici orchestrali in Haydn e i tempi di riverbero più o meno lunghi delle sale da concerto frequentate dal compositore (si veda per questo e altri esempi [8], [9] e [3]).

È però possibile considerare anche casi in cui siano gli eventi sonori ad influenzare la percezione del luogo in cui questi avvengono; o, esprimendosi con le parole del titolo di questo paragrafo, casi in cui è il suono a "disegnare" lo spazio.

Innanzitutto, ciò accade ogni volta che, grazie alla tecnologia, viene restituita un'immagine tridimensionale della posizione e del movimento di eventi sonori, senza che gli stessi eventi sonori si trovino dove vengono effettivamente percepiti.

Un banale esempio è quello della cosiddetta sorgente fantasma, ottenuta quando lo stesso suono viene emesso contemporaneamente da due diffusori posti di fronte all'ascoltatore e alla stessa distanza da esso, restituendo appunto la sensazione che il suono provenga da una sola sorgente posta esattamente a metà fra i due diffusori.

Un effetto che si può facilmente sperimentare, normalmente ascoltando la parte vocale, pressoché di qualunque canzone di musica leggera con un impianto stereo.

Si tratta insomma del risultato a livello percettivo, la rappresentazione nella mente dell'ascoltatore, di quello spazio sonoro creato artificialmente, che Vidolin indica come *spazio virtuale* [7].

Inoltre, il suono può influenzare la percezione di un luogo agendo a livello estetico o emozionale, contribuendo talvolta non poco, a rendere più o meno piacevole la fruizione del luogo stesso.

Basti pensare alla musica di sottofondo nei supermercati e in generale nei cosiddetti non luoghi; musica che non è lì per essere ascoltata, ma piuttosto con la funzione di rendere più familiare l'ambiente, al fine ad esempio di stimolare gli acquisti.

Anche tutti gli eventi acustici (musica ed effetti sonori) per il cinema o per i videogiochi possono considerarsi esempi di intervento del suono sullo spazio, in quanto metodologie atte a modificare in una qualche misura la percezione dello spazio reale in cui si trova il fruitore del film o del videogioco, in quello virtuale da questi ultimi rappresentato.

## USARE LO SPAZIO

Un esempio forse inconsapevole, ma certamente efficace di legame tra musica ed edificio ospite, senza l'impiego di elementi numerici, può essere visto in *4'33''*: il famoso brano silenzioso di John Cage.

In quest'opera, a suonare non sono gli strumenti tradizionali ma lo stesso sala da concerto con i rumori casuali che vi vengono prodotti, i quali diventano allo stesso tempo materiale e prodotto della composizione.

Qualcosa di simile, si può trovare nell'esplorazione dell'effetto Larsen realizzata in *Modes of Interference* da Agostino Di Scipio.

In questo lavoro, il rientro nei diffusori del segnale microforato proveniente dagli stessi, e opportunamente controllato, diventa ancora una volta prodotto e materiale di partenza della composizione, la quale viene così ad essere attivamente modellata dall'acustica della sala, e non più semplicemente ospitata da quest'ultima [10].

In un altro lavoro di Di Scipio, *Interactive Island (Sea Light and Colors)* assistiamo ad un anche maggiore intervento dell'architettura ospite.

In questo caso infatti, un computer analizza le differenze tra i suoni di sintesi e quelli, captati con dei microfoni, presenti nella sala. I valori ricavati che dunque rappresentano le alterazioni timbriche dell'ambiente sui suoni prodotti, vengono a loro volta usati come segnali di controllo per la generazione sonora [11].

È questo dunque un chiaro esempio di musica che sfruttando l'influenza dall'ambiente in cui avviene l'esecuzione ne evidenzia le caratteristiche acustiche, realizzando un forte legame suono - ambiente.

Insomma per concludere riassumendo quanto affermato negli ultimi due paragrafi, siamo portati a credere che più che sfruttando analogie di tipo numerico, è tramite l'impiego dell'elemento spaziale che possiamo realizzare un legame più stretto, più interessante tra musica ed architettura.

A questo fine, almeno due aspetti vanno inoltre considerati come fondamentali:

1. Sfruttare, in senso musicale, le alterazioni che ogni edificio inevitabilmente reca agli eventi acustici presenti al suo interno.
2. Sviluppare strategie estetiche, che tengano conto della contestualizzazione degli eventi sonori in funzione dell'ambiente col quale vanno a relazionarsi. E, ripensando a tutte le tematiche relative agli studi sul paesaggio sonoro, ricercare tra suono e ambiente un rapporto ecologico e non schizofonico<sup>2</sup>.

<sup>2</sup> Per schizofonia si intende "la frattura tra un suono originale e la sua riproduzione elettroacustica" [12] e per estensione la dissociazione tipica nel caso della "muzak" tra la musica ed il luogo in cui viene riprodotta.

## ECHI TRA LE VOLTE

Da queste considerazioni ha avuto inizio il progetto *Echi tra le volte*.

La versione qui presentata è stata commissionata da Medium Size s.r.l. per essere realizzata nelle chiese, ed è stata eseguita nella chiesa di Santa Maria del Popolo a Vigevano (PV) dal 5 al 25 novembre 2007.

Il lavoro è organizzato in due distinte fasi: la generazione delle altezze e l'esecuzione vera e propria.

Nella prima fase un algoritmo genetico produce svariate sequenze di numeri, corrispondenti ad intervalli musicali.

Nella seconda, i dati precedentemente generati vengono tradotti in valori di frequenza (in Hertz), calcolandoli a partire da una nota base ricavata dalle dimensioni della chiesa, per essere inviati ad un generatore di impulsi sinusoidali.

Infine, l'uscita audio dal generatore viene distribuita nell'ambiente da un sistema quadrifonico.

## LA GENERAZIONE

Per la fase generativa è stato utilizzato il software Open Music (<http://recherche.ircam.fr/equipement/repmus/OpenMusic/>), con il supporto di GA: una libreria precedentemente sviluppata, per l'impiego dell'algoritmo genetico nella composizione assistita all'elaboratore [13].

Partendo da una popolazione iniziale costituita da valori casuali, sono state generati 7 altri passaggi evolutivi, per un totale di 8 fasi successive, sfruttando funzioni di idoneità ricavate dall'analisi della struttura intervallare del Canto Gregoriano [14].

Tutto il processo si evolve passando dalla casualità all'ordine, attraverso la progressiva aggregazione di gruppi via via più complessi di note, alla creazione di sequenze, dapprima semplici linee sviluppate attorno ad una *corda di recita*, entro lo spazio intervallare di una terza, sino a raggiungere uno stile più ornato o, in qualche modo, *melismatico*.

Tutto ciò progressivamente evitando tritoni o altri intervalli proibiti o troppo ampi.

Uno specifico operatore di mutazione è stato opportunamente realizzato per la creazione di gruppi di note sulla falsa riga dei neumi gregoriani.

Una volta raggiunta la massima complessità di "fioritura" delle sequenze, un'ultima fase evolutiva senza funzione di idoneità riporta la popolazione ad una situazione di disordine.

Così eventualmente, il processo può ripartire da una nuova popolazione casuale, per generare nuove sequenze.

## SINTESI E REALIZZAZIONE

I valori ottenuti nella fase generativa sono a questo punto caricati in 8 diverse tabelle di una patch del software Pure Data (<http://www-crcra.ucsd.edu/~msp/index.htm>), e poi convertiti in Hertz per essere inviati all'oscillatore sinusoidale.

Poiché di fatto, i valori di cui sopra corrispondono ad intervalli musicali definiti come semitoni da una frequenza base, per ottenerne la conversione in Hertz, si è semplicemente moltiplicato questa frequenza base per 2 elevato alla  $n/12$ , dove con  $n$  si intende il numero di semitoni, e cioè ognuno dei valori generati da Open Music e caricati nelle tabelle di PD.

La frequenza base è stata fissata a 135 Hz, valore ottenuto in funzione della massima distanza fra due diffusori (25 metri, in questo caso). Data una temperatura media di 10 °C e dunque una velocità del suono pari a 338 metri al secondo<sup>3</sup>, la frequenza corrispondente ad una lunghezza d'onda pari alla distanza fra i due diffusori è infatti 13.52 Hz.

I valori in frequenza così ricavati sono a questo punto pronti

<sup>3</sup> Si ricorda che la formula che lega la velocità del suono alla temperatura è  $C = 20.1\sqrt{273 + ^\circ C}$  [15]

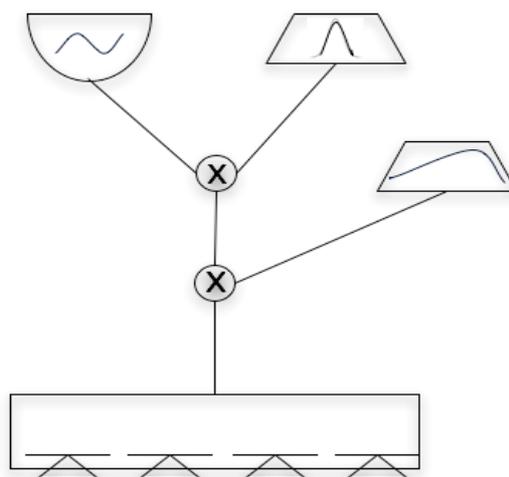


Figura 1. Schema della patch.

L'uscita dell'oscillatore sinusoidale (in alto a sinistra) è moltiplicata per il lettore continuo di tabelle (in alto a destra), dando così origine al treno di impulsi.

Un nuovo lettore di tabelle (in centro) definisce l'involuppo di ampiezza crescente.

Infine (in basso) si trova il distributore quadrifonico

per essere inviati ad un oscillatore sinusoidale.

Il treno di impulsi sinusoidali viene creato moltiplicando l'uscita dell'oscillatore per quella di un lettore continuo di tabelle (schema della patch - figura 1).

Il lettore di tabelle ha a disposizione 6 diverse curve: gaussiana, hanning, involuppo percussivo, ecc. e la scelta della curva avviene casualmente.

Ogni impulso così ottenuto avrà inoltre una sua specifica frequenza, risultante da una leggera alterazione (dell'ordine del 5-6 %) della frequenza relativa all'intero treno di impulsi, calcolata a sua volta, come abbiamo poc'anzi descritto, grazie alla moltiplicazione fra la frequenza base di 135 Hz ed uno dei valori presenti nelle 8 tabelle.

Ogni treno di impulsi ha una durata massima di 3 secondi; valore ottenuto calcolando il tempo di riverbero (RT60) in 4 punti diversi della chiesa, gli stessi dove andranno poi posizionati i diffusori, ottenendo appunto un valore medio pari a 3 secondi.

Poiché per definizione l'RT60 è il tempo necessario affinché il livello sonoro decada di 60 dB, un RT60 di 3 secondi corrisponde ad una velocità media di decadimento pari ad 1 dB ogni 50 millisecondi.

Pertanto, la durata di ogni impulso viene fatta variare casualmente fra i 40 ed i 60 millisecondi (dando dunque un valore medio pari a 50 millisecondi); mentre lo spazio intergranulare, il tempo cioè fra un impulso ed il successivo, è esattamente la metà del precedente, vale a dire varia casualmente fra i 20 e i 30 millisecondi.

A questo punto, all'intero treno di impulsi viene assegnato un involuppo di ampiezza, a curva crescente (si veda ancora figura 1) tramite una nuova moltiplicazione con un altro lettore di tabelle, ovviamente non continuo.

Tutto il processo si riassume dunque semplicemente nella trasformazione di ogni valore della fase di generazione, in un treno di impulsi con involuppo crescente e frequenza leggermente "instabile", che agendo da eccitatore del riverbero naturale dell'edificio, produrrà l'effetto percettivo di una nuvola sonora, nota di uno canto gregoriano virtuale.

Per quanto riguarda la velocità di esecuzione delle note di questo canto gregoriano sintetico, corrispondente in pratica alla velocità di lettura dei valori provenienti da Open Music e caricati nelle 8 tabelle, onde realizzare i treni di impulsi, si è

seguito un criterio evolutivo che segue quello realizzato durante la fase generativa.

All'inizio PD legge mediamente un valore ogni 30 secondi, una situazione quindi molto rarefatta, corrispondente alla lettura della popolazione di partenza che, lo ricordiamo, è costituita da valori casuali.

Poi, lentamente, il tempo fra una nota e l'altra inizia a diminuire, dapprima generando piccoli gruppi di 2 o 3 suoni, come ad imitare l'apparizione di casuali figure neumatiche.

Successivamente i neumi cominciano ad avvicinarsi dando origine a delle sequenze; prima relativamente brevi, come già accennato attorno ad una *corda di recita*, per poi lentamente evolversi in lunghe sequenze melismatiche dove le note si susseguono anche piuttosto velocemente (sino a 7-8 note al secondo).

Per evitare troncamenti data la velocità di lettura raggiungibile in questa fase (si ricordi che ogni nota può durare anche 3 secondi), vi sono più generatori di treni di impulsi, che uno speciale algoritmo distributore attiverà alla bisogna.

Parallelamente al ritorno alla casualità alla fine del processo generativo, giunti a questo punto la velocità di generazione delle note comincia a rallentare per riportare il "canto" alla situazione rarefatta di partenza.

Infine, per quanto riguarda la spazializzazione, il sistema è organizzato in modo da assegnare ogni singolo impulso ad una diversa coppia di diffusori.

Grazie alle dimensioni abbastanza ridotte della chiesa ospite, ed alla particolarità architettonica di non presentare la classica struttura a croce, praticamente eliminandone il braccio orizzontale, ma assumendo una forma vagamente ovoidale, è stato sufficiente impiegare solo 4 diffusori.

La posizione dei diffusori è così organizzata: due dietro l'altare, rivolti verso il fondo della chiesa, e dunque opposti rispetto alla posizione del pubblico, e altri due sulla balaustra sopra l'entrata, dove è situata la postazione dell'organista, anch'essi rivolti verso le pareti.

Questa versione di *Echi tra le volte* completa tutti i suoi passaggi in circa 7 ore. Tuttavia i valori generati da Open Music e caricati in PD non vengono aggiornati, sostituendoli con quelli di una nuova fase generativa, perciò tutto il processo ricomincerà con gli stessi valori corrispondenti alle altezze delle note.

Questo non rappresenta comunque un problema in quanto tutti i valori relativi ai tempi (durate delle note, degli impulsi, velocità e momento di produzione dei suoni, ecc.) invece varieranno.

Infatti è stato possibile lasciare in funzione *Echi tra le volte* ininterrottamente per circa 32 ore (dal venerdì mattina sino al sabato notte successivo) sempre ricevendo giudizi positivi da parte dei visitatori.

## CONCLUSIONI

Con *Echi tra le volte* è stato possibile realizzare il diretto intervento dell'edificio sul suono estendendo al di fuori del computer la catena di sintesi, sin dentro l'edificio, il quale viene effettivamente "suonato" come uno strumento musicale. Inoltre, l'impiego di materiale musicale derivato dal canto gregoriano, permette l'aggiunta di una nuova dimensione, un nuovo livello nella ricerca di un legame tra musica ed architettura.

Questa versione di *Echi tra le volte* è stata presentata insieme ad una mostra che partendo da pensatori quali Pitagora e Severino Boezio, ripercorre brevemente la relazione tra musica, geometria, matematica e astronomia, o in generale fra arte e scienza, come una storia del "pensiero unificante".

Opponendosi alla cosiddetta separazione dei saperi, e focalizzando l'attenzione sull'uomo e sulla percezione umana, l'intento è quello di suggerire un approccio olistico ed

ecologico.

Similmente, *Echi tra le volte* offre l'occasione di ripensare al rapporto tra musica ed architettura come ad una sinergia dove, spostando il fuoco dell'indagine dall'elemento misurabile alla dimensione della percezione umana, l'elemento visivo e quello sonoro traggono reciproco vantaggio dalla loro mutua influenza, restituendo un'esperienza fruitiva più ricca.

## RINGRAZIAMENTI

Innanzitutto vorrei indirizzare uno speciale ringraziamento al M° Giovanni Cospito del Conservatorio di Como, che ho avuto l'onore e il privilegio di avere come Relatore di questo lavoro conclusivo, ma soprattutto come Docente e Guida durante tutto il percorso di studio nella classe di Musica Elettronica.

Un sentito ringraziamento deve anche andare al Direttore M° Luca Bassetto, al corpo docente, al personale del Conservatorio e a tutti i colleghi e amici che ho avuto il grande piacere di conoscere durante questi anni a Como.

Allo stesso modo vorrei ringraziare il M° Antonio Bologna e l'architetto Fortunato D'Amico, di Medium Size, per averci fatto reso possibile la realizzazione di *Echi tra le volte* a Vigevano e per aver allestito la mostra; nonché le associazioni culturali "Le Zolle" e "Centro Ricerche per l'Ecologia Acustica" per tutto l'aiuto durante il periodo di allestimento.

## RIFERIMENTI

1. Piergiorgio Oddifreddi, *Questa è matematica per le mie orecchie*, 1997  
<http://www.extramuseum.it/giosci14/appr11.html>
2. Angela Pintore, *Il simbolismo musicale nell'opera di Leon Battista Alberti: dal De re aedificatoria al Sepolcro Rucellai*, Nexus Network Journal, Vol 6 n. 2, Autumn 2004  
<http://www.emis.de/journals/NNJ/Pintore-it.html>
3. Marta Michelutti, *Tra musica e architettura*, Conservatorio di Musica di Milano, a.a. 2002/2003  
<http://digilander.libero.it/initlabor/musica-architettura-michelutti/musica-architettura-marta1.html>
4. Steven Holl, "Stretto House", in *Pamphlet Architecture 16 – Architecture as a Translation of Music*, E. Martin, Ed. New York: Princeton Architectural Press, 1994, pp. 56-59
5. Mario Livio, *La sezione aurea storia di un numero e di un mistero che dura da tremila anni*, Milano, RCS, 2003
6. Kourosh Mavash, "Site + Sound : Space" in *Resonance essays on the intersection of Music and Architecture*, vol. 1, M. W. Muecke, M. S. Zach, Eds. Ames, IA: Culicidae Architectural Press, 2007, pp 53 -75
7. Alvise Vidolin, "Suonare lo spazio elettroacustico", in FAVARO, Roberto, (a cura di), *Musica e tecnologia domani Convegno internazionale sulla musica elettroacustica Teatro alla Scala 20-21 novembre 1999*, Lucca, LIM, 2002, pp. 143-156
8. Carmine Inannello, Sale da concerto e teatri d'opera, in, Sergio Cingolani e Renato Spagnolo (a cura di), *Acustica musicale e architettonica*, Torino, Utet, 2005, (pp. 663-704)
9. Steen Eiler Rasmussen *Ascoltare l'architettura*, (trad. a cura di Paolo Bettini),  
<http://www.unich.it/progettistidivanta/TRADUZIONI/Rasmussen-ASCOLTARE.pdf>
10. Agostino Di Scipio, *Using PD for Live Interactions in Sound. An Exploratory Approach*, 2006,

[http://lac.zkm.de/2006/presentations/lac2006\\_agostino\\_di\\_scipio.pdf](http://lac.zkm.de/2006/presentations/lac2006_agostino_di_scipio.pdf)

11. Agostino Di Scipio, "Iterated Nonlinear Functions as a Sound-Generating Engine", in *Leonardo*, vol. 34 N° 3, Cambridge (MA), MIT Press, 2001, pp. 249-254
12. Murray Schafer, *Il paesaggio sonoro*, Lucca, LIM editrice, 1985
13. Andrea Taroppi, *Composizione Musicale con gli Algoritmi Genetici GA: una libreria Open Music per la composizione musicale assistita basata su tecniche derivate dagli algoritmi genetici*, Tesi del Triennio di Musica Elettronica e Tecnologie del Suono, Conservatorio di Musica di Como, a.a. 2003/2004
14. Paolo Ferretti, *Estetica Gregoriana*, Roma: Pontificio Istituto Di Musica Sacra, 1934
15. Umberto Nicolao, *Acustica applicata per la diffusione sonora*, Segrate (MI), Editrice il Rostro, 2003



# Photos of GHOSTS

## (Photos of Grooves and HOles, Supporting Tracks Separation): Conservazione attiva di dischi fonografici per mezzo di immagini digitali

**Sergio Canazza**  
Laboratorio AVIRES  
Dipartimento di Scienze  
Storiche e Documentarie  
Università degli Studi di Udine

**Giovanni Ferrin**  
Laboratorio AVIRES  
Dipartimento di  
Filosofia  
Università degli Studi di Udine

**Lauro Snidaro**  
Laboratorio AVIRES  
Dipartimento di Matematica e  
Informatica  
Università degli Studi di Udine

### SOMMARIO

Nel campo delle memorie audio, la conservazione si articola in passiva<sup>1</sup> (difesa del supporto dagli agenti ambientali, senza alterarne la struttura) e attiva (trasferimento dei dati nel dominio digitale). Poiché il disco fonografico è caratterizzato da un'aspettativa di vita relativamente bassa – se confrontata con quella di altri monumenti – la conservazione passiva risulta insufficiente. Durante il trasferimento A/D del segnale audio viene persa molta informazione contenuta nel documento originale; solitamente queste informazioni vengono (almeno parzialmente) vicariate per mezzo di fotografie e schede catalografiche, memorizzate nella copia d'archivio assieme al segnale audio digitale [1], [2]. In questa sede si propone di minimizzare questa perdita d'informazione mediante il trasferimento dei dati nel solo dominio video. La copia d'archivio sarà quindi costituita solo dalle riprese fotografiche del disco fonografico. Per creare copie d'accesso è stato implementato un algoritmo in grado di estrarre il segnale audio dal documento fotografico.

### Parole Chiave

Conservazione attiva di documenti sonori; dischi fonografici; codicologia del suono.

### INTRODUZIONE

La conservazione delle opere musicali memorizzate su disco pone problemi nuovi rispetto a quelli tradizionali della musica scritta o stampata su carta.

Il *disco piatto* inventato nel 1887 da Emil Berliner offriva prestazioni molto scadenti: il suono registrato presentava un rapporto segnale/disturbo estremamente basso, era afflitto da una forte distorsione e possedeva una risposta in frequenza estremamente limitata. Fino al 1925 tutte le registrazioni vennero effettuate in modo meccanico: un diaframma catturava mediante un corno le variazioni di pressione dell'aria trasformandole in movimenti meccanici paralleli della testina

<sup>1</sup> La conservazione passiva si suddivide a sua volta in indiretta – che non comporta il coinvolgimento fisico del disco – e diretta, nella quale il disco viene trattato, senza comunque alterarne struttura e composizione. Nella conservazione passiva indiretta rientrano: la prevenzione ambientale (che si esplica attraverso il controllo dei parametri ambientali che sono, in ordine decrescente di pericolosità per i dischi: umidità relativa, temperatura, inquinamento, luce), la formazione del personale addetto alla conservazione, l'educazione dell'utente. La conservazione passiva diretta comprende gli interventi di: realizzazione di custodie di protezione; spolveratura delle raccolte; disinfestazione degli archivi con gas inerti.

di incisione con la quale si realizzava un solco sulla superficie del disco. Solo nel 1925 Maxfield e Harrison [3], attraverso l'uso combinato del primo microfono e del *tubo sottovuoto*, resero possibile la conversione del suono in impulsi elettrici che venivano utilizzati per comandare la testina di incisione destinata a solcare la superficie del disco. Innovazioni fondamentali furono poi l'introduzione della modulazione multicanale a opera di Arthur Keller nel 1929, e l'invenzione, nel 1931, del riproduttore acustico ai cristalli da parte di Sawyer. Nell'arco di un secolo si è quindi passati dalle incisioni – prima meccaniche, poi elettriche – su disco a 78 giri (o con velocità di rotazione non standard, da 16 a quasi 100 rpm), al *Long Playing* in vinile a 33 giri, al 45 giri, all'introduzione – verso la fine degli anni '50 del Novecento – della codifica stereo e successivamente anche di quella quadrifonica, sebbene quest'ultima non abbia avuto la diffusione sperata. Hanno avuto una grande importanza storica anche i dischi – sempre analogici – registrati tramite la tecnologia magnetica<sup>2</sup>. Dagli anni '70 hanno poi goduto di una larghissima diffusione i dischi digitali magnetici (floppy-disk, mini-disk, hard-disk, hard-disk array), magneto-ottici e ottici (CD, Laserdisk, DVD, HD-DVD, Blue-ray Disc).

Questo eterogeneità di formati, tecniche di registrazione e supporti pone una serie di problematiche che rendono veramente complesse le operazioni di conservazione del supporto (conservazione passiva), di accesso ai dati e di digitalizzazione (conservazione attiva).

*Automatic text scanning* e *optical character recognition* vengono largamente usate nelle Biblioteche: diversamente da quanto accade nel testo, la conversione A/D del segnale audio inciso nei dischi fonografici viene solitamente eseguita per mezzo di un processo invasivo.

Com'è noto, esistono *laser turntable* in cui il tradizionale *pickup* è sostituito da un raggio laser. In questo modo il supporto non viene fisicamente toccato durante la lettura. Questi apparecchi soffrono di alcuni svantaggi: 1) sono molto sensibili ai graffi nel supporto e alle eventuali ondulazioni anomale del solco; 2) la capacità di riflessione della superficie del disco deve essere ottima; 3) sono vincolati all'informazione presente in un ridotto *laser spot*.

È interessante applicare strumenti sviluppati nel campo dell'*image processing* per estrarre i dati audio (informazione contestuale e segnale audio) dal documento fonografico.

<sup>2</sup> Basti pensare agli esperimenti, condotti da Pierre Schaeffer alla fine degli anni '40 del secolo scorso, di analisi del suono tramite *ascolto ridotto*.

Questo approccio permette: a) di conservare tutta l'informazione presente del supporto (sviluppandone un modello virtuale, in 2D o – indispensabile nel caso di dischi a incisione verticale – 3D); b) una lettura completamente non invasiva; c) la conservazione attiva di supporti le cui corrottele ne renderebbero impossibile una lettura tradizionale (Figura 1); d) di avviare un processo automatico e su larga scala di conservazione attiva da parte degli archivi discografici. In letteratura sono presentate alcune soluzioni [3], [5], [6] che fanno uso di HW molto costoso e/o non comprendono elaborazioni (*de-noise*, equalizzazioni, *de-wowed*) del segnale audio ricostruito.

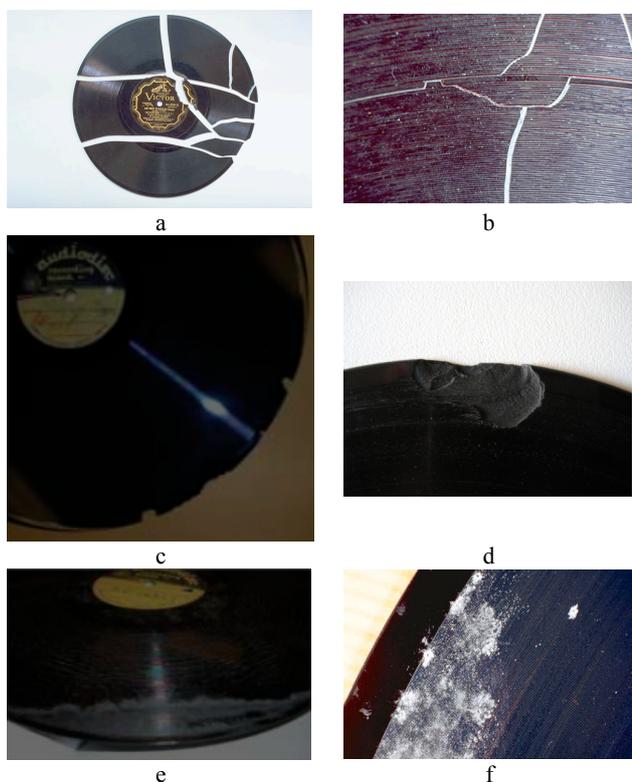


Figura 1. Esempi di corrottele in grado di rendere illeggibile il supporto fonografico. L'umidità è il fattore più pericoloso: mentre i dischi in gommalacca e in vinile sono meno inclini all'instabilità idrolitica, molte tipologie di *instantaneous discs* sono in grande pericolo di estinzione a causa dell'idrolisi (e). Tutti i dischi possono essere colpiti da micro-organismi (funghi e/o batteri) che crescono in ambienti caratterizzati da livelli di umidità superiori a 65 % RH (f). Inoltre, la base di metallo o di vetro dei dischi in gommalacca e in acetato li rende particolarmente fragili (a+d).

Dopo una breve introduzione delle metodologie di conservazione passiva dei dischi in funzione delle diverse composizioni del supporto, sono presentate alcune metodologie innovative per l'estrazione delle informazioni contestuali e del segnale audio. Saranno quindi mostrati due esempi di estrazione del segnale audio dal documento fotografico: il primo viene confrontato con una lettura tramite puntina; il secondo riguarda un disco spezzato, che non può essere letto con sistemi tradizionali. Entrambi i supporti considerati provengono dall'archivio personale di uno degli autori.

## TIPOLOGIA DEI DISCHI FONOGRAFICI

Il disco fonografico è stato, durante tutto il Novecento, il supporto più utilizzato per memorizzare segnali audio. Purtroppo, nonostante alcune tipologie di dischi si siano rivelati particolarmente robusti rispetto all'usura del tempo, molti fondi sono a rischio di scomparsa, soprattutto a causa dell'adozione, da parte di diversi archivi, di protocolli di conservazione non corretti. Una classificazione – funzionale rispetto alle diverse problematiche di conservazione – deve considerare (almeno) le seguenti categorie [7], [8], [9]:

- 1) *Shellac Discs*
- 2) *Instantaneous Discs*
- 3) *Microgroove Discs*

### *Shellac Discs*

Prodotti tra il 1898 e la metà degli anni '50 del Novecento<sup>3</sup>, ne esistono ancora oltre dieci milioni di esemplari. Questi supporti sono costituiti da una base (in alluminio, vetro o carta) accuratamente deumidificata, sulla quale è steso un sottile e uniforme strato di gommalacca<sup>4</sup> polverizzata in ogni lato in cui si vuole registrare l'audio<sup>5</sup>. Nonostante la rigidità che caratterizza questo supporto (che lo rende soggetto a rotture, v. Figura 1a), non sono emersi sistematici problemi di instabilità del materiale.

### *Instantaneous Discs*

Prima dell'introduzione dei nastri magnetici (avvenuta negli anni '40), gli *instantaneous discs* rappresentavano l'unico medium per la registrazione sonora che poteva venire letto immediatamente dopo l'incisione, evitando i lunghi processi industriali richiesti per la fabbricazione dei dischi industriali. Ne esistono ancora oltre tre milioni di esemplari: ognuno di questi è in copia unica e molti rivestono una grande importanza storica, sociale e culturale.

La maggior parte di questi supporti – i dischi in acetato – sono a grande rischio di scomparsa. Sono costituiti da una base in alluminio (o vetro, o acciaio, o carta), sulla quale è steso uno strato di nitrato o acetato di cellulosa. Col tempo questi supporti possono facilmente subire un processo idrolitico che fa restringere lo strato in acetato, oltre a renderlo fragile. La conseguenza è una delaminazione del disco che lo rende illeggibile (Figura 1e).

### *Microgroove Discs*

Dal 1948 i microsolfi in vinile hanno rimpiazzato i dischi in gommalacca e ancora oggi costituiscono uno dei supporti preferiti dagli audiofili. Il patrimonio di microsolfi è stimato in oltre 30 milioni di esemplari. I dischi in vinile sono composti da cloruro di polivinile (PVC da *polyvinyl chloride*) e da una percentuale (inferiore al 25%) di elementi secondari quali stabilizzatori, coloranti e sostanze antistatiche. Per assicurare all'impasto le proprietà plastiche necessarie, si effettua una copolimerizzazione dell'acetato di vinile con il cloruro di vinile. Il cloruro di polivinile è soggetto a fenomeni degenerativi se esposto alla luce ultravioletta o a sbalzi di temperatura. Quando un disco in vinile viene sottoposto a rapide variazioni di temperatura, il PVC che lo compone va incontro a un irreversibile processo di degrado termico, rilasciando dell'acido cloridrico (HCl). Per contrastarlo, in fase di produzione dei dischi, al cloruro di polivinile si aggiungono degli stabilizzatori chimici, che permettono, se non di evitare il processo, almeno di tenerlo sotto controllo, consumando l'acido cloridrico a mano a mano che viene rilasciato. I dischi

<sup>3</sup> Alcuni paesi (India e Pakistan tra gli altri) hanno in realtà continuato a fabbricarne sino alla metà degli anni '60.

<sup>4</sup> Macromolecola di origine animale prodotta dall'insetto *Coccus Laccae*.

<sup>5</sup> Alcuni dischi venivano infatti incisi solo su una facciata.

in vinile, non essendo un prodotto multistrato, non sono soggetti al fenomeno della delaminazione (contrariamente ai dischi in gommalacca o in acetato), provocato dalla diversa risposta dei singoli componenti alle condizioni ambientali. Il vinile si è dimostrato il più stabile tra tutti i materiali utilizzati nella produzione di supporti per la registrazione analogica dei suoni (nastri magnetici compresi).

Per una descrizione dettagliata delle diverse tipologie di dischi fonografici si vedano almeno [1], [2], [8], [9].

### Cause di deterioramento

Le principali cause che concorrono a rendere instabile e difficilmente leggibile il supporto fonografico possono essere riassunte in:

- 1) umidità e temperatura
- 2) deformazioni meccaniche
- 3) polvere e scorie

L'umidità è il fattore più pericoloso. In particolare, i dischi in acetato sono molto sensibili all'idrolisi. I funghi possono aggredire facilmente i supporti se il livello di umidità cresce oltre il 65% RH (Figura 1f).

Temperature oltre i 40° C sono pericolose, soprattutto per i dischi in vinile. In generale la temperatura determina la velocità delle reazioni chimiche come l'idrolisi e dovrebbe quindi essere mantenuta bassa e soprattutto stabile (per evitare modifiche nelle dimensioni del supporto).

È di particolare importanza controllare in modo simultaneo la temperatura e l'umidità. Ad esempio, è gravemente sbagliato raffreddare l'ambiente senza contemporaneamente deumidificarlo: diversamente si rischierebbe un inaccettabile innalzamento dell'umidità relativa che può incoraggiare la crescita di funghi. La Tabella 1 riporta le temperature e l'umidità relativa consigliate per la conservazione e l'accesso dei dischi fonografici.

Tabella 1. temperature e l'umidità relativa consigliate per la conservazione e l'accesso dei dischi fonografici.

	°C	±/24h	±/anno	RH	±/24h	±/anno
Conservazione	5÷10	±1	±3	30	±5	±5
Accesso	20	±1	±3	40	±5	±5

È imperativo evitare che graffi e altre deformazioni causate da utilizzi impropri dell'apparato di lettura pregiudichino l'integrità meccanica del supporto. In questo senso, è di grande importanza seguire corretti protocolli di conservazione passiva indiretta (in particolare: addestramento del personale). In generale, i dischi fonografici devono essere archiviati verticalmente: l'unica eccezione è costituita da alcune varianti degli *instantaneous discs*.

Polvere e piccoli detriti deviano la puntina dal corretto percorso di lettura, causando rumori impulsivi in banda audio. Le impronte digitali costituiscono un collante ideale per le scorie. È quindi essenziale che l'ambiente dove sono archiviati i dischi e l'area operativa (*access area*) siano puliti e privi di polvere.

Per una descrizione dettagliata delle cause di deterioramento e relativi protocolli di conservazione, si vedano almeno [10], [11], [12], [13], [14], [15].

L'obsolescenza dei sistemi di lettura tradizionali ha già ridotto le possibilità di recuperare le informazioni memorizzate in alcune tipologie di *instantaneous discs*. Poiché non è lecito ipotizzare che tutti gli archivi siano dotati di sistemi di lettura professionali, non è forse totalmente priva di interesse la proposta di un sistema per l'estrazione del segnale audio e

dell'informazione contestuale dal documento fotografico. In questo modo, considerato la semplicità d'uso, il basso prezzo e la grande diffusione di apparecchi in grado di acquisire immagini digitali (scanner e fotocamere), tutti gli archivi sono messi in grado di intraprendere immediatamente azioni di conservazione attiva assolutamente non invasive.

### ESTRAZIONE DELL'INFORMAZIONE CONTESTUALE

Per l'estrazione automatica dell'informazione contestuale sono stati considerati strumenti sviluppati nel campo della *computer vision*. Per calcolare l'eccentricità del disco (che causa oscillazioni del *pitch* o distorsioni del suono) è stata sfruttata la letteratura prodotta nel campo dell'*iris detection* (sistemi per il riconoscimento dell'iride) [16]. In particolare, il nostro sistema utilizza l'operatore integrodifferenziale sviluppato per il rilevamento dei confini della pupilla e dell'iride [17]. L'operatore funziona come rilevatore di discontinuità circolari, calcolando le coordinate del centro e la dimensione del raggio della circonferenza contenuta nell'immagine. In questo contesto, viene utilizzato per estrarre i contorni del disco e quindi per rilevare la circonferenza del foro centrale (Figura 2). L'immagine viene acquisita garantendo il parallelismo tra il piano focale e il disco. Il sistema calcola quindi il raggio, il centro del disco e la deviazione tra questo e il centro del foro.

Per separare automaticamente le tracce viene invece usato un rilevatore a soglia sulla funzione intensità luminosa calcolata tra lo *specchio* e il margine esterno del disco fonografico.

Questi dati vengono utilizzati per l'estrazione dei campioni audio e quindi memorizzati nella copia d'archivio per eventuali elaborazioni future (correzione del *pitch* sul segnale audio digitale).



Figura 2. rilevamento automatico dei contorni del disco e del foro centrale. In questo caso, la deviazione tra i centri delle due circonferenze è di 0.22 cm.

### ESTRAZIONE DEI CAMPIONI AUDIO

Il sistema HW/SW prototipato dagli autori possiede le seguenti caratteristiche: a) separa automaticamente le tracce presenti nel disco; b) l'intervento dell'utente è ridotto al minimo; c) utilizza HW *low-cost*; d) è robusto rispetto a polvere, graffi e ondulazioni anomale dei solchi; e) utilizza innovativi algoritmi di riduzione del rumore e di *de-wowed* [18]; f) applica una curva di equalizzazione, opportunamente scelta (in funzione della data di incisione del disco, dell'etichetta discografica, ecc.) da una banca dati creata appositamente.

Il sistema impiega uno scanner opportunamente modificato al fine di mantenere, durante la ripresa fotografica, un allineamento ottimo della lampada rispetto al supporto (irradiazione della luce coassiale al disco) per tutti i solchi, senza utilizzare luce ultravioletta. Lo scanner usato (HP

ScanJet 4890 Photo) è impostato con una profondità di 8 bit (a livelli di grigio) e con una risoluzione di 4800 dpi, senza correzione digitale.

Il software riconosce automaticamente il centro del disco e il raggio (v. sezione “Estrazione informazione contestuale”): informazioni necessarie per creare un modello *rettificato* dei solchi. Utilizzando un rilevamento a soglia sul grafico della funzione intensità luminosa di una sezione dell’immagine del disco (tra lo *specchio* e il bordo) sono rilevate le diverse tracce. Il solco è modellato utilizzando la curva di intensità luminosa dei *pixels* dell’immagine digitale. Per ottenere i campioni sonori è sufficiente seguire – con una velocità lineare calcolata sulla base della velocità angolare stimata del disco – i modelli dei solchi: l’ampiezza sonora è proporzionale alle oscillazioni del solco.

L’uscita del sistema consiste nei file audio (formato BWF) contenenti l’informazione delle diverse tracce contenute nel disco, su cui è possibile applicare una curva di equalizzazione e diversi algoritmi di *denoise*. La Figura 3 riassume il processo di acquisizione.

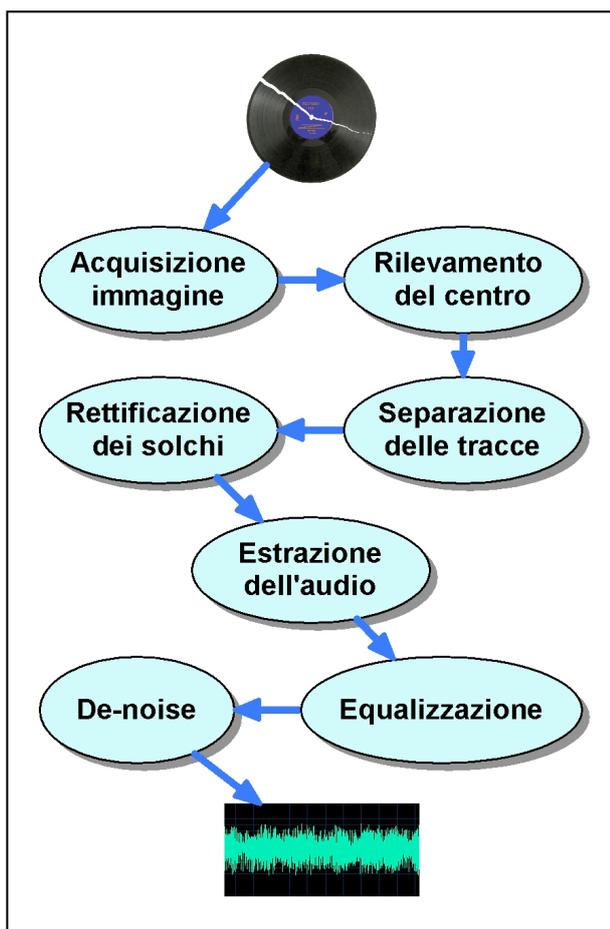


Figura 3. Schema a blocchi del sistema di acquisizione.

Il software è stato prototipato in linguaggio Matlab (versione utilizzata: 7.4.0): l’estrazione del segnale audio da un disco 78 rpm (durata: 3’ e 30”) impiega circa 3 ore (utilizzando un Apple MacBook Pro equipaggiato con processore Intel Core 2 Duo a 2.2 GHz, 2 GB di memoria), raggiungendo un fattore, approssimativamente, 50 volte superiore al tempo reale. La frequenza di campionamento reale è di 96 kHz: il segnale viene ricampionato a 44.1 kHz tramite la funzione *resample* di Matlab.

## DUE STUDI DI CASO: *STRANGE FRUIT* E *I-FEEL-LIKE-I'M-FIXIN'-TO-DIE RAG*

**Strange Fruit** (MT 117). New York, 20 Aprile 1939, Frank Newton & Café Society Band (Commodore): Kenneth Hollon (ts), Sonny White (p), Jimmy McLin (g), John Williams (b), Eddie Dougherty (d), Billie Holiday (v).

*Strange Fruit* è probabilmente la più famosa canzone di condanna al razzismo americano. Il testo è mutuato da una poesia scritta da Abel Meeropol (musicata dallo stesso autore) e pubblicata nel *The New York Teacher* (1937) con lo pseudonimo Lewis Allan (dal nome dei suoi due figli deceduti in tenera età) [19]. Meeropol intendeva esprimere l’orrore che gli ha provocato la foto scattata da Lawrence Beitler ([http://www.thehypertexts.com/Mysterious\\_Ways/Mysterious\\_Ways\\_Lynching\\_of\\_James\\_Cameron.htm](http://www.thehypertexts.com/Mysterious_Ways/Mysterious_Ways_Lynching_of_James_Cameron.htm)), che era venduta come souvenir per turisti a 50 centesimi di dollaro e che ritraeva il linciaggio di Thomas Shipp e Abram Smith eseguito nel Marion, Indiana, il 7 agosto 1930.

La versione della canzone qui considerata è stata inserita nel Grammy Hall of Fame nel 1978<sup>6</sup> e inclusa nella classifica *Songs of the Century*, redatta dalle Recording Industry of America e National Endowment for the Arts.

La qualità sonora della registrazione – a causa, in particolare, della ripresa microfonica – è al di sotto della media del tempo. Probabilmente la Commodore, nelle sue prime sedute di registrazioni, non poteva permettersi lo stesso equipaggiamento di cui erano dotate la CBS o la RCA.

La copia del 78 rpm considerata (disco in gommalacca a 78 rpm, 1939, Commodore, 526-A, WP24405A, Figura 4) è molto rovinata. Entrambe le superfici sono affette da graffi. Inoltre alcune zone sono notevolmente più scure di altre: si può ipotizzare che questa corruzione sia dovuta a lavaggi (effettuati prima che il supporto fosse acquisito dall’archivio personale di uno degli autori) in cui sono state impiegate sostanze aggressive. Le corrotture producono evidenti distorsioni durante la riproduzione mediante puntina. Il trasferimento è stato effettuato utilizzando un giradischi Rek-O-Kut-Rodine 3; il segnale è stato acquisito a 44.1 kHz mediante la RME Fireface 400. Non sono state applicate curve di equalizzazione.

Contemporaneamente il segnale è stato estratto dall’immagine fotografica del disco, acquisita a 4800 dpi, con una profondità di 8 bit (a toni di grigio), senza inserire alcuna correzione digitale. Sono stati rimossi i click e ridotto il rumore a larga banda mediante un algoritmo basato sul Filtro di Kalman Estes (EKF) [18]. Il segnale è stato quindi ricampionato a 44.1 kHz.



Figura 4. Etichetta del disco in esame (526-A).

In Figura 5 è mostrata una stima della densità spettrale di potenza di entrambi i segnali audio effettuata mediante il calcolo del periodogramma nell’intervallo 13s-18s (finestra di analisi di 2048 campioni). In banda audio non si notano

<sup>6</sup>[http://www.grammy.com/Recording\\_Academy/Awards/Hall\\_Of\\_Fame/#s](http://www.grammy.com/Recording_Academy/Awards/Hall_Of_Fame/#s)

artefatti introdotti dall'algoritmo di estrazione del segnale audio dal documento fotografico; contemporaneamente è evidente il miglioramento in termini di rapporto segnale rumore ottenuto per mezzo dell'EKF.

La Figura 6 presenta le evoluzioni temporali dei due segnali audio. È evidente la grande diminuzione di rumori impulsivi nella versione restaurata.

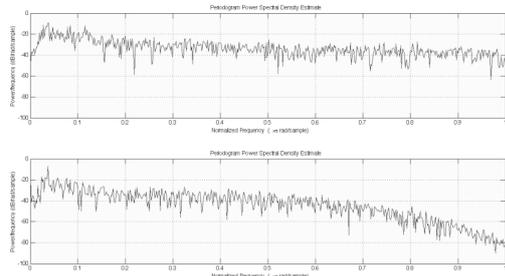


Figura 5. Periodogramma dei segnali audio rispettivamente acquisito mediante giradischi (sopra) ed estratto dal documento fotografico (sotto), relativi al disco 526-A, WP24405A. Finestra di analisi: 2048 punti. Intervallo di analisi: da 13s a 18s.

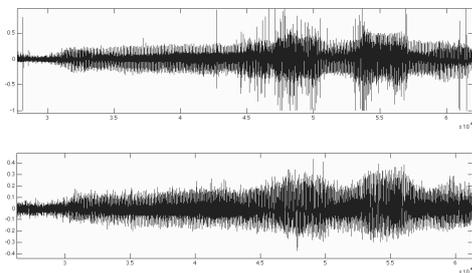


Figura 6. Evoluzione temporale dei segnali audio rispettivamente acquisito mediante giradischi (sopra) ed estratto dal documento fotografico (sotto), relativi al disco 526-A, WP24405A. Intervallo di analisi: da 12.5s a 23s.

***I-Feel-Like-I'm-Fixin'-to-Die Rag***, tratto da *Talking Issue Rag Baby No. 1: songs of opposition*, (Volume 1, Issue A: RAG-1001, 1965). Lato 1. Carl Schrager (washboard, kazoo), Bill Steele (b, washtub b), Barry Milton (g), Joe McDonald (v, g), Mike Beardsle (v). Lato 2. Pete Krug (v, g).

*I-Feel-Like-I'm-Fixin'-to-Die Rag* è probabilmente stata la più popolare canzone di protesta contro la guerra in Vietnam. È la *title track* del secondo *Long Playing* (Vanguard, 1967) del gruppo pop Country Joe and the Fish<sup>7</sup> e ha conosciuto un immenso successo internazionale grazie alla performance eseguita da Joe McDonald sabato 16 agosto 1969, durante il festival *Woodstock Music and Art Fair* tenuto nella Max Yasgur presso Bethel, New York. La prima edizione era però contenuta in un Extended Play edito nel 1965 (v. sopra). La confezione originale (una busta di dimensioni 7" e 1/2 x 10", v. Figura 7a) contiene: a) un disco microsolco a 33 rpm e 1/3 di

7"; b) un allegato di tre pagine con i testi e lo spartito delle canzoni.

La copia del microsolco in possesso di uno degli autori è rotta (il disco si presenta spezzato in due parti, v. Figura 7b) e quindi non è possibile riprodurre il documento per mezzo di un giradischi. Si è quindi estratto il segnale audio dal documento fotografico, acquisito a 4800 dpi, con una profondità di 8 bit (a toni di grigio), senza inserire alcuna correzione digitale. Il segnale estratto è stato equalizzato utilizzando una curva RIAA (l'incisione è del 1965). Sono stati rimossi i click mediante un algoritmo basato sul Filtro di Kalman Estes (EKF) [18]: poiché il disco – in vinile – è di buona qualità (almeno rispetto ai dischi in gommalacca) e, fatta eccezione per la grave corrottela che ne ha provocato la rottura, in buone condizioni, non è stato ridotto il rumore a larga banda. Il segnale è stato quindi ricampionato a 44.1 kHz.

La Figura 8 presenta una stima della densità spettrale di potenza del segnale audio effettuata mediante il calcolo del periodogramma nell'intervallo 21s+26s (finestra di analisi di 2048 campioni); in Figura 9 è raffigurata l'evoluzione del segnale audio. Da queste figure si evince che la corrottela non ha causato artefatti rilevanti nel segnale audio estratto dalla ripresa fotografica.



Figura 7. Fotografia della busta originale (a) e del disco rotto (b).

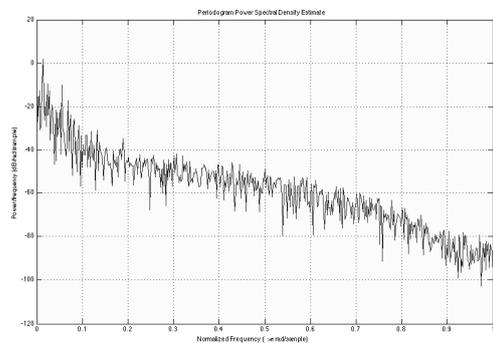


Figura 8. Periodogramma del segnale audio estratto dal documento fotografico relativo al disco RAG-1001. Finestra di analisi: 2048 punti. Intervallo di analisi: da 21s a 26s.

<sup>7</sup> Il nome del gruppo deriva da due locuzioni in uso nella cultura politica comunista statunitense: *Country Joe* era un popolare soprannome per Josif Stalin usato in America sin dagli anni '40 del Novecento, mentre *the fish* si riferisce a una frase – molto citata in USA – di Mao Tse-Tung, in cui il pensatore cinese sosteneva che il vero rivoluzionario doveva sapersi muovere attraverso la classe contadina come un pesce nell'acqua.

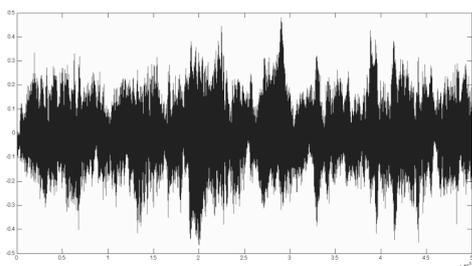


Figura 9. Evoluzione temporale del segnale audio estratto dal documento fotografico relativo al disco RAG-1001. Intervallo di analisi: da 12.5s a 23s.

## CONCLUSIONE

Negli archivi di dischi fonografici in cui sono state avviate azioni per la conservazione attiva dei documenti è già pratica comune memorizzare nella copia conservativa, assieme al segnale audio digitale, le fotografie del supporto (oltre che della copertina e degli eventuali allegati) e opportune schede catalografiche. Non è quindi improponibile (sia dal punto di vista dei costi, sia da quello della formazione del personale) pensare di acquisire documenti fotografici almeno di media qualità (da 4800 a 9600 dpi). La copia d'archivio sarà quindi costituita dalle fotografie del disco fonografico.

Gli archivi discografici che sinora non hanno avviato attività di trasferimento A/D possono in questo modo facilmente creare copie d'archivio dei dischi fonografici in loro possesso: diversamente dall'equipaggiamento professionale necessario per il trasferimento A/D del segnale audio, l'apparecchiatura per effettuare la ripresa fotografica è poco costosa e semplice da utilizzare.

Inoltre diventa finalmente possibile leggere dischi a) afflitti da gravi corrotture e b) di diversi formati senza per questo dover modificare l'equipaggiamento di lettura (come accade utilizzando i giradischi, dove è necessario regolare di volta in volta la velocità di rotazione, cambiare la puntina, modificare il peso del braccio, compensare la forza di *skating*, ecc.), con un evidente abbattimento dei costi per l'equipaggiamento tecnologico e per la formazione del personale.

L'algoritmo qui presentato – funzionante su hardware a basso costo – può essere vantaggiosamente utilizzato per la creazione di copie d'accesso estraendo il segnale audio e alcune informazioni contestuali direttamente dall'immagine dei solchi dei dischi.

Un'interessante ipotesi di sviluppo futuro potrebbe consistere nella realizzazione di uno strumento in grado – nei casi in cui nella copia conservativa è memorizzato anche il segnale audio, oltre al documento fotografico – di combinare l'informazione presente in entrambi i segnali audio (1-digitalizzato mediante puntina; 2-estratto dall'immagine), al fine di documentare in modo automatico le corrotture del supporto, che possono essere associate ai punti di minima somiglianza (secondo un qualche criterio) tra i due segnali.

## RIFERIMENTI

[1] Sergio Canazza. Note sulla conservazione attiva dei documenti sonori su disco. In Atti del Convegno annuale del Laboratorio per la Divulgazione Musicale (Ladimus). *Il suono riprodotto: storia, tecnica e cultura di una rivoluzione del Novecento*, EDT, Torino, 87-107, 2008.

[2] Sergio Canazza. Conservazione attiva e restauro audio dei 78 giri. Un caso di studio: Eternamente, In Canazza, S. e Casadei Turronei Monti, M. (a cura di), *Ri-mediazione dei documenti sonori*, pp. 695-715, Forum, Udine, 2006.

[3] Joseph Maxfield e Henry Harrison. Methods of high quality recording and reproduction of music and speech based on telephone research, *The bell system technical journal*, vol. 5, pp. 493-523, 1926.

[4] Stefano Cavaglieri, Ottar Johnsen e Frederic Bapst. Optical Retrieval and Storage of Analog Sound Recordings. *Proc. of AES 20th International Conference*, Budapest, Hungary, October 5-7, 2001.

[5] Vitaliy Fadeyev e Carl Haber. *Reconstruction of Mechanically Recorded Sound*. Lawrence Berkeley National Laboratory Technical Report 51983, 2003.

[6] Sylvain Stotzer, Ottar Johnsen, Frederic Bapst, Christoph Sudan e Rolf Ingol. Phonographic Sound Extraction Using Image and Signal Processing. *Proc. of ICASSP* Maggio, pp. 17-21, 2004.

[7] Marie-France Calas e Jean-Marc Fontaine. *La Conservation des documents sonores*. CNRS éditions, Ministère de la culture, Paris, 1996.

[8] Sarwan Khanna. Vinyl Compound for the Phonographic Industry. *Journal of the Audio Engineering Society*, 10-11, pp. 724-728, 1977.

[9] Leah Burt. Chemical Technology in the Edison Recording Industry. *Journal of the Audio Engineering Society*, 10-11, pp. 712-717, 1977.

[10] A.G. Pickett e M.M. Lemcore. *The Preservation and Storage of Sound Recordings*. Library of Congress, Washington D.C., 1959. Ristampato da ARSC, 1991.

[11] Jeremy Silver e Lloyd Stickells. Preserving Sound Recordings at the British Library National Sound Archive. *Library Conservation News*, 13, 1987.

[12] Dietrich Schüller. Preservation of Audio and Video Materials in Tropical Countries. *IASA Journal*, 7, pp. 35-45, 1996.

[13] Gilles Laurent. *The Care of Cylinders and Discs*. Milton Keynes: Technical Coordinating Committee, 1997.

[14] ARSC. *Audio Preservation: A Planning Study*. Association for Recorded Sound Collections (ARSC), Silver Springs, Maryland, AAA Committee, 1987.

[15] IASA TC-03. *The Safeguarding of the Audio Heritage: Ethics, Principles and Preservation Strategy*. Version 3, 2005.

[16] John Daugman. How iris recognition works. *IEEE Transactions on circuits and systems for video technology*, Vol. 14, n. 1, pp. 21-30, 2004.

[17] John Daugman. New methods in iris recognition. *IEEE Transactions on system, man, and cybernetics –part B: cybernetics*, vol. 37, n. 5, pp. 1167-1175, 2007.

[18] Sergio Canazza. Tecniche di filtraggio per il restauro audio: modelli a confronto. In Canazza, S. e Casadei Turronei Monti, M. (a cura di), *Ri-mediazione dei documenti sonori*, pp. 259-338. Forum, Udine, 2006.

[19] David Margolick. *Strange Fruit: Billie Holiday, Café Society, and an Early Cry for Civil Rights*. Running Press, Philadelphia, 2000.

# ***Atomi distratti* di Mario Garuti (2005)**

## **Analisi musicale e tecnologica ed elaborazione di un algoritmo di sintesi in ambiente Max/MSP per l'esecuzione della parte di tastiera elettronica**

**Marco Marinoni**

Conservatorio "Benedetto Marcello"  
di Venezia

### **SOMMARIO**

Questo articolo è focalizzato su *Atomi distratti*, un lavoro per ensemble e live electronics del compositore italiano Mario Garuti, ed è organizzato su tre differenti livelli:

1. analisi poetico-estesica del brano;
2. presentazione dell'ambiente esecutivo del live electronics approntato per la prima esecuzione;
3. elaborazione di un algoritmo di sintesi in ambiente Max/MSP per l'esecuzione della parte di tastiera elettronica.

Dopo una breve introduzione in cui si chiarisce l'idea all'origine delle scelte musicali operate nel brano, viene precisata la metodologia di approccio all'analisi.

Il primo livello include un'analisi dei cambiamenti contrastivi in cui viene mostrato come, in *Atomi distratti*, a un'organizzazione gerarchica *debole* (Imberty) corrisponda un'esperienza del tempo musicale coesa e unitaria, attraverso cui la rappresentazione simbolica e fantasmatica sottesa al brano (proprio la frammentazione del tempo interno, *mise en abyme* dell'ipotesi stessa) viene trasmessa integra ed intelligibile. Al fine di ottenere un tracciato delle immagini linguistiche, poetiche, musicali ed extra-musicali all'origine del brano viene utilizzato il mezzo dell'intervista, in cui il compositore è chiamato in prima persona a indagare il proprio lavoro.

Il secondo livello include due sotto-livelli: in primo luogo viene data descrizione del materiale pre-registrato e del suo ruolo nell'economia del brano, quindi una descrizione tecnologica dell'ambiente esecutivo del live electronics di *Atomi distratti* si accompagna a un'osservazione comparata longitudinale delle opere di Mario Garuti che includono in organico il mezzo elettronico, al fine di chiarire il ruolo del live electronics all'interno del brano e l'evolversi del rapporto che il compositore modenese intrattiene con l'elettronica nei suoi lavori. Anche in questo caso l'intervista è lo strumento privilegiato per accedere al mondo interno del compositore.

Il terzo livello è strutturato a partire da una *traduzione-interpretazione* in termini tecnologici delle descrizioni, registrate utilizzando il mezzo dell'intervista, che il compositore fornisce dei due preset di tastiera elettronica presenti in partitura. Di seguito viene descritto il progetto per lo strumento di sintesi, della cui realizzazione in ambiente Max/MSP si dà descrizione nel paragrafo conclusivo.

### **Parole Chiave**

Analisi musicologica; live electronics; algoritmo di sintesi.

### **1. INTRODUZIONE**

*Atomi distratti* è un lavoro di Mario Garuti per ensemble e live electronics che nasce, nelle parole dell'autore, dall'idea di comporre "una danza sospesa: questo comporta una forma di stroficità che permette di regolare la relazione tra prevedibilità e imprevedibilità. L'aspetto circolare di questa situazione implica variazioni interne di un oggetto sonoro in sé comunque riconoscibile" [1]. L'utilizzo del termine *danza* non è in questo caso da rapportare alla forma storicamente connotata, bensì riflette una particolare immagine metaforica e poetica dell'opera: un'idea musicale di partenza che ha informato e guidato il processo creativo fornendo prima un punto di convergenza verso cui indirizzare le strategie compositive messe in atto durante la stesura, quindi una lettura interpretativa possibile dei materiali musicali e della loro organizzazione.

### **2. METODOLOGIA**

Michel Imberty individua tre possibili direzioni su cui impostare il processo di segmentazione di un brano musicale:

1. percezione di cambiamenti qualitativi più o meno pregnanti nel flusso del tempo musicale;
2. attraverso modelli di riferimento appresi;
3. le suddivisioni sono determinate dalla struttura profonda dell'opera e testimoniano le modalità attraverso cui lo stile è stato percepito e decodificato dal soggetto [2].

La terza ipotesi ha come obiettivo un'indagine percettiva circa le differenti modalità di fruizione dell'organizzazione del tempo musicale da parte dei soggetti e attiene alla psicologia cognitiva e sperimentale. La seconda ipotesi non può parimenti essere applicata ad *Atomi distratti*, come del resto a gran parte della musica prodotta nel nostro tempo, poiché si basa sul confronto con forme storicamente connotate come ad esempio le forme classiche imposte dalla sintassi tonale. Sebbene parte della produzione contemporanea oggi privilegi la ripresa e rielaborazione di forme derivanti dalla tradizione, l'assenza di teorizzazioni precise ed esaustive circa la presenza e la riconoscibilità di uno o più *stili propri* all'interno della musica del nostro tempo (l'individuazione dello stile è operazione che non può essere compiuta nel darsi di una particolare corrente o epoca poiché necessita di una visione complessiva e *a posteriori*) rende impossibile qualsiasi osservazione di tipo comparativo se non con singole opere appartenenti allo stesso periodo. La prima ipotesi, basata sul cambiamento contrastivo, è quella che più si presta a essere presa come motore di organizzazione per questo tipo di osservazione analitica.

Il cambiamento contrastivo, per essere percepito, necessita che l'io percepisca non solo gli stati A e B ma la *transizione* da A

a B. Il passaggio costituisce la realtà percettiva della relazione tra le parti. B deve presentare una *qualità diversa* rispetto ad A. Il cambiamento introduce una *discontinuità* nel tessuto temporale attraverso due possibili modalità: *gerarchia* e *giustapposizione*.

Se lo stile è “gerarchia percettiva dei cambiamenti e delle unità che essi delimitano” [2], un brano in cui nessuna organizzazione gerarchica a livello macro-formale è individuabile, bensì solo una giustapposizione di parti estremamente differenti, ha come obiettivo e tratto distintivo la disgregazione dello stile. Attraverso un’analisi dei cambiamenti contrastivi mostreremo come *Atomi distratti*, in cui il grado di differenziazione tra le parti è massimo e qualsiasi forma di gerarchia è assente, risponda perfettamente alla definizione di *gerarchia debole*, la quale implica “numerosi cambiamenti tutti ugualmente pregnanti immediatamente percettibili, insieme a mutamenti più debolmente percepibili e in numero relativamente più grande, sia numerosi cambiamenti la cui organizzazione è incerta nel senso che non vi appare nessun contrasto chiaro o nessuna rottura netta” [2].

### 3. ANALISI MUSICALE

Il manoscritto è in formato A3 (cm 42 x 29.7). La stesura della partitura strumentale è stata ultimata in data 15 Dicembre 2005 a Sesto S. Giovanni. La progettazione e realizzazione della parte elettronica è stata ultimata in data 18 Gennaio 2006 (date e luogo sono indicate sull’ultima pagina del manoscritto, lateralmente a battuta 243). La prima pagina della partitura in notazione reca il titolo *Atomi distratti* su cui si innesta in modo (quasi) speculare *Amori distrutti* (vedi Appendice 1). In seconda di copertina si trova la dedica al Divertimento Ensemble. La prima esecuzione ha avuto luogo il 29 gennaio 2005 a Milano, Palazzina Liberty, all’interno della stagione concertistica Milano Musica 05, con i seguenti esecutori:

- Divertimento Ensemble
- Mario Garuti, tastiera elettronica
- Marco Marinoni, tracce audio e live electronics

Di seguito viene precisato l’organico strumentale:

- flauto alto / flauto basso
- oboe
- clarinetto in sib / clarinetto basso
- arpa
- tastiera elettronica
- glockenspiel / vibrafono (1° percussionista)
- marimba / 4 gongs (F#1, G1, A1, D2) (2° percussionista)
- pianoforte

In partitura viene inoltre indicata la possibilità assegnare a due esecutori la parte del primo percussionista. La durata riportata è 12’ e differisce di pochi secondi dalla reale durata della prima esecuzione (11’55’’).

A seguire le note all’esecuzione con le indicazioni relative al flauto e al pianoforte, quindi il progetto del live electronics, di cui si dirà in seguito.

Tra il progetto del live electronics e la partitura musicale, è riportato l’acrostico:

Attimidesideranti  
Trattinquieti  
Ondeterne  
Motoetero  
Instabilecerchio

Dolentel’ente  
Incrinature  
Stille  
Tramamata  
Retinfinite  
Amoridistrutti  
Tensegrità  
Trattinquieti  
Ineluttabile

Si tratta di quattordici parole organizzate in due gruppi, le cui iniziali lette in senso verticale riproducono il titolo del brano. Queste termini individuano e circoscrivono l’immaginario di *Atomi distratti*, rimandando a “strutture atomiche che durante il pezzo hanno atteggiamento centrifugo, perdono la loro funzionalità all’interno della danza sospesa. Ogni lettera delle parole ‘atomi distratti’ è iniziale di una considerazione verbale” [1].

Al fine di rendere queste informazioni accessibili e di chiarire le immagini che sono alla base delle operazioni simboliche e linguistiche attraverso le quali il brano è stato composto, ho chiesto all’autore di commentarle una per una: le considerazioni che seguono sono state raccolte in data 24 Aprile 2008 a Sesto S. Giovanni.

*Attimidesideranti*, (ovvero *attimi desideranti*) è il “tendere verso qualcosa che non conosci, un concetto di nostalgia legata a un desiderio sconosciuto, *sehnsucht*. E’ differente da un concetto di nostalgia verso un luogo conosciuto, un luogo verso cui si vuole tornare: questo implica direzionalità irreversibile e porta a una ipotetica ripartizione *A B C* del pezzo in cui a uno stato apparentemente autosufficiente segue la sua incrinatura e una zona conseguente inquieta, cioè *L’ente dolente*, pezzo per flauto basso e percussioni, incluso nei *Sei pezzi anestetici* (2003) che ho integrato in *Atomi distratti*, aggiungendo alcuni altri elementi: questo costituisce l’elemento B. La terza parte invece è totalmente caratterizzata dalla traccia audio AT\_20.wav e dall’elaborazione in tempo reale della stessa.

Questa tripartizione è motivata da un grado di omogeneità interna e presenta parti caratterizzate in modo totalmente differente”.

*Trattinquieti* (ovvero *tratti inquieti*) “si riferisce ad aspetti locali che costituiscono soprattutto la prima parte (parte A) in cui risulta evidente la fragilità degli elementi stessi. Tecnicamente parlando gli eventi sono composti da uno o pochi suoni isolati da silenzio, disposti a incastro. il gioco di ogni linea e anche l’incastro sono molto fragili”.

*Ondeterne* (ovvero *onde eterne*) “può essere rapportato all’ultima parte più elettronica”.

*Motoetero* (ovvero *moto etero*) e *Instabilecerchio* (ovvero instabile cerchio) “rimandano a una circolarità che si incrina. La parola *atomi* determina quello che è *L’ente dolente*”.

*Dolentel’ente* (ovvero *dolente l’ente*) “è la prima considerazione legata alla *D* di *distratti*. La distrazione implica una deviazione del proprio percorso con conseguente perdita di identità. Segue infatti la parola *incrinature*”.

*Incrinature* “è legata a *Trattinquieti*, che ritornerà”.

*Stille*, “in italiano *gocce* (elemento atomico, concluso in se stesso: puntillismo, anche se non weberiano), in tedesco è *tranquillità, silenzio*”.

*Tramamata* (ovvero *trama amata*) “ha in sé un verbo al passato, rimanda alla prima parte del pezzo ormai incrinata. C’è all’interno di *Atomi distratti* un decorso narrativo”.

*Retinfinite* (ovvero *reti infinite*): “la possibilità da parte del compositore di scegliere al momento qualsiasi possibile *estensione* che rimandi a certe necessarie *intensioni*. Devo agire su una certa necessità di cambiamento (decorso) ma, se l’idea è quella della fragilità, essa può essere espressa in modi

differenti. Concetto di campo unificato. Technè: penso che sia possibile relazionare tutto con tutto. Il problema è capire quali punti (o punto) ci sono in comune tra due oggetti che relazionano tra di loro. Ad esempio: 'poltrona' e 'ombrello'. Subito posso pensare interno/esterno o concavo/convesso. La poltrona ha forma convessa e riceve. L'ombrello ha forma concava pur ricevendo anch'esso (la pioggia dall'alto). Entrambi ricevono ma con due *opposizioni distintive formali*: sulla ricezione uno è morbido e riceve corpi pesanti, l'altro è teso e riceve pioggia o raggi solari, entità più eteree. Musicalmente parlando c'è questa *modalità di accostare* in senso sintagmatico parti formali essenzialmente differenti (le parti A, B e C della macro-forma) cioè si possono accostare sintagmaticamente oggetti differenti a livello paradigmatico all'interno della stessa unità formale o addirittura in sincrono (ad es. il rapporto tra gli elementi elettroacustici inclusi nelle tracce audio e gli oggetti sonori strumentali)".

*Amoridistrutti* (ovvero *amori distrutti*): "decadendo l'aulicità della *danza sospesa*, gli aspetti vitali sono meno direzionati secondo una costruzione formale e tendono più esplicitamente a una sorta di invasività del significante: la parte C che è una sorta di *metastasi dionisiaca* che si oppone alla *grazia* della prima parte. E' anche il continuo del titolo: volendo il pezzo si chiama *Atomi distratti / Amori distrutti* come da titolo inserito a specchio nella prima pagina. Uno specchio con anomalia".

*Tensegrità* "rimanda al concetto di *tensegrity* dell'architetto Buckminster Fuller che costruiva forme geodetiche basate sul triangolo equilatero, forma geometrica più coesa in assoluto che permette di costruire strutture ampie e stabili con la minor presenza di materia. Concetto che rimanda anche alla *tensegrità* di Castaneda e tutto l'immaginario che ne deriva".

*Trattinquieti* (ovvero *tratti inquieti*), "posta in posizione speculare rispetto alla sua presenza nella prima parola (*atomi*), ribadisce quello che è poi il tratto distintivo del pezzo".

*Ineluttabile*. "Non-reversibilità, nella formulazione di Giorgio Agamben. Rimanda anche al mio pezzo *Così è* (1990) per flauto in sol, sax alto e corno inglese, un lavoro molto statico, e a *...Con lume arsa* (1992), leggi *bruciata dalla ragione*. Ineluttabile quindi come qualcosa che non puoi controllare. Ovviamente è in relazione diretta con *Quantunque* (1998), titolo anche posto all'inizio del primo capitolo (*Qualunque*) del testo di Agamben, che parla della singolarità 'che scioglie dal falso dilemma che obbliga la conoscenza a scegliere fra l'ineffabilità dell'individuo e l'intelligibilità dell'universale' [3]".

Alcune *linee di comportamento* emergono dalle considerazioni dell'autore. Si tratta di *direzionalità* che si intersecano e si sovrappongono fino a formare una *texture di disposizioni interne* all'origine del brano. Di seguito sono elencate quelle particolarmente significative ai fini di questa osservazione:

- tripartizione formale del brano (tipo A B C), motivata dal grado di omogeneità interna delle parti, tra loro molto differenti.
  - A: uno stato apparentemente autosufficiente. Fragilità interna. Eventi composti da uno o pochi suoni isolati da silenzio, disposti a incastro. Condizione di *grazia* forse già in partenza ricordata, e idealizzata. Tema della perdita (lutto fantasticato) e dell'idealizzazione conseguente.
  - B: incrinatura della condizione A (*L'ente dolente*, già in *Sei pezzi anestetici*, 2003)
  - C: zona inquieta conseguente a B. Predominanza dell'elettronica persistente e con carattere di a-temporalità (*onde*

*etere*). *Metastasi dionisiaca* in opposizione alla *grazia* della parte A

- rottura della circolarità (non-reversibilità). Deviazione dal proprio percorso → perdita d'identità (distrazione);
- concezione atomica dell'oggetto sonoro. Il singolo evento è stabile e concluso in se stesso. L'instabilità (inquietudine) nasce apparentemente dalla relazione tra gli oggetti (ma è possibile che sia già in essi contenuta quale tratto ontologico proprio, *a priori*) e nel tempo porta gli stessi a degenerare in un'oscillazione paradossale (parte C);
- presenza di un decorso narrativo. Tema del ricordo. Ciò che è andato perduto è ormai sconosciuto, altro da sé. Tensione / nostalgia (*sehnsucht*) verso qualcosa che non si conosce (più) → non-reversibilità dei moti formali;
- tecnica compositiva: *action-painting* formalmente direzionato, istintuale ma consapevole che parte da un'analisi connotativa (basata sui tratti di *intensionalità* degli oggetti) e assume come terreno d'azione l'infinito possibile;
- attenzione all'equilibrio inteso come il massimo grado di coesione (formale) con il minor utilizzo di materia (sonora).

### 3.1 ANALISI DELLA MACRO-FORMA

*Atomi distratti* si divide in tre parti, collegate tra loro da zone ibride dette "di transizione". La suddivisione macro-formale è indicata in Tabella 1.

Parte	Battute	Indicazioni in partitura
A	1 – 122	semiminima = 96 ÷ 100
A→B	123 – 147	semiminima = 72
B	148 – 160	semiminima = 124; flauto in sol → flauto basso ( <i>L'ente dolente</i> ); tast. elettr. <i>Preset 1</i> → <i>Preset 2</i>
B→C	161 – 243	Start AT_18.wav, LE, cronometro
C	243 – fine	Fine partitura in notazione; start AT_20; traccia audio e LE soli

Tabella 1. Suddivisione macro-formale

#### 3.1.1 PARTE A

Non c'è frattura tra A e A→B, un unico arco narrativo di materiali sonori in continua trasformazione si conclude a battuta 147. Allo stesso modo la parte A non presenta in partitura ulteriori suddivisioni formali evidenti né interruzioni nell'omogeneità dei materiali, tuttavia è possibile individuare in essa quattro sotto-arcate narrative distinte, le prime tre di durata simile (circa 33-35 battute), la quarta di durata dimezzata (16 battute: manca l'arco narrativo discendente verso la zona di rarefazione che contraddistingue ogni transizione alla sotto-arcata successiva) con funzione di connessione alla zona di transizione A→B in cui i materiali subiranno una radicale intensificazione, già suggerita da alcune brevi zone di addensamento e maggiore instabilità presenti negli archi narrativi precedenti. Queste sotto-arcate sono a loro volta connesse mediante due zone di transizione (tra la prima e la seconda sotto-arcata e tra la terza e la quarta) individuate da battute vuote e una zona cadenzale costituita da due battute e individuabile nella regione caudale della seconda sotto-arcata. In Tabella 2 sono riportate le suddivisioni interne della parte A.

Battute	Durata (n° batt.)	Funzione
1 – 33	33	Sotto-arcata A1
34	1	Battuta vuota. Transizione alla sotto-arcata A2
35 – 68	34	Sotto-arcata A2
69 – 70	2	Regione cadenzale
71 – 105	35	Sotto-arcata A3
106	1	Battuta vuota. Transizione alla sotto-arcata A4
107 – 122	16	Sotto-arcata A4

Tabella 2. Parte A: suddivisioni interne

E' interessante notare come le zone di giunzione tra sotto-arcate siano costruite come degli *specchi con anomalia* (ritorna questa immagine, utilizzata dal compositore per commentare la considerazione *Amoridistrutti*, che oltre a essere inclusa nell'acrostico iniziale è anche lettura rovesciata e *obliqua* del titolo, forse addirittura come si ipotizzerà più avanti *contenuto fantasmatico latente* del brano: quindi un tratto distintivo ricorrente): infatti gli ultimi suoni di una sotto arcata sono i primi della sotto-arcata seguente e da subito innescano un vettore di modificazione / evoluzione dei materiali. La sotto-arcata A1 si conclude con i suoni mib, mi - si, do (arpa, clarinetto in sib – pianoforte, glockenspiel) e la sotto-arcata A2 inizia con i suoni mib, mi - si, do (arpa, clarinetto in sib – tastiera elettronica – piano, glockenspiel – tastiera elettronica). Al centro del chiasmo si trova un intervento dell'elettronica (traccia audio AT\_8). In maniera simile, gli ultimi suoni della sotto-arcata 2 (si, [la, mib, si, mi]) sono ripresi all'inizio della sotto-arcata A3 (mib, si, la) arricchiti da altre altezze (reb, mi) che introducono un graduale spostamento del piano armonico, e da figure ricorrenti (intervallo di settima maggiore si do alla tastiera elettronica in unisono con glockenspiel e parzialmente con il pianoforte). Similmente alla sotto-arcata A1, la sotto-arcata A3 si conclude sui suoni mi, mib, si, do (arpa, clarinetto in sib – glockenspiel e pianoforte all'unisono) mentre la sotto-arcata A4 inizia con suoni mib, mi - si, do (arpa, clarinetto in sib – pianoforte, pianoforte).

Le zone di addensamento e maggiore instabilità possono coinvolgere più strumenti (es. batt. 109 – 110), sono in genere caratterizzate da regioni ad articolazione complessa (es. oboe e flauto in sol da batt. 114; pianoforte, batt. 38 – 39) o agglomerati accordali (es. vibrafono, marimba e pianoforte, batt. 97 – 100) relativamente isolati (es. vibrafono, batt. 55; pianoforte, batt. 59 – 60), orchestrati da più strumenti (es. pianoforte, marimba e vibrafono, batt. 23 – 24; tastiera elettronica, marimba e pianoforte, batt. 28; arpa, vibrafono e marimba, batt. 29) anche organizzati direzionalmente (es. vibrafono, batt. 118 – 124: direzionalità obliqua discendente; pianoforte, batt. 118 – 125: direzionalità obliqua discendente. Questi due eventi complessi assolvono anche a funzione di ponte tra A e A→B).

Le batt. 41 – 43 presentano un elemento iterativo omoritmico lineare a-direzionato a densità minima (arpa) cui corrisponde una regione ad articolazione complessa (flauto, oboe, clarinetto in sib) e una zona a simmetria imperfetta (di nuovo l'immagine dello specchio con anomalia) con elemento speculare dell'elettronica (AT\_9 a batt. 41, AT\_10 a batt. 43, batt. 42 vuota) già in sé deformato a causa della parziale eterogeneità delle due tracce audio che partendo da suoni di origine comune (pletro o unghia sfregati sulla corda del pianoforte a produrre cluster di parziali) sono elaborate staticamente la prima dinamicamente la seconda, asse di simmetria de-focalizzato (gong e pianoforte sfasati di pochi istanti) e elemento speculare imperfetto del gong (fa# grave

batt. 41 e 44). In maniera simile, molte altre zone della parte A possono essere delimitate e osservate in un certo senso come oggetti musicali a se stanti (es. la prima pagina, batt. 1 – 4, che sarà oggetto di osservazione dell'analisi micro-formale).

### 3.1.2 TRANSIZIONE A→B

La prima zona di transizione conserva i tratti linguistici grammaticali (in termini figurati, di orchestrazione e scelta di materiali musicali) propri della parte A, modificandone la sintassi attraverso una graduale intensificazione della scrittura cui corrisponde un percettibile rallentamento (25% c.ca della velocità precedente) del tactus. Questo atteggiamento sembra essere finalizzato a suggerire un'impressione di anamorfosi e rivela un atteggiamento ambivalente di sincronica affermazione e negazione del cambiamento, quindi dello scorrere del tempo musicale. Lo stesso tipo di atteggiamento duale (scisso) è individuabile nella scelta dei materiali musicali: elementi lineari di durata medio-lunga o lunga, spesso de-focalizzati nel dominio del tempo (trilli e ribattuti) o in quello della frequenza (tremoli) mediante *trilli all'unisono* (es. marimba, batt. 125 – 128; vibrafono, batt. 127 – 131), *tremoli* (interpretabili rispetto ai trilli all'unisono come un grado successivo di de-focalizzazione dell'oggetto) (es. marimba, batt. 141 – 144; pianoforte, batt. 141 – 144), *flatterzunge* (de-focalizzazione timbrica che coinvolge sia tempo-ritmo che frequenza-altezza e spettro) (es. oboe, batt. 127 – 128), *perturbazioni dello spettro armonico sul transiente d'attacco* (es. flauto, batt. 128), *parziale e irregolare discontinuità interna* (in cui i gradienti di eterogeneità ritmica e imprevedibilità sono maggiori e più randomici rispetto a quelli dei trilli) (oboe, batt. 126; flauto, batt. 131).

Rispetto alla parte A in questa zona di transizione la dimensione obliqua assume una rilevanza determinante attraverso *glissandi* (es. clarinetto in sib, batt. 132; oboe, batt. 138; flauto, batt. 140), *elementi scalari rapidissimi ascendenti* (es. oboe, batt. 138), *discendenti* (flauto, batt. 139; pianoforte, fine batt. 133) e a *direzionalità mista* (es. flauto, batt. 132 – 133), *cluster glissati* (es. pianoforte, batt. 138 – 139).

In A→B assistiamo a una progressiva e violenta *contrazione* nell'organizzazione dei materiali che si accompagna a una *deriva centrifuga* caratterizzata da sempre maggiore de-focalizzazione degli oggetti, aggregazione istantanea in composti eterogenei instabili subito distrutti (condizione simile al movimento caotico delle molecole nei gas all'aumentare della pressione o della temperatura) e sfocia in un *collasso* del tempo musicale (batt. 141 – 147) contraddistinto da tremoli tra altezze lontane prima (marimba e pianoforte, batt. 141 – 142) e lontanissime infine (pianoforte, batt. 143): questo è il punto di massima de-focalizzazione, che innesca la deflagrazione del tessuto musicale attraverso gli aggregati accordali organizzati in senso discendente di pianoforte e vibrafono (batt. 144 – 146), una sorta di discesa nelle zone infere (parti inferiori, *inconscie*) che disgrega l'articolazione dei materiali di flauto, oboe e clarinetto in suoni tenuti, indifferenziati e a-direzionali (batt. 147 – 150), e quindi in un silenzio abitato da echi inquieti (la parte B).

La AT\_17 che parte a batt. 134 e dura 41 secondi c.ca è l'unico intervento dell'elettronica in questa zona di transizione. Si tratta di materiale altamente significativo poiché in esso è racchiusa una lettura metaforica e poetica di ciò che sta avvenendo al materiale strumentale (e al tempo musicale del brano). Essa è formata da due tracce monofoniche lineari in dinamiche basse con un lungo fade-in sino alla parte centrale dell'oggetto, dove sono percorse da un evento brevissimo verticale su entrambi i canali in dinamica *fff*; quindi una sola traccia monofonica lineare in dinamica bassa in fade-out lento. Ancora una volta l'immagine è quella dello *specchio con anomalia*: in questo caso l'evento centrale è asse di simmetria e la deformazione consiste nella perdita di una delle due tracce

iniziali, evento che ci dice di una non-reversibilità già espressa dal compositore nel termine *Ineluttabile* contenuto nell'acrostico iniziale e che individua uno dei tratti distintivi di *Atomi distratti*, caratterizzando, come si vedrà in seguito, il nucleo fantasmatico alla base del brano, espresso dalla constatazione *Amoridistrutti*, falsamente speculare al titolo. I materiali presenti nella AT\_17 sono omogenei: si tratta di frammenti della durata di 50ms c.ca (short attack, 40ms sustain, 10ms decay) iterati omoritmicamente (con periodo pari a 100ms) il cui contenuto non è riconoscibile e non permette alcuna inferenza circa il materiale di partenza.

Adottando un approccio ermeneutico, possiamo dire che la transizione A→B segna la perdita e l'abbandono di un tempo musicale adulto, complesso, individuato (nel senso psicoanalitico di *non fuso* con oggetti primordiali o con parti scisse dell'Io) e coeso in favore di un tempo idealizzato e indistinto, falsamente immobile che scorre al rallentatore, come "una pallina lanciata al rallentatore nell'acqua" (cf. Cap. 7), agghiacciato nella negazione del principio di realtà: fuga nell'immobilità di una idealizzazione impossibile e *grottesca* che rivelerà il suo potenziale tanatoico nella parte C.

### 3.1.3 PARTE B

Il titolo *L'ente dolente* è indicativo dei contenuti della parte B. la struttura formale è semplicissima: una linea proto-melodica (mancano elementi propriamente melodici, non vi sono mai più di due suoni per grado congiunto; al contrario suoni singoli, bicordi e tricordi si alternano a pause in modo apparentemente casuale) al glockenspiel con dinamica *f* e suoni lunghi tenuti con dinamiche basse a flauto basso, oboe e clarinetto in sib. Alcuni interventi locali dei gongs. A livello macro-formale la regione centrale (batt. 156 – 163) configura un *centro vuoto* differentemente caratterizzato in cui il glockenspiel tace: questa regione può essere considerata un ponte tra la parte B e la transizione B→C, esattamente come le battute 118 – 125 svolgono funzione di ponte tra A e A→B. Anche in questo caso, considerando sia la parte B che la transizione successiva, si viene a creare l'immagine dello specchio con anomalia, anamorfoosi con regione ponte contraddistinta da movimento obliquo ascendente simile ai pitch bend ascendenti della tastiera elettronica, cui il compositore assegna un ruolo straniante: "...addirittura un maggior senso di rallentando verso la fine, mediante il pitch bend ascendente" (vedi Cap. 7). In tale regione possiamo osservare la ripresa dei materiali proto-melodici del glockenspiel da parte della tastiera elettronica. Nessuna traccia audio viene attivata.

La parte B è attimo di massima sospensione, istante puro d'immobilità e meraviglia in cui la regressione idealizzante funziona e cancella il mondo esterno mentre il mondo interno è interamente contenuto in una sola istanza o immagine, organizzata secondo il massimo grado di semplicità intorno al nucleo fantasmatico che informa l'idealizzazione stessa (gli *Amori* che sono andati *distrutti*).

### 3.1.4 TRANSIZIONE B→C

La transizione B→C è caratterizzata da tripartizione, con suddivisioni interne come indicato in Tabella 3.

Arcata (batt.)	Sotto-arcata (batt.)	Durata (n° batt.)	Morfologia
BC1 (161 – 163)		3	code strumentali parte precedente; silenzio; start AT_18 batt. 161; start cronometro
BC2 (164 – 233)	BC2a (164 – 205)	42	<i>glockenspiel</i> , proto-melodia; <i>fl. basso, cl. basso sib, oboe, gongs</i> , echi; <i>tast.el.</i> , permutazione 5 suoni <i>pp</i>
	BC2b (205 – 233)	28	<i>glockenspiel</i> , proto-melodia; <i>fl. basso, gongs</i> , echi
BC3 (234 – 243)	BC3a (234 – 238)	5	<i>fl.basso</i> , lento glissando ascendente 1tono <i>ppp</i> ; <i>gong</i> , tremolo <i>ppp</i>
	BC3b (239 – 243)	5	<i>gong</i> , tremolo <i>ppp</i> cresc. a <i>mf</i> batt. 243 (start AT_20)

Tabella 3. Transizione B→C: suddivisioni interne

L'arcata centrale BC2 ha durata pari a 70 battute ed è suddivisibile in due sotto-arcate di durate lievemente differenti che ricordano le suddivisioni interne della parte A. Anche qui l'organizzazione formale ruota intorno a periodi di 35 battute circa; poiché questa è l'ultima sezione di *Atomi distratti* in cui l'ensemble strumentale agisce (è anche l'ultima a essere notata in partitura; la parte C infatti, come accennato in precedenza, è costituita di sola elettronica e non è riportata in partitura) sarebbe possibile ipotizzare una *tripartizione di secondo livello* tra la parte A, la transizione B→C e le regioni intermedie organizzate differentemente.

La traccia audio AT\_18 viene attivata all'inizio della sezione BC1, cinque battute prima della traccia AT\_19. Questa ultima, della durata di 10 secondi circa, è un'elaborazione ulteriore dei materiali inclusi nella traccia AT\_17: in questo caso si tratta di frammenti della durata di 100ms c.ca (short attack, 100ms sustain, short decay) iterati omoritmicamente (con periodo pari a 300ms) e distribuiti su entrambi i canali del file stereofonico secondo zone di maggiore o minore densità, con variazioni dinamiche in crescendo e diminuendo. Anche nella AT\_19 il contenuto non è riconducibile a suoni strumentali o concreti e non permette alcuna inferenza circa il materiale di partenza. La traccia AT\_18 dura 5 minuti, permane nella parte C ed è costituita suoni di sintesi organizzati in fasce orizzontali, con un fade in di 2'30" e nessun fade out. Al suo interno sono individuabili 8 brevissime perturbazioni, di cui 7 costituite da rumore bianco (0'42"; 1'30"; 2'23"; 2'25"; 2'42"; 3'30"; 4'18") e 1 da suoni distorti (1'45"). La traccia AT\_20 viene attivata 2'41" dopo la AT\_19.wav, poco dopo la fine del fade in.

La transizione B→C può essere interpretata come fase di rallentamento e coazione in cui il meccanismo dell'idealizzazione inizia a "incepparsi", a rivelare la sua natura di falsificazione della realtà e prepara il terreno per il ritorno del principio di realtà, abitato da fantasmi di colpa e angosce di morte. Coerentemente, se il tactus, il battito, tende ingannevolmente a rallentare (il valore di riferimento passa dalla semiminima alla minima), il tempo musicale subisce una continua contrazione: si passa da semiminima = 54 (batt. 156) a minima = 62 (batt. 164), quindi a minima = 68 (batt. 192) e infine a minima = 74 (batt. 213). La linea eseguita dal glockenspiel diviene sempre più frammentata, le interruzioni sono sempre più numerose, quasi a significare una intensificazione del ritmo respiratorio e i materiali eseguiti dagli altri strumenti subiscono una intensificazione con aumento dei livelli di articolazione e complessità timbrica tra

batt. 210 e batt. 220 per poi rarefarsi nuovamente tra batt. 221 e batt. 233 e terminare con un glissando ascendente (1 tono) del flauto basso che dura dall'inizio di batt. 234 (con attacco tongue-ram in levare) sino alla fine di batt. 238 su un tremolo tenuto *pppp* del gong della durata di quindici battute, una minima e una croma.

### 3.1.5 PARTE B

L'ensemble strumentale tace e permangono le tracce audio AT\_18 (non processata in tempo reale) e AT\_20 (processata in tempo reale). Si tratta di materiali opposti che coesistono e terminano nello stesso istante. Il contrasto tra la fissità agghiacciata della AT\_18 e il raptus dionisiaco della AT\_20, insieme al silenzio dell'ensemble testimoniano dell'avvenuta scissione tra tensione idealizzante (principio di piacere) e angosce di morte (principio di realtà), esitanti in un blocco delle attività dell'Io, uno stato catatonico di non-azione in cui la coesione dell'Io e del pensiero vengono meno.

### 3.2 ANALISI DELLA MICRO-FORMA

L'analisi micro-formale si limita, in questa sede, a un'osservazione della regione iniziale al fine di individuare alcuni tratti distintivi che ricorrono nel brano e contribuiscono alla sua strutturazione morfo-sintattica di base.

La stroficità iniziale (parte A) è costituita da elementi *antecedenti* e *conseguenti*. Il pattern iniziale, di 2 battute circa, è caratterizzato da una *semifrase di riferimento* internamente articolata secondo una flessuosità costituita da tutti gli strumenti che agiscono in modo frammentato. All'interno delle prime 2 battute è presente un elemento antecedente diacronico etereo su estensione medio acuta, costituito dalla traccia audio AT\_1 e dai materiali della tastiera elettronica e del flauto in sol, una granulazione e dispersione su un campo armonico (Fig. 1) e una *chiusura con intensionalità di cadenza* affidata a tastiera elettronica sulle note SI5 (DO4 = do centrale) e DO5, glockenspiel e pianoforte su SI7 e DO7 (Appendice 1, riquadri con numero 4), caratterizzata da *sincronicità*. La marimba introduce un suono grave (RE3 – SOL4) (Appendice 1, riquadro con numero 5) interno al campo armonico dell'*elemento antecedente* ma in sincrono con le note SI5 e SI7.

L'*elemento antecedente – parte a* (Appendice 1, riquadri con numero 1) agisce su tessitura medio-acuta (tastiera elettronica sulle note RE6, FA6 e SOL6 sincronicamente alla traccia AT\_1 e al RE5 del flauto in sol, anch'esso in battere ma con transiente d'attacco lungo) e possiede tre differenti piani di *sfasatura temporale*: la traccia audio AT\_1 è istantanea con decay rapidissimo, il flauto in sol esegue un soffiato che parte dal niente quindi viene percepito come successivo alla traccia audio, mentre la tastiera esegue un tricordo tenuto con attacco rapido e tempo di decadimento lungo. A seconda dell'interpretazione il pitch bend sulla porzione caudale del tricordo può o meno cadere sulle note dell'*elemento conseguente*.

L'*elemento antecedente – parte b* (Appendice 1, riquadri con numero 2) è costituito dal vibrafono che articola materiale (MI6, SIb4, FA6) sul quale l'arpa "risponde" (FA6, MIb6), chiudendo quindi insieme all'*elemento conseguente*, e dai materiali del flauto in sol (LA3) e del clarinetto in sib (RE6) alla fine di battuta 1.

L'*elemento conseguente* (Appendice 1, riquadri con numero 3) è costituito dai materiali di flauto in sol (DO6, DO4), arpa (FA5, LA4, SIb4), oboe (D5) e vibrafono (DO5 – RE6).

In Figura 1 viene mostrata una tavola riassuntiva dei campi armonici per la prima semifrase di riferimento.



Figura 1. Semifrase di riferimento (batt. 1-2): campi armonici

Il principio di variazione applicato alla parte A agisce sulla *posizione degli oggetti* mutandone la funzione all'interno della flessuosità fraseologica. Ad esempio: il secondo suono del flauto è un LA (batt. 1) che ha funzione di fondo a ridosso dell'aspetto cadenzale SI – DO; lo stesso LA alla batt. 3 sembra invece avere una funzione conclusiva di eco dell'evento. Il DO del flauto in sol a batt. 2 innesca la reale conclusione della semifrase di riferimento: le note RE5, LA3, DO6 del flauto (batt. 1, come elemento composto a[a-b-c]) vengono riprese poi a batt. 3 – 4 nella sequenza b-a-c (elemento composto a'[b-a-c]). Si tratta quindi di una speculazione sul concetto di *binarietà* (la binarietà è l'aspetto numerico più legato alla nostra fisiologia: l'essere umano esteriormente è binario e simmetrico sebbene le due parti in realtà differiscano. Il nostro modo di camminare è binario. Al contrario, internamente non siamo simmetrici: tendenzialmente gli organi interni, per come sono posti, rispondono a una maggiore complessità) e di simmetria anomala (specchio con anomalia).

La differente posizione dei tre suoni del flauto in sol rimanda a una duplicazione della prima semifrase, ottenendo le 4 battute iniziali (*prima frase*), che può essere compresa applicando alla seconda semi-frase una metodologia di osservazione congruente con quella utilizzata per la prima.

Procedendo lungo tutta la parte A, *Atomi distratti* deteriora in modo chirurgico questo *decorso strutturalmente binario* attraverso meccanismi di micro-variazione (anticipazione, scissione, frantumazione, ecc.). Il diverso posizionamento dei suoni comporta infinite possibilità di relazioni con gli altri suoni. Ad esempio, il rapporto che lega il LA3 del flauto in sol con il MIb6 dell'arpa quasi sincrono: a batt. 1 il LA3 anticipa il MIb6 di una semicroma mentre a batt. 3 lo anticipa di una croma. Ogni volta che torna la stessa nota non è mai identica poiché la sua funzione varia al variare del suo posizionamento così come, al contrario, l'aspetto cadenzale cerca più di ogni altro frammento di mantenere la propria identità lungo tutta la parte A, anche se i due elementi che lo compongono sono spesso *isolati* (es. uno solo di essi è presente a batt. 9 e a batt. 12) o tra di loro *posti in modo differente* (es. pianoforte, batt. 1 – 2, batt. 5, batt. 13, batt. 19, ecc.). Le battute 29 e 30 (pianoforte) invece esprimono un'altra tecnica: *iterazione della cellula* finalizzata a una sorta di *sospensione* del decorso narrativo.

Tutti questi procedimenti sono volti a costruire l'*instabilità fondamentale* che percorre la *danza sospesa* nella parte A e che, crescendo e diffondendosi, porterà la stessa a degenerare in A → B.

### 4. LE TRACCE AUDIO

In partitura è prevista l'esecuzione in tempo reale di 20 tracce audio stereofoniche. I valori di ampiezza relativa di ciascun audio file sono stati calibrati in termini di livello di pressione sonora (vedi Tab. 4, colonna "PEAK LEVEL"). I valori della colonna "DIN." mettono a confronto i livelli SPL con i valori di dinamica strumentale per calibrare il livello di emissione, al fine di integrare la proiezione sonora delle tracce audio con il suono strumentale. Il file AT\_20 corrisponde alla dinamica

massima (*fff*), mentre il file AT\_8 (prima parte) corrisponde alla dinamica minima (*ppp*).

AUDIO TRACK	START (batt.)	DUR. (sec)	PEAK LEVEL (dB)	DIN. ( <i>ppp</i> ÷ <i>fff</i> )
AT_1	1.1	0.557	-5.1	<i>f</i>
AT_2	4.1	0.150	-10.6	<i>mf</i>
AT_3	6.4	7.941	-3.0	<i>mf</i>
AT_4	11.4	0.266	-4.0	<i>f</i>
AT_5	13.1	0.981	-5.0	<i>mf</i>
AT_6	16.3	14.070	-3.0	<i>ff, f, mf</i>
AT_7	28.1	8.244	-21.7	<i>mp</i>
AT_8	35.1	8.627	-1.5	<i>ppp, f, mp</i>
AT_9	41.1	4.440	-6.0	<i>mf</i>
AT_10	43.1	2.475	-5.0	<i>ff</i>
AT_11	46.3	3.523	-1.5	<i>ff</i>
AT_12	51.4	1.162	-3.9	<i>ff</i>
AT_13	54.2	0.133	-2.0	<i>ff</i>
AT_14	54.3	50.393	-3.0	<i>pp, mf, mp, f</i>
AT_15	95.1	33.773	-18.9	<i>mf</i>
AT_16	99.1	0.905	-2.9	<i>f</i>
AT_17	134.1	41.308	-0.9	<i>mf, ff</i>
AT_18	161.1	299.400	-3.3	<i>ff, fff</i>
AT_19	166.1	10.782	-3.0	<i>mf</i>
AT_20	AT_18 + 2'41''	138.673	-1.0	<i>fff</i>

Tabella 4. Caratteristiche delle tracce audio

Per quanto riguarda l'istante di attivazione delle tracce audio, il numero dopo il punto nella colonna "START" indica il movimento di riferimento all'interno della battuta (es. 6.4 indica l'ultima semiminima della battuta 6).

Quando viene indicata una sola dinamica, essa si riferisce al punto di massima pressione sonora della traccia audio. Quando viene indicata più di una dinamica la traccia audio è composita, cioè costituita da elementi eterogenei, morfologicamente differenti.

Per l'attivazione della AT\_20 è indispensabile l'uso di un cronometro, che deve essere attivato sincronicamente alla AT\_18.

I materiali inclusi nelle tracce audio possono essere inviati direttamente ai due diffusori oppure subire un trattamento in tempo reale. Il trattamento live è necessario per quanto riguarda la AT\_20.wav, non è auspicabile per la AT\_18, è possibile per tutte le altre tracce e in particolare, nella prima esecuzione (a cui fanno riferimento le indicazioni riportate in partitura) è stato applicato alle tracce AT\_6, AT\_13, AT\_14, AT\_15, AT\_17, AT\_19.

La pertinenza degli interventi e l'intensità della trasformazione del materiale sonoro sono lasciate all'interpretazione dell'esecutore. In partitura sono riportati gli istanti di attivazione delle tracce AT\_1 - 19. Per quanto riguarda la traccia AT\_20, nessuna indicazione è data in partitura in quanto l'istante di attivazione è calcolato mediante il cronometro.

Nel capitolo seguente vengono descritti gli algoritmi implicati nel live electronics.

## 5. IL LIVE ELECTRONICS

Il trattamento in tempo reale è applicato solamente alle tracce audio, e solamente ad alcune, come descritto. Non è richiesta alcuna modificazione per quanto riguarda i materiali

strumentali (a parte l'amplificazione del flauto basso a partire da battuta 156 e la possibilità di un'amplificazione trasparente di tutti gli strumenti acustici, laddove le peculiarità architettoniche e acustiche della sala in cui ha luogo l'esecuzione lo richieda). In Figura 2 viene riportato lo schema generale del live electronics di *Atomi distratti*.

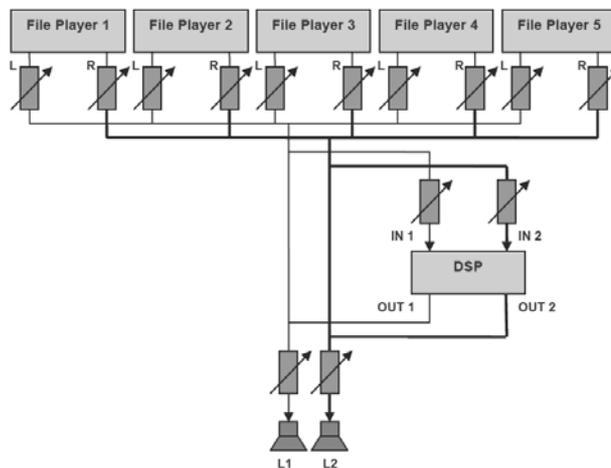


Figura 2. Live electronics, schema generale

L = Diffusore (Loudspeaker)

L'utilizzo in parallelo di 5 file player stereofonici permette la sovrapposizione delle tracce audio dove richiesto in partitura. Il modulo DSP si compone di tre unità, come riportato in Figura 3.

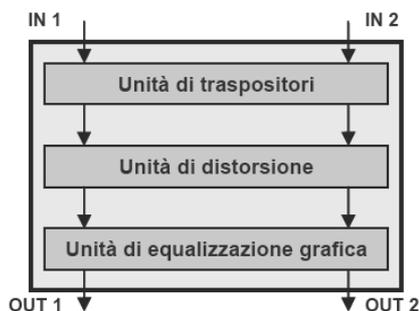


Figura 3. Modulo DSP

Di seguito vengono definite le singole unità che formano il modulo DSP.

### 5.1 UNITA' DI TRASPOSITORI

L'unità di trasposizione nel dominio della frequenza utilizza 15 traspositori organizzati come indicato in Figura 4.

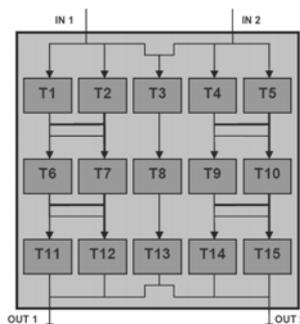


Figura 4. Unità di traspositori

T = Traspositore

Ogni traspositore è strutturato come in Figura 5.

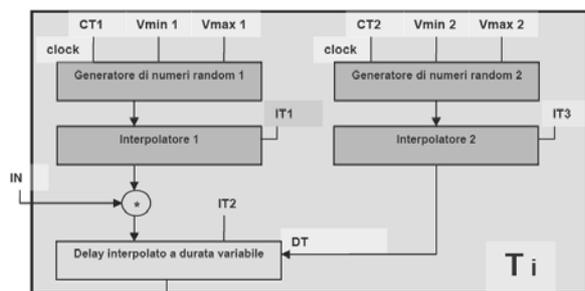


Figura 5. Traspositore

Il segnale in entrata nel traspositore viene moltiplicato per un valore di ampiezza randomico variabile interpolato, quindi viene inviato in una unità di ritardo a tempo variabile con interpolazione, in cui la variazione in tempo reale del parametro tempo, a causa dell'effetto Doppler, produce un fattore di trasposizione del parametro frequenza nel dominio del tempo. Il segnale in uscita è inviato all'unità di distorsione. In Tabella 5 sono indicati i valori dei parametri di controllo applicati ai quindici traspositori.

T	CT1	Vmin1	Vmax1	IT1	IT2	IT3	CT2	Vmin2	Vmax2
1	1875	-48	+157	1250	500	1937	625	0	58.08
2	1437	-48	+157	625	375	375	1062	0	29.04
3	4875	-48	+157	2875	625	625	3875	0	77.44
4	2875	-48	+157	1437	437	1250	2750	0	116.16
5	937	-48	+157	500	250	437	1375	0	96.80
6	2875	-48	+157	1437	437	937	1625	0	116.16
7	937	-48	+157	500	625	1625	375	0	77.44
8	1875	-48	+157	2875	250	3750	3875	0	29.04
9	1437	-48	+157	625	750	1125	2750	0	58.08
10	2125	-48	+157	1250	375	625	1375	0	84.70
11	1937	-48	+157	750	750	1437	2062	0	87.12
12	1625	-48	+157	625	250	375	2750	0	58.08
13	1875	-48	+157	1437	875	625	1875	0	116.16
14	1375	-48	+157	875	375	875	1625	0	29.04
15	1937	-48	+157	437	937	750	1250	0	77.44

Tabella 5. Valori assegnati ai parametri di ogni traspositore

- CT1 = frequenza della generazione (clock) [ms]
- Vmin1 = minimo numero random generato [int]
- Vmax1 = massimo numero random generato [int]
- IT1 = tempo di interpolazione [ms]
- IT2 = tempo di interpolazione dell'unità di ritardo [ms]
- IT3 = tempo di interpolazione [ms]
- CT2 = frequenza della generazione (clock) [ms]
- Vmin2 = minimo numero random generato [float]
- Vmax2 = massimo numero random generato [float]
- DT = dealy time [ms]

## 5.2 UNITA' DI DISTORSIONE

Unità di distorsione a 2 canali (linked) che implementa una tabella simile a quella riportata in Figura 6.

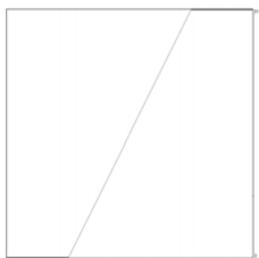


Figura 6. Tabella implementata nell'unità di distorsione

Attraverso il parametro "fattore di distorsione" è possibile variare il numero di punti utilizzati sull'asse delle ordinate al fine di aumentare o diminuire il livello di distorsione del segnale in uscita.

## 5.3 UNITA' DI EQUALIZZAZIONE GRAFICA

Equalizzatore FFT-based a 512 bande. Gain range -24 ÷ +6 dB. Frequency range 0 ÷ 22050 Hz. La tabella modificabile implementata per il controllo grafico è simile a quella riportata in Figura 7, in cui la curva raffigurata riproduce quella utilizzata per la prima esecuzione.

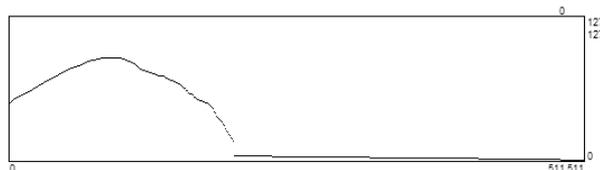


Figura 7. Tabella 512x128 implementata per il controllo grafico dell'equalizzatore

## 5.4 RUOLI E FUNZIONI DEL LIVE ELECTRONICS

Per comprendere il ruolo dell'elettronica in *Atomì distratti* è necessario indagare il rapporto del compositore con il mezzo elettronico nei lavori precedenti a questo. L'elettronica, in varie forme, è presente in sette lavori di Mario Garuti, a partire dal 1988:

- *Obliqua* per viola amplificata e traccia audio (1988)
- *Ybor* per contrabbasso amplificato e live electronics (1989)
- *For Francis Bacon* per pianoforte, orchestra e traccia audio (2001-02)
- *Crinale Amoreale* versione per flauto, percussioni, tastiera elettronica e traccia audio (2004)
- *Ermetika Emetika* versione per flauto in sol, percussioni, tastiera elettronica e traccia audio (2004)
- *Alapurplea* per ensemble con amplificazione e traccia audio (2004)
- *Miroir + Remake* installazione per video e traccia audio (2006)

Ho ritenuto utile interrogare il compositore su ognuno di questi lavori, in un gioco di libere associazioni guidate dalla memoria. Le considerazioni riportate di seguito sono state raccolte durante una conversazione con Mario Garuti avvenuta in data 16 maggio 2008 a Sesto S. Giovanni.

*Obliqua*. "Traccia audio costituita da 3 sovrapposizioni di frammenti differenti suonati dalla viola. E' un pezzo che potrebbe anche essere eseguito dal vivo con 4 viole".

*Ybor*. "Il live electronics è basato sull'idea dell'unità di ritardo. Ci sono processi legati a una de-focalizzazione del suono mediante delay molto ravvicinati; questo è il primo tipo di processo. L'altro processo è implementato attraverso un'unità di ritardo con tempi di delay molto più lunghi (c.ca 1 minuto) che permette di mandare in loop il materiale strumentale localizzato principalmente nella parte iniziale del pezzo (i primi 30 secondi)".

*For Francis Bacon*. "La traccia audio è presente nella terza parte, con un suo nome particolare: *A Desire for Flesh*. Il terzo movimento si chiama infatti *Figure Turning / A Desire for Flesh*. Il materiale è di natura eterogenea rispetto a quello dell'orchestra, centinaia e centinaia di frammenti a carattere

ritmico-impulsivo (esplosivo) temporalmente compressi con la funzione di enfatizzare la parte culminante dell'orchestra. Materiale: frammenti minimi (durata massima 1 sec) tratti da dischi rock”.

*Ermetika Emetika e Crinale Amoreale*. “La traccia audio è un mondo parallelo, materiale che non ha niente a che vedere con il materiale strumentale. Viaggia a velocità propria. È un puro accostamento di materiali eterogenei. Idea musicale: il materiale acustico è connotato da massima staticità quindi la non direzionalità del percorso musicale permette l'accostamento di oggetti sonori differenti come se l'idea fosse quella di arredare una stanza”.

*Alapurpurea*. “Il concetto è quello di accostare al massimo gli aspetti acustici a quelli elettroacustici, in una specie di morphing continuo. Gli strumenti cercano di avvicinarsi alla morfologia degli eventi musicali inclusi nella traccia audio, che in questo caso è stata composta prima della partitura strumentale. Esiste infatti anche una versione di *Alapurpurea* per nastro solo. I materiali presenti nella traccia audio sono frammenti di pianoforte e violino, entrambi suonati da me, e frammenti estrapolati dischi rock (il procedimento è simile a quello messo in atto nel terzo movimento di 'For Francis Bacon'”).

*Miroir + Remake*. “Si tratta di un'installazione in cui del materiale pre-registrato (12' c.ca) va in loop. Il materiale comprende suoni di tastiera elettronica eseguiti da me, suoni lunghi tenuti di chitarra elettrica che definiscono il piano armonico e i suoi cambiamenti, oltre a campioni di voce umana che si integrano nel tessuto ma al contempo sono come delle anomalie al suo interno: quando appaiono sembrano portare il brano verso qualcosa d'altro ma subito spariscono”.

Per quanto riguarda il ruolo dell'elettronica in *Atomi distratti*, l'idea alla base dell'accostamento tra materiali pre-registrati, elaborati in studio e trattati in tempo reale, e oggetti sonori strumentali è quella di definire un rapporto di *alterità*, secondo una logica binaria di differenziazione e contrasto che subisce durante il decorso formale del pezzo un processo di annullamento nell'identità, man mano che la traccia audio e il live electronics ad essa applicato prevalgono sul tessuto strumentale e, nell'ultima parte, lo sostituiscono. In *Atomi distratti*, diversamente da quanto accade in lavori come *Alapurpurea* o *For Francis Bacon*, in cui il nastro ha funzione *mimetica* con il tessuto strumentale (la traccia audio è *in funzione* di quest'ultimo in *For Francis Bacon* mentre in *Alapurpurea* avviene l'opposto, il tessuto strumentale viene composto *a partire da* una traccia audio e sembra quasi assolvere a una funzione di *completamento timbrico* nei confronti di questa) la traccia audio è “altro rispetto alla parte strumentale. I suoi interventi sono intesi come fratture all'interno del percorso acustico sino alla fase finale in cui la lacerazione diventa *il nuovo spazio d'azione*, univocamente elettronico. Come se si aprissero delle finestre, poi alla fine l'anomalia iniziale, il *virus* dilaga e si mangia tutto: il mondo viene risucchiato all'interno di questo *altro*” [4].

Dall'osservazione dei brani in questione e da quanto afferma il compositore circa gli stessi, sembra emergere la coesistenza di due tipi di atteggiamenti compositivi differenti:

1. *contrasto-rottura*. In brani come *Ermetika Emetika*, *Crinale Amoreale* e *Atomi distratti* il ruolo dell'elettronica è di tipo contrastivo, i materiali pre-registrati e le loro eventuali modificazioni in tempo reale sono finalizzati a creare fratture nel materiale strumentale o a contrastare con esso.
2. *mimesi-integrazione*: In brani come *Obliqua*, *Ybor* e *Alapurpurea* l'elettronica ha funzione mimetica con il tessuto strumentale; in *For Francis Bacon* la

traccia audio pre-registrata ha funzione di integrazione timbrica e supporto dinamico; in *Miroir + Remake* non è presente materiale strumentale prodotto durante l'esecuzione, tuttavia viene operata una distinzione tra i materiali strumentali e i materiali vocali pre-registrati: questi ultimi, pur integrandosi nel tessuto timbrico sottostante, hanno un rapporto ambivalente con esso, in quanto la loro funzione potrebbe essere interpretata come di temporanea e sottile rottura, precedente a un ripristino dell'equilibrio. Anche in questo caso emerge la tematica di uno stato idilliaco, apparentemente inerte che viene turbato e sempre (compulsivamente potremmo dire, dato che il brano prevede esecuzione in loop) ripristinato: la reiterazione compulsiva del ripristino della condizione di ordine può essere accostato a quella più sottile *nostalgia verso qualcosa di sconosciuto* (vedi Cap. 3), di cui in *Atomi distratti* si mette in scena una possibile drammatica resa.

## 6. CONCLUSIONI

Come è emerso dalla analisi, né le tre parti A, B e C in cui si suddivide a livello macro-formale *Atomi distratti*, né le sezioni e sotto-sezioni in cui ogni parte è articolata, né tanto meno le singole frasi e semifrasi che costituiscono, a livello micro-formale, i *mattoni* su cui è costruita la *texture* del brano (parte A e transizione A→B), sono organizzate tra loro in senso gerarchico. Riprendendo la definizione di *gerarchia debole* operata da Imberty, notiamo come sia possibile descrivere l'organizzazione di *Atomi distratti* attraverso di essa, in quanto l'estrema differenziazione tra i materiali presenti nelle tre parti principali A, B e C è causa di tre grandi “cambiamenti ugualmente pregnanti e immediatamente percettibili”, mentre i meccanismi compositivi emersi dall'analisi della micro-forma danno vita a “mutamenti più debolmente percepibili e in numero relativamente più grande”; infine le regioni di transizione A→B e B→C sono internamente costituite da “numerosi cambiamenti la cui organizzazione è incerta nel senso che non vi appare nessun contrasto chiaro o nessuna rottura netta” [2].

Le strategie compositive messe in atto per gestire le variabili prevedibilità-imprevedibilità, tensione-distensione, iterazione-mutamento, addensamento-rarefazione integrano una progettazione formale a priori che è stata evidenziata dall'analisi con tecniche di action-painting direzionato, istintuale ma consapevole che muovono da un'interrogazione sulle qualità connotative (intensionali) degli oggetti musicali e assumono come terreno d'azione l'intero *range* delle possibilità morfologiche e meccaniche delle risorse messe in atto.

Una lettura interpretativa del brano ha confermato quanto emerso durante le interviste al compositore, il quale descrive l'universo poetico di *Atomi distratti* come pervaso da irreversibilità, inquietudine pervasiva, perdita di una precedente condizione idilliaca, instabilità: tratti che ruotano intorno al meccanismo dell'idealizzazione. *Nostalgia culturale di un mondo originario sconosciuto, relazione fortemente fusionale e narcisistica ed esito drammatico nella scissione* sono i tre termini che meglio descrivono rispettivamente le parti A, B e C di *Atomi distratti*.

La *distrattione* è vista come causa del fallimento dell'idealizzazione, come risulta anche dalla frase “la *distrattione* implica una deviazione del proprio percorso con conseguente perdita di identità” [1] emersa durante l'intervista. Si tratta di un termine investito pulsionalmente al punto da essere incluso nel titolo del brano, che reca in sé un paradosso: “atomo” è per sua natura e definizione “indivisibile”, “distratto” significa “disgiunto”. Infatti il verbo latino

“distrahere”, che è all’origine della parola “distrazione” come del resto del verbo italiano “distrarre”, è traducibile con “separare, tirare qua e là, disgiungere” [5]. Si tratta di termini che descrivono perfettamente ciò che accade nei casi di scissione, e anche ciò che accade, in termini musicali, in *Atomi distratti*.

La presenza di un decorso narrativo rimanda al concetto di tempo musicale. Per quanto il brano sia significazione simbolica di frammentazione, distruzione e scissione, tecnicamente parlando il suo decorso narrativo è unitario, coeso e integrato, positivamente finalizzato alla *mise en abyme* del processo di distruzione dell’Io ad opera delle angosce di morte, conseguentemente al fallimento dell’idealizzazione di un tempo mitico.

Dall’osservazione dei precedenti lavori di Mario Garuti in cui viene utilizzato il mezzo elettronico sono emerse due linee di comportamento opposte che tuttavia condividono la attribuzione all’elettronica di ruolo costruttivo in un divenire diacronico in cui il tempo musicale non viene mai negato, ma è sempre caricato di una funzione espressiva ed esperito nella sua unità e non-reversibilità, anche quando viene chiamato a essere simbolo della propria distruzione. Se l’utilizzo di tracce audio pre-registrate o trattamenti in tempo reale sottende una scelta estetica, al contrario l’amplificazione degli strumenti, nei lavori di Mario Garuti, è sempre una scelta pragmatica, operata quando gli strumenti non raggiungono le dinamiche richieste. In questo senso tuttavia, persino la decisione di aggiungere una traccia audio nel terzo movimento di *For Francis Bacon* potrebbe in parte essere ricondotta alla necessità di incrementare l’impatto dinamico del pieno orchestrale, aggiungendo colori timbrici che non sarebbe stato possibile produrre con gli strumenti dell’orchestra: di qui la scelta del materiale musicale, ovvero esclusivamente suoni *elettrici*, estrapolati da dischi rock o techno.

L’unico lavoro composto appositamente per nastro solo, *Alapurplea* nella versione acustica, è stato in seguito utilizzato come materiale di partenza per la composizione di un brano per strumenti ed elettronica: un atteggiamento simile a quello alla base di alcuni pezzi di G. Verrando, che ci dice dell’estrema flessibilità e pragmaticità attraverso cui Mario Garuti (e molti compositori appartenenti alla sua generazione, da Fausto Romitelli a Gabriele Manca, da Carlo Ambrosini a Michele Tadini, ecc.) si pone in relazione con il mezzo elettronico, il quale viene preferibilmente utilizzato *in funzione di* (contrasto, mimesi, rinforzo dinamico, materiale di partenza per trascrizioni strumentali, ecc.), raramente *in sé*, come invece accadeva più spesso nelle avanguardie storiche. *Multimedialità e contaminazione*, del resto, rappresentano la *facies* e il principale tratto distintivo del nostro tempo.

## 7. LA TASTIERA ELETTRONICA

E’ previsto che un esecutore esegua dei suoni di sintesi in tempo reale mediante una tastiera elettronica posizionata sul palcoscenico insieme agli strumenti acustici. I suoni di sintesi includono due differenti tipologie timbriche (preset 1, 2) organizzate secondo quattro tipologie morfologiche:

- **preset 1 – tipologia morfologica a.** Suoni in notazione non misurata (non sono indicate né la durata né la precisa collocazione all’interno della battuta) con tempo di attacco lungo (6 sec c.ca). La rampa di attacco ha durata predefinita ma la pendenza e il tipo di curvatura sono controllate in tempo reale dall’esecutore. Durata complessiva dell’evento: 8 sec c.ca. Dinamica a discrezione dell’esecutore;
- **preset 1 – tipologia morfologica b.** Suoni in notazione non misurata (non sono indicate né la durata né la precisa collocazione all’interno della

battuta) con tempo di attacco lungo indicato graficamente da una forcilla tratteggiata posta sotto il suono. La durata della rampa di attacco, la pendenza e il tipo di curvatura sono performate in tempo reale dall’esecutore mediante *volume knob*, pedale MIDI o altra implementazione. Notazione con testa romboidale (simile a quella comunemente usata per indicare i suoni armonici). Dinamica a discrezione dell’esecutore;

- **preset 1 – tipologia morfologica c.** Suoni in notazione ordinaria. Attacco morbido ma *in tempo*, dinamica intorno a *mf*. Il valore d’ampiezza deve essere calcolato confrontandolo con le dinamiche degli altri strumenti;
- **preset 2 – tipologia morfologica a’.** Suoni in notazione non misurata (non sono indicate la durata né la precisa collocazione all’interno della battuta) con tempo di attacco medio-lungo (3 sec c.ca). La rampa di attacco ha durata predefinita ma la pendenza e il tipo di curvatura sono controllate in tempo reale dall’esecutore. Durata complessiva dell’evento: 5 sec c.ca. Dinamica a discrezione dell’esecutore.

In Figura 8 viene riportata la morfologia generale dei due preset.

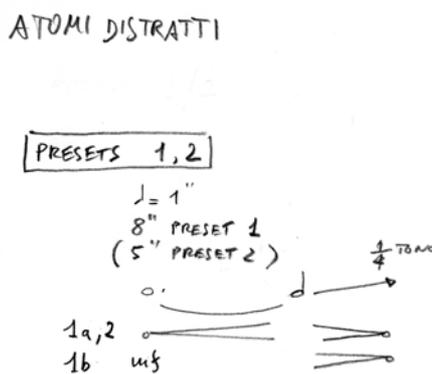


Figura 8. Morfologia generale dei suoni di sintesi in *Atomi distratti* (disegno di Mario Garuti)

Attualmente in partitura (manoscritto), per quanto riguarda la descrizione dei suoni di sintesi, sono riportate solamente le indicazioni relative alla prima esecuzione, con i suoni individuati sulla tastiera Korg 01/W fd Music Workstation. La volontà di rendere *Atomi distratti* riproducibile a lungo termine, su piattaforme e con implementazioni differenti da quelle utilizzate a Milano nel 2005, mi ha spinto a chiedere al compositore di chiarire, in termini descrittivi, i tratti distintivi principali dei due preset, al fine di realizzare una implementazione in ambiente Max/MSP di tali suoni.

Le due descrizioni che seguono sono state raccolte in data 16 Maggio 2008 a Sesto S. Giovanni.

**PRESET 1:** “Suono lungo fluttuante, di altezza lentamente instabile con un leggero glissando ascendente (approssimativamente un quarto di tono) verso la coda. Morbido ma asettico. Blu-grigio. Come una pallina lanciata al rallentatore nell’acqua, che dia addirittura un maggior senso di rallentando verso la fine, mediante il pitch bend ascendente. Un suono acquatico con un punto culminante dinamico verso la metà della sua durata, appena prima (a ridosso) del pitch bend, con il quale inizia anche il fade out”.

**PRESET 2:** "È molto simile al preset 1 ma qui è più evidente una nota metallica, quindi c'è una leggera luce che prima mancava. C'è un riflesso, un colore giallo insieme al blu".

L'assegnazione temporale dei due preset a sezioni del brano durante l'esecuzione è indicata in Tabella 6.

Preset	Assegnazione (battute)
1	1 - 155
2	156 - fine

Tabella 6. Assegnazione dei preset

### 7.1 ELABORAZIONE DI UN ALGORITMO DI SINTESI

Nelle Tabelle 7 e 8 sono state individuate per i due preset le determinanti necessarie alla costruzione di uno strumento di sintesi coerente con le caratteristiche morfologiche dei suoni richiesti dal compositore (colonna di sinistra) e la trascrizione-interpretazione tecnologica operata al fine di costruire tale strumento (colonna di destra).

PRESET 1	
Descrizione (M. Garuti)	Interpretazione tecnologica
Suono lungo.	Morfologia lineare orizzontale.
Con un punto culminante dinamico verso la metà della sua durata, appena prima (a ridosso) del pitch bend, con il quale inizia anche il fade-out.	Informazioni relative alla forma dell'involuppo.
Fluttuante. Suono acquatico.	Leggera riverberazione.
Di altezza lentamente instabile.	Vibrato con valori di frequenza dell'oscillatore modulante molto bassi (intorno a 1Hz).
Con un leggero glissando ascendente (approssimativamente un quarto di tono) verso la coda.	Graduale trasposizione nel dominio della frequenza, a partire dalla fine del punto culminante dinamico, lungo tutto il fade-out.
Morbido ma asettico.	Forma d'onda sinusoidale (asettico) tipo Hanning (morbido).

Tabella 7. Trascrizione tecnologica dei tratti distintivi del suono da generare (preset A)

PRESET 2	
Descrizione (M. Garuti)	Trascrizione tecnologica
E' più evidente una nota metallica, quindi c'è una leggera luce che prima mancava. C'è un riflesso, un colore giallo.	La forma d'onda generata è differente, più ricca di armoniche (la luce, il riflesso assente nel preset 1) ma sottile, tagliente (la nota metallica). La forma d'onda scelta è GEN9 [triplets of harmonic ratios, amplitudes and phases]: 1 0.2 0 - 8 0.5 0 - 8.01 0.5 0.2

Tabella 8. Trascrizione tecnologica dei tratti distintivi del suono da generare (preset B)

Di seguito è riportato lo schema a blocchi degli algoritmi coinvolti nell'implementazione dello strumento di sintesi.

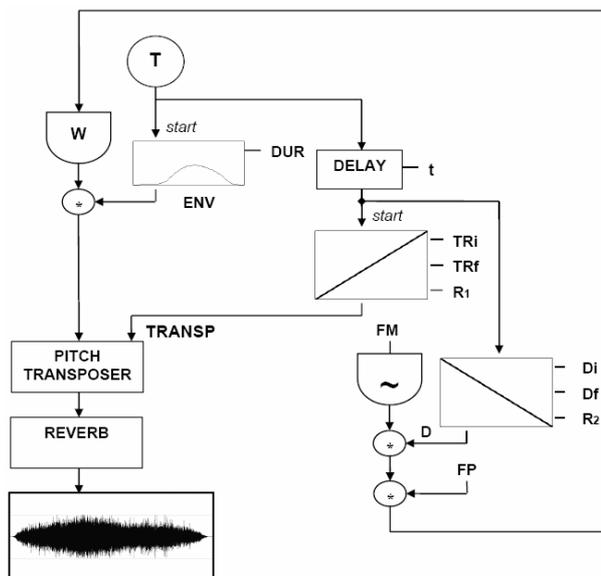


Figura 9. Schema a blocchi dello strumento di sintesi

- W** = waveform. Le due forme d'onda sono specificate in Figura 11
- T** = trigger. Inizializza la lettura dell'involuppo ENV
- ENV** = involuppo applicato al segnale in uscita dall'oscillatore. Le tipologie sono indicate in Figura 12
- DUR** = durata dell'involuppo. PRESET 1 = 8000 ms; PRESET 2 = 5000 ms
- t** = delay time. PRESET 1 = 4000 ms; PRESET 2 = 2500 ms
- TRANSP** = intervallo di trasposizione nel dominio della frequenza (fattore di moltiplicazione)
- R1** = ramp time 1. PRESET 1 = 4000 ms; PRESET 2 = 2500 ms
- FM** = frequenza modulante. PRESET 1 = PRESET 2 = 0.96 Hz
- FP** = frequenza portante. La nota indicata in partitura.
- D** = picco di deviazione
- R2** = ramp time 2. PRESET 1 = 4000 ms; PRESET 2 = 2500 ms
- Di** = picco di deviazione iniziale. PRESET 1 = PRESET 2 = 7 Hz
- Df** = picco di deviazione finale. PRESET 1 = PRESET 2 = 1 Hz
- TRi** = fattore di trasposizione iniziale (fattore di moltiplicazione). PRESET 1 = PRESET 2 = 1
- TRf** = fattore di trasposizione finale (fattore di moltiplicazione). PRESET 1 = PRESET 2 = 1.03

### 7.2 IMPLEMENTAZIONE IN AMBIENTE MAX/MSP

Ho elaborato uno strumento di sintesi in grado di generare in tempo reale una polifonia di suoni generati secondo uno dei due preset e una delle quattro tipologie morfologiche precedentemente descritti, coerentemente con la lettura interpretativa operata sulle descrizioni fornitemi dal compositore. Si tratta di un'implementazione di base, flessibile a modifiche per quanto riguarda il numero di suoni che possono essere generati insieme, i tempi di interpolazione degli involuppi ad essi applicati, i valori dei parametri numerici e la durata dei suoni stessi, nelle loro porzioni di attacco, mantenimento, decadimento e rilascio. Nelle Figure 10 - 15 sono riportate le rappresentazioni grafiche dei patch e subpatch implicati nello strumento.

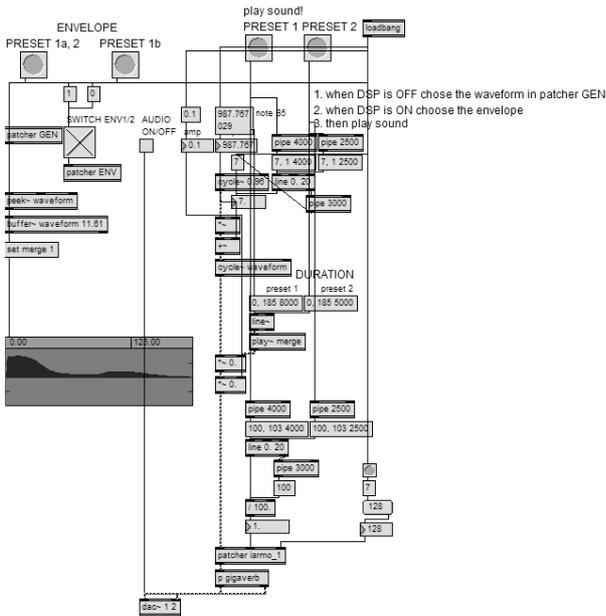


Figura 10. Strumento di sintesi: main patch



Figura 11. Strumento di sintesi: patch GEN

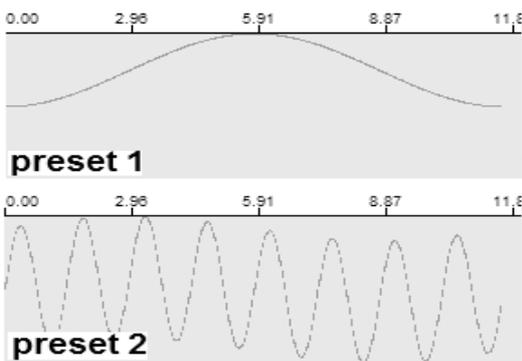


Figura 12. Strumento di sintesi: forme d'onda

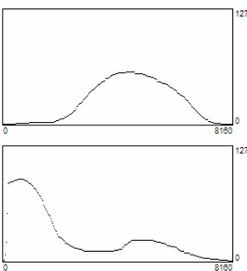


Figura 13. Strumento di sintesi: involucri (sopra: preset 1a, 2; sotto: preset 1b)

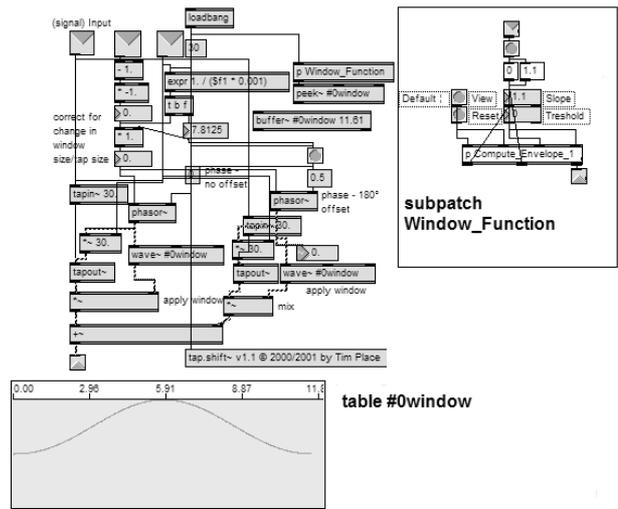


Figura 14. Strumento di sintesi: patch iarmo\_1 e subpatch Window\_Function

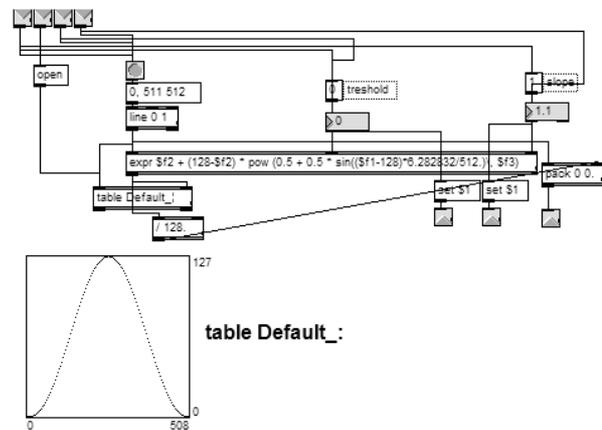


Figura 15. Strumento di sintesi: patch Compute\_Envelope\_1 e table Default\_

Come unità di riverbero è stato implementato l'oggetto `gigaverb~` [a mono in, stereo out reverb implementation by Juhana Sadeharju (kouhia@nic.funet.fi). Max/MSP-ified by Olaf Matthes. Version 1.0test3, (c) 1999-2004 Juhana Sadeharju / Olaf Matthes].

### CONSIDERAZIONI FINALI

In *Atomi distratti* la partitura strumentale è scritta in maniera deterministica e i punti d'entrata delle tracce audio sono indicati con precisione. Al contrario, le tracce audio subiscono un trattamento in tempo reale secondo processi non lineari solo parzialmente deterministici. L'intensità del processo è lasciata alla creatività dell'esecutore della parte elettronica. Si configura, in questo senso, una dissimmetria interna ai processi compositivi.

Attraverso l'introduzione nel sistema di un quantum di instabilità e di non-predicibilità abbiamo cercato di restituire alla prassi compositiva quello che Di Scipio aveva chiamato la "dose di creatività dell'aión" [6], senza per questo venire meno all'idea che tale prassi riguardi non solamente la morfologia dell'oggetto composto, ma anche e soprattutto le sue qualità intensionali e strutturali.

Se l'idea che il tempo sia come una lavagna vuota su cui scrivere simboli e relazioni, architetture e organismi astratti [7] può sembrare semplicistica, rimane che il lavoro compositivo consiste nel disporre oggetti più o meno specificamente

desiderati, pre-visti e focalizzati lungo questa linea, ovvero in qualcosa che ricorda il lavoro di un arredatore: arredare una stanza vuota, calcolando gli accostamenti, le distanze, le sovrapposizioni, ecc. La relazione tra gli eventi, intesa come oggetto-ponte tra reale e razionale, è il punto d'accesso al *tempo proprio* dell'opera, quindi al suo statuto formale; per questo è oggetto privilegiato dell'osservazione analitica, e su di essa si è focalizzata l'analisi musicologica di *Atomi distratti*. La tipologia di sguardo e la scelta dei punti di osservazione sono stati strutturati partendo da considerazioni extra-musicali attraverso le quali il compositore ha localizzato la rete di relazioni intensionali alla base della sua opera, fornendo una sorta di imprinting per l'analisi del brano; si è quindi cercato di ritrovare il senso di queste immagini all'interno dell'osservazione compiuta e infine di fornire un'interpretazione coerente e dotata di rilevanza in relazione alle intenzioni, al dato emerso dall'osservazione e agli investimenti oggettuali congetturati.

La considerazione di ordine teorico che ha motivato questo atteggiamento riguarda la *necessità di un approccio ermeneutico all'analisi musicologica*. Il senso di un'opera musicale che ne manifesta l'intenzione creatrice non può emergere semplicemente attraverso un'osservazione sperimentale, bensì richiede un *atto interpretativo*.

L'esistenza della *verità scientifica* non è messa in dubbio e tuttavia "non dobbiamo più chiedere ciò per cui essa è verità", ma limitarci ad assumere un atteggiamento "che non riposi più sul fantasma dell'oggettività come indipendenza da colui che descrive e che escluda insieme l'arbitrario, la soggettività" [8]. Prigogine estende questa necessità a domini relativamente lontani dalla fisica quali l'economia, la psicologia e le scienze sociali, senza spingersi ad estenderlo sino alla musica e all'arte in generale. Tuttavia non possiamo non riconoscere la congruenza tra ciò che egli individua come condizione di intelligibilità dei risultati empirici della meccanica quantistica prima e della fisica dei sistemi dissipativi in seguito con l'*atto ermeneutico soggettivo* da cui il senso di un'opera d'arte può sprigionare, se le fantasmatiche sottostanti sono indagate e comprese. L'approccio ermeneutico qui proposto si fonda sulla psicodinamica classica. Questo perché l'arte è una forma di comunicazione e, di conseguenza, c'è da aspettarsi che abbia origine nelle pulsioni erotiche dell'inconscio, come rileva anche Esiodo nella Teogonia, attribuendo a Eros non solo i poteri di dio dell'amore ma – cosa più importante – quelli del Demiurgo "che ha plasmato il mondo dalla fusione degli elementi in conflitto e che incarna dunque le tendenze creative e artistiche dell'umanità" [9]. Si tratta di una delle infinite letture possibili. In fatto di materia artistica, Freud e Jung si trovano in accordo quando scrivono che la psicoanalisi "non può far nulla per chiarire la natura del talento artistico, né per spiegare i mezzi con i quali l'artista opera" [10] e "l'impulso creativo, che trova nell'arte la sua espressione più chiara, è di natura irrazionale e alla fine si farà gioco di tutti i nostri sforzi razionalistici. Tutti i processi psichici consci potrebbero benissimo essere spiegati causalmente ma l'atto creativo, radicato nell'immensità dell'inconscio, eluderà per sempre i nostri tentativi di comprensione" [11]. Di qui la decisione di spingere il compositore a esprimersi sulla propria opera attraverso l'intervista: le sue affermazioni, integrate con l'osservazione analitica e musicologica longitudinale, sono divenute il punto di partenza per il lavoro ermeneutico.

I principali atteggiamenti compositivi emersi alcuni anni fa dall'osservazione di *Quantunque* (1998) erano tre: *defocalizzazione*, *significazione attraverso contrasto* e *strutturazione vincolata* [12]. I primi due atteggiamenti sono anche alla base di *Atomi distratti*: si tratta di strategie compositive finalizzate a un risultato musicale immediatamente esperibile dai sensi, mentre il terzo è una scelta di ordine meta-compositivo, in quanto informa e

direziona le strategie compositive. Quest'ultimo in *Atomi distratti* non è stato utilizzato, a favore di una organizzazione più intuitiva, sebbene controllata e attentamente calibrata in ogni sua porzione. Se in *Quantunque* a *figura* era ancora il principale organizzatore in termini micro-formali, in *Atomi distratti* essa scompare, disciolta in un unico vettore di trasformazione (processo osservabile chiaramente nella parte A), internamente omogeneo, in cui i singoli oggetti musicali, nel tempo della fruizione, perdono la loro identità individuale. L'idea di *flusso*, ibridata (secondo relazioni intensionali geometriche e sotteraneamente cicliche) con quella di *anamorfosi*, sono a mio avviso i principali *marcatori* dell'attuale evoluzione del pensiero compositivo di Garuti. Il suo sforzo di superare la visione post-weberiana di *Quantunque*, l'influenza donatoniana dei lavori del primo periodo e di parte degli anni novanta, fino a *Il demone Meridiano*, in cui l'importanza della figura come organizzatore formale era già subordinata a un vettore di trasformazione percettibile, e ferneyhoughiana (la complessità e imprevedibilità delle fantasie per strumento solo, da *Il giardino delle Esperidi* [1983-84] a *Geometriche amnesie* [1990-92]; in quest'ultimo, il tema dell'"amnesia" diviene organizzatore formale e motore di trasformazione della materia musicale, esattamente come in *Atomi distratti* lo è il tema della "distrazione") muove da una parte verso orizzonti di derivazione minimalista (*Mym* [1990], *Reich floating on the beach* [1994], *Cono in equilibrio sul suo vertice* [2000], *Flowering in her melting snow* [2006]), dall'altra verso l'idea di un "flusso cangiante" internamente e formalmente strutturato secondo un decorso trasformativo. La disposizione ricorrente in entrambe queste derive è la presenza di un *vettore di trasformazione*, che può agire in termini analogici (*Aseptic box* [2000-01], *Atomi distratti* – parti A e A→B, parte C), digitali (*Atomi distratti*, la macro-forma e parte B) o, come accade più spesso, ibridi. Negli ultimi lavori (*L'icône fanée* [2008], *Flowering in her melting snow*, *For Francis Bacon*, *L'automate purifié* [2007]), a partire da *Aseptic box*, tutti questi atteggiamenti compositivi coesistono nella chiara volontà di distillare dei processi di trasformazione in grado di vincere la forza di gravità nell'ascoltatore, creando una condizione di sospensione in cui lo scorrere del tempo non viene percepito, malgrado esso sia fortemente affermato nella sua freccia di irreversibilità attraverso tutti le strategie compositive messe in atto. Con le debite eccezioni (*Miroir/remake* e, in generale, i lavori di impronta "ambient").

## RIFERIMENTI

1. Da una conversazione con Mario Garuti trascritta a cura dell'autore in data 24 Aprile 2008. Materiale non pubblicato.
2. Michel Imberty. *Le scritture del tempo*, Ricordi, Milano, 1990.
3. Giorgio Agamben, *La comunità che viene*, Einaudi, kkk, 1998.
4. Da una conversazione con Mario Garuti trascritta a cura dell'autore in data 16 Maggio 2008. Materiale non pubblicato.
5. <http://www.etimo.it/?term=distrazione>.
6. Agostino Di Scipio, "Kairós: sulla morfologia dinamica del suono e del tempo", in *Sonus – Materiali per la musica contemporanea*, Fascicolo n. 12, Aprile 1994.
7. Iannis Xenakis, *Formalized music*, Stuyvesant, N.Y. 1992.
8. Ilya Prigogine, *La nuova alleanza*, Longanesi, Milano, 1981.
9. Arturo Schwartz, *L'immaginazione alchemica*, La Salamandra, Milano, 1979.
10. Sigmund Freud, *An autobiographical study*, 1925.
11. Carl Gustav Jung, *Psychology and literature*, 1930.
12. Marco Marinoni, "Intension und Komplexität in Mario Garuti's *Quantunque*", in *Musik & Ästhetik*, 5. Jahrgang, Heft 18, April 2001, Klett-Cotta Stuttgart.



# AudioDrome

Alessandro Petrolati

[www.alessandro-petrolati.it](http://www.alessandro-petrolati.it)

[info@alessandro-petrolati.it](mailto:info@alessandro-petrolati.it)

## SOMMARIO

*AudioDrome* è una raccolta di codice scritto in linguaggio C++, il namespace fornisce le API (*Application Programming Interface*) per lo sviluppo di DSP (*Digital Signal Processing*) *open-source*. Progettato per una facile “portabilità”, il *framework* è formato da un nucleo di classi primitive e da classi *wrapper* per le librerie di terze parti. Gli “oggetti” interagiscono attraverso un’interfaccia flessibile e semplice. Questo articolo presenta le caratteristiche principali del lavoro e descrive in generale, le problematiche legate alla programmazione di *software* per l’elaborazione del segnale digitale. Il *framework AudioDrome* si impiega come un sintetizzatore *code-based*, simile a *Csound* o *Super Collider*.

## DIGRESSIONE SU CSOUND

Un dispositivo digitale per l’elaborazione e per il trattamento del suono, deve necessariamente comprendere i seguenti punti:

### I/O

#### Process

#### Control

**I/O** è l’ingresso e l’uscita del DSP (*Process*), lo *stream* audio può essere gestito dall’OS (*Operating System*) o da specifici *drivers* software: *CoreAudio*, *Jack*, *ALSA*, *OSS*, *MME*, *DirectX*, *ASIO* etc..., l’I/O *hardware* della scheda audio (*sound board*), è chiamato ADC (*Analogic Digital Converter*) e DAC (*Digital Analogic Converter*). Anche la lettura/scrittura di uno specifico *file* audio è I/O.

**Process** è il nucleo software per l’elaborazione e la generazione. I *Csound opcodes* e le classi *AudioDrome* sono oggetti per il *Processing*.

**Control** è un I/O specializzato, lavora sui protocolli di controllo come per esempio il MIDI (*Music Instrument Digital Interface*) e OSC (*Open Sound Control*), permette il controllo dei parametri di sintesi dei DSP.

Un esempio è la seguente *patch* di *Csound*:

```
instr 1
  iAmp = ampmidi 7000
  iBaseCps = cpsmidi

  aSig oscili iAmp, iBaseCps, 1
  out aSig
endin
```

Lo *strumento* implementa un oscillatore *table-lookup* interpolato, controllato in ampiezza e in frequenza da messaggi MIDI. L’esempio indica chiaramente la semplicità di programmazione e la straordinaria potenza e flessibilità di *Csound*. I codici operativi (*opcodes*) sono “funzioni-oggetti”, *ammidi* e *cpsmidi* leggono dal *buffer* del dispositivo MIDI (in questo caso solo in *input*), *out* scrive i campioni su un *file* audio e/o in *real-time* sul DAC della scheda audio. L’*opcode oscili* (*Interpolate Oscillator*) è il modulo d’elaborazione del segnale (*Processing*).

L’alto livello d’astrazione, non ci permette di comprendere quello che avviene tra l’*opcode out* e il DAC. *Csound* e *AudioDrome* gestiscono l’I/O audio tramite la libreria *cross platform Portaudio*. Generalmente il *file orchestra* di *Csound*, contiene l’*header* per il *setup software/hardware* del sistema:

```
sr = 44100
kr = 4410
ksmps = 10
```

S’impone come *sr* (*Sampling Rate*), una velocità di campionamento e come *kr* un sotto-campionamento; *ksmps* esprime il rapporto *sr/kr* che dev’essere intero. L’*header* è costante (*const*), pertanto non può essere cambiato a tempo d’esecuzione (*runtime*), la frequenza di campionamento (SR) è il dato di riferimento assoluto. Ogni campione con velocità *kr* (o *sr*), è separato da un intervallo di tempo di  $1/kr$  secondi, in questo margine sono chiamate tutte le funzioni di calcolo degli *opcodes*. Nel caso di quelli che ritornano identificatori audio (che cominciano con la lettera ‘a’), sono chiamate *ksmps* volte le funzioni di processamento restituendo un *buffer* per ogni *kr* campione. L’esempio è preso da un sorgente di *Csound*:

```
int nn;
nn = csound->ksmps;

while (nn-->0) {
  ...
  xd = //opcode result
  *(ar++) = (MYFLT) xd;
}
```

la variabile d’appoggio locale *xd* è aggiornata col risultato di calcolo della funzione dell’*opcode*, viene incrementalmente copiata nel *buffer-array*, il modello è la base del sottocampionamento in *AudioDrome*.

## SPECIFICHE

Le classi *AudioDrome* sono ADT (*Abstract Data Type*) che modellano specifici DSP, l’istanza delle classi appare molto simile agli *opcode* di *Csound*. Le funzioni d’interfaccia di classe *setQualcheCosa()* impostano i parametri di sintesi del DSP, la frequenza di campionamento è un dato assoluto. Questa è gestita staticamente e accessibile in *read/write* solo dalle funzioni *getAbsoluteSR()* e *setAbsoluteSR()*, ma a differenza di *Csound* tutti i parametri di configurazione dei DSP, la SR e il fattore di sotto-campionamento possono variare a tempo d’esecuzione infatti, quando varia a *runtime* tutti gli oggetti vengono notificati e si ri-configurano automaticamente. La specifica più importante di *AudioDrome*, permette agli oggetti di avere un proprio **sotto-campionamento** espresso in campioni, questo determina una **SR locale per ciascun oggetto**. Un sotto-campionamento di 4 campioni, significa una risoluzione di  $SR/4 = 44.100/4 = 11025$  campioni e quindi un segnale con profilo a scalini (aggiornato ogni 4 campioni). Sotto-campionare gli oggetti di

controllo e/o secondari, agevola il lavoro della CPU e incrementa la *performance* (figura 1).

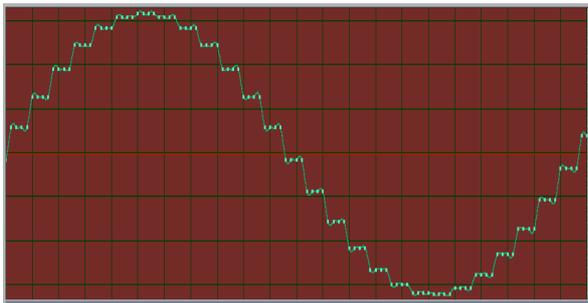


Figura 1. Sinusoide sotto-campionata di 4.

Ogni classe contiene la variabile **\_localSR**, qualificata *private/protected*, l'aggiornamento del valore *\_localSR* di classe, è gestito automaticamente dal meccanismo di notifica *Subject/Observer*, ogni volta che cambia nel sistema *\_absoluteSR* o il fattore di *downsampling* per l'oggetto.

Le specifiche principali sono qui riassunte:

- Notifica "Sampling Rate is Changed" a runtime
- Down Sampling (sotto-campionamento) degli oggetti
- Notifica "Down Sampling is Changed" a runtime
- Facilità di connessione degli oggetti.

## TASSONOMIA E INTERFACCIA

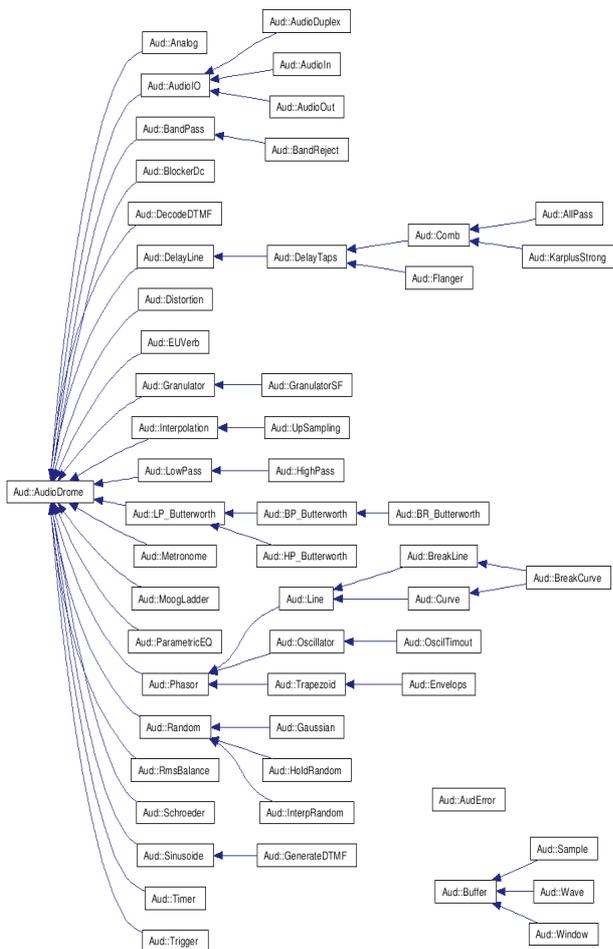


Figura 2. Le classi *AudioDrome*.

Nella gerarchia ereditaria (figura 2), ci sono modelli primitivi per la generazione e per l'elaborazione del segnale, tutte le classi DSP discendono da *AudioDrome* che contiene i metodi d'interfaccia. La "superclasse" inoltre, implementa le procedure di notifica "The Sampling Rate is Changed" e "Down Sampling". Ogni ramo corrisponde a uno specifico ambito operativo, ci sono classi per composizione e/o per aggregazione, l'argomento è approfondito di seguito.

Le classi derivate implementano le funzioni *set/get*, che forniscono l'interfaccia per l'oggetto mentre le funzioni "imperative" sono ereditate da *AudioDrome*, queste controllano il **PROCESSING** e l'I/O dei campioni dei DSP. I motori sono implementati localmente nelle sottoclassi, attraverso la procedura *void Processing()* qualificata *virtual protected* gestita da *public: PROCESSING()* di *AudioDrome*.

Un esempio delle funzioni *get/set*:

```
Oscillator obj; //creo un oggetto
obj.setFrequency(440); //imposto una frequenza
obj.getAbsoluteSR(); //imposto una Sampling rate
//oppure:
AudioDrome::setAbsoluteSR(48000);
```

l'interfaccia principale ereditata dalla "superclasse", queste funzioni **devono essere chiamate al campione**:

```
obj.PROCESSING();
obj.Input();
obj.Output();
```

necessariamente chiamate nel *loop* principale (ogni istante di campionamento) o nella *callback* (spiegata di seguito).

Un esempio del ciclo principale, nella funzione *main*:

```
long P3 = AudioDrome::getAbsoluteSR() * 5;
for (long i = 0; i < P3; i++)
{
    obj.PROCESSING();
    dac.PROCESSING();
    dac.Input( obj.Output() );
}
```

Il ciclo *for* di 5 secondi, invoca  $SR * 5 = 44.100 * 5 = 220.500$  volte la procedura *PROCESSING()*. Il risultato è accessibile "al campione" (*sample*), tramite la funzione *Output()* che ritorna il valore (*typedef double AudFloat*). La funzione *Input()* dell'oggetto *dac* prende in ingresso l'uscita dell'oggetto *obj*. La specifica di sotto-campionamento è implementata in *PROCESSING()*, la procedura controlla il numero di chiamate delle "polimorfiche" *Processing()* di classe. Il fattore di sotto-campionamento s'impone con:

```
obj.setDownSampling(10);
```

come detto, le funzioni *Input()* e *Output()* accedono alle relative variabili di classe alimentate da *PROCESSING()*. Alcuni oggetti sono multi-canale, quindi dovranno implementare le necessarie funzioni per l'I/O e conterranno le relative variabili d'appoggio:

```
AudFloat Output_3() { return _output_3; };
AudFloat Output_4() { return _output_4; };
```

etc..

```
Input_3(const AudFloat & val) { _input_3=val; };
Input_4(const AudFloat & val) { _input_4=val; };
```

etc...

Anche se la logica “orientata al campione” è più semplice e intuitiva, si preferisce gestire il segnale a *bufferFrame* di campioni, oppure con funzioni *callback* multi-thread. Si rimanda ad una trattazione approfondita dell’argomento.

### THE SAMPLING RATE IS CHANGED!

*AudioDrome::Phasor* è un modello primitivo importante, implementa un movimento di “*phase*” (rampa) normalizzato (da 0 a 1). Il numero di rampe per secondo è espresso dalla funzione *setFrequency()*, ciascuna rampa ha durata 1/freq. Un campionamento a intervalli di tempo regolari, tiene conto degli istanti di tempo che separano l’informazione, per calcolare valori d’incremento bilanciati (in relazione) sul numero di chiamate della funzione per secondo e sul valore di frequenza desiderato. *AudioDrome class* contiene la variabile *static \_absoluteSR*, qualificata *private*. Le sotto-classi accedono al dato con *setAbsoluteSR()* e *getAbsoluteSR()*, secondo i canoni della OOP (*Object Oriented Programming*). In genere il valore di SR è impiegato nel calcolo delle variabili stazionarie di classe: incrementi, coefficienti, valori temporali etc... usate da *Processing()*. Questo compito è assolto dalle funzioni *sets*, che sollevano *Processing()* da operazioni inutili, non necessarie a tempo di campionamento. Le funzioni *sets*, aggiornano le variabili di classe soltanto quando invocate.

Com’è calcolato l’incremento:

```
void Phasor::setFrequency(const AudFloat &freq)
{
    _freq = freq;
    _incr = _freq / _localSR;
}
```

*\_localSR* dipende dal fattore di sottocampionamento e da *getAbsoluteSR()*, *\_incr* è utilizzata da *Processing()*:

```
void Phasor::Processing()
{
    _output_1 = _phi += _incr;
    if (_phi >= 1.0)
    {
        _phi -= 1.0;
    }
    else if (_phi < 0.0)
    {
        _phi += 1.0;
    }
    _output_1 -= _incr;
}
```

La procedura è richiesta da *PROCESSING()* nel ciclo principale SR volte per secondo, mentre *setFrequency()* solo per modificare la frequenza d’esecuzione delle rampe.

Cosa accade se, cambia la frequenza di campionamento (SR)?

La risposta è “relativamente” semplice: *\_incr* non è aggiornato sulla base della nuova SR, la frequenza audio risultante dell’*object Phasor* sarà inversamente proporzionalmente al cambiamento della SR.

Serve pertanto, un meccanismo capace di aggiornare tutte le variabili di classe che dipendono da *getAbsoluteSR()* e/o dal numero di sotto-campionamento. L’implementazione della specifica si appoggia sul *Pattern Design* “*Subject/Observer*”, la tecnica di programmazione proposta dal **GoF** (*Gang of Four*) è semplificata grazie alla gerarchia ereditaria unitaria (tutte le sottoclassi discendono da *AudioDrome*). Il costruttore della classe *AudioDrome*, registra ogni istanza delle sottoclassi, il meccanismo di notifica chiama le procedure *srUpdate()* di tutti gli oggetti in memoria, localmente ciascuna *srUpdate()* inoltra una serie di chiamate alle funzioni membro che utilizzano il valore SR (generalmente le funzioni *sets*) e

alla *srUpdate()* della classe genitrice. Per contro, vanno considerate alcune regole quando si programmano nuove classi e quando si creano classi che contengono istanze di altri classi *AudioDrome*.

### COMPILAZIONE E DIPENDENZE

Il *package* è un **Xcode 3.0 Project** GCC/GNU sotto **Macintosh**, l’unica dipendenza facoltativa è la **Portaudio library** che serve per il trattamento dell’I/O audio. La libreria permette una “portabilità” (da cui il nome) del codice in quasi tutti i sistemi operativi, con la *pre-processor* `__WITH_PORTAUDIO__` macro, si abilita la compilazione dei *files* necessari al supporto audio, le classi *AudioIO*, *AudioIn*, *AudioOut* e *AudioDuplex* sono *wrappers* di *Portaudio*, interfacciano le API normalizzandola secondo la specifica *AudioDrome*. Omettendo la macro si rimuove la dipendenza, per esempio includendo i sorgenti in progetti VST (*Virtual Studio Technology*), in cui l’I/O audio è trattato direttamente dall’*host*. La cartella contiene le librerie statiche **STK** (*Synthesis ToolKit*) di Gary P. Scavone/Perry R. Cook e **Portaudio** di Ross Bencina, compilate. Xcode IDE (*Integrated Development Environment*) è configurato per i seguenti *Targets*: *Static-Library* (*libaudiiodrome.a*), *Shell* (compila un *exec* da terminale) e *Application* (compila e crea un *bundle* per Mac OSX). La sottocartella **Dependencies**, contiene le librerie e gli *header files* delle dipendenze. La *pre-processor* macro `__DOUBLE_PRECISION__`, definisce il tipo di dato *AudFloat*, questo è un *alias* per i tipi primitivi *double* (virgola mobile a doppia precisione 64 bit) o *float* (virgola mobile a singola precisione a 32 bit) secondo la codifica IEEE-754. La doppia precisione esprime il range  $-10^{-324} \div 10^{308}$  contro  $-10^{-45} \div 10^{38}$  dei *float*. *AudFloat* rappresenta la precisione di calcolo dei DSP (variabili di ritorno, locali, temporanee, etc...). Se *AudioDrome Library* (*libaudiiodrome.a*) è compilata in modalità `__DOUBLE_PRECISION__`, i progetti che la includono dovranno dichiarare la macro, per avere una corretta corrispondenza (*match*) tra gli *headers* e *libaudiiodrome.a*

### TUTORIAL

#### 1 - Hello grano!

```
/* Main.cpp */
#include "AudioDromeAll.h"
using namespace Aud;

int main(int argc, char ** argv)
{
    AudioDrome::setAbsoluteSR(44100);
    AudioOut dac;
    Granulator grn;

    ul P3 = AudioDrome::getAbsoluteSR() * 7.f;
    for (ul i = 0; i < P3 ; i++)
    {
        grn.PROCESSING();
        dac.PROCESSING();

        dac.Input( grn.Output() );
        dac.Input_2(grn.Output_2() );
        //or simply:
        //dac << grn;
    }
    dac.abortStream();
    return 0;
}
```

L’inclusione di *AudioDromeAll.h* comporta l’inclusione di tutti gli *headers files* delle classi, queste appartengono al *namespace Aud*. La direttiva *using* impone al compilatore di scavalcare il controllo, tuttavia in progetti di grandi dimensioni è sconsigliata a favore di una *scope resolution* esplicita.

La prima istruzione della funzione *main*, imposta il valore SR generale e poi s'istanzano gli oggetti per le classi *AudioOut* e *Granulator*. Dentro il ciclo *for*, si "attivano" i motori di *grn* e *dac*, quindi interconnessi mediante le relative funzioni I/O.

```
dac.Input( gr.Output() );
```

l'*overloading* degli operatori del C++, permette connessioni facili e intuitive:

```
dac << gr;
```

## 2 - Modulo riverbero in tempo reale:

```
/* Main.cpp */
#include "AudioDromeAll.h"
using namespace Aud;

int main(int argc, char ** argv)
{
    AudioDrome::setAbsoluteSR(44100);
    AudioIn adc;
    AudioOut dac;
    EUVerb euv;
    euv.setT60 (1.8);
    euv.setInperpolation(LINEAR);
    euv.setReverberatedLevel(0.3);
    euv.setDirectLevel(0.7);
    euv.setDirectDelay(0.9);

    while (true)
    {
        adc.PROCESSING();
        dac.PROCESSING();
        euv.PROCESSING();

        /* RealTime IO */
        euv.Input( adc.Output() );
        dac.Input( euv.Output() );
        dac.Input_2( euv.Output_2() );
        //or simply:
        //dac << euv << adc;
    }
    adc.abortStream();
    dac.abortStream();
    return true;
}
```

L'esempio presenta un riverbero basato sul modello proposto da Eugenio Giordani (LEMS), la classe è realizzata per composizione di oggetti *AudioDrome*, in questo caso il DSP prevede un solo *input* e restituisce un segnale pseudo-stereofonico. Ecco le connessioni alternative::

```
dac << euv << adc;
```

l'*output* di *adc*, ossia l'ingresso microfonico della scheda audio, è indirizzato sull'*input* del riverbero *euv*. L'uscita del riverbero sull'*input* del *dac*. L'operatore di direzione del flusso audio (*left shift*), collega gli oggetti in modalità STEREO per *default*, in questo caso l'*output* del riverbero è stereofonico.

Gli operazionali sovraccaricati semplificano ulteriormente la gestione del flusso del segnale:

```
dac << euv/2 << adc*0.5;
```

L'esempio di sopra divide l'uscita del riverbero e somma un *off-set* all'ingresso microfonico. Sono possibili le operazioni direttamente sugli oggetti:

```
dac << distort << (sine*0.5 + noise/10) - 1.0;
```

dove *sine*, *noise* e *distort* sono *AudioDrome* DSP.

## 3 - Full duplex audio stream mediante funzione call-back:

```
/* Main.cpp */
#include "AudioDromeAll.h"
using namespace Aud;

int MyCallback( const void *input, void *output,
               unsigned long frameCount,
               const PaStreamCallbackTimeInfo* timeInfo,
               PaStreamCallbackFlags statusFlags,
               void *userData )
{
    (void)(userData);
    // Prevent unused variable warning.

    PaFloat *out = (PaFloat*) output;
    PaFloat *in = (PaFloat*) input;

    for( unsigned int i = 0; i < frameCount; i++ )
    {
        *out++ = *in++;
        *out++ = *in++;
    }
    return 0;
}

int main(int argc, char ** argv)
{
    AudioDrome::setAbsoluteSR(44100);
    AudioDuplex io;
    io.setStreamCallBack(&MyCallback, NULL);

    io.sleep(60000);
    io.abortStream();
    return 0;
}
```

I parametri formali della funzione *MyCallback*, sono specificati da *Portaudio*, *AudioDuplex* è una classe *wrapper*, si noti l'assenza del ciclo *for* nel *main*, sostituito da *io.sleep(60000)*; *sleep* "congela" il *thread* corrente per un minuto (60000 Ms) mentre un'altro *thread* gestisce *MyCallback*. Nell'esempio la funzione collega semplicemente l'*input* con l'*output*.

## 4 - Riverbero in tempo reale in full duplex audio stream con funzione call-back:

```
/* Main.cpp */
#include "AudioDromeAll.h"
using namespace Aud;

int MyCallback( const void *input, void *output,
               unsigned long frameCount,
               const PaStreamCallbackTimeInfo* timeInfo,
               PaStreamCallbackFlags statusFlags,
               void *userData )
{
    EUVerb* euv = static_cast< EUVerb*>(userData);
    PaFloat *out = (PaFloat*) output;
    PaFloat *in = (PaFloat*) input;

    for( unsigned int i = 0; i < frameCount; i++ )
    {
        euv->PROCESSING();
        euv->Input(*in++);

        *in++;
        //not righth channel input for the EUVerb

        *out++ = euv->Output();
        *out++ = euv->Output_2();
    }
    return 0;
}
```

```

int main(int argc, char ** argv)
{
    AudioDrome::setAbsoluteSR(44100);
    EUVerb euv;
    AudioDuplex io;
    io.setStreamCallBack(&MyCallback, &euv);

    io.sleep(60000);
    io.abortStream();
    return 0;
}

```

Simile all'esempio precedente ma in questo caso viene passato a *MyCallback* l'oggetto *AudioDrome::EUVerb*, trasferito come puntatore a *void*, occorre un un *cast* esplicito per ottenere l'oggetto *EUVerb*:

```
EUVerb* euv = static_cast< EUVerb*>(userData);
```

Il *loop for* incrementa i puntatori del *buffer* sul dispositivo *Portaudio* I/O, il segnale d'ingresso del riverbero è il canale sinistro dello *stream audio In*:

```
euv->Input(*in++);
```

l'uscita del riverbero è assegnata allo *stream audio Out*:

```
*out++ = euv->Output();
*out++ = euv->Output_2();
```

## 5 - Variazione a tempo d'esecuzione della SR:

```

/* Main.cpp */
#include "AudioDromeAll.h"
using namespace Aud;

int main(int argc, char ** argv)
{
    AudioDrome::setAbsoluteSR(44100);
    AudioOut dac;
    Wave* saw = new Wave(4096);
    saw->genSaw();
    Oscillator osc(saw);
    osc.setFrequency(440);

    Metronome mt(false);
    mt.setFrequency(1);

    ul P3 = AudioDrome::getAbsoluteSR() * 5;

    for (ul i = 0; i < P3; i++)
    {
        dac.PROCESSING();
        mt.PROCESSING();
        osc.PROCESSING();

        if(mt.Output())
            AudioDrome::setAbsoluteSR (11025);

        dac << osc;
    }
    delete saw;
    dac.abortStream ();
    return true;
}

```

*Absolute Sampling Rate* è inizializzata a 48 Khz, nel *loop for* l'istruzione *if* verifica se *mt* (*Metronome*) ritorna un *trigger true*, quindi dopo il primo secondo (poiché *mt* ha una frequenza di 1 Hz) il tasso di campionamento si aggiorna a 11 Khz. Quando il *trigger* invoca *setAbsoluteSR()*, tutti gli oggetti in memoria vengono notificati, ri-configurandosi a *runtime*. L'istanza *dac* di *AudioOut* deve necessariamente chiudere lo *stream audio* per poi riaprirlo sulla nuova SR.

Si noti il costruttore dell'oscillatore che accetta il puntatore *Wave*, ossia ad una tabella *look-up* di 4096, contenente il profilo a "dente di sega" usato come forma d'onda:

```

Wave* saw = new Wave(4096);
saw->genSaw();
Oscillator osc(saw);

```

## 6 - DownSampling (sotto-campionamento):

```

/* Main.cpp */
#include "AudioDromeAll.h"
using namespace Aud;

int main(int argc, char ** argv)
{
    AudioDrome::setAbsoluteSR(44100);
    AudioOut dac;
    Oscillator lfo;
    lfo.setAmplitude(0.5);
    lfo.setFrequency(3);
    lfo.setDownSampling(50);

    Oscillator osc;
    osc.setFrequency(440);

    ul P3 = AudioDrome::getAbsoluteSR() * 5;
    for (ul i = 0; i < P3; i++)
    {
        dac.PROCESSING();
        lfo.PROCESSING();
        osc.PROCESSING();
        osc.setAmplitude( lfo[LEFT] + 0.5 );
        dac << osc;
    }

    dac.abortStream();
    return true;
}

```

L'esempio sopra tratta una semplice AM (*Amplitude Modulation*). L'*lfo* (*Low Frequency Oscillator*) è sottocampionato 50 campioni, i vantaggi in termini di velocità di calcolo sono evidenti. Un procedimento più sofisticato interpola il segnale sotto-campionato, attraverso l'oggetto *UpSampling* che viene inserito nella catena. Questo riceve in ingresso l'uscita dell'*lfo* e restituisce il segnale interpolato progressivamente sul numero di sotto-campionamento:

```
UpSampling up(lfo);
```

Il costruttore di *UpSampling*, configura *up* sulle basi di *lfo* (oggetto da interpolare), il costruttore accetta "per riferimento" un *AudioDrome object*, o direttamente il valore di sotto-campionamento.

```

for (ul i = 0; i < P3; i++)
{
    dac.PROCESSING();
    lfo.PROCESSING();
    osc.PROCESSING();
    up.PROCESSING();

    up << lfo; //lfo è l'input di up
    osc.setAmplitude( up.Output() + 0.5 );
    dac << osc;
}

```

Il modulo d'interpolazione *UpSampling*, permette di spingere il sotto-campionamento a 500 senza degradazione del segnale, il sottocampionamento non si sottrae alla legge di Nyquist, se 44.1 Khz permettono fino 22050 Hz di banda, allora  $44100/500 = 88.2/2 =$  permettono soltanto 44,1 Hz di banda.

## PROGRAMMARE UNA NUOVA CLASSE

Il codice di *myClassTemplate* è il modello per programmare nuove classi nella gerarchia, può ereditare da *AudioDrome* o da una sottoclasse dipendentemente dal livello d'astrazione del nuovo DSP. Il modello include la funzione *myMemberFunction()* che sarà rinominata come funzione membro specifica per l'astrazione, a titolo d'esempio è stata contraddistinta come */\*! Sampling Rate Change SLOT \*/* per indicare che utilizza, relativamente al DSP, la variabile *\_localSR* e/o *getAbsoluteSR()*. La funzione è quindi chiamata da *srUpdate()* notificata dal processo "Sampling Rate is Changed". La classe appartiene al namespace *Aud*.

```
/*! \class myClassTemplate */
#include "AudioDrome.h"

//begin namespace
namespace Aud
{
    //inherits from AudioDrome or a sub-class
    class myClassTemplate : public AudioDrome
    {
    public:
        /*! Public constructor
        myClassTemplate();

        /*! Public destructor
        virtual ~myClassTemplate();

        /*! Sets or gets ...
        /*!
        * if you need to upgrade
        * the class variables,
        * dependent on _localSR
        */
        /*! Sampling Rate Change SLOT */
        void myMemberFunction(const AudFloat &val){}

        * You can read or write the sample
        * in the DSP with the virtual Output()
        * or Input() functions.
        * Also you can redefined.
        */
        virtual AudFloat Output();
        virtual AudFloat Output_2();

        virtual void Input(const AudFloat& in);
        virtual void Input_2(const AudFloat& in);

protected:
        /*! Processing the audio sample
        /*!
        * ...
        */
        virtual void Processing();

        virtual void srUpdate()
        {
            /*! Sampling Rate Changing SLOT */
            this->myMemberFunction();
        }
        //my class variables
    };
} //end of namespace Aud
```

Il modello di classe, deve considerare tre punti principali

- 1) La classe implementa la procedura *Processing()*
- 2) La classe implementa la procedura *srUpdate()*
- 3) Se contiene oggetti *AudioDrome*, ridefinisce la procedura *PROCESSING()*

*Processing()* è l'ingranaggio di calcolo locale del DSP, in realtà potrebbe non essere ri-definita, ereditando la *Processing()* della sotto-classe. La procedura *srUpdate()* dev'essere ridefinita solo se la nuova classe utilizza la variabile *\_localSR* o *getAbsoluteSR()*, la *\_localSR* è la frequenza di campionamento dell'oggetto, calcolata come *\_absoluteSR / \_downSampling*. In particolare è necessario ridefinire *PROCESSING()* nelle classi contenitrici, infatti la gestione del *downSampling* è responsabilità della classe contenitrice, per esempio la classe *Granulator* contiene istanze di *OscilTimeout* e *Envelops*, il sotto-campionamento della sola istanza di *Granulator* porta a conseguenze drammatiche. Nelle classi contenitrici di oggetti, il sotto-campionamento dev'essere unitario, *Granulator::Processing()* chiama le *PROCESSING()* di calcolo degli oggetti contenuti, questi non sono configurati correttamente è necessario quindi inoltrargli la richiesta di sottocampionamento. In questo caso però si determina un **sotto-campionamento doppio**, quello attuato dalla classe contenitrice e quello dagli oggetti contenuti. La soluzione è semplice: **si inoltra il sotto-campionamento a tutti gli oggetti contenuti, ma non si sotto-campiona la classe contenitrice**. Un caso più complesso è quando la classe contenitrice chiama le *PROCESSING()* delle istanze degli oggetti contenuti e la *Processing()* della classe da cui eredita. In questo caso, il sottocampionamento dovrà essere gestito inoltrando la richiesta anche alla sola classe genitrice.

```
/*! Sampling Rate Change SIGNAL */
virtual void setDownSampling(const ul& samples)
{
    AudioDrome::setDownSampling(samples);
    _trig->setDownSampling(samples);

    for (unsigned int i = 0; i < _overlap; i++)
    {
        _osc[i]->setDownSampling(samples);
        _env[i]->setDownSampling(samples);
    }
}
```

L'esempio estrapolato da *Granulator*, ridefinisce la funzione *setDownSampling()*, la procedura inoltra la richiesta di sotto-campionamento a tutti gli oggetti contenuti.

```
virtual void PROCESSING()
{
    this->Processing();
}
```

La classe ridefinisce anche *PROCESSING()*, al fine di scavalcare il controllo *downSampling* della *AudioDrome::PROCESSING()*, questa chiama semplicemente *this->Processing()*. Per eliminare l'*overhead* di chiamata di funzione, *this->Processing()* potrebbe direttamente essere definita come *this->PROCESSING()*.

In fase di sviluppo, ho scelto il *pattern design*, come meccanismo di notifica della "Sampling Rate is Changed", al posto della tecnica *Signal/Slots*, implementata dalle *Boost C++ Library*. Nel caso del *pattern design*, le procedure *srUpdate()* devono richiamare se stesse lungo il ramo ereditario fino alla "superclasse" infatti, ogni classe potrebbe aver bisogno di aggiornare le proprie variabili. Nel caso dei *Signal/Slots* i costruttori di tutte le classi devono registrare la propria procedura *srUpdate()*, inserendo in un *vector* il puntatore a funzione. Entrambe le tecniche funzionano correttamente, ma non sollevano l'utente programmatore che sviluppa una nuova classe, dalla responsabilità di **ricordarsi la corretta implementazione di un particolare**. L'attuale soluzione appare concettualmente corretta e non troppo complicata:

```
virtual void srUpdate()
{
    Oscillator::srUpdate();
    this->setTime(_dur);
}
```

l'esempio è implementato nella classe *OscilTimeout*, la procedura richiama la funzione membro *setTime()* e nello stesso tempo chiamata la *Oscillator::srUpdate()* della classe genitrice, la stessa cosa fanno *Oscillator* e *Phasor* risalendo il ramo ereditario.

## UN PROGETTO VST

Per sviluppare un plugin VST è necessario procurarsi, dal sito ufficiale della **Steinberg**, il VST SDK (*Software Development Kit*). La compilazione del plugin dipende dalla piattaforma: in Windows è una DLL (*Dynamic Link Library*), in Mac OSX è una *bundle* con estensione *.vst*, ossia una cartella che contiene il file *plist.info* con la chiave *CFBundleName* per la definizione del nome. Su BeOS e SGI (MOTIF, UNIX) un plugin VST è una *shared Library*. Si rimanda alla documentazione in rete per studiare e approfondire l'argomento, in particolare segnalo il sito **www.parravicini.org** con la documentazione necessaria in italiano. Nel progetto si devono includere l'SDK VST e la *audiodrome library* (*libaudiodrome.a*) o direttamente i sorgenti, senza il supporto per l'audio in quanto gestito dall'*host*. Se la libreria è compilata in `__DOUBLE_PRECISION__`, è necessario ripetere la definizione al pre-processore nel progetto VST.

Dal sito [www.alessandro-petrolati.it](http://www.alessandro-petrolati.it) si può scaricare un *Xcode project* del seguente codice, si tratta di un *delay* stereofonico a interpolazione cubica:

```
// myPlugin.h
#include "AudioEffectX.h"
#include "AudioDromeAll.h"

using namespace Aud;

enum{
    // Global
    kNumPrograms = 1,
    // Parameters Tags
    kMaster=0, kDelayL, kDelayR, kNumParams
};

//! myPlugin estende la classe AudioEffectX
class myPlugin : public AudioEffectX
{
public:
    //! Costruttore pubblico
    myPlugin (audioMasterCallback audioMaster);

    //! Distruttore pubblico
    ~myPlugin(){};

    //! Implementazione della logica di Processing
    virtual void processReplacing (float **inputs,
                                  float **outputs,
                                  VstInt32 sampleFrames);

    //! Inoltra la Sampling Rate dell'Host
    //! agli oggetti AudioDrome
    /*!
    * La funzione è chiamata quando il tasso
    * di campionamento dell' Host cambia,
    * lavora in modo analogo alla srUpdate()
    * nell'architettura AudioDrome.
    */
    virtual void setSampleRate(float sampleRate);

    //! Chiamata quando cambia un parametro
    void setParameter(VstInt32 index, float value);
```

```
    //! Ritorna il valore del parametro
    virtual float getParameter(VstInt32 index);

    //! Imposta una stringa di rappresentazione
    //! del valore ("0.5", "-3", "PLATE", etc...)
    virtual void getParameterDisplay(VstInt32 indx,
                                      char* text);

    //! Imposta una stringa di rappresentazione del
    //! valore ("Time", "Gain", "RoomType", etc...)
    virtual void getParameterName(VstInt32 indx,
                                   char* text);

    //! Etichetta per il parametro
    //! ("sec", "dB", "type", etc...)
    virtual void getParameterLabel(VstInt32 indx,
                                    char* label);

protected:
    float fMaster;
    float fDelayL;
    float fDelayR;
private:
    DelayTaps delayL;
    DelayTaps delayR;
};
```

L'enumeratore (*enum*) è importante, imposta il numero di programmi in questo caso uno. I programmi sono i *presets* disponibili per il *plugin* anche se l'esempio non implementa i *presets*, il parametro è richiesto dal costruttore della classe genitrice *AudioEffectX* che stiamo estendendo. Le funzioni *sets* (virtuali) sono ridefinite perchè chiamate dall'*host*, l'enumeratore contiene la lista di tutti i parametri contemplati dal DSP, per ognuno è creato un controllo (*widget o slider*) nella GUI (*Graphic User Interface*) del *plugin*. *kNumParams* è passato quindi al costruttore di *AudioEffectX*, la classe deve contenere le variabili associate a ciascun controllo, in questo caso sono tre (*fMaster*, *fDelayL*, *fDelayR*) e gli oggetti DSP di *AudioDrome*. Di seguito l'implementazione delle funzioni della classe nel file *.cpp*:

```
// myPlugin.cpp
#include "myPlugin.h"

myPlugin::myPlugin(audioMasterCallback audioMaster) :
AudioEffectX(audioMaster, kNumPrograms, kNumParams)
{
    //default settings
    setParameter(kMaster, 0.5);
    setParameter(kDelayL, 0.3);
    setParameter(kDelayR, 0.6);
    setNumInputs(2); // stereo input
    setNumOutputs(2); // stereo output
    delayL.setInterpolation(CUBIC);
    delayR.setInterpolation(CUBIC);
}

la lista d'inizializzazione costruisce AudioEffectX, il costruttore di myPlugin configura lo stato iniziale dei controlli del plugin tramite le funzioni setParameter(), imposta inoltre il numero di canali I/O dell'host e il tipo d'interpolazione per gli AudioDrome objects.

void myPlugin::setSampleRate(float sampleRate)
{
    AudioEffectX::setSampleRate(sampleRate);
    AudioDrome::setAbsoluteSR( (AudFloat) sampleRate
)
}
```

In particolare si osservi la funzione *setSampleRate()*, chiamata dall'*host* quando cambia la SR, è ridefinita per inoltrare la richiesta a *AudioDrome::setAbsoluteSR()*, facendo scattare la

notifica “*Sampling Rate is Changed*”.

Le due funzioni che seguono sono l’I/O tra la classe DSP e la GUI del *plugin*:

```
void myPlugin::setParameter(VstInt32 index,
                            float value)
{
    switch (index)
    {
        case kMaster:
            fMaster = value;
            break;
        case kDelayL:
            fDelayL = value;
            delayL.setDelayTime(value);
            break;
        case kDelayR:
            fDelayR = value;
            delayR.setDelayTime(value);
            break;
    }
}

float myPlugin::getParameter(VstInt32 index)
{
    switch (index)
    {
        case kMaster:
            return fMaster;
        case kDelayL:
            return fDelayL;
        case kDelayR:
            return fDelayR;
    }
}
```

Quando dalla GUI si agisce sui controlli del *plugin*, l’*host* chiama *setParameter()* sul relativo indice di parametro, il costruito *switch* canalizza il valore sulla relativa variabile di classe e/o la passa direttamente agli oggetti *AudioDrome*. In realtà il meccanismo è più complesso: le funzioni *set* e *get*, sono chiamate di conseguenza, infatti è la seconda che aggiorna la posizione dello *slider* sul nuovo valore.

Il *plugin* invia i dati al DSP attraverso *setParameter()*, che implicitamente invoca *getParameter()*, in altri termini *queste* sono funzione *callback mutex* (a mutua-esclusione) chiamate quando cambiano i valori dei parametri nella GUI.

```
void myPlugin::processReplacing (float** inputs,
                                float** outputs,
                                VstInt32 sampleFrames)
{
    float* in1 = inputs[0];
    float* in2 = inputs[1];
    float* out1 = outputs[0];
    float* out2 = outputs[1];

    while (--sampleFrames >= 0)
    {
        delayL.PROCESSING();
        delayR.PROCESSING();
        delayL.Input(*in1++);
        delayR.Input(*in2++);

        (*out1++) = (float) delayL.Output() * fMaster;
        (*out2++) = (float) delayR.Output() * fMaster;
    }
}
```

La funzione *processReplacing()* è una classica *callBack*, si rimanda al **Tutorial** di questo articolo per chiarimenti. La funzione *main*, per i *plugins* VST è sostituita dalla seguente inclusa nel file *myPlugin\_Main.cpp*:

```
// myPlugin_Main.cpp
#include "myPlugin.h"

AudioEffect* createEffectInstance(audioMasterCallback
                                  audioMaster)
{
    return new myPlugin(audioMaster);
}
```

ATTENZIONE, la documentazione VST assume che:

*"All parameters - the user parameters, acting directly or indirectly on that data, as automated by the host, are 32 bit floating-point data. They must always range from 0.0 to 1.0 inclusive [0.0, +1.0], regardless of their internal or external representation."* Per mappare i parametri su valori diversi, è necessario implementare opportune funzioni, strutture dati o classi.

## RIFERIMENTI

1. Curtis Roads. Computer Music Tutorial, The MIT Massachusetts Institute of Technology, 1996
2. B. Stroustrup. The C++ Programming Language. Addison-Wesley, Reading, Massachusetts, 2000
3. Bruce Eckel, Thinking in C++ Volume One (2nd Edition)
4. Carlo Pescio, C++, Manuale di Stile, Infomedia, 1995
5. Paolo Marotta , C++, una panoramica sul linguaggio Copyright 1996-1999, <http://www.tutorialpc.it/c++menu.asp>
6. Cipriani/Bianchini, Il suono Virtuale, ConTempo,1998
7. R. Bencina and P. Burk, PortAudio - an Open Source Cross Platform Audio API. 2007 <http://www.portaudio.com>
8. The Synthesis ToolKit in C++ (STK) Perry R. Cook & Gary P. Scavone Music Technology, Faculty of Music McGill University/Departments of Computer Science & Music <http://ccrma.stanford.edu/software/stk/>
9. The Sound Object Library (SndObj) Victor Lazzarini Music Technology Laboratory National University of Ireland, Maynooth <http://music.nuim.ie//musictec/SndObj/main.html>
10. Pattern Design, GoF [http://en.wikipedia.org/wiki/Design\\_pattern\\_\(computer\\_science\)](http://en.wikipedia.org/wiki/Design_pattern_(computer_science))
11. Steinberg 3rd party developer support [http://www.steinberg.net/en/company/3rd\\_party\\_developer.html](http://www.steinberg.net/en/company/3rd_party_developer.html)
12. Boost C++ Library [http://www.boost.org/doc/libs/1\\_36\\_0/libs/libraries.htm](http://www.boost.org/doc/libs/1_36_0/libs/libraries.htm)
13. DoxyGen Source code documentation generator tool <http://www.doxygen.org>
14. Emanuele Parravicini, VST plugin - Getting Started <http://www.parravicini.org>
15. Eugenio Giordani LEMS (Laboratorio per la Musica Sperimentale), Conservatorio G. Rossini di Pesaro. <http://www.eugenio.giordani.it>

# Score Following in Tempo Reale basato su Hidden Markov Models

Nicola Montecchio  
Dipartimento di  
Ingegneria dell'Informazione  
Università di Padova

## SOMMARIO

E' stato sviluppato un sistema per l'allineamento automatico in tempo reale di un'esecuzione musicale alla corrispondente partitura, basato su modelli Markoviani nascosti. Le novità introdotte nel modello permettono delle buone prestazioni nel caso di allineamenti di musiche caratterizzate da una complessa polifonia.

## Parole Chiave

Score Following, Hidden Markov Models, Real Time, Allineamento

## INTRODUZIONE

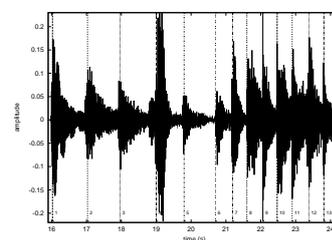
Il termine *score following* fa riferimento alla capacità di un sistema di allineare, in tempo reale, l'esecuzione di un brano musicale ad una versione digitale della corrispondente partitura (*score*); più precisamente, per esecuzione si intende un segnale audio digitale proveniente dalla registrazione di strumenti musicali acustici o elettronici, e per partitura una successione nel tempo di eventi quali note, accordi e pause. L'operazione è rappresentata in Figura 1.

La principale applicazione di tali sistemi è da sempre stata quella di generare un accompagnamento automatico in grado di "seguire" (*following*) un musicista umano, di risincronizzarsi cioè nel caso di cambiamenti di velocità e supportare eventuali errori da parte dell'esecutore. Un'altra possibilità offerta è quella di automatizzare la sincronizzazione di effetti di *live electronics* in esecuzioni di musica contemporanea, pratica sviluppata all'IRCAM a partire dal 1986.

Lo sviluppo di sistemi di score following è cominciato nel 1984, grazie a due articoli (indipendenti) di R. Dannenberg [2] e B. Vercoe [9]. I primi approcci erano generalmente basati su tecniche di programmazione dinamica, in particolare quella del *Dynamic Time Warping* (DTW) [8], inizialmente proposto nel campo del riconoscimento vocale e applicato in seguito anche all'allineamento audio. La disponibilità di potenza di calcolo e l'evoluzione delle tecniche di signal processing hanno consentito in seguito di poter utilizzare modelli statistici, in particolare modelli Markoviani nascosti (*Hidden Markov Models*), inizialmente proposti in [6] e successivamente perfezionati in [4] con la modellazione esplicita di eventuali situazioni di errore. Recenti ricerche fanno uso di modelli Markoviani nascosti gerarchici (*Hierarchical Hidden Markov Models*) e di modellazione degli strumenti basata su campionamenti [1], oppure di reti Bayesiane per l'apprendimento da esecuzioni passate [7].



(a)



(b)

Figura 1. Allineamento spartito-esecuzione

## MODELLO DEL SISTEMA

Il sistema presentato accetta in ingresso una partitura in formato MIDI ed un flusso audio digitale. Inizialmente la partitura è trasformata in un grafo adatto ad essere usato dal modello Markoviano. Durante l'esecuzione, il flusso audio è diviso in finestre di lunghezza fissa (*frame size*) ad intervalli regolari (*hop size*): ogni finestra corrisponde ad un intervallo temporale del modello. Le probabilità di osservazione per i vari stati del grafo sono calcolate a partire da caratteristiche spettrali estratte da ogni intervallo temporale, e l'allineamento è calcolato usando algoritmi classici descritti nella letteratura dei modelli Markoviani nascosti. Una panoramica del funzionamento del sistema è illustrata in Figura 2. Di seguito sono riassunte le caratteristiche principali del modello (si faccia riferimento a [3] per una descrizione dettagliata).

## Struttura del grafo

Si possono distinguere due livelli di astrazione nel grafo: il *livello superiore* corrisponde alla sequenza di eventi nello spartito, mentre il *livello inferiore* modella ognuno di questi eventi. L'uso di due livelli di astrazione riflette la separazione concettuale tra due possibili sorgenti di errore: il livello superiore si riferisce ad errori nell'esecuzione o nello spartito (note sbagliate, saltate, aggiunte), mentre il livello inferiore

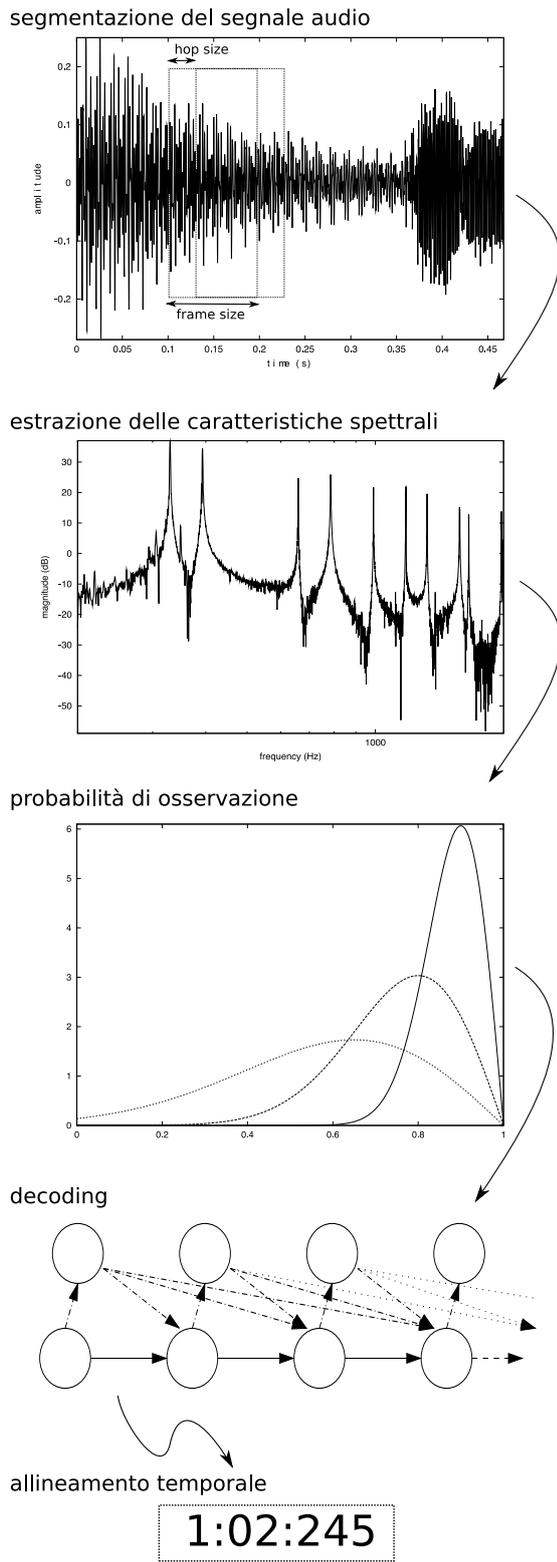


Figura 2. Funzionamento del sistema

re è legato al fatto che le caratteristiche acustiche di un'esecuzione reale possono variare drasticamente, a seconda della tecnica di esecuzione, dell'acustica della stanza, degli strumenti usati e dell'attrezzatura di registrazione, e quindi si prestano ad essere modellate statisticamente.

#### Livello superiore

Il livello superiore, nella sua versione più semplice, modella la sequenza di note e accordi (eventi) come una serie di stati (*event states*) collegati a formare una catena (Figura 3(a)), in cui tutte le probabilità di transizione sono pari a 1. Charamente questo approccio non tiene conto di possibili errori che potrebbero causare incongruenze nell'allineamento. Per ovviare al problema è stato introdotto un nuovo tipo di stato, denominato *ghost state*. Nella nuova struttura del grafo, ad ogni event state è associato un ghost state, il quale è a sua volta connesso ai successivi event states (Figura 3(b)). Le probabilità di transizione in uscita dai ghost states sono decrescenti relativamente alla distanza dagli event states, perché si suppone che l'errore di esecuzione sia *locale*. In caso di esecuzione corretta, il percorso più probabile estratto dall'algoritmo di decoding sarà contenuto nella catena inferiore, mentre in presenza di errori locali potrà passare per uno o più ghost states e riallinearsi in seguito.

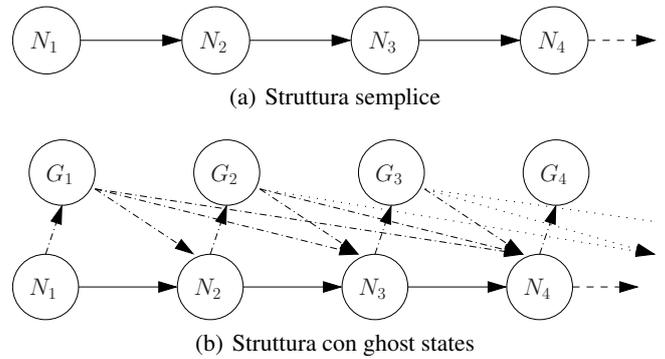


Figura 3. Livello superiore del grafo

#### Livello inferiore

Il livello inferiore modella le caratteristiche attese del flusso audio in entrata. Ogni event state è modellato come una catena di *sustain states*, ognuno caratterizzato da una probabilità costante di autotransizione, seguita da un *rest state* (Figura 4); i primi sono legati alle caratteristiche armoniche del segnale in entrata, mentre i secondi modellano la possibile presenza di silenzio alla fine di ogni evento, che può essere dovuto ad effetti stilistici come lo staccato. È facile dimostrare che il numero di istanti temporali trascorsi in tale catena segue una distribuzione negativa binomiale:

$$P(t) = \binom{t-1}{n_s-1} p^{t-n_s} (1-p)^{n_s}$$

dove  $n_s$  è il numero di *sustain states* e  $p$  la loro probabilità di autotransizione; mentre  $n_s$  è di solito fissato,  $p$  è calcolato per fare in modo che il valore atteso del tempo trascorso corrisponda alla durata nominale dell'evento riportato nello spartito.

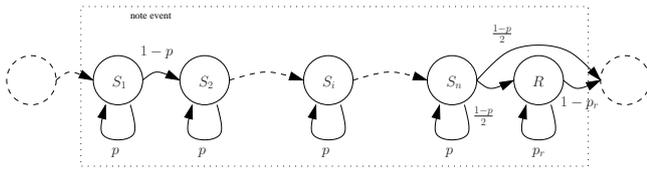


Figura 4. Struttura di un event state

### Probabilità di osservazione

Idealmente l'informazione più utile sul segnale audio in ingresso sarebbe una descrizione delle frequenze corrispondenti alle varie note suonate in ogni istante; sfortunatamente i vari algoritmi proposti finora nella letteratura per il riconoscimento delle frequenze non sono sufficientemente precisi per segnali polifonici. Il segnale quindi non è analizzato, ma è invece confrontato con le caratteristiche attese corrispondenti ai vari stati; per ogni tipo di stato il calcolo delle probabilità di osservazione è differente:

- *sustain states* - ad ogni stato è associato un banco di filtri passa banda, centrati sulle frequenze corrispondenti alle armoniche della nota/accordo a cui lo stato fa riferimento. La probabilità di osservazione  $b_i$  per l' $i$ -esimo stato è calcolata come  $b_i = f\left(\frac{E_i}{E_{tot}}\right)$  dove  $E_{tot}$  è l'energia del segnale nell'intervallo temporale corrente,  $E_i$  è l'energia del segnale filtrato e  $f(\cdot)$  è la funzione di densità di una distribuzione di probabilità continua;
- *rest states* - la probabilità di osservazione è calcolata come una funzione del rapporto tra l'energia del segnale nell'intervallo temporale corrente e una soglia di riferimento;
- *ghost states* - la probabilità di osservazione è calcolata come una media pesata delle probabilità di osservazione relative ai *sustain states* a cui lo stato è collegato. Poiché si ipotizza che l'errore sia locale, la distribuzione dei valori di peso è decrescente all'aumentare della distanza;

Per le distribuzioni di probabilità relative all'osservazione di *sustain* e *ghost states* si usano due funzioni di densità, in particolare *unilatera esponenziale* e *Rayleigh rovesciata*. La distribuzione unilatera esponenziale, usata in passato in sistemi simili a quello presentato, è una versione rovesciata e scalata della classica distribuzione esponenziale, ed è limitata all'intervallo  $[0, 1]$ ; è data da

$$f(x) = \frac{e^\lambda}{e^\lambda - 1} \lambda e^{\lambda(x-1)} \quad 0 \leq x \leq 1$$

ed è raffigurata in Figura 5(a). Similmente, la distribuzione di Rayleigh rovesciata è una variante della distribuzione di Rayleigh; è data da

$$f(x) = \frac{-(x-1)}{\sigma^2} e^{-\frac{(x-1)^2}{2\sigma^2}} \cdot C_\sigma \quad 0 \leq x \leq 1$$

ed è raffigurata, con diversi valori di  $\sigma$ , in Figura 5(b) (il fattore di scala  $C_\sigma$  è circa 1 per tipici valori di  $\sigma$ ). La distribuzione di Rayleigh rovesciata è stata introdotta per porre rimedio ad un problema che si verifica con la distribuzione

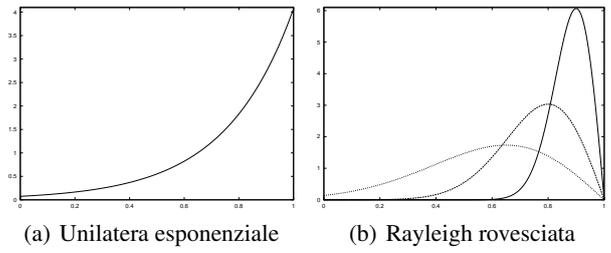


Figura 5. Distribuzioni di probabilità per l'osservazione



Figura 6. Situazione problematica per la distribuzione esponenziale unilatera

unilatera; si consideri un accordo seguito da una serie di note che fanno parte dell'accordo stesso (Figura 6): in questa situazione lo stato associato all'accordo risulterà sempre il più probabile se è utilizzata una distribuzione crescente su tutto l'intervallo  $[0, 1]$ , in quanto l'insieme delle armoniche dell'accordo è un sovrainsieme delle armoniche delle altre note; una distribuzione come quella di Rayleigh invece, decrescente dopo un certo punto, permette anche a note singole di risultare più probabili di eventi più densi.

### Decoding - Allineamento

In ogni intervallo temporale le probabilità di osservazione viste sopra sono usate per calcolare la probabilità di ogni stato di essere attivo, ed estrarre quindi la posizione più probabile nello spartito. Il calcolo è effettuato tramite un algoritmo standard nella letteratura dei modelli Markoviani nascosti [5], basato sulle cosiddette *forward variables*:

$$\alpha_1(i) = \pi(i)b_1(i)$$

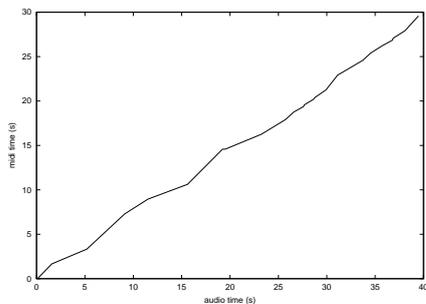
$$\alpha_{t+1}(i) = \left[ \sum_{j=1}^N \alpha_t(j)a_{ij} \right] b_{t+1}(i)$$

dove  $\pi(i)$  è la probabilità che l' $i$ -esimo stato sia uno stato iniziale,  $\alpha_t(i)$  rappresenta la probabilità di trovarsi nello stato  $i$  all'istante  $t$ ,  $a_{ij}$  è la probabilità di transizione dallo stato  $i$  allo stato  $j$ ,  $b_t(i)$  è la probabilità di osservazione relativa allo stato  $i$  al tempo  $t$ . Di solito  $\pi(i)$  è inizializzato ad un valore positivo per gli stati corrispondenti alla nota (accordo) iniziale dello spartito. L'uscita del sistema al tempo  $t$  è il tempo nominale (dello spartito) associato allo stato più probabile, caratterizzato cioè dall'indice

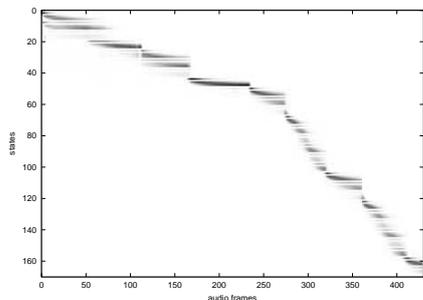
$$q_t = \arg \max_j \alpha_t(j)$$

### RISULTATI SPERIMENTALI

Il sistema descritto permette delle buone prestazioni nell'allineamento di brani caratterizzati da una complessa polifonia, tra i quali brani di musica orchestrale. Purtroppo al



(a) Allineamento temporale



(b) Evoluzione delle forward variables

Figura 7. Allineamento: Haydn, sinfonia n. 104

momento non esiste una collezione di coppie esecuzione-partitura allineate manualmente da esperti, quindi non è possibile valutare in modo oggettivo le prestazioni di un sistema di score following, ma solo qualitativamente. In [3] sono presentati diversi esempi di allineamenti, relativi anche alle possibili situazioni che mettono in difficoltà il sistema, e alcuni test su larga scala per valutare le differenze tra i vari parametri usati. In Figura 7 è presentato l'allineamento relativo all'incipit della Sinfonia n. 104 di F. J. Haydn, e l'evoluzione delle forward variables da cui l'allineamento temporale è stato ricavato. Come è possibile vedere dai grafici, il sistema produce un allineamento che è molto simile ad una linea retta; le deviazioni da questa linea ideale riflettono le variazioni della velocità di esecuzione che sono sempre presenti in una esecuzione umana. Dall'evoluzione delle forward variables si nota l'assenza di incertezze da parte del sistema (in ogni istante i valori positivi delle forward variables sono distribuiti su pochi stati).

### CONCLUSIONI E SVILUPPI FUTURI

È stato presentato un sistema in grado di allineare esecuzioni di musiche polifoniche alle rispettive partiture in tempo reale. Sebbene le prestazioni siano decisamente soddisfacenti, il lavoro di sviluppo ha evidenziato alcune aree di ricerca che appaiono molto interessanti:

- utilizzo della trasformata *wavelet* al posto della FFT; il sistema utilizza infatti la trasformata di Fourier per i calcoli delle energie nelle varie bande di frequenza, e questo comporta una risoluzione lineare in frequenza che male si adatta alla distribuzione esponenziale delle frequenze propria della scala musicale. L'utilizzo di una trasforma-

ta *wavelet*, caratterizzata da una risoluzione logaritmica in frequenza, comporterebbe invece una risoluzione costante su tutte le ottave dello spettro considerato; tuttavia si pongono nuovi problemi come la gestione dei differenti ritardi nelle varie bande di frequenza e la scelta delle funzioni di base da utilizzare;

- utilizzo di più sistemi operanti contemporaneamente; durante lo sviluppo del sistema, sono state individuate alcune misure di confidenza del sistema, le quali potrebbero essere usate per pesare l'influenza di ciascun sistema sull'allineamento finale;
- reintroduzione nel modello di caratteristiche specifiche per diversi strumenti, quali la brillantezza dell'attacco, la struttura armonica e in generale tutti i parametri che possano contribuire alla distinzione di eventi difficilmente separabili tramite la sola analisi armonica, quali la ripetizione di una stessa nota nel caso di strumento singolo o la presenza di note armonicamente correlate suonate da strumenti diversi nel caso di musica orchestrale.

### RIFERIMENTI

1. A. Cont, *Realtime Audio to Score Alignment for Polyphonic Music Instruments Using Sparse Non-negative constraints and Hierarchical HMMs*, IEEE International Conference in Acoustics and Speech Signal Processing (ICASSP), 2006
2. R. Dannenberg, *An On-Line Algorithm for Real-Time Accompaniment*, Proceedings of the International Computer Music Conference, 1984
3. N. Montecchio, *Real Time Score Following based on Hidden Markov Models*, Tesi di laurea, Università di Padova, 2008
4. N. Orio, F. Déchelle, *Score Following Using Spectral Analysis and Hidden Markov Models*, Proceedings of the International Computer Music Conference, 2001
5. L. R. Rabiner, *A Tutorial on Hidden Markov Models and Selected Application*, Proceedings of the IEEE, 1989
6. C. Raphael, *Automatic Segmentation of Acoustic Musical Signals Using Hidden Markov Models*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999
7. C. Raphael, *A Bayesian Network for Real-Time Musical Accompaniment*, Advances in Neural Information Processing Systems, MIT Press, 2002
8. D. R. Stammen, B. Pennycook, *Real-Time Recognition of Melodic Fragments Using the Dynamic Timewarp Algorithm*, Proceedings of the International Computer Music Conference, 1993
9. B. Vercoe, *The Synthetic Performer in the Context of Live Performance*, Proceedings of the International Computer Music Conference, 1984



## Attribuzione-Non Commerciale-Condividi allo stesso modo 2.5 Italia

### Tu sei libero:

- di riprodurre, distribuire, comunicare al pubblico, esporre in pubblico, rappresentare, eseguire e recitare quest'opera
- di modificare quest'opera

### Alle seguenti condizioni:

- **Attribuzione.** Devi attribuire la paternità dell'opera nei modi indicati dall'autore o da chi ti ha dato l'opera in licenza e in modo tale da non suggerire che essi avallino te o il modo in cui tu usi l'opera.
- **Non commerciale.** Non puoi usare quest'opera per fini commerciali.
- **Condividi allo stesso modo.** Se alteri o trasformi quest'opera, o se la usi per crearne un'altra, puoi distribuire l'opera risultante solo con una licenza identica o equivalente a questa.
- Ogni volta che usi o distribuisce quest'opera, devi farlo secondo i termini di questa licenza, che va comunicata con chiarezza.
- In ogni caso, puoi concordare col titolare dei diritti utilizzi di quest'opera non consentiti da questa licenza.
- Questa licenza lascia impregiudicati i diritti morali.

**Le utilizzazioni consentite dalla legge sul diritto d'autore e gli altri diritti non sono in alcun modo limitati da quanto sopra.**

Questo è un riassunto in linguaggio accessibile a tutti del Codice Legale (la licenza integrale):  
<http://creativecommons.org/licenses/by-nc-sa/2.5/it/legalcode>