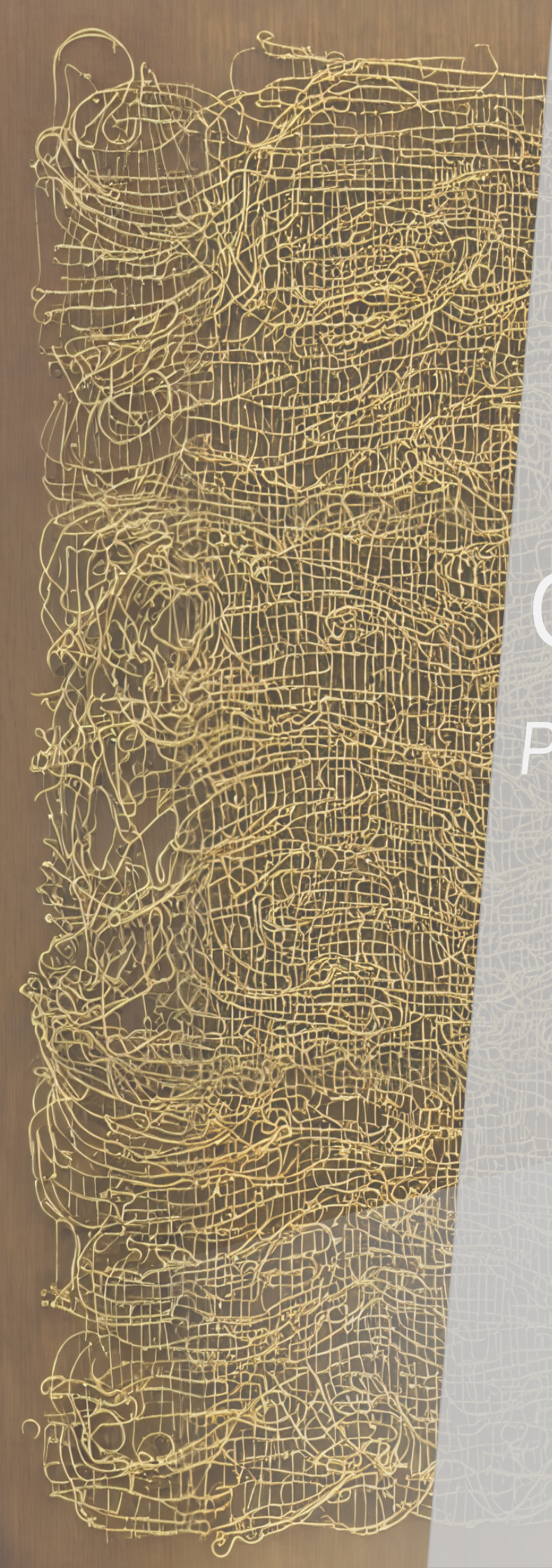


CORPI FISICI

PHYSICAL BODIES

XXIII Colloquio di Informatica Musicale

XXIII Colloquium on Music Informatics

Ancona 25-28 Ottobre 2022

Ancona, October 25-28, 2022

XXIII Colloquio di Informatica Musicale
23rd Colloquium on Music Informatics
Atti della Conferenza
Conference Proceedings

Ancona, 25-28 Ottobre 2022 *Ancona, October 25-28, 2022*

AA.VV.
CORPI FISICI | PHYSICAL BODIES
XXIII Colloquio di Informatica Musicale
23rd Colloquium on Music Informatics

A cura di:
Editors:
Leonardo Gabrielli

©AIMI - Associazione Informatica Musicale Italiana I diritti degli articoli pubblicati rimangono ai rispettivi autori *The copyright of the published papers remain with the respective authors*

ISBN: 9788890341366
ISSN: 2611-7355
Per gentile collaborazione di Davide Rocchesso e Stefano delle Monache
Courtesy of Davide Rocchesso and Stefano delle Monache

Pubblicato da:
Publisher:
DADI - Dip.Arti e Design Industriale, Università IUAV di Venezia

In co-edizione con:
Co-published with:
Università Politecnica delle Marche

Cover image artificially generated from Stable Diffusion with the following prompt: "acoustic piano soundboard with strings, colored in gold, view from top"

```
RIFERIMENTO BIBTEX | BIBTEX REFERENCE:
@proceedings{XXIICIM,
  Editor = {Leonardo Gabrielli},
  Organization ={AIMI - Associazione Informatica Musicale Italiana},
  Publisher = {DADI - Dip.  Arti e Design Industriale.  Università IUAV di Venezia},
  Title = {Corpi Fisici | Physical Bodies, Atti del {XXIII} Colloquio di Informatica
Musicale},
  Year = {2023}
}
```

Quest'opera è distribuita con Licenza Creative Commons Attribuzione - Non commerciale - Non opere derivate 4.0 Internazionale.
Copia della licenza è disponibile presso http://creativecommons.org/licenses/by-nc-nd/4.0/.
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.
To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

organized by:

AIMI

ACUSMATIQ

UNIVERSITÀ POLITECNICA DELLE MARCHE

with the friendly support of:

ernst von siemens music foundation



Comune di Ancona

sponsored by:

ORASTRON

Elk

Audio Modeling

VISCOUNT THE ART OF SOUND

studiologic TRUST YOUR FEELING

FATAR KEY ACTION TECHNOLOGIES SINCE 1956

PSOUND paolo principi

ARTinoise

INKAL VIDEO ART

dissemination partners:

SoundMiT

immagine in movimento CINEMATICA

Risme digitali

XXIII COLLOQUIO DI INFORMATICA MUSICALE
23rd COLLOQUIUM ON MUSIC INFORMATICS

CORPI FISICI

La XXIII edizione dei CIM è stata concepita sotto la necessità di riempire quel vuoto tra assenza e presenza, passato e presente, astratto e fisico. Una ovvia interpretazione del tema principale della conferenza è lo stop forzato imposto alle arti che ha dematerializzato il corpo fisico del performer e dello strumento musicale in un flusso di dati. Questa problematica sociale ha avuto un impatto sulla comunità della computer music: da un lato generando nuovi trend nelle performance musicali mediate dalla rete e stimolando nuove domande di ricerca sul futuro delle arti, dall'altro mostrando quanto è fragile una cultura priva di interazioni corporee. La nostra intenzione è quella di tornare ad accogliere fisicamente la comunità e porre il corpo fisico "in presenza" come oggetto del dibattito artistico e sociale: il corpo del performer, il corpo del suo strumento.

E dunque, lo strumento musicale è la seconda chiave di interpretazione del tema della conferenza. La ricerca nella computer music è in una continua dialettica con la corporeità dello strumento musicale: a tratti deriva verso una tendenza "information-centric" volta a rimuovere ogni vincolo fisico (suono come algoritmo), per poi tornare periodicamente alla fisicità del corpo umano, della voce e dello strumento musicale con nuove tecniche di augmentation o embodiment del linguaggio corporeo (suono da un corpo fisico).

Nuovamente, la nostra intenzione è quella di tornare a discutere lo strumento musicale come corpo vivo e oggetto di indagine della fisica, ponendo la tradizione di Ancona e della Regione Marche sotto il riflettore: conosciuta per la lunga tradizione nella manifattura di strumenti musicali, la nostra regione ha ospitato il suo ultimo CIM nel 1983: un'assenza troppo lunga! Per colmare questo vuoto vogliamo portare nuovamente gli strumenti e i loro esecutori sul palco, con un particolare riguardo verso quelli della tradizione marchigiana: la fisarmonica, l'organo e il sintetizzatore.

PHYSICAL BODIES

The XXIII CIM was conceived from the urge of filling gaps between absence and presence, past and present, abstract and physical. One interpretation of this topic is obvious: the forced stop that affected the arts and dematerialized the physical body of the performer and the musical instruments into a stream of data. This social issue had an impact on the computer music community, driving new trends in network-mediated performance, fostering new research questions on the future of the arts, but also showing how fragile is culture without physical interactions. Our intention here is to physically gather the community again to put the in-presence physical body as the object of artistic and social debate: the body of the performer, the body as an instrument and the body of the instrument.

Indeed, the musical instrument, with its physical presence and its resonant body, is the second key of interpretation of the conference topic. Computer music is in a continuous dialectic with the body of the musical instrument: as a discipline, it strives to be information-centric and remove any physical constraint (sound from the abstract), however on the other side, it periodically returns back to the physicality of the musical instrument with novel augmentation techniques (sound from the physical). Our intention is to discuss the musical instrument with its physicality, and doing so prompts us to put Ancona and the Marche Region under the spotlight. Known for its tradition of musical instruments making, our region hosted its last CIM in 1983, a long absence. To close this gap we aim to bring musical instruments on the stage again, especially those that have their roots here: the accordion, the organ and the synthesizer.

Leonardo Gabrielli, Gianpaolo Antongirolami

Coordinamento generale:
Leonardo Gabrielli, Università Politecnica delle Marche
Gianpaolo Antongirolami, Conservatorio di Perugia

Coordinamento scientifico:
Stefano Delle Monache, Delft University of Technology
Federico Fontana, Università di Udine
Stefano Squartini, Università Politecnica delle Marche
Luca Turchet, Università di Trento

Coordinamento musicale:
Nicola Casetta, Conservatorio di Pesaro
Alessandro Petrolati, Apesoft

Comitato di programma:
Federico Avanzini, Università di Milano
Stefano Delle Monache, Delft University of Technology
Federico Fontana, Università di Udine
Giorgio Klauer, Conservatorio di Venezia
Stefano Papetti, Zurich University of the Arts
Anna Terzaroli, Conservatorio di Benevento

Comitato locale:
Michela Cantarini, Università Politecnica delle Marche
Emanuele Storti, Università Politecnica delle Marche

Comitato di revisione:
Andrea Agostini
Nicoletta Andreuccetti
Massimo Vito Avantaggiato
Stefano Bassanese
Tonino Battista
Giovanni Bedetti
Alberto Boem
Alessio Brutti
Michele Buccoli
Nicola Buso
Diego Capoccitti
Nicola Cappelletti
Maura Capuzzo
Luigi Ceccarelli
Manfredi Clemente
Samuele Cornell
Stefano D'Angelo
Nadir Dalla Pozza
Riccardo Dapelo
Amalia de Götzen
Giovanni De Poli
Yuri de Pra
Fabio De Sanctis De Benedictis
Balandino Di Donato
Anthony Di Furia
Agostino Di Scipio

Comitato di revisione (continua):
Roberto Doati
Carlo Drioli
Michele Ducceschi
Alessandro Fiordelmondo
Enrico Francioni
Nicola Frattegiani
Roberto Fusco
Alessio Gabriele
Leonardo Gabrielli
Michele Geronazzo
Daniele Ghisi
Maurizio Giri
Andrea Gulli
Robert Jack
Giorgio Klauer
Sergio Lanza
Giacomo Lepri
Pasquale Lisena
Luca Andrea Ludovico
Eric Maestri
Marcella Mandanici
Raul Masu
Davide Andrea Mauro
Alessandro Ilic Mezza
Luigi Monteanni
Sandro Mungianu
Federico Ortica
Paolo Pachini
Stefano Papetti
Mattia Parisse
Pietro Polotti
Niccolò Pretto
Alessandro Giuseppe Privitera
Luca Richelli
Davide Rocchesso
Antonio Rodà
Cristina Rottondi
Mario Salvucci
Dario Sanfilippo
Sylviane Sapir
Francesco Scagliola
Nishal Silva
Giuseppe Silvi
Simone Spagnol
Emanuele Storti
Anna Terzaroli
Matteo Tomasetti
Stefano Tomassetti
Andrea Valle
Luca Vignati
Federico Visi
Gualtiero Volpe
Stefano Zambon
Roberto Zanata
Victor Zappi
Laura Zattra
Paolo Zavagna

INDICE | *SUMMARY*

Programma Scientifico | *Scientific Program* 6
Programma Musicale | *Music Program* 12
Abstract dei paper | *Papers Abstracts* 16
Abstract delle opere | *Musical Works Abstracts* 27

SCIENTIFIC SESSION 1: *Streams of Data*
Francesca Magro, Federico Fontana and Hanna Järveläinen
Il ruolo dell'Informazione Uditiva nelle Associazioni Verbali Multisensoriali 37
Fabio De Sanctis De Benedictis
Openmusic e Modalys per Nicola Baroni 44
Marco Tiraboschi and Federico Avanzini
SAMPLE: a Python Package for the Spectral Analysis of Modal Sounds 50

SCIENTIFIC SESSION 2: *Physical Bodies/Physical Instruments pt.1*
Luca Befera
Staging a Contemporary Hierophany: Strategies of Rhythm and Algorithmic Violence in "Δnfang" by the Fronte Vacuo Collective 56
Bleiz Macsen Del Sette, Simone Gregis, Stefano Mancuso and Sylviane Sapir
The Prague Experiment: il Suono dell'Immaginazione 64
Stefano Bassanese, Guglielmo Diana, Matteo Marson, Luca Martone and Andrea Valle
Organum Pineale. Design and Implementation of a Site-Specific Sound Installation 72
Andrea Cera, Corrado Canepa, Nicola Ferrari, Paolo Coletta, Simone Ghisio and Antonio Camurri
Interactive Sonification of Expressive Gesture: the DanzArTe - Emotion Wellbeing Technology Project 79

SCIENTIFIC SESSION 3: *Territories and Spaces*
Michele Geronazzo
L'Audio Egocentrico negli Ambienti Virtuali 87
Leonardo Gabrielli
The Italian Electronic Musical Instruments Manufacturers: a Historical Perspective and Open Research Questions 92
Federico Martusciello
The Reality of the Reproduction. Aesthetics of an Approach to Sound Design in The Soundscape Composition: a Case Study 101
Stefano Catena
Concepts and Approaches in Analysing Spatial Sound Movements: a Link between Mozart and Acousmatic Music 105

SCIENTIFIC SESSION 4: *Physical Bodies/Physical Instruments p.2*
Stefano Papetti, Hanna Järveläinen and Eric Larrieux
Uno Studio sul Controllo di Intonazione di un Sintetizzatore tramite Gesti di Spinta e Trazione 112
Silvia Lanzalone, Emanuela Mentuccia and Imma Battista
Memoria Solis, a Sustainable and Adaptive Musical Installation 118
Stefano Silvestri
Una scheda di Elaborazione Musicale Riprogrammabile e Controllabile attraverso un Ambiente di Live Coding 124

WORKSHOP: *Technologies for Music Education and Rehabilitation*
Marcella Mandanici, Simone Spagnol, Federico Avanzini, Adriano Baratè and Luca Andrea Ludovico
Una Tassonomia Multidimensionale delle Applicazioni per l'Educazione Musicale 132
Marcella Mandanici and Stefano Delle Monache
Technology Integration for the 21st Century Musician 140
Maurizio Botti, Maura Castelnovo, Ottavia Marini, Sara Valente and Marcella Mandanici
Move: un Ambiente Multimodale Interattivo per l'Apprendimento 148
Davys Moreno
Music for All: Finding Digital Resources 156

STUDENT POSTER SESSION

Matteo Sacchetto, Cristina Rottondi, Antonio Servetti, Louena Shtrepi, Marco Masoero and Andrea Valle
HiFiReM: un Approccio Unificato, Web e Nativo, per la Didattica Musicale Remota 165
Domenico Stefani
Riconoscimento in Tempo Reale di Tecniche Espressive per Chitarra su Embedded Computers 169
Luca Vignati
Deep Packet Loss Concealment per Sistemi di Interazione Remota in Tempo Reale 173
Alessio Mastrorillo, Giuseppe Silvi and Francesco Scagliola
Implementing UHJ Stereo in the Envelop for Live Suite 177
Paolo Marrone
A Web Playground for Ciaramella 180
Matteo Tomasetti
CIMIL VR Music System: un Ambiente Multiutente Condiviso per Comporre e Suonare Insieme nella Realtà Virtuale 184
Michele Rossi
Riconoscimento Automatico di Catene di Effetti per Chitarra Elettrica 188
Alberto Rizzo
On Making Feedback Instruments Lively - a Case Study 192



PROGRAMMA DEL COLLOQUIO

CONFERENCE PROGRAM

PROGRAMMA SCIENTIFICO | *SCIENTIFIC PROGRAM*

TUESDAY 25 OCTOBER

17.00	**Bootstrap of the Re.corder Hackaton with ARTINOISE (Aula Didattica)** **More info: https://fb.me/e/3SzgnxTa3**
21.00	**Opening Concert (Auditorium)**
	Brevi saluti istituzionali e cerimonia di consegna degli attestati del corso di perfezionamento "Ingegneria degli Strumenti Musicali" UNIVPM Concerto: – Kaija Saariaho Jardin Secret II (1984-86) harpsichord and tape – Arne Nordheim Dinosauros (1971) accordion and tape – James Tenney SAXONY (1978) saxophone quartet and live electronics (MARS) – Luigi Nono Post-prae-ludium n. 1 per Donau (1987) tuba and live electronics (MARS) Sylviane Sapir, live electronics Nicola Casetta, sound direction Giancarlo Schiaffini, tuba Roberta Pandolfi, harpsichord Raffaele Damen, accordion Sidera Saxophone Quartet (Gianpaolo Antongirolami, Michele Selva, Michele Bianchini, Daniele Berdini, saxophones)

WEDNESDAY 26 OCTOBER

9.30	**Opening Ceremony**
10.15	**Scientific Session 1: *"Streams of data"***
	— Francesca Magro, Federico Fontana and Hanna Järveläinen *"Il ruolo dell'Informazione Uditiva nelle Associazioni Verbali Multisensoriali"* — Fabio De Sanctis De Benedictis *"Openmusic e Modalys per Nicola Baroni"* — Marco Tiraboschi and Federico Avanzini *"SAMPLE: a Python Package for the Spectral Analysis of Modal Sounds"* — Emanuele Storti *"Towards an Integrated and Explorable Representation of FAIR Music Content: a Knowledge Graph Approach"* Chair: Emanuele Storti
11.15	**Break**
11.45	**Keynote Talk**
	— Elvira Brattico *"Making Sense of Music: Insights from Cognitive Neuroscience"* Chair: Federico Avanzini
12.45	**Sessione A.A.A... AIMI!**
13.00	**Lunch Break**
14.00	**Listening Session 1 (Auditorium)**
15.00	**Scientific Session 2: *"Physical Bodies/Physical Instruments pt.1"***
	— Luca Befera *"Staging a Contemporary Hierophany: Strategies of Rhythm and Algorithmic Violence in "Δnfang" by the Fronte Vacuo Collective"* — Bleiz Macsen Del Sette, Simone Gregis, Stefano Mancuso and Sylviane Sapir *"The Prague Experiment: il Suono dell'Immaginazione"* — Stefano Bassanese, Guglielmo Diana, Matteo Marson, Luca Martone and Andrea Valle *"Organum Pineale. Design and Implementation of a Site-Specific Sound Installation"* — Andrea Cera, Corrado Canepa, Nicola Ferrari, Paolo Coletta, Simone Ghisio and Antonio Camurri *"Interactive Sonification of Expressive Gesture: the DanzArTe - Emotion Wellbeing Technology Project"* Chair: Michele Geronazzo
16.00	**Workshop: *"A Trip to Mars"***
	Sylviane Sapir will illustrate the MARS, a real-time music computing platform from the 1990s, and compare it to current languages such as Max/MSP. She will discuss aspects of the compositions played with the MARS during the opening concert and explain how their performing environments were prepared. Chair: Leonardo Gabrielli In collaborazione con il Museo del Synth Marchigiano
21.00	**Concert (Auditorium – Selected works)**

THURSDAY 27 OCTOBER

9.30	**Scientific Session 3: *"Territories and Spaces"***
	— Michele Geronazzo *"L'Audio Egocentrico negli Ambienti Virtuali"* — Leonardo Gabrielli *"The Italian Electronic Musical Instruments Manufacturers: a Historical Perspective and Open Research Questions"* — Sridhar Sampath, Udhayakumar Ravirose, Devi Neelamegarajan, Kamalakannan Karuppaiah and Rashmi Eraiah *"Role of Temporal Fine Structure and Envelope in Identifying Indian Ragas - A Prospective Study for Temporal-Based Raga Classifier"* — Federico Martusciello *"The Reality of the Reproduction. Aesthetics of an Approach to Sound Design in The Soundscape Composition: a Case Study"* — Stefano Catena *"Concepts and Approaches in Analysing Spatial Sound Movements: a Link between Mozart and Acousmatic Music"* Chair: Stefania Cecchi
10.45	**Break**
11.00	**Industrial Presentation 1**
	— FATAR/Studiologic e Psound: "Studiologic: dall'algoritmo allo strumento musicale" Chair: Leonardo Gabrielli
11.45	**Keynote Talk**
	— Laura Zattra *"Lo Strumento Computer: Metodi Misti per Studiarne Passato e Presente, tra Archivi, Fonti Disperse e Storia Orale"* Chair: Stefano Papetti
12.45	**Poster Craze**
13.00	**Lunch Break**
14.00	**Listening Session 2 (Auditorium)**
15.00	**Student Poster Session (Foyer Auditorium)**
	— Matteo Sacchetto, Cristina Rottondi, Antonio Servetti, Louena Shtrepi, Marco Masoero and Andrea Valle *"HiFiReM: un Approccio Unificato, Web e Nativo, per la Didattica Musicale Remota"* — Domenico Stefani *"Riconoscimento in Tempo Reale di Tecniche Espressive per Chitarra su Embedded Computers"* — Luca Vignati *"Deep Packet Loss Concealment per Sistemi di Interazione Remota in Tempo Reale"* — Alessio Mastrorillo, Giuseppe Silvi and Francesco Scagliola *"Implementing UHJ Stereo in the Envelop for Live Suite"* — Paolo Marrone *"A Web Playground for Ciaramella"* — Matteo Tomasetti *"CIMIL VR Music System: un Ambiente Multiutente Condiviso per Comporre e Suonare Insieme nella Realtà Virtuale"*

	– Michele Rossi "*Riconoscimento Automatico di Catene di Effetti per Chitarra Elettrica*" – Alberto Rizzo "*On Making Feedback Instruments Lively - a Case Study*" Chair: Andrea Valle
16.00	**Workshop: *"I Generi della Musica Elettronica"***
	Questa sessione vuole essere un punto di partenza per una riflessione approfondita e duratura finalizzata a sensibilizzare la comunità AIMI/SMC sui temi di genere nella Musica Elettronica. Partecipano persone provenienti da tutti gli ambiti delle nostre professioni: dal momento dell'apprendimento alle posizioni senior, nei campi della composizione, della ricerca e della musicologia. Durante la tavola rotonda verranno toccati temi quali le figure storiche di riferimento; la presenza nei programmi di concerti, festival, convegni, luoghi di lavoro; l'avvicinamento alla tecnologia; i cambiamenti generazionali e la differenza tra Italia ed estero. La sessione viene introdotta da una prolusione di Sylviane Sapir. Chair: Sylviane Sapir, Anna Terzaroli, Laura Zattra
17.30	**Aimi Meeting**
21.00	**Concert (Auditorium – Selected Works)**

FRIDAY 28 OCTOBER

9.30	**Scientific Session 4: *"Physical Bodies/Physical Instruments p.2"***
	— Stefano Papetti, Hanna Järveläinen and Eric Larrieux *"Uno Studio sul Controllo di Intonazione di un Sintetizzatore tramite Gesti di Spinta e Trazione"* — Silvia Lanzalone, Emanuela Mentuccia and Imma Battista *"Memoria Solis, a Sustainable and Adaptive Musical Installation"* — Stefano Silvestri *"Una scheda di Elaborazione Musicale Riprogrammabile e Controllabile attraverso un Ambiente di Live Coding"* Chair: Michela Cantarini
10.15	**Industrial Presentations 2**
	— Audio Modeling: SWAM - strumenti acustici virtuali a modellazione: la sintesi supera il campionamento in espressività e controllo CAMELOT - il MIDI come non l'avete mai visto: la patchbay MIDI software più evoluta al mondo alla portata di tutti. — Orastron: "DSP Scalabile" Chair: Claudia Rinaldi
12.00	**Keynote Talk**
	— Henna Tahvanainen, *"The Resonant Body of the String Instrument"* Chair: Davide Rocchesso
13.00	**Lunch Break**
14.00	**Listening Session 3 (Aula Didattica)**
15.00	**Workshop: *"Technologies for Music Education and Rehabilitation"***
	— Marcella Mandanici, Simone Spagnol, Federico Avanzini, Adriano Baratè and Luca Andrea Ludovico *"Una Tassonomia Multidimensionale delle Applicazioni per l'Educazione Musicale"* — Marcella Mandanici and Stefano Delle Monache *"Technology Integration for the 21st Century Musician"* — Maurizio Botti, Maura Castelnovo, Ottavia Marini, Sara Valente and Marcella Mandanici *"Move: un Ambiente Multimodale Interattivo per l'Apprendimento"* — Davys Moreno *"Music for All: Finding Digital Resources"* Chair: Davys Moreno
16.30	**Awards And Closing Ceremony**

ALL DAYS

	Exhibit: Instruments and books (Sala Polveri/Foyer)
	– Museo del Synth Marchigiano e Italiano: rare vintage instruments and synthesizers – Luca Richelli, Mirco Trentin: "Demo Synth" display of various self-made synthesizers – Alessandro Cipriani: books "Musica elettronica e sound design" – Leonardo Gabrielli books: "Wireless Networked Music Performance" and "Developing Virtual Synthesizers with VCV Rack" – Atti di precedenti CIM

PROGRAMMA MUSICALE | *MUSIC PROGRAM*

25/10/2022 - h. 21:00 | Auditorium - opening concert

Kaija Saariaho - Jardin Secret II (1984-86)
harpsichord and tape

Arne Nordheim - Dinosauros (1971)
accordion and tape

James Tenney - SAXONY (1978)
saxophone quartet and live electronics (MARS)

Luigi Nono - Post-prae-ludium n. 1 per Donau (1987)
tuba and live electronics (MARS)

Sylviane Sapir, live electronics
Nicola Casetta and Alessandro Petrolati, sound diffusion
Roberta Pandolfi, harpsichord
Raffaele Damen, accordion
Sidera Saxophone Quartet (Gianpaolo Antongirolami, Michele Selva, Michele Bianchini, Daniele Berdini, saxophones)
Giancarlo Schiaffini, tuba

26/10/2022 - h. 14:00 | Auditorium

Federico Martusciello - Impressioni In-Naturali (2021)
digital support and live performer

Antonio D'Amato - Un certain vertige (2019)
fixed media [Italian premiere]

Riccardo Tesorini - Rather Than Fall (2021)
A/V

Mattia Parisse - Brulicautoma (2021)
fixed media

26/10/2022 - h. 21:00 | Auditorium

Giulio Colangelo - Organismo Aperto No.3 (2018)
fixed media

Cesare Saldicco - Spire II (2017-19 rev. 2022)
prepared piano and live electronics

Alessandro Cipriani, Giulio Latini - The Last Shipwreck (2021)
A/V

Matteo Tundo - Equinoziale (2021)
augmented piano

Roberto Doati - Studio VI (2021)
fixed media [World premiere]

Julian Scordato - Study for a cosmic city* (2019)
graphic sequencer and electronics [Italian premiere]

Andrea Nicoli - Le rive di Mykines (2021)
alto saxophone and electronics

Mauro Diciocia - Amusia Crescendo II (2021)
fixed media

Nicoletta Andreuccetti - ">=< CPD In Two minds" (2016 rev.2022)
soprano saxophone, synthesis sounds and live electronics

Luigi Ceccarelli - Neuromante x 4 (1993-2022)
saxophone quartet and fixed media [World premiere]

Roberta Pandolfi, piano
* Arazzi Laptop Ensemble (Giorgio Klauer, Luca Richelli, Julian Scordato, Paolo Zavagna)
Michele Bianchini, alto saxophone
Gianpaolo Antongirolami, soprano saxophone
Sidera Saxophone Quartet (Gianpaolo Antongirolami, Michele Selva, Michele Bianchini, Daniele Berdini)

27/10/2022 - h. 14:00 | Auditorium

Andrea Laudante - Perceptions of Prakṛti (2021)
fixed media [Italian premiere]

Paolo Montella - Ossa (2022)
fixed media

Marc Estibeiro, David Cotter - Latent (2021)
two guitars* and electronics

Dario Buccino - Ero già a me n. 1 (2004)
fixed media

Michele Del Prete - Quasi Andante (2022)
digital pipe organ, fixed media and live electronics [Italian premiere]

* Marc Estibeiro and David Cotter, guitars
Aleksey Vylegzhanin, organ - Davide Gagliardi, electronics

27/10/2022 - h. 21:00 | Auditorium

Nicola Cappelletti - Dissipatio HG (2020)
bass clarinet and electronics

Marco Marinoni - GIÀ CORALLO SONO LE SUE OSSA (2019)
fixed media [World premiere]

Cesare Saldicco - Spire VII (2018)
soprano saxophone and electronics

Luigi Ceccarelli - PIC (2011)
A/V

Roberto Zanata - Meccanica II (2017)
accordion and live electronics

Stefano Mancuso - La Pura Coscienza (2022)
accordion and live electronics [World premiere]

Fabio De Sanctis De Benedictis - Fabula (2017-2022)
baritone saxophone and live electronics [World premiere]

Antonio Mazzotti - SUSTAIN (2021)
laptop, mixer, video-sound diffusion system [World premiere]

Andrea Veneri - Iperventilazione (2020)
accordion, prerecorded sounds and digital sound processing

Matteo Tomasetti - Njörðr (2021)
fixed media

Raffaella Palumbo, bass clarinet
Gianpaolo Antongirolami, saxophone
Raffaele Damen, accordion

28/10/2022 - h. 14:00 | Aula Didattica

Dariusz Mazurowski - Dossier of Oblivion (2015)
fixed media [Italian premiere]

Mattia Olgiati - Chants croisés (2022)
fixed media [World premiere]

Jean-Basile Sosa - Liberté chérie (2020)
fixed media [Italian premiere]

Damián Anache - Por Mi Culpa (2022)
A/V

Francesco Altilio - Hidden Sounds (2022)
fixed media [World premiere]

Touchy Toy Collective - Errore di Matrice (2022)
A/V [World premiere]



ABSTRACT DEI PAPER

PAPERS ABSTRACTS

ABSTRACT DEI PAPER | *PAPERS ABSTRACTS*

REGULAR TRACK

Francesca Magro, Federico Fontana and Hanna Järveläinen | *The role of auditory information in multi-sensory verbal associations* (ITA: Il ruolo dell'informazione uditiva nelle associazioni verbali multisensoriali)

This work explores the results of a research conducted in 2021 by Federico Fontana, Hanna Järveläinen and Maurizio Favaro, inspired by studies on cross-modal responses that occur when listening to non-words such as Takete and Maluma, and by the existence of a cross-modal correspondence between the visual perception of a trajectory and the subjective production of movement; to investigate the association between auditory perceptions and verbalizations they chose to use auditory stimuli, not related to words but to the physical event of rolling a ball. Their results showed an apparent auditory effect on these associations: there was a bias towards Takete in the audio-visual condition. In the present research, a binary response test was performed (Takete/Maluma): twenty-four participants were presented with a random sequence of sound or silent videos or audio files representing the animation of a smooth or deformed ball rolling along a curved or zig-zag path on metal or plastic. This exploratory experiment showed that a rolling ball event can be perceived as Takete or Maluma, depending on the type of path and the auditory or visual modality of the event: specifically, the auditory one seems to be perceived more as Takete.

Fabio De Sanctis De Benedictis | *Open Music and Modalys for Nicola Baroni* (ITA: Open Music e Modalys per Nicola Baroni)

The Modalys library for Open Music shows significant points of interest, but does not appear sufficiently documented for early stage users. After some preliminary considerations, this paper examines a specific case, the creation of audio material for "Reversio", for cello and live electronics, composed for Nicola Baroni, proceeding from the problems faced in learning the implementation of Modalys in Open Music, until the presentation of two Modalys instruments among those built for the composition in question. We conclude with brief considerations on future works and implementations.

Marco Tiraboschi and Federico Avanzini | *SAMPLE: a Python Package for the Spectral Analysis of Modal Sounds*

We present SAMPLE, a Python package of tools for spectral analysis and modal parameters estimate. The core of the package is an implementation of the "Spectral Analysis for Modal Parameters Linear Estimate" (SAMPLE) algorithm. This includes a custom implementation of a Sinusoidal Analysis algorithm based on Spectral Modelling Synthesis. Our custom implementation is specifically designed for modal tracking. We also included utilities for automatically tuning the algorithm parameters, using a Bayesian optimization method based on Gaussian

Processes. For this purpose, we implemented efficient routines for computing perceptual audio representations for loss functions, such as the multiscale-spectrogram, the mel-spectrogram and the cochleagram. The package also comes with a Graphical User Interface, which allows to load and trim audio inputs, set the algorithm parameters, run the algorithm, listen to a resynthesis of the input, and export the results. The GUI is distributed both as an extra for the Python package and as a standalone executable.

Luca Befera | *Staging a Contemporary Hierophany: Strategies of Rhythm and Al-gorithmic Violence in "△nfang" by the Fronte Vacuo Collective*

The present paper studies the most recent performances of the Fronte Vacuo collective, which merge ubiquitous digital elements with the performers' body dimension. Specifically, I will consider the case of △nfang, the first and prototypical piece of the Humane Methods cycle, delving into the algorithm processing and staging. The software determines repetitive sound and light patterns based on the AI underlying learning process. Its real-time evolution continuously redefines the piece's multimedia content, also influencing human beings' behaviour. Naked and speechless performers enact obsessive actions aimed to express the software constraints and hegemony. The dramaturgy and overall structure are hence fixed, whereas internal parameters continuously change. Within this framework, I hypothesise that both the algorithm and performers might be conceived as instruments: the former, as built by the authors to autonomously play; the latter, as controlled and irrevocably influenced by the machine. Repetition stands as the fundamental parameter that permeates gestures, bodies, digital artefacts, and symbolic meaning towards a ritual attitude, enacted as an unavoidable as much as sterile practice.

Bleiz Macsen Del Sette, Simone Gregis, Stefano Mancuso and Sylviane Sapir | *The Prague experiment: the sound of imagination* (ITA: The Prague Experiment: il suono dell'immaginazione)

The following article describes the design and implementation of the sound elements of The Prague Experiment, a performative installation on the generative mechanisms of human imagination. The installation was created for the 2019 edition of the Prague Quadrennial, an international event focused on the role of scenography in the performing arts and architecture. The "sound of imagination'" was created by sonifying the visitors' biosignals through a musical automaton capable of producing a unique two-minute audio instance for each performance. The article aims to present the architecture of the device and the technical and aesthetic-artistic choices that influenced its design and implementation, reflecting on the artistic and educational results achieved.

Stefano Bassanese, Guglielmo Diana, Matteo Marson, Luca Martone and Andrea Valle | *Organum pineale. Design and implementation of a site-responsive sound installation*

The paper describes the sound installation Organum pineale, that has been commissioned by the National Museum of Cinema of Torino in the occasion of the "#FacceEmozioni" exhibition, dedicated to the relation between face expressiveness and emotions. The installation is site-responsive and exploits the peculiar acoustic features of a part of the Museum building, the

so-called "Orecchia". On one hand, the installation was meant to provide a sonic analogon of the main theme of the exhibition, in particular in relation to Descartes' conceptualization of passions. On the other, its design aimed at emphasizing the unusual acoustic features of the Orecchia, a visually empty space, but very rich on the aural side. Organum pineale features a 28 + 4.1. audio system, delivering sound materials originating from a selection of films that have been tagged in relation to Descartes' six main passions. The overall sound process involves spatialized audio and a generative feedback system exploiting the room acoustics.

Andrea Cera, Corrado Canepa, Nicola Ferrari, Paolo Coletta, Simone Ghisio, Antonio Camurri and Alberto Pilotto | *Interactive Sonification of Expressive Gesture: the DanzArTe - Emotion Wellbeing Technology Project*

This paper presents an interactive sonification model of full-body human expressive movement, adopted in the DanzArTe - Emotion Wellbeing Technology project. The model is grounded our Slow Mood and Aesthetic Resonance paradigms as well as on several principles and guidelines discussed in the paper. DanzArTe is a treatment protocol and an interactive system mainly for older people at risk of fragility, grounded on the active experience and real-time processing of visual arts content (classical religious paintings, image manipulation and 3D modelling) and interactive sonification, to support full-body physical activity and cognitive exercise of memory training.

Michele Geronazzo | *The Egocentric Audio in Virtual Environments* (ITA: L'audio egocentrico negli ambienti virtuali)

Relations between listeners, the real world, and virtual environments (VEs) must not only inspire the design of intuitive multimodal interfaces that can be traced back to human natural everyday behavior but also make sense of the mediating action of virtual reality (VR) technologies. This paper aims to describe a theoretical framework for the sonic interaction in virtual environments (SIVE) research area, called egocentric audio, and provides an inclusive vision of current and future challenges. In the VE using immersive audio technologies, the role of VR simulation must be active for a participatory exploration aimed at creating meaning in a network of actors, both human and non-human. The guardian of such a locus for expression of action is the auditory digital twin that promotes intra-actions between humanity and technology, capable of redefining dynamically and fluidly all configurations that are the determinants of an immersive and coherent experience. The idea of entanglement theory is here mainly declined in its egocentric and spatial connotation. It is linked to the emergence of a consciousness of the perceptual ability to listen through an actively transformative relationship orchestrated by a digital twin, creating movement, transparency, and provocative activity in VEs.

Leonardo Gabrielli | *The Italian Electronic Musical Instruments Manufacturers: a Historical Perspective and Open Research Questions*

Italy has gained international recognition in many industrial and handcraft fields, including the manufacturing of musical instruments such as violins and acoustic pianos. In the last years,

with the increasing interest in vintage electronic musical instruments (EMI), international hype has also grown around Italian instruments manufactured in the past decades, which fueled popular interest in the history behind these instruments. Many of the most sought ones were manufactured in Le Marche region in a unique era of economic growth and transition from a mostly rural to an industrial economy. This paper reports preliminary data about musical instrument manufacturers, some historical evidences gathered from informal interviews that may stand as a starting point for future research works and historical investigations related to the musical instrument industry in Le Marche region. Several considerations related to industrial practices and social factors are put forward.

Federico Martusciello | *Aesthetics of an approach to sound design in the soundscape composition: a case study*

In this paper the processes that led to the creation of the composition of the piece Impressioni In-Naturali (for dig- ital support and live performer) will be exposed. This allowed me to focus on a series of relevant issues including performance practice as a construction of the listening ex- perience in an electroacoustically configured space and as a form of ecological awareness. It was also necessary to define an acousmatic form characterized by a form that is not fixed but adaptive to the context, underlining an ac- tive attitude highlighted mainly in live electronics. First, the theoretical assumptions relating to the genesis of the composition will be illustrated and then an analysis of the choices of selection and structuring of the material at the microscopic level (realization of the synthesis sounds) and macroscopic (structuring of the generated material) and the practice of diffusion in relation to an informatic perfor- mance environment.

Stefano Catena | *Concepts and approaches in analysing spatial sound movements: a link between Mozart and acousmatic music*

There are many studies which identify the compositional exploration and exploitation of space throughout history; in the twentieth century — and particularly in acousmatic music — the spatial experience has become one of the central musical concerns. The qualities of sound moving in space might be considered distinct in instrumental or voice-based music if compared to electroacoustic and acousmatic music, where spatialisation technologies and reproduction systems have given a great range of possibilities to composers. By introducing a coupling-decoupling process for the analysis of spatial motions and their sounding parameters, it is possible to abstract movements in space from their intrinsic timbral characteristics that might be applicable in the studies of a range of music, both instrumental and electroacoustic. The concept of spatial sonorous object is introduced and used as basic unit for the analysis of the Notturno in D major K.286 by W.A. Mozart and identification of spatial motion archetypes. In this discussion, a preliminary analysis of 'traditional' spatial sound movements is proposed together with a process of coupling-decoupling, an approach in linking spatial movements from different music contexts and their repurpose in the creative practice.

Stefano Papetti, Hanna Järveläinen and Eric Larrieux | *A study on the pitch control of a synthesizer with push and pull gestures* (ITA: Uno studio sul controllo di intonazione di un sintetizzatore tramite gesti di spinta e trazione)

An experiment is described, conducted with experienced musicians, in which the accuracy of push and pull gestures exerted by a finger to control the pitch of a synthesizer against a reference tone was studied. To this end, an internally developed digital music interface capable of measuring forces in three dimensions applied to its top panel was used. The effect of different types of vibrotactile feedback provided by the interface during the execution of the assigned task was also tested. Finally, some subjective evaluations were collected from the participants regarding the interactive experience.

Silvia Lanzalone, Emanuela Mentuccia and Imma Battista | *Memoria Solis, a sustainable and adaptive musical installation*

The paper describes the work Memoria Solis, a sustainable and adaptive musical installation, that transforms its sounds and develops its musical form in relation to sunlight and shadows projected on its surface, composed of a solar panel. The system uses sunlight both for the production of music and for the production of electricity, making the work environmentally sustainable. The music of Memoria Solis is made starting from the real-time processing of voice and natural sounds, which follow each other over time and evolve based on the amount of sunlight detected on the solar panel. Memoria Solis was conceived and realized in 2018 by Imma Battista and Silvia Lanzalone, with the collaboration of Emanuela Mentuccia and the voice of Eleonora Claps. Memoria Solis is currently exhibited at Città della Scienza in Naples, Italy.

Stefano Silvestri | *A programmable music processing board controlled with a live coding environment*

The article describes the design of a low cost development board dedicated to the synthesis and processing of lo-fi music signals and its control using a programming environment written specially for live coding. The project uses the Arduino microcontroller platform as a digital signal processor and it can be used for a modular structure composition with elementar re-programmable operators both the sound synthesis that the controls parameters connectable via patch. Using a simple hardware computer interface can be connected the modules to the programming environment written in the SuperCollider and Java languages for the developed of live coding management operations for musical structures according to a JavaScript syntax

WORKSHOP: TECHNOLOGIES FOR MUSIC EDUCATION AND REHABILITATION

Marcella Mandanici, Simone Spagnol, Federico Avanzini, Adriano Baratè and Luca Andrea Ludovico | *A multidimensional taxonomy of music education applications* (ITA: Una tassonomia multidimensionale delle applicazioni per l'educazione musicale)

Several changes occurred in the practice of music education throughout the past 70 years for what concerns educational curricula, available technologies, and different teaching strategies and approaches. In order to outline a complete picture of this evolution, the paper presents a database of scientific papers in the field of music education exploiting one or more technological tools. Based on a number of tools for the analysis and organization of the selected contributions, we developed a multidimensional taxonomy of music education applications as a powerful and flexible tool useful for research, analysis, and evaluation. Furthermore, a web platform for consulting the database and analyzing the various dimensions of the taxonomy has been created and is publicly available at http://techmusicedu.lim.di.unimi.it/.

Marcella Mandanici and Stefano Delle Monache | *Technology Integration for the 21st Century Musician*

This paper examines the theme of technology integration in the context of professional musicians training. Whilst some important guidelines are derived from the analysis of frameworks and theories for technology integration at K-12 level, the "Technologies for Music Education" Master in Music Degree program is presented and discussed as a proposal for the design of novel educational paths. Following the TPACK theory, the didactic plan combines content, technological and pedagogical knowledge with the aim of responding to the needs of the music educator of the 21st century.

Maurizio Botti, Maura Castelnovo, Ottavia Marini, Sara Valente and Marcella Mandanici | *MOVE, an Interactive Multimodal Environment for Learning*

This article presents a learning experience in which MOVE - an interactive multimodal environment designed by the students of the Master on Technologies for Music Education in Brescia Conservatory - was used. The system currently includes four different games based on the user's movements tracked by a ceiling-mounted camera. These games were tested in a lower secondary school. During the teaching activities observation techniques were used for the identification of some important characteristics related to the use of MOVE, such as inclusion, involvement and development of memory and creativity.

Davys Moreno, Ana Maia and Paulo Jacob | *Music for all: finding digital resources*

The present study aims to find technological solu-tions to facilitate the inclusion of children with Spe-cial Needs in learning music. To achieve this end, a systematic literature review was conducted, between the years 2017 and 2022, with the keyword "accessible digital musical instruments" in different search engines and databases. In addition, a questionnaire survey was carried out to the musicians integrating the musical groups 5º Punkada and Ligados às Máquinas belonging to APC-Coimbra in Portugal. The results of the questionnaires provide us with information about the difficulties, the life stories, the dreams, the supports, and the technologies that the elements of the two groups use. The results obtained in the literature review refer to the presentation of a set of ADMI, adapted musical instruments, software, etc., that can be used by different users with Special Needs. In conclusion, it is referred the actuality

of this theme, the expensive cost and lack of access to accessible digital musical instruments, the need for further studies, the importance of the training of those involved in the use of these resources. We be-lieve that it is possible, with the participation of the whole community, to continue promoting "music for all".

STUDENTS TRACK

Matteo Sacchetto, Cristina Rottondi, Antonio Servetti, Louena Shtrepi, Marco Masoero and Andrea Valle | *HiFiReM: a unified platform, both web-based and native, for remote music education*

Professional low-latency audio transmission applications aimed at supporting online music sessions have traditionally been developed as native software, i.e., specific to each operating system, with consequent installation, configuration, and use issues. This problem has narrowed its use to a niche of computer-literate musicians. This study proposes a web-based platform, implemented both as a native application and as a web application, aimed at music education in order to facilitate networked music performances with low transmission delay and high audio quality.

Domenico Stefani | *Riconoscimento in Tempo Reale di Tecniche Espressive per Chitarra su Embedded Computers*

In the last decades of the past century, innovations in analog electronics led to the creation of a series of guitar-based instrument-controllers that could produce a wide range of sounds by directly controlling audio synthesizers. Despite the various improvements brought to these systems over the years, they have always been limited to tracking the frequency and intensity over time of the notes played, failing to consider the subtle expressive techniques generally used by guitarists to mutate the sound of the instrument. In this paper I present my research, which aims to use the most modern techniques of deep learning to enable real-time recognition of the expressive technique used on a guitar. Particular attention will be paid to proposing implementations on embedded devices and considering the limitations of these, in a manner that can enable the creation of new intelligent instruments where signal analysis and synthesis of new sounds is carried out in a self-contained manner. Along with a description of the research and methodology used, the results obtained so far in deep learning for real-time classification of audio, classification of expressive techniques, and onset detection are presented. Finally, the lines of research that will be pursued in the near future are presented.

Luca Vignati | *Deep Packet Loss Concealment for Real-Time Remote Interaction Systems* (ITA: Deep Packet Loss Concealment per Sistemi di Interazione Remota in Tempo Reale)

Recent decades, including the recent pandemic, have induced the still-growing need for remote interactions between people. Although in most cases this can be met with the now popular video calling services, some use cases remain excluded from remote interactions. This is the

case for musicians, whose sensitivity to time and sound quality poses a major obstacle for those attempting to develop a solution. The very infrastructure on which the interactions take place, namely the public Internet, is not designed to allow interactions with such stringent requirements. For this reason, any proposed solution must bend to an inconvenient reality: protocols suitable for real-time transmission are "best effort," meaning they provide no guarantees as to whether and when the packets sent will arrive at their destination. This gives rise to Packet Loss Concealment (PLC), or the attempt to guess the contents of a lost packet. The current state of the art in this area is achieved with traditional Digital Signal Processing (DSP) techniques, but recent developments in Deep Learning suggest that better results can now be achieved. In this paper I present the current status of my doctoral research in the area of PLC, with some partial results.

Alessio Mastrorillo, Giuseppe Silvi and Francesco Scagliola | *Implementing UHJ stereo in the Envelop for Live suite*

In this work we introduce a new Ambisonics decoder, an implementation of UHJ stereo in the Envelop for Live suite (E4L). This format was developed by Michel Gerzon and is an alternative to conventional stereo and binaural decoding. We also present a sixth-order quadrature filter made with a cascade of Biquad all-pass filters, necessary to derive the UHJ stereo format from the First Order components. This decoder is developed in gen codebox and is available as open-source repository.

Paolo Marrone | *A Web Playground for Ciaramella*

Ciaramella is a domain specific programming language for audio DSP. It experiments a fully declarative syntax and the Synchronous Data Flow computational model, featuring high modularity and composability. We implemented a web playground for coding in Ciaramella, compiling and instantly generating working web audio plugin prototypes. In order to obtain that, its compiler, Zampogna, which is written in JavaScript, has been embedded in a web page, and it has been augmented for the production of JavaScript code. We developed a simple graphical user interface targeting both domain experts and newcomers. Finally, this work lays the foundations for future WebAssembly support as target code.

Matteo Tomasetti | *CIMIL VR Music System: a shared multi-user environment for composing and playing together in virtual reality* (ITA: CIMIL VR Music System: un ambiente multiutente condiviso per comporre e suonare insieme nella realtà virtuale)

This article describes the research questions and the objectives underlying my doctoral degree, which mainly focuses on the scientific and artistic field of musical practices within virtual reality environments and, more generally, of "extended reality", a field identified with the name "Musical XR". The main goal is to develop a multi-user music system in a shared virtual reality environment where it is possible to compose and perform electronic music. Moreover, I outline the new design guidelines I will develop regarding the project design of some categories of musical parameters control in virtual reality environments that can be helpful and fundamental

to all the "Musical XR" researchers. Furthermore, I describe the methodologies and research projects that I will carry out in the coming years and the work plan I will adopt to finalize and implement the various projects.

Michele Rossi | *Real-Time Classification of Expressive Techniques for Guitar on Embedded Computers* (ITA: Riconoscimento in Tempo Reale di Tecniche Espressive per Chitarra su Embedded Computers)

In the last decades of the past century, innovations in analog electronics led to the creation of a series of guitar-based instrument-controllers that could produce a wide range of sounds by directly controlling audio synthesizers. Despite the various improvements brought to these systems over the years, they have always been limited to tracking the frequency and intensity over time of the notes played, failing to consider the subtle expressive techniques generally used by guitarists to mutate the sound of the instrument. In this paper I present my research, which aims to use the most modern techniques of deep learning to enable real-time recognition of the expressive technique used on a guitar. Particular attention will be paid to proposing implementations on embedded devices and considering the limitations of these, in a manner that can enable the creation of new intelligent instruments where signal analysis and synthesis of new sounds is carried out in a self-contained manner. Along with a description of the research and methodology used, the results obtained so far in deep learning for real-time classification of audio, classification of expressive techniques, and onset detection are presented. Finally, the lines of research that will be pursued in the near future are presented.

Alberto Rizzo | *On making feedback instruments lively - a case study*

The recursive circulation of sound from speakers to microphones have been extensively used as a method to design a new aesthetic for music compositions from the '60s, and it's still an actual research topic nowadays, especially in the NIME community. Feedback systems that involve and live of only air and electricity with can be configured as musical ecosystems for their ability to imitate nature in its manner of operation and because they can become autonomous entities. On the other hand, feedback instruments usually require a haptic trigger for their operation and provide an augmentation of a traditional musical instrument. At the same time, several feedback systems involve the coupling of the performer with the environment and a non-haptic interface. Despite these sensory differences, a feedback musician exposes himself to the same issues: balancing, negotiating control and un-control and finding new pockets of sonic magic. While conventional methods of machine listening are commonly used in the NIME community, I present a novel research with the aim to make a feedback instrument more lively using AI, exploiting the Feedback Shadow instrument as a case study. The instrument mentioned above uses ultrasonic waves as a way to both control gesture and the composition matter taking advantage of some physical properties of sound into the air. Due to its complex nature, the system is contextualised in the optic of cybernetic language, leading to a a better comprehension of the system and its behaviour. The creative output of the research exposed in Shadow Pattern shows the mutated relationship between human-machine, using ML as a tool to maintain homeostasis creating at the same time more complex sonic transitions and keeping a physical and satisfactory engagement with the performer. The visuals reflect this attempt to

engage a ML-mediated synaesthesia between sound and motion through the process of visual listening.



ABSTRACT DELLE OPERE

MUSICAL WORKS ABSTRACTS

ABSTRACT DELLE OPERE | *MUSICAL WORKS ABSTRACTS*

Federico Martusciello / Impressioni In-Naturali

In-Natural Impressions is a piece that focuses on issues related to the composition of the soundscape, operating on the mediation of the relationship between sound and space through a targeted use of simple mimetic sound synthesis forms and involving a performer how to actively interpret the diffusion space. The space as the subject of the composition is thematized starting from the principles of construction and organization of sound coherent with the recording context and through a hyper-realistic approach that raises awareness of the mediation of the acoustic experience. Starting from a phonography of real anthropic silence, in fact, the synthesis sounds are defined coherently with the elements present and / or applicable to the context. The structure of the work plays on the comparison between the "real" and the reconstructed context: the perceived sound environment is real, but the juxtaposition between sound and image is often deliberately decorrelated or ecologically impossible with surrealistic results. This calls for a form of attention that places the listener within the communicative process. The performance practice requires the performer to dwell on the relationship between space and sound in the context of diffusion in which the construction of information and the impression on listening takes place. The way of situating sound in space itself becomes a form of creativity that underlines an acousmatic form that is not fixed (or passive), not reproductive but rather adaptive to the performative context.

Mattia Parisse / Brulicautoma

"Brulicautoma" is a composition for fixed media, based on the sound materials produced by the self-made instrument "Gurdy Gurdy" (electro-mechanical instrument built with recovery materials and based on the operation of the medieval Hurdy Gurdy). The imperfections and errors inherent in the construction materials, as poor materials, imply a proportional spectral richness of the instrument. This complexity of sound has been fertile ground for processing and digital synthesis of sound, that led to the generation of a sound world consistent with that of the instrument. The material obtained is explored as a complex organism, and investigated through spectral and qualitative analysis of the timbral and spectro-morphological characteristics. The exploration extends within it in: a "microphonic" scale revealed the components and processes that give life to the sound object, presenting themselves as new elements, with their own morphological, "physical" and perceptive spectrum characteristics, existing in a totally different physical and temporal scale. The perceptive analysis of the sound material has guided the possible transitions between the various sounds, in a continuous and organic transformation.

Antonio D'Amato / Un certain vertige [Italian premiere]

This piece was originally composed expressly for the directional rotating speaker developed by Peter Färber at Zürcher Hochschule der Künste and a quadraphonic setup. Un certain vertige is a piece and a composition study based on the idea of blurring boundaries between two related but contrasting sound materials, that are progressively processed in order to merge in a unique fluxus. A large amount of processes that corrode and degrade the sound are implemented in order to create subtle organic textures with dense harmonics. Harsh timbres coalesce with smooth sound objects, that again are scraped in turn, in a sort of layered klangfarbenmelodie. In this version the mono track originally intended for the rotating speaker is adapted and processed in order to be reproduced by a static central speaker, a sort of reminiscence of the concerto-grosso dialogue that tries to merge the "soloist" and the quadraphonic setup in a seamless panorama where they respond each another.

Riccardo Tesorini / Rather Than Fall

Rips, sharp lacerations, erosions and burns. Fragments of film that go beyond space and time, in a non-linear, intermittent path. A call to the other side of the screen, which first casts the eye into the depths of the abyss to then make it re-emerge, gracefully floating in a world of ether. The image reveals itself instead of disappearing in what it makes us see, and the different linguistic systems explode into a thousand pieces uncovering unexpected connections between irreconcilable worlds, to the point of creating new immaterial auras.

Giulio Colangelo / Organismo Aperto No.3

corpus sine materia - elettronica sola Lavoro conclusivo di una serie incentrata sull'autodeterminazione della forma, sul contenuto che diventa incontenibile, prende "corpo senza ragione", e sviluppa e dilata in immediate variazioni. Diluendosi e riorganizzandosi in differenti stati, è un organismo aperto che rivolto all'esterno rivela la sua densità; non avendo più ragione di essere avvolto cerca di avvolgere il tutto.

Cesare Saldicco / Spire II

Spira Mirabilis is a research and writing project started in 2009 and dedicated to the solo instrument with or without electronics. All aspects - from formal processes, to syntax elements, through the idiomatic language - revolving around the concept of spiral, both as a mathematical representation as well as inspiration and aesthetic speculation. Spirals become formal metaphors of different musical gestures, especially in terms of timbre and directionality, that through reading and re-reading of the sound material lead the listener in a dimension of incessant research, as suggested by the Latin phrase eadem mutata resurgo, exploring the multiple labyrinths of the psyche, ego, consciousness, spirit, selfness.

Alessandro Cipriani, Giulio Latini / The Last Shipwreck

This the third part of an audiovisual work on the theme of Shipwrecks (both the physical ones in the Mediterranean, and the metaphorical one, pertaining social aspects in the present times. The music consists of electroacoustic variations on the first 16 bars of Matthäus-Passion by J.S. Bach using Cipriani's blocks technique described in the book "Electronic Music and Sound Design vol. 2". The video - inspired by the painting "The Raft of the Medusa" by Théodore Géricault - floating restlessly between latitudes and longitudes of the now useless geographical maps, fragments of lost bodies close to surrendering, wooden wrecks, ropes, sails, white and red cloths, with indelible traces of the Crucifixion merge and mingle with each other. The art book "Shipwrecks/Naufragi" was published in 2022

Matteo Tundo / Equinoziale

Da diversi anni sono interessato alle cose di confine ed in particolare ai confini temporali. In genere si è abituati a conferire ai confini caratteristiche di inizio o di fine, ecco che dunque momenti come l'equinozio divengono simbolicamente l'inizio di una stagione e la fine di un'altra. Ritengo però che sia più interessante pensare a questi momenti di confine come attimi temporali con una propria identità che porta il peso del tempo passato ed è in potenza di divenire qualcosa di nuovo. Questa possibilità del divenire racchiude in sé i princìpi naturali dell'evoluzione attiva, lasciando intravedere le strade possibili, come ombre di un pensiero finalmente realizzato. Il pianoforte in questo brano, con le sue preparazioni ed aumentazioni, è uno strumento vivo, organico ed in potenza, si potrebbe dire uno strumento di confine, che lascia trasparire le possibilità sonore di una storia evolutiva.

Roberto Doati / Studio VI [World premiere]

My Studi I-VIII (2020-21) are inspired by Karlheinz Stockhausen's Klavierstücke I-VIII (1952-55). I tried to recreate the electronic sound of those years: in its main morphology so close to that of piano sounds (due to sharp cuts in the magnetic tape) and in its 'color'. In Studio VI each sound is the sum of a three-fold 'image' obtained with convolution (long decay piano sounds) of Julia set spectrum with 'echoes' of vocal music, squarewaves, piano attacks. The long decay piano sounds are actually the deconvolution of original piano tones with selected short piano attacks; the result is a kind of RM piano. All the piano sounds are sampled from the David Tudor III version of Klavierstück VI. Following the Stockhausen's idea of 'satellites' and main sounds, I wrote an algorithm to serially generate all the parameters and formal organization in a sort of 'bad copy' of the original.

Julian Scordato / Study for a cosmic city (graphic sequencer and electronics, 2019) [Italian premiere]

Study for a cosmic city is an audiovisual work inspired by a utopian urban planning proposal exposed by Iannis Xenakis in his essay entitled La Ville Cosmique (1965). As an attempt to relate computer graphics to the formalization of sound, the structures that characterize the utopian city are designed using superquadratic curves capable of describing reasonable variations of sound parameters, such as amplitude and pitch. Parameter values extracted from buildings and pathways are sent to a networked ensemble of electroacoustic performers via OSC messages. Performers are free to map the score data to the variables of their favorite devices and algorithms. The aim is to preserve an autonomous design of digital musical instruments within a common environment based on data sharing. Performative gestures are left to the sensitivity of each musician. In the graphic representation of the score, each building transmits continuous messages relating to the position of a running cursor. Moreover, paths on the ground activate messages at road intersections and define the macro-formal articulation of the score. The graphic project and management of control parameters have been implemented exploiting advanced features of IanniX (a poly-temporal and multi-formal graphic sequencer based on Xenakis's UPIC) such as recursion, cursor acceleration patterns, and complex curves described by mathematical equations.

Andrea Nicoli / Le rive di Mykines

Often one is fascinated by places one has never visited. So it is for me the coast near Mykines, a small village in the Faroe Islands. But, in the piece I wrote for Michele Bianchini, I don't try to give a description of the place, but rather the suggestions and reflections that, hypothetically, could arise from the contemplation of those landscapes (by now a recurring theme in me). The tape contains only samples of the same instrument processed with different techniques and with different software, such as to generate an imaginary landscape that dialogues, contrasts or integrates with the real instrument.

Mauro Diciocia / Amusia Crescendo II

Secondo movimento di uno studio acusmatico più largo, dedicato al deficit neuropsicologico selettivo conosciuto come 'amusia', disturbo psichico che consistente nella perdita della capacità di riconoscere il valore dei suoni musicali. Sublimando il lemma a metafora - spogliandolo dal mero riferimento clinico - il brano mutua la forza evocativa del termine per interrogare la nostra percezione rispetto al senso narrativo-musicale di quei materiali sonori che le nostre orecchie percepiscono come rumore, interferenza, o semplicemente superflui.

Nicoletta Andreuccetti | >=< CPD In Two minds

In >=< CPD In Two minds, the subtle, high-pitched sounds of the soprano saxophone create a disharmonic and tense sound universe. Convergent Parallel Divergent are the development patterns that structure the relationship between saxophone and electronics, exploring various forms of interaction. The overlapping of sound planes and the resulting intersection of different trajectories give rise to a dynamic development in which electronic treatments emphasize and multiply instrumental gestures and spatial explorations.

Luigi Ceccarelli | Neuromante x 4 [World premiere]

Il titolo prende spunto dal primo romanzo cyberpunk di William Gibson. Gli universi cyberpunk degli anni '60 sono stati la prima rappresentazione della profonda simbiosi tra vita reale e vita virtuale, e i personaggi che popolano questi mondi si muovono con disinvoltura tra questi due stati vivendo costantemente in una realtà ibrida. L'insieme dei quattro esecutori e della parte elettronica forma un unico denso agglomerato elettromeccanico ibrido e indistinguibile, una immagine di quel mondo sonoro denso e complesso che solo oggi è possibile creare. In questo brano convivono al massimo livello tecnico ed espressivo sia la componente virtuosistica dell'esecutore che quella (altrettanto virtuosistica) dell'editing digitale, dando luogo ad una alta tensione performativa, impossibile da raggiungere sia da una esecuzione strumentale pura che da una esecuzione puramente digitale. Neuromante x 4 è una elaborazione di una precedente versione dal titolo "Neuromante" per sax contralto.

Andrea Laudante | Perceptions of Prakṛti [Italian premiere]

According to the Samkyha philosophy system, Prakṛti is the first cause of the manifest material universe—of everything except the purusa (pure consciousness). Prakṛti accounts for whatever is physical, both mind and matter-cum-energy or force. Perceptions of Prakṛti is a part of a cycle of compositions that try to explore two fundamental features present in every aspect of life, in both micro- and macrocosm: repetition and randomness. In my research I always try to identify something that is universal, something that could have express the fundament of life. I found in these two behaviours a way to decode the mechanics of what is around and inside me. Repetition and randomness manifest themselves through interconnection. Every single element in life has its own cycle and everything has an influence on all the rest. All is strictly intricate. Everything is interconnected.

Paolo Montella | Ossa

"In this world which we enter, appearing from a nowhere, and from which we disappear into a nowhere, Being and Appearing coincide". Hannah Arendt, Life of the mind Trying to sever the link between the acoustic phenomenon - even acousmatic one - and its typical essence - concretely understood - by making Being and Appearance coincide. Appearance is the generator of infinite Semblance that need refer to nothing other than the individual and its reproductions. Field-recording, as an operation determined by rituals, times and techniques, qualifies for its non-exhaustive character by releasing the need to be artistically varied. Thus escaping the source into the source itself. In this perspective, the world seems to arrange itself as a true musical paradigm that dictates its syntax, its connections, its structural functions. To borrow Schoenberg's words on theme and melody, a sort of chromatic total that asks the phenomenon and thus the appearance to be structured. In this work, field-recording is grafted onto the practice of dozens of interviews conducted between June and December 2021 with women from small towns in southern Italy. A., 11 years old girl, shows us her field - stark - at the centre of tense and contradictory forces, between her being a child, wanting to become a

woman, her aspirations, and social impositions. Her strength is disarming, the tone of her voice exceptionally beautiful.

Marc Estibeiro, David Cotter / Latent

Latent is a composition for two guitars and electronics which explores the boundaries between the physical presence of the acoustic instruments and the intangibility and abstraction of the electronic part. The environment for the electronic part has been created using the SuperCollider programming language; it uses the live input from the acoustic instruments as the source material throughout. The software also incorporates pitch tracking algorithms in order to offer the guitarists control over the electronics. Using the existing skillsets of the musicians and the natural sound of the instrument, rather than any external controllers or sensors, serves as a way to unite the contrasting traditions and performance practices associated with note-based instrumental music and texture-based acousmatic music. The score for the piece embeds traditional western notation into a graphical environment in order to encourage improvisation and experimentation. The piece exposes and explores the differences and similarities between digital instruments and acoustic instruments. Factors which may encourage or inhibit classical guitarists from engaging with electronics, such as interface design, notation, or choice of processing, are also highlighted in the composition. The guitarists interact with the electronics to enhance, contrast with, or subvert the acoustic, which in turn acts as a point of familiarity in an electroacoustic soundscape. The piece was initially developed during the later stages of the global Covid-19 pandemic. The composers were forced to work remotely and thus had to confront the inherent latency that is present in online collaborations. This gave rise to the title of the work.

Dario Buccino / Ero già a me n. 1

Ero già a me n. 1 is rooted on the HN System® and the jo-ha-kyū principle. The HN System is a musical system I have created, based on the parameterisation of body and mind performative processes and aimed at intensifying the experience of the hic et nunc (here and now), HN. Jo-ha-kyū is a fundamental principle in traditional Japanese arts, according to which nature is governed by cycles of three phases: jo (序), ascension of energy; ha (破), fracture, complexity; kyu (急), apex, descent. Ero già a me nº 1 reduces to a minimum any audial perception of objective spatial reality in order to immerse the listener in a completely mental spatial realm. The mind, deprived of the external reality, inhibits its own cortex activity (conscious) thereby amplifying the sub-cortex activity (subconscious and unconscious). The hypnotic aspect of music or, in other words, its ability to feed the spirit and touch the soul, grows more powerful. The aspiration is to 'play the listener'.

Michele Del Prete / Quasi Andante [Italian premiere]

My work is a reference to Mozart's Andante K616, a piece for mechanical organ. Mozart complained in a letter that this instrument was too childish to him as it could not play as a large organ: it lacked the pedal, a large plenum, and an 8' foundation (it was in fact a small 4' instrument). I wrote this piece for the 1773 Schwarz organ of the Leechkirche in Graz, which is a 4' instrument. In my composition I explore the mimetic relation between an organ that plays automatically (clockwork organ/fixed media), an organ that is played by an organist in a traditional sense of this word, and some interaction of these levels (live electronics). The relation I investigate is not a sheer mimetic one but hypermimetic, i.e. it does not follow the original/copy logic but it rather explores the mise an abyme of entities that refer to each other beyond any metaphysical primacy.

Nicola Cappelletti / Dissipatio HG

Dissipatio HG is an homage to the same title novel by G. Morselli (1977), published posthumously and written shortly before his death by suicide, which is also the theme of the novel: the dissipation of the entire human race is the sudden and unexpected context in which the suicide loses meaning. Only few mechanical devices remains to testify the past existence of mankind, in a mediated new relation with the world where the concepts of time, space and language are newly defined paradigms. The piece reveals then a sound dramaturgy, a search for a new language in which the voice of the bass clarinet tries to reconstruct his self-awareness in a context of illegible signals and disregarded responses, where the electronics shapes the perception of time, and the accidental synchronicities are nothing more than a lack of mimesis of residual and no longer present sounds.

Marco Marinoni / Già corallo sono le sue ossa [World premiere]

Dedicato a Jean-Claude Risset con materiali sonori tratti da "Resonant Sound Spaces" (2001-02). Sotto cinque tese di mare tuo padre giace. Già corallo sono le sue ossa e sono perle, quelli che erano i suoi occhi. Tutto ciò che di lui deve dissolversi viene cambiato dal mare in qualcosa di diverso e più ricco. Ad ogni ora le ninfe del mare una campana fanno rintoccare. [W. Shakespeare, La Tempesta, Atto I, Sc. II]. Contaminazione tra forme di vita organica e cristalli. Ciò che è organico diviene cristallo, nel tempo. L'acqua accoglie i cristalli e li guida verso la rinascita. Osiride, corpo frammentato in parti, le parti di un tutto che vive dei singoli grani di sé. Il mare è memoria e liquido amniotico, si oppone al dissolversi di ogni cosa che muore. Eco di vita, memoria di amori divenuti cristallo.

Cesare Saldicco / Spire VII

Spira Mirabilis is a research and writing project started in 2009 and dedicated to the solo instrument with or without electronics. All aspects - from formal processes, to syntax elements, through the idiomatic language - revolving around the concept of spiral, both as a mathematical representation as well as inspiration and aesthetic speculation. Spirals become formal metaphors of different musical gestures, especially in terms of timbre and directionality, that through reading and re-reading of the sound material lead the listener in a dimension of incessant research, as suggested by the Latin phrase eadem mutata resurgo, exploring the multiple labyrinths of the psyche, ego, consciousness, spirit, selfness.

Luigi Ceccarelli / PIC

The protagonist of this video is the painter Achille Perilli and the soundtrack, recorded in 1989, is an elaboration of his reading of a text that was freely adapted from the Futurist parolibere (freewords) of Fortunato Depero. The video footage was shot in his studio in Orvieto (Umbria, Italy) in the Spring of 2009 and it was then digitally edited in close connection to the soundtrack, using a technique of rhythmical fragmentation typical of the video clip. Achille s voice was divided and sampled into individual phonemes in order to obtain some fundamental rhythmic elements. These sounds then became the basis of complex polyrhythmic structures, which were precisely written out as a musical score. The score is played by two percussionists who, instead of using normal percussion instruments, use their drumsticks to activate two sets of sensors (Octapads) that reproduce the individual phonemes. The timbre of the resulting sounds is therefore that of Achille s voice, which the percussionists play, while creating the rhythmical structure in a live performance.

Roberto Zanata / Meccanica II

Clusters pieni coesistono con improvvise sezioni più distese e dialogiche. I clusters, sporcati da quasi impercettibili movimenti intervallari, sono pensati con il fine di creare una tensione crescente ed esasperata, e costituiscono il fulcro e la struttura del discorso musicale. La dissonanza sta alla base della tessitura di Meccanica II: si tratta di una dissonanza risonante e tangibile, che ha lo scopo di liberare energia triviale, senza mai essere brutale. Il live electronics si propone di infittire il tessuto timbrico e compositivo, dando vita a spazi sonori nuovi e inaspettati.

Stefano Mancuso / La Pura Coscienza [World premiere]

Il brano prevede l'interazione tra due elementi del pensiero umano: il pensiero conscio, interpretato dalla fisarmonica, e l'inconscio, rappresentato dal Live Electronics. Si presenta come una meditazione che porta il pensiero conscio ad interagire con l'inconscio, a scontrarsi con esso e, grazie all'utilizzo di un mantra, a fondersi raggiungendo lo stato di Pura Coscienza. Il fisarmonicista è in grado di modificare in tempo reale la composizione della sezione elettronica, tramite l'analisi FFT dei suoni ripresi ed un sistema di mapping. In particolare, l'elettronica reagisce all'altezza delle note eseguite, alla dinamica, ai registri scelti e al manuale utilizzato. Tutto il materiale elettronico viene generato a partire dalla registrazione di frammenti di fraseggi eseguiti al momento, ripreso tramite due microfoni direzionali, uno per manuale, ed un microfono centrale. Si è scelta una visualizzazione non sempre lineare del tempo, in modo da aiutare l'interprete alla fisarmonica a comprendere la natura dell'esecuzione. La partitura è da intendersi come una traccia di massima con indicazioni sugli eventi, ma è possibile modificare le varie scelte timbriche e di registro se l'interprete dovesse sentirne la necessità. Il brano indaga sulle potenzialità dell'elaborazione elettronica in tempo reale e sulla possibilità di ampliare lo spettro dello strumento acustico, mantenendo un forte rapporto con l'esecuzione e la gestualità del musicista. Infatti l'interprete decide, così come fa con il suo strumento, come deve evolversi l'elettronica. Al contempo però gli algoritmi cercano di influenzare l'interprete alla fisarmonica che quindi deve reagire adeguatamente, non limitandosi all'esecuzione della "parte scritta".

Fabio De Sanctis De Benedictis / Fabula [World premiere]

Fabula (storia, narrazione, in latino) per sax baritono, suoni elettronici e live electronics, è dedicata a Giampaolo Antongirolami, nasce dall'incontro interprete-compositore. Nel brano sono celati elementi biografici: un omaggio a Risset, una canzone popolare delle Marche, l'ultimo terremoto in questa regione italiana, e così via. I principali materiali audio sono originati da multifonici registrati dal sax baritono di Antongirolami, modificati e lavorati da patch di OpenMusic e di Max. Le note risultano dal nome e cognome dell'esecutore; analogamente la struttura del brano. Ne risulta una forma libera, narrativa, non simmetrica, con alcuni ritornelli, ed episodi in cui l'esecutore improvvisa ad ogni performance con sonorità sempre nuove. Una Fabula, appunto.

Antonio Mazzotti / SUSTAIN [World premiere]

"SUSTAIN" was designed with the Computer-Aided Algorithm Video-Music Composition system. It was conceived as a study for the computational model to produce musically meaningful results.The errors, imperfections, and limitations of the particular compositional media are the central constituting elements of the piece. system components are divided into abstractions of musical materials, abstractions of musical procedures, reconfigured by the generative qualities of the unforeseen error. 'SUSTAIN' was implemented in Mathematica, Csound and Kyma, that uses the Pacarana as audio accelerator and Processing for the rendering video. Despite the fact that both the set of symbols and the set of rules are finite, the number of expressions you can

generate by combining the symbols is infinite. In this composition, the errors, imperfections, and limitations of the particular compositional media are the central constituting elements of the pieces. In addition to music, this glitch aesthetic is also exploited in the visual arts. I set up situations in which errors are able to emerge and be exploited in the art making process. In this type of work the artist reconfigure and exploit the generative qualities of the unforeseen error. The composition Audio-Video was conceived as fixed media.

Andrea Veneri / Iperventilazione

Iperventilazione is a composition for accordion and prerecorded sounds and digital sound processing. This piece is inspired by the state of hyperventilation or the increase in the frequency of respiratory acts in conditions of rest which brings the amount of carbon dioxide above the normal amount produced by the body. This condition can lead to states of agitation, throbbing, dizziness, blurred vision, lightheadedness and syncope. A piece that, therefore, has as its pivotal point a fact that is outside the musical sphere but which is conceptually linked to the form, timbre and compositional choices.

Matteo Tomasetti / Njörðr

Njörðr is an electroacoustic composition for multi-channel format that represents a real meeting between human imagination and nature, with one of its most powerful symbols: the sea, which is alluded to through the sounds of waves, storms, drops and its structural metamorphosis. The composition presents an exploration of the essence of these environmental sounds, which are modified, recreated with different materials and amalgamated with traditional processed instrumental sounds and synthetic sound objects, in an attempt to create a surreal and evocative perceptive atmosphere. The initial idea of the piece is based on sound recordings made in Portugal, at the Praia Do Banho and Praia Do Meco, on the south coast of Lisbon. A key feature of the piece is its octophonic spatialization and the resulting spatial images, a symbol of movement and fluidity of the sound sources, which interconnect with the very nature of the soundscapes used. Contrasts and synthetic sound combinations are substituted for ambient and instrumental sounds, alternating the perception between abstract and realistic environments. The manipulations of a spectral and movement type are combined to create dreamlike soundscapes always defined by a sort of articulation of transformations and juxtapositions.

Dariusz Mazurowski / Dossier of Oblivion [Italian premiere]

Dossier of Oblivion is the second part of Hidden Dimensions and may be performed as a separate piece also. It's a very personal artistic expression, dealing with my thoughts of the flow of time and evanescence. The music was composed and recorded at the De eM Studio, between January 2014 and April 2015. The definitive version is a 6-channel audio projection (5.1). Main audio sources include complex analog, digital and hybrid synthesizer patches, various textures created with the use of phase vocoder technology, analysis and resynthesis of various source samples, physical modeling, complex analog-digital processing. For the multichannel mix various advanced software processors were used to obtain a faithful spatial diffusion of the sounds. Premiere: October 2, 2015, Cross-Art Festival.

Mattia Olgiati / Chants croisés [World premiere]

Australia e Mongolia: mondi geograficamente distanti ma accomunati da sonorità affini. In "Chants croisés" questi due mondi si incontrano in una incontrovertibile convoluzione sonora: il suono del didgeridoo e del canto armonico mongolo si intrecciano creando un timbro unico e quasi inscindibile. A coadiuvare questo incontro contribuisce il mare, quasi come una sorta

di connessione wi-fi tra due mondi lontani: risonanze erranti che ci portano in una dimensione percettiva familiare seppur, a tratti, onirica.

Damián Anache | Por Mi Culpa

Rhythm, lights and shadows; micro-details, grains and noises; overlapped layers and textures in a simulated space. Por Mi Culpa (Spanish for mea culpa, a Latin phrase translated as through my fault) rest beneath an oscillation between non-narrative abstract images and the documentation of a body's interaction with the audiovisual technology that produces and records the process of manufacture itself. The image is composed only of video footage and lighting effects generated by computer controlled lamps. The soundtrack is a live performance with no further editing.

Jean-Basile Sosa | Liberté chérie [Italian premiere]

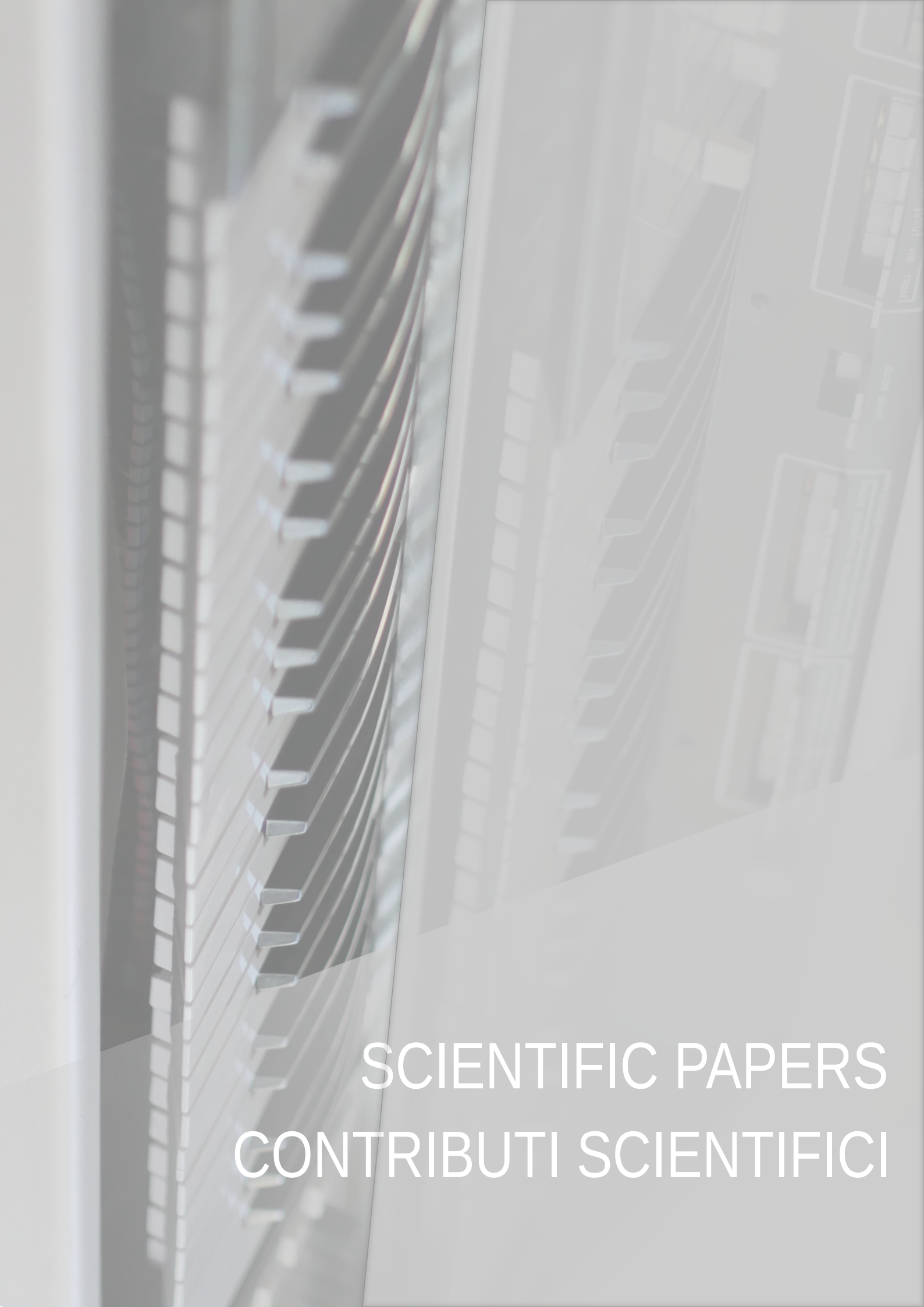
The notion of freedom runs through "Liberté chérie". This notion is thus exploited at different levels of the composition: freedom of rhythmic thought, where rhythm is perceived as a constantly renewed sound movement; freedom in the use of tonal and harmonic components for structural purposes; freedom of musical form dedicated to the expressiveness of electronic sounds; freedom, finally, to treat a perhaps obsolete poetic theme, given that it was a question of bearing witness to the feeling of nostalgia, and of a lost, cherished, and always fraternal childhood.

Francesco Altilio | Hidden Sounds [World premiere]

Tutti i giorni siamo circondati da una molteplicità di suoni. Alcuni sono evidenti, altri meno, molti altri quasi impercettibili. Questo lavoro si concentra su questi ultimi, attraverso un processo di captazione prima, e amplificazione poi. I "suoni inauditi" diventano forti, percettibili. La situazione cambia: in questo brano saranno loro i protagonisti. I suoni a cui mi riferisco provengono da dispositivi di uso: computer, smatphone, schermi, etc. Si tratta di suoni di natura elettromagnetica, generati da questi dispositivi in maniera continua. Le pulsazioni che ascolterete all'interno del brano, per esempio, provengono dalla testina di un hard disk che, sovraccaricata dal lavoro, genera un ritmo ostinato e frenetico. Inserendoli all'interno di un disegno compositivo, tali materiali sono stati elaborati in modo da farli coesistere e coabitare con quelli realizzati da strumenti musicali acustici ed elettronici.

Touchy Toy Collective | Errore di Matrice [World premiere]

The Touchy Toy collective proposes an audiovisual composition based on the violent and compulsive alternation of video patterns. As the title itself suggests, the video is realised by showing a matrix, generated through the use of digital sound waves and noise stored within matrices constructed with the basic Jitter objects within the Max/MSP development environment. Based on the different sound sources, a video output is generated, which is then processed with logical/arithmetic operations between several matrices. The sound is generated through the error due to the transposition from video to audio, resulting in glitch and noise sounds. The audio-video elements are based on glitch and the work manages to explore interesting shapes, gestures and textures. The Touch Toy Collective is composed by: Andrea Lupo, Gabriele Petrillo, Matteo Camerini.



SCIENTIFIC PAPERS
CONTRIBUTI SCIENTIFICI

IL RUOLO DELL'INFORMAZIONE UDITIVA NELLE ASSOCIAZIONI VERBALI MULTISENSORIALI

Francesca MAGRO
Department of Mathematics, Computer Science and Physics (DMIF), University of Udine, Udine, Italy

Federico FONTANA
Department of Mathematics, Computer Science and Physics (DMIF), University of Udine, Udine, Italy
federico.fontana@uniud.it

Hanna JÄRVELÄINEN
Institute for Computer Music and Sound Technology (ICST), Zurich University of the Arts, Zürich, Switzerland
hanna.jarvelainen@zhdk.ch

ABSTRACT

Questo lavoro approfondisce i risultati di una ricerca condotta nel 2021 da Federico Fontana, Hanna Järveläinen e Maurizio Favaro, ispirati dagli studi sulle risposte cross-modali che si verificano a seguito dell'ascolto di non-parole come *Takete* e *Maluma* e dall'esistenza di una corrispondenza cross-modale tra la percezione visiva di una traiettoria e la produzione soggettiva del movimento. Il presente studio contribuisce alla ricerca volta a chiarire il ruolo del suono nelle associazioni verbali multisensoriali attraverso un test di risposta binaria (*Takete*/*Maluma*): una sequenza casuale di video sonori, muti o solo file audio è stata presentata a ventiquattro partecipanti; questi stimoli consistevano in un'animazione di una palla liscia o deformata che rotolava lungo un percorso curvo o a zig-zag su una superficie di metallo o plastica. Questo esperimento esplorativo ha mostrato che l'evento di una palla che rotola può essere percepito come *Takete* o *Maluma* a seconda sia del tipo di percorso che della modalità uditiva o visiva in cui l'evento si presenta: nello specifico, quello uditivo sembra essere percepito più come *Takete*. I nostri risultati mostrano che lo stimolo uditivo ha svolto un ruolo importante nel prendere decisioni non casuali in associazioni verbali multisensoriali, confermando le ipotesi avanzate dallo studio precedente.

1. INTRODUZIONE

Nel 1929, a Tenerife, Köhler [1] ideò l'esperimento fonestesico *Takete*/*Maluma* [2] che consisteva nel chiedere ai partecipanti di attribuire il nome *Takete* o *Maluma* a immagini di forme angolose e spezzate oppure curve e continue: nella maggior parte dei casi i soggetti (adulti spagnoli) associavano la parola *Takete* alle prime e *Maluma* alle seconde. Questo risultato suggeriva che le persone fossero effettivamente portate ad associare, in modo spontaneo, determinate tipologie di nomi a specifiche tipologie di forme e oggetti. Il fenomeno osservato è un esempio classico di corrispondenze cross-modali, ovvero associazioni di domini di caratteristiche in diverse modalità sensoriali [3], e in quel caso avveniva tra suoni di parole senza senso che venivano mappati su forme visive. Le corrispondenze cross-modali, infatti, si verificano anche nel linguaggio umano.

Negli anni successivi sono state portate avanti ricerche a sostegno del simbolismo fonetico, della sua universalità, e della sua derivazione sia da caratteristiche acustiche che articolatorie, relative quindi ai movimenti dei muscoli della bocca coinvolti nell'articolazione delle parole. Per quanto riguarda il simbolismo sonoro, nel 2012 Walker [4] ha mappato la frequenza uditiva su diverse dimensioni semantiche: toni più alti venivano giudicati come più nitidi, spigolosi e luminosi, toni più bassi, invece, venivano associati all'oscurità e giudicati come più morbidi.

Nel 2001 Ramachandran e Hubbard [5] hanno attribuito il bias degli abbinamenti relativo al fenomeno *Takete*/*Maluma* alla coattivazione, con un meccanismo simile alla sinestesia, delle aree motorie o somatosensoriali del cervello umano che si attivano quando le persone articolano le vocali, e delle aree visive del cervello deputate alla percezione del contorno visivo della forma. Questi studiosi hanno suggerito che l'attivazione cross-modale che spiega questo effetto si basa sui neuroni specchio e l'area di Broca; nello specifico, la simulazione interna del gesto motorio articolatorio che corrisponde allo stimolo uditivo viene mappata su specifiche inflessioni fonemiche che vengono poi associate a determinate forme in modo, dunque, non arbitrario [6].

La risposta cross-modale stimolata dall'ascolto di non parole come *Takete* e *Maluma* va oltre le associazioni visive; l'effetto *Takete-Maluma è* legato anche alla sensazione di movimento: nel 2013, Fontana [7] ha condotto un esperimento in cui, tramite braccio robotico con terminazione a forma di penna, la mano di ciascun

Copyright: © 2022 Magro et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

partecipante veniva guidata facendole percorrere una traiettoria arrotondata e una frastagliata (ipotizzando che venissero associate rispettivamente a *Maluma* e *Takete* e ottenendo la conferma di ciò a ricerca ultimata) provando l'esistenza di una connettività tra non-parole e traiettorie tattili.

Sulla base di questi studi e sull'esistenza di un'associazione cross-modale tra la produzione soggettiva del movimento e la percezione visiva della forma definita dalla traiettoria del movimento stesso, Federico Fontana, Hanna Järveläinen e Maurizio Favaro [8] nel 2021 hanno condotto un test in cui sedici partecipanti dovevano abbinare alle parole *Takete* e *Maluma*, video sonori o muti di una pallina (di superficie liscia o dentellata) che rotola lungo un percorso (curvo o seghettato).

Questa ricerca esplorava, per la prima volta, il legame tra parole e feedback uditivo (legato a eventi dinamici) e hanno trovato che la presenza dello stimolo uditivo influenzava in modo significativo il bias delle scelte su *Takete* rispetto alla condizione solo visiva. Hanno ipotizzato, a seguito dei risultati ottenuti, un significativo ruolo del suono nella cognizione di eventi audiovisivi tale da suscitare nelle persone una sensazione di angolosità.

L'obiettivo del presente studio è quello di proseguire la ricerca precedente sul ruolo dell'informazione uditiva nelle associazioni verbali multisensoriali tra non-parole ed eventi audio-visivi, verificando che il bias su *Takete* venga confermato nella condizione solo uditiva e audio/visiva rispetto alla modalità solo visiva, e in caso positivo, la ricerca di eventuali spiegazioni a tale fenomeno.

È stato riprogettato l'esperimento dello studio precedente del 2021, introducendo tre modifiche sostanziali:

- l'inserimento di una condizione solo uditiva, oltre a quella visiva e audiovisiva;
- l'utilizzo di due materiali diversi per la superficie di rotolamento (il metallo e la plastica);
- La creazione di un'animazione tramite grafica 3D che mantiene il percorso e relativi bordi invisibili.

2. METODO

L'esperimento è stato svolto in una stanza silenziosa di uno degli autori, priva di fattori che potessero distrarre i partecipanti.

2.1 Partecipanti

Ventiquattro partecipanti (13 femmine e 11 maschi, di età M=29,04, sd=9,5 anni), tutti con una vista e un udito normali (tre di loro portavano gli occhiali per correggere la miopia), si sono offerti volontari per l'esperimento[1]. Tre di loro avevano una conoscenza pregressa del fenomeno *Takete/Maluma*.

2.1 Setup e stimoli

È stato condotto un test di risposta binaria (*Takete/Maluma*): è stata presentata una sequenza casuale di video sonori, video silenziosi o solo file audio, che riproducevano l'animazione di una palla liscia e una deformata che rotolavano lungo un percorso curvo o a zigzag su superficie di metallo o plastica. I percorsi sono stati intagliati in due tappetini di spugna e la loro altezza è di 69 cm. Le due lastre di metallo e plastica poste alla base non sono state tenute in verticale ma obliquamente, con un'inclinazione di circa 30 gradi, per far rotolare la palla lentamente. Le due palline, in legno, pesano circa 4 g. (fig. 1).

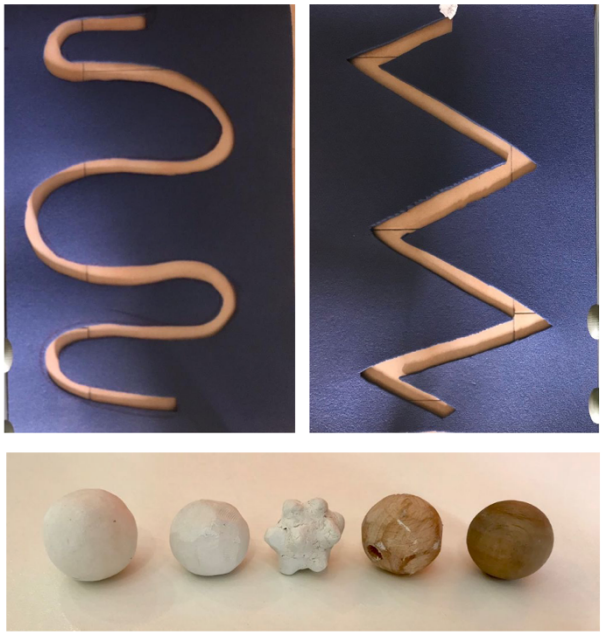


Figura 1. Materiali utilizzati per la realizzazione di percorsi e palline (le prime tre sono di prova, la quarta e la quinta sono invece quelle in legno utilizzate).

Per la registrazione dei suoni del rotolamento dati dalle diverse combinazioni dei materiali (utilizzando percorsi e palle costruiti fisicamente) il microfono è stato posto alla fine del percorso e rialzato rispetto a questo, in un punto di ascolto ideale in cui potesse essere percepita la dinamica del rotolamento e di avvicinamento della palla durante la sua discesa. Con l'utilizzo

[1] A causa della chiusura dell'Università per motivi legati al Covid-19, l'esperimento è stato svolto senza seguire il protocollo standard e il consenso di ciascun partecipante è stato espresso verbalmente e non dietro firma del modulo apposito.

del software *Maya 3D* sono stati ricreati i due percorsi e la discesa della palla in un'animazione 3D. Nello sviluppo dell'animazione si è voluto evidenziare un contrasto della palla con lo sfondo, e sono stati scelti colori neutri come il beige per la pallina, dovuto alla texture del legno, e il blu per lo sfondo (fig 2). È stato dimostrato infatti che il colore rosso migliora le prestazioni nei compiti orientati ai dettagli che richiedono un'attenzione mirata (in questo esperimento però lo scopo non è indirizzare l'attenzione verso una specifica componente), mentre colori più neutri e sui toni del blu si adattano meglio a situazioni normali [9]. Ciascun video ottenuto ha risoluzione di 3840 x 2160 pixel, 16:9, e frame rate di 29.97 fps.



Figura 2. Palla liscia (fatta corrispondere a *Maluma*) a sinistra, palla deformata (fatta corrispondere a *Takete*) a destra.

La fase di editing dei suoni ottenuti con le registrazioni è stata fatta usando il software *Audacity* ed è consistita in una leggera riduzione del rumore di fondo; il rumore è stato successivamente portato in tutti i campioni ad un livello quanto più simile a quello contenuto nel campione meno rumoroso, per pareggiarlo su tutto l'esperimento. Successivamente al montaggio di audio e video, per sottoporre il test ai partecipanti è stata creata un'applicazione da visualizzare su browser (fig 3).



Figura 3. Parte relativa ai pulsanti di una schermata d'esempio dell'applicativo usato per sottoporre i partecipanti all'esperimento: il pulsante 'Play' (per l'avvio della riproduzione del file multimediale), i due pulsanti di scelta 'Takete' e 'Maluma' e il pulsante 'Next' (per il passaggio al file multimediale successivo).

2.2 Procedura

Ciascun partecipante si è seduto davanti a un computer (*MacBook* con display retina da 15") e ha indossato un paio di cuffie chiuse (modello *beyerdynamic DT - 770 Pro*) una volta iniziato il test. La procedura di preparazione all'esecuzione del test è stata standardizzata con l'ausilio di slide esplicative. I partecipanti sono stati divisi in due gruppi: le sequenze di presentazione dei file multimediali erano diverse, così come la posizione a destra o sinistra dei pulsanti di scelta *Takete* e *Maluma*, per eliminare eventuali bias cognitivi non inerenti alle condizioni dell'esperimento. A ciascuno era permesso guardare o ascoltare un video/audio più volte; una volta presa una decisione dovevano cliccare sul pulsante *Takete* o sul pulsante *Maluma* e passare al file multimediale successivo, senza poter poi tornare indietro. I video/audio sono stati inseriti in una sequenza di prove bilanciata in modo casuale, per un totale di 2 palline {t, m}, 2 percorsi {T, M}, 2 materiali {metallo, plastica}, 3 modalità {A, AV, V}, 2 ripetizioni = 48 prove. Per permettere a ciascun partecipante di fare eventuali pause, i 48 video/audio sono stati suddivisi in tre diversi step successivi da 16 file ciascuno. Secondo le rispettive associazioni ipotizzate, la traiettoria, curva o a zig-zag, è stata contrassegnata rispettivamente con M (*Maluma*) o T (*Takete*) e la pallina, liscia o deformata, è stata contrassegnata rispettivamente con m (*Maluma*) o t (*Takete*). Ogni sessione è durata circa 10/15 minuti. Alla fine, ogni partecipante ha lasciato commenti sull'esperienza appena fatta e sull'eventuale conoscenza pregressa del fenomeno.

3. RISULTATI

Da una prima analisi descrittiva (*figg. 4-5*) si evince che, per quanto riguarda la superficie in metallo, le combinazioni incongruenti di percorso e palla (Tm e Mt) portano a risposte generalmente a favore di *Maluma* in condizioni solo video e a favore di *Takete* in condizioni solo audio.

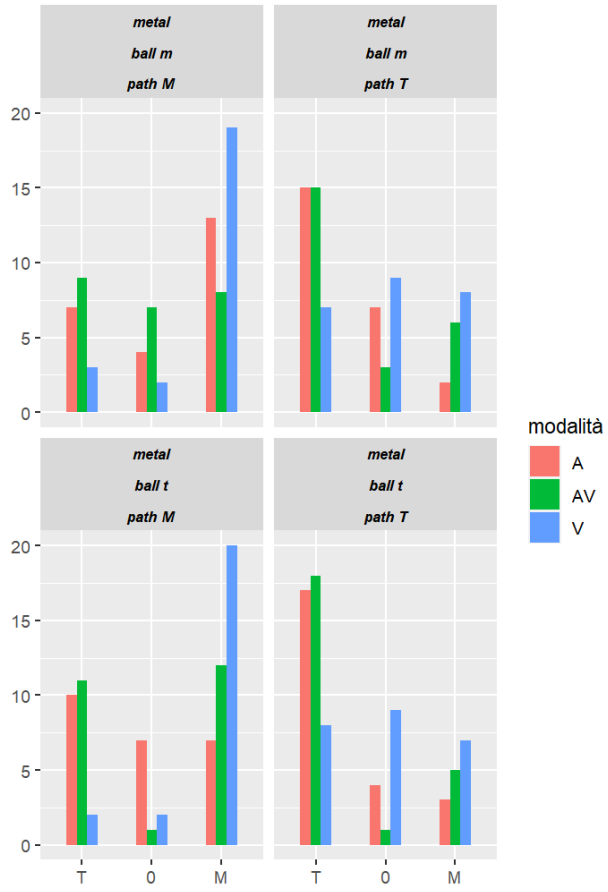


Figura 4. Risultati. Asse y = conteggio delle risposte (T: Takete; 0: incertezza, corrispondente a due associazioni discordanti per uno stesso stimolo da parte del partecipante; M: Maluma) per le tre modalità (A: audio; AV: audio-video; V: video) per ogni combinazione (**metallo**; palla M: palla liscia; palla T: palla deformata; percorso M: percorso curvo; percorso T: percorso a zig-zag).

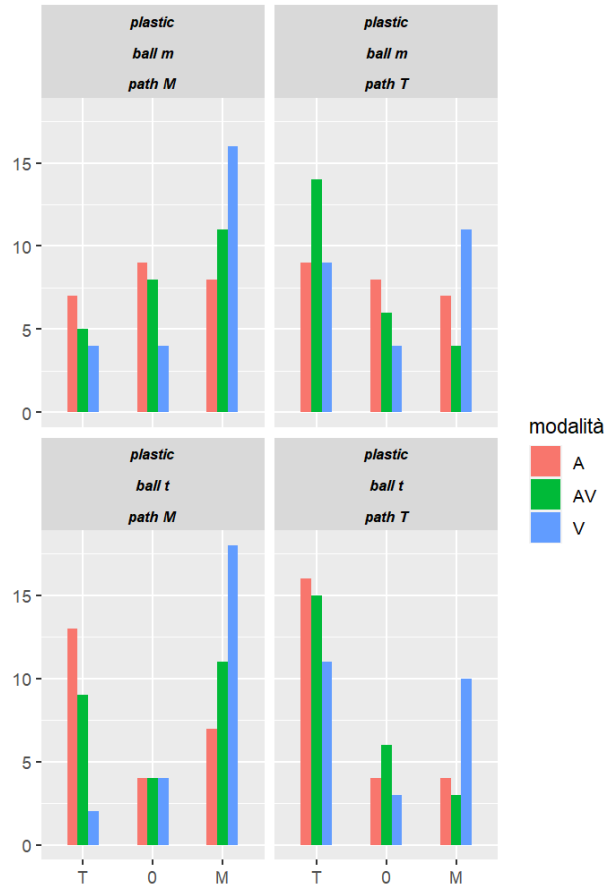


Figura 5. Risultati. Asse y = conteggio delle risposte (T: Takete; 0: incertezza, corrispondente a due associazioni discordanti per uno stesso stimolo da parte del partecipante; M: Maluma) per le tre modalità (A: audio; AV: audio-video; V: video) per ogni combinazione (**plastica**; palla M: palla liscia; palla T: palla deformata; percorso M: percorso curvo; percorso T: percorso a zig-zag).

Nelle combinazioni congruenti di percorso e palla (Mm e Tt) le risposte sono distorte rispettivamente su *Maluma* e *Takete* come previsto, e la condizione solo video sembra rafforzare il bias su *Maluma* mentre la condizione solo audio sembra rafforzare la distorsione su *Takete*. Le risposte *Takete* sono quindi generalmente favorite in condizioni audio, quelle *Maluma* invece in condizioni solo video, tanto che anche nella combinazione congruente di palla T e percorso T presentata in modalità solo video, la differenza tra la scelta di *Maluma* o *Takete* non sembra significativa. La modalità audio-video sembra influire significativamente sulla scelta di *Takete* solo nei due casi in cui il percorso è T.

Per i risultati nelle condizioni in cui il materiale della superficie è la plastica, possono essere fatte considerazioni simili con un'unica differenza: nel caso di percorso T e palla M la modalità audio sembra avere un ruolo meno significativo sulla scelta di *Takete*, a differenza della condizione audio-video che continua ad incidere fortemente sulla scelta di *Takete*.

Per l'analisi statistica è stato assunto un modello di regressione logistica; i coefficienti del modello sono stati stimati con metodi frequentisti utilizzando il programma R-Studio. Le variabili prese in esame sono: la variabile dipendente Y che assume valore 0 quando la risposta è *Takete* e valore 1 quando la risposta è *Maluma*; la variabile binaria "percorso" che assume valore M o T; la variabile binaria "palla" che assume valore m o t; la variabile binaria "materiale" che assume valore met o pla; la variabile categoriale "modalità" che assume valore A quando la modalità presentata è audio, AV quando la modalità presentata è audio-video e V quando la modalità presentata è video. All'interno del modello sono stati inizialmente inseriti tutti i possibili termini di interazione, successivamente eliminati in quanto risultati non significativi. La variabile risposta segue la distribuzione di Bernoulli, prendendo il valore 1 con una probabilità sconosciuta p e il valore 0 con probabilità 1-p:

$$f(k;p) = \begin{cases} p & if\ k = 1 \\ p - 1 & if\ k = 0 \end{cases} \quad (1)$$

La probabilità sconosciuta p del risultato *Maluma* viene predetta da un modello di regressione logistica dato da

$$log\frac{p}{1-p} = \beta_0 + \sum_{i=1}^{m} \beta_i \cdot x_i \quad (2)$$

dove m=4 é il numero di predittori, x_i sono i predittori (percorso, palla, materiale, modalità), β_0 è l'intercetta e i β_i sono i coefficienti di regressione stimati relativi alle rispettive variabili esplicative.

Per verificare un'eventuale influenza, nei risultati del modello, dell'associazione tra le risposte di uno stesso partecipante, è stato stimato un modello logistico normale con intercetta casuale su ogni intervistato. In questo secondo modello si considerano le variabili "percorso", "modalità", "materiale", "palla" come effetti fissi.
Sia Y_{ij} l'i-esima variabile risposta relativa al j-esimo soggetto intervistato dove $Y_{ij} = 1$ indica *Maluma*. Il modello assume

$$logit\left(Pr(Y_{ij} = 1|u_i)\right) = \alpha_j + \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \beta_4 x_{i4} + u, \quad (3)$$

con $u_i \sim N(0, \sigma^2_u)$, x_{i1}, x_{i2}, x_{i3}, x_{i4} le variabili indipendenti, i = 1,...,48 e j = 1,...,24. La varianza stimata degli effetti casuali è 0.002202 e quindi nel modello non sembra essere presente un'associazione tra le risposte relative allo stesso soggetto. Anche in questo caso, infatti, la variabile "percorso" e la variabile "modalità" risultano essere le uniche due variabili significative. Pertanto, viene utilizzato il primo modello in quanto meno complesso.

3.1 Interpretazione dei risultati

Le variabili significative risultano essere la modalità e il percorso. Il coefficiente relativo alla variabile "percorso" è negativo quindi la probabilità che la variabile risposta sia uguale a 1 (*Maluma*) è maggiore in corrispondenza del percorso M (questo perché la variabile "percorso" assume valore 1 in corrispondenza di T e valore 0 in corrispondenza di M). Per quanto riguarda la variabile "modalità" risultano significative solo la modalità audio e la modalità video: quando viene presentata al partecipante la modalità audio, la probabilità che quest'ultimo scelga *Takete* è maggiore, viceversa, quando viene presentata la modalità video, la scelta propende su *Maluma*. La modalità audio-video, invece, non sembra avere un effetto significativo sulla scelta.

4. DISCUSSIONE

La possibilità, contemplata nell'esperimento precedente a questo, che l'inclinazione sulle risposte *Takete* fosse dovuta, nella condizione audio-visiva, ad un effetto di *enhancement* multisensoriale [10], vista la costante congruenza tra informazioni uditive e informazioni visive presentate ai partecipanti, viene esclusa dai nuovi risultati: la modalità audio-video non è risultata significativa nella scelta di *Takete* o *Maluma* e il bias evidente verso *Takete* si registra in presenza dello stimolo solo uditivo, confermando che l'evento uditivo sembra essere percepito più come *Takete*.

Un recente studio [11] ha definito il rumore come un suono ad alto contenuto di ruvidità uditiva. Quest'ultima, infatti, ha mostrato un evidente collegamento con il design della forma: suoni più ruvidi sono stati associati a forme più appuntite. La loro tesi è che il rumore, definito dalla ruvidità uditiva, porti ad una mappatura audio-visiva tra quest'ultimo e la spigolosità nel dominio visivo e che questa proprietà psicoacustica funzioni indipendentemente da altri correlati noti dell'angolosità, come l'altezza del tono. Se i suoni presentati nel presente esperimento contenessero effettivamente un alto contenuto di ruvidità uditiva, queste ricerche potrebbero contribuire a spiegare perché l'evento uditivo in generale nelle nostre indagini è stato associato più a *Takete*[2], a prescindere dalla tipologia di percorso curvilineo o spigoloso (i percorsi "morbidi" sono stati comunque percepiti più come *Takete* nell'evento uditivo). Questa possibilità non viene approfondita nel

[2] In "Takete" sono presenti le consonanti stridenti /t/, /k/ mentre in "Maluma" le consonanti sonore /L/, /m/. «Le prime sono stabilite come consonanti di arresto in cui la voce (cioè la vibrazione delle corde vocali) segue il rilascio della consonante, creando un breve impulso di rumore turbolento aspirato prima dell'inizio della vibrazione stabile delle corde vocali per il suono vocale successivo." Le consonanti sonore sono consonanti continue "che vengono prodotte simultaneamente con la vibrazione delle corde vocali coinvolta nel suono vocale successivo, il che rende la componente consonante del suono relativamente simile a una struttura vocale.» (Cfr. Nielsen, A., Rendall, D., 2011[12]).

La distinzione affettivo-semantica tra suoni percepiti come rumorosi e stridenti o come relativamente melliflui e sonori nel linguaggio, e la conseguente associazione a diverse tipologie di forme visive, potrebbe essere trasposta su un piano in cui suoni di eventi naturali (come il rotolamento di una palla) percepiti come rumorosi vengano associati alla parola *Takete* piuttosto che a *Maluma*.

presente studio e resta un eventuale spunto di riflessione per ricerche future.

4.1 La forte influenza dell'informazione uditiva nell'esperimento

La generale tendenza a esprimere giudizi corrispondenti a *Maluma* in condizioni solo video è stata riscontrata sia nel precedente esperimento che nelle indagini del presente studio. È noto che gli esseri umani mostrano preferenza visiva per gli oggetti curvi [13]; in questo esperimento i percorsi, che insieme alla modalità sono risultati essere le variabili significative, sono invisibili: non appare la figura curva, ma solo il processo del movimento della palla lungo la traiettoria. Nonostante ciò, i risultati precedenti sono stati confermati. Per questo motivo viene esclusa l'ipotesi della precedente ricerca che parte dell'inclinazione verso *Maluma* nei video silenziosi sia dovuta a un'attenzione maggiore verso le componenti curve presentate visivamente. Nell' interpretazione del bias verso *Maluma* in condizioni di solo video, non possiamo escludere che un ruolo sia stato giocato da un principio di *good conrinuation* nell' interpolazione percettiva visuale tra un frame e il successivo; ciò potrebbe rappresentare un ulteriore spunto di ricerca per le fasi future di questo studio.

È stata confermata la forte influenza del canale uditivo sulla percezione dell'evento come *Takete*: il suono tra le due modalità sembra prendere il sopravvento sulla percezione dell'evento e la conseguente decisione associativa; si osserva infatti che nella modalità solo visiva con percorso T, la scelta propende su *Maluma*, mentre con l'introduzione della modalità audio la scelta propende su *Takete*.

Si prenda ad esempio il caso del metallo, nei due casi dei percorsi curvo e angoloso: nel caso in cui i percorsi siano curvi, la forte pendenza su *Maluma* trovata nella condizione video, in situazione audio video sembra perdersi come se l'informazione visiva perdesse di importanza al comparire dell'audio; nel caso dei percorsi angolosi, il bias in modalità solo audio è schiacciante su *Takete*, si verifica un bias forte su *Takete* anche in condizione audio-video, mentre nella condizione solo video c'è incertezza.

Supponendo, come nello studio precedente, che l'effetto riscontrato nella condizione audio-visiva dipenda dall'influenza uditiva, resta la domanda di come sia riuscito il canale uditivo ad assumere un ruolo equipotente a quello del canale visivo. Alcuni studi hanno portato prove di una dominanza uditiva in particolari circostanze: nell'elaborazione delle informazioni temporali [14] e in situazioni in cui il canale uditivo è portatore di una quantità maggiore di informazioni, come accade per la musica. I risultati dello studio di Shams et al. [15] sull'illusione visiva indotta dal suono, mostrano che, sebbene dipenda dal tipo di stimolo, la percezione visiva può qualitativamente essere alterata dal suono anche quando lo stimolo visivo non è ambiguo. La ricerca sulla dominanza sensoriale, sebbene ancora incerta, ha riscontrato l'evidenza di un dominio visivo generale: il nostro cervello contemporaneamente inondato da stimoli provenienti da più modalità sensoriali, non attribuisce lo stesso peso a tutte le modalità e sembra che le informazioni visive ricevano più frequentemente un'elaborazione preferenziale [16]. È stata dimostrata l'evidenza di un dominio visivo generale: un esempio della supposta dominanza visiva sugli stimoli uditivi è l'effetto Colavita, che consiste nel fatto che, quando si deve rispondere contemporaneamente ad un segnale uditivo e ad un segnale visivo, i partecipanti tendono a rispondere solo al segnale visivo, in sostegno della tesi del predominio visivo nei processi attenzionali [17]. Il fenomeno si verifica sia con stimoli semplici come lampi di luce, o toni uditivi, sia con stimoli che sono più complessi, come immagini e suoni che si riferiscono a oggetti realistici, sia con stimoli relativi al linguaggio visivi e uditivi [18]. Assodato che l'attenzione selettiva ad una determinata modalità sensoriale possa modulare, o in alcuni casi eliminare, la dominanza visiva [19], è stato dimostrato anche che le manipolazioni dell'attenzione (esogene o endogene) verso una o l'altra modalità non siano sufficienti per invertire l'effetto Colavita [20]. L'attenzione potrebbe dunque aver avuto un ruolo nella grande influenza che la modalità uditiva ha dimostrato ma non sono chiari gli elementi che potrebbero essere stati decisivi per un eventuale spostamento dell'attenzione al canale uditivo e di conseguenza per la modulazione o eliminazione della dominanza visiva.

5. CONCLUSIONI

In conclusione, è possibile affermare che l'informazione uditiva nella nostra indagine ha avuto un ruolo importante nelle percezioni dei partecipanti per prendere decisioni non casuali nell'associazione tra percezioni uditive e verbalizzazioni, consolidando e provando così l'ipotesi avanzata dallo studio precedente a questo. È prevista una futura analisi dei suoni utilizzati nei nostri test che miri a chiarire quali elementi di questi suoni abbiano causato specifici segnali associativi, e cosa abbia contribuito a questa apparente forza del canale uditivo nella competizione multisensoriale.

6. RIFERIMENTI BIBLIOGRAFICI

[1] Köhler, W. (1929). *Gestalt psychology*. New York: Liveright.

[2] Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Liveright.

[3] Shen, Y.-C., Chen, Y.-C., & Huang, P.-C. (2022). Seeing Sounds: The Role of Vowels and Consonants in Crossmodal Correspondences. *I-Perception*. *https://doi*. org/10.1177/20416695221084724

[4] Walker P. (2012). Cross-sensory correspondences and cross talk between dimensions of connotative meaning: visual angularity is hard, high-pitched, and bright. *Attention, perception & psychophysics*, *74*(8), 1792–1809. https://doi.org/10.3758/s13414-012-0341-9

[5] Ramachandran, V. S., Hubbard, E. M. (2001). Synaesthesia - A window into perception, thought and language. *Journal of Consciousness Studies*, 8(12), 3-34. https://doi.org/10.1111/1468-0068.00363

[6] Karthikeyan, S., Rammairone, B., & Ramachandra, V. (2016). The Bouba-Kiki Phenomenon Tested via Schematic Drawings of Facial Expressions: Further Validation of the Internal Simulation Hypothesis. *i-Perception*, *7*(1), 2041669516631877. https://doi.org/10.1177/2041669516631877

[7] Fontana, F. (2013). Association of Haptic Trajectories to Takete and Maluma. *Haptic and Audio Interaction Design*, 60-68. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41068-0_7

[8] Fontana, F., Järveläinen, H. & Favaro, M. (2021). Is an Auditory Event more Takete?. *18th Sound and Music Computing Conference*. https://doi.org/10.5281/zenodo.5038640

[9] Mehta, R., & Zhu, R. J. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science (New York, N.Y.)*, *323*(5918), 1226–1229. https://doi.org/10.1126/science.1169144

[10] Stein, B. E., London, N., Wilkinson, L. K. & Price, D. D. (1996). Enhancement of perceived visual intensity by auditory stimuli: a psychophysical analysis. J. Cogn. Neurosci., 8(6), 497–506. https://doi.org/10.1162/jocn.1996.8.6.497

[11] Liew, K., Lindborg, P., Rodrigues, R., & Styles, S. J. (2018). Cross-modal perception of noise-in-music: Audiences generate spiky shapes in response to auditory roughness in a novel electroacoustic concert setting. *Frontiers in Psychology*, 9(178). https://doi.org/10.3389/fpsyg.2018.00178

[12] Nielsen, A., Rendall, D. (2011). The sound of round: Evaluating the sound-symbolic role of consonants in the classic Takete-Maluma phenomenon. *Canadian Journal of Experimental Psychology/Revue Canadienne De Psychologie Expérimentale*, 65(2), 115-124. https://doi.org/10.1037/a0022268

[13] Munar, E., Gomez-Puerto, G., Call, J., & Nadal, M. (2015). Common Visual Preference for Curved Contours in Humans and Great Apes. *PLoS One*, 10(11). https://doi.org/10.1371/journal.pone.0141106

[14] Repp, B. H., & Penel, A., (2002). Auditory dominance in temporal processing: New evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 28*(5),* 1085–1099. http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.28.5.1085

[15] Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions: What you see is what you hear. *Nature, 408*, 788-788. https://doi.org/10.1038/35048669

[16] Huang, S., Li, Y., Zhang, W., Zhang, B., Liu, X., Mo, L., & Chen, Q. (2015). Multisensory Competition Is Modulated by Sensory Pathway Interactions with Fronto-Sensorimotor and Default Mode Network Regions. *Journal of Neuroscience,* 35(24), 9064-9077. https://www.jneurosci.org/content/35/24/9064

[17] Colavita, F. B. (1974) "Human sensory dominance," Percept. Psychophys., 16(2), 409–412. http://link.springer.com/10.3758/BF03203962

[18] Spence, C., Parise, C., & Chen, Y.-C. (2012). The Colavita Visual Dominance Effect. *The Neural Bases of Multisensory Processes*, 27. http://www.ncbi.nlm.nih.gov/pubmed/22593876

[19] Sinnett, S., Spence, C. & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception & Psychophysics* 69**,** 673–686. https://doi.org/10.3758/BF03193770

[20] Koppen, C., Spence, C. (2007). Seeing the light: Exploring the Colavita visual dominance effect. *Experimental Brain Research*, 180, 737–754. https://doi.org/10.1007/s00221-007-0894-3

OPENMUSIC E MODALYS PER NICOLA BARONI

Fabio De Sanctis De Benedictis
ISSM "P. Mascagni"
fabio.desanctis@consli.it

ABSTRACT

La libreria Modalys per OpenMusic mostra notevoli punti di interesse, ma non appare sufficientemente documentata per gli utenti allo stadio iniziale. Dopo alcune considerazioni preliminari in questo articolo si prende in esame un caso specifico, la creazione di materiale audio per *Reversio*, per violoncello e *live electronics*, composto per Nicola Baroni, procedendo dai problemi affrontati nell'apprendimento dell'implementazione di Modalys in OpenMusic, sino alla presentazione di due strumenti tra quelli costruiti per la composizione in oggetto. Si conclude con brevi considerazioni su lavori e implementazioni future.

1. INTRODUZIONE

La richiesta di Nicola Baroni di scrivergli un brano per violoncello ed elettronica, *Reversio*, terminato di comporre nel 2020, è stata l'occasione per approfondire lo studio di due tecniche: il riconoscimento di suoni da parte dell'elaboratore elettronico e l'uso di Modalys all'interno di OpenMusic. Nel corrente scritto ci occupiamo di quest'ultimo argomento. Dopo alcune considerazioni preliminari saranno descritti il percorso di approccio e di approfondimento di Modalys nella sua implementazione in OpenMusic, e poi due strumenti creati ed utilizzati per la produzione di materiale audio, uno in versione semplice e l'altro in versione ibrida.

2. CONSIDERAZIONI PRELIMINARI

Propedeutico è stato l'incontro con l'interprete. Si è presa visione delle sue pratiche di programmazione e di improvvisazione/composizione mediante l'utilizzo del riconoscimento interattivo di materiale sonoro del violoncello, e conseguente reazione da parte del *software*. A ciò ha fatto seguito lo studio della tesi di dottorato di Nicola Baroni e di una sua pubblicazione, [1] nonché delle *patch* di Max/MSP delle composizioni descritte nella tesi.

[1] La tesi di dottorato [1] e il materiale informatico ad essa collegato sono scaricabili da questa pagina: `https://era.ed.ac.uk/handle/1842/25498`, mentre la pubblicazione citata è [2]. Tutti i collegamenti Internet di questo articolo sono stati verificati in data 20 maggio 2022.

Copyright: ©2022 Fabio De Sanctis De Benedictis. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Durata (sec.)	150	30	90	60	60	90	30	150
Sequenza determinata	X		X		X		X	
Sequenza indeterminata		X		X		X		X
Matrice timbrica	1		2		3		4	

Table 1. La forma di *Reversio*.

Data la destinazione del brano e la sua occasione, nella pianificazione della composizione sono confluiti alcuni elementi quali l'uso di determinate figure, principalmente timbriche, anche come fonti di attivazione di eventi, l'uso di schemi paradigmatici per la distribuzione di tali figure, e una concezione legata al determinato e all'indeterminato, che ha portato a stabilire la forma descritta nella Tabella 1.

Le sequenze determinate partono da una situazione più timbrica per terminare con una prevalenza del suono ordinario, le sequenze indeterminate procedono esattamente al contrario, da cui il titolo dell'opera, che ci riconduce all'analoga figura retorica. Nelle sequenze indeterminate avviene la captazione e il riconoscimento degli eventi acustici del violoncello da parte di Max/MSP, utilizzando l'oggetto `gf` della libreria MuBu dell'Ircam e la libreria zsa.descriptors. [2]

Le durate e i transitori di attacco di ogni sezione o sottosezione sono stati per lo più determinati da una semplice *patch* di OpenMusic (vedi Figura 1), attraverso l'uso di `num2rtms-size`, una delle funzioni della mia libreria per OpenMusic o PWGL. [3]

Seguendo una prassi consueta per le mie composizioni, la parte elettronica si è avvalsa soprattutto di suoni derivati dal violoncello, o sintetizzati su suo modello, in questo caso utilizzando il software Modalys dell'Ircam, [4] nella sua implementazione all'interno di OpenMusic.

3. L'APPROCCIO A MODALYS

L'apprendimento del *software* di sintesi per modelli fisici è stato inizialmente rallentato da una documentazione non completamente esaustiva per quanto riguarda l'uso all'inter-

[2] Si veda `https://forum.ircam.fr/projects/detail/mubu/` per la libreria MuBu e `http://www.e--j.com/` per la libreria zsa.descriptors.

[3] Su questa libreria rimandiamo a quanto descritto in [3].

[4] `https://forum.ircam.fr/projects/detail/modalys/`. Modalys è un *software* per la sintesi secondo modelli fisici, più esattamente, come specificato nella documentazione, sintesi modale.

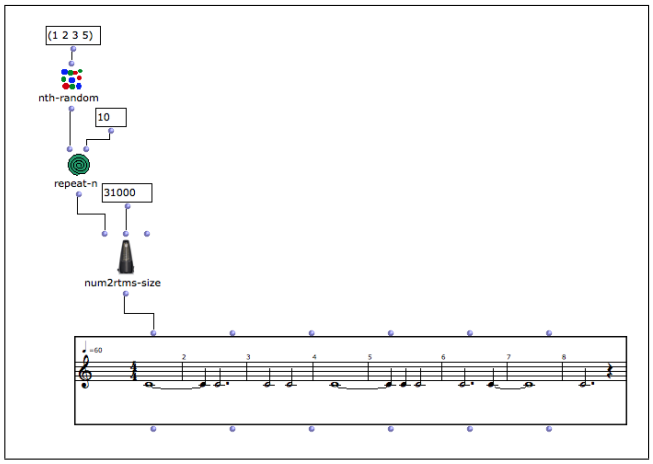


Figure 1. La *patch* di OpenMusic usata per articolare durate e punti di attacco delle sezioni di *Reversio*.

no di OpenMusic: i *tutorial* della libreria Modalys per OpenMusic mostrano tre cartelle: "classical examples", "finite elements" e "mauro lanza's examples". Le *patch* delle prime due cartelle appaiono non commentate e l'"example 8" utilizza una sotto-*patch* in realtà vuota, per cui non funziona. Gli esempi di Mauro Lanza sono maggiormente commentati, e, a parte i primi due, funzionano senza errori.[5] Tuttavia i modelli proposti sono di livello già abbastanza avanzato e non accompagnano l'utente alle prime armi sin dai primi passi. L'ottima documentazione *online* ha poi come oggetto il linguaggio di programmazione, ma non la sua implementazione nella libreria per OpenMusic. Lo stesso vale per alcuni tutorial su YouTube.[6] Per farla breve, si è proceduto quindi attraverso vari stadi:

- la completa traduzione in italiano della documentazione *online*
- lo studio delle funzioni che interessavano la composizione di *Reversio*
- l'iniziale traduzione di alcuni esempi redatti in linguaggio Lisp, acclusi alla versione a linea di comando di Modalys, in *patch* per OpenMusic

Per chiarire meglio, in Figura 2 si riporta la traslazione in OpenMusic del primo esempio in codice Lisp, con in rosso le linee del codice corrispondenti agli oggetti della libreria per OpenMusic. Il risultato finale è quello di una corda pizzicata. Il codice originario è contenuto nel file "ex1 - plucked string.lisp" all'interno della cartella "/Applications/Modalys/Examples/Lisp" della versione per computer Apple, unica piattaforma supportata attualmente.

Procedendo in questa maniera è stato possibile comprendere meglio l'utilizzo delle funzioni della libreria di Modalys per OpenMusic e iniziare a configurare strumenti propri. Sono state consultate anche altre fonti, stimolanti sotto molti aspetti, ma non del tutto illuminanti sul lato pratico-applicativo, vuoi per inadeguatezza dello scrivente su alcuni degli argomenti trattati, vuoi per gli argomenti stessi, focalizzati su aspetti differenti da quelli ricercati.[7]

4. GLI STRUMENTI UTILIZZATI

Gli strumenti utilizzati sono stati: una corda o una barra di metallo sollecitate da un arco; una membrana circolare, percossa oppure pizzicata; un piatto circolare non vincolato, anch'esso percosso oppure pizzicato. La creazione di strumenti ibridi – semplificando possiamo descriverli come la miscelazione di due strumenti in uno – duplicando ogni tipo di strumento ha permesso l'esecuzione di glissando, passando dinamicamente senza soluzione di continuità da uno strumento all'altro, intonati con altezze differenti.[8]

In Figura 3 la *patch* principale di utilizzo dello strumento violoncello, ossia una corda sollecitata da un arco. Lo strumento è racchiuso all'interno del `loop` denominato '`vc`' e sarà esaminato in seguito. La *patch* è divisa in due parti, una per la generazione di un file audio mono, sulla sinistra, l'altra per la generazione di un file audio quadrifonico, sulla destra. Il materiale musicale di ingresso consiste in una `chord-seq`, ossia una sequenza di note. In una versione alternativa di questa *patch* è possibile renderizzare singoli accordi. Il `loop` denominato '`vc`' contiene il motore di sintesi che utilizza appunto Modalys, ossia uno strumento relativo ad una corda sfregata con un arco; il `loop` denominato '`delay`' ritarda ogni suono generato da '`vc`' di una quantità temporale corrispondente agli *onset* della `chord-seq`, di modo che ogni suono sia posizionato nel suo corretto punto di attacco; infine la sotto-*patch* denominata '`mix`' miscela il tutto in un unico file monofonico. I parametri di '`vc`' sono quelli indicati nel commento in figura; la sotto-*patch* 'varie impostazioni per const' permette di scegliere, attraverso un `pop-up-menu`, un parametro collegato a coefficienti che simulano l'effetto della pece sull'arco, qui configurati secondo le scelte 'a piacere', ossia assolutamente casuali all'interno di gamme di valori predefiniti dall'utente, 'nth-random', ovvero una selezione all'interno di una lista di possibilità impostate dall'utente, 'suoni ordinari' (come in 'nth-random', ma con diverse impostazioni di selezione delle soglie minima e massima) e '(2 10 5 4)', impostazione predefinita; la sotto-*patch* 'vari inviluppi' permette, aprendola, di scegliere un inviluppo tra diversi, relativi alla velocità dell'arco. Quindi, in sintesi, il `loop` '`vc`' legge la `chord-seq` e renderizza, secondo i parametri assegnatigli, un insieme di suoni inviati a '`delay`' che li riposiziona sul loro punto di attacco, per avere poi la miscelazione finale in '`mix`'.

Nella seconda parte della *patch*, quella di destra, i suoni generati da '`vc`' sono inviati a `vbap.discrete`, della libreria OMPrisma, versione 2.4.0, che si occupa di collocarli in uno spazio quadrifonico. La disposizione può

[5] Il tutto è stato testato su un computer MacBook Pro mid 2012, processore 2,5 GHz Intel Core i5, 8 GB di Ram DDR3 a 1600 MHz, OsX El Capitan versione 10.11.6. Le versioni utilizzate dei programmi sono state le seguenti: Modalys 3.5.1.rc4 e OpenMusic 6.16, che erano le più recenti al momento della composizione del brano.

[6] La documentazione di Modalys è accessibile al seguente indirizzo: `http://support.ircam.fr/docs/Modalys/3.4.0/co/publication-web.html`. I *tutorial* su YouTube si trovano qui: `https://www.youtube.com/playlist?list=PLqxIlrZFcLJG_GXVir6v1We-M40g2NMlw`.

[7] Si vedano in particolare: [4]; [5]; [6]; [7]; [8]; [9].

[8] Ci siamo ispirati a quanto riportato nella documentazione *online* in questa pagina: `http://support.ircam.fr/docs/Modalys/3.4.0/co/object_hybrid_melt.html`.

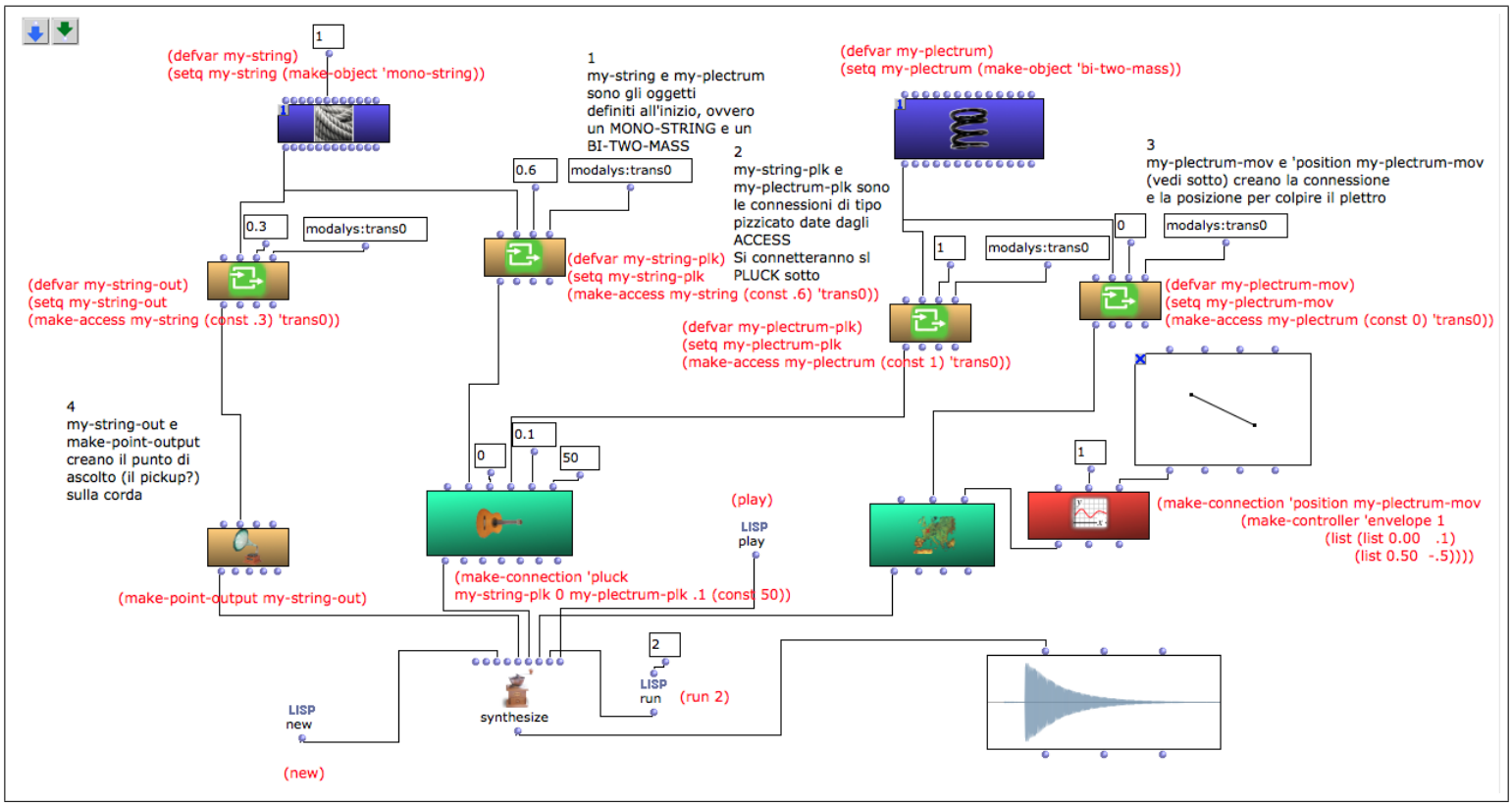


Figure 2. Il primo esempio di codice letto nel programma ModaLisp, versione a linea di comando di Modalys, tradotto in una *patch* di OpenMusic.

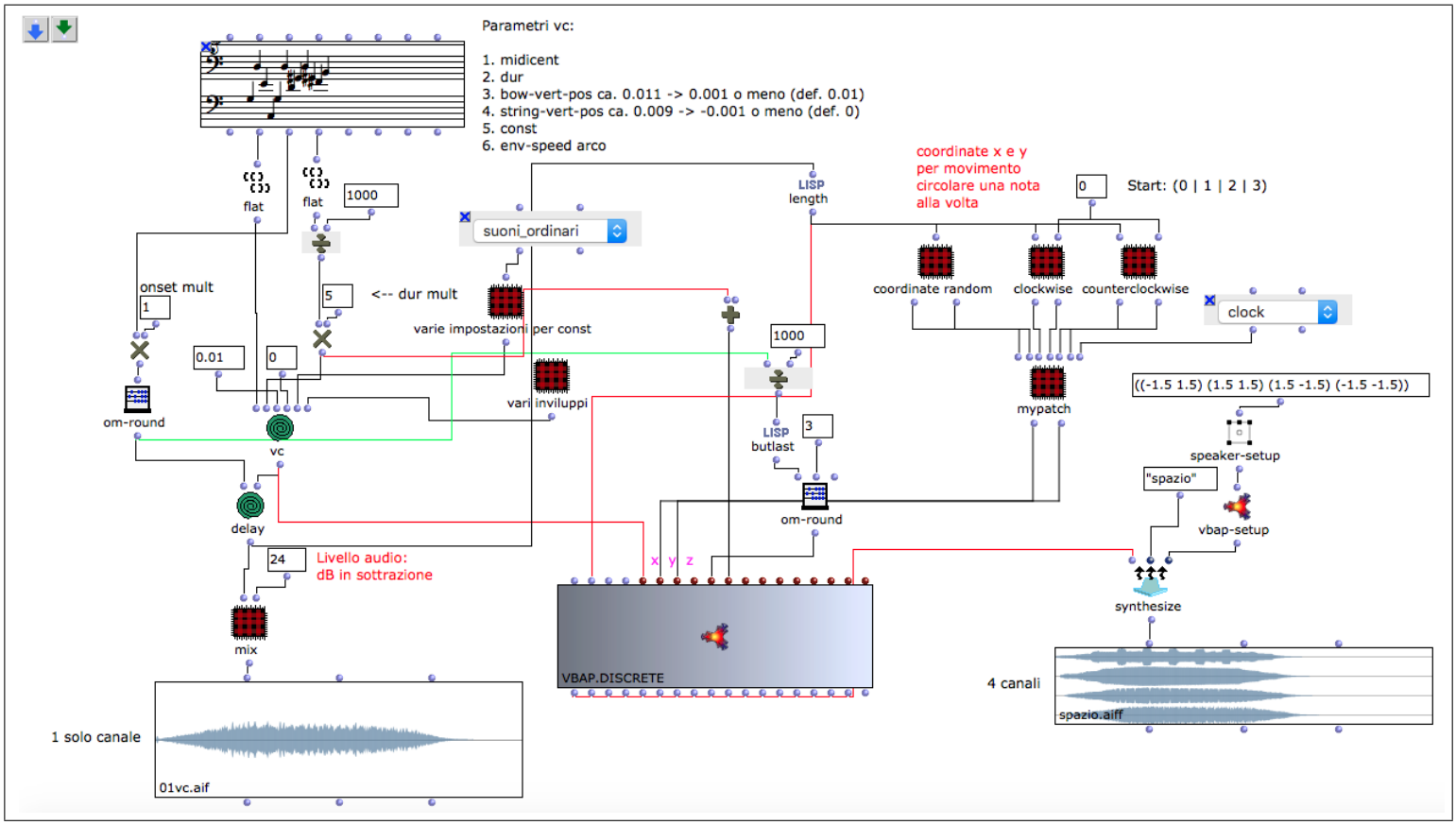


Figure 3. La *patch* principale di utilizzo dello strumento violoncello.

avvenire in maniera casuale oppure in senso progressivamente orario o antiorario, attraverso gli algoritmi contenuti nelle sotto-*patch* in alto a destra, selezionabili mediante il `pop-up-menu` con le voci 'rnd', 'clock' e 'counterclock', ovvero casualmente, in senso orario o antiorario. Valutando quindi l'oggetto `sound` corrispondente al risultato desiderato si ottiene il relativo file audio mono o quadrifonico. [9]

In Figura 4 si riporta l'interno del loop '`vc`', contenente appunto lo strumento Modalys.

In alto a sinistra i primi due ingressi del loop immettono le altezze e le durate della `chord-seq`. Gli altri ingressi afferiscono ai parametri dello strumento. Le varie `bpf` sono relative al tipo di movimento (verticale) di contatto dell'arco con la corda e sono riscalate corrispondentemente alle durate delle singole altezze. È come se, seguendo il profilo della `bpf`, l'arco calasse sulla corda e rimanesse aderente per la durata della nota, sino a quando il profilo della `bpf` non si 'alza' nuovamente.

In alto a destra, in colore blu, abbiamo gli oggetti `bi-two-mass` e `bi-string`. Il primo corrisponde a un modello bi-direzionale consistente di due masse collegate da una molla, con movimento possibile lungo due direzioni ortogonali nel piano. La massa minore è connessa all'oggetto eccitato, la grande viene manipolata dall'utente. `bi-string` è una corda vincolata agli estremi, che può vibrare orizzontalmente e verticalmente. I parametri permettono di specificare le caratteristiche fisiche degli oggetti. Ad esempio i parametri di `bi-string` sono stati cambiati nella versione corrispondente alla barra di metallo sollecitata dall'arco, per ottenere il materiale differente. I valori occorrenti sono stati dedotti partendo da quelli riportati nella manualistica del *software*, relativi a varie sostanze, e variati poi sperimentalmente.

Gli oggetti denominati `access` pongono in collegamento le vari componenti dello strumento. Così `bi-two-mass` è collegato a `position-connection`, `speed` e `bow`. Quest'ultimo corrisponde all'arco, `speed` alla sua velocità e `position-connection` dovrebbe corrispondere al punto di contatto sulla corda.
`bi-string` è collegato sia a `bow` che a `point-output`, ossia il punto in cui viene captato il suono sulla corda, come se avessimo un *pick-up*. Quindi il *loop* genera tanti suoni quante sono le altezze della `chord-seq`, i quali saranno poi riposizionati nella *patch* principale da '`delay`' e miscelati in un unico file audio da '`mix`' per la versione mono, oppure inviati a `vbap.discrete` per la versione quadrifonica.

In Figura 5 è possibile osservare una *patch* relativa ad uno strumento ibrido, variante del precedente, composto da due corde differenti suonate con l'arco. Il `pop-up-menu` permette di scegliere il tipo di miscelazione dei due strumenti, ossia attraverso un inviluppo oppure in base ad un rapporto costante. È una variante della *patch* di Figura 3: la differenza consiste nel fatto che nel *loop*, all'interno di una sotto-*patch*, sono contenuti due oggetti `bi-string` (vedi Figura 6). La seconda `bi-string`, posta in alto più a destra, riceve come frequenza un valore impostato da un fattore di trasposizione gestito nella *patch* principale. Il funzionamento complessivo è identico a quello dello strumento della *patch* di Figura 4, con alcune differenze, oltre a quella già rilevata di intonazione:

- le due `bi-string` si connettono a `melt-hybrid`, che crea materialmente l'ibridazione dei due oggetti, in questo caso identici se non per la diversa intonazione
- ognuno dei tre oggetti `access` di ogni `bi-string` si connette ad uno dei due ingressi dei tre oggetti `hybrid-access`, sorta di 'miscelatori di accessi', per così dire, i quali seguono poi gli stessi percorsi degli oggetti `access` nella versione a strumento singolo, ossia si collegano ai due ingressi di `bow` e a `point-output`

In uscita abbiamo quindi un suono generato dalla miscelazione delle due `bi-string` differenti, sollecitate da un unico arco e captate in un unico `point-output`. [10] Modalys permette l'ibridazione anche di tre strumenti differenti, ma in questo lavoro ci siamo limitati a due soli.

5. CONCLUSIONI E POSSIBILI DIREZIONI FUTURE

L'utilizzo di Modalys all'interno di OpenMusic ha permesso la creazione di *texture* sonore particolarmente dense e complesse, potendo sfruttare tutte le potenzialità algoritmiche offerte da OpenMusic per la creazione di `chord-seq` molto articolate. Allo stesso tempo Modalys ha permesso di mantenere una nitidezza timbrica, un controllo e una definizione non sempre raggiungibili con altri sistemi dediti alla sintesi per modelli fisici.

Per quanto riguarda le possibili direzioni future, personalmente restano senz'altro da comprendere e utilizzare al meglio altre funzioni, nonché approfondire lo studio della documentazione e degli esempi acclusi. Un campo di indagine promettente appare l'uso degli strumenti ibridi, per la creazione di gesti timbrici più dinamici e complessi, particolarmente efficaci, a giudizio dello scrivente, nell'esplorazione delle zone di confine tra suono ordinario e non ordinario. Il tutto, trasferibile su Max/MSP attraverso l'implementazione di Modalys in quell'ambiente, apre possibilità di interazione in tempo reale, cosa questa non ancora matura in OpenMusic, nato e utilizzato per altri scopi. Al momento in cui scriviamo Modalys non appare implementato in om#, il nuovo arrivato della famiglia OpenMusic, esito di OM7. [11] Se ciò accadrà si potranno aprire ulteriori direzioni di indagine e composizione verso l'utilizzo in tempo reale.

[9] Si ricorda che la valutazione avviene selezionando l'oggetto e premendo il tasto 'v' sulla tastiera alfanumerica dell'elaboratore elettronico.

[10] Modalys permette l'uso di più `point-output`, così come la moltiplicazione di altre caratteristiche degli strumenti costruiti, strategia qui non adottata.

[11] Si veda `https://cac-t-u-s.github.io/om-sharp/`

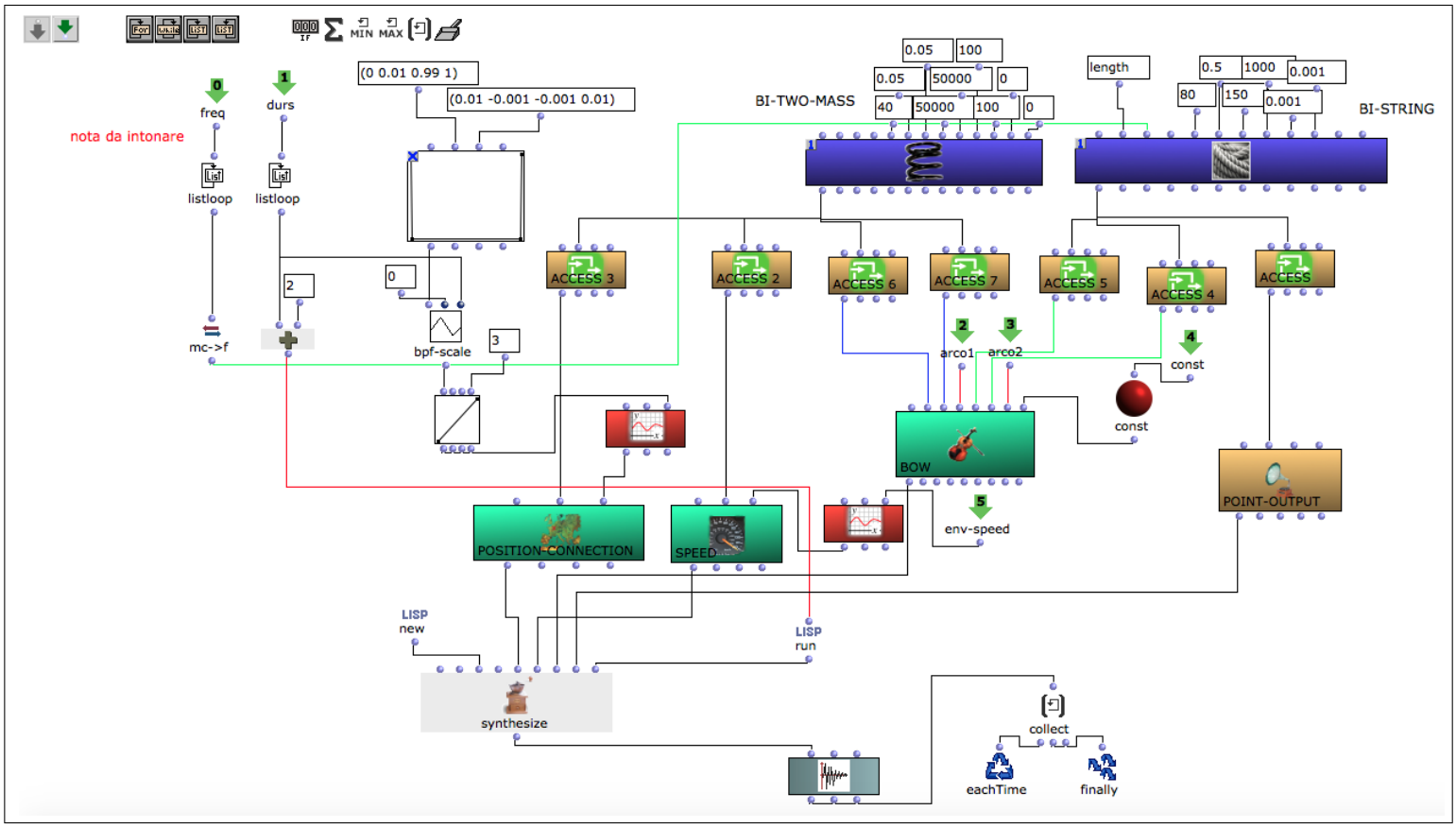


Figure 4. Il motore di sintesi audio: una corda sfregata da un arco ricostruita attraverso le funzioni di Modalys.

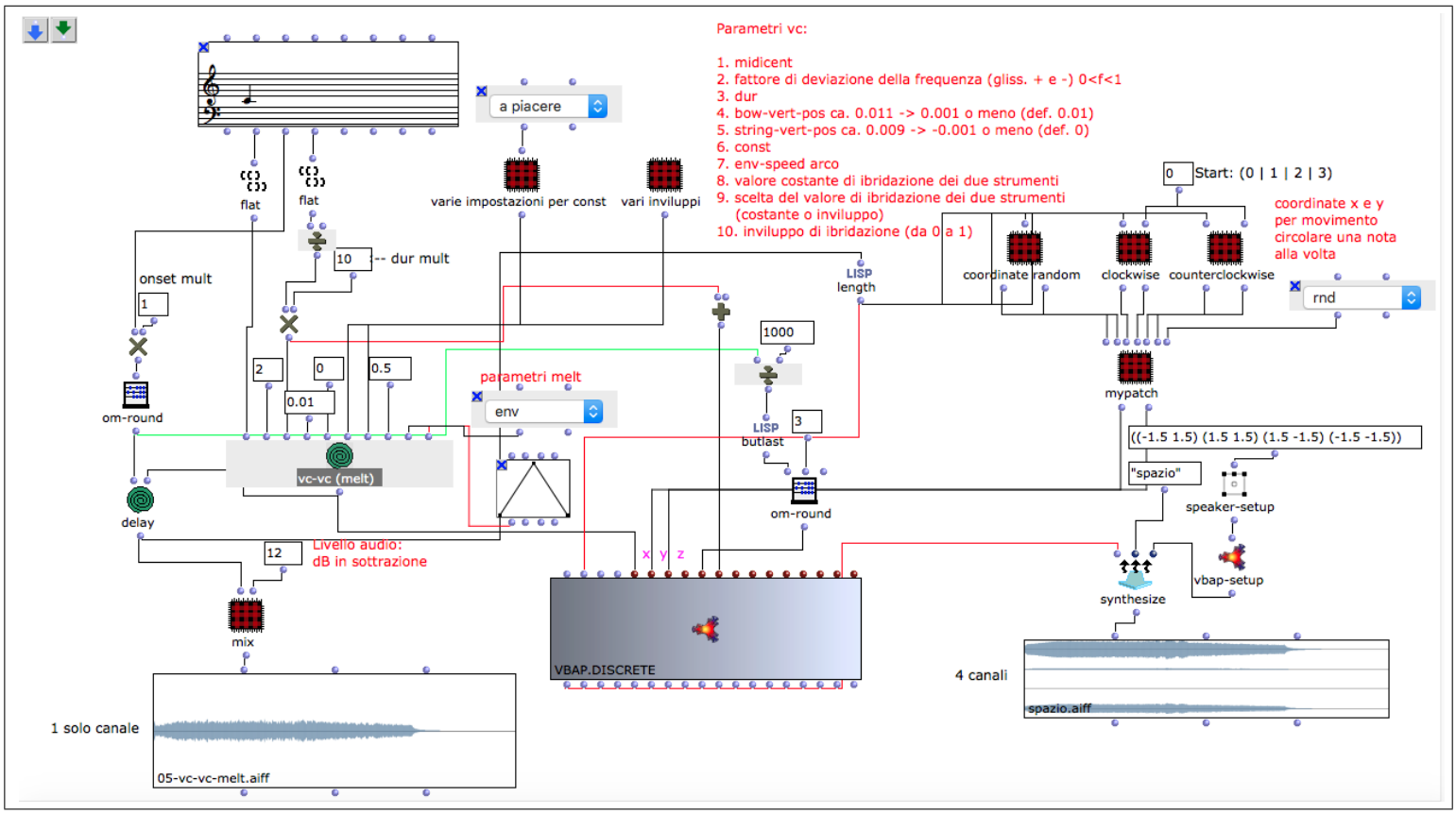


Figure 5. *Patch* relativa ad uno strumento ibrido composto da due corde differenti suonate con l'arco.

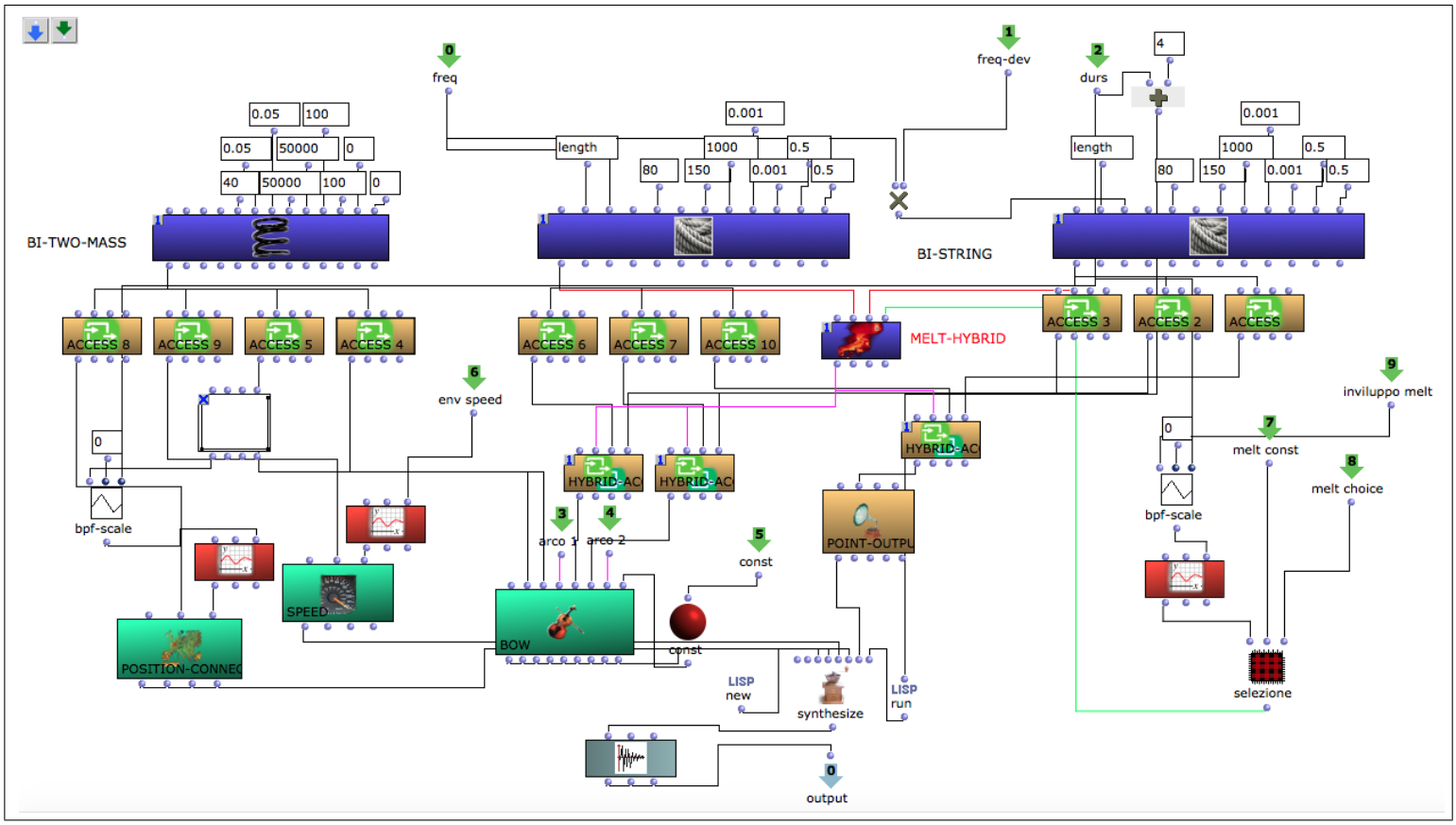


Figure 6. Lo strumento ibrido.

6. REFERENCES

[1] N. Baroni, *Gesture and listening: towards a social and eco-systemic hyperinstrument composition*. University of Edinburgh., 2015.

[2] N. Baroni and C. Benzi, "Sound Gesture and Rhetoric. Hyper-cello as an Algorithmic Composer," in *Proceedings of the Electroacoustic Music Studies Network Conference Electroacoustic Music Beyond Performance, Berlin*, June 2014.

[3] F. De Sanctis De Benedictis, "FDSDB_XXth_CT: una libreria di composizione algoritmica per Open Music e PWGL dedicata alle tecniche compositive del Novecento," in *Machine Sounds, Sound Machines. Atti del XXII CIM – Colloquio di Informatica Musicale*, 2019.

[4] P. Livorsi, "Some Applications of Open Music in Connection with the Program Modalys," in *The OM Composer's Book. Volume One*, 2006.

[5] *Ircam documentation. Modalys. Introduction, Third Edition*. Ircam-Centre Pompidou, 1997.

[6] *Ircam documentation. Modalys. Reference, Third Edition*. Ircam-Centre Pompidou, 1996.

[7] *Ircam documentation. Modalys. Tutorial, Third Edition*. Ircam-Centre Pompidou, 1997.

[8] J. B. N. Ellis and R. Caussé, "Modalys Demonstration," in *Proceedings of InternationalComputer Music Conference*, 2005.

[9] H. P. S. Tegljærg and T. Goepfer, "Virtual Bass-Clarinet in Modalys," in *Proceedings of the 2010 International Computer Music Conference, ICMC*, 2010.

SAMPLE: a Python Package for the Spectral Analysis of Modal Sounds

Marco Tiraboschi Federico Avanzini
Laboratory of Music Informatics (LIM)
Department of Computer Science
University of Milan, Italy
marco.tiraboschi,federico.avanzini@unimi.it

ABSTRACT

We present SAMPLE, a Python package of tools for spectral analysis and modal parameters estimate. The core of the package is an implementation of the "Spectral Analysis for Modal Parameters Linear Estimate" (SAMPLE) algorithm. This includes a custom implementation of a Sinusoidal Analysis algorithm based on Spectral Modelling Synthesis. Our custom implementation is specifically designed for modal tracking. We also included utilities for automatically tuning the algorithm parameters, using a Bayesian optimization method based on Gaussian Processes. For this purpose, we implemented efficient routines for computing perceptual audio representations for loss functions, such as the multiscale-spectrogram, the mel-spectrogram and the cochleagram. The package also comes with a Graphical User Interface, which allows to load and trim audio inputs, set the algorithm parameters, run the algorithm, listen to a resynthesis of the input, and export the results. The GUI is distributed both as an extra for the Python package and as a standalone executable.

1. INTRODUCTION

Modal synthesis [1] is an approach to sound synthesis based on physical or pseudo-physical models. It is used to generate sounds as results of interactions between objects, such as impacts and frictions [2], and it is becoming more and more relevant with the recent advancements in audio for videogames and extended-reality.

In many cases, technical issues stand between the sound designer and physical models. In modal synthesis, many are the parameters that the sound designer must finely tune: they have to decide dozens of modal frequencies, amplitudes, and decay times.

We introduce the SAMPLE Python package. It includes the implementation of several algorithms aimed at the automatic estimate of modal parameters from audio examples, including SAMPLE [3] and BeatsDROP.

It is available via the Python Package Index, on GitHub and on Zenodo [4]. In can be installed as a Python package, or used via the GUI as a standalone executable.

Copyright: ©2022 Marco Tiraboschi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

The package is thoroughly tested with unit tests using the Python builtin module *unittest* [5]. Code coverage is 100% on all the code base, excluding the GUI module.

2. SAMPLE

Our main goal for this package was to implement the "Spectral Analysis for Modal Parameters Linear Estimate" (SAMPLE) algorithm [3] in such a way that it would be easily available, distributable, and usable.

Most components of the package gravitate around the *SAMPLE* class. We designed this class following the patterns of the *scikit-learn* API [6, 7]. Every SAMPLE object includes a sinusoidal model and a regressor, that can be either a linear regressor or a semi-linear regressor.

Instances of the SAMPLE class can be fitted to audio examples to estimate the modal amplitudes, frequencies and decay times of the recorded objects. The input audio should be a sound generated hitting the object.

2.1 Sinusoidal Model

The *SinusoidalModel* class implements the analysis algorithm of *Sinusoidal Modelling Synthesis* (SMS) [8, 9, 10]. This includes both the spectral peak detection and the spectral peak continuation algorithms. Processing audio backwards is also supported. This is common in analysis techniques for additive synthesis [11].

The ModalModel class inherits from SinusoidalModel and adds custom functionalities for modal sounds. During the analysis step a ModalModel can check additional constraints. A ModalModel can discard trajectories if the modal frequencies are out of the desired boundaries, e.g. infrasounds or ultrasounds. It will, also, discard ill-behaved partials whose magnitude increases with time, instead of decreasing. When a trajectory stops, its starting amplitude is estimated and it can be discarded if it is too quiet. Finally, the stopped trajectory can be merged with a previous trajectory if their frequencies are similar.

2.2 Hinge Regression

The *HingeRegression* class implements the regression algorithm for the modal amplitudes and decay times. It is based on the Rectangular Trust Region Dogleg Approach (DogBox) [12] for non-linear least-squares optimization, implemented in SciPy [13].

HingeRegression fits an amplitude trajectory output by the sinusoidal model to a hinge function $h_{k,q,\alpha}(t)$, which

is a function that is linear for $t < \alpha$ and then continues as a constant.

$$h_{k,q,\alpha}(t) = k \cdot \min(t, \alpha) + q \tag{1}$$

HingeRegression improves the parameters estimate with respect to linear least-squares when the noise floor is high. In this case, linear regression is biased towards lower values for both the amplitude and the decay time [3].

2.3 BeatsDROP

The package also includes the implementation of the "Beats Duality for the Resolution Of Partials" (BeatsDROP) algorithm. The *DualBeatRegression* class implements a regression model for two modal partials with very similar frequencies. In this case, the sinusoidal model would not find two different trajectories, but it would output a single trajectory, in which the two partials interfere.

DualBeatRegression fits the modal frequencies, amplitudes and decay times of both partials to the sinusoidal model trajectory. The loss function includes the differences for both the amplitude and the frequency of the trajectory. This algorithm exploits the fact that beats of uneven amplitudes produce modulations of both amplitude and frequency. This regessor can be used instead of HingeRegression, to find two sets of modal parameters from one trajectory.

3. GRAPHICAL USER INTERFACE

We designed a Graphical user Interface to enable non-developers to access to our methods. It allows to load an audio file, apply the SAMPLE model, and export the results as a JSON file. Figure 1 displays a series of screenshots of the GUI in action.

3.1 Load Audio

The first pane of the GUI is the *"Load Audio"* pane, which allows the user to load and trim a target audio file.

The *"Load"* button opens a file browser for selecting the audio file to load. After loading a file, the GUI displays the audio waveform. We initialize the region-of-interest (ROI) using an algorithm for onset detection [14] and highlight it in a different color.

The SAMPLE algorithm will only process audio in the selected ROI. The users can adjust the ROI by specifying the desired *"start"* and *"stop"* timestamps in the text input boxes. Alternatively, they can modify the start and the stop timestamps by clicking on the waveform display: when the user clicks down on the display the GUI decides whether to affect the start or the stop timestamp (depending on which one is closer), while the specific timestamp value is determined on mouse release. The user can zoom in and out the waveform display for a finer selection of the ROI.

The user can listen to the currently selected ROI of the audio by clicking on the *"Play"* button. We handle audio playback for multiple platforms using PyGame [15].

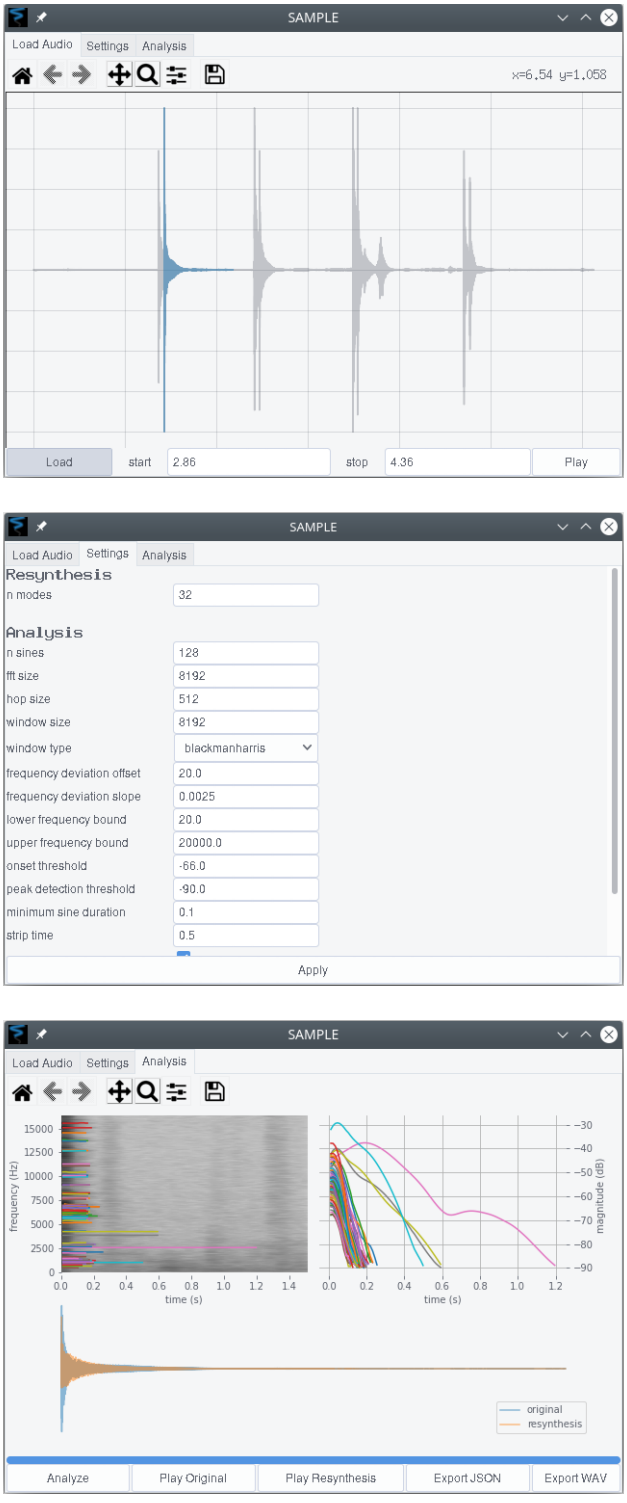


Figure 1. The three panes of the GUI, as they render on Ubuntu 20.04 with the *arc* theme. In the *"Load Audio"* pane, the user can load, trim and play the audio input. In the *"Settings"* pane, the user can tweak the algorithm parameters. In the *"Analysis"* pane, the user can run the algorithm, play an audio resynthesis and export the estimated modal parameters.

3.2 Settings

The second pane of the GUI is the *"Settings"* pane, which allows the user to specify most of the parameters for the SAMPLE algorithm.

- `n modes` controls the maximum number of modes that will be resynthesized by the model. Only the modes with the highest energy will be synthesized. The user can use this parameter to explicitly set the *level of audio detail* (LOAD). This parameter only affects the resynthesis, and not the analysis.
- `n sines` controls the maximum number of sinusoidal peaks that SAMPLE will track in each STFT frame.
- `fft size` is the dimension of the FFT for each frame. The user input is automatically rounded up to the next power of 2.
- `hop size` is the distance between STFT frames on the time axis, in samples.
- `window size` is the dimension of the STFT analysis window, in samples. It doesn't have to be a power of 2 and it is capped at the current `fft size`.
- `window type` is the name of the window function. It can be any of the window functions supported by SciPy [13] which doesn't need any more arguments than the window size.
- `frequency deviation offset` is the threshold for the peak continuation at 0 Hz, in hertz. Peaks at a frequency difference below the threshold will be considered the continuation of one another.
- `frequency deviation slope` determines the increment of the threshold for the peak continuation with frequency. The threshold at frequency f is

$$\tau(f) = \tau_{\text{offset}} + \tau_{\text{slope}} \cdot f \tag{2}$$

- `lower frequency bound` is the minimum frequency for accepting a trajectory, in hertz.
- `upper frequency bound` is the maximum frequency for accepting a trajectory, in hertz.
- `onset threshold` is the initial amplitude threshold for accepting a trajectory, in dBFS. If the linear fit of the amplitude trajectory returns an intercept below the threshold, the trajectory is discarded.
- `peak detection threshold` is the threshold for detecting a peak in the STFT, in dBFS.
- `minimum sine duration` is the minimum trajectory length for accepting a trajectory, in seconds.
- `strip time` is the maximum onset time for accepting a trajectory, in seconds. Trajectories starting later than this time will be discarded.
- `reverse` controls whether to process the audio backwards or not.
- `gui theme` is the name of the GUI theme, as defined in `ttkthemes` [16]. The new theme is applied the next time the GUI starts. When the theme is changed, the user is prompted whether they want to reload the GUI or not.

3.3 Analysis

The last pane of the GUI is the *"Analysis"* pane. The *"Analyze"* button runs the SAMPLE algorithm on the target audio. A progress bar displays how much of the target audio has been analyzed by the algorithm.

Once the analysis process has finished, the GUI populates the visual display. The top left subplot shows the spectrogram of the target audio in grayscale. The frequency trajectories are overlayed in different colors. The top right subplot displays the amplitude trajectories. Finally, the bottom subplot shows the waveform of the target and of the resynthesized audio. The resynthesis is a simple additive synthesis, where all partials have exponential amplitude envelopes.

$$\hat{x}(t) = \sum_{i=1}^{\text{n_modes}} a_i \exp\left(-\frac{2}{d_i}t\right) \cos\left(2\pi\nu_i t + \phi_i\right) \tag{3}$$

Where a_i is the modal amplitude, d_i the decay time, and ν_i the modal frequency of the i-th mode. Phase values ϕ_i can be chosen arbitrarily, at random, or set to zero.

The user can listen and compare to the target audio and the resynthesis, by clicking on the *"Play Original"* button and *"Play Resynthesis"* buttons, respectively.

The *"Export JSON"* button saves the inferred modal parameters as a JSON file. The *"Export WAV"* button saves the resynthesised audio as a WAV file. A file browser opens to make the user choose where to save the file.

4. AUTOMATIC OPTIMIZATION

The SAMPLE algorithm has many hyperparameters and it is not always obvious a-priori what the best values might be for the specific target sound. For this reason, we implemented a module to apply automatic hyperparameter optimization to SAMPLE. Figure 2 shows the loss function values estimated by the minimization process while optimizing the analysis window size, the number of peaks, the peak threshold and the minimum trajectory duration.

The class *SAMPLEOptimizer* allows to define the optimization problem with great flexibility. The developer can specify the value for any number of hyperparameters, which will not be optimized. A good practise is to specify a maximum number of modes to use in resynthesis, to avoid overfitting, because having a high number of modes often causes noise to be modelled as sinusoids even for a small decrease in the loss function. The specific value depends on the complexity of the input sound, but we generally obtained satisfactory results setting the number of modes between 16 and 128.

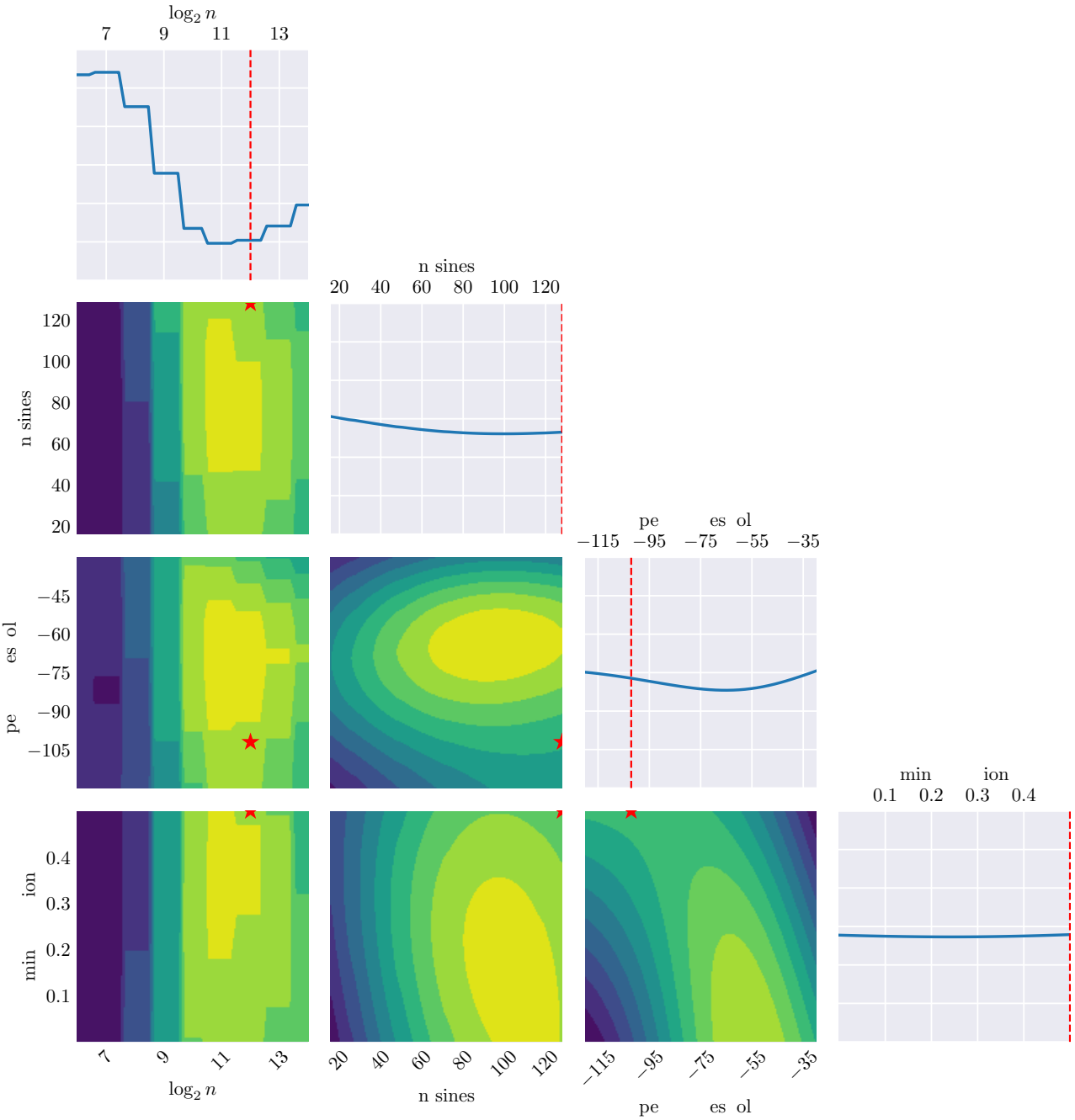


Figure 2. Partial Dependence plots for SAMPLE hyperparameter optimization on a synthetic modal sound. They show the dependence of the cochleagram loss function on each individual or pair of hyperparameters. A partial dependence plot (PDP) shows "the dependence between the target response and a set of input features of interest, marginalizing over the values of all other input features" [6]. Visualizing a function of more than two inputs would be impossible, so PDPs visualize different *views* of the same function. The line plots on the diagonal show the dependence of the loss function on one hyperparameter at a time. The contour plots in the lower triangle show the dependence of the loss function on two hyperparameters at a time.

Alternatively, the developer can specify hyperparameters as Scikit-Optimize [17] dimensions. These hyperparameters will be optimized by a Bayesian optimization algorithm based on Gaussian Processes (skopt.gp_minimize).

Some arguments of the *SAMPLE* class are not suitable for optimization. For example, the analysis window is a high-dimensional vector: optimizing all the values in the vector would be absurd. Instead, we would like to find the best window size and type. A remap function can be specified to transform hyperparameters into valid arguments for the constructor. The default remap function allows to specify the logarithm of the FFT size, to define the window with a name and a window size (as a fraction of the FFT size) and to specify the window overlap as a fraction of the size, instead of specifying the hop size in samples.

The loss function defines the optimization objective. It accepts the two arrays of original samples and resynthesized samples and returns a score, where lower is better. By default, it is a multi-scale spectral loss.

4.1 Loss Functions

We implemented several time-frequency audio representations with the aim of defining perceptually-based loss functions.

The first loss function we implemented is the multi-scale spectral loss. This is defined as the weighted sum of the L_p distances (usually with $p = 1$) between the spectrograms of the two sounds, with linear amplitude and in decibel, for different window sizes [18]. Given a set of spectrogram functions with different resolutions Σ, the loss function is defined as

$$\ell_{\Sigma,\alpha,p}(x, y) := \sum_{S_i \in \Sigma} \|S_i(x) - S_i(y)\|_p + \\ + \alpha \sum_{S_i \in \Sigma} \|\log S_i(x) - \log S_i(y)\|_p \quad (4)$$

To make this function more efficient, we optionally allow developers to specify a number of jobs or a process pool to compute the different spectrograms in parallel.

We also implemented a fast cochleagram function. A cochleagram is a time-frequency representation whose time and frequency resolution change with frequency and mimic human perception [19]. Because of this property, there is no need to compute cochleagrams at different resolutions to define a perceptually-based loss function. We implemented the cochleagram as the convolutions of the audio signal with a set of impulse responses (IRs) of *gammatone filters*. Traditionally, after the convolution a simple non-linearity would be applied, such as half-wave rectification. We allow the option of convolving the signal with the analytic signals of the IRs: this is equivalent to computing the analytic signal of the output [20], but faster. In this case, the output is complex and the default non-linearity is the absolute value, so that the cochleagram is real.

The resulting cochleagram has the same sampling rate as the input. This is usually excessive, especially for the purpose of defining a loss function, and the cochleagram is downsampled to the desired sample rate. This is wasteful, because we compute a high-resolution output and then discard a great portion of it. We implemented a custom function for what is commonly known as a *strided* convolution: the output for a strided convolution with stride $s \in \mathbb{N}^+$ at sample i is the output for a convolution at sample $s \cdot i$.

$$(x *_s y)[i] := (x * y)[si] \quad (5)$$

For a sufficiently big stride, our implementation is faster than using the best of all of the convolution methods available in SciPy [13], which are: *direct* (convolution in the time-domain), *fft* (product in the frequency-domain), *auto* (automatically determines whether to use *direct* or *fft*), and *overlap-and-add* [21]. Figure 3 shows the average speed-up for different strides and input sizes.

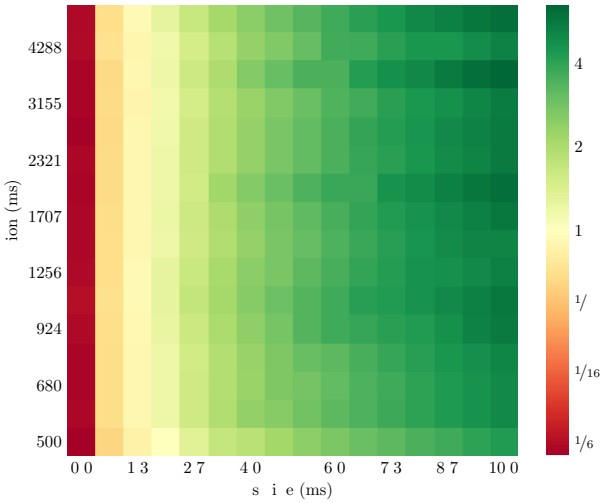


Figure 3. Speed-up of cochleagram method using the custom implementation for strided convolution. Speed-up is the ratio between the run-time of the custom implementation and the run-time of the best of four convolution methods (*direct*, *fft*, *auto*, and *overlap-and-add*). Speedup is evaluated for different input sizes and stride lengths (0 ms means 1 sample of stride) at 44.1 kHz averaging the run-times of 32 trials per case.

5. CONCLUSION

We presented the SAMPLE package for Python and its main components. The package is tested with unit tests with full code coverage.

We provided a general overview of the classes that implement the SAMPLE [3] and the BeatsDROP algorithms for the analysis of modal sounds.

We presented the Graphical User Interface for the SAMPLE method and how it constitutes a sufficient environment for preprocessing target audio files, configuring the algorithm hyperparameters, running the algorithm, inspecting and exporting the results.

Finally, we presented a module for interfacing SAMPLE with a Bayesian hyperparameter optimization function based on Gaussian Processes.

In the future, we plan to keep adding to the package any method related to SAMPLE that we may develop.

6. REFERENCES

[1] J.-M. Adrien, *The Missing Link: Modal Synthesis*, pp. 269–298. Cambridge, MA, USA: MIT Press, 1991.

[2] F. Avanzini, M. Rath, and D. Rocchesso, "Physically-based audio rendering of contact," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME2002)*, vol. 2, (Lausanne, Switzerland), pp. 445–448, IEEE, 2002.

[3] M. Tiraboschi, F. Avanzini, and S. Ntalampiras, "Spectral Analysis for Modal Parameters Linear Estimate," in *Proceedings of the 17th Sound and Music Computing Conference* (S. Spagnol and A. Valle, eds.), SMC, (Torino, Italy), pp. 276–283, Sound and Music Computing Network, Axea sas/SMC Network, 6 2020.

[4] M. Tiraboschi, "SAMPLE – Python package." https://doi.org/10.5281/zenodo.6536419, 2021.

[5] P. Hamill, *Unit test frameworks: tools for high-quality software development*. "O'Reilly Media, Inc.", 2004.

[6] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[7] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, R. Layton, J. VanderPlas, A. Joly, B. Holt, and G. Varoquaux, "API design for machine learning software: experiences from the scikit-learn project," in *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pp. 108–122, 2013.

[8] X. Serra, *A system for sound analysis/transformation/synthesis based on a deterministic plus stochastic decomposition*. PhD thesis, CCRMA Department of Music, Stanford University, 1989.

[9] X. Serra and J. Smith, "Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition," *Computer Music Journal*, vol. 14, no. 4, pp. 12–24, 1990.

[10] X. Serra, *Musical sound modeling with sinusoids plus noise*, ch. 3, pp. 91–122. Lisse, the Netherlands: Swets & Zeitlinger Publishers, 1997.

[11] J. Moorer, "Signal processing aspects of computer music: A survey," *Proceedings of the IEEE*, vol. 65, no. 8, pp. 1108–1137, 1977.

[12] C. Voglis and I. Lagaris, "A rectangular trust region dogleg approach for unconstrained and bound constrained nonlinear optimization," in *WSEAS International Conference on Applied Mathematics*, (Boca Raton, FL, USA), p. 4, CRC Press, 2004.

[13] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[14] S. Bock and G. Widmer, "Maximum filter vibrato suppression for onset detection," in *Proceedings of the 16th International Conference on Digital Audio Effects*, (Maynooth, Ireland), 2013.

[15] M. Von Appen *et al.*, "Pygame - a free and open source python programming language library for making multimedia applications." https://pypi.org/project/pygame. Version 2.1.2.

[16] RedFantom *et al.*, "Ttkthemes - a group of themes for the ttk extensions of tcl." https://pypi.org/project/ttkthemes. Version 3.2.2.

[17] T. Head, M. Kumar, H. Nahrstaedt, G. Louppe, and I. Shcherbatyi, "Scikit-optimize," Oct. 2021.

[18] J. Engel, L. Hantrakul, C. Gu, and A. Roberts, "Ddsp: Differentiable digital signal processing," in *International Conference on Learning Representations*, 2020.

[19] D. Wang and G. J. Brown, *Computational auditory scene analysis: Principles, algorithms, and applications*. Wiley - IEEE press, 2006.

[20] J. O. Smith, "Analytic signals and Hilbert transform filters," in *Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications*, W3K Publishing, Second ed., 2007.

[21] R. G. Lyons, *Understanding Digital Signal Processing*, ch. 13.10. Pearson, Third ed., 2011.

STAGING A CONTEMPORARY HIEROPHANY: STRATEGIES OF RHYTHM AND ALGORITHMIC VIOLENCE IN *ΔNFANG* BY THE FRONTE VACUO COLLECTIVE

Luca Befera
Università di Torino
luca.befera@unito.it

ABSTRACT

The present paper studies the most recent performances of the Fronte Vacuo collective, which merge ubiquitous digital elements with the performers' body dimension. Specifically, I will consider the case of *Δnfang*, the first and prototypical piece of the *Humane Methods* cycle, delving into the algorithm processing and staging. The software determines repetitive sound and light patterns based on the AI underlying learning process. Its real-time evolution continuously redefines the piece's multimedia content, also influencing human beings' behaviour. Naked and speechless performers enact obsessive actions aimed to express the software constraints and hegemony. The dramaturgy and overall structure are hence fixed, whereas internal parameters continuously change. Within this framework, I hypothesise that both the algorithm and performers might be conceived as instruments: the former, as built by the authors to autonomously play; the latter, as controlled and irrevocably influenced by the machine. Repetition stands as the fundamental parameter that permeates gestures, bodies, digital artefacts, and symbolic meaning towards a ritual attitude, enacted as an unavoidable as much as sterile practice.

1. INTRODUCTION

Contemporary musical performance has progressively entailed different media over the last decades, which extended staging and sensory perspectives [2,10]. The involvement of tactile, visual, and gestural dimensions has been strictly connected to the employment of various analogue media – such as sensors, lights, and visuals – and digital technologies – such as software used to compose automated or interactive processes [12,33,46]. These tools not only assumed a scenic function but became a pluralistic instrumental apparatus usually accessed through graphic interfaces [4,37]. The implications of the second and third waves of human-computer interaction [6,35] have then determined a new relationship with technology implying the predominance of digital environments and their merging with broad socio-ocultural contexts. The usage of artificial intelligence [30,44] and virtual or augmented realities [24,29] within performances and their growing real-time interactivity related to this perspective, manifesting the ubiquitous and pervasive nature of information technologies [40]. Such approaches were also parallel to the technical evolution that took place in the theatrical field [32], showing ongoing similarities to such scenic representations also in the centrality of bodies [39,48]. However, authors such as Marko Ciciliani, Jenifer Walshe, Alexander Schubert, Brigitta Muntendorf, Samson Young and Stefan Prins have still maintained a prominent use of the musical dimension. Parallelly, they are pursuing an experimental approach questioning the role of digital media, in which theatre "rediscovers itself as a part of society, that is, a laboratory of the social, in a very different sense than before: not as a workshop for innovation and not as a place for individual experience, but as an epistemological device in which the question of collectivity can be posed as a problem of government" [36]. In other words, they aim to unravel pervasive social dynamics rooted in information technologies, revealing or exploiting them within a theatrical environment.

Works by the Fronte Vacuo collective, founded in 2019 by Marco Donnarumma, Margherita Pevere, and Andrea Familari, are posited within this environment, merging the expertise of each author towards a particular aesthetic approach. Even if all dealing with intermedia composition, its members singularly studied muscular sensors and mechanical prosthetics [13], biological and biotechnological matter [41], and audiovisual performance tools [20]. The three works of the collective realised so far, all belonging to the *Humane Methods* cycle, brought on stage these different facets. Also, they focus on violence in algorithmic societies [22], namely discrimination carried by wide-spreading algorithms that impose partial accounts and pre-existing biases. Indeed, the AI automated systems work "in and through the relations of selves to selves, and selves to others, as these relations are manifest in the clusters and attributes of data"; consequently, it necessarily discriminates by

Copyright: © 2022 Befera. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

affording "greater degrees of recognition and value to some features of a scene than they do to others" [3].[1]

In a posthuman [27], feminist [9], and cooperative [26] perspective, Fronte Vacuo considers the performance as a body composed of *various* organic bonds which encompass humans and non-humans, living and non-living beings [16,42]. Rhythm stands as crucial since, in a conceptual view, since both sounds and matter are composed of repetitive patterns that define a distinctive milieu [11,16]. The different layers exploited in *Humane Methods* regard naked and speechless performers with repetitive and jerky movements, interacting within the sonic and light settings extemporaneously defined by the algorithm. Some of these aspects are not new in theatre and dance art production, as stated by the authors themselves [15,14]. The reflection on body and technology refers to the prosthetic tools in pieces such as *Exoskeleton* (1999) by Stelarc and *Epizoo* (1994) by Marcel·lí Antúnez Roca; the intermedial performance to works by the Santasangre collective such as *Sigradi* (2008); nudity and choreography to dance performances by Maria Donata D'Urso such as *Collection Particulière* (2006); costumes to the living sculptures by Olivier de Sagazan such as *Transfiguration* (1998); symbolism and rituality to theatrical pieces by the Societas Raffaello Sanzio such as *Tragedia Endogonidia* (2003).

On the other hand, other recent performances also dealt with algorithmic surveillance and computational dominance. In *Algorithms* (2015) by the Turbo Pascal collective, for example, spectators are arranged according to instructions given by the software and delivered by speakers and performers [36]. In *Sight Machine* (2017) by Trevor Paglen, the Quartett and audience members are constantly monitored by a camera while the show goes by, and their shapes are projected on a background canvas [50]. Both still differ from *Humane Methods*, especially because the former does not involve real-time processing, and AI is only assumed; the latter does not imply a ritual dramaturgy but a canonical concert venue, eventually augmented by digital media. Finally, sound and light pulsation patterns, further analysed, also recall the approach of minimalist composers such as Steve Reich – in homorhythmic loops and phase shifting [19] – and Fluxus authors such as Earle Brown – in their aleatoric assemblage according to pre-defined sections [25]. Nevertheless, the extemporaneous management of audiovisual material is not assigned to the performer but to the automated computation of the algorithm. The distinctive perspective of the play is indeed rooted in the AI dramaturgical implications, as always entwined with living beings on stage. The machine interplay is perceivable through the audiovisual outputs rendered by analogue media and paired with performers' actions through recursive gestures occurring within clearly defined narrative cycles. Also, molecular processes of onstage biological matter – as fungi and plants – conceptually recall the sequential renovation of the software, whereas in-motion mechanical devices – as self-standing prothesis – the performers' movement.

In dealing with artificial intelligence, the aesthetics of *Δnfang*, and generally of the *Humane Methods* cycle, considers the different issues already discussed in the academic literature about algorithm problematics [23,32,47,45], generally encompassing: 1) automation, as machines replace human agency through pre-set processes based on datasets; 2) architecture, as automated digital models influence the structure of commercial and public organisations; 3) predictive models, concerning the employment of quantitative knowledge for anticipating people behaviour. In this article, I will explore these issues in relation to *Δnfang*, hypothesising that the dramaturgical intent revolves around AI as a ubiquitous – or even sacred – entity which slowly and imperceptibly consumes human identity and body. Rhythm and repetition outline a ritual attitude that human beings cannot avoid, resulting as aligned with as contaminated by the recursive methods of the computational artefact.[2] Or, as described on the piece webpage, "it counts, they worship, they touch, they sense, it counts, they isolate, they hate, they attack" [21].

2. *ΔNFANG*: BODIES, INSTRUMENTS, AND PROCESSING

Δnfang (2019) is the first production of Fronte Vacuo and stands as a prototype for the following works, which developed from the same human-machine interaction but with different procedures.[3] It enacts the relationship between two main characters (interpreted by Marco Donnarumma and Margherita Pevere) evolving through recursive loops, where each scene restarts with the initial setting but also provides a slight change of the plot during its development. In the first cycle, the performers appear on stage dressed in a full-length robe and a cloth to cover the face with holes at the height of the eyes. They search for a stool and, once found, they grab it and sit down. Then, they perform the same mechanical gesture over and over, which recalls the cross sign ending with a shot in the forehead (Fig. 1). Meanwhile, sonic and light patterns play according to numeric outputs automatically generated by the AI learning

[1] *Δnfang* was premiered on 4 October 2019 at Romaeuropa Festival, Rome; *ʊr*, 22 September 2020 at Touch Me Festival, Zagreb; *Σxhale*, 5 February 2022 at CTM Festival, Berlin. The main differences between the pieces regard the number of characters – from one in *ʊr* to six in *Σxhale* – and the algorithm function – dealing with only lights in *ʊr* or lights and sounds in the others but working with autonomous inputs in *Δnfang* and with performers' motion data in *Σxhale*.

[2] Besides the collective website [22], this aspect was also expressed in the meeting I had with the authors in Vienna after the *Σxhale* performance on the 13 of March 2022.

[3] Except for the article drafted by the authors themselves [8], there are no scientific writings about the performance. Therefore, I will report data especially gathered from the collective website [22], the online interview with Marco Donnarumma and Andrea Familari occurred on May 9, 2022, and the analysis of the entire video of the premiere. I am kindly grateful to them and Margherita Pevere for the shared data and the attention devoted to me.

process. Over the cycles, the male figure on the right gets ever closer to the female one on the left in a devious way, but his actions are always interrupted: excluding the eighth and last sequence slowly fading out before the coda, the previous seven ones end abruptly with other performers, previously acting and watching in the background, stopping him while music and lights switch to a sinusoidal drone and a stable and warm neon (Fig. 2). The plot development also provides the secondary characters gradually undressing the main ones and, around the end, limiting the male freedom by tying his left leg. The climax occurs on the eighth repetition when the constricted male character enacts a silent scream. It follows a march of seven naked performers accompanied by a pre-set musical track where they simultaneously and slowly move in front of the stage with a jerky walk, until falling to the ground one by one.[4] Hence, this episode is a coda that breaks the loop repetition and brings to the performance end.

As here inferred, the hierarchical dynamic between AI and performers is enacted through the plot evolution, implying an obsessive and constrictive framework which ends with human identity deprivation. Both parties show the inability to reach their goal, as always partially controlled by someone or something else. Therefore, I posit here that different and never fulfilled agencies – as the "satisfying power to take meaningful action and see the results of our decisions and choices" [34] – are employed: performers, who are supposed to be consciously acting, emerge as mechanical and subjugated entities; the algorithm, which manifests the ability to evolve, is constituted as an inanimate computational artefact. To highlight this conflictual and infertile relationship, I propose to treat both parties not only as agents but as instruments which are built or enacted to be used – by the software or by the authors, always by rhythm centrality. The performance emerges as a dark ritual that is still enacted within the theatrical milieu to show compulsive behaviour and mortification of the body happening within a digital hegemonic context.



Figure 1. *Δnfang*, Romaeuropa Festival, snapshot showing the first episode.

[4] For the sake of argumentation, I will not address the symbolic meaning of the beginning and the end scenes employing the figure of the deer. Even if significant, this aspect is considered subsidiary for the present purposes.



Figure 2. *Δnfang*, Romaeuropa Festival, snapshot showing the first intermezzo.

2.1 Algorithm-Instrument and Digital Agency

The algorithm is the core digital aspect of the performance not only for the overall concept but also because it defines in real-time a substantial part of the scenic setting. It implies reinforcement learning whereby rewards are given to assist the achievement of certain results [46].[5] In this case, it is a Deep Q-learning implemented in Python and programmed through the Keras library. The task is to get as close as possible to an array of ten decimal numbers from 0 to 1 arbitrary assigned. Positive rewards are automatically given when getting closer and negative vice versa. The algorithm moves in parametric space and prints its values on a UDP server. Then, it gets the rewards and changes its status by adding or subtracting a constant in relation to the neural network processing. This process restarts from scratch for each episode. Indeed, the software is not programmed to fulfilling the task but to constantly try and start over [8].[6] Its behaviour is made intelligible by means of analogue devices on stage: the numerical data generated by the machine are converted to OSC protocol and then processed through Pure Data – used by Donnarumma for sounds production – and TouchDesigner – managed by Familari for light control. The two automated dynamics are therefore distinct but generated from a common matrix referring to the AI outputs. The main values employed to manage sounds and lights, as visible in Figure 3, are not only the numbers computed by the algorithm (on top) but also the distances (on bottom) – namely, the difference between the given values and the values outputted by the algorithm, approximated as integers from 0 to 2 – and the two rewards (in the centre).[7]

[5] The source code – developed by Baptiste Caramiaux like all the pieces of the *Humane Methods* cycle – is freely accessible on the programmer's GitHub account [7]. I also gathered more data from the email exchange with the programmer occurred on 29 August and 13 September 2022.

[6] Besides the lack of a long-term memory, Donnarumma also stated in the email exchanges occurred on 13 June 2022 that the goal is set to be almost impossible to reach.

[7] As Donnarumma and Familari stated in the email exchanges respectively occurred on 13 and 14 June 2022, even if the range is reported between 1 and 3 in the article about *Humane Methods* [8], both Pure Data and TouchDesigner process values from 0 to 2 for technical reasons. Basically, they are the same 3 values but with an offset of 1

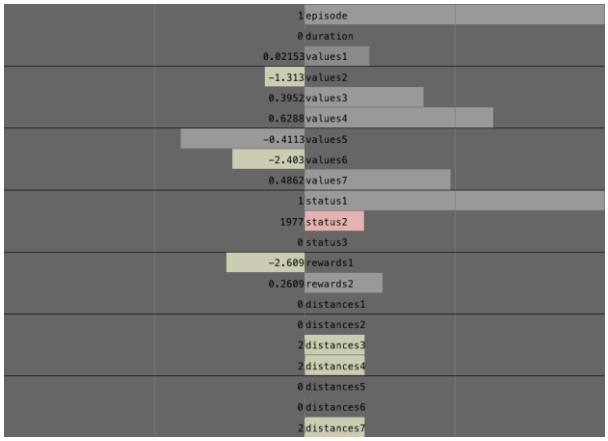


Figure 3. *Anfang*, excerpt from the TouchDesigner patch showing the algorithm outputs.

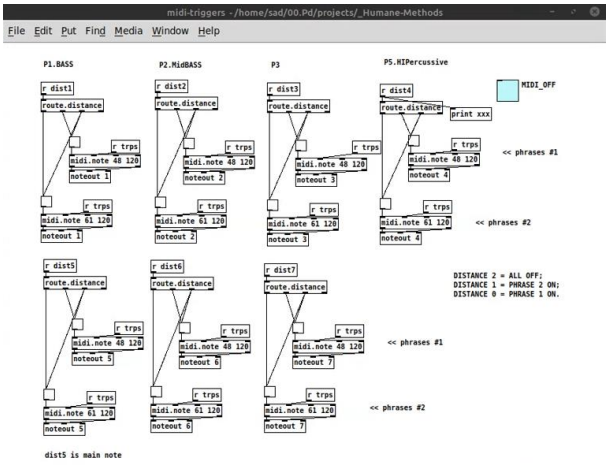


Figure 4. *Anfang*, excerpt from the Pure Data patch showing the 14 sound patterns.

Sounds are managed according to two pattern sets called “phrase 1” and “phrase 2” in the Pure Data patch visible in Figure 4, each consisting of seven elements. As shown in the note on the right, each of these phrases is triggered by the distance values: e.g., if the algorithm prints a number whose distance to the first given value is approximated as equal to zero, the first pattern of the first phrase will be activated; if equal to one, the second: if equal to 2, none of them. Patterns are played at 300 bpm and are replaced each 64 beats according to the algorithm re-inputted values. Figure 5 shows the transcription of the 14 patterns. The first index number refers to the pairs of patterns shown in Figure 4; the second one to the phrase. Hence, the first two staves correspond to the block on the top-left corner of the patch. Each block employs the same piano sample obtained from Renoise, thus implying seven different piano timbres.[8] Each quaver is equal to one beat, so that the two bars reported in Figure 5 are played eight times before changing (8 quavers × 8 repetitions = 64, which is the number of overall beats before other values are inputted). It is possible to overview the heterogeneity of the different patterns in regard to pitches and accents, which allow a certain variety within an overall homorhythmic tendency. Additionally, pitches do not all belong to the same harmonic series, often causing a certain acoustic nuisance which sides the obsessive patterns. Moreover, another slight variation is related to the sound triggering, as the AI values are gathered in Pure Data in a span of 500 milliseconds instead of 50 – which would be enough to guarantee correct interoperability. The higher value makes certain glitches occur, thus breaking the rhythm regularity with unknown, but not too disruptive, events (e.g., adding an initial phase shifting between patterns).

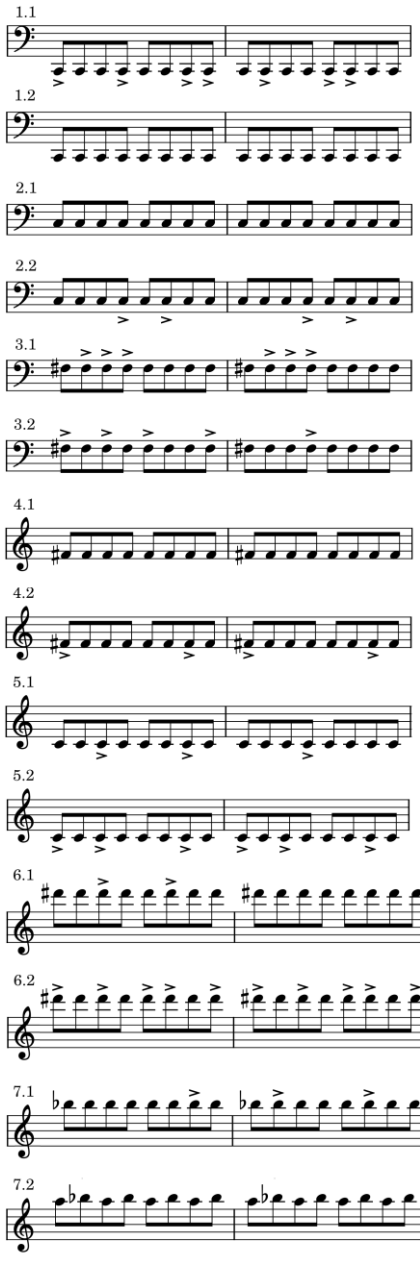


Figure 5. Transcription of the 14 sound patterns.

position. Note that only seven out of the ten values are employed for these calculations, and status and direction have never been used.

[8] The distinction between bass and percussion reported in the figure, vaguely recalled also by the pitch of the patterns, might suggest the influence of electronic popular music belonging to Donnarumma’s background [17].

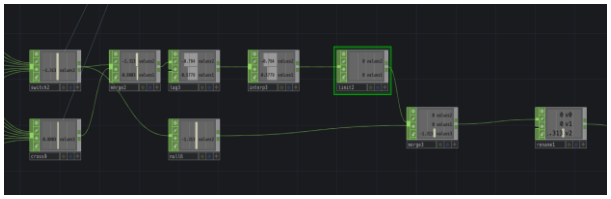


Figure 6. *Anfang*, excerpt from the TouchDesigner patch showing the lights starting management.

Lights, on the other hand, are triggered every second, resulting in the perception of a continuous flow still based on rhythmic patterns of pulsating intensity. This medium mainly employs the values outputted by the algorithm, not the distances. The excerpt of the TouchDesigner patch in Figure 6 highlights the prior transformation applied to these values to reproduce three different rhythms. The seven numbers are initially filtered by a switch module (on top-left) – which determines the simple alternation between one value and another – and a cross module (on bottom-left) – which also provides the transition throughout all the intermediate values between the seven numbers given by the algorithm. These functions are then reported in the sequence of modules at the centre: on the top, both are firstly merged, then slowed and interpolated to be smoothed – where the limit excludes negative or too high values for better management during the live performance; on the bottom, the switch remains as it is, also including the negative values. It follows the three oscillation patterns on the right: the interpolated and positive switch ("v0"); the interpolated and positive cross ("v1"); the cross as it is originally provided ("v2"). Further calculations are applied to associate each pattern to lights, which are pre-set and change depending on the scenic episode. Therefore, lights appear from ever-changing sources and in multifaceted oscillations – still not belonging to extended patterns as for music. Another cross further relates the intensity to learning, as lights become more stable as the distance to the given values decreases or, conversely, more dynamic**.**

The outlined techniques emulate AI learning through sound patterns and light oscillations which make the process perceivable on stage.[9] The algorithm, hence, is conceived "not for its capacity to achieve a certain goal, but rather for the particular computational behaviour driving its learning mechanism […], interpretable only by observing its choices and subsequent actions. In this sense, it became observable as an actor" [8]. Still, lacking an aesthetic intent, the software refers to pre-set statistical rules, so that its prerequisites are not detached from the authors' will [43]. It thus manifests a digital agency, namely the "capability of machines to act autonomously, but on behalf of humans" [1], even if implying a certain ability to evolve, act, and make choices. In summary, the algorithm stands halfway between an actor – as it is enacted to be – and an automated entity – as it only processes data over the whole piece, and its behaviour is expressed by balancing a certain mechanical homogeneity – e.g., concerning tonality of lights and regularity of patterns – with slight organic changes, towards a stochastic and obsessive atmosphere shaped in its likeness.

2.2 Performer-Instrument and Ritual Enactment

If the algorithm works through commands given by humans, performers are themselves influenced by the computational environment that, as staged, governs their behaviour. Indeed, they act as mechanical beings not able to realise their goals and enclosed in sterile relationships. Their subjugation is enacted within a ritual framework constituted by two main aspects: performers' connotation and gestures; the algorithm ubiquity suggested by the sound and light design.

Regarding the first aspect, ritualisation is "embedded within the dynamics of the body defined within a symbolically structured environment" [5], firstly recalled by the actors' rough clothing or nudity already mentioned. Then, performers repeat "a hybrid form of gestural prayer" each time with a variation, thus mirroring "the episodic learning of the algorithm" [8]. Figure 7 shows an example from the scene in the middle of the performance, where a red liquid – strongly recalling blood – is poured on the female character's belly after the male and the other actors gather around her. In the following episode, the male is forcibly undressed and tied. The disruptive deprivation hence brings to the manifestation of the character's naked body, presumably revealing his inner nature as a limited being. The event snapshotted in the figure is liminal: from then on, he and all the other characters will progressively collapse, showing the negative facet of the ritual event.

Concerning the algorithm, it lacks a clear physical embodiment but is depicted as ubiquitous through an ever-perceivable audiovisual environment.[10] Thus, it stands as "Unseen", namely as "an entity of imagination [which] ontological status is not known" and which, as such, gains normative values or even transcendent meanings [38]. At the same time, the mechanical repetition implies "the 'power, reliability and immortality of the machine'", as already explored by minimalist composers [28]. Automation and normative/transcendent values converge in a mechanical entity lying in another reality that, at the same time, is accessible through audiovisual signs. Through this ambivalence, the authors enact the underlying power relationship of the machine, as humans perform actions driven by computational factors as tangible as out of their consciousness.

[9] Another method also implies that music gets quieter and light dimmer as AI approaches the target values over the episodes [8]. Fading into darkness and silence might thus be compared to human collapsing which progressively occurs.

[10] The only physical manifestation regards the two protheses fixed on the iron structures, as stated on the performance website [21]. Still, these elements are not directly relatable to the algorithm because always in the background, working only in the last part of the play and not synchronized with the audiovisual stimuli.

It should also be noticed that it remains unclear if performers are aware or not of the algorithm stimuli during the play since they are not explicitly synchronised to what it does. The audiovisual outputs can be considered either diegetic – where beings on stage move their ritual actions precisely from the perception of lights and sounds in the background – or extradiegetic – where AI is instead manifested only to the audience who can overview its implicit and pervasive existence. As an appendix to this argument, it can be observed that the final march, as a coda in which the meaning of the work is shown, is instead explicitly synchronised with the steps of the performers. In any case, regardless of performers worshipping the AI or another related entity, the algorithm exercises its power as "something of a wholly different order, a reality that does not belong to our world, in objects that are an integral part of our natural 'profane' world"; in other words, it represents a "hierophany" in any case [18].

Hierophanies, as referring to myth and religion, imply the conscious recognition of sacred elements as reference points for better understanding the world and getting oriented within it. In this case, instead, the AI algorithmic violence induces loss and subjugation, also because its deceitful dictates are passively suffered. Therefore, the rhythmical audiovisual environment portrayed by the algorithm constricts humans within a negative ritual dynamic that they are not able to avoid. Their body results as driven by its ubiquity and sacred power as an instrument apparently deprived of free will. Hence, it is exactly this distorted view of sacredness, mediated by computational factors, to reveal the intrinsic nature of enacted social relationships, as enclosed within a theatrical environment that shows an archetypical insight. In other words, the actors' unconsciousness is counterposed to the spectators' knowledge: AI still represents a reference point to whoever watches not only for understanding the performative dynamics but also a broader socio-cultural milieu where dark and mischievous issues occur.



Figure 7. *Δnfang*, Romaeuropa Festival, snapshots showing one of the middle episodes.

3. CONCLUSIONS

The analysis of the performance and the underlying computational models have highlighted a prominent ritual dynamic that takes place according to digital dictates. The algorithm is composed to reflect an autonomous entity apparently acting with its own agency and influencing the whole environment and performers' actions. Still, the milieu where the performance takes place is defined by the authors, who also set the algorithm as an instrument that automatically plays. On the other hand, humans recall primitive beings holding their rites and relationships while unconsciously controlled by the digital simulacrum. These sterile loops are rooted in rhythm as a fundamental motion associating bodies, gestures, and digital processes. The performers' unconsciousness of the digital pervasiveness implies that their actions occur after the algorithm audiovisual interplay, as something not willingly made but strongly driven. Consequently, both performers and the algorithm, as acted by someone or something else, might be conceived as instruments with their own affordance.

The ritual dynamic of *Δnfang* emerges from rhythm itself and it is also characterised by the various props and gestures. By dealing with algorithmic violence, it recovers the perceptual dichotomy between the corporeal and the digital, merging it with the opposition between human and divine. The performative venue encloses these dichotomies as a distinctive milieu, so far as the event is enacted. The audience, hence, can observe the whole dramaturgy and participate from the outside in the expressiveness of the work. To this extent, the epistemological goal lies not in the ritual event but in its staging, as a metaphor for contemporary social dynamics. Indeed, *Δnfang* deals with the three issues generally addressed to algorithmic society mentioned in the introduction, as human agency is strongly influenced by machine behaviour; the automated digital model articulates the overall structure of the piece; the employment of processed data somehow determines the dramatic development. It emerges a theatrical setting showing a ritual of digital power and human loss where, insofar as the obsessive repetitiveness and power of the machine become ubiquitous, humans emerge as impotent and unconscious. Insofar as "it counts, they worship, they touch, they sense, […], they isolate, they hate, they attack" [21], and, ultimately, collapse.

4. REFERENCES

[1] P. J. Ågerfalk, "Artificial Intelligence as Digital Agency," *Eur. J. Inf. Syst.*, vol. 29, no. 1, pp. 1–8, 2020.

[2] G. Albert, "Post-Music: L'Ibridazione delle Forme Audiovisive nel XXI Secolo tra Performance e Tecnologia," in *La Musica fra Testo, Performance e Media: Forme e Concetti dell'Esperienza Musi-*

cale, A. Cecchi, Ed. Roma: NeoClassica, 2019, pp. 55–88.

[3] L. Amoore, *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham: Duke University Press, 2020.

[4] D. Arfib, J.-M. Couturier, and L. Kessous, "Expressiveness and Digital Musical Instrument Design," *J. New Music Res.*, vol. 34, no. 1, pp. 125–136, 2005.

[5] C. Bell, *Ritual Theory, Ritual Practice*. Oxford: Oxford University Press, 2009.

[6] S. Bødker, "When Second Wave HCI Meets Third Wave Challenges," in *Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles*, 2006, pp. 1–8.

[7] B. Caramiaux, "Humane Methods," 2020. [Online]. Available: https://github.com/bcaramiaux/humane-methods. [Accessed: 5-Sept-2022].

[8] B. Caramiaux and M. Donnarumma, "Artificial Intelligence in Music and Performance: A Subjective Art-Research Inquiry," in *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*, E. R. Miranda, Ed. Springer International, 2021.

[9] R. Braidotti, *The Posthuman*. Cambridge: Polity Press, 2013.

[10] M. Ciciliani, "Music in the Expanded Field: On Recent Approaches to Interdisciplinary Composition," *Darmstädter Beiträge zur Neuen Musik*, no. 24, pp. 23–35, 2016.

[11] G. Deleuze and F. Guattari, *A Thousand Plateaus*. Minneapolis: University of Minnesota Press, 1987.

[12] S. Dixon, *Digital Performance: A History of New Media in Theater, Dance, Performance Art and Installation*. Cambridge: MIT Press, 2007.

[13] M. Donnarumma, "About," 2022. [Online]. Available: https://marcodonnarumma.com/about/biography/. [Accessed: 15-Jun-2022].

[14] M. Donnarumma, "Across Bodily and Disciplinary Borders: Hybridity as Methodology, Expression, Dynamic," *Perform. Res.*, vol. 25, no. 4, pp. 36–44, 2020.

[15] M. Donnarumma, "Body and Digitality: From Early Experiments to Theater-Making," 2020. [Online]. Available: https://www.youtube.com/watch?v=69ONaW31kl8&feature=emb_title. [Accessed: 15-Jun-2022].

[16] M. Donnarumma, "Configuring Corporeality: Performing Bodies, Vibrations and New Musical Instruments," Goldsmiths University, London, 2016.

[17] M. Donnarumma and F. Arri, "Specie Contemporanea: L'Evoluzione della Carne in Suono Attraverso i Device e i Rituali Evolutivi di Marco Donnarumma," 2020. [Online]. Available: https://www.osservatoriofutura.it/2462-2/. [Accessed: 15-Jun-2022].

[18] M. Eliade, *The Sacred and the Profane: The Nature of Religion*. New York: A Harvest Book, 1987.

[19] T. Evans, "Analysing Minimalist and Postminimalist Music: An Overview of Methodologies," in *The Ashgate Research Companion to Minimalist and Postminimalist Music*, K. Potter, K. Gann, and P. ap Siôn, Eds. Farnham: Ashgate, 2013, pp. 241–258.

[20] A. Familari, "About," 2022. [Online]. Available: http://www.famifax.com/fami/about/. [Accessed: 15-Jun-2022].

[21] Fronte Vacuo, "Δnfang," 2019. [Online]. Available: https://frontevacuo.com/works/anfang/.

[22] Fronte Vacuo, "Humane Methods," 2022. [Online]. Available: https://frontevacuo.com/humane-methods/. [Accessed: 15-Jun-2022].

[23] L. Floridi, *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality*. Oxford: Oxford University Press, 2014.

[24] G. Giannachi, *Virtual Theatres: An Introduction*. New York: Routledge, 2004.

[25] C. Gresser, "Earle Brown's 'Creative Ambiguity' and Ideas of Co-creatorship in Selected Works," *Contemp. Music Rev.*, vol. 26, no. 3/4, pp. 377–394, 2007.

[26] C. Groth, M. Pevere, K. Niinimäki, and P. Kääriäinen, "Conditions for Experiential Knowledge Exchange in Collaborative Research Across the Sciences and Creative Practice," *CoDesign Int. J. CoCreation Des. Arts*, vol. 16, no. 4, pp. 328–344, 2020.

[27] N. Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press, 1999.

[28] J. P. Jones, "Accommodating the Threat of the Machine: The Act of Repetition in Live Performance," in *The Ashgate Research Companion to Minimalist and Postminimalist Music*, K. Potter, K. Gann, and P. ap Siôn, Eds. Farnham: Ashgate, 2013, pp. 141–157.

[29] R. Klich, "Multimedia Theatre in the Virtual Age," University of New South Wales, 2007.

[30] E. R. Miranda, Ed., *Handbook of Artificial Intelligence for Music: Foundations, Advanced Approaches, and Developments for Creativity*. Cham: Springer, 2021.

[31] H.-T. Lehmann, *Postdramatic Theatre*. London: Routledge, 2006.

[32] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The Ethics of Algorithms: Mapping the Debate," *Big Data Soc.*, pp. 1–21, 2016.

[33] A. M. Monteverdi, *Leggere uno Spettacolo Multimediale: La Nuova Scena tra Videomapping, Interaction Design e Intelligenza Artificiale*. Rome: Dino Audino, 2020.

[34] J. H. Murray, *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*. New York: The Free Press, 2016.

[35] K. L. Norman and J. Kirakowski, Eds., *The Wiley Handbook of Human Computer Interaction*. Hoboken: John Wiley & Sons, 2018.

[36] U. Otto, "Theatres of Control: The Performance of Algorithms and the Question of Governance," *Drama Rev.*, vol. 63, no. 4, pp. 121–138, 2019.

[37] G. Paine, "Gesture and Morphology in Laptop Music Performance," in *The Oxford Handbook of Computer Music*, R. T. Dean, Ed. Oxford: Oxford University Press, 2009, pp. 214–232.

[38] R. F. Paloutzian, R. J. Seitz, and H.-F. Angel, "The Processes of Believing and Communicating with the Unseen," in *Human Interaction with the Divine, the Sacred, and the Deceased Psychological, Scientific, and Theological Perspectives*, T. G. Plante and G. E. Schwartz, Eds. New York: Routledge, 2022, pp. 213–133.

[39] J. Parker-Starbuck, *Cyborg Theatre: Corporeal/Technological Intersections in Multimedia Performance*. Houndmills: Palgrave Macmillan, 2011.

[40] S. Parry, *Science in Performance: Theatre and the Politics of Engagement*. Manchester: Manchester University Press, 2020.

[41] M. Pevere, "About," 2022. [Online]. Available: https://www.margheritapevere.com/about/. [Accessed: 15-Jun-2022].

[42] M. Pevere, "Recalcitrant by Nature: Qeering Death through Biological Art Practice," *Whatever. A Transdiscipl. J. Queer Theor. Stud.*, vol. 4, pp. 645–651, 2021.

[43] A. Pizzo, "Performing/Watching Artificial Intelligence On Stage," *Skenè J. Theatr. Drama Stud.*, vol. 7, no. 1, pp. 91–110, 2021.

[44] A. Pizzo, V. Lombardo, and R. Damiano, "Algorithms and Interoperability between Drama and Artificial Intelligence," *Drama Rev.*, vol. 63, no. 4, pp. 14–32, 2019.

[45] A. Rouvroy, "La Gouvernementalité Algorithmique: Radicalisation et Stratégie Immunitaire du Capitalisme et du Néolibéralisme?"," *La Deleuziana*, vol. 3, pp. 30–36, 2016.

[46] C. Salter, *Entangled: Technology and the transformation of performance*. Cambridge: MIT Press, 2010.

[47] M. Schuilenburg and R. Peeters, Eds., *The Algorithmic Society: Technology, Power, and Knowledge*. New York: Routledge, 2020.

[48] J. Walshe, "Ein Körper ist kein Klavier: Editorial zur Diskussion über die „Neue Disziplin"," *Musik-Texte*, vol. 149, pp. 3–5, 2016.

[49] K. Warwick, *Artificial Intelligence: The Basics*. Oxford: Routledge, 2012.

[50] J. Zylinska, *AI Art: Machine Visions and Warped Dreams*. London: Open Humanities Press, 2020.

THE PRAGUE EXPERIMENT: THE SOUND OF IMAGINATION

Bleiz Macsen Del Sette
Conservatorio di Milano
b.delsette@qmul.ac.uk

Simone Gregis
Conservatorio di Milano
simone.gregis@gmail.com

Stefano Mancuso
Conservatorio di Milano
stefanomancuso88@gmail.com

Sylviane Sapir
Conservatorio di Milano
sylviane.sapir@conspaganini.it

ABSTRACT

The following article describes the design and implementation of the sound elements of *The Prague Experiment*, a performative installation on the generative mechanisms of human imagination. The installation was created for the 2019 edition of the Prague Quadrennial, an international event focused on the role of scenography in the performing arts and architecture. The "sound of imagination" was created by sonifying the visitors' biosignals through a *musical automaton* capable of producing a unique two-minute audio instance for each performance. The article aims to present the architecture of the device and the technical and aesthetic-artistic choices that influenced its design and implementation, reflecting on the artistic and educational results achieved.

1. THE PRAGUE EXPERIMENT

The Prague Experiment (TPE) is an interdisciplinary project born from the collaboration between the Scenography School of the Brera Academy and the Electronic Music School of the Milan Conservatory with the aim of producing a performative installation on the theme of imagination. The result of this collaboration, which took place during the academic year 2018-2019, was presented at the 50th edition of the Prague Quadrennial [1] held in June 2019. During the event, the installation won the *Award for the Best Imagination Students' Exhibition*.

1.1 Genesis and concept

TPE is the result of over a year of design work carried out within the educational activities of the Brera Academy and the Milan Conservatory and was developed following the thematic constraints proposed by the 2019 Prague Quadrennial (PQ19): *Imagination*, *Transformation*, and *Memory*.

The themes of PQ19 sparked a broad debate among the students, leading to the choice of subscribing to Milo Rau's "Ghent Manifesto" [1]. Therefore, TPE chooses to root its performance on breaking the fourth wall and the performative dictates of the *Theatre of the Real* [2], of which Rau is a spokesperson in his manifesto. The performance no longer aims to "represent reality" but rather to make "the representation reality".

To implement TPE, a "research center" is created: the *Imaginometric Society* [2] (IS), consisting of the students and teachers participating in the project.

1.2 Performance and organisation

TPE puts into action the efforts of the IS, which through the means of a "scientific experiment" aims to carry out its*mission*: to demonstrate that imagination is what makes us human. The key concept from which all the key elements of the installation were developed, from the scenography to the costumes to the performative act, has been thus focused.

Thanks to the adherence to the *Ghent Manifesto*, the design act itself and the different actionable roles come together within the performative act. Every person involved in the project is called to take part in the performance, covering a role within the IS.

The *Project Fellows* (PF), students of the Brera Academy, are tasked with designing and implementing the performance and its scenography. The *Sonic Fellows* (SF), students from the Conservatory of Milan, are responsible for creating and operating the device capable of "capturing imagination" and transforming it into music. The *Communication Fellows* (CF), coming from various backgrounds, are instead responsible for managing the IS website and documenting all activities carried out.

1.2.1 Performative space and experimental protocol

The experiment of the IS takes place in the *Imaginometer*, a device created by the PF and outlined in Figure 1. The experiment aims to "measure" the imagination of each visitor in a 5x5 m^2 space inspired by industrial clean-rooms, following a very precise protocol:

- The performance itinerary begins in the *Check In* area with the acquisition of the visitor's biosignals.
- The visitor then moves into the *Performance* area, where they will station for two minutes. During the stay, they will listen to the **unique** sound track created using the data acquired previously. At the

[1] https://www.pq.cz/

[2] https://www.imaginometricsociety.com/

Copyright: ©2022 Bleiz Macsen Del Sette et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

same time, five performers stimulated by an "imaginogenic act" [3] , interact with the visitor through a performative improvisation.

- The visitor concludes their passage in the *Imaginometer* by stopping at the *Check Out* area where their biosignals are again acquired. At the end of the itinerary, they are given a postcard with an identification code that allows them to access, through the IS website, the auditory memory of their experiment.

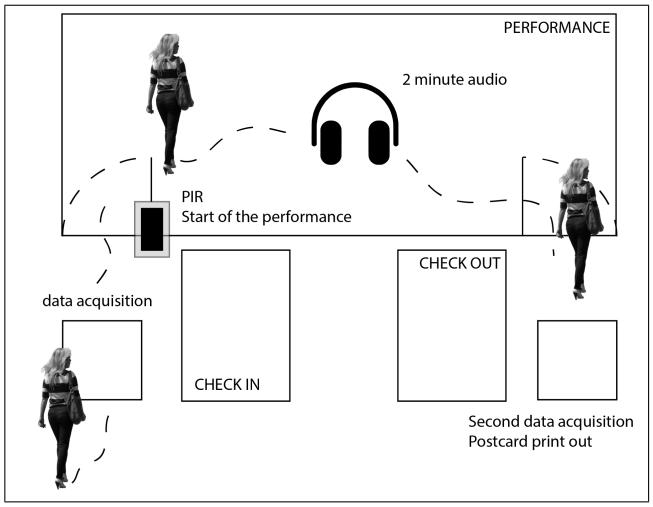


Figure 1. Diagram of the *Imaginometer* of TPE

2. THE SOUND OF IMAGINATION

[Imagination is] what makes us human, and it can only exist in the relationship between human beings. It multiplies in the act of reflecting on itself, as an entity that cannot be reproduced, but only recreated each time [4]

The principles underlying the artistic manifesto of IS have strongly influenced the design of the sound elements of the project, from the acquisition of biosignals to their sonification. "The sound of imagination", a unique soundtrack that each participant can listen to during the experiment, is created through a process that combines deterministic elements, stemming directly from the biosignals acquired for each visitor, with random elements introduced as part of the compositional process.

The goal is to preserve both the uniqueness of the musical result and its imperfect reproducibility, thus following the key principles on the execution of the performative act in a space that is at the same time co-temporal but unique in its reproductions contained in the *Ghent Manifesto*.

2.1 Sound materials and temporality

Most of the sound materials used in the sonification process come from environmental recordings made at the venues where the IS has carried out its previous experiments. These sound materials have been collected starting from the earliest preparatory iterations of the performance and reworked during the sonification process, contributing to the creation of a collective memory of the performative act. Recordings are made both by SF and by a new performative figure, the *Human Data Recorder* (HDR).

The link between imagination and memory is another central theme of TPE; each visitor has the opportunity to listen to a *sonic memory* of their own experiment through the IS website. Each memory is linked to the visitor's biosignals acquired during the actual performance but is influenced again by the random components contained in the composition, resulting in a new iteration of the visitor's "imaginogenic memory".

It is clear, therefore, how TPE creates a temporal narrative that, unlike many modern performative installations based on digital technologies, does not require real-time interactions between the data acquisition and sound production processes. The sound elements of TPE therefore comprises three fundamental and distinct aspects: the acquisition of the visitor's biosignals, the creation of a soundscape constructed from fragments of the ambient recordings through the sonification of the acquired data, and finally, the diffusion of the generated audio track through a bone conduction device.

The generative process of the "sound of imagination" has been conceived in such a way as to create an intimate relationship between the subject of the experiment and their imagination, giving ample space for their interactions with the performers.

2.2 Use of biosignals in artistic performance

In the late 1960s and early 1970s, the first artist to apply the process of sonification to the electrical signals produced by the human body was Alvin Lucier with his iconic work "Music for Solo performer", a seminal work not only in the author's career but in the history of electronic music [3].

Lucier's attempt to sonify brain waves can be framed within what is now called *Brain-Computer Music Interfacing*, which in turn is part of the larger field of *Human-Computer Interaction* that has deep roots in the artistic field. In the decades following "Music for Solo performer", the initial simple *audification* process of the electroencephalogram (*biofeedback*) has evolved into a complex analysis and processing of the signal itself (*biocontrol*) [4] acting as both a sound generator and a control signal [5].

Many biosignals have found a place in the world of performative art, especially with the introduction of wearable interfaces and increasingly less invasive monitoring systems [6]. Interfaces such as BioMuse [7] and its more recent iterations like the OpenSource EAVI project [8] have modified and evolved the relationship between performative art and the body, introducing the concept of *embodied musical interaction* [9].

Therefore, TPE is part of the growing interest in using the biological signals coming from our bodies in musical and installation practices, offering a further stimulus for artistic reflection, moving away from the dominant aspect of

[3] A phrase randomly extracted from a database of textual fragments concerning imagination, written both by members of the IS and by visitors

[4] https://www.imaginometricsociety.com/what-we-do-2/

real-time performance and the use of the body as a musical instrument.

2.3 The paradigm of the *biosoundscape*

The sonification process of TPE is devoid of interaction; signal acquisition and sonification occur in "deferred-time" and can also be reproduced an indefinite number of times through its *sonic memory*. TPE therefore differs from the real-time sonification metaphors that characterize much of contemporary artistic production related to biosignals, using a new paradigm: the *biosoundscape*.

The software created by the SF to generate the "sound of imagination", called the *musical automaton*, generates a layered audio track by modifying the sound materials collected by the HDR and synthesizing new ones through data extracted from the visitor's biosignals. The result can be compared to an "imaginogenic" soundscape, constantly in motion, strongly characterized by the sounds that compose it, but with unstable relationships between its elements and an unpredictable temporal development. For this reason, the ever-changing character of the sound track can be associated with the concept of "metabole" [5] [10].

Despite TPE's soundscape being entirely "imaginary", it can be easily described according to the key principles listed by Schafer [11] to define a soundscape. The "sound of imagination" is a landscape that contains tonic sounds, signals, and sound footprints, and its function is also to connect imagination to memories: a collective memory enclosed in the recorded audio materials and an individual memory hidden in the biosignals of each participant.

The bond between the elements of the Schaferian soundscape, the biosignals coming from the visitors' bodies, and the ever-changing character of the sound track thus contribute to creating the paradigm of the *biosoundscape* that characterizes the production of each of the audio tracks and subsequent *sonic memories* generated by the *musical automaton*.

2.4 TPE listening modes

So far, we have described the characteristics and design of the "sound of imagination" in its compositional sense, without focusing on the performative aspect contained in the implementation of TPE. In fact, the title of this paragraph does not refer to the four listening modes theorized by Schaeffer [12], but rather to the actual listening channels that the visitor can use to "understand" - this time in a Schaefferian sense - the soundscape of TPE.

During the performative act, the visitor is immersed in a soundscape composed of two separate parts:

- The audio track that represents their personal "sound of imagination"
- The actual sonic environment of the *Imaginometer* during their permanence in the *Performance* space.

[5] A perceptual sound effect that describes the unstable and metamorphic relationships between the elements that make up a sound cluster

The first part is transmitted through bone conduction headphones, that use the vibration of the temporal bone to conduct sound directly to the listener's inner ear [13]. The second part can therefore be perceived through the air (tympanic) pathway, since the listener's auditory channels are not obstructed. While this situation promotes interaction between the visitor and the performers, it also creates an "augmented" perceptual state that the visitor must deal with. Listening via bone conduction, in addition to being a completely new experience for many visitors, increases the multi-sensory aspect of listening, particularly the tactile aspect. Perceiving the vibrations of the skin and temporal bone induces a very intimate relationship with sound, and if the visitor is already surprised by a sonic representation of their imagination, they are even more surprised by this new form of listening.

Surprise generates emotions, thoughts, imagination, and while the visitor perceives sound through new and unexplored channels, in a form that could be defined as "augmented listening", they can more or less consciously select sound materials. This process can be attributed to the concept of "synecdoche", whereby the listener, in a complex sound environment, has the ability to select sound materials by valuing one or the other element at their choice [10], a process that complements the concept of "metabole" previously described.

In this way, the visitor constructs, and therefore partially composes, their own personal itinerary in the soundscape of their imagination. The integrity of TPE's soundscape is therefore placed in a space of co-creation with the visitor, in which listening makes it possible to participate in the composition of a unique soundscape, immersed in the performative space and ready to change with each iteration, conveyed through multimodal channels not only in its "production" phase but also and especially in its "reproduction" phase.

3. ARCHITECTURE OF THE SYSTEM

During the development process, great attention was given to the role that digital technologies and sound had to play within the collective performance of TPE. From the earliest stages of development, a non-invasive presence of technological devices was privileged in order to facilitate interaction with performers and the surrounding space. For example, the choice to use a microcomputer such as the Raspberry Pi or a bone conduction playback device was dictated not only by the needs of the sonification process, but also by the practicality and portability of the system itself, as well as the scenic requirements of the installation. The technologies used had to be as unobtrusive as possible, preferably hidden from the public, thus excluding the use of monitors, keyboards, and more "classical" control interfaces.

3.1 Hardware

In the setup for the Prague Quadrennial, three Raspberry Pi 3 B+ were used [6], along with two Arduino Uno

[6] https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/

Rev3 [7] for the biosignal acquisition, one router, one thermal paper printer [8], one external hard-drive, two AfterShokz Trekz Titanium bone-conduction headphones [9], one LED matrix, two DIY acquisition devices (Figure 3) equipped with three bioelectrical sensors (see 3.1.1) and two DIY control interfaces equipped with the necessary hardware components to control the Raspberry Pi in *headless* mode (see 3.1.2).

The Raspberry Pi boards run Raspbian Stretch Full 2019-04-08 (Debian 9.8), and the use of Arduino was necessary to manage the analog signal of the sensors, acting as an analog-digital converter and sending the data directly to Raspberry via USB. The limited battery life of the headphones (about 5 hours at full usage) forced the use of two pairs per day of work in order to ensure the continuity of the performance.

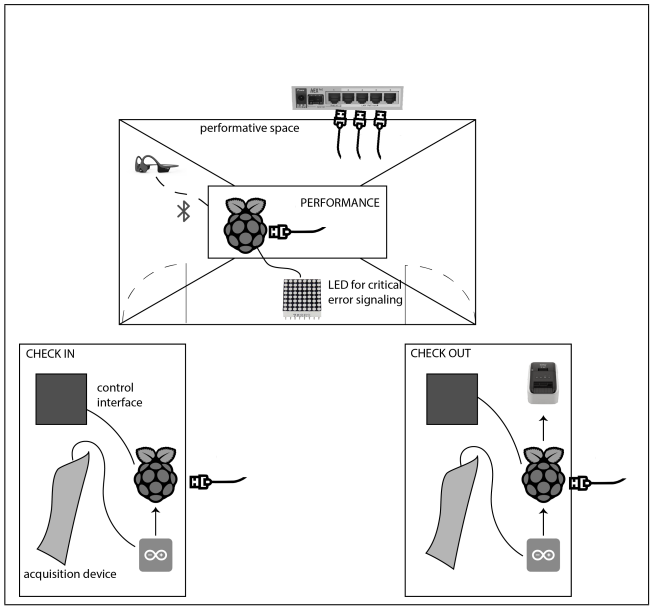


Figure 2. TPE hardware configuration. With respect to the diagram in Figure 1, connections and allocation of the hardware devices used in the three spaces *Check in*, *Performance* and *Check out*

3.1.1 Sensors

Three types of biosignals were chosen for acquisition: heart rate variability (HRV), electrodermal activity (EDA), and electromyogram (EMG) (Table 1).

HRV is often used in the field of affective computing for its link to emotions and stress, both in medical and artistic fields [14]. The sensor used for measurement is the Heart-Rate Sensor SKU SEN0203 [10], which detects volumetric changes in peripheral blood circulation.

EDA, which manifests itself with a variation in skin conductance, is sometimes also used in the medical field, despite its volatility, for emotion recognition purposes, thanks to its close link to arousal [15]. The Groove GSR Sensor [11] was used for measurement, which allows the detection of EDA through two electrodes placed on the fingers of one hand, measuring the variation of a low-voltage current applied between them.

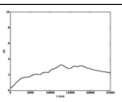
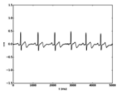
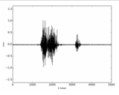
Signal	Profile	Bandwidth	Range
EDA		0.01 - 1 Hz	6 dB
HRV		0.01 - 250 Hz	60 dB
EMG		10 - 450 Hz	60 dB

Table 1. Biosignals used during TPE e their main characteristics

EMG, which describes the electrical activity of muscle fibers, can also be measured in the absence of movement and presents both a high-frequency and a low-frequency component from which it is possible to derive its envelope. This same envelope has often been used in musical and performative applications [16]. The MyoWare Muscle Sensor [12] was used to detect it; through three conductive electrodes, the sensor amplifies and measures the electrical phenomena resulting from muscle activity.

3.1.2 Control and Acquisition interfaces

The choice not to use monitors and keyboards made it necessary to develop a hardware interface with buttons and LEDs that would allow control of the main actions of the *Check-In* and *Check-Out* spaces: data acquisition, performance start-up, and error management. To signal the visitor's entrance into the performance space, a PIR (Passive InfraRed) sensor was used to automatically detect the visitor's passage from the *Check-In* space to the *Performance* space.

Furthermore, a specific device was developed to facilitate the data acquisition process for each visitor. In Figure 3 we can see an initial version of the neoprene glove used during TPE; in its final version, the HRV and EDA sensors were positioned on the thumb, middle and ring fingers, respectively, while the EMG sensor was placed on the forearm muscles. Adopting a DIY approach and not relying on other hardware already on the market, such as BiTalino or iCubeX, allowed us to create a practical and adaptable acquisition system for the numerous scenic and technical needs, without the need to integrate proprietary software for data acquisition and considerably reducing the budget investment.

3.2 Workflow

The main operation related to the *Check In* and *Check Out* spaces is the acquisition and storage of the data from the

[7] https://store.arduino.cc/arduino-uno-rev3
[8] https://bit.ly/2OutDqm
[9] https://docs.rs-online.com/61d1/0900766b815b9609.pdf
[10] https://media.digikey.com/pdf/Data%20Sheets/DFRobot%20PDFs/SEN0203_Web.pdf
[11] https://www.mouser.com/catalog/specsheets/Seeed_101020052.pdf
[12] https://cdn.sparkfun.com/datasheets/Sensors/Biometric/MyowareUserManualAT-04-001.pdf

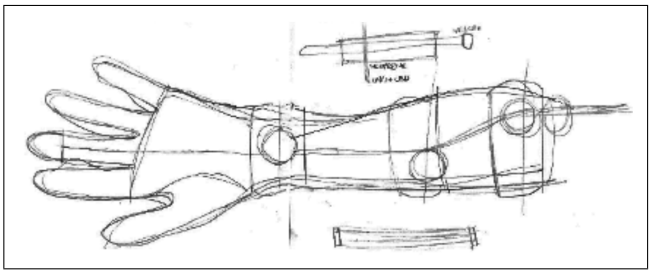


Figure 3. Paper prototype of the biosignal acquisition device. In this initial version, the HRV sensor was positioned on the visitor's wrist.

three biosignals detected for each visitor. In the *Performance* space, the *biosoundscape* is generated and played. Each space is equipped with its own Raspberry Pi, and the three microcomputers communicate and interact through a Python environment developed by us.

The data acquisition device is connected to the Arduino board, which sends the data to Raspberry Pi via USB at a baud rate of 230400 (∼1000Hz). At the start of the acquisition process, Raspberry Pi and Arduino are synchronized through a handshake operation to ensure constant and synchronous reading of the serial port. Data acquisition lasts 30 seconds and as soon as it is completed, the data is written to a text file and stored on a shared hard-drive on the network between the three spaces. When the visitor enters the performative space, their data is read to produce the *biosoundscape*, which is then played through the bone-conduction headphones. The audio engine chosen for the sonification is CSound [13] ; to use CSound within our Python environment, we referred to the ctcsound library [14] developed by François Pinot. The data collected for each visitor is subsequently manually transferred at the end of each day to the IS server, from which it can be reused to produce the *sonic memories*.

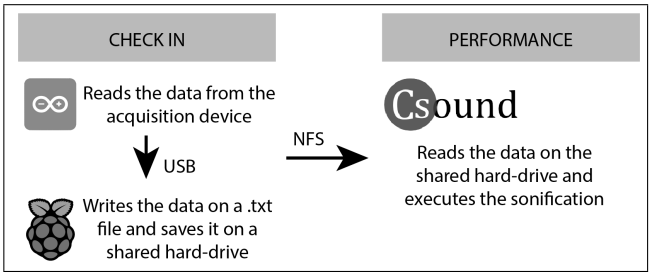


Figure 4. Description of the data flow between the data acquisition station in the *Check-in* space and the *Performance* space

3.3 Network and communication

To implement TPE it was necessary to build a communication system and synchronize the three Raspberry Pi used; the microcomputers were then plugged into a local network through an Ethernet router. The network presents itself with a star topology with the Rasberry Pi in the *Performance* space at its center, forwarding and receiving the messages through predefined static IP addresses. In addition, a remote control network was created via VNC [15] , and a laptop was additionally connected to the router in order to access the graphical interface of the three Raspberry Pis in case of blocks or malfunctions. The NFS [16] protocol, which allows reading and sharing of files and folders within a Client/Server architecture, was used for sharing and reading the recorded biosignal data. Real-time communication was handled using the OSC [17] and SSH [18] protocols, to manage the interaction between the three Raspberry Pi with the aim of dictating the execution times, communicating key information, and managing any errors during each phase of the performance (Figure 5). The selection of the correct pair of headphones to be used, identified by their unique MAC Address and corresponding ALSA profile, was carried out manually using a remote control system via OSC.

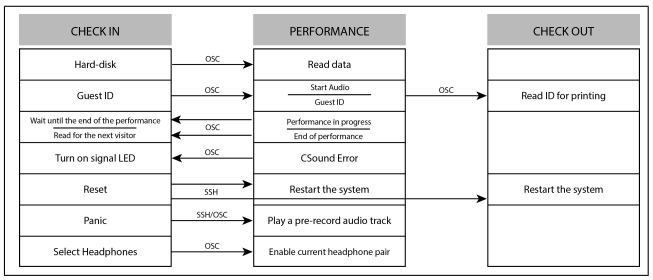


Figure 5. Exemplified diagram of the message flow and relevant transmission protocols between the three spaces of TPE

4. SONIFICATION

The sonification process used to generate the "sound of imagination" for each visitor, is the core of the artistic research of the project. The *biosoundscape* is indeed an artistic interpretation (not a scientific one) of the acquired biosignals; the goal is to stimulate the visitor's imagination and generate a unique experience both during and in the memory of their passage in the *Imaginometer*.

TPE does not require any kind of real-time interaction, but it was instead necessary to explore different solutions for using biosignals in order to create a musically relevant audio track. This process falls within the "parameter-based sonification" [17] definition, a sonification strategy rooted in the compositional tradition of Computer Music that easily allows experimenting with different solutions of arbitrary relationships between the parameter space of the data to be sonified and that of the process that generates the audio.

4.1 Artistic choices

The concept behind the sonification of the TPE is the generation of a *biosoundscape*, a 2-minute musical piece composed of various sound layers whose execution parameters

[13] https://csound.com/
[14] https://csound.com/docs/ctcsound/ctcsound-API.html
[15] Virtual Network Communication
[16] Network File System
[17] Open Sound Control
[18] Secure Shell

are algorithmically managed based on the previously measured biosignals, creating a unique sound experience for each visitor.

The structure of the piece, as well as the proposed sound materials, remain unchanged throughout the whole Experiment. This guarantees an optimal formal result on the musical level, which is only varied and modulated by the biosignal data. The ethereal and evocative nature of the *biosoundscape* experienced by the visitor accompanies the improvisational performance concurrently taking place in the same moment. The goal is to provoke the visitor's imagination and create points of connection between the performers' actions and the soundtrack that serves as the backdrop.

The absence of conscious control by the visitor over their own biosignals, combined with the natural errors of the acquisition device, clashes with the very notion of control taken in a strict sense. The biosignals must therefore be conditioned to prevent them from behaving as mere generators of random data. However, randomness is still preserved and intentionally introduced through processes of random selection of pre-recorded sound materials with a decidedly "Cagean" aesthetic: the *biosoundscape* as another version of the *Imaginary Landscape* [18].

4.2 Biosignals data processing

The acquisition devices designed for non-medical applications are not very precise and usually very sensitive to disturbances caused by low-quality electrodes, poor contact between electrodes and skin, defective cables, sudden movements, etc. This implies the need to perform a number of pre-processing operations before being able to use them.

In TPE's *musical automaton*, in order to use the biosignal data, some mathematical functions are applied that return a series of indexes, as well as convert the data themselves into tabulated functions. The data are normalized and filtered to smooth out their profile and eliminate any instantaneous measurement errors. The data processing is programmed in CSound and happens before the synthesis process (Figure 6).

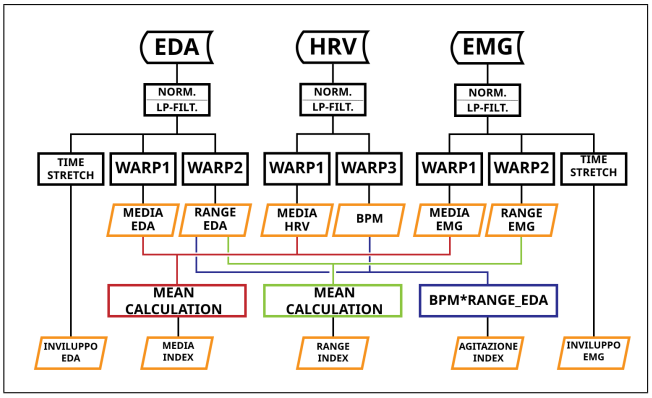


Figure 6. Exemplifying diagram of the data processing for each biosignal

After a first signal normalization, the application of a low-pass filter, and an additional transformation through non-linear functions (WARP), the signal averages are extracted. Moreover, to align the data acquisition duration (30 seconds) with that of the *biosoundscape* (2 minutes), a temporal scaling operation (TIME STRETCH) is applied. Nine indices and two envelopes are obtained from this data processing, and then used in the mapping procedure.

4.3 CSound Instruments and Mapping

The biosoundscape is composed of four main sound layers and three sound effects. Therefore, a CSound *orchestra* is created consisting of four instruments for the sound synthesis, plus three additional instruments to manage the sound effects and one instrument for the final mixing. Each synthesis instrument follows a fixed amplitude envelope, in order to dynamically fit with the other instruments and at the same time respect a predefined compositional structure, defining the shape of the *biosoundscape* (Figure 7).

4.3.1 First layer

The first sound layer is generated by a FOF [19] with three waveforms, whose fundamental frequency is derived directly from the *BPM* parameter (extracted from the heart rate frequency) and the *media_EMG* index. The amplitude follows an envelope linked to the *HRV Table* and *media_HRV* index, while the *range_EDA* and *media_EDA* indexes manage the spatialisation. In this section, a synthesised sound of a vocal nature is generated, but at the same time the synthesis process tries to mimic the visitor's heartbeat.

4.3.2 Second layer

The second sound layer is made up of an environmental audio sample randomly chosen from the pre-recorded samples (see section 2.1) to which an exaggerated time stretching effect is applied, followed by a bandpass filtering. The temporal expansion factor remains constant throughout the piece, but depends on the *media_index* parameter, while the cut frequency of the filter varies according to an LFO controlled by the *agitazione_index* parameter.

4.3.3 Third layer

The third instrument is a sort of pseudo-random Rompler [20] that plays short evocative samples using a system of phased pairs to generate a counterpoint. Although the choice of the sound samples is randomised, the playback frequency and density of the sounds are directly proportional to the *media_index* and *agitazione_index* parameters, and inversely proportional to the *BPM* index. The spatialization is managed through two LFOs whose phase modulation is connected to *media_EDA* and *media_HRV* values.

4.3.4 Fourth layer

The last sound layer is a granulator that uses as sound source another of the pre-recorded ambient sounds described in 2.1. Unlike the time stretching factor in the se-

[19] Fonction d'Onde Formantique, a voice synthesis technique developed in the 1980s [19]

[20] A combination of the words *ROM* and *sampler*, a type of synthesizer that plays pre-loaded sampled sounds

cond layer (4.3.2), which remains constant through its entire duration, this instrument is extremely variable and its numerous parameters are articulately controlled using the whole range of biosignal data leveraging a *many-to-many* mapping approach. In particular, the granulator uses *inviluppo_EDA* and *inviluppo_EMG* for density and amplitude modulations.

4.3.5 Sound effects and mixing

Of the three sound effects present in the piece, one is a long feedback filtered delay dedicated exclusively to the FOF layer (4.3.1); this effect has a purely compositional value and does not depend on any of the biosignal related parameters. The other two sound effects consist of a delay with a Moog Ladder filter [20], whose delay time is determined by the *media_HRV* index, and a Room reverb again with a purely compositional value in which the *size* parameter is randomized at each execution.

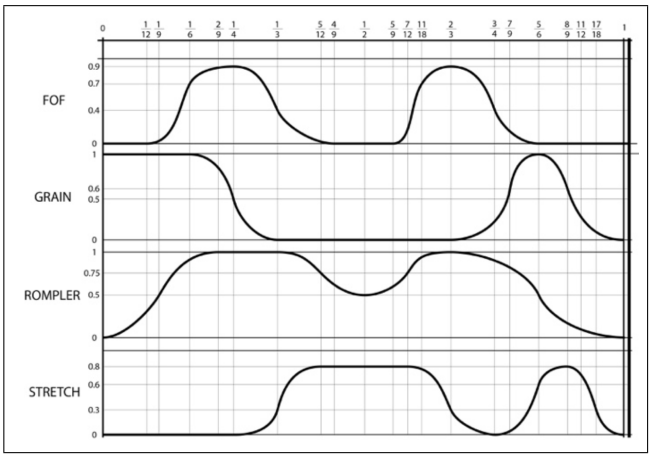


Figure 7. Depiction of the temporal evolution of the four sound layers in the *biosoundscape*

4.3.6 Compositional workflow

The compositional work was handled by various students from the Milan Conservatory, each of whom integrated and enriched the code with their own skills, while maintaining a common compositional guideline. The various blocks of code were then assembled in such a way as to be able to function organically thanks to an interlocking structure. Each sound layer follows a dynamic line able to bring out one element at a time, avoiding confusing overlaps and allowing the piece to have its own structural identity independent of the measured biosignals.

4.4 The *sonic memories*

Thanks to the use of the sound recordings made by the HDR during the IS experiments, the *musical automaton* feeds on the collective memory gathered from one experiment to another, and comes alive with the information extracted from the biosignals of the individual visitors, which in contrast represents its individual memory. Collective memory and individual memory together constitute TPE's *sonic memories*, that are stored in a database on the IS website.

To access their own *sonic memory*, each visitor can use their unique *visitor code*, generated in the *Check Out* space and delivered to them via a postcard at the end of their journey in the *Imaginometer*. The code allows access to a CSound server hosted on the IS website, which runs the same *musical automaton* used during TPE, retrieving the original biosignal data of the visitor from the database. This produces a *biosoundscape* similar to that of the experiment but with a very different listening and human experience.

The implementation of the server and the management of the database with the files containing the environmental recordings and the visitors' biosignals were carried out by Salvatore Zappalà and the CF.

5. CONCLUSIONS

TPE has been very successful with the public during PQ19; the participation of almost 2000 people in 10 days, and the recognition of an award from the organising committee, encouraged the IS to further develop the project and replicate again the "imaginogenic" experiment in a new venue.

In December 2019, at the Brera Academy of Fine Arts in Milan, The Milan Experiment (TME), an event promoted by Zona K [21] , once again attracted almost 1000 people in 5 days. In its Milan edition, the experiment involved almost a hundred performers of different age groups and was further supported by the inclusion of musicians from the Conservatory of Milan who regularly performed "Les Sonneries de Cantenac" by F. Lévy [22] to attract visitors. The employed hardware was also revised to simplify the acquisition process and improve the quality of the detected biosignals. Furthermore, the headless control device was removed and replaced with a more standard and easy to manage GUI controlled via monitor and keyboard, thanks to the new context and some scenography changes. A third iteration of the Experiment following TME was scheduled for 2020, but the outbreak of the COVID-19 pandemic forced their cancellation.

On the IS website, it is possible to view photos taken during the performances in Prague and Milan, which help to capture the spirit of this artistic and performative work, the reactions of the public and the large scenography implemented. To illustrate the result of the sonification, it is possible to listen to an example by accessing the sound database (https://www.imaginometricsociety.com/audio/) and entering the code **1906030105**.

The success of TPE and TME validates the intuition and artistic idea underlying the project and demonstrates once again how it is possible to implement a virtuous cycle that starts from an artistic work to stimulate more structured scientific research and then return to artistic applications, all within a multidisciplinary educational path.

[21] https://www.zonak.it/milan-experiment/
[22] https://www.fabienlevy.net/compositions-2/2-to-4-musicians/

6. ACKNOWLEDGMENT

We first want to thank the Institutions that financed and supported the TPE project, as well as the students of the Electronic Music School of the Conservatory who provided their fundamental contribution, in alphabetical order: Carlo Maria Amadio, Michael Barteloni, Giovanni Benda, Fabrizio Fogagnolo, Maurizio Gazzola, Edoardo Morelli, Matteo Tandelli, and Matteo Tomasetti. A heartfelt thanks also goes to Franco Ripa di Meana and all the students at the Brera Academy of Fine Arts with whom we had the pleasure of working and collaborating together. Finally, a special thanks goes to all the musicians and students of the Conservatory who participated in the second edition of the project in Milan with openness and enthusiasm.

7. REFERENCES

[1] M. Rau, "Ghent Manifesto," *Theater*, vol. 51, pp. 20–23, May 2021.

[2] L. De Vos, "A Tale of Truth: The Use of Brechtian Strategies for Non-Brechtian Artistic Purposes in Milo Rau's Theater of the Real," *Theater*, vol. 51, pp. 74–83, May 2021.

[3] V. Straebel and W. Thoben, "Alvin Lucier's *Music for Solo Performer* : Experimental music beyond sonification," *Organised Sound*, vol. 19, pp. 17–29, Apr. 2014.

[4] M. Ortiz, N. Coghlan, J. Jaimovich, and B. Knapp, "Biosignal-driven Art: Beyond biofeedback," *Ideas Sonica/Sonic Ideas*, vol. 3, no. 2, 2011.

[5] D. A. H. Williams and E. R. Miranda, "BCI for Music Making: Then, Now, and Next," in *Brain–Computer Interfaces Handbook*, pp. 193–206, CRC Press.

[6] R. Mcgee, Y.-Y. Fan, and R. Ali, "BioRhythm: a Biologically-inspired Audio-Visual Installation," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 80–83, 2011.

[7] H. S. Lusted and R. B. Knapp, "Biomuse: Musical performance generated by human bioelectric signals," *The Journal of the Acoustical Society of America*, vol. 84, pp. S179–S179, Nov. 1988. Publisher: Acoustical Society of America.

[8] B. Di Donato, A. Tanaka, M. Zbyszyński, and M. Klang, "EAVI EMG board," 2019.

[9] A. Tanaka and M. Donnarumma, "The Body as Musical Instrument," in *The Oxford Handbook of Music and the Body* (Y. Kim and S. L. Gilman, eds.), pp. 78–96, Oxford University Press, Oct. 2019.

[10] J. F. Augoyard and H. Torgue, *À l'écoute de l'environnement: répertoire des effets sonores*. Editions Parenthèses, 1995.

[11] R. M. Schafer, *The Soundscape: Our Sonic Environment and the Tuning of the World*. Simon and Schuster, Oct. 1993.

[12] P. Schaeffer, *Treatise on Musical Objects: An Essay Across Disciplines*. Univ of California Press, July 2017.

[13] S. Stenfelt, "Acoustic and Physiologic Aspects of Bone Conduction Hearing," *Advances in Oto-Rhino-Laryngology*, vol. 71, pp. 10–21, 2011.

[14] D. Borthakur, V. Grace, P. Batchelor, H. Dubey, and K. Mankodiya, "Fuzzy C-Means Clustering and Sonification of HRV Features," Sept. 2019. arXiv:1908.07107.

[15] D. Williams, C. Y. Wu, V. J. Hodge, D. Murphy, and P. I. Cowling, "A psychometric evaluation of emotional responses to horror music," p. 10, 2019.

[16] A. Tanaka and R. Knapp, "Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing," Jan. 2002.

[17] T. Hermann, A. Hunt, and J. Neuhoff, *The Sonification Handbook*. Jan. 2011.

[18] J. Cage, "Imaginary Landscapes," *Leonardo*, vol. 35, no. 5, pp. 498–498, 2002. Publisher: The MIT Press.

[19] X. Rodet, Y. Potard, and J.-B. Barrière, "The CHANT Project: From the Synthesis of the Singing Voice to Synthesis in General," *Computer Music Journal*, vol. 8, no. 3, pp. 15–31, 1984. Publisher: The MIT Press.

[20] R. A. Moog, "A Voltage-Controlled Low-Pass High-Pass Filter for Audio Signal Processing," Audio Engineering Society, Oct. 1965.

Organum pineale. Design and implementation of a site-responsive sound installation

S. Bassanese, G. Diana, M. Marson, L. Martone
SMET/Conservatorio di Torino
stefano.bassanese@conservatoriotorino.eu
guglielmo.s.diana@gmail.com
matteomarsonmusic@gmail.com
lucamartone1990@gmail.com

Andrea Valle
CIRMA-StudiUm/Università di Torino
andrea.valle@unito.it

ABSTRACT

The paper describes the sound installation *Organum pineale*, that has been commissioned by the National Museum of Cinema of Torino in the occasion of the "#FacceEmozioni" exhibition, dedicated to the relation between face expressiveness and emotions. The installation is site-responsive and exploits the peculiar acoustic features of a part of the Museum building, the so-called "Orecchia". On one hand, the installation was meant to provide a sonic *analogon* of the main theme of the exhibition, in particular in relation to Descartes' conceptualization of passions. On the other, its design aimed at emphasizing the unusual acoustic features of the Orecchia, a visually empty space, but very rich on the aural side. *Organum pineale* features a 28 + 4.1. audio system, delivering sound materials originating from a selection of films, that have been tagged in relation to Descartes' six main passions. The overall sound process involves spatialized audio and a generative feedback system exploiting the room acoustics.

1. INTRODUCTION

In the occasion of the "#FacceEmozioni" exhibition (Figure 1, see [1]), the National Museum of Cinema in Torino commissioned a sound installation to the Electronic Music School (SMET), Conservatorio di Musica "G.Verdi" of Torino, and CIRMA (Interdepartmental Research Center on Multimedia and Audiovideo), Università di Torino [2]. The Museum is located in the monumental building of Mole Antonelliana: a vast space, originally meant as a synagogue, that was the tallest brick building in Europe at the time of its construction (167.5 metres) [1]. The Museum is well-known for its spectacular multimedia setting.

The commission by the Museum was meant to exploit a specific, and unusual, room in the Mole, the so-called "Orecchia" (literally: ear). Placed at the side of the main building, the Orecchia serves as a connective space to move



Figure 1. #FacceEmozioni playbill.

inside the Mole and it crosses two stairways, that are open on the same side of the room, allowing to enter/exit the space (Figure 2). As a consequence, the Orecchia is a sort of autonomous body, characterized by almost cubic proportions. The cubic shape does not include the ceiling, as the room is surmounted by a dome vault. Walls are made of stone and concrete, with complex decorations in plaster (Figure 9). A specific feature of the Orecchia is that it offers a visual paradox: it is completely empty, but at the same time it gives the audience the possibility to observe from a wide angle the roof landscape of Torino: in fact, two of its sides are completely occupied by windows (Figure 3). So, properly in the Orecchia there is a lot to see, nothing of which is related to the contents provided by the Museum. From another perceptual perspective –i.e., the aural one– the geometry of the Orecchia results in a peculiar acoustic behavior, the main feature of which is a clearly delineated harmonic resonance. In other term, the room is a strongly "tuned" space, a sounding board with an almost instrument-like nature. The commission by the Museum was to create a sound installation inside the Orecchia in relation to the contents of the "#FacceEmozioni" exhibition. The latter was centered around physiognomy, that is, the relation between human facial expressions and passions, from 1500 to the contemporaneity. In relation to such a theme, the main issue at stake in the sound installation design was how to convey a sound *analogon* of passions in

[1] http://www.museocinema.it/en/museum-and-ma-prolo-foundation/mole-antonelliana

Copyright: ©2022 S. Bassanese, G. Diana, M. Marson, L. Martone et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



Figure 2. Entrance/exit stairways in the Orecchia.



Figure 3. Audience inside the Orecchia.

relation to their facial expressions as codified from Renaissance up to emoticons while taking into account the history of cinema.

2. BACKGROUND

The sound installation, titled *Organum pineale*, is conceived around the acoustics of the Orecchia and aims at setting up an audible scene of the representation of passions in relation to cinema. In these terms, there are five main conceptual keywords at the root of the project: space, instrument, passions, representation and listening.

2.1 Space

The Orecchia is a transitional space, as the audience is passing through it, properly without performing any action related to the Museum's contents. As noted by [3] these transitional spaces, exactly because of their supposed neutrality and the freedom they ensure to the audience (as visitors are not supposed to be bound to the contents) represent an opportunity for sound art –also on the political side– for "installing a space in another existing space, both physically and sensorially, and metaphysically and mentally (as a space for reflection and emotion) – an interior space in an exterior space" [3, p.108]. *Organum pineale* thus aims at opening an inner sonic theatre inside the external space of the Orecchia.

2.2 Instrument

As previously introduced, the Orecchia has a particular acoustic behavior as it is properly a highly resonant space, centered around three main frequencies and their various harmonic components (see Section 3). Rather than attenuating this harmonic feature, our purpose has been to emphasize it. The main concept was to treat the room as a gigantic resonator, i.e. to let the passions "resonate" inside it. An audio processing chain has been developed by means of various filter banks in order to emphasize such a "tuning of the space" (to cite Murray Schafer [4]). Another feature has involved sound distribution: the room has been equipped with a system of 28 speakers in order to let the listener clearly localize the sound sources in the space. In this sense, *Organum pineale* is "site-responsive". The latter term has been proposed by [5] in order to point out how in electro-acoustic practices, in order to deliver sound in a specific space, the sound artist has to continuously take into account the way the space responds to the sound itself, thus closing a feedback loop rather than assuming it as a fixed framework like in visual arts ("site-specific").

2.3 Passions

One of the fundamental texts of the modern history of passions is the essay by Descartes, *Les passions de l'Ame* ("The Passions of the Soul") [6, 7]. Here, Descartes introduces his classifications of six fundamental passions (wonder, love, hate, desire, joy, sadness). This basic organization stands at the core of the whole exhibition "#FaccEmozioni". As the Orecchia is the terminal place of the exhibition route, and therefore acts as its conclusion and recapitulation, even if in an acoustic form, the installation has its pivot exactly in the six main Cartesian passions. Descartes is also at the origin of the title of the work, *Organum pineale*. Notoriously, the pineal gland was conceived by Descartes as the place where *res cogitans* and *res extensa* were able to find a contact/exchange, that is, where soul and matter meet. Pineal gland thus plays a pivotal role in the establishing of passions. Descartes' body organ is the first reference for the title, but indeed a second reference comes from the meaning of "organ" as "musical instrument". It can be said that the Orecchia (both as the ear and –at the same time– as the architectural space) becomes the "organ" of the passions.

2.4 Representation

Passions, even if they have a biological substrate [7], are channeled into cultures, defining variable stereotypical typologies [8]. As shown by exhibition, they are obviously a crucial element of investigation for the performing actor, whose job mostly consists in bringing them on stage (see the many contributions in [1]). From Renaissance to 19th century, physiognomy has been an attempt to construct a code of passions, and, apart from its scientific results, has prompted a wide corpus of reflections on the actor's expressiveness, including both the face and the voice. In short, one can say that there is therefore an audible physiognomy of the passions that is conveyed by the voice [9,

10]. This is particularly interesting in relation to National Museum of Cinema, as the latter plays a pivotal role in witnessing, through the history of cinema, how the actor's craft has constantly changed over more than one century in relation to different cultural sensitivities.

2.5 Listening

The fruition space is called "ear" because of its shape, and, as mentioned, this shape also translates into a very particular acoustic response. In *Organum pineale* this architectural ear, so to say, is then thought as a theater of listening. As a consequence, technological interventions have been devised so to be as invisible as possible, so that the attention of the audience would not shift from listening to the sound to looking to the technical gear. Perceptually, the Museum has a notorious spectacular setting that, even if multimedia, is strongly focused on vision. Thus, while entering the Orecchia for *Organum pineale* as the final step of the exhibition, the visitor, until then mostly a viewer, is meant to be converted into a listener. In this perspective, our installation is meant as an example of (rare) purely acoustic art [3].

3. DEVELOPMENT

For the realization of *Organum Pineale* we have built an acoustic system powered by a ring of 28 little speakers, plus 4 other loudspeakers and a subwoofer, and two microphones. The design of the installation has been constantly monitored on the physical space, so that properly it can be said that it was not only implemented but also conceived in details as site-responsive (Figure 4). A first analytical step has required to analyze the resonance frequencies of the space. To do so, we proceeded empirically by generating a sweeping glissando signal in the range between 35 and 80 Hz, delivered it into the Orecchia through the subwoofer. Then, we recorded it and analyzed the signal resulting from the sinusoidal source plus the contribution of the room. We found three main resonant frequencies corresponding to 58 Hz, 65.5 Hz and 68.5 Hz, the latter having the highest intensity. Apart from fundamental low frequencies, other frequencies resulted as harmonic components of the same fundamental tones. The two loudspeaker groups perform different tasks. The little speakers are meant to convey film voice samples, that act as an acoustic foreground. The four loudspeakers are responsible instead for a sonic background that is generated in real time from a feedback system and includes samples from the film soundtracks (see next section), while the subwoofer is traditionally deputed to convey low frequencies. All sound processing is managed algorithmically by means of a Max/MSP patch installed on a computer with a dedicated 32-channel sound card (Antelope Orion 32), that accounts for the 28 + 4.1 setup (the subwoofer signal is derived from the 4 background channels). Seven amplifiers have been used for small passive speakers, each responsible for 4 channels. The 4.1 background system is active, and thus autonomously amplified. The whole technical setup is listed in Figure 5, while Figure 6 shows

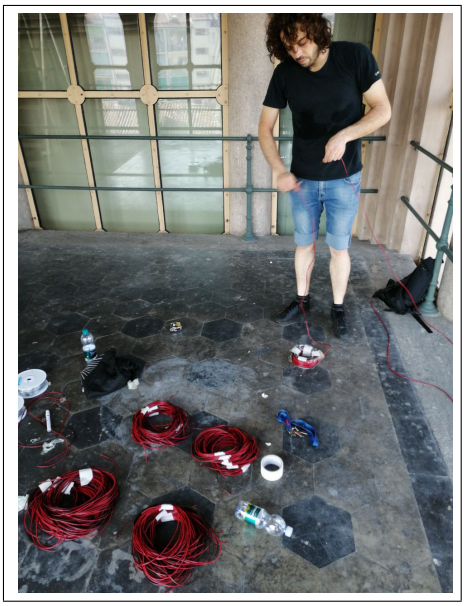


Figure 4. Cutting cables inside the Orecchia

Category	Product	Scheme short name	Quantity
Computer	Apple Mini-Mac	C	1
Audio Interface	Antelope Orion 32	AD-DA	1
Amplifiers	Auna C300 4 Channels 100W RMS	A	7
Mixer	Soundcraft EPM8	MX	1
Microphones	Movo LV8-D Omni-directional	M	2
Speakers	Array of 7 mini-speaker 3W	SA	4(7)
Loud Speakers	Alto TS 310	LS	4
Subwoofer	Alto TS 312s	SW	1

Figure 5. Technical setup.

the overall information flow. Details of Figure 6 are discussed throughout the next sections, the hardware setup is discussed later in relation to Figure 13.

3.1 Content organization

We selected a corpus of sound fragments (voice and soundtracks), extracted from films that we considered particularly significant with respect to the fundamental Cartesian passions. In cooperation with the scientific board of the Museum, 18 films have been taken into account, coming from various countries, genres, historical periods. For each, i) a soundtrack sample has been selected (see 3.2), and ii) various vocal samples have been extracted (for a total of 54) (see 3.3). The film sources are shown in Figure 7. As it can be seen, all samples have been tagged according to one of the six passions.

Samples are used as source materials to be processed and distributed in the space. Each passion is represented by a specific "tuning", both in spectral and spatialization terms, to be applied by the DSP algorithms. In particular, while vocal samples have been left unmodified in order to be clearly understandable, soundtracks samples have been processed by "tuning" them in relation to the Orecchia's

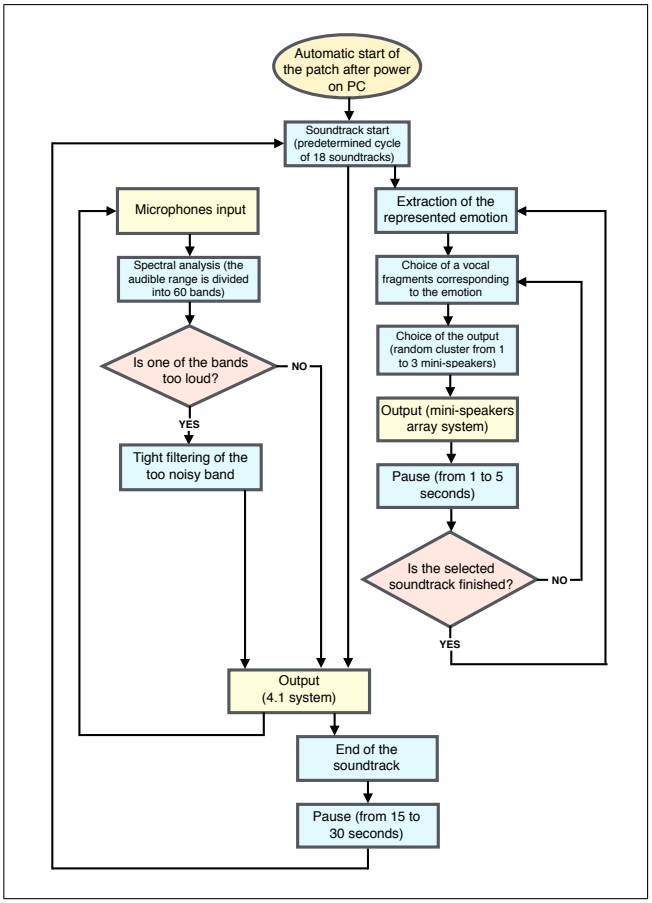


Figure 6. Logical flow of *Organum Pineale*'s control program.

Title	Director	Year	Emotion	Dialogues (54)
Breakfast at Tiffany's	Blake Edwards	1961	Love	2
When Harry Met Sally. . .	Rob Reiner	1989	Love	3
Gone with the Wind	Victor Fleming	1939	Love	6
The Bridges of Madison County	Clint Eastwood	1995	Desire	2
Call Me by Your Name	Luca Guadagnino	2017	Desire	3
Il racconto dei racconti	Matteo Garrone	2015	Desire	5
Forrest Gump	Robert Zemeckis	1994	Joy	2
It's a Wonderful Life	Frank Capra	1946	Joy	3
Risate di Gioia	Mario Monicelli	1960	Joy	3
The Wizard of Oz	Victor Fleming	1939	Wonder	2
Interstellar	Christopher Nolan	2014	Wonder	3
Totoro	Hayao Miyazaki	1988	Wonder	2
Kedmah	Amos Gitai	2002	Hate	3
Riget	Lars von Trier	1994	Hate	1
Seven	David Fincher	1995	Hate	2
Bonjour tristesse	Otto Preminger	1958	Sadness	4
Bram Stoker's Dracula	Francis Ford Coppola	1992	Sadness	5
Mission to Mars	Brian De Palma	2000	Sadness	3

Figure 7. Film selection and annotation.



Figure 8. PVC coulisse for the speakers.



Figure 9. A detail of the Ring.

resonant frequencies, with the aim of exploiting what is typically considered a negative acoustic feature of a listening environment (i.e. strong resonances). In *Organum pineale* all the sound materials are played cyclically in a pre-established order: this "installation score" proved to be empirically effective, and is the only element belonging to a classic composition side in the whole process.

3.2 The Ring

The Ring is deputed to the spatialization of vocal samples, that are left unprocessed for sake of understandability. It is made up of 28 small speakers. The latter are 8 ohm passive speakers with a frequency range between 200 Hz and 15 kHz, and 3 W of power. They are amplified by 4-channel car amplifiers, a cheap but effective solution, given that no hi-fi quality can be provided by the speaker themselves.

As can be seen in Figure 8, each speaker is expanded with a PVC cone with a little *coulisse*: this allows the speaker to project the sound more directionally and to tune it according to the space resonance. PVC has been instru-



Figure 10. Speakers mounted on aluminum supports.

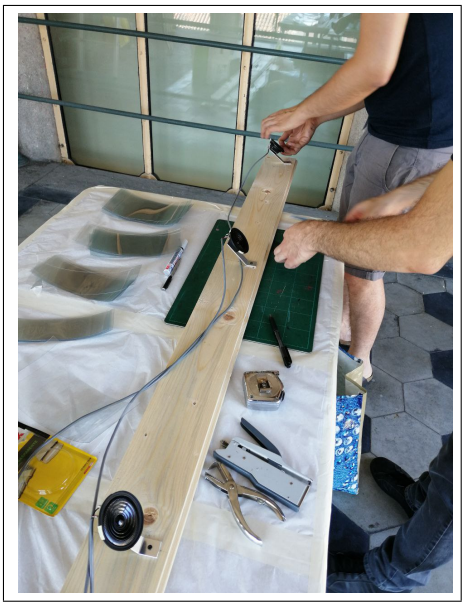


Figure 11. Fine tuning of speaker orientation.

mental in allowing an easy tuning as fine adjustments by trial and error could be done by cutting the sheets. This low-cost solution is inspired by electronic hacking tradition in electronic music [11] and has made possible to overcome budget issues related to dedicated, more expensive solution [12], such as Acouspade [2]. The 28 speakers are organized in a "Ring". The Ring has been placed on the ridge of the room, with 7 speakers for each side, organized into an array (Figure 9). These small speakers, partially hidden by plaster decorations, are the only visible equipment in the installation. Speakers have been placed on wooden boards, and hooked by means of aluminum bars (Figure 10). As aluminum is flexibile, it has been easy to fold the supports in order to set for each speaker the desired angle for audio projection (Figure 11). This fine tuning process was fundamental to obtain an effective spatialization. Overall control for the Ring is provided by the Max/MSP patch that determines in real time which vocal sample to use and where to place it in the space (Figure 6). While performing the passion sequence, the system selects a passion out of the six from Descartes. For the passion's defined duration (depending on available audio materials) the relative subset from the collection of vocal samples are loaded. An algorithm chooses randomly which speaker has to deliver the selected sample: the choice is constrained by an urn containing 28 values (representing each speaker), a matrix that indicates which speakers are free at the desired time (this available for sound delivery), and an index that allows the sound to be emitted by one or multiple speakers in a row. By means of this strategy, passing through the Orecchia allows the listener to ear voices coming from different angles of the vault. Every speaker is oriented toward a different point of the vault's surface. As the vault acts as an approximate parabola, the resulting effect for the listener is hearing voices speaking all around.

[2] ultrasonic-audio.com/

3.3 The feedback system

In *Organum pineale* a feedback system provides a sonic background for the whole installation. Feedback systems have a long tradition in electroacoustic music, be it analog or digital, as ways to generate complex, evolving sonic textures by exploiting the systems themselves [13]. From the hardware point of view, our system is made of four loudspeakers and a sub-woofer, plus two microphones, and it is completely invisible to the audience. The space of the Orecchia is substantially held upon columns, and the pavement has long, thin openings at its sides. Because of this, it was possible to hide the 4.1 channel diffusion under the pavement, in a service space, so that it can be heard without being seen. In order to emphasize the resonant features of the Orecchia, we designed a system that creates a controlled feedback from the room. The sources to be delivered in the Orecchia (over 4.1 channels) by this subsystem are samples from soundtracks. These same 4 audio signals are captured by two (hidden) omnidirectional microphones, together with vocal samples and the audio reactions by the audience in the Orecchia. The two signals resulting from the omnidirectional microphones feed an algorithm consisting of a bank of bandpass, adjacent filters with self-compensating amplitudes. As the output sound is fed back in the Orecchia by means of the 4.1 system, the result is a generative acoustic feedback system. Each filter instance is tuned around a region between two frequencies separated by an interval of semitone, with a slightly overlapping bandwidth: then, each frequency band is excited by a chorus algorithm that changes the amplitude envelope, each band's amplitude is then forced to oscillate between a minimum and a maximum amplitude value with a specific control frequency. As a result, the system dynamically increases and decreases each frequency amplitude over time. It must be observed that the microphones are placed in the Orecchia, so they also capture all the sound contributions from the audience passing by. In this sense,



Figure 12. The Ring.

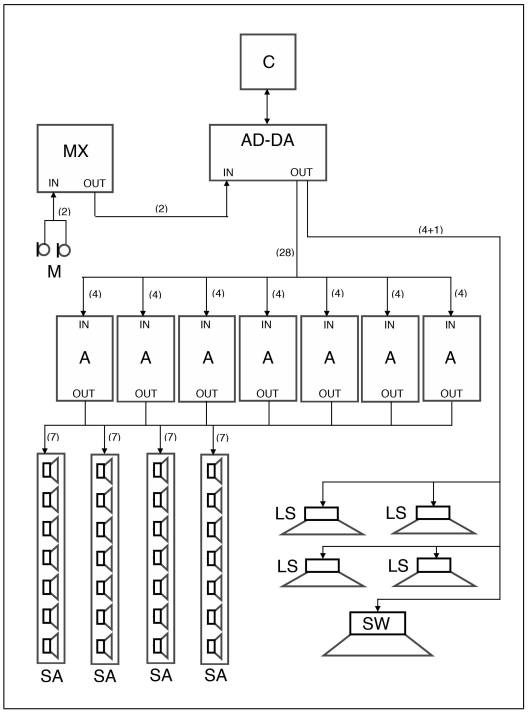


Figure 13. The hardware electroacoustic chain.

Organum pineale can also be thought as an audio-based interactive system.

The overall hardware setup is shown in Figure 13 (letters refers to Figure 5). To sum up, the two microphones (M) are connected to a mixer (MX), and their signals are sent to the Antelope sound card (AD-DA). Input signals are processed by a Mac Mini (C, computer) and multichannel audio is sent back together with the soundtracks and vocal samples to the soundcard. Then, 28 channels are amplified by 7 amplifiers (A) and delivered to the Ring (4x7 SA, speaker arrays). The feedback background is sent to the 4.1 system under the floor (LS, loudspeaker and SW, subwoofer). The microphones M capture the sound in the room, closing the feedback loop.

4. ROBUSTNESS

Organum pineale was open for six months [3] inside the Mole Antonelliana's Orecchia. Such a long opening has required us to create a very robust setup, as it would have been impossible to ensure an extensive technical support. For the same reason, the setup interface (startup and control) had to be very user-friendly, as it should be operated by the Museum staff. As a consequence, we designed *Organum Pineale*'s setup so that it could be switched on/off by simply pressing sequentially two switches in two electric plugs: then, with a two-minute delay an autolaunch program runs the application and no other procedure is required. Budget constraints have prompted us to create our own directional sound system, by means of small, cheap speakers surmounted by the PVC coulisses. The system has proven to be very effective, thus providing a low cost technical solution than can be fruitfully replicated. During the six months in which the installation was running, only few small problems have been reported: some coulisses fell down, and a small speaker had to be replaced. Robustness has proven to be a crucial feature in relation to two aspects: first, on the audience side, it was important that *Organum pineale* kept to seamlessly deliver sound output without interruptions; second, on the maintenance side, it was mandatory that interventions were to be reduced to the minimum, as any operation on the installation was to be supervised by the Museum staff for safety reasons, thus making it much more complicated to perform.

[3] From July 17th 2019 to January 6th 2020.

5. THE SITE, RESPONDING

The sonic result of *Organum pineale* is an immersive space, with people bathing in a continuum of sound streams. Informal reports we gathered from the audience (we hadn't the resources for collecting structured interviews) during the six months of its opening describes a sort of acoustic bubble punctuated by the incoming voices that acted as figurative, communicative pivots. A consistent part of the audience (youngsters mostly) discovered the feedback systems sending back their own processed voices and started experimenting with sound production to "test" the system. This feature underlines the "participatory" nature of the installation, even if it was not properly conceived as the result of a collective workshop [14]. Yet, another participatory feature lies in the shared design and implementation between teachers and students. As one of our team members wrote: "Working in this project has been a very deep experience in design and managing a complex artwork, thanks to our teachers who helped us understand the emerging issues, allowing us to pragmatically deal with them by exploring various possibilities; we have been able to learn how to work with institutions such as the National Museum of Cinema, handle technical material and experiment with it in relation to aesthetic choices". From reports coming from the Museum, the installation was very well received. In any case, an interesting issue has emerged. As we have discussed, the Museum is perceptually focused on vision. Moreover, the whole experience of the exhibition, culminating in the hall of the Mole, might be tiring for viewers. After traversing the hall, in order to exit the Museum, the audience had to go downstairs, thus passing through of the Orecchia. This final passage, so different from the previous aspects of the exhibition, has sometimes



Figure 14. Audience inside *Organum Pineale*.

proven to be difficult for the public, already tired or not willing to focus exclusively on hearing, as if they were under a sort of visual inertia because of the previous exhibition experience. Finally, the large windows offer a wide landscape, thus attracting the viewers' attention (Figure 14). This aspect seems to underline the more intrinsically difficult focus on aural perception than on visual one.

6. CONCLUSIONS AND FUTURE WORK

Organum pineale has proven successful both in terms of reception and in terms of shared working methodology. While strictly related by design to the space of the Orecchia, it has prompted more general design questions (purely aural art, site-specific construction, participatory feedback system) that we plan to address in the future. In fact, while it is not possible to strictly replicate *Organum pineale*, the format (i.e. exploitation of the space as an instrument to deliver some contents related to the space itself or to some specific theme) has proven very interesting, and the low cost technological solution very effective. In this sense, we plan to implement such a format in a possibly more contemplative, dedicated environment, be it cultural (e.g. church) or natural (e.g. cave).

7. CREDITS AND ACKNOWLEDGEMENTS

The concept for *Organum pineale* has been developed by Andrea Valle and Stefano Bassanese. Design and realization has been carried out by Guglielmo Diana, Matteo Marson and Luca Martone. *Organum pineale* is a production by SMET- Scuola di Musica Elettronica del Conservatorio di Torino in collaboration with CIRMA/StudiUm - Università degli Studi di Torino. Grazia Paganelli, Museo Nazionale del Cinema, has provided the film selection and annotation. We would like to thank Museo Nazionale del Cinema, and in particular curator Donata Pesenti Compagnoni for her constant support.

8. REFERENCES

[1] D. Pesenti Compagnoni and S. Arcagni, *Il volto delle emozioni. Dalla fisiognomica agli emoji*. Silvana Editoriale, 2019.

[2] A. Valle, *Il volto delle emozioni. Dalla fisiognomica agli emoji*, ch. Organum pineale, pp. 110–111. Silvana Editoriale, 2020.

[3] G. Klein, "Site-sounds: On strategies of sound art in public space," *Organised Sound*, vol. 14, p. 101, mar 2009.

[4] R. Murray Schafer, *The Tuning of the World*. New York: Knopf, 1977.

[5] L. Hayes, "From site-specific to site-responsive: Sound art performances as participatory milieu," *Organised Sound*, vol. 22, pp. 83–92, mar 2017.

[6] R. Descartes, *The Passions of the Soul*. Hackett Pub. Co, 1989.

[7] A. R. Damasio, *Descartes' Error. Emotion, reason and the human brain.* New York: Avon Books, 1994.

[8] A. J. Greimas and J. Fontanille, *Sémiotique des Passions*. Seuil, 1991.

[9] I. Fónagy, *La vive voix*. Payot, 1983.

[10] A. Valle, *Il volto delle emozioni. Dalla fisiognomica agli emoji*, ch. Fisognomica udibile, pp. 105–109. Silvana Editoriale, 2020.

[11] N. Collins, *Handmade Electronic Music. The art of hardware hacking*. New York–London: Routledge, 2006.

[12] J. Reis, "Short overview in parametric loudspeakers array technology and its implications in spatialization in electronic music," in *Proceedings of the International Computer Music Conference 2016*, (Utrecht), pp. 241–248, ICMC, 2016.

[13] D. Sanfilippo and A. Valle, "Feedback systems: An analytical framework," *Computer Music Journal*, vol. 37, pp. 12–27, 2013.

[14] A. Tanaka and A. Parkinson, *The Routledge Reseach Companion to Electronic Music*, ch. The Problems with Participation, pp. 156–177. London and New York: Routledge, 2018.

Interactive Sonification of Expressive Gesture: the *DanzArTe - Emotion Wellbeing Technology* Project

Andrea Cera
Casa Paganini-InfoMus, DIBRIS, University of Genova
andreawax@yahoo.it

Corrado Canepa
Casa Paganini-InfoMus, DIBRIS, University of Genova
corrado.canepa@dibris.unige.it

Nicola Ferrari
Casa Paganini-InfoMus, DIRAAS, University of Genova
nicola.ferrari@edu.unige.it

Alberto Pilotto
Dept. Geriatric Care, Orthogeriatrics and Rehabilitation, E.O. Galliera Hospital, Genova
Dept. of Interdisciplinary Medicine, "Aldo Moro", University of Bari
alberto.pilotto@galliera.it

Paolo Coletta
Casa Paganini-InfoMus, DIBRIS University of Genova
paolo.coletta@unige.it

Simone Ghisio
Casa Paganini-InfoMus, DIBRIS University of Genova
simoneghisio@gmail.com

Antonio Camurri
Casa Paganini-InfoMus, DIBRIS University of Genova
antonio.camurri@unige.it

Abstract

This paper presents an interactive sonification model of full-body human expressive movement, adopted in the *DanzArTe - Emotion Wellbeing Technology* project. The model is grounded on our Slow Mood and Aesthetic Resonance paradigms as well as on several principles and guidelines discussed in the paper. DanzArTe is a treatment protocol and an interactive system mainly for older people at risk of frailty, grounded on the active experience and real-time processing of visual arts content (classical religious paintings, image manipulation and 3D modelling) and interactive sonification, to support full-body physical activity and cognitive exercise of memory training.

1. INTRODUCTION

In this paper we present the interactive sonification model we designed for multimodal systems supporting the active experience of audiovisual cultural content, for cultural welfare and non-verbal affective applications. Multimodal non-verbal full-body affective and social emotions computing paradigms enable active group experience of content, i.e., they enable both individuals and groups of users to mould, augment, and recreate audiovisual content by means of their full-body individual as well as joint expressive movement. Active experience leverages on embodiment and enaction, on affect and emotion, on social interaction among users, stimulated to interact with the artistic audiovisual content and with each other.

1.1 The Slow Mood paradigm

Paul Valéry in "The Necessity of Poetry" [33] wrote that today every pleasure is destroyed by haste and nullified by velocity. Focusing on cultural heritage and visual art fruition, we propose the Slow Mood paradigm: we paradoxically aim to exploit the exciting rapidity of connections typical of contemporary technology, to rediscover the long time necessary for a fruitful social relation with contents belonging to our artistic history. In technology-enhanced museum experiences, as well as in movement therapy and rehabilitation, the Slow Mood paradigm can support the design of innovative, non-intrusive, and anticipatory/predictive forms of interaction, to enable an intimate dialogue with the artwork, in a slow pace, inducing optimal flow [6]. We propose a process articulated around a series of sessions, not in everyday exposure to technology as, for example, in Slow Computing or Calm Technology approaches [32]: the temporal scales involved in this investigation of aesthetic resonance and in specific contexts of use characterized by liminality [3] are one of the main differences with respect to the above mentioned approaches.

Copyright: © 2022 Cera et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1.2 Interactive sonification of human movement expressive qualities

In [19] an interactive system used in hospital to support physical exercises for Parkinson patients based on *aesthetically resonant* interactive soundscape and visual feedback is proposed: the real-time analysis of fluidity in full-body movement was used to enable the patient to paint using his/her own body on the surrounding walls and to create naturalistic soundscapes only when moving fluidly, using the body like a brush. Usually frustrating physical rehabilitation exercises are here transformed into a sort of artistic installation, enabling to interactively mould artistic content (visual and auditory). A similar approach is adopted in the exergames designed for the cognitive and motoric treatment and rehabilitation of children in the ARIEL-Augmented Rehabilitation Joint Lab with the G.Gaslini children hospital (ariel.unige.it). In the European funded project Horizon 2020 ICT DANCE (dance.dibris.unige.it), Casa Paganini-InfoMus collaborated with Virgilio Sieni in a series of public events (Atlante del Gesto Genova, https://it-

it.facebook.com/atlantedelgestoGenova/) involving over 120 non-dancer citizens participating to the (re)discovery of their own full-body expressive qualities concluding in a series of public collective performances [11]. In DANCE, the real-time measures and sonification of *fragility* and *lightness* expressive qualities of the full-body movement of one of the participants (the leader) had the goal to induce the other participants to entrain and express the same qualities in their full-body joint movements. In [23] interactive sonification is used in rehabilitation from chronic pain disease, and the interactive auditory stimuli have the role to enhance fluid and safe movements, and to "distract" the patient from possible chronic disease problems. In the EU Horizon 2020 ICT WHOLODANCE project (wholodance.eu), we developed interactive sonification of human movement qualities to support the learning process of movement in dance.

2. CASE STUDY: DANZARTE

Our Slow Mood and Aesthetic Resonance paradigms are the main pillars of our trans-disciplinary interaction design process of systems for the real-time full-body multimodal, non-verbal affective and social interaction and communication. DanzArTe, a project on cultural wellness and social inclusion, is the case study considered in this paper to explain our sonification model. DanzArTe aims to support an active experience, physical contact with a classical visual art painting, as a collective, autonomous amateur practice, enabling an embodied experience capable of exercising memory and physical movement emotionally. We focus on two main categories of users: older people living in residencies with "good" residual cognitive and physical capacities, and healthy older people (over 65yo) at risk of frailty.

Visual art content has the role of a guide to execute full-body physical activity, and is a reference for memory training. Interactive sonification provides cues to *nudge* [30] participants' movement towards the desired movement qualities: fluidity and entrainment among each other. Small groups of 4 participants in turn interact (individually or in dyads) with the visual content, which is displayed by a video projection visible only to one or two of the group: the other participants give their back to the video projection and imitate the active participants in their affective interaction with the visual stimuli.

DanzArTe includes a treatment protocol and an interactive system integrating real-time manipulation of visual art and interactive sonification content with the role of guide and stimulus for the users, as a sort of surrogate of artistic or clinical personnel. A supervisor of the system has the role of mediator and facilitator of the emotional experience among the participants.

The DanzArTe system is a scalable, low-cost, technological platform to support real time interaction with visual content (classical visual art paintings) and interactive sonification. A small group of up to 4 participants is engaged in a form of slow and safe movement sequence, inspired by a slow, gradual observation and discovery of a classical painting and related interactive sonification of the movement qualities of fluidity [1] and entrainment [13] in joint actions.

The choice of the paintings, of the movement and many other interaction design decisions are out of the scope of this paper. We only mention here our choice of classical religious paintings (by the Italian painter Luca Cambiaso)., familiar to older people, not inducing possible psychological issues in participants. Paintings were chosen also in relation to the affective movement they evoke, to avoid works characterized by static or difficult postures. The presence of the painting in the local museums and churches was also considered, to encourage participants to visit cultural institutions in the territory, or to bring such cultural institutions to the residency of older people where the system is adopted.

3. CRITERIA FOR THE DESIGN OF INTERACTIVE SONIFICATION OF HUMAN MOVEMENT

Our model of interactive sonification of human movement qualities is based on cross-modal mappings of movement qualities at different levels of abstraction and temporal scales [24, 25, 27, 28]. The main criteria and guidelines for our model are described in the following:

1. *Multiple Layers Cross-Modal Correspondence between movement and sound* [16, 25]: this is the main component of the sound architecture, and is grounded on previous research [11, 1, 24] identifying a four layers model for representing multimodal signals at different temporal scales: physical, low-, mid-, and high-levels of abstraction.
2. *Sonification integrates dynamically evolving environmental soundscapes following slowly evolving movement features related mainly to the context.* The higher-layers and temporal scales include strategies to modulate slowly varying features of background soundscapes. A sonification model at multiple temporal scales integrates simple sonic events up to complex sound streams and naturalistic simulated dynamically modulated soundscapes and modelling techniques (e.g. [4]).
3. *Slowness and continuity in dynamic curves*: movement fluidity is a mid-level quality [1, 24] characterized by a slowly varying response at a temporal scale in a range of half a second. The sonification of fluidity should reflect such slow pace and contribute to semi-conscious slow and fluid movements. Analogously, impulsiveness, rigidity, fragility are detected, but their mapping is subtractive, they are not sonified.
4. *Low-intrusiveness sonification* [5]: The goal of usual sound design approaches is to evoke explicit meaning, which, from an auditory scene analysis point of view, should be detectable and recognizable as clearly as possible by users. In our approach, the sonification should avoid perturbing the optimal flow

during the experience, remaining at a semi-conscious level: sonification emerges without startling, intrusive, or annoying effects. This constraint influences many features of the sonification model: it raises the need for a background sound stream at a very slow temporal scale, as a stable layer through which sonification events can emerge in a controlled way, keeping under control contrasts, surprises, ambiguities. This states the necessity of carefully controlling the pitched content and reinforces the importance of smooth and slow dynamics. Our approach aims to implicitly or semiconsciously evoke, encourage, nudge, and even elicit certain qualities of movement, including fluidity [11], or synchronization of behavior in a dyad or a small group of users [13]. The approach is grounded on the affordance enabled by the cross-modal mapping of movement qualities to sound. That is, interactive sonification of full-body movement qualities to induce the user(s) to the desired movement qualities: for example, to change the quality of a movement in a physical exercise from a fragile or rigid execution to fluid, continuous; or to induce a dyad performing a joint action to increase their level of synchronization [13].

5. *Subtractive design*: Saliency is created by subtraction, instead of addition of sound. A continuous soundscape, constantly present (slowly evolving following the user movement qualities at different temporal scales) from the beginning to the end of a session, can briefly disappear under certain circumstances, to create a salient moment and elicit attention.

6. *Silence*: a particular case of subtractive design. Music is the only human language including symbols for silence: the rest symbols are musical notation signs indicating the absence of a sound for a given duration. In our sonification architecture both foreground and background tend to be continuous, very soft in presence and supporting the flow of the experience, but punctuated with meaningful silences. Some of these silences happen in the foreground, others happen in the background. The use of an artificial background is also aimed at improving liminality [3] in terms of acoustic comfort, masking unwanted noise (e.g. coming from the external).

7. *Polyphony and orchestration techniques to the sonification of dyads and small groups*: in case of multiple users, or of a joint movement of a human with a virtual humanoid agent, orchestration techniques as well as timbral and evolution of the harmonic content have the objective to achieve clearly perceived distinct sonification streams of separate features, as well as to design a clearly perceived unique sonification of a group feature, e.g. the sonification of the level of entrainment between two users or of a user and a virtual agent. For example, in the DanzArTe project, the sonification of movements of the avatar and of the user have a clearly separated "signature".

4. THE DANZARTE INTERACTIVE COLLECTIVE EXPERIENCE

A room equipped with a video projector, a screen (or white wall) of 2x3m, a computer, a Kinect II and a stereo sound system. A mediator/facilitator (e.g. personnel of a residency for older people, or personnel from the museum or cultural institution) supervises the system and facilitates the social interaction among the participants. The participants have received a preparation about the whole experience, including an introductory session to familiarize with the protocol. The overall DanzArTe protocol is composed by four Sessions: one or two sessions per week, each session has a duration of about one hour. A preparatory session (Session 0) is performed the same day of Session 1. Sessions 1 and 2 are designed with one active participant, while Sessions 3 and 4 have two active participants interacting with each other aiming at synchronizing their behavior (entrainment). Participants in each group exchange their role and all have the possibility to be active participants. Active participants are in front of the video projection, while the others observe, imitate the active participants, and interact with suggestions, comments. In each session, each participant in their turn exchanges the role of active participant. For example, session 1 starts with the video projection of the painting by Luca Cambiaso "The Mother with the Child". The active user observes the painting for a few seconds, then the painting tends to disappear, except for the body of the Mother, which then loses the color, and fades out, replaced by a virtual agent in the same posture. The virtual agent is a 3D model derived by the preparatory drawings of the painter Luca Cambiaso, very essential and geometric. Now the painting has completely disappeared, and the agent starts to move from the posture of the Mother sitting with the Child on her lap, and slowly repeats a first component of the movement evoking the affective gesture of the Mother: the movement of the left arm to embrace the Child on her lap, then the other arm, the head inclination, one leg a little forward. If the participant imitates the agent correctly, the agent disappears and the part of the painting of the body part evoked by the movement starts to appear and gradually gets the original color from repetitions of the correct movement. A wrong quality of the movement (e.g., not fluid) causes the agent to return and show it again. The agent movements are based on motion capture recordings of performances by two dancers, who performed sequences generated from the postures depicted in the paintings, following guidelines from experts in science of movement and geriatrics. After the participant has completed each component of the movement, the agent shows the whole affective gesture, in a sort of final dance with all the parts of the body engaged. The participants starts this crescendo of movement repetitions and this gives life back to the original painting, that receives from the participants the

"energy" in terms of light of small candles that at the end illuminate the whole painting.
After the first sessions in collaboration with the Museo Diocesano di Genova, some feedback from participants were like "I could see the painting with new eyes", or "now when I look at a painting, I see more on it".
The participants are initially instructed (Session 0) to imitate the movements of the agent, and learn that their movement is causing sounds. So they understand that if they start moving correctly, the avatar slowly disappears, and a fragment of the painting emerges, as if it were lit by a candle, corresponding to the part of the body moved by the participant. This part of the painting regains its color by repeating the movement. The participants know that the picture will be "reconquered" according to the way they will be moving, imitating the affective gesture in the painting. It is important to notice that the system does not measure the posture exactly, but it mainly focuses on the following movement features: (i) the correct part of the body; (ii) Fluidity and Quantity of Movement of a part of the body [1]; (iii) level of synchronization among two participants. The interactive storyboard is controlled by the DanzArTe EyesWeb component: each session corresponds to a timed script depending on time and on the behavior of the users, thus determining the evolution of the interaction by comparing the participant's movements with the reference model of movement required at each moment of the storyboard.

5. THE INTERACTIVE SONIFICATION MODEL IN DANZARTE

A DanzArTe multimodal session integrates visual and auditory channels. Visual elements are present and crucial in every moment of a session, and carry two levels of information: the progressive discovery of the details of the original painting supports the narrative structure, guiding the users to the completion of the session. The agent is the guide, and shows the movements to be learned and performed, engaging physical and cognitive (memory) activation. Participants observe the painting, try to understand and imitate the movements of the agent emerging from the painting, remember the sequences of the learned movements, which are performed altogether in the final part of the session (the final dance). The design of the interactive sonification in DanzArTe obeys to three constraints derived from the interaction design criteria mentioned above: (i) to facilitate a state of serene concentration (ii) to create sonic embodied training, enhancing the perception of fluidity and slowness in the perceived and performed movements, or enhancing the synchronization of movement of two participants; (iii) to avoid unnecessary startling, intrusive or annoying effects that would destroy the flow and the liminality of the experience. We avoided the use of sound to communicate logical or verbal information, as well as we avoided the sonification of negative feedback (e.g., sounds of errors). We uniquely focused on the stimulation of positive emotional states, and adopted the subtractive strategy of cancelling sounds as an attentional cue in case of errors. In the following section we describe the sonification model using DanzArTe as a guideline.

5.1 General Sound Design Principles

The overall sound design architecture in DanzArTe is inspired by soundscape-based composition practices, and in particular by two structural functions, defined by Barry Truax as "keynotes" and "signals". The first "provide background continuity", while the second "provide foreground encoded and interpretable information", and "can become a keynote if heard frequently enough" [17]. These two functions are here named "background" and "foreground" for simplicity. The goal of the overall art-driven emotional wellbeing scenario consists of slowly transforming the user's movement quality, which usually starts from an initial hesitant, fragmented series of fragile movements, towards full-body movements characterized by fluidity and synchronization among participants. In a successful session, this slow transformation becomes audible: the "foreground" (fluidity sounds), initially rare and sparse, becomes gradually continuous, until it almost blends with the "background". This effect of fusion of "foreground" and "background" signals that full-body movement achieved fluidity or synchronization among participants.

5.2 The Background Component

The "background" is composed by loop-playback and real-time transformation of two stereo sound-files. At every new participant, the two sound-files are replaced by different ones, for the sake of variability and novelty. The first sound-file is always audible, from the beginning to the end of the session, but it can briefly disappear under certain circumstances, to create a salient moment and elicit attention [4]. The second sound-file slowly appears/disappears in relevant conditions. In accordance with the low-intrusiveness guidelines, these sounds appear and disappear following extremely slow dynamic curves (fades ranging between 5 and 20 seconds). Only in one case (detection of error in an important sequence) these sounds can disappear / appear with a faster fade out/in curve (1000 ms.).

Sound Materials

The sound-files used for the "background" are derived from various materials: field-recordings of natural soundscapes, synthesized winds (Skat-VG library [2]). These materials were treated with the aim of stabilizing their amplitude / timbre profile as much as possible (but stopping before completely destroying their recognisability); in order to render them mutually replaceable; and to place them inside our needed profile

of Spectral Flux (the change in magnitude between frequency bins of a FFT [29]) and Spectral Centroid (the weighted mean of the frequencies in a FFT bin), and average loudness. This is necessary to make sure that they don't collide with the "foreground" sounds, whose Spectral Flux and Spectral Centroid profile is described in a later section. To empirically validate these features we followed two methods: during the process of elaboration, we analyzed the computer's output in real-time, using Skat-VG audio analysis tools in MaxMSP. The second method consisted in analyzing the renderings of the sound-files in Sonic Visualizer, using the "BBC" and "libxtract" libraries. All "background" sound-files range inside the 300/800 Spectral Flux value (with BBC library and a window size 8192, a rhythmic track measures around 10000/20000 – this means that the "background" sound-files are stable, static, without percussive traits), and present a centroid between 1000 and 1500 Hz. (this means that they don't descend in the area which is mainly occupied by the "foreground" sounds). Their overall loudness averages around -32 dB (measured with Adobe Audition).

Slow-breath modulation

The "background" sounds are modulated following a slow cycle, which randomly oscillates between 0,14 Hz. and 0,16 Hz: this simulates a moderately slow breathing cycle of 8,5 to 10 breaths per minute in healthy subjects [15]. In the written instructions to the users/participants, the facilitators are encouraged to focus on this "breathing" in the sound as a target to which to tune one's own breath. This strategy is useful for older people, who tend to have a fast breathing rhythm. The modulation gradually pushes up and down the center frequencies of two resonant filters. The transformed signal is added to the original signal, modulating not only its timbre, but also its amplitude, further contributing to the sensation of "hearing a breathing".

5.3 The Foreground Component

The Foreground is composed of sounds ("sonifications") which accompany specific movements, underlining certain movement qualities. In a DanzArTe session there are two types of performers concurrently at work: human performers (one or two interacting participants, according to the session), and virtual performers (one or two video-projected agents, according to the session). Both types of performers execute movements, and both are sonified. The audio engines which create this real-time sonification are based on a granulation technique [20], applied to specific sound materials (see below). The engines dedicated to the participants are controlled by parameters (kinetic energy and smoothness) coming from the real-time video analysis, performed by an EyesWeb software component analysing input from Kinect II in real time. The engines dedicated to the avatars are controlled by the same set of parameters. Since the avatars' sequences are pre-recorded (and played back as video-files), there is no necessity to perform a CPU-expensive real-time analysis of their movement. The analysis parameters are metadata recorded previously, retrieved in correspondence with the particular frame of the sequence played by the avatar, and sent to the corresponding sonification engines.

Granulators technique

The granulation process used by the sonification engines is a classical technique consisting of using sound-files as raw material to be played back through multiple superposed, phase-shifted and cross-faded loops. It offers control of the portion of the sound-file to be accessed, and it enables pitch transpositions and timbral transformations based on position/transposition randomness. Further, it is possible to stretch the duration of any section of the base sound-file [20]. Each sonification engine is composed of a couple of granulators: a principal one, which uses a fairly long window (5s, i.e., the length of the superposed loops), and transposes the original sound-file down by one octave. A secondary granulator uses a shorter window (1s), without transposition, and its general amplitude is reduced by -10 dB. Both granulators' "heads" (i.e. the indication of the points on the corresponding sound-file to be played back) have the same address, but they apply a small, almost inaudible amount of randomness, both in position and transposition. Both granulators access the same sound-file, whose timbral content is characterized by fluidity and continuity (described in next section), to create a smooth, ceaseless sound. The difference between the two windows' length, and between the transpositions create a delicate effect of evanescence at each time the head moves across the sound-file. Participants' and agents' engines refer to structurally identical sound-files (length, internal subdivision, harmonic structure), but differentiated by timbre and pitch structure (explained in next section), in order to enable the users to better distinguish sounds produced by their own movements from movements produced by the avatars.

Interaction

Participants' data on movement qualities are analyzed in real-time by an EyesWeb module, while the movement of the agents are recorded metadata. Movement features are mapped on the parameters of the granulators according to the following steps:

1) the EyesWeb movement analysis module verifies if the correct quality is achieved (fluidity/smoothness or synchronization of movements in a dyad): this decision determines if the real-time sonification of the users' should be heard, corresponding to encouraging the users to continue with their movement. In case of correct movement quality, the movement data are sent as inputs to the sonification engines;

2) the sonification engine maps the kinetic energy on the output volume, the smoothness (or the degree of synchronization in case of a dyad joint movement) controls the cutoff frequency of a parallel low-pass

filter. Both parameters are treated in order to smooth out any undesired fast peaks or valleys, and to create slow temporal shapes, even at the expense of fast reactivity. The patch offers presets, characterized by different levels of reactivity (reverberant rooms can add a sense of increased slowness, and hence could need a more reactive sonification to balance this effect). Higher-order information, regarding the session's scenario advancement, determines changes in the "head" position, which translates in a new pitch set (i.e. a new harmonic area) being played. In general, when a new movement is proposed by the avatar, a new pitch set is generated as a signal of change of phrasing in movement. When the system is not producing the users' sonification, it outputs the "background" alone, in moments of pause of rest, or it can output the sonification of the avatar, when a new movement has to be shown to the users, or when it is necessary to show again an already seen movement that has not been performed in a satisfactory way. In these two last cases, the avatar sound starts with an additional high-pass filter, initially excluding sounds below 5000 Hz, and reconstituting the full sound in 12 seconds. This is an example of the subtraction strategy previously described.

Timbre, fluidity and stability

The sound-files used by the granulators contain a continuous, smooth, sustained chord. In the same way as for the "background", during the preparation of these sound materials we kept under control the Spectral Flux, which is connected with notions of smoothness and fluidity, as shown in previous experiments [1] [7], and the Spectral Centroid, which is connected with notions of dimension and mass [16]. The starting material for this series of sound-files is created with synthesis patches, mainly realized in Ableton Live, controlled with a real-time analysis patch and cross-checked in non-real time with Sonic Visualiser. The average Spectral Flux of these sound-files is higher than in the "background", but still quite low compared to rhythmic material. The control of Spectral Flux is due to the necessity of communicating a stable, fluid, smooth sound, which is cross-modally linked to smooth and fluid movements. We accepted a slightly higher value of Flux, for the necessity to modulate the sound-files, and create timbral differences between them, internal life, etc. The average Spectral Centroid of the sound-files ranges between 300 and 800 Hz, suggesting a low overall center of gravity (cross-modally associated with mass and stability). Moreover, one of the two granulators transposes the original sound one octave below, further lowering the perceived Centroid, and augmenting this sensation of stability. Another important reason for keeping low-pitched Centroid profiles is related to our low-intrusiveness paradigm. The frequency of the most emergent spectral component of a sound is directly correlated with its intrusiveness (e.g. sounds with higher spectral centroids risk to be the more intrusive) : to avoid intrusiveness it is preferred to use low-pitched sounds when possible [5]. As for the "background", for every new participant the sonification materials are replaced by different ones, for the sake of variability and novelty. Only the sound-files associated with the avatar remain the same, becoming a recognizable "voice" from the "teacher" across the sessions. To differentiate the avatar sound from the user's sounds, the respective chord structure is different. The sound-files for the users are based on harmonies of 3 or 4 notes. The sound-file for the avatar is based on a simpler harmonic template, a bi-chord. This makes it more emotionally neutral, less reassuring, but more adapted for the role of "correcting" teacher, a role which the avatar takes during the sessions. User sounds for sessions with a single active participant and for sessions with two active participants are the same, following the idea that correct movements in both cases are fluid, slow and continuous. After the first session, users learn to associate the type of timbre described in this section with correct movements. In the two-person sessions, users concentrate on being in synchrony with their partner, and "foreground" sounds play an important role in suggesting the right temporality and smoothness of movement.

Pitched materials and annoyance

The nature of this project demands that the sounds carry an information of reassurance and serenity. To do this, after some tests using only natural or abstract sounds (which resulted emotionally too neutral), we decided to make use of pitched material, choosing consonant harmonies which, coupled with smooth and slow dynamic profiles, could evoke a state of low arousal and positive valence. [8, 10, 12, 18] However, the use of pitched material has to be taken with extreme caution, when operating in a low-intrusiveness paradigm: pitches can be easily memorized, and become quickly annoying if heard frequently during a session, creating an unnecessary sense of recognition of musical elements, and distracting users [21, 22]. For this reason we designed a context where the pitched content not only doesn't use melodic profiles, but also slowly changes over time.

The "foreground" sound-files have the same structure: they are divided into 12 sections, each one of which contains a different chord. At the beginning of each section, one of the notes of the chord is abandoned and slowly fades away, while a new note slowly fades in. The logic of this harmonic drifting aims at smoothly cycling through all twelve possible transpositions. The choice of chord-types aims at introducing the minimum possible dissonance, using common notes as pivot to modulate. When these sound-files are played back by the granulators, the smooth transition already present inside the sound-file is added to the "evanescence" effect created by the granulators' settings: the migration from one chord to the other becomes almost unnoticed, and won't distract the users. This effect can be compared to a passage between "clouds" of harmonies,

or to a smooth change of light. When a session starts, the "heads" of the granulators are sent (randomly) to one of the 12 sections. From this moment on, at each time a new internal phase of the session starts (for instance, when a new movement is shown to the users), the "heads" will travel to the next chord section of the sound-file, creating a slow change of harmony. The internal phases of the session change in different ways, following the types of interaction between participant and avatar, and in general last around 15/20 seconds. In this way, the pitched content of DanzArTe constantly varies over time, and never fixates in the same chord. This creates a perpetual variation and modulation, and takes away the risk of annoyance effects arising from the insistence on a same set of pitches.

5.4 Mix

All sounds are collected, mixed, and sent to a stereo output. We avoided multi-channel setups to facilitate the scalability of the application in museums, cultural institutions, and residencies for older people. The maximum level difference between "foreground" and "background" is set by default at around 30 dB, modifiable via some presets. The nature of these sounds is not totally controllable: the interaction can generate different types of volumes and presence/absence of sonification; the kinetic energy of the user can change; the sounds have themselves an internal structure of subtle modulations which makes them oscillate in volume and presence. We proposed a tentative method to control the absolute output level of the patch, by a process of calibration using a test tone to be measured with a sound level meter: following this method, the "foreground" would ideally generate a maximum level of 50 dbA (fast), while the "background" would average at about 20 dBA (fast). The proposed calibration method aims at diffusing the "foreground" at a level similar to a voice quietly speaking, and the "background" at a level almost hidden in the real background noises of the room, or a little louder.

6. TESTS AND RESULTS

The DanzArTe project completed a pilot study, based on trials in two settings : a) a residency for older people (RSA Cardinal Minoretti) with 10 cases vs. 10 controls, and b) a number of persons selected by the museo Diocesano in Genova (12 cases vs. 12 controls). In both cases, the study included the development the standard treatment with 4 sessions: 2 per week for a total duration of 2 weeks treatment. Each session lasting about 60 minutes each, in groups of 4-5 persons. A paper in preparation will describe in details the positive results of the users' data analysis from this pilot project, based on the MPI (Multidimensional Prognostic Index) [31]. A dissemination and validation program with about 300 users is in progress, involving 4 residencies (RSA) in Liguria and Piemonte, in collaboration with Lavanderia a Vapore – Piemonte dal Vivo and with Museo Diocesano di Genova.

7. CONCLUSIONS

We described an interactive sonification and sound design architecture and a specific project in which it has been applied. The model is general and is a candidate for extensions towards future applications. The choice of granulators and looped sound-files offers a potentially infinite variation of themes and contents – while the general control architecture assures compatibility within the sessions' framework. The particular choices of sound-files in the DanzArTe project revolve around our general low-intrusiveness principle, the communication of positive and relaxed emotional states, and the embodiment of slow and fluid movement.

Possible improvements from this stage include:

- different EQ and mix profiles, in adaptation to different hearing limits of elderly participants;
- development of a 4 channels setup, to spatially separate participants' sonification from other sounds;
- further development of "subtraction" strategies (using the absence of sound to enhance attention when needed, instead of adding sound). A scientific validation process based on an international standard for analysing frailty [14], is ongoing and will be described in a paper in preparation. In the FET PROACTIVE EnTimeMent project (entimement.dibris.unige.it), we developed a conceptual framework to extend the approach described in this paper for the automated analysis, prediction, and sonification of individual as well as joint actions grounded on multiple temporal scales. Further documentation on DanzArTe technology can be found at this link :

https://andreacera.info/CP/CP.html

ACKNOWLEDGEMENTS

This initiative is carried out with the support of the Fondazione Compagnia di San Paolo as part of Well Impact, and of the EU Horizon 2020 FET PROACTIVE EnTimeMent project. The DanzArTe consortium is formed by Università di Genova, Casa Paganini-InfoMus (Coordinator) with partners Dipartimento Cure Geriatriche, Ortogeriatria e Riabilitazione | E.O. Ospedali Galliera (Genova); Lavanderia a Vapore - Collegno (Torino); Residenza per Anziani Cardinal Minoretti (Genova); Museo Diocesano (Genova); with the collaboration of AMEI (Associazione Musei Ecclesiastici Italiana), Fondazione Piemonte dal Vivo, Goethe Institut Genua – Turin, and SIGOT (Società Italiana Geriatria Ospedale e Territorio).

REFERENCES

[1] P. Alborno et al.: "Interactive Sonification of Movement Qualities - A case study on Fluidity.", Proceedings of ISon 2016, pp. 28-33, 2016.

[2] S.Baldan et al.: "The Sound Design Toolkit", SoftwareX, Volume 6, pp. 255-260, 2017.

[3] G.Bell: "Making Sense of Museums – the Museum as Cultural Ecology", Intel Labs, 2002.

[4] D. Botteldooren and B. De Coensel: "Informational masking and attention focussing on environmental sound.", NAG/DAGA, Proceedings, pp.399-402, 2009.

[5] A.Cera, N.Misdariis: "Intrusiveness, Annoyance and Sound Design.", LINKs-series 5-6, pp. 186-191, 2021.

[6] M.Csikszentmihalyi: "Flow: The Psychology of Optimal Experience", HarperCollins Publishers Inc., 2008.

[7] G.Dubus, R.Bresin: "Systematic Review of Mapping Strategies for the Sonification of Physical Quantities", PLoS ONE 8(12), e82491, 2013.

[8] T. Eerola and J.K. Vuoskoski: "A review of music and emotion studies: Approaches, emotion models, and stimuli.", Music Perception, 30(3), pp. 307–340, 2013.

[9] J.Y. Hong et al.: "Effects of adding natural sounds to urban noises on the perceived loudness of noise and soundscape quality.", Science of The Total Environment, 711, 134571, 2020.

[10] I.Lahdelma and T. Eerola: "Single chords convey distinct emotional qualities to both naïve and expert listeners", Psychology of Music, 44-1, pp.37-54. 2014.

[11] R. Niewiadomski et al.: "Does embodied training improve the recognition of mid-level expressive movement qualities sonification?", Journal on Multimodal User Interfaces, 13, pp. 191–203, 2019.

[12] R.Panda et al.: "Novel audio features for music emotion recognition", IEEE Transactions on Affective Computing, 11-4, pp. 614 - 626, 2018.

[13] J.Phillips-Silver and P.E.Keller: "Searching for roots of entrainment and joint action in early musical interactions.", Frontiers in human neuroscience, 6, 26, 2012.

[14] A.Pilotto et al.: "A multidimensional approach to frailty in older people", Ageing Res. Review, 60: 101047, 2020.

[15] M.A. Russo et al.: "The physiological effects of slow breathing in the healthy human.", Breathe, 13, pp.298–309, 2017.

[16] C.Spence: "Crossmodal correspondences: A tutorial review", Atten. Percept. Psychophys. 73, pp.971–995, 2011.

[17] B. Truax: "Soundscape Composition as Global Music: Electroacoustic Music as Soundscape.", Organised Sound, 13(2), pp. 103-109, 2008.

[18] F.Weninger et al.: "On the acoustics of emotion in audio: what speech, music, and sound have in common", Frontiers in Psychology, 4, 2013.

[19] A.Camurri et al.: "Application of multimedia techniques in the physical rehabilitation of Parkinson's patients", Journal of Visualization and Computer Animation, 14(5), pp. 269-278, Wiley Eds., 2003.

[20] C.Lippe: "Real-Time Granular Sampling Using the IRCAM Signal Processing Workstation.", Contemporary Music Review, Vol. 10, United Kingdom, Harwood Academic Publishers, pp. 149-155, 1994.

[21] E. G. Schellenberg and S.E. Trehub: "Good pitch memory is widespread.", Psychological Science, 14(3), pp.262–266, 2003.

[22] R.Moreno and R.E. Mayer: "A coherence effect in multimedia learning: The case for minimizing irrelevant sounds in the design of multimedia instructional messages.", Journal of Educational Psychology, 92(1), pp.117–125, 2000.

[23] A.Singh et al.: "Go-with-the-Flow: Tracking, Analysis and Sonification of Movement and Breathing to Build Confidence in Activity Despite Chronic Pain", Human–Computer Interaction, Taylor & Francis, 31(3-4), pp. 335-383, 2016.

[24] A.Camurri et al.: "The Dancer in the Eye: Towards a Multi-Layered Computational Framework of Qualities in Movement.", Proceedings of 3rd International Symposium on Movement and Computing (MOCO 2016), 2016.

[25] A.Camurri et al.: "Communicating expressiveness and affect in multimodal interactive systems.", IEEE Multimedia, 12(1), pp. 43-53, 2005.

[26] M.J. Vaessen et al.: "Computational feature analysis of body movements reveals hierarchical brain organization.", Cerebral Cortex, bhy228, 2018.

[27] Fraisse P.: "Perception and Estimation of Time", Ann. Rev. Psychol., 35, pp. 1-36, 1983.

[28] S.McAdams: "Psychological constraints on form-bearing dimensions in music", Contemporary Music Review, Vol. 4, pp. 181-198, 1989.

[29] S.Dixon: "Onset Detection Revisited.", International Conference on Digital Audio Effects, pp. 133–137, 2006.

[30] R.H.Thaler and C.R.Sunstein: "Nudge", Yale University Press, 2021.

[31] A.Pilotto et al.: "Using the multidimensional prognostic index to predict clinical outcomes of hospitalized older persons: a prospective, multicenter, international study", The Journals of Gerontology: Series A, 74(10), pp.1643-1649. 2019.

[32] L.Hallnäs and J.Redström : "Slow Technology – Designing for Reflection", Personal and Ubiquitous Computing, Volume 5, Issue 3, pp.201–212, 2001.

[33] P.Valéry: "La Necessité de la Poésie", CONFERENCIA, 32e ANNEE, N°4, 1938.

L'AUDIO EGOCENTRICO NEGLI AMBIENTI VIRTUALI

Michele Geronazzo
Dip. di Tecnica e Gestione dei Sistemi Industriali, Università degli Studi di Padova, Italia
Dyson School of Design Engineering, Imperial College London, UK
Dip. di Studi Umanistici e del Patrimonio Culturale, Università degli Studi di Udine, Italia
michele.geronazzo@unipd.it

ABSTRACT

Le relazioni tra ascoltatore, mondo reale e ambiente virtuale (*virtual environment*, VE) non solo devono ispirare il design di interfacce multimodali intuitive riconducibili al comportamento umano naturale e quotidiano, ma devono poter dare un senso all'azione mediatrice delle tecnologie di realtà virtuale (VR). Questo articolo vuole descrivere un primo framework teorico per l'ambito di ricerca *sonic interaction in virtual enviroments* (SIVE) denominato audio egocentrico e dotato di una visione inclusiva per le sfide attuali e future. In un VE che fa uso di tecnologie di audio immersivo, il ruolo della simulazione VR deve essere enattiva per una esplorazione partecipativa rivolta ad una costruzione di senso in una rete di attori, umani e non-umani. Il custode di tale *locus* per un'espressione d'azione è il digital twin uditivo che favorisce le intra-azioni tra uomo e tecnologia, in grado di ridefinire dinamicamente e in maniera fluida tutte quelle configurazioni determinanti per un'esperienza immersiva e coerente. L'idea di *entanglement theory* viene qui principalmente declinata nella sua connotazione egocentrica e spaziale, legata all'emergere di una conoscenza delle capacità percettive dell'ascoltare grazie ad una relazione attivamente trasformativa ad opera di un *digital twin* che crea movimento, trasparenza e attività provocatorie nei VE.

1. INTRODUZIONE

[1] La nostra esperienza uditiva quotidiana è caratterizzata dall'idea di immersione fin dall'inizio della nostra vita nel grembo materno, ascoltando attivamente i suoni che ci circondano da diverse posizioni nello spazio. Le informazioni uditive assumono la forma di un flusso continuo di messaggi alle orecchie, che ne delinea una rappresentazione compatta e omnidirezionale di apprendimento della nostra esistenza [2], [3]. L'attività temporale e spaziale dei suoni di interesse (ad esempio, dialoghi, allarmi, ecc.) ci permette di localizzare e codificare le informazioni contestuali e le intenzioni della nostra interazione sociale [4].

L'ipotesi che la nostra esperienza di ascolto quotidiano di oggetti sonori con determinate caratteristiche fisiche plasmi dinamicamente l'informazione acustica per le quali noi attribuiamo significato al nostro mondo uditivo è sostenuta dagli studi di acustica ecologica quali [3], [5], [6]. In particolare, l'idea di invariante acustico quale pattern complesso di cambiamento per un'interazione sonora nel mondo reale è fortemente collegata in primis all'apprendimento percettivo umano e ad una mediazione socio-culturale dettata dal mondo reale. Per alcune raccolte di studi classici sul tema dell'acustica ecologica si faccia riferimento a [7].

Da questo punto di vista, gli invarianti acustici si apprendono individualmente su base esperienziale. Da qui l'esigenza di trasferire questa conoscenza ad un sistema tecnologico che sia in grado di fornire una simulazione immersiva ed interattiva di un ambiente virtuale (*virtual environment*, VE) sonoro. Questo processo deve essere adattivo e dinamico per assicurare un livello di accoppiamento tra utente e sistema tale per cui l'esperienza di ascolto attivo sia considerata autentica.

Le tecnologie di realtà virtuale immersiva (qui genericamente denominata VR) permettono un'immensa flessibilità e crescenti possibilità per la creazione di ambienti virtuali con ontologie sia simili sia radicalmente diverse rispetto al mondo reale. Questo può essere detto in altre parole, facendo riferimento alla distinzione tra interazioni naturalistiche e magiche, dove le seconde possono venir considerate configurazioni sistemiche ed osservabili nel dominio delle illusioni artificiali, espandendo incredibilmente lo spettro delle esperienze digitali possibili [8], [9]. Il mio punto di partenza è lo schema di Slater relativo all'attributo *presence* per un sistema immersivo di VR [10] recentemente rivisitato da Skarbez [11]. I concetti di *plausibility illusion* e *place illusion* sono centrali per catturare gli stati interni soggettivi fondamentali a fornire credibilità ad una simulazione digitale sulla base dell'esperienza e delle aspettative individuali rispetto ad una struttura interna di riferimento per scene, ambienti ed eventi. Questo articolo raccoglie le basi per la definizione di un assetto teorico di riferimento per il campo di studio della disciplina Sonic Interactions in Virtual Environments (SIVE) [2].

[1] Questo articolo raccoglie i concetti alla base dell'idea di audio egocentrico nella loro forma originale. Per una trattazione estesa si faccia riferimento a [1].

[2] https://sive.create.aau.dk/

Copyright: © 2022 Geronazzo. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

2. MEDIAZIONE TECNOLOGICA

Il punto di vista che fornirò metterà al centro la relazione tra esperienza uditiva reale ed esperienza mediata dalla tecnologica immersiva con tecnologie di VR.

La prima è caratterizzata da alcuni elementi decisivi per conferire immersività in un mondo fisico. E' importante richiamare nuovamente il carattere di omnidirezionalità che permette all'ascoltatore di poter raccogliere le informazioni di contesto a 360°. Le peculiarità legate all'individualità del segnale uditivo sono il risultato delle trasformazioni acustiche ad opera della testa, orecchio, busto dell'ascoltatore che agiscono da impronta digitale spaziale per un segnale temporalmente complesso. La familiarità e quindi l'esperienza pregressa con sorgenti sonore ne condiziona la localizzazione nello spazio con alta intersoggettività. Infine, studi relativi alla plasticità neurale del cervello umano confermano un'adattabilità continua dell'ascolto rispetto a funzionalità fisiologiche compromesse in soggetti con perdite uditive e rispetto a stimolazioni elettriche, es. in persone dotate di impianto cocleare [12].

La seconda viene spesso indicata dalla peculiarità delle tecnologie VR nel fornire all'utente una controparte digitale chiamata avatar. Questo permette la creazione di un'esperienza situata in un preciso ambiente virtuale digitale che coinvolge il corpo. La letteratura scientifica supporta l'idea che la manipolazione di tali simulazioni possa indurre cambiamenti a livello cognitivo [13], come ad esempio effetti positivi in ambito didattico [14], riabilitativo [15] e terapeutico [16]. La capacità di mediazione nella relazione tra esperienza corporea e situata in un ambiente immersivo virtuale conferisce alle tecnologie VR la possibilità di cambiare la coscienza del sé [17].

3. AUDIO EGOCENTRICO

È quanto più necessario coniare, a livello terminologico, una nuova prospettiva che relazioni le due esperienze di ascolto, detta prospettiva di audio egocentrico. In particolare, facciamo riferimento al termine audio per identificare una componente sensoriale uditiva, richiamando implicitamente tutto un insieme di tecnologie che ne consentono la simulazione/rendering immersivo ed interattivo. Mentre il termine egocentrico richiama il sistema di riferimento percettivo e di acquisizione delle informazioni multi-sensoriali intrinsecamente legato alle tecnologie VR nonché un'accezione di soggettività ed individualità percettivo/cognitiva legata a tali esperienze che plasmano la consapevolezza di sé, della propria identità o una coscienza dell'individuo.

Questa caratterizzazione vuole estendere in qualche modo la disciplina del Sonic Interaction Design [18] tenendo in considerazione non solo l'importanza del suono come principale canale informativo per convogliare informazioni, significato, qualità estetiche ed emozioni, ma bensì una prospettiva egocentrica di intreccio tra soggetto percipiente e computer simulante l'ambiente percepito. Questo può essere descritto in prima battuta da istanze di personalizzazione, adattamento e mutua relazione per mantenere l'illusione di immersività, ma, come cercherò di argomentare in questa trattazione, è molto di più. Mi auguro che questa visione possa guidare lo sviluppo di nuove tecnologie di audio immersivo e un uso consapevole del design sonoro all'interno degli ambienti virtuali.

4. IL DITIGAL TWIN UDITIVO

Quindi il punto di partenza per le basi teoriche esposte è una prospettiva ecologicamente egocentrica. Il presupposto fondativo è quello di studiare un ente semovente con abilità di agire ed intenzionalità. Tale entità ha la possibilità di interagire con l'ambiente virtuale e accorgersi che lo sta facendo in maniera significativa grazie all'immersione in un campo acustico simulato che conferisce tridimensionalità all'esperienza uditiva.

E' doveroso quindi introdurre fin da subito una caratterizzazione terminologica di quello che è l'ascoltatore, non utente in questo contesto, in quanto un essere con un'esperienza pregressa e percezione uditive individualizzate. In stretta relazione vi è un *digital twin* uditivo, che si distingue dall'accezione comune di avatar. Mentre l'idea di avatar in VR prevede che un utente possa prendere il controllo di una qualsiasi forma di corpo virtuale all'interno di una simulazione digitale co-localizzato con oggetti, luoghi e altri avatar [19], il *digital twin*, così da me concepito, non può prescindere da una prospettiva egocentrica dell'ascoltatore per cui è creato. Le relazioni con l'ambiente virtuale non potranno prescindere dal considerare tecniche di personalizzazione del corpo virtuale in stretta relazione a quello del suo proprietario umano. Tale ente media l'interazione tra ascoltatore e tutti i suoni diegetici siano essi prodotti dai gesti dell'avatar o da sorgenti sonore dell'ambiente.

In tale contesto, l'elemento di immersività viene visto come una relazione dinamica tra azione fisica e significativa da parte dell'ascoltatore, nell'ambiente. In particolare, è l'aver effettuato una pratica corporea come camminare, sedersi, parlare, afferrare, ecc. che fornisce senso a luoghi, oggetti e avatar virtuali [20], i.e. nei VE. Quindi il senso di corporeità può essere considerato come un sentire interno soggettivo, espressione della relazione tra il proprio sé e tale ambiente. In particolare, Kilteni et al. [21] si focalizza sulla definizione di senso di corporeità per un corpo artificiale (avatar nel nostro caso) individuandola dall'emergere della mediazione tra le proprietà dell'avatar e la loro elaborazione come se fossero quelle della persona umana.

A tale proposito, mi ricollego all'idea di mediazione tecnologica ad opera di un *digital twin* uditivo custode e facilitatore di (i) un senso di collocamento spaziale, (ii) appartenenza del proprio corpo, e (iii) dell'abilità di agire per l'ascoltatore [22]. Una rappresentazione schematica si trova in Fig. 1. In prima battuta, una visione performativa potrebbe farci vedere la realtà come "fare", enattivamente compiendo azione pratiche [23]. In maniera analoga, l'ascoltatore e l'avatar non possono venir considerate entità fisse ed indipendenti che interagiscono, ma parti costitutive di fenomeni emergenti, multipli e dinamici risultato di elementi sociali, cognitivi e percettivi intrecciati. Questo agire intra-sistemico di elementi intrecciati costruisce dinamicamente identità e

proprietà dell'esperienza d'ascolto immersiva. La permanenza illusoria dell'immersività uditiva sta nel definire i confini tra elementi intrecciati in ogni situazione in maniera fluida e dinamica. Questa determinazione avviene appunto nel *digital twin* uditivo. Potremmo vedere tale entità come il luogo di confronto tra identità dell'ascoltatore e identità del corpo virtuale (avatar e ambiente), comunicante in maniera non-discorsiva (performativa) attraverso il suo manifestarsi secondo il livello di qualità della simulazione digitale.

Nell'ambiente virtuale immersivo, l'ascoltatore non può esistere senza il suo *digital twin* uditivo e viceversa. Attraverso la caratterizzazione del *digital twin*, i segnali acustici generati dall'ambiente virtuale vengono incanalati/filtrati esclusivamente per l'ascoltatore, secondo la sua capacità di estrarre informazioni significative da uno scenario omnidirezionale. E' quindi importante menzionare fin da subito la natura partecipativa del processo di intreccio tra ascoltatore e *digital twin*, come investigazione congiunta del processo attentivo dell'utente nel selezionare l'informazione significativa, es. l'effetto cocktails party [24]. In maniera speculare, potremmo considerare una simulazione che per catturare l'attenzione dell'ascoltatore si rivolga al *digital twin* affinchè fornisca il pattern migliore o cerchi di scoprirlo in modo da attrarre l'attenzione dell'ascoltatore. Il processo di elaborazione della decisione sarà quindi il risultato di intra-azione sistemica guidata dal *digital twin* uditivo.

5. UNA TASSONOMIA PER GLI AMBIENTI VIRTUALI

Ci sono molte classificazioni per descrivere gli spazi virtuali utilizzati per il suono e la musica, in relazione alla tecnologia VR immersiva. Riprendendo la recente formulazione in [25], possiamo mettere in luce tre aspetti fondanti per l'audio:

- **audio immersivo** – l'ambiente virtuale deve fornire la sensazione di essere circondati da un mondo.
- **audio interattivo**– l'ambiente deve permettere all'utente un influenzamento reciproco e significativo.
- **audio virtuale** - la realtà deve essere simulata dinamicamente.

Molte delle tassonomie esistenti per la VR [11], [26], [27] mettono spesso in luce le caratteristiche del sistema (o simulazione) e l'utente, non la stretta relazione che c'è con l'ascoltatore. In questa sezione, propongo una tassonomia centrata sull'audio che non vuole distinguere tra utente e sistema, ascoltatore e simulazione. Cercherò quindi di ripercorrere il ragionamento che ci porta a considerare tre dimensioni per valutare un fenomeno di mediazione tecnologica: immersione, coerenza ed entanglement. Qui ne esporrò una sintesi qualitativa, lasciando come studio futuro una descrizione e, in alcuni casi, quantificazione dei processi descrittivi qui introdotti.

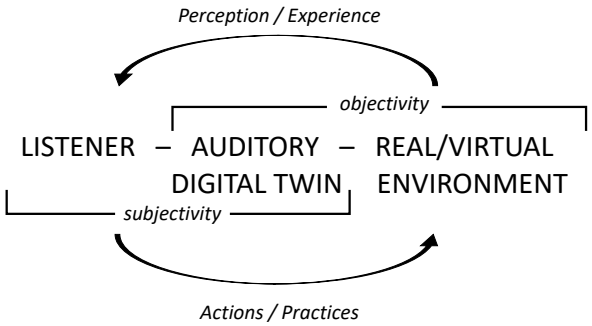


Figura 1. La mediazione tecnologica ad opera del *digital twin* uditivo (schema tratto da [28]).

Recentemente Skarbez e colleghi [11] hanno proposto una rivisitazione del virtual continuum di Milgran introducendo, in particolare, un elemento distintivo: l'idea di coerenza tra simulazione e comportamento umano. Essa tiene in considerazione la plausibilità e l'aspettativa del comportamento del sistema per l'utente in termini cognitivi, sociali e culturali. Tuttavia, le dimensioni proposte non possono e non hanno la pretesa di descrivere la relazione tra utente e sistema come sottolineato dagli autori nella loro tassonomia centrata sul sistema di VR. Possiamo a tutti gli effetti considerarla una rivisitazione di quella proposta da Milgram con un ribilanciamento, per così dire, tra le dimensioni già note. Il lavoro di Skarbez e colleghi è ancora una volta ancorato alla distinzione tra utente e sistema tecnologico che genera numerose difficoltà nell'inquadrare al lato pratico le intra-azioni di fattori in un'esperienza in VR.

Il processo di teorizzazione prende come tecnologia di riferimento quella per la VR. L'estensione di tale approccio alla realtà mista sarà oggetto di studio e confronto per un'eventuale rivisitazioni di tale framework teorico. Partendo dalle dimensioni Immersione [29] e Coerenza [30], suggerisco tre categorie principali che necessitano di un approccio interdisciplinare. Una rappresentazione schematica di tale proposta si trova in Fig. 2.

- **Immersione**: è l'informazione digitale relativa alla relazione ascoltatore-*digital twin* necessaria per permettere il supporto ad un numero via via crescente di azioni nell'ambiente virtuale. Misura del livello tecnologico del potenziale enattivo tra ascoltatore e *digital twin* uditivo.
- **Coerenza**: è l'informazione digitale relativa alla relazione digital twin-ambiente virtuale che permette di rendere plausibili un numero via via crescente di comportamenti nell'ambiente virtuale. Misura dell'efficacia del design dell'interazione sonora nell'ambiente virtuale.
- **Entanglement**: rappresenta l'efficacia complessiva della rete di attori e dei tagli agenziali che si creano dinamicamente, individualmente e adattivamente. Misura la partecipazione al *locus* di azione e una sua conseguente descrizione feno-

menologica. Il *digital twin* uditivo propone attivamente nuove relazioni, favorisce la ridefinizioni dei tagli agenziali in relazione all'azione trasformativa reciproca tra ascoltatore e tecnologia.

Sebbene queste tre categorie siano in linea con lo sviluppo dei concetti legati al design degli ambienti immersivi di VR, è opportuno osservare come siano in forte interconnessione l'una con l'altra: per esempio, risulta difficile raggiungere un elevato grado di Entanglement a fronte di un basso livello di Immersione. In riferimento alla Fig. 2.d, ho comunque optato per una rappresentazione semplificata dello spazio delle configurazioni in tre assi coordinati ortogonali, conscio del fatto che un movimento lungo un'unica dimensione appaia irrealizzabile o comunque realizzabile esclusivamente in specifiche circostanze. Ad esempio, il *digital twin* che osserva e non interviene nella creazione di nuove relazioni limita il livello di Immersione dell'esperienza uditiva nei VE già fortemente influenzata da un'individuazione a priori e statica delle configurazioni nella rete di attori.

6. CONCLUSIONI

Le basi teoriche proposte mettono al centro azione, comportamento e creazione di senso, perché a mio avviso è l'unico vero modo per avere un'esperienza uditiva "autentica" in VR. In particolare, la creazione di senso risulta la sfida più complessa. La diffrazione ed esplorazione delle differenze e di scoperte su base egocentrica richiede modalità inedite di indagine scientifica nel campo dell'acustica e del suono. Uno tra tutti riguarda il livello di personalizzazione che le tecnologie del futuro dovranno richiedere all'ascoltatore. Dovranno venir proposti nuovi paradigmi di interazione artificiale ed immersiva tra uomo e ambiente virtuale. L'abilità di agire per un *digital twin* verrà modellata come un effetto di rete che porterà con sé degli aspetti di natura etica importanti, oltre che ad una elevata complessità nella sua analisi.

Inoltre, come quantificare e classificare i vari sistemi di rete in queste dimensioni resta una sfida aperta di questa mia prima proposta di framework teorico. La visualizzazione e rappresentare delle transizioni e della creazione di tagli agenziali è un tema molto rilevante per avvicinarsi ad un'oggettività della descrizione dei fenomeni di mediazione. Le numerose possibilità, potremmo dire, di creazione di plurime ontologie avvengono attraverso metafore di interazioni "magiche" e una certa trascendenza dal reale in esperienze uniche dell'essere all'interno della VR. Per il momento, la VR non è ancora in grado di imitare completamente la realtà e non è detto che sia intrinsecamente in grado di farlo un giorno. Le sue attuali caratteristiche possono effettivamente permettere agli ascoltatori di fare ed essere cose altrimenti impossibili nel mondo reale. Mi piace pensare che la vera essenza del concetto di diffrazione nell'acquisizione di conoscenza sia la seguente: il *digital twin* deve esplorare delle direzioni che sono umana

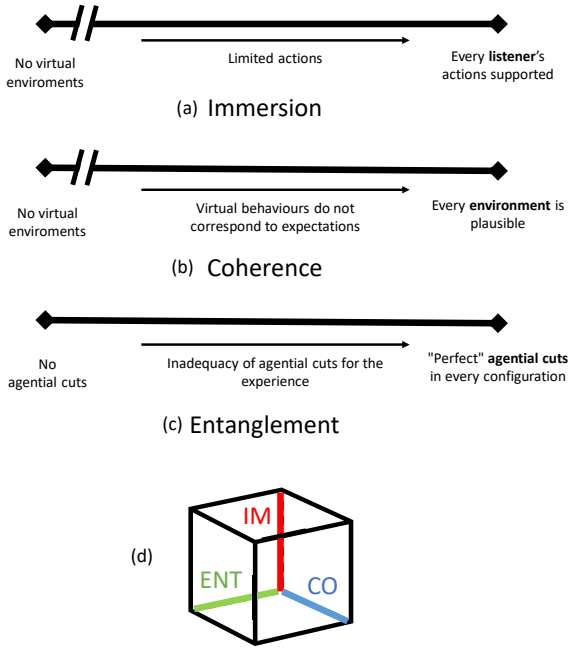


Figura 2. Tassonomia a tre dimensioni, (a) Immersione, (b) Coerenza (c) Entanglement, assieme alla loro relazione in (d). Questa figura è tratta da [28].

mente impossibili da testare nella realtà, con l'obiettivo di estrarre significato per l'ascoltatore.

Potremmo infine asserire che l'idea di super poteri e uomo tecnologicamente aumentato siano di particolare interesse per la nostra cultura occidentale. Il desiderio umano di aumentare le proprie capacità si riflette sulla tecnologie che creiamo e questa tendenza sta ricevendo sempre più attenzione grazie al movimento del post-umanesimo e miglioramento umano [31].

RINGRAZIAMENTI

Ringrazio i colleghi della Sezione Filosofia del Dip. di Studi Umanistici e del Patrimonio Culturale (Università degli Studi di Udine) per avermi dato l'opportunità di condurre una tale riflessione fondativa. Un grazie speciale a Stefania Serafin per l'entusiasmo che mi ha sempre trasmesso per gli ambiti di ricerca SID e SIVE.

7. BIBLIOGRAFIA

[1] M. Geronazzo and S. Serafin, Eds., *Sonic Interactions in Virtual Environments*, 1st ed. Springer International Publishing, 2022.

[2] D. E. Broadbent, *Perception and Communication*. Scientific Book Guild, 1958.

[3] W. W. Gaver, "What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[4] S. Adavanne, A. Politis, J. Nikunen, and T. Virtanen, "Sound Event Localization and Detection of

Overlapping Sources Using Convolutional Recurrent Neural Networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 13, no. 1, pp. 34–48, Mar. 2019.

[5] E. J. Gibson and A. D. Pick, *An Ecological Approach to Perceptual Learning and Development*. New York, NY: Oxford University Press, 2000.

[6] W. Młynarski and J. H. McDermott, "Ecological origins of perceptual grouping principles in the auditory system," *Proceedings of the National Academy of Sciences*, Nov. 2019.

[7] P. Polotti, D. Rocchesso, and D. R. Editors, *Sound to Sense , Sense to Sound A State of the Art in Sound and Music Computing*. Logos Verlag Berlin, 2008.

[8] J. J. LaViola Jr., E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice*, 2 edition. Boston: Addison-Wesley Professional, 2017.

[9] S. Serafin, C. Erkut, J. Kojs, N. C. Nilsson, and R. Nordahl, "Virtual Reality Musical Instruments: State of the Art, Design Principles, and Future Directions," *Computer Music Journal*, vol. 40, no. 3, pp. 22–40, Sep. 2016.

[10] M. Slater, "Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments," *Philosophical Transactions of the Royal Society B*, vol. 364, pp. 3549–3557, 2009.

[11] R. Skarbez, M. Smith, and M. C. Whitton, "Revisiting Milgram and Kishino's Reality-Virtuality Continuum," *Frontiers in Virtual Reality*, vol. 2, p. 27, 2021.

[12] B. Laback and P. Majdak, "Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates," *Proceedings of the National Academy of Sciences*, vol. 105, no. 2, pp. 814–817, Jan. 2008.

[13] H. Sauzéon *et al.*, "The use of virtual reality for episodic memory assessment: effects of active navigation," *Experimental Psychology*, vol. 59, no. 2, pp. 99–108, 2011.

[14] E. Degli Innocenti *et al.*, "Mobile virtual reality for musical genre learning in primary education," *Computers & Education*, vol. 139, pp. 102–117, Oct. 2019.

[15] A. Andreasen, M. Geronazzo, N. C. Nilsson, J. Zovnercuka, K. Konovalov, and S. Serafin, "Auditory feedback for navigation with echoes in virtual environments: training procedure and orientation strategies," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 5, pp. 1876–1886, May 2019.

[16] S. A. Osimo, R. Pizarro, B. Spanlang, and M. Slater, "Conversations between self and self as Sigmund Freud—A virtual body ownership paradigm for self counselling," *Scientific Reports*, vol. 5, no. 1, Art. no. 1, Sep. 2015.

[17] J. Vindenes and B. Wasson, "A Postphenomenological Framework for Studying User Experience of Immersive Virtual Reality," *Frontiers in Virtual Reality*, vol. 2, p. 40, 2021.

[18] K. Franinovic and S. Serafin, *Sonic Interaction Design*. MIT Press, 2013.

[19] U. Schultze, "The Avatar as Sociomaterial Entanglement: A Performative Perspective on Identity, Agency and World-Making in Virtual Worlds," *ICIS 2011 Proceedings*, Dec. 2011.

[20] M. S. A. Graziano, C. S. R. Taylor, and T. Moore, "Complex Movements Evoked by Microstimulation of Precentral Cortex," *Neuron*, vol. 34, no. 5, pp. 841–851, May 2002.

[21] K. Kilteni, R. Groten, and M. Slater, "The Sense of Embodiment in Virtual Reality," *Presence*, vol. 21, no. 4, pp. 373–387, Nov. 2012.

[22] K. Barad, *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press, 2007.

[23] D. Nyberg, "Computers, Customer Service Operatives and Cyborgs: Intra-actions in Call Centres," *Organization Studies*, vol. 30, no. 11, pp. 1181–1199, Nov. 2009.

[24] A. W. Bronkhorst, "The cocktail-party problem revisited: early processing and selection of multi-talker speech," *Attention, Perception, & Psychophysics*, vol. 77, no. 5, pp. 1465–1487, 2015.

[25] J. Atherton and G. Wang, "Doing vs. Being: A philosophy of design for artful VR," *Journal of New Music Research*, vol. 49, no. 1, pp. 35–59, Jan. 2020.

[26] P. Milgram and F. Kishino, "A Taxonomy of Mixed Reality Visual Displays," *IEICE Transactions on Information and Systems*, E77-D, pp. 1321-1329, 1994.

[27] P. S. D. Xiangyu Wang, "A user-centered taxonomy for specifying mixed reality systems for aec industry," *ITcon*, vol. 16, pp. 493–508, 2011.

[28] M. Geronazzo and S. Serafin, "Sonic Interactions in Virtual Environments: the egocentric audio perspective of the digital twin," in *Sonic Interactions in Virtual Environments*, Springer London, 2022, pp. 3–48.

[29] M. Geronazzo, S. Spagnol, and F. Avanzini, "Customized 3D Sound for Innovative Interaction Design," *in SMC-HCI Workshop - CHItaly 2011*, Alghero, Italy, Sep. 2011.

[30] R. Skarbez, N. F. Polys, J. T. Ogle, C. North, and D. A. Bowman, "Immersive Analytics: Theory and Research Agenda," *Frontiers in Robotics and AI*, vol. 6, 2019.

[31] S. Sadeghian and M. Hassenzahl, "From Limitations to Superpowers: A Design Approach to Better Focus on the Possibilities of Virtual Reality to Augment Human Capabilities," in *Designing Interactive Systems Conference 2021*, New York, NY, USA, Jun. 2021, pp. 180–189.

The Italian Electronic Musical Instruments Manufacturers: a Historical Perspective and Open Research Questions

Leonardo Gabrielli
Università Politecnica delle Marche
l.gabrielli@staff.univpm.it

ABSTRACT

Italy has gained international recognition in many industrial and handcraft fields, including the manufacturing of musical instruments such as violins and acoustic pianos. In the last years, with the increasing interest in vintage electronic musical instruments (EMI), international hype has also grown around Italian instruments manufactured in the past decades, which fueled popular interest in the history behind these instruments. Many of the most sought ones were manufactured in Le Marche region in a unique era of economic growth and transition from a mostly rural to an industrial economy.

This paper reports preliminary data about musical instrument manufacturers, some historical evidences gathered from informal interviews that may stand as a starting point for future research works and historical investigations related to the musical instrument industry in Le Marche region. Several considerations related to industrial practices and social factors are put forward.

1. INTRODUCTION

In the last years a renovated popular interest for synthesizers and electronic music has brought a wealth of anecdotes, documents and bits of information out in the public about these instruments, their designers and notable musicians. What is remarkable about this, is the emergence of an interest in the development of the instruments, the design philosophy and the industrial stories behind. A richer and complex picture of the electronic music of the XX century can now emerge, thanks to a novel interest in industrial aspects, rather than aesthetic and musical aspects. In fields related to information and communication technologies (ICT), it is now quite common to stimulate an interest in the general public thanks to a skilled story-telling. Examples of this trend can be found in the Silicon valley companies, that are able to raise curiosity in a field that would be otherwise very technical and of little interest to non-experts.

Another aspect that is raising interest in the history of the musical instruments, rather than in the composers and the performers, is the increasing level of technical preparation of the society. Many electronic music enthusiasts are intrigued by technical aspects or about the stories of the instrument developers rather than in the stories of their famous performers. Finally, the Internet allows a horizontal narration and the emergence of unkown facts (and factoids), allowing to share underground stories that otherwise would never emerge.

All the bits of information coming from informal sources (including but not limited to the Internet), tell that Italy has played a central role in the development of electronic music history, together with the US, the UK, Germany and Japan and it did on both the academic and the industrial frontiers. For what concerns the academic aspects it is well known that Italian composers of the XX century influenced contemporary music, and the research and music production centers of Padova, Milan and Rome attracted composers and scientists from all the western countries. The history of how Italian researchers and composers pushed towards innovation in music is well documented, since the Academia is used to produce written material that survives the years. Recently, essays have also been written about less known pieces of history such as women composers in electronics music [1] and Paolo Ketoff's early synthesizers [2].

For what concerns the musical instruments industry, this has produced less durable signs. Its history, if not unveiled now, risks to be lost forever. Italy had, and still has, a remarkable musical instrument manufacturing reputation abroad. At the forefront we have the traditional classical music instruments: Cremona's luthiers workshops and Fazioli's grand pianos. However, there is more: Castelfidardo's accordion workshops and the larger electronical musical instruments (EMI) industry of Le Marche. As a corollary to these industries come a large number of loudspeakers, live audio and digital audio companies, as well as many software houses developing virtual instruments distributed all along the country. While, e.g. the historical district of Cremona is supported by public institutions, and its historical legacy is preserved by textbooks, papers and a dedicated museum, there is nothing similar for Le Marche and its EMI tradition. Witnesses, tapes, photographs and printed material are now available but they need a careful archiving and restoration to survive the years. This material needs to be documented and analyzed by experts in the field, who can scientifically build a proper critical and historical analysis of the musical instrument industry.

The goal of a historical and archival work is not only

Copyright: ©2022 Leonardo Gabrielli et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

finalized to a sistematization of archaeological findings as if speaking of a dead and long gone story. Preserving this legacy can lead to several goals that stretch into the future: increase identification and build social relations between local citizens around their legacy; increase citizens self-awareness and self-confidence building on the successes of the past; create a global image and advertise it abroad to increase cultural tourism but also export activities and improve business to business appeal. Finally, the past history can inspire the future and enterpreneurs can learn from the errors conducted in the past to make wiser choices in the future. The musical instruments industry is still relevant and has the potential to grow if guided by knowledge and good business strategies.

Unfortunately, at the moment, documents, pictures and tapes are sparse in many private archives with little public access. Even more importantly, with this large amount of emerging knowledge, is the lack of verification and formal organization of the material, considering that most of it comes from the Internet or from anecdotal evidence. With historical instruments becoming now hype and sought for, musical instrument manufacturer and software houses are pushed to exploit their own take on the narrative to make profit. Their large outreach in the public can endanger the protection of a neutral narrative on the history and technicalities of electronic music. E.g. well known filter topologies have been re-branded, schematics from historical instruments have been re-engineered and intellectual property is not always correctly recognized by some international companies, damaging economically Italian manufacturers. As academics we have a role in building a corpus of historical and technical knowledge to protect information from manipulation or to establish the truth when several points of view emerge.

Within this framework, it is clear that there is a lot of room for researchers to conduct a field study to collect data from documents, interviews and witnesses.

This paper is intended as a first report of what has been done, thus far by non-experts, to document the Italian electronic music industry in informal terms. I inscribe myself among the non-experts, not being a historian or a economist. On the other hand, the engineering skills that characterize my research field are useful to validate some claims. However, the main skills that are required come from other research field spanning from history, to document preservation, to economic studies. For this reason, I encourage researchers from different disciplines to join this effort.

The outline of the paper follows. In chapter 2 the currently available data sources are presented together with their limitations. In chapter 3 a brief history of the EMI manufacturers in Italy is provided. In chapter 4 some technical considerations of interest for a history of engineering are reported. Chapter 5 sets some goals for future research, at least from my personal perspective. Finally, chapter 6 concludes the work.

2. HISTORICAL SOURCES

As discussed above, sources of historical information are currently very sparse. Currently, two important sources of information are the associations Associazione Culturale Farfisa and Acusmatiq MATME. The first, mainly curates the history of Farfisa, while the second deals more generally with the whole Italian industry. These initiatives have a role in gathering engineers and former employees in public talks, collecting material and instruments, creating public exhibitions of instruments. Other individuals and groups have collected and restored instruments and technical documents (see Claudio Capponi, Elettronica Musicale Italiana and Vintage Synthesizers Made in Italy), that are available online [1]. Acusmatiq MATME, through its Museo del Synth have gathered and publicly interviewed in the years eminent musicians and composers such as Vittorio Nocenzi and Patrizio Fariselli, engineers and researchers of the IRIS research center including Peppino Di Giugno and Sylviane Sapir, historical engineers and employees from Eko, ISELQUI, Crumar, Korg Italy, Roland Europe, Farfisa, Welson, Elka and Generalmusic. Most of the historical claims present in this paper are gathered from these interviews and written documents owned by the aforementioned associations and individuals (and the author himself).

More formal sources that provide evidence of the industrial activities described in this paper include published papers from conferences. For example, [3, 4] contain results from the development of MARS by Bontempi's IRIS research center. Technical reports of developments at ISELQUI are found in [5, 6, 7, 8]. Documents about the economic status of the musical industry in the early 1980s can be found in [9, 10]. Some relatively recent reviews of the music industry in the whole country are found in [11, 12, 13]. An analysis of an Italian synthesizer of the 1970s can be found in [14], that includes some historical considerations and comparison with synthesizers of the same age. A book that discusses vintage synthesizers and includes some Italian ones is [15]. Other written sources include a history of Giuseppe di Giugno [16] and [17], a review of computer music centers in 1995. Additionally, video documentaries can be found (see e.g. [18, 19] for a perspective on Farfisa and other local companies).

3. HISTORY OF MAJOR EMI MANUFACTURERS

3.1 Genesis

The EMI industry starts from the first accordion artisans from Castelfidardo, Ancona province, Italy. Several legends exist about how the accordion industry started [20], however they all revolve around the encounter between Paolo Soprani, a young boy from Castelfidardo, and a German pilgrim who was heading to Loreto, a town nearby Castelfidardo, and a place of worship. One of the well known accounts, e.g. reports that this pilgrim was hosted for one night by the Soprani family, wo where humble farmers. As a gift, he left his accordion to Paolo Soprani, a young boy in this family, who eventually managed to disassemble the accordion, and study its construction. What

[1] their links are, respectively www.capponiclaudio.it/, www.facebook.com/ElettronicaMusicaleItaliana/, www.facebook.com/vsmiproject/.

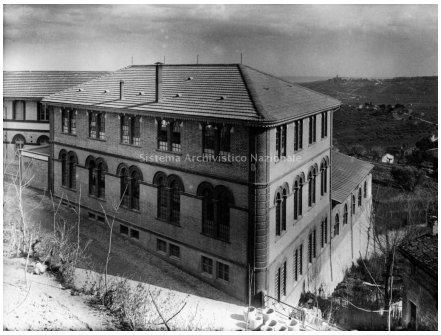


Figure 1: A part of the Soprani accordion factory in 1940s. Courtesy Sistema Archivistico Nazionale.

we know for sure, is that Paolo Soprani started his own accordion company in 1863, the first in the area, which eventually grew up extraordinarily. He became a generous entrepreneur, and built houses for workers around the Soprani factory, leading to the growth of the town. These houses and the factory are still present, and the factory has been restored, constituting a fine industrial architecture example (see Figure 1. The accordion business expanded in the years. Currently, Castelfidardo and the sorrounding towns (Osimo, Recanati) constitutes one of the biggest accordion manufacturing centers in Europe. Accordions are still manually built in workshops or small factory buildings by tens of small firms and brands, often family-owned even though materials and techniques evolved in the years.

Other accordion makers can be found outside the region. The Galanti family, in Mondaino, Rimini province, was in business at the end of the XIX century, and had a great success in the US in 1950s. This family later split into two different companies, Generalmusic and Viscount, both growing to be large digital instruments manufacturers.

3.2 The Analog Days

In the 1950s and 1960s, electronic appliances such as TV, radio and turntables started being common in middle-class homes. At the same time the accordion, as a popular instrument started declining, together with the spread of modern music genres from abroad. This was a reason for many accordion manufacturers to move from this product to electronic devices and electric and electromechanical musical instruments (electric guitar, fan-activated reed organs, etc.), electronic musical instruments (transistor organs), amplifiers and music listening appliances. The area was attractive also to foreign investors: in 1961 the swiss company Lenco opened a factory in Osimo to build turntables; at the end of the 1960s Belgian engineer Anton Parie moved his organ production to Castelfidardo, relying on the Gioielli family.

3.2.1 Organs and Amplification

Italian EMI were mainly conceived for home and amateur use. For this reason, with the development of digital solid state electronics, they also features accompanying sections. These were not uncommon in those days: a British instrument of the era, the Mellotron, was advertised for home use (a UK commercial can still be found on the Internet [2]) even though we now associate it to the best recordings of the Beatles or the early symphonic Progressive Rock bands.

Home organs could be sold worldwide, and accordion makers in Italy, saw this as an appealing market. In Camerano, Ancona province, the Farfisa group (FAbbriche Riunite di FISArmoniche) was started in 1946 from three separate companies: Settimio Soprani (Castelfidardo), Silvio Scandalli (Camerano) and Nazzareno Frontalini (Numana). Farfisa was, at first, a large accordion production facility, but later moved on from accordions to electric organs (with transistor generators) and turntables. Similarly, in Mondaino, Marcello Galanti founded RTE Electronics and started building home and church organs under the Viscount brand. Differently, Eko, in Recanati, Macerata province, grew in popularity for its guitars craft, following the growth of beat music. Vox and Thomas organs were largely produced in Italy at the EME factories, a venture project with Eko. In these factories Moog produced instruments such as the Satellite and the Cordovox.

By the beginning of the 1970s, several smaller companies between Ancona and Macerata were producing organs and string machines, mainly for home use. These instruments were less expensive than American or British alternatives, at least locally, due to reduced distribution and transportation costs. At the time, the availability of foreign instruments in Italy was reduced compared to current days of globalized market (and CO_2-intensive shipping). In the mid 1970s, the musical industry in Le Marche was flourishing. An incomplete list of manufacturers can be found in Table 1, showing a few details on these. In addition to musical instruments, Italian companies also produced amplifiers, mixers and public address systems. Several companies from Emilia Romagna region were/are well known for these devices: Davoli (Parma, est. 1957), Montarbo (Bologna, est. 1960), RCF (Reggio Emilia, est. 1949) and Lombardi (Castrocaro Terme, est. officially in 1971). LEM, a brand of Generalmusic was also very popular. In Le Marche region, FBT (Recanati, est. 1963) is still active in this business and has a reputation in public address for live, building installations, etc. Several large loudspeaker manufacturers were located in the Ancona area, including SIPE (now part of the ASK group), SICA and CIARE, still active. Maps of the status at the beginning of the 1980s are shown in Figures 2-3, compared to the current status.

Although the declared scope of the companies was on instruments for home use, some of these instruments were employed in professional live music. The most trivial example is the Farfisa Compact Duo, featured in Pink Floyd live performances, as seen in the "Live in Pompei" movie or the Ummagumma disc cover.

3.2.2 Synthesizers

In the 1970s, Italian organ companies started producing string machines and monophonic synthesizers, mainly compact ones, which were meant to stand upon the piano or the

[2] "The Mellotron: A Keyboard with the Power of an Orchestra", found at https://www.youtube.com/watch?v=HdkixaxjZCM

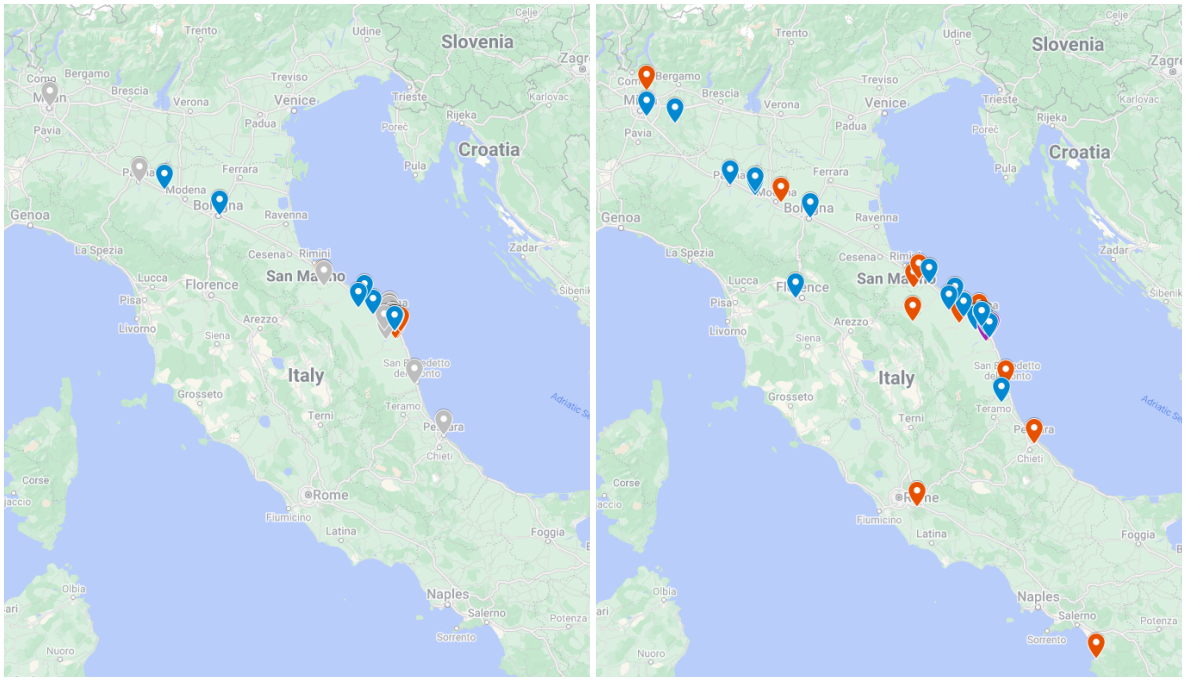


Figure 2: A (possibly incomplete) map of the companies present in 1980 (left) and in present days (right) in the fields of electronic musical instruments (red) and many of the satellite activities, including loudspeakers and digital audio (blue). In grey are shown those electronic musical instruments companies that are not active anymore. Historical data has been collected from the Museo del Synth, EMI, VSMI and Claudio Capponi web sources. Present data has been collected from the author's effort in noting down in the years all companies that had contacts with his institution or the Museo del Synth, or that participated to music fairs such as MusikMesse, NAMM show, MIR Rimini, SoundMit.

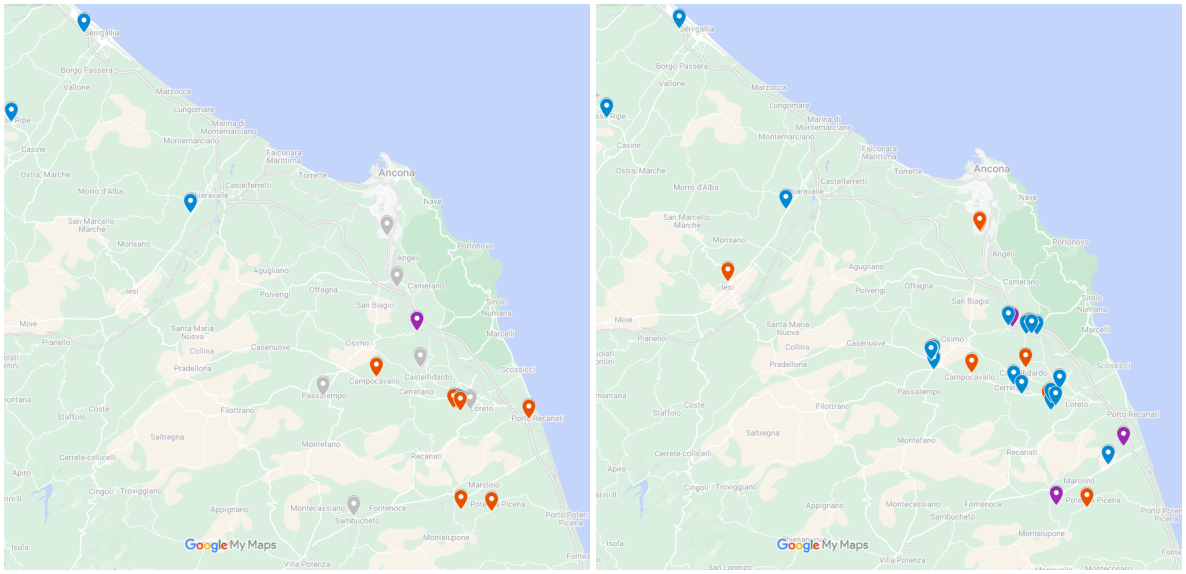


Figure 3: Close-up of Figure 2. Please note that many markers are overlapping, the map on the left has 20 markers, the map on the right has 32 markers.

Company	Established	Headquarter + Facilities	Instruments produced	Notes	End
Antonelli	1935	Osimo (AN) + Montefano (MC)	accordions, turntables (for Lenco), music toys	A new brand, *vociarmoniche*, is still active	-
Bontempi	1937	Potenza Picena (MC) + Martinsicuro (TE)	toy instruments, arrangers, organs	acquired Antonelli and Farfisa	2003?
CRB Elettronica	1949	Ancona	radio, tape recorders, guitar pickups for Eko, organs, synthesizers	Initially founded in Senigallia (AN) moved to Ancona in 1955. They started a joint venture with Solton (DE) in 1979, from which Ketron (still running in Ancona) was started	1982
Crumar	1971	Castelfidardo (AN)	synthesizers, string machines, organs		ca. 1987
Eko	1959	Recanati (MC)	guitars, organs, drum machines	with Thomas (USA) and Vox (UK) they founded EME in Montecassiano	still active
Elgam	1968	Recanati (MC), (Loreto (AN) after 1976)	organs with accompanyiment		1982
Elka	1965	Castelfidardo (AN)	organs, synthesizers		1989 (sold to GEM)
Farfisa	1946	Camerano (AN) + Castelfidardo (AN), + Sambucheto (MC), + San Benedetto del Tronto (AP)	accordions, then, record players, TV (for Metz), amplifiers, organs, synthesizers, acoustic pianos,		1998
Fatar	1956	Recanati (MC)	keybeds and keyboards, currently digital pianos and MIDI controllers (Studiologic)		still active
FBT	1963	Recanati (MC)	PA, mixers, tape echo, synthesizers		still active
Logan	N/A	Porto Recanati (MC)	string machines, organs	Some instruments rebranded Vox and Hohner. After the 1970s acoustic pianos for Baldwin	N/A
Orla	1965	Castelfidardo (AN)	parts for accordions, then organs, nowadays digital instruments		still active
Siel	1976	Acquaviva Picena (AP)	organs, synthesizers	after collaborating with ARP and Sequential Circuits they were acquired by Roland corp. and became Roland Europe	2013 (later acquired by Proel and rebranded as Dexxibel)
Welson	1969	Montefano (MC) + Castelfidardo (AN)	guitars, organs, synthesizers, amplifiers		1984

Table 1: Partial list of major EMI manufacturers located around the Castelfidardo area operating during the 1970s. Other large companies located outside the region are not listed.



Figure 4: The Eko Computerhythm (1972), the first programmable drum machine.

organ of a rock keyboard player, in the style of Progressive Rock keyboardists. These synthesizers were usually compact and meant for pop or rock music, not aimed at experimental electronic music. Following the success of American ARP Pro Soloist many Japanese and Italian companies followed with similar *preset synthesizers*, where the term preset refers to the availability of factory-made hardwired patches that could be recalled by pressure of a button to change timbre.

Notable exceptions in the analog days were the monphonic synthesizers Crumar DS-1 and DS-2, with a panel rich of controls and many novel design choices, the Crumar Spirit, designed by Robert Moog. In the 1980s, with the advent of digital electronics, the Elka Synthex and the Crumar Bit One were capable synthesizers with many features, including MIDI control.

Other successful products were string machines such as the Logan String Melody and the Elka Rhapsody, a string machine featured, e.g., in Klaus Schulze's studio.

3.2.3 The First Programmable Drum Machine

The first drum machine in history to be programmed with a step-sequencer was the Eko Computerhythm invented in Recanati in 1972. It had a matrix of 16 switches by 6 rows for the activation of the individual steps and a punch-card reader for fast programming. The Eko CR was featured by electronic music artists such as Jean-Michelle Jarre. Unfortunately, due to its weight and cost it never really entered mass production, notwithstanding the interest generated at Frankfurt Musikmesse in 1972. In that occasion, reportedly, Japanese engineers took pictures of it, possibly influencing the later Roland TR-808 (1980). The name itself of the Eko drum machine could have influenced the name of a prior Roland rhythm machine series, the CompuRhythm (1978). While in 1972 musicians were not used to produce music with sequencers, by the end of the 1970s times were mature for drum machines and sequencers.

3.2.4 International relations

In those years, there were a lot of interactions with the US and the UK. EME (Elettronica Musicale Europea) was a manufacturing firm founded by Thomas, Vox and Eko. This facility was in charge of producing many instruments for Thomas and Vox, such as amplifiers, organs and effect units. It also produced Robert Moog's monophonic Satellite. Another collaboration occurred between Robert Moog and Crumar for the development of the Spirit, a monophonic synthesizer. A small synthesizer based on the CEM3340 VCO and the SSM2040 filter was produced by Crumar for Nyle Steiner, to be used in conjunction with his EVI breath controller. Crumar also ventured into an American subsidiary in New York, responsible for the development of the GDS (General Development System), a computer based system for 32-oscillators additive synthesis used, e.g. by Wendy Carlos for Disney's Tron in 1982 [15].

3.2.5 Social Benefits

The music industry would employ at its peak some tens of thousands of employees (numbers vary but at the moment an official statistical number is still to be found). This was a small industrial revolution for Le Marche, modifying the structure of society. Before that, many workers, especially female, had "a cottimo" jobs, i.e. they did some work at home, like tuning accordion reeds and without being structured in a company as a worker. With the advent of larger manufacturing plants, companies started hiring women, giving them an extra income and more freedom from their husbands, although unionization required years of battles to improve working conditions [3].

Besides that, music brought modern culture in society. Many pictures of famous Italian singers and international bands can be found during their periodic visits to LEM, FBT, Lombardi and Eko factories to try their favorite instruments and equipment. The presence of national icons such as Mina, Adriano Celentano, Gianni Morandi, etc. also was a social aggregation factor. Youth flew to the factories to try see their idols and sometimes they could even listen to them and talk to them (as the founder of VSMI reported orally to me in an interview). Culturally, this would positively impact these relatively small towns, that had otherwise less chances to listen to famous singers and bands.

3.2.6 Engineering Specialization

During these times the companies turned from small family-owned companies into large enterprises, often with hundreds of employees. Development was conducted mainly by electrical or electronics specialists with a technical school degree. These specialists were well prepared in discrete analog and early digital electronics. The engineering faculty of the University of Ancona was founded in 1969 only, and electronic engineers started flowing into the musical instruments industry only in the late 1970s. Before that, only few engineers were employed in the music industry, often coming from other regions. One notable example is Giulio Viezzoli (1925-2014). He had an electronics degree from the University of Padua and worked for several years in Farfisa. He later founded his own company, Aethra, a large communication technology company.

3.3 The Digital Switch

At the turn of the 1980s digital technology was emerging in the music instrument market. With the advent of micro-

[3] see e.g. the history about Lenco at www.pensionati.cisl.it/public/pdf/pdf_1122_ponti_invisibili_completo.pdf

processors, as found in Dave Smith's Prophet V, or VLSI designs such as Yamaha FM synthesis chips, the music industry faced revolutions that the know-how of the Italian companies did barely have. In a 1982 contribution to the CIM [9] three issues for the EMI industry were highlighted: product reliability, lack of R&D facilities, a weak commercial network abroad. In an effort to overcome this, ISELQUI (Istituto Elettronico per la Qualità industriale), a part-private, part-public company was founded, to perform development and quality tests for the industry. As an example, the Crumar Bit One and Bit99, were partly developed by ISELQUI in conjunction with Crumar. The analog electronics were developed by Gervasio Pannelli, a brilliant expert who worked at the company and had found a smart mechanism for digitally controlling the tuning of the instrument. Carlo Jura designed the panel, called in by his brother Luciano, also involved in the project. Initially it was more sleek and provocative than the one went into production. The firmware of the instrument was developed by a young engineer from Castelfidardo called Sandro Gabrielli, my father, who worked at ISELQUI at the time. The instrument sold well, and was going to be followed by the Example, the first Italian sampler, to sell against the Akai S950 sampler. By the time the product was ready, Crumar was bankrupt and was sold to GEM, which rebranded it as LEM and produced a pre-production batch of about 50 pieces, making it nowadays very rare to find.

During the 1980s many of the companies listed above were closed down, were sold or left the musical instruments market. The few remaining were those with stronger R&D teams and production quality. At this point large Japanese companies found the opportunity to take over this know-how and make it their own. This is the case of Siel, founded in 1976 in Acquaviva Picena (AP), by four former Farfisa employees. The company had about 100 employees at the beginning of the 1980s and produced their own instruments, as well as the ARP Quartet and Sequential Circuits Prelude and Fugue. They were early adopter of the MIDI protocol and developed modern digital synthesizers. In 1987 the company made a joint venture with Roland, and finally became Roland Europe in 1988. In 1995, Korg Italy, a division of Korg Japan was founded in Osimo from three experts working in the music industry and Tsutomu Katoh, founder of Korg. Japanese companies knew that the territory offered skilled engineers, know-how and direct contact with companies producing plastic components (such as the keys), electronic circuits. As an example, FATAR, founded by Lino Ragni in 1956 has grown in the years to be the most important provider of keybeds and mechanics for keyboard instruments. Nowadays the company also assembles and produces OEM instruments and has its own digital keyboards brand.

3.4 The IRIS research center

A chapter of large interest for the computer music community is that of the IRIS research center. Around 1988, the Bontempi-Farfisa group started a new research center for the development of custom DSP ASICs and algorithms, to empower new generations of musical instruments for professional use. At the time, Bontempi was well known for educational and amateur instruments, but they aimed at reaching the professional market with new products.

The IRIS (Istituto di Ricerca per l'Industria dello Spettacolo) was founded in Paliano (FR), Italy, in a villa surrounded by a large green park. The center was located outside of Le Marche to take advantage of funds reserved to the southern Italian regions. Soon after its foundation, Giuseppe di Giugno, agreed to direct the research activities and collected a rich team of engineers and researchers.

The direction at Bontempi allowed them to conduct a few years of research to acquire the experience for later development of products. The experiments turned out to lead to the development of what we may be seen as the European equivalent of the Kyma Capibara. Stemming from Di Giugno's experience with music computing platforms, and Sylviane Sapir's work at the CSC in Padova with the 4I (a re-design of the 4X), the MARS system was designed on a hardware platform based on a custom ASIC and a real-time operating system, the RT20M.

Several generations of chips were developed, as well as hardware platforms and software releases. The large amount of documents found and stored by the Acusmatiq-MATME association suggests that several research directions were undertaken in the years. The MARS was first designed to be programmed using an Atari computer [3]. Five papers were published by the IRIS group at the 1992 ICMC related to the platform. Later, in the second half of the 1990s it was re-designed to be programmed by a Windows personal computer [4]. The IRIS team also developed a powerful spatialization DSP well ahead of the times. The VLSI ICs designed at IRIS, such as the X20 were used in some of Bontempi products, such as the Farfisa F1 (Farfisa was acquired by Bontempi after its shutdown).

3.5 Current Outlook

Nowadays, Le Marche is the first region in Italy in terms of sales for musical instruments production [4]. Musical instruments are still developed and produced, while audio equipment companies have grown notably, due to the larger fields of application of audio technologies [5]. A number of companies has survived several economical crises, and new ones have been founded in the years, adding musical software, modular synthesizers and audio devices to the portfolio of this region's products. The EMI companies still rely on many other satellite companies for mechanical parts, circuit board printing, wood works, silkscreen and veneering. Manual labor is currently very reduced, since production of electronic devices is mainly automated. However, parts of the instruments still involve handmade

[4] Revenues of Italian companies can be accessed freely online and can be grouped by their classification code (in this case ATECO code 32.2 "Fabbricazione di Strumenti Musicali"). To avoid anomalies due to the Covid-19 pandemics, revenues from 2019 have been considered. 12 out of the 15 companies with revenue over 1 million are located inside Le Marche. 62% of the overall revenue for the market segment is from companies in Le Marche. Please note that this data is incomplete since some companies may have different ATECO code. This is the case of wholesale dealers that also have an internal production of instruments. The largest of these is located in Le Marche and has the largest revenue of the lot.

[5] This can be seen from the number of blue markers in Figure 3.

parts and processes. Therefore, although employees are not as many as in times were production was mainly manual, the music industry and its satellite activities still involves hundreds of people and a very rare know-how.

The DSP skills are so advanced that many companies have embedded products and virtual instruments based on physical modelling. Some companies develop algorithms for larger companies abroad and some foreign companies subcontract local companies for the manufacturing. Some companies have offices in the US and other countries. Overall, the competences are so high that most companies that survived the decline of the 1980s have also survived the financial crisis of the 2007-2011, differently from other industrial sectors of the region, that notwithstanding the large public funding shut down their activities or were acquired by larger multinational groups (see e.g. the household appliances district).

4. AVENUES FOR FUTURE RESEARCH

Thus far, I have provided a short history of the EMI manufacturers up to now, from the informal material we have at the moment. As already discussed, to make this research more structured and robust, the material must be handed to experts for filing, classification, archiving and open sharing. The material owned by Museo del Synth, e.g., spans several media and includes: commercial brochures and fliers; technical documents (e.g. wiring diagrams, maintenance and repair, schematics); emails and meeting minutes; pictures of factories, music fairs, and products; music tapes; video tapes and commercials; recordings of interviews; tools and parts from factories; electronic equipment and components; etc. I hope that more researchers will consider working on this material.

From what was said above, many open questions for research arise:

- Building an open archive of documents that can be freely accessed by researchers, including a complete list of companies and all the related data. This may present privacy issues for companies that are still active.
- How is the EMI industry currently performing? What is required to face future challenges?
- Is the local manufacturing an added value or a dead weight for development in global times? What if globalization would decline?
- Is the EMI industry immune to the crisis of other local manufacturing districts (e.g. home appliances, apparel) triggered by localization?
- Can the historical knowledge increase Le Marche citizens self-awareness and self-confidence?
- Can we build an image abroad to improve perception of the cultural value of local musical instrument manufacturing and improve business to business relations?
- Some instruments of the past had reliability issues, poor mechanical and electronics design that endangered serviceability. These issues are better identified with the time passing by. Can we build some knowledge to prepare engineers in university classes to take better design choices?

Finally, I want to put forward a thesis that must be validated through quantitative data and econometrics. After the decline of many manufacturers, in the 1980s, satellite companies in Le Marche adapted to new vertical industrial sectors. Loudspeaker companies moved to TV and automotive. Many companies continued produce printed circuit boards, electrical devices, transformers and supplies, plastic keys and molds. Others moved to hardware and firmware design, acoustics and communications technologies. The thesis I am going to put forward follows: it is quite probable that the development of the electronic musical instruments industries of the 1960s and 1970s pushed the region towards a high technology industry, thus laying the foundations for the subsequent expansion of local industries towards electronics, industrial automation, communication technologies and informatics, and other high technology fields. From this perspective, the shift towards the information and communication fields that the music industry triggered, has gained Le Marche region a safe position with respect to other regions bound to traditional industrial consumer products.

Furthermore, the companies that still design and produce musical instruments, with the value added by craftmanship and artistic sensitivity that a musical instrument requires, are surely on a safer position than other industrial districts such as that of home appliances or fashion, that have suffered in the last years the effect of global economic crises.

5. CONCLUSIONS

This paper provided an overview on the history of electronic musical instruments production in Le Marche and other Italian regions, with some hints at further resarch topics related to economics and society. The data provided has been gathered following a very informal approach, just for the sake of providing a starting point for other researchers with more expertise in history and economic studies. I have not investigated, e.g. social aspects related to labor unions, work organization, relations between the company owners and the employees. In some cases, from the witnesses I have heard, the companies have enriched their territory and towns, providing job and wealth and enriching the local cultural life.

Currently, the region maintains the leadership in Italy in the musical instruments field and is second only to Lombardia in the digital audio field. Studies need to be conducted to indicate how to exploit the knowledge about the past and inspire new avenues for the future.

6. ACKNOWLEDGMENTS

I want to express all my gratitude to Paolo Bragaglia, Riccardo Pietroni and Daniele Marziali for their thorough work

and searches.

7. REFERENCES

[1] Johann Merrich, *Breve storia della musica elettronica e delle sue protagoniste*, Arcana, 2018.

[2] Luigi Pizzaleo, *Il liutaio elettronico. Paolo Ketoff e l'invenzione del Synket*, Aracne, 2014.

[3] Paolo Andrenacci, Emanuel Favreau, Nicola Larosa, Angelo Prestigiacomo, Claudio Rosati and Sylviane Sapir, *MARS: RT20M/EDIT20. Development tools and graphical user interface for a sound generation board*, in Proc. International Computer Music Conference (ICMC), pp. 340-343, 1992.

[4] Paolo Andrenacci, Fabio Armani, Renato Bessegato, Andrea Paladin, Patrizio Pisani, Angelo Prestigiacomo, Claudio Rosati, Sylviane Sapir and Mauro Vetuschi, *The new MARS workstation*, in Proc. International Computer Music Conference (ICMC), 1997.

[5] Antonio Bosetto and Enzo Pezzani, *L'attivit di informatica musicale all'ISELQUI*, in Atti del 5 Colloquio di Informatica Musicale, Ancona, 1983

[6] Gemma Avio, Paolo Danti and Sandro Gabrielli, *Il data base relazionale di suoni realizzato in ISELQUI*, in Atti del 5 Colloquio di Informatica Musicale, Ancona, 1983

[7] Antonio Bosetto and Eugenio Guarino, *Un sistema di sintesi basato sull'uso di "semicustom" sviluppati in Iselqui*, in VI Colloquio di Informatica Musicale, 1985, ISBN 88-7061-915-X

[8] Eugenio Giordani and Sandro Gabrielli, *Applicazioni della Tecnica di Distorsione Non Lineare a Funzioni di Ingresso in Regime Non Sinusoidale*, in Atti del VII Colloquio di Informatica Musicale, 1988.

[9] Domenico Corradetti, *Situazione e prospettive dell'industria di strumenti musicali*, in Informatica: Musica/Industria, Quaderni di Musica/Realt 1, Edizioni UNICOPLI Milano, 1982, ISBN 88-7061-880-3

[10] Mauro Bignami, *Musica, scuola e industria degli strumenti musicali*, in Informatica: Musica/Industria, Quaderni di Musica/Realt 1, Edizioni UNICOPLI Milano, 1982, ISBN 88-7061-880-3

[11] Luca Barbarito and Antonella Ardizzone, *The Italian Music Industry*, in "L'industria, Rivista di economia e politica industriale" 3/2010, pp. 423-458, doi: 10.1430/32502

[12] Fabio Antoldi, Ilaria Macconi, Silvia Gaiardi, *La produzione e l'export di strumenti musicali in Italia. Il ruolo delle piccole imprese artigiane*, in "Quaderni di ricerca sull'artigianato, Rivista di Economia, Cultura e Ricerca Sociale" 3/2017, pp. 297-322, doi: 10.12830/88741

[13] Fabio Antoldi, Chiara Capelli, Ilaria Macconi, *Territori che suonano. I fattori critici di successo della produzione italiana di strumenti musicali*, in "Quaderni di ricerca sull'artigianato, Rivista di Economia, Cultura e Ricerca Sociale" 3/2017, pp. 323-350, doi: 10.12830/88742

[14] Leonardo Gabrielli, Stefano D'Angelo and Luca Turchet, *Analysis and Emulation of Early Digitally-Controlled Oscillators Based on the Walsh-Hadamard Transform*, in Proc. Digital Audio Effects Conference (DAFx), 2019.

[15] Mark Vail *Vintage Synthesizers*, Backbeat Books, 2000.

[16] Alberto Giordano, *Peppino di Giugno, genio italiano*, Edizioni Musicali Esarmonia, 2020, ISBN 978-88-6371-143-1

[17] *Il Complesso di Elettra. Mappa ragionata dei centri di ricerca e produzione musicale in Italia*, Edited by CIDIM Comitato Nazionale Italiano Musica CIM-UNESCO, CEMAT, ISBN 88-85765-05-X

[18] Farfisa: per una storia della fisarmonica di Marco Alessi e Gianluca Cingolani. 2004

[19] Rai Teche Ritratti di citt: Ancona 1968 https://www.teche.rai.it/1968/02/ritratti-di-citta-ancona/

[20] Nazzareno Carini, *Castelfidardo e il suo museo della fisarmonica*, Indepently published on Amazon by his heirs, 2020.

AESTHETICS OF A CONSCIOUS APPROACH TO SOUND DESIGN IN THE SOUNDSCAPE COMPOSITION: A CASE STUDY

Federico Martusciello
Conservatorio Alfredo Casella
federicomartusciello@gmail.com

ABSTRACT

In this paper the processes that led to the creation of the composition of the piece *Impressioni In-Naturali* (for digital support and live performer) will be exposed. This allowed me to focus on a series of relevant issues including performance practice as a construction of the listening experience in an electroacoustically configured space and as a form of ecological awareness. It was also necessary to define an acousmatic form characterized by a form that is not fixed but adaptive to the context, underlining an active attitude highlighted mainly in live electronics. Will be illustrated an analysis of the choices of selection and structuring of the material at the microscopic (realization of the synthesis sounds) and macroscopic (structuring of the generated material) level and the practice of diffusion in relation to an informatic performance environment.

***Keywords* —** Soundscape composition, auditory perception, algorithmic composition, procedural audio, sound hyper-realism, phonorealism, relationship between sound and space, anthropic silence, acoustic ecology

1. INTRODUCTION

Impressioni In-Naturali is a piece for digital support and live performer that focuses on issues related to the composition of the soundscape, operating on the mediation of the relationship between sound and space through a targeted use of simple mimetic sound synthesis forms and involving a performer how to actively interpret the diffusion space. The space as the subject of the composition is thematized starting from the principles of construction and organization of sound coherent with the recording context and through a hyper-realistic approach that raises awareness of the mediation of the acoustic experience. Starting from a phonography of real anthropic silence, in fact, the synthesis sounds are defined coherently with the elements present and / or applicable to the context. The structure of the work plays on the comparison between the "real" and the reconstructed context: the perceived sound environment is real, but the juxtaposition between sound and image is often deliberately decorrelated or ecologically impossible with surrealistic results. This calls for a form of attention that places the listener within the communicative process. The performance practice requires the performer to dwell on the relationship between space and sound in the context of diffusion in which the construction of information and the *impression* on listening takes place. The way of situating sound in space itself becomes a form of creativity that underlines an acousmatic form that is not fixed (or passive), not reproductive but rather *adaptive* to the performative context.

Copyright: ©2022 First author et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

2. THE COMPOSITION OF *IMPRESSIONI IN-NATURALI*

The title of the piece is polysemic and refers, according to the association of words to the theoretical presuppositions and to the methods of structuring the material. The *In-Natural* pun is intended in the forms *in Nature, natural and unnatural*. *In nature* it is a clear reference to the starting material but also to the selection of the synthesis material whose task is to bring the listener's attention to the nature of sound through the sound of nature; a conscious approach to sound design oriented to the primary characteristics of sound, not only from the timbral point of view, but also from an ecological and ethological point of view, so as to make it perceptually justified for listening.

Natural / unnatural expresses the idea behind the montage that focuses on the contrast between synthetic but realistic sound and real recorded sound that appears as "unreal" due to electroacoustic mediation and constructive-performance practice in an electroacoustically configured space.

2.1 Sound material: a "phonorealistic" approach

The sound synthesis approach that I am going to illustrate below is not based on the accuracy of a physical modeling of sound sources, but rather on stylized approximations (or "sound icons") and almost deliberately artificial characteristics of certain sounds, such as we will see. Both aesthetically and conceptually this is connected to a photorealist approach, or if we want "phonorealist", in which the precise representation of reality in all its details is not as important as the representation of the essential, the "truth more than reality" [1]. This involves a reduction in information compared to the intrinsic qualities of a sound compared to the resonance of the original source. We must ask ourselves to what extent this reduction influences listening as a construction of sound reality and, consequently, how much it is actually necessary for a sound to be identifiable

with the actual acoustic event, in a context of experience decorrelated from the original source.

In a conscious approach to sound design, sound is seen as a process rather than a given. This requires an operational strategy called "procedural audio". In this approach, sound as a process counts, as opposed to sound as a product [2]. I believe that this allows us to study the realism of the reproduction and perception of the elements according to the needs related to the listening context, both from a microscopic and macroscopic point.

Starting from a phonography it is possible to obtain a very high number of useful information to define sound design parameters at multiple levels (depth, directionality, rhythmic patterns, reverberation time, etc.). Phonorealist ethics dictates that the composition must strictly adhere to the information found in the phonography.

2.1.1 Field recording: selection criteria of materials for synthesis

The recordings were made at the Sirente Velino Regional Natural Park, a place where it was possible to find an interval of anthropic silence of more than 15 continuous minutes [3]. The choice of selecting a recording without anthropogenic sounds focuses attention on ecological issues and allows a multi-layer stratification without, and therefore usable in different reproduction contexts, of over-masking due to "lo-fi" sounds [4]. The synthesis processes (described later) were guided precisely by the possible geophones and biophones present and / or applicable (as in the case of weather conditions) in line with the recorded environment. Some elements have also been selected or mounted in an ethological form (referring to species-specific behaviors related to the shooting context) of biological antithesis to introduce a further inconsistency detectable by listening.

The sound environments were recorded with two AKG C414 B-XLS cardioid polar figure microphones in ORTF configuration at a height of 180 cm above the ground (equivalent to my listening point).

It was necessary to listen carefully with headphones in an attempt to better grasp the overall perception of the acoustic field. In particular, it was necessary to pay attention to the positioning of the microphones in relation to one's listening position and to maintain adequate phase coherence.

2.1.2 Microstucture: synthetic material

The synthesis of sound took place through an empirical perceptual modeling and an emulation of the processes derived from the listening characteristics of the sounds taken as reference.

Synthetic sound is thought of as a process that conceals the behavioral characteristics of sound elements belonging to the context, in a convincing way to listen. Synthetic imitation need not be identical to imitated sound. For this reason the algorithms used reproduce the significant characteristics through an empirically inspired approach to the physical (ecological) processes that govern them according to an approach exemplified by Andy Farnell [5]. The synthesis material is composed of:

- Rain (2 types of synthesis processes)
- Wind on the leaves
- Thunder
- Birds (7 types of synthesis processes)
- Crickets (2 types of synthesis processes)
- Frog
- Fly

In the following I report one example of algorithm developed. All patches have been implemented in the Max MSP 8 programming environment except for the rain in Pure Data 0.51-4.

- **PD example: rain synthesis algorithm**

Rain is made up of semi-spherical water particles, approximately 1 mm to 3 mm in diameter, which fall at an almost constant speed. All droplets are approximately the same size, resulting from their formation during precipitation under nominally uniform conditions. What gives the rain an uneven sound, a diverse range of amplitudes and timbres, is the striking material [5].

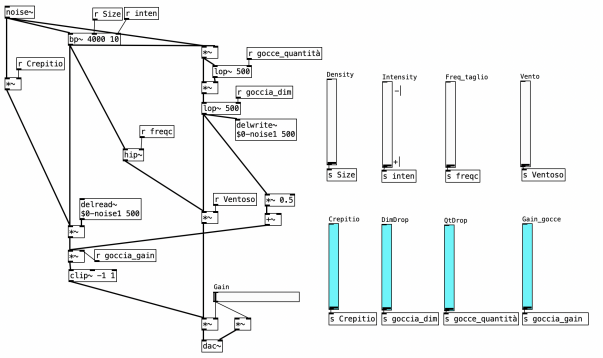


Figure 1. Pure Data complete Rain algorithm

- `Rain density`

The drops are generated from a filtered and modulated white noise source. The density parameter, intended as the number of drops in the unit of time (or quantity), is determined by the central frequency of the resonant band pass filter. The higher the center frequency of the filter, the greater the feeling of increasing the number of drops over the time interval. As we will see, another factor that affects the perception of density overall is the change in the cutoff frequency of the low pass filter. The perception of the events will be directly proportional to the increase in the cutoff frequency of the filter [r drops density].

- `Droplet size or impact material`

The sensation relative to the impact material is determined by the filtering of a low pass filter. The lower the cutting frequency, the lower the consistency of the impact material will appear; with a value of 9, for example, you will have the sensation of impact on the snow. With slightly higher values we are able to recall the impact of the drops on a surface covered with leaves.

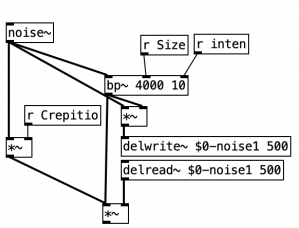


Figure 2. Rain density algorithm

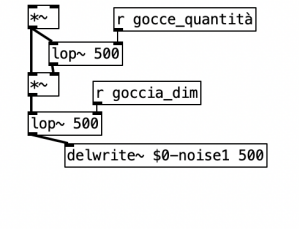


Figure 3. Droplet size or impact material algorithm

- `Rain/wind density`

When it rains and there is wind there is a perceptive feeling that the two phenomena are proportionally correlated. For this reason I decided to opt for a control that affects both the wind speed and the increase in the number of drops at the same time. By varying the central frequency of the band pass filter using [r Size] we will therefore vary the number of events in the unit of time but also the frequency range of the wind generation noise.

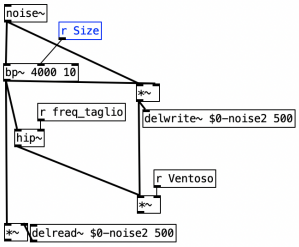


Figure 4. Rain/wind density algorithm

- `Rain/wind density`

The Q variation of the band pass filter affects both the droplet signal and the wind at the same time. By tightening the Q we obtain a decrease in the intensity of atmospheric phenomena, by enlarging it both events intensify. Furthermore, the larger the Q factor, the more we will have the sensation of an increase in the intensity of the collision impact of the rain.

2.1.3 Macrostructure: articulation and assembly of material

The overall editing aims at combining the synthesis sounds with the starting phonography. At the base of the choices

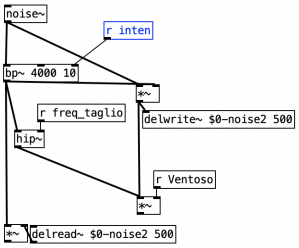


Figure 5. Rain/wind intensity algorithm

related to the selection of the starting material is the maintenance of an ecological coherence with respect to the shooting context: we have followed the information defined starting from the chosen phonography.

From a macrostructural point of view, the editing is characterized by three sections: the first is characterized entirely by synthetic sounds in which a totally virtual soundscape is presented. The continuous and information-rich spectra (broadband) of the environmental conditions (rain and wind) placed at the beginning mask the recognizability of a plausible recording and do not allow a recognition of the formal qualities of the single synthetic elements, leading the perception to accept that environment as hypothetically "real". In the second, the sounds of the environmental shots that intersect with the virtual soundscape enter into fade-in; the real and synthetic elements meet to create a perceptive masking game based on psychoacoustically sensible choices. The interruptions (discontinuities) created through editing techniques favor the processes of segregation between synthetic and real sources. Finally, in the third, all the sounds that make up the piece are superimposed at the same time, without any cutting or editing, so as to leave the performer the freedom to construct the soundscape based on his/her own immersive listening experience.

2.2 Audio mixing

The mixing of the audio material (phonography and synthesis sounds) was carried out on the basis of one's own impression on listening, of course, but seeking coherence in the perspective positioning (in terms of intensity and lateralization) of the individual sources with respect to the elements of the environmental recording. The level of the latter has remained unchanged (and must be left sent) with respect to the taking on the field.

2.3 Diffusion choices

It is preferable that the piece be performed outdoors, even better if in a context devoid of anthropogenic sounds, so as to superimpose an additional layer of overlap and potential masking between synthetic sounds, reference phonography and performance context, raising awareness of the related ecological issues.

The performance of the piece must take place at a level equal to the sound pressure found in the field. The reference level SPL (sound pressure level), as measured in the field, should be between 40 and 48 dB on average (A-

weighting). This choice is determined a priori by comparing the level being listened to with headphones compared to the real soundscape. This allows you to "render" the perspective reality of the elements with respect to your auditory perception of a certain place, in an attempt to replicate the perceptual conditions within a different place. This decontextualization and recontextualization will propose a new sound space at each performance, allowing the listener to project himself into the experience of the relationship between different spaces, real or virtual.

The diffusion of the piece must take place in stereo or, in conditions of a larger room, in quadraphony, octophony etc. Each element of the synthesis, with the exception of the atmospheric elements (Rain and wind on the leaves) and the original recording which must maintain a stereophonic configuration, can be assigned to 1 or more loudspeakers and spatialized at will.

2.3.1 The construction of the listening space in an electroacoustically configured space

Sound is always relational: it refers to the relationships between physical interactions (production) and the environment in which they occur (transmission and reception of sound) [6]. Through a sound-realistic approach the recording context can provide us with a model capable of assuming various forms of relationships that can be descriptively characterizing. The expressive freedom that is exercised through the techniques of reproduction and spatialization in an electroacoustically configured space cannot be easily regulated through a composed "logic". The possibility of improvisation linked to the control of the sources will give rise to momentary forms characterized by a "gesture" capable of revealing a broader structural form and thought. The gestures are transformed into significant "moments" and in alterations in the flow of information acquisition that vary the listening experience. It is precisely the denial of the "familiarity" of certain sounds in their contexts that can influence the way in which a listener understands and draws meanings from what he is listening to.

2.4 Real-time performance environment in Max MSP

The performer is provided with 16 independent tracks already dosed in terms of intensity with respect to the original take so that, with the same gain, the elements find a sensible place in the construction of the starting space.

Since the experience of space is mediated by the electroacoustic apparatus, the performer is left with the possibility of acting on the individual levels, the equalization and positioning of each voice on the basis of personal perceptual choices and listening considerations, taking into account the characteristics of the performance place.

The performer will also have the task of planning or improvising, where he deems it most necessary, both in the first and second sections of the piece, the modifications of the parameters entrusted to him but always bringing attention back to the mediated nature of the experience.

2.4.1 Acousmatic or live electronic music?

In *Impressioni In-Naturali* the performative situation is seen as an integral element of the compositional conception. For this reason I will focus on the differentiation that is commonly associated between two forms of performance practice: the acousmatic one in which a sort of passivity is implied with respect to the form of the "performing gesture" as opposed to the active attitude associated with live electronics.

A hyper-realistic performance practice becomes a way of taking the acousmatic dimension more seriously, allowing the performer to dwell on the problems that exist in the relationship between space and sound in a context in which, precisely on the basis of this relationship, the construction of information takes place for the impression upon listening. The way of situating sound in space itself becomes a form of creativity that underlines an acousmatic form that is not fixed (or passivity), not reproductive but rather *adaptive* to the performative context.

3. CONCLUSION

With this work I wanted to try my hand at a form of composition that thematized space as a compositional subject, not only as a re-contextualization of an existing place, but by making the place also the principle of determining the structural elements from a microscopic point of view (synthesis sound) and macroscopic (structuring choices). The initial environmental recording serves to interpret the subjective image of an internalized reality, rather than an objective one. Composing in sound, intended as a process rather than a product, has allowed me to deepen the sounds of nature and their acoustic interactions within the context of belonging. Furthermore, a sound acquires its meaning only within its context of origin. When detached from its cause it can be used in modeling new relationships within a new context.

4. REFERENCES

[1] K. Stremmel, *Realism*. Taschen, 2004.

[2] A. Farnell, "An introduction to procedural audio and its application in computer games," in *Audio mostly conference*, vol. 23, pp. 1–31, Citeseer, 2007.

[3] G. Hempton and J. Grossmann, *One square inch of silence: one man's search for natural silence in a noisy world*. Simon and Schuster, 2009.

[4] R. M. Schafer and N. Ala, *Il paesaggio sonoro*. Unicopli: Ricordi, 1985.

[5] A. Farnell, *Designing sound*. Mit Press, 2010.

[6] A. Di Scipio, "Mediazione e responsabilità, nel suono," *Mediazione e responsabilità, nel suono*, pp. 69–83, 2010.

[7] R. B. Altman, *Sound theory, sound practice*, vol. 2. Psychology Press, 1992.

Concepts and approaches in analysing spatial sound movements: a link between Mozart and acousmatic music

Stefano Catena
MTI^2 - Music, Technology and Innovation - Institute for Sonic Creativity
De Montfort University
Leicester LE1 9BH
stefano.catena23@gmail.com

ABSTRACT

There are many studies which identify the compositional exploration and exploitation of space throughout history [1, 2]; in the twentieth century — and particularly in acousmatic music — the spatial experience has become one of the central musical concerns [3]. The qualities of sound moving in space might be considered distinct in instrumental or voice-based music if compared to electroacoustic and acousmatic music, where spatialisation technologies and reproduction systems have given a great range of possibilities to composers. By introducing a *coupling-decoupling* process for the analysis of spatial motions and their sounding parameters, it is possible to abstract movements in space from their intrinsic timbral characteristics that might be applicable in the studies of a range of music, both instrumental and electroacoustic.
The concept of *spatial sonorous object* is introduced and used as basic unit for the analysis of the *Notturno in D major K.286* by W.A. Mozart and identification of spatial motion archetypes. In this discussion, a preliminary analysis of 'traditional' spatial sound movements is proposed together with a process of *coupling-decoupling*, an approach in linking spatial movements from different music contexts and their repurpose in the creative practice.

1. INTRODUCTION

Space and sound location have always been considered in musical practice by composers and artists: even the choice of performance space will have an impact on the musical result of a given piece, and some compositions would work better (or worse) in particular spaces. However, space has taken its place as true compositional parameter much later in music history than, for example, pitch, rhythm and timbre, and it is still being much discussed today.
Contemporary technologies for spatialisation and virtual reality permit sound to be positioned in an acoustic space with remarkable precision. However, an established language for articulating spatial experience in creative practice remains underdeveloped [4], despite a very active field [5] of related technical research spanning music and contemporary media (cinema, virtual reality and video games). Moreover, no clear and defined framework for analysing spatial music exists: the idea of a *spatial sonorous object* as basic unit for the identification and categorisation of spatial movements tries to address this gap. By linking the spectromorphology of sound to its spatial changes, it is possible to identify, study and exploit the *spatial sonorous object*: this process is called *coupling*. However, spatialisation is an extremely broad concept that encompasses several aspects of music: from fixed-media composition to live acousmatic diffusion, from site specific installations to virtual reality and many more. Not all spatialisation approaches can be analysed with the introduced method, but only those regarding *illusionary* approaches, rather than *architectonic* ones. The *illusionary* approach to spatialisation tries to trick the listeners' into perceiving that one or more recognizable sources move throughout the space in a coherent manner: alternatively, the *architectonic* approach exploits space by positioning differently sounding sources in various locations and creating an antiphony between them.
Interestingly, a piece that implements illusionary approaches to spatialisation is Mozart's *Notturno in D Major K. 286*: by separating four small orchestras in space, the Austrian composer moves the sound material between the instrumental groups, creating echo effects, distance and depth. By identifying the *spatial sonorous objects* and their purpose in the music, it is clear that Mozart's spatialisation has a structural relevance and is more than a simple compositional embellishment.
Once a *spatial sonorous object* has been identified, it is useful and creatively fascinating to use its identity and properties into a different context, for example in acousmatic music composition. In order to do so, the concept of *decoupling* from the sounding parameters is introduced.

2. BACKGROUND

Space has taken its place as true compositional parameter much later in music history than, for example, pitch, rhythm and timbre, and it is still being much discussed today. The first works in which space is used actively as a musical parameter date from between the 10th and 15th centuries with vocal music, particularly with *antiphonal psalms* [6]. The direct descendant of antiphonal psalms is the *polychoral style* where two or more choirs are ac-

Copyright: ©2022 Stefano Catena et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

tively used, also called *cori battenti* or *cori spezzati* (broken choirs). This compositional school was born in Venice in the 14th-15th centuries and its music was primarily performed in the Saint Mark's church with composers like Willaert, Andrea and Giovanni Gabrieli [1].
Later on, in the Classical and Romantic period, when the orchestra role became more prominent, composers tried to exploit the spatial location of small instrumental groups to create spatial movement and effects: one example is Mozart's Notturno in D Major K. 286. Apart from few examples, spatiality was not very often used: most of the time composers were much more interested in exploring other musical factors such as harmonic relations or timbral effects, leaving composed spatialisation out of the common practice [6]. Towards the end of the 19th century composers like Mahler (*Symphony n. 2*) and Ives (*Unanswered Questions*) wrote works to further explore spatiality as a musical parameter.
However, it was only after the World War 2 that space was integrated more comprehensively into compositional practice. Several composers used spatiality in creative ways, with or without the help of the electronic medium: Karlheinz Stockhausen integrated space in his serialist practice, for example *Gruppen* or *Carré*, works from Iannis Xenakis, the very large repertoire of spatial music from Henry Brant, Murray Schafer or Alvin Curran.
On the other hand, within the electroacoustic domain, spatiality and spatialisation have been explored and exploited much more deeply, also thanks to the introduction of the loudspeaker. Both in Paris with *musique concrète* and in Cologne with *Elektronische Musik*, spatialisation have been considered a compositional parameter and explored as such, for example with Pierre Henry's *potentiometre d'espace* or Stockhausen's multichannel piece *Gesang der Junglinge* [7].
Many more examples of spatial experimentation in the electroacoustic and acousmatic domain were produced: from Xenakis' Brussel Pavillion and Concrete PH [8], to Francois Bayle's Acousmonium for acousmatic diffusion [9], from Stockhausen's Oktophonie [10] to Chowning's illusions of moving sources [11].
Along with new technologies for the spatialisation and diffusion of sound, new theories about spatiality emerged, particularly those regarding the *spatiomorphology* [12]. With this term we intend the spatial properties and spatial changes in sound, modifying and altering the sonic experience. In the case of acousmatic and electroacoustic music, the terms *spatiomorphology* and *spectromorphology* are closely related and influence each other continuously.

3. SPACE AND SPATIAL SONOROUS OBJECTS

3.1 What space?

Space and spatialisation is a broad topic that connects many different fields in the artistic domain: from acousmatic music to instrumental music, from soundscape composition to audiovisual installation. The focus of this study will be the so-called *composed space* [12]. This is where the space is 'scored': this involves a compositional process or action that formalizes spatial movements into a score (for instrumental music) or onto fixed media (for acousmatic music). In acousmatic music this is mostly done in the studio, but there isn't yet a definitive 'scoring' for space in the electronic arts domain.
There have been many stylistic approaches to spatialisation [5], but two main notions can be identified and used as two extremes of a continuum: *organic* vs *architectonic* [13]. In the organic approach, the spatialisation aim is to blend the sound between loudspeaker, creating the illusion of a spatial motion, while the architectonic approach uses the loudspeaker (or an instrument/group of instruments) as sound source and is intended to be perceived this way. The two terms are the 'yin and yang' of spatialisation approaches, the two strongest oppositions: however, this doesn't exclude a continuum between them, where both techniques are used at the same time, or in creative ways to express contrasts.
The term *architectonic* is a direct reference to the exploitation of the acoustic space by means of sound sources and the creation of 'spatial dialogues' in such space. The positioning of the sources is crucial, rather than the movement of the sound between them: most of the time the musical material is extremely different to enhance the static spatial location of each source. The 'architecture' of the space becomes one the focus points of the musical work.
The term *organic* does indeed refer to change and changing behaviors (in this case spatial behaviors), but it is also related to a 'natural' motion, while in acousmatic music it is very likely that surreal or uncommon movements may be deployed. Moreover, while the term *architectonic* is a direct reference to acoustic space, buildings and to a possible compositional intent, the same cannot be said for the term *organic*. The term *illusionary* might be interesting to consider, given the fact that the spatialization achieved by blending sound between speakers is the illusion of moving sound source. 'Spatial illusion' is also a term already used by Barrett: "In the spatial illusion the perceived space appears real, but we are listening to an illusion in stereo or multichannel space produced through the phantom images from two or more loudspeakers" [14]. Moreover, 'illusion' refers both to a spatial image produced by the spatialization technique to 'trick' the listener to perceive a sound source in a certain location, but also to an 'illusionist', the composer, that actively creates the sonic experience.

3.2 Spatial sonorous objects

In the *illusionary* approach to spatialisation, it is possible to recognize spatial motions that may be categorized and studied for their structural role in a composition. This is not the case in the *architectonic* as no such movement can be identified, but rather a dialogue and an antiphony comes into play.
Consequently, it is possible to define a *spatial sonorous object* as any coherent and continuous sound movement that retains a musical function and that can be identifiable by its sounding properties. One can say that it forms the basic unit of a spatial musical structure. Together with its musical function and its movement in the listening space, the

spatial sonorous object is defined by the spectromorphology of the sound(s) it is based upon. The topic of spectromorphology is extremely vast, so it will only be a secondary concern when analysing spatial movements, but it is important to note its correlation to the movements and that most of the time spectral change can be linked to spatial change.
Unlike other musical parameters, space 'needs' sound in order to exist and vice versa. Sound and space are intertwined irreversibly: for example, whenever sound is recorded, the space where the sound 'happens' is recorded as well, inevitably being perceived with the spectral properties of the source [3]. In order to analyse musical *spatial sonorous objects*, this concept needs to be taken into consideration: the 'sounding' properties of the studied motion must be examined in conjunction with its spatial movements in order to fully understand their purpose and meaning. This is why the use of the expression '*sonorous*' is crucial: it refers to the latin word '*sonorus*', from '*sonor*' and '*sonare*', meaning '*to sound, to make a noise*' [1] .
The relevance of the sounding parameters can also be attributed to the 'Something to hold on to factor' [15]: this concept regards the significance of a given musical aspect to the listener and it was intended for composers to be aware of the need of inexperienced listeners to have navigation tools. Even though the 'something to hold on to factor' was originally introduced as a 'strategy to better understand timbral composition' [15], I believe that its usefulness can also spill into the topic of music analysis. Some categories of features that a listener might 'hold on to' can be relevant and be paired and analysed together with the spatial movements of the work. This process of identifying the pertinent factors to analyse space with can be called *coupling*.
The *spatial sonorous object* is conceptually similar to the Schaefferian 'Sound object' and Godøy's 'Gestural sonorous object': in a way, its conception comes from a need of identifying a spatial atomic unit that can be perceived as an 'intentional unit' [16]. This means that a *spatial sonorous object* needs to be recognized and constituted in our mental activity. While sound objects generally last only a few seconds or less, a spatial motion can range from less than a second to even minutes: for example, very slow spatial trajectories can even last an entire composition, but still being perceived as the same movement with the same spectromorphological properties.
We can therefore say that the *spatial sonorous object*: a) concerns spatial settings and *illusionary* spatialisation; b) has implicit sound properties and sounding characteristics that, paired with its spatiomorphology, define its identity; c) can be perceived as a single unit that retains musical function in a compositional practice. With this theoretical framework and background, it is possible to identify and study spatial motions, understanding their structural role (if any) and their musical function inside a piece.

[1] https://www.etymonline.com/word/sonorous

4. MOZART'S NOTTURNO D MAJOR K. 286

Mozart's 'Notturno in D Major for four orchestra K. 286' is a very uncommon composition for the 18th century: with this work he explicitly expressed his desire to incorporate space and sound movements inside the music. This piece from the earlier works of the composer takes advantage of the spatial parameter by placing four ensembles in different locations, each comprising a strings section (two violin parts, viola and bassi) and a horns section. This work has been chosen specifically for this study due to its simplicity and straightforwardness, both in terms of formal structure and its use of the spatial parameter: it is also very easy to understand the use of space simply by looking at how the score is organised.
The positions of the ensembles are not clearly stated by the composer, and the only references to space are the names for each instrumental group: while the first one could be seen as 'primary', the rest are named echoes, as if they were literal duplications of the primary group, and are numbered incrementally in the score: *erstes*, *zweites* and *drittes*. For convenience the groups will be renamed here Orchestra 1, 2, 3 and 4 just as they are identified in the analysed score [17]. The lack of an indication from the composer on how to arrange the four ensembles could lead to creative interpretations for their spatial placement and the resulting gestural effect. Placing the second to fourth groups behind the primary one would lead to a primitive echo effect but arranging them in a circle around the audience or in different layouts based on the concert hall would enhance the spatial experience of the work. However, in a performance of the work reviewed by the Times Magazine [1], Orchestra 1 was placed on stage, Orchestra 2 and 3 on the floor on both sides of the stage and Orchestra 4 almost out of sight in order to create the illusion of distance. To better recreate this distance effect, strings from Orchestra 2, 3 and 4 would use mutes: only half for Orchestra 2, three-quarters for Orchestra 3 and the whole group from Orchestra 4. The use of mutes was not a direct indication from the composer, but the reduction in volume and the progressive change in timbre would clearly enhance the spatial sonic experience of this interpretation [1]. This interpretation will not be directly considered in the analysis, but its peculiar use of the acoustic space and tone altering devices is thought-provoking.

4.1 The use of space

While the formal structure and style of the piece is quite canonical, the use of the spatial component in conjunction with tonal harmony, melodic phrasing and orchestration defines the organisation of the piece. The Notturno is divided into three parts: 'Andante', 'Allegretto grazioso' and 'Menuetto con Trio'. Only the Trio – which is performed by the primary group alone - does not involve all the four orchestras. The work is a typical 18th century tonal piece, following the traditional harmonic rules of the classical period: the orchestration is conventional as well, with violins generally performing the main melody, violas and bass accompanying with occasional doublings and the

horns reinforcing the harmony and rhythm. Mozart moves the sound between the groups similarly to an echo effect, where the phrase from the primary orchestra is repeated equally or shortened; the repeated echo always contains a part of the original phrase.

Analysing the first phrase (Fig.1) demonstrates the use of the spatial component that is predominant throughout the whole piece: the main musical idea is stated in the first four bars by Orchestra 1 and is repeated almost identically by Orchestra 2. The only difference is represented by the pause on the first beat in bar 4, with the horns that follow the harmonic resolution together with the strings. On the other hand, the first violin in Orchestra 2, which plays the main theme, remains the same, acting as link between the two orchestral groups: Mozart differentiates between the 'spaces' by linking them with a single A note (dominant), highlighting the tonal relationships in the material. This 'link between spaces' happens every time there is a shift from one orchestral group to another: a pause would represent a strong closure (between the first and second main phrases, bars 10-11, transitioning from Orchestra 4 to 1), while a common note or chord (vertical) between groups would represent a more fluid spatial transition. These links can also be melodic (horizontal), where two orchestras move a melody between them without any harmonic resolution, no vertical transition and no pauses.

In summary, there are three types of 'link between spaces'

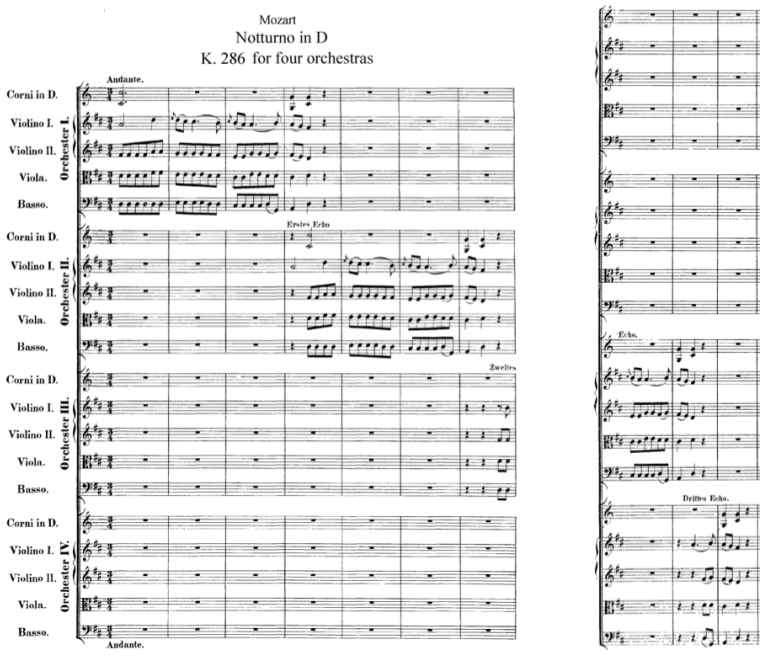


Figure 1. The opening phrase of Mozart's Notturno K. 286, bars 1-9

that Mozart uses to connect the different groups in space:

- *Vertical link*: a common note or chord between the orchestral groups that denotes a spatial movement, resulting in a fluid transition. This transition blends in with the harmonic evolution of the material and its rhythmic component. A vertical link is generally comprised of a perfect cadenza (V-I) in the first orchestra denoting its end and the passage to a new group (Fig.2).

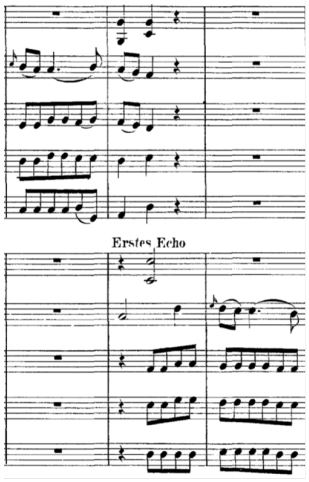


Figure 2. A vertical link, bars 2-4

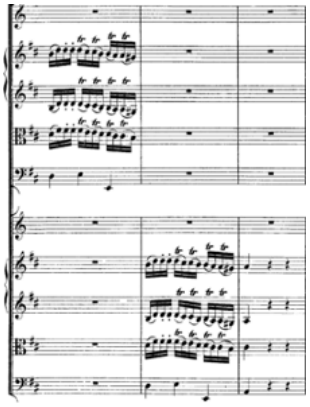


Figure 3. A horizontal link, bars 32-34

- *Horizontal link*: a melodic line that is passed between orchestras. This involves no overlapping notes and there are no pauses: the melody simply moves between the orchestras, for example in bars 30-31, where there's also a shortening of the material (Fig.3).

- *Pause link*: one the orchestras end their part with a pause, with no overlapping during the transition. This represents a harmonic and structural closure, being the end of a section or a switch between a 'full' echo to a shortened one (Fig.4).

4.2 Identifying spatial sonorous objects

In Mozart's Notturno a *spatial sonorous object* can be identified any time there is a continuous and coherent transition from one orchestral group to another, where each orchestral groups contains the same thematic and harmonic material.

But what are the other aspects to consider? As previously identified, the space itself is not sufficient, as no acoustic

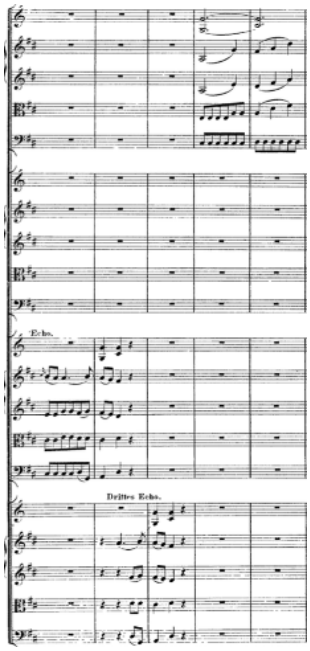


Figure 4. A pause link, bars 8-12

space exists without a sound source, but it must be analysed in conjunction with the other relevant parameters. For instance, what do we hold on to in the score at Fig.1?

- The spatial component: the sound movement between the orchestras and their linking process.
- The tonal harmony and melodic content.
- The echoing spatial rhythm that shortens the material.

These components make up the *spatial sonorous object*, an archetype of the spatial movement (linked with its spectromorphological properties) used in the composition. In Mozart's example, despite the continuous evolution of the music in terms of pitch, rhythm, harmony etc... whenever there is a spatial movement between the groups, no new material is introduced, shifting the listeners' focus to the gestural motion, and highlighting what to hold on to. This interaction, balanced by the composer, is the focus point of the whole work.
Mozart's Notturno contains two prevalent *spatial sonorous objects* that are used throughout the whole work, both inspired by real acoustic phenomenon: *echo* and *reverb*.

- The *echo* object is the most used spatial motion: the main material (melodic and harmonic) is introduced by the first orchestra, repeated faithfully or shortened by the groups one after the other, giving the impression of a call/response performed in different spaces. There is little or no overlapping between the orchestras, so that they may sound as a single repetition of the echoing material. For instance, the opening theme (Fig.1) is played by Orchestra 1 and lasts four bars; at bar four the second orchestra intervenes playing almost the same theme, only with a small variation at the beginning. The true difference starts when in Orchestra 3 the material is being shortened, diminishing from four to three bars: similarly, Orchestra 4 plays only two bars, closing the opening theme. This manipulation of the material, in conjunction with the spatial location of the groups, results in a spatial accelerando or allontanando (distancing), where the sound echoes faster in each orchestra before disappearing. In this case we hold on to not just the spatial manipulation of the sound, but also to the harmonic and thematic material that is comprised in the movement. By moving the melodic content in space, we can better appreciate both the spatial motion and the actual phrases and themes being performed.

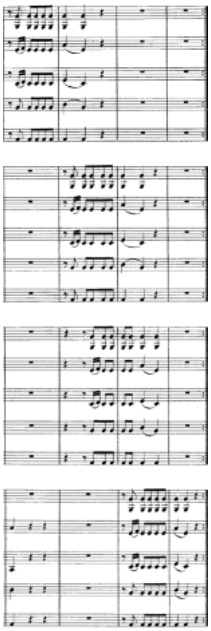


Figure 5. A reverb object, bars 34-37

- The *reverb* object is mostly used in cadenzas, in ending a section and is strongly related to closing V-I harmonic resolutions. This object is like the echo, but in this case there is a considerable overlapping of the orchestral groups, creating a faster 'decay' of the original sound (Orchestra 1); furthermore, no shortening of the material is present. For instance, the ending of the first section of the Andantino is denoted by the 'repeat' sign. The last spatial motion is a reverb object, and its whole duration is just four bars, where each of the orchestral groups has a two-bar part to play, indicating how much they overlap. In this case there is no real theme, so the spatial movement relates much more to the harmonic and rhythmic content rather than to its melodic phrasing; we can therefore affirm that we hold on to spatial motion, rhythmic and harmonic articulation in order to identify this gesture. An example of the reverb object is shown in Fig.5, and it is apparent how much the overlap and the rhythmic component of the orchestration affect this closing gesture.

These two movements represent the archetypes of the *spatial sonorous objects* used by Mozart: their implementation in the score is generally always similar throughout the work and so is their structural role.

5. LINKING INSTRUMENTAL AND ACOUSMATIC MUSIC

After classifying *spatial sonorous objects* in Mozart's work in relation to their identities and their role in the musical structure, several core questions are raised in relation to the use of this information in future works, particularly acousmatic music.

- Are instrumental spatial movements and acousmatic spatial movements comparable and interchangeable?
- Can they be abstracted and analysed only based on their intrinsic parameters?

5.1 The coupling-decoupling process as creative tool

This process is a theoretical framework for the analysis and repurpose of spatial motions in creative environments and is made up of four stages: *coupling, analysis, decoupling and creative use*. The main idea behind this procedure is the linking and consequent abstraction of the sounding properties from a studied spatial motion. This enables a more varied use of spatialisation in music and is a tool for composers to spatially organize their work from a different perspective. Moreover, the exploration of compositional techniques derived from 'traditional' music can be a new way of 'preserving the fire' rather than 'worshipping the ashes', as Gustav Mahler would allegedly say.
The stages that are involved in the process are four:

- *Coupling*: the step of identifying the sounding properties of the considered spatial movement. By *linking* the movement with its sonorous identities, it is possible not only to better understand and perceive the motion itself, but also to learn about its musical significance and structural role. As previously noted, the 'Something to hold on to factor' can be helpful in considering the sounding parameter: its perceptual approach is more interesting than a scientific one. In this stage it is possible to categorize *spatial sonorous objects*.
- *Analysis*: the step of identifying and contextualizing *spatial sonorous objects* in the larger frame of the musical composition. Once that the sounding properties of the spatial motions are determined, it is possible to understand their function and the function of the movements inside the structure of the composition. In this process, the spatial sonorous objects will be identified, studied, categorized and contextualized by their use and purpose.
- *Decoupling*: once that the *spatial sonorous objects* have been analyzed, it will be possible to translate their *spatial function* and *spatial properties* (defined in the 'creative use' bullet point) to other musical domains: for example, translating *echo* objects found in Mozart into acousmatic music composition. In order to do this, it is crucial to *abstract* the sounding properties from the spatial movement itself: without its spectromorphology, it is possible to obtain the pure spatial motion (with its related functions and properties) without the 'interference' of the sound. This abstraction is essential, as the interest lies in the *spatiomorphology* rather than the *spectromorphology* itself; however, the nature of the spatiomorphology might be contingent upon the spectromorphology that accompanies it. For example, a sound that rushes past a listener may be reliant on a series of rising then falling spectra (mimicking a Doppler effect) to make it spatially meaningful.
By abstracting the movements from their underlying and incomparable features (instrumentation, aesthetic considerations, etc..), they can be analyzed similarly, regardless of the origin of sound. For example, an *echo* object found in Mozart's Notturno could be applied to the spatial movement of a sound masses in acousmatic music by isolating the space parameter from pitch, orchestration, rhythm and so on. A concrete example of this application can be Normandeau's *Le renard et la rose* (1995) where the composer largely uses rhythmic, pulsing and echoing materials [2] : translating Mozart's spatial gestures onto Normandeau's work would provide a new and unique interpretation of this acousmatic opera.
Ideally, in a composed space, the *coupling-decoupling* would be integrated in the compositional process, rendering it an integral part of the work, rather than an interpretative tool.
- *Creative use*: the use of the *spatial functions* and spatial *properties* in a creative manner and, possibly, in different musical domains. By *spatial function* we mean the structural role of the motions inside a composition: for example, the closure function of the *reverb* object in Mozart's K.286. By *spatial properties*, instead, we mean the morphology of the movements and their relation to the sounding parameters: for instance, the spatial 'allontanando' of the *echo* object, its movement pattern (always from Orchestra 1 to 4 in numeric order) or the strong V-I harmony (tension-resolution) in the *reverb* object.

By following this process, it is possible to compare and interchange spatial movements from two different domains. Instrumental music might not always be the starting point of these spatial operations: some movements could be translated from the acousmatic to instrumental domain, or even adapted in pieces coming from the same macro-category. In this case, only the 'instrumental to acousmatic' approach is firstly investigated: however, it can be interesting to consider the problematics in other techniques, as it is more complicated to spatialise sound in the instrumental realm.

[2] https://electrocd.com/en/presse/6244 for a review of *Le renard et la rose*

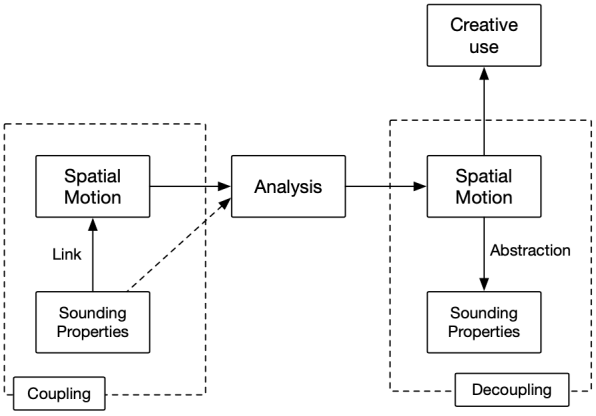


Figure 6. The coupling-decoupling process diagram

6. CONCLUSIONS AND FURTHER WORKS

A new theoretical framework for the analysis of spatial motions and their role and function in a musical structure has been presented, alongside a case study on Mozart's 'Notturno in D Major for four Orchestras K. 286'. The introduction of the *spatial sonorous object* can be a useful tool for spatial analysis of music, not just instrumental pieces, but also in the broader domain of electroacoustic and acousmatic music. Although only on a theoretical level, the process of coupling-decoupling tries to link two different musical domains, instrumental and acousmatic, by 'translating' spatial movements from one to the other.
This idea can be a powerful tool to spatially organize a piece and a new way to enlarge the vocabulary of spatial expression used in both the considered worlds. In this way, a multitude of approaches to spatialisation can be used creatively rather than exclusively: the Mozart analysis has provided two different spatial archetypes that can be exploited compositionally.
Future works will use this presented approach to create listening tests for the study of the reception of the categorized movements. The listening tests' results will be used as foundation for the generation of a taxonomy and a new vocabulary of spatial motions.

7. REFERENCES

[1] J. W. Solomon, "Spatialization in music: The analysis and interpretation of spatial gestures," 2007.

[2] E. Bates, "The composition and performance of spatial music," 2009.

[3] D. Smalley, "Space-form and the acousmatic image," *Organised Sound*, vol. 12, pp. 35–58, 2007.

[4] G. S. Kendall, "Spatial perception and cognition in multichannel audio for electroacoustic music," *Organised Sound*, vol. 15, pp. 228–238, 12 2010.

[5] N. Peters, G. Marentakis, and S. McAdams, "Current technologies and compositional practices for spatialization: A qualitative and quantitative analysis," *Computer Music Journal*, vol. 35, pp. 10–27, 3 2011.

[6] C. Roads, *Composing Electronic Music: A New Aesthetic*. Oxford University Press, 8 2015.

[7] K. Stockhausen, "Music in space," *Die Reihe*, vol. 5, pp. 67–82, 1961.

[8] V. Lombardo, F. Nunnari, J. A. Padget, S. Weinzierl, A. Arghinenti, A. Valle, H. H. Vogel, J. Fitch, R. Dobson, J. Padget, K. Tazelaar, S. Benser, S. Kersten, R. Starosolski, W. Borczyk, W. Pytlik, S. Niedbała, and F. Kommunikations-Wissenschaft, "The virtual electronic poem (VEP) project," 2005.

[9] S. DeSantos, C. Roads, and F. Bayle, "Acousmatic morphology: An interview with François Bayle," *Computer Music Journal, MIT*, vol. 21, pp. 11–19, 10 1997.

[10] M. Clarke and P. Manning, "The influence of technology on the composition of Stockhausen's Octophonie, with particular reference to the issues of spatialisation in a three-dimensional listening environment," *Organised Sound*, vol. 13, pp. 177–187, 12 2008.

[11] J. Chowning, "The simulation of moving sound sources," *Computer Music Journal*, vol. 1, 06 1977.

[12] D. Smalley, "Spectromorphology: explaining soundshapes," *Organised Sound*, vol. 2, pp. 107–126, 1997.

[13] M. Strassmüller, "The parameter of space: Can instrumental and acousmatic music share the same approach to space?," 2008.

[14] N. Barrett, "Spatio-musical composition strategies," *Organised Sound*, vol. 7, pp. 313–323, 2002.

[15] L. Landy, "The 'Something to hold on to factor' in timbral composition," *Contemporary Music Review*, vol. 10, pp. 49–60, 1994.

[16] R. I. Godøy, "Gestural-sonorous objects: Embodied extensions of Schaeffer's conceptual apparatus," *Organised Sound*, vol. 11, pp. 149–157, 2006.

[17] W. A. Mozart, "*Notturno in D major, K.286/269a*," 1777. Mozarts Werke, Serie IX: Cassationen, Serenaden und Divertimente, Breitkopf und Härtel, pp 293-324, vol. 10.

Uno studio sul controllo di intonazione di un sintetizzatore tramite gesti di spinta e trazione

Stefano Papetti **Hanna Järveläinen** **Eric Larrieux**

Institute for Computer Music and Sound Technology
Zurich University of the Arts
name.surname@zhdk.ch

ABSTRACT

Viene descritto un esperimento, condotto con musicisti esperti, in cui è stata studiata l'accuratezza dei gesti di spinta e trazione esercitati da un dito per il controllo di intonazione di un sintetizzatore rispetto a un tono di riferimento. A questo fine, è stata utilizzata un'interfaccia musicale digitale sviluppata internamente, capace di misurare forze nelle tre dimensioni applicate al suo pannello superiore. È stato inoltre testato l'effetto di diversi tipi di feedback vibrotattile forniti dall'interfaccia durante l'esecuzione del compito assegnato. Infine sono state raccolte alcune valutazioni soggettive dai partecipanti riguardo all'esperienza interattiva.

1. INTRODUZIONE

Le innovazioni incorporate nelle più recenti interfacce musicali digitali hanno consentito di mappare gesti continui e multidimensionali su diversi parametri musicali, accrescendo così il loro potenziale espressivo per il controllo di sintetizzatori e strumenti virtuali [1, 2]. Alcuni esempi degni di nota, attualmente disponibili come prodotti commerciali, sono: ROLI Seaboard RISE [1] , Madrona Labs Soundplane [2] , Roger Linn Design Linnstrument [3] , così come varie interfacce grafiche basate su touchscreen progettate per dispositivi quali Apple iPad.

Tra i controlli espressivi più rilevanti musicalmente, la manipolazione dell'intonazione (intendendo con questo sia *pitch modulation* che *pitch bending*) è stata tradizionalmente possibile nella pratica musicale con diversi strumenti, per esempio quelli a corda e perfino alcune percussioni. In tali strumenti, l'intonazione può essere modificata deformando meccanicamente la sorgente sonora (es. tirando le corde di una chitarra o premendo la membrana di un tabla) con la stessa mano o le stesse dita impiegate per suonare. Al contrario, negli strumenti musicali elettronici e digitali – specialmente quelli provvisti di tastiera – il *pitch bending* è generalmente delocalizzato e assegnato ad una leva con ritorno a molla o a una rotella che vengono tipicamente controllate con la mano sinistra. Questo risulta in un'ergonomia molto diversa da quella non mediata offerta dagli strumenti musicali classici. Oggigiorno alcune interfacce musicali digitali innovative – quali quelle menzionate poco sopra – hanno però reintrodotto la possibilità di modulare l'intonazione in modo più diretto e naturale, ad esempio scorrendo un dito verso l'alto o il basso sulla stessa superficie che viene suonata. Sebbene queste interfacce offrano la possibilità di adattare alcune caratteristiche della funzione che mappa un gesto al *pitch bending*, il suo controllo accurato pone comunque delle difficoltà, in parte poiché utilizzando tali dispositivi il/la musicista non riceve un feedback aptico (es. di forza) paragonabile a quello prodotto dagli strumenti musicali classici [3].

In letteratura si trova che la capacità di giovani soggetti normo-udenti di discriminare differenze di intonazione per via uditiva è di circa 0.5% in un ampio spettro di frequenze, mentre i musicisti possono perfino raggiungere lo 0.1% [4, 5].

Sebbene il senso del tatto sia molto più limitato in frequenza rispetto all'udito (le mani possono percepire vibrazioni fino a circa 1 kHz, con un picco di sensibilità tra i 200 e i 300 Hz), anche gli stimoli tattili possono essere percepiti come intonati, e questo in dipendenza sia della loro frequenza che dell'intensità [6, 7]. La capacità del tatto di discriminare differenze di intonazione è stata misurata tra il 18% e il 3% [8, 9].

È stato dimostrato che esistono interazioni tra l'udito e il tatto per quanto riguarda la percezione di consonanza, intensità, e intonazione di uno stimolo sonoro e uno vibratorio [10, 11]. Tuttavia rimane da chiarire se e come uno stimolo vibrotattile rumoroso possa interferire con il controllo dell'intonazione, o se viceversa uno stimolo tattile coerente con il feedback uditivo possa essere di supporto.

In uno studio connesso a quello qui descritto, è stato analizzato l'utilizzo di feedback vibrotattile in uno strumento musicale digitale per la navigazione di uno spazio di intonazione [12]. Si è verificato che il feedback tattile può dare un contributo positivo alla capacità di suonare in modo intonato quando lo strumento è nascosto alla vista, tuttavia con un impatto negativo sulla performance temporale.

Il presente lavoro riguarda diversi aspetti: la progettazione di interfacce musicali basate sul gesto, il controllo

[1] https://roli.com/products/seaboard/rise2
[2] https://madronalabs.com/soundplane
[3] https://www.rogerlinndesign.com/linnstrument

Copyright: ©2022 Stefano Papetti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

dell'intonazione di uno strumento musicale, e infine i possibili effetti cross-modali che suono e tatto possono avere sulla discriminazione dell'intonazione. Nell'esperimento descritto di seguito, l'intonazione di un sintetizzatore doveva essere controllata fino a corrispondere a quella di un tono di riferimento. Il controllo avveniva premendo leggermente un dito sul pannello superiore di un dispositivo interattivo, e poi spingendo in avanti (incremento in frequenza) o indietro (decremento in frequenza). Il pannello poteva inoltre riprodurre due diversi tipi di vibrazione in contemporanea all'esecuzione del compito assegnato. L'introduzione di feedback tattile permette di chiudere il loop di interazione a livello multisensoriale, coinvolgendo vista, udito e tatto, similmente a quanto accade nella pratica musicale con strumenti tradizionali.

2. APPARATO SPERIMENTALE



Figure 1. Il setup sperimentale.

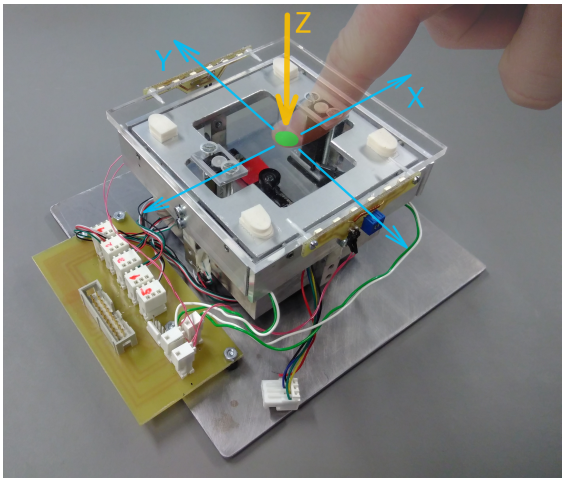


Figure 2. Il TouchBox.

Il dispositivo utilizzato nell'esperimento, chiamato TouchBox, è la più recente implementazione di un progetto precedentemente pubblicato in forma open-access [4]. Il TouchBox, mostrato in Fig. 2, dispone di un pannello superiore in Plexiglas che misura le forze applicate ad esso in tre dimensioni (lungo la direzione normale, longitudinale e trasversale), e può riprodurre feedback vibrotattile di varia natura, essendo pilotabile da generici segnali audio. Si rimanda alla pubblicazione originale per i dettagli tecnici [13] [5].

Come mostrato in Fig. 1, i partecipanti all'esperimento sedevano ad un tavolo su cui era posizionato il TouchBox e premevano un dito sul suo pannello superiore. Per rendere l'azione più confortevole, essi potevano appoggiare l'avambraccio su di un supporto e regolare l'altezza della seduta.

Per evitare che il dito scivolasse durante l'azione di spinta e trazione a causa della superficie liscia in Plexiglas, è stato incollato al centro del pannello superiore del TouchBox una piccola porzione ($2 \times 2\mathrm{cm}$) di carta vetrata fine.

Lo stimolo sonoro – generato da un sintetizzatore software realizzato in Max – è stato fornito tramite cuffie di tipo chiuso (Beyerdynamic DT 770 Pro), connesse ad una interfaccia audio (MOTU M4). Il sintetizzatore riproduceva una semplice forma d'onda a dente di sega processata da un filtro passa-basso per ottenere un suono più piacevole in vista del suo uso prolungato.

Lo stimolo vibratorio – sempre generato in Max sotto forma di segnale audio – è stato riprodotto pilotando un attuatore a bobina fissato al di sotto del pannello del TouchBox. A questo scopo potevano essere utilizzati due diversi segnali: lo stesso segnale generato dal sintetizzatore sonoro, oppure un segnale rumoroso. Nel primo caso, l'obbiettivo era quello di simulare quanto avviene nella pratica musicale con uno strumento acustico o elettro-acustico, in cui la sorgente sonora e quella di vibrazione coincidono; l'obbiettivo della seconda configurazione era invece quello di fornire una vibrazione del tutto scorrelata dal segnale sonoro. Essendo il picco di sensibilità alle vibrazioni tra 200-300 Hz [6], al fine di produrre feedback vibrotattile uniformemente efficace, l'estensione in frequenza del sintetizzatore sonoro è stata limitata a ± 2 semitoni intorno a DO4 = 261.6 Hz, risultando così nell'intervallo di frequenze 233.08-293.66 Hz (SIb3-RE4). Sebbene ristretto, tale intervallo in frequenza non ha limitato la validità dell'esperimento in quanto la capacità di discriminazione in frequenza degli esseri umani è stabile in un ampio intervallo da circa 200 Hz fino a diversi kHz [14]. Similmente, il feedback vibratorio rumoroso è stato processato da un filtro passa-banda impostato a 40-360 Hz al fine di massimizzare la sensazione vibrotattile. In entrambi i casi l'intensità delle vibrazioni è stata equalizzata ad un valore di accelerazione pari a 120 dB RMS (re $10^{-6}\,\mathrm{m/s^2}$), di modo da essere sempre chiaramente percepibile [6].

Le possibili emissioni acustiche indesiderate generate dall'attuatore sono state sufficientemente mascherate grazie all'uso di cuffie.

Un interruttore a pedale, connesso via MIDI all'interfaccia audio, consentiva ai partecipanti di registrare come risposta il valore attuale in frequenza del suono di sintetizzatore, e successivamente di procedere nel test.

[4] https://doi.org/10.5281/zenodo.1204970

[5] https://doi.org/10.5281/zenodo.3929455

Il tavolo infine ospitava uno schermo e un mouse, tramite i quali i soggetti potevano assegnare delle valutazioni, come descritto di seguito.

3. ESPERIMENTO

Controllando un sintetizzatore sonoro tramite il TouchBox, l'esperimento richiedeva di riprodurre quanto più accuratamente possibile l'intonazione di un suono di riferimento. L'esperimento ha quindi affrontato le seguenti questioni:

1. I gesti di spinta e trazione del dito risultano in una diversa accuratezza di intonazione?

2. Il feedback vibrotattile può influenzare il controllo di intonazione?

Inoltre è stato chiesto ai partecipanti di valutare i seguenti aspetti soggettivi relativi al compito assegnato: senso di controllo (in seguito riferito solo come *controllo* o *agency*), fiducia nella correttezza della propria risposta (*autostima* o *confidence*), gradevolezza dell'esperienza (*gradimento* o *pleasantness*).

3.1 Design sperimentale e stimoli

Il task sperimentale è stato eseguito mediante due fattori incrociati: gesto e feedback vibrotattile. Due opzioni erano disponibili per il gesto: spingere in avanti (spinta o *push*) o tirare all'indietro (trazione o *pull*) il pannello superiore del TouchBox. I due gesti venivano rispettivamente richiesti per aumentare o diminuire la frequenza del tono di partenza del sintetizzatore in direzione del tono di riferimento. Tre opzioni erano disponibili per il feedback vibrotattile: nessuna vibrazione, vibrazione rumorosa (*noise*), e vibrazione intonata (*pitch*, ovvero identica al segnale usato per il feedback sonoro). La vibrazione, se presente, veniva comunque prodotta in contemporanea al suono di sintetizzatore, che era invece sempre presente.

Le frequenze target sono state randomizzate in un intervallo continuo ± 2 semitoni intorno alla nota DO4 = 261.6 Hz. Il tono iniziale – ovvero prodotto all'inizio di ogni ripetizione del test – era da 1 a 3 semitoni al di sopra o al di sotto della frequenza target.

L'esperimento ha misurato le seguenti grandezze: accuratezza relativa di intonazione (misura oggettiva), e punteggi assegnati dai partecipanti relativamente gli attributi *controllo*, *autostima* e *gradimento* (misura soggettiva).

3.2 Soggetti e procedura sperimentale

Hanno partecipato all'esperimento 31 soggetti normoudenti con formazione musicale professionale (14 F, 17 M, età 20-39, m=24.6; anni di pratica strumentale: m=16). All'inizio di una sessione sperimentale, il/la partecipante veniva istruito/a su come eseguire correttamente il compito assegnato, usando un dito a scelta della mano dominante durante tutto l'esperimento.

La durata di una singola sessione, comprendente 72 ripetizioni, è stata di circa 20 minuti, includendo anche tre brevi pause.

All'inizio di ciascuna ripetizione veniva suonato un tono di riferimento per 3 s, dopo di che il soggetto doveva regolare la frequenza del suono di sintetizzatore fino a corrispondere con quella del tono di riferimento.

Al fine di minimizzare confusione e interruzioni, le ripetizioni sono state organizzate in 2 blocchi in base al gesto da utilizzare, in ordine casuale; inoltre lo stesso tipo di feedback vibrotattile è stato fornito per una sequenza di 4 ripetizioni successive.

L'assegnazione dei punteggi per le valutazioni soggettive è avvenuta dopo ciascuna sequenza di 4 ripetizioni. A questo scopo, i partecipanti dovevano utilizzare il mouse messo a disposizione per muovere 3 cursori – uno per attributo visualizzati a schermo – che offrivano una scala visuo-analogica mappata internamente sull'intervallo continuo [0, 1] (0 = molto basso, 1 = molto alto).

4. RISULTATI

4.1 Analisi descrittiva

I risultati sono presentati in Fig. 3 per l'accuratezza di intonazione, e in Fig. 4 per le valutazioni soggettive. I dati grezzi ivi mostrati suggeriscono che il gesto di spinta ha causato errori generalmente maggiori. Allo stesso tempo, i soggetti hanno coerentemente assegnato punteggi migliori ai 3 attributi soggettivi quando utilizzavano il gesto di spinta. Va anche osservato che l'esecuzione del compito è stata leggermente più veloce con la spinta, essendo il tempo medio per ripetizione pari a 10.6 s contro i 12.0 s della trazione. Ad ogni modo, come visibile in Fig. 3, i soggetti hanno dimostrato una tendenza leggermente maggiore al miglioramento (*learning effect*) con la spinta, in quanto verso la fine delle sessioni sia spinta che trazione si avvicinano ad una accuratezza dello 0.5%.

4.2 Analisi statistica

L'analisi statistica dei dati è stata effettuata utilizzando il software R e il pacchetto brms [15, 16]. Sui dati di accuratezza di intonazione e i punteggi soggettivi è stato eseguito un fit di modelli statistici, e i rispettivi parametri sono stati stimati tramite inferenza Bayesiana.

L'errore di controllo di intonazione (espresso in centesimi, dove centesimo = $\frac{1}{100}$ di un semitono) [6] è stato predetto usando un modello Gaussiano dai fattori gesto, feedback vibrotattile e dalla loro interazione. Per ogni soggetto è stata utilizzata una intercetta individuale.

Solo il fattore gesto ha avuto un effetto significativo sull'errore. Gli errori con segno stimati mostrano una chiara differenza: si nota un *overshoot* di 8.74 centesimi (+0.53% rispetto alla frequenza target) per la trazione, e un *undershoot* di -12.1 centesimi (-0.68% rispetto alla frequenza target) per la spinta. Per studiare la significatività di tale differenza, senza considerare l'effetto del segno, è stato fatto il fit di un secondo modello per valori di errore relativi al gesto di spinta simmetrici attorno allo zero. La stima di tale errore risulta maggiore di 3.4 centesimi per la spinta,

[6] L'errore è stato calcolato come $1200 \cdot \log_2(\frac{\text{frequenza selezionata}}{\text{frequenza target}})$, dove 1200 rappresenta l'intervallo di 1 ottava in centesimi.

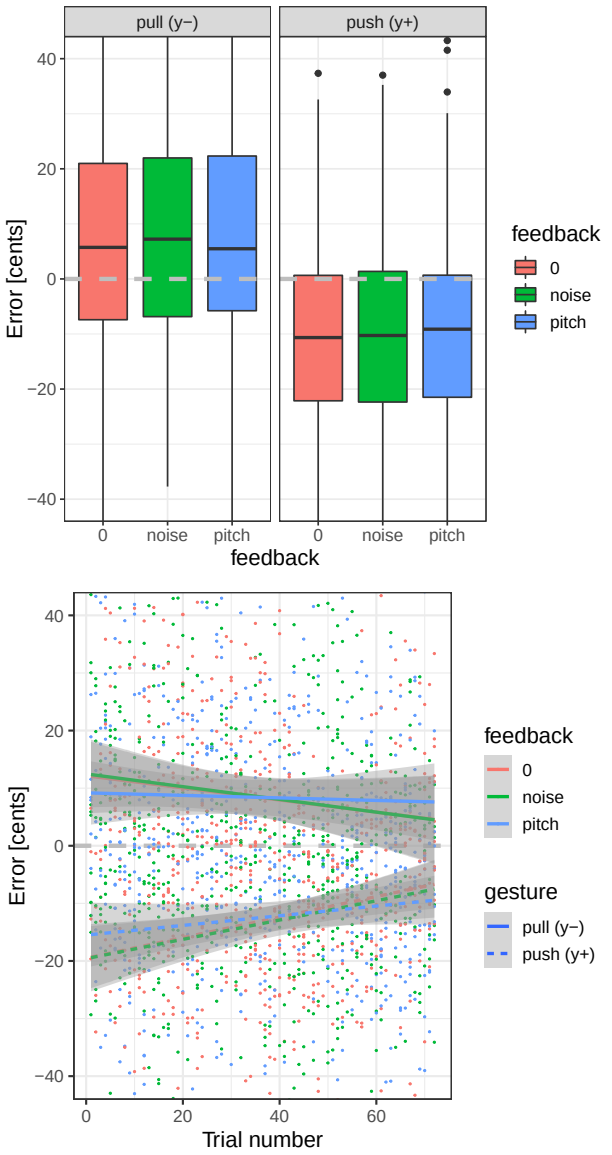


Figure 3. Sopra: accuratezza di intonazione (dati grezzi); Sotto: accuratezza come funzione del numero di ripetizione (dati grezzi e trend lineare).

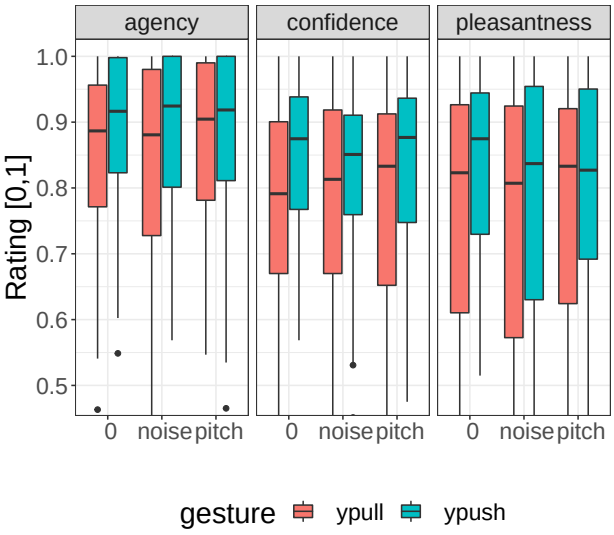


Figure 4. Risultato delle valutazioni soggettive (dati grezzi).

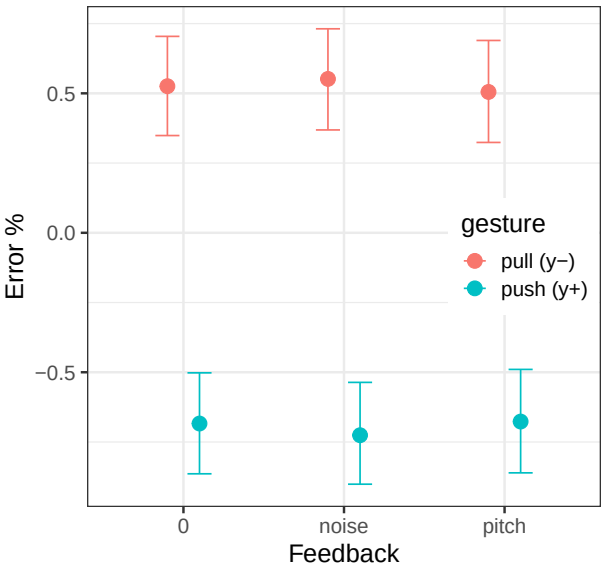


Figure 5. Effetti stimati di gesto e feedback vibrotattile, con intervalli di credibilità al 95%.

anche se con una probabilità di effetto non nullo leggermente inferiore al 95%. Gli effetti stimati sono presentati in Fig. 5.

Similmente, i punteggi di *controllo*, *autostima* e *gradimento* sono stati predetti utilizzando un modello additivo per i seguenti fattori: attributo soggettivo, feedback vibrotattile e gesto, con una intercetta individuale per ogni soggetto. Sulla base della correlazione tra gli attributi nei dati grezzi, questi sono stati modellati come condizioni per le valutazioni soggettive invece che come dimensioni percettive separate.

Anche a causa dell'intervallo limitato, i risultati su scala visuo-analogica hanno prodotto una deviazione dalla norma. Per questo motivo è stata utilizzata una distribuzione beta *zero-one inflated* (ZOIB), la quale modella l'intervallo aperto $(0, 1)$ come una distribuzione beta e gli estremi $\{0, 1\}$ come una una risposta binaria secondo la distribuzione di Bernoulli [17, 18, 19]. Come mostrato in Fig. 6, secondo questi modelli la vibrazione rumorosa ha avuto un effetto negativo sui punteggi, così come l'hanno avuto gli attributi di *autostima* e *gradimento*. Al contrario, il gesto di spinta ha avuto un effetto positivo sui punteggi.

5. DISCUSSIONE

L'accuratezza di intonazione è stata mediamente dello 0.6% – un risultato in linea con le soglie di discriminazione di intonazione per toni con spettro complesso per adulti normo-udenti [4], sebbene i musicisti possano raggiungere un'accuratezza dello 0.1% [5]. Tuttavia, le condizioni di tipo *ecologico* adottate nel presente studio non possono essere comparate a delle condizioni tipiche da laboratorio: il nostro obbiettivo era infatti quello di studiare l'effica-

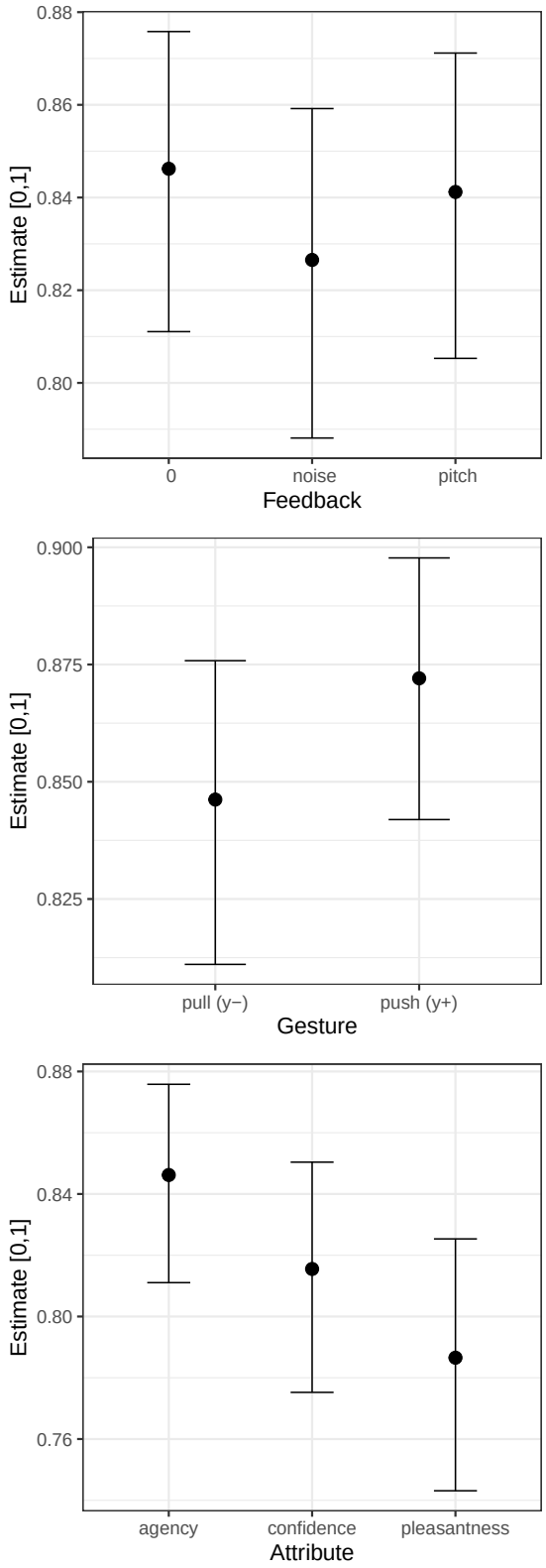


Figure 6. Stima degli effetti condizionali sulle valutazioni soggettive, e relativi intervalli di confidenza al 95% per i fattori feedback vibrotattile, gesto e attributi.

cia di alcuni gesti comuni per il controllo di intonazione in un contesto musicale, e non quello di determinare risultati puramente psicofisici.

Di rilevanza generale è stata la differenza in accuratezza tra spinta e trazione, però accompagnata da un compromesso tra velocità e accuratezza: sebbene leggermente meno accurata, l'esecuzione del compito tramite spinta ha richiesto meno tempo ed è risultata in punteggi maggiori per *controllo*, *autostima* e *gradimento*. Inoltre, gli errori medi si sono equivalsi verso la fine delle sessioni, in quanto i soggetti hanno dimostrato una tendenza all'apprendimento maggiore per la spinta che per la trazione. Alla luce di tutto ciò, il gesto di spinta può essere considerato come più efficiente per compiti che vengono eseguiti frequentemente e richiedono un'esecuzione rapida, come appunto quello di controllare un'interfaccia musicale. Al contrario, se il compito è nuovo o poco frequente, l'accuratezza massima sembra ascriversi al gesto di trazione. In ogni caso, utilizzando gesti unidirezionali il target non viene praticamente mai raggiunto con precisione assoluta. Si può immaginare che una combinazione di entrambi i gesti potrebbe risultare in una accuratezza maggiore. In futuro sarebbe anche interessante associare il gesto di spinta al decremento in frequenza, o viceversa la trazione all'incremento di frequenza [20, 21].

Non sono stati misurati effetti significativi del feedback vibrotattile sull'accuratezza. Rispetto alla condizione in assenza di vibrazioni, la vibrazione intonata ha prodotto errori leggermente inferiori e quella rumorosa leggermente superiori, tuttavia i rispettivi effetti non sono risultati credibilmente non nulli secondo i dati raccolti e i modelli adottati. La verifica di tali tendenze richiederebbe un esperimento con numero di partecipanti di molto maggiore. Ciononostante, la vibrazione rumorosa ha avuto un effetto credibilmente negativo sui punteggi di *controllo*, *autostima* e *gradimento*. A causa del grande numero di parametri da stimare nella distribuzione ZOIB non vi sono stati dati sufficienti per includere possibili interazioni tra gli attributi nel modello statistico delle valutazioni soggettive. I dati grezzi suggeriscono un effetto positivo della vibrazione intonata per il gesto di trazione per tutti e tre gli attributi (Fig. 4); questo potrebbe servire da ipotesi di partenza per un futuro esperimento.

Si sottolinea come nel presente esperimento si siano voluti utilizzare segnali vibrotattili semplici, la cui natura fosse perfettamente correlata con quella del segnale sonoro (vibrazione intonata), o del tutto scorrelata con esso (vibrazione rumorosa). Strategie più complesse sono ovviamente possibili, sebbene in un certo senso più arbitrarie e meno facilmente riconducibili alla natura del segnale sonoro [12]. Inoltre il nostro esperimento ha coinvolto soltanto musicisti con competenze avanzate: una scelta motivata dal nostro obbiettivo generale di fornire basi scientifiche per la progettazione di interfacce musicali di nuova generazione, le quali eventualmente integrino feedback vibrotattile. Tuttavia, la capacità di tali musicisti di controllare l'intonazione può essere stata già sufficiente da consentire loro di eseguire il compito basandosi sul solo feedback sonoro. Rimane da verificare se, proponendo lo stesso compito a studenti di strumento alle prime armi, bambini o persone inabili a percepire l'intonazione, si possano misurare benefici del feedback vibrotattile maggiori: è infatti noto che il rinforzo multisensoriale è più efficace quando l'informa-

zione unisensoriale è poco affidabile [22]. Per i musicisti professionisti potrebbe comunque verificarsi un effetto positivo nelle esecuzioni di musica d'insieme, o nei contesti con elevato rumore di fondo.

6. CONCLUSIONI

Il controllo dell'intonazione tramite spinta e trazione del dito, misurata su giovani musicisti adulti, è risultata dello stesso ordine di accuratezza di quella sulla discriminazione dell'intonazione precedentemente misurata su giovani adulti normo-udenti in esperimenti di laboratorio. Si è riscontrato un leggero *trade-off* per i due gesti analizzati: da un lato errori più contenuti per la trazione e dall'altro maggiore velocità di esecuzione per la spinta. I vari tipi di feedback vibrotattile interattivo hanno avuto solo effetto sulle valutazioni soggettive, ma non sull'effettiva performance. In generale, i nostri risultati possono trovare applicazione nella progettazione più efficace di interfacce musicali controllate tramite gesto.

7. REFERENCES

[1] D. Schwarz, W. Liu, and F. Bevilacqua, "A Survey on the Use of 2D Touch Interfaces for Musical Expression," in *New Interfaces for Musical Expression (NIME)*, (Birmingham, United Kingdom), 2020.

[2] V. Zappi and A. P. McPherson, "Dimensionality and appropriation in digital musical instrument design," in *14th International Conference on New Interfaces for Musical Expression, NIME 2014, London, United Kingdom, June 30 - July 4, 2014*, pp. 455–460, 2014.

[3] S. Papetti and C. Saitis, "Musical Haptics: Introduction," in *Musical Haptics* (S. Papetti and C. Saitis, eds.), pp. 1–7, Springer, Cham, 2018.

[4] B. C. J. Moore and R. W. Peters, "Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity," *J. Acoust. Soc. Am.*, vol. 91, p. 2881, 1992.

[5] M. F. Spiegel and C. S. Watson, "Performance On Frequency-discrimination Tasks By Musicians And Nonmusicians," *J. Acoust. Soc. Am.*, vol. 76, pp. 1690–1695, dec 1984.

[6] R. T. Verrillo, "Vibration sensation in humans," *Music Perception: An Interdisciplinary Journal*, vol. 9, no. 3, pp. 281–302, 1992.

[7] J. A. Harris, E. Arabzadeh, A. L. Fairhall, C. Benito, and M. E. Diamond, "Factors affecting frequency discrimination of vibrotactile stimuli: Implications for cortical encoding," *PLoS One*, vol. 1, no. 1, 2006.

[8] H. Pongrac, "Vibrotactile Perception: Differential Effects of Frequency, Amplitude, and Acceleration," in *2006 IEEE Int. Workshop on Haptic Audio Visual Environments and their Applications (HAVE 2006)*, pp. 54–59, IEEE, 2006.

[9] O. Franzén and J. Nordmark, "Vibrotactile frequency discrimination," *Percept. Psychophys.*, vol. 17, pp. 480–484, 1975.

[10] J. M. Yau, A. I. Weber, and S. J. Bensmaia, "Separate mechanisms for audio-tactile pitch and loudness interactions.," *Front. Psychol.*, vol. 1, p. 160, 2010.

[11] R. Okazaki, T. Hachisu, M. Sato, S. Fukushima, V. Hayward, and H. Kajimoto, "Judged Consonance of Tactile and Auditory Frequencies," in *2013 World Haptics Conf.*, pp. 663–666, 2013.

[12] R. Jack, T. Stockman, and A. McPherson, "Navigation of Pitch Space on a Digital Musical Instrument with Dynamic Tactile Feedback," in *Proc. TEI '16 Tenth Int. Conf. Tangible, Embed. Embodied Interact.*, (New York, NY, USA), pp. 3–11, ACM, feb 2016.

[13] S. Papetti, M. Frohlich, and S. Schiesser, "The TouchBox: an open-source audio-haptic device for finger-based interaction," in *IEEE World Haptics Conf.*, pp. 491–496, 2019.

[14] H. Dai and C. Micheyl, "Psychometric functions for pure-tone frequency discrimination," *J. Acoust. Soc. Am.*, vol. 130, no. 1, pp. 263–272, 2011.

[15] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021.

[16] P.-C. Bürkner, "brms: An R package for Bayesian multilevel models using Stan," *J. of Statistical Software, Articles*, vol. 80, no. 1, pp. 1–28, 2017.

[17] J. K. Kurschke, *Doing Bayesian data analysis - A tutorial with R, JAGS, and Stan*. Academic Press, 2nd edition ed., 2014.

[18] F. Liu and Y. Kong, "zoib: An R Package for Bayesian Inference for Beta Regression and Zero/One Inflated Beta Regression," *R J.*, vol. 7, no. 2, pp. 34–51, 2015.

[19] R. Ospina and S. L. P. Ferrari, "Inflated beta distributions," *Stat Pap.*, vol. 51, pp. 111–126, 2010.

[20] D. Casasanto, W. Phillips, and L. Boroditsky, "Do we think about music in terms of space? metaphoric representation of musical pitch.," in *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2003.

[21] Z. Eitan and R. Timmers, "Beethoven's last piano sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context," *Cognition*, vol. 114, pp. 405–422, 2010.

[22] C. Spence, "Multisensory perception, cognition, and behavior: Evaluation the factors modulating multisensory integration," in *New Handbook of Multisensory Processing* (B. E. Stein, ed.), MIT Press, 2012.

MEMORIA SOLIS, A SUSTAINABLE AND ADAPTIVE MUSICAL INSTALLATION

Imma Battista
Salerno State Conservatoire of Music 'Giuseppe Martucci', Italy
imma.battista@consalerno.it

Silvia Lanzalone
CRM-Centro Ricerche Musicali, Rome
Latina State Conservatoire of Music 'Ottorino Respighi', Italy
silvia.lanzalone@crm-music.it
lanzalone@conslatina.it

Emanuela Mentuccia
CRM-Centro Ricerche Musicali, Rome, Italy
emanuela.mentuccia@crm-music.it

ABSTRACT

The paper describes the work *Memoria Solis*, a sustainable and adaptive musical installation, that transforms its sounds and develops its musical form in relation to sunlight and shadows projected on its surface, composed of a solar panel. The system uses sunlight both for the production of music and for the production of electricity, making the work environmentally sustainable. The music of *Memoria Solis* is made starting from the real-time processing of voice and natural sounds, which follow each other over time and evolve based on the amount of sunlight detected on the solar panel.

Memoria Solis was conceived and realized in 2018 by Imma Battista and Silvia Lanzalone, with the collaboration of Emanuela Mentuccia and the voice of Eleonora Claps at CRM - Centro Ricerche Musicali, Roma. *Memoria Solis* is currently exhibited at Città della Scienza in Naples, Italy.

1. CONCEPT

Memoria Solis was created in 2018 starting from an idea by Imma Battista, who wanted to create a sound installation based on the sun and on the possibility of obtaining data from sunlight, to produce sound and music, as well as the electricity necessary for the life of the opera.

The idea was stimulated by the great environmental problems arising from global warming and the increasingly urgent need to find solutions on a global level for the sustenance of humanity. Art cannot remain impassive in the face of such problems and is shaped around them by proposing solutions in its own processes of interpretation of a changing reality and expression of a universal feeling.

The work had to be able to propose an eco-sustainable system and, at the same time, to communicate hope and confidence in a new vision. The use of a solar panel, in its most typical form, constituted not only a functional element, but also a symbolic one, to be inserted in the work in its entirety, without any manipulation that would alter its structure and recognizability, precisely to strengthen the link with the ecological and ideal purpose that would characterize the work).

Copyright: © 2022 First author's last name et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1.1 The Poem

The connection with the sun and with the several cultural and symbolic references related to it, made indispensable the presence of the word as vehicle and expression of a stratification of meanings, otherwise destined to remain hidden to the user.

The creation of a poem by Imma Battista was therefore a useful choice to make explicit and strengthen the symbolic and expressive contents of the work. The poem had to be able to counterbalance the geometric structure of the solar panel and sublimate the conceptual aspects of the idea behind the work, to bring out evocative elements that could come from the reference to sunlight. It was therefore decided to establish anchorages that could refer to the ancient cultures of the Mediterranean, to its most typical landscapes, to the sensuality of a mythology now disappeared, but still very present in our memory.

The words of the poem are integrally reported in the various contexts in which the work is presented, as in the opening ceremony, in the information panel, or in the demonstration video, but above all they have been manipulated and integrated in the process of composition of the music, as will be described later.

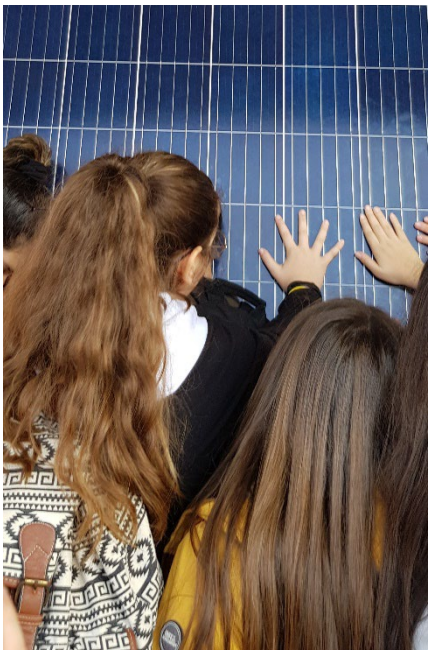


Figure 1: *Memoria Solis* at Città della Scienza in 2018, Neaples, Italy.

2. MUSIC

The music of *Memoria Solis* is made starting from the real-time processing of voice and natural sounds, which follow each other over time and evolve based on the amount of sunlight detected on the solar panel. The musical score, the audio signal processing algorithm, the rescaling, and mapping system of the data related to the amount of sunlight, the sound diffusion project, have been conceived and realized by Silvia Lanzalone.

The aesthetic and technological choices have been made by the author on the basis of the general idea of the work and converge towards the realization of a unitary system, both from the point of view of integration of its parts and from the expressive point of view. The system is able to evolve over time and to adapt to environmental conditions without losing its identity, indeed retaining a strong characterization.

The music has been conceived in such a way as to be able to convey a clear analogy with the characteristics of the sun in summer, the period of maximum presence of sunlight: the dazzling brightness and the intense heat. Both conditions cause a sort of sensory estrangement and alteration of the normal perceptive condition, leading to a different relationship with one's body and one's faculties of thought. Fragments of words taken from the poem confirm a strong analogy with the physical and inner state evoked by it.

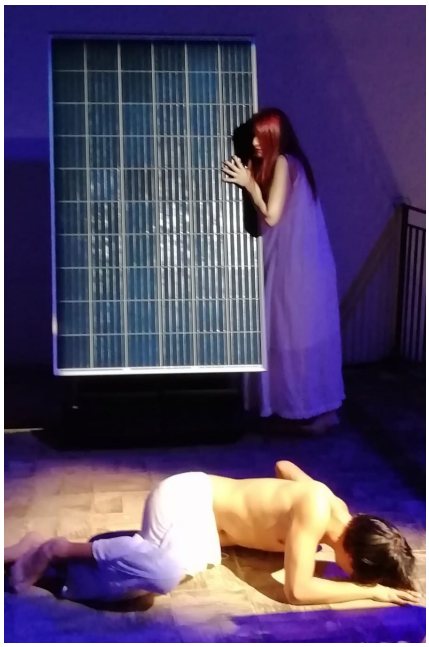


Figure 2: *Memoria Solis,* performative version at opening, Salerno 2018. E. Claps, voice; P. Lara, dance.

The opening of *Memoria Solis* was made in 2018 at Third Biennial of Contemporary Art "Carta Bianca" of Salerno[1] [1] with a performative version of the work, which includes a specific score for voice and live electronics by Silvia Lanzalone. The performer at the voice was Eleonora Claps, and a choreography was created for the occasion by Eleonora Santoro for the dancer Pierpaolo Lara. Both the singer and the dancer interacted with the sounds of the installation, with the lights and a scenographic setting created by the architect Emanuela Mentuccia (Figura 2).

2.1 Musical score and algorithm

The musical score is constantly but slowly renewed through a set of compositional rules that adjust timbres (digital signal processing) and sound articulations (temporal relationships) to the characteristics of light, according to an adaptive criterion. The slow evolution of the sunlight is clearly predominant during most of the time that the light persists and the formation of shadows due to the passage of the audience, or the passage of other moving elements in the surroundings of the work (eg: birds, airplanes, clouds, etc..), allow to obtain small variations and transformations of the sound material.

The digital data resulting from the conversion of the variation of the electrical voltage output from the solar panel are rescaled and reinterpreted with appropriate mapping functions. The data, thus modified, are associated with different parameters of sound processing and music generation. These variations correlate to the emergence of fragments of declaimed poetry, which welcomes the passage of the user as if to "awaken" him from the "dazzling" condition of the sound.

The score is contained in the calculation processes of a Max patch[2], and consists of a complex system of rules (Figure 3). The Max patch is implemented for sound design and music, and is composed by many algorithms of digital signal processing, and provides the following controls:

1. Time, date, and seasons;
2. Sensitivity of the photovoltaic system;
3. Threshold of the detected data;
4. Choice of concrete sounds;
5. Choice of sound processing type;
6. Choosing the amount of sound processing.



Figure 3: *Memoria Solis,* Max patch (main) for sound design and music (score).

[1] Third Biennial of Contemporary Art "Carta Bianca", Palazzo Fruscione, Salerno, October 6th, 2018.

[2] The Max patch is implemented with Max, a software by Cycling '74.

The concrete sounds subjected to selection and processing consist of 24 soundscapes, grouped into 4 morphological categories of 6 sounds each, established on the basis of the degree of internal discontinuity and the level of density:

1. Impulsive sounds (6);
2. Less impulsive sounds (6);
3. Medium continuous sounds (6);
4. Continuous sounds (6).

Voice sounds, on the other hand, are taken from the poem and constitute a fifth category, which is itself divided into different types. The voice interpreter was directed to recite the poem with different interpretations, often to highlight certain sounds of certain phonemes and certain expressions. The typologies are made based on the segmentation on the text and the different interpretations required of the vocal interpreter:

1. Words (320);
2. Verses (32);
3. Stanzas (16);
4. Poems (4).

Both score and sounds are contained and processed through a complex calculation algorithm implemented in MaxMSP. Sound elaborations are directly or inversely proportional to the amount of light detected and involve all categories of soundscapes except voice, which must remain intelligible:

1. Table look-up;
2. Time stretching;
3. Pitch shift;
4. Delay with feedback;
5. Comb filter.

Further parameters concern adjustments inherent to the polyphony of the soundscapes, the flow chart of sound elaborations, the choice of the category of vocal fragments to be heard in the different parts of the day and in the different periods of the year.

The work also decreases in intensity starting from sunset and remains silent during the night hours to then awaken at dawn, at the first appearance of the sun.

2.2 Sound diffusion

The creative application of specific acoustic phenomena allowed to the creation of art sound installations using original technologies based on vibrating surfaces which generate plane waves. This original technology allows the integration of acoustic features and electronic devices and represents one of the main research topics at CRM since 1998. Planofoni® are flat surfaces studied by the CRM of Rome to spread the sound with uniform propagation characteristics in the space, to make the music perceptible and controllable musically vibrational qualities of bodies [2, 3, 4, 5].

The research made at CRM and further experiments made on the producible sonorities using vibrating panels, leads Silvia Lanzalone to creating works for Planofoni® of different materials, such as *Tracciati*, for percussions on Planofoni® and electronics (1998 - 2012), or works for Planofoni® of terracotta sound sculptures of different shapes produced by the visual artist Debora Mondovì, such as *Foglie di Terra e di Suono*, sound sculpture with terracotta plates and electronic sounds (2009) and *Voci d'amore*, electroacoustic art installation on love poems (2014) [5, 6, 7].

The sound diffusion of *Memoria Solis* occurs through the solar panel itself, put into vibration according to the technology used at CRM-Centro Ricerche Musicali for the Planofoni®. The sound is diffused directly from the surface of the solar panel. The audience can feel the music vibrating as their hand touches the work, perceiving its intense heat and causing shadows on its surface. The shadows, detected by computer, in turn, cause the words of the recited poem to emerge.

The panel is vibrated using four electro-dynamic exciters distributed over its surface (Figure 4). The four exciters, designed for stimulating bending waves on plates, have the following technical specifications: Maximum power 50W; Nominal impedance 8Ω; Voice coil diameter 35,6mm; Net weight 0,162kg.

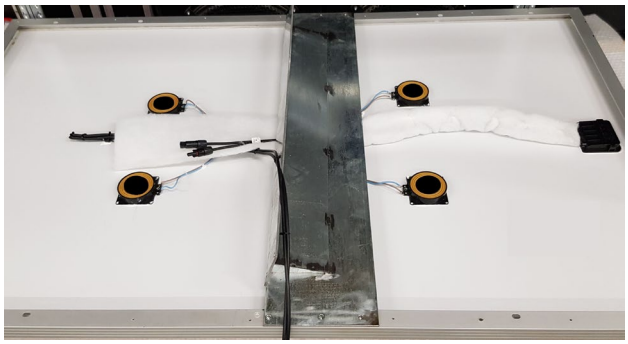


Figure 4: *Memoria Solis,* back of solar panel with four electro-dynamic exciters.

The two mono amplifiers connected to the four exciters have the following technical specifications: Output power 40W RMS/55W peak@8Ω; Frequency response 10 KHz-20KHz/-1 dB; Voltage gain 26dB; Input impedance 20kΩ; THD plus noise 0.03%; Slew rate 19V/µs; Signal-to-noise ratio 92dB.

3. SYSTEM

The system uses sunlight both for the production of music and for the production of electricity, making the work environmentally sustainable.

The photovoltaic system was achieved by the engineering company Sigma Engineering s.r.l., both as an engineering project and as an implementation of hardware components (Figure 5). The production of electricity from the solar panel had to ensure the operation of the computer system and the electroacoustic system, and was achieved by implementing a typical photovoltaic system, with some modifications and additions. The management of the switching on and off of the work during periods of inactivity, or of maintenance, was achieved by means of a special controller external to the computer with a timer adjustable by the user. The controller, built on a PIC system, controls relays which in turn switch the computer and amplifiers on and off. A battery system provides a

reserve supply of energy in case of any problems, or in cases of particular power requirements when there is a lack of sunlight, such as during any nocturnal switching on of the work.

The AD conversion of the DC voltage values of the solar panel was carried out using an Arduino system that sent the data to the serial port for acquisition by the MaxMSP program. Before sending the voltage to the Arduino, it was necessary to adapt the electrical signal to reduce the voltage from about 30 volts supplied by the solar panel to the 5 volts accepted by the Arduino. A cooling system using fans ensures proper ventilation of the technology compartment, which is indispensable especially during the hottest periods, since the work must always be exposed to the sun.

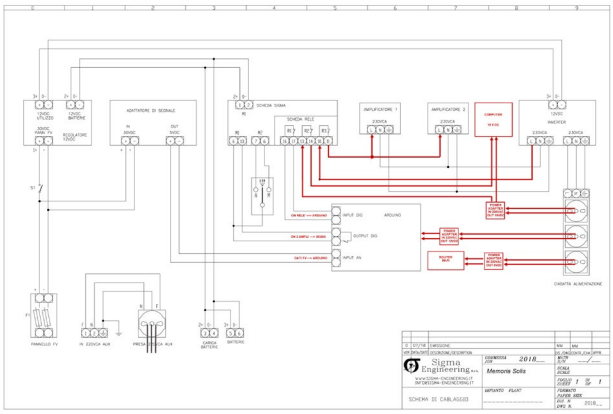


Figure 5: *Memoria Solis,* photovoltaic system wiring diagram, Sigma Engineering s.r.l.

4. VISUAL DESIGN

As in a Rothko painting, the color and the structure of the opera are inseparable. The color itself constitutes the form of the work: "...art to me is anecdote of the spirit, and the only means of making concrete the purpose of its varied quickness and stillness" (Mark Rothko).

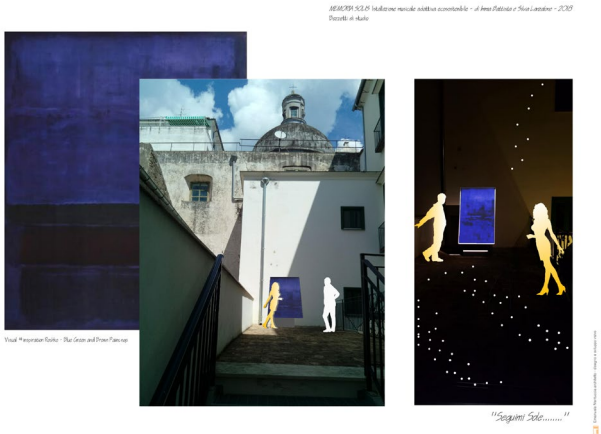


Figure 6: *Memoria Solis,* study sketches by Emanuela Mentuccia. Visual inspiration from Mark Rothko, *Blue Green and Brown* (1952).

The inspiration for the visual development of the work *Memoria Solis* came from the painting by Mark Rothko intitled *Blue Green and Brown* (1952) and from the suggestions and reflections made on his writings (Figure 6): the intimacy and humanity he sought are sublimated by the combination of color, texture and space that together create a unique presence. The sound and the word emerge in the foreground, and at the same time are layered on the surface, enveloping the viewer who is instinctively driven to tactile contact. In this regard, several preliminary studies of the project have been carried out, to analyze some problems related to the use and ergonomics: the photovoltaic panel used has standard dimensions 1650x1000mm, and a thickness equal to 50mm.

The approach to the panel had to be pleasant, not aggressive, and attractive. The inclination given, in addition to functional reasons related to solar radiation, had to contrast its sharp geometric shape. Various hypotheses have been made regarding the height of the positioning, considering the correct transmission of the sound waves transmitted by the Planofono®, to obtain the best listening condition, and the best vision. From the various empirical and practical tests carried out, it was found that the panel, raised 230mm from the ground and with an inclination of 15°, favored both listening and vision. This position allowed the rear part of the opera, containing the technologies, to remain in the background and barely visible from the front.

The work consists of two main elements: a photovoltaic panel, provided with a metal cover behind containing audio technologies (size 1000x1650mm and weight about 25kg) and a support in painted metal carpentry with opening door (size 900x450x750mm) firmly anchored to a base (size 90x80cm), where it is possible to insert internal metal ballasts. The base is also prepared for the outflow of rainwater (Figure 7).



Figure 7: *Memoria Solis* at Città della Scienza in 2019, Neaples, Italy.

The photovoltaic panel is provided with a metal cover behind containing audio technologies and a support in painted metal carpentry with opening door firmly anchored to a base.

The support of the photovoltaic panel in metal carpentry, also acts as a container for the electrical cabinet and is suitably raised from the ground to avoid flooding and the accidental intrusion of small animals. For this purpose, the holes necessary for the air circulation have been suitably screened with narrow mesh nets. The total weight of the complete structure is 165kg and for its formal characteristics and weight distribution is self-supporting and resistant to horizontal and vertical pressures typical of outdoor environments (wind and snow), and accidental pressures of things or people.

The structure has been realized with materials suitable for outdoor installation. The metalwork has been subjected to a galvanizing treatment first, and then to an industrial oven-baked painting, to give it a further superficial resistance to atmospheric agents. The color of the structure was chosen to be an appropriate shade of gray, which would be neutral in the context.

The engineering aspect and the development of the electrotechnical project carried out in collaboration with Sigma Engineering has strongly conditioned the dimensional choices of the container. To contain all the necessary technologies, the electrical cabinet of watertight type has reached conspicuous dimensions. The "heart" of the work resides there.

However, the position of the opera in space was determined above all by the possibility of being adequately affected by the sun. The metallic sheen of the crystalline silicon is deep and recalls light. Both the light and the position of the work with respect to the light are fundamental. The angle of inclination of the photovoltaic panel, and that of the rays emitted by the sun, determine how the electromagnetic radiations strike the surface of the photovoltaic cells of the solar panel. The exposure determines the percentage of electrical energy and the variations of sound parameters.

Memoria Solis is an eco-sustainable opera and somehow intended as an "autonomous organism" placed and collocable in natural contexts, or in artificial built spaces, it had to be adequately protected from any external "attacks" and structurally stable, as mentioned above, and with adequate characteristics to resist shocks, bad weather, extreme temperature variations (very hot and very cold), intense rain, strong wind. It must also have characteristics of durability, ease of cleaning and maintenance. Conceived for the Third Biennial of Contemporary Art of Salerno "Carta Bianca" and placed in an outdoor terrace of the historic Palazzo Fruscione, the opera required a preliminary study of the site according to the exposure and the possibility of being irradiated for enough hours. The ancient palace is, in fact, located in a dense urban mesh in which the sun is not always guaranteed. Moreover, the opera, included in an exhibition route, had to be accessible, protected from environmental noises, and guarantee a correct diffusion of the music coming from it.

Memoria Solis was awarded, in October 2018, the “Città della Scienza" Prize, dedicated to works that combine artistic creativity and technological advancement, and transported, at the beginning of 2019, to the Fondazione Idis - Città della Scienza, Institute for the diffusion and enhancement of scientific culture, in Naples, Italy. A preliminary study of the positioning in outdoor spaces was also carried out in Naples, to give adequate visibility to the work in the Foundation's exhibition itinerary, the most suitable listening space, and the best solar exposure conditions.

The work *Memoria Solis* is designed to be transportable: it can be disassembled and reassembled in its individual parts. It can also be installed in any environment where it can be reached by natural light and monitored remotely via an internet connection.

5. CONCLUSIONS

Scientific and technological progress deeply pervades society and culture, including art in all its manifestations, enriching it with previously unseen expressive possibilities. The convergence of different art forms for the creation of multisensory works leads to a renewal of the ways of enjoyment of the public, which from simple spectator becomes an active participant.

Memoria Solis is an eco-sustainable adaptive musical installation that interprets sunlight and transforms it into music. The music, a ductile, iridescent, and immaterial element, can adapt to the environmental conditions and to the actions of the audience, evolving with them in order to take on changing but coherent forms, thanks to an adaptive algorithm and an integrated electroacoustic system. *Memoria Solis* is an adaptive art installation, since it is able to condition the evolution of music, preserving its own formal and expressive identity, even in view of elaborations occurred over time, and the interactions with the visitors and surroundings. The movement of shadows in the surrounding environment produces the emergence of a voice that declines fragments of poetry, full of memory and suggestions. The "survival" of the opera, from the point of view of the electrical sustenance necessary for the functioning of the technological apparatus, is allowed thanks to its eco-sustainable conception that self-generates energy.

In 2020, during the lockdown for the Covid-19 Pandemic, *Memoria Solis* continued to spread sound and fragments of poetry in the empty, almost metaphysical space, as an autonomous organism, like a tree or any other natural element. The space, which in that moment was deprived of the human being, was filled with the humanity poured into the opera, that became, in that moment, a symbol of hope. Art that looks towards the future is able to bring man towards a healthy relationship with the environment through technology and the use of renewable sources, restoring an "ethical" dimension to aesthetics.

6. ACKNOWLEDGMENTS

Memoria solis was conceived and realized in 2018 by Imma Battista (concept idea and poem), Silvia Lanzalone (adaptive music and system), with the collaboration on Emanuela Mentuccia (project and visual design), and the vocal contribution of Eleonora Claps (recorded voice). The production was made possible thanks to the contribution of Conservatorio Statale di Musica "G. Martucci" di Salerno (Salerno State Conservatoire of Music), CRM-Centro Ricerche Musicali, Rome (Centre for Music Research), Sigma Engineering S.r.l. (Ltd).

Memoria Solis was created for Terza Biennale di Arte Contemporanea di Salerno 2018 "Carta Bianca" (Salerno Contemporary Art Exhibition 2018) and received the prize "Città della Scienza" 2018 (City of Science), devoted to the works that merge artistic creativity and technological advancement.

Memoria Solis was exhibited in the Futuro Remoto 2019 event at Città della Scienza in Naples, Italy (Figure 1).

Memoria Solis is currently exhibited at Fondazione Idis, Città della Scienza in Naples, Italy.

7. REFERENCES

[1] AA.VV., Terza biennale d'arte contemporanea di Salerno 2018 "Carta bianca". Catalogo. Ed. Crosslink, Salerno, Italy, 2018.

[2] M. Lupone, L. Bianchini, S. Lanzalone, and A. Gabriele. "Research at Rome's Centro Ricerche Musicali on Interactive and Adaptive Installations and on Augmented Instruments", in Computer Music Journal, MIT Press, Cambridge USA, 2021. 44 (2-3), pp.133–158.

[3] M. Lupone, L. Bianchini, S. Lanzalone, and A. Gabriele. "The Research at CRM - Centro Ricerche Musicali: Interactive and Adaptive installations and Augmented Instruments (Feed-Drum®, SkinAct, Windback, ResoFlute)", in Proceedings of ICMC - International Computer Music Conference, New York, USA, 2018.

[4] L. Seno. "Nuova liuteria: caratterizzazione acusto-vibrazionale dei Planofoni®: un nuovo strumento (e concetto) per la musica e l'arte contemporanee", in Proceedings of Annual Meeting of the Acoustical Engineering Society. Como, Italy, 2005.

[5] S. Lanzalone. "Strumenti aumentati", in Acustica. Fondamenti e applicazioni, a cura di R. Spagnolo, UTET, Torino, 2015.

[6] S. Lanzalone. "Suoni Scolpiti e Sculture Sonore: alcuni esempi di installazioni d'arte elettroacustica", in Proceedings of XVII CIM - Colloquio di Informatica Musicale. Venezia, Italy, 2008.

[7] S. Lanzalone. "Suoni scolpiti / sculture sonore. Un percorso (est)etico tra re-incarnazione e inte(g)razione", in Le arti del suono, semestrale, Numero 2 / 2009, installazioni sonore e altri ascolti, pp. 109-132, Edizione Orizzonti Meridionali, Cosenza, Italy, 2009.

Una scheda di elaborazione musicale riprogrammabile e controllabile attraverso un ambiente di live coding

Stefano Silvestri

Dip. Nuove Tecnologie e Linguaggi Musicali,
Conservatorio di musica Giuseppe Martucci - Salerno
stefano.silvestri@consalerno.it

SOMMARIO

L'articolo descrive la realizzazione di una scheda di sviluppo a basso costo dedicata alla sintesi ed elaborazione di segnali musicali lo-fi e il suo controllo mediante un ambiente di programmazione appositamente scritto per il live coding. Il progetto sfrutta la piattaforma a microcontrollore Arduino come processore per l'elaborazione digitale e consente l'assemblaggio di una struttura modulare con elementi collegabili via patch e riprogrammabili come operatori sia di sintesi che di controllo del suono. Mediante una semplice interfaccia computer è possibile connettere i moduli all'ambiente di programmazione sviluppato per definire operazioni di gestione delle strutture musicali in live coding con SuperCollider e secondo una sintassi JavaScript.

1. INTRODUZIONE

Come è noto, nei principali centri storici di sperimentazione musicale elettroacustica e durante il secondo periodo di sperimentazione dell'informatica musicale, alcuni sistemi di generazione e trasformazione del suono vengono progettati seguendo criteri di modularità: a partire dagli apparati analogici [1] e con la successiva introduzione di sistemi ibridi, formati da generatori interfacciati ad unità di calcolo [2,3,4], emergono i presupposti per la realizzazione delle prime architetture digitali dedicate alla computazione dei campioni sonori in tempo reale [5,6]. Sistemi strutturati secondo una logica simile sono di riferimento, tra l'altro, per gli attuali sintetizzatori modulari e per la sperimentazione di nuovi dispositivi digitali a basso costo controllabili "dal vivo" con moderni criteri di programmazione interattiva.

In tal sede il concetto di live coding, ampiamente discusso e analizzato sotto il profilo tecnico e musicologico in [7,8], viene applicato come front-end per la gestione di generatori sonori configurabili con modalità alternative rispetto agli standard definiti dall'industria dei sintetizzatori hardware modulari[1]. L'obiettivo è di presentare un possibile schema di unità hardware riprogrammabili e una modalità astratta di controllo interattivo via codice sorgente.

Copyright: © 2022 Stefano Silvestri. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

2. LA PIATTAFORMA "SYNTHBOARD"

Il sistema proposto di seguito[2] sfrutta la scheda Arduino nelle versioni Pro Mini e NANO come unità di elaborazione digitale basandosi quindi su microcontrollore AVR Atmel[3] modello ATmega 328P-AU. Di seguito si riporta una tabella riepilogativa delle caratteristiche di queste schede a confronto con la versione principale Arduino UNO Rev.3:

Arduino	µCU	CPU	I/O	Dim (mm)	Digital in /ADC pin /PWMout	Flash [kB]	UART/ USB
UNO	328P-PU	16 MHz	5V	68.6x 53.4	14/6/6	32	2.0B
NANO	328P-AU	16 MHz	5V	43.1x 18.5	14/8/6	16 32	Mini-B
Pro-mini	328P-AU	8/16 MHz	5/3.3 V	33x 17.8	14/8/6	32	no

Tabella 1. Caratteristiche delle schede Arduino

Dopo alcuni test con prototipi realizzati in vari formati si è scelto di impiegare un Arduino NANO come compromesso intermedio tra costo/prestazioni e proprietà dell'hardware. Grazie alla disponibilità di una libreria DSP[4] per algoritmi musicali [9], molto efficiente e pratica da impiegare in ambiente Arduino[5], è possibile generare segnali audio a bassa qualità con frequenze di campionamento di 16.384KHz o 32.768KHz su uscita pseudo-analogica in modulazione PWM, in base alle specifiche dell'hardware AVR a 8-bit. Diversi parame-

[1] Per un'ampia trattazione dell'ingegneria degli strumenti musicali elettronici si vedano le "Electronotes" di Bernie Hutchins (Cornell University, 1972) alla pagina: https://sndmt.link/electro-notes. Son inclusi contributi di: Robert Moog, Don Buchla, Serge Tcherepnin, etc.

[2] Il progetto è rilasciato con licenza open-source (CC per la parte hardware).

[3] Famiglia di µCU (1996, Atmel; 2016, Microchip Technology) basata su architettura ad Harvard modificata, RISC 8-bit, SRAM dati, EEPROM per dati permanenti e flash di memoria programma.

[4] Si fa riferimento al link: https://sensorium.github.io/Mozzi/.

[5] L'Arduino IDE è disponibile al link: https://www.arduino.cc/. Il software è realizzato in java e deriva del linguaggio di programmazione Processing e dal progetto Wiring.

tri come: modalità audio, sample e control rate, output mono/stereo, abilitazione dell'ingresso, sono impostabili nel file di definizione "mozzi_config.h" della libreria. Il progetto qui esposto, denominato "SynthBoard", è accompagnato da una utility software C/C++ scritta per facilitare l'I/O utente ed è principalmente strutturato sulle possibilità della scheda Arduino NANO congiuntamente a quelle del core audio[6].

2.1. Specifiche del sistema di conversione audio

La scheda digitale realizza una conversione audio analogica con tecnica PWM permettendo di generare segnali entro una banda che va da ~1Hz fino a ~8kHz impostando una frequenza di campionamento di 16384Hz. Essendo la frequenza di interrupt e la profondità di bit del campione interdipendenti a causa del modo in cui viene utilizzata la PWM per l'uscita audio, si ottiene un buon compromesso di elaborazione definendo i seguenti valori dei parametri nel file di configurazione "mozzi_config.h" della libreria:

```
#define AUDIO_MODE STANDARD_PLUS
#define AUDIO_RATE 16384
```

Diversamente da un DAC a rete di resistori pesati la risoluzione del convertitore PWM è in tal caso funzione del duty-cycle dell'onda rettangolare generata attraverso i Timer del micro e segue la relazione [10]:

$$\frac{f_{CLK}}{f_{PWM}} = 2^N \qquad (1)$$

La libreria audio occupa il Timer1 che fa capo ai pin 9 e 10 di Arduino e la frequenza della sorgente di clock del timer, utilizzata come riferimento di incremento del contatore, è f_{CLK}=16MHz fornita dall'oscillatore a cristallo esterno. La frequenza di PWM dipende dalla frequenza base f_{BASE} del contatore e dall'impostazione del valore del prescaler, nel caso considerato, sul relativo registro TCCR1 del Timer1[7], si ha f_{BASE}=31250Hz secondo le specifiche del datasheet [11]. Con un valore di prescaler PR_{CLK}=1, seguendo la relazione (1), si ha:

$$2^N = \frac{F_{CLK}}{f_{BASE}/PR_{CLK}} = \frac{16\,MHz}{31250\,Hz} = 512 \qquad (2)$$

ovvero il segnale di uscita sarà definito da un insieme di 512 valori possibili di quantizzazione rappresentabile con un numero di bit pari a:

$$N_{bit} = \lfloor \log_2(2^N) \rfloor = \lfloor \log_2(512) \rfloor = 9\,bit \qquad (3)$$

Considerando un valore picco-picco della tensione di uscita V_{pp}=5V la risoluzione teorica del convertitore risulta:

$$V_{RES} = \frac{V_{pp}}{2^{N_{bit}} - 1} = \frac{5\,V}{511} = 0.00978\,V \qquad (4)$$

Ovvero l'intervallo tra due livelli consecutivi di tensione di uscita vale all'incirca 9.78mV. Verranno sfruttate queste caratteristiche per la quantizzazione dei segnali.

Secondo il modo STANDARD_PLUS dalla libreria i valori di uscita dal processo updateAudio() sono rappresentati da interi compresi nell'intervallo di computazione [-244, 243] quindi con 488 valori complessivi di risoluzione dei campioni si realizza una headroom artificiale al di sopra degli 8-bit di risoluzione impiegati dalle tabelle degli oscillatori per la wavetable lookup.

2.2. Realizzazione della parte hardware

Lo schema elettrico della sezione principale di elaborazione della scheda è il seguente:

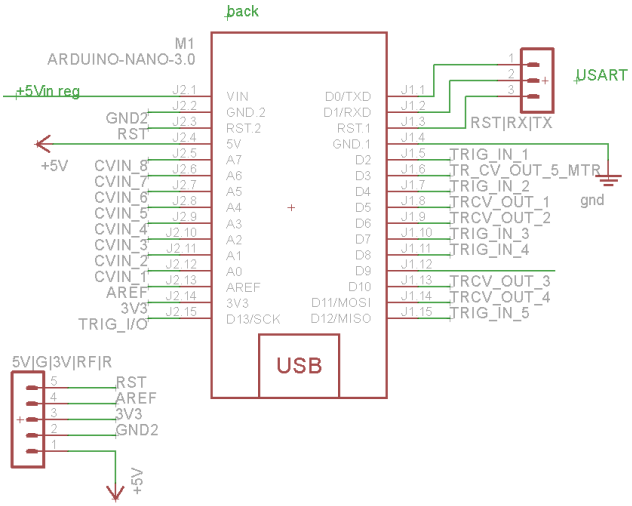


Figura 1. Unità di elaborazione basata su Arduino NANO V3.0 – 5V.

La fig.1 mostra i collegamenti con le sezioni di I/O digitale, definite come GATE/TRIG di input/output, e con quelle PWM, definite come Control Voltage (CV) a 5V in ingresso e uscita. Sono inoltre mostrati i collegamenti con il circuito di output audio (pin D9) e con quello di alimentazione (pin VIN).

Il circuito di input CV supporta un segnale di controllo nell'intervallo [0,5V] definibile dai potenziometri KNB1–8 oppure mediante la sezione CVIN1–8 accessibile dal pin header sul pannello frontale del circuito stampato. Di seguito il dettaglio di uno degli otto circuiti analoghi di CV input:

[6] La libreria Mozzi consente una modalità di computazione audio chiamata STANDARD_PLUS che impiega il Timer1 a 16-bit, quindi 2^{16}-1=65535 cicli di conteggio possibili, con uscita audio PWM sul pin 9. Per ridurre il rumore di fondo e aumentare la precisione di conversione è previsto il modo HIFI che sfrutta Timer1 e Timer2 a 8-bit mentre l'uscita audio è sui pin 9 e 10 secondo la tecnica Dual PWM. Per non bloccare il processore la classica funzione analogRead() dell'ADC a 10-bit è sostituita con mozziAnalogRead() nel ciclo updateControl() dello sketch.

[7] Per default il valore di prescaler è 64. Essendo la frequenza base 62500Hz (T=16µs) per il Timer0 e 31250Hz (T=32µs) per i Timer1 e Timer2 si ottengono rispettivamente frequenze di PWM pari a 62500Hz/64≈976.56Hz sui pin ~5 e ~6 mentre si ha 31250Hz/64≈488.28Hz sui pin ~9,~10 e ~3,~11. I divisori impostabili sono 1,8,64,256,1024 per i pin 5,6,9,10 con aggiunta di 32,128 sui pin 3,11.

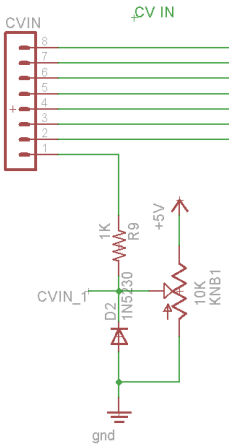


Figura 2. Circuito posto agli ingressi di CV input per ogni potenziometro di controllo e su CVIN1–8.

Il diodo zener 1N5230 da 4.7V e il resistore da 1K formano un circuito di protezione per tensioni diverse da 5V. La sezione che può ricevere segnali di Gate e di Trigger dall'esterno è simile:

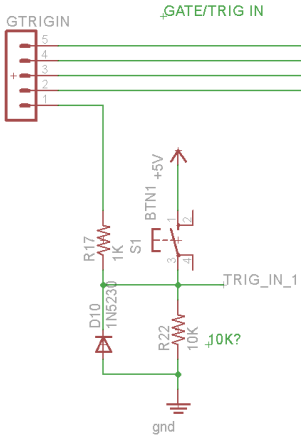


Figura 3. Circuito posto sugli ingressi di Trigger/Gate TRIGIN_1–5 e su ogni pulsante BTN1–5 della scheda.

Il segnale elettrico esterno definibile in ingresso ai TRIG_IN_1–5 consiste in un livello logico alto +5V o basso 0V impostabile anche sulla scheda attraverso l'header GTRIGIN_1–5. Segue la sezione dedicata alle uscite PWM nel range 0-5V oppure, in misura definibile dall'utente, per segnali Gate/Trigger digitali di I/O e che fa capo anche ai LED1–5 sul PCB:

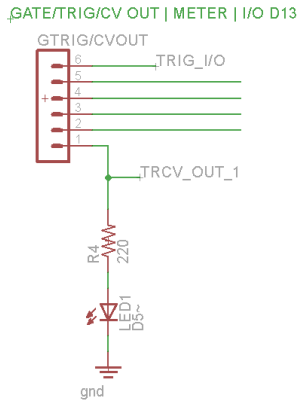


Figura 4. Circuito per segnali di Gate/Trigger in output e input digitale o per realizzare uscite di controllo PWM da TRCV_OUT_1 a TR_CV_OUT_5_MTR.

Il pin D3~ di Arduino è connesso al LED 5 e consente opzionalmente di realizzare un semplice indicatore visivo del livello del segnale di output relativo al pin audio D9~. Collegato a quest'ultimo pin si trova il circuito dell'uscita audio costituito da un filtro passivo di ricostruzione e da un buffer adattatore di impedenza realizzato con un LM358N funzionante con alimentazione singola [12]:

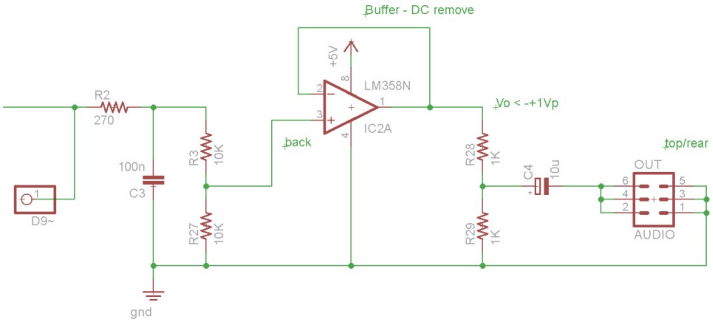


Figura 5. Circuito dell'uscita audio realizzato con una rete di filtraggio RC ed uno stadio buffer ad operazionale.

Il filtro passivo con risposta di tipo passa-basso del 1°ordine è dimensionato al fine di ridurre il contributo della portante a 16384Hz del segnale di campionamento. Con un valore R2=270Ω e C3=0.1μF la frequenza di taglio risulta:

$$f_{cut} = \frac{1}{2\pi RC} = \frac{1}{2\pi\, 270\,\Omega \cdot 100\,nF} =$$

$$= \frac{1}{2\pi\, 270\,\Omega \cdot .0000001\,F} \approx 5894.627\,Hz \quad (5)$$

Ovvero un valore a -3dB di circa 6kHz con pendenza della curva di risposta del filtro di -6dB/ottava. Per possibili operazioni di controllo il segnale di uscita è prelevabile anche prima della rete di filtraggio e del buffer. La funzione del buffer, realizzato con un integrato a doppio operazionale, è quella di creare un segnale "true analog" in uscita mediante un singolo OpAmp in configurazione non invertente[8] che, attraverso i partitori resistivi sui pin 3 e 1, regola il livello del range [0,5V] riportandolo al valore limite di $1V_p$ al fine di rendere il segnale compatibile con il livello operativo standard per ingressi di linea[9]. Attraverso questo circuito e con l'aggiunta del condensatore di disaccopiamento C4 in uscita, il segnale viene centrato sullo 0V e si elimina il DC-offset.

Nel caso di collegamenti multipli dell'I/O è opportuno ricordare che il fan-in/out del micro non supera 40mA con una impedenza di ~20KΩ sui diversi pin.

Si conclude la descrizione dello schema con un circuito tale da garantire un'alimentazione stabilizzata, soprattutto nel collegamento di più schede in parallelo:

[8] Un voltage follower non invertente a guadagno unitario.

[9] Diversamente, senza prevedere un adattamento simile, il segnale 0÷5V dell'uscita PWM di Arduino può compromettere i circuiti preamplificatori dei canali di ingresso al mixer o gli stadi attivi di crossover ed i trasduttori dei diffusori acustici.

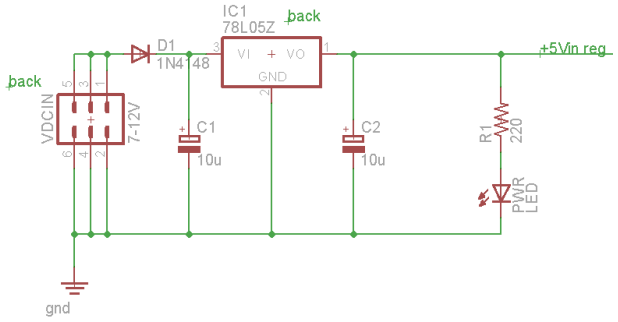


Figura 6. Circuito di regolazione dell'alimentazione a 5VDC con regolatore lineare 78L05, diodo 1N4148 contro le inversioni di polarità e LED di segnalazione dell'accensione.

La singola scheda può essere alimentata in modo ottimale con una tensione di 9V per un assorbimento di pochi mA (<500mA). Il connettore multiplo, posto fisicamente sul retro, consente di creare un ponte di alimentazione per altre schede del sistema modulare.

La figura seguente mostra un test di generazione prodotto con un semplice algoritmo di sintesi additiva statica a nove oscillatori:

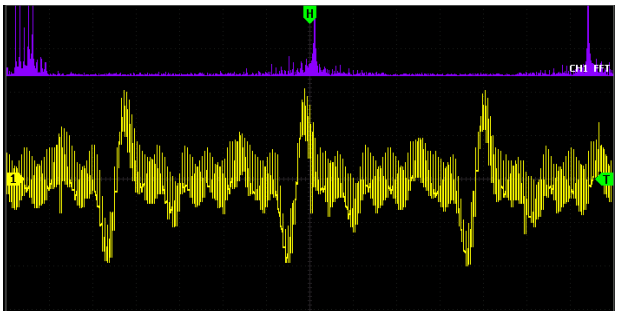


Figura 7. Analisi all'oscilloscopio del segnale risultante prodotto dall'algoritmo di sintesi additiva a nove oscillatori sinusoidali.

L'esempio è stato realizzato con il listato "additive_static_9_osc.ino" la cui sezione principale del codice è riportata di seguito:

```
Oscil <SIN2048_NUM_CELLS, AUDIO_RATE> osc1(SIN2048_DATA);
...
Oscil <SIN2048_NUM_CELLS, AUDIO_RATE> osc9(SIN2048_DATA);
int out;
void setup() { startMozzi(CONTROL_RATE); }
void loop() { audioHook(); }
void updateControl() {
  int f0=500; float k=knb1();//0.3876;
  osc1.setFreq(f0*2*k); osc2.setFreq(f0*3*k);
  osc3.setFreq(f0*4*k); osc4.setFreq(f0*5*k);
  osc5.setFreq(f0*6*k); osc6.setFreq(f0*7*k);
  osc7.setFreq(f0*8*k); osc8.setFreq(f0*9*k);
  meter(out);
}
int updateAudio() {
  out = (osc1.next() + osc2.next() + (osc3.next()>>1)
    + osc4.next() + osc5.next() + (osc6.next()>>2)
    + (osc7.next()>>2) + (osc8.next()>>2)
    + (osc9.next()>>3));
  return out>>2;
}
```

Figura 8. Frammento del listato Arduino che consente di programmare la scheda come un banco di nove oscillatori digitali limitati in banda audio.

Le classi audio derivano da `#include <MozziGuts.h>` mentre funzioni come `meter()` e `knb1()` sono parte dell'apposita libreria `lowerUtils.cpp/.h` la quale facilita l'accesso allo stato e al valore di: potenziometri, pulsanti, LED, input/output CV in PWM e segnali di Gate/Trigger da e verso altre schede analoghe.

In fig.7 il livello V_{pp} del segnale di uscita è circa ~791mV e l'ampiezza delle componenti armoniche pari e dispari ($\forall$ k≥1, k∈ℕ; escluso f_0) viene riscalata mediante l'operatore bitwise di right shift sul valore restituito dai metodi *.next()* nell'*updateAudio()*. Gli oscillatori wavetable impiegano delle tabelle di campioni precaricati nel file `<tables/sin2048_int8.h>`. Il `meter()` calcola `abs(sig)` con PWM indipendente su D3~ definito in `"setPwmFrequency.h"`.

Di seguito si mostrano un modello 3D del PCB dual layer della scheda[10] e la sua realizzazione (fig.10):

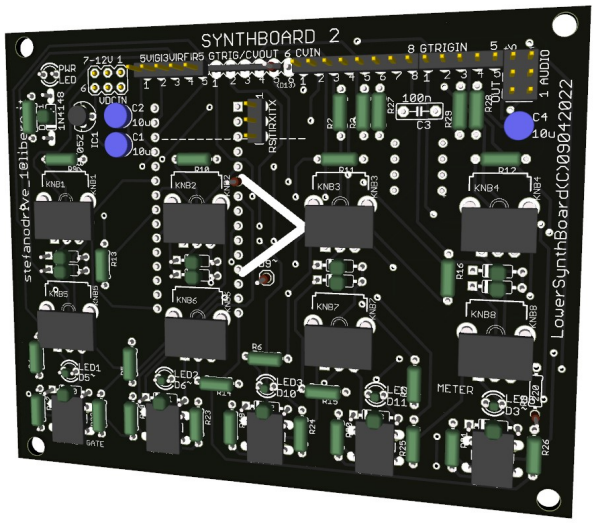


Figura 9. Modello in 3D della scheda "SynthBoard v.2" ricavato a partire dal file Gerber di EagleCAD.

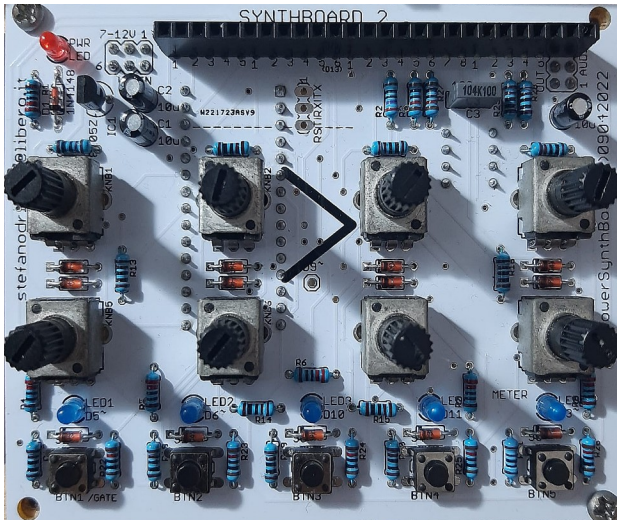


Figura 10. Scheda "SynthBoard v.2" con socket per Arduino NANO montato sul retro.

[10] Lo schema elettrico è realizzato in EagleCAD: https://www.autodesk.com/products/eagle; la visualizzazione 3D del circuito stampato è in 3D Gerber Viewer: https://www.zofzpcb.com/, versioni freeware.

In fig.10 sono evidenti le strip di pin per la realizzazione dei collegamenti del patch con cavi jumper e gli alloggiamenti dei componenti per i controlli manuali. Le specifiche tecniche di questa versione di scheda sono riassunte di seguito:

- 8 potenziometri
- 8 input CV nel range 0÷5V
- 5 pulsanti
- 5 input Gate/Trig con stati 0 o 5V
- 5 LED
- 1 LED impostabile come Meter di output
- 5 Trig/CV PWM out
- 1 Gate/Trig configurabile come I/O (pin 13)
- 3 Audio out (pin D9~)
- 3 Connessioni di alimentazione
- 1 Strip USART TX/RX/RESET
- 1 Strip 5V/GND/3.3V/AREF/RESET

Ogni controllo è programmabile dall'utente nei limiti delle possibilità di elaborazione della scheda e delle combinazioni possibili realizzate per collegamenti tra più schede.

3. L'AMBIENTE "STRING"

L'ambiente di sviluppo denominato "STRING"[11] è stato originariamente pensato per controllare un generatore sonoro hardware da computer, attraverso processi di live coding [13]. Parallelamente alle sperimentazioni con la parte hardware sono state integrate numerose funzionalità che rendono il software un tool di supporto scalabile e utile per la programmazione del codice dal vivo in ambiente SuperCollider[12]. L'idea alla base del progetto è quella di definire un framework e uno script DSL (Domain-Specific Language) aperto ed integrabile con nuove funzionalità per la scrittura di codice sorgente in tempo reale orientato a vari livelli della struttura musicale.

3.1. Indicazioni sul concetto di Live Coding

E' possibile definire il live coding musicale come una forma di improvvisazione, legata alle pratiche di composizione algoritmica, dove l'elaborazione in tempo reale e la scrittura stessa degli algoritmi avvengono allo stesso tempo. Si parla in tal caso anche di programmazione "on-the-fly" [14] indicando uno stile in cui il compositore modifica le istruzioni mentre il programma è in fase di esecuzione[13]. L'idea è che la manipolazione degli algoritmi divenga "un mezzo ulteriore di espressione musicale". Alcune indicazioni utili al progetto di un ambiente di live conding e all'uso di linguaggi di programmazione come strumento musicale sono fornite in [15,16,17,18] mentre alcune fonti che delineano un possibile percorso storico riguardano: le prime performance di computer music di Pietro Grossi [19], i sistemi in tempo reale di Giuseppe Di Giugno e del gruppo Ac.El. [5], in cui era implicito un controllo a run-time, la relazione tra pratiche del live coding e strumenti elettroacustici di Hugh Davies [20], le sperimentazioni dei brani *Audio Wave* (1980) di John Bischoff e *Waterfront*[14] (1985) di Ron Kuivila [7].

La costruzione di modelli algoritmici in live coding fa riferimento alle prime sperimentazioni di Gottfried Michael Koenig, Iannis Xenakis, Laurie Spiegel, etc. dove il processo formale è applicato come principio generatore della forma musicale. Le strategie possibili per il design di un modello compositivo si basano quindi su regole matematiche e su processi formali di rappresentazione[15]. In particolare nell'interessante lavoro di Spiegel[16] si trova la definizione di pattern [21] inteso come una sequenza generica di informazioni che può essere soggetta a trasformazioni[17].

Attualmente i linguaggi di live coding disponibili[18] implementano differenti paradigmi orientati alla descrizione, combinazione e composizione di pattern, correlati alla definizione del codice di programmazione audio. L'ambiente di sviluppo proposto in tal sede implementa le suddette operazioni di strutturazione dei parametri del suono, tuttavia orientandosi verso ulteriori paradigmi che possano estendere, in diverse direzioni, alcuni concetti oltre quelli classicamente ricorrenti di: "infinite spiral timeline", "cyclical pattern" [22], "generative event/time-based content" e "Threnoscope" [23,24].

3.2. Realizzazione del software

La realizzazione del programma di live coding ha fatto riferimento alle seguenti necessità: piena compatibilità con il linguaggio SuperCollider, sintassi di programmazione JavaScript like, opzioni di ridefinizione della sintassi, compressione del codice per una maggior rapidità e modularità, composizione al livello della macrostruttura e microstruttura sonora, accesso a periferiche hardware esterne. Il sistema basa il suo funzionamento sulla struttura client-server di SuperCollider e necessita della classe "SSC.sc" appositamente scritta per la definizione dell'interprete, sfruttando la proprietà di reflection del linguaggio[19]. Una volta incluso questo file dal menu Edit/Preferences/Interpreter nell'ambiente

[11] Il nome rimanda alla sua semplice funzione di inviare stringhe di testo, via protocollo OSC, ad un interprete scritto in linguaggio SuperCollider.

[12] Pagina ufficiale: https://supercollider.github.io/.

[13] Ciò comporta la necessità di un meccanismo non bloccante per possibili interruzioni di compilazione come errori logici e di sintassi.

[14] Realizzato con il sistema Mountain, costituito da personal computer e un banco di 16 oscillatori analogici, rappresenta la prima performance di live coding documentata, peraltro interrotta da un crash del sistema operativo.

[15] Si tratta di: modelli deterministici, serie, gruppi, processi combinatori, modelli stocastici, distribuzioni di densità di probabilità, probabilità condizionata, rapporti di peso statistico, maschere di tendenza, catene di Markov, processi di interpolazione/estrapolazione, principi cronologico-associativi, strutture a blocchi.

[16] E' significativo il suo impegno presso i Bell Laboratories alla parte di controllo algoritmico per il sistema GROOVE (Generated Real-time Output Operations on Voltage-controlled Equipment) di Max Mathews e F. Richard Moore.

[17] Operazioni fondamentali di: sfasamento, inversione, trasposizione, rotazione, riscalamento, interpolazione, frammentazione, estrapolazione, sostituzione interparametrica, combinazione, iterazione.

[18] Si vedano i principali ambienti di live coding audio/video e code sharing: TidalCycles, Sonic-pi, FoxDot, Limut, Gibber, Ixi lang, etc.

[19] Si sfrutta la capacità di SuperCollider di eseguire codice che ha come oggetto istruzioni di elaborazione definite nello stesso linguaggio SuperCollider.

Scide di SuperCollider è possibile abilitare la comunicazione OSC[20] tra l'interprete "SSC.sc" e STRING. Viene di seguito riportato il diagramma della struttura complessiva dell'ambiente software:

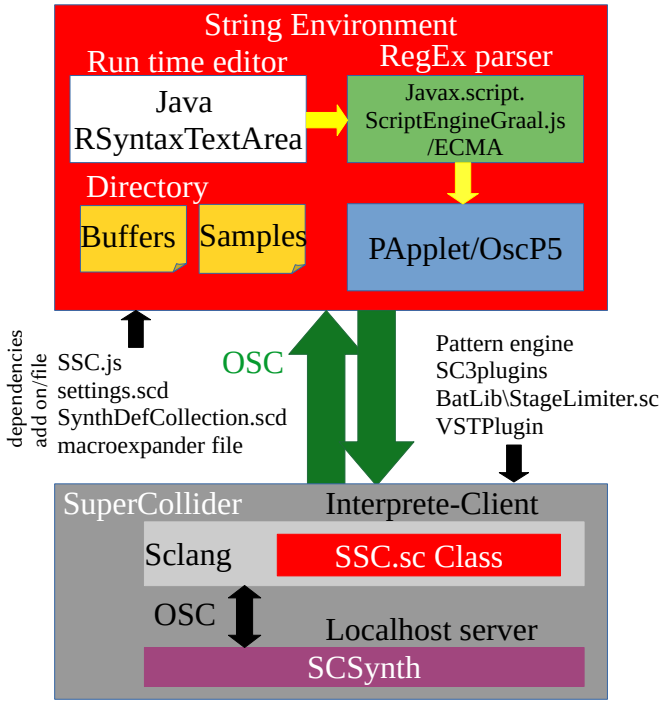


Figura 11. Schema della struttura complessiva dell'ambiente di programmazione STRING_v5.

Il progetto è scritto in Java[21] e si compone di una classe principale "String_IDE.java" che impiega la versatile classe `RsyntaxTextArea`[22] per la realizzazione dell'editor implementato nel file "RSTA_Editor.java". Attraverso la combinazione di tasti "Ctrl+e" le istruzioni di un blocco di codice, delimitato da righe vuote, subiscono un processo di parsing mediante l'uso delle regular expression, sostituendo le parole chiave del linguaggio di script "String" con corrispondenti function in linguaggio JavaScript definite nel file "SSC.js". Tali funzioni possono essere ridefinite dall'utente e supportano le operazioni principali di gestione degli Eventi, NodeProxy e PatternProxy in ambiente SuperCollider.

In origine era stato creato un motore di gestione dei pattern facendo riferimento a Pattern lato client, ProxySpace e altre funzionalità della JITLib[23], combinando gli Event Stream Pdef e Pbind[24]. Successivamente, come pattern engine, è stato importato un Quarks scritto dallo sviluppatore software canadese Glen Fraser[25] e ispirato al linguaggio Tidal Cycles, utile per descrivere schemi e frasi in una notazione compatta, compatibile inoltre con le extension VSTPlugin e ChordSymbol.

Diversamente dalle prime versioni del progetto[26], il riquadro "ScriptEngineGraal.js/ECMA" (fig.11) si occupa di valutare codice in standard ECMAScript 262, diviene quindi possibile combinare istruzioni JavaScript con i comandi opportunamente tradotti dal parser sintattico e definiti attraverso il motore di gestione dei pattern per SuperColldier. L'invio degli algoritmi e la ricezione dei risultati avvengono dinamicamente dato che ogni esecuzione è istanziata in thread separati di StringBuffer Object di Java. Le istruzioni vengono così inviate via protocollo OSC impiegando la libreria oscP5 del linguaggio Processing[27]. Il flusso delle informazioni è quindi il seguente: *codice String→parsing→link alle istruzioni del pattern engine→invio OSC→interpreter in SC*. La figura seguente mostra l'interfaccia dell'ambiente testuale di programmazione, con la relativa MenuBar e una PostWindow contenente il risultato della conversione del codice string nelle istruzioni SC:

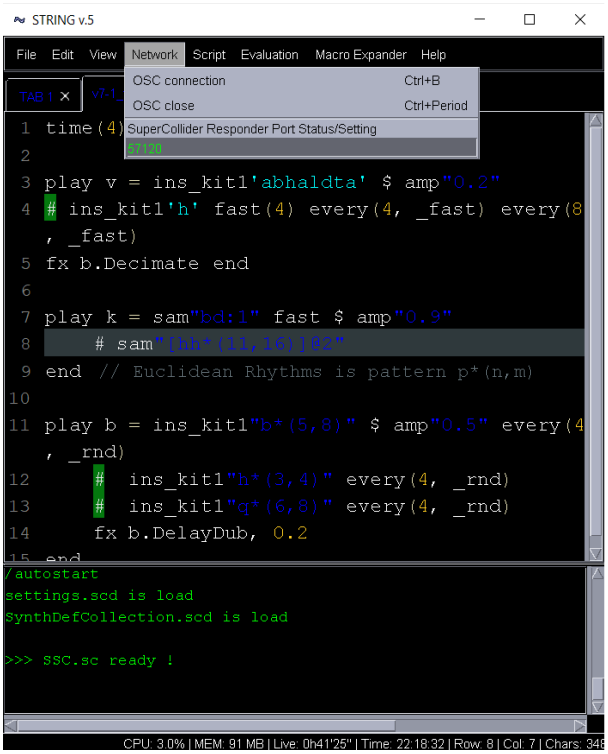


Figura 12. Schermata della GUI dell'ambiente STRING con JavaScript/ECMAScript 2020.

[20] Per le specifiche del protocollo di comunicazione si veda la pagina: https://ccrma.stanford.edu/groups/osc/index.html.

[21] Nello specifico Java SE-11 in Eclipse IDE 2021-03: https://www.eclipse.org/eclipseide/.

[22] Si tratta di un text editor Swing evoluto che offre syntax highlighting, colorazione delle keyword, stringhe di autocompletamento mediante sottoclassi AutoCompletion e DefaultCompletionProvider. Si veda il link: https://github.com/bobbylight/RsyntaxTextArea.

[23] SuperCollider è un linguaggio dinamico e permette la modifica di un programma in esecuzione. La libreria Just In Time estende e unifica tali possibilità fornendo placeholders astratti (proxy) utilizzabili nei calcoli e aggiornabili in fase di esecuzione, semplificando operazioni di: conversational/interactive/on-the fly programming in cui la scrittura del codice è parte integrante della pratica musicale.

[24] Si rimanda al tutorial di H. James Harkins: https://doc.sccode.org/Tutorials/A-Practical-Guide/PG_01_Introduction.html

[25] Progetto disponibile al link: https://github.com/totalgee/bacalao.

[26] La prima versione denominata "Liver" sfruttava la libreria Processing Quarks place/Qscript per la valutazione di espressioni matematiche, equazioni e algoritmi di scheduling. In quella successiva "SCL" (SuperColldier Live Script) è stata impiegata la classe ScriptEngine, un esecutore di codice JavaScript in ambiente Java, demandando i thread dei pattern a funzioni SuperColldier. A partire dal JDK 11 lo ScriptEngine Nashorn viene deprecato e quindi sostituito da "GraalVM/graal.js", un motore di script moderno compatibile (Marzo 2020) con la specifica "ECMAScript 2019" e successive. Si vedano: https://tc39.es/ecma262/ e https://www.graalvm.org/.

[27] Links: https://processing.org/; https://sojamo.de/libraries/oscP5/.

Una volta avviato l'IDE[28] è possibile valutare il comando `SSC.start` in Scide quindi se il collegamento e il caricamento dei file di dependencies (fig.11) hanno successo la post window di SC ritorna il messaggio:

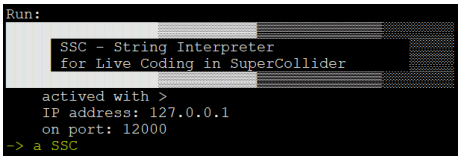


Figura 13. Post window di Scide/SuperCollider-3.11.2 al termine del collegamento con STRING.

Il file "settings.sc" del sistema è modificabile dall'utente e contiene: le opzioni del device audio, le liste dei file sonori, i dizionari dei character patterns, la definizione di funzioni per "fx" e altre specifiche relative al pattern engine. Il file aggiornabile "SynthDefCollection.sc" contiene le definizioni dei processi di: sintesi, modulazione, campionamento, controllo della porta COM, etc. implementati in SynthDef e costrutti di SuperCollider.

3.3. Struttura dello script

Gli algoritmi STRING estendono la gestione dei pattern con la possibilità di definire in poche righe di codice: eventi continui, cluster di eventi, pattern multipli, stream crossfading, etc. Con riferimento alle generalizzazioni dei processi temporali esposte in [25], considerata una generica funzione di elaborazione S composta da i simboli $\{a,b,c\} \in S$ di durata Δ, con k valori di pause τ, un generico pattern P è rappresentato formalmente da:

$$P = \sum_{0}^{k} \sum_{0}^{i} \Delta_k S_i + \tau_k \quad \text{per} \quad \tau \geq 0 \qquad (6)$$

Lo schema generale di implementazione di P è definito, nel caso di base, dalla seguente sintassi dello script:

```
play <var> = <P ref> <arg> $ <P>
<arg> $ ... $ <P> <arg>, interonset P
      # <P> <arg> $..., // parallel P
      # ...
      fx <type> $...,level // process
      ...,          // parallel fx bus
      set <P> <arg> $ ..., // fx args
      ...,          // other arg slots
end
```

dove il processo `play` associato alla variabile `<var>` è la somma di coppie di simboli dei pattern `<P>` e argomenti `<arg>` di trasformazione dei pattern, assegnati ai parametri di un Synth. Essi sono riferiti al pattern iniziale `<P ref>` e delimitati da simboli `$` di concatenamento dei pattern e simboli `#` di stratificazione dei pattern. I tag `fx` e `set` consentono di dichiarare rispettivamente bus di elaborazione audio su `P` e pattern multipli di controllo dei loro parametri [26]. L'esempio di codice in fig.12 segue questa sintassi. Tra i vantaggi di STRING vi sono la possibilità di: personalizzare, comprimere, definire con macro[29] e semplificare il codice di programmazione per un rapido accesso alle istruzioni.

4. IL SISTEMA MODULARE

In aggiunta alla gestione dei pattern elaborati a partire da: SynthDef, functions, samples audio e VST, l'ambiente può rispondere dinamicamente ai comandi per il controllo di hardware esterno via seriale, per una codifica dal vivo associata alle schede "SynthBoard" (fig.15) secondo la seguente struttura modulare:

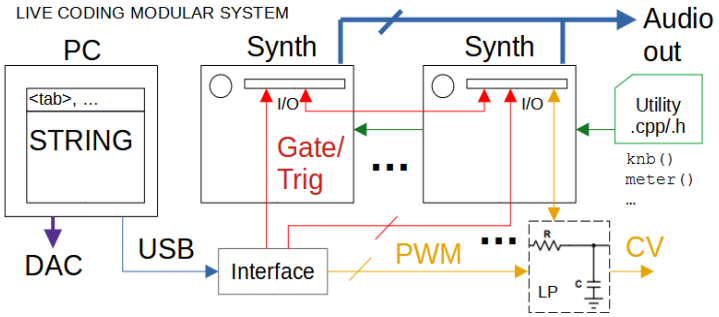


Figura 14. Schema della configurazione modulare.

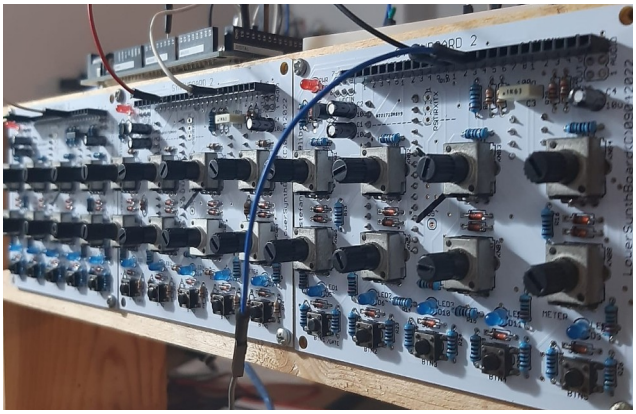


Figura 15. Sistema modulare realizzato con tre schede "SynthBoard v.2" interfacciate all'ambiente STRING tramite la scheda a microcontrollore Arduino MEGA.

L'Object `SerialPort` di SC ha permesso di definire ad alto livello segnali digitali di CV nel range 0-255 e impulsi di trigger HIGH/LOW attraverso l'UART di una interfaccia PC realizzata con Arduino MEGA[30].

5. SVILUPPI

E' previsto un repository on-line[31] contenente: documentazione e dettagli di realizzazione delle componenti software, guida alla riproducibilità del setup hardware,

[28] Test condotti su SO Windows 10 a 64-bit.

[29] Raggruppamenti di codice arbitrario organizzato e invocato nell'ambiente per parole chiave e autocompletamento del codice. A tal proposito il file "macroexpander" in fig.11 sfrutta la serializzazione java per salvare String Object corrispondenti al codice sorgente.

[30] Nel firmware è utile adottare alti valori di baud rate (es. 115200 baud) per non creare ritardi nella sincronizzazione tra pattern audio generati da PC e pattern di controllo dell'hardware esterno. Con una scheda Arduino MEGA come interfaccia sono disponibili: 54 pin di I/O digitali, di cui 15 output PWM, 16 input analogici e una µCU molto performante con sigla ATmega2560.

[31] Si veda la pagina: https://elettronicamusicale.wordpress.com/.

inoltre esempi e strumenti di programmazione validi anche a scopo didattico. Oltre al refactoring del codice dell'ambiente è possibile dotare la scheda di input audio per elaborare segnali esterni ad es. con firmware: Modulation, StateVariable filter, WaveShaper, Delay, etc. A tal scopo occorre escludere il KNOB1 con un microswitch che commuti il segnale direttamente sul pin A0 modificando lo schema in fig.2. Tale campionamento comporta un aggravio ulteriore delle capacità di calcolo, quindi una riduzione della qualità audio e dell'SNR.

6. BIBLIOGRAFIA

[1] M. Novati, AA.VV. *Lo studio di Fonologia, un diario musicale 1954-1983*, Ricordi, 2009

[2] L. Tarabella, B. Bertini, A. Caioli, A. Guerra: *Informatica e musica*, Jackson libri, pp.78–94, 1992.

[3] G. Bertini: *Il terminale musicale polifonico TAU2 il progetto, la realizzazione e il suo ruolo nella computer music*, Musica/Tecnologia 1, Firenze University Press, pp. 97–131, 2007.

[4] C. Roads, M. Mathews: *Interview with Max Mathews*, Computer Music Journal, Vol. 4, No. 4, The MIT Press, pp. 15–22, 1980.

[5] S. Cavaliere: "Il gruppo di informatica musicale nell'istituto di fisica sperimentale dell'università di Napoli" in AA.VV. *Antonio De Santis, pioniere dell'informatica musicale*, Edizioni del conservatorio San Pietro a Majella, p.37, Napoli, 2016.

[6] A. Giordano: *Peppino Di Giugno, genio italiano, macchine digitali per la musica in tempo reale*, Edizioni musicali Esarmonia, pp.44–116, 2020.

[7] A. Di Scipio: *Circuiti del tempo, un percorso storico-critico nella creatività musicale elettroacustica e informatica*, Libreria Musicale Italiana, pp.275–523, 2021.

[8] G. Mori: *Live Coding? What does it mean An Ethnographical Survey on an Innovative Improvisational Approach*, Aracne editore, Roma, 2020.

[9] T. Barrass: "Mozzi: Interactive sound synthesis on the Open Source Arduino Microprocessor" *ICMC 2013 Computer Science, Art*, 2013.

[10] B. Nisarga, Application Report PWM DAC Using MSP430 High-Resolution Timer, MSP430 Applications, *Texas Instruments*, p.4, 2021.

[11] Datasheet Atmel ATmega328P 8-bit AVR Microcontroller with 32K Bytes In-System Programmable Flash. *Atmel Corporation*, p.33, 2015.

[12] B. Edstrom: *Arduino for musicians a complete guide to Arduino and Teensy Microcontrollers,* Oxford University Press, pp.202,235,272, 2016.

[13] S. Silvestri: "Paradigmi di programmazione per il live coding e progettazione di un nuovo linguaggio basato su SuperCollider e Java", atti del *ESCOM-Italy — European Society for the Cognitive Sciences Of Music*, università Campus Bio—Medico di Roma, 2018.

[14] G. Wang, P. R. Cook: *On-the-fly Programming: Using Code as an Expressive Musical Instrument*, Proceedings of the 2004 International Conference on New Interfaces for Musical Expression — NIME, 2004.

[15] A. Blackwell, S. Aaron: *Craft Practices of Live Coding Language Design*, Proceedings of the First International Conference on Live Coding, Leeds, UK, 2015.

[16] S. Aaron, A. Blackwell, R. Hoadley: *A principled approach to developing new languages for live coding*, Proceedings of the International Conference on New Interfaces for Musical Expression, 30 May — 1 June, Oslo Norway, 2011

[17] R. Bell: *An Interface for Realtime Music Using Interpreted Haskell*, Linux Audio Conference, p.115, 2011.

[18] A. Blackwell, N. Collins: *The programming language as a musical instrument*, Proceedings of PPIG05 – Psychology of Programming Interest Group, 2005.

[19] G. Mori: *Pietro Grossi's Live Coding. An Early Case of Computer Music Performance*, International Conference on Live Coding, University of Leeds, 2015.

[20] J. Mooney: *Hugh Davies's Electroacoustic Musical Instruments and their Relation to Present-Day Live Coding Practice: Some Historic Precedents and Similarities*, Proceedings of the First International Conference on Live Coding, ICSRiM, University of Leeds, UK, 2015.

[21] L. Spiegel: *Manipulations of Musical Patterns,* Proceedings of the Symposium on Small Computers and the Arts, Oct., 19-22. 1981.

[22] A. McLean: *Making Programming Languages to Dance to: Live Coding with Tidal*, Proceedings of the 2nd ACM SIGPLAN International Workshop on Functional Art, Music, Modelling and Design, 2014.

[23] T. Magnusson: *Scoring with Code: Composing with algorithmic notation*, in Organised Sound, Vol.19, Special Issue 3: Mediation: Notation and Communication in Electroacoustic Music Performance, Cambridge University Press, pp.268–275, 2014.

[24] T. Magnusson: *The threnoscope: a musical work for live coding performance*, First International Workshop on Live Programming in conjunction with ICSE 2013, San Francisco May 18th–26th, 2013.

[25] S. Aaron, D. Orchard, A. Blackwell: *Temporal semantics for a live coding language*. Published in FARM, pp.37-47, Gothenburg, Sweden, 2014.

[26] A. Du Bois, R. Ribeiro: *Combining Effects in a Music Programming Language based on Patterns*, 17th Brazilian Symposium on Computer Music – SBCM, 2019.

UNA TASSONOMIA MULTIDIMENSIONALE DELLE APPLICAZIONI PER L'EDUCAZIONE MUSICALE

Marcella Mandanici
Conservatorio "L. Marenzio", Brescia
marcella.mandanici@consbs.it

Simone Spagnol
Università Iuav di Venezia
sspagnol@iuav.it

Federico Avanzini, Adriano Baratè, Luca Andrea Ludovico
Università degli Studi di Milano
federico.avanzini@di.unimi.it

ABSTRACT

Negli ultimi settant'anni molti sono stati i cambiamenti che si sono susseguiti nella pratica dell'educazione musicale, coinvolgendo vari aspetti quali i curriculi educativi, le tecnologie disponibili e le strategie di insegnamento. Al fine di delineare un quadro completo di tale evoluzione, questo articolo presenta un database di pubblicazioni scientifiche nel campo dell'educazione musicale che utilizzino uno o più strumenti tecnologici. Basandosi su strumenti per l'analisi e l'organizzazione dei vari contributi, si propone una tassonomia multidimensionale quale strumento utile per la ricerca e la valutazione di applicazioni relative all'educazione musicale. Inoltre, viene presentata una piattaforma web per la consultazione del database e l'analisi delle dimensioni della tassonomia, accessibile all'URL http://techmusicedu.lim.di.unimi.it.

1. INTRODUZIONE

Gli ultimi settant'anni hanno visto importanti cambiamenti tecnologici nella fruizione e nella creazione della musica. La disponibilità di editor musicali e di workstation audio digitali (DAW) dove visualizzare dati audio e MIDI ha offerto la possibilità di produrre musica a categorie di utenti prima completamente escluse da tali attività. Fra questi i bambini e gli studenti delle scuole primarie e secondarie hanno potuto finalmente avvicinarsi alla musica rompendo le barriere legate alla conoscenza della notazione musicale. Inoltre, la diffusione del World Wide Web e delle reti cellulari hanno reso possibile l'ascolto della musica e lo svolgimento di pratiche musicali in qualsiasi momento e luogo, favorendo la condivisione e le pratiche collaborative.

Questi due soli esempi danno un'idea dell'impatto che la tecnologia digitale ha avuto, oltre che sull'industria e sulla cultura musicale, anche sull'educazione musicale. Anche se è probabilmente presto per valutare l'impatto della recente crisi sanitaria conseguente all'epidemia di Covid-19 sulle tecnologie per l'educazione in generale e sull'educazione musicale in particolare, le tecnologie musicali hanno senza ombra di dubbio ottenuto durante gli ultimi due anni un'attenzione particolare da parte di insegnanti, studenti e operatori educativi grazie anche alle possibilità offerte dalla didattica a distanza. Tuttavia, anche se la ricerca sui materiali digitali è andata velocemente evolvendo, essa si è sviluppata in maniera asimmetrica rispetto alle pratiche educative che, per varie ragioni, ancora stentano ad inserire le tecnologie informatiche nei curriculi.

Partendo da queste premesse, il presente articolo si propone di presentare una tassonomia delle applicazioni e dei materiali digitali legati all'educazione musicale. Lo scopo è non solo ottenerne una panoramica aggiornata, ma anche valutare l'oggetto della nostra analisi rispetto a specifici obiettivi formativi (teorie dell'apprendimento), alle tecnologie impiegate e al tipo di attività.

Nella Sezione 2 l'articolo presenta i principi impiegati per la creazione del database di applicazioni per l'educazione musicale che ha portato alla definizione della tassonomia. I domini di cui tale tassonomia si compone vengono analizzati nella Sezione 3 insieme con i relativi esempi applicativi. Infine, la Sezione 4 presenta un'interfaccia web per esplorare il database e visualizzare grafici comparativi sfruttando le dimensioni e i nodi della tassonomia.

2. IL DATABASE

Per la creazione del database è stata condotta una vasta ricerca di pubblicazioni scientifiche riguardanti le tecnologie e i materiali digitali per l'educazione musicale. La scelta di concentrare l'attenzione esclusivamente sulle pubblicazioni scientifiche (articoli su rivista, atti di conferenze, libri, capitoli di libro e, in via eccezionale, tesi di laurea e di dottorato) è stata dettata dalla necessità di ottenere informazioni affidabili sui materiali digitali utilizzati, l'uso che se ne propone e le esperienze educative connesse.

La costruzione del database di pubblicazioni scientifiche è avvenuta seguendo diverse fasi. La prima fase ha previsto la ricerca di articoli utilizzando parole chiave generiche quali "music learning applications", "music education technology", "computer aided music education" attraverso i più popolari motori di ricerca e database di letteratura scientifica, tra cui Google Scholar, CiteseerX, ACM Digital Library, IEEE Xplore Digital Library, JSTOR e Scopus. Una volta raccolta la prima tranche di articoli, è stata effettuata una seconda ricerca basata sulla bibliografia. A

Copyright: ©2022 Marcella Mandanici et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

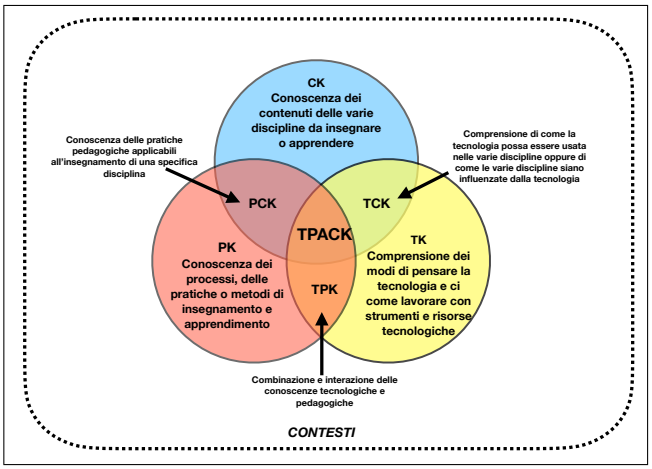


Figura 1. Schematizzazione del paradigma TPACK - Technological Pedagogical and Content Knowledge (figura adattata da `http://www.tpack.org`).

questo punto, è emersa una lista delle principali riviste dedicate ai materiali digitali per l'educazione musicale. Sono stati passati in rassegna gli archivi di tali riviste, tra cui British Journal of Educational Technology, British Journal of Music Education, Computers & Education, International Journal of Educational Research, Journal of Music, Technology and Education e Music Education Research. Inoltre, una volta definiti i nodi della tassonomia (vedi Sezione 3), è stata condotta un'ulteriore ricerca basata sulle parole chiave associate.

Infine, sono stati esclusi tutti i contributi non dedicati esplicitamente all'educazione e gli articoli scritti prima dell'anno 2000, in quest'ultimo caso per evitare di includere applicazioni o sistemi obsoleti. La raccolta, tuttora in fase di sviluppo, ammonta attualmente a un totale di 136 articoli.

3. LA TASSONOMIA MULTIDIMENSIONALE

Secondo Bauer [1], ai primordi dell'introduzione della tecnologia nelle scuole si pensava che gli insegnanti fossero in grado di trovare autonomamente il modo di integrare le potenzialità dei nuovi strumenti nei loro curriculi. Molto presto, però, questo pensiero si rivelò sbagliato. L'utilizzo della tecnologia in un contesto educativo richiede, infatti, non solo la comprensione del suo funzionamento, ma anche l'abilità del docente di ricercare e interpretare le relazioni dinamiche tra tecnologia, contenuti e conoscenza pedagogica. Tali relazioni vengono visualizzate in Figura 1, che si riferisce in particolare al paradigma TPACK (*Technological Pedagogical and Content Knowledge*). Questo fu proposto da Mishra e Koehler [2] nel 2006, partendo dal precedente lavoro di Shulman [3]. Shulman rigetta l'idea che la competenza dell'insegnante debba essere costituita da due aree mutuamente esclusive, quella pedagogica e quella propria della materia insegnata (contenuto), introducendo la nozione di conoscenza pedagogica del contenuto (*Pedagogical Content Knowledge*, PCK). Mishra e Koehler aggiungono una dimensione tecnologica al paradigma di Shulman, in modo che i contenuti delle tre forme primarie di conoscenza (tecnologica – TK, pedagogica – PK, contenuto – CK) formino tre forme secondarie (PCK, TCK e TPK), nonché una conoscenza data dalla loro completa integrazione, denominata appunto TPACK.

LIVELLO 1 DOMINI	**LIVELLO 2** *Dimensioni*	**LIVELLO 3** Nodi
METADATI	*Materiale*	Nome
	Contributo	Valutazione Presentazione Caso di studio Rassegna Pratiche
	Data	20xx
TECNOLOGICO	*Applicazione*	Desktop VR, AR Informatica tangibile Informatica mobile Ambienti intelligenti
	Tecnologie abilitanti (input)	Mouse e tastiera MIDI Dispositivi mobili Tangibili e indossabili Videocamera Microfono Strumenti elettrici
	Risposta del sistema	Audio Grafica Video Aptica Dati grezzi
MUSICALE	*Attività*	Creare Eseguire Rispondere
PEDAGOGICO	*Teoria dell'apprendimento*	Comportamentismo Cognitivismo Costruttivismo
	Utenti	Scuola d'infanzia, primaria Scuola secondaria Conservatorio, università Adulti, musicisti e non Non definiti
	Luogo	Classe Laboratorio Ovunque Web

Tabella 1. I tre livelli della tassonomia con i relativi domini, dimensioni e nodi.

Bauer [1] propone inoltre una tassonomia delle attività musicali, basata sui tre processi artistici denominati *creare*, *eseguire* e *rispondere*. Nonostante ogni attività musicale preveda l'inestricabile compresenza di questi tre processi, tale tassonomia prevede la categorizzazione di tutte le tipologie di attività musicali, inclusi gli aspetti tecnologici e la fruizione di materiali digitali [4]. Esplicitando un punto di vista dettagliato e completo delle tipologie di attività che compongono un curriculum musicale, tale tassonomia può rappresentare il punto di partenza per classificare le attività musicali abilitate dalla tecnologia.

D'altro canto, come evidenziato dal paradigma TPACK, le attività musicali da sole non sono sufficienti per classificare le applicazioni e i materiali digitali. Si propone, quindi, una tassonomia composta da un certo numero di dimensioni appartenenti a diversi domini [5] e organizzata in tre livelli. Il primo livello include i tre domini corrispondenti alle forme primarie di conoscenza del paradigma TPACK (TECNOLOGICO, MUSICALE, PEDAGOGICO) oltre a un dominio METADATI. Il secondo livello comprende dieci dimensioni, mentre il terzo individua tutti i nodi derivanti dall'analisi della letteratura scientifica delineata in Sezione 2. La tassonomia, schematizzata in Tabella 1, è analizzata in dettaglio nelle sezioni 3.1, 3.2, 3.3 e 3.4.

3.1 Il dominio dei metadati

Il dominio dei metadati nella tassonomia proposta include le seguenti tre dimensioni:

- *Materiale* – Specifica il nome dell'applicazione, se indicato esplicitamente nella pubblicazione; in caso diverso, si assegna per convenzione un nome autoesplicativo seguito dall'etichetta "(attr.)".
- *Contributo* – Indica la tipologia di pubblicazione scientifica. I relativi nodi sono: Valutazione (descrizione dell'applicazione più risultati di test), Presentazione (solo descrizione), Caso di studio (descrizione dell'uso dell'applicazione in un determinato contesto), Rassegna (l'applicazione è descritta insieme ad altre simili), Pratiche (esperienze o metodi di insegnamento);
- *Data* – Indica l'anno di pubblicazione. Il database include solo pubblicazioni edite dall'anno 2000 in poi.

3.2 Il dominio tecnologico

Dal punto di vista dell'utente, l'aspetto tecnologico più rilevante riguarda l'interazione con il sistema alla base del funzionamento dell'applicazione; per questo motivo le dimensioni del dominio tecnologico si focalizzano su aspetti di interazione uomo-macchina (HCI). Nella loro tassonomia dei gesti nella HCI, Karam & Schraefel propongono una classificazione in quattro livelli: domini applicativi, tecnologie abilitanti, risposta del sistema e stili gestuali [6]. Tale struttura permette una definizione chiara del tipo di applicazione e delle tecnologie impiegate per veicolare i dati in ingresso e in uscita. Quindi la tassonomia di Karam & Schraefel è stata rivita sulla base delle caratteristiche specifiche delle applicazioni per l'educazione musicale, preservando in particolare i primi tre livelli. Il dominio tecnologico include, pertanto, le seguenti tre dimensioni:

- *Applicazione* – Si riferisce alla tipologia di artefatto digitale impiegato durante l'apprendimento, ovvero l'oggetto dell'azione da parte dell'utente. I nodi di questa dimensione sono:
  - Desktop – Include i software eseguibili su un personal computer, tra cui ad esempio sequenziatori, DAW ed editor audio (ad es. *Audacity*), sistemi di tutoraggio [7], giochi [8], ambienti di programmazione (ad es. *Max*);
  - Realtà virtuale (VR), Realtà aumentata (AR) – Applicazioni che mettono a disposizione un ambiente di realtà virtuale o aumentata. Esempi tipici di AR per l'educazione musicale sono quei sistemi per l'apprendimento di uno strumento musicale che sfruttano la proiezione di informazioni sullo strumento stesso [9, 10];
  - Informatica tangibile – Include i sistemi capaci di ricevere ed elaborare informazioni date dalla manipolazione di oggetti fisici o da sensori tattili o indossabili. Esempi notevoli sono *Reactable* [11] e, più recentemente, *Kibo* [12];
  - Informatica mobile – Copre le applicazioni per tablet e smartphone, le quali, rispetto alle tradizionali lezioni scolastiche, permettono un apprendimento svincolato da requisiti di spazio e tempo. Tra gli esempi citiamo *JamMo*, un software per la creazione di musica pensato per i bambini [13] e le molteplici app musicali per iPad [14];
  - Ambienti intelligenti – Si tratta di ambienti interattivi di larga scala in grado di rispondere a movimenti e/o gesti dell'utente, nei quali quindi l'interazione si basa sul movimento corporeo, la propriocezione e la cinestesia [15]. *Sound Maker* [16] per la creazione di musica e *Harmonic Walk* [17] per l'armonizzazione delle melodie rappresentano due esempi significativi.
- *Tecnologie abilitanti (input)* – Raggruppa le tecnologie utilizzate per immettere dati nel sistema, ad esempio audio o movimenti dell'esecutore. I nodi di questa dimensione sono:
  - Mouse e tastiera – I tradizionali dispositivi di interazione con il computer vengono utilizzati anche nel campo musicale, ad esempio per la manipolazione di elementi grafici nelle DAW e nei sequenziatori, oppure per il controllo di note e durate negli editor di partiture [18];
  - MIDI – Acronimo di Musical Instrument Digital Interface, è il protocollo standard per l'interazione tra e con gli strumenti musicali elettronici. In un contesto educativo, i dati prodotti dai dispositivi MIDI possono essere usati come input per i software di notazione musicale [19], i sequenziatori [20] e i sistemi di tutoraggio per le performance assistite [21];
  - Dispositivi mobili – Tablet e smartphone possono essere usati come elementi di controllo distribuiti di eventi audio, come nel progetto *Soundcool* [22]; come superfici per performance percussive (ad es. l'applicazione *Rhythm Workers* [23]); o come videocamera e/o sensori inerziali per fornire input a varie applicazioni [24];
  - Tangibili e indossabili – L'interazione musicale con oggetti fisici sensorizzati può avvenire attraverso la loro manipolazione creativa [25]. I dispositivi indossabili fungono invece da rilevatori di movimento, come in *MusicJacket* [26], o da marker per un sistema di visione computerizzata [27];
  - Videocamera – L'utilizzo più comune di una videocamera è nell'insegnamento a distanza di tecniche strumentali [28]. Inoltre, registrazioni di performance musicali possono essere utilizzate per insegnare ad esempio le arcate nelle esecuzioni violinistiche [29] o la diteggiatura [30];
  - Microfono – La facoltà di immettere segnali audio digitali in un computer è fondamentale,

non solo per l'utilizzo di DAW e sequenziatori, ma anche ad esempio per giochi e coach vocali [31, 32];

- Strumenti elettrici – Questi strumenti possono fornire direttamente segnali audio a un computer, ad esempio per applicazioni per la valutazione automatica dell'espressione in un'esecuzione musicale [33].

- *Risposta del sistema* – La terza dimensione si riferisce alla tipologia di dati prodotti dall'applicazione musicale. I nodi di questa dimensione sono:
  - Audio – Ovvero il risultato di una registrazione, di una combinazione di dati MIDI ed elaborazione digitale, o di un programma (scritto ad esempio in *Csound* o *Max*) [34];
  - Grafica – Il feedback visivo è molto comune nei sistemi di tutoraggio per l'apprendimento di uno strumento (ad esempio, il pianoforte [35]), per l'allenamento della voce [36] o dell'orecchio (ad es. *IMUTUS* [7]);
  - Video – Analogamente, un feedback video in tempo reale può essere usato per la valutazione di una performance vocale o strumentale [37];
  - Aptica – Output tattile fornito da specifici attuatori: ad esempio, in *PianoTouch* [38] viene utilizzato un guanto aptico per guidare il movimento delle dita sulla tastiera di un pianoforte;
  - Dati grezzi – Ovvero dati che, prevedendo un'ulteriore elaborazione da parte del computer, non vengono comunicati all'utente.

3.3 Il dominio musicale

Il dominio musicale contiene una sola dimensione:

- *Attività* – Ovvero i tre processi artistici definiti in [1]:
  - Creare – Raggruppa tutte le applicazioni che riguardano attività musicali che spaziano dalla composizione tradizionale [39] all'improvvisazione [40] e alla programmazione [41];
  - Eseguire – Ovvero le applicazioni per l'apprendimento di tecniche strumentali, vocali e di direzione [42] oltre che di aspetti teorici quali il ritmo [43], la melodia [44] e l'armonia [45];
  - Rispondere – All'interno di questo processo artistico, Bauer [1] elenca attività quali *"Ascoltare e descrivere"*, *"Analizzare"* e *"Valutare"*. Figurano quindi in questo nodo applicazioni per l'educazione dell'orecchio [46], la percezione della musica [27] o, più in generale, per l'educazione musicale [47].

3.4 Il dominio pedagogico

Il dominio pedagogico si riferisce agli aspetti di usabilità didattica che caratterizzano le applicazioni e include le seguenti tre dimensioni:

- *Teoria dell'apprendimento* – La prima dimensione classifica le applicazioni in base alle teoria dell'apprendimento che ne ha ispirato la progettazione. Le tre teorie dell'apprendimento sono:
  - Comportamentismo – Secondo questa teoria l'apprendimento è un processo di risposta a stimoli esterni. Questo nodo comprende quindi applicazioni in cui i contenuti sono organizzati in unità sequenziali che prevedono una risposta da parte dell'utente e una valutazione da parte del sistema (istruzione programmata [48]). Esempi di applicazioni sono quelle per lo sviluppo di abilità quali la lettura a prima vista [8] e l'educazione dell'orecchio [49];
  - Cognitivismo – Questa teoria considera i processi interni che consentono all'individuo di selezionare, acquisire, organizzare e richiamare alla memoria le informazioni. Alcuni di questi meccanismi sono l'apprendimento enattivo e l'apprendimento attraverso modelli o esempi eleborati [50]. Fanno quindi parte di questo nodo le applicazioni che impiegano rappresentazioni visive o multimediali e/o funzionalità di tutoraggio. Un esempio è *Music Paint Machine* [51], sistema che produce una rappresentazione visiva dei movimenti di un esecutore che viene impiegata per sviluppare meglio le abilità esecutive dei bambini;
  - Costruttivismo – Diversamente dalle precedenti teorie, il costruttivismo non considera l'apprendimento come l'acquisizione di un sapere definito a priori, ma bensì come qualcosa che il discente stesso deve costruire attraverso un suo personale sviluppo. Ambienti ispirati al costruttivismo devono offrire ricchi stimoli che consentano l'esplorazione e la manipolazione degli elementi alla ricerca di soluzioni originali. Un esempio è *Continuator* di Pachet [52] dove l'intelligenza artificiale è usata per rispondere ad una sequenza musicale eseguita su una tastiera MIDI. L'attivazione di meccanismi di risposta a ciò che il sistema è in grado di elaborare costituisce un ottimo esempio di come un agente intelligente possa stimolare la creatività degli utenti.

- *Utenti* – La seconda dimensione identifica quali sono gli utenti finali per cui l'applicazione è stata progettata. I suoi nodi sono:
  - Scuola d'infanzia, primaria – Applicazioni pensate per un uso semplice e intuitivo, che ben si adatti ad un approccio giocoso e accattivante. Ad esempio nell'ambiente *DrumStep* elementi grafici e animazioni permettono di superare la notazione musicale standard offrendo un approccio ludico allo studio del ritmo [53];

  - Scuola secondaria – Applicazioni con un grado superiore di complessità, finalizzate soprattutto alla composizione musicale [54] o all'esecuzione strumentale [55];
  - Conservatorio, università – Applicazioni per la creazione e l'esecuzione musicale a livello superiore o professionale [56, 57];
  - Adulti, musicisti e non – Applicazioni testate da adulti, musicisti o non, o più genericamente intese per utenti adulti.
  - Non definiti – Nodo di default per applicazioni non dirette a un pubblico specifico.

- *Luogo* – L'ultima dimensione tiene conto dei possibili spazi in cui l'attività educativa può avere luogo. I suoi nodi sono:
  - Classe – Aula scolastica adattata alla specifica attività musicale, con l'utilizzo di tecnologie quali personal computer, strumenti musicali elettronici e/o cuffie, oltre che microfoni, altoparlanti, proiettori e servizi web;
  - Laboratorio – Tutte le applicazioni che utilizzino tecnologie meno popolari quali sensori, attuatori, oggetti fisici sensorizzati, sistemi per il tracciamento del movimento richiedono un ambiente più affine a un laboratorio;
  - Ovunque – Le applicazioni che necessitano solamente di un dispositivo mobile ricadono in questa categoria;
  - Web – Il web è un luogo virtuale per applicazioni distribuite e collaborative, principalmente per la creazione di musica, ma anche per la performance (ad es. *JamMo* [13]).

4. UNA PIATTAFORMA WEB PER ESPLORARE LA TASSONOMIA

Nelle intenzioni originarie, la tassonomia non dovrebbe solamente rappresentare uno strumento teorico per classificare materiali, ma soprattutto uno strumento pratico per aiutare gli educatori a identificare gli approcci più appropriati a seconda delle proprie necessità. A tal proposito, il database è stato pubblicato su Zotero, [1] dove continuerà a venire aggiornato. Tutte le pubblicazioni presenti nel database sono state etichettate seguendo la tassonomia proposta, associando dunque a ognuna di esse un sottoinsieme di nodi sia in maniera diretta (quando esplicitato nella pubblicazione stessa) sia per inferenza. Inoltre, è stata sviluppata e rilasciata pubblicamente un'interfaccia web per la consultazione interattiva del database.

L'interfaccia, liberamente accessibile all'URL `http://techmusicedu.lim.di.unimi.it`, supporta l'esplorazione gerarchica della tassonomia. La navigazione prende avvio dal livello del dominio, attraversa il livello della dimensione per arrivare a quello dei nodi, e infine visualizza le pubblicazioni etichettate come previsto del nodo selezionato. Inoltre, è stata implementata un'opzione di ricerca grazie alla quale l'utente può: (i) selezionare singoli nodi (ad es., Cognitivismo); (ii) selezionare più nodi della stessa dimensione (ad es., Classe e Web); (iii) selezionare più nodi di dimensioni diverse (ad es., MIDI, Eseguire e Scuola d'infanzia, primaria). La ricerca restituisce gli articoli etichettati contemporaneamente secondo tutti i nodi selezionati.

Infine, l'interfaccia web mette a disposizione dell'utente strumenti di visualizzazione per analizzare quantitativamente il numero di pubblicazioni nel database a seconda delle etichette. Le tre visualizzazioni supportate sono:

- Grafico delle distribuzioni – Grafico a torta che riporta la distribuzione degli articoli in base al dominio e alla dimensione selezionati;
- Grafico comparativo per assi – I dati vengono visualizzati sotto forma di cerchi su una griglia bidimensionale, dove gli assi sono selezionabili tra i domini e le dimensioni disponibili. La dimensione di ogni cerchio è proporzionale al numero di articoli etichettati con i corrispondenti nodi di entrambi gli assi;
- Grafico comparativo per anno – Simile al precedente, ma l'asse x è riservato alla dimensione *Data* al fine di monitorare l'evoluzione temporale di una data dimensione.

4.1 Analisi delle dimensioni

A dimostrazione dell'efficacia della tassonomia e delle potenzialità dell'interfaccia web, si riportano a titolo di esempio alcuni approfondimenti. Per motivi di spazio non si mostrano i grafici corrispondenti alle ricerche, rimandando il lettore all'interfaccia web per la loro consultazione.

Focalizzandosi sul dominio TECNOLOGICO e in particolare sulla dimensione *Applicazione*, si nota che il nodo Desktop è il più rappresentato, con il 65% delle pubblicazioni aventi questa etichetta. Ciò non deve sorprendere, poiché si tratta del tipo di applicazione più tradizionale; è inoltre l'unico nodo rappresentato tra il 2001 e il 2005. A partire dal 2014, la rilevanza di questo nodo inizia percentualmente a decrescere, principalmente a favore di applicazioni AR/VR.

Per quanto riguarda il dominio MUSICALE, la dimensione *Attività* indica come la maggioranza delle applicazioni implementi attività appartenenti ai processi Creare o Eseguire in proporzioni simili, mentre il processo Rispondere è più raro, con solo il 14% delle pubblicazioni. Tuttavia, si può anche notare che prima del 2008 non è stata pubblicata alcuna applicazione relativa a tale processo, mentre gli anni seguenti mostrano un andamento positivo, e quindi una maggiore attenzione al processo Rispondere.

Per quanto riguarda invece il dominio PEDAGOGICO, nonostante la dimensione *Luogo* dimostri una chiara prevalenza di applicazioni pensate per la classe o il laboratorio, il nodo Ovunque inizia a presentare andamento positivo a partire dal 2010 circa, grazie alla crescente disponibilità di dispositivi portatili a basso costo. D'altro canto, si evidenzia un trend sorprendentemente negativo per il nodo Web a partire dagli stessi anni.

[1] `https://www.zotero.org/`

Informazioni utili possono anche essere derivate dalle relazioni tra dimensioni. A tale proposito, si evidenzia un interessante risultato dato dall'intersezione delle dimensioni *Attività* e *Teoria dell'apprendimento*. Le connessioni più frequenti sono tra i nodi Eseguire e Cognitivismo (35%) e tra i nodi Creare e Costruttivismo (45%). Non sorprende invece l'intersezione nulla tra i nodi Comportamentismo e Creare, data la difficoltà – se non l'impossibilità – di ricondurre un'attività creativa a una serie di domande chiuse.

5. CONCLUSIONI

Questo articolo ha presentato un approccio originale all'interpretazione e alla valutazione di materiali digitali legati all'educazione musicale. Partendo dal paradigma TPACK come strumento per l'organizzazione della conoscenza, è stata sviluppata una tassonomia multidimensionale delle applicazioni tecnologiche per la didattica della musica. Uno dei vantaggi più importanti di questo tipo di rappresentazione sta nella possibilità di analizzare gli stessi dati da diverse angolazioni. Inoltre, si tratta di uno strumento flessibile che si presta facilmente a modifiche ed estensioni. Il processo di etichettatura del database di pubblicazioni attraverso i nodi della tassonomia ha rappresentato una validazione dal basso (*bottom-up*) dell'approccio teorico proposto; infatti, tale attività ha permesso di evidenziare limiti e difetti della tassonomia, e quindi di raffinare iterativamente il processo di progettazione della stessa che in futuro dovrà evolvere attraverso fasi di valutazione qualitativa e di usabilità da parte dei potenziali utenti.

Gli obiettivi del lavoro sono ambiziosi. In primo luogo, si vuole dare la possibilità agli insegnanti di orientarsi nel campo delle tecnologie per l'educazione musicale con uno strumento informativo di facile utilizzo basato su solide basi scientifiche. In secondo luogo, si punta a condividere con la comunità scientifica uno strumento utile per l'analisi dello stato dell'arte nel campo, contribuendo agli avanzamenti nei settori scientifici coinvolti grazie all'approccio multidisciplinare offerto dalla tassonomia. Infine, si vuole guidare le future azioni degli innovatori nel campo dell'educazione musicale, con la speranza che ambiscano alla sperimentazione di nuovi curriculi rivolti alla piena integrazione tecnologica nei processi educativi.

6. BIBLIOGRAFIA

[1] W. I. Bauer, *Music Learning Today: Digital Pedagogy for Creating, Performing, and Responding to Music*. New York, NY, USA: Oxford Univ. Press, 2014. doi: 10.1093/acprof:oso/9780199890590.001.0001.

[2] M. Koehler and P. Mishra, "What is technological pedagogical content knowledge (TPACK)?," *Contemp. Issues Technol. Teacher Educ.*, vol. 9, no. 1, pp. 60–70, 2009. doi: 10.1177/002205741319300303.

[3] L. S. Shulman, "Those who understand: Knowledge growth in teaching," *Educ. Res.*, vol. 15, pp. 4–14, Feb. 1986. doi: 10.3102/0013189X015002004.

[4] W. Bauer, J. Harris, and M. Hofer, "Music learning activity types." `http://activitytypes.wm.edu/MusicLearningATs-June2012.pdf`, Jun 2012. Accessed May 10, 2021.

[5] K. S. Law, C.-S. Wong, and W. M. Mobley, "Toward a taxonomy of multidimensional constructs," *Acad. Manage. Rev.*, vol. 23, pp. 741–755, Oct. 1998. doi: 10.2307/259060.

[6] M. Karam and M. C. Schraefel, "A taxonomy of gestures in human computer interactions," project report, University of Southampton, 2005.

[7] S. Raptis, A. Chalamandaris, A. Baxevanis, A. Askenfelt, E. Schoonderwaldt, K. F. Hansen, D. Fober, S. Letz, and Y. Orlarey, "IMUTUS: An interactive system for learning to play a musical instrument," in *Proc. Int. Conf. Interactive Computer Aided Learning*, (Villach, Austria), Sept. 29–30 2004.

[8] K. H. Smith, "The effect of computer-assisted instruction and field independence on the development of rhythm sight-reading skills of middle school instrumental students," *Int. J. Music Educ.*, vol. 27, pp. 59–68, Feb. 2009. doi: 10.1177/0255761408099064.

[9] M. Löchtefeld, S. Gehring, R. Jung, and A. Krüger, "guitAR: Supporting guitar learning through mobile projection," in *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, (Vancouver, Canada), pp. 1447–1452, May 7–12 2011.

[10] K. Rogers, A. Röhlig, M. Weing, J. Gugenheimer, B. Könings, M. Klepsch, F. Schaub, E. Rukzio, T. Seufert, and M. Weber, "P.I.A.N.O.: Faster piano learning with interactive projection," in *Proc. 9th ACM Int. Conf. Interactive Tabletops and Surfaces*, (Dresden, Germany), pp. 149–158, Nov. 16–19 2014.

[11] A. Xambó, E. Hornecker, P. Marshall, S. Jordà, C. Dobbyn, and R. Laney, "Let's jam the Reactable: Peer learning during musical improvisation with a tabletop tangible interface," *ACM Trans. Comput.-Human Interact.*, vol. 20, Dec. 2013. doi: 10.1145/2530541.

[12] M. D. Amico and L. A. Ludovico, "Kibo: A MIDI controller with a tangible user interface for music education," in *Proc. 12th Int. Conf. Computer Supported Education*, (Setúbal, Portugal), pp. 613–619, May 2–4 2020. doi: 10.5220/0009805206130619.

[13] P. Paananen and M. Myllykoski, "JamMo: Developmentally designed software for children's mobile music-making," in *Proc. 7th Triennial Conf. European Society for the Cognitive Sciences of Music*, (Jyväskylä, Finland), pp. 391–400, Aug. 12–16 2009.

[14] H. Ruismäki, A. Juvonen, and K. Lehtonen, "The iPad and music in the new learning environment," *Eur. J. Soc. Behav. Sci.*, vol. 6, no. 3, pp. 1084–1096, 2013. doi: 10.15405/ejsbs.85.

[15] R. J. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum, "Reality-based interaction: A framework for post-WIMP interfaces," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, (Florence, Italy), pp. 201–210, Apr. 6–9 2008.

[16] A. N. Antle, M. Droumeva, and G. Corness, "Playing with the sound maker: Do embodied metaphors help children learn?," in *Proc. 7th Int. Conf. Interaction Design and Children*, (Chicago, IL, USA), pp. 178–185, June 10–13 2008. doi: 10.1145/1463689.1463754.

[17] M. Mandanici, A. Rodà, and S. Canazza, "The harmonic walk: An interactive physical environment to learn tonal melody accompaniment," *Adv. Multim.*, vol. 2016, 2016. doi: 10.1155/2016/4027164.

[18] M. Gall and N. Breeze, "Music and ejay: An opportunity for creative collaborations in the classroom," *International Journal of Educational Research*, vol. 47, no. 1, pp. 27–40, 2008.

[19] S. T. Schroth, J. A. Helfer, and R. Dammers, "Using technology to assist gifted children's musical development," *Gifted Child Today*, vol. 32, pp. 54–61, Apr. 2009. doi: 10.4219/gct-2009-877.

[20] J. McDowall, "Music technology: New literacies in the early years," in *Proc. IFIP Working Group 3.5 Conf.: Young children and learning technologies*, (Parramatta, Australia), pp. 83–88, July 2003.

[21] I. Barakonyi and D. Schmalstieg, "Augmented reality agents in the development pipeline of computer entertainment," in *Proc. 4th Int. Conf. Entertainment Computing*, (Sanda, Japan), pp. 345–356, Sept. 19–21 2005. doi: 10.1007/11558651_34.

[22] N. Berbel-Gómez, A. Murillo-Ribes, J. Sastre-Martínez, and M. E. Riaño Galán, "Sound creation and artistic language hybridization through the use of the collaborative creation system: Soundcool," *Turk. Online J. Educ. Technol.*, pp. 997–1009, Nov. 2017.

[23] V. Bégel, A. Seilles, and S. Dalla Bella, "Rhythm Workers: A music-based serious game for training rhythm skills," *Music Sci.*, vol. 1, Oct. 2018. doi: 10.1177/2059204318794369.

[24] M. Rusiñol, J. Chazalon, and K. Diaz-Chito, "Augmented songbook: An augmented reality educational application for raising music awareness," *Multim. Tools Appl.*, vol. 77, pp. 13773–13798, June 2018. doi: 10.1007/s11042-017-4991-4.

[25] S. Bakker, E. Van Den Hoven, and A. N. Antle, "MoSo tangibles: Evaluating embodied learning," in *Proc. 5th Int. Conf. Tangible, Embedded, and Embodied Interaction*, (Funchal, Portugal), pp. 85–92, Jan. 23–26 2010. doi: 10.1145/1935701.1935720.

[26] J. Van Der Linden, E. Schoonderwaldt, J. Bird, and R. Johnson, "MusicJacket — Combining motion capture and vibrotactile feedback to teach violin bowing," *IEEE Trans. Instrum. Meas.*, vol. 60, pp. 104–113, Jan. 2011. doi: 10.1109/TIM.2010.2065770.

[27] V. Farinazzo Martins, L. Gomez, and A. G. Dionísio Corrêa, "Teaching children musical perception with MUSIC-AR," *EAI Endorsed Trans. e-Learn.*, vol. 2, Mar. 2015. doi: 10.4108/el.2.5.e3.

[28] J. Palazón Herrera and A. Giráldez Hayes, "Vodcasting and instrumental practice in secondary school music classes," *J. Music Technol. Educ.*, vol. 7, pp. 163–180, Oct. 2014. doi: 10.1386/jmte.7.2.163_1.

[29] J. Van Der Linden, E. Schoonderwaldt, and J. Bird, "Towards a real-time system for teaching novices correct violin bowing technique," in *2009 IEEE Int. Work. Haptic Audio Visual Environments and Games*, (Lecco, Italy), pp. 81–86, Nov. 7–8 2009. doi: 10.1109/HAVE.2009.5356123.

[30] A. Goodwin and R. Green, "Key detection for a virtual piano teacher," in *Proc. 28th Int. Conf. Image and Vision Computing New Zealand*, (Wellington, New Zealand), pp. 282–287, Nov. 27–29 2013. doi: 10.1109/IVCNZ.2013.6727030.

[31] P. Hämäläinen, T. Mäki-Patola, V. Pulkki, and M. Airas, "Musical computer games played by singing," in *Proc. 7th Int. Conf. Digital Audio Effects*, (Naples, Italy), pp. 367–371, Oct. 5–8 2004.

[32] M. Pérez-Gil, J. Tejada, R. Morant, and A. Pérez-González De Martos, "Cantus: Construction and evaluation of a software solution for real-time vocal music training and musical intonation assessment," *J. Music Technol. Educ.*, vol. 9, pp. 125–144, July 2016. doi: 10.1386/jmte.9.2.125_1.

[33] J. Karlsson, S. Liljeström, and P. N. Juslin, "Teaching musical expression: Effects of production and delivery of feedback by teacher vs. computer on rated feedback quality," *Music Educ. Res.*, vol. 11, pp. 175–191, May 2009. doi: 10.1080/14613800902924532.

[34] V. J. Manzo, *Max/MSP/Jitter for Music: A Practical Guide to Developing Interactive Music Systems for Education and More*. New York, NY, USA: Oxford Univ. Press, 2011.

[35] D. Hackl and C. Anthes, "HoloKeys - An augmented reality application for learning the piano," in *Proc. 10th Forum Media Technology and 3rd All Around Audio Symp.*, (St. Pölten, Austria), pp. 140–144, Nov. 29–30 2017.

[36] G. F. Welch, E. Himonides, D. M. Howard, and J. Brereton, "VOXed: Technology as a meaningful teaching aid in the singing studio," in *Proc. Conf. Interdisciplinary Musicology*, (Graz, Austria), Apr. 15–18 2004.

[37] J. Yin, Y. Wang, and D. Hsu, "Digital violin tutor: An integrated system for beginning violin learners," in *Proc. 13th Annu. ACM Int. Conf. Multimedia*, (Hilton, Singapore), pp. 976–985, Nov. 6–11 2005. doi: 10.1145/1101149.1101353.

[38] K. Huang, E. Y. Do, and T. Starner, "Pianotouch: A wearable haptic piano instruction system for passive learning of piano skills," in *2008 12th IEEE International Symposium on Wearable Computers*, pp. 41–44, 2008.

[39] J. Savage and M. Challis, "Dunwich Revisited: Collaborative composition and performance with new technologies," *Brit. J. Music Educ.*, vol. 18, pp. 139–149, July 2001. doi: 10.1017/S0265051701000237.

[40] A. Brown and S. Dillon, "Networked improvisational musical environments: Learning through on-line collaborative music making," in *Music Education with Digital Technology* (J. Finney and P. Burnard, eds.), ch. 8, pp. 95–106, London, U.K.: A. & C. Black, 2007.

[41] L. A. Ludovico, D. Malchiodi, and L. Zecca, "A multimodal LEGO®-based learning activity mixing musical notation and computer programming," in *Proc. 1st ACM SIGCHI Int. Work. Multimodal Interaction for Education*, (Glasgow, U.K.), pp. 44–48, Nov. 13 2017. doi: 10.1145/3139513.3139519.

[42] D. Hollinger and J. Sullivan, "The effects of technology-based conducting practice on skill achievement in novice conductors," *Res. Issues Music Educ.*, vol. 5, pp. 1–6, Sept. 2007.

[43] C.-H. Chou and Y.-L. Chu, "Interactive rhythm learning system by combining tablet computers and robots," *Appl. Sci.*, vol. 7, Mar. 2017. doi: 10.3390/app7030258.

[44] C. Dittmar, E. Cano, J. Abeßer, and S. Grollmisch, "Music information retrieval meets music education," in *Multimodal Music Processing* (M. Müller, M. Goto, and M. Schedl, eds.), vol. 3, pp. 95–120, Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2012. doi: 10.4230/DFU.Vol3.11041.95.

[45] M. Mandanici, A. Baratè, L. A. Ludovico, and F. Avanzini, "A computer-based approach to teach tonal harmony to young students," in *Proc. 11th Int. Conf. Computer Supported Education*, (Heraklion, Greece), pp. 271–279, May 2–4 2019. doi: 10.5220/0007700402710279.

[46] A. Portowitz, K. A. Peppler, and M. Downton, "In Harmony: A technology-based music education model to enhance musical understanding and general learning skills," *Int. J. Music Educ.*, vol. 32, pp. 242–260, May 2014. doi: 10.1177/0255761413517056.

[47] F. Frosini, N. Mitolo, P. Nesi, and M. Paolucci, "Collaborative solution for music education," in *2008 Int. Conf. Automated Solutions for Cross Media Content and Multi-Channel Distribution*, (Florence, Italy), pp. 71–78, Nov. 17–19 2008. doi: 10.1109/AXMEDIS.2008.22.

[48] B. F. Skinner, "Teaching machines," *Science*, vol. 128, pp. 969–977, Oct. 1958. doi: 10.1126/science.128.3330.969.

[49] C. S. Loh, "Mona Listen: A web-based ear training module for musical pitch discrimination of melodic intervals," in *Proc. E-Learn 2004*, (Washington, DC, USA), Nov. 1–5 2004.

[50] S. Kalyuga, P. Chandler, and J. Sweller, "Incorporating learner experience into the design of multimedia instruction," *J. Educ. Psychol.*, vol. 92, pp. 126–136, Mar. 2000. doi: 10.1037/0022-0663.92.1.126.

[51] L. Nijs, B. Moens, M. Lesaffre, and M. Leman, "The music paint machine: Stimulating self-monitoring through the generation of creative visual output using a technology-enhanced learning tool," *J. New Music Res.*, vol. 41, pp. 79–101, Apr. 2012. doi: 10.1080/09298215.2011.650180.

[52] A. R. Addessi and F. Pachet, "Experiments with a musical machine: Musical style replication in 3 to 5 year old children," *Brit. J. Music Educ.*, vol. 22, pp. 21–46, Mar. 2005. doi: 10.1017/S0265051704005972.

[53] C. McCarthy, J. Bligh, K. Jennings, and B. Tangney, "Virtual collaborative learning environments for music: Networked drumsteps," *Comput. Educ.*, vol. 44, pp. 173–195, Feb. 2005. doi: 10.1016/j.compedu.2004.08.004.

[54] L. Mellor, "Creativity, originality, identity: Investigating computer-based composition in the secondary school," *Music Educ. Res.*, vol. 10, pp. 451–472, Dec. 2008. doi: 10.1080/14613800802547680.

[55] M. T. Hopkins, "Pilot-testing of new software for measuring string players' instrument tuning skills," *J. Music Technol. Educ.*, vol. 7, pp. 5–21, May 2014. doi: 10.1386/jmte.7.1.5_1.

[56] S. Airy and J. M. Parr, "MIDI, music and me: Students' perspectives on composing with MIDI," *Music Educ. Res.*, vol. 3, no. 1, pp. 41–49, 2001. doi: 10.1080/14613800020029941.

[57] T. K.-L. Ho, H.-S. Lin, C.-K. Chen, and J.-L. Tsai, "Development of a computer-based visualised quantitative learning system for playing violin vibrato," *Brit. J. Educ. Technol.*, vol. 46, pp. 71–81, Jan. 2015. doi: 10.1111/bjet.12124.

Technology Integration for the 21st Century Musician

Marcella Mandanici
Dept. of Music Education
Music Conservatory "L. Marenzio"
Brescia, Italy
marcella.mandanici@consbs.it

Stefano Delle Monache
Faculty of Industrial Design Engineering
Delft University of Technology
Delft, Netherlands
s.dellemonache@tudelft.nl

ABSTRACT

This paper examines the theme of technology integration in the context of professional musicians training. Whilst some important guidelines are derived from the analysis of frameworks and theories for technology integration at K-12 level, the "Technologies for Music Education" Master in Music Degree program is presented and discussed as a proposal for the design of novel educational paths. Following the TPACK theory, the didactic plan combines content, technological and pedagogical knowledge with the aim of responding to the needs of the music educator of the 21st century.

1. INTRODUCTION

In his book "The Digital Musician", Andrew Hugill attempts to outline the characteristics and skills of this new kind of musician, starting from the analysis of various cultural backgrounds and genesis of musical training [1]. Hugill identifies 3 types of musicians (classical, pop–rock and electronic) and observes that even the electronic musicians may not fit the characteristics of the digital musician, as they mainly engage with the traditional electroacoustic chain of tools still, from the microphone to the loudspeaker, and are possibly unaware of the innovations introduced by the digital age. On the other hand, every musician is experienced in some kind of technology (i.e., his own instrument) be it an ancient, acoustic or electric instrument. The key point is that nowadays the digital technology affects every kind of music activity and represents much more than a simple expansion of musicianship. As pinpointed by A. Brown [2, pp. 3–15] computers can be certainly used as plain tools to support traditional music tasks (i.e. writing scores or recording music) or, in a more meaningful way, as rich design materials to shape and transform the inherent nature of the sonic matter at hand [3].

Especially, the development of a digital musicianship is not only concerned with the awareness and knowledge of the possibilities offered by the technologies of sound (e.g., sound recording and production, archiving and sharing platforms). The digital musician may well be an acoustic instrument performer or composer who deploys digital means to inform her activity creatively, e.g., to master instrumental techniques [4], to improve the students' abilities, to manage music performances and deliver them to large audiences [5]. Yet, more importantly, the digital musician is concerned with the societal aspects of the *how* and the *why* making music [6, 7]. This involves a critical attitude towards the music conventions and the clear separation of roles: The digital musicianship rather describes a fluid asset of music abilities which can rapidly shift and evolve (i.e. from performer to creator; from composer to sound artist or designer; from computer musician to sound engineer; from music teacher to educational applications designer, and so forth).

In the contemporary society, the digital literacy is considered as much important as being able to write and count [8]. In this respect, the institutional stakeholders are putting a great effort to endowing the 21st century citizen of knowledge and tools to cope with the increasing challenges of technology (e.g., security, privacy, awareness) [9]. Therefore, being a *digital musician* can no longer be considered an option, yet it is becoming a more and more relevant aspect of the musicians' professional background.

This article aims at providing food for thought regarding the theoretical tools, the educational guidelines and strategies to facilitate a process of technological integration in the academic curriculum of musicians. Section 2 provides the theoretical background and frameworks for the technological integration, and sets the point of departure for Section 3, wherein a brief analysis of the curriculum of professional musicians in Italian Conservatories is reported. An example of music curriculum and didactic plan inspired by technology integration is presented in Section 4. Final considerations are drawn in Section 5.

2. TECHNOLOGY INTEGRATION

According to J. Dockstader technology integration is *"...using computers effectively and efficiently in the general content areas to allow students to learn how to apply computer skills in meaningful ways."* [10, p. 2], that is essentially having the curriculum driving the technology and not the way around. In this respect, the author stresses 7 main reasons to integrate technology in the students' curriculum:

1. *Correctly designed, more depth into the content-area curriculum is possible*. The technology acts as a

Copyright: ©2020 Marcella Mandanici et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

magnifying glass on the knowledge within a specific domain. This is certainly a desirable effect, yet it may have possible drawbacks on school administration and the role of teaching, as pinpointed by B.F. Skinner, the father of the programmed instruction approach to learning through computers [11]: In particular, the main risk is to downgrade the role of the teacher from a trustworthy figure of knowledge to a simple tutor or organizer of activities.

2. *In the information age there is an intrinsic need to learn technology.* Although this is a 20 year old statement, employing technology in the learning process has an autotelic value, that is it safeguards that nobody can graduate without having acquired a basic technological literacy.

3. *Students are motivated by technology, thus increasing academic engagement time*. While studies confirm that students' engagement increases with the use of technology [12], the effect on learning outcomes and performance is much dependent on the technology introduced and the media orchestration skills of the teacher [13]. Yet, the concept of engagement is extremely complex and several measures have been proposed in the literature [14].

4. *The possibility of reaching knowledge in an easier way gives to the students the possibility of reflecting and analyzing information*, that is the higher amount of information available may facilitate critical thinking processes. This is one of the most appealing promises offered by technology, yet students need to be prepared, guided and introduced to problem-solving, information gathering and filtering [15].

5. *Students learn where to find information in an information rich world.* This means that students become aware of the large set of information offered by the web and develop methods and strategies to find what they need. Information-related abilities include data filtering, source ratings, content evaluation and organization [16].

6. *Computer skills should not be taught in isolation*. This point underpins the social nature of computational learning, where knowledge is boosted mainly by experience sharing, communication and cooperation in problem solving.

7. *Students develop computer literacy by applying diverse computer skills as part of the learning process.* The level of computer literacy implied in every day life can be increased to a higher level, when technological means are embodied in learning activities, especially if these occur in highly specialized contexts.

2.1 Framing Technology Integration

Several frameworks, models and theories have been proposed in the last 20 years, with the aim of facilitating the process of technology integration in the teaching environment [17].

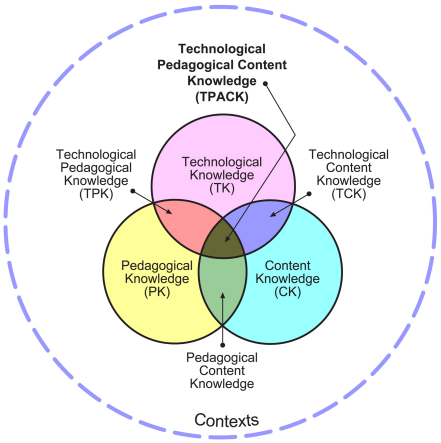


Figure 1. Representation of the TPACK framework. Reproduced by permission of the publisher, ©2012 by `tpack.org`.

In 2016, the International Society for Technology in Education (ISTE) published a framework for innovation in education, in the form of standards to maximize the benefits of using technology in the learning process [1] .

The ISTE standards describe the digital age knowledge and skills, including computational thinking competencies, addressed to students, educators, education leaders, and coaches. These standards set the road map to empower young population to become aware citizens of the information and technology society.

The Technological and Pedagogical Content Knowledge Paradigm (TPACK), shown in Figure 1, tackles the core problem of integrating technologies in the learning process from the perspective of 3 primary forms of knowledge [18]. Content Knowledge (CK) is domain-dependent, it is the teacher's knowledge about the matter to be taught; Technological Knowledge (TK) refers to the information and technological resources, methods and tools to be used in pedagogical contexts; Pedagogical Knowledge (PK) includes instructional strategies, lesson planning and classroom management skills. The 3 secondary forms of knowledge result from the overlap of CK and TK (TCK), TK and PK (TPK) and CK and PK (CPK). Finally, the Technological Pedagogical Content Knowledge emerges at the intersection of the 3 secondary forms (TCK, TPK, and CPK) and is situated in the various contexts where it happens (levels of instructions, places, didactic systems, etc.)

The kernel of the TPACK theory is that it does not results from the simple addition of the 3 primary forms of knowledge as separate sources, but emerges through subsequent integrations between the various areas. TPACK expresses a completely new knowledge field that dynamically integrates the original 3 primary forms and that takes in account all the transformative potential of technology.

On the same line, the SAMR model, shown in Figure 2, re-frames knowledge as a hierarchical network of

[1] `https://www.iste.org/`.

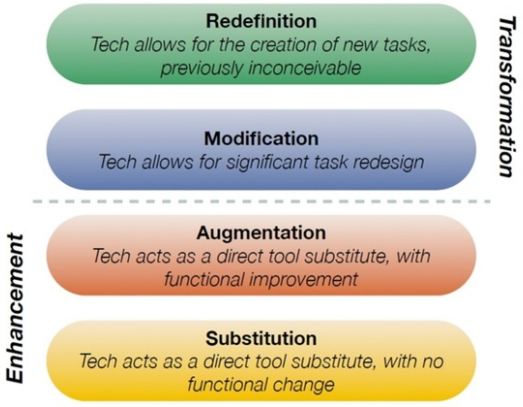


Figure 2. Representation of the SAMR Model.

changes implied in the use of technology in learning processes [19]. This model organizes the technology integration in 4 levels. At the lower level (*Substitution*), technology acts as replacement tool, with no functional changes. That is the case of music sequencers for instance, in which the MIDI representation acts as a mere substitute of the score. Although acting still as replacement tool, *Augmentation* technology introduces functional improvement: The MIDI score can be played at different tempos, be transposed, and eventually be split into sections to ease practicing and rehearsal. It can also be stored, printed and shared through the web. These first two levels form the *Enhancement* stage: The content is essentially untouched, despite the empowerment of the functions.

In the *Transformation* stage, the use of technology in learning processes does affect the content, that can be modified or even redefined. The *Modification* allows some degree of task redesign, for instance in the MIDI metaphor, the score can be reinterpreted by modifying the dynamics and the timbres (i.e., program change) of the various tracks. The last level (*Redefinition*) takes advantage of all the possibilities offered by the technology to redefine the content. For instance, the MIDI score can be deeply manipulated in creative ways, thus allowing the emergence of compositional processes previously inconceivable.

The *Enhancement* and *Transformation* stages in the SAMR model respectively resonate with Brown's analysis on the use of the computer as a *musical tool* or as a *musical medium* [2]. Brown's third category of computer as a *musical instrument* refers to the rather traditional activities of computer music, from digital sound processing to computer assisted composition and live electronics performance. These could be re-framed as a further extension of the SAMR model, thus stressing the artists' ability to practice creativity in the technological environment at their disposal [2].

Finally, the Technology Integration Matrix (TIM), developed by the Florida Center for Instructional Technology (FCIT) [3], represents a further framework for targeting the use of technology to enhance learning. In the TIM framework, 5 interdependent characteristics of meaningful learning environments (*Active, Collaborative, Constructive, Authentic*, and *Goal-Directed*) are associated to 5 levels of technology integration (*Entry, Adoption, Adaptation, Infusion* and *Transformation*), thus forming a $5x5$ squares matrix where each cell is provided with specific setting descriptors, learning objectives, teaching methods, and sample lessons [20].

3. TECHNOLOGY INTEGRATION IN THE MUSICIANS PROFESSIONAL CURRICULUM

Although the frameworks discussed in the previous section are mainly addressed to K-12 level of instruction, some general principles can be well adapted and expanded to higher education and professional training in the music domain. Indeed, if technology integration is a fundamental step at K-12 level, *a fortiori* it must be considered an indispensable element in the professional practice. On the other hand, it may require lots of effort to introduce a proper technology integration in a field which is extremely established and grounded in the crystallized concept of tradition, which is the case of music teaching in the Italian Conservatories of Music.

Nevertheless, a technology integration cannot be realized by simply summing technological courses to the more traditional subjects. Instead, it requires to revise the curricula in depth and creatively, by conceiving novel training courses that respond to the needs of the information and communication age. The following subsections attempt to disclose actions and fields of intervention to effectively integrate technology in the Italian system of music education, by leveraging the analysis of the B.Mus. and M.Mus. in Music – Instruments curricula available at the Music Conservatory of Brescia.

3.1 The Musicians Professional Curriculum: The case of the Music Conservatory of Brescia

The current music education system in Italy derives from the application of the Bologna process [4] and from the reform set by Law No. 508 of December, 21^{st}, 1999 [5].

The education system of professional musicians is currently organized in a framework of Bachelor of Music's (B.Mus.) Degree [6] and Master of Music's (M.Mus.) Degree [7]. The general framework is common to all the conservatories, although some degree of discretion is left to the institutions. In this respect, authors report here the organization of the B.Mus. and M.Mus. in Music – Traditional Instruments curricula of the Music Conservatory of

[2] This ability, that is the advancement and adaptation of musical practices and techniques around music technologies, is deeply rooted in the origin of music making (e.g., the development of keyboard instruments from the harpsichord to the fortepiano and the piano, and the technical and aesthetic changes derived, see `https://schubert.org/museum/keyboard-evolution/`).

[3] `https://fcit.usf.edu/matrix/`.

[4] `https://bit.ly/2Sq4eTU`.

[5] `http://www.miur.it/0006Menu_C/0012Docume/0098Normat/1128Riform_cf4.htm`.

[6] Law Decree No. 124 of September, 30^{th}, 2009 (`http://attiministeriali.miur.it/anno-2009/settembre/dm-30092009-n-124.aspx`).

[7] Ministerial Decree No. 14 of January, 9^{th}, 2018 (`https://shorturl.me/BL8U3cD`).

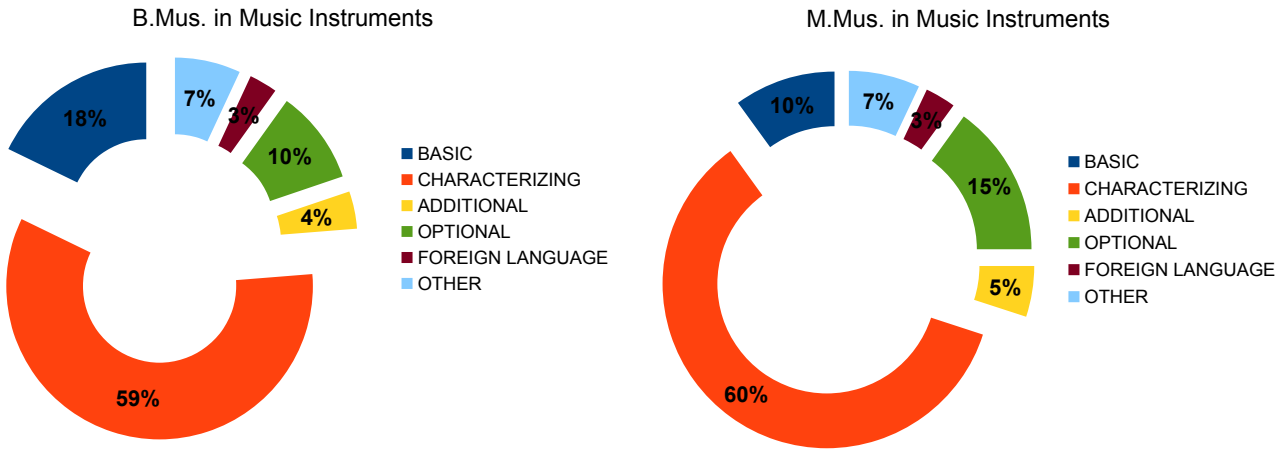


Figure 3. Percentages of ECTS assigned to the various types of training activities in the B.Mus. (left) and M.Mus. (right) in Music – Instruments curricula of the Music Conservatory of Brescia.

Brescia [8], as representative case study.

The curriculum of both degrees organizes the subjects into 4 main categories:

1. Basic/foundational: these courses, aimed at providing the basic knowledge on music, include music theory, ear training, choral practice, music history, analysis, philology, etc.;

2. Characterizing/vocational (and related): these training activities represent the core of the curriculum, and essentially include instrumental practice (solo and ensemble), chamber music and orchestral practice;

3. Additional: these activities mix up various subjects such as acoustics, music informatics, bodily awareness, and others;

4. Optional: these activities can be chosen by students among other subjects offered by the Conservatory (e.g. music analysis, vocal technique, improvisation, music listening, choir conducting, music philosophy, etnomusicology, artistic research, etc.)

+ Foreign language: mainly English (B2);

+ Other: activities related to the production of the final thesis.

Figure 3 shows the distribution of the subjects of B.Mus. and M.Mus. in Music – Traditional Instruments [9], within the 4 categories. Almost the 60% of the curriculum consists of characterizing subjects, namely those related to music practice and training in the various instruments. The basic and additional subjects are shared to all the instrument curricula and represent the actual common ground for the development of the professional musicianship. The technological courses (i.e., acoustics and musical informatics) are included in the additional category and weight respectively 2 and 3 ECTS (less than the 3% of the overall B.Mus. 180 ECTS). Yet, it must be noted that these technological subjects are compulsory only in the B.Mus curriculum.

In this respect, the traditional instrument curricula are shaped as a grid of vertical structures of characterizing subjects, specific per each instrument, linked horizontally by means of basic and additional training activities (which are common to all the curricula). The rationale of such verticality of the curriculum is to train proficient professional instrumentalists, yet it must be observed that the same approach has been adopted in the other less traditional fields of professional training such as Music Education, Music Technology, and Jazz.

3.2 Actions for Technology Integration

Starting back from the instrument, which represents the readily available technological environment of any instrumentalist, a first meaningful intervention is the systematic and progressive introduction of music technologies in the musical practice, from the rather traditional performance with live electronics, to the creative exploration of instrumental techniques combined with digital sound processing. It must be said however, at least in the experience of the authors, that such practices are already informally taking place among the youngest population of students.

As far as concerns the horizontal structure of the curricula (i.e., the basic and additional subjects), there exist already several technological tools to support ear training [10], music analysis [11], and theory [12].

A more robust music technology curriculum should include mandatory subjects and content such as music programming, recording techniques, and audio mastering, and other technology-oriented matters as optional activities. Here

[8] https://www.consbs.it/offerta_formativa/triennio-accademico-di-i-livello/.

[9] The Jazz and Pop curricula are hence not considered.

[10] https://blog.landr.com/best-ear-training-apps/

[11] http://web.mit.edu/music21/

[12] https://www.musictheory.net/tools

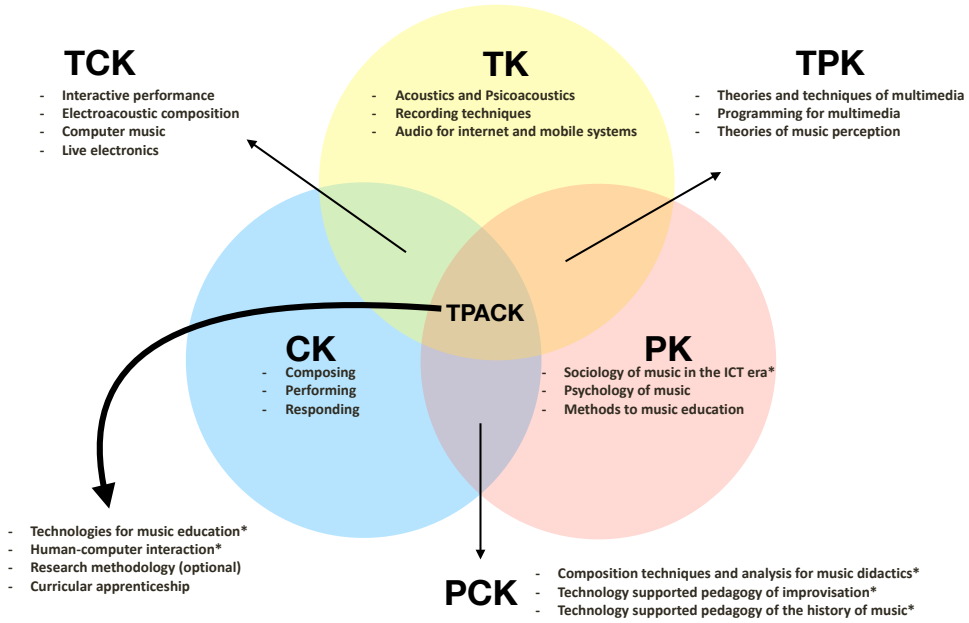


Figure 4. The TDM didactic plan, represented according to the TPACK paradigm. The asterisk indicates a slight modification of the subject's title compared to the ministerial nomenclature.

authors do not claim any exhaustive solution, nor do they intend to analyze the complex phenomenon of knowledge integration that could potentially affect Music Electronics courses when they recognize the need to train teachers of music technologies or other professionals. Rather here the focus is on the urgent need to reconsider music pedagogy in the light of the pervasive societal changes sparked off by technological development. And yet, any effort to reform the current music education system cannot help but tackle a vision on the musician's role in the information and communication society. In the next Section, authors introduce as a case study the M.Mus. degree in Technologies for Music Education, recently launched at the Music Conservatory of Brescia. The curriculum represents the attempt to respond to the needs of new professional profiles, within a novel educational path, inspired by the TPACK paradigm of technology integration.

4. THE "TECHNOLOGIES FOR MUSIC EDUCATION" MASTER IN MUSIC DEGREE

The M.Mus.in "Technologies for Music Education" (TDM) belongs to the master degrees of the Department of Music Education of the Music Conservatory of Brescia. The TDM [13] has been approved by the Italian Ministry of Education, University and Research (MIUR) in 2018, and the first academic year was launched in 2019. Although the M.Mus. in TDM is highly innovative, it is not the result of an official revision process. Rather, it represents the attempt to find the bearings and soften the vertical framework of Masters degrees in Music, as defined by the Ministerial Decree No.14 of January, 9^{th}, 2018, in order to shape novel didactic approaches and contents. To this end, the subjects introduced in the Ministerial Decree of July, 3^{th}, 2009 [14] are reported with their original name, the code of their disciplinary artistic sector, and their modifications (where introduced).

4.1 The didactic plan

The TDM degree aims at training professionals in music education, expert in the use of educational technologies, learning environments and approaches, and capable of developing coherent training paths through research and didactic experimentation.

The TDM curriculum combines diverse subjects on music technology, informatics and music pedagogy. The curriculum rationale softens the vertical paradigm, by informing the organization of subjects according to the TPACK framework discussed in Section 2.1. As shown in Figure 4, the TDM's didactic plan is arranged in a Music Technology Knowledge area (TK), a Music Pedagogy Knowledge area (PK), and a Music Knowledge area (CK), where CK is represented by the background knowledge and musicianship developed during the Bachelor studies. From the overlap of the 3 areas, the secondary forms of knowledge (TCK, TPK and CPK) emerge, and the corresponding subjects are reported as well.

[13] https://shorturl.at/eqR48

[14] http://attiministeriali.miur.it/anno-2009/luglio/dm-03072009-(5).aspx

4.2 The three primary forms of knowledge

Content knowledge (CK) – The music content knowledge is represented by the essential background knowledge and music literacy of the incoming students. In this respect, the 2014 NAfME (National Association for Music Education) standards on music literacy [15] set the essential abilities that student should develop, in order to carry out the three artistic processes of "creating, performing and responding". These processes represent in a synthetic way the actual entry knowledge that any incoming TDM musician must have (regardless the bachelor degree of origin).

Technological Knowledge (TK) – The essential contents are derived from the disciplinary sector of electronic music (artistic disciplinary code COME, in the Italian system). The available subjects include 1) *Acoustics and Psicoacoustics* (COME/03) that provide knowledge on the physics of sound and its perceptual mechanisms; 2) *Sound Recording Techniques* (COME/04) concerned with the literacy on electroacoustic systems, tools, and sound capture and reproduction techniques; 3) *Audio for Internet and mobile systems* (COME/06), to provide basic knowledge and abilities to design and exploit music and audio applications for the network.

Pedagogical Knowledge (PK) – This area draws from the sector of music pedagogy (artistic disciplinary code CODD). Subjects include: 1) *Sociology of music in ICT* (derived from *Fundamentals of Music Sociology*, CODD/06) tackles the processes of music production, distribution and reception in the realm of streaming services and music sharing platforms; 2) *Psychology of music* (CODD/04) provides empirical knowledge on musical behaviour and experience at large; 3) *Methods for music education* (CODD/04) provides an overview of several approaches to music teaching.

4.3 The three secondary forms of knowledge

As shown in Figure 4, the dual overlap between the primary areas of knowledge generates 3 secondary forms of knowledge integration.

Technological Content Knowledge (TCK) – This area draws on several COME subjects (i.e., electronic music) and namely, 1) *Interactive performance* (COME/01) focuses on multimodal interaction and systems for music performance; 2) in *Electroacoustic composition* (COME/02), the history of electroacoustic music production is analyzed, and several means for sound organization are tackled; 3) on the same line, the course on *Computer music* (COME/01) stresses the music creation from the algorithmic perspective; 4) *Live Electronics* (COME /01) combines all these aspects, and the strategies for musical interaction, liveness and performativity are put in the foreground.

This corpus of subjects aims at providing the students with a solid technological awareness on the various approaches to music technology education.

Technological Pedagogical Knowledge (TPK) – 3 subjects from the COME/06 area are included: 1) *Theories and techniques of multimedia* and 2) *Programming for multimedia* explore multimedia communication and cognition, and provide strategies and means to exploit different media in the design of effective interactive educational tools; 3) *Theories of music perception* complements the other two subjects, and focuses on the pedagogical aspects by providing an overview on auditory perception and cognition, musical meaning-making and emotions.

Pedagogical Content Knowledge (PCK) – 3 subjects taken from the music pedagogy area CODD are available: 1) in *Composition techniques and analysis for music didactics* (CODD/02), the students learn the principles of music composition addressed to children ensembles (Orff instruments [16]); 2) *Technology supported pedagogy of improvisation* (derived from *Pedagogy of music improvisation*) (CODD/02), combines Orff instruments and technological music tools to practice music improvisation; 3) *Technology supported pedagogy of history of music* (derived from *Pedagogy of the history of music*) (CODD/06) proposes novel approaches and tools to teach the history of music. Taken together, the subjects aim at providing the students with a sound background in music teaching practices and opportunities.

4.4 The Technological, Pedagogical and Content Knowledge (TPACK) area

This area includes 4 core subjects of knowledge integration. *Technologies for music education* (adapted from *Music pedagogy*, CODD/04) provides an in-depth knowledge on tools, systems and application for music education, ranging from the GUIDO (Graded Units for Interactive Dictation Operations) system paradigm of programmed instruction for music, to artificial intelligence and machine learning techniques [21], in the wider scope of *i)* learning theories, including behaviorism, social cognitivism and constructivism, *ii)* approaches to formal and informal learning, such as connectivism, on line and blended learning, and gamification [22], and *iii)* the established standards for technology integration.

Human-computer interaction (adapted from *Informatica musicale*, COME/05) provides students with the basic knowledge and skills on the interaction paradigms between humans and digital tools and representations, ranging from post-WIMP interfaces, to embodied strategies to sensing and actuating technologies [23, 24]. A relevant part is devoted to the understanding and exploitation of the sound-gesture relationships and their computational augmentation in the educational setting and applications.

Research Methodology is instrumental to provide the future educator with the indispensable knowledge and skills to better comprehend and assess the learning mechanisms involved in the integration of technologies in music teaching, especially when educational applications are not ready to hand and some sort of technological development must

[15] https://nafme.org

[16] Orff instruments are a set of simple percussive instruments (pitched and not pitched) created by the German composer Carl Orff (1895-1982). They are used in the context of pedagogical activities based on movement and improvisation.

be undertaken. Although this course is optional and yet highly recommended, it provides the student with a basic literacy on the scientific approach to research, to wide the reach of her future collaborations in the academic field.

Curricular apprenticeship. The Music Conservatory of Brescia benefits of a large network of schools at various levels, that can host the didactic experimentation. The apprenticeship represents the final outcome of the M.Mus. in TDM, and more importantly the real benchmark of TPACK, whose effectiveness can thus be assessed in real settings.

4.5 Report of TDM Activities

At the time of writing this report TDM has completed its first full two years cycle with 4 students who will support their thesis defense probably in March 2023.

From the teachers point of view there are at least two major challenges in the realization of the TDM program. Firstly a great effort in the preparation of the courses, many of which were brand new for the Conservatoire. Consider for instance the "Technologies for Music Education" (TDM) course, whose content is derived directly from research and the "Human-Computer Interaction" course, not a novelty in Universities but a completely new research field for music education in Conservatories. Secondly, while much cooperation and interest came from colleagues in the Department of Electronic Music, an almost complete closure and unwillingness for change has been registered from all the colleagues (but one) of the Music Education Department.

From the point of view of the students, when asking for information many of them get scared of the test programming required for admission to TDM, and give up without even giving it a try. Students currently attending report that one of the main benefits of TDM is the knowledge of many different ways to use technology in the classroom and, consequently, the possibility of choosing the best solution according to the context. Technology is not only a tool, but rather it shapes the learning environment with creative solutions for making music and reflecting. However, TDM is also very challenging because it requires excellent basic knowledge and the ability to benefit from cooperation with other students.

5. CONCLUSION

Technology integration does not happen simply by adding technological subjects to the curriculum, but rather by conceiving novel paths to include technology in everyday teaching. This goal is extremely difficult to achieve in highly established contexts, such as the Italian music education system (i.e., music conservatories), which devotes the major efforts to the instrumental practice. However, no music profession nowadays can be successful without a minimum degree of digital literacy, which is poorly present in the traditional instruments music curricula.

In addition, the sector of electronic music studies shows similar limitations. According to Sapir et al. (2018), one of the main employment opportunities for a student of electronic music is teaching [25]. If that is the case, the question comes into play, why not providing students with the appropriate knowledge and skills to tackle their future profession in the best way? This is where technology integration actually hits the spot. Indeed a more integrated approach to the design of the various curricula may offer several advantages:

1. *The improved and closed connection to the needs of real world professions.* For instance music technology applications for people with special needs is a potentially relevant subject in the M.Mus. in Music Therapy (active in Brescia, as well as in other music conservatories [17, 18]). In a similar way the technological, aesthetic and perceptual background of the professional sound production and design in various application fields, including interactive media and video games may be extended also to other Conservatory courses, beyond the Electronic Music courses;

2. *The possibility to prepare the ground for academic research opportunities in music conservatories, and to establish practices, protocols, ways of knowing in the artistic domain.* In this respect, a large body of work is readily available in the communities of Sound and Music Computing (SMC), Sonic Interaction Design (SID), and New Interfaces for Musical Expression (NIME).

3. *The introduction of a research mission in perspective makes the establishment of Ph.D courses much more achievable and credible.* In this respect, the Music conservatory of Brescia and the LIM [19] (Laboratory of Music Informatics) of the University of Milan are already working on a research partnership for the prototyping and evaluation of music education applications and other relevant research projects.

Finally, authors hope with this contribution to sparkle a debate on the technology integration in music education, and a critique on the current rigid top-down structure of professional curricula, with the aim of envisioning shared strategies to reform and update the cultivation of the digital musicianship.

6. REFERENCES

[1] A. Hugill, *The digital musician*. Routledge, 2018.

[2] A. R. Brown, *Computers in music education: Amplifying musicality*. Routledge, 2012.

[3] A. Vallgårda, M. Winther, N. Mørch, and E. E. Vizer, "Temporal form in interaction design," *International Journal of Design*, vol. 9, no. 3, 2015.

[4] G. Volpe, K. Kolykhalova, E. Volta, S. Ghisio, G. Waddell, P. Alborno, S. Piana, C. Canepa, and R. Ramirez-Melendez, "A multimodal corpus for technology-enhanced learning of violin playing," in *Proceedings*

[17] http://www.conservatorioverona.it/it/musicoterapia/

[18] https://www.consaq.it/musicoterapia-home.html

[19] http://www.lim.di.unimi.it/

of the 12th Biannual Conference on Italian SIGCHI Chapter, CHItaly '17, (New York, NY, USA), Association for Computing Machinery, 2017.

[5] A. M. Webb, C. Wang, A. Kerne, and P. Cesar, "Distributed liveness: Understanding how new technologies transform performance experiences," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work Social Computing*, CSCW '16, (New York, NY, USA), p. 432–437, Association for Computing Machinery, 2016.

[6] L. Kardos, "Making room for 21st century musicianship in higher education," *Action, Criticism, and Theory for Music Education*, vol. 17, no. 1, pp. 1–14, 2018.

[7] O. Lähdeoja, "Composing the context: Considerations on materially mediated electronic musicianship," *Organised Sound*, vol. 23, no. 1, p. 61–70, 2018.

[8] J. M. Wing, "Computational thinking," *Communications of the ACM*, vol. 49, no. 3, pp. 33–35, 2006.

[9] R. Vuorikari, Y. Punie, S. C. Gomez, G. Van Den Brande, *et al.*, "Digcomp 2.0: The digital competence framework for citizens. update phase 1: The conceptual reference model," tech. rep., Joint Research Centre (Seville site), 2016.

[10] J. Dockstader, "Teachers of the 21st century know the what, why, and how of technology," *THE journal*, vol. 26, no. 6, pp. 73–75, 1999.

[11] B. Skinner, "Programmed instruction revisited.," *Phi Delta Kappan*, vol. 68, no. 2, pp. 103–10, 1986.

[12] T. Laird and G. Kuh, "Student experiences with information technology and their relationship to other aspects of student engagement," *Research in Higher education*, vol. 46, no. 2, pp. 211–233, 2005.

[13] L. F. M. G. Pedro, C. M. M. de Oliveira Barbosa, and C. M. das Neves Santos, "A critical review of mobile learning integration in formal educational contexts," *International Journal of Educational Technology in Higher Education*, vol. 15, no. 1, p. 10, 2018.

[14] C. R. Henrie, L. R. Halverson, and C. R. Graham, "Measuring student engagement in technology-mediated learning: A review," *Computers & Education*, vol. 90, pp. 36–53, 2015.

[15] K. Kereluik, P. Mishra, C. Fahnoe, and L. Terry, "What knowledge is of most worth: Teacher knowledge for 21st century learning," *Journal of digital learning in teacher education*, vol. 29, no. 4, pp. 127–140, 2013.

[16] R. E. Anderson, "Implications of the information and knowledge society for education," in *International handbook of information technology in primary and secondary education*, pp. 5–22, Springer, 2008.

[17] E. R. Hamilton, J. M. Rosenberg, and M. Akcaoglu, "The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use," *TechTrends*, vol. 60, no. 5, pp. 433–441, 2016.

[18] M. Koehler and P. Mishra, "What is technological pedagogical content knowledge (TPACK)?," *Contemporary issues in technology and teacher education*, vol. 9, no. 1, pp. 60–70, 2009.

[19] R. Puentedura, "Building transformation: An introduction to the SAMR model [blog post]," *Retrieved February*, vol. 4, p. 2019, 2014.

[20] J. C. Harmes, J. L. Welsh, and R. J. Winkelman, "A framework for defining and evaluating technology integration in the instruction of real-world skills," in *Handbook of research on technology tools for real-world skill development*, pp. 137–162, IGI Global, 2016.

[21] S. Holland, "Artificial intelligence in music education: A critical review.," *Readings in Music and Artificial Intelligence*, vol. 20, pp. 239–274, 2000.

[22] J. Lowyck, "Bridging learning theories and technology-enhanced environments: A critical appraisal of its history," in *Handbook of research on educational communications and technology*, pp. 3–20, Springer, 2014.

[23] R. J. Jacob, A. Girouard, L. M. Hirshfield, M. S. Horn, O. Shaer, E. T. Solovey, and J. Zigelbaum, "Reality-based interaction: a framework for post-wimp interfaces," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 201–210, 2008.

[24] J. A. Jacko, *Human computer interaction handbook: Fundamentals, evolving technologies, and emerging applications*. CRC press, 2012.

[25] S. Sapir, G. Cospito, M. Marchi, R. Sinigaglia, and R. Tajé, "Per l'istituzione di corsi di formazione alla ricerca nei settori artistico-disciplinari della musica elettronica." https://www.comecrescendo.it/wp-content/uploads/2018/11/sapir-milano.pdf. Accessed: 2020-03-02.

MOVE: UN AMBIENTE MULTIMODALE INTERATTIVO PER L'APPRENDIMENTO

Maurizio Botti, Maura Castelnovo, Ottavia Marini, Sara Valente e Marcella Mandanici
Conservatorio "L. Marenzio" - Brescia
botti.maurizio, castelnovo.maura, ottavia.marini, sara.valente, marcella.mandanici@conservatorio.brescia.it

ABSTRACT

Questo articolo presenta un'esperienza didattica in cui è stato utilizzato MOVE, ambiente multimodale interattivo progettato nell'ambito del Biennio di Tecnologie per la Didattica della Musica del Conservatorio di Brescia. Attualmente il sistema comprende quattro diversi giochi basati sul movimento dell'utente all'interno di una porzione di pavimento rilevata da una telecamera montata a soffitto. Questi giochi sono stati oggetto di una sperimentazione presso una scuola secondaria di primo grado. Durante le attività didattiche sono state utilizzate tecniche di osservazione che hanno portato all'individuazione di alcune importanti risultati legati all'utilizzo di MOVE quali l'inclusione, il coinvolgimento e lo sviluppo della memoria e della creatività.

1. INTRODUZIONE

La definizione di ambiente multimodale interattivo non è univoca in letteratura e varia in dipendenza dell'ambito di studi. Secondo Jaimes e Sebe un sistema multimodale è tale se risponde alle sollecitazioni in ingresso appartenenti a più di una modalità o canale di comunicazione (parole, gesti, testo, etc.) [1]. Diversamente, la classica definizione di ambiente multimodale interattivo di Moreno e Mayer - ricavata dal lavoro di A. Paivio [2] - prevede invece la rappresentazione della conoscenza attraverso due tipi di canali, verbale e non-verbale [3]. Unificando queste due visioni provenienti rispettivamente dagli ambiti di studio dell'interazione persona-macchina e della psicologia dell'apprendimento, si può dire che gli ambienti multimodali interattivi si riferiscono genericamente a sistemi che prevedono la combinazione di diverse modalità sia in ingresso (ossia modalità di interazione) sia in uscita (ossia rappresentazione del contenuto) [4]. Parte fondante del sistema MOVE è la sua modalità di interazione, che è basata sul movimento dell'intero corpo su una superficie compresa nel raggio visivo di una telecamera. La superficie così delimitata diventa uno spazio interattivo che ha il potere di unire la percezione del corpo fisico (cinestesia, orientamento, propriocezione) con la ricchezza e la varietà del mondo digitale. L'utente interagisce attraverso il movimento e vive lo spazio essendo cosciente di trovarsi in un ambiente reale regolato dalle leggi della fisica elementare. Allo stesso tempo genera feedback sonori e visivi che non provengono dallo spazio fisico e che lo coinvolgono percettivamente e cognitivamente. Ambienti di questo tipo si prestano a diversi utilizzi e possono essere impiegati efficacemente nell'apprendimento [1], nell'intrattenimento [2] e nella riabilitazione [3]. L'interazione ludica, il movimento fisico e, a volte, la manipolazione di oggetti tangibili vengono combinati per creare un'esperienza coinvolgente e inclusiva, che può dare luogo a diverse interrelazioni fra il giocatore e coloro che assistono all'attività [5].

Copyright: ©2022 First author et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1.1 Ambienti Multimodali Interattivi per l'apprendimento

Partendo dai primi esperimenti di Myron Krueger nel 1971, gli ambienti multimodali interattivi basati sul movimento si sono sviluppati utilizzando varie tecnologie [6]. Anticipando le teorie dell'*embodied cognition*, lo stesso Krueger ha evidenziato come l'alto potere di coinvolgimento di questi ambienti sia proprio dipendente dall'uso del movimento fisico come modalità di interazione. I bambini, quindi, potrebbero apprendere meglio se viene data loro l'opportunità di costruire conoscenza attraverso il fare, avendo il pieno controllo del corpo in un ambiente di gioco. Un esempio è il sistema "SMALLab" [4], un ambiente sviluppato da un team di ricerca dell'Arizona State University. Il sistema utilizza videocamere per il tracciamento del movimento e un videoproiettore che fornisce un feedback visivo sul pavimento stesso. Gli alunni interagiscono muovendosi all'interno di scenari di gioco dinamici che facilitano l'interazione sociale e l'apprendimento collaborativo [7]. Allo stesso modo "Wizefloor" [5], sviluppato dal centro di ricerca danese Alexandra Instituttet A/S [6], propone un'area interattiva in cui vengono proiettati quiz, giochi di memoria e di categorizzazione [8]. Si tratta in entrambi i casi di sistemi commerciali che permettono una facile installazione in contesti scolastici offrendo esperienze allineate ai contenuti disciplinari. Con finalità diverse è stata progettata la "Stanza Logomotoria" [7], realizzata presso l'Università di Padova nel 2008. Il sistema, simile ai precedenti per la tecnologia utilizzata, è basato sul modello

[1] https://www.breezecreative.com/dynamic-floor
[2] https://www.design-io.com/projects/connectedworlds
[3] https://www.snoezelen.info/
[4] https://www.smallablearning.com/
[5] https://www.wizefloor.dk
[6] https://alexandra.dk/
[7] http://smc.dei.unipd.it/logomotoria.html/

della “memoria risonante”, ovvero la capacità dell’utente di memorizzare un suono legato ad una precisa posizione spaziale [9]. Il bambino prima esplora i suoni e poi cerca di abbinarli ad una narrazione proposta dal sistema stesso. Altre applicazioni basate sulla rappresentazione spaziale di concetti musicali o finalizzate alla riabilitazione di soggetti non vedenti sono state realizzate da Mandanici et al. [10, 11].

Dopo aver contestualizzato il sistema MOVE sia nell’ambito dell’interazione persona-macchina, sia nell’ambito della multimodalità come pratica educativa, il resto di questo articolo presenta le componenti hardware e software del sistema (Sezione 2.1), specificando successivamente le caratteristiche di ognuno dei quattro giochi implementati (Sezione 2.2). La metodologia e le finalità della sperimentazione incentrata sui quattro giochi del sistema MOVE sono descritte nella Sezione 3, unitamente alla presentazione dei progetti didattici che hanno fatto da contesto all’utilizzo dei giochi stessi. La Sezione 4 riporta gli esiti dell’osservazione effettuata durante la sperimentazione, mettendo in risalto e interpretando i comportamenti ricorrenti degli alunni e le loro reazioni. La Sezione 5 riassume i risultati ottenuti, prefigurando gli sviluppi futuri sia dal punto di vista del sistema sia da punto di vista del suo utilizzo didattico.

2. MOVE

Il sistema MOVE è un ambiente multimodale interattivo che prevede quattro giochi musicali basati sul movimento del corpo. Il sistema è semplice, economico e facile da installare e da trasportare. Queste caratteristiche lo rendono fruibile in qualunque ambiente scolastico.

2.1 Architettura di sistema

La parte hardware del sistema, illustrata in Figura 1, è costituita da una telecamera ancorata al soffitto, due diffusori audio necessari per l’ascolto stereofonico, un monitor o un proiettore e un computer a cui vengono collegati fisicamente questi dispositivi. Viene impiegato anche un tablet in grado di comunicare con il computer attraverso una comune connessione Wi-Fi. Il tablet ha una funzione di interfaccia utente implementata attraverso ”Mira-web”, un software che permette il controllo e la riproduzione di una patch su un qualsiasi browser [8] . Il computer invece esegue il software di gioco che si compone di tre parti essenziali:

1. **Il sistema di motion tracking**, implementato utilizzando l’ambiente MAX [9] e in particolare la libreria di J.M. Pelletier [10] [12] e gli oggetti nativi di Jitter per il calcolo medio dei valori dei piani di una matrice.

2. **Il sistema di output audio**, anch’esso implementato in MAX, che in funzione del gioco selezionato, impiega audio preregistrato, sintesi audio o MIDI.

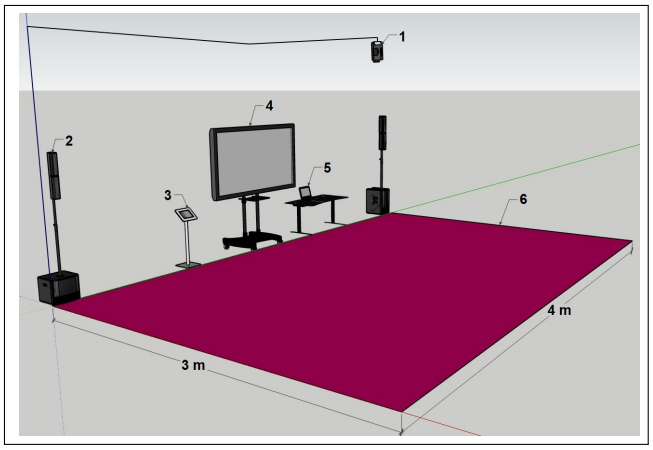


Figure 1. Architettura hardware del sistema MOVE composta da (1) videocamera, (2) diffusori stereo, (3) tablet, (4) monitor, (5) computer, (6) tappeto rosso in etilene vinil acetato antiriflesso.

3. **Il sistema di output video** sviluppato utilizzando Processing 3.0 [11] che gestisce la parte grafica durante le diverse fasi di gioco.

La telecamera a soffitto traccia il movimento dell’utente all’interno dell’area di gioco, corrispondente alla porzione di pavimento coperta dal raggio della telecamera. L’area tracciata dipende quindi dal tipo di telecamera e dalla distanza della stessa con il pavimento. Nel nostro caso la telecamera C270 HD Webcam [12] , in una stanza di circa 3m di altezza, produce un’area di gioco pari a 4x3 metri. Per il corretto funzionamento degli algoritmi di tracciamento del movimento quest’area dev’essere opaca, di colore uniforme e in contrasto con il colore dell’abbigliamento dell’utente. Per questo motivo sul pavimento è stato posizionato un tappetino in etilene vinil acetato di colore rosso ed è stato chiesto agli utenti di indossare vestiti scuri. Anche l’illuminazione dell’ambiente deve essere controllata in modo che la luce proveniente dall’esterno non interferisca col rilevamento della sagoma dell’utente. Il sistema di tracciamento del movimento riceve come input la ripresa in tempo reale della telecamera e converte la posizione dell’utente in coordinate cartesiane. I dati acquisiti vengono utilizzati in maniera differente in funzione della progettazione dei vari giochi.

2.2 I quattro giochi di MOVE

MOVE è costituito da quattro esperienze di gioco. Nonostante i giochi siano focalizzati su elementi musicali, in essi sono rilevabili numerosi elementi multidisciplinari che proiettano queste attività giocose oltre i limiti della singola disciplina. I giochi sono individuali ma prevedono comunque varie forme di interazione con chi assiste all’attività. L’interfaccia di controllo riprodotta in Figura 2 consente la selezione del gioco desiderato.

[8] https://cycling74.com/articles/content-you-need-miraweb
[9] https://cycling74.com/
[10] https://jmpelletier.com/cvjit/
[11] https://processing.org/
[12] https://www.logitech.com/it-it/products/webcams/c270-hd-webcam.960-001063.html

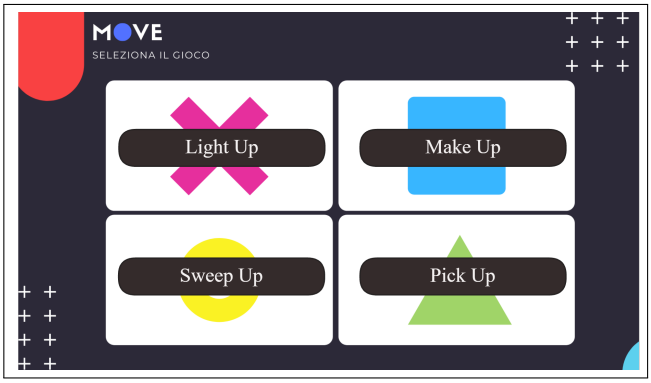


Figure 2. Schermata *Home* dell'interfaccia utente per la selezione del gioco.

2.2.1 *Light-up*

In questo gioco [13] si propone all'utente di esplorare e riconoscere gli elementi caratteristici di un quadro, in questo caso "Giallo, Rosso, Blu" di W.Kandinsky, illustrato nella parte superiore di Figura 3 [13]. Nel gioco sedici particolari pittorici significativi del quadro sono associati ad elementi sonori caratterizzanti, corrispondenti a sedici file audio appositamente creati secondo la "Teoria del Colore" di Kandinsky [14]. Come si può vedere nella Figura 3 il sistema segue la posizione dell'utente con un alone luminoso che svela soltanto una piccola porzione del dipinto, lasciandone il resto oscurato. L'utente entra nell'ambiente e cerca di individuare i sedici elementi. Una volta completata la ricerca il quadro si svela al giocatore nel suo insieme unitamente ad una sonorizzazione che include tutti i suoni associati agli elementi caratteristici. Il gioco permette di sperimentare la sinestesia fra suono e immagine in modo interattivo e dinamico, stimola la memoria e la coordinazione nello spazio e avvicina gli utenti ai linguaggi espressivi contemporanei.

2.2.2 *Sweep-up*

In questo gioco viene richiesto all'utente di porre attenzione ad un suono, in particolare alla sua frequenza (altezza), e di ricercarlo attraverso il movimento fisico nello spazio basandosi solo su informazioni uditive (il gioco si esegue da bendati) [14]. Il sistema genera una sinusoide che varia in ampiezza e in intensità in funzione della distanza delle coordinate dell'utente rispetto a un punto precedentemente inizializzato dal gioco in modo randomico. Come si può osservare in Figura 4 l'output visivo mostra il punto da raggiungere, rappresentato dal cerchio nero, e la posizione del giocatore nello spazio, rappresentato da una mosca. Il giocatore non può vedere lo schermo, ma questo funge da monitoraggio del gioco da parte dei restanti alunni della classe che possono guidare il compagno nel raggiungimento dell'obiettivo. "Sweep-up" è un gioco multidiscplinare in quanto permette attraverso il suono di avvicinare gli utenti a concetti scientifici, come in questo caso alla fisica del suono. Il movimento del corpo è strettamente legato all'ascolto, aumentando così il coinvolgimento degli utenti e

[13] https://youtu.be/qB_jAXJBd1k
[14] https://youtu.be/lB9MXWvAJTI



Figure 3. Nella parte superiore il quadro "Giallo, Rosso, Blu" di W.Kandinsky, nella parte inferiore l'alone luminoso che svela soltanto una piccola porzione del dipinto a seconda della posizione dell'utente.

la comprensione di concetti musicali come l'altezza e l'intensità. Ai giocatori vengono richieste abilità di ascolto, coordinazione, orientamento spaziale e memorizzazione.

2.2.3 *Pick-up*

Questo gioco è basato su registrazioni multi traccia che consentono all'utente un ascolto selettivo della composizione. Infatti ad ogni traccia corrisponde una diversa funzione musicale (ad esempio basso, melodia, accompagnamento ecc). [15]. Un approccio simile è stato seguito da F.Pachet che nel 2000 ha proposto un ascolto dinamico di un mix in base al movimento del mouse su un'interfaccia grafica [15]. Lo spazio di gioco di "Pick-up" è suddiviso in quattro settori di uguali dimensioni a ognuno dei quali corrisponde uno strumento presente nel mix. L'utente può quindi aumentare il volume di uno specifico strumento spostandosi nel settore appropriato, identificando tale settore tramite il video. Sullo schermo vengono infatti proiettate le immagini stilizzate dei quattro strumenti presenti nella composizione, e quando l'utente si posiziona nell'area corrispondente ad un determinato strumento questo assume una colorazione diversa dagli altri. Il gioco permette la navigazione in diversi brani di stili differenti, consentendo così anche un'associazione tra stile e strumento. Ad esempio uno dei brani previsti dal sistema è il "Quartetto per archi n.17 in si bemolle maggiore K 458" di Mo-

[15] https://youtu.be/ZUFLLUSb3_c

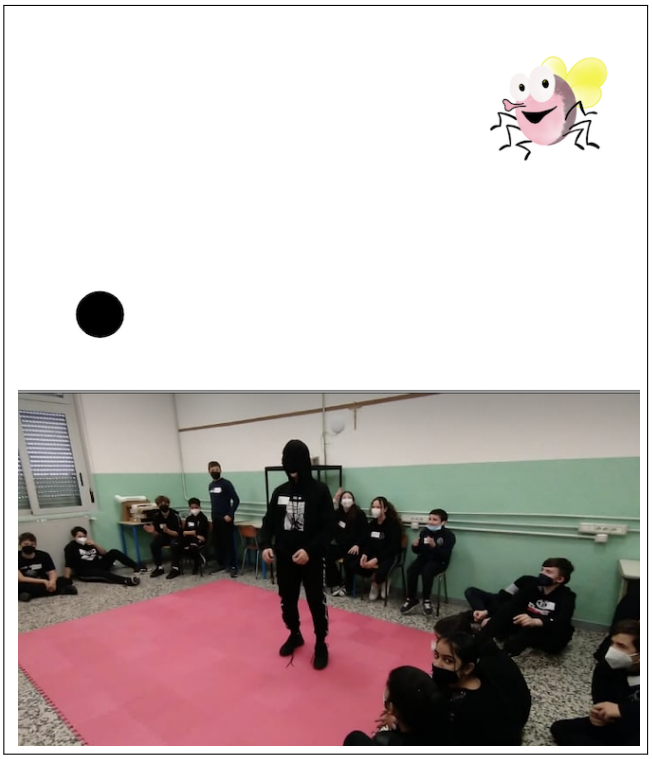


Figure 4. Nella parte superiore l'output video, relativo al gioco "Sweep-up", dove il punto da raggiungere è rappresentato dal cerchio nero, mentre la mosca rappresenta la posizione dell'utente. Nella parte inferiore dell'immagine è rappresentata un'esperienza di gioco in classe.

zart [16] con il quale l'utente può entrare in confidenza con la formazione del quartetto d'archi e focalizzare l'attenzione su uno dei quattro strumenti coinvolti portandolo in primo piano.

2.2.4 Make-up

Il gioco "Make-up" [17] invita l'utente ad improvvisare e creare un prodotto sonoro partendo dall'esplorazione dei modelli tratti dalla composizione "In C" di T. Riley [16]. La partitura, infatti, si presenta già suddivisa in modelli ritmico-melodici, facilitando la trasposizione dei frammenti nello spazio di gioco. Questo è stato tratto dalla struttura a griglia descritta da Mandanici ed al. per l'applicazione "Jazz Improvisation" [17]. I diversi modelli sono distribuiti sulla superficie attraverso una griglia 5x4, per un totale di 20 caselle. Ogni modello è riprodotto tramite l'esecuzione di un file MIDI caratterizzato da diversi timbri, eventualmente manipolabili in maniera interattiva. La presenza o l'assenza dell'utente sulle caselle attiva o disattiva la riproduzione in loop dei vari modelli. Le abilità in gioco sono molteplici: ascolto, memoria, abilità di classificazione e catalogazione e creatività.

3. SPERIMENTAZIONE

Il sistema MOVE è stato impiegato per una sperimentazione svoltasi nel mese di Marzo 2022, nella Scuola Secondaria di primo grado "Don Gaetano Portioli", situata ad Ostiano, in provincia di Cremona. La scuola propone un progetto educativo di ispirazione montessoriana in cui l'apprendimento nasce da una libera esplorazione delle risorse disponibili, comprese quelle tecnologiche. L'esperienza significativa del fare permette una costruzione della conoscenza che si concretizza nella produzione di artefatti di varia natura. In questo contesto la tecnologia non è strettamente legata alla disciplina, ma piuttosto diventa strumento trasversale con cui lavorare in maniera collaborativa per il raggiungimento dell'obiettivo [18]. La scuola, attraverso spazi innovativi, inclusivi e dinamici, incoraggia e sostiene un approccio laboratoriale alle attività. Le tecnologie sono presenti in ogni ambiente attraverso strumenti come LIM [18], Big Pad [19], Stampanti 3D [20]. Inoltre l'istituto adotta il protocollo BYOD (Bring Your Own Device) che permette agli alunni di utilizzare i propri dispositivi personali come supporto alle attività in modo competen-

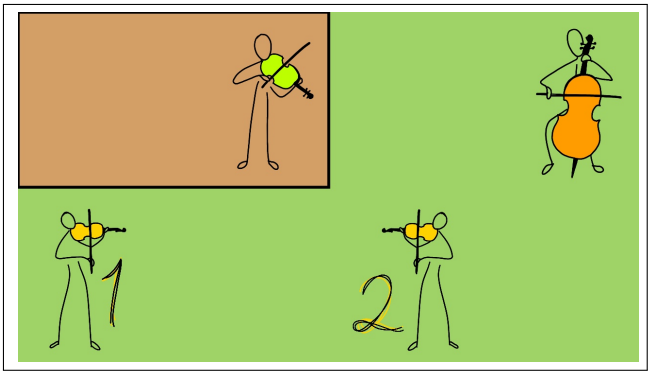


Figure 5. Output video relativo al gioco "Pick-up". L'immagine rappresenta i quattro strumenti presenti nella composizione "Quartetto per archi n.17 in si bemolle maggiore K 458" di Mozart nel momento in cui lo studente si trova nell'area corrispondente alla viola.



Figure 6. Fase di gioco relativa a "Make-up". Un alunno sperimenta l'attività circondato dai compagni di classe, mentre sullo schermo vengono visualizzate in azzurro le caselle che rappresentano le caselle attive e in rosa quelle inattive.

[16] https://youtu.be/LH2_0q5UnLw
[17] https://youtu.be/mydg26Ewox8
[18] https://www.campustore.it/ambienti-digitali/lim/lim-campusboard.html
[19] https://www.sharp.it/cps/rde/xchg/it/hs.xsl/-/html/monitor-interattivi-e-big-pad.htm
[20] https://www.campustore.it/elettronica-e-fablab/stampanti-e-scanner-3d.html

te e responsabile. Le possibilità di apprendere sono così ampliate sia per la frequentazione di ambienti digitali condivisi, sia per l'accesso alle informazioni, coinvolgendo le diverse discipline in un'educazione globale e trasversale.

3.1 I Progetti didattici

Le attività didattiche progettate intorno al sistema MOVE si sono integrate perfettamente con il modus operandi della Scuola. L'approccio multidisciplinare, già presente nei diversi giochi, ha rappresentato il punto di partenza per la progettazione di quattro percorsi didattici ispirati da ogni gioco. Sono state coinvolte quattro classi (due prime e due seconde), ognuna delle quali ha sperimentato uno dei quattro percorsi didattici durante due incontri di cinque ore ciascuno.

3.1.1 Light-up: esplora l'arte attraverso il suono

Il percorso didattico che impiega "Light-up" è volto ad esplorare il legame tra arte pittorica e suono tramite le tecnologie, aumentare la capacità di ascolto, stimolare la creatività nel creare legami sinestetici tra forma-colore e suono ed infine ampliare le competenze digitali degli alunni. Partendo dalla "Teoria del colore" di W.Kandinsky, gli alunni sono stati guidati nella creazione di una personale sonorizzazione del quadro "Giallo, Rosso, Blu" (Figura 3), attraverso un lavoro sia personale sia collaborativo. Gli obiettivi sono stati saper reperire file audio dal web (utilizzando ad esempio il sito Freesound [21]), o crearli autonomamente tramite registrazioni fatte con il proprio smartphone, saper gestire semplici processi di editing audio [22] e organizzare gallerie di suoni. Aver ampiamente sperimentato il gioco prima di strutturare una sonorizzazione del quadro ha permesso agli alunni di familiarizzare con linguaggi estetici molto lontani dalle loro routine in modo immediato e coinvolgente.

3.1.2 Sweep-up: alla scoperta del suono

Il percorso didattico costruito sul gioco "Sweep-up" è volto ad indagare le caratteristiche fisiche del suono (altezza, intensità) attraverso l'impiego di tecnologie digitali, promuovendo lo sviluppo del pensiero computazionale. Il gioco ha rappresentato uno stimolo iniziale per confrontare l'esperienza manipolativa del suono attraverso il corpo con le possibilità espressivo-gestuali del Theremin, il primo strumento musicale elettronico basato su variazioni di frequenza e ampiezza controllate dal movimento delle mani dell'esecutore [19]. Gli alunni hanno approfondito le caratteristiche dell'onda sonora, analizzando i risultati percettivi di diverse frequenze attraverso esempi visivi e uditivi. Successivamente, attraverso un lavoro collaborativo, hanno svolto una ricerca sulle caratteristiche e il funzionamento del Theremin. In una seconda fase gli alunni hanno costruito il circuito elettrico necessario per la realizzazione di un prototipo dello strumento con "Arduino Student Kit" [23] , e ragionato sul codice. In questo modo è stato possibile confrontare le possibilità espressivo-gestuali dell'artefatto con l'esperienza fisica nell'ambiente interattivo.

3.1.3 Pick-up: ascolta i protagonisti della musica

Il progetto didattico legato al gioco "Pick-up" è basato sull'esplorazione degli stili musicali e sulla classificazione degli strumenti secondo Hornbostel-Sachs [20]. In una prima fase è stata introdotta la suddivisione degli stili musicali sulla base della tricotomia *musica colta, musica tradizionale e popular music* introdotta dal musicologo P.Tagg [21], focalizzando l'attenzione sugli strumenti impiegati. In secondo luogo è stata spiegata la tassonomia degli strumenti musicali secondo Hornbostel-Sachs fornendo diversi esempi di ascolto. L'obiettivo di questo percorso è, oltre all'acquisizione delle nozioni relative alla tricotomia e alla classificazione degli strumenti, anche quello di permettere agli alunni di familiarizzare e di riconoscere i timbri strumentali più diffusi negli stili citati. L'ambiente interattivo è stato utilizzato con due finalità differenti: in un primo momento l'obiettivo è stato l'esplorazione dei timbri strumentali presenti nei brani proposti, mentre in un secondo momento lo spazio interattivo è stato impiegato con finalità di verifica. È stato infatti richiesto agli alunni di identificare solo attraverso l'ascolto lo strumento portato in primo piano dall'alunno che interagiva con l'ambiente stimolando così, oltre alle abilità di ascolto, anche capacità di interazione sociale.

3.1.4 Make-up: compositore per un giorno

Il percorso didattico proposto promuove la composizione musicale in ambiente informatico, utilizzando ognuno i propri mezzi (computer, tablet, cellulare) come strumenti didattici creativi, capaci di facilitare il pensiero critico e la riflessione. In una prima fase sono stati esplorati software per l'editing audio (Audacity [24] e Soundation [25]) promuovendo il lavoro collaborativo e cooperativo e l'utilizzo consapevole di applicazioni digitali musicali. In una seconda fase del progetto sono stati manipolati e creati materiali sonori da utilizzare per il progetto compositivo finale. "Make-up" è stato introdotto come strumento capace di stimolare un confronto fra composizione al computer e composizione attraverso il coinvolgimento totale del proprio corpo in un'esperienza fisica-sensoriale globale. Infatti attraverso l'esperienza di gioco di "Make-up" la performance fisica diventa una solo cosa con il lavoro intellettivo complesso, supportando lo sviluppo di abilità cognitive come l'attenzione, la memoria di lavoro, l'elaborazione, la prontezza di riflessi e la capacità di pianificazione.

4. RISULTATI

Nel corso delle attività di sperimentazione sull'utilizzo del sistema MOVE sono stati raccolte informazioni attraverso osservazioni dirette sul comportamento di 72 alunni totali. Le classi coinvolte sono state: 1A (18 alunni, 9 maschi e 9 femmine); 1B (18 alunni, 11 maschi e 7 femmine di cui 2 con bisogni educativi speciali); 2A (16 alunni, 12 maschi

[21] https://freesound.org/
[22] https://123apps.com/it/
[23] https://www.arduino.cc/
[24] https://www.audacityteam.org/
[25] https://soundation.com/

Temi emersi	Light Up	Sweep Up	Pick Up	Make Up	Totale
Coinvolgimento	4	4	4	4	16
Relazione con i pari in contesto educativo	4	4	4	4	16
Inclusione	4	3	4	4	15
Confidenza con l'ambiente	4	3	3	3	13
Sviluppo della memoria spaziale, visiva e uditiva	2	3	3	3	11
Orientamento nello spazio e coordinazione dei movimenti	3	3	3	2	11
Comprensione e rispetto delle regole	3	4	4	0	11
Benessere	3	2	4	2	11
Relazione con i tutor in contesto educativo	3	4	3	0	10
Creatività	1	0	0	3	4

Table 1. Punteggi attribuiti ai temi emersi dall'osservazione presentati in ordine di rilevanza

e 4 femmine di cui 3 con bisogni educativi speciali); 2B (20 alunni, 14 maschi e 6 femmine di cui due con bisogni educativi speciali). Ogni autore ha seguito un progetto ed una classe. [26]

Nel caso specifico sono stati annotati gli atteggiamenti emersi sia durante le lezioni sia a posteriori. Le osservazioni sono state raccolte attraverso relazioni, diari e liste di controllo. Particolare attenzione è stata rivolta alle dinamiche osservate con l'ausilio di videoriprese e fotografie.

Il materiale raccolto è stato poi sottoposto ad un'analisi tematica, in cui i dati sono stati esaminati per identificare temi comuni e ricorrenti [22].

4.1 Analisi Tematica

Dall'osservazione sono emersi diversi temi, rappresentati in Tabella 1 in ordine di rilevanza. Ogni autore ha valutato i temi riscontrati, relativi al gioco sperimentato, esprimendo un punteggio da un minimo di 0 a un massimo di 4, al fine di evidenziare le potenzialità di MOVE.

- **Coinvolgimento**. In tutti i giochi la componente ludica di MOVE ha motivato gli alunni a partecipare attivamente, favorendo maggiore attenzione e concentrazione rispetto alla lezione frontale. In modo particolare gli elementi giocosi di "Sweep-up" hanno incentivato gli alunni a sfidarsi nel completare più velocemente il gioco.

- **Relazione tra pari in contesto educativo**. Anche in un contesto di gioco individuale il sistema ha promosso le interazioni sociali favorendo la cooperazione tra pari per il raggiungimento dell'obiettivo. In "Light-up" gli alunni che non partecipavano in prima persona aiutavano il giocatore a trovare gli elementi del quadro non ancora individuati.

- **Inclusione**. Il sistema ha consentito a tutti gli alunni, compresi quelli con bisogni educativi speciali (disturbo dello spettro autistico, difficoltà linguistiche, ritardo cognitivo) di partecipare all'esperienza di gioco in modo efficace. Di particolare interesse è stato il comportamento osservato in un alunno, affetto da disturbo dello spettro autistico, che senza bisogno di incoraggiamento ha partecipato al gioco "Pick-up" dimostrando un particolare benessere derivato dall'attività ludica. Sicuramente non tutte le attività sono inclusive per ogni disabilità. Nel caso di quelle sensoriali e motorie i giochi possono essere fruiti parzialmente, adottando delle strategie compensative da parte del docente.

- **Confidenza con l'ambiente**. Per motivi di praticità e di economicità è stata scelta la proiezione frontale rispetto alla proiezione a pavimento. Nonostante ciò non sono stati notati particolari disagi o mancanza di concentrazione da parte degli alunni coinvolti. La confidenza con l'ambiente si è sviluppata attraverso la pratica, favorendo così una maggiore consapevolezza dell'esperienza di gioco. Probabilmente un'attività più prolungata nel tempo avrebbe portato una crescente familiarità con l'ambiente interattivo.

- **Sviluppo della memoria spaziale, visiva e uditiva**. Come si evince dai punteggi riportati in tabella, la maggiore criticità nello sviluppo della memoria riguarda il gioco "Light-up". Probabilmente questo dato deriva dalla complessità e dall'elevato numero di elementi sonori da ricercare nel quadro.

- **Orientamento nello spazio e coordinazione dei movimenti**. L'orientamento nello spazio dipende dalle capacità degli alunni di sfruttare l'output visivo e/o uditivo per coordinare il movimento e raggiungere gli obiettivi di gioco. Nel caso di "Sweep-up" il giocatore dispone esclusivamente di un feedback sonoro, ciò implica dunque maggiori abilità di orientamento spaziale in funzione dell'ascolto. In "Make-up" il punteggio rilevato risulta basso; questo si spiega con la scelta degli alunni di non seguire le regole di gioco.

- **Comprensione e rispetto delle regole**. Ogni gioco presenta una schermata iniziale dove vengono esplicitate le regole. Anche se queste appaiono chiare, e non hanno bisogno di ulteriori spiegazioni da parte del docente, alcuni alunni hanno scelto di utilizzare il gioco in modo libero, esplorativo, prefiggendosi obiettivi diversi da quelli del gioco stesso. In parti-

[26] 1A: Light Up - Ottavia Marini; 1B: Pick Up - Maurizio Botti; 2A: Make Up - Maura Castelnovo; 2B: Sweep Up - Sara Valente. Ogni progetto è stato supervisionato da Marcella Mandanici.

colare in "Make-up" le regole del gioco, pur essendo state comprese dagli alunni, sono state ignorate in funzione del raggiungimento di un obiettivo ludico diverso: azionare tutte le caselle nel minor tempo possibile.

- **Benessere**. La maggior parte degli alunni ha interagito in prima persona con il sistema dimostrando particolare entusiasmo. Contrariamente, alcuni alunni hanno dimostrato timidezza probabilmente per timore di sbagliare o di essere giudicati.

- **Relazione con i tutor in contesto educativo** L'ambiente ha incentivato un momento di condivisione meno formale rispetto alla lezione in classe, portando ad una nuova dimensione la relazione tra studente e tutor. Soltanto nel caso del gioco "Make-up" il valore di questa osservazione risulta pari a zero: gli alunni, infatti, dopo aver ricevuto le informazioni riguardanti le regole dell'attività, non si sono avvalsi del supporto del tutor.

- **Creatività**. E' la componente principale di "Make-up", l'unico gioco che permette di creare effettivamente una propria versione combinando il materiale sonoro proposto. Gli altri giochi hanno finalità diverse che non contemplano la creazione di materiali nuovi, se non in attività appositamente ideate dal docente.

5. CONCLUSIONI E LAVORI FUTURI

Questo articolo presenta alcuni progetti educativi basati sull'utilizzo di MOVE, ambiente multimodale interattivo legato al movimento del corpo. A seguito delle osservazioni effettuate durante la sperimentazione si possono trarre alcune conclusioni. La tecnologia, già ampiamente utilizzata nel contesto scolastico in cui si è operato, ha permesso di integrare le classiche strategie di insegnamento con attività ludiche legate al movimento corporeo. I giochi hanno messo alla prova gli alunni in attività interessanti da svolgere autonomamente o in gruppo. A tal proposito MOVE rappresenta un prezioso strumento per il docente e, inserito in una progettazione didattica, permette di sviluppare non solo le competenze specifiche disciplinari ma anche quelle trasversali. I percorsi didattici proposti hanno favorito la collaborazione, la comunicazione, lo sviluppo del pensiero critico e la creatività, competenze indispensabili per il cittadino del ventunesimo secolo [23].

La sperimentazione di MOVE a scuola ha evidenziato alcuni possibili sviluppi futuri volti ad aumentarne le potenzialità. Per alcuni giochi ad esempio potrebbero essere implementate nuove funzionalità. In "Ligh-up" sarebbe interessante introdurre altri quadri da esplorare o, in alternativa, prevedere una modalità che permetta al giocatore di ri-sonorizzare il quadro scelto; in "Sweep-up" si potrebbero introdurre altri parametri del suono verso cui sensibilizzare l'attenzione degli alunni, incrementando contemporaneamente elementi tipici dei giochi come classifiche, livelli e partite temporizzate.

Inoltre i temi emersi dall'analisi osservativa possono essere utilizzati per indagare ulteriormente l'effettiva valenza formativa derivante dall'interazione con MOVE o altri sistemi simili. Queste ricerche, d'altra parte, richiedono un approccio specifico che è stato impossibile attuare nelle prime sperimentazioni descritte in questo lavoro. Per lo stesso motivo, in queste occasioni, non è stato nemmeno possibile raccogliere le impressioni degli alunni partecipanti attraverso questionari strutturati. Le motivazioni e le strategie per l'inserimento di tali esperienze nei contesti educativi necessitano infatti di essere supportati da dati sperimentali solidi e oggettivi, tali da giustificare i profondi cambiamenti nella pratica didattica che l'uso di queste tecnologie comporta.

6. RINGRAZIAMENTI

Si ringraziano il Dirigente scolastico Dott.ssa Paola Bellini, dell' "Istituto Comprensivo Ugo Foscolo" (CR), la Coordinatrice di plesso Prof.ssa Simona Pilotti e il corpo docente per la loro attiva collaborazione nelle attività sperimentali. Un particolare ringraziamento agli alunni partecipanti.

7. REFERENCES

[1] A. Jaimes and N. Sebe, "Multimodal human–computer interaction: A survey," *Computer vision and image understanding*, vol. 108, no. 1-2, pp. 116–134, 2007.

[2] A. Paivio, *Mental representations: A dual coding approach*. Oxford University Press, 1990.

[3] R. Moreno and R. Mayer, "Interactive multimodal learning environments," *Educational psychology review*, vol. 19, no. 3, pp. 309–326, 2007.

[4] F. Karray, M. Alemzadeh, J. Abou Saleh, and M. N. Arab, "Human-computer interaction: Overview on state of the art," *International journal on smart sensing and intelligent systems*, vol. 1, no. 1, 2017.

[5] M. Mandanici, F. Altieri, A. Rodà, and S. Canazza, "Inclusive sound and music serious games in a large-scale responsive environment," *British Journal of Educational Technology*, vol. 49, no. 4, pp. 620–635, 2018.

[6] M. W. Krueger, "Responsive environments," in *Proceedings of the June 13-16, 1977, national computer conference*, pp. 423–433, 1977.

[7] D. Birchfield, T. Ciufo, and G. Minyard, "Smallab: a mediated platform for education," in *ACM SIGGRAPH 2006 Educators program*, pp. 33–es, 2006.

[8] K. Grønbæk, O. S. Iversen, K. J. Kortbek, K. R. Nielsen, and L. Aagaard, "Igamefloor: a platform for co-located collaborative games," in *Proceedings of the international conference on Advances in computer entertainment technology*, pp. 64–71, 2007.

[9] A. Camurri, G. Volpe, S. Canazza, C. Canepa, A. Roda, S. Zanolla, and G. L. Foresti, "The 'stanza logo–motoria': an interactive environment for learning and communication," in *Proceedings of SMC Conference*, vol. 51, n. 8, 2010.

[10] M. Mandanici, A. Rodà, and S. Canazza, "The harmonic walk: An interactive physical environment to learn tonal melody accompaniment," *Advances in Multimedia*, vol. 2016, 2016.

[11] M. Mandanici, A. Rodà, and M. Ricca, "The task of walking straight as an interactive serious game for blind children," *EAI Endorsed Transactions on Game-Based Learning*, vol. 5, no. 16, p. e4, 2018.

[12] J.-M. Pelletier, "Sound and sight: Composition for image analysis systems," *Master's Thesis, Institute of Advanced Media Arts & Sciences*, 2004.

[13] V. V. Kandinskij, "Giallo, rosso, blu", olio su tela 127x200cm, *"Centre national d'art et de culture Georges Pompidou, Parigi*," 1925.

[14] V. V. Kandinskij and E. Pontiggia, *Lo spirituale nell'arte*. Se, 1989.

[15] F. Pachet, O. Delerue, and P. Hanappe, "Dynamic audio mixing.," in *ICMC*, 2000.

[16] T. Riley', "In c," 1964.

[17] M. Mandanici, "Sounding spaces for ubiquitous music creation," in *Procedings of the 14th Proc. Symposium on Computer Music Multidisciplinary Research*, pp. 675–684, 2019.

[18] M. Valle, *La pedagogia Montessori e le nuove tecnologie: Un'integrazione possibile?* Il leone verde, 2018.

[19] L. S. Theremin and O. Petrishev, "The design of a musical instrument based on cathode relays," *Leonardo Music Journal*, vol. 6, no. 1, pp. 49–50, 1996.

[20] D. Lee, "Hornbostel-sachs classification of musical instruments," *Knowledge Organization*, vol. 47, no. 1, pp. 72–91, 2019.

[21] P. Tagg, "Analysing popular music: theory, method, and practice," *Reading pop: Approaches to textual analysis in popular music*, vol. 1982, p. 71, 2000.

[22] M. Susanna, "La ricerca sul campo in educazione. i metodi qualitativi," *MONDADORI, MILANO*, 1998.

[23] P. Mishra and K. Kereluik, "What 21st century learning? a review and a synthesis," in *Society for Information Technology & Teacher Education International Conference*, pp. 3301–3312, Association for the Advancement of Computing in Education (AACE), 2011.

MUSIC FOR ALL: FINDING DIGITAL RESOURCES

Davys Moreno
CIDTFF – University of Aveiro
davys.moreno@ua.pt

Ana Maia
CRTIC- Aveiro
crticeeeixo@gmail.com

Paulo Jacob
APC - Coimbra
pjrjacob@gmail.com

ABSTRACT

The present study aims to find technological solutions to facilitate the inclusion of children with Special Needs in learning music. To achieve this end, a systematic literature review was conducted, between the years 2017 and 2022, with the keyword "accessible digital musical instruments" in different search engines and databases. In addition, a questionnaire survey was carried out to the musicians integrating the musical groups *5ª Punkada* and *Ligados às Máquinas* belonging to APC-Coimbra in Portugal. The results of the questionnaires provide us with information about the difficulties, the life stories, the dreams, the supports, and the technologies that the elements of the two groups use. The results obtained in the literature review refer to the presentation of a set of ADMI, adapted musical instruments, software, etc., that can be used by different users with Special Needs. In conclusion, it is referred the actuality of this theme, the expensive cost and lack of access to accessible digital musical instruments, the need for further studies, the importance of the training of those involved in the use of these resources. We believe that it is possible, with the participation of the whole community, to continue promoting "music for all".

1. INTRODUCTION

Portugal with the implementation of the Decree-Laws No. 54/2018 and No. 55/2018, both of 6 July [11,12], hopes to contribute to expanding, improving, and cementing an inclusive school, which considers the characteristics, abilities and needs of children and the environments where they interact [11,12,26]. With this new legislation, the care previously focused on the student, becomes focused on the school organisation. The importance of categorising difficulties is no longer given, and a restricted conception of support measures only for students with "Special Educational Needs" is abandoned.

Copyright: © 2022 Moreno et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

It is assumed that any child can have Special Needs (SN) and need measures to support learning, throughout their school career, according to the uniqueness of each one. The approach becomes multi-level centred on the organisation of an integrated set of support measures: i) universal measures; ii) selective measures; iii) additional measures. The mobilisation of the different measures is made according to the specific needs of each child. Furthermore, the importance of mobilising other community resources and the involvement of parents in the educational process is reinforced. Therefore, the importance of categorising difficulties is no longer given, so that the needs of educational measures are given to any child with SN [28,29,30,31].

The Arts Education Programmes (AEP) courses in Portugal [10], aim at developing the aptitudes of students with artistic vocation, in the fields of visual arts, dance and/or music, with the purpose of being able to exercise a profession in an artistic area in the future. These courses, in this context, and despite educational inclusion being considered fundamental, as advocated by the Basic Law of the Educational System [26], and current legislation [10,11,12], are not yet prepared to welcome some children with SN, namely with motor limitations.

An example of this lack of preparation happened in the year 2018, when the mother of a child with motor limitations arising from Cerebral Palsy (CP) wanted to enrol her 7-year-old son in an Art School in the First Year of the First Cycle of the AEP of Music. Unfortunately, the guardian of this child ended up giving up the enrolment because he considered that the Art School did not meet the necessary conditions for his child to attend this type of education [28,29].

The child in question has a diagnosis of Dysgraphic CP, with a Severe Dyskinetic form in the upper limbs and moves with a wheelchair. He has a good cognitive development. He uses assistive devices for communication and learning: laptop, GRID 3 software, and the PC Eye Mini gaze access tool [28,29].

This situation triggered an investigation, with the purpose of promoting the inclusion of children with SN in the AEP of Music using digital technologies. For this reason, one of the Centres for Information and Communication Technology and Resources (CRTIC) was contacted, to request collaboration to find technological solutions to promote the inclusion of this child in the AEP of Music. It was from this need, in collaboration

with the CRTIC, that the need to carry out the present study arose.
The 25 existing CRTIC, constitute a network of services of the Portuguese Ministry of Education, whose main aim is to assess the needs of children and young people in terms of Technology/Support Products, to facilitate their access to the curriculum and their participation in school. They are constituted as a prescribing entity of Support Products for Learning [13], and with the application of the new legislation [11,12], they started to be defined as part of the Specific Organisational Resources in favour of the Inclusive School [11,12].
The present work arises both from the need to understand what has been done in Portugal to respond to people with CP to learn music, as well as update the information compiled from the different literature reviews already conducted on the different existing Accessible Digital Instruments (our authors).
In this context, we will first talk about the Associations of CP, institutions that work with people with CP, some of the work done by them, which includes examples of success, to then define the purpose of the study and the Methodological Processes used in this research.

2. ASSOCIATIONS OF CP

The Association of Cerebral Palsy (APC) originated in 1960 in the city of Lisbon, Portugal, and were founded by a group of people with disabilities, parents, friends, and technicians who sought to defend the right to rehabilitation of children, young people, and adults with CP through the maximum development of their abilities and the right to essential technical aids to improve their functionality and their social, family, school and professional inclusion (https://www.fappc.pt/). In addition, PCAs seek to: (i) raise public awareness of the problem of CP, its prevention and the need for early detection and intervention; (ii) defend the right of the Person with disabilities to education, vocational training and work, as well as to their affective fulfilment and social inclusion - with participation in equal opportunities; (iii) develop recreational, sports, cultural and leisure activities; (iv) fostering the creation of Support Units for the severely disabled and dependent; (v) developing training actions for technicians, parents and carers; (vi) fostering research and specialisation in the area of CP and; (vii) promoting its membership in similar national and foreign associations, and fostering the exchange of experiences at national and international level. Over the years, other Associations appeared and developed their action all over the country through the Regional Centres of the South (Lisbon), North (Oporto), Centre (Coimbra), Viseu, Faro, Vila Real, Braga, Évora, Guimarães, Madeira, Alentejo Litoral (Odemira), Leiria, Viana do Castelo, Castelo Branco and Beja. As an example of Good Practices, we will situate it in the development of the work that APC Coimbra (http://www.apc-coimbra.org.pt/) carries out on behalf of people with CP.

APC Coimbra was established as a non-profit organisation in 1975 and its mission is to promote the social inclusion of disadvantaged people, especially those with disabilities. The range of work of the organization involves people with CP, neurological problems, and various other disabilities. In this context we find that this association has successfully developed different works in music, such as the musical groups *5ª Punkada* and *Ligados às Máquinas*.

The group *5ª Punkada*, is a group made up of 5 musicians belonging to APC Coimbra that has already performed over 300 concerts. In December 2021 they released their first album entitled "Are we Punks or Not?", which featured both Surma and Victor Torpedo, with production by Rui Gaspar (https://www.youtube.com/watch?v=Wk6eSDY_neI).
5ª Punkada is internationally recognized for its originality and musical energy, with which they intend to "... get everyone dancing and start a revolution!" - characteristic expression of this band that accompanies them to all the locations where they perform.

The musical group *Ligados às Máquinas* (https://www.youtube.com/watch?v=3tkpK_AOe3A) is formed by 9 persons of APC Coimbra. This group conjugates resources that Digital Technologies (hardware and software) provide to make music. The work developed for the formation of this musical group can be characterised in 5 phases: (i) phase of promotion of autonomy, process of customization of switches, initiation and adaptation to the musical software to be used; (ii) phase of development of contents, process of selection of sounds and subsequent composition; (iii) phase of customization of contents, in which the intervenient are challenged to conceive a personal list of sounds and musical ideas to develop and integrate in the own musical program; (iv) phase of experimentation and development of the musical software to be used in a personalized way; (v) phase of integration of the intervenient already in a group situation, using other instruments and/or personalized software).

Next, we will define the Purpose of the Study and the Methodological Processes used in this research.

3. OBJECTIVE

The purpose of this study is to provide a set of useful information for all institutions and schools, as well as for individual music education professionals, special education teachers and especially for AEP of Music teachers, who wish to complement their tools and expand their possibilities regarding accessibility issues in music education. We also wanted to know in practice what has been the experience of people with disabilities, who already participate as musicians, in musical groups and what have been the digital technologies, support products and ADMI generally used by these musicians.

Considering our purpose, we thought of questions to guide our search for information (Table 1). In this sense, we specified the following objectives that we made correspond to the questions to be asked:

Table 1. Objectives and questions to be asked.

Objective	Questions
To know who he is, what he does, what his professional experience is and, how he started to develop in music	Tell us a little about yourself: your life, who you are, your professional experience and who introduced you to music?
To know his perception of what can facilitate or hinder his musical artistic development	What have been your greatest difficulties in achieving your goals in music?
Knowing if you use or have felt the need to use digital technologies or support products in your profession and in music	Have you used or felt the need to use digital technologies or support products in the development of your profession? And in music?
Knowing how you can promote the activities you do so that they can be replicated by others with SN	Are the artistic-musical activities you carry out accessible to other people with Special Needs? Please specify.
Whether you use ADMI for your musical artistic activities and which ones	Have you ever felt the need to use ADMI? Which ones? In which contexts?
Open question for any issues that interviewees may consider relevant and that were not covered by the questions above	Are there other aspects that you consider important to mention?

For this purpose, we conducted 12 interview surveys to musicians with SN participating in successful Portuguese musical groups such as *5ª Punkada* and *Ligados às Máquinas.*

In view of our objectives, we conducted: i) an update of the literature reviews already conducted by two of the authors of this article; ii) an analysis of the results obtained from the answers to a questionnaire with a set of open questions, related to accessibility and the use of digital technologies, which were conducted to the members of the musical groups *Ligados às Máquinas* and *5ª Punkada*, from APC Coimbra.

Next, we will explain the methodological procedures considered for the development of our research, namely the Methodology and the process of application and data analysis.

4. METHODOLOGY

This exploratory study, conducted between January and September 2022, has an interpretative nature with qualitative designation. It was built in a descriptive logical way [1,4], with the purpose of obtaining practical knowledge about the existing and available technological resources in music that can be used by users with SN and those used by the groups under study.

Considering the questions listed (Table 1), we developed an interview script that was validated by professional experts in the areas of technology, music education, school inclusion and health. Using these scripts, in January, February and March 2022, we interviewed twelve members of the musical groups *5ª Punkada* and *Ligados às Máquinas*. The interviews were later transcribed in Word and their transcription validated by the respective interviewees. With the information in Word, we started the search for inferences, through the reading of the answers obtained, in an inductive way.

The answers to each question were read, looking for patterns and comparing data. As the interview script was made up of open questions, the information gathered in each question was not only directed to that question in question but considered as a whole. In this way, categorisation was carried out by text segments in an inductive manner. This led us to a thematic organisation according to the information compiled and not according to the order of the questions. Then, an analysis of the frequencies of the categories related to each theme was performed. From the frequency of the categories, we moved on to "Questioning", making comparisons between the answers obtained by the musicians in each of these groupings.

On the other hand, we observed that the systematic literature reviews [30,31], already carried out recently by two of the authors of this article, were already outdated, probably due to the technological and digital development. We therefore decided to update the information compiled. Due to this, we made a new literature review, between the months of March and May 2022, by means of the Keyword ADMI - Accessible Digital Musical Instruments. The concept of ADMI, arises from the development of Digital Musical Instruments (DMI - Digital Musical Instrument), focused on the musical domain and performance, designed for people with disabilities (Accessible), differentiating it from a solely therapeutic perspective, as is the case with Music Therapy [5,6,7,8,9,16,17,18]. In this review, we searched for papers published between the years 2017 and 2022. The decision to choose this range of years was motivated by the search for updated information, since in Accessibilities and Technologies, these quickly become obsolete. The Search Engines Google Scholar (88 papers found) and b-on (15 papers found) and the Databases Scopus (12 papers found), ERIC (19 papers found) and Web of Science (9 papers found) were used. As exclusion crite-

ria, Books, theses and all repeated communications between Search Engines and Databases were not considered. We selected only the papers related to the theme of our study, i.e., all articles referring to ADMI, obtaining a total of 29 papers. In each one of them, we selected the name, the mode of access, and a brief description of each ADMI.

Once the objectives and the methodology used were defined, we will now present the results and their discussion.

5. DISCUSSION OF RESULT

Regarding the literature review, all the authors of the selected papers, talk about the need to work as a team to be able to design and manufacture adaptable, accessible, and customised music technologies, providing technical support and ongoing maintenance [2,3,5,6,7,8,9,14,16, 17,18,19,20,21,22,23,24,25,27,32,33,34,35,36,37,38,39, 40,41]. They also mention, the importance of sharing resources, knowledge, and skills, with other researchers and/or interdisciplinary development teams, encompassing user participation, for the sake of science [5,6,7,8,9,14,16,17,18,27,38,39]. Some of these authors, consider that this teamwork, should be complemented with the need to develop a clear policy agenda, which should include social, ethical, and cultural considerations, to disseminate studies and technological trends outside communities, as for example are the New Interfaces for Musical Expression - NIME [16,17,18,24,27,34,35,36]. In addition, references were found to ADMI prototypes designed to work with people with NE, developed by interdisciplinary teams, with user participation and iterative prototyping [3,5,6,7,8,9,14,15,19,20,23,37,41]. Ten different types of interfaces were identified through key concepts suggested by Emma Frid [16,17,18]: concrete controllers, non-contact controllers, Musical Brain-Computer Interfaces (BCMIs), adapted instruments, wearable controllers or prosthetic devices, mouth controllers, audio controllers, gaze controllers, touchscreen controllers and mouse-controlled interfaces [16,17,18]. Some of these adapted instruments can serve as facilitators of interaction in different contexts with children, youth and/or adults with neurological limitations [36,38,39,40,41], others are musical interfaces that can aid, for example, interaction with users with Autism Spectrum Disorders [16,17,18,37,38,39], others constitute themselves as musical technologies for children, young people and adults with hearing impairment [3,16,17,18] or blindness and/or visual impairment [16,17,18,34,34,35]. Furthermore, a wide variety of interfaces can be successfully created that collaborate with the development of personalised, tailor-made Instruments, such as some accessible Instruments, technologically designed to be like traditional Instruments [16,17,18,19,20,21]. These Instruments can empower the participation of their users with NE, in ensembles and musical groups, promoting true inclusion [21,33,38,39]. "Accessible digital musical instruments (ADMI) can provide access to active music-making to a broad range of people, including those who may experience different access barriers due to impairments and exclusive social practices" (Forster 2022, Page, 1).

For Harrison & McPherson [20] accessible instruments refer to two categories of instruments designed for people with disabilities: (i) therapeutic devices - which have few barriers, allowing non-musicians or those with severe limitations, to produce expressive music without the need for practice and (ii) performance-centred instruments - instruments that may require a lot of training and practice, but allow the user a similar level of expression and virtuosity as with a traditional instrument.

ADMI most often fall into the second category (performance-centred instruments) [5]. but can also be used from a therapeutic or experimental perspective [16,17,18]. For people with differentiated access needs to musical instruments, their relationship with music is often only positioned in terms of treatment or therapy [16,23]. In contrast, ADMI uses go further, enabling engagement with music in enriching ways and promoting performative, artistic and inclusive musical practice.

The field of ADMI, dedicated to instruments designed for people with disabilities, is growing rapidly, with the recognition that adaptations of existing Digital Musical Instruments, can promote inclusive music making [5,6,7,16,17,18]. ADMI offer opportunities for engagement with a wider range of sounds than those that acoustic instruments can provide. Some ADMI use eye-tracking via head movements [5,6,7,8,9], freeing the user from reliance on interfaces such as keyboard and mouse to interact with sounds. Furthermore, some authors have implemented a simple study method, based on a set of exercises, to aid the use of ADMI [5,6], thus providing greater opportunities for exploration, improvisation, empowerment, and fluidity in music production, for users with or without SN [5,6,7,8,9,16,17,18,23].

It will be the promotion of developed music technologies, including musical instruments created by both the NIME community and others, which can collaborate positively with the "democratisation of music" [7,9,32].

In the future, methods related to virtual reality and machine learning seem to be promising to bring important approaches in this area [33,38,39].

Next, we will present a list of instruments found, which can be used to facilitate access to users with some special needs. Afterwards we will move on to analyse the results obtained from the interviews conducted with the musicians who are members of the ensembles *Ligados às Máquinas* and *5ª Punkada* which belong to APC Coimbra in Portugal.

Based on the information collected in the selected articles, we identified a set of ADMI that can support the teaching of music to children and young people with NE, which on September 14, 2022, were active, which we will present below. We identified for each of them, the forms of access that allow their use, namely access by hands, fingers, gaze, movement, body/touch,

shake/turn, handles and joystick. After that, we will present the analysis of results obtained by interviews with musicians from the bands *Ligados às Máquinas* and *5ª Punkada*.

L'ORGUE SENSORIEL
(https://www.orguesensoriel.com/).
This instrument makes it possible to adapt to the gestural capacities of disabled people and their ability to translate their gestures into music. The instrument captures any form of gesture, sign of expression, finger flexion, chin rest, rotation of the or breathing, to collect the expressive possibilities of each person and translate them into music. In this way even people with the greatest difficulties can make music in an active way.

SOUNDBEAM (https://www.soundbeam.co.uk/).
This is a musical instrument, which focuses exclusively on the dimension of touchless playing. Designed by Edward Williams to produce avant-garde dance music, the Soundbeam was quickly used for people with disabilities due to the limited physical movement that is required for it to function properly. It is a MIDI hardware and software that uses ultrasonic sensors, detects movement, and transforms it into sound. It contains a series of pre-installed music tracks and can be used by moving.

SOUNDBOX AND MINIBOX
(https://www.soundbeam.co.uk/vibroacoustics).
These are two vibro-acoustic boxes that allow the user to experience sounds and music with their body. It was designed so that deaf people could make music using Soundbeam. It can be used by using the body and/or movement.

ARCANA INSTRUMENT
(https://arcanainstruments.com/).
Inspired by video game culture and controller design. Arcana is an instrument adaptable to various forms of access (switches, joystick, etc.), very easy to use and accessible to all. It is designed to be an easy-to-play instrument. It allows the user to play the "guitar" in a simple and ergonomic way. It offers a variety of musical expression methods, uses various melody scales and harmonic scales, and includes a wide range of notes. To imitate guitar playing, its mechanisms allow movements and sensations of fingering with one hand, while, on the other hand, pressing the keys, as if they were strings, to produce the sound of the strings. This instrument allows the user to play together, with others or individually using the hands, the fingers, the knobs and/or the joystick.

BRUTBOX
(https://reso-nance.org/wiki/projets/brutbox/accueil)
This is a collective musical practice device designed to overcome the possible difficulties that people with special needs have. Based on the free software Pure Data (https://puredata.info/), it allows all users to perform a wide variety of musical gestures, regardless of the potential or motor or learning difficulties the user may have, they can always find an active place in the collective musical game. BRUTBOX has a series of intuitive interfaces that compensate for the possible motor and intellectual difficulties of users by promoting the development of everyone's strengths.

MAKEY MAKEY (https://makeymakey.com/).
This is an electronic card that does not require computer programming and allows endless scenarios to be invented to trigger sounds or play video games. It is an ideal tool for working with people with disability because any user becomes the creator of his or her own device, which can be modified according to needs, disabilities and motor and intellectual abilities. Makey Makey works via USB, so no additional software is required. It can be connected to a PC or tablet, allowing programmed circuits, plus it allows objects to be turned into switches. The kits run on Windows, Mac OSX and Linux and no programming knowledge is required to use them, moreover, for instrumental execution they can be used from a finger, any body part to hands, fingers, knobs, joysticks, and other objects.

ALPHASPHERE ME
(https://alphasphere.com/alphasphere-me/).
This instrument functions as a MIDI, synthesizer and/or sequencer, connected to the PC. It consists of tactile pads, sensitive to pressure, around a spherical surface. The hands are used for its use.

ARTIPHON INSTRUMEN 1
(https://artiphon.com/pages/instrument1).
This is a MIDI controller that adapts to the way you play. It allows fingering, sliding, playing, and tapping any sound in a single interface. Hands, fingers, shake and/or turn can be used for its use.

BEAMZ SYSTEM (https://thebeamz.com).
This is an interactive system that uses lasers to trigger sounds, sound effects and even songs. This instrument can be used to teach musical concepts and skills. It connects to any device and hands can be used to operate it.

CMG CELL MUSIC GEAR (http://cmg.tokyo/en/).
This musical tool uses a set of 3D sensors. It can detect all kinds of pressure information on a 2D surface, such as a touch pad. It can emit a certain pitch, noise, timbre, etc. It can also use various styles of sound. Hands, feet, fingers, other body parts and even a single finger can be used to perform it if necessary.

JAMBOXX (https://www.jamboxx.com).
This is an electronic instrument that can be used by breathing. This instrument allows one to change the sound of various musical instruments and, also to play notes simultaneously. The user can sonar this instrument by means of blowing.

MIDIWING (http://www.midiwing.com).
This is a MIDI controller that adapts to those who use it, reducing the technical barriers related to playing an instrument. It can be played by more than one person, namely in an orchestra or a band. It is simple to operate and accepts many types of switches and control devices such as a joystick or a mouse, just one way in which the instrument can be customised to any type and any level of ability. The switches and other control devices can be placed anywhere that is convenient for the user. They can be operated by either hand, eye, knob, or joystick.

MOTIONCOMPOSER
(https://motioncomposer.de/en/).
This is a device that analyses movements and gestures and transforms them into computer data. It uses video cameras connected to a PC and algorithmic composition

software, making it possible to associate notes, correcting errors to maintain the rhythm. It can be used by means of movement.
NUMARK ORBIT
(https://www.numark.com/product/orbit).
This is a MIDI controller, wireless, fully adaptable, and easily usable with finger or.
ORBA (https://artiphon.com/pages/orba-by-artiphon).
This is a synthesizer, looper and controller that allows you to create songs using intuitive gestures such as tapping, swiping, and waving. It allows the user to play parts of a song on a musical instrument or to switch this between other musical instruments. The hands, shake or turn can be used for its use.
OSMoSIS (http://sro.sussex.ac.uk/id/eprint/96488/).
This is a system that converts body movements into sounds, using an interface (system based on Microsoft Kinect) that allows full body tracking. It is used through movement.
OTOMATONE (https://otamatone.jp/en/lineup/).
This is an eighth note synthesizer. It works with one hand holding and squeezing the "head", while the other hand controls the pitch of the melody on the stem. It can be connected to iOS applications. Its form d use is the fingers (can be one hand) or the hands.
PHONOTONIC (https://www.f6s.com/phonotonic).
This is a small musical object that, when connected to an application, transforms movement into music. Tempo and melody can be controlled through movement.
SKOOG (https://skoogmusic.com).
This is an easy to use tactile musical instrument with accessibility features. It has the function of a MIDI and can be carried, rotated, shaken, or connected to the iPad or computer. For its use, hands, shaking or turning can be used.
SKWITCH (https://skoogmusic.com/skwitch/).
This is a Gadget & App that connects to the iPhone. It uses sensors, which when pressed can be used as a musical instrument, due to which, it produces musical notes. It can be used with different parts of the body, from a finger to hands or feet.
SOUND CONTROL
(http://soundcontrolsoftware.com).
This is easy to use software and can be used to create other "musical instruments". It uses movement, detected with a variety of ready-made devices, this includes webcam, Leap Motion, micro:bit, mouse, microphone, etc. We can access this instrument using our hands, our eyes or the joystick.
NETYTAR (https://vision.unipv.it/research/netytar/).
This is a virtual digital musical instrument. It can be used by people with motor disability. This instrument is controlled by an eye control device and a switch. The way to access and use it is by looking.
After presenting the list of musical instruments found, we will move on to present the results obtained in the interviews with the members of *Ligados às Máquinas* and *5ª Punkada*. In this context, we observe that the analysis of the totality of the categories formulated made us realise that all the interviewees referred to the same characteristics, whose theme referred to five large groups of information (see fig. 1).

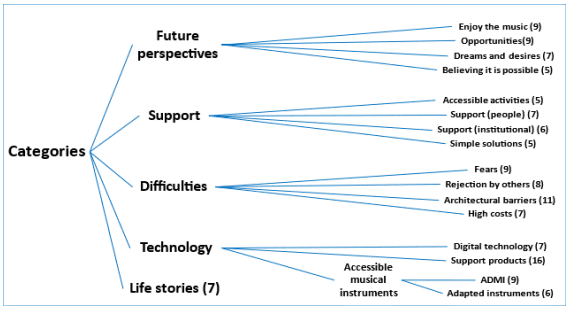


Figure 1. Categories - 5 major groups of information

As can be seen in Fig. 2 the information compiled in these groups of information is related to categories that refer to aspects related to the life experience that respondents have as musicians, the technologies used, their difficulties, their fears, their stories, their support, and their life experiences.
Among the most frequent categories related to the Support that respondents need to have been Accessible Activities (5 categories); Supports (people) (7 categories); Supports (institutional) (6 categories); Simple Solutions (5 categories). Examples: "I realise that the activities we are doing we can only do them because they are accessible to us"; "If it wasn't for Tina and António I wouldn't be able to do what I do, thank you"; "APC Coimbra is an institution that supports us a lot, it does everything to see us well"; "Sometimes it is just a matter of moving a chair, the solution is so simple, but not everyone can see that".
Among the most frequent categories related to the Difficulties that respondents report having, are Fears (9 categories); Architectural Barriers (11 categories); Feeling the Rejection of others (8 categories); High Costs (7 categories). Examples: "I am afraid that this will end, that it is just a dream"; "Every time I go on stage, I am afraid that things might go wrong"; "We have been invited to play in many places, we warn that we are people with mobility problems and many times they forget that we must go to the bathroom and, just imagine, we don't have baths around, accessible or adapted and me, dying to urinate"; "They forget that we can't climb stairs and many times we can't enter the stage of the theatres because these don't have access"; "Whenever we go to concerts in other places we warn that we can't climb stairs, that we are people with mobility problems and when we get to the hotel, it has no access, no ramps, no lift and they booked us the room on the top floor, so people reject us a lot"; "When we go out to eat, I get very sad with people, they reject me, they don't want to sit next to me because I eat weird"; "Everything is so expensive and hard to get, the musical instruments we use are very expensive".
Among the most frequent categories related to Technology and its use, respondents describe the Accessible musical instruments, which are subdivided into ADMI (9 categories) and Adapted instruments (6 categories); we also have the Support Products they use (16 categories) and the Digital Technology at their reach (7categories). Examples: "The adaptation that Professor Paulo Jacob made for me to be able to play the key-

board was perfect, look, I am playing"; "Soundbeam and Garageband are fantastic, with them I can make music at will"; "we make music by means of the computer, it is incredible"; "there are so many switches and cables that I must use, but, it does not matter, these are the products that help me to be part of this musical group"; "The Makey - Makey together with the Soundplant and the articulated arms, the tripods of microphone, etc. are magnificent, we can make music at will, I love it"; "The software used to make our music is incredible"; "The computer is very important for us, we can communicate and even make music with it"; "I love being a professional DJ, I love putting on different songs and if it wasn't for the Digital Technology I use I wouldn't be able to do what I do".

Among the most frequent categories related to Future Prospects, respondents tell us about their Taste for music (9 categories); their Dreams and desires (7 categories); the Opportunities they have access to (9 categories) and in the importance of Believing that it is possible (5 categories). Example: "I like what I do, I really like music"; "everyone likes the songs we play"; "When we are on stage making music, I like to see the audience's happy face enjoying what we do"; "I want to keep making music all my life"; "I wish I could make a living from music"; "We have the opportunity to do something different"; "Thanks to the musical work that Paulo Jacob does with us, we have the opportunity to be somebody, people believe in us, they look at us with other eyes, we have the opportunity to be somebody"; "Yes, it is possible", "We believe that it is possible and we are doing it".

Finally, among the most frequent categories related to Life Stories (7 categories), respondents report us some anecdotal examples of situations experienced by them. For example: "you can't imagine the number of jokes we have lived through, you could write a book"; "I love working with them, in the years of service I have there are many stories to tell"; "I have many stories and it is true that sometimes they only come to mind because we talk".

After the presentation of the information collected, we will move on to the final reflections.

6. FINAL REFLECTIONS

A first interesting aspect to mention is the timeliness of the bibliography, as most of the articles are from 2020 to 2022. Therefore, we can assume that issues related to the inclusion of children and young people with Special Needs in music learning is a very current concern, which leads us to think that we are probably moving towards a more inclusive society in all areas, including music. It is important to note that the key concepts for a successful ADMI are based on adaptability and customisation, interdisciplinary development teams, user participation and iterative prototyping [16, p.15]. It is also important, to work in a community with common goals by sharing resources, knowledge, and skills [27]. As we found in the questionnaire responses from the music groups *5ª Punkada* and *Ligados às Máquinas*, adaptations were customised and were essential to being able to make music.

Unfortunately, access to ADMI is still not for everyone. We dare mention that many of the instruments/devices named in this exploratory study are expensive for most people in Portugal. As musicians from the musical groups interviewed stated, adapted instruments and the necessary technologies are very expensive. In addition, ADMI are not considered in the list of Support Products which can be prescribed by CRTIC. It should be noted that some of the CRTIC have developed, in an experimental way, some low cost "homemade" musical devices. According to the survey carried out by the different CRTIC, there were no assessment requests from these centres regarding the need for accessibility in music. This makes us reflect on whether this situation is due to the inexistence of needs at this level, or whether this need is not even felt, either due to ignorance about the existing accessible instruments/devices, or due to ignorance about the existence of CRTIC.

We believe that with the information compiled, we already have some possible resources to be used with the child with CP who triggered the study, namely all those that are accessible through the eyes, on the computer. However, we consider that making the resources available is not enough, recognizing the need for training that the use of some of them implies, both for the child and for all those involved in his/her learning. Furthermore, it will be important to make it possible for CRTIC to prescribe these resources, providing specific training to their elements to enable their replication with the probable users.

It will be pertinent to carry out more in-depth and wider studies on this theme, to assess and evaluate the usability of each ADMI, relating each one of them to their possibilities of use, whether linked to therapeutic purposes, musical experimentation, and expression and/or music teaching (in music education and/or Arts Education Programmes of Music). Finally, it would be interesting to relate the potentialities of each ADMI with the curricular contents in music learning, which would collaborate with a greater democratization of teaching, thus promoting "music for all". As Samuels & Schroeder, [38] refer, it is not enough to create or have access to ADMI, it is important that they are used in a sustainable way, without the need for constant support from others (whether these are researchers, designers, or teachers), so that their users can progress, acquiring greater skill and technique over time, like any musician. In this field as we can observe in the answers to the questionnaires, there is still a long way to go. The need for support from others is still essential to enable the access to adapted instruments and ADMI.

7. ACKNOWLEDGEMENTS

Work financed by FCT – Foundation for science and Technology, within the scope of the PhD scholarship with reference 2020.07331.BD, and by National Funds through FCT - Foundation for Science and Technology, IP, within the scope of the UIDB/00194/2020 project,

referring to CIDTFF - Research Center in Didactics and Technology in the Training of Trainers.

We are grateful for the collaboration of CRTIC (ICT Resource Centers, Portugal) in the identification of some resources related to the learning of music, which can be used with children / young people with NE.

We are grateful for the collaboration of APC Coimbra and the musicians of *5ª Punkada* and *Ligados às Máquinas* for all the support given to this research work.

8. REFERENCES

[1] Amado, J: Manual de Investigação Qualitativa em Educação 3ª edição. Imprensa da Universidade de Coimbra, Coimbra University Press, 2017.

[2] Bell, A. P: "(dis) Ability and Music Education: Paralympian Patrick Anderson and the Experience of Disability in Music," *Action, Criticism & Theory for Music Education*, 16(3), 2017.

[3] Cavdir, D, and Wang, G: "Felt sound: A shared musical experience for the deaf and hard of hearing." *Proceedings of the Inter-national Conference on New Interfaces for Musical Expression (nime.org)*, pp. 176-181, 2020.

[4] Coutinho, C. P.: Metodologia de investigação em ciências sociais e humanas: Teoria e Prática, Coimbra, Almedina, 2018.

[5] Davanzo, N, and Avanzini, D: "Resin: a Vocal Tract Resonances and Head Based Accessible Digital Musical Instrument". *In Audio Mostly '21 Proceedings. ACM*, New York, pp- 1-4, 2021.

[6] Davanzo, N, and De Filippis, M, and Avanzini, F: "Netychords: An Accessible Digital Musical Instrument for Playing Chords using Gaze and Head Movements." *Proceedings of the 5th International Conference on Computer-Human Interaction Research and Applications (CHIRA 2021)*, pages 202-209, 2021.

[7] Davanzo N, and Avanzini F: "Experimental Evaluation of Three Interaction Channels for Accessible Digital Musical Instruments." *In: Miesenberger K., Manduchi R., Covarrubias Rodriguez M., Peňáz P. (eds) Computers Helping People with Special Needs. ICCHP 2020. Lecture Notes in Computer Science*, vol 12377. Springer, Cham, 2020.

[8] Davanzo, N, and Avanzini, F: "A Method for Learning Netytar: An Accessible Digital Musical Instrument." *International Conference on Computer Supported Education*, Scitepress. pp. 620-628, 2020.

[9] Davanzo, N, and Avanzini, F: "Hands-Free Accessible Digital Musical Instruments: Conceptual Framework, Challenges, and Perspectives." *IEEE Access*, 8, pp. 163975-163995, 2020.

[10] Decreto Lei N.º 344/1990 do Ministério de Educação e Ciência: Diário da República: 1.ª série — N.º 253, 1990.

[11] Decreto Lei N.º 54/2018 do Ministério de Educação e Ciência: Diário da República: 1.ª série — N.º 129, 2018.

[12] Decreto Lei N.º 55/2018 do Ministério de Educação e Ciência. Diário da República: 1.ª série — N.º 129, 2018.

[13] Despacho Nº 5291/2015 do Ministério da Educação e Ciência: Diário da República: 2.ª série - N.º 98, 2015.

[14] Förster, A: "Accessible Digital Musical Instruments in Special Educational Needs Schools – Design Considerations Based on 16 Qualitative Interviews with Music Teachers," *International Journal of Human–Computer Interaction*, 2022.

[15] Förster, A, and Komesker, C, and Schnell, N: "SnoeSky and SonicDive-Design and Evaluation of Two Accessible Digital Musical Instruments for a SEN School," *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 83-88, 2020.

[16] Frid, E, and Lindetorp, H, and Hansen, K. F. and Elblaus, L, and Bresin, R: "Sound Forest: Evaluation of an accessible multisensory music installation," *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1-12, 2019.

[17] Frid, E: "Accessible digital musical instruments—a review of musical interfaces in inclusive music practice," *Multimodal Technologies and Interaction*, 3(3), 57, 2019.

[18] Frid, E: "Correction: Frid, E. Accessible Digital Musical Instruments—A Review of Musical Interfaces in Inclusive Music Practice, Multimodal Technologies and Interaction 2019, Vol. 3, Page 57," *Multimodal Technologies and Interaction*, 4(3), pp. 34, 2020.

[19] Harrison, J, and Chamberlain, A, and McPherson, A. P: "Accessible instruments in the wild: Engaging with a community of learning-disabled musicians," *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1-6, 2019.

[20] Harrison, J, and McPherson, A. P: "Adapting the bass guitar for one-handed playing," *Journal of New Music Research*, 46(3), pp. 270-285, 2017.

[21] Holloway, l, and Goncu, C, and Ilsar, A, and Butler, M, and Marriott, K: "Infosonics: Accessible In-fographics for People who are Blind using Sonification and Voice," *CHI Conference on Human Factors in Computing Systems (CHI '22)*, April 29-May, New Orleans, LA, USA. ACM, New York, pp. 1-13, 2022.

[22] Horton Patric W, and Horton, P: "Tensions in designing tablet-based musical instruments for stu-

dents with disabilities," *International Journal of Child-Computer Interaction*, Volume 33, 2022.

[23] Ilsar, A, and Kenning, G: "Inclusive improvisation through sound and movement mapping: from DMI to ADMI," *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1-8, 2020.

[24] Jack, R. H, and Harrison, J, and McPherson, A. P: "Digital Musical Instruments as Research Products," *Proceedings of the 20th International Conference on New Interfaces for Musical Expression*, 2020.

[25] Kondo, E, and Tabei, K, and Okuno, R, and Akazawa, K: "Case Report: Accessible Digital Musical Instrument Can Be Used for Active Music Therapy in a Person With Severe Dementia and Worsening Behavioral and Psychological Symptoms: A Case Study Over a Year and a Half," *Front. Neurol. 13:831523*, 2022.

[26] Lei de Bases do Sistema Educativo: Lei nº 46/86, Diário da República n.º 237/1986 de 14 de outubro (Iª Série), pp. 3067–3081, 1986.

[27] Lucas, A, and Schroeder, F, and Ortiz, M: "The Longevity of Bespoke, Accessible Music Technology: A Case for Community," *Proceedings of the 20th International Conference on New Interfaces for Musical Expression* , pp. 243-248, 2020.

[28] Moreno, D, and Moreira, A, and Tymoshchuk, O, and Marques, C: "Análise de conteúdo utilizando o webqda: opção metodológica para caracterizar uma criança com paralisia cerebral," *New Trends in Qualitative Research*, 2, pp. 687-702, 2020.

[29] Moreno D, and Moreira A, and Tymoshchuk O, and Marques C: "A Child with Cerebral Palsy in Arts Education Programmes: Building Scaffoldings for Inclusion," *Costa A.P., Reis L.P., Moreira A., Longo L., Bryda G. (eds), Computer Supported Qualitative Research. WCQR 2021. Advances in Intelligent Systems and Computing, vol 1345. Springer, Cham*., pp.172-190, 2021.

[30] Moreno, D, and Maia, A: "Accessible music for everyone: discovering resources." *Perspectives and Trends in Education and Technology*, Springer, Singapore, pp. 883-893, 2022.

[31] Moreno, D, and Moreira, A, and Tymoshchuk, O, and Marques, C: "Studying Inclusion in Music Education - An Integrative Literature Review as a Support in the Choice of Methodology, Using WebQDA," *Costa, A.P., Moreira, A., Sánchez Gómez, M.C., WaMbaleka, S. (eds) Computer Supported Qualitative Re-search. WCQR 2022. Lecture Notes in Networks and Systems, vol 466. Springer, Cham., pp. 175-193*, 2022.

[32] Morreale, F, and Bin, A, and McPherson, A, and Stapleton, P, and Wanderley, M: "A NIME of the Times: Developing an Outward-Looking Political Agenda For This Community," *Proceedings of the International Conference on New Interfaces for Musical Expression*, 2021.

[33] Parke-Wolfe, S. T, and Scurto, H, and Fiebrink, R: "Sound Control: Supporting custom musical interface design for children with disabilities," *Proceedings of the 19th International Conference on New Interfaces for Musical Expression*, 2019.

[34] Payne, W, and Ahmed, F, and Gardell, M, and Luke DuBois, R, and Hurst, A: "SoundCells: Designing a Browser-Based Music Technology for Braille and Print Notation," *19th Web for All Conference (W4A'22), April 25–26, 2022, Lyon, France*, ACM, New York, pp. 1-12, 2022.

[35] Payne, W. C, and Xu, A. Y, and Ahmed, F, and Ye, L, and Hurst, A: "How Blind and Visually Impaired Composers, Producers, and Songwriters Leverage and Adapt Music Technology.," *The 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, pp. 1-12, 2020.

[36] Peñalba, A, and Valles, M. J, and Partesotti, E, and Sevillano, M. Á, and Castañón, R: "Accessibility and participation in the use of an inclusive musical instrument: The case of MotionComposer*," Journal of Music, Technology & Education, 12(1)*, pp. 79-94, 2019.

[37] Ragone, G: "Designing Embodied Musical Interaction for Children with Autism," *The 22nd International ACM SIGAC-CESS Conference on Computers and Accessibility*, pp. 1-4, 2020.

[38] Samuels, K, and Schroeder, F: "Performance without Barriers: Improvising with Inclusive and Accessible Digital Musical Instruments," *Contemporary Music Review*, 38(5), pp. 476-489, 2019.

[39] Samuels, K: "The Drake music project Northern Ireland: Providing access to music technology for individuals with unique abilities," *Social Inclusion*, 7(1), pp. 152-163, 2019.

[40] Wan, L, and Crawford, R, and Jenkins, L: "Digital Listening Tools to Facilitate Children's Self-Regulation of Instrumental Music Practice," *Journal of Research in Music Education*, pp. 1–24, 2022.

[41] Wright, J: "The Appropriation and Utility of Constrained ADMIs," *Proceedings of the 20th International Conference on New Inter-faces for Musical Expression*, pp.564-569, 2020.

HiFiReM: un approccio unificato, web e nativo, per la didattica musicale remota

Matteo Sacchetto[1], **Cristina Rottondi**[1]
Antonio Servetti[2]
[1]DET - [2]DAUIN, Politecnico di Torino
name.surname@polito.it

Louena Shtrepi, Marco Masoero
DENERG, Politecnico di Torino
name.surname@polito.it

Andrea Valle
StudiUm/CIRMA - Università di Torino
name.surname@unito.it

ABSTRACT

Le applicazioni professionali di trasmissione audio a bassa latenza finalizzate a supportare sessioni musicali in rete sono state tradizionalmente sviluppate come software nativi, specifici per ciascun sistema operativo, con conseguenti difficoltà di installazione, configurazione e utilizzo. Ciò ne ha ristretto l'uso ad una nicchia di musicisti esperti di informatica. Questo studio propone una implementazione fruibile via browser, realizzata sia come applicazione nativa sia come applicazione web, rivolta in particolar modo alla didattica musicale, al fine di facilitare lo svolgimento di sessioni musicali in rete con basso ritardo di trasmissione ed elevata qualità audio.

1. INTRODUZIONE

L'interesse per la realizzazione di performance musicali a distanza (Networked Music Performance, in breve NMP), spinto dalle necessità derivate dal confinamento imposto durante la recente emergenza pandemica, ha messo in evidenza l'esistenza di numerose soluzioni software per suonare in rete sviluppate in ambito accademico, ma il cui uso è ancora limitato all'interno di nicchie di utilizzatori altamente specializzati, soprattutto per quanto riguarda le conoscenze tecnico-informatiche.

Per contro, le soluzioni più adottate nella pratica, soprattutto per la didattica musicale remota, si sono avvalse di strumenti per videoconferenza già largamente diffusi e conosciuti (ad es. Skype, Google Meet, Zoom) i quali presentano evidenti problemi di prestazioni e qualità.

Quanto accaduto, preferire soluzioni di minore qualità, ma dall'utilizzo più immediato, ha evidenziato la necessità e l'urgenza di migliorare la facilità d'uso delle soluzioni per NMP già esistenti. In tale ottica, il progetto HiFiReM descritto in questo articolo si pone due obiettivi fondamentali: i) realizzare un'applicazione per la didattica musciale remota utilizzabile via web, senza necessità di installazione e fruibile direttamente via browser; ii) sviluppare una soluzione nativa (basata su una box hardware con l'applicazione integrata) che soddisfi esigenze professionali di alta fedeltà audio, ma sempre facilmente utilizzabile tramite la stessa interfaccia web.

Copyright: ©2022 Matteo Sacchetto et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Dopo una breve panoramica sullo stato dell'arte nella Sezione 2, segue nella Sezione 3 la descrizione dell'architettura della soluzione proposta, mentre i risultati ottenuti durante una fase di validazione preliminare sono riportati in Sezione 4. La Sezione 5 conclude l'articolo.

2. STATO DELL'ARTE

La quasi totalità delle applicazioni per NMP (ad es. SonoBus [1], JackTrip [2], SoundJack [3]) sono applicazioni native, ossia software sviluppati per un dispositivo specifico utilizzando un linguaggio di programmazione anche esso specifico per la piattaforma ove vengono eseguiti (macOS, Windows, Android). Un'applicazione web, invece, è scritta in un linguaggio più "generico" che viene interpretato ed eseguito da un browser, pertanto può essere utilizzata su ogni piattaforma senza necessità di adattamento o installazione.

L'applicazione web paga il vantaggio della semplicità di utilizzo in termini di maggiori limitazioni nell'ottimizzazione delle proprie prestazioni, in particolare per quanto riguarda la riduzione della latenza di acquisizione, elaborazione e trasmissione dei dati audio, perché è forzata ad utilizzare soltanto le funzionalità che il browser le mette a disposizione. Tali limitazioni hanno spinto, fino ad oggi, a preferire la modalità nativa per applicazioni "esigenti" quali quelle di NMP [4].

Tuttavia l'implementazione degli standard Web Audio e Web RTC nei maggiori browser [1] ha permesso di migrare sul web anche applicazioni per l'elaborazione e la trasmissione audio: per primi i programmi di videoconferenza, in seguito anche software per intere digital audio workstation (es. Audiotool [2]).

Allo stato attuale, tuttavia, non sono ancora presenti soluzioni web in grado di trasmettere audio a basso ritardo e senza compressione, dal momento che l'implementazione base di WebRTC non lo permette. Soluzioni web come Google Meet, Jitsi, LiveLab e altre, soffrono di ritardi bocca-orecchio dell'ordine del centinaio di millisecondi e della riduzione di qualità dovuta alla compressione audio, soprattutto nelle alte frequenze [5]. In aggiunta, gli algoritmi di elaborazione audio pensati per le comunicazioni vocali possono introdurre alterazioni del suono a causa dei meccanismi di aggiustamento automatico del guadagno

[1] Il livello di supporto delle funzionalità nei vari browser può essere verificato online: *https://caniuse.com/audio-api*, *https://caniuse.com/rtcpeerconnection*

[2] *https://www.audiotool.com/*

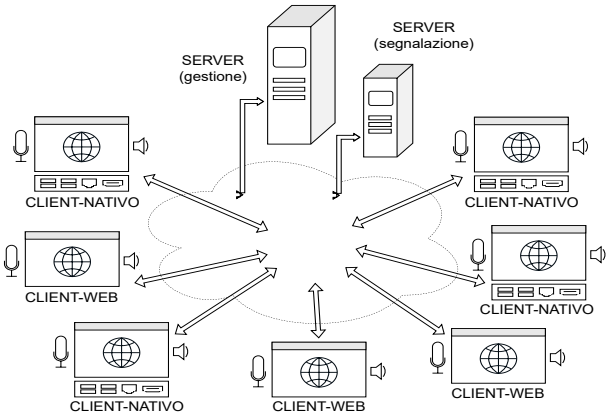


Figura 1. Architettura di HiFiReM con le componenti server e le due tipologie di client (web e nativo) unificate dall'utilizzo della stessa interfaccia nel browser.

che modificano i livelli sonori. Inoltre, gli algoritmi di cancellazione del rumore tendono a smorzare suoni stazionari sostenuti, mentre gli algoritmi di adattamento della velocità di riproduzione alle fluttuazioni di banda del canale possono modificare a tratti altezza e timbro dei suoni.

HiFiReM ambisce a superare le limitazioni sopra citate e costituisce il primo approccio unificato, web e nativo, per la trasmissione di audio stereo ad alta qualità (non compresso) per NMP e, nello specifico, finalizzata a supportare conservatori e scuole di musica nell'erogazione della didattica musicale remota.

3. ARCHITETTURA

Il sistema HiFiReM è strutturato in due componenti fondamentali, una server e una client. La componente *server* comprende due funzionalità: di *segnalazione*, per gestire la comunicazione tra i diversi client; di *segreteria*: per gestire le necessità di organizzazione didattica. La componente *client* è anche essa suddivisa in due parti/implementazioni, una web ed una nativa. Indipendentemente da quale implementazione venga utilizzata, in entrambe l'interazione utente avviene tramite il browser per mezzo della stessa GUI (Graphical User Interface), come illustrato in Fig.1.

3.1 Componente server

La componente server per la gestione della didattica (*segreteria*) permette di integrare nell'interfaccia web un ambiente in cui, a partire dalla definizione dei ruoli di amministratore, segreteria, maestri/docenti e allievi, gli utenti possono essere organizzati in classi. Qui, ciascuna classe può riunirsi in una "stanza" per suonare insieme da remoto. Le stanze rappresentano il concetto di aule didattiche e sono configurate in modo da poter gestire scenari differenti come lezioni individuali, lezioni a gruppi, prove d'insieme (soltanto tra allievi senza il docente) e lezioni con docenti esterni (masterclass).

La componente server per la *segnalazione* si occupa invece di mettere in comunicazione i vari client. Questa componente tiene traccia delle varie stanze attive e si occupa di selezionare il protocollo di segnalazione corretto in base al tipo di stanza, astraendo dunque il fatto che si stiano utilizzando due stack tecnologici completamente diversi tra la componente client–nativo e quella client–web [3] .

3.2 Componente client–web

La componente client–web rappresenta al momento l'unica soluzione web per la trasmissione a bassisimo ritardo con audio originale, senza le alterazioni causate dalla compressione digitale e dall'elaborazione audio intrinseca del browser.

L'implementazione è resa possibile dall'utilizzo innovativo degli standard Web Audio e WebRTC Invece di affidare la gestione del flusso audio ai componenti convenzionali *MediaStream* ed *RTCPeerConnection* si è provveduto a "dirottare" lo stream audio sull'*RTCDataChannel* (pensato convenzionalmente per la trasmissione di dati, non di flussi audio) prelevando i campioni dal MediaStream tramite un *AudioWorklet*. Per i dettagli tecnici si rimanda a [6].

Nell'implementare questa soluzione, che permette di poter controllare a più basso livello il flusso di trasmissione secondo le proprie esigenze, è stato inoltre possibile ridurre l'overhead di gestione tramite l'utilizzo degli *SharedArrayBuffers* così da non introdurre ritardi addizionali e ottenere una latenza molto ridotta [5].

3.3 Componente client–nativo

La componente client–nativo è stata realizzata per gli scenari più esigenti, non limitati all'interazione del docente con il singolo allievo, ma in cui più soggetti suonano contemporaneamente. Questa soluzione, realizzata su *box* Raspberry Pi 4B, è visibile in Fig. 2: l'audio è acquisito, trasmesso e riprodotto direttamente tramite la box, con latenza compatibile con altri ambienti per NMP [7], mentre l'interazione dell'utente (configurazione, connessione, controllo) avviene tramite browser riutilizzando, in modo integrato, la stessa interfaccia dell'implementazione del client web.

[3] Si noti però che, a causa di sostanziali differenze in termini di ritardi, riportati in Sez. 4.2, non è al momento ragionevole connettere a una stessa stanza client web e nativi perché gli utenti web sarebbero temporalmente disallineati rispetto agli utenti nativi

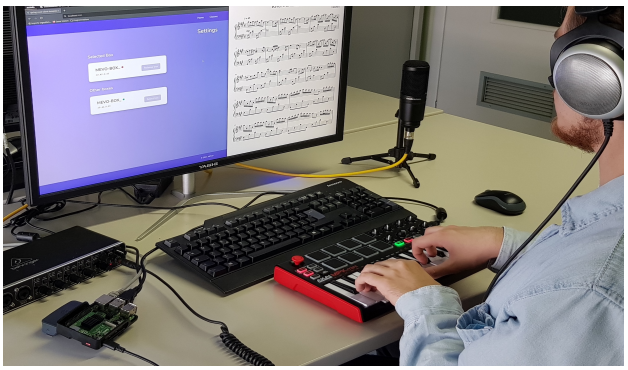


Figura 2. Utilizzo del sistema HiFiReM tramite applicazione web (finestra lilla sul monitor) o box Raspberry Pi 4B (in basso a sinistra) sempre per mezzo dell'interfaccia unificata nel browser.

4. VALIDAZIONE E RISULTATI

Dal punto di vista delle funzionalità del sistema il tratto caratterizzante della soluzione proposta, specifico per la didattica musicale remota, è rappresentato dalla concezione delle *stanze* per l'esecuzione musicale come ambienti didattici. A differenza delle classiche soluzioni per NMP dove i musicisti interagiscono tra di loro da pari, qui viene data al docente (o ad un tecnico audio terzo) la possibilità di controllare le impostazioni dei dispositivi degli allievi e provvedere al mixaggio audio personalizzato per docente e studenti.

Dal punto di vista, invece, delle prestazioni del sistema l'implementazione software è stata confrontata con le prestazioni di altri software per NMP al fine di validare i due fattori di merito principali in questo contesto: a) la qualità del segnale audio ricevuto, b) il ritardo globale bocca-orecchio. Per entrambe queste verifiche, al fine di poter valutare le prestazioni ottime ottenibili dai differenti software al netto delle caratteristiche della rete, si è scelto di lavorare in uno scenario di rete locale cablata a 1 Gb/s dove l'influenza di banda di trasmissione, ritardo e perdite fosse trascurabile. Per la stessa ragione il buffer di dejitter è stato dimensionato al valore minimo possibile in modo che i) compensasse il jitter della rete locale rendendo il numero di pacchetti scartati al ricevitore trascurabile, e ii) fosse (grossomodo) identico tra le diverse applicazioni confrontate: dell'ordine di 6–9 ms per le soluzioni native e dell'ordine di 20 ms per le soluzioni web [4] .

4.1 Qualità audio

La scelta di trasmettere audio non compresso adottata dalla soluzione proposta e da molti software nativi per NMP garantisce in sè (salvo perdite) la massima qualità audio e non necessita di confronti tra gli applicativi.

Nel caso della soluzione web si è provveduto comunque a svolgere un test di ascolto "indicativo" per confrontare la soluzione proposta (con audio non compresso) con una soluzione WebRTC tradizionale (con audio compresso tramite Opus, controllo del guadagno, riduzione del rumore e cancellazione dell'eco) tra quelle spesso utilizzate per la didattica musicale remota durante l'emergenza COVID-19.

Il test è stato svolto in ambiente controllato con livello di rumore di fondo in accordo con le raccomandazioni della ITU-R BS.1116-1 da parte di sette soggetti che si sono autodichiarati normoudenti e che hanno valutato quattro stimoli corrispondenti a differenti tipologie di audio selezionati tra il materiale già utilizzato in un precedente studio [8]: violino (6,7s), cembalo (6,9s), soprano (2,8s), soprano vibrato (8,1s). Per ogni stimolo, si è confrontato l'audio non compresso trasmesso dalla soluzione proposta con quello compresso trasmesso da una applicazione WebRTC tradizionale tramite un test di tipo ABX [9] presentato in cuffia (Sennheiser 600HD) ad ogni partecipante.

	Violino		Cembalo		Soprano		Soprano v.	
Soggetto	P	T	P	T	P	T	P	T
#1	×	✓	×	×	×	×	✓	✓
#2	×	✓	×	✓	×	✓	×	✓
#3	×	✓	✓	×	✓	✓	×	×
#4	×	✓	×	✓	×	×	×	✓
#5	×	×	✓	✓	×	✓	✓	×
#6	×	✓	×	✓	×	✓	×	✓
#7	✓	✓	×	✓	×	✓	×	✓
Totale	1	6	2	5	1	5	2	5

Tabella 1. Risultati del test d'ascolto di tipo ABX, con indicazione, per ciascun soggetto, dei riconoscimenti corretti (✓) e di quelli errati (×) per ciascuna delle due soluzioni, quella proposta (P) e quella terza (T).

In questa procedura vengono presentati all'ascoltatore tre stimoli: lo stimolo A e lo stimolo B, che hanno una differenza nota, e lo stimolo X. Il compito dell'ascoltatore è d'identificare se X è uguale ad A o a B. Se non vi è alcuna differenza udibile tra i due segnali, le risposte dell'ascoltatore dovrebbero essere distribuite in modo binomiale in modo tale che la probabilità di rispondere $X = A$ sia uguale alla probabilità di rispondere $X = B$, ovvero il 50%. Questo punteggio viene interpretato come un'indicazione dell'assenza di differenze percettive tra A e B.

Il numero minimo di risposte corrette necessarie per indicare una differenza percettiva può essere dato dalla probabilità cumulativa inversa di una distribuzione binomiale, basata sul numero di prove, livello di confidenza e probabilità di risposta corretta. Per le condizioni del test effettuato il numero minimo di risposte corrette necessarie per indicare una differenza percettiva è 6.

I risultati sono presentati in Tab. 1. Nel caso dell'audio trasmesso dal sistema HiFiReM , in media solo 1,5 soggetti su 7 hanno individuato correttamente l'audio non originale nelle varie prove d'ascolto, mentre nel caso dell'audio trasmesso da una applicazione web terza in media ben 5,25 persone su 7 hanno individuato correttamente l'audio non originale. [5]

4.2 Ritardo bocca-orecchio

L'analisi del ritardo bocca-orecchio, i cui risultati sono riportati in Fig.3, ha fornito risultati significativamente diversi nel caso della componente nativa e della componente web. La componente nativa si è dimostrata essere al pari, se non migliore, delle altre implementazioni con cui ci si è confrontati: SonoBus e JackTrip. La componente web ha evidenziato un ritardo bocca-orecchio molto più significativo, tra i 100 e i 150 ms, quando eseguita su sistemi operativi Windows o Ubuntu. Ha ottenuto invece prestazioni decisamente migliori su macOS con browser Firefox, dove la latenza misurata è stata di circa 30-40 ms (ritardi tra 50-60 ms si sono misurati utilizzando invece il browser Chrome su macOS). Seppur non confrontabili con il ritardo delle soluzioni native, questi risultati preliminari indi-

[4] Non è possibile impostare un valore preciso al millisecondo di questo parametro perché spesso è calcolato come multiplo della dimensione di un frame audio, che dipende dalle specifiche scelte implementative dell'applicazione in uso.

[5] Sebbene il test non possa reputarsi statisticamente significativo rappresenta un riscontro promettente sulla qualità del sistema, volendo svolgere successivamente prove più sistematiche su un più ampio numero di soggetti e contenuti audio.

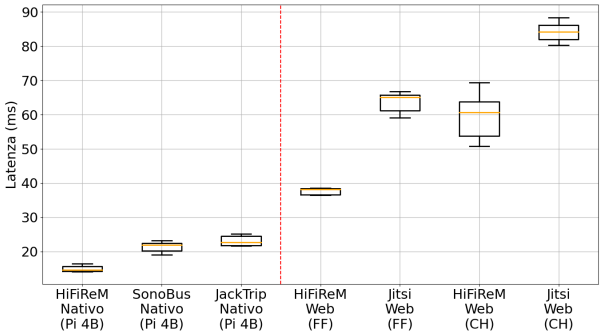


Figura 3. Ritardo bocca-orecchio (latenza) delle varie soluzioni analizzate, sia native su Raspberry Pi 4B (a sinistra), sia web su macOS con differenti browser, Mozilla Firefox v.97 (FF) e Google Chrome v.98 (CH) (a destra).

cano un significativo guadagno rispetto al ritardo di altre implementazioni web con compressione audio, che si aggira, nelle stesse condizioni, intorno ai 60 ms per Firefox e agli 80 ms per Chrome su macOS (su Windows e Ubuntu i ritardi eccedono i 150 ms). Il motivo del guadagno in ambiente macOS è verosimilmente dovuto alla migliore integrazione del browser nel sistema di gestione audio: su Windows non è possibile far lavorare il browser con i driver audio ASIO che assicurerebbero un minor ritardo, su Linux, sebbene esistano server audio a bassa latenza come Jack, nelle più comuni distribuzioni non è possibile collegare direttamente i browser a questi ultimi, ma è necessario connetterli attraverso i server audio di default che hanno maggior ritardo [6] .

5. CONCLUSIONI E SVILUPPI FUTURI

La realizzazione di un sistema unificato, web e nativo, per la Networked Music Performance presenta al momento ancora dei limiti tecnici da superare. Infatti, mentre non si riscontrano difficoltà rilevanti nel migrare la maggior parte delle applicazioni in ambiente web, nel caso di applicazioni real-time la limitata integrazione del browser con il sistema operativo costituisce ancora un ostacolo per il soddisfacimento dei requisiti di bassa latenza. L'implementazione corrente ha comunque ottenuto il risultato di poter garantire, con ritardi ridotti rispetto alle soluzioni esistenti, un ambiente web per la trasmissione in tempo reale di audio non compresso ad altà fedeltà adeguato per la didattica musicale.

L'integrazione della componente web e della componente nativa in un contesto applicativo indirizzato alla didattica musicale pone le basi per una successiva sperimentazione sul campo a beneficio delle scuole di musica e dei conservatori italiani. Permettere l'esecuzione di queste attività da remoto significa facilitare la partecipazione degli studenti indipendentemente dal luogo di residenza; limitare gli spostamenti riducendo i costi in termini di denaro e tempo, mitigando nel contempo l'inquinamento dovuto alle emissioni dei mezzi di trasporto.

[6] Una soluzione interessante per il futuro è l'integrazione dentro le distribuzioni Linux del server audio PipeWire (https://pipewire.org/) che è specificamente orientato a fornire bassa latenza per l'audio real-time.

6. RINGRAZIAMENTI

La parte web di questo progetto è stata avviata come tesi di laurea [10] e pubblicata online su GitHub [11]. Il progetto si è poi evoluto, includendo anche la parte nativa, come attività di dottorato. Si ringraziano le persone che a vario titolo hanno collaborato alle ricerche qui descritte, in particolar modo (in ordine alfabetico): Carlo Barbagallo, Chris Chafe, Leonardo Severi, Luigi Pirisi, Massimiliano Zanoni, Paolo Gastaldi.

Queste ricerche sono state parzialmente finanziate dal Fondo Integrativo Speciale per la Ricerca (FISR) del Ministero dell'Università e della Ricerca (MUR), Programma FISR COVID 2020, tramite il progetto (cod. FISR2020IP_-01206) "Piattaforma ad alta fedeltà per didattica musicale remota e concerti distribuiti su rete Internet".

7. BIBLIOGRAFIA

[1] J. Chappell, "SonoBus." https://sonobus.net/, Accesso: giu. 2022.

[2] J.-P. Cáceres and C. Chafe, "JackTrip: Under the hood of an engine for network audio," *Journal of New Music Research*, vol. 39, pp. 183–187, nov. 2010. https://dx.doi.org/10.1080/09298215.2010.481361.

[3] C. Hoene, I. Howell, and A. Carôt, "Networked Music Performance: Developing Soundjack and the Fastmusic Box During the Coronavirus Pandemic," in *AES International Audio Education Conference*, lug. 2021.

[4] L. Vignati, S. Zambon, and L. Turchet, "A comparison of real-time linux-based architectures for embedded musical applications," *J. Audio Eng. Soc*, vol. 70, no. 1/2, pp. 83–93, 2022.

[5] M. Sacchetto, P. Gastaldi, C. Chafe, C. Rottondi, and A. Servetti, "Web-Based Networked Music Performances via WebRTC: a Low-Latency PCM Audio Solution," *Journal of the Acoustical Society of America*, 2022. (in stampa).

[6] M. Sacchetto, A. Servetti, and C. Chafe, "JackTrip WebRTC: Networked Music Experiments in a Web Browser," in *6th International Web Audio Conference (WAC)*, (Barcellona, Spagna), lug. 2021.

[7] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, vol. 4, pp. 8823–8843, dic. 2016. https://doi.org/10.1109/ACCESS.2016.2628440.

[8] I. Howell, K. J. Gautereaux, J. Glasner, N. Perna, C. Ballantyne, and T. Nestorova, "Preliminary Report: Comparing the Audio Quality of Classical Music Lessons Over Zoom, Microsoft Teams, VoiceLessonsApp, and Apple FaceTime." [online] https://www.ianhowellcountertenor.com/preliminary-report-testing-video-conferencing-platforms, mag. 2020.

[9] W. Munson and M. Gardner, "Standardizing Auditory Tests," *Journal of the Acoustical Society of America*, vol. 22, no. 4, pp. 675–675, 1950. https://doi.org/10.1121/1.1917190.

[10] M. Sacchetto, "JackTrip-WebRTC - Networked Music Performance with Web Technologies," in *Tesi di Laurea Magistrale in Ingegneria Informatica presso il Politecnico di Torino*, 2020. http://webthesis.biblio.polito.it/id/eprint/16659.

[11] M. Sacchetto, "JackTrip-WebRTC." https://github.com/JackTrip-webrtc/JackTrip-webrtc, 2020. Accesso: giu. 2022.

RICONOSCIMENTO IN TEMPO REALE DI TECNICHE ESPRESSIVE PER CHITARRA SU EMBEDDED COMPUTERS

Domenico Stefani
Creative, Intelligent & Multisensory Interactions Laboratory
Università di Trento
domenico.stefani@unitn.it

ABSTRACT

Negli ultimi decenni dello scorso secolo, le innovazioni nel campo dell'elettronica analogica hanno portato alla creazione di una serie di strumenti-controller basati sulla chitarra che potevano produrre un'ampia gamma di suoni controllando direttamente dei sintetizzatori audio. Nonostante i vari miglioramenti portati a questi sistemi negli anni, essi sono sempre stati limitati al tracciamento della frequenza e intensità nel tempo delle note suonate, mancando di considerare le sottili tecniche espressive generalmente usate dai chitarristi per mutare il suono dello strumento. In questo documento presento la mia ricerca, che si propone di utilizzare le tecniche più moderne del deep learning per permettere il riconoscimento in tempo reale della tecnica espressiva usata su di una chitarra. Particolare attenzione verrà dedicata a proporre implementazioni su dispositivi embedded e a considerare le limitazioni di questi, di maniera da poter permettere la creazione di nuovi strumenti intelligenti dove l'analisi del segnale e sintesi di nuovi suoni venga svolta in maniera auto-contenuta. Assieme ad una descrizione della ricerca e della metodologia utilizzata, vengono presentati i risultati ottenuti fino ad ora nel deep learning per classificazione in tempo reale di audio, classificazione di tecniche espressive e onset detection. Infine, vengono presentate le linee di ricerca che verranno seguite nel futuro prossimo.

1. INTRODUZIONE

L'evoluzione della tecnologia audio degli ultimi cinquanta anni ha reso possibile l'uso della chitarra per produrre un'ampia gamma di nuovi suoni, generati controllando direttamente sintetizzatori audio. Questa famiglia di strumenti-controller ha potuto beneficiare da diverse innovazioni elettroniche, che hanno portato ad un miglioramento dei processi di tracciamento dell'*intonazione* ed evoluzione temporale delle note suonate dal chitarrista. Questo ha poi permesso di controllare dei sintetizzatori audio usando la chitarra al posto di una tastiera. Tuttavia, codificare la performance di un chitarrista tramite questi unici due parametri risulta limitante, in quanto manca di considerare le differenze timbriche introdotte tramite l'uso di varie tecniche espressive, proprie delle performance di ogni chitarrista. Alcuni esempi sono l'uso di diverse posizioni del plettro, gli armonici, mutare le corde o il vibrato. La mancanza di riconoscimento automatico di queste tecniche porta ad avere un controllo ridotto sulla generazione del suono sintetico, oltre che ad un'interazione atipica tra il chitarrista e strumento, dove l'uso di ogni tecnica espressiva non ha effetto sul suono e viene generalmente evitato. Questo porta gli utenti di questo tipo di controller a dover ricorrere interfacce fisiche convenzionali di sintetizzatore per modificare il suono sintetico, staccando le mani dallo strumento invece di utilizzare le tecniche espressive che sono abituati ad utilizzare.

Questo articolo presenta la mia ricerca, sullo sviluppo di modelli di deep learning che possano riconoscere in tempo-reale le tecniche espressive usate da un chitarrista, così che questa informazione possa essere riutilizzata per controllare algoritmi di sintesi sonora o diversi sistemi per performance audio-video o effettistica da palco come sistemi di illuminazione. I requisiti principali di questi sistemi sono l'accuratezza del riconoscimento, la ridotta latenza nell'esecuzione di questi algoritmi in tempo reale. In aggiunta, la mia ricerca si concentra principalmente sull'esecuzione di questi sistemi in dispositivi embedded come piccoli single-board computers che possano essere integrati direttamente all'interno di strumenti musicali, risultando in nuovi strumenti musicali con intelligenza "incorporata", come la famiglia degli smart musical instruments [1].

2. LAVORI CORRELATI

Questa ricerca si svolge nel campo del Music Information Retrieval (MIR), campo che si concentra sull'estrazione di informazioni dalla musica. Alcuni esempi di sfide che vengono affrontate nel campo del MIR includono la trascrizione automatica di performance musicali [2], beat-detection [3], onset detection [4], e classificazione del genere musicale [5].

Nonostante questo, la ricerca in questo campo si è sempre principalmente concentrata nell'ideazione di metodi offline (quindi non-in-tempo-reale) che vendono applicati all'analisi di grandi basi di dati, dove la principale metrica di valutazione è l'accuratezza del sistema creato, mentre generalmente non viene considerato il tempo di esecuzione. Sebbene questo permetta di sperimentare potenti metodi e algoritmi che non debbano necessariamente essere veloci o completare l'esecuzione in ridotti tempi prestabiliti, l'utilizzo di sistemi prettamente offline può essere applicato

Copyright: ©2022 Domenico Stefani et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

solo a contesti offline come l'analisi di basi di dati, mentre ogni tipo di sistema sviluppato per performance musicali deve necessariamente essere eseguito in tempo reale.

Il campo del MIR comprende anche una serie di metodi per il riconoscimento delle tecniche espressive, in particolare della chitarra, proposti negli anni [6], [7], [8], [9]. Tuttavia, in maniera simile alla gran parte dei sistemi di MIR esplorati in ricerca, anche i vari approcci sopracitati sono basati sull'esecuzione offline, quindi inadatti all'utilizzo in tempo reale, dove il riconoscimento della tecnica espressiva deve essere ottenuto in pochi millisecondi dall'inizio (anche detto *onset*) della nota relativa [10]. In particolare, Moore [11] descrive come $30ms$ sia il massimo intervallo fra due toni complessi, oltre il quale il sistema uditivo umano possa percepire i due come distinti. Questo offre un massimo limite empirico alla latenza di un sistema controller-sintetizzatore.

Con alcune eccezioni limitate (es. Reboursiere et al. [12]), il riconoscimento in tempo reale delle tecniche espressive per chitarra resta un campo largamente inesplorato, dove i recenti avanzamenti nel campo del deep learning possono dare risultati utili. In aggiunta, le piattaforme embedded per sistemi audio rese disponibili negli ultimi anni ([13], [14]) hanno aperto la possibilità di ideare strumenti intelligenti [1] che, tra le varie caratteristiche, sono definiti dalla presenza di *embedded intelligence*. Questa si riferisce alla capacità di processare l'audio dello strumento all'interno dello strumento stesso.

3. PROGRESSO E RISULTATI

Il problema affrontato in questa ricerca, ovvero il riconoscimento in tempo reale e su dispositivi embedded delle tecniche espressive per chitarra, offre una serie di sfide. Prima di tutto, in maniera simile agli approcci offline, il riconoscimento della tecnica deve essere accurato. Tuttavia, il riconoscimento deve anche avvenire con un ritardo quanto più minimo da quando il musicista suona una nota con lo strumento, di modo che sia possibile riutilizzare questa informazione per sintetizzare suoni che vengano percepiti come contemporanei alla nota originale. Questo limita la gamma di algoritmi che possono essere utilizzati ai più veloci, e richiede di trovare un compromesso con l'accuratezza del riconoscimento. Infine, eseguire simili sistemi di riconoscimento su di dispositivi embedded pone ulteriori limiti sulla quantità di operazioni che possono essere eseguite, rispetto a più potenti personal computers.

3.1 Primi studi e risultati

La parte iniziale di questa ricerca è stata concentrata sulla raccolta di dati audio per le tecniche espressive più comuni e lo sviluppo di una pipeline di classificazione a più stadi.

La pipeline, o catena di classificazione è composta da un onset detector, una serie di estrattori di proprietà timbriche del segnale (features) e infine una rete neurale che assolve il compito di classificatore, processando le features estratte dal segnale e calcolando a quale tecnica espressiva corrisponde la nota suonata. L'estrazione separata delle features richiede una ridotta potenza computazionale, contrariamente alle pratiche adottate nell'offline end-to-end deep learning, e permette di ottenere un sistema eseguibile in tempo reale. Questo passo si rivela fondamentale per un sistema che deve essere eseguito in tempo reale con limiti imposti dalla piattaforma hardware designata. Le features utilizzate includono Mel-frequency cepstral coefficients (MFCC), Bark-frequency cepstral coefficients (BFCC) e Real Cepstrum. Figura 1 mostra lo schema a blocchi del sistema di classificazione e alcune possibili applicazioni.

Per l'addestramento e il testing del classificatore si è resa necessaria la registrazione di un dataset di tracce audio, registrate direttamente dai pickup interni di cinque diverse chitarre, grazie ad altrettanti chitarristi con esperienza. È stato scelto di registrare una serie esaustiva di suoni (oltre 20 mila note in totale) con dodici diverse tecniche espressive suonate sulla chitarra acustica. Inoltre, le tecniche prese in considerazione sono in parte tecniche "convenzionali' suonate sulle corde (es. palm-mute, armonici, diverse posizioni del plettro), e in parte tecniche percussive che comportano l'uso del corpo della chitarra acustica come strumento percussivo.

Una demo preliminare del sistema è stata presentata in [15], mentre il primo classificatore è stato presentato in [16]. Data la complessità introdotta dai requisiti di bassa latenza ed esecuzione embedded, il problema è stato ridotto alla classificazione di singole note (monofonia).

Il sistema creato è in grado di distinguere tra tecniche percussive e convenzionali (classificazione binaria) con un' accuratezza del 99.2%, mentre ha raggiunto una simile accuratezza di 99.1% nel distinguere quattro diverse tecniche percussive e una macro-classe comprendente tutte le tecniche convenzionali in analisi. Al contrario, il problema di classificazione completo, quindi la categorizzazione tra dodici tecniche, si è rivelato considerevolmente più complesso a causa delle più sottili differenze timbriche nell'attacco delle tecniche suonate sulle corde. I risultati di questo primo sistema con il problema completo hanno infatti raggiunto solo un'accuratezza del 56.5%.

Tutti i risultati di classificazione sono stati ottenuti con una latenza media di $30.7\ ms$, vicina ai minimi requisiti discussi da Moore [11], ma migliorabile in un'iterazione successiva del sistema. La latenza media totale è il risultato dalla somma delle latenze di onset detection, allineamento della finestra di estrazione, computazione delle features, e tempo di esecuzione del classificatore, che corrispondono rispettivamente a $19.0\ ms$, $7.77\ ms$, $0.78\ ms$, e $3.15\ ms$.

Il valore aggiunto del lavoro presentato consiste nell'implementazione dell'intero sistema in un Raspberry PI 4: un single-board computer che per l'occasione è stato fornito del sistema operativo open-source per processing di audio in tempo reale, Elk OS [17]. Figura 2 mostra i dispositivi che compongono il sistema. Il segnale audio in input viene campionato a 48 kHz e raggruppato in buffer di 64 campioni per una ridotta latenza.

3.2 Raffinamento della pipeline di classificazione

Nonostante alcune delle limitazioni e i modesti risultati presentati in [16], la pipeline di classificazione ha dimostrato di essere funzionale, eseguibile in tempo reale e ra-

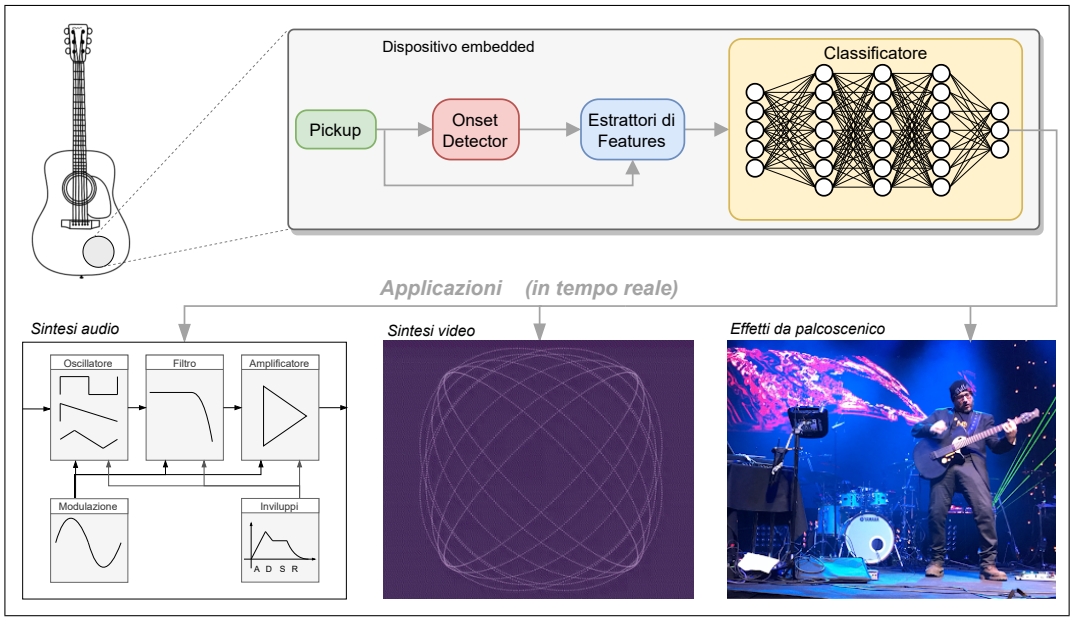


Figure 1. *Sistema di classificazione della tecnica espressiva in tempo reale e possibili applicazioni. Il segnale audio registrato dai pickup dello strumento viene analizzato da un onset detector, che attiva il calcolo di una serie di features timbriche quando viene identificata una nota nel segnale. In seguito, le features calcolate vengono utilizzate da una rete neurale profonda per categorizzare la tecnica espressiva usata. L'immagine mostra alcune possibili applicazioni in tempo reale che non si limitano alla sintesi audio, ma includono la sintesi video e il controllo di effettistica da palco come l'illuminazione.*



Figure 2. *Prototipo del classificatore di tecniche espressive. Il segnale audio dei pickup della chitarra viene passato ad un Raspberry PI 4 tramite un interfaccia audio (Elk PI Hat). Il risultato della classificazione viene comunicato tramite le cuffie con dei semplici toni (sinusoidali) con una diversa frequenza in base alla tecnica espressiva usata.*

gionevolmente accurata con il problema più semplice di classificazione percussiva, anche a fronte di una regolazione poco fine dei parametri di tutta la catena di classificazione.

Il lavoro di ricerca è quindi continuato con il raffinamento dei vari stadi della classificazione, a partire dal riconoscimento dell'inizio di una nota in tempo reale (Onset Detection). In [18] abbiamo presentato un approccio basato su algoritmi genetici per ottimizzare i parametri iniziali di vari onset parametrici, di maniera da raggiungere il compromesso ottimale tra l'accuratezza dell'individuazione dell'onset e la sua latenza. Questo ha permesso contemporaneamente di ridurre gli errori di individuazione (falsi positivi e negativi) e di ridurre la latenza di individuazione, ma soprattutto la sua variabilità. Ridurre la variabilità della latenza di individuazione dell'onset ha permesso di apprezzare un miglioramento sostanziale della performance del classificatore neurale, derivata da un allineamento più accurato della finestra di estrazione delle features, e di un aumento della qualità di queste ultime.

In secondo luogo, l'estrazione delle features è stata abbinata ad un algoritmo di selezione automatica di features. Questa ha permesso di mantenere solo i coefficienti più importanti e ridurre notevolmente il rumore delle features passate al classificatore neurale, che a sua volta ha permesso di ridurre le dimensioni della rete neurale, rimuovendo una grande parte di neuroni che altrimenti avrebbero dovuto essere dedicati dalla rete alla selezione delle features.

Infine, un lavoro più fine di progettazione e raffinamento del classificatore neurale è attualmente in corso, con soddisfacenti risultati preliminari superiori al 90% di accuratezza con una media di 20 millisecondi di latenza tra onset e risultati (10 millisecondi inferiore al lavoro presentato in precedenza).

La necessità di ottimizzare la latenza della classificazione ha portato anche alla comparazione di quattro diversi sistemi di esecuzione di modelli di Deep Learning su dispositivi con risorse computazionali limitate [19].

4. CONCLUSIONI E LAVORI FUTURI

Questo documento ha riportato il progresso della mia ricerca diretta al riconoscimento automatico in tempo reale delle tecniche espressive per chitarra su dispositivi embed-

ded. I risultati ottenuti fino ad ora hanno aperto la possibilità per alcune linee di ricerca per il futuro che verranno ora discusse brevemente.

Il lavoro attuale sul miglioramento della classificazione monofonica delle tecniche espressive per chitarra sarà completato con una serie di test tecnici e percettivi per verificare l'accuratezza del sistema e l'impercettibilità del ritardo tra le note suonate dal musicista e i suoni sintetici generati secondo la tecnica classificata.

In seguito, questa ricerca mirerà a coprire le limitazioni del sistema attuale, in particolare la classificazione monofonica. L'estensione alla classificazione polifonica della tecnica espressiva sarà affrontata tramite l'uso di un pickup esafonico, che permette di avere un segnale audio separato per ogni corda, su cui poter effettuare la classificazione.

Inoltre, sperimenterò le possibilità offerte dai più recenti dispositivi di accelerazione hardware come Tensor Processing Unit (TPU), le quali potrebbero permettere di eseguire il modello di classificazione più velocemente, oppure di avere modelli più complessi senza introdurre latenza eccessiva.

In maniera simile, sarà possibile utilizzare la tecnica del transfer-learning di modo da adattare un modello generico di classificazione della tecnica ad una chitarra specifica.

5. RIFERIMENTI

[1] L. Turchet, "Smart Musical Instruments: vision, design principles, and future directions," *IEEE Access*, vol. 7, pp. 8944–8963, 2019.

[2] S. Sigtia, E. Benetos, and S. Dixon, "An end-to-end neural network for polyphonic piano music transcription," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, pp. 927–939, 2016.

[3] S. Böck and M. Schedl, "Enhanced beat tracking with context-aware neural networks," in *Proc. 14th Int. Conf. on Digital Audio Effects (DAFx-11)*, pp. 135–139, 2011.

[4] J. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. Sandler, "A tutorial on onset detection in music signals," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 5, pp. 1035–1047, 2005.

[5] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, no. 5, pp. 293–302, 2002.

[6] T. H. Özaslan, E. Guaus, E. Palacios, and J. L. Arcos, "Attack based articulation analysis of nylon string guitar," in *Proc. 7th Int. Symposium on Computer Music Modeling and Retrieval (CMMR)*, 2010.

[7] C. Kehling, J. Abeßer, C. Dittmar, and G. Schuller, "Automatic tablature transcription of electric guitar recordings by estimation of score- and instrument-related parameters," *Proc. 17th Int. Conf. on Digital Audio Effects (DAFx 2014)*, pp. 1–8, 2014.

[8] Y. P. Chen, L. Su, and Y. H. Yang, "Electric guitar playing technique detection in real-world recordings based on F0 sequence pattern recognition," *Proc. 16th Int. Society for Music Information Retrieval Conference (ISMIR)*, pp. 708–714, 2015.

[9] J. Abeßer, H. Lukashevich, and G. Schuller, "Feature-based extraction of plucking and expression styles of the electric bass guitar," in *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2290–2293, 2010.

[10] A. McPherson, R. Jack, and G. Moro, "Action-sound latency: Are our tools fast enough?," in *Proc. Int. Conference on New Interfaces for Musical Expression*, pp. 20–25, 2016.

[11] B. C. J. Moore, *An introduction to the psychology of hearing*. Brill, 2012.

[12] L. Reboursiere, O. Lähdeoja, R. Chessini, T. Drugman, S. Dupont, C. Picard, and N. Riche, "Guitar as controller," in *Journal of The Acoustical Society of America*, vol. 80, pp. 41–54, 01 2011.

[13] L. Vignati, S. Zambon, and L. Turchet, "A comparison of real-time linux-based architectures for embedded musical applications," *Journal of the Audio Engineering Society*, vol. 70, no. 1/2, pp. 83–93, 2022.

[14] E. Meneses, J. Wang, S. Freire, and M. Wanderley, "A comparison of open-source linux frameworks for an augmented musical instrument implementation," in *Proc. Int. Conf. on New Interfaces for Musical Expression*, pp. 222–227, June 2019.

[15] D. Stefani and L. Turchet, "Demo of the timbreid-vst plugin for embedded real-time classification of individual musical instruments timbres," in *Proc. 27th Conf. of Open Innovations Association (FRUCT)*, vol. 2, pp. 412–413, 2020.

[16] D. Stefani and L. Turchet, "On the Challenges of Embedded Real-Time Music Information Retrieval," in *Proceedings of the 25-th Int. Conf. on Digital Audio Effects (DAFx20in22)*, vol. 3, pp. 177–184, Sept. 2022.

[17] L. Turchet and C. Fischione, "Elk Audio OS: an open source operating system for the Internet of Musical Things," *ACM Transactions on the Internet of Things*, vol. 2, no. 2, pp. 1–18, 2021.

[18] D. Stefani and L. Turchet, "Bio-Inspired Optimization of Parametric Onset Detectors," in *Proc. 24th Int. Conf. on Digital Audio Effects (DAFx20in21)*, vol. 2, pp. 268–275, Sept. 2021.

[19] D. Stefani, S. Peroni, and L. Turchet, "A Comparison of Deep Learning Inference Engines for Embedded Real-Time Audio Classification," in *Proceedings of the 25-th Int. Conf. on Digital Audio Effects (DAFx20in22)*, vol. 3, pp. 256–263, Sept. 2022.

DEEP PACKET LOSS CONCEALMENT PER SISTEMI DI INTERAZIONE REMOTA IN TEMPO REALE

Luca Vignati
Creative, Intelligent & Multisensory Interactions Laboratory
Università di Trento
luca.vignati@unitn.it

ABSTRACT

Gli ultimi decenni, inclusa la recente pandemia, hanno indotto la necessità, ancora in crescita, di interazioni remote tra le persone. Anche se nella maggior parte dei casi questa può essere soddisfatta con gli ormai popolari servizi di videochiamata, alcuni casi d'uso rimangono esclusi dalle interazioni remote. Questo è il caso per i musicisti, la cui sensibilità nei confronti del tempo e della qualità sonora pone un grosso ostacolo per coloro che tentano di sviluppare una soluzione. L'infrastruttura stessa sulla quale le interazioni hanno luogo, ovvero la rete Internet pubblica, non è progettata per consentire interazioni con requisiti così stringenti. Per questo motivo, ogni soluzione proposta deve piegarsi ad una scomoda realtà: i protocolli adatti alla trasmissione in tempo reale sono "best effort", ovvero non forniscono garanzie su se e quando i pacchetti inviati arriveranno a destinazione. Da qui nasce il Packet Loss Concealment (PLC), ovvero il tentativo di indovinare il contenuto di un pacchetto perso. L'attuale stato dell'arte in questo ambito è raggiunto con le tradizionali tecniche di DSP (Digital Signal Processing), ma recenti sviluppi nell'ambito del Deep Learning suggeriscono che sia ora possibile raggiungere risultati migliori. In questo articolo presento lo stato attuale della mia ricerca di dottorato nell'ambito del PLC, con alcuni risultati parziali.

1. INTRODUZIONE

Il feedback uditivo per la musica dal vivo, le esecuzioni musicali remote via Internet e le applicazioni di teleconferenza di alta qualità sono alcuni esempi di applicazioni che impongono requisiti molto stringenti sulla qualità dell'audio e, soprattutto, sulla latenza di trasmissione [1]. Gli standard odierni per lo streaming audio funzionano con latenze di andata e ritorno dell'ordine di qualche centinaio di millisecondi, ma gli scenari sopracitati richiedono latenze dell'ordine di poche decine di millisecondi [2]. Esperimenti psico-acustici hanno dimostrato che, affinché due musicisti geograficamente distanti possano suonare insieme in modo sincrono, la latenza totale per entrambe le parti deve rimanere al di sotto dei 20-30 millisecondi [3, 4]. Inoltre, il jitter (cioè la variabilità della latenza) è un aspetto particolarmente critico per i musicisti e non deve superare i 2-3 millisecondi.

La globalizzazione e la recente pandemia hanno evidenziato la necessità dei musicisti di suonare insieme in tempo reale da luoghi diversi via Internet. Tuttavia, le soluzioni attuali non sono in grado di soddisfare contemporaneamente gli stringenti requisiti di latenza impercettibile e di audio ad alta qualità necessari per catturare le sfumature dell'esecuzione. È importante sottolineare che questi problemi riguardano sia i musicisti professionisti che quelli amatoriali, indipendentemente dallo strumento suonato.

Molti segmenti di mercato emergenti (come l'automotive, l'IoT e gli eSport) utilizzano lo stato dell'arte del networking e fanno sempre più affidamento su di esso. Questi settori ampliano i confini di ciò che è possibile fare tramite la rete Internet, ponendo requisiti sempre più ambiziosi agli operatori e ai fornitori di rete. A loro volta, queste nuove capacità di rete aprono le porte a molte nuove applicazioni un tempo ritenute impossibili, come le esecuzioni musicali remote via Internet. Nell'ultimo decennio, ci sono stati alcuni sforzi di ricerca incentrati sullo streaming audio a bassa latenza per queste applicazioni sulle attuali tecnologie di rete, come Ethernet e IEEE 802.11 WiFi [4].

Questo campo di ricerca comprende molti argomenti: Packet Loss Concealment (PLC), Forward Error Correction (FEC), gestione della coda di jitter, sincronizzazione temporale e compressione con perdita. L'uso combinato di tecniche provenienti da questi argomenti e le capacità delle nuove tecnologie di rete, come la 5G NR URLLC (New-Radio Ultra-Reliable Low-Latency Connection), ha il potenziale di rendere possibile lo streaming audio in tempo reale e ad alta qualità via Internet.

Questo articolo presenta la mia ricerca nell'ambito del Packet Loss Concealment, che tra gli argomenti elencati sopra è il problema con il maggiore impatto sulla qualità dell'esperienza offerta dai sistemi di interazione remota in tempo reale. In particolare, la mia ricerca punta all'utilizzo dello stato dell'arte del Deep Learning per creare un algoritmo di PLC che sia in grado di offrire una qualità superiore operando in tempo reale.

2. LAVORI CORRELATI

Questa ricerca si colloca nell'intersezione tra l'ambito delle Networked Music Performances (NMP) [3], del Music Information Retrieval (MIR) e del Deep Learning. Al momento della scrittura di questo articolo, invece, lo stato dell'arte inquadra questo problema nell'intersezione tra

Copyright: ©2022 Luca Vignati et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

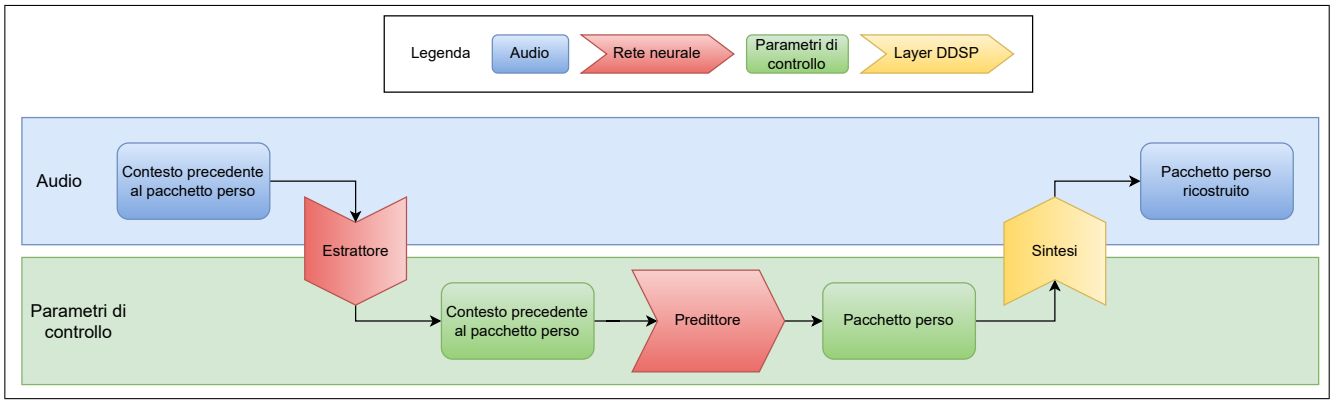


Figura 1. *Architettura dell'algoritmo di PLC. Su sfondo azzurro è rappresentato il dominio dell'audio, su sfondo verde il dominio dei parametri di controllo. L'estrattore e la sintesi sono i due componenti che consentono la transizione da un dominio all'altro. Il predittore lavora esclusivamente nel dominio dei parametri di controllo.*

Digital Signal Processing (DSP) e Time Series Prediction, poiché le consolidate tecniche tradizionali di DSP offrono il miglior risultato con il costo computazionale più basso [5, 6]. Inoltre, i più recenti risultati che adottano tecniche basate su Deep Learning per PLC sono indirizzate e ottimizzate per il parlato invece che per gli strumenti musicali [7, 8, 9].

Nonostante la letteratura offra alcuni studi che adottano tecniche basate su Deep Learning per PLC ottimizzate per strumenti musicali [10], le nuove possibilità sbloccate da nuove tecnologie come Differentiable Digital Signal Processing (DDSP) [11] e Temporal Belief Memory [12] suggeriscono che sia ora possibile raggiungere risultati superiori.

3. PROGRESSO E RISULTATI

Questa ricerca punta a risolvere il problema di creare un algoritmo di PLC usando il Deep Learning per migliorare lo stato dell'arte. Ciò presenta due sfide principali:

- L'**ideazione** di una o più reti neurali necessarie a svolgere i compiti di predizione e generazione del pacchetto perso, inclusa la relativa raccolta o generazione di un adeguato dataset per l'addestramento e la validazione del modello.
- L'**ottimizzazione** pre e post addestramento dell'architettura della rete per consentirne l'esecuzione in tempo reale.

In questo articolo si discutono i progressi e i risultati parziali ottenuti sul fronte del primo punto.

3.1 Architettura

La parte iniziale di questa ricerca si è concentrata sull'ideazione dell'architettura dell'algoritmo di PLC. Gli ottimi risultati ottenuti dal gruppo Magenta [11, 13, 14] con la loro libreria DDSP hanno ispirato un'architettura composta da due componenti come mostrato in Figura 1.

L'idea di base consiste nell'avere una rete neurale specializzata nella predizione di serie temporali multi-variate (per esempio una rete neurale ricorrente o RNN) che utilizzi l'input di un contesto audio precedente alla perdita del pacchetto per predire e generare il pacchetto perso. La libreria DDSP entra in gioco nel semplificare il compito del predittore: questa infatti consiste in una serie di componenti del DSP tradizionale (oscillatori, filtri, riverberi ecc.) resi differenziabili (ovvero compatibili con la procedura di *backpropagation*, dal momento che consentono al gradiente di attraversarli) così da poterli inserire nell'architettura di una rete neurale. Questo rende l'architettura più interpretabile e, in questo caso, solleva il predittore dalla responsabilità di imparare a generare audio coerente. Infatti, come mostrato in Figura 1, il predittore opera interamente nello spazio dei parametri di controllo predicendo alcune feature, che grazie a DDSP sono interpretabili e comprensibili; queste rappresentano i parametri di controllo di un sintetizzatore formato da un banco di oscillatori (controllabili indipendentemente in frequenza e ampiezza) e da un rumore filtrato (controllabile nell'ampiezza delle singole bande).

Per consentire al predittore di operare esclusivamente su queste feature diventa necessario pre processare l'audio in input al sistema per estrarre le feature da dare in input al predittore. Il componente che si occupa di questo è l'estrattore mostrato in Figura 1, che estrae dallo spettrogramma in input le feature che rappresentano le tracce di frequenze e ampiezze delle armoniche del segnale nel tempo, oltre che alle ampiezze delle bande del rumore filtrato nel tempo.

Al momento della stesura di questo articolo, l'unico componente sviluppato e testato è l'estrattore. Come mostrato in Figura 2 e 3, la sua architettura si sviluppa attorno ad una rete ResNet e varia in base al tipo di addestramento:

- **Addestramento con supervisione**: questo è l'addestramento principale, molto lungo e richiede grandi quantità di dati. Per questo viene usato un dataset sintetico generato usando gli stessi strumenti della libreria di DDSP così che l'audio sintetico possa essere dato in input alla rete e i parametri usati per generare il dataset possano essere usati come ground truth durante l'addestramento. La loss function uti-

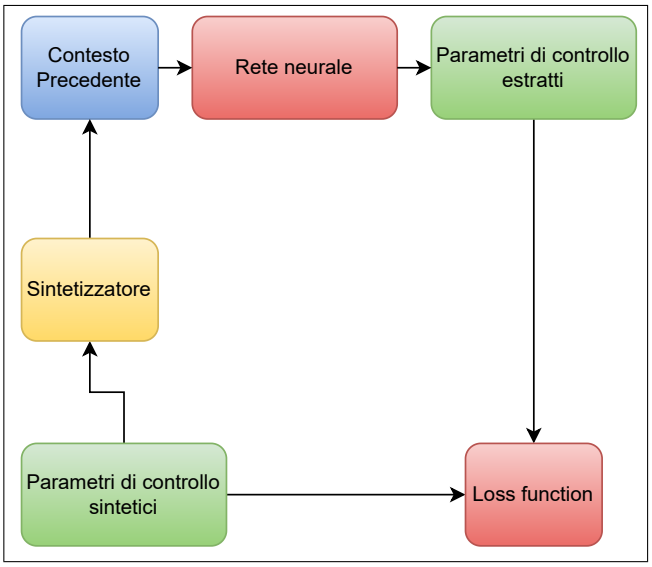


Figura 2. *Schema di **addestramento con supervisione** dell'estrattore. L'audio in input è sintetizzato partendo da dei parametri di controllo generati pseudo-casualmente. Questi parametri di controllo sono utilizzati anche come ground truth dalla loss function.*

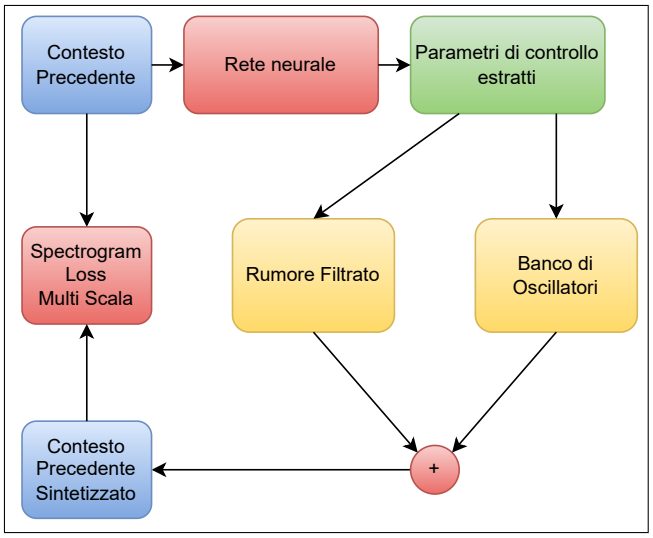


Figura 3. *Schema di **addestramento senza supervisione** dell'estrattore. I componenti gialli rappresentano i layer DSP differenziabili della libreria DDSP. Questi rendono possibile questo schema di addestramento in cui i parametri di controllo estratti vengono usati per risintetizzare l'audio, che a sua volta viene usato per calcolare la loss function contro l'input della rete.*

lizzata sfrutta la stima kernel di densità (KDE) [15] per calcolare la coerenza tra l'output e la ground truth.

- **Addestramento senza supervisione**: questo raffinamento dell'addestramento richiede molto meno tempo e dati e permette di rendere il modello più efficace nei confronti di uno specifico sottoinsieme di problemi (come il suono di un singolo strumento o di una famiglia di strumenti). In questo caso l'output che viene considerato è quello in uscita dai layer generatori della libreria DDSP, quindi si tratta di audio. La loss function utilizzata confronta gli spettrogrammi dell'audio in ingresso e quello in uscita usando diverse dimensioni delle finestre (da 64 a 2048) impiegate per calcolare la Short-Time Fourier Transform (STFT). Questa tecnica permette di ridurre la penalità derivante dal trade-off tra la precisione nel tempo e in frequenza imposta dalla STFT. In questo caso, la rete assume la forma di un autoencoder, dove l'encoder è costituito dalla ResNet, lo spazio latente sono i parametri di controllo e il decoder è rappresentato dai layer DDSP che sintetizzano l'audio a partire dai parametri di controllo estratti.

Per raggiungere il massimo livello di accuratezza si può quindi procedere all'addestramento della rete utilizzando entrambi i tipi di addestramento: prima quello con supervisione e poi quello senza supervisione. Questo approccio consente di sfruttare i vantaggi di entrambi gli approcci, utilizzando un dataset sintetico (virtualmente illimitato in dimensione e caratteristiche) per compiere la parte più significativa dell'addestramento e un dataset reale (con esempi più accurati e realistici ma di più difficile reperibilità) per rifinire l'addestramento.

3.2 Risultati parziali

Lo stato embrionale di questa ricerca implica che gli unici risultati disponibili al tempo della stesura di questo articolo siano quelli relativi all'addestramento con supervisione dell'estrattore presentato nel capitolo precedente.

In Figura 4 è mostrato un esempio tratto dal set di validazione del modello, dove sono messi a confronto l'input e l'output dell'estrattore. Come si può osservare, il modello ha raggiunto un buon livello di accuratezza; le discrepanze più evidenti sono dovute principalmente al limitato numero di oscillatori a disposizione del banco di oscillatori per sintetizzare le armoniche del segnale. Il modello però compensa questa limitazione prediligendo le armoniche con ampiezza maggiore, riducendo quindi la differenza timbrica percepita.

I risultati mostrati sono l'esito dell'addestramento con supervisione e sono tratti dal dataset sintetico usato per l'addestramento e la validazione. L'ulteriore addestramento senza supervisione svolto su un dataset composto da audio reale è in grado di aumentare ulteriormente l'accuratezza del modello per uno specifico ambito timbrico.

4. CONCLUSIONI E LAVORI FUTURI

I risultati presentati nella sezione precedente sono incoraggianti dal momento che mostrano che l'estrattore è in grado di estrarre una rappresentazione più significativa dallo spettrogramma che riceve in input. Tuttavia resta da valutare l'impatto che l'utilizzo di questa rappresentazione ha sulle prestazioni del predittore. Infatti, l'intera architettura del predittore necessita ancora di essere messa alla prova per individuare la configurazione ottimale. Inoltre, in questa prima fase è stato trascurato l'aspetto dell'esecuzione

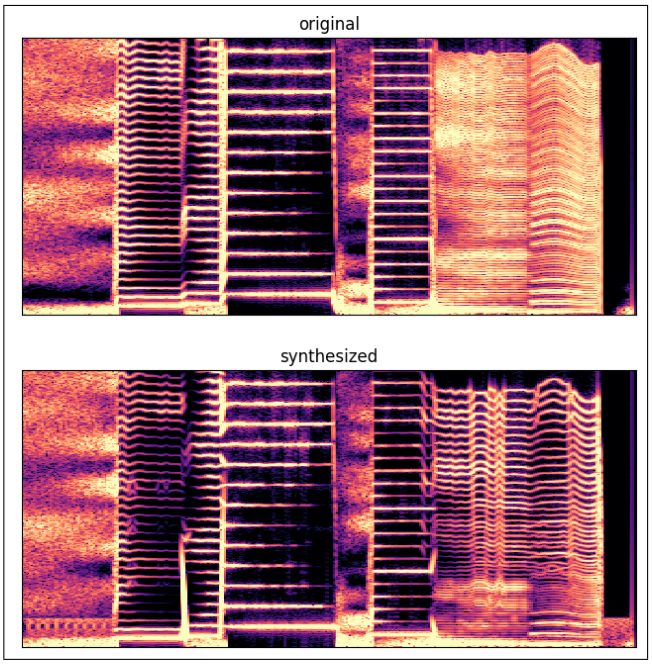


Figura 4. *Confronto tra input (sopra) e output (sotto) dell'estrattore alla fine del training con supervisione. I parametri in output sono stati usati per sintetizzare l'audio visualizzato nell'immagine sotto. Gli spettrogrammi raffigurati rappresentano un esempio di audio estratto dal dataset sintetico usato per la validazione del modello durante l'addestramento.*

in tempo reale. Durante l'ottimizzazione dell'architettura verrà misurato il costo computazionale per garantirne il funzionamento in tempo reale. Infine, è fondamentale affrontare il problema della minimizzazione delle discontinuità tra il pacchetto sintetizzato e quelli a lui adiacenti poiché eventuali discontinuità nel segnale riprodotto si traducono in artefatti molto percepibili. Per valutare il risultato a livello percettivo sarà necessario condurre dei test di ascolto con musicisti.

5. BIBLIOGRAFIA

[1] L. Turchet, C. Fischione, G. Essl, D. Keller, and M. Barthet, "Internet of Musical Things: Vision and Challenges," *IEEE Access*, vol. 6, pp. 61994–62017, 2018.

[2] X. Jiang, H. Shokri-Ghadikolaei, G. Fodor, E. Modiano, Z. Pang, M. Zorzi, and C. Fischione, "Low-latency networking: Where latency lurks and how to tame it," *Proceedings of the IEEE*, vol. 107, no. 2, pp. 280–306, 2018.

[3] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.

[4] L. Gabrielli and S. Squartini, *Wireless Networked Music Performance*. Springer, 2016.

[5] Y. HUANG, *Packet Loss Concealment in Networked Music Performance using Auto-Regressive Models*. PhD thesis, 2022.

[6] M. Fink, *Enhancements for Networked Music Performances*. PhD thesis, Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg, 2018.

[7] J.-M. Valin, A. Mustafa, C. Montgomery, T. B. Terriberry, M. Klingbeil, P. Smaragdis, and A. Krishnaswamy, "Real-time packet loss concealment with mixed generative and predictive model," *arXiv preprint arXiv:2205.05785*, 2022.

[8] L. Diener, S. Sootla, S. Branets, A. Saabas, R. Aichner, and R. Cutler, "Interspeech 2022 audio deep packet loss concealment challenge," *arXiv preprint arXiv:2204.05222*, 2022.

[9] S. Deng and C. Bao, "Phase unwrapping based packet loss concealment using deep neural networks," *Speech Communication*, vol. 138, pp. 88–97, 2022.

[10] P. Verma, A. I. Mezza, C. Chafe, and C. Rottondi, "A deep learning approach for low-latency packet loss concealment of audio signals in networked music performance applications," in *2020 27th Conference of Open Innovations Association (FRUCT)*, pp. 268–275, IEEE, 2020.

[11] J. Engel, L. Hantrakul, C. Gu, and A. Roberts, "Ddsp: Differentiable digital signal processing," *arXiv preprint arXiv:2001.04643*, 2020.

[12] Y.-J. Kim and M. Chi, "Temporal belief memory: Imputing missing data during rnn training.," in *In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-2018)*, 2018.

[13] C. Hawthorne, I. Simon, R. Swavely, E. Manilow, and J. Engel, "Sequence-to-sequence piano transcription with transformers," *arXiv preprint arXiv:2107.09142*, 2021.

[14] J. Engel, R. Swavely, L. H. Hantrakul, A. Roberts, and C. Hawthorne, "Self-supervised pitch detection by inverse audio synthesis," in *In Proceedings of the 37 th International Conference on Machine Learning*, 2020.

[15] G. R. Terrell and D. W. Scott, "Variable kernel density estimation," *The Annals of Statistics*, pp. 1236–1265, 1992.

Implementing UHJ stereo in the Envelop for Live suite

Mastrorillo Alessio
Conservatorio N. Piccinni, Bari
mastrolex@gmail.com

Silvi Giuseppe
Conservatorio N. Piccinni, Bari
grammaton@me.com

Scagliola Francesco
Conservatorio N. Piccinni, Bari
francesco.scagliola@gmail.com

ABSTRACT

In this work we introduce a new Ambisonics decoder, an implementation of UHJ stereo in the Envelop for Live suite (E4L). This format was developed by Michel Gerzon and is an alternative to conventional stereo and binaural decoding. We also present a sixth-order quadrature filter made with a cascade of Biquad all-pass filters, necessary to derive the UHJ stereo format from the First Order components. This decoder is developed in *gen~* codebox and is available as open-source repository.

1. INTRODUCTION

1.1 The UHJ format

Ambisonics C-format, also called UHJ from *Universal UD-4*, *Matrix H* and *System 45J*, is a development of the Ambisonics technique specially designed to be compatible with mono and stereo *(2-channel)* reproduction systems. It is stereo compatible without further decoding, while giving the listener a stereo image that is significantly wider than conventional stereo, both listening with headphones and stereo speakers. It conveys audio information from the entire Ambisonics First-Order (FOA):

... The front-stage material is reproduced with sharply defined images, occupying virtually the whole of the stereo stage, and some sound positions can appear marginally beyond the loudspeakers. Rear-stage sounds appear with rather less well-defined images between the stereo loudspeakers [1].

It represents an alternative to binaural, which keeps the entire sonic sphere but only works on headphones, and stereo decoding that is made by positioning virtual microphones in the sphere, thus excluding audio information coming from other directions. This format works with the four *WXYZ* components and derive four *LRTQ* components that can recreate the entire sonic sphere, where *LR* are two signals ready for stereo reproduction. For UHJ stereo applications, the first three components *WXY* are used, thus covering the entire horizontal plane.

Ambisonics UHJ can also be mixed down with classic stereo material: an example of UHJ stereo encoded material is present in many commercial album releases such as The Alan Parsons Project's *Stereotomy*, Paul McCartney's *Liverpool Oratorio*, Frank Perry's *Zodiac* and in many symphonic recordings [2].

Copyright: © 2022 Mastrorillo Alessio et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 Unported License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1.2 The Envelop for Live suite

Envelop for Live (E4L) is an open-source audio production framework for spatial audio composition and performance. Envelop for Live operates within the music production environment of Ableton Live 10+ and Max for Live. Envelop for Live is designed to be a highly modular, flexible platform for artists to compose and perform spatial audio, and for developers to create new kinds of audio effects for the Ambisonics domain [3].

Since no UHJ decoder was available in this suite, the author has collaborated with the community of developers in *Envelop Sound*, by developing a brand new decoder, which has been therefore included in the suite [4]. This decoder has been used for the author 1's electroacoustic music composition *Trippin' on the edge of time* and *Emptiness of the hanging* stereo master, premiered and selected by Electric Sound, SMC, ICMC, eviMus and Sin[x]Thésis.

2. THE GERZON'S MATH

In Gerzon's *Ambisonics in Multichannel Broadcasting and Video* [1], the formula to convert B-format into UHJ has a 90° shifted *j* component, as follows:

$$\Sigma = 0.9397W + 0.1856X \tag{1}$$

$$\Delta = j(-0.3420W + 0.5099X) + 0.6555Y \tag{2}$$

while the Left and Right signals can be derived with the following equations:

$$L = 1/2(\Sigma + \Delta) \tag{3}$$

$$R = 1/2(\Sigma - \Delta) \tag{4}$$

This equation takes in the three ambisonic B-format *WXY* components and put out Σ and Δ, from which the stereo left and right signals can be derived. In Fig. 1 are shown the polar patterns of the Σ and Δ intermediate components.

The *j* component is obtained by quadrature. A sinusoid can be expressed as the sum of a sine function (phase zero) and a cosine function (phase $\pi/2$). If the sine part is the *in-phase* component, the cosine part can be called the *phase-quadrature* component [5]. In general, phase-quadrature means '90° out of phase' [6]. The mathematical procedure to compute the phase quadrature is known as Hilbert Transform [7].

When computing such quadrature, the output will be decomposed in two components, with a relative phase distance of 90° between the two.

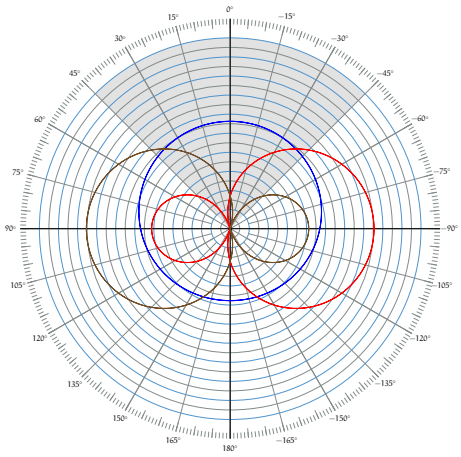


Figure 1. Σ in blue in the centre, Δ and $-\Delta$ in red and brown in the sides.

3. IMPLEMENTING THE NEW DECODER AND QUADRATURE FILTER

The decoder is written in *gen~* codebox which is an extension of the MaxMSP patching environment, allowing the code to compute at the sample-rate frequency, *'per-sample'*, for minimum latency and accurate timing. The coefficients in this *gen~* algorithm are calculated with *double-precision*.

The filter for the signal quadrature is made by a cascade of three Biquad all-pass filters with equation:

$$y[n] = b0w[n] + b1w[n-1] + b2w[n-2] \quad (5)$$

$$w[n] = x[n] - a1w[n-1] - a2w[n-2] \quad (6)$$

for the real part and three more for the imaginary part, as in Fig. 2. This sixth-order filter can be configured as desired, and the order can scale to higher values by adding more Biquads to the cascade.

The E4L suite internally uses AmbiX format with ACN channel order and SN3D weighting, thus a $-3db$ gain is applied to the W component to convert the content to FuMa format, as required by Eq. 1 and 2.

The signal goes in two filters as in Fig. 2. The coefficients for these filters are calculated considering the values of control frequency, quality factor and sample-rate. The three *WXY* components are filtered to make the *Wr, Xr* and *Yr* real parts and the *Wi* and *Xi* imaginary parts. In the Eq. 1 and 2 the real parts are used in both Σ and Δ, while the imaginary part is only used in the *j* arguments in Δ.

3.1 The search for the control frequency and quality factor values

The signal generated by 31 sinusoidal oscillators with amplitude 1, with frequencies subdivided by third of octave from 20Hz to 20000Hz, has been connected to the same

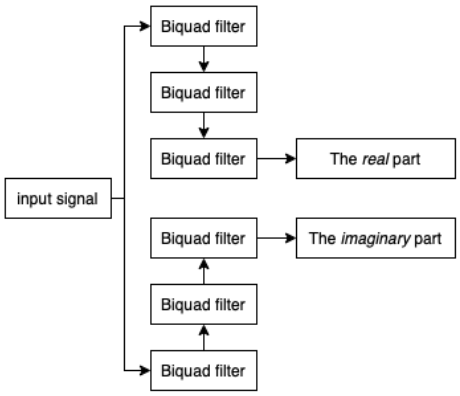


Figure 2. The quadrature filter

amount of filters for quadrature. The control frequency and quality factor in each filter are searched observing the output of the complex equation:

$$Re(x)^2 + Im(x)^2 = 1 \quad (7)$$

where $Re(x)$ and $Im(x)$ are the real and imaginary part of the signal x. The values discovered are shown in Tab. 1.

frequency Hz	**Q factor**
for the *real* part	
374.1	0.1093
666.8	0.4210
17551	0.5750
for the *imaginary* part	
35.61	0.2571
3723	0.3464
6786	0.1200

Table 1. The values for control frequency and Q factor

The output values of the quadrature filters, tested with the Eq. 7 are shown in Tab. 2. The values of Min and Max are the minimum and maximum values in the output oscillation, where a smaller deviation from 1 implies smaller error in quadrature. A greater value, as at 10Hz or above 16000Hz, means that the phase shift between the imaginary and real part is lower than 90°, thus reducing the stereo width in the lowest and highest frequencies.

3.2 Setting the filter

The implementation of the Biquad filters and the Gerzon's equation is straightforward, stable and light in computational cost. The filter setting, on the other hand, took many working hours to be configured. The use of 31 simultaneous oscillators and filters is expensive in terms of memory and CPU time, exponentially increasing when testing the filter response at higher sample-rates. It is anyway necessary, since observing the filter behavior at different frequencies at the same time is essential to set it correctly. The control frequency and quality factor of every single filter has been set while observing all the 31 filters output, starting with the lower frequencies first.

The set of values shown in Tab. 1 have been discovered with this procedure, and are the most precise and stable out of hundreds of different attempts.

Frequency Hz	Min	Max
10	1.021	0.981
20	1.013	0.989
25	1.008	0.993
31.5	1.004	0.995
40	1.002	0.999
50	1.003	0.996
63	1.007	0.997
80	1.010	0.990
100	1.002	0.996
125	1.016	0.984
160	1.003	0.990
200	1.021	0.996
250	1.022	0.982
315	1.008	0.996
400	1.006	0.987
500	1.001	0.994
630	1.006	0.996
800	1.015	0.971
1000	1.016	0.993
1250	1.001	0.998
1600	1.013	0.998
2000	1.005	0.990
2500	1.014	0.998
3150	1.001	0.997
4000	1.011	0.986
5000	1.008	0.979
6300	1.007	0.990
8000	1.012	0.985
10000	1.025	0.964
12500	1.011	0.992
16000	1.148	0.901
20000	1.244	0.719

Table 2. Result of the test with Eq. 7

4. CONCLUSION

This decoder has been used in the decoding of Ambisonics recordings with sample-rate up to 192kHz, returning great sonic details and an enveloping sound. When decoding Ambisonics material to 2-channel mixing, the choice on the decoder depends on the final reproduction system, that can be stereo speakers or headphones. With the UHJ format, this choice can be overcome, with the advantage of preserving the entire horizontal plane. Head rotation can be implemented, along with controls for balancing the omnidirectional W and directional XY components. The inclusion of this decoder in the E4L suite could be the opportunity for further development of the quadrature filter, with a higher-order and a wider frequency response.

The Ambisonics UHJ has proved to be a strong tool in music composition. The production and execution of author 1's *Trippin' on the edge of time* and *Emptiness of the hanging* has been the opportunity to successfully evaluate the entire environment, from the production step to the concert hall.

Acknowledgments

I would like to thank my supervisor Giuseppe Silvi for letting me discover the Ambisonics world, and Francesco Scagliola for the support in implementing this technique in music composition.

5. REFERENCES

[1] M. Gerzon, "Ambisonics in multichannel broadcasting and video," *Journal of the Audio Engineering Society*, p. 31, 1983.

[2] "Surround Discography website," www.surrounddiscography.com, accessed: 2022-06-12.

[3] "Envelop for Live webpage," https://www.envelop.us/software, accessed: 2021-01-09.

[4] "UHJ decoder for E4L," https://github.com/EnvelopSound/EnvelopForLive/blob/master/patchers/decoder/e4l.decoder.uhj.maxpat, accessed: 2021-11-08.

[5] M. Gast, *Wireless Networks: The Definitive Guide*. O'Reilly, 2005.

[6] Julius O. Smith III, *Mathematics of the Discrete Fourier Transform (DFT), with Audio Applications - Second Edition*. W3K Publishing, 2007.

[7] C. B. Gold, A. V. Oppenheim, "Theory and implementation of the discrete Hilbert Transform," 1969.

A Web Playground for Ciaramella

Paolo Marrone
Orastron s.r.l.
Università di Udine
paolo.marrone@orastron.com

ABSTRACT

Ciaramella is a domain specific programming language for audio DSP. It experiments a fully declarative syntax and the Synchronous Data Flow computational model, featuring high modularity and composability. We implemented a web playground for coding in Ciaramella, compiling and instantly generating working web audio plugin prototypes. In order to obtain that, its compiler, Zampogna, which is written in JavaScript, has been embedded in a web page, and it has been augmented for the production of JavaScript code. We developed a simple graphical user interface targeting both domain experts and newcomers. Finally, this work lays the foundations for future WebAssembly support as target code.

1. INTRODUCTION

Ciaramella is a recent domain specific text based programming language for the description of audio DSP systems featuring an high-level of abstraction. It supports experimental approaches like declarative syntax and Synchronous Data Flow model (SDF) [1], whose combination is new to the audio programming languages field. Indeed, the existing audio programming languages, which share a similar domain with Ciaramella, do adopt different paradigms. Faust [2], for example, is purely functional; Kronos [3] is defined as a *metalanguage* and experiments interactive graphical programming and just-in-time compilation; Max Gen [4] focuses on visual programming.

In this work we explore the possibility to deploy programs written in Ciaramella as audio applications directly on the web.

The possibilities and popularity of web programming have been continuously increasing during the years. Concerning audio, W3C developed the Web Audio API [5], which is now a standard adopted and implemented by all major web browsers. Web audio provides a routing graph mechanism capable of connecting independent audio nodes, multiplexing signals and handling parameters. Some common audio nodes are provided by default, but custom ones can be user-defined. In particular, the DSP algorithm can be written in pure JavaScript or in WebAssembly (using JavaScript as glue code). The WebAssembly option allows the audio developer to use any programming language (e.g. C, C++, Rust) as far as it is supported by a compiler that targets WebAssembly. Also, it is particularly appealing for its near-native performances.

When it comes to high-level languages such as Ciaramella, Faust or Kronos, a desired web application for the end user is an online playground which permits to code, compile, instantly listen to the produced plugin, and eventually export the compiled files. Faust, Kronos and Soul, for example, have developed their own playgrounds [1] [2] [3].

Faust's compiler is written in C++ and it has been compiled to WebAssembly via the Emscripten toolchain [4] to make it embeddable in a web page[6]. Its playground comes with an auto-completion text editor, compilation, on-the-fly interactive graphical plugin creation and export features. Kronos provides a similar environment, but it is focused on a graphical composition system rather than a text editor.

The paper is organized as follows. Section 2 recalls the main features of Ciaramella and its compiler, Zampogna. Section 3 describes the new Ciaramella web playground.

2. CIARAMELLA AND ZAMPOGNA

Ciaramella is an audio specific programming language aimed at describing DSP systems. Its declarative nature and SDF-based model result in a light and unconstrained syntax which guarantees an high-level of modularity and flexibility. Even more interestingly, it permits the description of delay-free feedbacks between subsystems. This makes it possible for known models that are hard to code such as wave digital filters (WDF) to be programmed easily [7].

As simplicity and minimalism are Ciaramella's design goals, a small set of programming abstractions is provided which reflect the SDF concepts and make the description of DSP systems possible. In particular, the fundamental components are:

- the **block**, that encapsulates an operation and it may be simple (atomic) or composed by other blocks (composite);
- input or output **ports**, attached to a block and representing connection endpoints;
- the **connection**, that defines a directed flow of data (e.g. a signal) between two ports.

Copyright: ©2022 Paolo Marrone et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

[1] https://faustide.grame.fr/
[2] https://kronoslang.io/veneer/
[3] https://soul.dev/lab/
[4] https://emscripten.org/

The language supports some built-in atomic blocks for the basic arithmetic operations and for the delay operation, i.e. the **unit delay** block, which is necessary to produce computable loops.

2.1 Syntax

Since Ciaramella is designed to cope with high-level manipulation of data streams, every variable or expression refers to entire flows rather than instantaneous values. For example the

```
a = b + c
```

code specifies that, for every temporal sample n, $a_n = b_n + c_n$.

A Ciaramella program is made of a list of statements, which can be assignments or composite block definitions. An assignment is in the simple form:

```
id1, id2 = expression
```

and it also acts as a declaration of the `id1` and `id2` variables. Indeed, in Ciaramella a variable can be declared and assigned only once and the order of the statements is not relevant.

The syntax for composite block definition is:

```
y1, y2 = composite_block_id (x1, x2, vol) {
        tmp = x1 + x2
        y1 = y2 * vol
        y2 = tmp / 2
}
```

It defines a reusable block (like a function), referable as `composite_block_id`, which has 3 input ports (`x1`, `x2` and `vol`) and 2 output ones (`y1`, `y2`). The body contains the assignments and expressions for calculating the values of the output. The output ports must be always assigned and can be used as normal variables, while the input ones are considered external and cannot be re-assigned.

A composite block can be used following a C-like function call syntax. This operation is called "composite block instantiation". For example:

```
t1, t2 = composite_block_id (in1, in2, 0.3)
```

instantiates a `composite_block_id` and redirects its outputs to the newly created `t1` and `t2` variables.

The built-in unitary delay block is of particular interest for the creation of computable feedbacks. The following example implements an iteration counter and shows its syntax:

```
counter = delay1(counter) + 1
@counter = 0
```

The `delay1(counter)` expression returns the value of `counter` at the $n-1$th iteration, while the special `@` symbol sets the initial value of `counter`, which is needed for the first iteration.

A typical Ciaramella program consists in a list of composite block definitions and assignments of constants. Each composite block has its local scope, while constants are globally scoped and can be accessed anywhere within the code.

The following example is the implementation of three composite blocks representing some trivial low-pass filtering functions.

Listing 1. Low-pass filters in Ciaramella

```
b = 0.1
y = lp (x) {
        y_z1 = delay1(y)
        y = y_z1 + b * (x - y_z1)
        @y = 0
}
y = lp3 (x) {
        y = lp(lp(lp(x)))
}

yL, yR = lp3_stereo (xL, xR, volume) {
        yL = lp3(xL) * volume
        yR = lp3(xR) * volume
}
```

For a more advanced example that implements a WDF based low-pass filter, check [7].

2.2 Compiler

We implemented a compiler for Ciaramella called Zampogna, written in JavaScript. Its entire codebase consists of about 1500 lines of code. Zampogna is able to produce C++ with VST2 wrapper, MATLAB, and lately JavaScript with Web Audio wrapper.

The compilation process consists of several steps, from the Ciaramella code parsing to the target code generation. A distinctive step is the production of an intermediate graph representation (IG) starting from the abstract syntax tree (AST). The IG reflects the SDF process network which in turn corresponds to the DSP system described by the Ciaramella code. The IG gets flattened, in order to reduce the system to atomic blocks, then optimized, and finally scheduled for sequential execution. The SDF formalism allows the scheduling to be accomplished statically (at compile time) [8], which is fundamental to maintain high performances of the DSP algorithms. The last phase of the compiler is the production of target code. To accomplish this, it makes use of the doT templating library [9], which enhances modularity allowing to easily add support for further target languages.

Zampogna can be used via the `zampogna-cli.js` command line interface. For instance, to compile the previous low-pass filters example (listing 1), assuming the code is saved in the `lp.crm` file, the following command can be used:

```
zampogna-cli.js -i lp3_stereo -t cpp
-c volume lp.crm
```

The `-i lp3_stereo` option specifies that `lp3_stereo` composite block acts as *initial block*, that is like the typical C *main* function. The `-t cpp` option selects C++ as target language and `-c volume` communicates to the compiler that the `volume` input is an user input control; the other inputs are assumed to be normal audio data flows (audio rate).

We highlight that writing the compiler purely in JavaScript with almost no external dependency, makes it natively embeddable in any web page.

3. WEB IDE/PLAYGROUND

We developed a fully functional Ciaramella web playground. In order to accomplish it, we first added JavaScript as target language to Zampogna. Then, we generated the web compatibile version of Zampogna, `zampogna-web.js`, through *browserify* [5] . A `zampogna-cli.js` is also available as command line interface to be used via Node.js [6] . Finally, we created the HTML web page that acts as an all-in-one text editor, compiler and execution environment.

Both the language and the web playground are at an experimental stage. Zampogna source code is available at https://github.com/paolomarrone/Zampogna, while the web playground can be tested at https://ciaramella.dev/.

3.1 User Interface

The code text editing section features tabbing to allow working on more projects simultaneously. Alongside the text editor, input areas for the necessary compiler options are present (mainly: initial block and control inputs information). A read-only console box has been placed below the "Compile" button to retrieve the compilation messages from Zampogna. This section is showed in Figure 1.

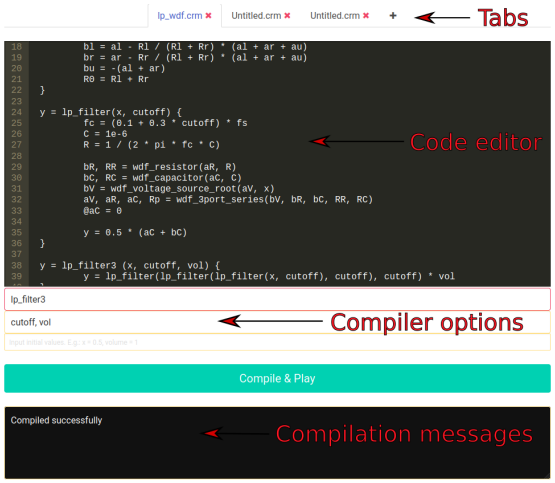


Figure 1. Ciaramella web editor

In case of valid input and successful compilation, the produced JavaScript code serves as input for the execution section. This section consists in a small Web Audio nodes network composed by an audio source node, the custom node that incorporates the output of the compilation, and an output node that represents the environment audio system. As source node, it is possible to choose between the system microphone (*MediaStreamAudioSourceNode*), or a file from both server or client side (*AudioBufferSourceNode*). The custom node is, instead, an *AudioWorkletNode*.

[5] `https://browserify.org/`
[6] `https://nodejs.org/en/`

The execution section dynamically generates the graphical elements for the manipulation of the user controls starting from the compilation options specified by the user. They are simple HTML input elements of range type. Figure 2 shows an example of this section.

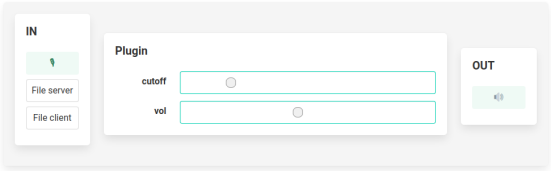


Figure 2. Ciaramella web player

The last section of the page is devoted to the plugin export feature (Figure 3), allowing to build plugins in C++, MATLAB, and JavaScript language. It is possible, then, to view and download them.

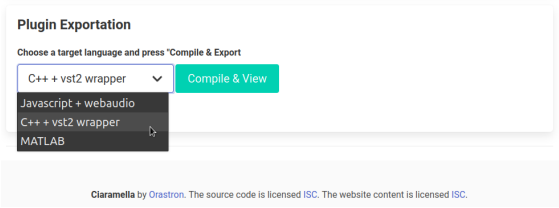


Figure 3. Ciaramella web plugin exportation

We did use the Bulma CSS library [7] for the styling. The whole webpage, including custom styling and scripts, amounts to about 400 lines of code.

3.2 Implementation Details

The output of the compilation for the execution section is a JavaScript program in form of string (`procJsStr`). It contains the code of the audio processing algorithm and an *AudioWorkletProcessor* wrapper.

Typically, `procJsStr` would be saved in a file (commonly `processor.js`) for later inclusion as a module to the audioWorklet. This procedure is not viable in our case because `procJsStr` is generated dynamically at client-side and we want the server that hosts the webpage to be completely unaware of what the user does. Our solution relies on the Blob JavaScript object [8] , that is more abstract than the File one. Indeed, the File interface is based on Blob. If `ctx` is the *AudioContext* and the custom *AudioWorkletProcessor* is registered as `PluginProc`, the code to create the custom node is:

```
let scriptUrl = URL.createObjectURL(new Blob(
[procJsStr], {type: "text/javascript"}));

await ctx.audioWorklet.addModule(scriptUrl);

let customNode = new AudioWorkletNode(ctx,
"PluginProc", { outputChannelCount: [1] });
```

[7] `https://bulma.io/`
[8] `https://developer.mozilla.org/en-US/docs/Web/API/Blob`

In a realistic scenario, the compilation is repeated many times, and the "module substitution" problem arises: the *AudioWorklet* provides the *addModule()* method but no method for deleting or editing of the existing ones. Consequently, modules accumulate over and over and, for long sessions, they may encumber the memory. There are at least two ways to handle this problem.

The first one is the destruction and recreation of the whole *AudioContext* object: this is the simplest and drastic solution.

In the second solution, instead, there exist only one context and one module, but the latter has to provide a mechanism to change its inner audio processing part. In particular, the output of the compilation needs to be sent via the *AudioWorkletNode* port messaging system to the *AudioWorkletProcessor*, which then uses the built-in JavaScript *eval()* function [9] to substitute and evaluate its own inner code. This solution, unfortunately, has some efficiency problems because *eval* constrains the JavaScript interpreters or engines from performing ahead of time optimizations [10]. More specifically, modern browsers tend to convert JavaScript scripts to machine code, losing variable naming informations; consequently the not-optimized code passed to *eval* causes long and expensive lookups within the machine code. Moreover, it leads to security risks based on the ease for bad actors to execute arbitrary code.

Ultimately, since *eval* is generally discouraged [10] and since we require the code to be as fast as possible, we chose to implement the first solution.

4. CONCLUSIONS

We developed a fully functional web playground for Ciaramella. It is now possible for both domain experts and newcomers to try out Ciaramella on the fly without being limited by installation and cumbersome testing processes. We showed how JavaScript was a convenient choice for the compiler code: we obtained both small codebase size and native execution on the web. Furthermore, this work is a platform for future development: the next target language we aim to support is WebAssembly, which will be used for the audio processing part of the web audio plugins. It will grant near-native performances and make plugins usable in real applications rather than for prototyping purposes.

5. REFERENCES

[1] E. A. Lee and D. G. Messerschmitt, "Synchronous data flow," *Proceedings of the IEEE*, vol. 75, no. 9, pp. 1235–1245, 1987.

[2] Y. Orlarey, D. Fober, and S. Letz, "Syntactical and semantical aspects of FAUST," *Soft Computing*, vol. 8, no. 9, pp. 623–632, 2004.

[3] V. Norilo, "Kronos: a declarative metaprogramming language for digital signal processing," *Computer Music Journal*, vol. 39, no. 4, pp. 30–48, 2015.

[4] "Max/gen." `https://docs.cycling74.com/max8/vignettes/gen_topic`. Accessed: 2022-09-14.

[5] "Web audio API." `https://www.w3.org/TR/webaudio/`. Accessed: 2022-05-27.

[6] S. Letz, Y. Orlarey, and D. Fober, "Compiling faust audio DSP code to webassembly," in *Web Audio Conference*, 2017.

[7] P. Marrone, S. D'Angelo, F. Fontana, G. Costagliola, and G. Puppis, "Ciaramella: a synchronous data flow programming language for audio DSP," in *Sound and Music Computing Conference*, SMC, 2022.

[8] E. A. Lee and D. G. Messerschmitt, "Static scheduling of synchronous data flow programs for digital signal processing," *IEEE Transactions on computers*, vol. 100, no. 1, pp. 24–35, 1987.

[9] "dot - the fastest + concise javascript template engine for node.js and browsers.." `https://olado.github.io/`. Accessed: 2022-05-27.

[10] S. H. Jensen, P. A. Jonsson, and A. Møller, "Remedying the eval that men do," in *Proceedings of the 2012 International Symposium on Software Testing and Analysis*, pp. 34–44, 2012.

[9] `https://tc39.es/ecma262/multipage/global-object.html#sec-eval-x`

[10] `https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Global_Objects/eval`

CIMIL VR MUSIC SYSTEM: UN AMBIENTE MULTIUTENTE CONDIVISO PER COMPORRE E SUONARE INSIEME NELLA REALTÀ VIRTUALE

Matteo Tomasetti
Creative, Intelligent & Multisensory Interactions Laboratory
Università di Trento
matteo.tomasetti@unitn.it

ABSTRACT

In questo articolo vengono descritte le domande di ricerca e gli obiettivi alla base del mio dottorato, il quale è incentrato principalmente nel campo scientifico ed artistico delle pratiche musicali all'interno di ambienti di realtà virtuale, e più in generale di "realtà estesa", un campo identificato con il nome di "Musical XR". L'obiettivo principale è quello di sviluppare un sistema musicale multiutente in un ambiente condiviso di realtà virtuale dove si può comporre e suonare musica elettronica. Inoltre, vengono delineate le nuove linee guida di progettazione che andrò a sviluppare in merito al design progettuale di alcune categorie di controllo di parametri musicali in ambienti di realtà virtuale che siano proficue per i ricercatori e le ricercatrici di "Musical XR". Vengono, inoltre, descritti le metodologie e le ricerche che andrò a compiere nei prossimi anni e il piano di lavoro che andrò ad adottare per la finalizzazione e implementazione dei vari progetti.

1. INTRODUZIONE

Nell'ultimo decennio si è verificata una notevole espansione della ricerca artistica e scientifica nello scenario della musica nella realtà estesa (XR). Una nuova era di esperienze immersive legate alla realtà virtuale (VR) e accessibili al grande pubblico è emersa grazie ai moderni sviluppi nella tecnologia dei display a bassa persistenza e alle più recenti tecniche di tracciamento spaziale [1].
Questa espansione ha portato allo sviluppo di una nuova area di ricerca denominata "Musical XR", che racchiude tutti i processi musicali sviluppati per ambienti di realtà virtuale, mista e aumentata. Le interazioni musicali legate alla tradizione musicale sono state completamente stravolte dalle tecnologie della realtà estesa, le quali hanno permesso agli esecutori e al pubblico di interagire musicalmente con oggetti, agenti e ambienti virtuali, rappresentando un fondamentale cambio di paradigma [2].
In effetti, la VR è diventata un mezzo popolare per i compositori e i sound designer, poiché i contenuti multimediali possono essere distribuiti rapidamente e gli ambienti virtuali possono essere sviluppati con una forte attenzione verso il suono immersivo [3].

Copyright: ©2022 Matteo Tomasetti et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

All'interno di un'esperienza di VR, i suoni e la musica possono essere utilizzati come elementi di sottofondo o come elementi interattivi e controllabili dall'utente [2]. Inoltre, i recenti sviluppi della tecnologia VR hanno riportato l'interesse per le tecniche legate all'audio spaziale, come l'Ambisonics [4] e l'audio binaurale [5], i quali offrono vantaggi significativi per la progettazione dell'audio immersivo [6]. Atherton et al. [7] sottolineano che il design dell'audio nelle esperienze VR dovrebbe essere generato dinamicamente, contenere forme di interazione ed essere immersivo.
La VR è un'esperienza immersiva e il senso di immersione è tratto in parte dalla fenomenologia dell'*embodiment* [8] [9], una teoria legata allo studio della percezione e incentrata sulla nozione di un corpo unificato che fissa l'esperienza umana. Oltre alle esperienze e alle teorie legate alla percezione, ciò che emerge dalle ultime ricerche e dalla letteratura è un'immagine di "Musical XR" come un campo artistico e scientifico piuttosto frammentato a livello tecnologico, dove artisti, compositori e ricercatori si sono spesso concentrati sullo sviluppo di sistemi su misura e di singole tecnologie [2]. Con il miglioramento delle attuali tecnologie XR in termini di risoluzione, velocità e capacità di creare spazi in grado di coinvolgere i nostri sistemi percettivi umani ad un livello vicino a quello della nostra realtà fisica, il compito dei designers, creatori e musicisti di comprendere e presentare la percezione in modo significativo e credibile è diventato sempre più complesso.
L'immersione e la presenza non sono semplici conquiste tecnologiche basate sulla densità di pixel e sul frame-rate, ma richiedono invece nuove direzioni di ricerca interdisciplinari che abbracciano psicologia cognitiva, percezione, arte, design, scienza e filosofia [1].
Gli obiettivi principali della mia ricerca riguardano lo sviluppo del *CIMIL VR Music System*, un nuovo sistema musicale multiutente condiviso con diverse tipologie di interazione audiovisiva e la realizzazione di nuove linee guida di progettazione (*design guidelines*) relative al controllo di parametri musicali in ambienti di "Musical XR" con il fine di supportare il design e il progredire di esperienze musicali, performative e compositive multiutente all'interno di ambienti condivisi di VR.
La sigla CIMIL si riferisce al *Creative, Intelligent and Multisensory Interactions Laboratory* dell'università di Trento all'interno del quale sto svolgendo il mio PhD.
Nei sistemi VR multiutente condivisi è la rete Internet a collegare le persone collocate in aree geografiche diverse nello stesso ambiente VR. Questo concetto rientra nell'a-

rea di ricerca delle *Networked Music Performance* (NMP), in cui diversi musicisti connessi alla rete si esibiscono insieme da diverse località geografiche. Nelle NMPs, il flusso audio viene trasmesso attraverso Internet, il che può introdurre ritardi di trasmissione che possono influire negativamente sulla sincronizzazione in tempo reale nella performance dei musicisti [10]. Tuttavia, l'efficienza della rete e della trasmissione dei dati odierna ha superato questi problemi di latenza audio e video, consentendo di realizzare NMP in VR, dove ogni persona connessa all'ambiente virtuale può essere rappresentata sotto forma di avatar. Loveridge [11] ha esaminato lo stato dell'arte per quanto riguarda la relazione tra VR e NMP, riferendo che la VR potrebbe consentire nuovi tipi di collaborazione condivisa in ambienti musicali. Tuttavia, ad oggi, sono state condotte poche ricerche in questo campo.
Le domande di ricerca del mio dottorato sono le seguenti:

1. Come definire le linee guida di progettazione in modo da semplificare il processo di composizione e performance musicale all'interno di ambienti VR?
2. Quali strumenti musicali possono essere considerati utili e da sviluppare per i compositori di musica elettronica nella VR?
3. Che tipologie di interazioni multiutente con il suono si possono creare in ambienti condivisi di VR?

Per rispondere a queste domande ho iniziato ad indagare le molteplici questioni interdisciplinari legate alla realtà virtuale, come la percezione, l'interazione uomo-macchina, le tecnologie disponibili, la dimensione tecno-somatica [12], l'*embodiment* [8] [9], le *affordances* [13], avvalendomi di numerosi apporti scientifici, tra cui quello dell'esperienza utente (UX), della progettazione di interfacce grafiche tridimensionali (3DUI), dei *Virtual Reality Musical Instruments* (VRMIs) [14] e delle *New Interfaces for Musical Expression* (NIME), con un focus specifico verso l'audio tridimensionale e la possibilità di creare strumenti specifici per la gestione interattiva della spazializzazione immersiva del suono. Alcune delle mie ricerche saranno finalizzate a stilare una o più *design guideline* funzionali ad aiutare ricercatori e ricercatrici di "Musical XR" nel design progettuale di alcune categorie di controllo di parametri musicali e di pratiche compositive e performative in ambienti di VR, spiegati più dettagliatamente nella sezione 3.
Sono consapevole che questa proposta di ricerca è piuttosto ambiziosa, ed è per questo che tengo a chiarire che è mia intenzione condurre solo una ricerca di base in un'area di ricerca così inesplorata. Precisamente, mi propongo di implementare un sistema VR di base, che getti le fondamenta per possibili estensioni future (che, idealmente, potrebbero trasformarsi in applicazioni commerciali). Con lo sviluppo di questo sistema, indagherò le opportunità e le sfide relative agli ambienti virtuali specifici per la composizione e la performance musicale, poiché è piuttosto difficile creare da solo un sistema musicale VR multiutente condiviso completo entro tre anni. Tuttavia, il livello di base dell'implementazione tecnica del *CIMIL VR Music System* proposto mi permetterà di concentrarmi sugli aspetti scientifici sopra indicati.

2. METODOLOGIA DELLA RICERCA

Il percorso triennale del mio dottorato e la mia proposta di ricerca segue un ordine crescente di complessità. Ciò significa che il mio PhD mira ad affrontare diversi progetti e diverse sfide legate ai sistemi di composizione ed esecuzione musicale sviluppati in VR. Ogni progetto e la relativa sfida sussegue il successivo in ordine di complessità, con un approccio di pianificazione e progettazione dal basso verso l'alto (*bottom-up*).
I diversi software e i relativi linguaggi di programmazione che sto utilizzando per la realizzazione dei diversi progetti sono Unity [1] , utilizzato come motore grafico multipiattaforma per il rendering in tempo reale, Pure Data [2] , Max [3] , Faust [4] e Wwise [5] utilizzati per la sintesi e gestione del suono nell'ambiente virtuale e nei VRMIs, Atmoky Ears [6] , IEM Plug-in Suite [7] utilizzati per la gestione della tecnologia Ambisonics e dell'audio binaurale, e Photon Fusion [8] , il quale è formato da un insieme di prodotti software, infrastrutture e architetture di rete che lavorano con Unity e che vengono utilizzate per la gestione di ambienti virtuali condivisi e multiutente.
Per la realizzazione degli scopi del mio dottorato descritti precedentemente mi avvarrò di diverse metodologie: nell'arco del primo anno mi sto basando principalmente sui metodi sviluppati e impiegati dai diversi ricercatori che lavorano nel campo scientifico e di design della UX e dell'interfaccia utente (UI).
I metodi che sto utilizzando si possono dividere in due categorie, entrambe basate su studi con gli utenti (*user studies*):

1. Ricerca qualitativa, realizzata attraverso tecniche di osservazione dei comportamenti degli utenti di sistemi XR coinvolti negli esperimenti, analisi tematiche delle domande aperte, questionari e altre metodologie di feedback;
2. Ricerca quantitativa, conseguita tramite l'analisi statistica dei segnali di input registrati da sistemi di "Musical XR" nella fase degli esperimenti con gli utenti.

Innanzitutto, per raggiungere gli obiettivi e rispondere alle domande di ricerca del mio dottorato sto indagando le esigenze e le necessità degli utenti tramite delle interviste.
I diversi partecipanti coinvolti nelle mie ricerche sulla base delle mie conoscenze personali sono principalmente musicisti professionisti (con minimo dieci anni di esperienza musicale) e che lavorano attivamente nel campo della VR. I metodi qualitativi che sto utilizzando per la validazione dei miei progetti includono test di usabilità e di qualità dell'esperienza, valutati e analizzati principalmente tramite questionari, quali "System Usability Scale" (SUS),

[1] https://unity.com/
[2] https://puredata.info/
[3] https://cycling74.com/products/max/
[4] https://faust.grame.fr/
[5] https://www.audiokinetic.com/en/
[6] https://atmoky.com/
[7] https://plugins.iem.at/
[8] https://www.photonengine.com/en-US/Fusion

"Quality of Experience" (QoE), "User Safisfaction Evaluation Questionnarie" (USEQ), "Questionnaire for User Interaction Satisfaction" (QUIS), "ITC-Sense of Presence Inventory" (ITC-SOPI), "Creativity Support Index" (CSI), "NASA Task Load Index" (NASA TLX), "Motion Sickness Assessment Questionnaire" (MSAQ).
I metodi quantitativi che sto utilizzando, invece, riguardano principalmente l'analisi, per ciascun utente coinvolto negli *user studies*, dei file di log dei sistemi di "Musical XR" creati. Questi, per esempio, possono riportare il tempo che viene impiegato per compiere una determinata interazione musicale, gli errori che vengono commessi nel controllare determinati parametri musicali nell'ambiente virtuale e la precisione nel controllo dei parametri e delle interazioni.
Per la realizzazione del *CIMIL VR Music System* nell'arco del secondo e terzo anno, invece, la metodologia principale che andrò ad utilizzare è quella del Co-Design [15] [16] [17]. Ciò significa che organizzerò diversi laboratori (*workshops*) con compositori ed esecutori esperti che lavorano nel settore della musica elettronica e della VR, nei quali sarà necessario utilizzare strumenti di prototipazione rapida precedentemente sviluppati nei software descritti, come VRMIs, tecniche di interazione e controllo musicale in VR, e diverse 3DUIs. Durante i diversi laboratori lascerò che i musicisti coinvolti scelgano, selezionino e discutano i migliori strumenti utili all'implementazione del *CIMIL VR Music System*. Quest'ultimo, quindi, presenterà le esigenze dei compositori (*users' needs*) raccolte. Con questa metodologia, ogni utente che verrà coinvolto nei diversi laboratori contribuirà al processo di progettazione e implementazione del sistema VR.
Queste metodologie di Co-Design, *user studies* e di ricerca qualitativa e quantitativa, unite assieme agli approcci di analisi attitudinali e comportamentali, mi permettono di validare scientificamente i miei progetti e di migliorare l'esperienza utente e gli strumenti che sto sviluppando in base ai feedback degli utenti coinvolti.

3. PIANO DI LAVORO

Per raggiungere gli obiettivi e rispondere alle domande di ricerca del mio dottorato triennale (2021-2024) ho suddiviso il piano di lavoro in questo modo:

- **Primo anno (2021/2022)**:
  - Analisi approfondita della letteratura (*state-of-the-art*) riguardante il campo di ricerca della "Musical XR";
  - Indagine, tramite interviste a compositori di musica elettronica e sound designer che lavorano con la VR, sulle esigenze e sulle necessità di cui hanno bisogno per comporre e suonare (live performance) in un ambiente di realtà virtuale, adottando un approccio al processo di progettazione incentrato completamente sull'utente;
  - Primo progetto di ricerca finalizzato alle NMPs, cioè all'atto di suonare a distanza, e che riguarda il confronto in cuffia tra un sistema binaurale con *head tracker* e un sistema stereofonico. Il design sperimentale e lo *user studies* sono stati condotti durante i mesi di aprile e maggio 2022;
  - Secondo progetto di ricerca che riguarda la progettazione di una 3DUI per la valutazione delle migliori tecniche di input per il controllo continuo dei parametri musicali in VR, quali *slider*, *knob*, *pad XY*. Le tecniche di input che verranno confrontate riguardano *hand tracking*, *raycasting* con il movimento fisico del controller per sistemi VR come Oculus Quest controller, *raycasting* con movimento del *thumbstick* e *grabbing*.
    Ogni tecnica di input utilizzata per il controllo degli *slider*, *knob* e *pad XY* implementati nella 3DUI controlla, a sua volta, dei parametri musicali (frequenza di taglio di un filtro passa basso, tonalità e ampiezza) di un sintetizzatore sottrattivo digitale.
    Lo *user studies* verrà effettuato nel mese di novembre 2022 e verranno coinvolti 20 musicisti. Questo studio è finalizzato ad ottenere, per tutti i ricercatori e le ricercatrici di "Musical XR", una linea guida di progettazione utile per il controllo musicale di oggetti con valori continui in VR.
- **Secondo anno (2022/2023)**:
  - Terzo progetto di ricerca che riguarda l'implementazione del *CIMIL VR Music System*, un sistema compositivo e performativo per la musica elettronica con 3DUI in VR.
    Questo sistema includerà le esigenze e le necessità dei musicisti emerse dalle interviste condotte nel primo anno e le differenti implementazioni sviluppate nei diversi laboratori condotti nel secondo anno, come descritto nella sezione 2.
    Per sistema si intende una serie di strumenti e VRMIs personalizzati controllabili dall'utente. Questo sistema beneficierà, inoltre, della precedente ricerca sull'audio spaziale nelle NMPs (primo progetto) e sulle tecniche di input (secondo progetto), in quanto verrà utilizzato il risultato migliore dello *user studies* riguardo l'audio spaziale del primo progetto e la tecnica di input che risulterà migliore per il controllo continuo dei parametri musicali in VR.
    Per validare scientificamente questo sistema verranno condotti degli *user studies*.
- **Terzo anno (2023/2024)**:
  - Quarto progetto che riguarda l'implementazione e l'espansione del *CIMIL VR Music System* in un contesto multiutente e con un ambiente condiviso di VR.
    Anche per questo progetto verrà utilizzato l'approccio metodologico del Co-Design, con l'in-

clusione di 10 utenti *stakeholders* esperti di composizione musicale elettronica e di XR che parteciperanno attivamente alla fase di *Design Thinking* e prototipazione/estensione del *CIMIL VR Music System* verso un nuovo sistema multiutente e con un ambiente condiviso di VR. La finalizzazione del nuovo sistema permetterà di eseguire NMPs in VR.
Inoltre, durante questo quarto progetto verranno condotti degli studi riguardanti le diverse possibilità di interazione multiutente con il suono in ambiente condiviso di VR;

– Finalizzazione della tesi di dottorato incentrata sugli argomenti di ricerca indagati nel triennio 2021-2024.

4. CONCLUSIONI

In questo articolo ho presentato le principali attività di ricerca che ho iniziato a svolgere nel mio dottorato (2021-2024), il quale indaga le pratiche compositive e performative di musica elettronica all'interno del campo scientifico e artistico denominato "Musical XR".
Inoltre ho presentato le domande di ricerca, gli obiettivi, le varie metodologie delle quali mi avvarrò e, infine, il piano di lavoro. In generale, il goal del mio dottorato si può riassumere in:

- Creazione del *CIMIL VR Music System*, un ambiente multiutente condiviso di realtà virtuale per performance e composizione musicale;
- Stilare delle linee guida di progettazione utili e relative alle tecniche di input utilizzate nei sistemi XR per il controllo di determinati parametri musicali e compositivi.

5. REFERENCES

[1] A. Çamcı and R. Hamilton, "Audio-first vr: New perspectives on musical experiences in virtual environments," *Journal of New Music Research*, vol. 49, no. 1, pp. 1–7, 2020.

[2] L. Turchet, R. Hamilton, and A. Çamci, "Music in extended realities," *IEEE Access*, vol. 9, pp. 15810–15832, 2021.

[3] Z. Buckley and K. Carlson, "Towards a framework for composition design for music-led virtual reality experiences," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1497–1499, IEEE, 2019.

[4] F. Zotter and M. Frank, *Ambisonics: A practical 3D audio theory for recording, studio production, sound reinforcement, and virtual reality*. Springer Nature, 2019.

[5] J.-M. Jôt, S. Wardle, and V. Larcher, "Approaches to binaural synthesis," in *Audio Engineering Society Convention 105*, Audio Engineering Society, 1998.

[6] A. Çamcı, "Some considerations on creativity support for vr audio," in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1500–1502, IEEE, 2019.

[7] J. Atherton and G. Wang, "Doing vs. being: A philosophy of design for artful vr," *Journal of New Music Research*, vol. 49, no. 1, pp. 35–59, 2020.

[8] M. Merleau-Ponty, *The visible and the invisible: Followed by working notes*. Northwestern University Press, 1968.

[9] G. Corness, "The musical experience through the lens of embodiment," *Leonardo Music Journal*, vol. 18, pp. 21–24, 2008.

[10] C. Rottondi, C. Chafe, C. Allocchio, and A. Sarti, "An overview on networked music performance technologies," *IEEE Access*, vol. 4, pp. 8823–8843, 2016.

[11] B. Loveridge, "Networked music performance in virtual reality: current perspectives," *Journal of Network Music and Arts*, vol. 2, no. 1, p. 2, 2020.

[12] G. Paine, "Interaction as material: The techno-somatic dimension," *Organised Sound*, vol. 20, no. 1, pp. 82–89, 2015.

[13] J. J. Gibson, *The ecological approach to visual perception: classic edition*. Psychology Press, 2014.

[14] S. Serafin, C. Erkut, J. Kojs, N. C. Nilsson, and R. Nordahl, "Virtual reality musical instruments: State of the art, design principles, and future directions," *Computer Music Journal*, vol. 40, no. 3, pp. 22–40, 2016.

[15] D. Schuler and A. Namioka, *Participatory design: Principles and practices*. CRC Press, 1993.

[16] C. Spinuzzi, "The methodology of participatory design," *Technical communication*, vol. 52, no. 2, pp. 163–174, 2005.

[17] M. J. Muller, *Participatory design: the third space in HCI*. CRC press, 2007.

RICONOSCIMENTO AUTOMATICO DI CATENE DI EFFETTI PER CHITARRA ELETTRICA

Michele Rossi
Creative, Intelligent & Multisensory Interactions Laboratory
Dipartimento di Ingegneria e Scienza dell'Informazione, Università di Trento
michele.rossi-2@studenti.unitn.it

ABSTRACT

L'obiettivo di questo progetto consiste nell'individuare se una traccia di chitarra è stata processata con una catena di effetti, indicandone la tipologia. Per ottenere tale risultato, sono state impiegate tecniche basate sul Deep Learning (DL) facendo utilizzo, in particolare, di una Convolutional Neural Network (CNN). L'architettura di tale rete non è stata definita a priori, ma ha subito un processo di ottimizzazione dei suoi iperparametri. Questo è avvenuto grazie all'implementazione di algoritmi evolutivi, ed in particolare algoritmi genetici (GAs) [1], in grado di far mutare una rappresentazione compatta che codifica le principali caratteristiche della rete neurale stessa. Una volta ottenuta la neural network più performante, la sua implementazione è stata utilizzata per la fase di classificazione, con l'impiego di un nuovo dataset appositamente creato per questo progetto, ottenendo un'accuratezza del 92% sul test set.

1. INTRODUZIONE

Riconoscere quali effetti sono stati applicati ad un suono di chitarra elettrica costituisce un campo di ricerca che presenta, ad oggi, pochi contributi nella letteratura. Eppure, un'accurata ricostruzione della catena di effetti presenti su una traccia di chitarra può essere utile in diversi ambiti inerenti al Music Information Retrieval (MIR). Alcuni campi che possono beneficiare di tali informazioni sono, ad esempio, la classificazione di generi musicali, i sistemi di trascrizione automatica, o, più in generale, la produzione musicale.

Il riconoscimento degli effetti utilizzati per processare il suono della chitarra elettrica è un ambito che ha visto da poco l'utilizzo delle reti neurali ed in particolare del deep learning. Precedenti lavori (ad esempio [2]) sono focalizzati principalmente sull'estrazione e la selezione di features, successivamente impiegate nell'algoritmo di classificazione (ad esempio Support vector machines (SVM)). Questo approccio risulta concettualmente distante da quello tipicamente impiegato nel deep learning, dove la fase di feature extraction, se presente, è tipicamente meno onerosa. Alcuni ambiti del MIR hanno già sperimentato, con ottimi risultati, l'impiego di reti neurali deep, focalizzandosi, ad esempio, sulle convolutional neural network per la classificazione di generi musicali [3, 4, 5, 6].

Tra gli studi maggiormente collegati a quello proposto troviamo [2]. Uno dei principali limiti di questo articolo consiste nell'utilizzo delle SVM per la fase di classificazione, che necessitano di una fase molto onerosa di feature extraction e feature selecion. Inoltre, il progetto si basa su un dataset elaborato dallo stesso autore in cui è presente, per ogni file audio, una singola nota di chitarra (o un insieme di note suonate in contemporanea) per la durata di due secondi. Tale dataset risulta molto semplificato rispetto ad un caso reale, che consiste tipicamente in un vero e proprio assolo o fraseggio di chitarra, riducendo così la generalità dell'algoritmo ottenuto. Inoltre nel dataset sono presenti alcune tracce di chitarra provenienti dallo stesso suono (non processato) che viene elaborato con il medesimo effetto, ma con una diversa intensità. Questo aspetto semplifica la corretta classificazione di tali elementi qualora finissero nel test set, in quanto una versione quasi identica di essi si troverà con un alta probabilità presente nel training set.

Un altro articolo che affronta il problema di classificazione dell'effetto utilizzato (e anche di stima dei suoi parametri) è [7]. In questo progetto viene implementato un algoritmo basato su una convolutional neural network (CNN) ed è utilizzata la matrice associata al mel power-spectrogram [8] di ogni traccia audio come input della rete neurale. Anche in questo caso il dataset di partenza (con i segnali di chitarra non processati) coincide con quello presente in [2], con le stesse problematiche di mancanza di generalità precedentemente accennate. In questo lavoro, inoltre, non tutti i principali plugin di chitarra vengono presi in considerazione, ma soltanto gli effetti non-lineari, quali distorto, overdrive e fuzz.

Un articolo con l'obiettivo di classificare e stimare i parametri dei principali effetti di chitarra è [9]. Anche in questo caso viene utilizzato come dataset di partenza (come tracce di chitarra non processate) quello proposto da [2]. Inoltre, anche questo progetto presenta una fase di feature extraction e di feature selection molto onerosa, necessaria per effettuare, successivamente, la classificazione tramite l'impiego delle SVM. Per quanto riguarda la stima dei parametri, vengono impiegate semplici reti neurali che differiscono molto tra loro in base al particolare effetto considerato.

Infine, M. Stein [2] presenta un ulteriore articolo incen-

Copyright: ©2022 Michele Rossi et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

trato sulla stima di catene di effetti. Anche in questo caso è presente un elaborato preprocessing così come l'utilizzo delle SVM. Per questo progetto vengono utilizzate solamente catene di elaborazione con un massimo di tre effetti. Inoltre sono preimpostati sia l'ordine dei plugin (prima effetti non lineari, poi di modulazione ed infine di ambiente), sia la scelta di quale plugin possa occupare una particolare posizione. La scelta è limitata infatti a due tipi di effetto per ogni posizione disponibile.

La metodologia proposta in questo report è volta a superare alcune delle principali limitazioni presenti negli articoli sopra citati. In particolare è stato assemblato un nuovo dataset che presenta dei file audio contendenti del brevi assoli di chitarra che meglio generalizzano le tipiche tracce di chitarra reali. Inoltre, l'approccio utilizzato si basa su una convolutional neural network e sull'utilizzo del mel-spectrogram come input della rete. Un'ulteriore fase consiste nell'ottimizzazione dell'architettura delle CNN sfruttando algoritmi evolutivi. Questo approccio è stato esplorato in letteratura principalmente nell'ambito di classificazione di immagini [2, 10, 11, 12, 13, 14].

Nell'ambito di questo progetto è stato creato un dataset sia per la classificazione di singoli effetti, sia per la classificazione di catene di effetti. In questo report verrà elaborato solamente quest'ultimo problema, poiché rappresenta il reale contributo di questo progetto.

2. DATASET

Il nuovo dataset assemblato è composto da 400 tracce di chitarra della durata di 11.5 secondi ciascuna (10 secondi di audio effettivo). Le registrazioni sono state effettuate con quattro diverse chitarre, considerando due posizioni di pickup ciascuna e due tecniche della mano destra (plettro o dita) per ogni posizione. Il dataset consiste in brevi assoli principalmente basati sulla scala pentatonica (la scala statisticamente più utilizzata dai chitarristi) suonati considerando tutte le tonalità disponibili. La registrazione è stata ottenuta in presa diretta con l'impiego della Audient ID22 come scheda audio e di Reaper come Digital Audio Workstation (DAW). Il sampling rate scelto è di 44100 [Hz], mentre la bit depth 16 [bit]. Gli effetti con cui sono state processate le tracce sono stati selezionati seguendo la classificazione di [2] considerando quindi feedback delay, slapback delay, riverbero, chorus, flanger, phaser, tremolo, vibrato, distorto e overdrive. La scelta dei plugin è ricaduta su quelli nativamente presenti in Pedalboard (libreria prodotta da Spotify per l'elaborazione di segnali audio con Python) e su alcuni vst gratuiti, per permettere una semplice riproducibilità delle tracce processate. Un ulteriore contributo di questo progetto consiste nella creazione di un dataset dove le tracce sono state processate con più effetti in cascata. In particolare, ad ogni registrazione di chitarra sono stati applicati fino ad un massimo di 5 effetti. Per rendere il problema gestibile considerando l'enorme numero di combinazioni, il nuovo dataset è stato processato con una catena di effetti che non variano i propri parametri, ma mantengono una configurazione di default. L'applicazione di tali effetti, se presenti, rispetta inoltre un ordine prestabilito, scelta che ricalca il reale utilizzo dei pedali di chitarra

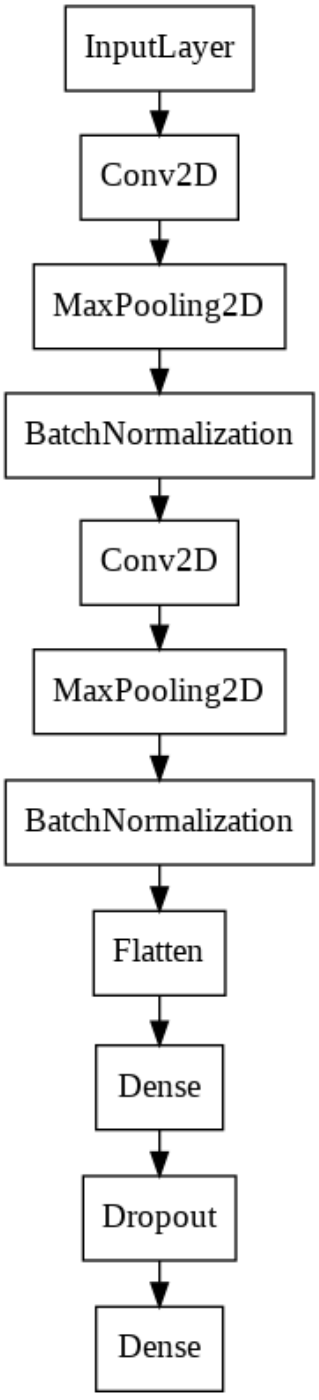


Figure 1. Esempio di convolutional neural network ottenuta in seguito all'impiego di algoritmi evolutivi

dove ogni effetto deve essere collocato (o risulta comunque più opportuno) in una precisa posizione della catena. La dimensione finale del dataset per la multilabel classification, considerando tutte le possibili combinazioni, è pari a 12800, considerando che ad ognuna delle 400 tracce sono state applicate in totale 32 combinazioni di effetti.

3. METODOLOGIA

In questa sezione verranno presentate le principali fasi che compongono l'algoritmo sviluppato nell'ambito di questo progetto.

3.1 Preprocessing

Durante questa fase ogni traccia di chitarra viene divisa in 5 segmenti della durata di 2 secondi ciascuno. Ogni segmento viene poi convertito nel corrispondente mel-spectrogram, che consiste in una rappresentazione compatta del segnale audio, in forma di matrice, che verrà utilizzata come input della rete neurale.

3.2 Architettura della CNN

La CNN ha una struttura variabile che dipende dalla scelta dei suoi iperparametri. In generale, è possibile dividere il modello in due blocchi principali, uno contenente la parte di convolutional layer e l'altra composta dai dense layer. La prima si compone di un numero variabile di convolutional layer, ognuno dei quali è seguito da un maxpooling layer. La seconda è composta da alcuni dense layer seguiti da un output layer composto da 5 neuroni, che corrispondono alle cinque classi presenti nella formulazione del problema come multi-label classification. I due blocchi sono collegato da un flatten layer che riceve in ingresso un input multidimensionale e produce in output un vettore, che verrà poi processato dai successivi layer. Nel modello sono implementati sia la Batch-Normalization sia il dropout (presente in coda ad ogni dense layer).

Come funzione di attivazione è stata scelta la Rectified Linear Unit (ReLU) considerando le buone performance che riesce a produrre associata ad una CNN e la sua capacità di ovviare al problema del vanishig gradient. Solo nell'ultimo layer (output) è stata implementata una sigmoide, in quanto ogni neurone, in questo ultimo step, restituisce come valore di uscita la probabilità che un particolare effetto sia presente nella traccia audio considerata. Come loss function è stata scelta la binary crossentropy, mentre come ottimizzatore è stato impostato il noto Adam.

3.3 Algoritmi Genetici

Gli algoritmi genetici sono utilizzati per permettere l'esplorazione dei parametri della CNN. Ogni modello neurale è codificato all'interno di un vettore di 33 elementi, che rappresenta quindi un modello compatto della rete neurale stessa. Di seguito sono riportati alcuni del parametri, specificando il range entro il quale possono variare.

- Numero di convolution layer presenti (min: 2, max: 5)
- dimensione orizzontale dei kernel del primo conv. layer (min: 2, max: 7)
- dimensione verticale dei kernel del primo conv. layer (min: 2, max: 7)
- dimensione orizzontale del filtro associato al primo maxpool layer (min: 1, max: 2)
- dimensione verticale del filtro associato al primo maxpool layer (min: 1, max: 2)
- (gli stessi parametri ma associati algi altri conv. laye se presenti)
- numero di dense layer (min: 1, max: 3)
- numero di neuroni del primo dense layer (min: 2, max: 64)
- probabilità associata al dropout che segue il primo dense layer
- (parametri associati agli altri dense layer se presenti)

Le dimensioni dei kernel associati ai convolutional layer e ai dense layers sono ottimizzate considerando separatamente la dimensione orizzontale e verticale. Tale scelta deriva dal fatto che le righe e le colonne dello spettrogramma presentano due significati molto diversi. La dimensione orizzontale rappresenta l'informazione associata al tempo (quantizzato in frames), mentre in verticale troviamo l'informazione relativa a quanto ciascuna frequenza è presente in quello specifico frame.

Gli algoritmi genetici sono stati implementati da zero in Python. Di seguito le principali funzioni con una breve descrizione del loro funzionamento.

- Fitness function. Questa funzione fornisce il punteggio associato ad un particolare vettore di parametri. Crea una CNN, la allena su un dataset ridotto e restituisce un valore di accuratezza sul test set.
- Initialize Population. Funzione che crea randomicamente la prima popolazione di vettori, ognuno dei quali rappresenta la codifica compatta di una CNN.
- Population evaluation. Questa funzione salva il valore della fitness associata ad ogni vettore in un python dictionary. Inoltre calcola e salva la probabilità che un particolare individuo (vettore) venga selezionato per la generazione successiva.
- Selection. Seleziona due vettori basandosi sul roulette-weel principle. Utilizza le probabilità calcolate dalla funzione precedente.
- Crossover. Performa un single-point crossover fra due vettori.
- Mutation. Questa funzione ha un diverso comportamento in base al gene considerato. Per quanto riguarda i valori associati al numero di layer, le dimensioni dei kernel o dei filtri utilizza un generatore di numeri casuali (entro i limiti di quel particolare

parametro). Per qunto riguarda gli altri parametri, la funzione modifica leggermente il valore precedente (aggiungendo del rumore). Solo nel caso della probabilità di dropout la funzione lavora con numeri reali e non interi.

4. RISULTATI

L'algoritmo è stato allenato con l'utilizzo di Google Colab Pro+, con una GPU Tesla P100. La popolazione iniziale (del GAs) è stata impostata a 10, con un numero di generazioni pari a 30. Le probabilità di crossover e mutazione utilizzate sono rispettivamente 0.8 e 0.2. L'algoritmo ha ottenuto un'accuratezza del 92% nel test set, considerando come corretta predizione il caso in cui l'intero set di effetti presente su un particolare sample fosse correttamente riconosciuto. Il test set corrisponde al 20% dell'intero set.

Un ulteriore risultato riguarda il secondo setup dell'esperimento, in cui il trainig e l'evaluation delle performance sono state eseguite su chitarre differenti. In questo caso, infatti, tre chitarre sono servite per allenare la CNN, mentre una quarta per la valutazione. In questo caso è stata raggiunta un'accuratezza pari al 70%. Questo risultato evidenza come la rete neurale abbia saputo generalizzare rispetto al tipo di chitarra coinvolto, ottenendo un'accuratezza elevata nonostante la maggiore difficoltà del task a cui era sottoposta.

5. LIMITI

L'approccio proposto presenta alcuni limiti. Ad esempio, per quanto riguarda la fase di creazione del dataset, gli effetti sono stati applicati senza far variare i loro parametri. Inoltre, considerando dei casi reali, per ogni tipologia di effetto da applicare alla chitarra sono presenti diversi pedali (e quindi plugin) e per ognuno di essi diverse case costruttrici. Non è quindi possibile una stima precisa e completa di ogni plugin esistente da parte dell' algoritmo. Il progetto proposto ha anche il limite di lavorare con simulazioni dei pedali di chitarra, quando invece la maggior parte delle canzoni attualmente disponibili è stata registrata sfruttando catene di effetti analogici.

Ulteriori fattori quali l'utilizzo di diversi amplificatori per chitarra, diversi microfoni e differenti preamplificatori della scheda audio utilizzata per registrare, fanno sì che il suono finale della chitarra (nelle registrazioni reali) venga ulteriormente influenzato. La categorizzazione proposta in questo progetto consiste quindi in un'approssimazione che non tiene conto di molti altri fattori che influiscono sul suono finale della chitarra processata.

6. REFERENCES

[1] S. Forrest, "Genetic algorithms," *ACM Computing Surveys (CSUR)*, vol. 28, no. 1, pp. 77–80, 1996.

[2] M. Stein, J. Abeßer, C. Dittmar, and G. Schuller, "Automatic detection of audio effects in guitar and bass recordings," in *Audio Engineering Society Convention 128*, Audio Engineering Society, 2010.

[3] W. Zhang, W. Lei, X. Xu, and X. Xing, "Improved music genre classification with convolutional neural networks.," in *Interspeech*, pp. 3304–3308, 2016.

[4] M. Dong, "Convolutional neural network achieves human-level accuracy in music genre classification," *arXiv preprint arXiv:1802.09697*, 2018.

[5] Y. M. Costa, L. S. Oliveira, and C. N. Silla Jr, "An evaluation of convolutional neural networks for music classification using spectrograms," *Applied soft computing*, vol. 52, pp. 28–38, 2017.

[6] R. L. Aguiar, Y. M. Costa, and C. N. Silla, "Exploring data augmentation to improve music genre classification with convnets," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, 2018.

[7] M. Comunità, D. Stowell, and J. D. Reiss, "Guitar effects recognition and parameter estimation with convolutional neural networks," *arXiv preprint arXiv:2012.03216*, 2020.

[8] K. Choi, G. Fazekas, K. Cho, and M. Sandler, "A tutorial on deep learning for music information retrieval," *arXiv preprint arXiv:1709.04396*, 2017.

[9] H. Jürgens, R. Hinrichs, and J. Ostermann, "Recognizing guitar effects and their parameter settings," in *Proceedings of the 23rd International Conference on Digital Audio Effects (DAFx2020)*, 2020.

[10] Y. Sun, B. Xue, M. Zhang, and G. G. Yen, "Evolving deep convolutional neural networks for image classification," *IEEE Transactions on Evolutionary Computation*, vol. 24, no. 2, pp. 394–407, 2019.

[11] B. Wang, Y. Sun, B. Xue, and M. Zhang, "Evolving deep convolutional neural networks by variable-length particle swarm optimization for image classification," in *2018 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1–8, IEEE, 2018.

[12] L. Xie and A. Yuille, "Genetic cnn," in *Proceedings of the IEEE international conference on computer vision*, pp. 1379–1388, 2017.

[13] B. Wang, Y. Sun, B. Xue, and M. Zhang, "A hybrid differential evolution approach to designing deep convolutional neural networks for image classification," in *Australasian Joint Conference on Artificial Intelligence*, pp. 237–250, Springer, 2018.

[14] E. Real, S. Moore, A. Selle, S. Saxena, Y. L. Suematsu, J. Tan, Q. V. Le, and A. Kurakin, "Large-scale evolution of image classifiers," in *International Conference on Machine Learning*, pp. 2902–2911, PMLR, 2017.

On making feedback instruments lively - a case study

Alberto Rizzo
University of Sussex
a.rizzo@sussex.ac.uk

ABSTRACT

The recursive circulation of sound from speakers to microphones has been extensively used as a method to design a new aesthetic for music compositions from the '60s, and it's still an actual research topic nowadays[1], especially in the *NIME* [1] community. Feedback systems that involve and live of only air and electricity can be configured as musical ecosystems for their ability to imitate nature in its manner of operation and because they can become autonomous entities[3][4]. On the other hand, feedback instruments usually require a haptic trigger for their operation and provide an augmentation of a traditional musical instrument[5]. At the same time, several feedback systems involve the coupling of the performer with the environment and a non-haptic interface. Despite these sensory differences, a feedback musician exposes himself to the same issues: balancing, negotiating control and un-control and finding new pockets of sonic magic. While conventional methods of machine listening are commonly used in the NIME community, I present novel research with the aim to make a feedback instrument more lively using *AI* [2], exploiting the *Feedback Shadow* instrument as a case study. The instrument mentioned above uses ultrasonic waves to control gesture and composition matter, taking advantage of some physical properties of sound in the air. Due to its complex nature, the system is contextualised in the optic of cybernetic language, leading to a better comprehension of the system and its behaviour. The creative output of the research exposed in *Shadow Pattern* shows the mutated relationship between human-machine, using ML as a tool to maintain homeostasis creating at the same time more complex sonic transitions and keeping a physical and satisfactory engagement with the performer. The visuals reflect this attempt to engage ML-mediated synaesthesia between sound and motion through the process of visual listening[6] [3] .

[1] NIME: The International Conference on New Interfaces for Musical Expression gathers researchers and musicians from all over the world to share their knowledge and late-breaking work on new musical interface design[2]

[2] AI: acronym for Artificial Intelligence

[3] Alberto Novello defines this process as "a deeper way of understanding sound through its visualisation through light" in a synchronous way using Lissajous patterns

Copyright: ©2022 First author et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 Unported, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

1. INTRODUCTION

A minimal definition of feedback, despite spans different fields in academia, takes into account the configuration of a system, provided with input and output, in which some kind of transformation occurs as the output is fed back to the input after a delay[7][8]. When approaching this phenomenon applied to musical instruments, it is beneficial to provide a historical background to contextualise the discourse. From the '60s, the use of systems based on feedback gained popularity due to the success of cybernetics[8][9], and experimentation emerged, from devices implementing self-regulation mechanisms based on feedback loops to the investigation of the Larsen [4] phenomenon. While microphones and loudspeakers are designed to remain invisible and sonically transparent[1], composers like Stockhausen, Tudor and Lucier, among others, started to implement acoustic feedback in their composition practice, exploiting systems of transducers as sonic sculptures and instruments themselves and experimenting interactions with materials[11], space[12] and movement [13]. Lucier's configuration for *Bird and Person Dyning (1975)*, in particular, shows a different approach to the player mode interaction. The performer moving around the performing space wearing microphones in a feedback loop with an electronic bird and a couple of loudspeakers shows a unique composition environment in which the act of moving the head shapes the sound in the room and is informed by listening. While no haptic gestures are employed, the performer's physical presence in the space and the proprioceptive perception[14] is essential to tame the sound ecosystem. In the attempt to classify this particular kind of instrument with these characteristics, it has been helpful to define a new category of *open-air feedback instruments*, with this definition exploits some similarities with the open-air controller[15] of a *DMI*. [5]

The main characteristic of these systems, as opposed to feedback augmented instruments like the *self-resonating vibrotactile feedback instruments*[16], is the lack of the tactile feedback loop, Additionally, they usually do not derive from an existing instrument of the classical organology tradition since they mostly lack a resonant body.

[4] Larsen: Søren Absalon Larsen (April 5, 1871 – January 2, 1957) was a Danish physicist who worked in the field of electroacoustics and is best known for giving his name to the Larsen effect[10], also called audio feedback

[5] DMI: acronym for digital musical instrument

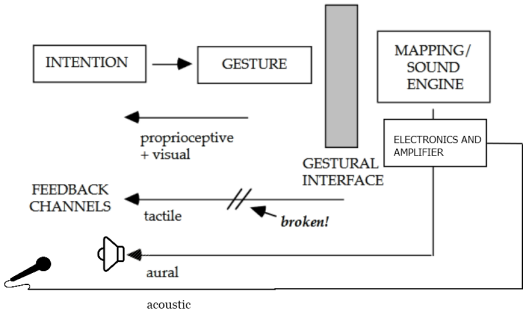


Figure 1: Abstract representation of the open-air feedback instrument

In this particular context, I argue that feedback instruments have no clear mapping relationship with sonic parameters, but they have an intrinsic mechanism whose complexity is challenging to break down, and at the same time, it is often something not desirable for a feedback musician.

2. CASE STUDY: FEEDBACK SHADOW INSTRUMENT

Feedback Shadow is an *open-air feedback instrument* characterised by the recursive circulation of sound streams constrained to frequencies higher than the range of human hearing bounced into the physical environment. The name of this instrument refers to the performer's bodily influence in producing an acoustic shadow [6] attenuating the sound wave propagation in the air. The physical performing interface is composed of two pairs of ultrasonic microphones and speakers in which the re-circulation of the ultrasonic streams is generated and amplified digitally through the SuperCollider [7] software environment that also is responsible for the signal processing [8] . The extreme directionality of these particular sound waves produce very localised streams that bounce everywhere in the environment and can be used with gestural movements to redirect and feed the feedback loop increasingly.

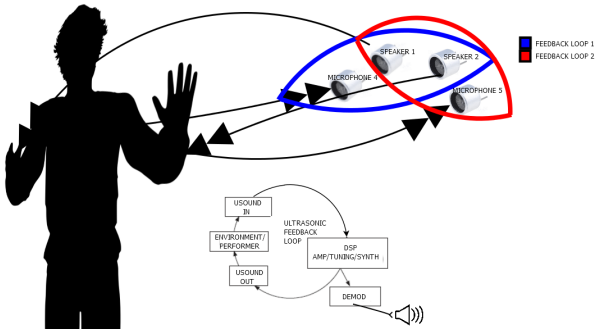


Figure 2: Performer interaction with Feedback Shadow

[6] Acoustic shadow: is an area through which sound waves fail to propagate, due to topographical obstructions or disruption of the waves via phenomena such as wind currents, buildings, or the human body[17]

[7] SuperCollider: is a platform for audio synthesis and algorithmic composition, used by musicians, artists, and researchers working with sound[18]

[8] The signal processing of Feedback Shadow includes demodulation, creative tuning and creation of the feedback network with its equalisation, distortion and dynamics processes

The so-called ultrasound waves inherit some sound properties that are used in other electronic music instruments like the Theremin[19]. *Audibility* is made possible through the *heterodyning* [9] process managed in software via the ring modulation of the ultrasonic microphones inputs with a reference 40kHz cycle wave emitted by a couple of ultrasonic speakers.

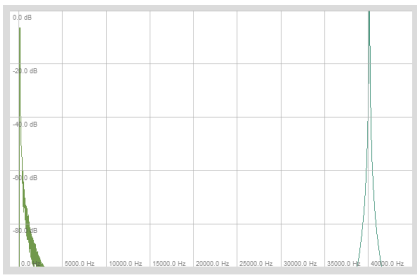


Figure 3: distortion products

Additionally, gestures around the speakers produce an audible frequency shift due to the doppler effect [10] The first documented use of the combined methods in a performing practice is for Lucinda Childs *Vehicle* performance, presented for the 1966 *9 Evenings: Theatre and engineering* event [11] . It is important to mention the fundamental input in the discovery of the creative use of ultrasounds by Ronald Kuivila, an American composer that pioneered the use of ultrasound and built a stratified experience over 40 years of practice and agency that can be seen in his last solo piece *Listening to the air*[22]. John Driscoll is another important composer who re-configured the concept of sonic sculpture acquired in his experience with the Tudor's Rainforest ensemble[23] creating an inaudible feedback loop shaped by tidy motions in the air. In Feedback Shadow the noise produced boosting the opportunely tuned microphone provides the essential sound source which is constantly played, re-injected in the loop and demodulated in the audible domain while the ultrasounds waves are mediated by the air, the environment and the performer gestures.

I would argue that composing with ultrasound in a feedback network means considering multiple aspects of the musician's relationship with the acoustic shadow field and the environment in a man - ambient - machine ecosystemic design[24]. The performer assumes the role of a composer, crafting the streams of sound that interacts with the physical space. The space itself, in which audible music comes to life through the ether, is also the listener's space and becomes the interface in which objects, motion and gesture continuously shape the feedback loop.

[9] Heterodyning: is the process of multiplying a weak signal by a strong sinusoidal carrier, sometimes called the local oscillator, to shift the frequency of the signal in such a way that the information carried by the signal is preserved[20], allowing the demodulation of the signal in a lower range

[10] Doppler effect: is an alteration in the observed frequency of a sound due to motion of either the source or the observer

[11] The 70 kHz doppler sonar system developed by Peter Hirsch allowed the transposition in the audible range of the reflected signals resulting from the movement of three buckets inside the original ultrasonic beams[21]

3. FRAMING FEEDBACK SHADOW USING CYBERNETIC LANGUAGE

In feedback Shadow, as for his non-linearity behaviour, the complexity results in the impossibility to understand the behaviour of the instrument by isolating every single component. For this reason, it is beneficial to describe the behaviour of the instrument using the cybernetic language, focusing on asking what all the possible system behaviours are, but not necessarily investigating every single cause. In the feedback Shadow instrument, we can mainly observe a positive feedback configuration in which a trajectory of the system in a specific direction will cause a further shifting in the same spot[25], leading to a magnification of the effect caused by the stimuli[9]. We can think of this shifting between different points (defined as "attractors") as the system's reaction to a complex event. Considering the attractors as sonic areas of further exploration is a common practice and joy among feedback musicians, where the golden area consists of the threshold between a state and another one [12] Feedback Shadow issue of navigating freely into different sonic textures, like other similar systems, is the chance of falling into them[26], without any way to escape from the saturating feedback [13] , While rehearsing with Feedback Shadow, it is helpful to think about the different modes of interactions:

A Intrinsic reaction to physical stimuli in the air defined as *internal variables*

B Hardcoded mapped *control parameters* of the Supercollider sound mechanism managed through MIDI

Pitch and timbre space, in particular, may be altered in different ways:

A Gestural movement of the performer and far to near approach in the ultrasound field may lead to a momentary or increasing glissando to a new sonic area.

C Pinching the circuit with the hands connecting different transducers leads to external noise and areas of maximum and minimum saturation.

In all the situations, the system often reaches a bifurcation point, in which, considering the basic trajectory from state A to B, it is inherently impossible to determine in advance which direction change will take[27]. The absence of movement generally leads to a rapid transition that takes to a fixed attractor.

The variation of a hard-coded control parameter produces different sonic and timbral changes: The inaudible signal detected by the feedback microphone, in particular, is duplicated in the signal flow and can be tuned to precise frequencies through two ringing filters, allowing the creation of beat patterns. Due to the characteristic of the resonators to create "centers of gravity" in the spectrum, this cause an overall phase response change in the system and, consequently, in the feedback loop. The analysis of Feedback Shadow in cybernetic language ultimately lead to consider it as a system very open to energy [14] but closed to control[9]. After this consideration, two different, but connected goals have been set for this research:

A Adding some sonic complexity to the system.

B Facilitating the interaction of the performer with the instrument, interrogating, as the early feedback pioneer Elian Radigue did[28], on all the possibility of keeping the instrument in a playable zone.

In recent years, NIME community found different methodologies to achieve those results involving machine listening, the audio feature extracting and adaptive filtering in order to detect and force the feedback to resonate at a frequency that aligns with the natural resonances of the acoustic instrument augmented with the feedback itself[29][30]. Other research on feedback instruments leads to control directly the overall system gain:

- Kiefer, Overholt et al CoFlo algorithm uses a measure of complexity, to monitor the sound circulating in the feedback loop[26] avoiding strong basins of attraction.
- Other approaches, including the first iteration of Feedback Shadow and the augmented double bass[31], involve manual control of the microphones amplification to feed the network.

Differently from the *augmented double bass*, *Feedback Shadow* lacks of the haptic interaction, leading to vaster areas of fixed attraction and shorter transitions to such an extent that it is possible to navigate only two states. Investigating the use of machine learning has been considered to avoid unnecessary complexity to the system and for its potential flexibility in terms of possible sonic outcomes.

4. NN AS A CREATIVE TOOL

Machine Learning has been used by creative practitioners since at least the early 1990s[32] when artificial neural networks has been employed to control sound synthesis. Since then, researchers like Rebecca Fiebrink democratised[33] and brought ML tools for music to liberate the creative use of interfaces and sensors mapped to gestures to music or sound control parameters.

In the last few years, different artists collaborated in the development of these new ML tools integrating them into their practice, with the Fiebrink-Sonami collaboration giving new insights to the creative uses of machine learning while designing a *DMI*, viewing ML support into mapping as a way to explore new sonic territories instead building very accurate and tight models from training examples[34]. While ML learning has been scarcely used in conjunction with feedback instruments, *Fluid corpus manipulation* [15] project commissions led to some interesting examples of

[12] This area is usually characterised by a rich sonic harmonic texture.

[13] The saturating feedback is mainly characterised by a dominant resonant frequency, high amplitude, and fewer harmonics.

[14] the feedback loop can be stopped only wrapping the feedback microphone in different layers of very tight cloth

[15] `https://flucoma.org`

the use of NNs as a way to cleverly couple two systems affecting each other.[35] While experimenting with different solutions with NN's, I contextualised Ashby's design methodologies[36] coupling Feedback shadow with a neural network. Considering Feedback shadow as a black box, completely ignoring its content and dominated by another machine whose activities and values design the behaviour of the other machine led to an important insight: What if the dominating machine could learn how to affect the dominated in a very specific way on the basis of its spectral characteristic over time? This idea of coupling two systems with feedback lead to the use of the regression approach to dynamically shape the feedback shadow control parameters [16] based on the ultrasonic feature extraction of the instrument itself.

The practical implementation, after some attempts to find a suitable spectral descriptor for the quite unpredictable high-frequency signal, in the last iteration uses 20 bands of MFCC to monitor the speaker output in the range of 38 to 45 kHz. These data are the input of the NN, that allows the adaptive mapping of an opportunely scaled three-dimensional control space composed of linear, exponential gain and the volume level of a sine wave oscillator used as a pedal tone.

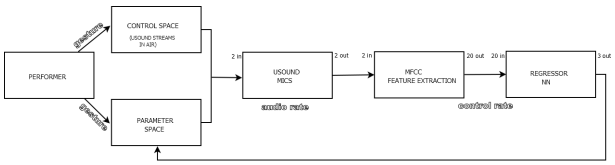


Figure 4: representation of the ML interaction with Feedback Shadow

When approaching this tool, initially a consideration has been made in order to use it to open a window of creative and sonic imagination and thinking about how this can change the way of interaction with the instrument and the composition structure. In that sense, trying to associate specific sonic textures to gestures and building very accurate models from the ultrasonic trainings it turned out to be extremely difficult and sonically not captivating. Additionally, thinking about a specific sound and training a model on top of it is non-relevant with the ultrasonic dataset since it is very difficult for the performer to reproduce afterwards the exact sonic situation without affecting manually the control parameters.

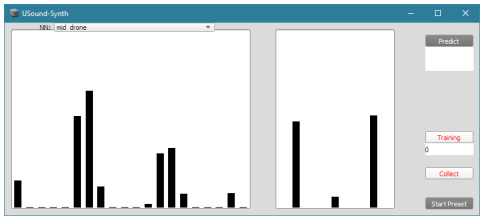


Figure 5: Neural network GUI

[16] The control parameters are usually changed by the performer through a MIDI device or through the selection of different presets

In the last iteration I considered each trained model a different machine configuration composed of 3 states and consequently being able to detect the following performer actions:

Ac1 player gestures in the air.

Ac2 interrupting the circuit.

Ac3 the variation of specific control parameters.

While training the neural network each of these actions has been considered as a way to interrupt the transition A-B and be able to find captivating sonic areas in the middle of those. Additionally, for different sonic presets the variation of specific MFCC bands to a specific gesture has been taken into account.

5. CREATIVE ROLE OF MACHINE LEARNING IN THE PERFORMANCE

In *Shadow Patterns* the intention was to explore and feel the airflow, and at the same time reproduce a very slow transition in feedback without having to deal with strong in-harmonic attractors [17] . The preparatory phase required five trained models that react to the above-mentioned musician's actions. The performer, after selecting the preset representing the starting point of the sonic exploration can create a trajectory between 3 above-mentioned different states of each model, changing continuously the mode of interaction with the instrument and eventually the model. In some way, this approach is similar of what is depicted by Sonami's use of ML for her *Spring Spyre*: Feedback Shadow it's not an instrument that you can play in a conventional sense, and at the same time, a specific set of gestures does not lead to a very precise sequence of sounds[34]. For this specific reason, dimensional reduction tries to navigate the three-dimensional spaces focused on the attempt to trace trajectories between the stored sonic states and anything in between. In this sense, the result of having created a trajectory between the three sonic states resulted in a slow pattern transition in which new auditory territories and data loops can be discovered every time. The represented A-B trajectory of states can be reconfigured in the flow diagram depicted in Figure 6. At the end of this practice-based research, I argue that the importance of ML in feedback instruments can be seen as its reconfiguration as a homeostasis [18] tool that mediates continuously the behaviour of the feedback loop, creating a trajectory and shifting between novel intermediate sonic states determined by different performer inputs. Future research will focus on the field of sound perception and spatial audio, finding possibilities of exploiting structured regressor neural networks and their capabilities of recognising trained patterns from a live and physical input to render audible sound spots in three-dimensional space by focusing ultrasonic waves shaping our soundscape[37].

[17] *Sonic Un-seen* exploits the usual models, re-purposing the pre-recorded ultrasonic materials. Some additional whale fixed media are employed in a spectral morphing with the Feedback Shadow sound output. The aim is to raise awareness and protect life in our waters from the destructive effects of oceanic artificial noise pollution.

[18] Homeostat: It is one of the first devices capable of adapting itself to the environment

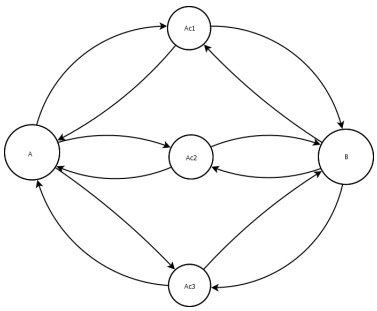


Figure 6: ML as a self-regulator between man and instrument

6. APPENDIX

Shadow Patterns - for ML-mediated feedback instrument (12:00). [19]
Sonic Un-seen - for ML-mediated feedback instrument and fixed media (08:00) [20] .

7. REFERENCES

[1] V. E. C., *Between Air and Electricity: Microphones and Loudspeakers as Musical Instruments.* New York: Bloomsbury Academic, 2017.

[2] "New Interfaces for Musical Expression." http://www.nime.org, 1998. Last accessed 2021-10-09.

[3] Y. Nakai, "How to imitate nature in her manner of operation: Between what john cage did and what he said he did," *Perspectives of new music*, vol. 52, pp. 141–160, 2014.

[4] A. Di Scipio, "'sound is the interface': from interactive to ecosystemic signal processing," *Organised Sound*, vol. 8, no. 3, p. 269–277, 2003.

[5] D. Sanfilippo and A. Valle, "Feedback systems: An analytical framework," *Computer music journal*, vol. 37, pp. 12–27, 2013.

[6] A. Novello, "Media archaeology-based visual music - accepted for publication on." https://oajournals.fupress.net/index.php/mt/, 2022.

[7] J. de Rosnay, "Feedback. principia cybernetica web (1997)." http://pespmc1.vub.ac.be/FEEDBACK.html. Last accessed 2021-10-21.

[8] D. Sanfilippo and A. Valle, "Feedback systems: An analytical framework," *Computer Music Journal*, vol. 37, no. 2, pp. 12–27, 2013.

[9] N. Wiener, *Cybernetics; or control and communication in the animal and the machine.* John Wiley., 1948.

[10] "Søren Absalon Larsen." https://dbpedia.org/page/Soren_Absalon_Larsen. Last accessed 2021-10-01.

[11] H. T., "Tudor, interview with teddy hultberg, düsseldorf, 17 – 18 may 1988." https://davidtudor.org/Articles/hultberg.html. Last accessed 2021-10-21.

[12] "Bird and person dyning (1975)." http://www.alvin-lucier-film.com/bird.html. Last accessed 2021-10-21.

[13] van Eck C., *Between Air and Electricity: Microphones and Loudspeakers as Musical Instruments.* New York: Bloomsbury Academic, 2017.

[14] C. Cadoz, "Le geste canal de communication homme/machine: la communication "instrumentale"," *Revue des Sciences et Technologies de l'Information - Série TSI : Technique et Science Informatiques*, vol. 13, no. 1, pp. 31–61, 1994.

[15] J. Rovan and V. Hayward, "Typology of tactile sounds and their synthesis in gesture-driven computer music performance," *Editions IRCAM*, 01 2000.

[16] A. Eldridge, C. Kiefer, D. Overholt, and H. Ulfarsson, "Self-resonating vibrotactile feedback instruments ——: Making, playing, conceptualising :," 06 2021.

[17] "Acoustic shadow." https://en.wikipedia.org/wiki?curid=5435674. Last accessed 2021-09-01.

[18] "Supercollider." https://supercollider.github.io/. Last accessed 2021-11-11.

[19] M. Kennedy and J. B. Kennedy, "theremin," 2007.

[20] H. E. Rowe, "Heterodyne principle.." https://doi.org/10.1036/1097-8542.316000, 2020. Last accessed 2021-11-01.

[21] "Motion music machine (sonar)." https://www.fondation-langlois.org/9evenings/e/tool/mmm.html. Last accessed 2021-10-21.

[22] D. Pozzi, "Listening to the air." https://www.researchcatalogue.net/view/386118/432464, 2019. Last accessed 2021-05-10.

[23] J. Driscoll, "Resonance: From the Architectural to the Microscopic," *Leonardo Music Journal*, vol. 22, pp. 25–33, 12 2012.

[24] C. Anderson, "Dynamic networks of sonic interactions: An interview with agostino di scipio," vol. 29, pp. 11–28, 2005.

[25] C. Boner and C. R. Boner, "Behavior of sound system response immediately below feedback," *Journal of The Audio Engineering Society*, vol. 14, pp. 200–203, 1966.

[26] C. Kiefer, D. Overholt, and A. Eldridge, "Shaping the behaviour of feedback instruments with complexity-controlled gain dynamics," September 2020.

[19] https://youtu.be/JjQq8ljP2qU
[20] https://youtu.be/9DCpnZ1Y77Q

[27] J. E. Earley, "Order out of chaos: Man's new dialogue with nature," *Process Studies*, vol. 14, no. 3, pp. 204–205, 1985.

[28] "K. molleson. interview: Eliane radigue." `http://katemolleson.com/interview-eliane-radigue/`. Last accessed 2021-10-21.

[29] J. Snyder, M. Mulshine, and R. Erramilli, "The feedback trombone: Controlling feedback in brass instruments," 06 2018.

[30] H. Ulfarsson and A. Melbye, "Sculpting the behaviour of the feedback-actuated augmented bass: Design strategies for subtle manipulations of string feedback using simple adaptive algorithms," 09 2020.

[31] T. P. Liontiris, "Low frequency feedback drones: A non-invasive augmentation of the double bass.," in *NIME*, pp. 340–341, 2018.

[32] M. Lee, A. Freed, and D. Wessel, "Real-time neural network processing of gestural and acoustic signals," in *Proceedings of the International Computer Music Conference*, pp. 277–277, INTERNATIONAL COMPUTER MUSIC ASSOCIATION, 1991.

[33] F. Bernardo, M. Zbyszyński, R. Fiebrink, and M. Grierson, "Interactive machine learning for end-user innovation," 03 2017.

[34] R. Fiebrink and L. Sonami, "Reflections on eight years of instrument creation with machine learning," 2020.

[35] "feedbackfeedforward." `https://www.flucoma.org/commissions/`. Last accessed 2021-11-20.

[36] W. R. Ashby, *An Introduction to Cybernetics*. John Wiley., 1956.

[37] C. Rajguru, M. Obrist, and G. Memoli, "Spatial soundscapes and virtual worlds: challenges and opportunities," *Frontiers in Psychology*, vol. 11, p. 569056, 2020.

