

Atti Proceedings



XXIV CIM – Colloquio di Informatica Musicale
24th CIM – Colloquium on Music Informatics

Atti della Conferenza
Conference Proceedings

Torino, 30 Settembre – 2 Ottobre 2024
Torino, September 30th – October 2nd 2024

Davide Andrea Mauro, Simone Spagnol and Andrea Valle, a cura di/*eds.*

AA.VV.

MEMORIE PROIETTIVE/PROJECTING MEMORIES

XXIV Colloquio di Informatica Musicale

24th Colloquium on Music Informatics

A cura di/eds.:

Davide Andrea Mauro, Simone Spagnol, Andrea Valle

©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: 9788890341373

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 - Dipartimento Arti e Design Industriale, Università IUAV di Venezia

In co-edizione con/co-published with:

CIRMA – Università degli Studi di Torino

AIMI – Associazione Informatica Musicale Italiana

<https://aimi-musica.org/>

BibTeX:

```
@proceedings{24CIMConf,  
  Editor = {Davide Andrea Mauro, Simone Spagnol and Andrea Valle},  
  Organization = {AIMI - Associazione Informatica Musicale Italiana},  
  Publisher = {DADI - Dip. Arti e Design Industriale. Università IUAV di Venezia},  
  Title = {Memorie proiettive/Projecting Memories. Atti del {XXIV} Colloquio di Informatica Musicale/  
  Proceedings of the {XXIV} Colloquium on Music Informatics},  
  Year = {2024}}
```

Copyright

Memorie proiettive/Projecting Memories. Atti del XXIV Colloquio di Informatica Musicale/Proceedings of the XXIV Colloquium on Music Informatics
© 2024 by AA.VV. è distribuita con Licenza CC BY-NC-ND 4.0

Copia della licenza è disponibile presso <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Memorie proiettive/Projecting Memories. Atti del XXIV Colloquio di Informatica Musicale/Proceedings of the XXIV Colloquium on Music Informatics

© 2024 by AA.VV. is licensed under CC BY-NC-ND 4.0

To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



Cover drawing: algorithmic radial arborescences generated with the Shoebot Python Library (<http://shoebot.net/>) by Andrea Valle
Typeset with ConTeXt (<http://wiki.contextgarden.net/>) by Andrea Valle

With the support of:



**UNIVERSITÀ
DI TORINO**



**CONSERVATORIO
STATALE DI MUSICA
GIUSEPPE VERDI
TORINO**

**TO)))
LISTEN
(((TO**

SMETI Scuola di Musica Elettronica
del Conservatorio
di Torino



ART IN MED
L'Arte nella divulgazione
delle Scienze Mediche



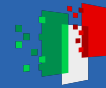
Finanziato
dall'Unione europea
NextGenerationEU



Ministero
dell'Università
e della Ricerca



Italiadomani
2022-2025



**UNIVERSITÀ
DI TORINO**

General chairs

Stefano Bassanese, Conservatorio di Torino
Andrea Valle, Università di Torino

Scientific chairs

Davide Andrea Mauro, Paderborn University
Simone Spagnol, Università IUAV di Venezia
Andrea Valle, Università di Torino

Music chairs

Andrea Agostini, Conservatorio di Torino
Stefano Bassanese, Conservatorio di Torino
Daniele Ghisi, Conservatorio di Torino

Program committee

Andrea Agostini, Conservatorio di Torino
Giacomo Albert, Università di Torino
Carlo Barbagallo, Conservatorio di Torino
Stefano Bassanese, Conservatorio di Torino
Francesco Bianchi, Conservatorio di Torino
Leonardo Gabrielli, Università Politecnica delle Marche
Daniele Ghisi, Conservatorio di Torino
Francesco Giomi, Tempo Reale
Luca Guidarini, Ensemble Collettivo 21
Davide Andrea Mauro, Paderborn University
Simone Spagnol, Università IUAV di Venezia
Anna Terzaroli, Conservatorio di Adria
Luca Turchet, Università di Trento
Andrea Valle, Università di Torino
Paolo Zavagna, Conservatorio di Venezia

Scientific Reviewers

Giacomo Albert, Università di Torino
Adriano Baratè, Università di Milano
Alessio Brutti, FBK - Fondazione Bruno Kessler
Luca Comanducci, Politecnico di Milano
Stefano D'Angelo, dangelo.audio
Stefano Delle Monache, IRCAM
Michele Ducceschi, Università di Bologna
Davide Fantini, Università di Milano
Vanessa Faschi, Università di Milano
Federico Fontana, Università di Udine
Leonardo Gabrielli, Università Politecnica delle Marche
Michele Geronazzo, Università di Padova
Riccardo Giampiccolo, Politecnico di Milano
Luca Andrea Ludovico, Università di Milano
Nicola Orio, Università di Padova
Giorgio Presti, Università di Milano
Niccolò Pretto, Libera Università di Bolzano
Alessandro Giuseppe Privitera, Università di Udine
Davide Rocchesso, Università di Milano
Antonio Rodà, Università di Padova
Cristina Rottondi, Politecnico di Torino
Sylviane Sapir, Conservatorio di Milano
Stefania Serafin, Aalborg University
Giuseppe Silvi, Conservatorio di Bari
Federico Simonetta, Gran Sasso Science Institute
Domenico Stefani, Università di Trento
Anna Terzaroli, Conservatorio di Adria
Marco Tiraboschi, Università di Milano
Gualtiero Volpe, Università di Genova
Stefano Zambon, MIND Music Labs
Victor Zappi, Northeastern University
Laura Zattra, IRCAM / Conservatorio di Rovigo e Vicenza
Paolo Zavagna, Conservatorio di Venezia

Music Reviewers

Alessandro Anatrini, Hochschule für Musik und Theater Hamburg
Nicoletta Andreuccetti, Independent
Massimo Avantaggiato, Conservatorio di Milano
Lorenzo Ballerini, Conservatorio di Trapani
Carlo Barbagallo, Conservatorio di Torino
Angelo Benedetti, Conservatorio di Perugia
Nicola Bernardini, Conservatorio di Roma
Francesco Bianchi, Conservatorio di Torino
Nicola Buso, Conservatorio di Trieste
Francesco Canavese, Tempo Reale
Diego Capocitti, Independent
Nicola Cappelletti, Conservatorio di Perugia
Maura Capuzzo, Conservatorio di Venezia
Nicola Casetta, Conservatorio di Pesaro
Daniela Cattivelli, Conservatorio di Avellino
Luigi Ceccarelli, Edison Studio
Antonino Chiamonte, Conservatorio di Frosinone
Walter Ciancusi, Conservatorio di Sassari
Fabio Cifariello Ciardi, Conservatorio di Perugia
Pasquale Citera, Conservatorio di Bologna
Giorgio Colombo Taccani, Conservatorio di Torino
Diego Dall'Osto, Conservatorio di Vicenza
Riccardo Dapelo, Conservatorio di Piacenza
Maria Cristina De Amicis, Conservatorio di L'Aquila
Fabio De Sanctis De Benedictis, ISSM "P. Mascagni" di Livorno
Stefano Delle Monache, IRCAM
Agostino Di Scipio, Conservatorio di L'Aquila
Roberto Doati, Conservatorio di Genova
Giulia Francavilla, Independent
Nicola Frattegiani, Conservatorio di Frosinone
Francesco Giomi, Tempo Reale
Luca Guidarini, Università di Pavia
Giorgio Klauer, Conservatorio di Trieste
Sergio Lanza, Conservatorio di Alessandria
Silvia Lanzalone, CRM, Roma
Ilaria Lemmo, Conservatorio di Torino
Giulia Lorusso, Independent
Marco Marinoni, Conservatorio di Latina
Matteo Marson, Conservatorio di Torino
Damiano Meacci, Conservatorio di Firenze
Alessandro Olla, TiConZero
Federico Ortica, Independent
Claudio Panariello, KTH, Stockholm
Mattia Parisse, Independent
Pietro Polotti, Conservatorio di Trieste
Alessandro Ratoci, Conservatorio di Mantova
Luca Richelli, Conservatorio di Vicenza
Silvia Rosani, Goldsmiths University of London
Dario Sanfilippo, Independent
Andrea Sarto, Independent
Francesco Scagliola, Conservatorio di Bari
Julian Scordato, Conservatorio di Padova
Carlo Siega, 'A. Bruckner' Privatuniversität, Linz/Conservatorio di Trieste
Giuseppe Silvi, Conservatorio di Bari
Anna Terzaroli, Conservatorio di Adria
Matteo Tundo, Independent
Gianluca Verlingieri, Conservatorio di Cuneo
Giovanni Verrando, Scuola Civica "C. Abbado" di Milano
Roberto Zanata, Conservatorio di Cesena
Paolo Zavagna, Conservatorio di Venezia

Table of Contents

| | | |
|--|-----|---|
| Davide Andrea Mauro, Simone Spagnol and Andrea Valle | vii | Prefazione/Preface |
| | ix | Special session: Polifonia del ricercare - Suono e musica nei progetti di ricerca |

KEYNOTES

| | | |
|---------------------------|-----|---|
| Pierre-Alexandre Tremblay | xv | All In: The wager of entangled musicking research(es) |
| Marinos Koutomichalis | xvi | Vectors of care & musical informatics. From sinewaves to AI: Tactics to synthesize audio with computers |

ABSTRACTS

| | |
|--|-------|
| Paper Abstracts | xviii |
| Abstract dei lavori musicali/Musical Works Abstracts | xxv |

SESSION 1 – PROJECTING MEMORIES

| | | |
|---|----|--|
| Fabio De Sanctis De Benedictis | 2 | Memoria come composizione, composizione come memoria: Strumenti algoritmici in Poisson Trio e Anamniseis |
| Stefano Catena and Enrico Dorigatti | 8 | Missing memories: Why we need to analyse spatialisation |
| Daniele Giuseppe Annese, Francesco Vitucci, Anthony Di Furia, Francesco Scagliola, and Giuseppe Silvi | 16 | Archeotopologie: Implementazione critica di memorie senza colore |
| Daniel Scorrane and Agostino Di Scipio | 24 | Gli 'sciami di glissando' in Diamorphoses. Ricostruzione mediante Digital Morphophone Environment |

SESSION 2 – COMPOSITIONAL PROCESSES

| | | |
|----------------------------------|----|---|
| Cristiano Bocci and Andrea Valle | 32 | Fully generalized Fibonacci series modulo n as music sequence generators |
| Paolo Paradiso | 40 | La sperimentazione vocale nell'epoca dell'intelligenza artificiale |
| Marco Matteo Markidis | 48 | Mediation process in a computer music interpretation: An ecosystemic approach |

POSTER SESSION

| | | |
|---------------------------------|----|---|
| Barbara Grosso and Andrea Valle | 56 | Sonification of Edoardo Sanguineti's poems |
| Giovanni Sparano | 63 | Software di supporto all'esecuzione musicale in MiraWeb: Due casi di studio |

| | | |
|--|----|---|
| Davide Commone | 70 | Manifold Voyager: Itinerario sperimentale tra forme musicali ricorsive |
| Giuseppe de Benedittis | 77 | Partiture di Voltaggio: Metodo di notazione per sintetizzatore modulare |
| Alexandrina Bargan, Maximiliano Romero, and Simone Spagnol | 83 | Studio di un'interfaccia musicale per pazienti con demenza di Alzheimer lieve |
| Daniel Scorrane | 91 | Riproposizione digitale di uno strumento pionieristico. Digital Morphophone Environment |
| Costantino Rizzuti and Fabrizio Rizzuti | 99 | Controllo remoto via OSC di un prototipo di synth basato su Raspberry Pi |

SESSION 3 – TOOLS AND PLATFORMS

| | | |
|--|-----|--|
| Riccardo Ancona | 107 | Una prospettiva critica sui dataset per la sintesi text-to-audio |
| Leonardo Gabrielli | 115 | Considerazioni su VCV Rack come piattaforma didattica per l'ingegnere e il musicista |
| Andrea Gulli, Federico Fontana, Hanna Järveläinen, and Michele Geronazzo | 123 | A mobile game app for adaptive assessment of pitch discrimination in children with different hearing ability |
| Alessandro Anatrini | 129 | Wavepilot: Framework multidimensionale per l'esplorazione dello spazio parametrico di strumenti digitali |

SESSION 4 – ANALYTICAL APPROACHES

| | | |
|--------------------------------------|-----|---|
| Angelica Speroni and Alessandro Bile | 136 | Il potere dell'intermedialità nella creazione della memoria artistica: il caso di enigma 33 di Lucia Romualdi |
| Claudia Rinaldi and Marco Pennese | 143 | A Pierre. Dell'azzurro silenzio, inquietum (1985). Analisi formale |

Prefazione/Preface

Per noi organizzatori, questo CIM è stato un lavoro lungo 4 anni.

Quando nel 2020 è scoppiata la pandemia, siamo stati costretti a cambiare sostanzialmente la conferenza CIM/SMC, spostando tutto online e abbandonando la sede di Torino. È stato quindi un grande piacere avere la conferenza CIM a Torino quest'anno. Festeggiamo quindi a 14 anni di distanza (invece che a 10) il precedente CIM torinese.

Quest'anno abbiamo ricevuto un totale di 90 proposte: 35 contributi scientifici e 55 contributi musicali. Delle 35 proposte scientifiche, 13 sono state accettate per la presentazione orale e 7 per la presentazione di poster, con un tasso di accettazione del 57%. Per la traccia musicale sono state accettate 19 proposte, con un tasso di accettazione del 35%.

XXIV CIM ha avuto l'aiuto di 93 revisori che hanno esaminato tutte le proposte per compilare un programma finale. Sulla base delle loro raccomandazioni, insieme ai comitati di programma e ai presidenti, sono state prese le decisioni finali organizzando il lavoro in 4 sessioni di paper, 1 sessione di poster e 3 concerti.

Per questo CIM abbiamo ospitato due relatori principali (Pierre Alexandre Tremblay e Marinos Koutsomichalis), entrambi capaci di oltrepassare i confini della ricerca scientifica e artistica.

Il programma musicale di quest'anno è stato particolarmente ricco, grazie alla collaborazione con la terza edizione del Festival dell'ascolto sperimentale To Listen To (<https://www.to-listen-to.it/>).

Il XXIV CIM si è concluso con la tradizionale attribuzione del premio Teresa Rampazzi e del premio Aldo Piccialli per i migliori contributi rispettivamente al programma musicale e al programma scientifico.

Premio Teresa Rampazzi – *Non è un compendio di etologia numerico-digitale* di Totalee (Andrea Laudante, Paolo Montella, Giuseppe Pisano), per la qualità sonora di un'opera musicale che riflette un pensiero elettroacustico articolato e denso di riferimenti a espressioni, culture e ambiti musicali multiformi.

Premio Aldo Piccialli – Daniel Scorrane, per la capacità di coniugare ricerca artistica, scientifica e storica in un lavoro di ampio respiro che ben rappresenta lo spirito del Colloquio di Informatica Musicale.

For us as organizers this CIM has been 4 years in the making. When during the 2020 the pandemic struck we were forced to substantially change the CIM/SMC Conference, moving everything online and abandoning the Turin location. It has been then a great pleasure to have the CIM conference in Turin this year. We are therefore celebrating 14 years later (instead of 10) after the previous CIM in Turin.

This year we received a total of 90 submissions: 35 scientific contributions, and 55 music contributions. Out of the 35 scientific submissions 13 have been accepted for oral presentation and 7 for poster presentation, with an acceptance rate of 57%. For the music track 19 submissions have been accepted, with an acceptance rate of 35%.

XXIV CIM had the help of 93 reviewers that examined all the submission in order to compile a final program. Based on their recommendations, together with the Program Committees and Chairs, final decisions were made organizing the work in 4 Paper Sessions, 1 Poster Session, and 3 Concerts.

For this CIM we featured two keynote speakers (Pierre Alexandre Tremblay and Marinos Koutsomichalis), both members crossing the boundaries of the scientific and artistic research.

The music program this year was particularly rich, thanks to the collaboration with the third edition of Festival dell'ascolto sperimentale To Listen To (<https://www.to-listen-to.it/>).

The XXIV CIM concluded with the traditional awarding of the Teresa Rampazzi award and the Aldo Piccialli award for the best contributions to the music program and the scientific program, respectively.

Teresa Rampazzi award – *Non è un compendio di etologia numerico-digitale* by Totaleee (Andrea Laudante, Paolo Montella, Giuseppe Pisano), for the sonic quality of a musical work that reflects an articulated electroacoustic meaning full of references to multifaceted musical expressions, cultures, and fields.

Aldo Piccialli award – Daniel Scorrane, for the ability to combine artistic, scientific, and historical research in a wide-ranging work that well represents the spirit of the Colloquium of Musical Informatics.

Detmold/Venezia/Torino, Ottobre 2024
Davide Andrea Mauro, Simone Spagnol and Andrea Valle
XXIV CIM Scientific Chairs

Special Session

Polifonia del ricercare - Suono e musica nei progetti di ricerca

Lorenzo Porcaro – AA4MD. Algorithmic Auditing for Music Discoverability

The evolution of online platforms over the past decades has radically transformed the way people discover music, and nowadays thanks to social media and music streaming services listeners have access to an ever-increasing amount of tracks and artists. Within these platforms, one of the goals of Recommender Systems (RS) is to help users discover music without making them feel overwhelmed while exploring the huge catalogues available. However, RSs have come under scrutiny due to their potential negative societal impact, notably with regard to issues of fairness, non-discrimination, inclusion and diversity. Algorithmic auditing has emerged as a tool to analyse the problematic behaviours exhibited by RS, and to offer remedies that can limit their negative impact. The AA4MD project aims to advance this area of research by crafting auditing techniques tailored specifically for music RS. The objective of the project is to demonstrate how the involvement of end-users in the auditing process can contribute to the recognition, analysis, and mitigation of problematic behaviours which may arise while discovering music. Ultimately, the AA4MD project seeks to contribute to addressing the challenges posed by music RS behaviours and to foster a more inclusive and diverse environment for music discovery within the digital landscape.

Giorgio Stefano Gnecco – MOTUS. Automated Analysis and Prediction of Human Movement Qualities

Innovative applications of human movement analysis, e.g., for mitigating/slowing down certain pathological conditions, have recently emerged from the modeling and automated measurement of expressive mid-level individual and group movement qualities, at a higher complexity level than movement qualities derived directly from physical signals. The availability of automated analysis techniques of mid-level expressive movement qualities can contribute to interaction design incorporating performance practices inspired by artistic theories in dance and music. This work, as part of the project MOTUS and in partnership with EuroMov, investigates how such practices and techniques can support embodied interaction design by enabling automated measuring of cues of leadership, cohesion, and fluidity in group movement. In particular, the dance-inspired scientific approach, the data collection protocol, and techniques of analysis for assessing leadership and cohesion within the group and fluidity of the dancers' movement are described.

Matteo Leonardi, Federico Simonetta, Angelica Vomera – LAUDARE. The Italian Lauda: Disseminating Poetry and Concepts Through Melody (12th-16th centuries)

La lauda è un genere di poesia musicale sviluppatosi in Italia a partire dal XII secolo. Era principalmente utilizzata da laici al di fuori della liturgia ufficiale per trasmettere messaggi religiosi e non religiosi a un pubblico per lo più analfabeta e non istruito. Di conseguenza, essa rappresenta una ricca risorsa storica. Il progetto ERC AdG LAUDARE mira a raccogliere e digitalizzare i manoscritti italiani contenenti Laude dal XII al XVI secolo per studiarne il funzionamento interno in termini di composizione, trasmissione, musica e influssi della tradizione orale. Il database ad accesso aperto risultante consentirà indagini su testi e musica, oltre che sui modelli di diffusione geografica. A tal fine, il gruppo di ricerca sta sviluppando strumenti di AI, formati file e protocolli di trascrizione per processare manoscritti e stampe poetico-musicali provenienti da un periodo storico lungo 4 secoli. I risultati del progetto saranno pubblicati in varie forme e presentati in workshop e conferenze.

Stefano Giacomelli – Musical Metaverse: An Inclusive Extended Reality Platform for Networked Musical Interactions

Il Metaverso implementa un universo digitale virtuale immersivo che può affiancarsi al mondo tangibile combinando elementi di Realtà Virtuale (VR) e Realtà Aumentata (AR) e nel quale ogni partecipante interagisce con gli altri attraverso un avatar personalizzato. Il Metaverso Musicale è una parte emergente del Metaverso dedicata alle attività musicali. Attualmente, le opportunità e le sfide tecniche, artistiche e di inclusività sociale del MM sono poco esplorate: le attuali piattaforme di Extended Reality (XR) non supportano con efficacia interazioni in tempo reale tra musicisti in locazioni remote, pochi hanno studiato cosa significa agire musicalmente nell'XR, non vi sono studi sui fattori di diversità (genere, disabilità, ecc.) nel settore. Il progetto mira a definire le linee-guida necessarie per creare più credibili e flessibili ambienti di interazione nel Metaverso Musicale, seguendo tre obiettivi interdipendenti: 1) studiare le esigenze dei musicisti e progettare con loro metodi adeguati; 2) sviluppare soluzioni per l'interconnessione sincrona efficiente in ambienti XR collaborativi; 3) sviluppare servizi per applicazioni musicali. Seguiamo un paradigma metodologico iterativo in tre fasi parzialmente sovrapposte (Design-Develop-Evaluate), ognuna ripetuta due volte per scopi di verifica, operando in diversi contesti di esperienza (composizione, esecuzione, didattica) e ambito culturale (musica classica, sperimentazione, linguaggi di fruizione popolare).

Federico Avanzini – SONICOM. Transforming auditory-based social interaction and communication in AR/VR

Simulating spatially located sounds in virtual or augmented reality (VR/AR) must be done in a unique way for each individual and currently requires expensive and time-consuming individual measurements, making it commercially unfeasible. Furthermore, the impact of immersive audio beyond perceptual metrics such as presence and localisation is still an unexplored area of research, specifically when related with social interaction, entering the behavioural and cognitive realms. SONICOM is designing a new generation of immersive audio technologies and techniques, specifically looking at personalisation and customisation of the audio rendering. Using a data-driven approach, it explores, maps, and models how the physical characteristics of spatialised auditory stimuli can influence observable behavioural, physiological, kinematic, and psychophysical reactions of listeners within social interaction scenarios. The developed techniques and models are evaluated in an ecologically valid manner, exploiting AR/VR simulations as well as real-life scenarios, and developing appropriate hardware and software proofs-of-concept. Finally, in order to reinforce the idea of reproducible research and promoting future development and innovation in the area of auditory-based social interaction, the SONICOM Ecosystem is being created, which will include auditory data closely linked with model implementations and immersive audio rendering components.

Andrea Gulli – S-TWIN. The auditory digital twin of a cochlear implant: framework requirements for immersive sonic interactions with children

This research project investigates the potential of an "auditory digital twin" (DT) to enhance rehabilitation for children who use cochlear implants (CIs). Traditional CI rehabilitation, while effective, faces limitations due to cost and therapist availability. The proposed DT system aims to supplement existing therapy by assisting young children with CIs in learning to identify and locate sounds in their surroundings. Specifically, the project focuses on preverbal children aged 3-6 and aims to create engaging tasks that encourage learning through movement. By analyzing data

on how children move in response to the sounds, the DT would personalize the experience and dynamically adapt to their individual needs.

Davide Rocchesso – MAHATMA. Multiscale Analysis of Human and Artificial Trajectories: Models and Applications

Trajectories are paths followed by objects in motion through space, as functions of time. Our senses continuously detect and interpret trajectories for surviving in environments populated by human, animal, and artificial moving agents. In general, trajectories unfold in space and time, with possible degenerate cases when one of the two domains is collapsed to a single point in space (e.g., point-like tactile stimulation) or time (e.g., the trace left by a snake on the sand). The project is aimed at providing consistent representations of trajectories at different scales of the egocentric space, through a variety of technological means that exploit the different senses at the scale they are most effective. Applications are developed as proofs-of-concept for the trajectory models, in the areas of sport and performing arts, with implications in a variety of fields, including culture- and art-enabled motor reactivation and rehabilitation, and navigation help for sensory-impaired people.

Keynotes

Keynote 1

Pierre-Alexandre Tremblay

All In: The wager of entangled musicking research(es)

Over the past decades, definitions and (dis)agreements on what is valid research in musicking have been in flux, with anchors of various flavours: auto-ethnographic, techno-scientific, strictly aesthetic, wider cultural critic... but what happens when a researcher's musicking straddles across many fields, in an un-disciplined approach? Furthermore, as all attendees of this conference will certainly know, music/sound practice/research with/through/against technology is a fantastic Petri dish for rich interdisciplinary tensions; therefore, what could we learn from our many multi-faceted endeavours, and the various communities in/from which they are happening? Even more daringly, what if we gamble all in, a wager on a monistic view of musicking research where all divisions are considered perspectival, for the benefit of ever-evolving hybrids that challenge stagnating disciplinary boundaries and centres? This talk will present and question the author's holistic assumptions candidly, towards an emerging claim that a return to music itself, in all its forms, should be at the centre of the various episteme of music research, mixed with a healthy challenge to the primacy of the word and the various institutions that depend on it.

About the speaker

Pierre Alexandre Tremblay (Montréal, 1975) is a composer and a performer on bass guitar and sound processing devices, in solo and within various ensembles. He is a member of the London-based collective Loop, and his music is released on Empreintes DIGITALes and Ora.

He formally studied composition with Michel Tétrault, Marcelle Deschênes, and Jonty Harrison, bass guitar with Jean-Guy Larin, Sylvain Bolduc, and Michel Donato, analysis with Michel Longtin and Stéphane Roy, studio technique with Francis Dhomont, Robert Normandeau, and Jean Piché. Pierre Alexandre has been Professor in Composition and Improvisation at the University of Huddersfield (England, UK), he anchored the ERC funded Fluid Corpus Manipulation project. He is now Research Professor in Composition at the Conservatorio della Svizzera italiana, Lugano. He previously worked in popular music as producer and bassist, and has a keen interest in creative coding.

<https://www.pierrealexandretremblay.com/>

Keynote 2

Marinos Koutsomichalis

Vectors of care & musical informatics. From sinewaves to AI: Tactics to synthesize audio with computers

Intended as an overview of the state of the art, this keynote accounts for historical and topical tactics vis-à-vis the synthesis of audio with computers. It more deeply elaborates on specific examples (also drawing on the speaker's practice over two decades) aiming at bridging canonical with niche tactics and experimental with hackneyed ones. In tandem with this narrative, it also outlines affairs of care in the arts and highlights related areas of concern as well as opportunities in the broader field of Musical Informatics.

About the speaker

Marinos Koutsomichalis is an artist, scholar, and creative technologist. He is broadly interested in the materiality of self-generative systems, (post-)digital objecthood, sound, image, data, electronic circuitry, perception, selfhood, landscapes/environments, and the media/technologies we rely upon to mediate, probe, interact, or otherwise engage with the former. He is an Assistant Professor in Creative Multimedia at the Cyprus University of Technology (Limassol, CY) where he co-directs the Media Arts and Design Research Lab. He has administrated three National and three EU funded projects either as the Principle Scientific Coordinator or the Leader of some individual work package. These include the COST Action 'Toolkit of Care' (Action Chair), the HORIZON-MSCA Staff Exchange project 'Epistemology in Science, Technology, Engineering, Arts, and Maths' (WP Leader), the E+ KA220 project 'Aesthetics and Ecology in Technological Education' (PR Leader) and others.

<https://marinoskoutsomichalis.com/>

Paper Abstracts

Fabio De Sanctis De Benedictis – Memoria come Composizione, Composizione come Memoria: strumenti algoritmici in Poisson Trio e Anamniseis

This paper examines the algorithmic compositional techniques implemented in two of the writer's compositions, recalling composers and techniques of the past.. After a brief introduction, we move on to very briefly describe the author's algorithmic composition library, in which memory plays a main role, being based on the formalization of compositional techniques from the twentieth century onwards.. Subsequently, some algorithmic functions used in two of own compositions are analyzed: Poisson Trio, for B flat clarinet, alto saxophone and bassoon, and Anamniseis, a quadriphonic electroacoustic work, inspired by the figure of Xenakis. Specifically, it illustrates how the matrix used by Xenakis in Achorriopsis is used to control rhythmic density and polyphony in Poisson Trio, while pitches are determined by sieve theory. In Anamniseis we illustrate some audio generation methods in which some functions of the algorithmic composition library intersect with the OpenMusic libraries based on Csound. We thus have banks of filters, synthesis tools from physical models (Modalys), creation of OMChroma classes from Csound code, algorithmic spatialization.

At the end of this, brief conclusions follow, with considerations on the role of memory in composition, on the relationship between past, present and future in the field of compositional techniques, and some reflections on tradition and avant-garde in note.

Stefano Catena and Enrico Dorigatti – Missing memories: why we need to analyse spatialisation

This paper aims to foster a discussion over the necessity of developing an analysis and classification framework of the spatial parameter in music, especially within the electroacoustic and acousmatic music contexts. As the exploration of space as an artistic feature is not novel in music, our discourse builds up from a review of the field, investigating its origins and identifying previous key experiences. This paper considers the state of the art of spatial analysis, comparing it with issues currently affecting electroacoustic and acousmatic music musicology, namely the lack of an agreed notation system and the technological fragmentation of the field. The following discussion proposes spatialisation as an experiential feature and recommends a perspective that contrast the dominant philosophy which interprets space from a geometrical and technological perspective. From this novel standpoint, we advocate that spatialisation should be studied independently from the technology employed, focusing only on aural implications. By inviting a conversation on this matter, we also suggest that the proposed perspective could work retrospectively, improving the analysis of electroacoustic pieces belonging to the historical repertoire.

Daniele Giuseppe Annese, Francesco Vitucci, Anthony Di Furia, Francesco Scagliola and Giuseppe Silvi – Archeotopologie: implementazione critica di memorie senza colore

In the study of electronic music, the critical observation of archetypal algorithms and computer architectures is a technical and stylistic exercise comparable to studying passages from the repertoire for acoustic musical instruments. Specifically, it involves practising signal processing skills for educational purposes and the sustainability of musical and cultural heritage.

The reference literature relates to the emergence of digital reverberation techniques through the construction and relationships of filters. The texts by Manfred R. Schroeder express moments of mathematical and physical insights, the re-construction of which offers inevitable technical, computer, and musical reflections.

Implementation through textual computer languages and appropriate signal analysis models allows access to historical literature in a contemporary understanding, ensuring future continuity of the text and opening perspectives for new algorithmic construction.

Daniel Scorrane and Agostino Di Scipio – Gli ‘sciami di glissando’ in Diamorphoses. Ricostruzione mediante Digital Morphophone Environment

Iannis Xenakis' first electroacoustic work, *Diamorphoses* (for magnetic tape, 1957), has been studied and analyzed by various scholars, usually in analysis approaches based on listening and perceptual features. Information concerning the production means is limited and uncertain. Some authors mention the use of the phonogène as well as of the morphophone, magnetophonic processing systems developed in the '50s at GRM in Paris, where the piece was made. Here, based on the historical sources and clues gathered, we put forth the hypothesis that the morphophone was used to produce very peculiar sonorities sometimes described as "glissando swarms" or in other, similar ways. To test the hypothesis, in concrete operational manners, we resort to the Digital Morphophone Environment, a recent audio digital application written in Max. By way of reconstructing the accumulation process of innumerable short glissando sounds, we prove the hypothesis to be essentially correct. The paper also provides hints as to how Xenakis might have tried to manage the 'density' of his thick sound textures.

Cristiano Bocci and Andrea Valle – Fully generalized Fibonacci series modulo n as music sequence generators

In this paper we introduce fully generalized Fibonacci sequences as useful tools for the generation of integers that can be interpreted in the musical context for algorithmic composition. In particular, we take into account the modulo operation on Fibonacci sequences resulting in various periodic behaviors that we interpret in the pitch class domain. First we introduce Fibonacci sequences and generalized Fibonacci sequences, then we discuss the modulo operator applied to sequences. We propose various interpretations of the resulting sequences in terms of pitches and pitch classes, and describe some possible operations. Finally, we introduce fully generalized Fibonacci sequences and describe a possible implementation in an algorithmic composition environment.

Paolo Paradiso – La sperimentazione vocale nell'epoca dell'Intelligenza Artificiale

This essay will take into consideration some innovative creative possibilities offered by digital technologies with artificial intelligence (hereinafter AI) in the field of vocal experimentation. Given the transformative scope of these technologies, it is extremely necessary to distinguish what about AI is a reality today and what, instead, falls within the scope of speculation and science fiction. AI challenges important social, political and cultural issues on several fronts. Furthermore, it poses philosophical questions about the human entity itself: does the traditionally understood definition of "human" need to be rethought in light of increasingly advanced technological progress? Are we heading towards a "posthumanism"? And what consequences can all this have in music? These and other questions will be addressed in this essay, not always claiming to provide univocal and definitive answers, but rather opening the way to diversified discussions, with the aim of trying to bring greater clarity on what it means to make music using AI today. The case study analyzed concerns the experimental composition *ULTRACHUNK*, for voice and AI, by the composer Jennifer Walshe, which allows us to convene reflections on the philosophical theories of posthumanism and apply them to the musical world.

Marco Matteo Markidis – Mediation Process in a Computer Music interpretation: an Ecosystemic Approach

This paper presents an interpretation of Audible Ecosystemics no. 3a, a piece for solo live electronics by Agostino Di Scipio. The project implements the digital signal processes transcribed in the score to realize the piece in an audio synthesis environment, independently from the original computing environment. The work consists of two parts: (1) an implementation phase based only on schemes and scores written by the composer and (2) a second phase of direct testing with the original audio synthesis environment. Finally, the project aims to execute this version in a concert and to distribute implemented libraries.

Barbara Grosso and Andrea Valle – Sonification of Edoardo Sanguineti's poems

Edoardo Sanguineti's poems are characterized by a very idiosyncratic style, which includes not only aural but also visual features, as poems show peculiar structures, sound repetitions, extra- and multi- alphabetical components, and multilingualism.

The paper introduces a sonification system that allows the parallel display of these stylistic aspects. After identifying a corpus of 20 poems representative of the aforementioned features, a preliminary design step (which includes scenario analysis, task analysis and data analysis) has led to a sonification scheme in which each single character is associated with a sound, in particular an earcon.

In the sonification algorithm, the text is processed linearly and values for sound parameters are retrieved for each character so to generate an earcon stream based on different synthesis techniques. We discuss implementation in SuperCollider and introduce a graphical interface, which enables playback management and parameter control, so that the features of each poem can be explored interactively.

Giovanni Sparano – Software di supporto all'esecuzione musicale in MiraWeb: due casi di studio

The paper describes the implementation and the operating instructions of software tools that support musical performance, designed in Max/MSP with the aid of MiraWeb. These tools have dynamic graphic interfaces, updated in real time, and are served over the network to the performers. The software are developed as client-server systems based on websocket technology. Two case studies are presented, concerning the digital adaptations of the scores and the creation of the performance instructions for Variations II by John Cage, and Frame by the author. Both of them are for variable ensembles. The software are open source and distributed freely, upon request to the author, under the CC BY-NC-SA 4.0 license.

Davide Commone – Manifold Voyager, Itinerario sperimentale tra forme musicali ricorsive

This paper briefly presents the compositional method used for creating a piece for trombone trio, fixed media, and live electronics. It also discusses the musical, scientific, and technological inspirations that contributed to the composition. The work is proposed as a musical experiment, aiming to formulate a compositional method where sound distribution over time is managed by an algorithm. This algorithm, drawing from data extracted from soundscapes recordings and employing recursive procedures, generates a grid of instructions on how to arrange musical materials over time. The focus on diverse thematic areas – such as music theory, the study of soundscapes, and information theory – primarily arises from three different suggestions: the influences of repetition

and transformation in Steve Reich's musical production, studies linking soundscapes to their musical properties, notably those by Murray Schafer, and research on information theory and formal logic in relation to certain artistic production, as conducted by cognitive and computer scientist Douglas Richard Hofstadter.

Giuseppe de Benedittis – Partiture di Voltaggio, Metodo di notazione per sintetizzatore modulare

The reproducibility of a musical work, in the absence of the author or the machines/instruments on which it was conceived, but especially its notation, are two problems in electroacoustic music. Over the years, musicologists, performers, and composers have addressed these issues but have failed to reach agreement on a common system. This is mainly due to the willingness and interest of the composer to have his or her work performed but also to the use of specific instruments, which are often difficult to access and for which there is no codified writing system to define parameter values over time. Therefore, it is not difficult to be faced with the creation of personal forms of notation that very often are not easy to interpret or too sketchy.

Through observation and comparison of the work done in this regard by great artists, a method of notation for modular synthesizers through the use of voltage control (CV) is proposed here.

Alexandrina Bargan, Maximiliano Romero and Simone Spagnol – Studio di un'interfaccia musicale per pazienti con demenza di Alzheimer lieve

Non-pharmacological therapies currently represent the set of interventions that focus on the well-being and quality of life of people with cognitive impairment, in particular those affected by Alzheimer's dementia. Among the most common interventions is music therapy, whose benefits - that can be seen starting from the initial phase of the disease - are the maintenance and enhancement of those skills that are still present. Within this context, the objective of the article is to describe the process that led to the (still ongoing) development of an interface to create frequent opportunities for musical activity, even outside of purely therapeutic sessions. It is equipped with a series of interactive handles through which participants, in the presence of professionals or volunteers, can control the playback of musical tracks and soundscapes with high gestural freedom.

Daniel Scorrane – Riproposizione digitale di uno strumento pionieristico. Digital Morphophone Environment

This paper introduces a digital rendering of the morphophone, a complex magnetophonic device (tape loop delay) developed in the 1950s within the laboratories of the GRM (Groupe de Recherches Musicales) in Paris. The analysis, design and implementation methodologies underlying the Digital Morphophone Environment are discussed. Through a detailed analysis of historical sources and available documentation, including a limited amount of literature and mostly repertory images, the essential processes of morphophone were modeled in the Max visual programming environment. The goals for this include, on the one hand, the opportunity to study and make available a now obsolete, no longer available tool and, on the other hand, the opportunity for possible explorations in computer music and research.

Costantino Rizzuti and Fabrizio Rizzuti – Controllo remoto via OSC di un prototipo di synth basato su Raspberry Pi

SynthBerry Pi is the prototype of a standalone synthesizer based on the Raspberry Pi mini computer and the Pure Data development environment. The prototype is equipped with a control surface made up of eight potentiometers connected to the computer which allow to change in real time

the parameters of the sound generation processes. Furthermore, by sending Open Sound Control (OSC) messages over a wireless network, the Pure Data patches can be controlled remotely also using mobile devices such as tablets or smartphones.

Riccardo Ancona – Una prospettiva critica sui dataset per la sintesi text-to-audio

Text-to-audio synthesis (TTA) promises sound generation mediated solely by natural language. Like all generative deep learning techniques, it derives its world memory and semantic boundaries from the dataset with which it is trained. Following a humanistic tradition of critical algorithm analysis, this article examines the datasets employed by various text-to-audio synthesis models and how they are designed and manipulated by researchers. Datasets are not inert informational structures, but socio-technical objects involving specific symbolisations mediated by cultural, political, and industrial perspectives. A comparative study of label datasets, caption datasets, and algorithmically augmented datasets reveals the technical and ethical limitations associated with a quantitative approach to data collection. In contrast to a purely computational dataset evaluation paradigm, a qualitative analysis methodology is proposed to interpret the connotative capability of models and suggest alternative practices in the collection of information.

Leonardo Gabrielli – Considerazioni su VCV Rack come piattaforma didattica per l'ingegnere e il musicista

Academic institutions teaching subjects like digital sound processing are small in number and take different approaches depending on the type of training required: from engineering to electroacoustic music composition. Nevertheless, the variety of languages and software platforms useful for teaching is extremely diverse, making it difficult for the teacher preparing new teaching materials to choose. This article discusses the potential adoption of a relatively recent software, VCV Rack, as a tool that can serve a good number of different educational contexts. Rack is an open-source software developed by VCV that emulates a Eurorack modular system where users can load plugins written in C++ in addition to those provided by the community. Programming such plugins is ideal for engineers who want to delve into DSP module programming, while the presence of a modular workspace makes the software suitable for musical improvisation, synthesis, and random composition. Some of the author's experiences are reported along with suggestions and an overview of the advantages and disadvantages.

Andrea Gulli, Federico Fontana, Hanna Järveläinen and Michele Geronazzo – A mobile game app for adaptive assessment of pitch discrimination in children with different hearing ability

Recent advancements in audiological testing and rehabilitation prioritize utmost personalization and minimal stress in patients, especially the young ones. Hearing health assessments integrate innovative approaches, emphasizing ecological listening scenarios and patient engagement. Driven by these principles, a mobile application tailored for pitch discrimination in children has been designed. The app implements intuitive game mechanics with a captivating graphical interface and harnesses machine-learning algorithms to adapt sound pressure levels to individual comfort levels. It utilizes simple yet effective acoustic stimuli obtained from second-order digital resonators, ensuring a more ecological approach. The pitch discrimination threshold is obtained with adaptive psychometric techniques to guarantee reliable and faster measurements. Preliminary qualitative evaluations involving normal hearing and single-sided deaf with cochlear implant children yield promising outcomes. The resulting perceptual thresholds align with established literature, envisioning the app's efficacy in delivering accurate assessments. The presented tool paves the

way for the use of gameplay in young hearing-impaired individuals' rehabilitation after treatment with cochlear implants.

Alessandro Anatrini – WavePilot: framework multidimensionale per l'esplorazione dello spazio parametrico di strumenti digitali

The increasing spread of deep learning techniques for multimedia content generation has opened new possibilities in sonic and visual exploration and human-computer interaction (HCI). In this context, I present WavePilot, a framework designed to facilitate the exploration and manipulation of Digital Multimedia Instrument (DMmI) parameters. WavePilot employs a Variational Autoencoder (VAE) to translate the values of one or more DMmIs into a multidimensional representation of their parameter space. The primary goal is to enhance interaction by providing the user with a high-level graphical user interface (GUI) in the form of a navigable virtual space, simplifying the programming of DMmIs. Unlike more conventional approaches that are limited to the multidimensional representation of a Digital Music Instrument's (DMI) timbral space and rely on analyzing the perceptual characteristics of audio, this approach focuses exclusively on parametric values to offer an immediate visual representation context, agnostic to the type of devices applied, be they audio generators, video generators, or effects.

This approach aims to stimulate critical reflection among various stakeholders in the musical and technological fields, such as musicians, researchers, developers, and designers. The focus is on rethinking interaction modalities with artistic material in light of the new possibilities offered by AI-based devices. This paper provides the motivations behind the development of WavePilot, along with an analysis of its architecture and usage methods.

Angelica Speroni and Alessandro Bile – Il potere dell'intermedialità nella creazione della memoria artistica: il caso di enigma 33 di Lucia Romualdi

In the contemporary arts scene, intermediality is the fundamental concept that constitutes the very fabric of works. Intermediality, understood as an artistic methodology that dialectically involves different expressive languages, not only dissolves traditional boundaries between disciplines, but also acts as a potential catalyst for memory. In this context, Lucia Romualdi's work stands out for its ability to weave together musical, visual, and conceptual elements, resulting in works that transcend the conventional boundaries of art by creating a memory in the viewers. This talk aims to explore the power of intermediality in the construction of artistic memory, focusing on the eloquent example of enigma 33, the artist's latest work.

This is presented as a contemporary epic, an enchanted fable written in captivity in the Nazi camps that unravels through the meanders of historical memory and imagination, an installation to which the projected images, Claudio Jacomucci's spatialized accordion and Franco Mazzi's narrator's voice contribute. The very fact that it is conceived as an intermedial work determines its structuring as memory, and for it to survive its intermedial quality will have to be preserved.

Similarly, biological mnemonic processes construct memories by relating different districts of memory, and where their communication is prevented the created memory is lost.

Claudia Rinaldi and Marco Pennese – A Pierre. Dell'azzurro silenzio, inquietum (1985). Analisi Formale

This study provides a formal analysis of the piece "A Pierre. Dell'azzurro silenzio, inquietum" composed by Luigi Nono in 1985, which stands as an emblematic work within the electroacoustic tradition. Employing an interdisciplinary approach that integrates attentive listening with score

analysis, we delve into the sonic layers and formal structures that characterize this composition. Additionally, we contextualize the work within the history and aesthetic evolution of the composer, focusing on his later works to grasp the creative context and influences that have shaped this masterpiece.

Abstract dei lavori musicali/Musical Works Abstract

Concerto inaugurale/Inaugural Concert (September 30th 2024)

Karlheinz Stockhausen, Klavierstücke I-IV (1952-53)

Roberto Doati, Studi I-IV (2021)

Karlheinz Stockhausen, Klavierstück V (1954)

Roberto Doati, Studio V (2020)

Karlheinz Stockhausen, Klavierstück VI (1955-1961)

Roberto Doati, Studio VI (2021)

Karlheinz Stockhausen, Klavierstück VII (1954)

Roberto Doati, Studio VII (2021)

Karlheinz Stockhausen, Klavierstück VIII (1954)

Roberto Doati, Studio VIII (2021)

per pianoforte e elettronica

Ciro Longobardi pianoforte, Roberto Doati elettronica

I Klavierstücke I-VIII di Karlheinz Stockhausen ruotano intorno all'esperienza elettronica degli Elektronische Studie I e II. Se i Klavierstücke I-IV (1952-53, dedicati a Marcelle Mercenier) rappresentano una sorta di schizzo dei pezzi elettronici che verranno, i Klavierstücke V-VIII (1954-55, dedicati a David Tudor) rivelano una nuova attenzione al fattore temporale che nel contempo 'dilatata' la forma secondo "criteri statistici" e consente all'autore di costruire timbri diversi (quasi in competizione con quelli elettronici su cui aveva lavorato per 18 mesi) che emergono dal costante uso di risonanze prodotte dalla pressione silenziosa dei tasti. Inoltre le due serie rappresentano il passaggio dal serialismo integrale al controllo aleatorio delle strutture, dalla musica per 'punti' (Punkte) alla musica per 'gruppi' (Gruppen). Nella nostra esecuzione, la serie V-VIII verrà amplificata secondo le indicazioni dello stesso Stockhausen che voleva 'immergere' l'ascoltatore nelle risonanze dello strumento. Gli Studi elettronici I-VIII di Roberto Doati prendono ispirazione dai Klavierstücke I-VIII. L'autore ha voluto ricreare il suono elettronico degli anni '50: nella sua morfologia principale tanto simile a quella dei suoni di pianoforte (dovuta a tagli netti del nastro magnetico) e nel suo 'colore' ottenuto anche grazie alla convoluzione con la risposta all'impulso del riverbero a piastra EMT 140, quello usato da Stockhausen per Kontakte. Ogni studio adotta generazioni spettrali e articolazioni diverse in una sorta di 'mala copiatura' dei Klavierstücke, ma sempre concependo ogni suono come una momentform la cui durata e istante di inizio siano imprevedibili, ed entro cui talvolta è possibile sentire l'eco appena accennata di una composizione strumentale.

Concerto/Concert I (October 1st 2024)

Arazzi Laptop Ensemble (Luca Richelli, Giovanni Sparano, Julian Scordato, Paolo Zavagna) – All vivid noise (2023, rev. 2024)

per organico variabile di strumenti acustici e laptop

Lorenzo Di Marco flauto, Alessandro Mazzacane violoncello, Daniele Nava pianoforte

Il brano presentato, per organico variabile di strumenti acustici e laptop, è un omaggio ad Alvis Vidolin. Le lettere del titolo vengono utilizzate come dispositivo generatore di materiali e articolazioni sonore nel tempo, principalmente attraverso permutazioni di lettere/note componenti il nome. Si tratta di una realtime composition che si sviluppa e realizza nel momento dell'esecuzione grazie a una partitura grafica generativa. Quest'ultima viene poi tradotta sia in notazione per tutti gli esecutori, sia in dati di controllo liberamente interpretabili dagli esecutori ai laptop.

Stefano Catena – Travelling Without Moving (2023)

per fixed media

“Travelling Without Moving” is a journey in acousmatic composed spaces and environments, from granular rainy textures, to sonic trajectories circling around the listener. The concepts of “journey” and “motion” are investigated musically: the experience “moves” between soundscapes, leading to always-changing sonic worlds, both from the spectral and spatial perspectives. The piece was composed during a residency at the Royal College of Music in Stockholm, and most of the atmospheres are inspired by the acoustic experiences I had during my stay in the city. More importantly, the spatial component of the music is paramount, as the piece was composed for the 29.4 partial dome system hosted in Lilla Salen.

Luca Spanedda – Riti: room is the instrument (2023)

per live electronics

Questo brano si basa su un sistema complesso in grado di manifestare comportamenti emergenti e caotici, dove la personalità acustica di un ambiente (digitale o fisico) diviene oggetto di studio per la performance e viene riflessa in termini di variazione del comportamento del sistema stesso. La prima versione del brano è stata concepita durante il periodo di Laurea Biennale al Conservatorio Santa Cecilia di Roma. Il sistema è costruito utilizzando delle soluzioni di equazioni differenziali di sistemi caotici (L'Attrattore di Lorenz, L'oscillatore di Duffing) come motore di sintesi del suono, opportunamente influenzate da banchi di filtri che hanno il compito di simulare tramite un approccio di modellazione acustica, permettendo nella performance il manifestarsi di risonanze modali che ricordino il suono uno strumento musicale (o altro). Nella versione qui presentata, l'analisi di alcune note di violoncello viene utilizzata per influenzare degli oscillatori forzati di Duffing, i cui parametri sono modificati in tempo reale dall'esecutore che ne esplora le transizioni fra le risonanze modali dello strumento e la soglia del caos. L'idea dell'acronimo RITI evoca l'articolo “Sound is the interface” (SITI) di Agostino Di Scipio, che presenta una prospettiva musicale con capacità di un sistema di autoosservarsi tramite l'ambiente circostante

Totaleee (Andrea Laudante, Paolo Montella, Giuseppe Pisano) – Non è un compendio di etologia numerico-digitale (2023)

per fixed media

In this work we wanted to challenge the idea that acousmatic music composition is an individual activity. We decided to act collectively, to share practices and tools, to let our respective works and sounds influence each other, and to find solutions that would otherwise be unexpected. The work was carried out first remotely then the final composition and spatial mixing process was done in a multichannel studio with a 24-loudspeaker dome.

Agostino Di Scipio – Audible Ecosystemics n. 3a (2003) / realizzazione di Marco Matteo Markidis (2024)

per live electronics

Audible Ecosystemics n. 3a explores the structural relationships that intervene between a digital signal processor, an electroacoustic chain composed of an analog mixer, condenser microphones, and loudspeakers, and the environment, understood as the set of acoustic space, electronic interpreter, and audience. And the sonic material that such an environment creates during the execution. Indeed, this composition uses the background noise of the environment and amplifies it, distorts it, and creates sonic movements and polyphonies while trying to regulate and auto-regulate

its dynamics, growth, and, finally, its own annihilation. This new version realized in Pure Data programming language is an original attempt to create a realization starting directly from the score provided by the composer and to finally have a version that behaves systematically as the score expects but keeps its distinct behavior of its own, new version, untied from the original one. Indeed, the version does not try to replicate the exact behavior of the original implementation, but it tries to have an its own original identity, within its characteristics and peculiarity.

Audible Ecosystemics n. 3a is a piece for solo live electronics, the third one of the series Audible Ecosystemics, that the Italian composer Agostino Di Scipio wrote dating from the beginning of the 2000s. This version, entirely written in Pure Data, aims to create a new version of such a piece, while mediating between the score, a confront with the original environment, when needed, and the actual environment.

Stefano Mancuso – 3 Haiku [Omaggio a Matsuo Bashō] (2024)

per violoncello e live electronics

Alessandro Mazzacane, violoncello

Il lavoro è suddiviso in tre sezioni, ciascuna relativa ad un diverso componimento del poeta giapponese Matsuo Bashō (1644 – 1694). Nell'ottica del tema "Memorie Proiettive", l'incredibile attualità della filosofia zen che permea i componimenti citati suggerisce la possibilità di sperimentare paesaggi sonori creati attraverso la trasmutazione in tempo reale del suono del violoncello, nel piano temporale, frequenziale e spaziale. Il lavoro è suddiviso in tre sezioni, ciascuna relativa ad un diverso componimento.

Le nubi di tanto in tanto ci danno riposo mentre guardiamo la luna:

La melodia eseguita viene scomposta e ricomposta in vari strati, sovrapponendosi al suono naturale del violoncello, concetto che si lega al dualismo tipico della filosofia orientale.

L'allodola canta per tutto il giorno ed il giorno non è lungo abbastanza:

Metafora della vita, il violoncello esegue frammenti melodici di carattere ripetitivo, sempre più energicamente e velocemente fino a raggiungere un apice, dal quale inesorabilmente rallenta la sua corsa ammirando i frutti della sua vita, i grani accumulati nel corso dell'esecuzione.

Tracce d'un sogno di guerrieri nell'erba d'estate:

Il violoncellista recita un proverbio samurai, un canto di guerra. Queste parole riecheggiano in un ambiente dapprima instabile, ma etereo, attraverso l'uso di un vocoder. Il tappeto generato tende via via a stabilizzarsi su rapporti stabili, perturbato dalla melodia del violoncello trasmutata in tempo reale. Ciò che resta nel tempo del canto dei guerrieri è solo erba estiva.

Riccardo Tesorini – HyperReal (2021)

per fixed media

A light beating of wings, from which microcosms, very distant worlds, abrupt tugs spring up like a domino effect. Temporary windows, silences and brutal sound walls.

Excruciating screams that rip and tear until they find an unexpected peace. Writhing in a thousand folds, until it settles placidly in a teeming magma. Until it goes out among lapilli of light.

Nicola Casetta – Waking cloud (2024)

per live electronics

Dynamically controlled feedback systems for solo live electronics

"Waking Cloud" is a dynamic live-set performance, centering on intricate internal feedback processes.

The piece unfolds through a series of meticulously curated feedback states, offering a delicate

balance between controlled parameters and improvisational freedom. Its temporal evolution is characterized by a remarkable flexibility, capable of seamlessly adapting to the unique dynamics of each performance setting. The title of the performance, "Waking Cloud", draws inspiration from James Joyce's seminal work, "Finnegans Wake". In Joyce's novel, time experiences a feedback state, reflecting the cyclical and recursive nature of human experience. The title alludes to this theme of temporal feedback, suggesting a parallel between the fluid, ever-shifting nature of time in Joyce's narrative and the dynamic, flexible temporal development of the performance. "Waking Cloud" aims to create an immersive sonic experience where time itself becomes a fluid and malleable element, shaped by intricate feedback processes and improvisational exploration.

Lidia Zielinska – Insektarion (2024)

per fixed media

"Insektarion" is a piece about Wrocław - one of the biggest cities in Poland. "My" Wrocław is not so much the city's audiosphere, but rather a mental sketch of memory. The work established anew the atmosphere, sound logos, symbols and sonic emblems recalling different situations and people in Wrocław. I used a field recordings made by myself and from the archive of the Laboratory of the University of Wrocław.

Jacopo Cenni – H U N T (2022)

teatro sonoro per performer e lampade a incandescenza

La specie umana cerca da sempre di comprendere la complessità del mondo in cui vive. Creando strumenti d'analisi sempre più sofisticati è arrivata a ottenere risposte via via più rigorose. Nonostante ciò, non è mai riuscita a catturare la luce che va cercando.

HUNT parla dell'incessante ricerca della Verità. La performance prevede due personaggi sul palco: il performer, che impersona l'Umano, e una serie di lampade a incandescenza, che rappresentano il concetto di Verità.

Al centro del palco il performer sta in piedi, immobile, nell'oscurità quasi totale. Inizia a esplorare lo spazio che lo circonda, sperimentando con gli strumenti a sua disposizione. Acquisisce confidenza, ma presto perde il controllo delle sue azioni: questo eccesso innesca la luce, della quale inizialmente ha paura, ma che successivamente lo incuriosisce.

L'oscillazione tra il dominio dell'Umano e la perdita di controllo sulla luce determina lo sviluppo della performance, mettendone in evidenza la caparbia e la crescente frustrazione. Gli stati d'animo dell'Umano, in cui si alternano egomania e stati di delirio, sono amplificati dal suono. Quest'ultimo, inoltre, rende evidente la relazione cacciatore- preda tra i due personaggi, delineando due identità sonore distinte, rispettivamente concreta e sintetica, materica e astratta.

L'utilizzo di tecniche di spazializzazione del suono immersive ribalta la fruizione dello spettatore, collocando quest'ultimo al centro dell'esperienza e trasformandolo nell'oggetto stesso dell'opera. L'autore si propone di ricostruire l'inesauribile sete umana di conoscenza attraverso materiali teatrali e musicali, allestendo una performance di teatro sonoro sempre più tesa e disperata.

Concerto/Concert II (October 2nd 2024)

Paolo Montella – Cairo backwards (2023)

per fixed media

On the Muqattam hill of Cairo, we believed we had discovered an ideal spot to experience the call to Fajr prayer, the dawn adhan. From the rooftop of the Arab Foundation for Digital Expression, the day before, we had marveled at a breathtaking view of the city, and the acoustics seemed

promising. However, upon ascending to the rooftop, we find something unexpected. A low, resonant and rumbling hum from an air conditioner would have accompanied our descent into the field where 300 minarets would echo their own psalmody.

Even when lacking a poetic breath, a narrative is always a subjective representation of the world. Consequently, a representation of the world in the form of art shifts reality from a plane of absolute integrity and adherence to itself to a metaphorical and symbolic plane. In the context of Field Recording, this composition analyzes the ritual of field recordings as a completed process that forces us to endlessly vary the discrete representation of a continuous appearance. Through the recorder as a medium, we are willing to perceive it as a musical paradigm that dictates its syntax, connections, and structural functions. The composition also introduces possible developments on the role that the electroacoustic composer plays in this poethics and on the political implications that derive from the degree of adherence to the reality of this specific world construction.

This composition almost exclusively of recordings made in Cairo during the month of February 2023, using only the Zoom H2n recorder.

João Pedro Oliveira – Singularity (2019)

per violoncello e fixed media

Alessandro Mazzacane, violoncello

A singularity is a phenomenon that relates to several areas of knowledge. In cosmology, the singularity lies at the center of a black hole (resulting from a collapsing star), where matter compresses into an unimaginably small region whose density becomes infinite. Everything that passes within a certain proximity of a singularity will be inexorably attracted and can never escape this attraction. This piece operates with very heavy densities in the electronics part in opposition to the instrument. The instrument tries continually (and without success) to escape the weight and force transmitted by the electronics.

Antonio Forastiero – Animated grains / Reflections (2023)

per fixed media

Animated grains is a compositional experience that starts from the electronic processing of sound events obtained by manipulating small mechanical games. Sounds related to movement and rotation are processed and combined in particular sound combinations capable of generating a particularly dynamic sound result. A timbral exploration that leads to the development of complex and articulated sound situations, rich in sound elements of different nature, linked to sensations of temporal and spatial movement.

Reflections is an electro-acoustic work in which undefined reverberated sound spaces develop, contrasted by non-reverberated dynamic sound situations. Within this environment, recurring sound elements, subject to multiple reflections, through different levels of reverberation, recreate particular multidimensional and variable sound interweaving, giving life to a continuous exchange between movement and stillness.

Ekaterina Khmelevskaya – East West Jungle Disco (2023)

per flauto ed electronics

Lorenzo Di Marco, flauto

The piece reveals the nature of an ancient musical instrument. The voices of flutes from different parts of the planet combine in a driving and slightly mystical dance cocktail.

The piece is a kind of an open score inviting the performer to collaborate with the composer. The piece consists of a live performance and a fixed soundtrack. The performer can play solo or in an ensemble (1-4 performers are possible). The performer plays several instruments of the flute family. It is preferable to use not only an alto flute and piccolo, but also ethnic varieties of flutes. The performer independently determines which fragments from the score and which instruments he will play. The rest of the parts of the score must be pre-recorded by the performer, or it is possible to use a ready-made soundtrack provided by the composer.

The soundtrack was created in stereo format. The performance requires microphone amplification of the flute and signal processing using hall and delay effects.

Andrea Laudante – 12th perception of Prakṛti (2023)

per fixed media

This piece is the 12th piece of the series “Perceptions of Prakṛti”, a work in which I explore repetition and randomness. The 12th perception of Prakṛti was composed using samples from the instruments of Productions Totem Contemporain in Montreal, Quebec.

Aurés Moussong – Abîmes de Motashem* (2016)

per violoncello e fixed media

Alessandro Mazzacane, violoncello

Composition for cello and tape, which is the result of a process of hybridization and exchange of particles, explored through the “composition in abyss”, an artistic or literary technique in which an element within a work represents the entire work or contains repeated versions of itself in a recursive structure, like a miniature universe found within the same creative work. One of the most extraordinary representations of this phenomenon is found in the dreamlike carpets of Kashan, a region in Iran where Motashem Kashan carpets are manufactured, named after the 16th-century Persian poet.

The electronics are primarily made with samples of Iranian cello, setar, and santoor, the latter two performed and recorded by the composer. These samples were processed and transformed to create various sonic dimensions and a narrative within the tape that engages and interacts with the soloist.

Antonio Scarcia – Grainstone (2010)

per fixed media

Grainstone è un lavoro che si snoda attraverso un approccio integrato di composizione e micro-composizione, mediante tecniche di sintesi basate sul campionamento e sul trattamento dei materiali sonori nei domini del tempo e della frequenza applicate a combinazione di gesti tesi ad articolare una contrapposizione tra classi di suono visti come antagonisti. La realizzazione, con tecniche che richiamano quelle del classico tape studio in chiave digitale, è il risultato di un processo che ha integrato Common Music ed elaborazione tramite Supercollider.

Andrea Fabris e Chiara Vitofrancesco – Dimora inquieta (2023)

per grancassa, live electronics e live video

DIMORA INQUIETA nasce dallo studio del concetto di alterazione dell’immaginazione dovuta allo stato di deprivazione sensoriale in una condizione di isolamento. Ciò causa la caduta della reattività dei sensi e aberrazioni

percettive. La monotonia degli stimoli in una condizione di isolamento porta alla perdita dei punti di riferimento rendendoci vulnerabili al desiderio di simulazione di quel qualcosa che non

è disponibile nell'ambiente impoverito in cui siamo, come accade ai marinai che in balia del mare, percepiscono cose altre. È un progetto in duo di Andrea Fabris (Grancassa e Live Electronics) e Chiara Vitofrancesco (VJ) derivato da azioni audiovisive che nascono dall'incontro di due discipline ad oggi complementari: musica post acustica e la videoarte che trovano terreno comune nella pratica dell'improvvisazione. È espressione di potenziale immaginifico che ha lo scopo di tenere insieme vari mondi e mostrare le connessioni tra essi: i due performer si pongono in una reciproca posizione di ascolto allo scopo di rispondere e mettere l'altro nella condizione di reagire creando un andamento dinamico, un corpo/forma impermeabile come luogo di scambio continuo, inquieto, mai stabile in uno solo dei due ambiti.

In tale luogo la realtà viene descritta attraverso l'intensità gestuale generata dagli strumenti adoperati dai due performer.

Contributions

Memoria come Composizione, Composizione come Memoria: strumenti algoritmici in *Poisson Trio* e *Anamniseis*

Fabio De Sanctis De Benedictis
Conservatorio "P. Mascagni" – Livorno
fabio.desanctis@conslit.it

ABSTRACT

In questo scritto si prendono in esame le tecniche compositive algoritmiche messe in opera in due composizioni dello scrivente, richiamando alla memoria compositori e tecniche del passato. Dopo una breve introduzione si passa a descrivere molto sinteticamente la libreria di composizione algoritmica dell'autore, nella quale la memoria gioca un ruolo principale, essendo basata sulla formalizzazione di tecniche compositive dal Novecento in avanti. In seguito si analizzano alcune funzioni algoritmiche utilizzate in due composizioni proprie: *Poisson Trio*, per clarinetto in si bemolle, sassofono contralto e fagotto, e *Anamniseis*, opera elettroacustica quadrifonica, ispirate alla figura di Xenakis. Al termine di ciò seguono brevi conclusioni, con considerazioni sul ruolo della memoria nella composizione, sul rapporto tra passato, presente e futuro nel campo delle tecniche compositive, e alcune riflessioni su tradizione e avanguardia in nota.

1. INTRODUZIONE

Parlare di memoria implica discutere del tempo. È ben noto ed evidente, sin dal passato,¹ il ruolo della memoria nella fruizione musicale, considerando l'essenza intrinsecamente temporale di questa forma artistica. Non sarà quindi necessario dilungarci più di tanto su questo concetto: Jeanne Hersch mette bene in evidenza la natura della musica, temporale da un lato e atemporale dall'altro, richiamando più o meno direttamente il ruolo della memoria nella ricostruzione del senso;² la funzione delle aspettative nella teoria di Meyer non potrebbe sussistere senza operazioni mnemoniche: possiamo prevedere lo svolgimento della musica che stiamo ascoltando in base alla nostra competenza stilistica, acquisita attraverso precedenti esperienze di ascolto, ovvero grazie alla nostra memoria [3]; l'analisi di Imberty della *Sequenza* per voce di Luciano

¹ Possiamo rimandare ad esempio a [1], in cui si dibatte tra le altre cose il ruolo della memoria nella produzione e nella fruizione musicale.

² «La musica si svolge e si avvolge in un unico movimento. Non sappiamo come, ma essa è contemporaneamente successiva e simultanea. Le note e i ritmi si susseguono e si dissolvono, si susseguono dissolvendosi, ma al tempo stesso non si dissolvono, ognuno implica i precedenti e i successivi, e deve loro un senso.» [2], p. 76.

Copyright: ©2024 Fabio De Sanctis De Benedictis et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Berio mostra il ruolo della memoria a breve e lungo termine per la ricostruzione della forma dell'opera [4]. Questi sono pochi esempi dei tanti che potrebbero essere fatti. Per quanto ci riguarda vorremmo occuparci di un aspetto particolare, quello della memoria nella composizione musicale, riportando un paio di casi della ricerca artistica/compositiva dello scrivente, ricerca nella musica e con la musica.³ Si procederà quindi richiamando preliminarmente alcune informazioni relative alla libreria di composizione algoritmica dell'autore, strumento chiamato in causa e utilizzato negli esempi che saranno discussi, procedendo poi alla illustrazione di due esperienze compositive, pervenendo infine alle conclusioni.

2. LA LIBRERIA DI COMPOSIZIONE ALGORITMICA FDSDB_XXTH_CT

Altrove abbiamo già descritto le caratteristiche principali di FDSDB_XXth_CT, la libreria di composizione algoritmica per OpenMusic realizzata dallo scrivente [6, 7]. Per gli scopi del presente contributo possiamo semplicemente ricordare che detta libreria sviluppa tecniche compositive di differenti compositori dal Novecento in avanti, realizzandole in forma di funzioni per OpenMusic. Al momento sono compresi processi compositivi di circa venti compositori, oltre a procedimenti legati alla *Pitch-Class Set Theory* sviluppata in direzione compositiva – principalmente seguendo il lavoro di Robert Morris,⁴ nonché varie utilità e operazioni per lo sviluppo di durate e altezze. In totale abbiamo circa 219 funzioni. Uno degli scopi o delle conseguenze possibili di questa libreria è quello di poter tramandare le tecniche compositive dei compositori che ci hanno preceduto, quindi in qualche modo anche la loro memoria. Ogni funzione ha come icona una rappresentazione, foto o quadro, del compositore di riferimento, in modo che il suo ricordo sia richiamato anche dal registro visivo. Memoria quindi come passaggio di testimone dal passato al presente, e possibilmente al futuro, nell'ottica donatoniana della possibilità di risvegliare la creatività attraverso l'insegnamento della tecnica.⁵

³ Recentemente si è distinta la ricerca artistica "nell'arte" e "con l'arte", si veda [5], pp. 57-58.

⁴ Gli strumenti compositivi ottenuti dalla *Pitch-Class Set Theory* si sono ricavati principalmente dai seguenti testi: [8], [9] e [10].

⁵ «Può accadere, tuttavia, che l'inventività assopita di un giovane abbia bisogno di una qualsivoglia disciplina per poter risvegliarsi. [...] qualcosa che somiglia a una preanalisi dei procedimenti operativi.» [11], pp. 83-84. Vorremmo poi ricordare le parole di Sylviane Sapiro, che al XXII CIM di Udine, svoltosi dal 20 al 23 novembre 2018, ha citato Jean-Claude Risset, il quale affermava il nostro ruolo di "traghettoni", dal passato al futuro.

3. DUE CASI DI STUDIO

Nell'anno del centenario della nascita di Xenakis si sono composti due brani, uno strumentale e l'altro elettroacustico, legati alla figura del compositore greco. Si tratta di *Poisson Trio*, per clarinetto in si bemolle, sassofono contralto e fagotto, e *Anamniseis* (Reminiscenze), opera elettroacustica quadrifonica. Il primo è stato selezionato a seguito di un bando di composizioni per la *VII MusMat Conference*, svoltasi a Rio De Janeiro dal 18 al 21 ottobre 2022. All'indirizzo <https://youtu.be/MwsGFViyZy8?feature=shared> è possibile ascoltare l'esecuzione in concerto dell'*Atrai Ensemble*.⁶ Il secondo è stato selezionato a seguito di un bando di composizioni elettroacustiche per la rassegna *MusLAB electroacústica* e per la rassegna *Atemporánea Festival*, è stato eseguito rispettivamente a Guayaquil, Ecuador e a Buenos Aires, Argentina, e registrato in forma stereo/binaurale su un CD dell'etichetta Cero Records.⁷ È stato programmato anche nella stagione concertistica *Suoni Inauditi* del Conservatorio di appartenenza. In queste due opere il riferimento a Xenakis, sorta di omaggio alla sua opera, si è concretizzato in forme diverse: una traslitterazione in chiave personale di una tecnica utilizzata dal nostro in *Achorripsis*; l'utilizzo della teoria dei setacci⁸ sia per la composizione sia per la generazione di suoni elettronici; la memoria di ascolto di alcune opere strumentali ed elettroacustiche di Xenakis come principio di ispirazione per la composizione di *Anamniseis*, ricreandone alcune caratteristiche nelle trame sonore utilizzate o ricostruendo algoritmicamente procedimenti di generazione audio; le acquisizioni di una recente analisi di *Aroura*, per dodici archi solisti, di recente pubblicazione [12].⁹ Entrando maggiormente nel dettaglio, si procederà di seguito ad illustrare alcuni esempi di strumenti algoritmici utilizzati per la realizzazione delle due composizioni in discussione.

3.1 Poisson Trio

Poisson Trio è una breve composizione per clarinetto in si bemolle, sassofono contralto e fagotto, quasi un piccolo Divertimento, il cui titolo rivela le origini compositive, ossia la rivisitazione dell'utilizzo della distribuzione di Poisson che Xenakis ha impiegato per *Achorripsis*. L'idea alla base prende infatti avvio da un oggetto della libreria di composizione assistita dello scrivente, dedicato alla generazione di matrici di numeri distribuiti secondo la funzione di Poisson.¹⁰ In Figura 1 abbiamo una *patch* di OpenMu-

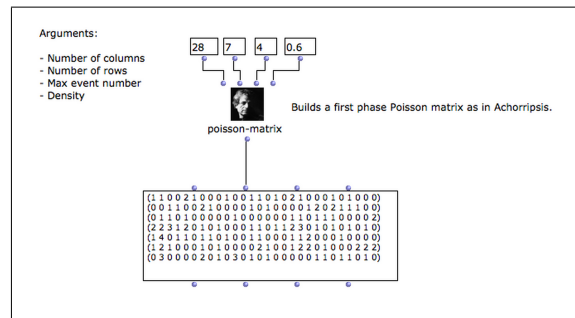


Figure 1. La funzione `poisson-matrix` per la generazione di matrici di numeri distribuiti secondo la funzione di Poisson.

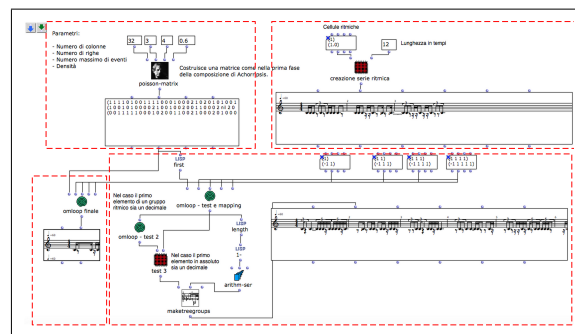


Figure 2. La *patch* di OpenMusic per la generazione di successioni ritmiche dalla matrice di Poisson.

sic tratta dalla documentazione della libreria, relativa alla funzione in oggetto.

Per la composizione di questo trio di fiati siamo partiti dalle successioni ritmiche, con la *patch* illustrata in Figura 2, in cui la parte in alto a destra ha verificato preliminarmente il meccanismo di selezione casuale di cellule ritmiche da riserve predefinite dall'utente. In alto a sinistra è possibile invece osservare la funzione della libreria con la generazione della matrice, sostanzialmente il contenuto della Figura 1. Il meccanismo di mappatura dei numeri della matrice di Poisson in cellule ritmiche è stato dapprima verificato su una linea sola della matrice, ovvero su una singola voce. Si vedano i due `omloop` e la sotto-*patch* nella parte inferiore della figura, sotto la matrice, un po' a destra. Infine, risolti alcuni problemi di cui sarà discusso tra poco, si è proceduto a incorporare questo meccanismo in un `omloop` finale, in basso a sinistra nella figura, potendo così procedere ad una mappatura iterativa che ha permesso di ottenere successioni ritmiche polifoniche. In sostanza ogni linea della matrice diviene una voce, ogni numero diventa invece una cellula ritmica, decisa preliminarmente dell'estensione di un tempo di battuta; il numero di linee della matrice corrisponde così alla polifonia, il numero di colonne alla lunghezza espressa in tempi di battuta. Nell'esempio rappresentato abbiamo perciò come

pagina, la lezione tenuta presso la classe di Composizione di Marco Stroppa alla *Staatliche Hochschule für Musik und Darstellende Kunst* di Stoccarda, dal 18 al 20 gennaio 2008 [14].

⁶ Qui invece il video dell'intero concerto: https://youtu.be/bV_gb5pinQI?feature=shared.

⁷ <http://www.cero-records.com/release/muslab-planeta-complejo-vol-1%2c%2b72%2c%2b73/>.

⁸ La teoria dei setacci, o dei crivelli, può essere generalizzata come un algoritmo di selezione che permette di creare scale di altezze o serie di impulsi ritmici progressivi nel tempo, da cui le durate. Si veda quanto discusso brevemente più avanti.

⁹ Ancora una volta possiamo individuare un esempio di reversibilità tra analisi musicale e composizione, come già tracciato in [13], p. 2.

¹⁰ Una spiegazione un po' più dettagliata del funzionamento della tecnica compositiva di Xenakis per l'utilizzo della funzione di Poisson nella composizione di *Achorripsis* è visionabile alla pagina "Conferenze e seminari" del sito dei corsi dell'autore (<https://sites.google.com/site/fdsdbmascagni/conferenze-e-seminari>), congiuntamente alla spiegazione dell'algoritmo di OpenMusic alla base dello sviluppo della funzione della libreria. Si consulti allo scopo, nella stessa



Figure 3. Uno dei possibili risultati della *patch* in Figura 2, importato in MuseScore e formattato per comodità di lettura.

risultato finale¹¹ una polifonia a tre voci della durata di $32/4=8$ battute (32 è appunto il primo parametro di ingresso della funzione *poisson-matrix*). Nell'esempio riportato le cellule ritmiche sono state configurate, nella riserva predefinita, in modo che al numero zero corrisponda una cellula con un impulso per movimento, al numero uno una cellula ritmica con due impulsi e così di seguito. Uno dei vantaggi di questo metodo è quello di poter gestire la densità ritmica in modo assai semplice, agendo sul numero massimo di eventi previsto per la generazione della matrice (terzo parametro di ingresso della funzione), e/o sul tipo di riserva ritmica preparata per la mappatura. Manovrando questi fattori è stato possibile gestire processi dinamici di densità nello svolgimento formale del brano.

Poco prima abbiamo accennato ad alcuni problemi che si sono dovuti risolvere nella preparazione di questo algoritmo compositivo. Uno di questi è stato relativo alla presenza di alcuni *nil* nella valutazione dei valori della matrice, e l'altro alla mappatura di una cellula ritmica iniziante con una nota legata ad una pausa precedente, cosa di per sé illogica.¹² Il primo problema è stato risolto all'interno di *omloop - test* e *mapping*, per cui ogni *nil* viene sostituito con un valore -1, corrispondente a una pausa di semiminima, prima ancora di effettuare la mappatura complessiva. Il secondo problema è stato risolto all'interno della funzione iterativa *omloop - test 2*, attraverso alcune semplici operazioni logiche: se il primo valore è minore di zero (ossia una pausa) e il secondo valore è un decimale (ossia una nota legata al valore precedente), allora quest'ultimo viene sostituito con una pausa della stessa durata. In Figura 4 viene riportato l'interno di *omloop - test 2*.

La sotto-*patch* finale, denominata *test 3*, risolve il

¹¹ Figura 2 – il risultato deriva dall'*omloop* finale in basso a sinistra, non interessa le altre partiture mostrate in figura.

¹² Ricordiamo che la rappresentazione ritmica di OpenMusic, denominata *tree*, prevede che le pause siano rappresentate da valori ritmici negativi, e le legature di valore da numeri interi decimali, ad esempio 1.0 al posto del normale 1, ad indicare che il valore in oggetto è legato a quello precedente. Per approfondimenti si rimanda all'aiuto online di OpenMusic, in particolare partendo da questa pagina: <https://support.ircam.fr/docs/om/om6-manual/co/RT.html>.

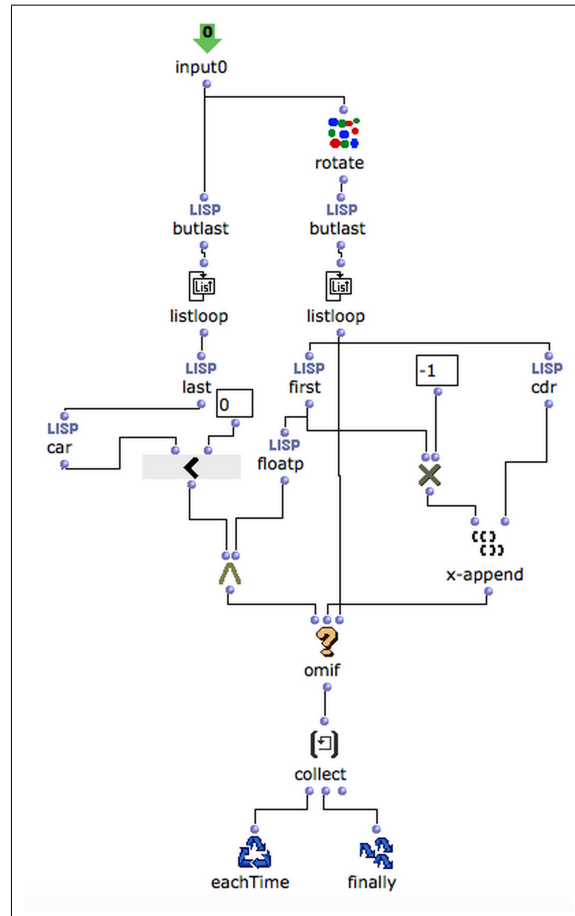


Figure 4. L'interno di *omloop - test 2*.

caso particolare in cui in una voce sia il primo valore in assoluto ad essere un numero decimale, ossia un valore ritmico legato ad una figura precedente, inesistente, non essendoci niente prima dell'inizio. Anche in questo caso il valore viene sostituito con una pausa. Senza questi accorgimenti i numeri decimali sarebbero stati trasformati in valori ritmici reali, ossia note, distorcendo la logica generale di mappatura. Una volta verificato il funzionamento su una sola linea della matrice, ossia su una singola voce, il tutto è stato incorporato nella funzione iterativa *omloop* finale per ottenere una polifonia rispondente al risultato della matrice di Poisson. Ecco che la tecnica compositiva viene richiamata alla memoria e trasformata, a generare nuovo materiale. Materiale che non risponderà in modo rigoroso alle intenzioni di Xenakis, d'altronde non avrebbe senso rifare *Achorripsis*, se non per fini di studio algoritmico, mentre ha senso, almeno per chi scrive, rinnovare la tradizione, anche e soprattutto quella delle avanguardie storiche, prolungandola in qualche modo verso il futuro. La distribuzione degli impulsi delle figure ritmiche selezionate all'interno delle singole riserve preserva comunque qualcosa della distribuzione di Poisson.

Una volta impostate le trame ritmiche, le altezze sono state selezionate liberamente da scale, anche microtonali,

realizzate con la tecnica dei setacci, attraverso le funzioni corrispondenti della libreria. Brevemente, i setacci utilizzati sono stati quelli relativi all'unione delle sequenze di multipli di alcune coppie di numeri primi tra loro: 3-5, 3-7, 5-7, 3-11, 5-11, 7-11. Ovvero, prendendo ad esempio la coppia 7-11, partendo da zero abbiamo la seguente successione che comprende tutti i loro multipli sino alla soglia del loro minimo comune multiplo 77: 0, 7, 11, 14, 21, 22, 28, 33, 35, 42, 44, 49, 55, 56, 63, 66, 70, 77. Stabilendo questi numeri come punti di una semiretta, è possibile trarne figure ritmiche rispetto ad un'unità di misura (ogni numero stabilisce la posizione di un attacco di nota), o altezze rispetto ad una unità intervallare. Infatti il delta risultante della successione numerica precedente sarà 7, 4, 3, 7, 1, 6, 5, 2, 7, 2, 5, 6, 1, 7, 3, 4, 7, per cui se si prende una croma come unità di misura si avrà in ordine: una minima puntata legata a una croma (totale 7 crome), una minima, una semiminima puntata e così di seguito. Se si assume il semitono come unità di misura, partendo ad esempio dal Do3, si avranno di seguito: Sol3 (sette semitoni sopra), Si3, Re4, La4 ecc. Con moduli inferiori al semitono abbiamo soluzioni infratonali. In *Poisson Trio* sono state utilizzate quattro unità di generazione di setacci di altezze, due semitonal e due quartitonal, sulle cui coppie possibili sono state applicate operazioni insiemistiche di unione, differenza, intersezione, per ottenere scale più varie e articolate. Lo spazio non ci consente di discutere più in dettaglio il lavoro svolto, si rimanda ad una recente *Masterclass* tenuta presso il Conservatorio di Livorno in data 13 novembre 2023, *Punti di vista sulla drammaturgia*, il cui materiale può essere consultato dalla pagina del sito dei corsi dell'autore riportata in nota 10.

3.2 Anamniseis

Il secondo esempio che viene trattato in questo scritto riguarda una composizione elettroacustica quadrifonica, in cui si sono impiegate differenti librerie di OpenMusic per la generazione di file audio mono e quadrifonici, poi montati insieme utilizzando il *sequencer* Reaper. In Figura 5 la schermata principale, relativa alla distribuzione dei file quadrifonici nel *sequencer* audio.¹³

Le librerie di OpenMusic utilizzate per la generazione audio sono state principalmente: Modalys, OMChroma, om4csound, OMPrisma, OMSpat, OM-SuperVP, OM-Chant. Per limiti di spazio daremo solo pochi esempi, relativi all'utilizzo di strumenti di sintesi insieme a funzioni della libreria di composizione algoritmica, con lo scopo di illustrare possibili modalità di raccordo tra i due ambiti.

In Figura 6 abbiamo due esempi di uno stesso banco di filtri realizzato attraverso funzioni di om4csound, libreria di Mauro Lanza. Un file audio viene filtrato da serie di altezze desunte da operazioni sulla teoria dei setacci, a sinistra, oppure da uno dei sette modi a trasposizione limitata di Messiaen distribuito su tre ottave consecutive, a destra. La libreria om4csound crea un listato di codice per Csound, richiamando poi questo programma come mo-

¹³ Ogni traccia è stata impostata su quattro canali di uscita, così come la traccia *Master*.

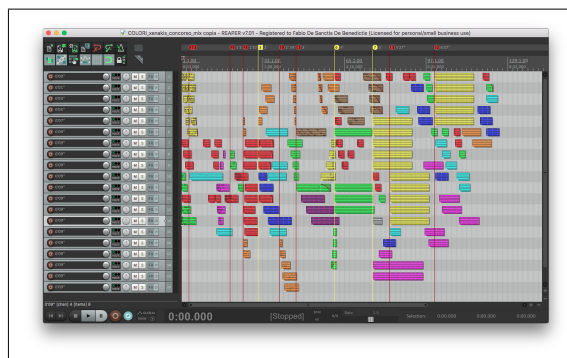


Figure 5. Il montaggio definitivo dei file audio in Reaper. I colori illustrano sommariamente il tipo di materiale utilizzato, solo come riferimento visivo, senza corrispondenze a lunga distanza.

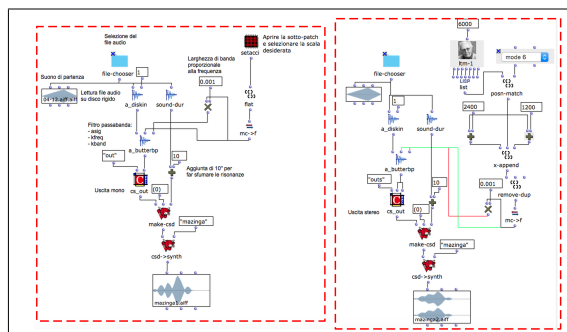


Figure 6. Banco di filtri.

tore di sintesi.¹⁴ È quindi possibile utilizzare, ove opportuno o richiesto, funzioni della libreria per la realizzazione di durate e altezze da dare in pasto alla renderizzazione audio.

In Figura 7 abbiamo una *patch* di OpenMusic basata su uno strumento della libreria Modalys, una membrana circolare, incorporata nella funzione iterativa omloop, che renderizza una sequenza di altezze e durate organizzate attraverso funzioni di maschera di tendenza della libreria.¹⁵ Rispetto ad approcci precedenti¹⁶ sono state utilizzate alcune sotto-*patch* prese dai *tutorial* di Mauro Lanza, per cui le varie note sono state sintetizzate all'interno del processo di Modalys, invece di esserle una alla volta e poi mixate successivamente. Il calcolo risulta così molto più efficiente. In omloop abbiamo lo strumento di sintesi, visualizzato in Figura 8: due membrane circolari all'unisono,

¹⁴ Utilizzando in questo caso, come visibile in figura, l'*opcode* *butterbp*. La generazione di dati per Csound e l'uso di questo programma come motore di sintesi è un processo comune a tutte le librerie di OpenMusic che utilizzano un programma esterno per la renderizzazione audio. Ulteriori esempi possono essere trovati in [15].

¹⁵ Il riferimento alle maschere di tendenza, con la loro conseguente formalizzazione all'interno della libreria di composizione algoritmica, è stato derivato richiamando dal passato alcuni processi presenti in Cmask, un linguaggio di generazione di partiture stocastiche per Csound (<https://github.com/kunstmusik/cmask> e <http://www.bartetzki.de/de/software.html>), e pure con riferimento a [16], pp. 223-224.

¹⁶ Si veda ad esempio quanto riportato in [17], in particolare gli strumenti descritti nelle figure 3 e 5.

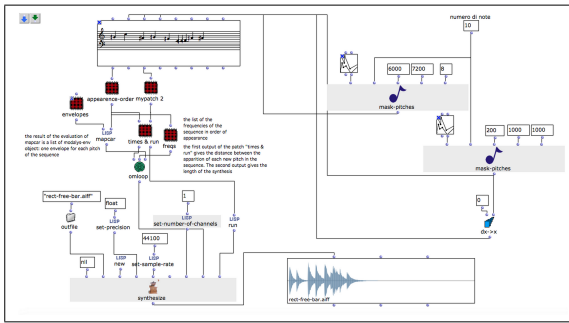


Figure 7. Realizzazione di una sequenza di suoni di membrana circolare, con altezze e durate governate da maschere di tendenza con lo stesso profilo.

con parametri scelti e differenziati casualmente, entro ambiti specificati dall'utente (funzione `om-random`). In un altro esempio, che non riportiamo, si è applicata una variazione per cui le maschere di tendenza filtrano selezioni di note in scale realizzate con la tecnica dei setacci.

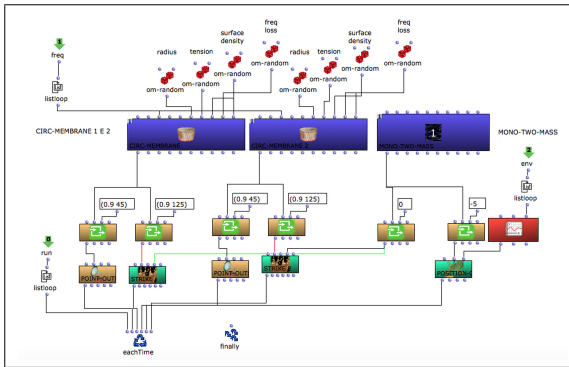


Figure 8. Lo strumento di sintesi in Modalys.

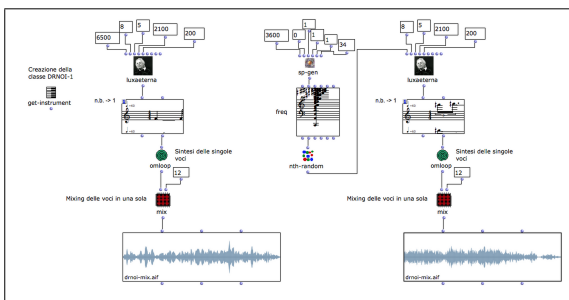


Figure 9. Renderizzazione audio di una polifonia come quella in *Lux Aeterna* di Ligeti.

In Figura 9 abbiamo l'applicazione di una polifonia di stampo ligetiano ad uno strumento di Csound creato appositamente come classe utente della libreria `OMChroma`. Valutando la funzione `get-instrument` viene iniziata la procedura di creazione della nuova classe, nella quale introdurre il codice di Csound richiesto. In questo caso si

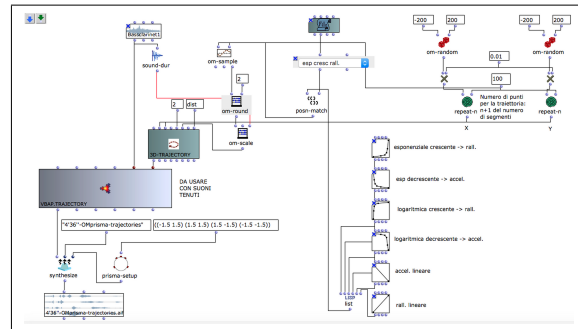


Figure 10. Spazializzazione automatica in quadrifonia.

è utilizzata una semplice, ma efficace, orchestra di Csound di Eric Lyon riportata nel testo storico di Boulanger [18], `drnoi.orc`, di cui si illustra il codice:

```
sr = 44100
kr = 4410
ksmps = 10
nchnls = 1
```

```
instr 1
ires1 = p4
ires2 = p5
ibf = .02
iamp = p6*32767 * 5
iseed = p7
ibw = p4*ibf
kenv linseg 0,p3/2,iamp,p3/2,0
anoise randi 1.0, 13000, iseed
ares reson anoise,ires1,ibw,1
ares1 reson ares,ires2,ibw,1
out ares1*kenv
endin
```

4. SPAZIALIZZAZIONE

Per la distribuzione dei suoni sui quattro diffusori si sono utilizzate alcune *patch* di OpenMusic basate sulle librerie `OM-Spat` e `OMPrisma`. Le traiettorie di spazializzazione sono state determinate casualmente, oppure utilizzando la documentazione di queste librerie per la realizzazione automatica di traiettorie circolari a varie velocità. In Figura 10 abbiamo la spazializzazione di un suono in modo casuale, con la possibilità di scegliere un decorso temporale in accelerando o rallentando, secondo varie curvature, disintegrando il suono tra i diffusori nel caso di traiettorie molto veloci e formate da numerosi segmenti. Nell'esempio riportato sono 101.

5. CONCLUSIONI

Nei due casi analizzati il ricorso alla memoria ha comportato il recupero dal passato di alcune tecniche compositive, principalmente riferite all'operato di Xenakis: la distribuzione di Poisson, la teoria dei setacci, maschere di tendenza. Tutte queste hanno permesso, inquadrate in un

nuovo contesto e variate opportunamente, la generazione di materiale musicale e audio. Xenakis non è stato l'unico compositore chiamato in causa, abbiamo visto anche materiali e processi desunti da Messiaen, Ligeti, Bartetzki. In questo modo il passato (possiamo utilizzare tranquillamente il termine tradizione senza il timore di suscitare contrapposizioni di storica memoria?) viene utilizzato in una nuova chiave, portato a conoscenza delle generazioni del presente e magari proiettato verso il futuro. Il confronto tra tradizione e avanguardia, se si passano questi due termini forse troppo semplicistici e non definiti preliminarmente, potrebbe anche rivelare delle sorprese, in particolare nei casi in cui pratiche compositive più recenti siano confrontabili con altre del passato. In una conferenza tenuta dallo scrivente diversi anni fa presso una scuola privata di Livorno, dal titolo esemplare *Tradizione e avanguardia*, sono stati accostati brani del passato e più recenti, mostrando caratteri più affini di quanto non potesse sembrare. Ecco quindi che *Partiels* di Grisey viene accostato al *Preludio de L'oro del Reno* di Wagner, il canone mensurale fiammingo ad alcuni processi ritmici de *Il Canto Sospeso* di Nono, il Madrigale *A un giro sol de' begl'occhi lucenti* di Monteverdi (prima sezione) a *Music for Airports* di Brian Eno, *Les cris de Paris* di Jannequin a *Cries of London* di Berio, solo per citare alcuni degli esempi mostrati allora. Sicuramente alcuni riferimenti saranno stati consapevoli nei compositori venuti dopo, altri forse no. Questo di certo non comporta, non deve comportare, una nuova accademia, nel senso peggiorativo del termine. Le tecniche scelte rispondono in primo luogo a un'esigenza espressiva, che precede ogni azione, e questo approccio non deve essere esclusivo, ma può suscitare naturalmente, diremmo anzi in modo spontaneo, l'invenzione di nuove soluzioni, poetiche o tecniche che siano. A nostro avviso si acquista così una consapevolezza e una coscienza storica di cui forse oggi avvertiamo un'urgente e nuova necessità. In ciò anche la musica di ricerca, se quella qui esposta può essere definita tale, può dare il proprio contributo, forse marginalmente quanto a disseminazione, ma non per questo meno efficacemente. Per il futuro ci si propone il mantenimento e l'aggiornamento del novero di compositori e tecniche compositive della propria libreria, cercando di incrementare il "traghetamento" di elementi del passato verso il futuro, nella composizione così come nella didattica della composizione.

6. REFERENCES

- [1] S. Agostino, *De Musica*. Sansoni, 1969.
- [2] J. Hersch, *Tempo e musica*. Baldini Castoldi Dalai editore, 2009.
- [3] L. B. Meyer, *Emozione e significato nella musica*. Il Mulino, 1992.
- [4] M. Imberty, "L'occhio e l'orecchio: Sequenza III di Berio," in *Il senso in musica*, pp. 163–186, 1987.
- [5] M. Doğan-Dack, *Artistic Practice as Research in Music: Theory, Criticism, Practice*. Routledge, 2015.
- [6] Fabio De Sanctis De Benedictis, "FDSDB_XXth_CT: una libreria di composizione algoritmica per Open Music e PWGL dedicata alle tecniche compositive del Novecento," in *XXII Colloquio di Informatica Musicale, Atti della Conferenza*, pp. 122–127, 2018.
- [7] F. De Sanctis De Benedictis, "Open Music come strumento didattico: il corso di fondamenti di composizione II," *d.a.t. [divulgazione audiotestuale]*, no. 5, pp. 3–45, 2019.
- [8] R. Morris, *Composition with Pitch Classes: A Theory of Compositional Design*. Yale University Press, 1987.
- [9] R. Morris, *Class Notes for Atonal Music Theory*. Frog Peak Music, 1991.
- [10] R. Morris, *Class Notes for Advanced Atonal Music Theory*. Frog Peak Music, 2001.
- [11] F. Donatoni, "Processo e figura," in *Il sigaro di Armando*, pp. 83–86, 1982.
- [12] F. De Sanctis De Benedictis, "Aroura: profezia di insensibilità?," in *Xenakis rimane*, pp. 244–290, 2023.
- [13] I. Bent, *Analisi musicale*. EDT, 1990.
- [14] A. Orcalli, *Fenomenologia della musica sperimentale*. Sonus Edizioni Musicali, 1993.
- [15] F. De Sanctis De Benedictis, "Algorithmic Composition with Open Music and Csound: two examples," in *Proceedings of the Fifth International Csound Conference*, pp. 102–108, 2019.
- [16] A. Cipriani and M. Giri, *Musica Elettronica e Sound Design, vol. 1*. Contemponet, 2009.
- [17] F. De Sanctis De Benedictis, "OpenMusic e Modalys per Nicola Baroni," in *Corpi fisici, Physical Bodies, Atti del XXIII CIM – Colloquio di Informatica Musicale*, 2022.
- [18] R. Boulanger, *The Csound Book*. The MIT Press, 2000.

Missing memories: why we need to analyse and notate spatialisation

Stefano Catena

De Montfort University
MTI²
Institute for Sonic Creativity
Leicester, UK
stefano.catena23@gmail.com

Enrico Dorigatti

University of Portsmouth
School of Creative Technologies
Portsmouth, UK
enrico@enicodorigatti.com

ABSTRACT

This paper aims to foster a discussion over the necessity of developing an analysis and classification framework of the spatial parameter in music, especially within the electroacoustic and acousmatic music contexts. As the exploration of space as an artistic feature is not novel in music, our discourse builds up from a review of the field, investigating its origins and identifying previous key experiences. This paper considers the state of the art of spatial analysis, comparing it with issues currently affecting electroacoustic and acousmatic music musicology, namely the lack of an agreed notation system and the technological fragmentation of the field. The following discussion proposes spatialisation as an experiential feature and recommends a perspective that contrast the dominant philosophy which interprets space from a geometrical and technological perspective. From this novel standpoint, we advocate that spatialisation should be studied independently from the technology employed, focusing only on aural implications. By inviting a conversation on this matter, we also suggest that the proposed perspective could work retrospectively, improving the analysis of electroacoustic pieces belonging to the historical repertoire.

1. INTRODUCTION AND BACKGROUND

The importance of using strong and widely adopted ontological systems for describing items within the same category is vital to ensure consistency in the field of interest. For example, music technology has witnessed this necessity with the development of wide databases of audio samples for automated sound recognition and categorisation: the organisation and classification of sounds is unequivocally a useful tool for sound artists and composers [1]. However, despite these recent advancements interesting the broader musical field and more specifically in the case of electroacoustic and acousmatic music, it still suffers the absence of effective frameworks for important parameters, one of which being spatialisation. Spatial cues are embedded into sounds whenever they are recorded or played back, and the human brain has developed in order to

be able to elaborate such information and understand what occurs around us. In fact, Blauert reminds us that there is no non-spatial hearing [2] and the spatial setting in which sounds 'happen' is as important as sound itself. This factor is a reminder that spatiality is an experiential aspect, difficult to study from an analytical perspective.

1.1 Historical approaches to spatialisation

The spatiality of sound has been a crucial point in the musical practice: certain peculiar spaces were chosen as musical venues even in prehistoric eras [3], but even in Greek and Roman periods, amphitheatres were built to harness architectural properties of the space. In the 15th century, instead, the first instances of 'spatialisation' can be observed: composers such as Gabrieli or Willaert would not just exploit reverberant spaces to create diffuseness, but would actively employ distant choirs to create antiphonal call-and-response gestures [4]. This musical style prevalent in Venice and performed mainly in Saint Mark's church was called *polychoral style*, or alternatively *cori spezzati*. Outside of Italy, composers such as Thomas Tallis with a 40-voice motet piece called *Spem in Alium* (1570) would employ a similar spatial approach with separated singers. In the classical period, instead, composers such as Mozart and Haydn would emulate echo effects with spatially separated orchestras: two notable examples are Mozart's *Notturno in D Major for four orchestras* (1977) and Haydn's *'Echo' Symphony n. 38* (1765-69) [3]. The explorations of spatialisation up until the 20th century are still relatively scarce, with a few exceptions of 'off-stage' musicians with Berlioz's *Symphonie Fantastique* (1830), Ives' *'Unanswered Questions'* (1908) or Giuseppe Verdi's *Requiem* (1874) [5]. Many contemporary composers have also experimented with spatialisation in the instrumental domain, mostly after World War II: Xenakis' *Terretektorh* (1966) or *Persephassa* (1969) [6] are two compositions that surround the listeners with musicians, both giving an immersive feeling and also removing the traditional division between audience and performer that has been in place for centuries. Composers like Stockhausen with *Gruppen* (1955-57) or *Carrè* (1959-60), Boulez, Brant or Schafer have used spatialisation actively in their instrumental musical practice [7]. A few notable examples include Brant's *Desert Forests* (2000) and *Ice Fields* (2001), Boulez's *Repons* (1981) or Schafer's *Third String Quartet* (1981). However, the true revolution of spatial exploration in music oc-

Copyright: ©2024 Stefano Catena et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

curs after the Second World War and the contextual introduction of the loudspeaker in the electroacoustic music practice. By placing loudspeakers in the desired position, composers were able to position and move electronic sounds with ease, making the spatialisation process much simpler than with acoustic instruments. Compositions such as Stockhausen's *Gesang der Junglinge* (1956), *Kontakte* (1960), *Spiral* (1968) during Osaka's Expo or his later *Oktophonie* (1991) [5] employed different layouts of speakers and various spatialisation techniques to achieve musical goals and a novel feeling of 'immersiveness'. Another notable example of a large multichannel system is Varese's *Poème Electronique* and Xenakis' *Concrete PH*, diffused in the Philips Pavilion during the World Fair of 1958 in Brussels; with between 300 and 450 speakers [8], the composer was able to create the illusion of spatial trajectories moving through the remarkable structure [4]. At the same time, the French school of *musique concrète* developed its own diffusion tools to 'enlarge' acousmatic compositions onto larger multichannel systems. Pioneered by Pierre Henry and Pierre Schaeffer in the 60s, the *potentiometre d'espace* utilized induction coils to manipulate the spatial distribution of sounds [4]. Later in the 70s, François Bayle developed the large and diversified multichannel system that we know today as the Acousmonium [9]: with many loudspeakers of different size and frequency response, it was possible to 'project' sounds in space with a mixing console, where an 'interpreter' controls their motion and placement.

1.2 Technological developments for spatialisation

The technological advancements have deeply influenced the use and exploration of spatialisation in the electroacoustic and acousmatic communities. During the 70s, several experiments with computer music and audio recording led to some key groundbreaking changes for spatial audio in the sonic arts. John Chowning artistically explored the possibilities of Max Mathew's *MUSIC* software concerning spatialisation in a quadrasonic environment and physicoacoustic phenomena such as the Doppler shift, for example through his iconic FM piece *Turenas* [10]. In particular, his research has led the way to 'spatial illusion': creating a sonic illusion of moving sounds through amplitude manipulation and global/local reverberation, alongside the Doppler effect [11]. Additionally, mathematician Michael Gerzon (and others) extended the technique of Mid/Side and coincident recording into what is known as First Order Ambisonics (FOA) [12]. The Ambisonics format does not carry speaker signals but rather contains a recording of the whole sound field called B-format, which is eventually decoded for an arbitrary speaker layout [12]. Another remarkable technology that allows for immersive spatial textures is Wave Field Synthesis: this spatialisation technique is capable of reproducing a physical wave field without the need for a sweet spot [13]. Differently from Ambisonics, Wave Field Synthesis can create any spatial image accurately for every listener regardless from their position in the room. However, Wave Field Synthesis systems require sophisticated software and a very large

number of tightly positioned loudspeakers, making it fairly complicated to access and, most importantly, effectively use in composition. First Order Ambisonics, instead, requires that the listening position is in the centre of the multichannel system, or otherwise the spatial image may be distorted or altered. This effect is mitigated by increasing the number of 'spatial harmonics' in the Ambisonics B-format, also called Higher Order Ambisonics: for fifth order (36 channels of audio), the sweet spot is large enough to almost encompass the whole area between any multichannel system [12].

Nowadays, new technologies such as the icosahedral loudspeaker (IKO) are used as technical and musical tools: first built in 2006 at IEM, it features 20 loudspeakers and Third Order Ambisonics, alongside early reflections, to create immersive soundscapes with one single instrument [14]. Dolby Atmos, introduced in 2012 and specifically developed for the cinema [15], uses an 'object based' ¹ approach to spatialisation, and has brought large multichannel spatialisation to commercial cinema and even home theatres, following standard layouts such as 5.1 and 7.1 [16]. While Dolby Atmos has become the standard in commercial applications, theaters and cinemas, it is still relatively unused in the electroacoustic artistic practice: it's not well suited for current HDLAs ², since it requires one of the aforementioned standard speaker layouts. This makes it difficult to integrate in spatial compositional workflows, where octophonic speaker setups are still a vast majority for acousmatic and electroacoustic reproduction.

However, in this paper we are only going to deal with the musical and the acousmatic: spatialisation techniques and developments for audiovisual medium will not be considered for sake of space and consistency.

2. SPATIAL ANALYSIS: STATE OF THE ART

Even though no shared framework for spatial analysis exists yet [17], understanding and categorising spatialisation and spatiality has always been a challenge for the musicological field of electroacoustic music. By spatialisation we mean the 'intentional placement (and motion) of sound within a given space' as suggested by Holbrook [18]. As described in section 1, between the 50s and the 80s, many experiments with spatialisation have been conducted to enhance the distribution and diffusion of sound through physical space. However, even though the interest in spatial composition has exponentially risen and the technological output of technological research in spatial audio has been remarkable, the musicological perspective on spatialisation did not witness a parallel interest. Nonetheless, some attempts to categorise and describe spatial events have been made.

However, before delving into existing spatial analysis frameworks, an important distinction must be made between electroacoustic approaches to spatialisation. Two macro-categories of spatialisation modes can be observed:

¹ 'object based' approach of spatialisation treats sound sources as independent objects, allowing audio to be rendered dynamically according to the speaker layout

² High Density Loudspeaker Arrays

live diffusion or 'fixed' mode of delivery [19].

- Live diffusion: this spatialisation approach historically enlarges a stereophonic composition onto a multichannel system by controlling the amplitude of signals sent to groups or single speakers [20, 21]. This allows an 'interpreter' to move or position sounds in space arbitrarily, following their own understanding of the sonic material. This is the mode of delivery that *musique concrète* has adopted since its beginning, through the multichannel system called Acousmonium [9]. However, several large multichannel systems like the Birmingham ElectroAcoustic Music Theatre (BEAST) have researched live diffusion and developed innovative ways and layouts to enhance and enrich the acousmatic musical experience [22].
- Fixed: in this spatialisation approach the spatial parameter is engrained in the composition—one could argue that it is *scored*—and written onto a fixed media. This way, the composer can define precisely how the spatial parameter will influence the music and the preferred method of reproduction. Smalley's *composed space* [23] refers to this mode of delivery.

While these modes of delivery relate to the same musical parameter, they are opposite both in technical and aesthetic concerns. Live diffusion entails a very subjective interpretation of the musical material, requires preparation for a performance and, in some cases, the rehearsing of chosen spatial 'gestures' that will reinforce the acousmatic writing [24]. Fixed media, instead, does not have any live component to the performance, but can be limited by the spatialisation technology, the multichannel layout, or the listening space itself. For example, using a 'channel based'³ approach rather than an 'object based' can make playback on a different multichannel system very complicated.

2.1 Existing spatial analysis frameworks

English composer Trevor Wishart discusses historical approaches to spatialisation in acousmatic music and proposes a large ontology of spatial motions, moving through nine 'distinguishable spatial positions', along with their related purpose and possible counterpoint with other motions [25]. However, this approach to space is inevitably 'geometrical', where sound motions follow specific lines or trajectories, and evidence shows how it is very complicated for the human perception to precisely track these complex movements through space [26]. Conversely, Emmerson and Smalley both propose a scene-based approach to define meaningful compositional aspects of space [27]. Smalley, in particular, defines and creates a large ontology of spatial constructs, from *personal spaces* to *microphone spaces*, from the *performed space* to the *perspective space* and much more [19]. Emmerson, instead, suggests the notion of *frames*, a 'defined spatial areas of interest', that are progressively applied from 'local' to 'field', from smallest

³ By 'channel based' it is intended a spatialisation mode that links a single channel of the final output to a specific speaker

to largest [28]. Kendall proposes a description of spatial dimensional attributes such as width, depth, distance and direction or immersive attributes such as envelopment and presence [17]. Catena, instead, tries to establish a basic unit for the categorisation and description of spatial events called *spatial sonorous object* [7], taking into account musical attributes of sound (called *sounding properties*) and only their spatial behaviour. Catena's analytical method, however, deals only with *composed space* in fixed media, which is a crucial distinction in the approach to spatiality. The rationale behind this choice is that the two modes described earlier, live diffusion and fixed media, must be analysed separately, and different parameters must be considered. For instance, in live diffusion, a crucial aspect to analyse is the gestural component of the performer, which is non-existent in fixed-media reproduction.

3. CURRENT CHALLENGES AND LIMITATIONS

The technologies resulting from the first experiments with spatialisation were isolated and unrelated to one another, resulting in several different approaches and, consequently, differences in spatial organisation. In other words, every composer would develop and employ their own spatialisation strategies and creative thinking, leading to a great diversification of spatialisation uses. This fragmentation, which we would argue is still somewhat active today [29], has complicated the analysis of spatialisation in electroacoustic music. Moreover, the lack of the framework this article advocates for—i.e. a basic system of shared parameters that defines the analytical characteristics of a spatial event—amplifies this fragmentation and reduces our musicological understanding of the *why* and *how* spatialisation is used.

3.1 Technological barriers

Technology, as described in section 1, had (and still has) a very strong influence on the way artists deal with spatialisation. The very introduction of the loudspeaker in the musical practice is an obvious example of technology shaping the compositional thought, not just in the spatial domain. The fact that loudspeakers could be placed and addressed individually, in stereo couples or in a single large sound field has led to heterogeneous spatial thinking, as well as compositional results. A beautiful example of spatial heterogeneity in the same domain is the whole *ReCycle* suite of works by Jonty Harrison. Each of the four pieces uses a slightly different but incompatible spatial layout of an 8-channel ring, making every piece unique while also showcasing the different ways to organise composition spatially. For instance, his work *FreeFall* uses a layout called 'double diamond', offering him five different planes in the front-rear axis [30]. Conversely, in *Internal Combustion* the eight speakers are managed in stereo pairs, referred to by Harrison himself as the 'shoe box' [31]. These examples show how there is a clear lack of standardization in managing multichannel arrays [31]. This is even more noticeable when investigating the broader technological approach to multichannel systems: the Acousmonium and the BEAST,

for example, have two completely dissimilar modes of delivery and spatial thinking. This lack of standardization may lead composers to approach these spaces similarly to 'site-specific' installations, similarly to how Varese treated the Brussels Pavilion [31].

On the other side of the issue, the same inconsistency in standards exists in the software domain. There is a saturation of free and commercial plugins and software for spatialisation, exacerbating the fragmentation of spatial thinking. Moreover, as stated in section 2, the proliferation of new tools for spatialisation has been exponentially faster than its musical side. This issue left a large gap in understanding the reasons behind spatialisation approaches, methods, and the function of the spatial parameter inside a compositional structure. Indeed, the field of music technology has been much more receptive to *how* we spatialise, rather than *why*, not to mention *for whom*.

3.2 A brief overview of the evolution of notation

The notation systems composers employed to write their music changed and evolved over time and strove to develop a new semiographic vocabulary [32]. Amongst all, this ongoing revolution would allow them to represent in the best way possible the new instruments introduced in their music and the new playing techniques required to the performers. These latter, commonly known as *extended techniques* [33, 34], in turn evolved, encompassing tasks going beyond the bare musical field, narrowing the gap between the musical performance and the theatrical one and bridging the world of musical (e.g. tuned) sounds with the scope of non-musical ones⁴ [32]. The ever-growing and non-consistent symbolic palette employed by composers, and the attempts to represent the music visually as a score, tried to keep pace with the ongoing aesthetic and artistic revolution. These efforts towards the exploration of new notation strategies, which also witnessed the creation of advanced tools such as the UPIC system developed by Xenakis [35], were not simply another side of such revolution happening in music, but were rather driven by the necessity composers had to represent their music and thus make it readable to others, such as performers. It is worth mentioning that, as part of the changes happening, the very essence of the compositional practice and the score was challenged and subverted, with many pieces relying not only on the extended techniques and non-musical sounds mentioned but also on processes—either similar to algorithms or purely chance-based—employed by the composer or extemporaneously by the performer. John Cage is well-known for the exploration of such processes in conceptual (e.g. in his well-known 4'33" piece) and practical terms, such as through the I-Ching mechanism [36], and his compendium of scores [37] offers a glimpse on how avantgarde composers started to intend the score, which, after having encompassed an extended palette of symbols and alphanumeric instructions, started to employ shapes and colours much as they were pieces of abstract visual art. Examples are Franco Donatoni's *Babai* for harpsichord solo, Cor-

⁴ Stockhausen's *Helikopter-Streichquartett*, despite being a quite late work, offers a good example of this new conception of music.

nelius Cardew's *Treatise*, or Joe Jone's *Five Pieces for Piano*.

3.2.1 Notation in electroacoustic music

One of the main recent milestones in the development of music was the introduction of electric and electronic instruments and technologies. Explored consistently from the fifties, they opened to the possibility of working with completely abstract, synthesised sounds on the one hand, and, on the other, with concrete, field-recorded samples.

In terms of notation, these technological developments opened to unexpected scenarios. In Smalley's words, "Electroacoustic music opens access to all sounds, a bewildering sonic array ranging from the real to the surreal and beyond" [23, p. 107] and the realisation of a notation system encompassing all these possibilities was far than trivial. Emmerson acknowledges the new role of the score in electroacoustic music, shifting from a support essential to study and performance, to an artefact acquiring descriptive importance:

Scores for electroacoustic music exist, of course, but have a variety of functions: for performance with instruments or over complex loud-speaker systems, or simply for reading and background information, but rarely for the re-creation of the work itself. The score moves away from prescription towards description. The emphasis is in all cases away from a reliance upon the written hieroglyph as a means to express, or at least to transmit, musical utterance. [38, p. 2].

In the context of live electronics or mixed performances, scores reporting the variations to be applied to specific music parameters by the electronics performer or sound engineer started to be employed, such as by Luigi Nono in *A Pierre. Dell'azzurro silenzio, inquietum* [39] or *Postprae-ludium n. 1 "per Donau"*. Such expedients, however, although portraying faithfully the state of the parameters in time, say little or nothing about the actual music and perceptive result. Other composers working with tape, thus fixed, music started to employ visual scores rigorously reporting on graph paper shapes suggesting the evolution in time of pitch and spectra, also giving information on the dynamics through colour intensity—examples of purely descriptive supports, no matter their resolution and details. Examples are Stockhausen's *Studie II* and Franco Evangelisti's *Incontri di fasce sonore*. However, it is necessary to observe that these pieces, regarded as the vanguard in the exploration of tape and pure music, were composed of a fairly low amount of sonic events happening at the same time, which would make their detection, notation, and interpretation a doable process. As electroacoustic music evolved, pieces became more intricate thanks the advent of more powerful technologies, amongst which computers and software, with enabled composers to craft and alter sound objects with surgical precision at a microsound level [40]. This would make it difficult, if not impossible, to graphically notate sound events happen-

ing in time, and it is also necessary to consider the virtual impossibility of giving any possible notation system a shared, widely acknowledged, and consistent meaning, as well as the ability to convey perceptive information—i.e. clues about the actual sonic result. Despite such difficulties, frameworks for sound-based (electroacoustic) music notation were prototyped and proposed over time, as discussed in [41]. With the advent of software supporting FFT⁵ analysis, and despite some scepticism due to its limitations [23], spectrograms became the unofficial standard in electroacoustic music notation, possibly accompanied by the waveform of the piece. This combination would allow analysts to spot changes occurring in dynamics and frequency content, providing some insights into the morphology of the sound object but still without providing any information about the timbre and other perceptual characteristics. Software such as the *Acousmographie*⁶ were specifically developed to make it possible for composers to draw symbols on the spectrogram of their pieces to highlight relevant musical passages. In a sense, this operation would coincide with realising and providing a personal notation layer over the surgical, objective one provided by the spectrogram-waveform pair. However, given the overall niche scope of electroacoustic music, tools such as the *Acousmographie* and similar often remain confined to a pedagogical or experimental, local dimension—i.e. hardly known beyond the network of the developers and scholarly conference attendees.

The surge in the exploration of notation systems through the research of the avantgarde composers ended with the advent of electronic music where, besides the impossibility of creating a system able to consistently represent the whole sonic array employed, the necessity of graphically notating a piece for performance purposes is marginal [38, 23]. And in such cases, the actions required to the electronics performer can be, most of the times, schematised through diagrams and timestamps, thus making a visual representation of the music itself not necessary at all.

3.2.2 Notating space in electroacoustic music

However, different is the situation when it comes to the musicological analysis of electroacoustic music, where the presence of a score would greatly benefit the analyst by bridging the gaps left open by the limitations of our auditory system. The development of a modern notation system for electroacoustic music, thus, should focus on the perceptive qualities of a piece, and amongst the usual musical parameters, it should make it possible to analyse and represent, through an ad-hoc framework, the spatial parameter. There is to consider, indeed, that modern electroacoustic composers, thanks to the advancements in the field, increasingly explore space as a musical parameter, incorporating sound diffusion strategies in the 3D environment as a compositional and aesthetic feature functional to the representation and aural embodiment of their artistic vision and idea. However, as no notation system has ever been

standardised and agreed upon to represent musical parameters in electroacoustic music—besides the spectrogram, which however falls short in returning a usable and informative picture of how space is employed within a composition—likewise nothing has been realised specifically for the spatial parameter. Some scholars and composers, by proposing their vision and taxonomies of spatial trajectories, somehow represent an exception to the previous statement. However, as seen in section 2.1, their proposals for the analysis and representation of space focus on the geometrical properties rather than on the perceptual qualities of spatial movements.

4. WHY, HOW AND FOR WHOM WE SHOULD ANALYSE SPACE

4.1 Why: missing memories

One of the most important argument advocating for the establishment of spatial analysis and notation is how we, as musicians and musicologists, can better keep track of compositional strategies for spatialisation. Many electroacoustic pieces are never analysed from a spatial perspective: many times the only focus of the analysis is spectromorphological, i.e. investigating the spectral development and relationship inside the composition, discarding spatial ones. This is true not just in electroacoustic music, but also in traditional and contemporary instrumental composition: while rarer before 1950 (as described in section 1), spatialisation of sound became common after World War II. However, many contemporary pieces are still only analysed from a 'sonic' perspective, discarding how space influences musical perception, even on a timbral level. We argue that for electroacoustic and even more acousmatic music—where the only medium is sound—this is not good enough. If we want to describe a piece of music and understand the composer's intentions and their techniques and compositional strategies, we must take into account the spatialisation approach from both the creator's and the listeners' perspectives.

Re-evaluating 'legacy' pieces with a new system of analysis and notation would be a valuable method to retrieve the missing memories of space. While many pieces in the electroacoustic arts make use of spatialisation, they are mere descriptions of technical means, but very few illustrate a musical end that such technologies are made to achieve. For example, Chowning's simulation of moving sound sources experiments [11], while being invaluable from a technical and technological perspective, do not describe any musical relationship between sound and spatial motions, other than deeming Lissajous trajectories as 'natural' [10]. By analysing and notating the piece more than 50 years later, a new perspective could be given to the music and the relationship between its sounds and spaces. This is true not just for Chowning's music, of course, but also for all composers that have investigated musically the active use of spatialisation.

⁵ Fast Fourier Transformation

⁶ <https://inagrm.com/en/showcase/news/203/acousmographie>. Last access: 08 July 2023.

4.2 How: a strive for simplicity

Understanding and categorising spatialisation is arguably not easy. Space is embedded in our everyday sonic experience, and it is impossible to separate sound from its spatial environment. Even sound recorded in an anechoic chamber communicates its lack of reverberation to the human brain, and we understand it and perceive it as such. This becomes even more complicated when musical factors come into play: composers manage and deploy sounds through space arbitrarily, using the technologies at their disposal. In section 2 we introduced examples of frameworks for the analysis of spatialisation in electroacoustic arts: however, none of them have been largely adopted or further developed into a more encompassing notation system for such analysis. Wishart's approach, for instance, only discusses spatial motions, not considering that statically positioned sounds may have a role in a compositional structure: many composers in the 19th century already demonstrated this applicability. Moreover, Wishart's trajectories are extremely complicated to precisely track aurally—e.g. recognising a front-to-back 'figure of 8' spatial motion is troublesome—making his framework difficult to employ practically. On the other hand, Smalley provides a very large ontology of 'space-forms': however, they are not always strictly relating to spatialisation. For example, his concept of 'Spectral space' is defined as "the impression of space and spaciousness produced by occupancy of, and motion within, the range of audible frequencies" [19]. It relates to space as recorded or generated, but it is an impression, not necessarily a conscious act by the composer to place or move sound through space. This ontology, while comprehensive, is not suitable to be a theoretical bedrock for the creation of a notation system that describes spatialisation. Following Catena's work [7], instead, a basic unit for the description of spatialisation is introduced. The *spatial sonorous object* is an arguably simple concept that only deals with three key components:

- sound movement or its placement, which Smalley calls 'spatiomorphology' [23];
- the relationship between sounding properties and spatial morphology;
- its compositional function;

In fact, this approach goes against the grain compared to other analytical frameworks: by defining spatial events with three quantifiable descriptors it is possible to encapsulate them into a single coherent unit. This method, moreover, does not take into account any technological issue: by excluding the practical methods of spatialisation from the analysis—discarding whether Ambisonics, amplitude panning, WaveField Synthesis etc.. or even which speaker layout is used in the process—we can focus only on the musically valuable information. The value can be found in what we can hear and perceive as listeners, which is also the foundation of electroacoustic music analysis [42].

As previously noted, we believe that the listeners' reception of spatial events is crucial. Even more so when talking about acousmatic music where no visual stimuli

are present, and the experience intended by the composer is mediated only through sound. We have also described how many composers have a very 'Euclidean' approach to spatialisation, where each sound source is described precisely by a set of geometric values representing space (x,y,z). We would take a more 'relativist' stand, where it is more important that sound *moves*, rather than its precise trajectory [43]. Instead of looking at every single motion, it would be more effective to take a holistic approach and look at a whole 'spatial scene', meaning a complete musical texture, and analyse its perceptual properties as received by the person experiencing the music. We believe that a phenomenological approach to creating a spatial vocabulary and notation established via the aural experience is the best way to provide a simple and accessible shared language grounded in perceptually relevant attributes.

4.3 For whom: the listeners' experience

Space and its reception is a very personal experience. While spatial listening is largely auditory and we hear sound in a specific manner due to our physical shape, a big portion of how we interpret sounds is cultural. Hearing a voice speaking in one's mother tongue will be a strong factor in understanding a compositional choice inside the music, for instance. One example of this cultural circumstance relating to space can be found in Trevor Wishart's *Encounters in the Republic of Heaven*. This is a work that is almost exclusively based on spoken words, speech, and vowel sounds. During the work, several transformations are applied to the vocal sounds, creating heterogeneous textures: these sounds are spatialised in an 8-channel environment to create a sense of immersiveness [44], but also to reference their original material. At 2:32 of the third act, a lady keeps repeating the phrase "round and round". The words are looped, and as soon as the listener hears that the speech is unnatural due to the looping effect, the sound begins to move around the listening space. The meaning of the English words is linked to a specific spatialisation choice. This is a fascinating compositional technique, but if non-English speakers were to listen to the piece, they would not be able to catch the reference, unless explained through dramaturgic information. Another very important aspect to consider in relation to listeners' experience is the aforementioned relationship between sounding properties and spatial morphologies. How do listeners of electroacoustic art connect the spatialisation of sound to its spectromorphology? Is there something about the spatial aspect of sound that may help or aid in better understanding the music? These examples show that the listeners' experience, in the broader sense of the term, is paramount when analysing spatialisation. However, it is also important to study the reception of spatialisation from listeners: while the composers may be clear in what they want to achieve by spatial means, it is not granted that a listener understands their intentions. Shifting the focus to the experiential aspect will be invaluable to better understand spatial organisation and improve the reach of electroacoustic and acousmatic music. It is crucial to point out how it is more relevant to understand and analyse these problems from a

musical perspective rather than a scientific one: this means to investigate *reception* (musical and experiential) rather than *perception* (scientific).

5. CONCLUSIONS

This paper aims to be a starting point for the discussion on the development of a shared framework for the analysis and notation of spatialisation in electroacoustic music. While many differences in genres can be observed among the several branches of electroacoustic music—and as such they need to be analysed—the focal point of this study must necessarily include listeners’ reception. We have provided an overview of spatialisation throughout music history and existing spatial analysis frameworks and notation systems but also discussed downsides and possible new approaches that may be significantly valuable in this field. Moreover, we have suggested how technological improvements have aided spatialisation techniques, but have been arguably challenging in studying spatial events coherently. The development of the proposed framework will aid both the understanding of current spatialisation approaches and their listeners’ reception, but also in re-evaluating historical pieces under a new perspective that includes their spatial organisation.

6. REFERENCES

- [1] D. A. Mauro and A. Valle, “Sound ontologies: Methods and approaches for the description of sound,” in *Foundations in Sound Design for Linear Media*, pp. 283–302, Routledge, 2019.
- [2] J. Blauert, *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.
- [3] B. Boren, “History of 3d sound,” in *Immersive sound*, pp. 40–62, Routledge, 2017.
- [4] J. W. Solomon, *Spatialization in music: the analysis and interpretation of spatial gestures*. PhD thesis, University of Georgia, 2007.
- [5] R. Zvonar, “A history of spatial music,” *eContact*, Vol. 7(4), *Canadian Electroacoustic Community*, 2004.
- [6] B. Hofmann, “Spatial aspects in xenakis’ instrumental works,” in *Proc. of the International Symposium Iannis Xenakis* (M. Solomos, A. Georgaki, and G. Zervos, eds.), pp. 193–201, 5 2005.
- [7] S. Catena, “Concepts and approaches in analysing spatial sound movements: a link between mozart and acousmatic music,” in *Proceedings of the XXIII CIM (Colloqui d’Informatica Musicale)*, pp. 105–111, 2022.
- [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] J. Chowning, “Turenas: the realization of a dream,” *Journées d’Informatique Musicale*, 5 2011.
- [11] J. M. Chowning, “The simulation of moving sound sources,” *Computer Music Journal*, pp. 48–52, 1977.
- [12] F. Zotter and M. Frank, *A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, vol. 19. Springer Cham, 1 ed., 5 2019.
- [13] M. A. Baalman, “Spatial composition techniques and sound spatialisation technologies,” *Organised Sound*, vol. 15, pp. 209–218, 12 2010.
- [14] F. Schultz, M. Zaunschirm, and F. Zotter, “Directivity and electro-acoustic measurements of the iko,” in *Audio Engineering Society Convention 144*, Audio Engineering Society, 2018.
- [15] D. Baxter, *Immersive Sound Production: A Practical Guide*. Focal Press, 2022.
- [16] G. Sergi, “Knocking at the door of cinematic artifice: Dolby atmos, challenges and opportunities,” *The New Soundtrack*, vol. 3, no. 2, pp. 107–121, 2013.
- [17] G. S. Kendall and M. Ardila, “The artistic play of spatial organization: Spatial attributes, scene analysis and auditory spatial schemata,” in *4th International Symposium on Computer Music Modeling and Retrieval (CMMR)*, pp. 125–138, Springer, 2007.
- [18] U. A. Holbrook, “Sound objects and spatial morphologies,” *Organised Sound*, vol. 24, pp. 20–29, 4 2019.
- [19] D. Smalley, “Space-form and the acousmatic image,” *Organised Sound*, vol. 12, pp. 35–58, 2007.
- [20] N. Barrett, “Spatio-musical composition strategies,” *Organised Sound*, vol. 7, pp. 313–323, 2002.
- [21] N. Stavropoulos, “Multi-channel formats in electroacoustic composition: Acoustic space as a carrier of musical structure,” in *Proceedings of the Digital Music Research Network Conference, London, UK*, 2006.
- [22] J. Harrison, “Sound, space, sculpture: some thoughts on the ‘what’, ‘how’ and ‘why’ of sound diffusion,” *Organised Sound*, vol. 3, pp. 117–127, 8 1998.
- [23] D. Smalley, “Spectromorphology: explaining sound-shapes,” *Organised Sound*, 2(2), pp. 107–126, 1997.
- [24] A. V. Gorne, “L’interprétation spatiale. essai de formalisation méthodologique,” December 2002. Université de Lille.
- [25] T. Wishart, *On sonic art*, vol. 12. Psychology Press, 1996.

- [26] F. Schumacher, V. Espinoza, F. Mardones, R. Vergara, A. Aránguiz, and V. Aguilera, “Perceptual recognition of sound trajectories in space,” *Computer Music Journal*, vol. 45, pp. 39–54, 4 2022.
- [27] E. Deleflie, *Music and Spatial Verisimilitude*. PhD thesis, University of Wollongong, 2013.
- [28] S. Emmerson, “Aural landscape: musical space,” *Organised Sound*, vol. 3, no. 2, p. 135–140, 1998.
- [29] 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.
- [30] J. Harrison, “Round the world in sixty minutes: approaches to the evocation of space, place and location in recent multichannel works,” *Ouvirouver - EDUFU - Editora da Universidade Federal de Uberlândia*, vol. 12, pp. 14–29, 7 2016.
- [31] J. Fielder, “A history of the development of multichannel speaker arrays for the presentation and diffusion acousmatic music.” Unpublished manuscript. Retrieved from <https://api.semanticscholar.org/CorpusID:198233525>, 2016.
- [32] N. Nedelcut, “Semiographic aspects in the music of the 20th century,” *Journal of Science and Arts. Supplement at*, no. 2, p. 13, 2010.
- [33] C. Dimpker, *Extended Notation: The depiction of the unconventional*, vol. 17. LIT Verlag Münster, 2013.
- [34] M. Solomos, *From music to sound: The emergence of sound in 20th-and 21st-century music*. Routledge, 2019.
- [35] G. Marino, M.-H. Serra, and J.-M. Raczinski, “The upic system: Origins and innovations,” *Perspectives of New Music*, pp. 258–269, 1993.
- [36] M. G. Jensen, “John cage, chance operations, and the chaos game: Cage and the” i ching”,” *The Musical Times*, vol. 150, no. 1907, pp. 97–102, 2009.
- [37] J. Cage, *Notations*. Something Else Press, 1969.
- [38] S. Emmerson, *The language electroacoustic music*. Springer, 1986.
- [39] F. Sallis, “A spectral examination of nono’s ’a pierre. dell’ azzurro silenzio, inquietum,” in *Live Electronics Music: Composition, Performance, Study*, Routledge, 2018.
- [40] C. Roads, *Microsound*. The MIT Press, 2001.
- [41] M. Sköld, “Notation as visual representation of sound-based music,” *Journal of New Music Research*, vol. 51, no. 2-3, pp. 186–202, 2022.
- [42] L. Zattra, “Analysis and analyses of electroacoustic music,” *Sound and Music Computing (SMC05)*, Salerno, Italy, vol. 36, 2005.
- [43] J. Harrison and S. Wilson, “Organised sound: Editorial,” 12 2010.
- [44] T. Wishart, “Encounters in the republic of heaven,” *Musica/Tecnologia*, pp. 9–33, 2020.

ARCHEOTOPOLOGIE: IMPLEMENTAZIONE CRITICA DI MEMORIE SENZA COLORE

Daniele Giuseppe Anese
Conservatorio “N. Piccinni” di Bari
dan.annese1@gmail.com

Francesco Vitucci
Conservatorio “N. Piccinni” di Bari
francescovitucci1@gmail.com

Anthony Di Furia
Conservatorio “N. Piccinni” di Bari
anthonydifuria.sound@gmail.com

Francesco Scagliola
Conservatorio “N. Piccinni” di Bari
francesco.scagliola@gmail.com

Giuseppe Silvi
Conservatorio “N. Piccinni” di Bari
grammaton@me.com

ABSTRACT

Nello studio della musica elettronica l’osservazione critica di algoritmi e architetture informatiche archetipiche è un esercizio tecnico, stilistico, paragonabile allo studio di passi del repertorio per lo strumento musicale acustico. Nello specifico, si esercitano capacità di elaborazione del segnale a fini didattici e di sostenibilità del patrimonio musicale e culturale.

La letteratura di riferimento a cui si attinge è quella relativa alla nascita delle tecniche di riverberazione digitale mediante costruzione e relazioni di filtri. I testi di Manfred R. Schroeder rappresentano momenti di approfondimento matematico e fisico, la cui ri-costruzione offre inevitabili riflessioni di carattere scientifico, informatico e musicale.

L’implementazione mediante linguaggi informatici testuali e opportuni modelli di analisi del segnale, consente l’accesso alla letteratura storica in una chiave di comprensione contemporanea, garantisce continuità futura del testo e apre prospettive di costruzione algoritmica nuova.

1. INTRODUZIONE

Ad un approccio superficiale, considerato l’impiego di strumenti tecnologici nati nel secondo novecento, può accadere che la letteratura informatica ed elettroacustica siano ritenute molto più giovani della letteratura in fisica acustica. Tuttavia, nonostante l’attenzione architettonica a fenomeni sonori sia stata per millenni parte integrante delle capacità progettuali di strutture, la fisica acustica ha ottenuto una base scientifica solida solo dai primi anni del novecento, ad opera di Wallace Sabine. [1]

L’osservazione dei fenomeni di riflessione e riverberazione, la loro descrizione scientifica, fino alla definizione controllata dei tempi di riverberazione da parte di Sabine sono punti di partenza anche nella letteratura sulla modellazione digitale del riverberatore, fin dal principio, nell’opera di Manfred Schroeder. [2, 3] A tale proposito è stata

realizzata una traduzione integrale del testo *Natural Sounding Artificial Reverberation*¹ [2] al quale il presente articolo fa riferimento per una implementazione critica che aiuti a comprendere la specificità dell’intervento di Schroeder. I due testi fanno da supporto didattico per la costruzione passo passo dei riverberatori con strumenti di analisi e grafici da confrontare con quelli proposti dall’autore e mirano alla riscrittura e all’ampliamento dei processi per dare vita a nuovi strumenti musicali. [4]

Come per Sabine, chiamato a risolvere problematiche di carattere acustico, Schroeder rende possibile un avanzamento scientifico legato al riverbero approssiando alla soluzione di alcuni problemi di stabilità e linearità in frequenza in relazione ai riverberi elettronici disponibili all’epoca. È consapevole delle criticità, che mette in chiaro fin dal principio: la diffusione di un riverbero necessita di un numero minimo di 1000 echi per secondo per non esprimere fastidiose fluttuazioni. Inoltre questa diffusione deve avvenire senza danneggiare il contributo timbrico della sorgente, cosa che accadeva con i riverberi dell’epoca:

1. Le loro *risposte in ampiezza e in frequenza* non sono piatte. Infatti, deviano da una risposta piatta così tanto che è possibile percepire una sgradevole “colorazione”. [...]
2. La *densità degli echi* (i.e., il numero di echi al secondo che l’output del riverberatore genera quando viene stimolato da un singolo impulso all’input) è troppo bassa comparata con la densità degli echi di un ambiente acustico reale. [...]

L’implementazione segue la struttura narrativa del testo di Schroeder costruendo i due blocchi basilari *Comb Filter* e *All-Pass Filter* e le successive elaborazioni ed interrelazioni tra questi. La scrittura su *Faust* delle funzioni consente una documentazione mediante diagramma a blocchi e l’analisi immediata della risposta all’impulso nei domini di tempo e frequenza. Il confronto di questi dati, con i materiali di analisi forniti dall’autore insieme alla documentazione matematica, costruiscono una visione ampia e robusta sui problemi di implementazione *sample-rate*.

Copyright: © 2024 Daniele Giuseppe Anese et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ <https://github.com/s-e-a-m/2024-CIM-MSNSAR/blob/main/rsrc/mnsar-it/build/2024-mnsar-it.pdf> - Daniele Giuseppe Anese, 2024.

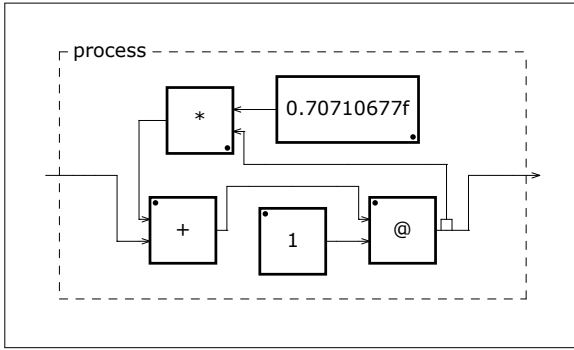


Figure 1. Diagramma a blocchi generato con la funzione `df1` in *Faust*, apparentemente identica a quella proposta da Schroeder.

2. IMPLEMENTAZIONE CRITICA DEI COMPONENTI PRINCIPALI

L'articolo [2] propone un'analisi matematica dei processi e descrive implementazioni informatiche *archetipiche* mediante diagrammi a blocchi e stampe di risposte all'impulso. *Faust* (*Functional Audio Stream*)² è un linguaggio di programmazione funzionale per la sintesi del suono e l'elaborazione audio. La sua sintassi consente di scrivere algoritmi nel dominio del tempo che vengono rappresentati come diagrammi a blocchi. La scrittura in ambienti *samplerate* come *Faust* produce quindi diagrammi immediatamente confrontabili con gli originali ma deve tener conto di problematiche di elaborazione numerica esperibili solo con un metodo di analisi dei risultati confrontabile con quello proposto dall'autore. A tal proposito si è scelto di elaborare i dati delle risposte all'impulso provenienti da *Faust* nell'ambiente di calcolo numerico *Wolfram Mathematica*³ al fine di ottenere rappresentazioni grafiche chiare e comprensibili. Ciascuna delle funzioni implementate è disponibile nell'omonima libreria *Faust* di *SEAM - Sustainable Electro-Acoustic Music*⁴, progetto fondato sul concetto di sostenibilità della pratica musicale elettroacustica dal vivo con l'idea di far crescere la consapevolezza dei problemi implementativi elettronici e informatici che questa pratica ha affrontato nel tempo.

2.1 Delay in Feedback Loop

Il percorso di costruzione di un sistema di riverberazione in grado di dare densità alle riflessioni e risposta timbrica a magnitudine lineare parte, per Manfred Schroeder, da una linea di ritardo, capace di restituire un singolo eco dopo un tempo di ritardo t . Un oggetto semplice, per cui ancora insufficiente a soddisfare gli obiettivi desiderati.

Per produrre echi multipli senza usare più delay (costosi), la linea di ritardo è inserita in un ciclo di feedback [...], con gain g , minore di uno (in maniera tale da mantenere stabile il ciclo).

²<https://faust.grame.fr>

³<https://www.wolfram.com/mathematica/>

⁴<https://github.com/s-e-a-m/faust-libraries>

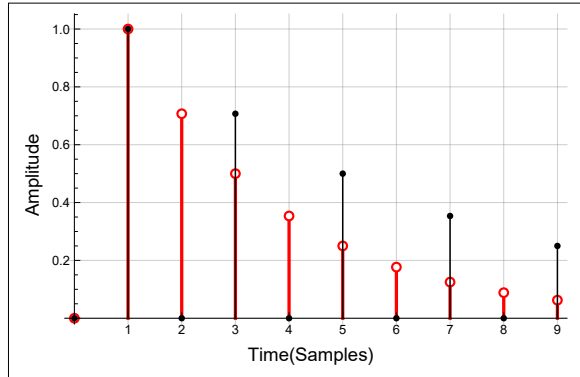


Figure 2. Comparazione tra le risposte all'impulso della funzione `df1` in *Wolfram Mathematica*. Il diagramma in rosso rappresenta la risposta originale di Schroeder, caratterizzata da riscaldamenti dell'impulso a periodo t , quello in nero rappresenta la risposta generata dal codice *Faust*, in cui i riscaldamenti avvengono a periodo $2t$.

La topologia che Schroeder descrive in queste righe è quella del più semplice filtro a struttura ricorsiva *IIR*. Il cuore attorno a cui ruota la linea di retroazione è un ritardo variabile, che cambia il comportamento temporale e quindi spettrale del filtro in funzione dell'unità di ritardo.

Schroeder battezza da subito questa struttura, dichiarandone l'individualità: *Delay in Feedback Loop* è il mattone elementare su cui si poggeranno tutte le architetture riverberanti successive.

$$df1(t, g) = (+ : @ (t)) \sim * (g);$$

Il codice *Faust* produce il diagramma a blocchi in Figura 1. Questo, coerente con il modello di Schroeder, può essere smontato e descritto in più parti per comprenderne la sintassi. Il blocco di ritardo è realizzato utilizzando la funzione `@`, una primitiva di *Faust* che esprime un tempo di ritardo t , intero e positivo, in quanto unità al campione. L'uscita del blocco di ritardo viene prelevata prima dell'uscita dal filtro mediante la composizione ricorsiva \sim e reindirizzata alla sua entrata, in somma con l'entrata del filtro; questo circuito di retroazione è controllato dal coefficiente $g < 1$. Il filtro prevede un valore di ritardo $t \geq 1$ applicato a tutti i campioni in entrata, situazione che può produrre diversi risultati sonori. La Figura 2 compara, sovrappoendole, le due risposte all'impulso: quella originaria di Schroeder e quella generata dalla funzione `df1`. Entrambe condividono la stessa tendenza, per cui ogni nuova occorrenza dell'impulso viene attenuata di un coefficiente g , e questo conferisce al segnale un decadimento esponenziale. Pur avendo un diagramma a blocchi apparentemente identico, le risposte all'impulso dei filtri si comportano in modo diverso, motivo per cui meritano uno sguardo di analisi e un'attenta riflessione.

La risposta all'impulso di Figura 2 è stata generata impostando le variabili $t = 1$ (un campione di ritardo) e $g = 1/\sqrt{2}$ (ampiezza scalata di $1/\sqrt{2}$ ad ogni ciclo di feedback). Ci si aspetta quindi un comportamento molto simile a quello descritto da Schroeder: per una successio-

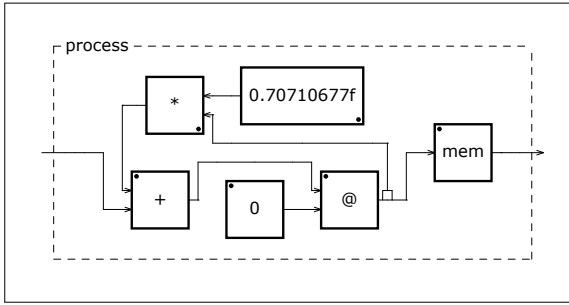


Figure 3. Diagramma a blocchi generato con la funzione `dflc` in *Faust*. La correzione temporale determina un $t - 1 = 0$ per $t = 1$. Il campione sottratto all'unità di ritardo viene ripristinato fuori dal ciclo di reiterazione che già conta, intrinsecamente, un campione di ritardo.

ne di multipli di 1 dovremmo avere il primo impulso nella posizione $n[1]$ con ampiezza $a = 1$ (non scalata, l'impulso non è ancor transitato nel ciclo di feedback). Il secondo impulso dovrebbe essere posizionato subito dopo il primo, nella posizione $n[2]$ con ampiezza $a = 1/\sqrt{2}$, e così proseguendo. Tuttavia, osservando l'indicizzazione dei campioni sull'asse delle ascisse si può constatare che in realtà il nostro filtro sta impiegando due campioni per ogni ciclo impulsivo in luogo di uno, restituendo gli impulsi a periodo $2t$. La motivazione di questa differenza sta nella natura degli ambienti di programmazione *samplerate* come *Faust*. Poichè qui il tempo scorre ad un intervallo minimo di un campione, qualsiasi tipo di operazione richiede almeno un campione per essere eseguita. Per cui, la composizione \sim produce, nel momento in cui la linea di retroazione ritorna in input, un inevitabile campione di ritardo. Per adeguare quindi, i risultati prodotti dall'implementazione *samplerate* del filtro a quelli numerici prospettati da Schroeder, occorre definire una nuova funzione `dflc`, corretta:

```
dflc(t,g) = (+ : @(t-1))~*(g) : mem;
```

Sottraendo il campione di ritardo, prodotto dalla ricorsione, alla variabile t con $t - 1$ si produrranno, per $t = 1$, zero campioni di ritardo all'uscita e un campione scalato di g all'entrata della somma in input. Questo, però genera un offset temporale in quanto la sequenza di campioni ritardati, ora a periodo corretto, si trova in uscita un campione in anticipo a causa del ritardo zero. Anche questa problematica è risolvibile inserendo un ulteriore campione di ritardo, utilizzando la primitiva `mem`, all'uscita del filtro, in modo da bilanciare l'intera sequenza temporale.

La Figura 4 mostra la risposta all'impulso del filtro corretto, ora coerente con le aspettative. Nel descrivere le caratteristiche del filtro nel dominio del tempo e della frequenza, Schroeder commenta:

La risposta in ampiezza-frequenza ha le sembianze di un pettine con massimi e minimi periodici [...]. È proprio questo sistema di picchi e valli che conferisce l'indesiderata qualità "colorata" al suono [...].

È da sottolineare il carattere sperimentale e pionieristico

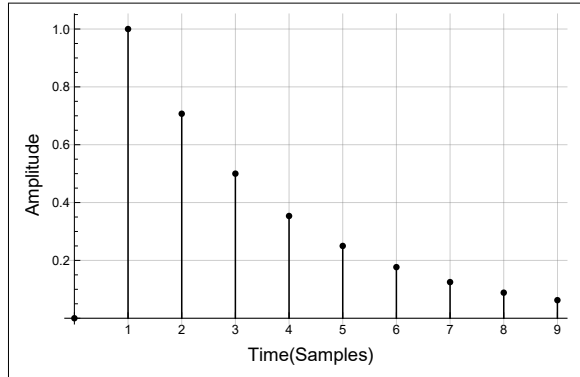


Figure 4. Risposta all'impulso della funzione `dflc` in *Wolfram Mathematica*, coerente con la rappresentazione di Schroeder, e caratterizzata da riscalamanti dell'impulso a periodo t .

delle implementazioni di Schroeder di quel periodo, al punto che la dicitura «ha le sembianze un pettine» poi si trasforma linguisticamente in *comb-filter*, filtro a *pettine*.

Il filtro appena costruito può quindi portare a diversi risultati sonori: dal semplice ritardo di una quantità di campioni (per $g = 0$) ad un filtraggio *low-pass* (con $t = 1$), ad un filtro *comb* (con $t > 1$).

2.2 All-Pass Filter

Sebbene il filtro *comb* sia sufficiente a generare echi multipli, la colorazione del segnale in uscita rappresenterebbe un ostacolo alla progettazione di un riverberatore artificiale con una risposta in frequenza piatta. Per eliminare questo problema Schroeder propone quindi di miscelare il segnale ritardato uscente dal filtro *comb* con una certa quantità di segnale diretto.

In altre parole, l'aggiunta di un percorso non ritardato opportunamente proporzionato, ha convertito il filtro *comb* [...] in un filtro *all-pass*⁵ [...]. Ciò non è un semplice risultato accademico. La conversione [...] è accompagnata da un miglioramento marcato nella qualità del suono riverberato dal carattere cavo del precedente alla qualità del successivo, perfettamente "incoloro".

Consultando i diagrammi a blocchi originali, è evidente l'intenzione di Schroeder, nella rappresentazione del filtro *comb* di predisporlo anche solo graficamente per essere inserito nel futuro *all-pass*. Per questo motivo, l'implementazione *Faust*, segue la stessa logica di stratificazione (tipica dei linguaggi di programmazione funzionale) per cui, la definizione di funzioni semplici (`dflc`) viene poi utilizzata all'interno di funzioni più complesse (`apf`).

Dal diagramma a blocchi di Figura 5: mediante la composizione divisiva `<:`, ciascun impulso in ingresso $n[x]$

⁵ *ndt* si usa il nome *all-pass* non tradotto per definire il tipo di filtro, ideato in quel periodo storico dallo stesso autore, con la caratteristica "passa tutto".

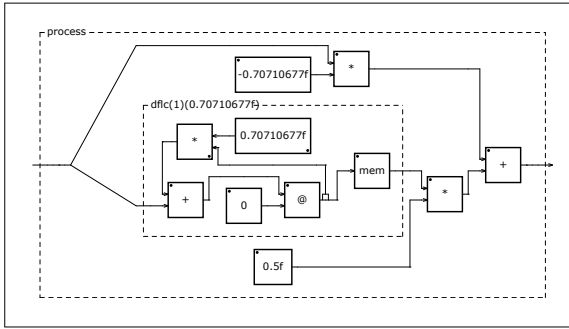


Figure 5. Diagramma a blocchi generato con la funzione `apf` in *Faust*, con aggiunta dei coefficienti di bilanciamento $-g$ per il segnale diretto e $1 - g^2$ per il segnale ritardato.

viene biforcuto in corrispondenza del punto di prelievo. Di questi, uno viene scalato di un coefficiente $-g$ (negativo, per cui viene anche invertito di fase), l'altro transita all'interno di `df1c`, dove, dopo ciascun ciclo di retroazione, a periodo t e coefficiente di riscaldamento g , viene ulteriormente attenuato di $1 - g^2$. Questi due processi, simultanei, vengono poi sommati prima dell'uscita del filtro.

```
apf(t, g) = _ <: *(-g)+(df1c(t, g)*(1-(g*g)));
```

Il che significa che a posizione $n[1]$, in cui il campione $x[1]$ non ha ancora percorso il ciclo di retroazione, $y[1]$ equivale a $x[1]$ scalato di $-g$; al ciclo successivo in posizione $n[2]$, il campione $x[2]$ viene sommato al precedente $x[1]$, uscente dal ciclo di retroazione. La somma di ogni valore $x[n]$ con il precedente $x[n - 1]$ procede fino ad estinzione del segnale e questo determina la condizione per cui

[...] il riverberatore all-pass [...] condivide con gli altri riverberatori che usano delay e feedback la proprietà del decadimento esponenziale dell'energia sonora come avviene negli ambienti acusticamente ben progettati.

La presenza simultanea di segnale diretto e segnale ritardato ha come risultato il fatto che il filtro *all-pass* si comporti in maniera differente al variare dei valori assunti dal coefficiente di bilanciamento g : per $g = 1$ la componente ritardata `df1c` viene completamente annullata limitando l'intervento del filtro all'inversione di fase del segnale in ingresso; per $g = 0$ la componente diretta e il coefficiente di feedback di `df1c` vengono annullate, limitando l'intervento del filtro al ritardo del segnale in ingresso; per $g = 1/\sqrt{2}$ le componenti diretta e ritardata contribuiscono a quantità bilanciate.

2.3 Tempo di Riverberazione T60

Circa sessant'anni prima della pubblicazione di *Natural Sounding Artificial Reverberation*, Wallace Sabine, padre della fisica acustica, definì il tempo di riverbero o *T60* come la quantità di tempo in cui, in un determinato ambiente chiuso, la densità di energia sonora della riverberazione

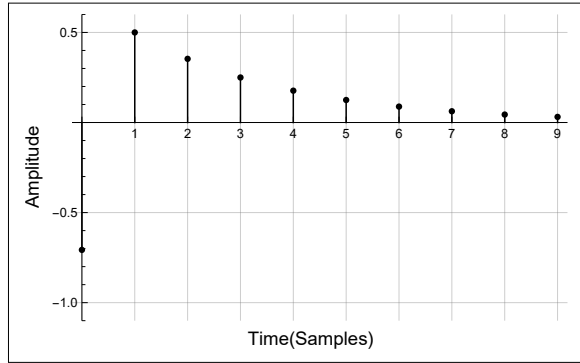


Figure 6. Risposta all'impulso della funzione `apf` in *Wolfram Mathematica*, caratterizzata dal riscaldamento e inversione di fase del primo campione in uscita e dal decadimento esponenziale dei successivi.

decade di $60dB$.

$$T = [60/(-20 \cdot \log|g|)] \cdot \tau = (3/\log|1/g|) \cdot \tau \quad (1)$$

Schroeder, di fatto, eredita la relazione in totale continuità con Sabine, introducendo, però, un discriminante fondamentale: quella che per Sabine era una formula di misurazione delle caratteristiche acustiche di determinato ambiente fisico, diventa per Schroeder un vero e proprio algoritmo di sintesi numerica nella riproduzione di quelle stesse caratteristiche acustiche, in un ambiente digitale.

Per facilitare l'implementazione dei componenti successivi, si è deciso di affiancare a questa relazione anche le relazioni inverse, per il calcolo di t e g .

$$\tau = [(-20 \cdot \log|g|) \cdot T]/60 \quad (2)$$

$$g = 10^{(60 \cdot t)/(-20 \cdot T)} \quad (3)$$

È possibile confrontare i risultati numerici delle tre relazioni facendo riferimento all'esempio presente nel repository.

```
process = sms.t60(11,1/sqrt(2)),
tau(219.247253417968,1/sqrt(2)),
gain(219.247253417968,11);
```

3. INCREMENTO DELLA DENSITÀ DI ECHI

Schroeder conosceva a pieno i requisiti di una buona riverberazione, ma era altrettanto consapevole che, nel raggiungimento di questi obiettivi, l'utilizzo di linee di ritardo in feedback, avrebbe generato diverse problematiche, di carattere computazionale (per restituire la densità di echi richiesta sarebbe servito un numero eccessivo di linee di ritardo *in parallelo*), ma anche di carattere qualitativo (connettere *in serie* anche solo due linee di ritardo in feedback avrebbe danneggiato il segnale in ingresso determinando una scarsa qualità del segnale riverberato).

Per quanto riguarda il problema della *bassa densità di echi*, abbiamo scoperto che servono approssimativamente 1000 echi al secondo per una riverberazione libera da fluttuazioni. [...] Sfortunatamente, le densità di echi

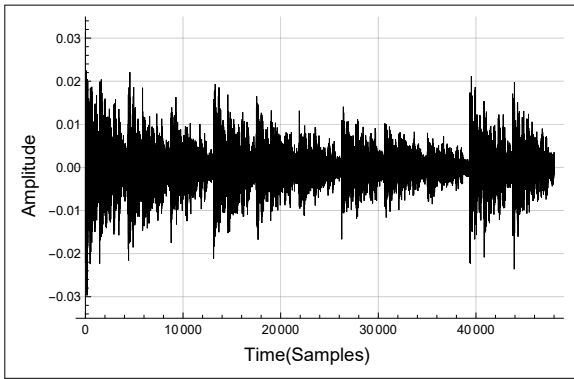


Figure 7. Risposta all'impulso della funzione `apfs` in *Wolfram Mathematica*.

di 1000 per secondo non sono facilmente raggiungibili da dispositivi di ritardo monodimensionali.

3.1 Serie di All-Pass

Il filtro *all-pass*, grazie alla sua risposta in frequenza piatta, permette di superare queste problematiche e costituisce di fatto una singola unità riverberante da poter collegare in serie per un numero variabile di volte.

[...] In questa maniera, ciascun'unità *moltiplicherà* effettivamente il numero di echi prodotti dalle unità precedenti. Supponendo che ogni impulso sia "propagato" in 3 di uguale ampiezza, il fattore di moltiplicazione per ogni unità addizionale è all'incirca 3. Ricominciando con un'unità che produce 25 echi al secondo, sono richieste solo all'incirca tra le 3 e 4 unità addizionali per raggiungere la desiderata densità di echi.

Sulla base di queste considerazioni teoriche, Schroeder costruisce così la serie: imposta il ritardo del primo *all-pass* $t_1 = 0.1$ sec e ognuno dei successivi a $1/3$ della sezione precedente; i coefficienti sono tutti uguali al valore di quadratura $g = 0.7$. Istanziandone 5 con queste impostazioni, ottiene 810 echi al secondo: il massimo raggiungibile con le tecnologie che ha a disposizione. Un risultato empirico, non identico a quello teorico desiderato, ma comunque sufficiente a soddisfare delle buone condizioni di riverberazione. Inoltre, per evitare cancellazioni e sovrapposizioni di echi, consiglia di scegliere i tempi di ritardo secondo rapporti incommensurabili, anziché secondo multipli interi.

La struttura sintattica di *Faust*, capace di mettere in relazione programmazione funzionale e algebra dei diagrammi a blocchi, facilita l'implementazione di questo algoritmo. Le iterazioni in *Faust*, funzionano in maniera analoga ai cicli `for(...)` negli altri linguaggi di programmazione. La funzione `seq`, itera espressioni connettendole in serie; accetta tre argomenti: il primo costituisce il nome della variabile contenente il numero dell'iterazione corrente (ana-

logamente alla variabile `i` nei cicli `for(...)`, indicizzata a partire da 0), il secondo costituisce il numero totale delle iterazioni espresso come una costante intera, il terzo costituisce l'espressione da iterare.

```
apfnp(md,t,m) = seq(i,8,sms.apfv(md,t*m^i : max(
  sff.np,1), 1/sqrt(2-(i/10))));
```

La funzione `apfnp` è appunto un'iterazione del filtro `apf`, mediante l'operatore `seq`. Questa, ha come variabili: N , l'ordine dell'iterazione (da istanziare a tempo di compilazione), md , la massima allocazione di memoria del blocco di ritardo, t_1 , il primo tempo di ritardo, m , la base della relazione esponenziale che regola il rapporto tra tempi di ritardo consecutivi della serie, $g = 1/\sqrt{2}$, il coefficiente di feedback degli *all-pass*, identico per ogni filtro della serie.

Per quanto riguarda la costruzione della serie si è deciso di assegnare il tempo di ritardo minimo t_1 al primo *all-pass* della catena e di calcolare i tempi crescenti successivi mediante la relazione impostata nell'iteratore `seq`. A partire da t_1 la relazione $t * m^i : \max(np, 1)$ restituisce numeri primi in proporzione, in maniera tale che, per ciascuna iterazione il tempo di ritardo sia un numero primo in rapporto di all'incirca 3 : 1 con il precedente.

```
np = ffunction(int next_pr(int), "../h/nextprime.h", "");
```

Dato un numero intero la funzione `np` restituisce il numero primo immediatamente successivo. È stata scritta utilizzando la primitiva `ffunction`, mediante la quale è possibile richiamare una funzione esterna scritta in linguaggio *C* indicandone il tipo di dato (`int`), il nome della funzione `nextpr_(int)` e il file include contenente il codice *C* sotto forma di stringa ("`nextprime.h`").

Le indicazioni fornite da Schroeder, insieme alle tecnologie di cui disponeva, sono strumenti su cui oggi poter ragionare creativamente. Per questo motivo, nell'implementazione *Faust* di questo algoritmo si è pensato di aumentare il numero di *all-pass* della serie superando quello da lui utilizzato. Il rapporto 3 : 1 però comporta una crescita esponenziale del tempo di ritardo che, per un valore di $t_1 = 2$ raggiunge il dato proposto da Schroeder (0.1 sec) già all'ottava unità. La Figura 7 indica la risposta all'impulso della serie.

3.2 Innesto di All-Pass

La serie di filtri *all-pass* illustrata nel paragrafo precedente costituisce già un algoritmo di riverbero autosufficiente. Nonostante questo, Schroeder propone diversi raffinamenti con lo scopo di conferire, al riverbero, un alto livello di realismo. Alcuni di questi sono:

- miscelazione del suono riverberato con suono diretto.
- introduzione di un intervallo temporale tra suono diretto e riverbero.

Questi raffinamenti, però, devono essere effettuati a condizione che nessuno dei due introduca alcun tipo di colorazione o fluttuazione del segnale riverberato, ovvero, che non violino il *principio all-pass*. Schroeder propone quindi

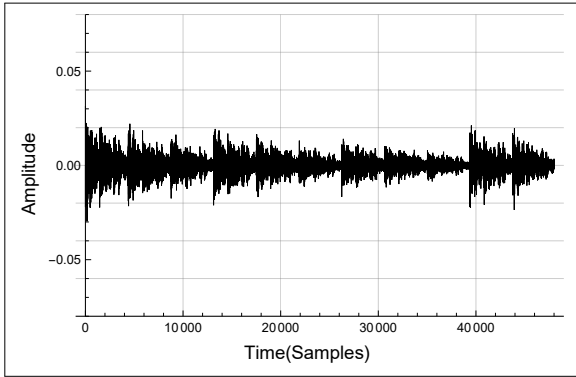


Figure 8. Risposta all’impulso della funzione `mdrned` in *Wolfram Mathematica*. L’incapsulamento della serie di *all-pass* all’interno di un *all-pass* determina il decadimento lineare anzichè esponenziale dell’impulso.

di incapsulare la serie di filtri *all-pass* in un’ulteriore filtro *all-pass*:

```
mdrned(t, g) = _ <: *(-g) + ((+ : de.delay(ma.SR
/2, int(t-1)) :
sms.apfnp(ma.SR, 2, 3)) ^*(g) : mem) * (1-(
g*g));
```

La miscelazione è ottenuta grazie ai coefficienti di bilanciamento ($-g$ per il segnale diretto, $1-g^2$ per il segnale ritardato) mentre, il ritardo tra suono diretto e suono riverberato (quello che nella letteratura successiva verrà definito *pre-delay*) è ottenuto mediante il blocco `delay`. Tra questo e il punto di prelievo del ciclo di retroazione viene inserita la funzione `apfnp`, responsabile del suono riverberato.

Un elemento che contraddistingue quest’architettura dalle precedenti (e dalle successive), evidente nella risposta all’impulso in Figura 8, è il decadimento lineare del segnale riverberato.

3.3 Rete di All-Pass e Comb

In ogni metodo scientifico, il feedback tra osservazione della realtà e produzione di modelli teorici è fondamentale. L’uno influenza l’altro in una relazione costante. Così Schroeder, sebbene abbia prodotto un modello teorico impeccabile, non perde di vista la constatazione empirica e pone un problema cruciale:

Ci si potrebbe chiedere: “Perchè insistere su risposte in frequenza perfettamente piatte se gli ambienti reali hanno risposte in frequenza altamente irregolari?”.

Dagli esperimenti psicoacustici svolti nei Bell Telephone Laboratories aveva appreso che, con la sufficiente densità di echi al secondo, le risposte frastagliate erano sostanzialmente indistinguibili dalle risposte piatte. Per un ambiente con un $T60 = 1$ sec ci sono 15 picchi in ogni intervallo di 100cps. La riflessione sull’esperienza, lo spinge a tornare a ragionare sul modello:

Quindi si potrebbe sperare che se un riverberatore artificiale avesse un numero simile di pic-

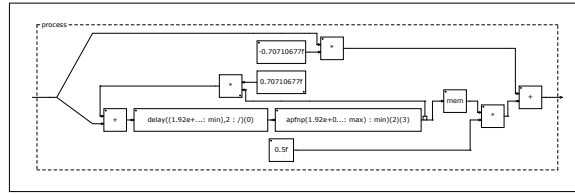


Figure 9. Diagramma a blocchi generato con la funzione `mdrned` in *Faust*, con incapsulamento della funzione `apfnp` all’interno di un filtro *all-pass*, nello specifico tra blocco di ritardo e punto di prelievo del ciclo di retroazione.

chi alla risposta, potrebbe suonare bene tanto quanto un ambiente reale.

Per ottenere un simile risultato, Schroeder propone di utilizzare diversi filtri *comb*. Poichè con un ritardo di 0.04 sec si generano 4 picchi ogni 100cps, per approssimare il numero di picchi alla risposta di una stanza con $T60 = 1$ sec richiede tra le 3 e 4 unità collegate in parallelo, ciascuna con tempi di ritardo incommensurabili e $g = 0.85$, per prevenire ogni tipo di fluttuazione. Ciascun filtro *comb* di questa configurazione produce 25 echi al secondo. Di conseguenza tutti 4 ne producono 100 (un decimo dei 1000 richiesti); poichè ogni unità riverberante *all-pass* moltiplica il numero di echi al secondo per 3, ne sono richieste due per arrivare quasi a quello richiesto in concordanza con le premesse teoriche dell’inizio.

```
dflapf = _ <: dflc(673, 0.89), dflc(1447, 0.78),
dflc(1811, 0.74), dflc(2111, 0.63) :>
apf(173, 1/sqrt(2)) : apf(229, 1/sqrt(2))
);
```

3.4 Tempo di Riverbero Dipendente dalla Frequenza

L’ultimo raffinamento che Schroeder propone per l’architettura illustrata nel paragrafo precedente è la dipendenza del tempo di riverbero dalla frequenza. Per realizzare questa condizione, ciascuno dei guadagni g dei filtri *comb* può essere reso dipendente dalla frequenza mediante l’aggiunta di un filtro *low-pass* ad un polo (in sostituzione del filtro RC da lui proposto) all’interno di ogni ciclo di feedback.

```
dflf(md, t, g, d) = (+ : de.delay(md, t-1)) ^sfi.
eavg2(d) * (g) : mem;
```

Qui, `eavg2` è controllato mediante coefficiente $0 < a < 1$, tuttavia la libreria dispone anche di un filtro `lp1p`, il cui controllo avviene mediante frequenza di taglio `fc`.

4. RIVERBERAZIONE AMBIOFONICA

Alla fine del racconto di Schroeder ci si trova a leggere e ragionare attorno ad un’ipotesi che oggi risulta piuttosto lontana dal senso comune di riverbero artificiale. Schroeder invoca un’ultima modifica dei suoi riverberatori affinché possano «produrre un realistico riverbero tridimensionale [...] Ovvero renderlo “ambiofonico”». Il termine ambiofonico viene riferito da Schroeder a Roelof Vermulen [5] ma il testo citato non riporta il termine. Risalendo alle fonti, appare l’introduzione del neologismo *ambiofonico* ad

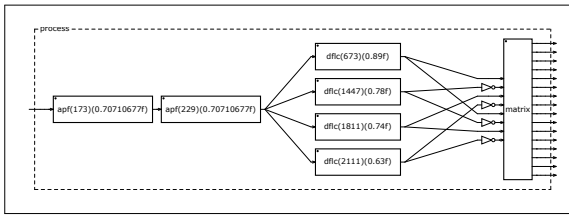


Figure 10. Diagramma a blocchi generato con la funzione ambiorev in Faust.

opera di Derk Kleis, ingegnere Philips di cui Vermeulen era Direttore della Ricerca Acustica dal 1947.

A new term is introduced to replace the formerly used “stereo-reverberation”. It appears that the latter is sometimes confused with stereophony - often abbreviated to stereo. The risk of confusion is all the greater since the advent of the stereophonic gramophone record which is making stereophony - or stereo - familiar to the general public. It is hoped the introduction of the term *ambiphony* will put an end to this confusion.

Il testo a cui si riferisce Schroeder [5] è una elaborazione di un precedente lavoro di Vermeulen [6] al quale Kleis fa riferimento introducendo il termine prima di descrivere la necessità di correggere l’acustica di ambienti il cui tempo di riverbero è troppo breve introducendo suono indiretto artificiale mediante, appunto, la tecnica ambiofonica.

Il dato importante della parte finale del testo di Schroeder è che chiarisce l’esigenza di introdurre passo passo i riverberatori digitali illustrati nelle sezioni precedenti al fine di poter entrare nel pieno della ricerca sulla riverberazione artificiale che infuocava le discussioni di quel periodo con una proposta completamente nuova: sostituire i riverberi elettromeccanici con i digitali per la conversione elettroacustica delle sale da concerto.

5. CONCLUSIONI

Il percorso di approfondimento critico nei confronti dei testi di Manfred Schroeder ha permesso una progressiva comprensione delle architetture dei processori di riverbero archetipici corroborando l’importanza della letteratura storica in quanto materiale vivo, attuale, accessibile.

La ri-costruzione di queste architetture mediante l’utilizzo di Faust e Wolfram Mathematica ha permesso la risoluzione delle problematiche relative a implementazioni in ambienti *samplerate* mediante una struttura di analisi solida corredata di diagramma a blocchi, supporto matematico e risposte all’impulso dettagliate.

La comprensione a basso livello dei principi di funzionamento ha stimolato implementazioni creative di questi stessi processori che sono state documentate in un ulteriore articolo dello stesso gruppo di ricerca. [4]

In ultimo, la redazione di una libreria di funzioni inserita nel progetto SEAM - Sustainable Electro-Acoustic Music fa sì che i risultati di questa ricerca contribuiscano alla

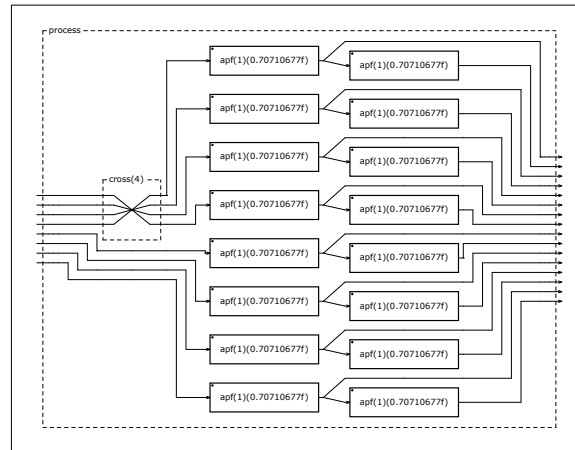


Figure 11. La matrice di resistenze indicata da Schroeder è stata sostituita con una rete di all-pass innestati che riproducono il movimento del fronte d’onda.

crescita di un ambiente condiviso, garante dell’accessibilità e della tutela di buona parte del patrimonio informatico elettroacustico ad oggi conosciuto.

6. REFERENCES

- [1] W. C. Sabine, *Collected papers on acoustics*. 1921.
- [2] M. R. Schroeder, “Natural sounding artificial reverberation,” *J. Audio Eng. Soc.*, vol. 10, no. 3, pp. 219–223, 1962.
- [3] M. R. Schroeder, “New method of measuring reverberation time,” *Acoustical Society of America*, 1964.
- [4] F. Vitucci, D. Annese, A. Di Furia, F. Scagliola, and G. Silvi, “Architettura aperta come strumento creativo: il riverberatore,” *Atti del XXIV Colloquio di Informatica Musicale*, 2024, under review.
- [5] R. Vermeulen, “Stereo-reverberation,” *J. Audio Eng. Soc.*, vol. 6, no. 2, pp. 124–130, 1958.
- [6] R. Vermeulen, “Stereo-reverberation,” *Philips Technical Review*, vol. 17, no. 9, pp. 258–268, 1956.
- [7] M. R. Schroeder, “An artificial stereophonic effect obtained from a single audio signal,” *J. Audio Eng. Soc.*, vol. 6, no. 2, pp. 74–79, 1958.
- [8] M. R. Schroeder and B. F. Logan, “‘colorless’ artificial reverberation,” *J. Audio Eng. Soc.*, vol. 9, no. 3, pp. 192–197, 1961.
- [9] M. R. Schroeder, “Listening with two ears,” *Music Perception: An Interdisciplinary Journal*, vol. 10, no. 3, pp. 255–280, 1993.
- [10] J. Dattorro, “Effect design, part 1: Reverberator and other filters,” *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 660–684, 1997.

A. LIBRERIA

```

1 declare name "SEAM Schroeder - Library";
2 declare version "0.2";
3 declare author "Daniele Annese";
4 declare author "Anthony Di Furia";
5 declare author "Giuseppe Silvi";
6 declare author "Francesco Vitucci";
7 declare license "CC3";
8
9 import("seam.lib");
10
11 //=====
12 //===== SCHROEDER - NATURAL SOUNDING ARTIFICIAL REVERBERATION - 1962 ===
13 //=====
14 sms = library("seam.schroeder.lib");
15 //===== DELAY FEEDBACK IN LOOP ---
16 // direct implementation, wrong impulse response
17 dfl(t,g) = (+ : @(t))~*(g);
18 //process = os.impulse : dfl(1,1/sqrt(2));
19
20 //----- DELAY FEEDBACK IN LOOP - FIXED ---
21 // delay compensation, right impulse response
22 dflc(t,g) = (+ : @(t-1))~*(g) : mem;
23 //process = os.impulse : dflc(1,1/sqrt(2));
24
25 //----- DELAY FEEDBACK IN LOOP - FIXED - VARIABLE ---
26 // runtime variable delay - max del required
27 dflcv(md,t,g) = (+ : de.delay(md, int(t-1)))~*(g) : mem;
28 //process = os.impulse : dflcv(ma.SR,1,1/sqrt(2));
29
30 //===== ALL-PASS FILTER ===
31 apf(t,g) = _ <: *(-g)+(dflc(t,g)*(1-(g*g)));
32 //process = os.impulse : apf(1,1/sqrt(2));
33
34 //----- ALL-PASS FILTER - VARIABLE ---
35 // runtime variable delay - max del required
36 apfv(md,t,g) = _ <: *(-g)+(dflcv(md,t,g)*(1-(g*g)));
37 //process = os.impulse : apfv(ma.SR,1,1/sqrt(2));
38
39 //===== T60 ===
40 t60(t,g) = (60/(-20*log10(g)))*t;
41 //process = msT60(0.1,1/sqrt(2)); // 2 seconds
42
43 //===== INCREASE OF ECHO DENSITY ===
44 // direct implementation of five all-pass in serie.
45 ied5 = apf(5507,1/sqrt(2)) : apf(1831,1/sqrt(2)) : apf(613,1/sqrt(2)) :
46       apf(199,1/sqrt(2)) : apf(67,1/sqrt(2));
47 //process = os.impulse : ied5;
48
49 //----- INCREASE OF ECHO DENSITY WITH VARIABLE TIMES ---
50 apfs(N,md,t,g) = seq(i,N,apfv(md,ba.take(i+1,t),g));
51 //process = os.impulse <: apfs(9,(spn.p3t),1/sqrt(2));
52
53 //----- INCREASE OF ECHO DENSITY WITH NEXT PRIME FUNCTION ---
54 // N = order
55 // t = smallest t
56 // m = base for m^i => step (3 for 1/3 ratio)
57 apfnp(md,t,m) = seq(i,8,sms.apfv(md,t*m^i : max(sff.np,1), 1/sqrt(2-(i/10))));
58 //process = os.impulse : apfnp(ma.SR,2,3);
59
60 //===== MIXING OF DIRECTSOUND AND REVERBERATION - NON-EXPONENTIAL DECAY ===
61 mdrned(t,g) = _ <: *(-g)+(((+ : de.delay(ma.SR/2, int(t-1)) :
62       sms.apfnp(ma.SR,2,3))~*(g) : mem)*(1-(g*g)));
63 //process = os.impulse : mdrned(1,1/sqrt(2));
64
65 //===== THE COMB FILTER APPROACH ===
66 dflapf = _ <: dflc(673,0.89), dflc(1447,0.78), dflc(1811,0.74), dflc(2111,0.63) :>
67       apf(173,1/sqrt(2)) : apf(229,1/sqrt(2));
68
69 apfdfl = apf(173,1/sqrt(2)) : apf(229,1/sqrt(2)) <:
70       dflc(673,0.89), dflc(1447,0.78), dflc(1811,0.74), dflc(2111,0.63);
71
72 //===== FREQUENCY-DEPENDENT REVERBERATION TIME ===
73 dflf(md,t,g,d) = (+ : de.delay(md,t-1))~sfi.eavg2(d)*(g) : mem;
74 //process = os.impulse : dflf(ma.SR,1,1/sqrt(2),0.9);
75
76 //===== AMBIOPHONIC-REVERBERATION ===
77 //----- NESTED ALL-PASS ---
78 apfn(0,t,g) = _;
79 apfn(1,t,g) = apf(t,g);
80 apfn(n,t,g) = apf(t,g) <: _, apfn(n-1,t,g);
81 //----- MATRIX ---
82 matrix = ro.cross(4), si.bus(4) : par(i,8, apfn(2,1,1/sqrt(2)));
83 //----- 16 CH AMBIOPHONIC REV EXAMPLE ---
84 ambiorev = apf(173,1/sqrt(2)) : apf(229,1/sqrt(2)) <:
85       dflc(673,0.89), dflc(1447,0.78), dflc(1811,0.74), dflc(2111,0.63) <:
86       par(i,4,_,0-_) : matrix;

```

GLI ‘SCIAMI DI GLISSANDO’ IN *DIAMORPHOSES*. RICOSTRUZIONE MEDIANTE *DIGITAL MORPHOPHONE ENVIRONMENT*

Daniel Scorrane

Dipartimento di Musica Elettronica
Conservatorio di L’Aquila
daniel.scorrane@gmail.com

Agostino Di Scipio

Dipartimento di Musica Elettronica
Conservatorio di L’Aquila
a.discipio@consaq.it

ABSTRACT

Il primo lavoro elettroacustico di Iannis Xenakis, *Diamorphoses* (per nastro magnetico, 1957) è stato commentato e analizzato da vari studiosi, di solito in approcci di analisi basati sull’ascolto. Le informazioni su processi e mezzi impiegati nella sua realizzazione sono poche e incerte. Talvolta si allude all’uso di *phonogène* e *morphophone*, sistemi di elaborazione magnetofonica sviluppati negli anni 1950 al GRM di Parigi, dove il brano fu realizzato. In base alle fonti storiche e agli indizi raccolti, qui consideriamo l’ipotesi che il *morphophone* sia stato usato per produrre sonorità assai peculiari, descritte talvolta come “sciami di glissando” [1] o in modi simili. Per verificarlo in modo concretamente operativo, ricorriamo al *Digital Morphophone Environment*, recente applicazione di elaborazione audio-numerica scritta in MAX. La ricostruzione del processo di accumulazione di brevi suoni glissati indica che l’ipotesi sia sostanzialmente corretta e fornisce suggerimenti su come Xenakis abbia provato a gestire la ‘densità’ delle sue fitte texture sonore.

1. INTRODUZIONE

1.1 *Diamorphoses*: elementi di contesto e ricezione

Iannis Xenakis (1922-2001) realizzò *Diamorphoses*, suo primo brano di musica elettroacustica, nel 1957, negli studi del *Groupe de Recherche en Musique Concrète* fondato da Pierre Schaeffer (nel 1958 ridenominato *Groupe de Recherches Musicales*, GRM). Il brano fu ascoltato la prima volta il 5 ottobre 1958 a Bruxelles (concerto GRM all’Expo 1958). Olivier Messiaen esprime subito la sua ammirazione per questo lavoro [2]. Più tardi, François Bernard Mâche lo avrebbe definito “un’opera di *musique concrète* che finalmente superava il livello sperimentale” [3]. *Diamorphoses* è considerato oggi uno dei titoli importanti del repertorio elettroacustico pionieristico, nonché un caso di musica concreta piuttosto singolare rispetto alle direttive seguite al GRM in quegli anni.

Al tempo, Xenakis era in realtà piuttosto interessato all’approccio ‘morfológico’ di Schaeffer, soprattutto in vi-

sta della possibilità di descrivere e realizzare oggetti sonori timbricamente dinamici. Il titolo stesso, *Diamorphoses* (dal greco antico *δια-μορφοσις*, cioè ‘formazione’), va appunto inteso come un formarsi e trasformarsi del suono secondo criteri contrastanti di continuità e discontinuità [4]. Avvicinandosi al GRM, il giovane compositore sperava anche di riprendere e ampliare i formalismi probabilistici che stava già sperimentando nella scrittura orchestrale (*Pithoprakta*, 1956; *Achorripsis*, 1957).

Nel corso degli anni, vari studiosi si sono occupati di *Diamorphoses*, prevalentemente con metodi d’analisi basati sull’ascolto [5, 6, 7, 8, 9, 10], di recente anche esaminando alcuni appunti manoscritti del compositore (negli Archives Xenakis alla Bibliothèque Nationale de France) [9, 11]. Tutti evidenziano una significativa varietà di materiali di origine genericamente concreta, magari non immediatamente riconoscibili, inclusi molti materiali sonori ‘ambientali’, presi probabilmente da archivi sonori accessibili nel contesto GRM (per esempio rumore di terremoti, tuoni e altro). Tutti gli studiosi tendono a segmentare l’arcata intera del brano (durata totale 6’55” ca.) in tre parti, pur con alcune differenze nella determinazione temporale della parte centrale. Segnaliamo però che Xenakis parlava di questo brano pensandolo strutturato in due parti [12].

1.2 Dispositivi magnetofonici: il ‘caso’ *morphophone*

Poche, sommarie e incerte sono le indicazioni esistenti su processi e mezzi di produzione. A parte le necessarie operazioni di editing e montaggio del nastro magnetico, alcuni autori [9, 11, 13] ritengono plausibile l’uso di *phonogène* e *morphophone*, dispositivi magnetofonici ideati da Schaeffer e Jacques Poullin nella prima metà degli anni 1950 (anche con l’apporto di Abraham Moles, per il *morphophone*). Oggi, quei sistemi risultano da tempo dispersi, se non perfino distrutti [14]. Lo studio e la ricostruzione delle pratiche produttive messe in atto mediante di essi, non può dunque fondarsi sulla possibilità di accedere a tali dispositivi.

Risulta plausibile – e viene anche dato per scontato [15] – l’uso del *phonogène à coulisse*, che permetteva di variare con continuità la velocità di rotazione del magnetofono, con conseguente trasposizione di frequenza e durata del segnale: un modo semplice di imporre al suono degli andamenti in *glissando*. Plausibile è forse anche l’uso del *phonogène a tastiera* (trasposizioni sui gradi discreti della scala cromatica, entro due ottave). È noto, in ogni caso,

Copyright: ©2024 Daniel Scorrane et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

che entrambi i *phonogène* furono utilizzati in modo normale e ricorrente nelle produzioni GRM almeno fino ai primi anni 1960.

Più problematico è il discorso sul *morphophone*, sistema magnetofonico sviluppato dal 1953 [16] e disponibile almeno dal 1954 [17]. Benché regolarmente elencato tra i più innovativi sistemi elettroacustici del suo tempo [18, 19, 20, 21], in retrospettiva non sembra esserci oggi alcuna certezza del suo impiego nell'insieme delle produzioni GRM.

Alcune testimonianze a nostra disposizione – tra cui scambi personali con Daniel Teruggi (direttore del GRM 1997-2017) – alludono al possibile uso del *morphophone* in brani realizzati a fine anni 1950 da Luc Ferrari e André Boucourechliev [22]. Altrove [1, 23] si assume che Schaeffer e Pierre Henry lo avessero usato già per la parte su nastro di *Orphée 1953* (per tre voci, clavicembalo, violino e nastro magnetico, 1953). Tali indicazioni però restano generiche e mai suffragate da precisi riscontri documentali. A volte si arriva a dubitare che il dispositivo sia mai stato impiegato musicalmente [16]. Xenakis non ne ha mai parlato esplicitamente in scritti o interviste in cui cita *Diamorphoses*. Ma ancor più significativo è che non ve ne sia menzione neanche nel *Traité des objets musicaux* o in altre delle principali pubblicazioni di Schaeffer [24, 25].

Sembra poi che il *morphophone* presentasse fragilità e malfunzionamenti tecnici tali da scoraggiarne l'uso [16]. Ciò si lega bene con la circostanza che Francis Coupigny (al tempo assistente di Poullin al GRM) nel 1960 ne progettò uno nuovo, che dobbiamo immaginare più affidabile e performante del precedente [26]¹. Resta il fatto che, secondo Teruggi [16], le forme di elaborazione del suono in teoria praticabili con tale dispositivo furono realmente adoperate al GRM solo più tardi, già in contesto tecnologico digitale, con risorse software quali i *GRM Tools*, a fine anni 1990 [27, 28].

1.3 Piccola anatomia funzionale del *morphophone*

In breve, si tratta di un anello di nastro magnetico (ca. 4 secondi di durata) montato su un meccanismo rotante, con testina di registrazione e testina di cancellazione in posizioni fisse e dieci testine di riproduzione riposizionabili lungo la circonferenza dell'anello (entro limiti). Il segnale di ogni testina di riproduzione arrivava a un filtro passabanda prima di essere amplificato e sommato agli altri, e poteva essere riportato all'ingresso del sistema (retroazione, *feedback*). La testina di cancellazione poteva essere rimossa, con conseguente sovra-registrazione e accumulazione di segnali (*overdub*). Il suono da elaborare era in genere registrato sull'anello di nastro, ma poteva anche provenire da altre sorgenti (microfoni o altro).

In retrospettiva, possiamo vederlo come una sorta di *multi-tap delay* con (o senza) *feedback*. A fine anni 1950 lo si poteva assimilare alla *delay wheel* dei laboratori di ricerca Philips di Eindhoven, sistema magnetofonico ideato appunto per ottenere ritardi multipli e riverberazione (ma

¹ In [13] si lascia intendere che il *morphophone* disegnato da Coupigny sia stato presumibilmente adoperato da Xenakis nel 1962 per *Bohor*.

in realtà anche e soprattutto per configurare sistemi di diffusione sonora multicanale [29]). Un accostamento tra i due sistemi fu fatto già nel 1958 [21].

Va detto tuttavia che nell'approccio del GRM il *morphophone* serviva non solo o non tanto per produrre effetti di eco e riverberazione, ma per “forgiare il profilo dinamico di un oggetto sonoro mediante ripetizione controllata di una cellula registrata” [19], per ottenere una “trasformazione della morfologia del suono mediante moltiplicazione” [30] (cfr. anche [31]). Così connotato, il dispositivo sicuramente può aver attratto l'attenzione di Xenakis.

1.4 Obiettivi

Ma con quali finalità e nella realizzazione di quali materiali sonori Xenakis vi avrebbe fatto ricorso, nel suo primo approccio ai mezzi elettroacustici? La strategia che proponiamo consiste nell'individuare - ad un ascolto attento e informato alle modalità tecniche di produzione - passaggi o sonorità di *Diamorphoses* di cui tentare una possibile ricostruzione, ripercorrendo con mezzi odierni alcune delle operazioni che il compositore può aver effettuato nelle fasi di lavorazione del brano. Ciò può fornire importanti indizi sulle pratiche produttive usate dal compositore nella realizzazione del brano, dunque sulla concreta musicalità di una prassi - allora del tutto nuova e non sempre ben documentata - il cui significato altrimenti sarebbe relegato al solo discorso teorico musicale ed estetologico.

Esporre ora criteri e scelte utili in tal senso, per poi passare a verifiche empiriche realizzate col *Digital Morphophone Environment*, applicazione software che replica il funzionamento del *morphophone* [32, 33] e ci consente di ricostruire con ragionevole approssimazione alcune delle sonorità caratteristiche di *Diamorphoses*.

2. CRITERI DI INDAGINE

2.1 *Sciami, masse e ragnatele di glissando*

Viste le funzioni attribuite al *morphophone* nel suo contesto storico tecnico-musicale, e considerato l'interesse di Xenakis per la creazione di trame sonore connotate da elevata 'densità' statistica (quantità media di eventi unitari in un dato tempo e un dato ambito di frequenza), si è deciso di portare la nostra attenzione su particolari complessi sonori presenti nella seconda e terza parte di *Diamorphoses*. In un caso – nel segmento da 2'22" a 4'05" – si tratta di sequenze ritmicamente irregolari (aperiodiche) di uno stesso suono, ripetuto e talvolta appena modulato in frequenza, simile al rintocco di una campanella o altro oggetto metallico piccolo e leggero, con un picco di energia intorno a 8.2 kHz. In Figura 1 vediamo il sonogramma complessivo di queste sequenze (isolate da suoni concomitanti), così come presenti nella registrazione disponibile².

² Ci serviamo della registrazione pubblicata su CD EMF-003 nel 1997. Il libretto del CD afferma che *Diamorphoses* è una composizione 'mono' anche se la traccia audio presenta leggere ma evidenti differenze tra i due canali. In generale, le fonti storiche non concordano sul numero di canali di questo lavoro: per motivi che qui non importa approfondire, esse variano nel definirlo 'mono', 'stereo' e perfino 'su quattro piste' [13, 34]. Qui ci atterremo all'odierno catalogo ufficiale delle opere di Xenakis [35], dove *Diamorphoses* è considerato 'stereo', come confermano recenti aggiornamenti [36].

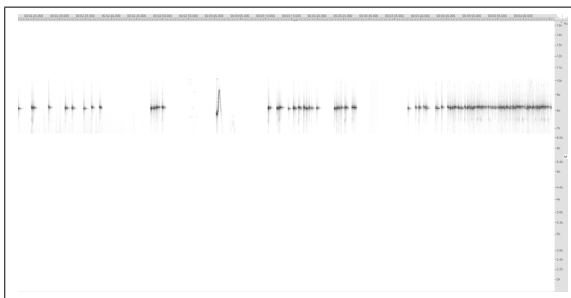


Figura 1. Ripetizioni e sovrapposizioni di “piccole campane”.

Nell’altro caso – nel segmento da 4’10” a 6’53” (parte finale del brano) – abbiamo materiali sonori altamente peculiari, texture in cui si ammassano innumerevoli brevi glissati in ambiti di frequenza elevati: (a) 7.5-10 kHz, (b) 6-12.5 kHz, (c) 1-7 kHz, (d) 1-12.5 kHz, (e) 4-12.5 kHz. La Figura 2 mostra il sonogramma di alcuni di tali materiali (isolati da suoni concomitanti presentati astraendo dall’effettiva posizione nella durata del brano). È assai probabile che Olivier Messiaen si riferisse a queste sonorità quando, nel 1959, parlò di “gigantesche e multicolori ragnatele sonore” [2]. Le quali sono state poi definite “sciame di glissando” (*Glissandoschwärmen* [1]), “masse di suoni glissati acutissimi” [37], “bande di glissando” [6] e “densi filamenti sonori” [38]. Insieme a pochi altri materiali, anche di morfologia diversa, questi “glissandi in grappoli” (*clusterartigen Glissandi* [1]) segnano tutta la terza e ultima parte del brano. Nella trascrizione grafica fatta da Makis Solomos nel 1993 [10], si tratta delle sonorità identificate dalle lettere g e f.

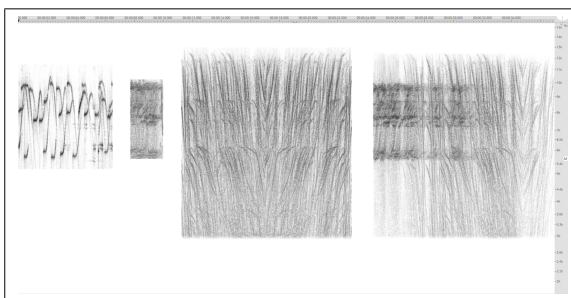


Figura 2. Alcuni degli “sciame di glissando”.

2.2 Costruzione del suono

Xenakis ha ricordato che tali passaggi di *Diamorphoses* furono “molto costruiti” e che l’elemento-base della costruzione fu il suono di una “piccola campana” [12]. Non ha però chiarito come ha davvero agito. È verosimile che abbia innanzitutto utilizzato il *phonogène à coulisse* per modulare il suono iniziale e imporgli una curva di glissando [15]. Ed è inoltre verosimile - suggeriamo noi - che abbia poi usato il *morphophone* per ‘moltiplicare’ (ripetere, accumulare) quel suono iniziale in grandi quanti-

tà, facendone texture prolungate e (quasi) continue. Tale suggerimento trova conforto in annotazioni di altri autori [1, 13].

2.3 Operatore di densità

Come si è detto, Xenakis era interessato a determinare la densità di queste texture sonore, cioè in pratica la loro “quantità media di suoni al secondo” [4]. Il compositore avrebbe approfondito questo aspetto nei due anni successivi, lavorando su *Analogique A et B* (per nove archi e nastro magnetico, 1958-59), nel quadro tecnico e teorico della ‘musica stocastica markoviana’ di cui scrisse più volte intorno al 1960 (questi scritti nel 1963 confluirono nel libro *Musiques Formelles* [39]). Specialmente *Analogique B* (la parte su nastro, inizialmente concepita come autonoma rispetto agli archi di *Analogique A*) richiese sforzi importanti al fine di forgiare suono e musica partendo da innumerevoli elementi minimi tempo-discreti (‘grani sonori’) e organizzandoli in agglomerati di densità variabile [40, 38, 4]. In quel contesto, la densità di eventi nell’unità di tempo fu (almeno in linea di principio) regolata su una scala logaritmica di base 3.

Ma già lavorando su *Diamorphoses*, Xenakis aveva intuito che la densità – o meglio la densità *percepita* – segue “una legge logaritmica con base tra 2 e 3”, come avrebbe poi scritto [39]. Vuol dire che essa “risulta aumentare di grado al raddoppiarsi o anche al triplicarsi dell’effettiva quantità di eventi sonori elementari” [38]. Come segnala Solomos, sulla scorta di documenti presenti negli Archives Xenakis [11], il compositore giunse a tale intuizione confrontandosi in particolare con Abraham Moles – uno degli artefici del *morphophone*.

Si può allora di nuovo ipotizzare l’impiego - molto empirico e forse anche difficoltoso, ma musicalmente efficace - del *morphophone*, non solo per trasformare un’unità sonora elementare in una trama articolata, ma anche per determinare la densità della trama stessa. Ciò in effetti era possibile, per esempio, mediante il feedback gestibile sul dispositivo stesso, oppure sottoponendo ad esso suoni precedentemente scaturiti dal suo utilizzo: la densità degli “sciame di glissando” risulterebbe allora da procedimenti di accumulazione, ripetuti in modi assimilabili alla tecnica di stratificazione di grani sonori su nastro magnetico escogitata da Xenakis due anni dopo per le ‘nuvole di grani sonori’ di *Analogique B* [38, 40, 41] (vi allude lo stesso compositore in un passaggio di [4]).

3. VERIFICHE COL DME

3.1 Breve introduzione al DME

Digital Morphophone Environment (DME) è una recente applicazione che ripropone i processi caratteristici del *morphophone* in ambito audio digitale [32, 33]. In base alle poche informazioni pertinenti disponibili, e studiando con molta attenzione schemi e immagini d’archivio, l’applicazione è stata progettata e scritta in MAX³ allo scopo di modellare le funzionalità del *morphophone* e le relative

³ versione 8.6.2 (<https://cycling74.com/products/max>).

modalità di elaborazione del suono. Il progetto tiene conto anche di aspetti caratteristici della tecnologia analogica, modellati in termini generici ma efficaci a emulare coerenti comportamenti analogici nel trattamento del segnale [42].

DME può operare in due modalità, ‘filologica’ ed ‘estesa’. La prima è aderente all’originale, la seconda sfrutta le risorse di calcolo per forme di elaborazione del segnale oggi comuni, ampliando quindi le possibilità di intervento. L’interfaccia grafica (in Figura 3) presenta il pannello principale coi moduli costitutivi del *morphophone* e fornisce rapido accesso ai parametri in gioco. Qui non possiamo approfondire scelte e criteri di implementazione (si veda [32, 33]), ma può essere utile riportare il diagramma a blocchi generale che schematizza gli stadi principali dei processi implementati (Figura 4).

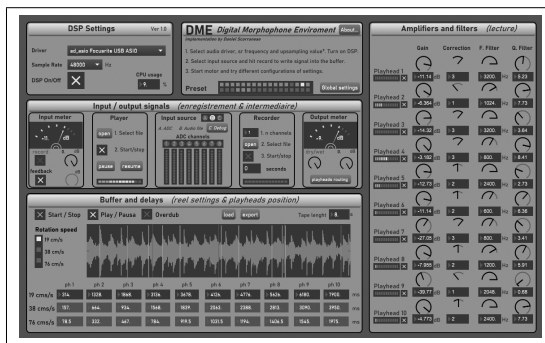


Figura 3. DME, pannello principale.

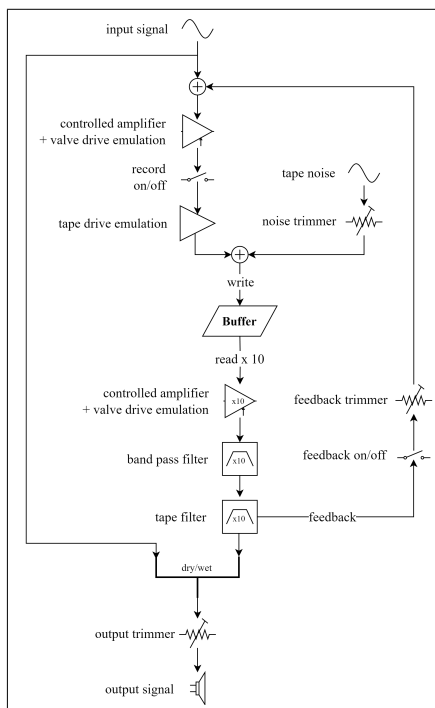


Figura 4. DME, diagramma a blocchi generale.

3.2 Modalità operative

Gli esempi illustrati di seguito e resi disponibili⁴ sono stati realizzati partendo da frammenti di *Diamorphoses* che possiamo considerare plausibili elementi-base di costruzioni sonore più estese e complesse. Sottoponendo tali elementi-base al DME, in modalità filologica, si è cercato di ricostruire il processo di ‘moltiplicazione’ e ‘densificazione’ da cui, secondo l’ipotesi da verificare, presero forma le sonorità texturali richiamate nella sezione 2.3 di questo paper. Gli esempi sono stati ottenuti in prima istanza senza ricorrere al filtraggio passa-banda e senza emulare il *flutter* del nastro e la preamplificazione analogica nei moduli del *morphophone*. Per ciascun esempio confronteremo il sonogramma della sonorità originale presa in esame (dalla registrazione segnalata in nota 2) col sonogramma della nostra ricostruzione. Data la natura statistica delle sonorità in gioco, gli esempi ricostruiti non mirano a riprodurre la loro esatta micro-organizzazione, ma ad avvicinarsi ad essa su base appunto statistica.

3.2.1 Primo esempio

Consideriamo il segmento tra 3’37” e 4’06” di *Diamorphoses*, in particolare la sonorità connotata dalla ripetizione di un picco di energia a 8.2 kHz (Figura 1). L’elemento base di tale segmento può essere costituito da un suono isolato presente a 3’10”, indicato appunto come quello di una piccola campana o altro piccolo oggetto metallico, presumibilmente già frutto di trasposizione in frequenza. Nel segmento considerato, l’elemento-base è ripetuto dapprima con ritmo discreto e irregolare, poi in modo più rapido con sovrapposizioni multiple tra istanze successive, fino a raggiungere un elevato grado di densità. Un test iniziale, molto semplice, consiste nell’approssimare la condizione ritmica che si ascolta tra 3’45” e 3’49” (dieci ripetizio-

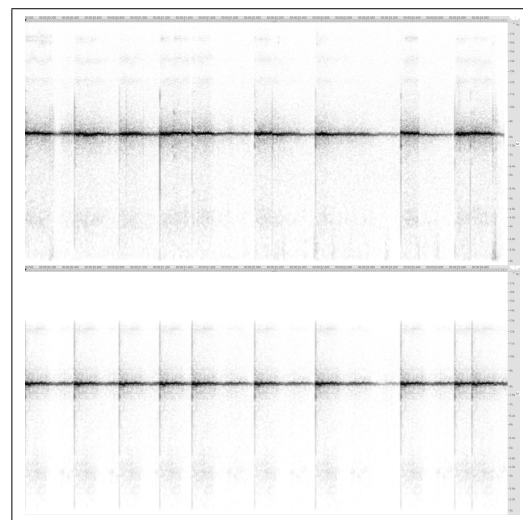


Figura 5. Segmento 3’45” - 3’49”. Originale (in alto) e ricostruzione (in basso).

⁴ https://github.com/danielscorranese/SCIAMIGLISSANDO_audioE_samples

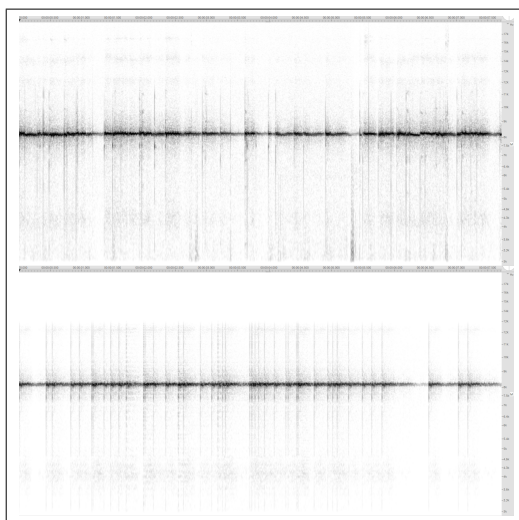


Figura 6. Segmento 3'49" - 4'06". Originale (in alto) e ricostruzione (in basso).

ni dell'elemento-base). Individuati gli intervalli temporali medi tra le ripetizioni, viene determinato il ritardo per ogni 'testina di lettura' del DME. I guadagni delle testine sono leggermente variati. Figura 5 mette a confronto il sonogramma dell'originale e quello del risultato da noi ottenuto.

Procediamo all'ulteriore densificazione presente da 3'49" a 4'06". Il suono sottoposto al DME sarà stavolta proprio la sequenza ottenuta col test iniziale ora visto. I valori di ritardo saranno leggermente sfasati rispetto alla configurazione precedente. Inoltre risulta opportuno ridurre il numero di testine di lettura coinvolte. Figura 6 mette a confronto il sonogramma di una porzione dell'originale e il risultato da noi ottenuto.

3.2.2 Secondo esempio

Consideriamo ora la texture avente energia concentrata fra 5.5 e 10 kHz, nel segmento da 4'26" a 4'49" (con interruzioni brusche) e poi nel segmento da 6'16" a 6'44" (con variazioni d'ampiezza gradualmente). Tale sonorità sembra effettivamente uno sviluppo del breve intreccio polifonico di suoni glissati presente fra 4'11" e 4'23", a sua volta realizzato presumibilmente con varie istanze della 'campanella' modulate col *phonogène à coulisse* (si veda la porzione di sonogramma più a sinistra in Figura 2). Oltre a condividere gli estremi di banda frequenziale, le due sonorità presentano una distribuzione di energia simile, con un avvallamento caratteristico intorno a 7.2 kHz. L'esempio da realizzare col DME prende dunque come elemento-base la piccola costruzione polifonica che inizia a 4'11".

Data l'elevata densità texturale del materiale in esame, è impossibile determinare i tempi medi di ripetizione: nel tentativo di ricostruzione, le linee di ritardo saranno perciò regolate in modo aperiodico e del tutto casuale (fino a un ritardo massimo di 4 secondi), mentre i livelli delle testine di lettura saranno posti tutti a -6 dB. L'esito così ottenuto però non sembra essere 'denso' abbastanza. Procediamo

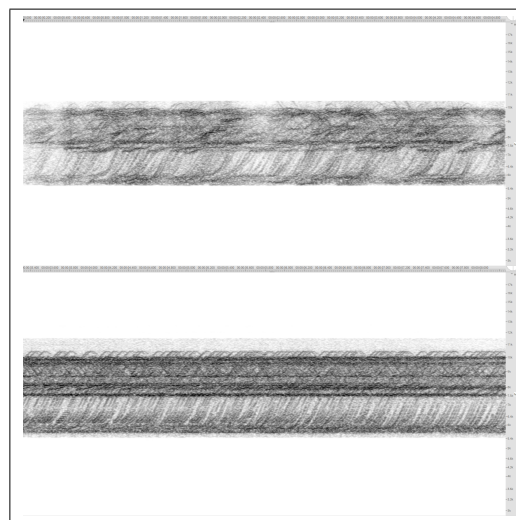


Figura 7. Porzione del segmento 6'16" - 6'44". Originale (in alto) e ricostruzione (in basso).

pertanto a sottoporre questo stesso risultato sonoro a una nuova elaborazione con DME, con valori di ritardo differenti ma secondo criteri simili. La densità raggiunta risulta ora più vicina all'originale. In Figura 7 si vede il sonogramma di un frammento dell'originale e della nostra ricostruzione (vengono provvisoriamente tralasciati i leggeri cali di ampiezza presenti nell'originale).

3.2.3 Terzo esempio

Consideriamo infine l'agglomerato di suoni glissati di diversa estensione e durata, ascendenti e discendenti nel registro di frequenza tra 1 e 7.3 kHz, nel segmento da 4'20" a 4'51". Concentriamoci in particolare sugli ampi glissandi ascendenti, che nel sonogramma (Figura 8, in alto) appaio-

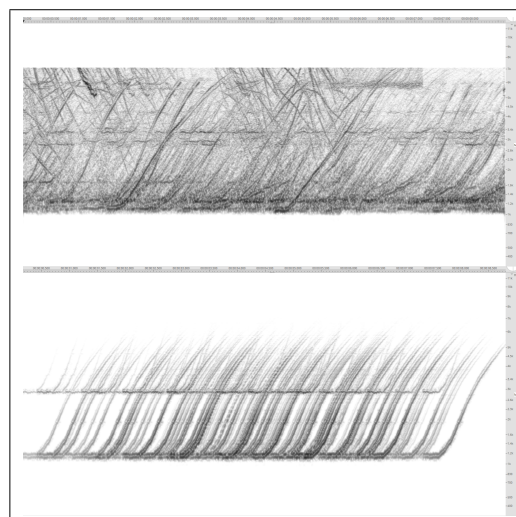


Figura 8. Porzione del segmento 4'20" - 4'51". Originale (in alto) e ricostruzione (in basso).

no intrecciati e sovrapposti ad altro materiale. L'elemento-base può essere individuato nella campanella con glissando ascendente presentata a 1'45", avente due componenti frequenziali in particolare rilievo (Figura 9). Come nel caso precedente, le linee di ritardo saranno regolate in modo aperiodico e casuale. Anche il livello delle testine di lettura stavolta assumerà valori diversi e casuali tra 0 e -30 dB. L'esito così ottenuto viene poi sottoposto di nuovo al DME, con valori di ritardo e livello diversi. Figura 8 mette a confronto il sonogramma di una porzione del segmento originale e della ricostruzione con DME. Naturalmente, nella ricostruzione non vi sono le molteplici striature spettrali da attribuire a suoni concomitanti (nell'originale) con quelli qui considerati.

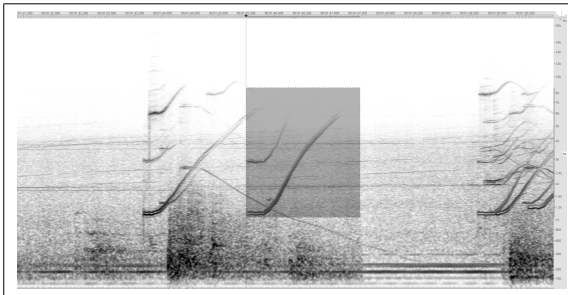


Figura 9. Il riquadro evidenzia il frammento a 1'45" (campanella modulata), usato come elemento-base per l'esempio in Figura 8.

3.3 Valutazioni

All'ascolto si rileva una forte congruenza qualitativa fra le sonorità di *Diamorphoses* prese in esame e la loro ricostruzione. Al netto di innumerevoli sfumature e di oggettive, inevitabili differenze nel dettaglio del segnale audio, gli esempi ricostruiti appaiono credibili istanze dei 'medesimi' processi da cui furono derivate le sonorità originali. Col DME in modalità filologica, l'eventuale simulazione di stadi di pre-amplificazione analogica può rendere in qualche modo più verosimile e 'antico' il suono, ma non sembra davvero necessaria per valutare il ruolo del *morphophone* nel produrre queste sonorità.

All'ascolto, inoltre, l'uso reiterato del *morphophone* risulta un plausibile espediente per graduare la densità della texture sonora. Ciò al momento non permette di fare luce circa le precise scelte fatte da Xenakis in tal senso, ma almeno consente di testarne empiricamente la concreta operatività.

4. CONCLUSIONI

L'indagine illustrata in questo paper rappresenta il tentativo di studiare aspetti rilevanti del repertorio elettroacustico storico con un approccio operativamente ricostruttivo di sonorità e strutture musicali, reso possibile da risorse software a loro volta concepite per modellare e in qualche modo 'ricostruire' le funzionalità di specifici mezzi tecnici (in questo caso, un particolare sistema di elaborazione

magnetofonica). Un approccio del genere richiama problematiche di ricerca e di prospettiva analitica più ampie. Qui però ci limiteremo a qualche conclusione specifica e circoscritta: gli esperimenti e le verifiche empiriche fatte col DME evidenziano che certi materiali sonori di *Diamorphoses* sono stati prodotti con forme di elaborazione del suono che nel 1957 al GRM di Parigi si potevano praticare solo col *morphophone*, o che non si sarebbero comunque praticate altrettanto efficacemente con altri mezzi e altre soluzioni. In breve, l'indagine offre evidenze indirette ma concrete e verificabili che spingono a ritenere accertato l'uso del *morphophone* nella lavorazione di *Diamorphoses*.

Abbiamo altresì osservato che il controllo sulla densità di trame sonore complesse fu affrontato da Xenakis assecondando la funzione 'moltiplicativa' o 'accumulativa' attribuita al *morphophone* da fonti vicine al GRM. In tal senso risulta plausibile che egli abbia usato il *morphophone* in feedback, oppure che lo abbia usato in modo iterativo, cioè sottoponendo ad esso suoni da esso stesso ottenuti in precedenza (oltre a suoni ottenuti dal *phonogène*). Questo in definitiva non permette un'effettiva misurazione della densità degli "sciame di glissando" e di altri materiali, però significa chiaramente che il ricorso al *morphophone* non fu estemporaneo o casuale: in quel particolare momento, Xenakis forse lo reputò un modo sistematico di graduare la densità di masse sonore nel contesto compositivo elettroacustico.

Infine, è musicalmente significativo che le trame prodotte col *morphophone* siano collocate, nel montaggio di *Diamorphoses*, sempre in un momento successivo (anche molto successivo) rispetto al momento in cui si presentano le unità sonore elementari da cui furono presumibilmente derivate. Questo suggerisce che, mettendo mano al *morphophone* (pur temporaneamente e sicuramente incontrando molteplici difficoltà tecniche), Xenakis colse il potenziale compositivo di una dinamica generativa e direzionata nel suono, interna ai processi di lavorazione. L'invenzione del suono (la 'micro-composizione', nei termini di Xenakis) si legava all'ideazione di schemi operativi e processi generativi. Quella prima esperienza va considerata dunque foriera di sviluppi importanti nel quadro della sua produzione (non solo) elettroacustica [13, 36, 38, 43].

5. RIFERIMENTI

- [1] M. Brech, "Formalisierte Kompositionweise in der frühen elektroakustischen Musik," in *Iannis Xenakis: Das elektroakustische Werk* (R. Paland and C. v. Blumröder, eds.), Wien: Der Apfel, 2009.
- [2] O. Messiaen, *Revue Musicale*. "Préface", vol. 244, 1959.
- [3] F.-B. Mâche, *Un demi-siècle de musique...et toujours contemporaine*, p. 153. Paris: L'Harmattan, 2000.
- [4] B. Varga, *Conversations with Iannis Xenakis*, pp. 94, 111. London: Faber and Faber, 1996.
- [5] S. De Stefano, "Spettromorfologie e articolazione strutturale in *Diamorphoses* (1957) di Iannis Xenakis,"

- in *Atti del Congresso di Didattica della musica elettronica* (M. De Amicis, ed.), L'Aquila: Istituto Gramma, 1998.
- [6] T. DeLio, "Diamorphoses by Iannis Xenakis," in *Electroacoustic music: Analytical perspectives* (T. Licata, ed.), Westport: Greenwood, 2002.
- [7] R. Frisius, "Probleme der Analyse elektroakustischer Musik von Iannis Xenakis, dargestellt am Versuch einer Höranalyse von Diamorphoses," in *Iannis Xenakis: Das elektroakustische Werk* (R. Paland and C. v. Blumröder, eds.), Wien: Der Apfel, 2009.
- [8] M. Malt, "Une proposition pour l'analyse des musiques électroacoustiques de Xenakis à partir de l'utilisation de descripteurs audio," in *Iannis Xenakis. La musique électroacoustique* (M. Solomos, ed.), Paris: L'Harmattan, 2015.
- [9] M. Solomos and B. Gibson, "Research on the First Musique Concrète: The Case of Xenakis's First Electroacoustic Pieces," International Electroacoustic Music Studies Network. Lisbon, June 2013.
- [10] M. Solomos, "À propos des premières œuvres (1953-69) de I. Xenakis: pour une approche historique de l'émergence du phénomène sonore," p. 269. PhD thesis, Université La Sorbonne, Paris, 1993.
- [11] M. Solomos, "Xenakis first composition in musique concrète: Diamorphoses," in *Proc. of the Xenakis International Symposium* (D. Exarchos, ed.), Goldsmiths: Univ of London, 2011. (www.gold.ac.uk/ccmc/xenakis-international-symposium/programme).
- [12] F. Delalande, *'Il faut être constamment un immigré': Entretiens avec Xenakis*, p. 39. Paris: Buchet/Chastel, 1997.
- [13] R. Friedl, *Towards a Philology of Electroacoustic Music - Xenakis's Tape Music as Paradigm*. PhD thesis, Goldsmiths, University of London, 2019.
- [14] D. Teruggi, "De phonogène en phonogène: cinquante années d'outils concrets," in *Du sonore au musical: Cinquante années de recherches concrètes (1948-1998)* (S. Dallet and A. Veitl, eds.), Paris: L'Harmattan, 2001.
- [15] M. Battier, "De la machine à l'oreille: Le paradoxe de la musique concrète," in *Du sonore au musical: Cinquante années de recherches concrètes (1948-1998)* (S. Dallet and A. Veitl, eds.), Paris: L'Harmattan, 2001.
- [16] D. Teruggi, "Technology and musique concrète: the technical developments of the Groupe de Recherches Musicales and their implication in musical composition," *Organised Sound*, vol. 12, no. 3, 2007.
- [17] E. Gayou, "The GRM: Landmarks on a historic route," *Organised sound*, vol. 12, no. 3, 2007.
- [18] L. Cross, "Electronic music, 1948-1953," *Perspectives of new music*, vol. 7, 1968.
- [19] Unknown, "Petit Lexique De Musique Experimentale Et Electronique," *Esprit*, no. 280, p. 122, 1960.
- [20] J. Poullin, "L'apport des techniques d'enregistrement dans la fabrication de matières et formes musicales nouvelles. Applications à la musique concrète," *L'onde Électrique*, no. 324, 1954. Reissued in *Ars sonora*, 9, 1999 (<http://www.ars-sonora.org/html/numeros/numero09/>).
- [21] V. A. Ussachevsky, "The Processes of Experimental Music," *Journal of the Audio Engineering Society*, vol. 6, 1958.
- [22] D. Teruggi. personal email, may 7th 2021.
- [23] B. Fox, "Schaeffer stands his ground: *Orfée 53* and evocative sound," in *Sound in multimedia contexts*, Montréal: Electroacoustic Music Studies Network (EMS), 2005. <http://www.ems-network.org/spip.php?article223>.
- [24] P. Schaeffer, *Traité des objets musicaux*. Paris: Seuil, 1966.
- [25] P. Schaeffer, *La musique concrète*. Paris: Presse Universitaires de France, 1967. (Reprinted 1973 and 2020).
- [26] F. Coupigny, "Sous la direction de Pierre Schaeffer: repères pour un parcours inventif," in *Du sonore au musical: Cinquante années de recherches concrètes (1948-1998)* (S. Dallet and A. Veitl, eds.), Paris: L'Harmattan, 2001.
- [27] E. Favreau, "Les outils de traitement GRM Tools," in *Actes - Journées d'Informatique Musicale*, Marseille: CNRS-LMA, 1998.
- [28] D. Teruggi and Y. Geslin, "Studio Report: The GRM—Groupe de Recherches Musicales, Institut National de l'Audiovisuel," in *Proceedings of the International Computer Music Association*, University of Miami, 2004.
- [29] R. Vermeulen, "Stereo Reverberation," *Philips Technical Review*, 1956.
- [30] INA-GRM, *Répertoire acousmatique 1948-1980*, p. 275. Paris: Institut National de l'Audiovisuel, 1980.
- [31] M. Chion and G. Reibel, *Les musiques électroacoustiques*, p. 225. Paris: INA-GRM, 1976.
- [32] D. Scorrane, "Morphophone. Anelli di suono tra analogico e digitale." BA thesis. Conservatory of L'Aquila. 2023.
- [33] D. Scorrane, "Riproposizione digitale di uno strumento pionieristico. Digital Morphophone Environment," paper presentation, XXIV Colloquium of Musical Informatics, Turin, 2024.

- [34] N. Matossian, *Iannis Xenakis*. Paris: Fayard, 1981.
- [35] *Xenakis Catalogue*. Paris: Salabert, 2001.
- [36] A. Di Scipio and M. Solomos, *La musique électroacoustique de Xenakis*. Paris: Salabert, in print.
- [37] M. Solomos, “Analysing the First Electroacoustic Music of Iannis Xenakis,” paper presentation, 5th European Music Analysis Conference, Bristol, 2002.
- [38] A. Di Scipio, “Stochastics and Granular Sound in Xenakis’ Electroacoustic Music,” in *Iannis Xenakis. La musique électroacoustique* (M. Solomos, ed.), Paris: L’Harmattan, 2015.
- [39] I. Xenakis, *Musiques Formelles*. (Revue Musicale, 253-254), 1963, p. 69. Iannis Xenakis, *Formalized Music*, Pendragon Press, Stuyvesant, 1992, p. 373.
- [40] A. Di Scipio, “Le nubi di suono ed i loro meccanismi. Uno studio di Analogique A et B,” in *Xenakis: Musicista scienziato architetto* (A. Melchiorre, ed.), Milano: Edizioni Scuole Civiche di Milano, 2006.
- [41] A. Di Scipio, “Nuvole sonore, composizione automatica, robot danzanti. La ricerca sonologica e musicale di Iannis Xenakis,” in *Xenakis rimane* (M. Giommoni and R. Cresti, eds.), Treviso: Diastema, 2023.
- [42] M. Camras, *Magnetic Recording Handbook*. Dordrecht: Springer Netherlands, 1998.
- [43] A. Di Scipio, “Compositional models in Xenakis’s electroacoustic and computer music,” *Perspectives of New Music*, vol. 32, 1998.

Fully generalized Fibonacci series modulo n as music sequence generators

Cristiano Bocci
Università di Siena
cristiano.bocci@unisi.it

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

ABSTRACT

In this paper we introduce fully generalized Fibonacci sequences as useful tools for the generation of integers that can be interpreted in the musical context for algorithmic composition. In particular, we take into account the modulo operation on Fibonacci sequences resulting in various periodic behaviors that we interpret in the pitch class domain. First we introduce Fibonacci sequences and generalized Fibonacci sequences, then we discuss the modulo operator applied to sequences. We propose various interpretations of the resulting sequences in terms of pitches and pitch classes, and describe some possible operations. Finally, we introduce fully generalized Fibonacci sequences and describe a possible implementation in an algorithmic composition environment.

1. INTRODUCTION

Use of the Fibonacci sequence has a long tradition in 20th Century Music. Kramer [1] provides an early account of references including Bartók, Stockhausen and Nono (on the latter see also [2], on Ferneyhough see [3]). In these cases, Fibonacci sequences are used to define time proportion, like in [4]. Mathematical and musicological discussions mostly deal with the relation between Fibonacci sequences and golden ratio (see e.g. [5, 6]). Mongoven [6] also reports contemporary uses while proposing an application to tuning systems.

In this paper we are mostly interested in pitch and pitch class interpretation of Fibonacci sequences, e.g. when numbers are mapped onto pitches and pitch classes. First, we discuss some general features of Fibonacci numbers, then we consider some musical applications to pitch domain, finally we generalize Fibonacci sequences and describe possible music developments.

2. MATHEMATICAL ASPECTS OF FIBONACCI SEQUENCES

The Fibonacci numbers are the integer numbers in the sequence (the Fibonacci sequence) defined by the recurrence

Copyright: ©2024 Cristiano Bocci et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

relation

$$F : \begin{cases} F_0 = 0 \\ F_1 = 1 \\ F_i = F_{i-1} + F_{i-2}. \end{cases} \quad (1)$$

The Fibonacci numbers have been widely studied and satisfy many and varied identities. Many of the identities involve both addition and multiplication and so the full ring structure of the integers is required to prove them.

In matrix representation, given

$$U = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$$

then

$$U^i = \begin{pmatrix} F_{i-1} & F_i \\ F_i & F_{i+1} \end{pmatrix}.$$

It is worth to mention two other well-known sequences.

- The Pell sequence

$$P : \begin{cases} P_0 = 0 \\ P_1 = 1 \\ P_i = 2P_{i-1} + P_{i-2}. \end{cases} \quad (2)$$

- The Lucas sequence

$$L : \begin{cases} L_0 = 2 \\ L_1 = 1 \\ L_i = L_{i-1} + L_{i-2}. \end{cases} \quad (3)$$

Hence the first 9 steps of the three sequences are

| i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ... |
|-------|---|---|---|---|----|----|----|-----|-----|-----|-----|
| F_i | 0 | 1 | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | ... |
| P_i | 0 | 1 | 2 | 5 | 12 | 29 | 70 | 169 | 408 | 985 | ... |
| L_i | 2 | 1 | 3 | 4 | 7 | 11 | 18 | 29 | 47 | 76 | ... |

(4)

Fibonacci, Pell and Lucas sequences are respectively sequences A000045, A000129 and A000032 in the On-Line Encyclopedia of Integer Sequences (OEIS)¹.

3. PISANO PERIODS

In the following we consider the sequence $F(\text{mod } n)$ obtained by taking the remainders of the Fibonacci sequence modulo n , where n is an integer. For example, if $n = 3$ one has

| i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | ... |
|--------------------|---|---|---|---|---|---|---|----|----|----|-----|
| F_i | 0 | 1 | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | ... |
| $F(\text{mod } 3)$ | 0 | 1 | 1 | 2 | 0 | 2 | 2 | 1 | 0 | 1 | ... |

(5)

¹ <https://oeis.org/>

| $\pi(n)$ | 0+ | 1+ | 2+ | 3+ | 4+ | 5+ |
|----------|----|----|-----|----|-----|-----|
| +1 | 1 | 10 | 16 | 30 | 40 | 72 |
| +2 | 3 | 24 | 30 | 48 | 48 | 84 |
| +3 | 8 | 28 | 48 | 40 | 88 | 108 |
| +4 | 6 | 48 | 24 | 36 | 30 | 72 |
| +5 | 20 | 40 | 100 | 80 | 120 | 20 |
| +6 | 24 | 24 | 84 | 24 | 48 | 40 |
| +7 | 16 | 36 | 72 | 76 | 32 | 72 |
| +8 | 12 | 24 | 48 | 18 | 24 | 42 |
| +9 | 24 | 18 | 14 | 56 | 112 | 58 |
| +10 | 60 | 60 | 120 | 60 | 300 | 120 |

Table 1. Pisano periods

$F(\bmod n)$ is a bi-infinite sequence, that is, given any two consecutive terms, we can find the terms preceding and following those terms. In other terms any pair of consecutive terms of $F(\bmod n)$ determines the entire sequence both forward and backward.

Note that there are only n^2 pairs of possible terms hence the serie repeats, and the recurrence of a pair results in recurrence of all following terms. By the recursive term in (1) and the definition of $F(\bmod n)$ we have that if $F_{t+1} \equiv F_{s+1} \bmod n$ and $F_t \equiv F_s \bmod n$ then

$$F_{t+1} \equiv F_{s+1}, \dots, F_{t-s+1} \equiv F_1 \bmod n$$

and $F_{t-s} \equiv F_0 \bmod n$. Hence we have the following

Theorem 3.1 ([7]). $F(\bmod n)$ forms a simply periodic series. That is, the series is periodic and repeats by returning to its starting values.

Denote by $\pi(n)$ the least positive integer k such that $F_k \equiv 0 \bmod n$ and $F_{k+1} \equiv 1 \bmod n$. Thus, $\pi(n)$ denotes the period of $F(\bmod n)$. For example, according to (5), $F(\bmod 3)$ has period 8, so $\pi(3) = 8$.

Definition 3.1.1. $\pi(n)$ is called Pisano period of the sequence $F(\bmod n)$.

The existence of periodic functions in Fibonacci numbers was noted by Joseph Louis Lagrange in 1774 [8]. In Table 1, the first 60 Pisano periods are shown. Here the rows represent units while columns represent decimals. Hence, to check the Pisano period of, for example, 43, it is enough to control the intersection between the column 4+ and the row +3. In Figure 1 there is the plot of the first 10,000 Pisano periods².

We can notice that $U^k \equiv I \bmod n$ precisely when $\pi(n)|k$, where I is the identity matrix³.

We now collect some results about Pisano periods.

Theorem 3.2 ([7]). For $n > 2$, $\pi(n)$ is even.

Proposition 3.2.1. $\pi(\text{lcm}(n_1, n_2)) = \text{lcm}(\pi(n_1), \pi(n_2))$, where lcm is the lowest common multipole.

² A clear visualization of Pisano periods in relation to periodicity has been proposed by YouTube user Jacob Yatsko here: <https://www.youtube.com/watch?v=01eLKODSCqW>

³ We remind that $|$ stands for “divides”.

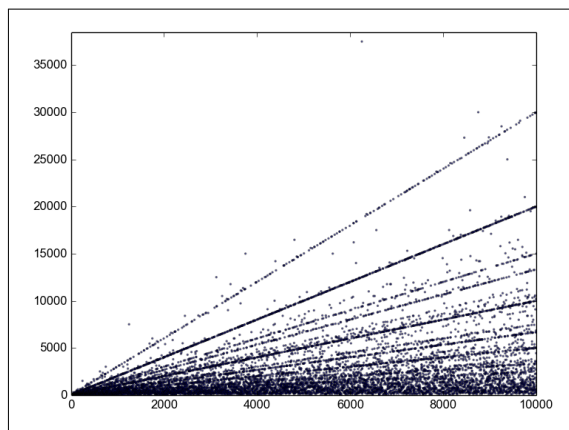


Figure 1. Plot of the first 10,000 Pisano periods;

Source: Robodile, CC BY-SA 4.0

<https://creativecommons.org/licenses/by-sa/4.0>, via Wikimedia Commons.

As a corollaries we get the following ones.

Corollary 3.2.1. if $k|n$ then $\pi(k)|\pi(n)$.

For example, $17|51$ and $\pi(17) = 36$ divides $\pi(51) = 72$. Observe, however that $51/17 = 3$ but $\pi(51)/\pi(17) = 2$. Similarly, $11|22$ and $\pi(11) = 10$ divides $\pi(22) = 30$, but this time, one has $22/11 = 2$ but $\pi(22)/\pi(11) = 3$.

Corollary 3.2.2. if n has prime factorization

$$n = p_1^{e_1} p_2^{e_2} \dots p_n^{e_n},$$

then

$$\pi(n) = \text{lcm}(\pi(p_1^{e_1}), \pi(p_2^{e_2}), \dots, \pi(p_n^{e_n})).$$

The following propositions give bounds for $\pi(n)$.

Proposition 3.2.2. $\pi(n) \leq 6n$, with equality holds if and only if $n = 2 \cdot 5^n$.

In Figure 1 we can notice that $\pi(250) = 1500$ and $\pi(1250) = 7500$ stand out.

Proposition 3.2.3 ([9]). Given n , let $t > 0$ such that $L_t \leq n$ where L_t is the t -th Lucas number. Then $\pi(n) \geq 2t$.

Since 0 represents the initial step, it is important to make some considerations about the appearance of 0.

Note that 0 is in the sequence (as $F_0 = 0$) and since $F(\bmod n)$ is periodic, we get that, for any integer n , infinitely many Fibonacci numbers are divisible by n .

Proposition 3.2.4. The zeros of the sequence $F(\bmod n)$ are evenly spaced, that is if F_s and F_t are congruent to 0 modulo n , then F_{s+t} and F_{s-t} are congruent to 0 modulo n .

This follows from the well-know equalities

$$F_{s+t} = F_{s-1}F_t + F_sF_{t+1} \quad (6)$$

and

$$F_{s-t} = (-1)^t (F_sF_{t+1} - F_{s+1}F_t). \quad (7)$$

3.1 Generalized Fibonacci sequences

A first kind of generalization of the Fibonacci sequence consists in considering different initial steps. To this aim we define the sequences $G_{[a,b]}$ of the form

$$G_{[a,b]} : \begin{cases} G_0 = a \\ G_1 = b \\ G_i = G_{i-1} + G_{i-2}. \end{cases} \quad (8)$$

for integers a, b . Clearly $G_{[0,1]}$ is the usual Fibonacci sequence and $G_{[2,1]}$ is the Lucas sequence.

Again, we can consider the sequence $G_{[a,b]}(\text{mod } n)$. In the following table we have the first steps of $G_{[a,b]}$ and $G_{[a,b]}(\text{mod } n)$ for $a = 3, b = 2$ and $n = 3$:

| i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | ... |
|----------------------------|---|---|---|---|----|----|----|----|----|-----|
| G_i | 3 | 2 | 5 | 7 | 12 | 19 | 31 | 50 | 81 | ... |
| $G_{[3,2]}(\text{mod } 3)$ | 0 | 2 | 2 | 1 | 0 | 1 | 1 | 2 | 2 | ... |

(9)

Proposition 3.2.5. *Let a, b, n be integers with $m \geq 2$.*

- (a) $G_{[a,b]}$ satisfies $G_i = F_{i-1}a + F_i b$.
- (b) $G_{[a,b]}(\text{mod } n)$ is periodic.

Part (a) of the previous proposition shows an important link between $G_{[a,b]}$ sequences and Fibonacci sequence F and part (b) tells us that sequences $G_{[a,b]}$ have Pisano period.

We denote by $\pi(a, b, n)$ the Pisano periods of the sequences $G_{[a,b]}(\text{mod } n)$. We have the following important result.

Proposition 3.2.6. *The Pisano period of $G_{[a,b]}(\text{mod } n)$ divides the Pisano period of $F(\text{mod } n)$, that is*

$$\pi(a, b, n) | \pi(n).$$

Also for $\pi(a, b, n)$ we have the same results of Proposition 3.2.1 and Corollaries 3.2.1 and 3.2.2 (see [7]).

Let $D = b^2 - ab - a^2$. The value D plays an important role to determine $\pi(a, b, n)$.

Theorem 3.3. *If $LCD(D, n) = 1$ then $\pi(a, b, n) = \pi(n)$.*

For example, the Lucas sequence has $D = -5$. Thus, for any n that is not a multiple of 5, the Pisano period of the Lucas sequence mod n is the same of $F(\text{mod } n)$.

Many papers on sequences $G_{[a,b]}$ mainly focus on the case $LCD(a, b) = 1$, that is a and b are coprime. However, for the topics of the following sections, also the case in which b divides a is of particular interest. In this context, we establish the following result.

Proposition 3.3.1. *If $a = kb$ and b divides n then all sequences $G_{[kb,b]}(\text{mod } n)$ give the same set of numbers, for all non-negative integers k .*

Proof. Write $n = t \cdot b$. Given any k , we have $k = qn + r$ for suitable integers q and r . Thus the sequence $G_{[kb,b]}$ can be written as

$$(qn+r)b, b, (qn+r+1)b, (qn+r+2)b, (qn+r+3)b, \dots$$

Thus, when we consider this sequence mod n we get

$$rb \text{ mod } n, b \text{ mod } n, (r+1)b \text{ mod } n, \dots$$

This shows that all elements of the sequence are all the multiple of b modulo n , independently of the value of r or, equivalently, of the value of k . \square

4. FIBONACCI SEQUENCES AND PITCH CLASSES

Pitch classes are defined in relation to Forte’s set theory for atonal music, where each pitch class (e.g. C representing the class of all possible Cs, regardless of octave) is indexed by an integer number (e.g. 0 for C) (see [10, 11, 12], in other contexts defined as chroma values). In this context, we assume: equal temperament; octave equivalence; and enharmonic equivalence. The first assumption says that the octave range ($[f, 2f]$, where f is frequency) can be divided into 12 equal steps where $f_{i+1} = \sqrt[12]{2}f_i$ (12 equal temperament, 12-ET), the second that each corresponding step in different octaves shares a perceptual "sameness" feature, the third that, in relation to Western harmony, alterations are not relevant, so there is properly no difference among e.g. $B\sharp, C, D\flat$, as they are all represented by pitch class index 0. These three constraints allow for integer notation, that is, each step in an equally tempered octave can be indexed by an integer in the range 0-11. As the pattern repeats itself, there are only 12 pitch classes [12].

The first application of Fibonacci numbers to the pitch domain can be traced back to Joseph Schillinger [13]. In the chapter devoted to *Theory of Melody* of his *System* he suggests to use Fibonacci sequences as a way to generate pitches, like in Figure 2 (omitting 0, 1). Here Schillinger is also proposing a specific interpretation of Fibonacci sequences as incremental positive intervals rather than pitch classes. This means that $[0, 1, 1, 2, 3]$ can be interpreted not as C, D \flat , D \flat , D, E \flat but rather as C, D \flat , D, E, G. It can be easily seen that, due to the accumulation process, numbers mapped onto pitches very soon exceed the available piano range. As a solution to this problem, Schillinger suggests to transpose each pitch to a viable octave, by means of "readjustment of the range" [13, p. 334] (Figure 3). This operation is indeed equivalent to apply mod 12 to the number. This modulo operation has relevant implications. Much more recently, Haek [14] has proposed an application of Fibonacci mod n to serial composition. In a 2023 YouTube video, Evanstein presents the application *Fibonacci Music Box* that deals with modulo application to Fibonacci series so to obtain pitch sequences⁴. In the following we discuss some more general features with number sequences interpreted as pitch class indices.

5. FIBONACCI MOD N AND PITCH CLASSES

In this section we propose an application to the pitch domain of Fibonacci sequences. Generalized Fibonacci sequences provide a very simple yet powerful formalism to

⁴ https://www.youtube.com/watch?v=_aIf4WUCNZU

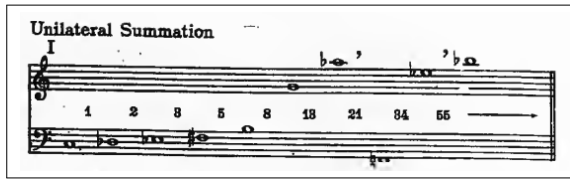


Figure 2. “Unilateral symmetry of Fibonacci series”. From [13, p. 334]

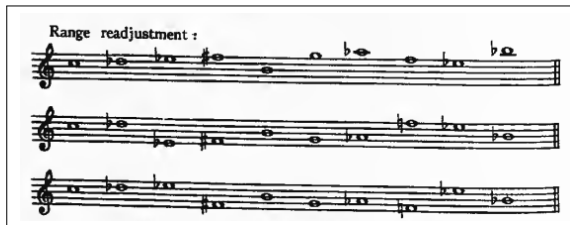


Figure 3. “Readjustment of the range”. From [13, p. 334]

generate infinite sequences of integers. Thus, they are rightful candidates to enter the algorithmic composer’s toolbox. As already observed by Schillinger, Fibonacci sequences rapidly grow outside a useful pitch range. Modulo application solves this issue but at the same time provides a new feature. When modulo n is applied, the result is a cyclic pattern that can be interpreted in relation to pitch domain. In this case, pitch sequences are interesting as they provide a non-uniform distribution of pitches while ensuring a variable periodicity. We can start by considering three cases:

1. $n < 12$
2. $n = 12$
3. $n > 12$

5.1 $n < 12$

While discussing the first case, we also introduce some general aspects.

If $n < 12$, then pitch classes $\geq n$ are necessarily missing from the resulting cycle. Figure 4 shows the pitch class pattern for $G_{[2,1]}$ with $n = 11$ (Lucas Sequence mod 11) (Figure 5 shows a line visualization for the melodic profile). It can be observed that pitch class 11 (= B) is missing. But while this holds true by definition, not all the other 11 pitch classes are present. It is possible to study various features of this pattern. Pitch classes in the sequence are only 7, the set being $\{0, 1, 2, 3, 4, 7, 10\}$. The Pisano period is 10. Finally, the number of occurrences for each pitch class is the following (in which the items in the array indicate occurrences for pitch classes 0 – 11): $[1, 1, 1, 2, 1, 0, 0, 3, 0, 0, 1, 0]$. This means that there are single occurrences of C, C \sharp , D, E, B \flat , 2 occurrences of D \sharp , 3 occurrences of G. Of course, there are no occurrences for B, but neither for F, F \sharp , G \sharp , A (Figure 6).

How to interpret the sequence in Figure 4? First, it can be seen as a specific melodic form. Second, it can be taken

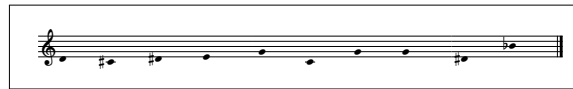


Figure 4. Pitch class sequence for $a = 2, b = 1, n = 11$.

into account as a generative cell, such as those used e.g. by Stravinskij ([15, 16]). Third, it can be exploited as a sort of weighted pitch class sequence: in the case of Figure 4 pitch class 7 (G) has more occurrences than all the other pitch classes, so it acts like a “modal” pivot for the whole sequence. By taking into account the resulting pitch class set, many analytical features can be explored. Ian Ring’s website provides an extensive analytical approach to scale patterns⁵. Ring has automatically computed an extensive set of analytical features from existing literature. In this context, a scale is an ordered sequence of pitch classes starting from 0 and including no intervals < 4 . A scale can be represented in bit form with the lowest bit (representing pitch class 0) at the right. The bit form of the set $\{0, 1, 2, 3, 4, 7, 10\}$ is thus $[0, 1, 0, 0, 1, 0, 0, 1, 1, 1, 1]$. This allows to compute the decimal form, that is, 1183. As a scale starts by definition from 0, the decimal is always odd. This number can be used as a pointer to a page in Ring’s website, so to explore the resulting scale pattern in relation e.g. to Forte number, Tonnetz, prime form, interval vector, chirality and many others⁶.

In short, generalized Fibonacci sequences mod n allows to generate simultaneously a scale *and* a repeating melody. A “Sequence Graph” like the one in Figure 7 shows in a compact way this twofold nature for the previously discussed $G_{[2,1]}$ with $n = 11$. The pitch sequence is $[2, 1, 3, 4, 7, 0, 7, 7, 3, 10]$. Nodes represent pitch classes (integers as labels), their radius being proportional to the number of occurrences (node 7 is the largest). Edges represent the sequencing order, as specified by each edge label. The graph is by definition directed and cyclic [17]. The dotted edge indicates the cycle loop and can also be used to quickly identify the starting node. Its label indicates the Pisano period.

5.2 $n = 12$

All the previous considerations apply to $n = 12$, with the difference that it is indeed possible to create a full chromatic pitch set (decimal notation: 4095). While this is theoretically possible, by considering all combinations of generalized Fibonacci sequences with $0 < a < 100$ and $1 < b < 100$ this does not happen, the best approximation being the pitch class $\{0, 1, 2, 3, 4, 5, 7, 8, 9, 10, 11\}$, that is, the chromatic set without F \sharp (decimal: 4031), in all cases with $\pi(n) = 24$. Figure 8 shows the patterns with pitch class set 4031 generated by lowest a (= 0) and b , respectively 1 (classic Fibonacci sequence) and 5 ($G_{[0,1]}$ and $G_{[0,5]}$ with $n = 12$). Even in case of the same pitch class set, still pitch sequences and occurrence counts are different.

⁵ <https://ianring.com/musictheory/scales/>

⁶ <https://ianring.com/musictheory/scales/1183>

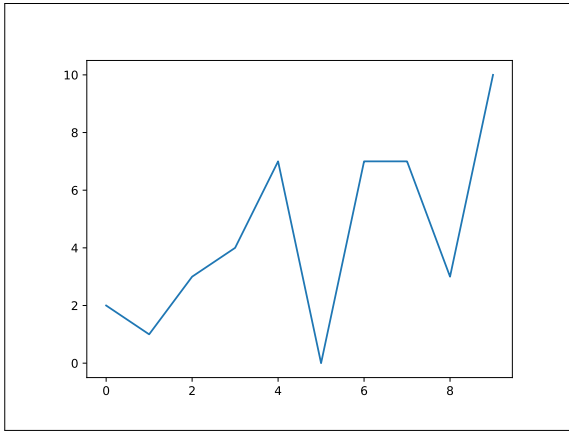


Figure 5. Melodic profile in $a = 2, b = 1, n = 11$.

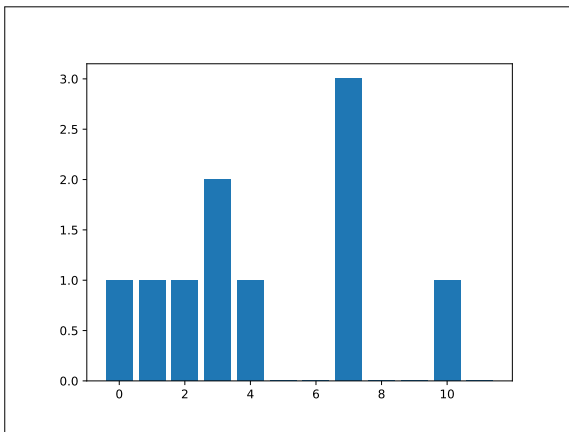


Figure 6. Occurrences for pitch classes in $a = 2, b = 1, n = 11$.

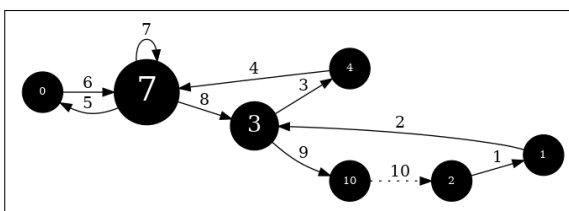


Figure 7. Sequence Graph for in $a = 2, b = 1, n = 11$.



Figure 8. Patterns for $a = 0$ and $b = 1$ (top) and $b = 5$ (bottom).



Figure 9. Patterns for $a = 2$ and $b = 1$ and $n = 12, 15, 18, 21, 24$.

If $b = 2$, then all sequences with $a = kb$ (where k is an integer), result in the pitch class set $\{0, 2, 4, 6, 8, 10\}$, with $\pi(n) = 24$, that is, a whole tone scale. Yet, pitch sequences are different, with different occurrence counts. In short, various hexatonic patterns can be generated. When $b = 4$, then all sequences with $a = kb$ result in the pitch class set $\{0, 4, 8\}$, with $n = 8$, that is, an augmented triad. Similarly, with $b = 3$ and $a = kb$, the pitch class set is $\{0, 3, 6, 9\}$, a diminished 7th chord.

In general, by Proposition 3.3.1, if b is a divisor of n , then all sequences $G_{[kb,b]}(\text{mod } n)$ result in the same set of pitch classes.

5.3 $n > 12$

If n is greater than 12 all the previous considerations apply but the obtained remainders result in a range of more than one octave. This feature can be exploited in order to distribute pitches over a larger span, in this sense privileging a pitch interpretation rather than a pitch class one. Figure 9 shows Lucas sequence ($G_{[2,1]}$) with five moduli ranging from $n = 12$ (top) to $n = 24$ (bottom). Meter is kept in the music notation so that different Pisano periods are clearly visible (respectively: 24, 8, 24, 16, 24). Extension of pitch range is clearly visible, and in particular top and bottom sequences ($n = 12$ and $n = 24$) show how the same pitches are distributed over 1 and 2 octaves respectively.

This approach leads to interesting results in terms of pitch sequences but does not allow to take into account the analytical features discussed in subsections 5.1 and 5.2, as those features apply not to pitches but to pitch classes. A second approach can thus be proposed so to exploit the richness of variable n while still placing the discussion at the pitch class level. The algorithm is the following:

1. Generate a generalized Fibonacci sequence G
2. Apply modulo $n > 12$ to G so to get G_1
3. Apply modulo $n_1 = 12$ to G_1 so to get G_2

G_2 represents a sequence of pitches in the range $[0, 1, \dots, 11]$, that is one octave, and actual pitches can be interpreted as pitch classes. Figure 10 shows the further application of modulo $n = 12$ to the patterns from Figure 9. Pitch patterns with $n = 12$ and $n = 24$ (top and bottom) result in the same pitch class sequence.



Figure 10. Patterns from Figure 9 with further $n_1 = 12$.

6. MORE GENERALIZATIONS

The classic Fibonacci sequence allows for several generalizations. An interesting case is given by the so-called k -Fibonacci sequence [18]:

$$F_k : \begin{cases} F_{k,0} = 0 \\ F_{k,1} = 1 \\ F_{k,i} = kF_{k,i-1} + F_{k,i-2}. \end{cases} \quad (10)$$

If $k = 2$, the Pell sequence appears. If $k = 3$ we get the following sequence

| | | | | | | | | | | |
|-----------|---|---|---|----|----|-----|-----|------|------|-----|
| i | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | ... |
| $F_{3,i}$ | 0 | 1 | 3 | 10 | 33 | 109 | 360 | 1189 | 3927 | ... |

(11)

The k -Fibonacci sequences satisfy nice properties. For example the following ones show useful relationship among the terms of the sequence:

- Catalan's identity:

$$F_{k,n-r}F_{k,n+r} - F_{k,n}^2 = (-1)^{n+1-r}F_{k,n}^2;$$

- Simson's identity:

$$F_{k,n-1}F_{k,n+1} - F_{k,n}^2 = (-1)^n;$$

- d'Ocagne's identity:

$$F_{k,m}F_{k,n+1} - F_{k,m+1}F_{k,n} = (-1)^n F_{k,m-n}.$$

More properties, with details of the proofs can be found in [19, 20, 21].

Also for this sequences we can talk of Pisano period thanks to the following result.

Theorem 6.1 ([18]). $F_k \pmod{n}$ is a simple periodic sequence.

In particular, the results of Corollaries 3.2.1 and 3.2.2 are valid also for k -Fibonacci sequences.

Taking a step further, as a new contribution, we define fully generalized Fibonacci as in the following:

$$H_{[a,b,k_1,k_2,s]} : \begin{cases} H_0 = a \\ H_1 = b \\ H_i = k_1H_{i-1} + k_2H_{i-2} + s. \end{cases} \quad (12)$$

where a, b, k_1, k_2, s are integers. Parameters a and b provide the initial steps, while k_1 and k_2 determine the size of the recursive steps and s is called the shift of the sequence. Obviously the Fibonacci sequence F corresponds to $H_{[0,1,1,1,0]}$, the sequences $G_{[a,b]}$ are obtained for $H_{[a,b,1,1,0]}$. Finally, with $H_{[0,1,k,1,0]}$ we get the k -fibonacci sequences F_k .

A larger amount of different sequences can be obtained by changing some parameters in the definition of Fibonacci recursion laws. By means of generalization a very large palette of numerical sequences becomes easily available to the music composer. Moreover, use of moduli introduces further richness. The previous discussion has shown how it may be complex to choose in advance the generator parameters that lead to a certain sequence without empirically verifying the results. In this sense, an experimental approach is the most viable approach to explore fully generalized Fibonacci sequences, as slight modifications of available parameters may have dramatic effects. As an example, let us consider the fully generalized Lucas sequence $H_{[2,1,3,5,3]}$, with $n = 23$ and $n_1 = 12$. The Pisano period is very large ($\pi(n) = 176$) and the sequence graph particularly complex, as it can be seen in Figure 11. In a computational environment for algorithmic composition, visualization and sonification displays allow to navigate these kinds of datascares.

7. IMPLEMENTATION

Implementation of fully generalized Fibonacci sequences is straightforward. Listing 1 shows two Python functions `fibor` and `fiboi` that compute a sequence of length `ln` by passing `a`, `b`, `k1`, `k2`, `s` as arguments. The first is recursive, the second iterative. Length must be empirically large enough to comprise Pisano period (here, 1000). Modulo can be applied by making use of list comprehension on the resulting `seq` sequence (line 9) by applying n and then $n_1 = 12$ so to get pitch classes. Here we are computing $H_{[2,1,3,5,3]}$ with $n = 23$ and $n_1 = 12$ as discussed in section 6.

```

1 def fibor(a=0, b=1, k1=1, k2=1, s=0, seq = [], i
  = 0, ln = 1000):
2     if i < ln:
3         c = (a*k1) + (b*k2) + s
4         seq.append(c)
5         fibor(b, c, k1, k2, s, seq, i+1, ln)
6     return [a,b]+seq
7
8 def fiboi(a=0, b=1, k1=1, k2=1, s=0, ln=1000):
9     seq = [a, b]
10    for x in range(ln):
11        a, b = b, (a*k1) + (b*k2) + s
12        seq.append(b)
13    return seq
14
15 n = 23
16 seq = [(x % n) % 12 for x in fiboi(2,1,3,5,3)]
    
```

Listing 1. Python example.

Functions for period estimation and for data analysis (e.g. occurrence count) and conversion (e.g. decimal notation of pitch class set) are similarly trivial. Jupyter Notebook⁷ provides a web-based computing platform that offers an interactive environment for Python development⁸. In the context of algorithmic composition, Jupyter Notebook allows to exploit the large ecosystem of Python libraries. As an example, after computing the decimal representation of a pitch class set, by means of the `webbrowser` module it is possible to automatically access Ian Ring's pages dedicated to that scale. Figures 5 and 6 have been generated using the standard `matplotlib` module for plotting. The sequence graph in Figure 7 has been generated with the `pydot` module, a Python interface for Graphviz [22]. The same module has been used to generate the sequence graph for $H_{[2,1,3,5,3]}$ with $n = 23$ and $n_1 = 12$ computed in Listing 1, shown in Figure 11. The Python Music21 package provides an integrated bridge towards music computation and notation [23]. It includes functionalities to analyse, plot and notate music data. In relation to music notation, it allows to generate MusicXML that can be further manipulated in music notation softwares (e.g. MuseScore⁹). Like the other previously mentioned modules, Music21 can be directly accessed via Jupyter Notebook, so that notation output can be displayed in the Notebook itself (via MuseScore backend) and results can be heard (via MIDI playback). All music examples (Fig. 4, 8, 9, 10), have been generated via Music21 and MuseScore. Further data sonification can be obtained by the direct usage of the SuperCollider¹⁰ environment into Jupyter Notebook by means of the `sc3nb` Python-interface [24].

8. CONCLUSIONS AND FUTURE DEVELOPMENTS

Fully generalized Fibonacci sequences are simple yet powerful generative devices that results in complex, unpredictable patterns that, once the modulo operation is applied, nevertheless lead to repeating organizations. These sequences can be interpreted in the music domain in various ways. In our proposal, we mapped integers onto indices for pitch and pitch class description. In the latter case, modulo 12 is the key operation that enables to lead back the obtained results to the vast music theory literature dedicated to the topic. It is worth mentioning that obtained sequences can indeed be manipulated further. As an example, transposition results in adding a factor t to all the integers in a sequence. As the operation is obvious, we have focused on sequence properties rather than on pitch classes in themselves (e.g. we have not taken into account that Fibonacci starts on C while Lucas on D, as pitch organization can be transposed at will). Fully generalized Fibonacci sequences favors an experimental, empirical approach, as through computational environments it is possible for each sequence to

explore pitch class set composition, occurrence distribution, Pisano period. In short, fully generalized Fibonacci sequences are useful tools for algorithmic composition.

The previous discussion focused on pitch classed based on 12-ET. We can think of extending the approach by taking into account non-integer numbers. As an example, $4 \bmod 2.5 = 1.5$. Fractional parts can be interpreted as semitone fractions, like in MIDI notation, where e.g. 60.5 indicates a middle C raised of a quarter tone. Indeed, other fractions can be obtained. Non-integers can be applied in fully generalized Fibonacci both to parameters a, b, k_1, k_2, s and to $\bmod n$. If holding the octave constraint, by further applying $\bmod n = 12$ it is possible to create non-12-ET pitch class sets. In relation to this, a different approach to be explored is to define a specific mapping function that maps the output integers from H onto specific sets of non-integers values. Finally, while we consider pitch and pitch class interpretation as particularly promising, fully generalized Fibonacci $\bmod n$ can be used to generate values to control arbitrary cycling parameters for sound and music computing. As a suggestion, periodic sequences can be used to fill looping wavetables for audio synthesis. In this sense, one can think of Figure 5 as a wavetable plot. An interesting feature is that these wavetables are amplitude limited, as the amplitude range is by definition in the range $[0, n - 1]$, where n is the last applied modulo.

9. REFERENCES

- [1] J. Kramer, "The Fibonacci Series in Twentieth-Century Music," *Journal of Music Theory*, vol. 17, no. 1, pp. 110–148, 1973.
- [2] J. Kochavi, *The Fibonacci Sequence as Metric Suspension in Luigi Nono's Il Canto Sospeso*, ch. Chapter 5, *Mathematical Music Theory*, pp. 83–108. WORLD SCIENTIFIC, 2018.
- [3] G. Albert, "Weakening Structures or Structuring Mistakes? Brian Ferneyhough's Manipulation of the Fibonacci Sequence in his Second String Quartet," *Mitteilungen der Paul Sacher Stiftung*, vol. 28, pp. 55–60, 2015.
- [4] E. L. Lowman, "An Example of Fibonacci Numbers Used to Generate Rhythmic Values in Modern Music," *The Fibonacci Quarterly*, vol. 4, no. 9, pp. 423–26, 1971.
- [5] A. Shannon, I. Klamka, and R. van Gend, "Generalized Fibonacci Numbers and Music," *JOURNAL OF ADVANCES IN MATHEMATICS*, vol. 14, pp. 7564–7579, 05 2018.
- [6] C. Mongoven, "A Style of Music characterized by Fibonacci and the Golden Ratio," *Congressus Numerantium*, 01 2010.
- [7] D. Wall, "Fibonacci series modulo m ," *The American Mathematical Monthly*, vol. 67, no. 6, pp. 525–532, 1960.

⁷ <https://jupyter.org/>

⁸ See the implementation on colab: <https://github.com/vanderaalle/FullyGeneralizedFibonacciSequences/blob/main/FibonacciModuloN.ipynb>

⁹ <https://musescore.org/>

¹⁰ <https://supercollider.github.io/>

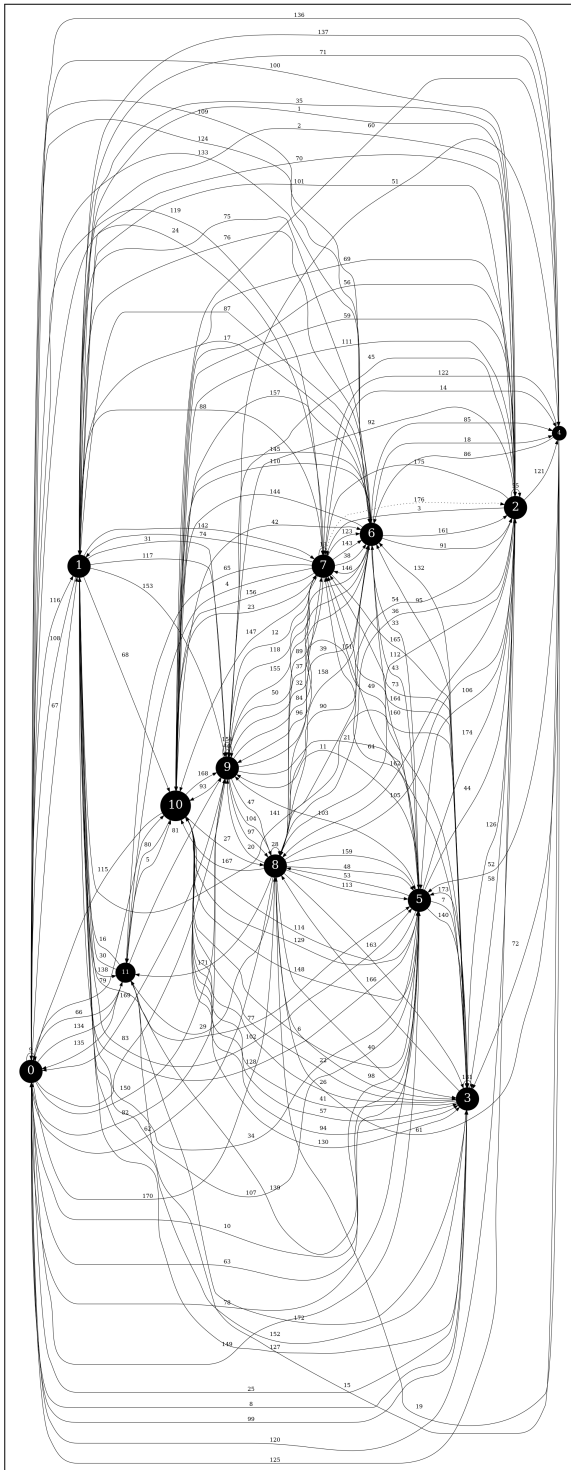


Figure 11. Graph for $H_{[2,1,3,5,3]}$ with $n = 23$ and $n_1 = 12$.

- [8] D. Fulton and W. L. Morris, "On Arithmetical functions related to the Fibonacci numbers," *Acta Arithmetica*, vol. 16, 1969.
- [9] A. Catlin, "A lower bound for the period of the Fibonacci series modulo m ," *Fibonacci Quarterly*, vol. 12, pp. 349–350, 1974.
- [10] A. Forte, *The Structure of Atonal Music*. Yale University Press, 1973.
- [11] J. Rahn, *Basic Atonal Theory*. Longman music series, Longman, 1980.
- [12] J. Straus, *Introduction to Post-tonal Theory*. Prentice Hall, 1990.
- [13] J. Schillinger, L. Dowling, and A. Shaw, *The Schillinger System of Musical Composition*. C. Fischer, Incorporated, 1946.
- [14] J. Haek, "Residue Cycles of Fibonacci Series Modulo m as Tools for Serial Composition," *Perspectives of New Music*, vol. 46, no. 2, pp. 33–58, 2008.
- [15] J. N. Straus, "A Strategy of Large-Scale Organization in Stravinsky's Late Music," *Integral*, vol. 11, pp. 1–36, 1997.
- [16] A. Rehdig, "Towards A 'Logic of Discontinuity' in Stravinsky's 'Symphonies of Wind Instruments': Hasty, Kramer and Straus Reconsidered," *Music Analysis*, vol. 17, no. 1, pp. 39–65, 1998.
- [17] R. Diestel, *Graph Theory*. Springer, August 2005.
- [18] A. Falcon and A. Plaza, " k -Fibonacci sequences modulo m ," *Chaos Solitons & Fractals*, vol. 41, pp. 497–504, 2009.
- [19] A. Falcon and A. Plaza, "On the Fibonacci k -numbers," *Chaos Solitons & Fractals*, vol. 32, no. 5, p. 1615–1624, 2007.
- [20] A. Falcon and A. Plaza, "The k -Fibonacci sequence and the Pascal 2-triangle," *Chaos Solitons & Fractals*, vol. 33, no. 1, pp. 38–49, 2007.
- [21] A. Falcon and A. Plaza, "On k -Fibonacci sequences and polynomials and their derivatives," *Chaos Solitons & Fractals*, vol. 39, pp. 1005–1019, 2007.
- [22] E. R. Gansner and S. C. North, "An open graph visualization system and its applications to software engineering," *Software: Practice and Experience*, vol. 30, no. 11, pp. 1203–1233, 2000.
- [23] M. S. Cuthbert and C. Ariza, "Music21: A Toolkit for Computer-Aided Musicology and Symbolic Music Data," in *ISMIR* (J. S. Downie and R. C. Veltkamp, eds.), pp. 637–642, 2010.
- [24] T. Hermann and D. Reinsch, "sc3nb: a Python-SuperCollider Interface for Auditory Data Science," in *Proceedings of the 16th International Audio Mostly Conference, AM '21*, (New York, NY, USA), p. 208–215, 2021.

LA SPERIMENTAZIONE VOCALE NELL'EPOCA DELL'INTELLIGENZA ARTIFICIALE

Paolo Paradiso

Università di Pavia

paolo.paradiso01@universitadipavia.it

ABSTRACT

In questo saggio verranno prese in considerazione alcune innovative possibilità creative offerte dalle tecnologie digitali con intelligenza artificiale (d'ora in avanti IA) nell'ambito della sperimentazione vocale. Data la portata trasformativa di queste tecnologie, risulta oltremodo necessario distinguere ciò che sull'IA è una realtà oggi e ciò che, invece, rientra nell'ambito della speculazione e fantascienza. L'IA sfida su diversi fronti importanti questioni sociali, politiche e culturali. Inoltre, pone quesiti filosofici sull'entità umana stessa: la definizione di "umano" tradizionalmente intesa necessita di un ripensamento alla luce dei progressi tecnologici sempre più avanzati? Si sta andando incontro ad un "postumanesimo"? E in musica, tutto ciò, quali conseguenze può avere? In questo saggio si tenterà di affrontare queste ed altre questioni, non sempre pretendendo di restituire risposte univoche e definitive, ma piuttosto aprendo la strada a discussioni diversificate fra loro, con l'obiettivo di provare a portare maggiore chiarezza su cosa significhi oggi fare musica servendosi dell'IA. Il caso di studio analizzato riguarda la composizione sperimentale *ULTRACHUNK*, per voce e IA, della compositrice Jennifer Walshe, che consente di convocare riflessioni sulle teorie filosofiche del postumanesimo e applicarle al mondo musicale.

1. INTRODUZIONE

L'universo che si apre con le nuove tecnologie permette di accogliere in maniera sempre più accurata le necessità di sperimentazione degli artisti e musicisti contemporanei. Attraverso l'elettronica si ha infatti la possibilità di superare i limiti tecnici e pratici del passato e intraprendere direzioni estetiche inusuali e sconosciute.

Copyright: © 2024 Paradiso. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Il digitale, il cyberspazio, la realtà virtuale e l'infosfera¹ sono dimensioni alternative alla realtà fisica che sollecitano innovative forme d'interazione tra umano e tecnologia. Questi assetti tecnologici: «permettono da un lato la "simulazione" di universi modellizzati, e percettivamente esperibili, ma anche di giungere alla creazione di spazi, ambienti, situazioni, dove il fattore estetico viene ad essere influenzato dalle nuove applicazioni elettroniche» [14]. In questo modo il corpo dell'artista si dissemina nello spazio (virtuale e/o reale), attraverso modalità differenti e raggiungendo molteplici esiti estetici.

Oltre a una disseminazione spaziale, anche la dimensione temporale viene influenzata da riprogrammazioni che prevedono la possibilità di riproduzione, archiviazione e manipolazioni di vario tipo. L'evento torna (in contrapposizione al documento e nell'ambito di un presente immerso in un'oralità terziaria²), ricoprendo un ruolo fondamentale in una logica performativa: «nel quale il corpo vive il suo rapporto con lo spazio, il tempo, gli oggetti, le protesi strumentali» [14]. In questo contesto la presenza sconfinata in telepresenza differita, in parte svolta nel cyberspazio e in grado di mobilitare "energia": «dove energia sta per combinazione di energia vitale ed energia artificiale» [14]. Ecco che sorgono le possibilità di creare voci nuove, voci macchiniche personalizzate, voci in grado di mettere in discussione la tradizionale concezione della voce umana stessa, innescando riflessioni di codeterminazione, influenza ed evoluzione reciproca.

All'interno di questa condizione di prossimità con la tecnologia, la pratica del *live electronics* rappresenta il dispositivo per eccellenza per l'espressione delle nuove istanze performative, definibili (come si specificherà più avanti) come "postumane". Il pensiero postumanista³, infatti, apre ad un decentramento dell'*anthropos* e

¹ Con questo termine ci si riferisce al concetto delineato da Luciano Floridi di ambiente informazionale formato da vari enti, con le loro caratteristiche e in relazione fra loro [12].

² Riprendendo l'analisi di Walter J. Ong su oralità primaria e secondaria, Derrick de Kerckhove, per oralità terziaria, intende l'oralità: «dei sistemi multimediali, della realtà virtuale e della Rete, è un'oralità elettronica, come la "seconda", ma si basa sulla simulazione della sensorialità, non più sulla sua trasmissione. Ad esempio, il sonoro dei "beep" delle piccole macchine elettroniche, dal computer al telefonino, è un linguaggio tattile che restituisce un feedback rispetto alle nostre azioni, è una simulazione organica» [5].

³ Il termine "postumano" possiede innumerevoli sfaccettature. In generale, racchiude la necessità di riflettere su cosa significhi essere

della sua soggettività dalla scena sociale. In questo modo, consente l'adeguato riconoscimento dell'agentività non-umana e nello specifico di quella macchinica dell'IA, che va così a ricoprire un ruolo attivo nel dispositivo performativo, immerso a sua volta in un ecosistema misto e vissuto dai rapporti rizomatici dei molteplici altri [4]. La voce si inserisce nel mezzo di questa performatività relazionale, in una posizione (sua per definizione) di *extimità* [9].

La voce "altra", quella altamente mediata dalla tecnologia, fa risaltare così la logica abduittiva/speculativa dell'IA, la sua specifica agency, nettamente distante da quella umana, ma che con quest'ultima può entrare in contatto producendo risultati inaspettati, ibridi. L'imitazione dell'umano diventa allora una delle possibili strade da percorrere con le tecnologie musicali munite di IA, ma non l'unica.

Come si vedrà in seguito, le sperimentazioni compositive di Jennifer Walshe rappresentano un caso emblematico del rilevante apporto dell'arte nell'esplorazione dell'IA senza finalità commerciali, solitamente più legate, invece, alla creazione di sistemi per imitare e automatizzare la produzione musicale, così da renderla più efficiente e veloce. Trattando il versante tecnologico come attante capace d'influenzare in maniera significativa la composizione finale, Walshe è in grado di sfruttare a proprio vantaggio l'IA e le sue peculiari e specifiche abilità rappresentative dei dati d'addestramento e di input in tempo reale: che sia tramite la rielaborazione macchinica della propria voce, messa in relazione al progredire della storia della musica occidentale (come nell'album *A Late Anthology of Early Music vol. 1. Ancient to Renaissance*); o in improvvisazioni estemporanee con uno suo "doppio artificiale" (come in *ULTRACHUNK*), l'agency macchinica viene messa sotto i riflettori, ponendo in questo modo nuove questioni estetiche in primo piano.

2. ULTRACHUNK

2.1 Per voce e IA

ULTRACHUNK è un brano musicale di Jennifer Walshe composto per voce e rete neurale artificiale, in collaborazione con il techno-artista Memo Akten. La performance è stata eseguita in diverse occasioni, ma per questo saggio verrà presa in considerazione l'esecuzione del 2018 come parte della rassegna Assembly della Somerset House Studios di Londra, che ha anche commissionato il progetto⁴.

umani nell'era del digitale (o nella quarta rivoluzione industriale [13]). Al suo interno troviamo pensieri, correnti e movimenti molto diversi e anche in contrapposizione fra loro. In questo articolo si è scelto di far riferimento al postumanesimo filosofico, in particolare attraverso la lente dell'analisi critica svolta da Rosi Braidotti, aperta alle prospettive post-antropocentriche e post-dualistiche [4, 11].

⁴ Grazie anche sul supporto e ai finanziamenti di The Case Foundation e del Goethe-Institut. La registrazione di questa performance è fruibile sul sito: <https://www.memo.tv/works/ultrachunk/> (ultima consultazione: 19/04/2024).

Walshe ha lavorato a lungo in sperimentazioni musicali che coinvolgono il machine learning. La peculiarità di quest'opera è di consistere in un'improvvisazione dal vivo tra Walshe in qualità di vocalist e il suo alter-ego artificiale, sottoforma di audio e video. In questa performance estemporanea, Walshe e la rete si auto-influenzano continuamente nella scelta e produzione dei suoni. Per l'addestramento della rete è stato necessario assemblare un corposo dataset di registrazioni d'improvvisazioni vocali di Walshe.

La varietà delle tipologie di vocalizzazioni eseguite da Walshe denota il suo estremo controllo e la sua profonda conoscenza del proprio apparato vocale. Prediligendo suoni non linguistici, Walshe comincia l'improvvisazione alternando suoni vocalici e consonantici senza significato semantico definito, in risposta immediata a un primo suono elettronico che apre la performance, inserendosi quasi come una sorta di tappeto sonoro e che fin dai primi interventi di Walshe evolve in conseguenza a questi. La performance inizia con una generale bassa intensità di volume, come a indicare una ricerca di interconnessione umano-tecnologia che può sorgere solamente in uno stato di concentrazione e ascolto attento nei confronti del partner, sottolineato dall'alternanza di momenti sonori e di silenzio. La destrezza vocale di Walshe si manifesta fin dai suoi primi interventi, dove molto agilmente dimostra grande disinvoltura nei cambi di registro. Oltre a suoni con coinvolgimento delle pliche vocali, Walshe esegue anche "rumori" corporei del tratto vocale, come possono essere: schiocchi di lingua e di guance; sibili, sospiri e forti ispirazioni ed espirazioni; colpi di glottide (che spesso anticipano attacchi fonatori bruschi); colpi di tosse; rumori di masticazione e faringei che sfruttano la saliva e il passaggio dell'aria e fischi. Altri degli effetti vocali che si possono citare sono: uso di growl; vocal fry; tongue trill; attacchi soffiati e acuti con vibrato rapido. La quantità degli effetti prodotti da Walshe presuppone una padronanza, oltre che della fonazione effettuata prettamente tramite le pliche vocali, anche dell'utilizzo espressivo di tutti gli altri elementi implicati: labbra, palato molle, lingua, mandibola e glottide. La gestione del fiato risulta fondamentale per l'esecuzione delle varietà delle tecniche articolatorie, fonatorie e di manipolazione della risonanza coinvolte.

2.2 Una voce autoriflessiva incontra la voce della macchina

L'insieme degli elementi citati nel paragrafo precedente testimonia la discendenza artistica di Walshe dalla sperimentazione della *vocal performance art* e il coinvolgimento di tecniche proprie della vocalità estesa. Continuando l'operato di una prolifica schiera di cantanti del XX secolo (che fecero della propria voce il punto di partenza della propria ricerca estetica), Walshe utilizza il suo strumento vocale in modo tale da produrre una voce: «autoriflessiva, in grado di andare al di là delle norme codificate della vocalità e capace di integrare nel registro del canto un universo di suoni vocali non ne-

cessariamente linguistici» [15]. La “riflessione” sulla propria voce riguarda allo stesso tempo sia una ricerca delle proprie possibilità di espressione sonora vocale, sia un ripensamento e un allargamento delle tradizioni vocali socialmente e culturalmente iterate nel tempo, un tentativo per superare: «the boundaries of established vocal codes and, at the same time, capturing the cultural unrest of the time» [16].

La voce viene allora attraversata da una portata comunicativa gestuale, per lo più slegata dalla dimensione simbolica del linguaggio⁵, che ha trovato terreno fertile proprio in quelle pratiche di vocal performance art (dalla seconda metà del Novecento in avanti), che presuppongono: «una presa di coscienza della [loro] trasversalità⁶ [...] del loro radicamento in un contesto sperimentale, [e] del loro carattere inevitabilmente mediato tanto dalla scrittura, quanto dalla tecnologia» [15].

Quello che rimane fondamentale, nella pratica vocale performativa di Walshe, è proprio l’apporto della mediazione tecnologia, con la quale ha rapporti artistici che si radicano profondamente nella sua estetica: «I regard my voice as an amplified voice, first and foremost, not an a cappella voice. I think of my voice as happening in tandem with the microphone, the PA, the space and the acoustic. It’s an amplified voice, designed to have a wide dynamic range and have everything picked up» [10].

La tendenza prevalente nella vocal performance art prevede una graduale emancipazione della figura dell’esecutore-cantante da quella del compositore: da modello (o “musa”) attorno cui costruire l’opera musicale, a vere e proprie funzioni collaborative attive (come possono essere state, solo per presentare degli esempi, Cathy Berberian a fianco di Luciano Berio, Carla Henius con Luigi Nono e Michiko Hirayama con Giacinto Scelsi).

Il passo successivo è la sovrapposizione del performer-cantante con quella del compositore. Si forma in questo modo il nuovo ruolo artistico del *VocalPerformerComposer*, che: «segna in maniera netta una sospensione del rapporto gerarchico tra compositore ed esecutore e costituisce una fase cruciale nel processo di emancipazione e di soggettivazione della voce che ha interessato tutti i generi nel Novecento» [15]. Questo non va confuso con l’affermarsi di un monolitico autore-esecutore chiuso nei propri progetti estetici. Anzi, proprio con il crescente coinvolgimento e utilizzo di sistemi elettronici per fare musica (come possono essere i software di machine learning), l’apertura verso collaborazioni multiple s’ispessisce. Nel live electronics (dove è prevista la manipolazione vocale dal vivo), è

infatti consueta la presenza dell’interprete elettronico a fianco del cantante, in una cornice performativa che diventa, pertanto, condivisa. La voce stessa viene distribuita tra diversi collaboratori, coinvolti in un processo di manipolazione dove: «la/il cantante non ha più il diretto controllo sull’esito; passa nelle mani del tecnico del suono [...]. Si instaura qui un importante rapporto di collaborazione» [22].

In aggiunta a questa nuova divisione del lavoro (ripartito quindi tra più soggetti), l’agentività tecnologica entra nell’ambito performativo come nuovo tipo di attante con ruolo attivo, alimentando e complessificando il paradigma della creatività distribuita: «Coinvolgendo agenti umani e non umani, [...] all’interno di questo orizzonte si collocano assetti di tipo strettamente collaborativo, dove la creatività è condivisa in modo intenzionale» [26] e dove i lavori di ciascun membro di questa rete giocano un ruolo fondamentale d’influenza reciproca.

La pervasività dello sviluppo tecnologico concorre poi a ricalibrare le definizioni di “voce”, già messe in crisi dalle inclinazioni generali delle pratiche vocali del Novecento, indirizzate alla de-soggettivazione della voce, presentata come voce-oggetto: «La tendenza alla depersonalizzazione amplia la dimensione della voce e mette in questione gli ordini e le gerarchie su cui è costruita la soggettività occidentale, spesso anche grazie soprattutto all’uso della tecnologia *live electronics*» [15]. La voce si riappropria così di emissioni prelinguistiche storicamente escluse dalle pratiche vocali stilizzate, alla ricerca di: «sonorità residuali sempre molto presenti nella voce [...] ma generalmente rimosse nelle pratiche di canto e di dizione. [...] Il tratto vocalico non è più esperito come “corpo” ma “spazio”» [8]. Si ha così una riscoperta delle possibilità fonatorie dimenticate e trascurate che scava all’interno della: «dimensione emancipatrice del semiotico» [15] e si lega alla tecnopoiesi nella ridefinizione del ruolo umano nell’arte e nella musica, non in senso dualistico (umano/tecnologia) ma continuativo [8].

Attraverso queste modalità si delinea in maniera più precisa la figura del “composer-technologist-vocalist”, o, più in generale, di “poliartista”, categoria di artista (alla quale si può inscrivere la stessa Walshe) in grado di gestire competenze tecnico-pratiche e artistico-culturali. Un’artista che: «agisce poeticamente utilizzando tutte le tecniche, tutti i supporti, tutti gli spazi, ma, principalmente, riconducendo all’ambito poetico il suo stesso corpo, quindi il suo gesto e la sua voce. [...] Una tendenza che [...] [fa] leva sulla contaminazione dei sistemi, sulla compenetrazione degli universi separati, sull’uso di nuovi media e di nuovi supporti, [...] attraverso una diversa concezione della materialità del linguaggio, sostenuto dalla voce» [14].

2.3 Voce clonata

In *ULTRACHUNK*, la voce artificiale viene generata estemporaneamente grazie alla rete neurale artificiale GRANMA (*Granular Neural Music & Audio*), creata

⁵ «Questo lavoro sulla voce, da una parte comporta una mortificazione del suo potenziale significante, dall’altra sortisce l’effetto di una riscoperta di una totalità della significanza corporea, storicamente neutralizzata dalla pratica che Barthes definiva feno-canto» [17].

⁶ Cathy Berberian parlò a tal proposito di “nuova vocalità”, una modalità per affrontare la voce aperta alle contaminazioni di stile e genere e accogliente nei confronti dei “rumori comunicativi” attraverso cui si può esprimere la voce [1].

da Memo Akten. Servendosi di un immenso dataset di ore ed ore di materiale audio e video della stessa Walshe mentre improvvisa con la sua voce davanti a una webcam e a un microfono, la rete produce delle inferenze probabilistiche in risposta agli input ricevuti durante l'improvvisazione della performance. In questo modo, la voce autoriflessiva guadagna un preziosissimo mezzo per la sua ricerca introspettiva: il suo doppio artificiale⁷.

Questa voce artificiale, che rappresenta l'alter-ego tecnologico di Walshe, ha come prima caratteristica quella di palesare in maniera evidente la sua natura elettronica all'ascolto. Produrre voci artificiali dalle sonorità quasi indistinguibili da quelle umane, è oggi una realtà, grazie ai metodi avanzati di machine learning. Questo non significa che, per scelte artistiche, espressive ed estetiche, non si possa optare per un timbro che vada, invece, proprio ad evidenziare l'origine digitale della voce, come avviene in *ULTRACHUNK*. Il concetto di "naturalità" della voce artificiale ha avuto, già dagli anni Settanta del Novecento, un grande peso nella ricerca e nello sviluppo dei metodi sintetici di produzione vocale, surclassando d'importanza il precedente obiettivo dell'intelligibilità: «Con l'imporsi di sistemi *model-based* atti a formare modelli computazionali dell'apparato fono-articolatorio umano e dell'atto del parlare, la naturalezza è divenuta l'obiettivo principale della sintesi vocale» [20]. Questo "parametro antropomorfo" di maggiore o minore fedeltà al modello umano può però essere utilizzato anche per la ricerca di un suono meno umano e più elettronico, che sottolinei le diverse origini vocali e in questo modo contribuisca a mettere in luce le peculiarità timbriche da un lato di Walshe e dell'altro del suo doppio tecnologico. Non nascondendo, ma anzi mettendo in primo piano, la fonte biologica e quella digitale delle due voci, ci si allinea ai principi di riconoscimento e apprezzamento delle specifiche agentività all'interno di un unico insieme cooperativo e interconnesso. A tal proposito Agostino Di Scipio parla della formazione di un'"agentività unitaria" che comprende al contempo umani, macchine e ambiente, dal carattere "ecosistemico": «e non genericamente interattivo [...]. Il dispositivo performativo sarà allora *ecosistema performativo*. E avrà carattere *adattivo* nel momento e nella misura in cui, adeguando sé stesso a potenzialità e limitazioni dello spazio circostante, agisce anche su tale spazio per adeguarlo a sé» [8]. Un ambiente, quindi, ibrido, nell'ottica postumana di ecosistema performativo [8]. È attraverso composizioni quali *ULTRACHUNK* che risulta, allora, possibile: «indagare il nesso di voce, corpo e spazio nella viva reciprocità dinamica di agenzie diverse [...], di cui le pratiche per-

⁷ L'insieme dei dati per l'addestramento della rete è stato usato per costruire l'alter-ego artificiale di Walshe anche per ricreare artificialmente l'immagine su un grande schermo posto alle spalle di Walshe (la quale è, invece, posizionata sul palcoscenico leggermente decentrata, munita di asta, microfono e due altoparlanti spia collocati a terra e rivolti verso di lei). L'effetto, nella manipolazione digitale visiva, è di continuo mutamento dell'aspetto del viso di Walshe e dello sfondo alle sue spalle, come una sorta d'ininterrotto morphing delle numerose varianti della stessa Walshe archiviate nel dataset.

formative *live electronics* restituiscono l'essenziale "continuità" con esiti di volta in volta singolari» [8].

È possibile riscontrare questa stretta cooperazione durante tutta la performance. Tuttavia, è possibile individuare momenti di maggiore evidenza relazionale d'influenza reciproca. Questo è rintracciabile nei numerosi richiami, imitazioni, variazioni, proposte musicali alle volte accolte, altre volte modificate, all'interno di un intreccio che rende difficoltoso comprendere con chiarezza chi, di volta in volta, abbia suggerito un'idea e chi l'abbia ripresa. Non emerge, pertanto, un conduttore e un imitatore o seguatore, quanto una compartecipazione alla pari. Gli ultimi minuti della performance, ad esempio, sono carichi di momenti di continua e vicendevole influenza musicale e individuarli tutti non è chiaramente possibile, considerando soprattutto che si tratta di un flusso ininterrotto d'improvvisazione estemporanea senza una struttura fissa, che quindi gode di una grande libertà, non dovendo seguire schemi, modi o ritmi precostituiti.

Consistendo di un insieme di dati ricavati dalla stessa Walshe, le aspettative di risposta da parte dell'IA sono ristrette al training set di partenza, che rappresenta il ventaglio di possibilità per la rete neurale artificiale e le sue combinazioni estemporanee⁸. La componente non prevedibile risiede, invece, proprio nella rielaborazione dal vivo del materiale usato nell'addestramento, in stretta simbiosi con i contributi improvvisati da Walshe.

A permettere la clonazione artificiale della voce di Walshe è l'approccio con deep learning, attraverso una computerizzazione delle caratteristiche fondamentali della sua voce. La rete neurale artificiale è in grado, in questo modo, di creare un suo: «*profilo vocale* sintetico [...] e lo impiega per generare nuovo suono vocale. [...] Questo approccio è [...] completamente guidato dai dati: l'algoritmo non riceve istruzioni su cosa sia la voce [...], piuttosto estrae quel tipo di conoscenza direttamente dall'analisi dei dati» [20]. L'algoritmo è impostato per riconoscere pattern e correlazioni ricorrenti, così da poterne estrarre delle previsioni per i successivi output. Grazie alla logica abduktiva e speculativa dell'IA basata su dati, è possibile programmare una rete neurale artificiale che: «non si limita a ricomporre frammenti dei campioni sonori, ma piuttosto estrae da quei campioni le caratteristiche che li rendono quello che sono [...]; poi utilizza quelle caratteristiche apprese per generare, sinteticamente, suoni simili ma nuovi» [20]. La si può considerare un'evoluzione della sintesi concatenativa, basata sulla ricomposizione dei dati di partenza. La differenza tra questo tipo di approccio e la linea statistica del machine learning è il diverso tipo di logica dell'archivio soggiacente. Nel caso della sintesi concatenativa, si segue la logica dell'assemblaggio o

⁸ Questo tipo di reti neurali artificiali: «are trained, they're not symbolic AI where somebody has coded every possible permutation. It's important to be aware of the limits of the dataset that you train these networks on, and that the network only understands this learning. Your own biases within that dataset, in bringing it together, are all fed in». [10].

del database. Per il machine e deep learning, vige invece la logica dell'addestramento (o, come già detto, "abduzione/speculativa"): «in cui la conoscenza deriva da processi di apprendimento macchinico e si dispiega come generalizzazione di nuovi elementi non presenti nel database di addestramento» [20]. Grazie alla peculiare "auto-ottimizzazione macchinica", l'IA a base dati può così includere nella sua versione di sintesi vocale gli aspetti più legati al corpo della voce (respiri, sibili, rumori causati da frizioni e movimenti delle labbra, masticazione, intercalari ecc.). La voce, infatti, non viene trattata come fonetizzazione della scrittura, ma come pura materia sonora [20].

3. PERFORMATIVITÀ POSTUMANA

3.1 Agentività macchinica e agentività umana

Il caso di *ULTRACHUNK* conduce a un'integrazione operativa tra l'agency umana e quella macchinica, che cooperano in uno spazio condiviso. Pertanto, si può affermare che sia un caso nel quale: «le dinamiche dell'uno attraversano quelle dell'altro, interdipendenti tra loro e insieme co-dipendenti dall'ambiente condiviso. [...] Si crea un unico dispositivo o *meta-organismo di performatività vocalica*, che è pur sempre "corpo vocalico", ma di costituzione mista – un corpo che la voce diffrange e proietta in uno spazio vocalico a sua volta reticolare e differenziato» [8].

Alla luce di quanto detto, è possibile interpretare un'ulteriore sfumatura del significato del titolo di questa performance. Attingendo direttamente alle scienze cognitive si comprende come: «a chunk is a collection of elements having strong associations with one another» [18]. Raccogliendo informazioni dal proprio ambiente, gli agenti coinvolti nella performance (Walshe e la rete) reagiscono ai dati in input per processare in maniera estemporanea degli output contestuali. Questo processo viene messo in atto, come è già stato evidenziato, in una costante interazione che prevede una selezione e gerarchizzazione di quanto ricevuto dall'esterno, inferendo le formulazioni musicali immediatamente successive. Il prefisso "ultra" connota la performance in quanto interazione che va "oltre", che va "al di là", della collaborazione improvvisativa esclusivamente umana, integrando in maniera attiva l'agentività macchinica non umana e creando un contatto tra l'organico e l'inorganico.

Inoltre, *ULTRACHUNK* rappresenta un caso d'interazione tra voce umana e voce artificiale che riprende e trasporta in campo musicale molti dei presupposti affrontati dalla prospettiva postumanistica, dove la soggettività umana non rappresenta più il fulcro e il modello perfetto a cui riferirsi, ma uno degli attori coinvolti nella realizzazione della performance, legato ad altri attori, umani e non [4, 11]. La controparte artificiale acquisisce così un ruolo attivo nella performance e

solamente adoperando un "ascolto postumano"⁹, è possibile riconoscere la specificità macchinica in tutto il suo valore e la sua reciproca appartenenza con la componente umana. Espandere l'idea di ascolto consente proprio di immaginare modalità attraverso cui altre agency non-umane possano interpretare (ovvero, nel caso dei sistemi informatici con IA, decodificare matematicamente) i suoni dell'ambiente condiviso, restituendo all'umano il risultato di questo processo.

A conferire alla macchina il mezzo per contribuire alla performance in maniera costruttiva è proprio la voce, con la sua forza seduttiva. Essa permette alla macchina di agire retroattivamente sulle azioni umane (nel caso analizzato, di Walshe) e diventare partner relazionale attivo e sociale nella performance. Questa condizione si rifà all'"effetto-voce" (nel senso dato da Mladen Dolar¹⁰), che persiste nella voce sintetica, acquisendo nuove sfumature perturbanti. Tale effetto è sempre frutto di relazioni difficilmente rappresentabili ed essenzialmente performative, così come lo è anche la specifica modalità degli algoritmi di trasformare i dati in informazione.

L'altro elemento fondamentale, per composizioni sperimentali come *ULTRACHUNK*, è l'applicazione dell'addestramento non supervisionato (o auto-apprendimento) per la preparazione della rete. Con questo metodo si può ridurre l'apporto umano nell'etichettamento dei dati di input, sebbene rimanga molto alto per quanto riguarda il calibro dei parametri della rete: «Unsupervised learning algorithms learn solely from unlabeled inputs x [...]. [They] typically produce generative models, which can produce realistic text, images, audio, and video, rather than simply predicting labels for such data» [25]. Così, non solo si riconosce una specifica agency macchinica, ma la si pone in primo piano, per esplorare le sue derive espressive ed

⁹ Ad esempio, come viene proposto da Dominic Pettman, attraverso il suo tentativo di decostruire la metafisica umanista della presenza tramite i concetti di: istante di ritornello, *punctum* aurale, voce della natura (o voce creaturale), *voix-là* e voce ecologica. Con questi termini l'autore struttura una concezione trasversale ed ontologica di voce, estesa a tutti gli esseri (o "molteplici altri" per usare l'espressione di Rosi Braidotti [4]): «We might call this the *vox mundi* – the voice of the world – [...] that cannot be reduced to a concept such as "communication" but nevertheless impacts and influences all those within hearing distance» [21].

¹⁰ Nel suo studio sulla voce Dolar tratta di alcuni degli effetti perturbanti e per altri versi affascinanti che portano con sé le macchine, una volta munite di voce. A creare un interesse particolare verso queste voci artificiali risiederebbe la loro capacità di ricreare il più umano degli effetti, quello di "interiorità". Riferendosi ad alcune macchine parlanti del XVII secolo, Dolar sostiene come: «la [loro] principale attrattiva stava nell'enigma di come qualcosa di così palesemente non-umano potesse produrre degli effetti umani» [9]. Questo "effetto-voce" viene creato non solo dalla causalità meccanica e senza senso, ma anche da una causalità che Dolar definisce "zoppicante", riprendendo il termine da Jacques Lacan: una causalità che rappresenta un eccesso sulla sua causa, creato dall'effetto-voce stesso. La medesima causalità può essere pertanto definita a sua volta "perturbata", in quanto presenta in sé un intoppo: «È come se l'effetto si emancipasse dalla sua origine meccanica e iniziasse a funzionare come un *surplus* – proprio come il fantasma nella macchina; come se fosse un effetto senza una causa propria, un effetto in eccesso sulla sua causa determinabile – e questa è una delle strane proprietà della voce». [9]

estetiche. L'obiettivo è proprio quello di sfruttare la componente meno prevedibile (e anche più distante dalla simulazione di comportamenti prettamente umani) delle reti neurali artificiali¹¹.

Lasciando un margine più ampio di agentività, l'IA non viene applicata per simulare musiche basate su regole preimpostate. È piuttosto l'algoritmo stesso che (entro certi limiti) costruisce una serie d'impostazioni da applicare per la generazione degli output¹². In questo modo, è possibile individuare una nuova estetica musicale, non più esclusivamente basata sul senso musicale umano, ma aperta al "senso musicale" dell'IA. Questo decentramento dell'impianto antropocentrico nell'arte permette l'esplorazione di questa particolare sfera della performatività postumana, in grado di comprendere le agency dei molteplici altri all'interno di una correlazione rizomatica continua e dinamica.

Il riconoscimento e l'analisi della performatività macchinica permette di concepire la voce più in qualità di oggetto (come inteso da Dolar), proprio grazie alla modalità dell'IA di ascoltare il suono della voce slegato dalla dimensione semantica, quindi non assimilandolo per forza di cose sempre a un gesto o a una citazione¹³. Così facendo, questo tipo d'indagine abilita a individuare un terzo livello analitico della voce, oltre quello semantico e vocalico, vale a dire l'impronta tecnologica: «Se il contenuto semantico del parlato corrisponde al significato e il contenuto acustico alla materia sonora, vi è un ulteriore livello di differenziazione analitica, che Ernst¹⁴ definisce sonico [...]. Con questo termine egli si riferisce alla mediazione tecno-culturale che porta un suono ad essere quello che è, ovvero all'"impronta" lasciata dagli strumenti utilizzati per registrarlo e riprodurlo, oppure per "sintetizzarlo"» [20].

¹¹ Oltre ad *ULTRACHUNK*, si possono citare altri progetti improntati su questo metodo. Per fare qualche esempio: *A Late Anthology of Early Music Vol. 1: Ancient to Renaissance* (del 2020, sempre di Walshe, questa volta in collaborazione con il duo Dadabots), *Text Score Dataset 1.0* (cominciato da Walshe, Ragnar Arni Ólafsson e David De Roure nel 2017), *IS IT COOL TO TRY HARD NOW?* (del 2017, di Walshe); *Tomomibot* (di Tomomi Adachi, Andreas Dzialocha e Marcello Lussana, del 2018) e *Destructured Bach* (di Oberlunar, del 2021).

¹² Per questo scopo vengono spesso utilizzate reti neurali ricorrenti (RNN). La peculiarità di queste reti risiede nel poter usufruire di una memoria nelle connessioni tra unità nascoste della rete, attraverso una serie di passi temporali: «Units may take as input a value computed from their own output at an earlier step in the computation. [...] This allows the RNN to have internal state, or memory: inputs received at earlier time steps affect the RNN's response to the current input» [25].

¹³ Il riferimento qui va a Luciano Berio, secondo cui la voce: «è sempre una citazione, è sempre gesto [...]: accende associazioni e porta sempre in sé un modello, naturale o culturale che sia» [2].

¹⁴ Wolfgang Ernst è uno dei principali esponenti della *German media theory*. Il suo approccio "radicale" all'archeologia dei media: «mette al centro l'archivio, inteso non solo come contenitore di documenti ma come dispositivo tecnico per la memorizzazione. Nella proposta teorico-epistemologica di Ernst [...] l'archeologia dei media non deve rivolgersi solo al "cosa" appartiene al passato, ma soprattutto al "come" il passato viene scritto, memorizzato e archiviato» [20].

3.2 Estensione del sé

In composizioni quali *ULTRACHUNK*, l'ibridazione umano-macchina viene rimarcata sia dall'elemento uditivo che da quello visivo. Si va così incontro a una performance che allinea, mette a contatto e mescola fra loro l'identità umana e l'agentività artificiale, per esplorarne i risvolti artistici, estetici e comunicativi. In queste (e altre) performance di live electronics, le voci attuano una profonda relazione d'ascolto reciproco, proponendo di volta in volta reazioni diverse a quanto sentito, attraverso emissioni sonore (di natura biologica o di origine elettronica).

ULTRACHUNK è un perfetto esempio nel quale la performer proietta elementi di sé verso l'esterno, diventando parte di un corpo esteso, un corpo-ambiente inserito in un rizoma d'interazioni continue. La stessa soggettività umana mette in mostra la propria "porosità", riscontrabile ancora più direttamente nelle tendenze odierne d'ibridazione biotecnologica. In tutto questo, la voce mantiene un legame con il corpo, attraverso le dinamiche interne di fonazione. Ma, al contempo, la voce clonata e trasferita nel dispositivo tecnologico complica il rapporto con il controllo della propria voce (e per estensione della propria identità¹⁵).

Trasferendo parte di sé all'esterno del proprio corpo, s'innescia un processo di estensione corporea, rappresentata dal trasferimento della voce e dalla sua clonazione su un dispositivo elettronico. A fianco alla voce incorporata, si definisce una "extracorporea" (per esempio la rete neurale artificiale in *ULTRACHUNK*, l'alterego di Walshe). Nell'incontro tra queste voci si individuano nuove tipologie performative, che abbracciano e accolgono le istanze postumaniste citate.

La voce, allora, si propone come mezzo privilegiato: da un lato, per il disincorporamento umano, sempre più trasferito in corpi esterni, digitali e protesi tecnologiche¹⁶; dall'altro per la formazione di un "corpo tecnologico", rappresentato dal suo corpo vocalico. Pertanto, la voce diventa uno strumento adatto per attuare l'ibridazione umano-macchina, esprimendosi specialmente attraverso le pratiche di live electronics, capaci di inglobare i nuovi mezzi tecnologici nel rapporto performativo con l'umano. In questo processo i concetti di "io" e "altro" vengono a fondersi fra loro, per cooperare ai risultati della performance [26].

¹⁵ La voce, infatti, si propone come un marcatore identitario molto personale, capace di rappresentare l'essenza di una persona che, grazie alla propria voce, si auto riconosce e diventa riconoscibile da altri [7].

¹⁶ Si ricorda, a tal proposito, l'aspirazione transumanista all'*uploading* mentale su supporto tecnologico esterno. Un versante tecnofilo che richiama i reconditi desideri umani di superamento della morte: «If I transfer my consciousness, and the consciousness of myself, either to someone else or to a hologram, I become an immortal. AI is a tiny facet of the historical journey towards achieving immortality through the process of artificialisation» [6].

4. CONCLUSIONI

L'avvicinamento ai principi fondamentali della teoria postumanista (che scardina l'ordine gerarchico vitruviano tradizionale, inserendo l'umano stesso all'interno dell'assemblaggio sociotecnico composto da agency umane e non-umane [4, 8, 11, 20]) ha permesso, di allargare il punto di vista e intraprendere una direzione non antropocentrica, così da considerare i processi operazionali dell'IA senza avere l'umano come modello a cui far riferimento. In questa cornice, la voce (intesa nella sua condizione sociotecnica [20]) si pone come mezzo privilegiato per l'interazione umano-macchina, per via del suo legame con l'identità, la personalità e la soggettività umana [7, 9]. Ecco che risulta utile l'uso di uno sguardo trasversale per intendere l'IA come fenomeno tecno-sociale e storico-culturale, di cui è possibile ricostruire e decostruire il nucleo e che co-esiste e co-opera alla realizzazione della realtà sociale in cui l'umano è inserito.

In questo modo, è stato possibile interpretare il ruolo dell'IA nella sfera creativa, artistica e musicale odierna. Essa va ad inserirsi nel complesso assemblaggio socio-tecnico che costituisce la creatività stessa, così come indagato dall'archeologia dei media di Ernst e Domenico Napolitano [20].

L'agency dell'IA risulta "aumentata", rispetto a quella degli strumenti tecnologici tradizionali: è un'agency interattiva e per certi versi autonoma e autoapprendente, che segna l'epocale passaggio dalla comprensione alla previsione. Per la prima volta, infatti, si possono scindere fra loro i concetti di: intelligenza, comprensione e intenzionalità. Inoltre, è proprio grazie a questo genere di tecnologie che si può parlare di: "*augmented intelligence*" o "*extended intelligence*" [24]; estensione dell'umano e "mente estesa" [19]; migrazione del pensiero umano che si fa disincarnato, nell'idea di "umanità aumentata" [3]. Declinazioni diverse per definire il desiderio umano di estendere la propria corporeità e le varie forme di trasposizione dell'umano nelle macchine, ovvero di oggettivazione della sua soggettività, così come era stato con la scrittura [3, 12].

Attingendo alla filosofia postumanista, l'idea di "umanità aumentata", lascia lo spazio per una riconfigurazione generale dei ruoli dell'IA e dell'umano in chiave ecologico-postantropocentrica e di co-relazione per quanto concerne la costruzione della soggettività e della realtà sociale, costitutivamente legate ad aspetti performativi e dinamici di attori umani e non. In questo modo è possibile considerare l'IA non come un'entità in contrapposizione alla specie umana, ma come oggetto tecnico avvolto da un "alone di socialità" [3]. Questo legame tra umano-macchina-ambiente sembra destinato a determinare la realtà futura, portando con sé tutta una serie di conseguenze per i tradizionali modi d'intendere la corporeità, l'individualità e la soggettività.

La velocità del progresso di questo versante tecnologico è enormemente più rapida della capacità di adatta-

mento ai cambiamenti delle società umane e questo è un altro degli elementi che differenzia l'IA dalle tecnologie che la precedono. Uno sforzo collettivo per una maggiore consapevolezza di come funziona l'IA e come conviene approcciarsi ad essa per il bene comune, diventa la strada più appropriata e insieme necessaria.

Il postumanesimo entra in questo quadro come metodologia per affrontare in maniera inclusiva, positiva ed ecologica l'enorme impatto dell'IA: questo approccio non si pone come alternativo o nemico della tecnologia, ma si dimostra capace di integrarla nello schema complessivo del mondo, ricalibrando i pesi degli agenti al suo interno in modo più inclusivo.

Il repertorio di Walshe è servito per inquadrare a livello pratico alcune delle possibilità d'esplorazione artistico-musicale con l'IA. Scegliendo come punto di riferimento la performance *ULTRACHUNK* si è potuto effettuare un confronto tra quanto esposto nei vari paragrafi e l'esperienza musicale diretta. L'IA rappresenta oggi un'immensa fucina creativa per gli artisti interessati ad indagarne i risvolti e gli utilizzi.

La corrente fase storica prevede l'introduzione dell'IA su larga scala in campo artistico. Questa operazione presenta alcuni aspetti di passaggio, di collegamento tra un prima e un dopo che caratterizzano il presente, forse distinguendolo dai futuri mutamenti estetici di quest'arte "postumana". Per adesso, infatti, l'arte generata da o con IA tende a distinguersi da quella umana: o per una esagerata "perfezione" dei suoi parametri (come si può notare nelle voci palesemente robotiche e prive di difetti), oppure mettendo in mostra errori "non-umani", bizzarri, persino perturbanti. Questa fotografia, però, potrebbe un domani trasformarsi, perseguendo la strada dell'esponenziale indistinguibilità tra prodotto umano e artificiale. Ecco perché la situazione attuale rappresenta una così rigogliosa epoca di potenziale sperimentale per l'arte e la musica: è ancora possibile esplorare la relazione tra la creatività umana e l'agency macchinica, senza il timore di confonderle. Finché la tecnologia palesa i propri effetti, questi possono rappresentare ricco e variegato materiale estetico ed espressivo da poter aggiungere e far interagire alle possibilità creative musicali e artistiche. Solamente concependo il rapporto che s'instaura tra umani e macchine in un ambiente comune (fisico, digitale o ibrido che sia), si possono comprendere le operazioni che stanno alla base di questo innovativo versante artistico. *ULTRACHUNK* rappresenta un emblema per la sperimentazione del tipo d'interazione che si può creare tra umano e macchina: un'improvvisazione vocale tra un performer umano col proprio alter-ego artificiale munito, quest'ultimo, di agentività specialmente attraverso la clonazione della voce del performer umano.

Die Zukunftsmusik (sia nel senso letterale che con l'accezione di cosa avveniristica) è alle porte e l'IA ne farà senz'altro parte: «Where do we go from here? AI. What is coming next? AI. What a time to be alive» [27].

5. RIFERIMENTI

- [1] C. Berberian, “Appendix: ‘La nuova vocalità nell’opera contemporanea’”, in P. Karantonis, F. Placanica, A. Sivuoja-Kauppalaa e P. Verstraete (a cura di): *Cathy Berberian. Pioneer of Contemporary Vocality*, Ashgate, Farnham-Burlington, 2014, p. 64.
- [2] L. Berio: *Un ricordo al futuro. Lezioni americane*, Einaudi, Torino 2006, p. 41.
- [3] R. Bodei: *Dominio e sottomissione. Schiavi, animali, macchine, Intelligenza Artificiale*, il Mulino, Bologna, 2023, pp. 298-301.
- [4] R. Braidotti: *Il postumano. La vita oltre l’individuo, oltre la specie, oltre la morte. Vol. 1*, DeriveApprodi, Roma, 2021.
- [5] A. Buffardi e D. de Kerckhove: *Il sapere digitale. Pensiero ipertestuale e conoscenza connettiva*, Liguori Editore, Napoli, 2011, p. 41.
- [6] M. Clancy, “The Future: Interview with Jacques Attali”, in Id. (a cura di): *Artificial Intelligence and Music Ecosystem*, Routledge, Abingdon-New York, 2000, p. 4.
- [7] S. Connor: *Dumbstruck. A Cultural History of Ventriloquism*, Oxford University Press, Oxford-New York, 2000, p. 3, 6.
- [8] A. Di Scipio, “In viva voce”, in Garda (a cura di): *La mediazione tecnologica della voce*, NeoClassica, Roma, 2023, pp. 155-190.
- [9] M. Dolar: *La voce del padrone, una teoria della voce tra arte, politica e psicoanalisi*, Orthodes, Napoli-Salerno, 2014, pp. 19-20.
- [10] M. Dyer, Z. Kanga e J. Walshe, “In Conversation with Jennifer Walshe: Performing with Intelligent Machines”, *Contemporary Music Review*, vol. 42, n. 3, novembre 2023, p. 392.
- [11] F. Ferrando: *Il Postumanesimo Filosofico e le sue Alterità*, ETS, Pisa, 2016.
- [12] L. Floridi e F. Cabitza: *Intelligenza artificiale. L’uso delle nuove macchine*, Bompiani, Firenze-Milano, 2021, pp. 24-25, 93-97, 134-138.
- [13] L. Floridi: *La quarta rivoluzione industriale. Come l’infosfera sta trasformando il mondo*, Raffaello Cortina Editore, Milano, 2004.
- [14] G. Fontana: *La poesia della voce e del gesto. Percorsi della vocalità nella poesia d’azione*, Editoriale Sometti, Mantova, 2004, p. 43-45.
- [15] M. Garda, “Arcipelago voce”, in L. Cardilli e S. L. Vallauri (a cura di): *L’arte orale. Poesia, musica, performance*, Accademia University Press, Torino, 2020, pp. 162-176, 180-184.
- [16] M. Garda, “How female is the voice? Conceptualisations and practices”, in S. Facci e M. Garda (a cura di): *The Female Voice in the Twentieth Century. Material, Symbolic and Aesthetic Dimensions*, Routledge, Abingdon-New York, 2021, p. 73.
- [17] M. Garda, “La traccia della voce tra poesia sonora e sperimentazione musicale”, in Id. e E. Rocconi (a cura di): *Registrare la performance. Testi, modelli, simulacri tra memoria e immaginazione*, Pavia University Press, Pavia, 2016, p. 83.
- [18] F. Gobet, P. C. R. Lane, S. Croker, P. C-H. Cheng, G. Jones, I. Oliver e J. M. Pine, “Chunking Mechanisms in Human Learning”, *Trends in Cognitive Sciences*, vol. 5, n. 6, giugno 2001, p. 236.
- [19] D. Montanaro, “Esplorando i confini dell’intelligenza”, in A. Ventura (a cura di): *Pensiero umano e intelligenza artificiale. Rischi, opportunità e trasformazioni sociali*, Le api, Roma, 2023, p. 99.
- [20] D. Napolitano: *La voce artificiale. Un’indagine media-archeologica sul computer parlante*, Editoriale Scientifica, Napoli, 2022, pp. 30-31, 78, 202, 204, 208.
- [21] D. Pettman: *Sonic Intimacy. Voice, Species, Technics (or, How to Listen to the World)*, Stanford University Press, Standford, 2017, p. 63.
- [22] I. Pustijanac, “Comporre con le soglie. Le dimensioni della voce nelle pratiche compositive di musica mista”, in M. Garda (a cura di): *La mediazione tecnologica della voce*, cit., p. 135.
- [23] V. Rizzardi, “‘Hear what I Feel’. Joan La Barbara, the 1970s and the ‘extended voice’”, in S. Facci e M. Garda (a cura di), *The Female Voice in the Twentieth Century*, cit., pp. 165, 170.
- [24] F. Rossi: *Il confine del futuro. Possiamo fidarci dell’intelligenza artificiale?*, Feltrinelli, Milano, 2019, pp. 34-35.
- [25] S. Russel e P. Norvig: *Artificial Intelligence. A Modern Approach. Fourth Edition*, Pearson, Harlow, 2022, pp. 823-827.
- [26] I. Sarno, “Io è/e un altro. Pratiche vocali ed elettronica dal vivo nel XXI secolo”, in M. Garda (a cura di): *La mediazione tecnologica della voce*, cit., p. 195.
- [27] J. Walshe, *Ghosts of the Hidden Layer. Talk Given at the Darmstädter Ferienkurse, 25.7.18*, Milker Corporation, <https://milker.org/ghosts-of-the-hidden-layer> (ultima consultazione: 19/04/2024).

Mediation Process in a Computer Music Interpretation: an Ecosystemic Approach

Marco Matteo Markidis

<https://markidis.space>

mm.markidis@gmail.com

ABSTRACT

This paper interprets *Audible Ecosystemics* no. 3a, a solo live electronics piece by Agostino Di Scipio. The project reproduces the digital signal processes outlined in the score, allowing the piece to be realized within a different audio synthesis environment than originally used. The work unfolds in two phases: (1) implementing the piece based solely on the composer’s schemes and scores, and (2) direct testing with the original audio synthesis environment. Ultimately, the project aims to perform this version in a concert and distribute the implemented libraries.

1. INTRODUCTION

At the beginning of the 2000s, Agostino Di Scipio composed a cycle of pieces called *Audible Ecosystemics*¹ for solo live electronics. Initially composed of three works², the cycle includes both concert versions and installations. The pieces were conceived as a study of the intrinsic sound material of an acoustic space, such as impulse response, Larsen tone, and background noise.

Considering the acoustic space of the concert hall as a container for the acoustic sources, Di Scipio conceived the three original pieces for solo live electronics, where the musical material is entirely generated within the same acoustic space without external sound sources³, such as fixed electronics or traditional performers. From this viewpoint, the pieces reveal their “ecological” nature, contrasting historically with a “colonialist” approach, where sound inside the acoustic space is generated by an external source to the concert hall.

Inspired by the works of American composer Alvin Lucier, Di Scipio creates audible systems capable of self-sustenance and possessing unique characteristics, properties, and dynamic laws, termed audible ecosystems. The environment where these systems operate is the source of

¹ Listenable on the attached CD to [1] and *Hörbare Ökosysteme: Live Elektronische Kompositionen 1993-2005*.

² In 2020, a fourth piece, *Silence Study*, was added to the cycle.

³ Only *Audible Ecosystemics* no. 1 requires active external intervention: synthetic or generated impulses are performed live.

Copyright: ©2024 Marco Matteo Markidis et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

the piece’s dynamic behaviour. The dynamic control signals are derived from environmental sound, and all processes are part of the overall feedback structure. Thus, in these systems, a structural coupling between electronics and the environment generates the sound material and its dynamics. These dynamics feed on themselves to continue the development and evolution of the piece, alternating phases of growth and interruption, resulting in self-regulating sound textures and musical forms.

The actual performance involves the active participation of an electronic interpreter, who primarily listens and, if necessary, acts as an integral part of the overall ecosystem process. The interpreter adapts to the situational set of structural conditions and acoustic and spatial aspects of the live performance. In this sense, the electronic interpreter is part of the environment, just like the audience present in the concert hall.

2. AUDIBLE (ECO)SUSTAINABILITY

The *Audible Ecosystemics* cycle was originally written for the Kyma audio synthesis workstation⁴. Over the years, in addition to the Kyma patch, the composer has also produced scores complete with graphic and textual notations for the performance of the pieces and the implemented electronics⁵. This allows performers to either use the Kyma patch provided by the composer or create their own version of the pieces independently⁶. Thus, the composer critically addresses one of the thorniest issues in electronic music: the sustainability of live electro-acoustic music [2]⁷.

By making the DSP score available under a Creative Commons license⁸, and in a high-level language not dependent on any specific implementation, the composer enables performers to create patches in an environment of their choosing. This approach is critical toward the sustainability of specific audio software and hardware, such as the external multiprocessor required by the Kyma audio synthesis environment [3]. Indeed, by providing a high-level DSP score, the composer abstracts the piece from its specific implementation network, albeit at the cost of an additional mediation layer. Consequently, given that the graphical transcription of digital signal processing (DSP)

⁴ <https://kyma.symbolicsound.com>

⁵ Currently not available for no. 1.

⁶ This independence applies both to the patch and to the audio synthesis workstation.

⁷ An Italian translated version of this article can be found in: *Musica/Tecnologia*, no. 13, 2019. In general, this volume discusses notating electronics and electroacoustic music.

⁸ BY–NC–ND, <http://www.creativecommons.it>

serves as the reference for the DSP itself, there is a genuine attempt to preserve the piece independently of the original hardware and software. This process creates a mediation layer between the original electronic creation, the graphical score abstracted from it, and its computer implementation in an audio synthesis environment based on the graphical score.

The mediation process is tripartite: although the composer’s idea for preserving the piece rests solely on the graphical score, this score should ideally result in a hyper-notated document. However, it appears the composer prefers not to explicit all the details, suggesting there may be times when a direct comparison with the original environment is necessary. Although, in some cases, the graphical score offers specific insights, some details still require in-depth analysis that only an impulse response or direct comparison with the original implementation can provide.

The primary risk is underestimating the musicality of the suggested implementation. Poorly developed details might lead to technically precise but musically inadequate implementations. This could result in correct technical execution that fails to preserve the dynamics and quality of the piece, compromising its musical essence and reducing it to a mere technical exercise.

3. PREVIOUS WORKS

Even though the author did not initially consult or review previous works related to the creation of their version of *Audible Ecosystemics*, other works were examined during the implementation process. Specifically, the author discussed with Giorgio Klauer his implementation of *Audible Ecosystemics* no. 2 and with Owen Green his implementation of *Audible Ecosystemics* no. 3. Giorgio Klauer directly provided the author with his Max/MSP patch, and discussions with Owen Green clarified how he analyzed certain details within the Kyma environment, particularly the creation of control signals and multi-threading handling.

Additionally, discussions have occasionally occurred on the Pure Data mailing list about porting the Kyma environment to Pure Data. However, no quoted articles, nor specific expertise or projects were found. Regarding the use of Pure Data for porting, some historical pieces for electronics and instruments have been realized [4], with a focus on score following for the realization of Pierre Boulez’s *Dialogue de l’Ombre Double*. In this context, the realization of Philippe Manoury’s *Pluton* is particularly intriguing, involving the import of a Max/FTS patch and considerations regarding the original use of 4X hardware. However, specific details on these processes are not provided.

4. FROM SCORE TO PD, THROUGH KYMA

Without delving into the specific choices the composer made in transposing the original patch into a graphical score, this section will analyze the interpretive and computational aspects that led to a new version of *Audible Ecosystemics* no. 3a [5]. After requesting and obtaining the score for this piece from the composer [6], and selecting a new audio

synthesis environment, the author focused on a straightforward mapping from the score⁹ to the new environment through a one-to-one translation of the various elements present in the graphical score. However, upon completing this phase, the listening test proved to be a failure. The initial results lacked dynamic range and variety, with all audio contributions stemming from the main loop (Figure 1). None of the gestures heard in the available recording were present¹⁰. Consequently, a discussion with the composer became necessary. The composer recommended starting with a more detailed analysis and sent two impulse responses of the control signal generation modules (Figure 2). The impulse responses were at a sampling rate of 44100 Hz, so the author decided to perform the tests at the same sampling frequency.

With the provided impulse responses, the author studied the impulse response (IR) of the Kyma integrator (Figure 3, left) and the more complex integrator chain, which includes also a feedback delay line and a $(1-x)^2$ mapping (Figure 3, right), where x represents the input sample. The resultant control signal, termed *InAmp0* (Figure 2, third signal from the top), became a focal point of the analysis. The audio synthesis environment chosen for the implementation was Pure Data (Pd) [7]¹¹.

4.1 Detailed Analysis of the *InAmp0* Chain

To analyze the *InAmp0* chain, the author first studied the terminology used for the signal processing functions in the score¹². Specifically, for the integrator, the score notes: “average of absolute values of input samples, over specific time windows (one may try RMS measures, instead)”. From the impulse response of the integrator, it was deduced that it operates as a filter with an infinite impulse response (IIR). This finding led to the immediate exclusion of finite impulse response (FIR) integrators, such as envelope followers (`[env~]`¹³) which are implemented as moving averages or exponential moving averages.

Referencing the *AmplitudeFollower* object in the Kyma manual¹⁴, the object is defined as “AmplitudeFollower follows the amplitude of the input, taking an average of the absolute values of individual input samples”, adding that “This is similar to the RMS but requires less computation”. Examining the reference for the *RMSSquared*¹⁵ object reveals the following finite difference equation:

$$output = input^2 * timeConst + prev * (1 - timeConst),$$

where *timeConst* is a constant value that controls the response time, *input* is the input sample, *output* is the output sample, and *prev* is the output sample at the previous time. Therefore, it can be inferred that the impulse

⁹ See Appendix A for details about the score.

¹⁰ Refer to the version on *Hörbare Ökosysteme: Live Elektronische Kompositionen 1993-2005*, RZ edition

¹¹ Current stable version 0.55.0, freely downloadable from <http://msp.ucsd.edu/software.html>.

¹² [6], page 5, paragraph *DSP unit*.

¹³ In Pd, control objects are noted as `[name]`, while audio objects are noted as `[name~]`.

¹⁴ Manual Body Kyma 4.5, page 221, paragraph *AmplitudeFollower*.

¹⁵ Manual Body Kyma 4.5, page 331, paragraph *RMSSquared*.

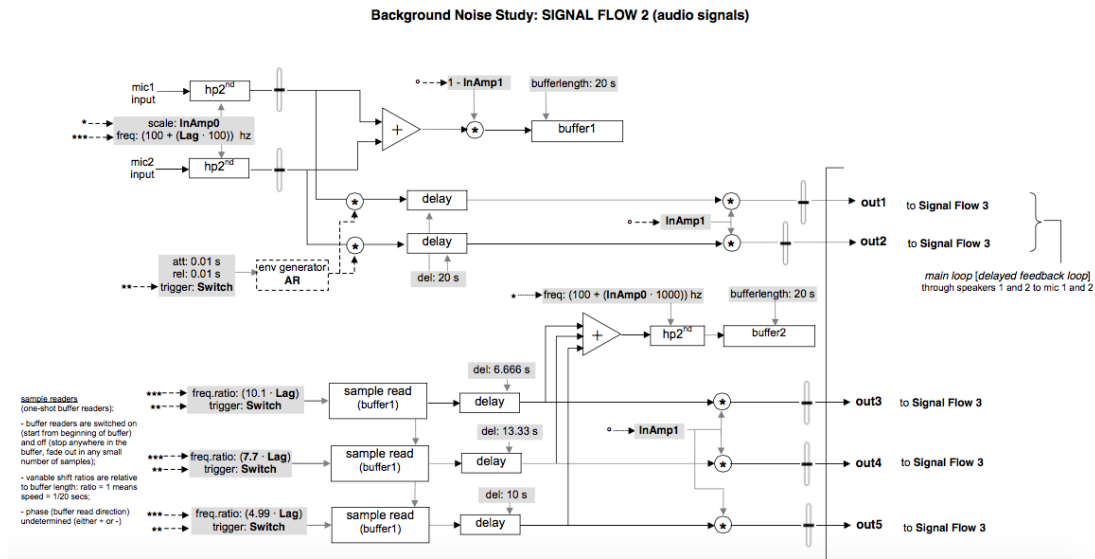


Figure 1. Representation of audio signals in *Audible Ecosystemics* no. 3a; showcasing the main cycle (top) and samplers (bottom). This bifurcated visual encapsulates the piece’s multi-layered audio structure, illustrating the core thematic interplay between cyclical patterns, sampled triggered audio fragments, and cascade buffer contribution. Adapted from [6].

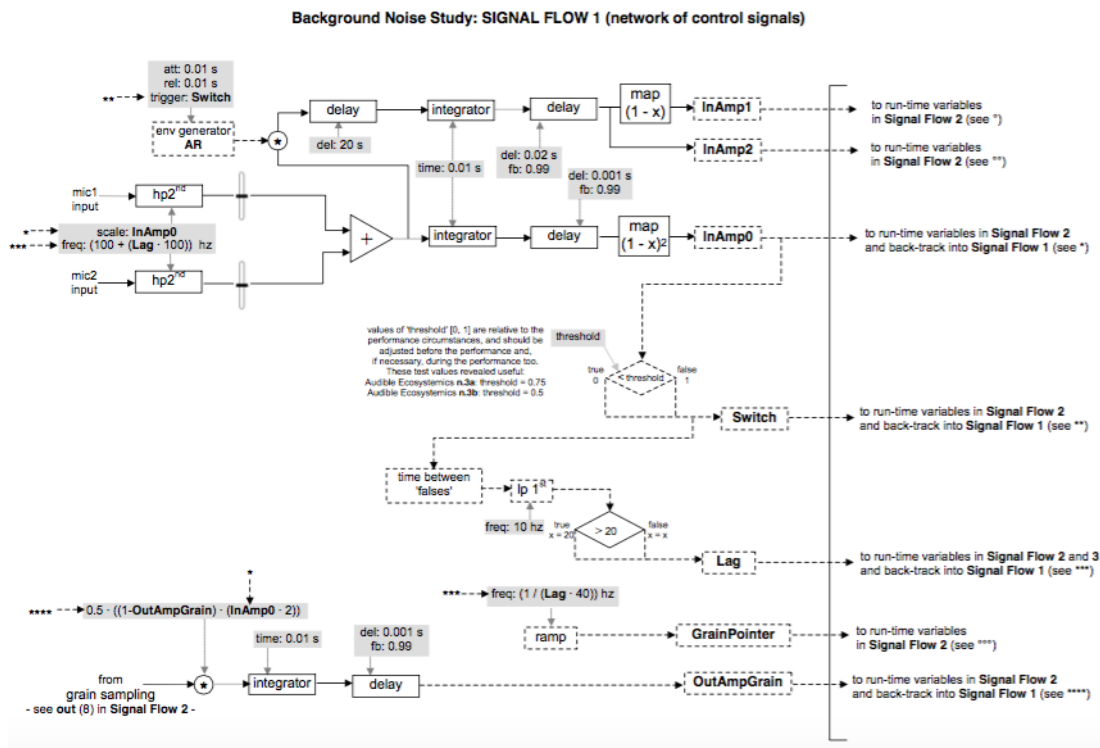


Figure 2. Representation of the network of control signals in *Audible Ecosystemics* no. 3a, showcasing the control signals derived from the microphones and their interactions (*Switch* and *InAmp0*, *Switch* and *Lag*), as well as the feedback loop within *InAmp0*. This visual illustrates the piece’s cascading control structure, highlighting principles of mutual dependencies and self-organization. Adapted from [6].

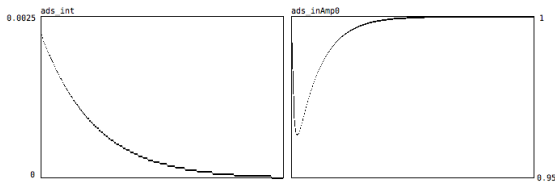


Figure 3. Impulse Responses Generated in the Kyma Audio Workstation. Left: Integrator applied to a 2000-sample table. Right: A complex chain combining an integrator, a feedback delay, and a $(1 - x)^2$ mapping function, controlling a 44100-sample table. In both instances, the Y-axis is appropriately scaled for visualization.

response for the desired integrator follows this equation:

$$output = |input| * timeConst + prev * (1 - timeConst),$$

where $|input|$ indicates the absolute value of the input sample.

This equation can be implemented in Pd in two ways: one by using an absolute value object (`[abs~]`) combined with a one-pole low-pass filter (`[lop~]`) (Figure 4, left), with a cut-off frequency proportional to the integration time reported in the score; and the other by using sample-time expression evaluation via the `[fexpr~]` object (Figure 4, right). Given the identical impulse response, the first option was chosen because the `[fexpr~]` object can be CPU expensive¹⁶ and in this case it requires a proper re-blocking strategy¹⁷.

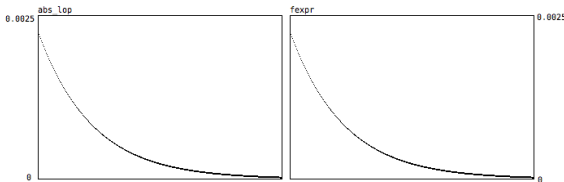


Figure 4. Impulse Responses Generated in the Pure Data Audio Environment. Left: Integrator combining `[abs~]` and `[lop~]` applied to a 2000-sample table. Right: Integrator using `[fexpr~]` applied to a 2000-sample table. In both instances, the Y-axis is appropriately scaled for comparison.

Next, the author proceeded to study the impulse response of a chain consisting of an integrator, a feedback delay, and a $(1 - x)^2$ mapping used as the control signal. For the delay module, only the feedback loop is employed without direct input. In the integrator chain, feedback delay, and $(1 - x)^2$ mapping, the feedback delay module requires a unitary blocksize, while the other two modules are independent of blocksize at the computation level (Figure 5).

¹⁶ See `[expr]` help patch, subpatch `pd [fexpr~] Examples`.

¹⁷ The `InAmp0` patch is computed at unitary blocksize, while the usage of `[fexpr~]` with the described syntax requires at least a blocksize of two samples.

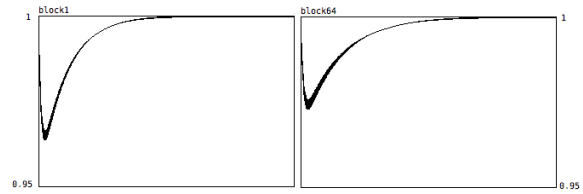


Figure 5. Comparative Analysis of Impulse Response by Blocksize Variation in a Feedback Delay Module within the Pure Data Audio Environment. Left: Detailed view of the impulse response from an integrator, feedback delay, and mapping chain, with the feedback delay calculated at a unitary blocksize. Right: Impulse response of the identical setup but with the feedback delay calculated at a blocksize of 64. Both scenarios are demonstrated on a table of length 44100 samples. The Y-axis is uniformly scaled for comparative analysis.

Finally, the author calculated the impulse response of `InAmp0`, which comprises a chain of an integrator, a feedback delay, and a $(1 - x)^2$ mapping used as the control signal. Figure 6 shows the resulting impulse response.

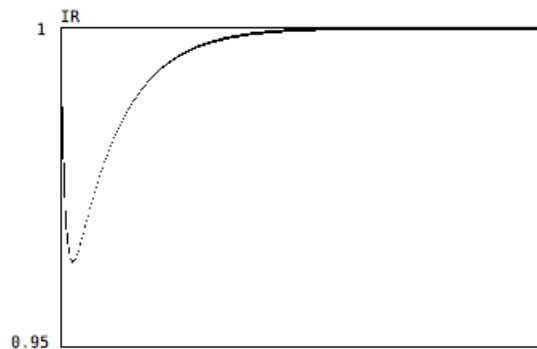


Figure 6. Final Impulse Response in the Pure Data Audio Environment, showcasing a chain composed of an integrator, feedback delay (computed at unitary blocksize), and the mapping function $(1 - x)^2$ as control signals. This response is generated over a 44100-sample table at a sampling frequency of 44100 Hz.

5. FROM KYMA TO PD, AND BACK

At this point, having developed impulse responses consistent with those provided by the composer, a new listening test was carried out. This time, the sonic result was much closer to expectations, and the system as a whole responded more appropriately. However, the system still lacked the robustness of the original; in terms of exploring the space of sound possibilities, this version appeared flatter and more static. This resulted in a greater uniformity of the explored sonic space and failed to generate new sound textures. These two aspects are inherently con-

nected. Given the design, the micro- and macro-musical forms are linked because the material generating both is the same. This relationship emerges in the programming through a dense interweaving of audio and control signals. Therefore, two different aspects had to be accounted for, delving even further into a lower level of comparison between the initial and final audio environments and studying the specificities of the individual environments.

5.1 Kyma

Several features of Kyma make it a complex environment to compare. Beyond the proprietary language, which restricts access to the source code, the use of external hardware with its unique architecture complicates the testing process further. Additionally, limited information is available on the language itself, making considerations such as scheduling and constructing the DSP¹⁸ graph challenging. Scheduling is particularly demanding as it occurs on a multi-processor architecture, where the compiled DSP graph must meet constraints such as memory usage, computation time, and correct ordering of processes within a directed acyclic graph¹⁹.

5.2 Control signals

In the Kyma environment, any signal within an audio generator, with frequency updated every 1000 Hz, can be used as control signal²⁰. In this sense, the structure of control signals is comparable to that of audio environments such as Csound and Supercollider [8], where the control signal is evaluated before the start of each control block, in which the audio samples are computed²¹. On the other hand, Pure Data has a sample rate audio stream, and control signals are managed by a double-precision scheduler that intertwines the computation of the control signal atomically at each audio block size.

Therefore, the strategy the author used, where possible, is to subsample with a sample and hold unit controlled by a ramp generator at a frequency of 1000 Hz. This approach reduces the effects of a non-unitary block size that emerge from the audio-to-control conversion and vice versa, and introduces a technique for generating control signals at low frequencies from audio signals. In this, the ramp generator phase is set to zero every time the DSP is restarted²².

By converting audio signals into control signals at regular intervals, like using Pd objects as a [meter~] connected to a [snapshot~], the process suffers from using a default block size of 64 samples for control signal generation. Assuming we want to subsample a 1000 Hz frequency signal with a given frequency of 44100 Hz, the

signal should be subsampled every 44 samples, whereas using a connected [meter~ 1] to one [snapshot~] would subsample every 64 samples, i.e., at every block size. This technique of switching from audio to control is, however, employed where a control signal is explicitly required by an object, such as for the graphical interface.

5.3 Internal feedback

In two separate sections of the score, an internal feedback mechanism is utilized. The first instance pertains to the creation of *InAmp0*, as illustrated in Figure 2. In this case, the control signal that manages *InAmp0* is scaled by *InAmp0* itself. A similar behavior is also present in the granulator, in the management of the control signal *OutAmpGrain*. This mechanism requires careful technical handling because it involves a feedback loop that cannot be directly implemented without creating a DSP loop²³. Therefore, a solution that breaks the audio DSP chain must be considered.

The chosen solution involves handling the feedback mechanism using a pair of [tabsend~] and [tabreceive~] objects. These objects continuously write a block of audio to an array and then receive it from the same array. By providing the same array to both objects, it is possible to break the audio DSP chain and, by properly managing the block size, ensure that the operational delay between them is set to one sample. This way, by writing the output of the desired signal, such as *InAmp0*, to [tabsend~] and then retrieving it with [tabreceive~], the signal is effectively feed-backed at the beginning of the chain where needed.

5.4 Granular synthesis

Analyzing the generation of audio signals in the score reveals the presence of “grain sampling,” or a granulator (Figure 7).

In the terminology notes, it states: “asynchronous granulation, i.e., read samples off buffer memory in successive short chunks of signal (‘grains’): allow for a kind of granular sampling that can be used for smooth time-stretching; should include ‘grain density’ controls and slight random variations on grain parameters (‘jitter’); no frequency shift requested”. The initial strategy for creating the granulator involved writing an asynchronous granulator, following a known implementation [9]. However, this approach did not produce the expected results. Apart from issues with sound quality, the implementation did not meet the requirements for this particular piece. It was unable to achieve the necessary grain density, both in terms of quantity and average density, or adequately displace the grains.

Therefore, a thorough reflection was undertaken regarding the necessary qualities of the granulator. The aim was not to replicate the behavior of the Kyma granulator exactly but to understand its contribution to the systemic dynamic. The granulator plays a specific role in contributing during certain sections of the piece. Thus, the focus

¹⁸ Audio signal processing graph.

¹⁹ The Kyma Language for Sound Design, Version 4.5. Manual Body page 79, paragraph *Scheduling on Multiple Processors*.

²⁰ The Kyma Language for Sound Design, Version 4.5. Manual Body, page 55, paragraph *Update Rate*.

²¹ <https://csoundjournal.com/issue10/CsoundRates.html>, paragraph *The Csound Processing Chain and Control Blocks*.

²² See *What's New in Kyma X.80 (6.80) – 11 April 2011*: For Sounds pasted into parameter fields as controls and then down-sampled (e.g. [aSound] L:10), the phase of the control is now set to zero on each replay. <http://www.symbolicsound.com/cgi-bin/bin/view/Know/KymaXDownloads>

²³ DSP loops in Pd are discussed at <http://pd-tutorial.com/english/ch04.html#id434824>.

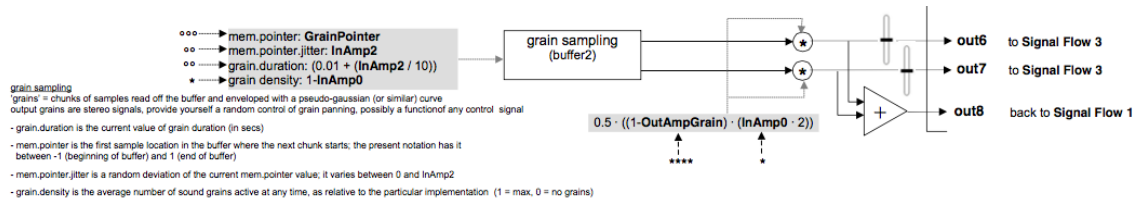


Figure 7. Detail of Granular Synthesis Parameters in the Score of *Audible Ecosystemics* no. 3a, highlighting the networks of audio signals. The main image displays the interdependent relations between granulator parameters and control signals, illustrating an auto-organization phenomenon. The bottom-left insert provides terminology notes related to the granulator. Adapted from [6].

was not primarily on the aesthetic quality of the granulator but on its effectiveness in meeting the systemic dynamic's needs. This involved considering the structural relationships and the network of interactions.

6. AELIB

In addition to the integrator chain, delay with feedback plus control mapping (Section 4.1), and the granulator (Section 5.4), the author developed other objects for the creation of the software. Notably, some of these, like a blocksize-time memory writer—which can be used in combination with one or more memory readers—are particularly useful in computer composition, specifically in ecosystemic composition types [1]. These abstractions have been collected into a library called *aeLib*, making them accessible independently of the main software. The library is distributed under an open-source license.

7. LIVE PERFORMANCE

In April 2024, the author presented the work at the Laboratorio ElettroAcustico Permanente (Rome, Italy)²⁴ during a workshop²⁵. At the end of the workshop, a public execution of the described version took place. Among the audience were the composer and the workshop participants. After the execution, a public listening guide and discussion were held. Both the author and the composer were satisfied with the listening experience, which was evaluated based on several criteria related to the nature of the piece²⁶. Audience feedback was positive. Interestingly, after the workshop, the participants were well-informed about the most important internal mechanisms of the piece, allowing them to proactively follow some events during the live performance.

During the workshop, the author explained step-by-step how the system should react to specific events, how it is supposed to internally organize certain gestures, and what the basic expectations are for particular events or circumstances. This detailed explanation facilitated active listening by the audience, leading to an engaging discussion at the end of the live execution.

²⁴ <https://l-e-a-p.github.io>

²⁵ <https://l-e-a-p.github.io/salti/2024-04-03-mmm-ae3/>

²⁶ See Section 9.

8. LAYERS OF MEDIATION

8.1 Motivation

After completing the technical aspects of the project, a reflective analysis allowed for the formulation of a comprehensive methodology termed the *layer of mediation*. This methodology emerged from practical experience and aims to provide a structured approach to problem-solving in complex translation tasks. The motivations for introducing this methodology are multifaceted:

- 1. Introduce a Coherent Methodology:** The goal is to establish a coherent methodology that can be applied to similar projects, ensuring consistency and repeatability in the approach.
- 2. Encourage Problem-Solving Through Layer Separation:** By clearly separating tasks into different layers, the methodology fosters a problem-solving approach that addresses specific challenges in a structured manner.
- 3. Handle Complex Tasks from the Score:** When complex tasks need to be implemented from the score, the methodology helps manage multi-layer interactions. For instance, a visual block on the score may correspond to several units in the chosen audio synthesis environment.
- 4. Enable Informed Decision-Making:** Once the translation tasks are clearly defined, decisions can be made more effectively within the target audio synthesis environment. This flexible approach surpasses the limitations of a one-to-one translation, particularly when the task's context is not fully understood.

8.2 Definition

To explain the methodology used in this project, the author introduces the concept of the *layer of mediation*. This term refers to a conceptual level where the original software's attributes—such as properties, functionalities, and performance characteristics—are translated, adapted, or reconciled with the capabilities, constraints, and paradigms of the target environment. This process involves several distinct layers, each dedicated to specific aspects of adapting software between two fundamentally different computing realms.

The term “layer of mediation” was chosen for several reasons:

- 1. Project Nature:** Unlike traditional porting [3], this project involves creating a unique version of an electronic piece based directly on the score, moving beyond a straightforward “same input — same output” logic.
- 2. Interpreter’s Role:** The interpreter plays a proactive role in this mediation process, interpreting the score and translating its requirements into a new, different computing environment. The goal is to create a new version, rather than simply an equivalent one.
- 3. Terminology Considerations:** Other terms like “adaptation layer” and “translation strategies” were considered but did not fully encompass the project’s scope.

To enhance understanding, we can review the core content of Section 4 and Section 5. The “layer of mediation” here involves identifying structural layers that can divide the project into different areas of interest and application.

8.3 Framework application

Within this conceptual framework, the mediation between computing environments is dissected into four distinct layers:

1. Conceptual, Design Pattern, and Paradigm Mediation

This layer addresses the foundational theories and methodologies driving the project’s approach, focusing on conceptual mediation aspects such as determinism and control signal generation, as explored in Section 5.2. While design pattern and paradigm mediation were considered, they were deemed too abstract for the scope of this work. By using the Pure Data audio environment, we accept the design patterns and paradigms already implemented in it, such as the dataflow paradigm.

- 2. Runtime Environment Mediation** This layer scrutinizes performance-related details, particularly aspects such as round-trip delay, which are vital for live performances. Although not the central focus of this article, these aspects are acknowledged for their importance. For instance, they play a crucial role in pieces based on audio feedback.

- 3. Hardware Mediation** Discussions pertinent to hardware, particularly the comparison between dedicated hardware and general-purpose CPUs and their scheduling differences, are covered in Section 5.1. This layer is critical for understanding the physical limitations and capabilities of each computing environment.

- 4. Syntax and Language Mediation** This layer dives into the translation of DSP score primitives, which are the foundational objects of our work. Key activities and examples are outlined in Section 4.1 and Section 5.4, showcasing the practical application of this mediation process.

While some aspects, like design pattern and paradigm mediation, were set aside due to their high level of abstraction, each of these layers plays a crucial role in bridging the original and target computing environments. This layered approach allows for a nuanced analysis, ensuring comprehensive coverage of the project’s technical and conceptual breadth.

8.4 Case Study: Multi-Layered Internal Feedback

An interesting observation can be made based on the analysis carried out in Section 5.3. In *Kyma*, feedback mechanisms are opaque to the user, which hides the complexity and necessary considerations like block size. Within the graphical user interface of an object, the field labeled “Scale” allows the user to insert the name of the object that should scale the output of the current object. The same principle applies in the score: the output of high-pass filters in Figure 2 is scaled by *InAmp0*, and the end of that chain is also named *InAmp0*.

Now, let’s focus on the proposed solution. By using a mediation layer methodology, this task can be viewed as multi-layered. The need for the audio graph to be acyclic is a form of *Conceptual Mediation* because it involves underlying principles and methodologies. The practical implementation necessitates knowledge of Pure Data primitives, making this a clear case of *Syntax and Language Mediation*. This contrasts with the approach outlined for control signal generation in Section 5.2. In that case, the combined use of a phasor and a sample-and-hold unit can be abstracted from the Pure Data context, which means it represents pure *Conceptual Mediation* without requiring specific expertise in Pure Data syntax.

9. SATISFACTION CRITERIA

When to stop is a natural question at this point. As mentioned in Section 7, satisfaction criteria based on listening are considered when analyzing the results of the new version. An important parameter is the variety that the version must demonstrate: several sonic patterns are structural to the composition, but exhibiting a variety of behaviors is a fundamental feature of *Audible Ecosystemics* no. 3a. For this reason, the interpreter can consider this characteristic while analyzing the work done.

Moreover, understanding an ecosystemic executive praxis is crucial for making fundamental evaluations, such as those discussed in Section 5.4. However, two consequences of the methodology described in Section 8 are that the version may contain systematic “inaudible” contributions and rare, not yet heard ones. To mediate the risks associated with these outcomes, the primary focus is on sound, which acts as the medium.

Only several live concerts in different environments — understood as the combination of the room, microphones, analog mixer, loudspeakers, audience, computer, audio drivers, and all other factors influencing the live execution — can confirm or, if necessary, draw the interpreter’s attention to needed corrections. These corrections can then mitigate unwanted scenarios.

Finally, it is important to underline that when we consider the written score as the reference for the piece, and not the electronics originally developed in Kyma, the quality of the version is determined by the audibility of the relationships indicated in the score. However, experience with the original patch and the Kyma audio workstation itself could act as a non-transparent medium, biasing some relationships.

10. NEXT STEPS

The author believes that further details warrant exploration, such as using audio signals to trigger grains in the granulator and testing different latencies in the electroacoustic chain to determine if any audible effects persist. Additionally, further concerts should expand the exploration of this version's possibilities and reveal whether additional improvements are needed to enhance its musical variety.

The author is also working on his own version of *Audible Ecosystemics* no. 2, the second piece in the cycle, applying the described methodology to a different work. If the framework provides satisfying results, it may be considered for other pieces, even those from various composers.

Furthermore, now that the implementation is complete, it becomes pertinent to compare this environment with the original one. Future work should focus on developing evaluation criteria and performing a comprehensive comparative analysis to identify potential advantages and assess the performance and suitability of each system.

11. CONCLUSION

The approach taken by the author facilitated addressing several major issues in contemporary electroacoustic music and computer music systems, such as the sustainability of electroacoustic works, scheduling differences, the construction of audio flow graphs, the generation of control signals, and the implementation of individual audio generators. This perspective enables the establishment of multiple starting points and outlines a strategy for retrospectively analyzing and mediating the risks associated with abstracting a DSP score from its native computing environment and adapting it to a different one. This interpretative view proposes a pathway to layer different computational aspects while focusing on a listening-based practice. The live execution held in April confirmed the success of the implementation: even though it does not validate the methodology in its entirety, it demonstrates at least a correct implementation of the network of relationships that is audibly perceptible.

12. ACKNOWLEDGMENT

The author would like to thank Giorgio Klauer and Owen Green for their valuable discussions about their projects; the Pd community for the continuous development of the software and the wealth of interesting details shared on the mailing list; Luca Spanedda for his expertise in filter design; Patrick Araujo for his support with hosting a GitLab server; Francesco Roberto Dani for his conversations about

granular synthesis and his review; LEAP and its members for their hospitality; and Agostino Di Scipio for this impressive piece of music.

13. REFERENCES

- [1] A. Di Scipio, *Polveri sonore. Una prospettiva ecosistemica della composizione*. (Roma, Italia): La Camera Verde, 2014.
- [2] N. Bernardini and A. Vidolin, "Sustainable live electroacoustic music," in *Proceedings Sound and Music Computing 2005 - XV CIM*, (Salerno, Italia), 2005.
- [3] S. Lemouton, "Computer music interpretation in practice," in *Proceedings, International Computer Music Conference*, (Utrecht, Netherlands), 2016.
- [4] M. Puckette, "New public-domain realizations of standard pieces for instruments and live electronics," in *Proceedings, International Computer Music Conference*, (San Francisco, USA), pp. 377–380, International Computer Music Association, 10 2001.
- [5] A. Di Scipio, "Listening to yourself through the otherself: On background noise study and other works," *Organised Sound*, vol. 16, no. 2, pp. 97–108, 2011.
- [6] A. Di Scipio, "Audible ecosystemics no. 3a," December 2016. Revision, Creative Commons License.
- [7] M. Puckette, "Pure Data," in *Proceedings, International Computer Music Conference*, (San Francisco, USA), pp. 224–227, International Computer Music Association, 1997.
- [8] O. Green, "Audible ecosystemics as artefactual assemblages: Thoughts on making and knowing prompted by practical investigation of Di Scipio's work," in *Contemporary Music Review*, 33(1), 2014.
- [9] F. Bianchi, *Inventare il suono con Pure Data*. Manuale introduttivo di musica elettronica, Torino, Italia, 2013.

A. SCORE

In this article, the score of *Audible Ecosystemics* no. 3a has been referenced multiple times. The version considered as the reference is the revision *December 2016*, which includes only version 3a. The score consists of nine pages, six of which are technical. These pages are organized as follows:

- Two pages of General Technical Setup
- Three pages of DSP score
- One page of Interpretative Criteria

In the article, the figures are taken from the DSP score pages; in Section 4.1, terminology is sourced from one of the General Technical Setup pages. The score is available under a Creative Commons license and can be requested directly from the assistant to the composer ²⁷.

²⁷ Instructions for obtaining the score are found on the composer's website: <http://www.agostinodiscipio.it/materials.html>.

SONIFICATION OF EDOARDO SANGUINETI'S POEMS

Barbara Grosso

Università di Torino

barbara.grosso@edu.unito.it

Andrea Valle

CIRMA - Università di Torino

andrea.valle@unito.it

ABSTRACT

Edoardo Sanguineti's poems are characterized by a very idiosyncratic style, which includes not only aural but also visual features, as poems show peculiar structures, sound repetitions, extra- and multi- alphabetical components, and multilingualism. The paper introduces a sonification system that allows the parallel display of these stylistic aspects. After identifying a corpus of 20 poems representative of the aforementioned features, a preliminary design step (which includes scenario analysis, task analysis and data analysis) has led to a sonification scheme in which each single character is associated with a sound, in particular an earcon. In the sonification algorithm, the text is processed linearly and values for sound parameters are retrieved for each character so to generate an earcon stream based on different synthesis techniques. We discuss implementation in SuperCollider and introduce a graphical interface, which enables playback management and parameter control, so that the features of each poem can be explored interactively.

1. INTRODUCTION

Sonification is the methodological framework that aims at converting data into sound to aurally communicate information to users [1] [2]. Sonification studies have proved that humans are able to identify patterns and structures within sound streams, so that auditory displays have been usefully proposed to convey information from various domains. In particular, time series and more generally linearly organized data have been successfully explored (e.g. heart-rate data [3]), as they can be aptly represented with sonification, given the temporal nature of sound. In the following study, we will focus on textual data, and in particular literary ones. As we will discuss, we will exploit the linear nature of literary text. Approaches to sonification dealing with literary texts have been proposed with different goals, which may include aesthetic enjoyment, analysis, creation of systems for automatic generation of music, communication of various textual elements. Sonification techniques may be based on various features, such as intonation and pronunciation [4], analysis of structures at different levels [5, 6], sentiment analysis [7, 8], lexicogrammatical properties [9].

Copyright: ©2024 Barbara Grosso et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

In this paper we present a sonification system specifically designed for Edoardo Sanguineti's poems. Our aim is not specifically musical, rather, while taking into account the musical results of data mapping, our focus is on data displaying via audio features.

Deeply linked to neo-avantgarde, Sanguineti (1930–2010) has been one of the most relevant Italian poets of the 20th Century. His poems show a complex intermingling of features, some properly related to language itself, others to various semiotic systems referring to written symbols (alternate writings, different alphabets, numeric elements) which can be appreciated by a reader rather than by a listener.

Our sonification is meant as an integrated approach capable of differentiating these different semiotic layers that constitutes structural elements in Sanguineti's approach to poetic writing. Starting from texts, we discuss four aspects characterizing Sanguineti's style that emerge from scenario analysis and, thus, led to the selection of a corpus of 20 poems: structures, sound repetitions, extra-alphabetical component relevance and multilingualism. Next, on the basis of task and data analysis outcomes, we define a set of earcons [2], one for each character. In the final section, we discuss the sonification algorithm, implemented in SuperCollider, as well as the requirements for a graphical interface to manage playback and manipulate some sound parameters. In the conclusions, we discuss effectiveness and limits of our sonification approach.

2. SCENARIO: SANGUINETI'S POEMS

Edoardo Sanguineti's career as a poet started in 1951, and in the next five decades his style changed depending on periods and occasions [10, 11, 12]. We performed a preliminary scenario analysis by taking into account the collected 821 poems [13, 14, 15]. We have identified four characteristic features of Sanguineti's poetic style, some more related to the visual domain, others to the aural one.

2.1 Structures

A first feature concerns structures: as an artist of neo-avantgarde, Sanguineti aimed at radically breaking with the traditional way of writing poetry, which was largely based on standardized metrics and specific stylistic choices (e.g. lexicon). To the purpose, he initially used very long free verses (often indented, so that a new line begins where the previous one ends), capable of including a variety of lexical choices. Later, Sanguineti has been able to reconsider traditional metrics: especially in occasional poems,

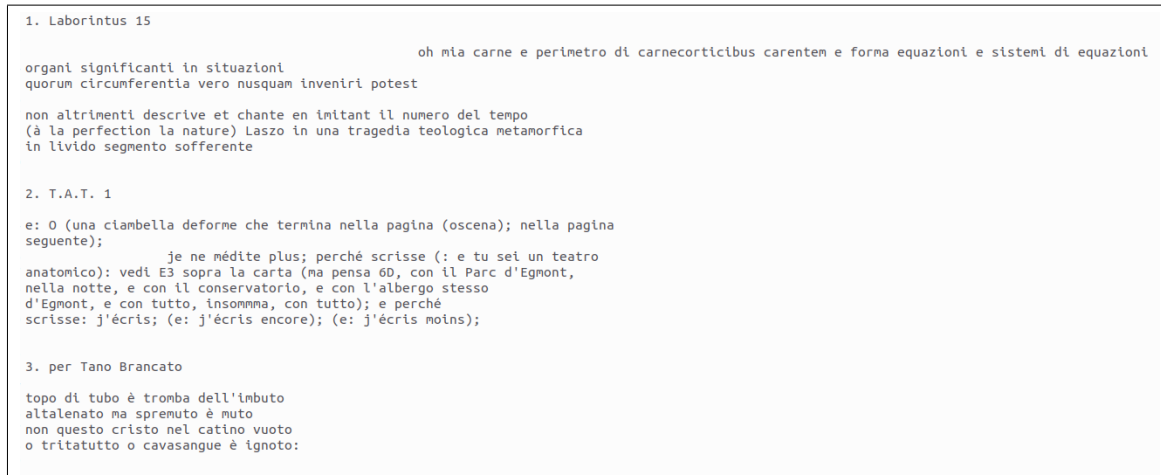


Figure 1. Three examples from Sanguineti's poems.

metrics has been used as a sort of reference to the past, while not always complying with traditional rules. Therefore, Sanguineti's poems may be written in a panoply of metrical forms, from free verse to e.g. hendecasyllable, including intermediate forms, e.g. following a newly invented metrical scheme.

2.2 Extra-Alphabetical Elements

The second feature is the relevance of extra-alphabetical elements, which include numbers, punctuation marks and other symbols. Sanguineti explained [12] that in his first poems he used no punctuation at all (or, as an alternative, a deliberately chaotic one). In a following phase of his work, he started to use punctuation in an apparently less arbitrary way: he employed almost only commas, parentheses, colons and, less often, question and exclamation marks. Yet, punctuation usage is still very far from standardized writing conventions. As an example, parentheses are used to add more layers to the discourse (and they may not be matched); a colon may be at the end of the poem meaning a non-conclusion (along with the lowercase character at the beginning of a sentence), as if that fragment was just an excerpt from a wider speech. Again, we can find poems with no extra-alphabetical characters (or only the essential ones, like apostrophes), other poems with a massive use of them, and some intermediate cases.

2.3 Sound Iteration

The third characteristic aspect is the repetition of sounds. Sanguineti paid an intense attention to the choice of words, often with a specific focus on their phonological features, while sometimes almost ignoring meaning: some very abstract poems make use of words almost freely associated by following a rhythmic criterium, so that incongruous images are evoked by means of phonological/prosodic relations. A side effect is that the written text is particularly hard to read. In general, when metric rules apply, Sanguineti is prompted to play with words in a more ex-

treme way (as himself declared¹). Occasional poems (e.g. homages to friends) are many times generated by a composition method focused on chaotic enumeration. This aspect is particularly relevant in some poems while absent in others, with a continuum of possibilities in between.

2.4 Multilingualism

A fourth feature is multilingualism. As Sanguineti clarified [12], at the beginning of his poetic enterprise he wanted to create a sort of negative poetic harmony. A way to sabotage poetic tradition was to mix languages: in particular, medieval Latin introduces more ambiguity effects, because of the distance and, at the same time, the similarity with Italian. In other cases, when the author writes poems linked to his travels, languages convey some kind of local flavor. Thus, among Sanguineti's poems, some contain words and/or sentences in one or more foreign languages, while others are exclusively written in Italian.

A general aspect in Sanguineti's works is the opposition between graphological and phonological dimensions. Phonological aspects depend exclusively on the linguistic dimension. Yet, Sanguineti uses a lot of abbreviations, acronyms, different ways to write the same information (e.g. "secondo" and "II", for "second"), words with the same spelling in different languages, and other similar ambiguities. Figure 1 includes three excerpts from *Laborintus 15*, *T.A.T. 1* and *per Tano Brancato*, meant to show how Sanguineti's writing techniques may vary.

From the complete collection, we selected 20 poems (listed in Table 1), meant to be representative of the previously discussed features. As we have observed before, features are distributed over a continuum, and they are not orthogonal (i.e. they are independent).

¹ See "Edoardo Sanguineti e la sua poesia", RavennaPoesia, <https://www.youtube.com/watch?v=mQwjy0K-dd8>

| POEMS | REGULAR STRUCTURES | EXTRA-ALPH. COMPONENT | SOUND ITERATION | MULTILING. |
|---|---------------------------|------------------------------|------------------------|-------------------|
| <i>Laborintus 15</i> | NO | NO | INTERMEDIATE | YES |
| <i>Laborintus 23</i> | NO | INTERMEDIATE | INTERMEDIATE | YES |
| <i>Erotopaegnia 4</i> | NO | INTERMEDIATE | INTERMEDIATE | NO |
| <i>Purgatorio de l'Inferno 1</i> | NO | INTERMEDIATE | INTERMEDIATE | INTERMEDIATE |
| <i>T.A.T. 1</i> | NO | YES | YES | INTERMEDIATE |
| <i>Reisebilder 23</i> | NO | INTERMEDIATE | INTERMEDIATE | YES |
| <i>Stracciafoglio 46</i> | NO | INTERMEDIATE | YES | NO |
| <i>Pièce en forme de cube</i> | NO | YES | NO | YES |
| <i>Rompilingua scioglitesta</i> | YES | NO | YES | NO |
| <i>Albedo</i> | YES | NO | YES | INTERMEDIATE |
| <i>per Tano Brancato</i> | YES | NO | YES | NO |
| <i>Alfabeto apocalittico - T</i> | YES | INTERMEDIATE | YES | NO |
| <i>Nespoleide 1</i> | INTERMEDIATE | YES | INTERMEDIATE | NO |
| <i>Opus 89</i> | NO | YES | INTERMEDIATE | YES |
| <i>Corollario 34</i> | NO | INTERMEDIATE | NO | YES |
| <i>La ginestra</i> | YES | NO | NO | NO |
| <i>Scrabble</i> | INTERMEDIATE | INTERMEDIATE | YES | NO |
| <i>Il suono del teatro (primo getto)</i> | YES | YES | YES | NO |
| <i>Omaggio a Dürer - M4</i> | YES | INTERMEDIATE | YES | NO |
| <i>Nove filastrocche per Luca - Girotondo</i> | YES | INTERMEDIATE | YES | NO |

Table 1. List of 20 selected poems: every text can be fully representative, not representative or partly representative for each of the four features of Sanguineti's production.

3. TASK AND DATA ANALYSIS

A basic design feature in our approach is to take into account the text as a sequence of characters. This approach allows to emphasize the variability of characters (such as punctuation marks, numbers, etc) which would be lost at the word level. It can also be observed that, due to the radical experimentalism of the author, the word level in itself is often complex to grasp.

According to TaDa approach [1], the next phase, preparatory for sound design, consists in Task analysis and Data analysis.

3.1 Task Analysis

The purpose of our sonification is the parallel aural display of the four stylistic features that we have discussed before. As the sonification is based on single characters, each of them should be associated with a sound event, to be collected into a stream. The user, thus, should be able to explore data by manipulating some sound parameters to understand what is the best configuration to highlight textual properties.

In relation to structures, essentially the sonification has to detect indented lines and partitions into verses and stanzas. In this case, horizontal tab and line feed are the relevant characters. The extra-alphabetical component is recognizable at the single character level. Repetitions should be directly heard from reoccurring characters that make up words, more or less distant into the text, and from pattern identification (i.e. subsets of elements). For what concerns multilingualism, the presence of special characters like ç, œ or ß, as well as accents within words or simply characters not belonging to Italian alphabet (j, k, w, x, y) has to be emphasized. More generally, languages are different because of characters absolute frequency (both single and in groups), so it should be possible to identify foreign words even from particular recurring character sequences.

3.2 Data Analysis

Data can be grouped into three types: alphabetical, extra-alphabetical and blank, and can be distinguished in ordinal and nominal data [1].

The alphabetical ones are Latin and Greek characters, with or without diacritical marks². In case of Latin characters, there are both uppercase and lowercase characters, plus ç, œ and ß (lowercase only); uppercase Greek characters are absent. These data can be thought as ordinal data as they are ordered (historically) by the alphabets (see [1]). Diacritical marks are nominal data. Numbers and other signs (punctuation marks, mathematical symbols, others) are extra-alphabetical characters: these are nominal data, while numbers are ordered from 0 to 9 with fixed increments. Blank characters are those corresponding to some kinds of cursor displacement; therefore, they are nominal data.

² Considered diacritics are grave accent, acute accent, circumflex, umlaut and tilde, for Latin characters; for the Greek ones, they are grave accent, acute accent, circumflex and acute accent with smooth breathing.

4. SOUND DESIGN

We decided to associate each character with an earcon, i.e. a specific sound event generated by defining a set of parameters. Earcons are synthesized from two sound sources and employ different techniques: periodic signal and additive synthesis for alphabetical characters; aperiodic signal and subtractive synthesis for all other non-alphabetical ones.

For periodic signals, we distinguish between Latin and Greek characters. We use sine waves for the former and saw waves for the latter, both amplitude-enveloped. Uppercase Latin characters are rendered by means of adding two quasi-harmonic frequencies to the fundamental of the relative lowercase ones. Fundamental frequencies for periodic signals are computed by taking into account MIDI notes on a major scale (starting from MIDI number 60 –middle C– for Latin alphabet and from 48 –one octave below– for Greek). Pitches are associated with characters following the alphabetical order as an index for scale steps. This mapping is in some sense traditionally addressing some fundamental aspects of Western music, thus ensuring a sort of basic “musicality” in the sonification scheme. We considered this aspect useful for addressing the recognizability of character mapping, while not being explicitly designed in relation to music composition. Special characters œ and ß are treated as two characters in one, so there are two sounds in the duration usually allocated for one. In the œ case, we also use a glissando between the o and e frequencies. Characters with diacritical marks are defined by adding a kind of final acciaccatura (depending on the diacritical mark itself) to the pitch associated with the basic character.

Sounds for all other non-alphabetical characters are synthesized filtering a pink noise. Numeric characters are sonified by taking into account the major scale, from MIDI note 48, and by taking one pitch every two steps. The sonification uses bandpass filtering to obtain a pitched, complex sound. Like uppercases, we have a larger harmonic content. This creates a link between uppercase characters and Arabic numerals, because Roman numerals are (also) uppercase characters. Note that we cannot clearly know without context if a capital character is to be meant literally or as a Roman number, e.g. “I” can be Italian plural article “I”, English pronoun or the Roman number for 1.

The other extra-alphabetical characters can be considered in relation to each other, or to other characters with shared elements. Their sound is shorter and more impulsive. In order to differentiate punctuation marks, we have devised a visual metaphor. Period is the basic character, made up of a single dot. If we consider the squared space in which every character can be inscribed in writing, period is placed at the bottom, so we decided that its pitch is low (MIDI note 48). This frequency controls a bank of resonant harmonically-related high-pass filters that ring modulate an impulsive signal. By following the same metaphor, comma means a shorter pause and it is located at the bottom too in the writing space, but its shape is curve: consequently, sound is shorter, pitch is similarly low (50), and a “curve” effect is rendered through a kind of acciaccatura. We define the other marks by combining and/or extending

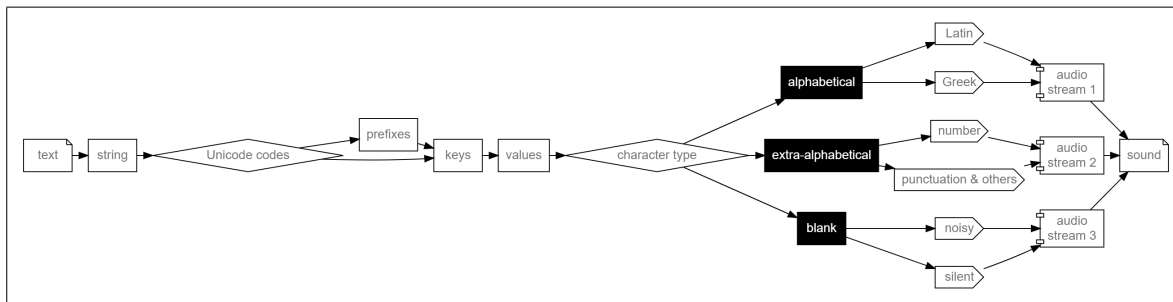


Figure 2. Poem sonification diagram.

the previous ones (period and comma). Colon is a double dot (two sounds in one time), as well as ellipsis is made up of three dots (three sounds in one time) and semicolon is composed of a dot and a comma (two sounds in one time, plus frequency changing and acciaccatura). Apostrophe is like a comma but at the top, so its pitch is higher (96). Quotation mark can be thought essentially as a double apostrophe. Parentheses are “curve” as the comma, but they fill all the graphical vertical space: thus, frequency increases from note 50 to 62 for left parenthesis, while decreasing from 62 to 50 for right parenthesis. Exclamation mark and question mark are made-up of two signs, one is a dot and the other fills the vertical space: frequency starts from the period note (48) and stops at a higher one (72 for the first, 60 for the second). An acciaccatura is added for question mark. Hyphen-minus can be seen as a long dot in the middle of vertical space, so its pitch is intermediate (60), and its duration is longer than period sonification. Without entering into details, some other indications follow. Equal sign is substantially a double hyphen-minus. Multiplication sign is written like a lowercase x, therefore, its duration is the same as other mathematical symbols and its pitch is the same as character x (MIDI note: 100). Plus sign pitch is intermediate among the hyphen-minus and the multiplication sign (80). Slash is a single sign that, without curve, fills space from the bottom to the top; it is also used to represent division, so its frequency starts from hyphen-minus note (60) and stops at the multiplication sign one (100). Percent sign is composed of three symbols (like ellipsis) and fills all vertical space, so its frequency grows from hyphen-minus note (60) to 84. Ampersand is often used as conjunction instead of character e: this case is similar to the multiplication sign one (thus, 67), and so is the degree sign one (“1°” with the character o, 84), too. Furthermore, these extra-alphabetical characters have the same number of harmonic bands as numeric characters, but the pink noise is filtered using a different filter, a bank of resonant high-pass ones. Blank characters sounds are generated by impulse modulating a single resonant high-pass filter. First, space character is left silent, so that word items can be separated, creating a specific rhythmic organization in the stream based on word-like elements. A new-line implies a sequence of carriage return and line feed, thus the former is silent to make the latter stand out; cut frequency associated with line feed is MIDI note 50 and

two sounds are generated to compose the earcon. Horizontal tab sound is characterized by a “bouncing” effect to recall cursor steps forward, so the number of tabulations is rhythmically mirrored by earcon repetitions, with frequency associated with MIDI note 110.

5. SONIFICATION ALGORITHM

The previously introduced sonification scheme is implemented in real-time in SuperCollider following the information flow shown in Figure 2.

First, when the user selects a poem from the menu on the interface (Figure 3), its path is passed to a function that reads the text (previously saved in TXT format) and returns a string. Then, when the user presses the PLAY button, a cycle begins: the string is read character by character, each of which is transformed into the relative Unicode code.

SuperCollider can only process ASCII characters, so the Unicode ones must be converted in their respective numeric codes: characters corresponding to valid ASCII values become positive numbers, while the others become negative numbers. These numbers are grouped through prefixes. ASCII characters are represented only by a single positive number. Unicode ones are represented by two or three negative numbers. To define sound parameters for each character, a data structure made up of dictionaries allows recursive selection in relation to characters by means of the first numbers used as prefixes. The selection process ends in keys, which are mapped onto sound parameters (values in Figure 2). In case of ASCII characters, the key can be immediately accessed (no prefix needed).

Earcons are organized in three internal streams, one for each character type (alphabetical, extra-alphabetical, blank, see black labels in Figure 2). For each type, two synthesis processes are available (including silence generation), for a total of six audio synthesizers. Switching among these three audio streams is controlled by the character type. This operation leads to the composition of the single final output stream (sound).

To highlight poem features through sound, we have developed a GUI for interactive parameter control (see Figure 3), which allows, after text choice, general playback management (play, stop, pause/resume), speed control, and panning and volume control for each stream. Additionally, the GUI offers a so-called “multilingual” mode (see MULTI button) in order to emphasize charac-

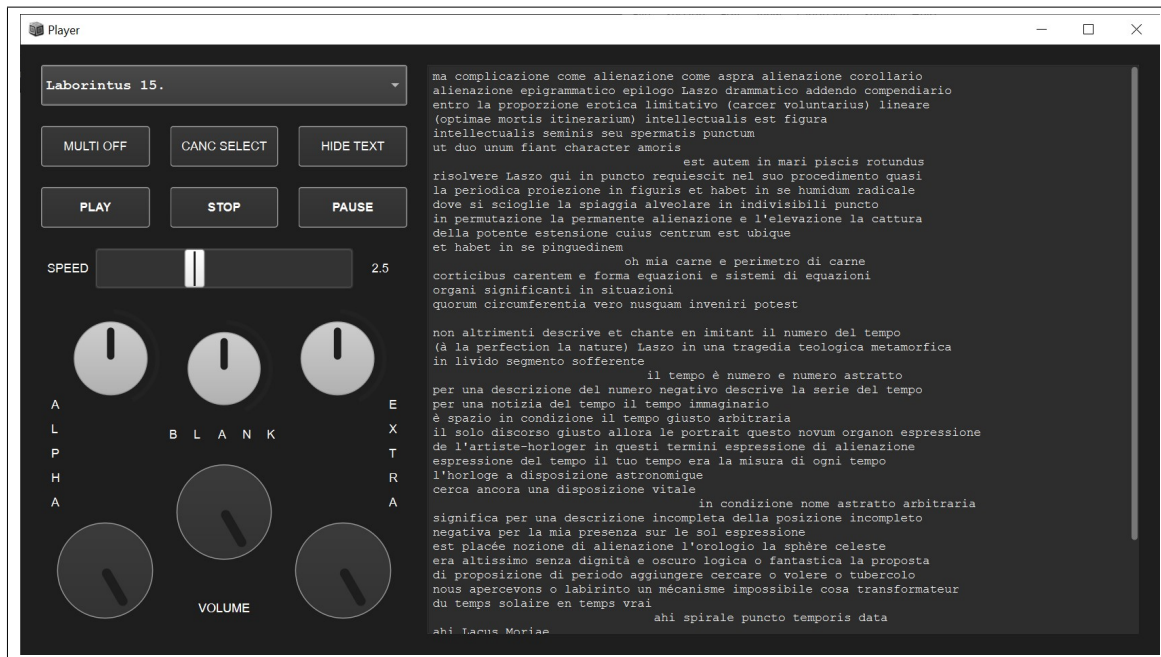


Figure 3. Graphical User Interface.

ters that are less frequently employed in Italian: this mode changes character order from the alphabetical to an order that depends on their absolute frequency in the Italian language³. In this way the association between characters and MIDI notes changes too, with higher pitches representing less frequent characters. It is also possible (via code) to change the basic musical scale. As an example, a hexatonic scale makes infrequent characters sound even more high-pitched, thus indicating the likely presence of foreign words. Finally, the chosen text can be shown (or hidden, HIDE TEXT); users can select characters or parts of the poem to listen to them in detail, and then cancel the selection by pressing a dedicated button (CANC SELECT).

6. CONCLUSION

In this paper we presented a sonification system designed for Edoardo Sanguineti's poems that allows the parallel aural display of four textual features, some of which mainly concerning hearing while others sight, especially in reference to structures, sound repetitions, extra-alphabetical component relevance and multilingualism. We have not yet performed an evaluation test with users. The following considerations are thus empirical and preliminary.

First of all, additive and subtractive synthesis techniques allow perceptual links among character types. While we have not addressed specifically the music composition aspects (rather focusing on aural data display), the mapping scheme, being largely based on Western scales, offer a basic musicality, resulting in a continuous diatonic

³ The new sequence is: e, i, a, o, n, r, l, t, s, d, c, u, p, m, g, v, h, f, b, q, z, j, y, x, k, w. See https://en.wikipedia.org/wiki/character_frequency

melodic/harmonic layering.

In relation to structures, we suggest that they can be generally perceived in sonification. It can be noted that there is a feedback effect between playback speed and sound design accuracy. If playback is sped up (to concentrate more characters over time) the discriminating sound design gets partly lost, even if, however, most characters remain identifiable. On the other hand, there is a clearer perception of the overall data shape. A possible solution could be to differentiate the three sound streams playback speed.

Sound repetitions are audible despite the high number of signals chained in the sonification stream. Some character combinations within words turn out to be particularly relevant in sound, and thus clearly recognizable. Poems composed with evident care to this aspect show a clearer "musical" structure.

The extra-alphabetical component can be easily distinguished, in particular if the user exploits channel separation, i.e. mapping alphabetical and non-alphabetical streams to L/R in the stereo front.

The multilingualism feature, instead, remains partially masked. Excluding special characters and the ones with diacritics, multilingual mode can highlight at most characters not included in the Italian alphabet but this case is not clearly distinguishable from Italian less frequent characters. Indeed, multilingualism cannot be solved merely through single characters succession.

As a final note, we suspect that a certain training is required. To catch all textual features as a whole, it is necessary to listen to sonification carefully many times.

The SuperCollider code and some audio examples are available here:

<https://github.com/vanderaalle/sanguinetiSonification>

7. REFERENCES

- [1] S. Barrass, *Auditory Information Design*. PhD thesis, The Australian National University and Computer Science, 1997.
- [2] T. Hermann, A. Hunt, and J. Neuhoff, eds., *The Sonification Handbook*. Logos Publishing House, 2011.
- [3] M. Ballora, B. Pennycook, P. C. Ivanov, L. Glass, and A. L. Goldberger, “Heart rate sonification: A new approach to medical diagnosis,” *Leonardo*, vol. 37, no. 1, pp. 41–46, 2004.
- [4] C. Huang, H. Lu, and J. Ren, “Algorithmic approach to sonification of classical Chinese poetry,” *Multimedia Tools and Applications*, vol. 61, pp. 489–518, 2012.
- [5] A. Coelho, P. Martins, and A. Cardoso, “A musical sonification of the Portuguese epopee,” in *Proceedings of the 4th International Workshop on Musical Metacreation*, jun 2016.
- [6] I. Emsley and D. D. Roure, ““It will discourse most eloquent music”: Sonifying variants of Hamlet,” *Journal of the Text Encoding Initiative*, no. 10, 2017.
- [7] H. Davis and S. Mohammad, “Generating music from literature,” in *Proceedings of the 3rd Workshop on Computational Linguistics for Literature (CLfL)*, pp. 1–10, apr 2014.
- [8] C. Stere and S. Trausan-Matu, “Poem sonification: from words to notes,” in *Proceedings of the 16th Conference on Human-Computer Interaction - RoCHI*, pp. 11–14, oct 2019.
- [9] E. Morales, K. James, R. Horst, Y. Takeda, and E. Yung, “The sound of our words: Singling, a textual sonification software,” in *Sounding Board: Proceedings of the 26th Annual International Conference on Auditory Display*, pp. 164–168, jun 2021.
- [10] A. Petrella, “Edoardo Sanguineti,” *Belfagor*, vol. 60, no. 5, pp. 543–556, 2005.
- [11] E. Sanguineti, “La poesia italiana del secondo Novecento,” *Comunicare Letterature Lingue*, vol. 1, p. 27–58, jan 2001.
- [12] E. Sanguineti, “La mia poesia,” *Comunicare Letterature Lingue*, vol. 1, pp. 59–90, jan 2001.
- [13] E. Sanguineti, *Segnalibro: Poesie 1951-1981*. Feltrinelli Editore, 2021.
- [14] E. Sanguineti, *Il gatto lupesco: Poesie 1982-2001*. Feltrinelli Editore, 2021.
- [15] E. Sanguineti, *Varie ed eventuali: Poesie 1995-2010*. Feltrinelli Editore, 2010.

Software di supporto all'esecuzione musicale in MiraWeb: due casi di studio

Giovanni Sparano

Conservatorio di Musica “B. Marcello”, Venezia

giovannisparano@gmail.com

ABSTRACT

L'articolo descrive l'implementazione e il funzionamento di strumenti software di supporto all'esecuzione musicale, realizzati in Max/MSP con l'ausilio di MiraWeb e dotati di interfacce grafiche dinamiche per gli esecutori, aggiornate in tempo reale, che vengono distribuite in rete grazie a un sistema client-server basato sulla tecnologia *websocket*.

Nei due casi di studio presentati, si illustrano gli adattamenti digitali delle partiture e la creazione delle indicazioni esecutive di *Variations II* di John Cage e *Frame* dell'autore, entrambi per organico variabile. I software sono *open source* e vengono distribuiti gratuitamente, su richiesta all'autore, con licenza CC BY-NC-SA 4.0.

1. INTRODUZIONE

Facendo seguito agli “sforzi per rendere più semplice eseguire alcune tra i più noti brani del repertorio della musica per live electronics, senza ricorrere a [...] software proprietario”¹ [1], l'autore ha realizzato l'adattamento digitale della partitura del brano *Variations II* di John Cage [2].

Scritto nel 1961, e dedicato come regalo di compleanno a David Tudor, “rappresenta il più alto grado di astrazione di un modello notazionale e compositivo che Cage abbia sviluppato nel periodo tra il 1958 e il 1961”², e uno tra i primi esempi di esecuzione al live electronics di Tudor, che concepisce il pianoforte amplificato, utilizzato per la sua realizzazione del 1961 del brano di Cage, come uno strumento elettronico [3].

Nell'implementare l'adattamento digitale, l'autore ha utilizzato MiraWeb³, un *package* aggiuntivo gratuito per Max, sviluppato dalla stessa Cycling '74⁴, reperibile nel *Package Manager*⁵, che permette di controllare da remoto alcuni controlli grafici di Max grazie alla tecnologia *websocket*. Inserendo un oggetto *mira.frame* all'interno di una *patch* viene definita un'area all'interno della quale disporre alcuni oggetti di Max, compatibili con Mira [4], e viene

¹ “We describe here our efforts to make it easier to perform certain well-known pieces from the live electronic music repertory, without resorting to [...] proprietary software”

² “[Variations II] represents the greatest degree of abstraction of a compositional and notational model that Cage developed over the period from 1958 to 1961”

³ <https://github.com/Cycling74/miraweb>

⁴ <https://cycling74.com>

⁵ https://docs.cycling74.com/max8/vignettes/package_manager

Copyright: ©2024 Giovanni Sparano et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

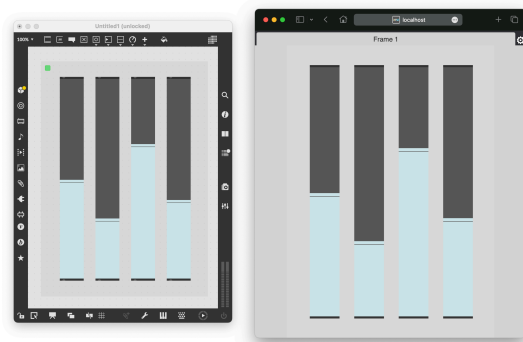


Figura 1. Mirroring di MiraWeb.

attivato un server Xebra⁶ che crea la connessione con i client, fornendo loro il *mirroring* della *patch* di Max (vedi Figura 1).

Due o più oggetti *mira.frame* all'interno di una stessa *patch*, si risolvono in altrettante distinte interfacce a *tab* nella pagina web generata da Xebra e servita ai client. MiraWeb [1] è utilizzabile con qualsiasi browser compatibile con la tecnologia *websocket*, su reti sia locali [5] sia geografiche. La comunicazione tra il browser e la *patch* remota è bilaterale, quindi è possibile controllare una *patch* che risiede sul server agendo sulle interfacce grafiche dei client, e/o inoltrare verso esse dati provenienti dal server.

Le due implementazioni di seguito illustrate sfruttano esclusivamente la comunicazione dal server verso i client, non essendo necessaria un'interazione da parte dei client.

2. VARIATIONS II E FRAME

Vengono presentati due casi di studio che illustrano gli adattamenti digitali delle partiture e la creazione delle indicazioni esecutive dei brani *Variations II* [2] e *Frame* dell'autore, composto nel 2022. Entrambi i brani sono concepiti per un organico variabile, sia di strumenti acustici, sia di strumenti elettroacustici. I software realizzati sono entrambi *open source* e vengono distribuiti gratuitamente, su richiesta all'autore, con licenza CC BY-NC-SA 4.0⁷.

2.1 Variations II

Una prima versione del software per la generazione della partitura e il controllo dell'esecuzione del brano di Cage è

⁶ <https://cycling74.github.io/xebra.js/>

⁷ <https://creativecommons.org/licenses/by-nc-sa/4.0/>

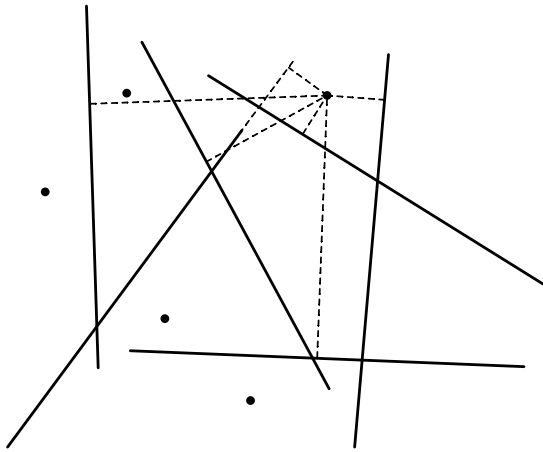


Figura 2. Esempio di sovrapposizione (punti e linee continue) e misurazione (linee tratteggiate).

stata utilizzata nell’ambito del Living Lab Music 8, evento organizzato dal Laboratorio SaMPL⁸ del Conservatorio di Musica “C. Pollini” di Padova, in collaborazione con l’associazione culturale Pase⁹, e svoltosi presso Pase Platform a Venezia¹⁰.

Variations II è la seconda di otto *Variations*, una serie di composizioni scritte tra il 1958 e il 1967 caratterizzate da procedure aleatorie per la determinazione dei parametri performativi¹¹, e della struttura formale [6].

La partitura di *Variations II* è composta da una serie di istruzioni verbali, che descrivono l’algoritmo generativo della performance, e da 11 fogli trasparenti: in 6 di questi è riportata per ognuno una linea, in 5 un punto. Secondo l’algoritmo formalizzato da Cage, questi fogli trasparenti devono essere sovrapposti casualmente a formare una grafico, come illustrato in Figura 2.

Le misurazioni delle perpendicolari, eseguite secondo una “qualsiasi regola”, tra ogni punto e ogni linea o “dove necessario, l’estensione di una linea”, restituiscono 30 determinazioni, 5 per ognuno dei 6 diversi parametri: frequenza, ampiezza, timbro, durata, punto di occorrenza e struttura dell’evento.

Sovrapponendo più volte i fogli trasparenti è possibile ottenere un “programma di qualsiasi lunghezza”, inteso come durata a piacere di un brano.

Cage domanda a queste sovrapposizioni e misurazioni le risposte a “domande [...] riguardo altre questioni o dettagli” nella creazione della partitura. Formulando la “domanda in modo che possa essere risposta dalla misurazione delle perpendicolari”¹² [2].

⁸ <https://sites.google.com/conservatoriopollini.it/sampl/home>

⁹ <https://pase-platform.com>

¹⁰ <https://vimeo.com/showcase/8496541/video/562377091>

¹¹ Alcune delle *Variations*, come ad esempio la terza, sono concepite come *happening*

¹² “[If] questions [arise] regarding other matters or details” e “the question in such a way that it can be answered by measurement of a dropped perpendicular.”

2.1.1 Opere correlate

I primi due esempi di digitalizzazione della partitura di *Variations II* sono rintracciabili in *Variations 10b*, versione interattiva del brano di Cage realizzata da Nicholas Knouf [7] con l’ausilio di Processing per la parte grafica e Python come interfaccia verso Csound per la parte sonora, e nell’installazione sonora *Variations II* di Pierpaolo Leo [8].

Entrambi gli adattamenti sono delle installazioni che, oltre a generare una versione digitale della partitura, hanno lo scopo principale di restituire una versione sonora del brano udibile dai visitatori che interagiscono con essa.

Un ulteriore esempio di digitalizzazione della partitura di Cage viene offerto dai membri del *Decibel New Music Ensemble*¹³, i quali hanno realizzato diversi adattamenti digitali per tutte le *Variations*, implementando dei software in Max integrato con Java [6] [5] e un’applicazione per iPad¹⁴ [9].

Diversamente da quanto fatto nei due esempi precedenti, l’ensemble ha realizzato le digitalizzazioni delle partiture con l’obiettivo di fornirsi dei “materiali per una performance dal vivo”¹⁵ [6] più agevolmente utilizzabili in fase esecutiva.

2.1.2 Digitalizzazione della partitura

A differenza delle realizzazioni orientate all’installazione d’arte di Knouf [7] e Leo [8] e al fatto che i software realizzati dal *Decibel New Music Ensemble* sono proprietari, non liberamente distribuiti e, nel caso dell’app per iPad, a pagamento, la versione dell’autore si pone come alternativa gratuita, riscalabile in quanto *open source* e utile ai fini di un’esecuzione musicale con una qualsiasi formazione, sia acustica sia elettroacustica.

Nel progettare la versione digitalizzata della partitura di *Variations II* sono stati arbitrariamente previsti 5 esecutori, pari al numero di persone che hanno testato in sede di concerto la prima versione del software¹⁶, ognuno dei quali con una diversa sovrapposizione degli 11 elementi grafici. Ogni sovrapposizione restituisce 30 determinazioni, ossia 5 per ognuno dei 6 parametri musicali, che vengono assegnate alle diverse sezioni di cui si compone il brano di Cage.

La struttura temporale di questa versione di *Variations II* è fissa, uguale per tutti gli esecutori e suddivisa di 5 sezioni di lunghezza uniforme: di default la durata totale è fissata simbolicamente in 9’6“, ossia il doppio della durata del celeberrimo 4’33” [10]. Il software permette, comunque, di modificare a piacere tale valore.

Al fine di rendere universale l’indicazione di altezza e non vincolarla a un preciso valore in Hz o a un’altezza in notazione tradizionale, il parametro frequenza è stato in-

¹³ <https://decibelnewmusic.com>

¹⁴ <https://decibelnewmusic.com/the-decibel-complete-john-cage-variations-scoreplayer/>

¹⁵ “[scored] materials for live performance”

¹⁶ Esecuzione a cura di Gabriel Bizzo, Alessandro Gambato, Mattia Pizzato, Christian Rapisarda, diretti da Julian Scordato. Inizialmente gli esecutori previsti erano 5

terpretato come registro e suddiviso in 5 regioni: basso, medio basso, medio, medio alto e alto¹⁷.

L'ampiezza viene invece espressa in 8 indicazioni dinamiche della notazione musicale, da più che pianissimo (*ppp*) a più che fortissimo (*fff*)¹⁸.

I timbri sono 5, indicati con valori interi da 1 a 5, e possono rappresentare una diversa tecnica strumentale, o un vero e proprio timbro differente ove lo strumento lo permetta.

I tre restanti parametri vengono espressi come valori reali tra 0 e 1, e riferiti alle indicazioni testuali di durata, tra breve e lungo, occorrenza, tra rarefatto e denso, e infine di struttura, tra semplice e complessa.

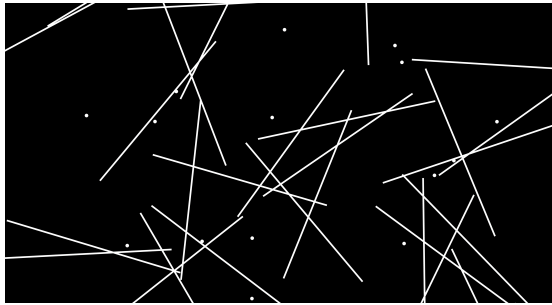


Figura 3. Videoproiezione delle sovrapposizioni.

La durata e l'occorrenza vengono interpretati come la durata di un singolo evento sonoro e la densità all'interno di una delle 5 sezioni. La struttura degli eventi viene interpretata come la complessità dell'articolazione degli eventi sonori, facendo riferimento alla differenziazione tra aggregati e costellazioni di suoni che fa Cage [2].

```
function lines(i)
{
  var a = Math.random() * 6.28318531;
  var s = Math.sin(a);
  var c = Math.cos(a);
  var px = 0;
  var py = 0.5;
  var xnew = px * c - py * s;
  var ynew = px * s + py * c;
  var cx = -xnew;
  var cy = -ynew;
  var tx = Math.random() - 0.5;
  var ty = Math.random() - 0.5;
  px = xnew + tx;
  py = ynew + ty;
  cx += tx;
  cy += ty;

  linex1[i] = cx;
  linex2[i] = px;
  liney1[i] = cy;
  liney2[i] = py;
  outlet(0, cx, cy, 0, px, py, 0);
}
```

Figura 4. Funzione Javascript per determinare le posizioni delle linee.

¹⁷ Nell'interfaccia grafica, le regioni sono abbreviate con B, MB, M, MA e A

¹⁸ Le restanti indicazioni dinamiche sono *pp*, *p*, *mp*, *mf*, *f* e *ff*

```
function dots(i)
{
  var px = Math.random() * 2 - 1;
  var py = Math.random() * 2 - 1;

  dotx[i] = px;
  doty[i] = py;
  outlet(1,px,py);
}

function bang(){
  for(i=0;i<6;i++){
    {
      lines(i);
      if(i<5)
        dots(i);
    }
  }
  for(i=0;i<5;i++){
    {
      var x = dotx[i];
      var y = doty[i];

      for(j=0;j<6;j++){
        {
          var x1 = linex1[j];
          var y1 = liney1[j];
          var x2 = linex2[j];
          var y2 = liney2[j];
          measures(x, y, x1, y1, x2, y2);
        }
      }
    }
  }
}
```

Figura 5. Funzione Javascript per determinare le posizioni dei punti.

La sovrapposizione degli elementi grafici e il calcolo delle distanze tra di essi avviene tramite un programma scritto in JavaScript¹⁹, che genera le coordinate casuali, in un piano cartesiano, delle linee e dei punti. I valori risultanti dalle misurazioni vengono successivamente normalizzati tra 0 e 1 e assegnati, con le eventuali opportune quantizzazioni, a uno dei 6 parametri musicali. In Figura 4, Figura 5 e Figura 6 vengono rispettivamente illustrate le funzioni per la determinazione delle posizioni delle linee, dei punti e per la misurazione delle perpendicolari tra quest'ultimi e le linee.

L'insieme delle linee e dei punti di tutte le sovrapposizioni viene trasformato, grazie all'utilizzo di Jitter²⁰, in oggetti grafici OpenGL in due dimensioni e visualizzati in una finestra videoproiettabile. Questa visualizzazione, illustrata in Figura 3, è utilizzabile come sfondo visivamente estetico durante la performance.

2.1.3 Interfaccia grafica

L'interfaccia grafica è divisa in due parti: una lato server per la generazione, il controllo e la distribuzione della partitura in rete, e una lato client che fornisce le indicazioni esecutive e il cronometro per la sincronizzazione dell'esecuzione. L'interfaccia server (vedi Figura 7) è visibile all'interno della *patch* di Max, mentre le 5 istanze dell'interfaccia client (vedi Figura 8) sono contenute in altrettanti *tab* della pagina html generata da MiraWeb.

¹⁹ Utilizzando l'oggetto *js* che permette di utilizzare JavaScript nell'ambiente Max

²⁰ Jitter è lo strumento incluso in Max per la generazione e il controllo


```

function measures(x, y, x1, y1, x2, y2) {
  var A = x - x1;
  var B = y - y1;
  var C = x2 - x1;
  var D = y2 - y1;

  var dot = A * C + B * D;
  var len_sq = C * C + D * D;
  var param = -1;
  if (len_sq != 0)
    param = dot / len_sq;
  var xx, yy;

  if (param < 0) {
    xx = x1;
    yy = y1;
  }
  else if (param > 1) {
    xx = x2;
    yy = y2;
  }
  else {
    xx = x1 + param * C;
    yy = y1 + param * D;
  }

  var dx = x - xx;
  var dy = y - yy;
  outlet(2, Math.sqrt(dx * dx + dy * dy));
}

```

Figura 6. Funzione Javascript per la misurazione delle perpendicolari tra linee e punti.

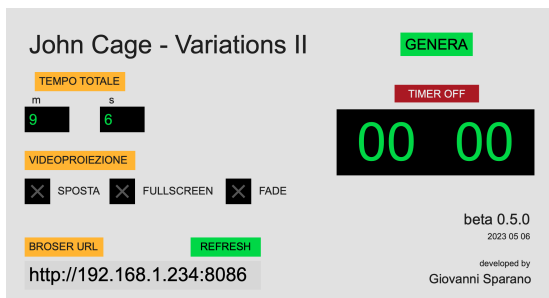


Figura 7. Interfaccia server.

Nell'interfaccia server sono presenti, quali elementi grafici di controllo, un pulsante per la generazione della partitura, 2 controlli numerici per l'inserimento del tempo totale della performance, un cronometro provvisto di interruttore per l'attivazione e spegnimento, 3 controlli per la gestione della videoproiezione, un display, munito di pulsante di refresh, per l'indicazione dell'indirizzo IP e della porta, da comunicare ai client per il collegamento.

Nell'interfaccia client sono presenti, quali elementi visivi di supporto all'esecuzione, un cronometro per la sincronizzazione, una barra di scorrimento temporale suddivisa graficamente nelle 5 sezioni previste, 3 *tab* per il registro, la dinamica e il timbro, infine 3 *slider* orizzontali per la durata, l'occorrenza e la struttura dell'evento sonoro. A ogni cambio di sezione, i valori dei parametri nell'interfaccia grafica vengono aggiornati.

di video e grafica 2D e 3D



Figura 8. Interfaccia client.

2.2 Frame

Frame è un brano dell'autore scritto per l'Arazzi Laptop Ensemble²¹ nel 2022 e liberamente ispirato a *Variations II*, dal quale eredita la didascalia di "per un qualsiasi numero di esecutori e per qualsiasi mezzo di produzione del suono" [2].

La prima esecuzione assoluta del brano, a cura del Pollini Electroacoustic Ensemble²², è avvenuta il 28 settembre 2023 presso la Sala Concerti del Conservatorio di Musica "B. Marcello" di Venezia, nell'ambito del 68° Laboratorio di musica e tecnologie digitali, evento organizzato dalle Scuole di Musica Elettronica dei Conservatori di Padova e Venezia e dal Laboratorio SaMPL²³.

2.2.1 Algoritmo generativo

L'algoritmo di generazione delle parti distribuite in rete per gli esecutori è stato implementato interamente in ambiente Max e può essere suddiviso in 6 parti, ognuna delle quali determina, rispettivamente, gli aspetti formali e parametrici relativi alle griglie temporali per le singole parti, le tipologie di suoni, la densità, le altezze, le dinamiche e le sezioni di pausa.

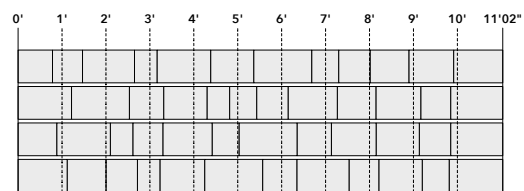


Figura 9. Griglia temporale d'esempio.

La durata del brano è definita dalla somma di 12 numeri primi in successione, che rappresentano le sezioni del brano in secondi. Il suo valore predefinito è 662 secondi, ossia 11 minuti e 2 secondi, dato dalla sequenza:

31 37 41 43 47 53 59 61 67 71 73 79.

Per durate totali diverse è possibile utilizzare 12 numeri primi consecutivi differenti, come ad esempio la sequenza:

²¹ <http://www.arazzilaptopensemble.net>

²² Esecuzione a cura di Simone Bastianello, Francesco Rasi, Federico Sconosciuto, Maria Sensolo e Tito Scutari

²³ <https://sites.google.com/conservatoriopollini.it/sampl/produzione>

11 13 17 19 23 29 31 37 41 43 47 53,

che ha una durata totale di 364 secondi, ossia 6 minuti e 4 secondi. Le griglie temporali per ogni esecutore sono ottenute mediante permutazione della successione di numeri primi prescelti. In Figura 9 viene mostrato un esempio di griglia temporale per 4 esecutori, utilizzando la successione di numeri primi predefiniti.

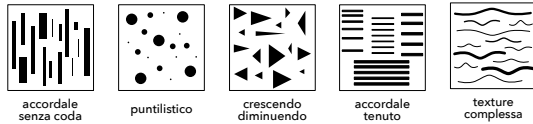


Figura 10. Tipologie di suoni.

Le tipologie di suoni descrivono le possibili articolazioni sonore e sono: accordale senza coda, puntillistico, crescendo decrescendo, accordale tenuto e texture complessa. Ognuna di esse è rappresentata con un disegno intuitivo, come illustrato in Figura 10, che suggerisce agli esecutori come caratterizzare l'esecuzione musicale. Viene generata dal software un lista di 12 determinazioni casuali di valori interi tra 1 e 5, con eventuali ripetizioni, che rappresenta le tipologie di suoni per ognuna delle sezioni.

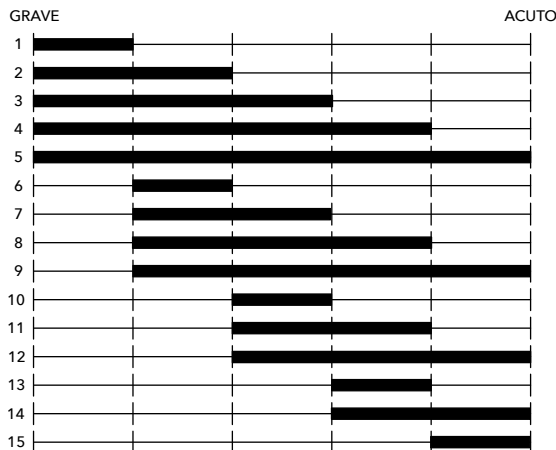


Figura 11. Schema degli intervalli di altezze.

Le densità degli eventi sonori, analogamente alle tipologie di suoni, si ottengono generando una lista di 12 determinazioni casuali di valori interi tra 1 e 5, con eventuali ripetizioni, che rappresenta i diversi gradi di densità per sezione, da rarefatto a denso.

Anche le altezze si ottengono generando una lista di 12 determinazioni casuali, ma di valori interi tra 1 e 15 che descrivono un intervallo di altezze, dal grave all'acuto, secondo lo schema riportato in Figura 11, che illustra come appare lo *slider* di questo parametro nell'interfaccia grafica, di seguito descritta.

Le dinamiche mutuano schema, sistema di generazione dalle altezze e modalità di rappresentazione nell'interfaccia grafica, ma nell'intervallo di indicazioni dinamiche tra *pp* e *ff*, secondo lo schema riportato in Figura 12.

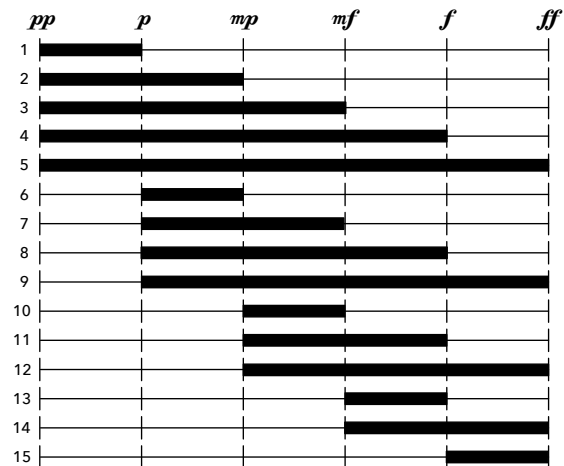


Figura 12. Schema degli intervalli di dinamiche.

Infine, si determina casualmente quale delle 12 sezioni sarà in pausa.

2.2.2 Interfaccia grafica

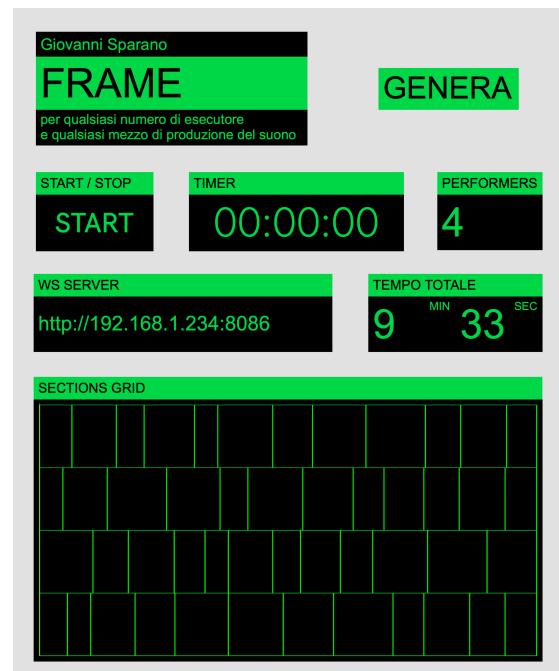


Figura 13. Interfaccia server.

L'interfaccia grafica di *Frame*, che rispecchia le scelte di strutturazione di base fatte per *Variations II*, è divisa in due parti, una lato server e una lato client, come illustrato in Figura 13 e in Figura 14. Uniche eccezioni sono l'assenza del controllo della videoproiezione lato server²⁴, il cronometro lato client, e la presenza sul server di un di-

²⁴ Non ci sono, in questo caso, elementi visivi da mostrare

splay per la visualizzazione della griglia temporale, simile allo schema temporale di Figura 9, e dotata di cursore di scorrimento quando il cronometro è attivo.

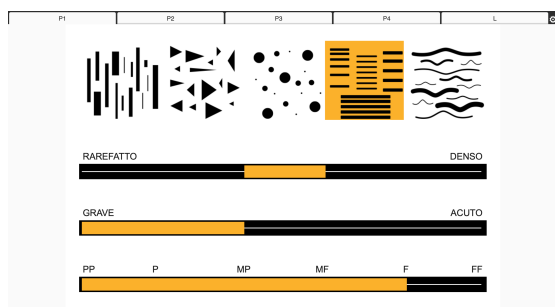


Figura 14. Interfaccia client.

Nell'interfaccia server sono presenti, quali elementi grafici di controllo, un pulsante per la generazione della partitura, 2 controlli numerici per l'inserimento del tempo totale della performance, un cronometro provvisto di interruttore per l'attivazione e spegnimento e un display, con collegamento ipertestuale in questo caso, per l'indirizzo IP e la porta.

L'interfaccia client è suddivisa in un numero di *tab* pari al numero degli esecutori impostati lato server, più un ulteriore *tab* nel quale troviamo un breve testo introduttivo, alcune indicazioni esecutive e una legenda (vedi Figura 15).

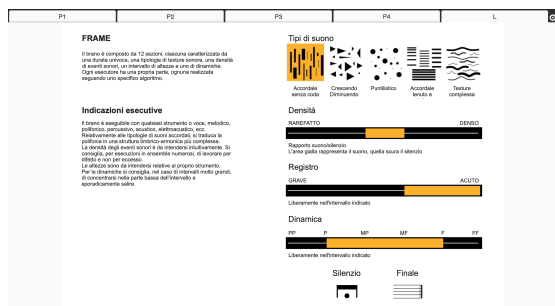


Figura 15. Tab della legenda.

Nell'interfaccia client sono presenti, quali elementi visivi di supporto all'esecuzione, un *tab* con le tipologie di suoni e 3 *slider* orizzontali che indicano, rispettivamente, la densità, il registro e la dinamica.

I valori dei parametri nell'interfaccia client vengono aggiornati a ogni cambio sezione, ma a differenza di quanto avviene per *Variations II*, le sezioni non cambiano in modo sincrono tra gli esecutori.

3. CONCLUSIONI E SVILUPPI FUTURI

Entrambi i software sono stati largamente testati sia dall'autore, sia da 13 persone terze, che hanno suggerito modifiche al software, alcune implementate a partire dalla seconda versione, altre programmate per futuri sviluppi. Nel-

la prima versione di *Variations II* non erano presenti nell'interfaccia client né il cronometro, né la barra di scorrimento temporale, utile invece a comprendere quando avviene un cambio di sezione temporale. Nella prima versione di *Frame*, gli esecutori non erano in grado di capire che erano giunti al finale dell'esecuzione, in quanto tale indicazione era visibile solo nell'interfaccia server. È stato quindi aggiunto un simbolo di pentagramma con la doppia stanghetta finale in tutte le interfacce client al momento della conclusione del brano.

Per entrambi i software, MiraWeb si è dimostrato uno strumento per il supporto all'esecuzione musicale molto affidabile, facilmente e rapidamente gestibile in sede di concerto.

Oltre alle prime utilizzazioni dei software di cui si è trattato in precedenza, *Variations II* è stato eseguito²⁵ il 1° ottobre 2021 presso la Sala Concerti di Palazzo Pisani di Venezia, in occasione del Festival Giovanni Morelli organizzato dalla Fondazione Ugo e Olga Levi di Venezia, in collaborazione con il Conservatorio di Musica "B. Marcello" Venezia²⁶. *Frame* è stato eseguito²⁷ il 25 novembre 2023 presso lo Spazio Biosfera di Padova, in occasione dell'evento *Come Terra*, organizzato dal Laboratorio SaMPL e dal Centro di Sonologia Computazionale dell'Università di Padova²⁸.

Gli immediati sviluppi futuri, suggeriti da alcuni degli utilizzatori del software, prevedono la possibilità di inserimento di un numero a piacere di esecutori per il software di *Variations II* e l'inoltro dei parametri esecutivi in rete, per entrambi i software, tramite protocollo OSC al fine di rendere l'esecuzione multimodale: in questo modo un esecutore informatico potrà suonare intuitivamente osservando l'interfaccia per l'esecuzione, oppure demandare al proprio ambiente esecutivo il controllo di alcuni o tutti i parametri musicali.

4. BIBLIOGRAFIA

- [1] M. Puckette, "New Public-Domain Realizations of Standard Pieces for Instruments and Live Electronics," in *International Conference on Mathematics and Computing*, 2001.
- [2] J. Cage, *Variations II*. Peters Edition, EP 6768, 1961.
- [3] J. Pritchett, "David Tudor as Composer/Performer in Cage's *Variations II*," in *Leonardo Music Journal*, vol. 14, pp. 11–16, 2004.
- [4] S. Tarakajian, D. Zicarelli, and J. Clayton, "Mira: Liveness in iPad Controllers for Max/MSP," in *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 421–426, 2013.

²⁵ Esecuzione a cura di Arazzi Laptop Ensemble: Jacopo Caneva, Giovanni Dinello, Giorgio Klauer, Nicola Privato, Nicola Raccanelli, Luca Richelli, Giovanni Sparano, Paolo Zavagna

²⁶ <https://www.fondazionelevi.it/event/un-ciolo-nascosto-festival-giovanni-morelli/>

²⁷ Esecuzione a cura del Pollini Electroacoustic Ensemble: Simone Bastianello, Francesco Rasi, Federico Sconosciuto, Maria Sensolo, Tito Scutari

²⁸ <https://sites.google.com/conservatoriopollini.it/sampl/produzione>

- [5] C. A. Hope, S. James, and L. Vickery, “New digital interactions with John Cage’s Variations IV, V and VI,” in *Proceedings of the Australasian Computer Music Conference*, 2012.
- [6] L. Vickery, C. Hope, and S. James, “Digital adaptations of the scores for Cage Variations I, II and III,” in *Proceedings of the International Computer Music Conference 2012*, 2012.
- [7] N. Knouf, “Variations 10b: a digital realization of Cage’s Variations II,” in *Proceedings of the 14th ACM International Conference on Multimedia*, MM ’06, (New York, NY, USA), p. 736–739, Association for Computing Machinery, 2006.
- [8] P. Leo, “VARIATIONS II: sound installation by Pierpaolo Leo.”
<http://www.pierpaololeo.it/interaction/variations.htm>, 2006.
- [9] C. Hope, L. Vickery, A. Wyatt, and S. James, “Mobilising John Cage: The Design and Creation of Score Generators for the Complete John Cage Variations I–VIII,” *Malaysian Journal of Music*, vol. 2, p. 34–45, Feb. 2013.
- [10] J. Cage, *4’33”*. Peters Edition, EP 6777a, 1952.

MANIFOLD VOYAGER

Itinerario sperimentale tra forme musicali ricorsive

Davide Commone

Conservatorio Benedetto Marcello - Venezia

videcado@protonmail.com

commone.davide.student@conservatoriovenezia.eu

ABSTRACT

In questo articolo viene esposto il metodo compositivo utilizzato per la realizzazione di un brano per trio di tromboni, fixed media e live electronics; ne vengono evidenziate le suggestioni musicali, scientifiche e tecnologiche che hanno contribuito alla sua realizzazione.

Il lavoro viene proposto come un esperimento musicale, il cui fine è tentare di formulare un metodo compositivo dove la distribuzione dei suoni nel tempo viene gestita da un algoritmo che, attraverso una serie di dati estratti da alcune registrazioni di paesaggi sonori e attraverso l'utilizzo di procedimenti implementati in una forma ricorsiva, genera un reticolo di istruzioni su come disporre nel tempo materiali musicali.

L'attenzione rivolta alle diverse aree tematiche non direttamente collegate tra loro – quali la teoria musicale, lo studio del paesaggio sonoro e la teoria dell'informazione – nasce principalmente dalle suggestioni suscitate dai fenomeni di ripetizione e trasformazione presenti nella produzione musicale di Steve Reich; dagli studi che mettono in relazione il paesaggio sonoro e le sue proprietà musicali, in particolare quelli condotti da Murray Schafer [1] e dagli studi sulla teoria dell'informazione e sulla logica formale, in rapporto a certa produzione artistica, condotti dall'informatico e scienziato cognitivo Douglas Richard Hofstadter [2].

1. FORME, STRUTTURE, MORFOLOGIE: STRUMENTI PER L'ORGANIZZAZIONE DEL SUONO NEL TEMPO

1.1 Definizioni e contesto

In questo capitolo verranno esposti brevemente i risultati di una ricerca preliminare dove sono state approfondite le possibili interpretazioni e le applicazioni nella composizione musicale di alcuni concetti chiave per la concezione di un metodo compositivo, al fine di poter utilizzare in modo consapevole questi concetti come fondamento del lavoro oggetto di questo articolo.

Copyright: © 2024 Commone. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Nell'ambito della composizione musicale, per *forma* si può intendere un insieme di procedimenti [3] che hanno come scopo lo svolgimento nel tempo di un materiale musicale, al quale viene *data forma* alterando e modificando le sue caratteristiche in base alle proprietà del sistema utilizzato per la costruzione del materiale; questa definizione generica del concetto di forma è stata maturata da Charles Rosen durante i suoi studi sulle forme Sonata del periodo classico e viene esposta nel saggio *Sonata Forms* con l'intento di chiarire la differenza tra una forma sonata nel senso di una prassi codificata a posteriori e le *forme sonata* intese come un eterogeneo insieme di procedimenti, che si rivelano essere anche alla base di altri generi.

Nel caso specifico, parlando del genere della Sonata del periodo classico, si possono definire materiali i cosiddetti *temi*, costruiti utilizzando le regole del sistema armonico tonale, sulla cui proprietà fondante, quale il rapporto tonica-dominante-tonica, si basa tutto il vasto insieme di procedimenti chiamati modulazioni; in sintesi si può quindi riassumere che la forma di una Sonata consiste nel modo in cui il compositore ha lavorato sui temi e, nello specifico, il percorso armonico che ha tracciato con le modulazioni.

Rosen, con questa teoria, vuole prendere le distanze dalla principale corrente analitica nata alla metà del XIX secolo, secondo la quale la *forma* di una Sonata consisterebbe nella scansione dei momenti di Esposizione, Sviluppo e Ripresa; d'accordo con quanto studiato da Rosen, questa visione parla più di una prassi codificata a posteriori e restituita in questo modo a scopo didattico, oppure di un possibile modo di dare una scansione, un'ordine strutturale alle diverse tipologie di procedimenti di elaborazione del materiale operati dal compositore nello stesso brano; in questo senso la divisione Esposizione-Sviluppo-Ripresa descrive soltanto uno dei possibili modelli cristallizzati di struttura, non può essere considerato una definizione per *forma*.

John Cage ha riassunto in modo estremamente sintetico ed astratto come la *struttura* può essere considerata *divisione dell'intero in parti* [4]; questa definizione può essere assunta sia come una direzione in cui procedere per un'analisi sia come direzione per un metodo compositivo che preveda un'organizzazione sistematica dei procedimenti; nel primo caso l'*intero* è già un brano la cui composizione può considerarsi ultimata e dove la

sua *divisione in parti* ha come finalità il tentativo di ricostruire i procedimenti organizzativi con cui ha operato il compositore, nel secondo caso *l'intero* rappresenta un progetto e la sua *divisione in parti* ha il compito di organizzare il lavoro del compositore nella realizzazione del brano e di dare una scansione ai procedimenti utilizzati.

La particolarità, che in alcuni casi diventa un aspetto problematico di queste categorie, è che dal punto di vista del loro significato letterale corrente, queste parole sono utilizzate per parlare di oggetti con i quali ci si confronta attraverso l'esperienza visiva, dove il dominio di qualsiasi considerazione sulla forma e la struttura di un oggetto è lo spazio. Ad esempio, nell'etimologia della parola forma, «Figura esteriore della materia [...] Indi, modo di disporre la materia nei lavori umani» [5], compare il concetto di disposizione della materia ma nel significato corrente più utilizzato, ovvero, “*L'aspetto esteriore con cui si configura ogni oggetto corporeo o fantastico*” [6].

Occuparsi della struttura e della forma nella composizione musicale significa occuparsi del modo in cui la nostra memoria riesce a percepire il rapporto tra il suono e il tempo. Alcune problematiche sorgono quando si cerca di traslare una caratteristica propria del dominio dello spazio così come inteso dalle arti figurative al dominio del tempo inteso dal punto di vista musicale; anche se può funzionare dal punto di vista teorico, questo isomorfismo non esiste nella percezione umana di questi due domini [7].

Nella letteratura teorica dell'ultimo secolo è stato utilizzato di frequente dai compositori, soprattutto coloro che si sono occupati di musica elettroacustica, il termine *morfologia* per descrivere alcuni aspetti del loro metodo compositivo e del modo in cui viene visto il rapporto tra il materiale sonoro e il tempo. James Tenney, in un articolo del 1969 [8], considera come *morfologia* il modo di svolgere e articolare nel tempo i singoli parametri che caratterizzano il materiale; collegando ciò alle precedenti interpretazioni di forma e struttura, possiamo dire che il concetto di morfologia può essere interpretato nella composizione musicale come una forma di basso livello.

Tracciando un'epistemologia del rapporto tra forma e struttura, utile per abbozzare un metodo compositivo, possiamo riassumere ciò: il punto di partenza è il materiale, che può essere originale o meno; devono essere note le proprietà del sistema con cui è stato costruito o acquisito quel materiale; in base alle sue peculiarità e alla direzione estetica che si decide di intraprendere, si decidono i procedimenti da utilizzare, si suddividono i procedimenti in tipologie, dando loro un'organizzazione per distribuirli nel tempo, creando quindi una struttura. Successivamente verranno esposti dei modelli teorici, secondo i quali il metodo appena descritto può essere applicato a diversi livelli di profondità durante la composizione di un brano, trattando i risultati ottenuti come oggetti attraverso cui costruire altre versioni di loro stessi.

1.2 Processi sistematici ed estetica

L'utilizzo di materiali musicali, generati o acquisiti attraverso l'utilizzo di tecnologie elettroacustiche o digitali, ha posto per alcuni compositori il problema di formalizzare in modo sistematico dei metodi compositivi e di basare il loro operato su interessi trasversali come la matematica, la statistica e la fisica acustica; attribuendo all'espressione *fare musica* il significato di costruire sistemi di simboli per organizzare i suoni. È necessario notare che talvolta, questa direzione è stata accompagnata da un manifestato disinteresse nei confronti del risultato estetico generato da questi procedimenti [9], nel lavoro oggetto di questo articolo l'interesse legato alla creazione di sistemi per l'organizzazione del suono non ha come scopo quello di intraprendere questa direzione.

Un lavoro di tipo sistematico sulla forma può convivere con considerazioni di tipo estetico sul risultato ottenuto dai procedimenti, quando nella ricerca del metodo compositivo viene considerato anche il problema della percezione, come dimostrato nel caso di Reich; un esempio sono tutti i procedimenti che nella composizione di musica elettroacustica considerano la diffusione del suono nello spazio come una caratteristica fondamentale dell'opera; in questo caso, ricerche sul modo in cui viene percepito dall'ascoltatore il movimento del suono nello spazio e quali sono le sue soglie percettive fanno da fondamento per metodi di spazializzazione esteticamente funzionali.

In alcuni dei procedimenti utilizzati da Trevor Wishart in *Red Bird: A Political Prisoners's Dream* [10] (1973-77), lo svolgimento dei procedimenti di trasformazione applicati ai suoni concreti utilizzati è preceduto da considerazioni estetiche che riguardano il modo in cui può essere percepita la forma dei materiali utilizzati [11]; ad esempio viene osservato come la somiglianza della forma occupata nello spazio di due oggetti come un libro e una porta coincide con la somiglianza – in un contesto di ascolto acusmatico – che si può riscontrare nel suono prodotto da un libro sbattuto su un tavolo e una porta chiusa con forza; sulle proprietà di questa tipologia di isomorfismo vengono basati i criteri di successione dei suoni concreti utilizzati e la forma delle trasformazioni.

Una parte interessante del modello teorico proposto da James Tenney in *Forms in 20th Century Music* riguarda l'implicazione di due o più livelli gerarchici ognuno dei quali – con propria forma e struttura – articola la forma del livello gerarchico superiore, che a sua volta attraverso le sue proprietà formali e strutturali articola il livello ancora superiore; nello specifico Tenney individua tre livelli gerarchici principali chiamati: *elementi, combinazioni e sequenze*.

Un elemento viene definito come un'unità indivisibile, le cui proprietà sono determinate dal sistema di organizzazione delle altezze (o dei suoni in generale, nel caso in cui gli elementi vengono rappresentati da suoni concreti o suoni generati attraverso tecniche di sintesi)

scelto dal compositore; nel livello gerarchico superiore, più elementi costituiscono delle *combinazioni*, paragonate da Tenney a quello che tradizionalmente viene definito come “frase”, in questo livello gerarchico la forma e la sua organizzazione si occupa di disporre nel dominio del tempo gli elementi, al fine di creare quello che viene chiamato *primary Gestalt*, la natura morfologica e statistica di questo livello gerarchico può influenzare in modo significativo i livelli gerarchici superiori, soprattutto se la forma prende in considerazione altri parametri oltre a quello del tempo, come ad esempio la dinamica, il timbro, o una meta-organizzazione delle altezze che costituiscono l'elemento. In una *sequenza*, ad un livello gerarchico superiore, una successione di *combinazioni di elementi* genera un gesto musicale più ampio e articolato; i procedimenti formali in questo caso riguardano l'elaborazione e i criteri di successione delle combinazioni; le caratteristiche che possono influenzare direttamente il risultato musicale dei procedimenti formali applicati a questo livello dipendono dalla varietà di combinazioni utilizzate; si possono quindi generare sequenze isomorfe (tutte le combinazioni usate nella sequenza hanno la stessa forma) o eteromorfe (tutte le combinazioni nella sequenza hanno una forma differente), oltre ad un utilizzo di durate irregolari delle singole *combinazioni*.

Da questo livello in poi il compositore può decidere se considerare la stratificazione terminata, e quindi ottenere un brano dalla successione di sequenze, oppure continuare ad articolare il procedimento compositivo salendo ancora su altri livelli gerarchici e avere quindi sezioni (successioni di sequenze), movimenti (successioni di sezioni), eccetera; il tipo di ampiezza di respiro della forma ricercato dal compositore e il tipo di contesto musicale che si vuole creare rappresentano quindi una variabile importante per definire cosa è un *elemento* e come esso viene percepito. Lo stesso oggetto può essere percepito come un'unità indivisibile in un contesto musicale estremamente denso, mentre può essere percepito come una combinazione di elementi in una testura musicale più distesa.

2. COSTRUZIONE DI UNA MACCHINA RICORSIVA

2.1 Implementazione dell'algoritmo

Nel seguente paragrafo verrà esposto il lavoro svolto nella realizzazione di un algoritmo, la cui idea e il cui funzionamento è nato dalle suggestioni delle tematiche trattate da R. Hofstadter in *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante*, in particolare l'osservazione e la creazione di fenomeni autoreferenziali chiamati *strani anelli*. Le seguenti applicazioni delle nozioni di logica formale non vengono proposte in questo lavoro con pretese di tipo scientifico o matematico, esse rappresentano solo delle speculazioni nell'ambito della composizione musicale, utilizzando le risorse e le conoscenze acquisite nell'ambito della Computer Music e in

particolare nella Computer Aided Algorithmic Composition (CAAC) [12].

La funzione per cui è stato scritto questo algoritmo è esclusivamente quella di fornire al compositore una sorta di canovaccio, in forma di notazione musicale tradizionale o in forma di schemi, attraverso il quale procedere alla composizione del brano; nelle modalità scelte per l'implementazione dell'algoritmo è stata volutamente esclusa la possibilità di far generare da quest'ultimo un brano completo o una partitura pronta per essere sottoposta a un esecutore; in questo la scelta finale riguardo a dettagli tecnici ed espressivi rimane compito del compositore. L'idea alla base è quella di produrre un meccanismo ricorsivo che riceve in input una sequenza di altezze (in questo caso rappresentate da note senza durata) e che restituisca come output una lista composta da un numero variabile di altre sequenze di altezze, sorelle della lista originale precedentemente inserita. Per realizzare questa idea sono stati utilizzati i procedimenti del *moto inverso* e *moto retrogrado* utilizzati in modo ricorsivo e reciproco: la lista di note viene retrogradata, la lista risultante viene a sua volta invertita ed in questa alternanza il procedimento va avanti fino a che l'algoritmo non individua nel nuovo risultato generato una copia della lista originale da cui si era partiti. Lo stesso tipo di operazione viene applicato in modo parallelo sulla stessa sequenza originale, partendo però dal procedimento di inversione.

Nel diagramma di flusso in Figura 1 viene illustrato il procedimento della parte dell'algoritmo che riguarda il procedimento ricorsivo appena descritto, con un esempio dei risultati. In termini logici, questo rappresenta un sistema formale con un numero finito di livelli gerarchici attraverso i quali si possono fare un numero limitato di *push* prima di ritrovarsi al livello da cui si era partiti. Il numero di risultati possibili che possono essere generati da una sequenza di altezze – ovvero il numero limitato di livelli in cui si può fare un *push* prima di tornare al livello di partenza – è dato dalla logica intrinseca delle relazioni intervallari che caratterizza il sistema temperato utilizzato in questo lavoro, conosciuto anche come *12-ET (12 equal temperament)* [13]. La proprietà in questione riguarda il numero limitato di passi che si possono compiere con uno stesso intervallo prima della ricorrenza della stessa nota dalla quale si era partiti, nel caso di *12-ET* sono l'intervallo di quinta giusta e di seconda minore, insieme ai relativi intervalli complementari di quarta giusta e settima maggiore; solo con questi intervalli si possono compiere 12 passi e percorrere l'intero totale cromatico.

Si può quindi riassumere che in base all'intervallo tra la prima e l'ultima nota della sequenza data in input (oppure considerando il suo complementare, ovvero l'intervallo tra l'ultima e la prima) viene generato un numero di specchiamenti uguale al doppio dei passi possibili per il determinato intervallo, dato che per ogni livello di *push* vengono generate due sequenze, contando anche quella originale da cui parte il procedimento. Nel caso della sequenza in Figura 1, con l'intervallo di

terza maggiore, essendo possibili un massimo di tre *push* vengono generate sei liste per ciascuno dei rami del procedimento.

Per costruire un algoritmo in grado di fornire un canovaccio con un reticolo di istruzioni più specifico e variegato, al fine di poterlo usare concretamente durante la composizione dei brani presentati in questo lavoro, i procedimenti riassunti nel diagramma in Figura 1, insieme ad altri procedimenti che verranno descritti di seguito, sono stati implementati [14] in un codice Python [15] con l'ausilio della libreria esterna *Abjad* [16], utilizzata come *API* del software di notazione musicale *Lilypond* [17].

Le funzioni precedentemente approfondite e riassunte nel diagramma di Figura 1 sono state definite nel diagramma di flusso di Figura 2 con le variabili *riro()* ed *irio()*, dal risultato di queste due funzioni – alle quali viene dato il medesimo input – si formano due percorsi paralleli nell'algoritmo, che successivamente rappresenteranno due sezioni differenti in ogni brano della composizione; i due percorsi sono utilizzati per generare istruzioni con cui assemblare e strutturare lo svolgimento delle due sezioni della composizione, nello specifico, vengono generate istruzioni sulle modalità di interpolazione graduale tra una sequenza e l'altra, sulle variazioni ritmiche e sul ruolo che ciascuno degli stru-

menti impiegati ha nella testura polifonica in ciascuno dei passaggi di interpolazione. Questi procedimenti, tuttavia, lasciano libertà al compositore per quanto riguarda molte scelte di strumentazione, lo scambio e l'incrocio di alcune parti per favorire l'esecuzione da parte degli strumentisti, le indicazioni di articolazione e di dinamica per dare una direzione verso dei *climax* di espressione, oltre alla libertà applicata nel comporre la sequenza di altezze data in input e il relativo modo di distribuire le altezze nelle parti strumentali.

Di seguito verranno esposti alcuni dettagli algoritmici del codice Python. La funzione definita come *combinatore()* si occupa di cambiare l'ordine degli elementi all'interno della lista di sequenze data in input, ovvero quella generata da *riro()* e da *irio()*; il criterio con cui viene eseguita questa operazione è il calcolo del modulo tra ogni elemento della chiave {"onsetfeat":} del dizionario *field_dataset*, costituito da descrittori numerici di registrazioni di paesaggi sonori, e la quantità di sequenze contenuta nella lista data in input; il risultato del calcolo di modulo si occupa di dare una nuova posizione ad ogni sequenza, che diverrà l'ordine con cui queste verranno utilizzate durante la composizione. La funzione definita come *grio()* si occupa di formulare un canovaccio di indicazioni che riguardano il modo graduale con cui si pas-

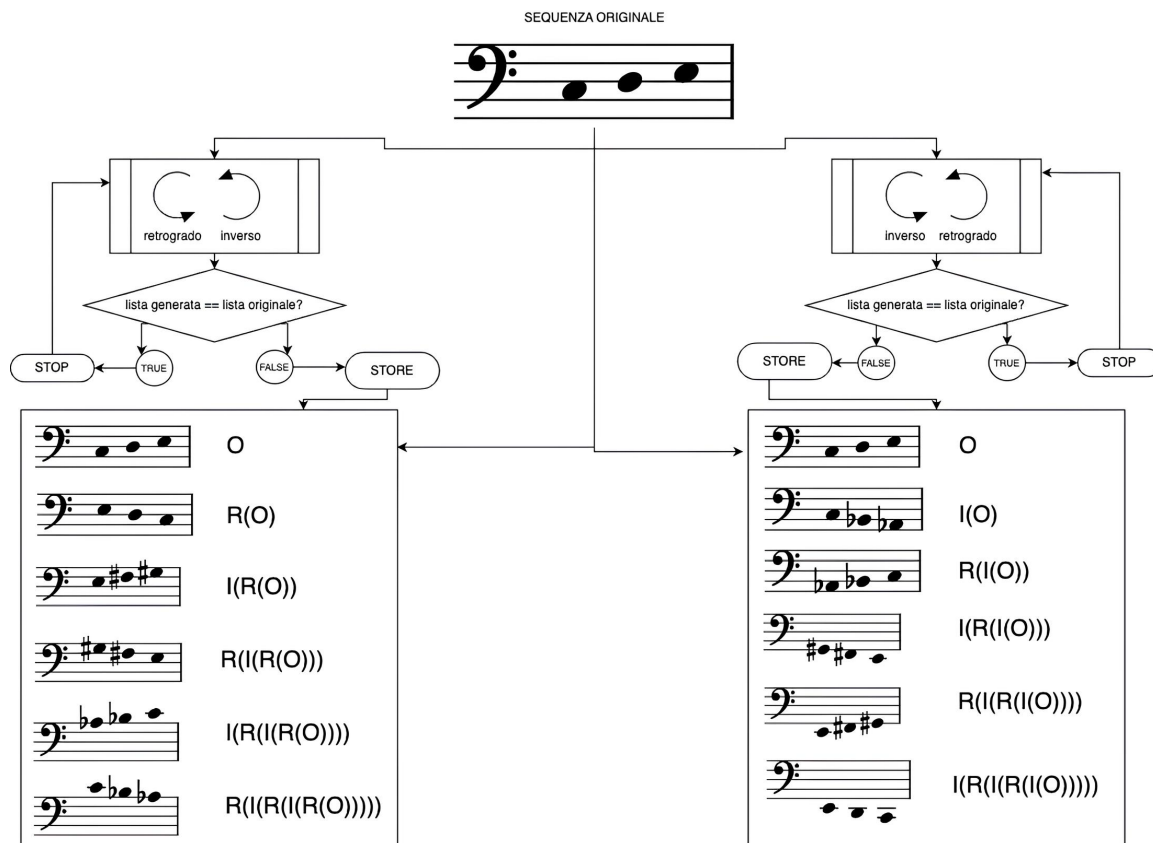


Figura 1. Diagramma di flusso con esempio dei risultati dei procedimenti di retrogrado ed inverso applicati in modo ricorsivo.

recursion_machine.py

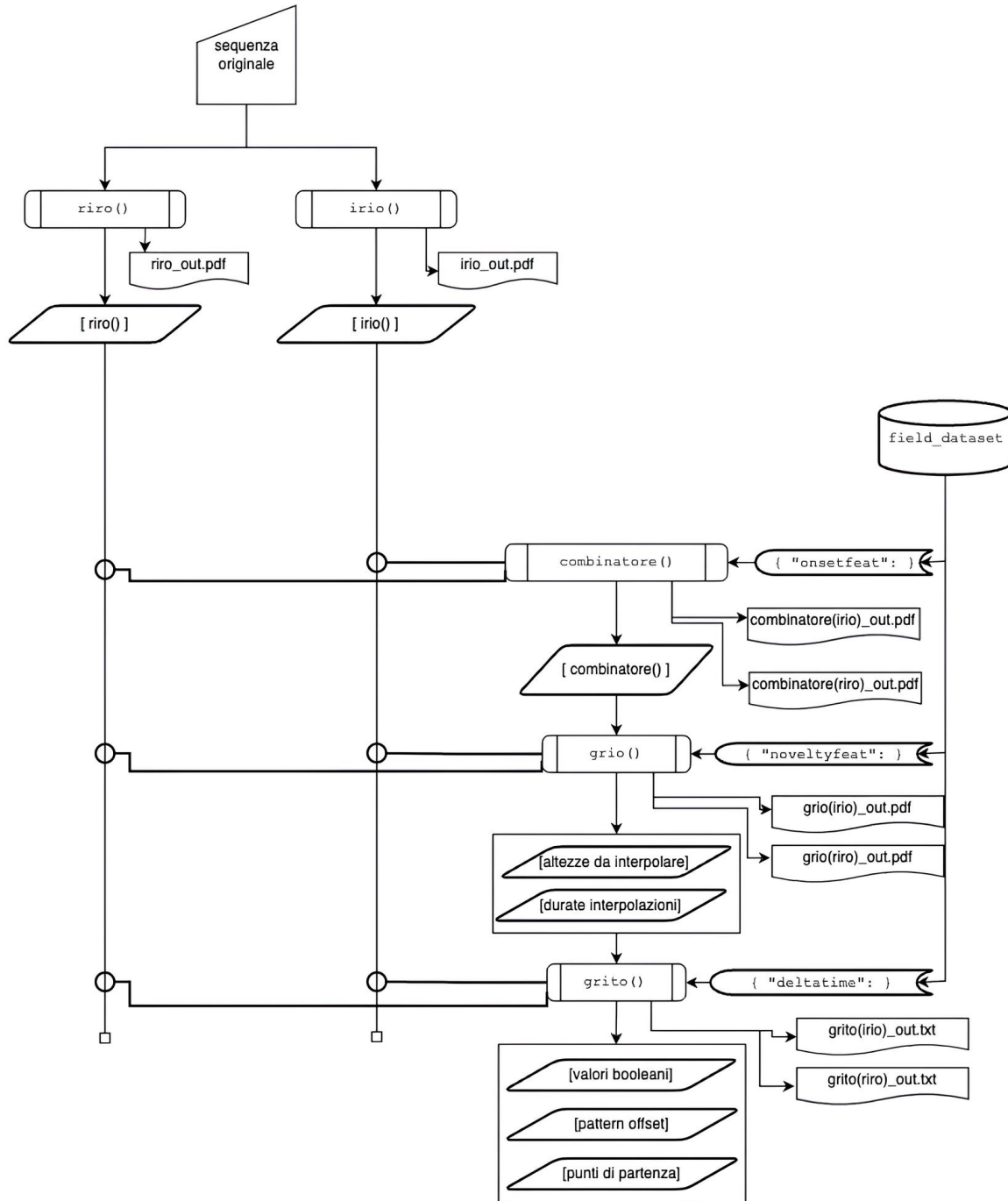


Figura 2. Diagramma di flusso (ANSI standard) del funzionamento generale del codice `recursion_machine.py`. Viene rappresentato il flusso di dati in input e in output delle principali funzioni.

serà da una sequenza all'altra della lista di sequenze riordinata dalla funzione `combinatore()`; attraverso un calcolo di modulo tra ogni elemento della chiave `{"noveltyfeat":}` del dizionario `field_dataset` ed il totale delle note contenute in ciascuna sequenza della lista `output` di `combinatore()` viene stabilito quali altezze saranno sottoposte al procedimento di interpolazione graduale; infine un ulteriore calcolo di modulo tra i descrittori della medesima chiave si occupa di stabilire la durata in battute dell'interpolazione graduale di ciascuna delle note selezionate. Nella Figura 2 è possibile osservare, nei riquadri dedicati all'output della funzione `grio()`, le note selezionate per il procedimento di interpolazione sono evidenziate in rosso e accompagnate dal numero di battute che occuperà l'interpolazione della nota selezionata con la nota che occupa la stessa posizione nella sequenza successiva. La funzione `grito()` si occupa di creare un canovaccio di istruzioni relative allo svolgimento della polifonia e del modo in cui vengono distribuite le altezze tra le parti nel corso delle sezioni di interpolazione; nello specifico, viene creata una collezione di liste contenenti tanti valori booleani (0 ed 1) quanti gli strumenti impiegati; per ogni battuta dell'interpolazione, al valore 0 corrisponde una battuta di pausa, al valore 1 corrisponde l'esecuzione del materiale prestabilito per quella voce. Un'ulteriore istruzione generata da `grito()` è quella che riguarda lo stato di rotazione del pattern ritmico attraverso il quale vengono presentate le altezze di ogni sequenza, quindi, attraverso i descrittori numerici della chiave `{"deltatime":}` del dizionario `field_dataset` viene generato per ogni sequenza un valore in quarti che rappresenta l'offset con il quale si presenterà il pattern. Infine, attraverso la stessa chiave, viene generata una lista di valori che stanno ad indicare la posizione dalla quale viene ripercorso il pattern ritmico durante le battute di interpolazione stabilite dalla funzione `grio()`.

I procedimenti delle funzioni `combinatore()`, `grio()` e `grito()`, sono nati da intenzioni estetiche e musicali che tendono verso un equilibrio tra un procedimento sistematico e quasi *meccanico*, ed un brano che possa coinvolgere l'ascoltatore senza renderlo vittima di un procedimento nel quale non è in grado di intravedere delle intenzioni musicali. In altri termini, non tecnici ma quasi fantascientifici, le funzioni `riro()` e `irio()` rappresentano delle macchine che attraverso lo specchiamento reciproco di un oggetto – che in questo caso è una sequenza di altezze – creano due multiversi costituiti da un numero finito di dimensioni nelle quali l'oggetto specchiato appare in copie ricorsivamente invertite di se stesso; il restante delle funzioni può quindi essere rappresentato come un dispositivo che ha il compito di tracciare un viaggio da compiere – basato sulle coordinate fornite da una raccolta di dati su dei paesaggi sonori – tra le dimensioni di ciascun multiverso.

Attribuire ai procedimenti formali un significato simbolico extra-musicale è stato un tentativo di creare

una sorta di *musica a programma* nella quale i personaggi, le azioni o le allusioni a oggetti extra-musicali non vengono evocati nell'opera attraverso figure musicali come i *Leitmotif* ma attraverso la costruzione di algoritmi che organizzano lo svolgimento e la struttura musicale della composizione.

2.2 Legame con i paesaggi sonori

Come materiale di partenza per la progettazione della composizione sono stati scelti tre frammenti di paesaggi sonori registrati dall'autore; l'interesse legato a questo tipo materiale è nato dalla possibilità di poter ricavare dal segnale audio descrittori numerici di alcuni parametri; nella teoria dei segnali, questa operazione prende il nome di *audio information retrieval*. L'idea di utilizzare descrittori numerici, estratti da registrazioni di paesaggi sonori o da altri tipi di suoni, come struttura per l'organizzazione dei parametri musicali è stata proposta ad esempio in opere come *En Plein Ear* [18] di Giorgio Klauer e *Les espaces acoustiques* [19] di Gérard Grisey; nel lavoro presentato in questo articolo, il legame originariamente soggettivo tra le varie tematiche coinvolte viene presentato come un esperimento musicale.

La tipologia di informazioni che è stata scelta di estrarre dal segnale audio riguarda principalmente il rapporto tra ampiezza e tempo utilizzando un algoritmo implementato in una patch *Pure Data* [20] con l'ausilio della libreria esterna *Fluid Corpus Manipulation* [21].

Per ciascuna traccia vengono ricavate tre liste di valori che riguardano rispettivamente il delta temporale tra i valori di *onset* (misurata in millisecondi) [22]; i valori in decibel relativi alla differenza di ampiezza tra la finestra di analisi nella quale viene individuato l'*onset* e quella che la precede; i valori della terza lista, compresi tra 0 e 1, riguardano la cosiddetta *novelty feature detection*, che viene definita nell'ambito del *machine learning* come un'operazione per il riconoscimento di anomalie in un flusso di dati rispetto ad un modello di normalità precedentemente stabilito e previsto per quel tipo di segnale [23].

I risultati di questa operazione sono stati inseriti, sotto forma di liste organizzate in un dizionario all'interno di un codice Python, che è stato successivamente utilizzato come un modulo da richiamare all'interno di `recursion_machine.py`.

2.3 Sviluppi futuri

Per quanto riguarda futuri sviluppi di questo progetto da parte dell'autore, è stato pensato di procedere in modo parallelo su due livelli. Il primo – non per importanza – è di tipo tecnologico e riguarda l'estensione dell'algoritmo attraverso l'implementazione di altre funzioni e procedimenti; la struttura di `recursion_machine.py` è già in parte modulare, ovvero, l'utente può decidere di utilizzare solo alcune funzioni o di organizzare il flusso di dati in svariate combinazioni, implementando altre funzioni verranno aumentate le possibilità in questa direzione.

Il secondo livello è di tipo musicale ed estetico; a seguito di un primo esperimento di applicazione musicale durante la composizione di un brano per trio di tromboni è stato riscontrato dall'autore che un cambiamento di priorità avvenuto durante la composizione – dove il lavoro che ha assorbito la maggior parte delle attenzioni è stata la formulazione di una procedura e la sua implementazione in un algoritmo – ha portato ad un risultato musicale non del tutto compatibile con le premesse e i desideri estetici da cui era partito il progetto. Questo è avvenuto probabilmente perché l'algoritmo – che doveva essere un mezzo, uno strumento - è diventato lo scopo; per confrontarsi con questo tipo di problematica è stato pensato di comporre dei brani di breve durata, da considerare *studi*.

3. BIBLIOGRAFIA

- [1] Murray Schafer, *The Tuning of the World*, New York, Alfred A. Knopf, 1977.
- [2] Douglas Richard Hofstadter, *Gödel, Escher, Bach: un'Eterna Ghirlanda Brillante*, Milano, Adelphi Edizioni, 1990.
- [3] Charles Rosen, *Sonata Forms*, New York, W.W. Norton & Norton, 1980, pp. 28-70 (*Aria*), pp. 71-97 (*Concerto*).
- [4] John Cage, *Silenzio*, Milano, ShaKe Edizioni, 2010, traduzione di Giancarlo Carlotti, p. 29 (*Composizione come processo - I. Cambiamenti*).
- [5] Francesco Bonomi, *Vocabolario Etimologico della Lingua Italiana*, <https://etimo.it/?term=forma&find>, [12-23].
- [6] <https://www.treccani.it/vocabolario/forma/>, [12-23].
- [7] Jeffrey Izzo, *Space, Time, and Memory: Examining the Disconnect between Looking at Contemporary Art and Listening to Contemporary Music*, «College Music Symposium», LX -II, 2020, p.6.
- [8] James Tenney, and others, *Form in Twentieth-Century Music*, in Larry Polansky, and Lauren Pratt (eds), «From Scratch: Writings in Music Theory», 1969-'70, p. 2.
- [9] Iannis Xenakis, *Formalized Music*, Stuyvesant NY, Pendragon Press, 1992, xiii-xiv.
- [10] Trevor Wishart, *Red Bird (A Political Prisoner's Dream)*, York Electronic Studios, UK, 1978.
- [11] Trevor Wishart, *On Sonic Art*, Edinburgh, Harwood Academic Publishers, 1996, p. 155.
- [12] Jose David Fernandez, Francisco Vico, *AI Methods in Algorithmic Composition: A Comprehensive Survey*, «Journal of Artificial Intelligence Research», XLVIII, 2013, pp. 513-582.
- [13] <https://en.xen.wiki/w/12edo>, Enciclopedia di xenarmonia. [02-24].
- [14] <https://github.com/davidecommone/R-ecursionMachine>
- [15] <https://www.python.org/>
- [16] <https://abjad.github.io/>
- [17] <https://lilypond.org/>
- [18] Giorgio Klauer, *En Plein Ear. Sonic postcards from a busy Summer*, «Musica/Tecnologia», XIV, 2020, DOI: 10.13128/music_tec-12347 – <http://www.fupress.com/m>, pp. 49-72.
- [19] Gérard Grisey, *Les espaces Acoustiques*, Kairos, 0012422KAI, Germany, 2016.
- [20] <https://puredata.info/>
- [21] Tremblay, P.A., Green, O., Roma, G., Bradbury, J., Moore, T., Hart, J., & Harker, A. (2022) The Fluid Corpus Manipulation Toolbox (v.1). Zenodo. doi: /10.5281/zenodo.6834643 [03-24]
- [22] <https://learn.flucoma.org/reference>
- [23] Marco A. F ed altri, *A Review Of Novelty Detection*, «Signal Processing», Volume 99, 2014, pp. 149-215.

Partiture di Voltaggio Metodo di notazione per sintetizzatore modulare

Giuseppe de Benedittis

Conservatorio di Musica “B. Marcello” – Venezia

g.d.benedittis@gmail.com

ABSTRACT

La riproducibilità di un'opera musicale, in assenza dell'autore o delle macchine/strumenti su cui essa è stata concepita, ma soprattutto la sua notazione, sono due problemi della musica elettroacustica.

Negli anni, musicologi, interpreti e compositori hanno affrontato tali argomenti non riuscendo tuttavia a raggiungere un accordo su un sistema comune. Ciò è dovuto principalmente alla volontà e all'interesse del compositore a far eseguire la sua opera ma anche all'utilizzo di strumenti specifici, che spesso sono di difficile accesso e per i quali non esiste un sistema di scrittura codificato per definire i valori dei parametri nel tempo. Non è difficile dunque trovarsi di fronte alla creazione di forme di notazione personali che molto spesso non sono di facile interpretazione o sono troppo sommarie.

Attraverso l'osservazione e il confronto del lavoro svolto in merito da grandi artisti, viene qui proposto un metodo di notazione per sintetizzatori modulari attraverso l'utilizzo del controllo in voltaggio (CV), o meglio tramite lo sviluppo di un sistema hardware in grado di trasformare dati numerici in tensione elettrica che andrà poi a modificare autonomamente i parametri di controllo di un sintetizzatore.

1. INTRODUZIONE

La vastità di possibilità e la complessità nel programmare un sintetizzatore modulare ci fa subito intendere la difficoltà nel notare come cambiano i valori dei parametri di tale strumento nel tempo. Alcuni artisti 'modulari' considerano la *patch* stessa come opera d'arte, un'ottica che conferisce un approccio unico alla preservazione del lavoro. Tale concetto ha guadagnato popolarità parallelamente al crescente interesse per l'arte generativa.

Un esempio di questa idea che pone al centro dell'attenzione il processo lo possiamo trovare nell'album del 1978 di David Chesworth, intitolato *The Unattended Serge* (fig. 1). Quest'opera è una raccolta di estratti provenienti da una registrazione ininterrotta di 24 ore di una *patch* eseguita dal sintetizzatore *Serge* presso l'Università La Trobe di Melbourne.

Copyright: ©2024 Giuseppe de Benedittis et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

David Chesworth
The Unattended Serge 1978 / Five Evolutionary Things 1979
early works for the Serge Tcherepnin synthesizer

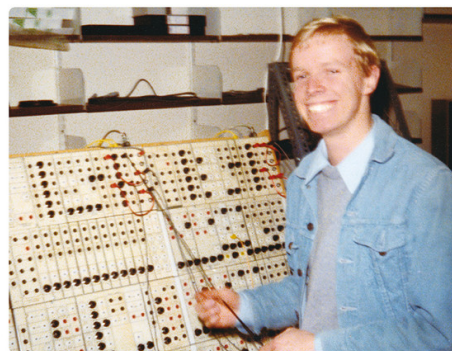


Figura 1. Copertina dell'album contenente *The Unattended Serge*.

2. PRIMI TENTATIVI DI NOTAZIONE

Molti sono stati comunque i tentativi fatti per la creazione di un modello, ma tutt'oggi non abbiamo un linguaggio universale. Popolari, soprattutto tra gli anni '60 e '70, furono i diagrammi di flusso o schemi a blocchi (fig. 2), il cui utilizzo all'epoca era principalmente ingegneristico e ben si sposava con la natura dello strumento.

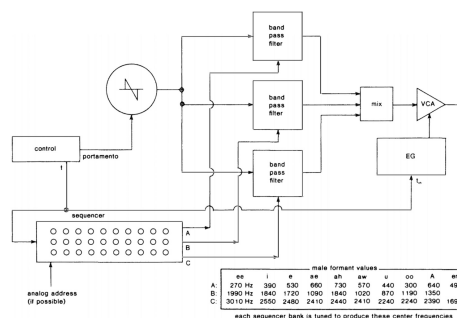


Figura 2. Diagramma di flusso tratto da [1].

Nella fig. 3 possiamo osservare il diagramma relativo al brano *Bromeliads* (25/12/1974) di Warren Burt, tratto dal CD *Harmonia Mundane* e presente nella serie *Serge-*

O-Voxes, dove il principale strumento era il *Serge Modular Synthesizer*.

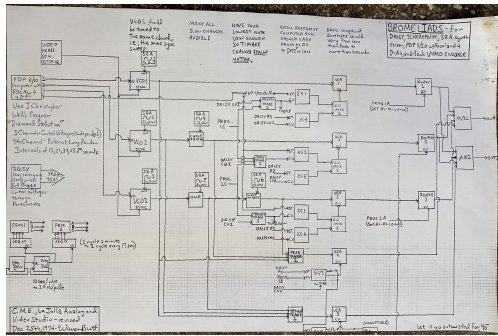


Figura 3. Partitura di *Bromeliads*, per gentile concessione di Warren Burt.

Come si può notare, oltre al diagramma, che ben espone la costruzione della *patch*, abbiamo alcune indicazioni su come il brano debba essere eseguito, indicazione che tuttavia mancano di precisione relativamente ai movimenti dei potenziometri, risultando dunque di scarsa utilità per l'esecuzione.

Invece che sui diagrammi, c'è chi si è concentrato su un aspetto più grafico, e magari più immediato, proponendo vere e proprie raffigurazioni di moduli e di come siano connessi tra di loro, come possiamo osservare in [2]¹ (fig. 4), o come nel moderno *Moog Mavis*² (fig. 5), o ancora come nel caso della compositrice francese Eliane Radigue, che ha sviluppato una propria tecnica di notazione su uno specifico sistema modulare da lei utilizzato, l'*ARP 2500*³ (fig. 6).

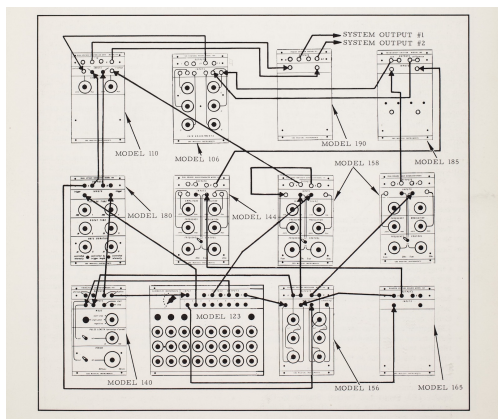


Figura 4. *Patch* di esempio tratta da [2, p. 4].

Possiamo osservare che la notazione è una rappresentazione grafica della macchina dove vengono annotate le po-

¹ Originariamente denominato *Buchla Music Box*, è un sistema ideato da Don Buchla nel 1963 su commissione di Ramon Sender e Morton Subnick.

² Nel mercato da giugno 2022, *Mavis* è un sintetizzatore analogico venduto dalla *Moog Music Inc.*® come kit per una costruzione *faidatè*.

³ Sintetizzatore analogico costruito a partire dal 1969 fino al 1981 dalla *ARP Instruments, Inc.*®



Figura 5. Esempi di cartoncini da posizionare sul sintetizzatore analogico modulare *Moog Mavis* per ottenere configurazioni particolari.

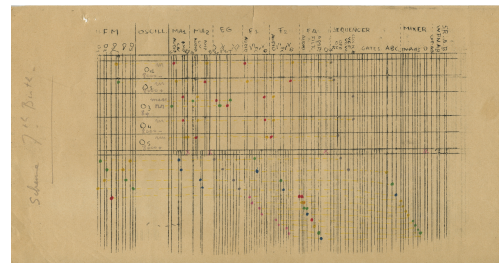


Figura 6. Rappresentazione grafica dell'*ARP 2500* a cura di Eliane Radigue.

sizioni dei fader. Quello che vediamo sono in realtà appunti preparatori datati 1970 del brano *7th Birth*, che uscirà nel 1971, fornendoci quindi un'istantanea delle impostazioni del sintetizzatore ma non una vera partitura.

2.1 Lacune esecutive

Una caratteristica comune tra le rappresentazioni grafiche di una *patch* e gli schemi a blocchi è l'attenzione posta alla riproducibilità della configurazione del sintetizzatore, piuttosto che alla documentazione delle specifiche tecniche esecutive che si sviluppano nel corso del tempo.

Potremmo dunque acquisire una comprensione dettagliata di come impostare un sistema modulare, ma rimarremmo all'oscuro delle dinamiche esecutive, come l'articolazione dei potenziometri o la manipolazione dei cursori durante una *performance*.

Inoltre, mancherebbe la descrizione dell'impiego di interfacce tattili come *joystick*, trasduttori capacitivi o altri sensori per plasmare il suono. Possiamo dunque concludere dicendo che tali scritture non forniscono informazioni sufficienti per la ricreazione esatta della *performance*. Questa mancanza rappresenta una differenza significativa rispetto alla notazione musicale tradizionale.

2.2 Il caso Ciani

Un più complesso esempio è [3]⁴, un *report* di Suzanne Ciani scritto al termine di uno studio sovvenzionato dal *National Endowment for the Arts* nel 1976, nel quale la compositrice cerca di fondere la scrittura tradizionale e quella sperimentale, seppur con qualche dubbio in merito alla sua utilità come documento per eventuali esecutori.

My work now with the Buchla is not in any way traditional and my scoring techniques or what I use to enable somebody else to perhaps recreate my composition on the Buchla is more of a documentation [...]

Well, as a matter of fact, I have a paper [...] that I called *The Buchla cookbook* that documents my early practice in performing the Buchla [...] It's quite interesting and I have used it myself. When I came back to performing on the Buchla, I consulted that paper.

What it does, it outlines the raw materials that are used in the performance. First, it gives you the tone rows for the sequencer (fig. 7). Then it gives you the patch diagram [fig. 8]. Then it gives you examples of moving from one circumstance to another, performance actions [fig. 9] [...]

Live performance is also... It's not like a piano piece where you are recreating a given set definition of the composition, the notes.

It is a little bit more free form, it's like jazz. What would you do if you wanted to reproduce a jazz composition? You'd start really with the fundamentals, you'd start with what I call the raw materials and nowadays, we have tools like just recording. You could say, "And here is the recording of that jazz performance." And you have a document. Does that allow you to reduce it by performing yourself?⁵



Figura 7. Toni per le linee del *sequencer*, tratto da [3].

⁴ https://www.finearts.uvic.ca/~aschloss/course_mat/MU307/MU307%20Labs/Lab3_BUCHLA/Suzanne%20Ciani/Ciani_BuchlaCookbook.pdf

⁵ Estratto da [4], <https://www.notion.so/blog/suzanne-ciani>

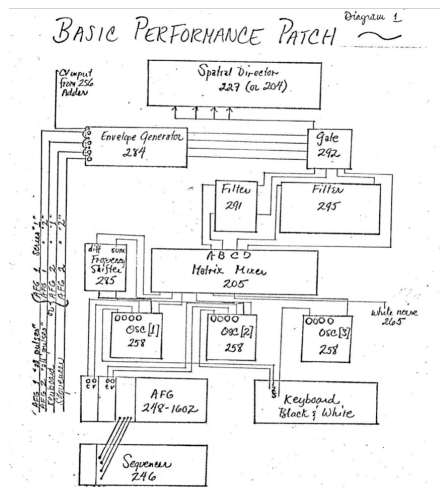


Figura 8. Patch d'esempio, tratto da [3].

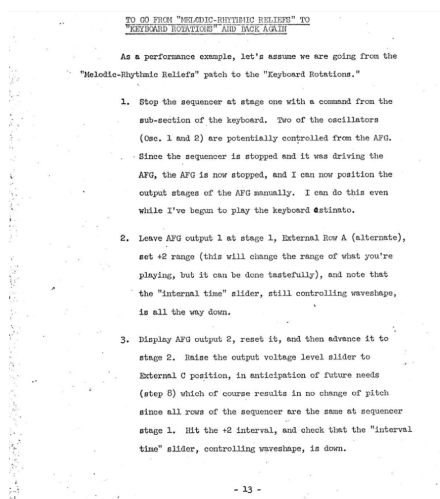


Figura 9. Istruzioni riguardo le azioni esecutive, tratto da [3].

Possiamo dunque identificare il sintetizzatore modulare come uno strumento ribelle che porta il suo esecutore ad abbandonare schemi compositivi predefiniti portandolo ad un' esplorazione basata su un ascolto profondo. Difatti per molti compositori che lavorano con un sistema modulare la preservazione e la notazione non suscitano spesso grande interesse, che invece si manifesta per l'imprevedibilità e l'instabilità dello strumento stesso.

3. NOTAZIONE MEDIANTE CV

Nei sintetizzatori modulari per controllare i vari parametri a disposizione ci si può servire del controllo in voltaggio, ovvero un controllo di tensione elettrica, permettendo così di automatizzare funzioni e comportamenti dei suoni generati dalla macchina.

Il metodo qui presentato si propone dunque di creare una partitura che sfrutti la generazione e il controllo di tale tensione elettrica.

3.1 Descrizione generale del progetto

Nel caso in questione è stato utilizzato come sintetizzatore un *Serge Modular* che utilizza connettori a banana. Attraverso l'utilizzo di un componente MCP4728 [5]⁶, un convertitore da analogico a digitale a 4 canali, interfacciato ad un microcontrollore come *Arduino* [6]⁷, abbiamo la possibilità di controllare ed inviare al sintetizzatore quattro differenti tensioni in quattro differenti ingressi dedicati al CV tramite l'utilizzo del computer.

Crea una semplice *patch* con il software *Max/MSP* (fig. 10) che consenta di inviare dati ad *Arduino* tramite porta seriale e una volta caricato lo *sketch*⁸ sulla *board*, che permette a sua volta di trasmettere i dati correttamente scalati e indirizzati ad ognuno dei quattro canali di *output* a disposizione, tramite l'utilizzo di cavi con connettore *jack* 3.5mm ad un'estremità e connettore a banana nell'altra, possiamo dunque inviare un controllo in voltaggio al sintetizzatore.

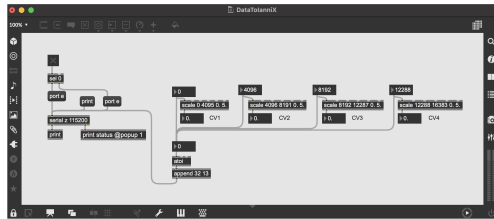


Figura 10. Patch di base per inviare quattro flussi di dati differenti tramite porta seriale.

3.2 Connessioni fisiche

Nella fig.11 vengono presentate schematicamente le connessioni fisiche tra *Arduino*, l'MCP4728 e il connettore *jack* femmina secondo questa codifica colore:

- colore nero : *ground*
- colore rosso : corrente
- colore verde : *clock* (SCL)
- colore blu : dati (SDA)

⁶ <https://ww1.microchip.com/downloads/en/devicedoc/22187e.pdf>

⁷ <https://docs.arduino.cc/resources/datasheets/A000066-datasheet.pdf>

⁸ Per consultare liberamente il codice contattare l'autore.

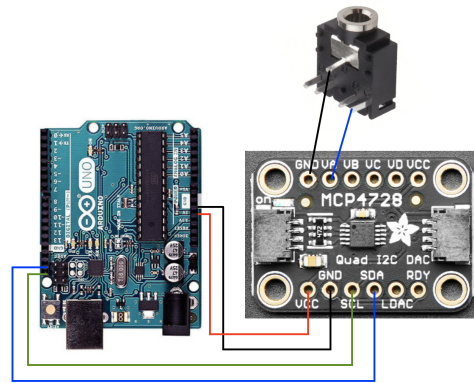


Figura 11. Connessioni fisiche

Tramite l'invio di un segnale di *clock* e di dati attraverso le porte SCL e SDA fornite da *Arduino*, stiamo dunque utilizzando il protocollo I2C [7].

3.3 Implementazione

Per un'ampliamento del numero di controlli in voltaggio possibili va aumentato il numero di MCP4728. Il protocollo I2C prevede l'utilizzo di fino a 8 *chip* sullo stesso *bus* dati con la condizione che ogni *chip* abbia un indirizzo diverso, ottenendo così la possibilità di avere fino a 32 diversi CV sullo stesso *bus*. La libreria MCP4728 non prevede però un cambiamento di indirizzo, così è stato utilizzato un codice che permetta di cambiare il numero dei *bit* che serve come identificativo del *chip* [8]⁹. Nella fig. 12 possiamo vedere un circuito con due *chip* riceventi dati sullo stesso *bus* I2C.

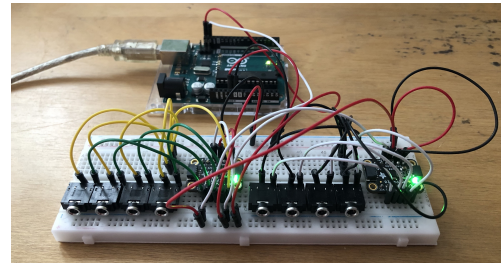


Figura 12. La *board* di *Arduino* collegata a due MCP4728 con 8 DAC.

4. PARTITURA GRAFICA E CV

Durante la fase di prototipazione per l'invio dei dati che, una volta trasformati, generano la tensione è stato utilizzato il *software Max/MSP*. Per la creazione di una partitura grafica e al contempo per la generazione di tensione in tempo reale è stato utilizzato il *software open source IanniX* [9, 10]¹⁰, che permette la generazione di messaggi tramite

⁹ https://github.com/jknipper/mcp4728_program_address

¹⁰ <https://www.iannix.org/en/>

l'utilizzo di cursori che leggono delle traiettorie disegnate dall'utente o l'utilizzo di *trigger*. Detto *software*, che prevede una comunicazione tramite porta seriale, ci permette di disegnare delle traiettorie la cui forma e lunghezza determina il valore che viene inviato alla *board*, attraverso la lettura di un cursore dedicato o di un *trigger*, per essere scalato e trasformato in CV.

4.1 Scalaggio flusso dati

Come indicato nel *Datasheet* del MCP4728 la sua risoluzione è di “ 2^{12} , meaning the DAC code ranges from 0 to 4095” ([5, p. 47]). Il controllo in voltaggio del sistema modulare utilizzato accetta un voltaggio che oscilla tra i valori di 0V e 5V. Pertanto *Arduino* riceverà da *Iannix*, su ogni canale preso in considerazione, un valore compreso tra 0 e 4095 che invierà a sua volta all'MCP4728 che si occuperà di inviare i dati ai canali corretti, fornendo così una tensione in uscita compresa tra 0V e 5V.

4.2 Programmazione IanniX

IanniX offre la possibilità di creare un proprio *mapping* dati grazie all'utilizzo di cursori che leggono temporalmente le traiettorie disegnate o tramite l'impiego di *trigger*, attivati sempre dalla lettura del cursore. Sfruttando il piano cartesiano abbiamo dunque la possibilità di disegnare delle traiettorie nello spazio alla quale possiamo attribuire dei valori che verranno in seguito letti dal cursore. Possiamo dunque definire il cursore come l'esecutore della partitura poiché nello svolgere il suo percorso sarà lui che invierà i dati alla *board*. *IanniX* ci permette dunque, tramite l'utilizzo del protocollo USB e sfruttando una comunicazione seriale, di inviare i dati letti dal cursore direttamente ad *Arduino*.

Il video esemplificativo¹¹ presenta un'esecuzione di una partitura dove sono stati controllati in voltaggio i seguenti parametri del sistema modulare, la cui *patch* è visibile in fig. 13:

- frequenza di 3 oscillatori;
- frequenza di taglio e Q di un filtro passa basso;
- *waveshape* di un oscillatore;
- *clock* di un *sequencer*;
- Parametro per il controllo della durata complessiva dei parametri *rise* e *fall* di un *envelope generator* che controlla l'apertura e la chiusura di un *Voltage Controlled Amplifier* (VCA).

Nell'esecuzione di tale *patch* non vi è alcuna interazione umana con la macchina, è quindi una pura esecuzione dei dati inviati al sintetizzatore attraverso la catena illustrata nell'articolo. Il video non ha nessuna ambizione artistica in quanto vuole unicamente illustrare il corretto funzionamento di quanto proposto.

¹¹ <https://www.youtube.com/watch?v=F7cQ0TdZs80&t=2s>

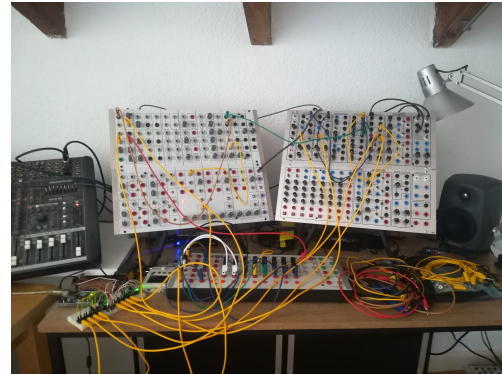


Figura 13. Sistema modulare che riceve gli otto controlli in voltaggio.

4.3 Conclusioni e sviluppi futuri

Dato il crescente interesse verso il *mondo modulare* degli ultimi anni, la creazione di sistemi e moduli negli ultimi anni è aumentata esponenzialmente favorendo la nascita di nuovi costruttori, siano essi grandi aziende o privati con una produzione limitata. Tale aumento dell'offerta in maniera congiunta alla natura stessa dei sistemi modulari rende ancora più difficile la determinazione di un linguaggio comune che permetta dunque una scrittura capace di comunicare in maniera efficace. Basti pensare quanto possa essere complesso e oneroso da un punto di vista economico ottenere lo stesso sistema per la riproduzione anche di un solo brano.

Il metodo proposto potrebbe avere una certa rilevanza per scopi didattici attraverso l'utilizzo di moduli fondamentali (oscillatori, filtri, amplificatori, generatori di inviluppo e randomici) e di partiture create appositamente data l'immediatezza e la semplicità con cui, tramite un sistema modulare, si possono apprendere concetti base della musica elettronica.

Tuttavia la totale automazione di un sistema modulare che permetta dunque una fedele riproducibilità dell'opera risulterebbe fine a sé stessa (oltre che limitata da un punto di vista tecnico) a meno che non sia riscontrabile una forte volontà, in questo senso, a livello compositivo.

Attraverso lo sviluppo di questo progetto, tenendo ben presente la storia e le capacità dello strumento, si è individuato dove l'utilizzo di questo metodo può assumere forme più interessanti e di rilievo artistico, ovvero in ambito della *performance* dal vivo, dove il musicista servendosi dell'*hardware* qui presentato in combinazione con il proprio ambiente esecutivo avrà la possibilità di creare delle vere e proprie automazioni, selezionabili ad esempio con un semplice *midi controller*, andando a creare un'estensione del sistema modulare che apre così la possibilità a forme di espressività sia più complesse sia più precise.

Per quanto riguarda l'aspetto grafico del metodo proposto si ritiene possa essere interessante, sempre a scopo didattico, sviluppare un codice grazie alla quale possano essere identificate determinate tipologie di comportamenti, magari attraverso la creazione di un software dedicato.

A livello performativo utilizzando la possibilità di automazione fornita, per l'indole improvvisativa dello strumento e di conseguenza per il tipo di interazione uomo/macchina, l'aspetto grafico non viene ritenuto rilevante in seguito allo studio condotto.

5. BIBLIOGRAFIA

- [1] A. Strange, *Electronic Music: Systems, Techniques, and Controls*. W. C. Brown Company, 2 ed., 1983.
- [2] H. Howe, *Buchla User Manual*. New York, s. d. ma 1960.
- [3] S. Ciani, *Report to National Endowment*, 1976. Conosciuto come *The Buchla Cookbook*, consultato 2024-05-05.
- [4] D. Zuegel and S. Ciani, "Suzanne ciani explains the composition of her sensory career," 1 2022. Consultato 2024-05-05.
- [5] *12-Bit, Quad Digital-to-Analog Converter with EEPROM Memory Datasheet*, 2010. <https://ww1.microchip.com/downloads/en/devicedoc/22187e.pdf>, consultato 2024-05-05.
- [6] *Arduino UNO R3 Product Reference Manual*, 2 2024. <https://docs.arduino.cc/resources/datasheets/A000066-datasheet.pdf>, consultato 2024-05-05.
- [7] *I2C-bus specification and user manual*, 10 2021. <https://www.nxp.com/docs/en/user-guide/UM10204.pdf>, consultato 2024-05-05.
- [8] J. Knipper, *MCP4728 program address*, 2020. https://github.com/jknipper/mcp4728_program_address, consultato 2024-05-05.
- [9] T. Coduys, G. Jacquemin, and M. Ranc, *IanniX*, 2010. <https://www.iannix.org/en/>, consultato 2024-05-05.
- [10] J. Scordato, *IanniX software documentation*, 2016. <https://www.iannix.org/download/documentation.pdf>, consultato 2024-05-05.

STUDIO DI UN'INTERFACCIA MUSICALE PER PAZIENTI CON DEMENZA DI ALZHEIMER LIEVE

Alexandrina Bargan
Dip. di Culture del Progetto
Università Iuav di Venezia
alexandrinabargan@icloud.com

Maximiliano Romero
Dip. di Culture del Progetto
Università Iuav di Venezia
mromero@iuav.it

Simone Spagnol
Dip. di Culture del Progetto
Università Iuav di Venezia
sspagnol@iuav.it

ABSTRACT

Le terapie non farmacologiche rappresentano allo stato attuale l'insieme di interventi che pongono in primo piano il benessere e la qualità di vita delle persone con un deterioramento cognitivo, in particolare quelle affette da demenza di Alzheimer. Tra gli interventi più comuni vi è la musicoterapia, i cui benefici, riscontrabili già a partire dalla fase iniziale della malattia, sono il mantenimento e la valorizzazione delle abilità ancora presenti. All'interno di questo contesto l'obiettivo dell'articolo è descrivere il processo che ha portato allo sviluppo (tuttora in corso) di un'interfaccia per incrementare le occasioni di attività musicale, ovvero al di fuori delle sessioni prettamente terapeutiche. Essa è dotata di una serie di maniglie interattive attraverso le quali i partecipanti, in presenza di professionisti o volontari, possono controllare la riproduzione di tracce musicali e paesaggi sonori con un'alta libertà gestuale.

1. INTRODUZIONE

Durante l'invecchiamento, lo spettro del declino cognitivo nelle persone anziane varia da quello che può essere definito come un normale declino dovuto all'invecchiamento fisiologico fino al deterioramento cognitivo lieve (*mild cognitive impairment*, MCI) [9]. Il progredire di tale condizione può infine condurre all'insorgere della demenza, principalmente nella forma del morbo di Alzheimer.

La demenza ha un impatto significativo non solo sugli individui ma anche sui caregiver, le famiglie, le comunità e le società. Le implicazioni della malattia neurodegenerativa sono molto differenziate da persona a persona. Vi sono un gruppo di otto sintomi comunemente presenti anche se non sempre compresenti: anosognosia, agnosia, afasia, aprassia, alterazioni della percezione, amnesia, apatia, deficit di attenzione.¹ Sulla base del peggioramento di questi sintomi è possibile distinguere tre fasi del disturbo: lo stadio lieve, modera-

to e avanzato. In Italia si contano circa 1.100.000 persone che soffrono di questa malattia, di cui circa 600.000 soffrono del morbo di Alzheimer.²

Intervenire fin dagli esordi è molto importante, sia con strumenti farmacologici che non. È possibile infatti stimolare le diverse capacità ancora presenti, un aspetto meglio noto come riserva cognitiva [2]. Dopo aver compreso quali aree del cervello siano più o meno compromesse dalla malattia, sarà l'operatore a proporre le terapie più appropriate. Nel caso di quelle non farmacologiche, l'obiettivo non è la guarigione o il ripristino della funzione perduta, bensì il rallentamento della velocità dello sviluppo della patologia e del livello di compromissione delle relative abilità individuali, ovvero la promozione di strategie di compenso. Si passa così "dalla cura della malattia alla cura della persona", ovvero il fine ultimo di questi interventi che pongono al centro anche storia, preferenze, abitudini e hobby del singolo [19].

La musicoterapia ne è un esempio emergente, grazie anche al crescente interesse all'interno della comunità scientifica [5,12,13,14,19]. Si evidenzia però che la tendenza è maggiormente accentuata in paesi come Olanda e Inghilterra. In Italia questa disciplina è regolamentata dalla Normativa UNI 11592:2015 per le Arti e Terapie, ma non vi è un albo di musicoterapisti. Vi sono tuttavia alcune associazioni come l'AIM (Associazione Italiana professionisti della Musicoterapia)³ che riuniscono i vari professionisti.

Il musicoterapista mette in pratica il programma terapeutico avendo come obiettivo quello di stabilire un rapporto terapeutico con i pazienti, affinché si creino le condizioni per una libera espressione attraverso la musica. Ogni produzione da parte del paziente presenta delle caratteristiche che il musicoterapista non solo sa interpretare, ma usa le conoscenze acquisite per adattare, seduta dopo seduta, l'intervento nella maniera più efficace possibile [23]. L'influenza della musica nelle sue dimensioni principali (come timbro, ritmo, melodia e armonia) sulle persone affette da demenza è significativa fin dall'esordio del disturbo, poiché offre una modalità di espressione attraverso una comunicazione non verbale, aspetto che diventa più rilevante con l'avanzare della malattia [4,13]. Il coinvolgimento significativo

Copyright: © 2024 Bargan et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ <https://ontariocaregiver.ca>

² <https://www.epicentro.iss.it>

³ <https://www.aim-musicoterapia.it>

della persona, che possa suscitare gli interessi e le abilità personali, deve comunque essere adeguatamente dimensionato in modo tale da non incontrare difficoltà; il rischio è una mancata aderenza alla terapia.

Durante la ricerca illustrata in questo articolo è emerso che spesso i Centri Sollievo (centri assistenziali per persone con diagnosi di decadimento cognitivo, nati all'interno della regione Veneto nel 2013) hanno una ridotta disponibilità di terapisti. Questo comporta che alcune attività dei pazienti non siano supervisionate. Anche se, nel caso delle attività basate sulla musica [18], solo quelle supervisionate da musicoterapisti possono essere considerate musicoterapia, sono emerse molte potenzialità delle attività musicali realizzate in autonomia. Nonostante questo, purtroppo oggi nei Centri Sollievo si offre solo circa un'ora di terapia alla settimana. Condizione che suggerisce una necessità, e quindi opportunità, per lo sviluppo di nuovi dispositivi, interfacce e/o altre soluzioni che possano stimolare l'inclusione dell'uso attivo della musica nella quotidianità dei pazienti.

Questo articolo parte dall'analisi dello stato dell'arte che ha permesso di comprendere le opportunità progettuali (Sezione 2). A queste sono state integrate le osservazioni ottenute in seguito ad alcune interviste semi-strutturate con musicoterapisti (Sezione 3). Da questa fase di ricerca si arriva alla successiva fase di progettazione (Sezione 4), partendo dall'esplorazione del concept, per poi passare allo sviluppo vero e proprio e alla prototipazione dell'interfaccia proposta. Questa iniziale esplorazione progettuale mette in luce nuove opportunità di partecipazione attiva nell'ottica della composizione musicale con la mediazione di uno strumento tecnologico. L'articolo si chiude con una serie di osservazioni utili per l'ulteriore sviluppo del progetto (Sezione 5).

2. STATO DELL'ARTE

Una ricerca degli studi presenti su ACM Digital Library (con le seguenti parole chiave: *Music therapy AND Alzheimer OR Dementia; Soundscape AND Dementia; Music AND Dementia; Music activity AND Dementia*), ampliata dalla consultazione in rete tramite motori di ricerca generalisti, ha permesso di selezionare una serie di dispositivi musicali particolarmente attinenti al tema qui affrontato. I progetti risultanti da questa ricerca sono stati catalogati attraverso schede di prodotto (si veda in Fig. 1 un esempio) e analizzati secondo una serie di parametri d'interesse come: livello di demenza dell'utente target (lieve, medio, avanzato), livello di sviluppo del progetto (concept, prototipo, prodotto di serie), numero di partecipanti coinvolti, grado di intervento del terapeuta e/o caregiver, caratteristiche sonore e musicali, tipologia di interazione, scopo dell'interazione, altri stimoli sensoriali, materiali, componenti elettroniche, stato di sviluppo del prodotto. In Tabella 1 sono elencati i dieci prodotti analizzati con il relativo stato di sviluppo.

| Tumbler: Interactive Soundscape Player | |
|--|--|
| Fonte | https://maartenhouben.be/tumbler |
| Contesto | Domestico |
| Target | Demenza (da lieve a media) |
| N. partecipanti | Pazienti e caregiver |
| Intervento terapeuta e/o caregiver | Interazione condivisa |
| Obiettivi attività | Stimolare conversazioni, emozioni e attività ricreativa. |
| Caratteristiche suono/musica | Suoni familiari e significativi nella quotidianità, registrati da pazienti e caregiver; racconti |
| Tipologia di interazione | Scuotere, ribaltare, ruotare |
| Scopo interazione | Comporre soundscapes attraverso l'esplorazione dei suoni |
| Altri stimoli sensoriali | Luminosi |
| Materiali | Stampa 3D in ABS, MDF |
| Stato | Prototipo funzionante |
| Componenti | Accelerometro, giroscopio, speaker, batteria (12h autonomia); base ricaricabile con ghiera per regolare volume, due bottoni per registrare storie attraverso il dispositivo, LED. |
| Considerazioni soggettive | La personalizzazione delle tracce audio permette anche di creare momenti durante i quali sia pazienti che caregiver hanno la possibilità di spostare l'attenzione sui suoni che solitamente passano inosservati. Si crea in questo modo una connessione affettiva già dal momento della selezione e registrazione: è un atto di condivisione. Questo è poi arricchito dalla riproduzione e combinazione su più livelli durante l'utilizzo del prodotto. L'esplorazione limita la possibilità di sbagliare interazioni: ogni movimento rappresenta una scoperta e stimola a provare varie combinazioni. |

Figura 1. Esempio di scheda di prodotto.

| Nome del progetto | Stato di sviluppo |
|-------------------------|---------------------------|
| MMM Duet System [3] | Prodotto open source |
| Dementia Soundboard [6] | Prodotto di ricerca |
| Tumbler [8] | Prototipo funzionante |
| Tune [22] | Prodotto open source |
| SAM [7] | Prodotto di ricerca |
| Memory box [15] | Prodotto commercializzato |
| SYNE [1] | Prototipo non funzionante |
| Sentic [21] | Prototipo non funzionante |
| Ichó [24] | Prodotto commercializzato |
| Sound of memory [11] | Prototipo funzionante |

Tabella 1. I progetti analizzati.

Per definire quali potrebbero essere eventuali mancanze nel mercato, sono state create delle mappe di posizionamento dei prodotti e/o progetti analizzati, riportate in Fig. 2 e Fig. 3. Esse sono costituite da assi cartesiani che rappresentano qualità opposte e permettono di evidenziare opportunità di sviluppo allo stato attuale. Il primo grafico è costituito da: riproduzione vs. composizione di tracce musicali (asse x) e attività individuale vs. attività di gruppo o in coppia (asse y), mentre il secondo da: difficoltà fissa dell'attività vs. difficoltà variabile (asse x) e stadio avanzato vs. stadio lieve di demenza (asse y).

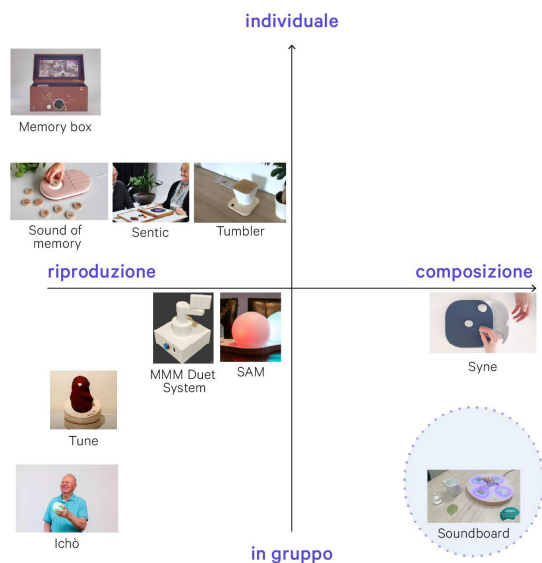


Figura 2. Mappa di posizionamento: riproduzione vs. composizione di tracce musicali (asse x); attività individuale vs. attività di gruppo o in coppia (asse y).

Nel primo caso (Fig. 2) è stato possibile riflettere sulla relazione tra la finalità dell'interazione e la modalità attraverso la quale essa potesse essere raggiunta. La composizione è un aspetto che permette maggior coinvolgimento, dal momento che rende l'attività più personalizzabile, con la complessità che può essere definita a seconda dello stadio della malattia neurodegenerativa [21]. Prevedere, per esempio, una melodia in sottofondo (analogamente alla tecnica del *grounding* utilizzata in musicoterapia) [23] alla quale si può contribuire individualmente interagendo e producendo ulteriori effetti sonori in un'ottica di composizione personale, potrebbe avere svariati effetti positivi. Innanzitutto, verrebbe meno l'ansia da prestazione dal momento che esiste un costante supporto alla base. In secondo luogo, in un'ottica di interazione di gruppo vi è un minor risalto nei confronti di chi non partecipa in egual misura all'attività, senza però lasciare la percezione di esclusione. Poche sono le soluzioni che prevedono un'interazione di gruppo, ma dal confronto con i musicoterapisti si è compreso come la connessione sociale che emerge attraverso il coinvolgimento di più persone

sia fondamentale anche per creare una rete di supporto reciproca tra i partecipanti all'attività [17].

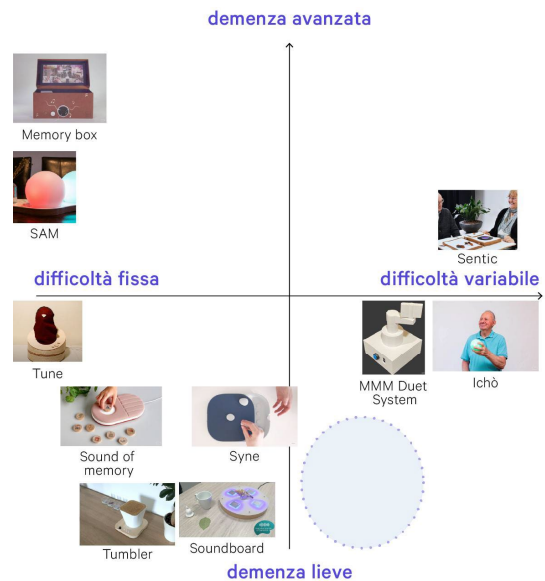


Figura 3. Mappa di posizionamento: difficoltà fissa dell'attività vs. difficoltà variabile (asse x); stadio avanzato vs. stadio lieve di demenza (asse y).

Nel secondo caso (Fig. 3), la difficoltà di interazione dipende sia dallo stadio della demenza che dalla percezione soggettiva delle proprie abilità e potenzialità [10,21]. Quest'ultima, ad esempio, si evidenzia di fronte ad un'offerta strumentale durante la terapia: utilizzare strumenti musicali che richiedano abilità tecniche può scoraggiare i partecipanti, lasciando una prima impressione negativa. Altre componenti che determinano la difficoltà di interazione possono essere la necessaria sequenzialità delle azioni per portare a termine un compito, particolari processi cognitivi coinvolti come la memoria a breve termine, l'attenzione selettiva, ecc. È preferibile prediligere interazioni di tipo esplorativo, che non portino alla possibilità di errore [10].

Dall'analisi di queste mappe di posizionamento è stata quindi rilevata la presenza di un'opportunità di sviluppo di un'interfaccia caratterizzata da una partecipazione attiva durante attività musicali tramite un approccio di tipo compositivo, ovvero una soluzione integrata all'interno di un contesto di gruppo fino a un massimo di sei persone. Questo numero deriva dalle seguenti riflessioni: garantire un numero limitato di tracce e dunque una maggiore chiarezza nella composizione e nella fruizione dell'interfaccia; agevolare la connessione sociale del gruppo [16].

3. INTERVISTE

Al fine di comprendere al meglio i requisiti dell'interfaccia sono state condotte tre interviste semi-strutturate, inizialmente con una musicoterapista che opera presso i Centri Sollievo tramite l'associazione di

volontariato AMAd della Pedemontana Veneta, e in seguito con altri due professionisti associati all'elenco Nazionale di Musicoterapisti, contattati al fine di arricchire ulteriormente le testimonianze. Il materiale comunicativo (render, immagini, video) utilizzato nelle interviste è risultato molto utile per coinvolgere i professionisti e spiegare loro in che modo sia possibile per un designer intervenire sulle necessità individuate. In questa fase non è stato previsto un coinvolgimento attivo di pazienti.

A livello generale è emerso un elemento che ha messo d'accordo i musicoterapisti: l'osservazione di un interesse importante, da parte dei pazienti, nei confronti della musica anche oltre le sessioni di musicoterapia e dunque la necessità di poter prevedere soluzioni adatte ad attività non supervisionate. Inoltre, la riproduzione di paesaggi sonori è spesso utilizzata come alternativa all'impiego di strumenti musicali: essi vengono proposti in modo tale da stimolare l'immaginazione e la riflessione sullo stato emotivo dei partecipanti, aspetto che secondo gli intervistati riveste una particolare importanza in relazione alle modalità espressive e comunicative dei pazienti.

Sono state quindi avanzate alcune idee su dispositivi e/o interfacce differenti da quelle che abitualmente vengono utilizzate come mezzi per svolgere attività musicali. In risposta, i professionisti hanno espresso entusiasmo di fronte alle novità, sia dei casi studio⁴ che sono stati loro presentati come esempi, sia del progetto di interfaccia qui proposto. Di conseguenza, la disponibilità di uno strumento per generiche attività musicali fruibile all'interno dei Centri Solievo o altri contesti è stata ritenuta il punto di partenza del progetto per lo sviluppo dell'interfaccia.



Figura 4. Esplorazione del contesto.

⁴ <https://omniphone776792973.wordpress.com/>

4. STUDIO DEL DISPOSITIVO

4.1 Brief

I risultati provenienti dall'analisi dello stato dell'arte e dalle interviste sono stati riassunti attraverso la creazione di un brief, ovvero l'insieme degli obiettivi da seguire durante la fase di ideazione di un progetto. Nello specifico, l'obiettivo principale del progetto è la realizzazione di un dispositivo per la composizione personalizzata di melodie e paesaggi sonori all'interno di un'attività di gruppo che tenga conto delle abilità residue dei partecipanti. Gli obiettivi secondari del progetto sono i seguenti: (1) evitare di banalizzare l'attività, ovvero semplificarla in modo eccessivo; (2) rendere facilmente fruibile l'interfaccia anche al di fuori delle sessioni terapeutiche; (3) limitare la possibilità di commettere errori, prevedendo modalità di fruizione chiare e attuabili in pochi step e quindi riducendo il carico cognitivo; (4) garantire la libertà espressiva in forma gestuale.

4.2 Concept

Da questo momento in poi è seguita l'esplorazione delle possibili soluzioni attraverso degli schizzi, partendo dall'esplorazione delle gestualità presenti durante le sessioni di musicoterapia (si veda la Fig. 4). L'obiettivo era comprendere aspetti che nella ricerca bibliografica non potevano emergere come la componente sociale, la prossemica e i gesti. I momenti di improvvisazione sono caratterizzati da un'ampia libertà espressiva che si esprime attraverso i movimenti, i quali vengono osservati dagli altri componenti del gruppo e possono diventare oggetto di imitazione e risposta [16] attraverso il linguaggio della musica. Sentirsi parte di un gruppo, di un insieme, pur mantenendo il proprio ruolo durante l'attività, è un aspetto molto significativo per le persone con una malattia neurodegenerativa. D'altronde, la consapevolezza di poter contribuire alla costruzione di un significato sia personale che sociale (all'interno del gruppo) è uno dei maggiori potenziali della musicoterapia [13].

L'opzione che meglio rispondeva ai criteri stabiliti inizialmente è stata quella di un'interfaccia la cui configurazione potesse comunicare, sia a livello fisico/spaziale che a livello percettivo, una collaborazione condivisa. In questo modo vi è una somiglianza con il setting delle sessioni di musicoterapia di gruppo (si veda di nuovo la Fig. 4). Si tratta dunque di un dispositivo centrale dal quale si diramano varie maniglie, una per ciascun utente. Dal punto di vista interattivo vi è un'ampia libertà gestuale nel movimento, realizzabile attraverso l'utilizzo delle mani grazie a sensori di movimento integrati all'interno delle maniglie in grado di raccogliere dati sulla configurazione spaziale da convertire in feedback sonoro. L'interfaccia permette la composizione di brani musicali o di paesaggi sonori a partire da singole tracce, ciascuna controllata da un singolo utente a seconda di varie possibili modalità e gestualità.

4.3 Sviluppo

A partire dalle ultime riflessioni relative al concept, il processo di sviluppo ha seguito inizialmente una fase di divergenza, caratterizzata da un processo di studio principalmente formale attraverso mockup e/o prototipi grezzi, per esplorare le varie opzioni percorribili. In seguito, nella fase di convergenza, sono stati riassunti gli aspetti che rispondevano meglio al brief definito in precedenza per arrivare a una soluzione progettuale finale. Durante lo sviluppo sono emerse due diverse versioni dell'interfaccia. La prima era caratterizzata da un dispositivo dalla forma compatta e circolare. Presentava sei maniglie di forma triangolare e disposte a ragiera, ciascuna collegata alla componente centrale tramite dei cavi nascosti all'interno del corpo principale.

A questo punto è emersa una riflessione su alcune criticità. Innanzitutto, la mancanza di un chiaro invito all'interazione e un volume delle maniglie eccessivo in relazione agli ingombri della tecnologia al loro interno. Inoltre, dal punto di vista interattivo, era necessario comprendere quale potesse essere la modalità di creazione della musica a partire da più tracce: esse potevano essere già preassegnate a ciascun utente oppure selezionate individualmente a seconda delle preferenze. Quest'ultima modalità è risultata la più coerente con quanto emerso nella rappresentazione dei prodotti già esistenti sulla mappa di posizionamento (si veda la Fig. 2), in quanto lasciare maggior controllo compositivo ai partecipanti è un'opzione tuttora poco esplorata.

Con la seconda (e definitiva) versione, lo sviluppo verticale della forma ha permesso di esaltare maggiormente la distinzione tra le due zone dell'interfaccia musicale: quella superiore, ovvero una piattaforma in legno che diventa base di appoggio delle maniglie, e quella inferiore, dotata di appositi alloggi per i 6 moduli che rappresentano le diverse tracce. La forma delle maniglie è stata modificata con il fine di ottimizzare l'impiego di materiale, pur mantenendo una presa ergonomica. Infine, un tasto per la regolazione del volume è stato implementato su ciascuna di esse.

Oltre allo sviluppo di modelli fisici per studiare ingombri e finalizzare la forma, è stata affiancata anche una fase di prototipazione attraverso l'utilizzo di Pure Data (Pd) e Arduino. Il primo è un ambiente e linguaggio di programmazione grafico open source, impiegato per la produzione e la performance di musica elettronica/elettroacustica e per la prototipazione. Il secondo, anch'esso con licenza libera, integra sia una parte software che hardware (microcontrollore) per interagire con il mondo reale tramite sensori ed attuatori.

Al fine di valutare il concept selezionato, questa volta dal punto di vista delle interazioni, si è deciso di testare diversi sensori interfacciati ad Arduino e Pd (si veda la Fig. 5). Dal momento che il movimento rotatorio di ciascuna maniglia deve poter controllare un suono o una traccia, si è deciso che il giroscopio fosse il sensore più adatto allo scopo. Il motivo della scelta risponde direttamente al brief, ovvero non limitare gli utenti

alla sola riproduzione, ma includere anche la possibilità di controllare alcuni parametri (frequenza, tempo) e di applicare effetti audio e filtri digitali in modo tale da stimolare la curiosità dei partecipanti a sperimentare con l'audio. Ci si è inoltre concentrati sui feedback, essenziali quando si interagisce con un'interfaccia per poter comprendere se le proprie azioni producano un effetto visibile in modo tale da poter adeguare nuovamente la propria risposta [20]. Pertanto, si è ipotizzato di includere, oltre ai feedback sonori e luminosi (striscia LED), un motore vibrotattile per segnalare la possibilità di iniziare a muovere le maniglie e, durante l'attività, far percepire che l'oggetto in mano è funzionante e risponde ai gesti dell'utente.

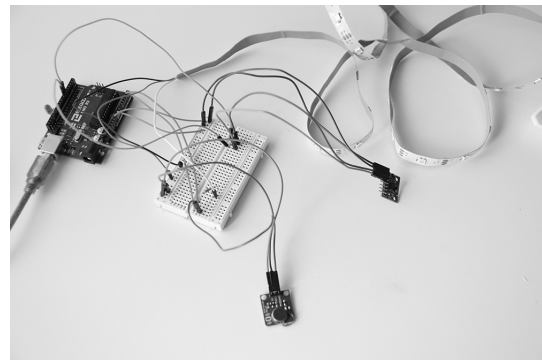


Figura 5. Prototipazione con giroscopio, motore vibrotattile e striscia LED.

L'interfaccia che proponiamo, di cui riportiamo un render 3D in Fig. 6 e un esploso in Fig. 7, è dunque composta da: una base in legno nella zona superiore con sei alloggi per le maniglie e una griglia centrale per l'uscita audio rivolta verso l'alto; un corpo centrale cilindrico, all'interno del quale è collocato un sistema di cassa a due vie, con woofer e tweeter coassiali; una base contenente sei tasti per attivare diverse tipologie di effetti o filtri digitali, e sei moduli circolari differenziati visivamente da simboli posti sul fronte e tag RFID sul retro. Ciascun modulo rappresenta una traccia differente che ogni partecipante può selezionare o scambiare durante l'attività posizionandolo nell'apposito alloggiamento. Qui, un lettore RFID potrà riconoscere il tag.

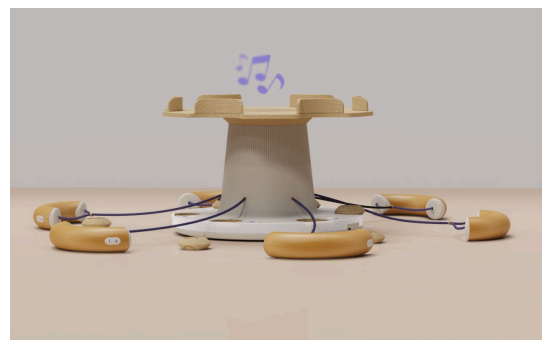


Figura 6. Render del dispositivo.

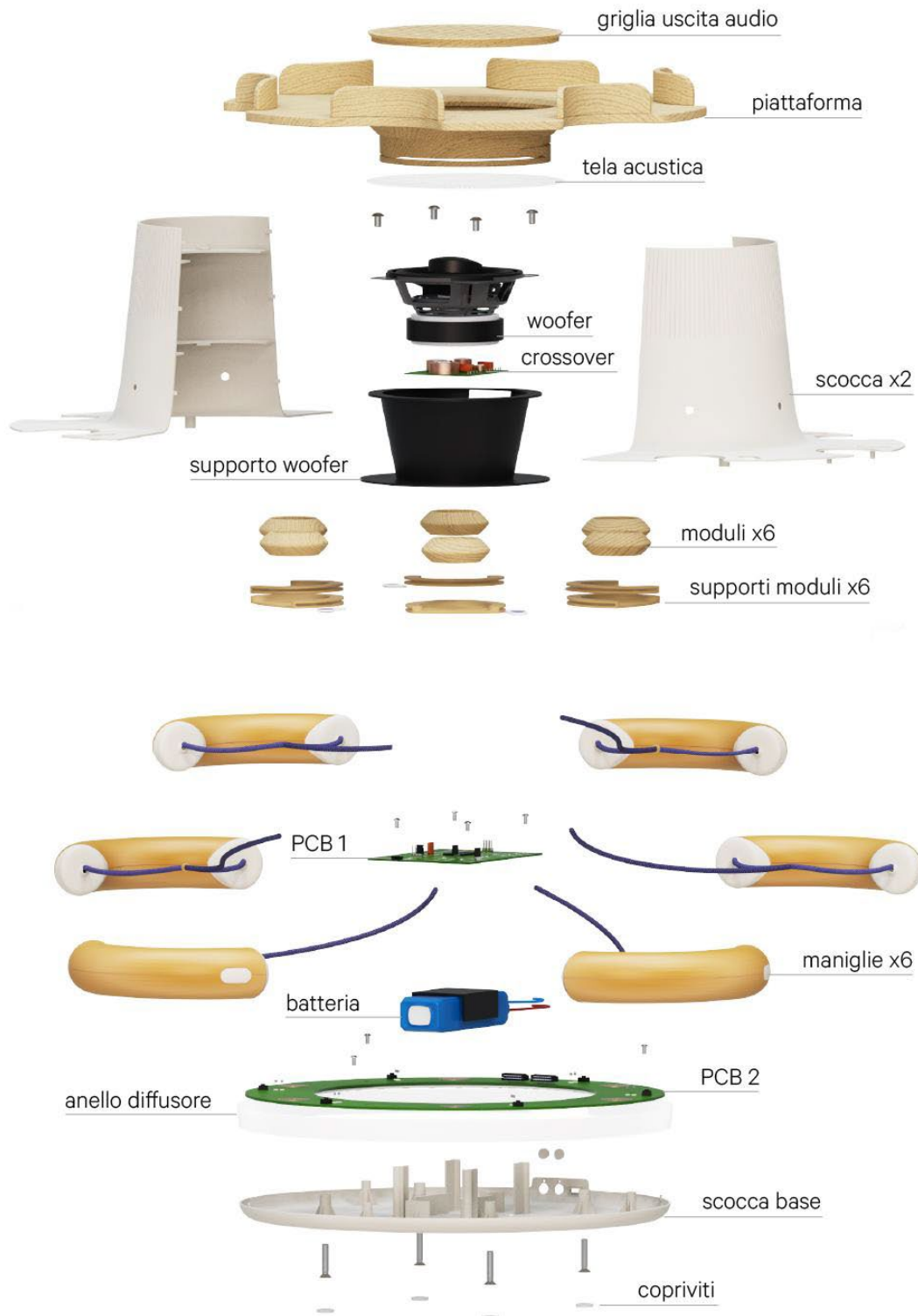


Figura 7. Esploso del dispositivo.

I feedback utilizzati sono divisi in tre tipologie: sonori, luminosi e vibrotattili. I primi due sono percepibili all'intero gruppo (compreso il professionista che ha proposto e segue l'attività), mentre il terzo è rivolto al singolo utente

5. CONCLUSIONI E SVILUPPI FUTURI

Questo articolo descrive un processo non ancora concluso, ma ha lo scopo principale di illustrare l'approccio sistemico del design del prodotto nel contesto del benessere della persona. Possiamo affermare che dall'analisi delle mappe di posizionamento dei progetti e prodotti esistenti abbiamo rilevato la presenza di spazi di progettazione significativi. Essi hanno rappresentato un'opportunità di sviluppo di un prodotto basato su un'attività di gruppo, con la possibilità di svolgere attività musicali non supervisionate dal terapeuta. Quest'ultima condizione apre le porte a un'interessante esplorazione verso alternative alla musicoterapia in senso stretto, ovvero attività basate sulla musica che non hanno necessariamente obiettivi terapeutici.

Ovviamente, in assenza di un prototipo completamente funzionante non è ancora possibile affermare con certezza i possibili benefici sulla sfera emotivo-comportamentale dei pazienti. Pertanto, come prossimi sviluppi, sarà innanzitutto necessario sviluppare il prototipo per condurre sperimentazioni a lungo termine sul prodotto in sé, studiando ad esempio la percezione da parte degli utenti di questo nuovo strumento, la forma ed ergonomia delle maniglie, ecc. Dal punto di vista delle interazioni si prevede di valutare la comprensibilità dei feedback sonori e vibrotattili, oltre al grado di interesse e coinvolgimento da parte degli utenti nel poter controllare i suoni attraverso una modalità espressiva data dal movimento tridimensionale delle maniglie.

Infine, ma non meno importante, dovranno essere coinvolte le figure professionali durante ogni fase dello sviluppo, in quanto riteniamo il contributo degli stakeholder fondamentale per la buona riuscita del progetto stesso.

6. BIBLIOGRAFIA

- [1] S. Altun: "SYNE: Audio Reminiscence Therapy", <https://www.sinanaltundesign.com/project/syne-audio-reminiscence-therapy> (ultimo accesso: 15/07/2024).
- [2] M. Boccardi: *La riabilitazione nella demenza grave. Manuale pratico per operatori e caregiver*, Erickson, 2007.
- [3] J. Christensen, S. Kauenhofen, J. D. Loehr, J. Lang, S. Peacock, and J. J. Nicol: "MMM Duet System: New accessible musical technology for people living with dementia," in *Proceedings of the 18th International Audio Mostly Conference*, pp. 38–44, Edinburgh, 2023.
- [4] R. Dowlen, J. Keady, C. Milligan, C. Swarbrick, N. Ponsillo, L. Geddes, and B. Riley: "In the moment with music: An exploration of the embodied and sensory experiences of people living with dementia during improvised music-making," *Ageing and Society*, vol. 42, no. 11, pp. 2642–2664, 2022.
- [5] M. Gómez-Gallego, J. C. Gómez-Gallego, M. Gallego-Mellado, and J. García-García: "Comparative Efficacy of Active Group Music Intervention versus Group Music Listening in Alzheimer's Disease," *International Journal of Environmental Research and Public Health*, vol. 18, no. 15, 2021.
- [6] M. Houben, R. Brankaert, S. Bakker, G. Kenning, I. Bongers, and B. Eggen: "Foregrounding Everyday Sounds in Dementia," in *Proceedings of the 2019 Conference on Designing Interactive Systems*, pp. 71–83, San Diego, 2019.
- [7] M. Houben, R. Brankaert, E. Dhaeze, G. Kenning, I. Bongers, and B. Eggen: "Enriching Everyday Lived Experiences in Dementia Care," in *16th International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 1–13, Daejeon, 2022.
- [8] M. Houben, R. Brankaert, G. Kenning, I. Bongers, and B. Eggen: "Designing for Everyday Sounds at Home with People with Dementia and their Partners," in *CHI Conference on Human Factors in Computing Systems*, pp. 1–15, New Orleans, 2022.
- [9] S. Jongsiriyanyong and P. Limpawattana: "Mild Cognitive Impairment in Clinical Practice: A Review Article," *American Journal of Alzheimer's Disease & Other Dementias*, vol. 33, no. 8, pp. 500–507, 2018.
- [10] A. Lazar, C. Edasis, and A. M. Piper: "A Critical Lens on Dementia and Design in HCI," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 2175–2188, Denver, 2017.
- [11] S. Lundberg: "Sound of memory", <https://sandrallundberg.com/sound-of-memory> (ultimo accesso: 02/02/2024).
- [12] R. A. MacDonald and G. B. Wilson: "Musical improvisation and health: A review," *Psychology of Well-Being*, vol. 4, no. 1, 2014.
- [13] G. Manarolo: *Musica e musicoterapia nella terza età*, FrancoAngeli, 2022.
- [14] A. M. Matziorinis and S. Koelsch: "The promise of music therapy for Alzheimer's disease: A review," *Annals of the New York Academy of Sciences*, vol. 1516, no. 1, pp. 11–17, 2022.
- [15] C. Meineck: "Music memory box", <https://studiomeineck.com/music-memory-box/> (ultimo accesso: 15/07/2024).

- [16] K. Morrissey, G. Wood, D. Green, N. Pantidi, and J. McCarthy: “«I’m a rambler, I’m a gambler, I’m a long way from home»: The Place of Props, Music, and Design in Dementia Care,” in *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*, pp. 1008–1020, Brisbane, 2016.
- [17] H. M. O’Rourke, T. Hopper, L. Bartel, M. Archibald, M. Hoben, J. Swindle, D. Thibault, and T. Whynot: “Music Connects Us: Development of a Music-Based Group Activity Intervention to Engage People Living with Dementia and Address Loneliness,” *Healthcare*, vol. 9, no. 5, 2021.
- [18] A. Raglio, S. Filippi, D. Bellandi, and M. Strambadiale: “Global music approach to persons with dementia: Evidence and practice,” *Clinical interventions in aging*, vol. 9, pp. 1669–1676, 2014.
- [19] A. Raglio, G. Manarolo, and D. Villani: *Musicoterapia e malattia di Alzheimer: proposte applicative e ipotesi di ricerca*, Edizioni Cosmopolis, 2001.
- [20] K. J. Schelle, C. Gomez Naranjo, M. Ten Bhömer, O. Tomico, and S. Wensveen: “Tactile Dialogues: Personalization of Vibrotactile Behavior to Trigger Interpersonal Communication,” in *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction*, pp. 637–642, Stanford, 2015.
- [21] M. Thoolen, R. Brankaert, and Y. Lu: “Sentic: A Tailored Interface Design for People with Dementia to Access Music,” in *Designing Interactive Systems Conference 2019 Companion*, pp. 57–60, New York, 2019.
- [22] Università Iuav di Venezia: “Tune”, <https://sites.google.com/iuav.it/lab-prodotti-intelligenti/aa-20202021/tune> (ultimo accesso: 15/07/2024).
- [23] T. Wigram: *Improvisation: Methods and Techniques for Music Therapy Clinicians, Educators and Students*, Jessica Kingsley Publishers, 2004.
- [24] “Ichó”, <https://icho-systems.de/demenz> (ultimo accesso: 15/07/2024).

RIPROPOSIZIONE DIGITALE DI UNO STRUMENTO PIONIERISTICO. *DIGITAL MORPHOPHONE ENVIRONMENT*

Daniel Scorrane

Dipartimento di Musica Elettronica
Conservatorio di L'Aquila
daniel.scorrane@gmail.com

ABSTRACT

Il presente contributo introduce una riproposizione digitale del *morphophone*, sistema complesso di eco magnetofonica (*tape loop delay*) sviluppato negli anni Cinquanta all'interno dei laboratori del GRM (*Groupe de Recherches Musicales*) di Parigi. In particolare vengono esplicate le metodologie di analisi, progettazione e implementazione che sono alla base del *Digital Morphophone Environment*.

Attraverso un'analisi approfondita delle fonti storiche e della documentazione disponibile, comprendente una quantità limitata di materiale bibliografico e soprattutto immagini di repertorio, si sono modellati i processi essenziali del *morphophone* in ambiente di programmazione visuale MAX¹. Tra gli obiettivi che hanno condotto alla realizzazione di questo lavoro vi è da un lato l'opportunità di studiare e rendere disponibile un dispositivo ormai obsoleto e non più utilizzabile, dall'altro emergono ipotesi di nuovo utilizzo in ambito musicale e di ricerca.

1. INTRODUZIONE

L'avvento della registrazione basata su nastro magnetico che ha caratterizzato i primi anni Cinquanta del Novecento, ha portato allo sviluppo di sistemi di elaborazione del suono mirati alla generazione di eco e riverberazione, ottenuti mediante l'ausilio di anelli di nastro e molteplici testine di registrazione e riproduzione del segnale. Tali dispositivi hanno trovato largo impiego in applicazioni musicali e in contesti tecnici diversi. Si ricorda, ad esempio, il caso della *delay wheel* della Philips, sistema realizzato presso i laboratori di ricerca di Eindhoven a fini di correzione acustica degli ambienti adibiti alla fruizione musicale [1].

Le attività avviate da Pierre Schaeffer (1910-1995), in particolare gli esperimenti sulla musica concreta presso la radio di Parigi a fine anni Quaranta, risultano rilevanti per definire l'*iter* creativo di infrastrutture tecnologiche pensate come ausilio alla composizione, tra le quali il *morphophone*. L'introduzione del magnetofono nel 1950 al GRM (in quel momento ancora chiamato *Club d'Essay*) "apre

sviluppi importantissimi per l'emancipazione dell'arte musicale concreta, in particolare abilita ad intervenire sul suono con nuove metodologie che consentono di trasfigurarlo, di farlo divenire realmente un materiale modellabile" [2].

Intervenendo sul normale funzionamento dei magnetofoni, Schaeffer e Jacques Poullin (1920-2014) concepiscono nuovi strumenti di elaborazione del suono. Nel 1952 Poullin realizza il *tri-pistes*, antenato dei magnetofoni multicanale, basato sullo scorrimento sincrono di tre nastri magnetici. L'anno seguente viene realizzato il *phonogène*, che utilizza una piccola tastiera musicale per modulare la velocità di scorrimento e trasporre il segnale registrato sul magnetofono. Seguendo lo stesso principio Poullin realizza una variante, il *phonogène à coulisse* che differisce dal primo in quanto consente di effettuare trasposizioni continue del segnale, indipendenti dalla scala cromatica temperata. Nel 1953 sempre ad opera di Poullin, questa volta con l'apporto di Abraham Moles (1920-1992), viene realizzato il *morphophone* di cui ci occuperemo in questa sede² (Figura 1).

Almeno a partire dagli anni Novanta, vi sono state numerose ricerche volte alla modellazione digitale di sistemi audio analogici (*Virtual Analog*): dalla creazione di modelli di circuitazione di apparati audio generici, a dispositivi di sintesi ed elaborazione del suono più specifici [3, 4, 5, 6, 7, 8, 9, 10, 11, 12]. Benché più contenuta, risulta di particolare interesse in questa sede la letteratura relativa alla modellazione di sistemi di elaborazione magnetofonica [13, 14, 15].

Il lavoro qui proposto da un lato intende analizzare e virtualizzare i processi di uno strumento ormai obsoleto, dall'altro renderne possibile l'uso anche a fini di nuovo utilizzo in ambito musicale e di ricerca.

2. MORPHOPHONE

2.1 Panoramica

Il dispositivo viene concepito al fine di costruire forme sonore complesse mediante molteplice ripetizione del segnale a intervalli di tempo differenti. Ciò è reso possibile tramite un anello di nastro rotante e 12 testine magnetofoniche disposte intorno al meccanismo di rotazione: una testina di registrazione, 10 di riproduzione e una di cancellazione. La posizione delle testine di registrazione e cancel-

Copyright: ©2024 Daniel Scorrane et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

¹ Versione 8.6.2 (<https://cycling74.com/>).

² Il *phonogène* a tastiera fu brevettato da Schaeffer e realizzato in collaborazione con l'azienda *Tolana*. I prototipi del *phonogène à coulisse* e del *morphophone* furono prodotti in collaborazione con un'azienda esterna, denominata *S.A.R.E.G.*

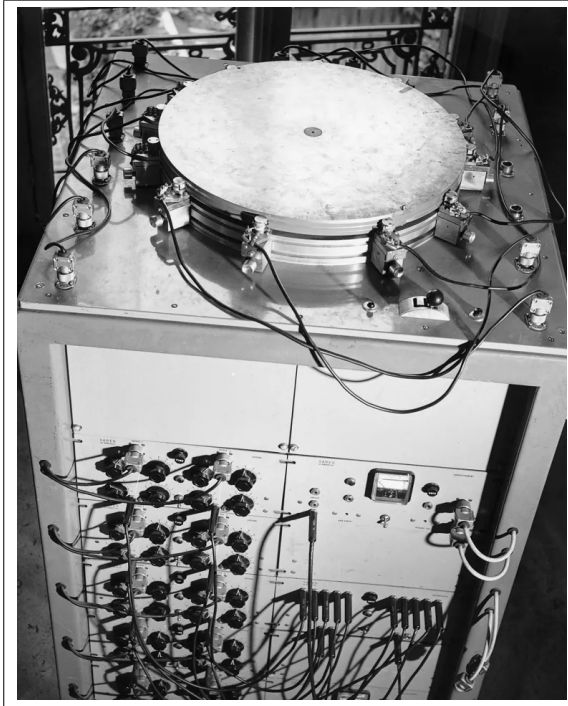


Figura 1. Panoramica del *Morphophone* [16].

lazione è fissa, mentre quella delle testine di riproduzione è variabile in modo da fornire un controllo sul tempo di ritardo della singola ripetizione, relativo alla distanza rispetto alla testina di registrazione. Inoltre, ogni testina di riproduzione invia segnale ad un proprio preamplificatore e un proprio filtro passa banda, consentendo di regolare in modo indipendente sia il livello che la risposta in frequenza di ogni singola ripetizione. I segnali preamplificati vengono poi miscelati e amplificati in uscita.

Il massimo tempo di ritardo ottenibile dipende da due fattori: il diametro complessivo della ruota e la velocità di rotazione. Le fonti a disposizione indicano che il diametro è di 50 cm circa [17, 18]. Le velocità di rotazione disponibili sono di 19, 38 e 76 cm/s. Quindi, con una bobina di diametro uguale a 50 cm (circonferenza: 157.8 cm) e una velocità di rotazione di 19 cm/s, il massimo tempo di ritardo ottenibile è di 8.3 secondi³.

Il segnale delle testine di riproduzione può essere riportato all'ingresso in modo da instaurare un processo di retroazione (*feedback*) utile per reiterare le ripetizioni. Il tempo di estinzione del feedback dipende dalle impostazioni dei preamplificatori e dei filtri passa banda delle testine di riproduzione. La testina di cancellazione può essere rimossa consentendo la sovraincisione del segnale (*overdub*), utile sia per introdurre ulteriori materiali sullo stesso anello di nastro, sia per ottenere *texture* musicali molto complesse.

³ In realtà, nel calcolo andrebbe considerata anche la porzione di nastro scoperta presente tra la testina di cancellazione e quella di scrittura (3 cm circa [18]) che si traduce in una riduzione della circonferenza complessiva.

2.2 Modalità di elaborazione

Il *morphophone* poteva prestarsi a diverse modalità di impiego: per produrre riverberazione artificiale, regolando tempi di ritardo e ampiezze delle ripetizioni in modo da simulare le riflessioni naturali di una sorgente in uno spazio chiuso; o come *multi-tap delay*, nella ripetizione di un dato evento sonoro con tempi di ritardo e timbri differenti. L'utilizzo del feedback in correlazione con eventuale overdub poteva dar luogo a materiali sonori densi e complessi, con fenomeni di accumulazione del segnale. Secondo le fonti a disposizione, il *morphophone* era stato pensato da Moles, Schaeffer e Poullin come un mezzo per modificare lo spettro di un suono prolungato; lo si evince, per esempio, in un articolo di Poullin che spiega: “un elemento omogeneo di dinamica stabile che passa più volte davanti alle testine di lettura potrà subire una modulazione dinamica e secondariamente, se si desidera, delle modifiche spettrali” [19]. Moles a sua volta lo descrive come “un apparecchio che ha lo scopo di fare evolvere lo spettro nel corso della durata, cioè di creare un oggetto sonoro finito, distinto [...] dalla forma sonora originale, dotato di un particolare *attacco*, un *corpo* e un' *estinzione*...”⁴ [20].

Non ci sono fonti certe sull'impiego del *morphophone* nella realizzazione di opere musicali, fatto probabilmente giustificato dalla difficoltà operativa del dispositivo: secondo Daniel Teruggi, infatti, “[t]his was a very complex and rich machine, however, it was almost impossibile to make it work. It was very difficult to stick a looped tape on the turning disk and since the magnetic heads had to be in close contact with the tape after a few turns, a head would detach the tape. Some experiments were done on the machine but it was never really musically used” [17]. Secondo alcuni autori [21, 22] sarebbe stato utilizzato, tra gli altri, da Iannis Xenakis in alcuni suoi lavori al GRM. Anche in questo caso le evidenze storiche sono incerte benché recenti esperimenti forniscano indicazioni positive in tal senso [23].

3. DIGITAL MORPHOPHONE ENVIRONMENT

Allo stato dell'arte non sembrano esistere riproposizioni che modellino il *morphophone* in ambito digitale. Vi sono applicazioni basate sul suo principio di funzionamento ma ripensate in maniera più o meno indipendente. Si citano a tal proposito *Delays*⁵, plugin incluso nei famosi *GRM Tools* [24] e *Morphophone*⁶, realizzato in ambiente REAKTOR⁷ da Jonathan Tremblay. In ogni caso, dopo una loro attenta analisi, nessuna di queste è sembrata direttamente utile per la modellazione qui proposta.

DME prevede due modalità di funzionamento, ‘filologica’ ed ‘estesa’. Si è ritenuto utile distinguere queste due modalità per ottenere da un lato una ricostruzione aderente

⁴ T.d.A (Traduzioni dal francese a cura dell'autore del presente articolo).

⁵ <https://inagrm.com/en/store/product/3/classic>.

⁶ <https://www.native-instruments.com/en/reaktor-community/reaktor-user-library/entry/show/8381/>.

⁷ <https://www.native-instruments.com/en/products/komplete/synths/reaktor-6/>.

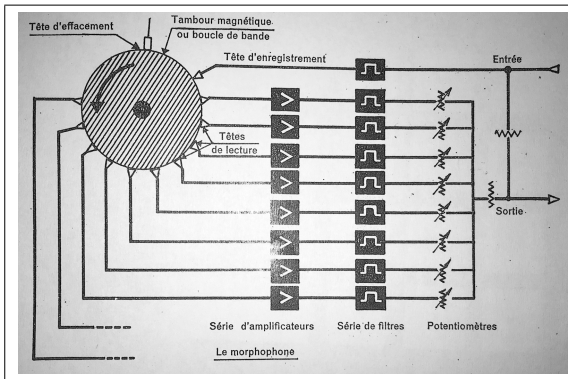


Figura 2. Diagramma di flusso del segnale nel *morphophone*.

all'originale⁸, dall'altro uno strumento capace di sfruttare risorse di calcolo e metodi di elaborazione attualmente disponibili per ampliarne le possibilità di intervento. Qui per semplicità considereremo solo la versione filologica.

Come accennato nell'introduzione, il lavoro di implementazione è stato preceduto e coadiuvato da un'analisi della documentazione disponibile, comprendente una quantità limitata di materiale bibliografico - con informazioni spesso contraddittorie - e soprattutto immagini di repertorio. Queste ultime, costituendosi come fonti primarie e dirette, si sono rivelate utili per comprendere alcune delle principali caratteristiche strutturali e di interfaccia offerte dal *morphophone*. Tra le fonti più importanti vi è un diagramma funzionale presente in un volume di raro reperimento di Abraham Moles [26] (vedi Figura 2)⁹.

3.1 Struttura e pannello di controllo del *morphophone*

La configurazione complessiva del *morphophone* rispecchia un approccio che può essere definito 'modulare'. Ogni modulo è messo in comunicazione con il resto del sistema attraverso cablaggi interni ed esterni. Quelli esterni consentono all'utilizzatore di adattare il sistema alle proprie esigenze, creando potenzialmente un elevato numero di diverse configurazioni.

Nella parte superiore del dispositivo è presente la ruota con intorno le 12 testine e i controlli di azionamento e velocità del motore di rotazione (Figura 1).

3.1.1 Moduli lecture: amplificazione e filtraggio

In Figura 3 vi è un dettaglio parziale del pannello, dove si vedono (a sinistra) tre dei cinque moduli di lettura (denominati *lecture*, indicati dalla lettera A), con preamplificazione dei segnali delle testine di riproduzione e controlli dei filtri passa banda. Ogni modulo riguarda due testine di riproduzione (5 moduli x 2 testine = 10 testine). Considerando il modulo dall'alto verso il basso troviamo l'ingresso del segnale della testina (1 - *entree*), potenziometri per

⁸ Naturalmente ci riferiamo qui alle fonti storiche disponibili laddove sembra assodato che il *morphophone* originale sia andato disperso o distrutto [25].

⁹ Si noti che la direzione della freccia che indica il senso di rotazione è errata.

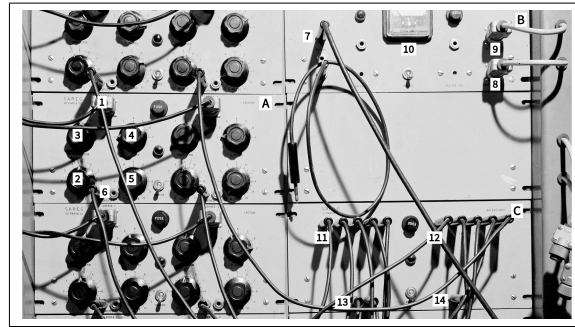


Figura 3. Dettaglio del pannello frontale del *morphophone*. A. *Lecture*: 1 - *entree*, 2 - *gain*, 3 - *correction*, 4 - *f.filtre*, 5 - *q.filtre*, 6 - *sorties*. B. *Enregistrement*: 7 - *entree*, 8 - *enregistrement*, 9 - *effacement*, 10 - *VU meter*. C. *Intermediaire*: 11 - *entrees 1*, 12 - *entrees 2*, 13 - *sorties 1*, 14 - *sorties 2*.

il controllo del preamplificatore (2 - *gain*) e dei filtri passa banda (3 - *correction*, 4 - *f.filtre*, 5 - *q.filtre*) e due prese di uscita per ciascuna delle due testine (6 - *sorties*).

Per quanto riguarda i filtri passa banda, il potenziometro *f.filtre* gestisce la frequenza centrale (su sette valori discreti possibili) e *q.filtre* la larghezza di banda (o 'campanatura'). Meno evidente è la funzione del parametro *correction*, di cui non è stato possibile trovare informazioni. Alcuni dettagli di un filmato girato nei laboratori del GRM nel 1958 [27], sembrano confermare che si tratti di un controllo che modifica la scala del potenziometro *f.filtre*. Inoltre sembra che tale controllo sia un potenziometro a scatti, suggerendo così un riscaldamento selezionabile su valori fissi, aspetto rintracciabile nei medesimi potenziometri presenti nel *morphophone*. Si è ritenuto perciò plausibile che i



Figura 4. Istantanea del filmato BBC raffigurante un modulo di filtraggio [27].

sette valori di f_{filtre} vengano riscalati dalle tre posizioni di *correction*. Rimangono tuttavia ignoti i valori assoluti di frequenza. Nell'implementazione di DME si è arbitrariamente deciso di riprendere i range visibili nel filmato citato (Figura 4), sebbene possano non corrispondere a quelle realmente disponibili nel *morphophone* (è comunque possibile modificare tali valori nell'apposita sezione di configurazione di DME). I tre range di frequenza centrale dei passa banda sono i seguenti (in Hz):

I 32, 64, 128, 256, 512, 1024, 2048;

II 37.5, 75, 150, 300, 600, 1200, 2400;

III 50, 100, 200, 400, 800, 1600, 3200.

Per quanto riguarda il fattore Q del filtro passa banda, nel *morphophone* viene indicata una scala di valori che va da 0 a 10, in egual modo riportati nell'implementazione, dove il valore 0 corrisponde alla disattivazione del filtro.

3.1.2 Modulo enregistrement: acquisizione del segnale

Sulla metà destra del pannello frontale del *morphophone* (Figura 3) vi sono due tipologie di moduli: due moduli *enregistrement*¹⁰ (B) e un modulo *intermediaire* (C), incaricati rispettivamente dell'acquisizione e dello smistamento del segnale audio (rappresentando di fatto la configurazione di un mixer). Ciascuno dei moduli *enregistrement* presenta tre ingressi, un'uscita per il segnale da indirizzare verso la testina di registrazione e un'uscita presumibilmente riservata alla testina di cancellazione (9 - *effacement*). Inoltre è presente un *VU meter* utile a monitorare il segnale in ingresso. La presenza di tre ingressi suggerisce la possibilità di acquisire più sorgenti contemporaneamente, soluzione necessaria qualora si voglia ricorrere a retroazione mentre uno degli ingressi è occupato per l'acquisizione di altro segnale.

3.1.3 Modulo intermediaire: flusso del segnale

In assenza di fonti certe, si può supporre che il segnale delle dieci testine venga mixato in due gruppi di cinque e smistato su due corrispondenti coppie di uscite (per tornare agli ingressi e/o per andare ad altri dispositivi). In generale, la doppia uscita della singola testina di riproduzione, i molteplici ingressi del modulo *enregistrement* e l'impiego del modulo *intermediaire* offrono nell'insieme la possibilità di realizzare molteplici configurazioni di feedback.

3.2 Implementazione e scelte metodologiche

Nell'implementazione dell'applicativo DME si è teso a realizzare un'architettura modulare, così da rendere ciascuna unità funzionalmente autocontenuta e collegabile agli altri moduli. A tal proposito è stata impiegata la funzione *bpatcher* presente in MAX, che consente di racchiudere il contenuto di una *patch* in una porzione d'interfaccia definita dal programmatore. Lo schema in Figura 5 riporta un

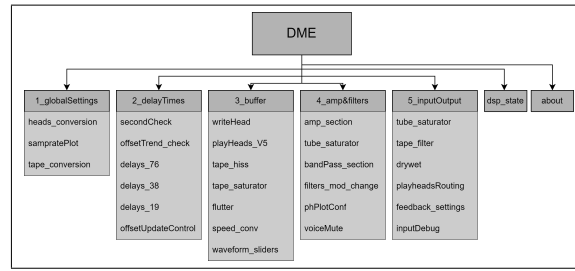


Figura 5. DME: schema riepilogativo delle patch.

quadro generale delle patch che nell'insieme compongono l'ambiente DME. Ciascun elemento grafico corrisponde ad una singola patch. Gli elementi con sfondo scuro sono gli elementi di interfaccia che richiamano al loro interno funzioni racchiuse in *subpatches* (sfondo più chiaro). Ad esempio, la patch *1_globalSettings* è presente nella patch principale DME (Figura 6) e al suo interno vengono istanziate le patch *heads_conversion*, *samplePlot* e *tape_conversion*. Le patch di interfaccia (*dsp_state* e *about* escluse) sono contrassegnate da un numero per mettere in risalto l'ordine con cui vengono inizializzati i dati contenuti in fase di avvio del software DME. Infatti per garantire il funzionamento del programma è estremamente importante che l'inizializzazione dei dati di configurazione segua un ordine ben preciso, per la corretta temporizzazione delle operazioni necessarie al calcolo delle variabili che regolano il funzionamento dell'algoritmo. Perciò si è ricorso all'impiego della funzione *ptrstorage* che consente, tra le altre cose, di definire l'ordine di inizializzazione delle variabili (*priority*). Le variabili, raccolte secondo una logica di *preset*, sono contenute in strutture dati fornite nel file *controlData.json*.

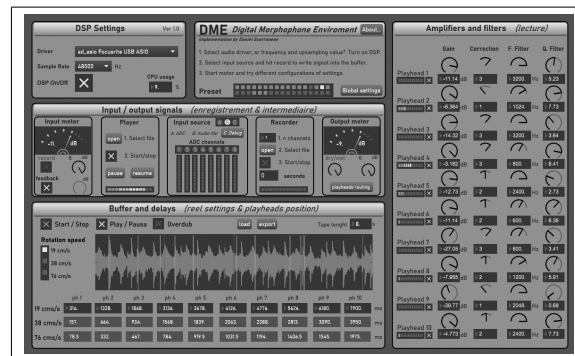


Figura 6. DME: pannello principale.

3.2.1 Ordine delle elaborazioni e funzioni extra

Per semplificare la comprensione dell'ordine dei processi di elaborazione, si faccia riferimento allo schema in Figura 7. È doveroso specificare che sono state introdotte alcune funzioni originariamente non presenti nel *morphophone*, sulla base di necessità sorte durante le fasi di sperimentazione. Esse comprendono un registratore multicanale, un controllo *dry/wet*, un selettore *mute* sull'ingresso,

¹⁰ Dalle foto a disposizione, il modulo *enregistrement* posto in basso risulta sempre scollegato.

un controllo di ampiezza sul segnale in uscita (*output trimmer*) ed un controllo di ampiezza complessiva sui segnali in retroazione (*feedback trimmer*).

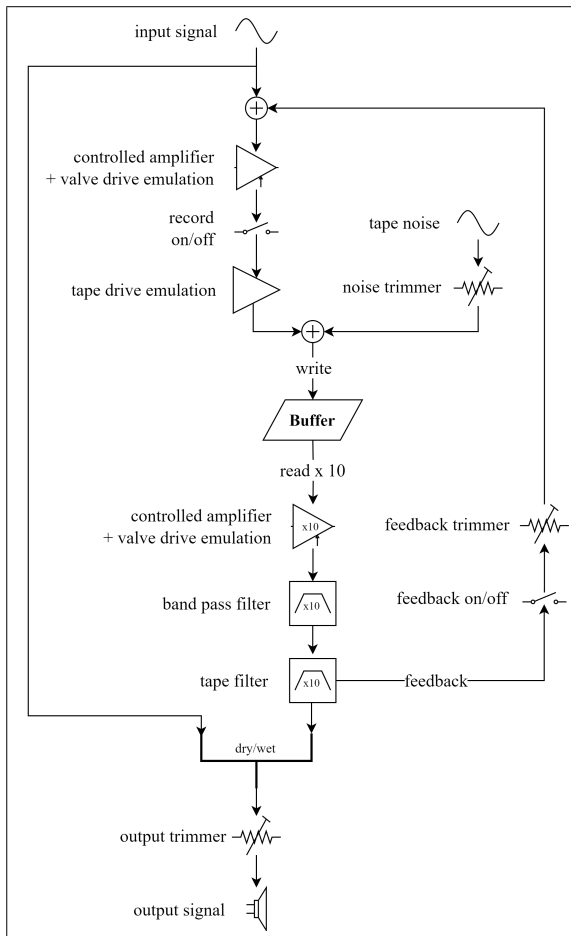


Figura 7. DME: diagramma a blocchi generale.

3.2.2 Gestione ingressi e uscite

L'implementazione prevede un elevato grado di libertà nella gestione di ingressi e uscite, anche in deroga ai limiti specifici del *morphophone* stesso. L'utente può selezionare simultaneamente molteplici sorgenti di elaborazione, sia accedendo agli ingressi dell'interfaccia audio in utilizzo (fino a otto canali di ingresso), sia impiegando un file audio. Poiché il segnale in ingresso al sistema di elaborazione deve risultare monofonico, qualora la sorgente risulti composta da più canali, questi verranno sommati.

Per quanto riguarda la gestione degli *output*, è possibile assegnare il segnale di ogni singola testina di riproduzione ad un qualsiasi canale di uscita (fino a 10), o inviarlo di nuovo all'ingresso per l'eventuale retroazione. Tali operazioni di *routing* vengono svolte attraverso un'interfaccia grafica. La selezione del canale di *output* non è esclusiva, per cui è possibile assegnare una testina a più *output* contemporaneamente e allo stesso modo un singolo canale di *output* può accogliere più sorgenti.

3.2.3 Modellazione della preamplificazione analogica

Nel tentativo di replicare (genericamente) gli aspetti caratteristici della tecnologia analogica del *morphophone*, si è ritenuto opportuno integrare il contributo di distorsione armonica che si verifica quando la preamplificazione risulta elevata. Nel *morphophone* si identificano tre punti in cui tale circostanza può verificarsi. In fase di acquisizione del segnale è presente un preamplificatore (presumibilmente valvolare), che con livelli di amplificazione elevati può portare alla saturazione del circuito ('schiacciamento' della forma d'onda e corrispondente distorsione armonica). In secondo luogo ciò avviene in fase di riproduzione, dove è presente un amplificatore (probabilmente di potenza) che consente di regolare il livello di ampiezza del segnale delle testine di riproduzione. Infine, in fase di registrazione, livelli elevati di corrente possono portare ad uno stato di saturazione sul nastro corrispondente all'allineamento totale dei domini magnetici nel campo magnetizzante, causando di nuovo distorsione armonica. È ben chiaro che tali fenomeni sono da considerarsi indesiderati, ma allo stesso tempo caratterizzano in maniera non trascurabile la risposta complessiva del sistema. Per questo è stato integrato un sistema di elaborazione mirato a simulare la distorsione nei tre stadi appena descritti¹¹. Ciò è stato fatto con un *external* che ripropone in ambiente *gen~* di MAX un *plugin* LADSPA¹² in grado di simulare tali processi mediante apposite funzioni di *waveshaping*. Si noti che il grado di saturazione magnetica è dipendente dal livello di guadagno del preamplificatore posto in ingresso, il quale, dunque, condiziona la risposta di due processi di saturazione differenti (valvolare e magnetica).

3.2.4 Anello di nastro e buffer circolare

Per simulare l'anello di nastro è stato utilizzato un *buffer* circolare, scritto e letto ciclicamente tramite appositi puntatori (corrispettivi delle testine di scrittura e di lettura). Gli indici di scrittura e lettura - implementati in MAX rispettivamente dalle funzioni *ipoke~*¹³ e *peek~*¹⁴ - sono pilotati da oscillatori a dente di sega, la cui funzione consente di accedere in modo ciclico e lineare al contenuto del buffer. La frequenza degli oscillatori definisce la velocità di accesso al buffer, la quale è strettamente dipendente dalla frequenza di campionamento in uso. Ad esempio, una frequenza di campionamento uguale a 44100 Hz implica che una frequenza di oscillazione uguale ad 1 Hz elabori 44100 campioni al secondo; quindi, un buffer costituito da 44100 campioni necessita di un oscillatore a dente di sega con frequenza uguale a 1 Hz per essere scritto e riprodotto una volta al secondo. Dimezzando o raddoppiando la frequenza degli oscillatori, è possibile rispettivamente raddoppiare o dimezzare la durata complessiva. In analogia con il *mor-*

¹¹ L'elaborazione si attiva e disattiva automaticamente in relazione alla modalità di funzionamento selezionata, filologica o estesa.

¹² Lo specifico *plugin* è *TabTubeWarmth*, incluso nella libreria *Tap-Plugins* di Tom Szilagyi (<https://tomscii.sig7.se/tap-plugins/ladspa/tubewarmth.html>). Il *porting* in ambiente *gen~* è stato realizzato da un utente anonimo (<https://cycling74.com/tools/stkr-waveshaping>).

¹³ <https://www.no-tv.org/MaxMSP/>.

¹⁴ Si noti che entrambe le funzioni implementano algoritmi di interpolazione dei campioni (lineare in scrittura, cubica in lettura).

phophone, le tre diverse velocità di rotazione del motore sono ottenute agendo sulla frequenza degli oscillatori. La dimensione del buffer, modificabile dall'utente, è impostata di *default* al corrispettivo di 4 secondi di durata (con velocità di rotazione pari a 38 cm/s), indipendentemente dalla frequenza di campionamento in utilizzo. La procedura di sovrascrittura ciclica del buffer (*overdub*), corrispondente all'operazione di rimozione della testina di cancellazione, è implementata dalla funzione *ipoke~*. Abilitando l'*overdub* e impiegando due velocità differenti per la scrittura e la lettura del buffer si altera la durata del segnale originario con conseguente trasposizione di altezza. È inoltre possibile sia importare nel buffer un *file* audio preregistrato, sia esportarne il contenuto su un nuovo *file* (gli stessi risultati si possono ottenere utilizzando rispettivamente il *player* e il registratore multicanale). Tali funzionalità riproducono alcune delle modalità operative tipiche della tecnologia a nastro magnetico, che consistono nell'impiego di nastri preregistrati come sorgenti di elaborazione o per lo svolgimento di operazioni di sovraincisione.

Il buffer e i puntatori di scrittura/lettura sono definiti all'interno di istanze *poly~*, scelta adottata per due motivi: da una parte consentono di replicare facilmente e gestire in maniera efficiente l'algoritmo inerente al puntatore di lettura per il numero di volte desiderato (10 in questo caso); dall'altra permettono di definire una frequenza di campionamento diversa per le elaborazioni presenti al loro interno attraverso operazioni di *upsampling* e *downsampling*. Nello specifico, operazioni di *upsampling* si sono rivelate necessarie per limitare le conseguenze di *aliasing* dovute alle modulazioni applicate alla frequenza dell'oscillatore che pilota il puntatore di scrittura. Tali modulazioni hanno luogo dal momento in cui viene variata la velocità di rotazione del motore, e nello specifico in fase di avvio e arresto del motore e in caso di selezione di una velocità di rotazione diversa. Infatti, il passaggio da una condizione di stasi ad una in movimento, così come l'assestamento del motore ad una diversa velocità di rotazione non avviene in maniera immediata, bensì richiede un tempo variabile che dipende dalle proprietà fisiche-realizzative del dispositivo considerato (caratteristiche del motore, inerzia del mezzo, ecc.). Tornando all'implementazione, per simulare all'occorrenza tale comportamento è stata applicata una variazione dinamica nel tempo alla frequenza degli oscillatori che generano gli indici di scrittura e di lettura, in modo che raggiungano un determinato valore in un tempo stabilito (modificabile dall'utente). Tuttavia, in fase di *debug* è stato riscontrato che la modulazione della frequenza del puntatore di scrittura introduce delle frequenze spurie, originariamente non presenti nel segnale memorizzato nel buffer. Impiegando una frequenza di campionamento elevata e/o ricorrendo ad *upsampling*, l'energia delle bande si riduce notevolmente, fino a divenire trascurabile nel caso di sorgenti sonore che presentano uno spettro complesso.

3.2.5 Simulazione dell'effetto flutter e del tape hiss

Modulazioni sulle frequenze degli oscillatori che pilotano gli indici di scrittura e lettura avvengono anche nel caso in cui venga abilitata la funzione *flutter*, elaborazione mirata

a simulare eventuali difetti del motore di rotazione che ne compromettono il mantenimento di una velocità costante. Le variazioni indotte dipendono dalla tipologia del problema che affligge il motore e possono risultare diverse, causando per esempio una variazione periodica della velocità di rotazione, così come fluttuazioni il cui andamento risulta meno prevedibile. La condizione considerata in DME replica il secondo caso descritto: valori di frequenza e ampiezza casuale sono periodicamente estratti da un generatore di numeri casuali. Tali valori descrivono il comportamento di un oscillatore sinusoidale, il cui modulo viene impiegato per influenzare la frequenza degli oscillatori che generano gli indici. Tuttavia è chiaro che il range di modulazione di tali fenomeni resta in ambito sub audio.

Si noti in Figura 7 la presenza di un generatore di rumore integrato per simulare eventuali condizioni di degradamento del nastro magnetico, le quali causano un incremento del rumore di fondo. La simulazione è stata ottenuta effettuando modulazione di ampiezza tra un generatore di rumore bianco con uno di rumore rosa. L'operazione viene reiterata una seconda volta impiegando come sorgenti il segnale risultante dalla prima modulazione ed il medesimo generatore di rumore rosa. Il segnale viene poi opportunamente regolato in ampiezza, filtrato con un passa basso e restituito in somma con la sorgente di elaborazione.

3.2.6 Linee di ritardo

Una volta acquisito ed opportunamente condizionato dai processi illustrati, il segnale viene registrato nel buffer e riprodotto dai dieci puntatori di lettura, ognuno regolato ad un tempo di ritardo differente corrispondente alla distanza che intercorre tra le testine di lettura e quella di scrittura. Per simulare il sistema di ritardo si è intervenuti sul valore degli indici dei puntatori di lettura in modo da creare una differenza tra questi ultimi e il valore attuale dell'indice di scrittura. Gli oscillatori che pilotano i puntatori di scrittura e di lettura devono operare in maniera sincrona (ovvero, è necessario che condividano i medesimi valori di frequenza e di fase), in modo che il valore degli indici dei puntatori sia uguale per tutti. Questo implica che un qualsiasi tipo di intervento volto a modificare lo stato del motore di rotazione (*start/stop*, *play/pausa* e variazione di velocità di rotazione) produca la stessa variazione di frequenza e fase di tutti gli oscillatori considerati. A partire da questa condizione iniziale, considerando un istante di tempo t , è possibile sottrarre al valore dell'indice di lettura una determinata quantità (espressa in campioni) corrispondente al tempo di ritardo desiderato, in modo da ottenere una differenza tra il valore attuale dell'indice di scrittura e quello di lettura. Tale differenza consente al puntatore di lettura di accedere al contenuto memorizzato dal puntatore di scrittura con un certo tempo di ritardo, restituendo in questo modo il risultato desiderato.

3.2.7 Gestione delle linee di ritardo

I tempi di ritardo, che nel *morphophone* sono indotti dalla posizione delle testine di riproduzione intorno alla ruota, in DME sono modificabili agendo su delle apposite caselle numeriche le quali esprimono il tempo (in millisecondi)

di) che intercorre tra la scrittura e la lettura del segnale, suddivise a loro volta per velocità di rotazione del motore. La libertà nella selezione dei tempi di ritardo è limitata da regole ben precise volte a replicare la gestione di questo aspetto nel *morphophone*. In primo luogo, l'escursione possibile di una singola testina di riproduzione è vincolata dalla posizione delle testine adiacenti, comportando che il tempo di ritardo ottenibile sia limitato entro un certo *range* di valori. In secondo luogo, è necessario considerare che la dimensione complessiva della testina occupa uno spazio definito lungo la superficie del nastro magnetico, implicando che tra due testine adiacenti vi sia una porzione di nastro 'scoperta' raggiungibile solo dal momento in cui varia la posizione di una delle due testine. Infatti, considerando una testina generica, il punto in cui avviene la trasduzione (da energia elettrica a energia magnetica nel caso delle testine di scrittura e cancellazione, viceversa nel caso di quelle di riproduzione) si trova solitamente al centro della struttura fisica complessiva della stessa. La porzione di nastro scoperta si traduce in uno 'scarto' temporale (*offset*) dipendente dalla dimensione della testina e dalla velocità di rotazione considerata. Ad esempio, assumendo che una testina occupi complessivamente la superficie di 1.9 cm, e considerando che la porzione di nastro scoperta corrisponda mediamente alla dimensione della testina (vedi Figura 8), ad una velocità di rotazione di 38 cm/s lo scarto temporale corrisponde a 50 ms (tempo = spazio / velocità). Di conseguenza, maggiore sarà la dimensione della testina, maggiore sarà lo scarto da considerare. È possibile allora definire il massimo tempo di ritardo ottenibile corrispondente alla grandezza del buffer meno il valore di *offset*; il tempo di ritardo minimo corrisponde invece al valore di *offset*.

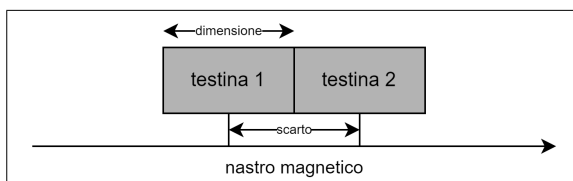


Figura 8. Porzione di nastro scoperta tra due testine adiacenti.

3.2.8 Filtro passa banda

Il segnale registrato nel buffer e riprodotto simultaneamente dai dieci puntatori di lettura, dà luogo alla generazione di dieci repliche del segnale originario, ritardate di un certo tempo, per ognuna delle quali è previsto un processo di amplificazione (con simulazione di saturazione valvolare) e filtraggio di tipo passa banda. Si è già discusso in 3.2.3 del sistema adottato per l'amplificazione del segnale riprodotto dalle testine di riproduzione. Per quanto riguarda il filtraggio è stata impiegata la funzione *reson~* presente in MAX, la quale implementa un filtro digitale basato sulle equazioni alle differenze finite, e nello specifico un filtro IIR con componente FIR, entrambe del secondo ordine. Di

seguito l'equazione implementata dalla funzione:

$$y[n] = a_0(x[n] - r \cdot x[n-2]) + b_1 \cdot y[n-1] + b_2 \cdot y[n-2] \quad (1)$$

dove r , b_1 e b_2 sono coefficienti calcolati sulla base della frequenza centrale e della larghezza di banda selezionati.

Si segnala la presenza di un filtro alla fine della catena di elaborazione inserito per simulare la risposta in frequenza generica di un dispositivo magnetofonico [28] che si abilita automaticamente dal momento in cui viene selezionata la modalità filologica, mentre in modalità estesa viene esclusa (ad ogni modo l'utente può decidere lo stato dell'elaborazione indipendentemente dalla modalità di funzionamento selezionata).

4. CONCLUSIONI E SVILUPPI FUTURI

DME si configura come un tentativo di digitalizzazione dei processi essenziali del *morphophone*, in grado di replicare modalità operative e caratteristiche tipiche del dispositivo. Particolare attenzione è stata posta su un'adeguata riproposizione dei processi di elaborazione offerti dal *morphophone*. Una modellazione accurata e fedele risulta ardua a causa dell'assenza di dati tecnico-costruttivi certi (come schemi di circuitazione), e dell'inattuabile valutazione sperimentale diretta del dispositivo. Sebbene recenti esperimenti con il software [23] sembrano dimostrare che il risultato dell'elaborazione sia coerente con l'esito previsto, sono possibili ulteriori integrazioni e migliorie al software nell'ottica di ottimizzare l'efficienza di calcolo, e fornire modelli più accurati (seppur generici) per la riproposizione delle peculiarità tipiche della tecnologia basata su nastro magnetico. Benché questo sforzo rimanga ancora in corso, DME è reperibile in rete ¹⁵ sulla piattaforma Github.

5. RIFERIMENTI

- [1] K. Tazelaar, *On the threshold of beauty: Philips and the origins of electronic music in the Netherlands 1925 - 1965*, p. 56. Rotterdam: V2 Publ, 2013.
- [2] F. Galante and N. Sani, *Musica espansa: percorsi elettroacustici di fine millennio*, p. 46. Milan: LIM/Ricordi, 2000.
- [3] J. O. Smith, "Physical Modeling Synthesis Update," *Computer Music Journal*, vol. 20, no. 2, pp. 44–56, 1996.
- [4] J. Lane, D. Hoory, E. Martinez, and P. Wang, "Modeling Analog Synthesis with DSPs," *Computer Music Journal*, vol. 21, no. 4, pp. 23–41, 1997.
- [5] D. Lowenfels, "Virtual Analog Synthesis with a Time-Varying Comb Filter," *Journal of the Audio Engineering Society*, 2003.
- [6] V. Välimäki and A. Huovilainen, "Oscillator and Filter Algorithms for Virtual Analog Synthesis," *Computer Music Journal*, vol. 30, no. 2, pp. 19–31, 2006.

¹⁵ <https://github.com/danielscorranese/DigitalMorphophoneEnvironment>.

- [7] V. Lazzarini and J. Timoney, “New Perspectives on Distortion Synthesis for Virtual Analog Oscillators,” *Computer Music Journal*, vol. 34, no. 1, pp. 28–40, 2010.
- [8] G. De Sanctis and A. Sarti, “Virtual Analog Modeling in the Wave-Digital Domain,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 4, pp. 715–727, 2010.
- [9] J. Pekonen and V. Välimäki, “The Brief History of Virtual Analog Synthesis,” in *Proceedings of Forum Acusticum*, 2011.
- [10] K. J. Werner, *Virtual Analog Modeling of Audio Circuitry Using Wave Digital Filters*. PhD thesis, Stanford University, 2016.
- [11] A. Bernardini and A. Sarti, “Towards Inverse Virtual Analog Modeling,” in *Proc. 22nd International Conference on Digital Audio Effects (DAFx 2019)*, pp. 1–8, 2019.
- [12] J. Wilczek, A. Wright, E. A. P. Habets, and V. Välimäki, “Virtual Analog Modeling of Distortion Circuits Using Neural Ordinary Differential Equations,” in *Proc. 25th International Conference on Digital Audio Effects (DAFx 2022)*, 2022.
- [13] S. Arnardottir, J. S. Abel, and J. O. Smith, “A Digital Model of the Echoplex Tape Delay,” *Journal of the Audio Engineering Society*, 2008.
- [14] V. Välimäki, S. Bilbao, J. O. Smith, J. S. Abel, J. Parkarinen, and D. Berners, “Virtual Analog Effects,” in *DAFX: Digital Audio Effects*, pp. 473–522, 2011.
- [15] J. Chowdhury, “Real-time Physical Modelling for Analog Tape Machines,” in *Proc. 22nd International Conference on Digital Audio Effects (DAFx 2019)*, 2019.
- [16] “The ‘Groupe de Recherches Musicales’ Pierre Schaeffer, Pierre Henry & Jacques Poullin, France 1951.” <https://120years.net/the-grm-group-and-rtf-electronic-music-studio-pierre-schaeffer-jacques-poullin-france-1951>. [last accessed april 15th 2024].
- [17] D. Teruggi, “Technology and musique concrète: the technical developments of the Groupe de Recherches Musicales and their implication in musical composition,” *Organised Sound*, vol. 12, no. 3, p. 218, 2007.
- [18] D. Teruggi. Personal communication, may 7th 2021.
- [19] J. Poullin, “L’apport des techniques d’enregistrement dans la fabrication de matières et formes musicales nouvelles. Applications à la musique concrète,” *L’onde Électrique*, no. 324, p. 10, 1954. Reissued in *Ars sonora*, 9, 1999 (<http://www.ars-sonora.org/html/numeros/numero09/>).
- [20] A. Moles and P. Schaeffer, “Vers une musique expérimentale,” *La revue musicale*, no. 236, pp. 121–122, 1957.
- [21] R. Friedl, *Towards a Philology of Electroacoustic Music - Xenakis’s Tape Music as Paradigm*. PhD, Goldsmiths, University of London. p. 68, 2019.
- [22] M. Solomos and B. Gibson, “Research on the First Musique Concrète: The Case of Xenakis’s First Electroacoustic Pieces,” p. 3, International Electroacoustic Music Studies Network. Lisbon, June 2013.
- [23] D. Scorrane and A. Di Scipio, “Gli ‘sciami di glissando’ in Diamorphoses. Ricostruzione mediante Digital Morphophone Environment,” paper presentation. XXIV Colloquium of Musical Informatics, Turin, 2024.
- [24] E. Favreau, “Les outils de traitement GRM Tools,” in *Actes - Journées d’Informatique Musicale*, Marseille: CNRS-LMA, 1998.
- [25] D. Teruggi, “De phonogène en phonogène: cinquante années d’outils concrets,” in *Du sonore au musical: Cinquante années de recherches concrètes (1948–1998)* (S. Dallet and A. Veitl, eds.), Paris: L’Harmattan, 2001.
- [26] A. A. Moles, *Les Musiques expérimentales : Revue d’une tendance importante de la musique contemporaine*, p. 74. Ed. du Cercle d’art contemporain. Paris, 1960.
- [27] Unknown, “Musique concrète performance.” <https://www.facebook.com/watch/?v=1786093964886230>, 1958. [last accessed april 15th 2024].
- [28] M. Camras, *Magnetic Recording Handbook*, p. 74. Dordrecht: Springer Netherlands, 1998.

Controllo remoto via OSC di un prototipo di synth basato su Raspberry Pi

Costantino Rizzuti

Artis Lab

costantinorizzuti@gmail.com

Fabrizio Rizzuti

Artis Lab

fabrizio.rizzuti@gmail.com

ABSTRACT

SynthBerry Pi è il prototipo di un sintetizzatore autonomo basato sul single-board computer Raspberry Pi e sull'ambiente di sviluppo Pure Data. Il prototipo è dotato di una superficie di controllo composta da otto potenziometri connessi al computer che consentono di modificare, in tempo reale, i parametri di controllo dei processi di generazione del suono. Inoltre, sfruttando la connettività WiFi di Raspberry Pi, le patch di Pure Data possono essere controllate in remoto, utilizzando dispositivi quali tablet o smartphone, mediante l'invio di messaggi Open Sound Control.

1. INTRODUZIONE

Negli ultimi venti anni la miniaturizzazione dei sistemi di calcolo e la crescente diffusione di piattaforme *open*, software e hardware, ha reso possibile l'accesso per artisti e designer a tecnologie fino ad allora a disposizione soltanto per tecnici e ingegneri attivi nei grandi centri di ricerca delle università o di grandi compagnie private.

Come afferma Noble [1], tutto ciò ha creato delle condizioni assolutamente nuove e mai sperimentate prima che hanno spinto alla nascita e allo sviluppo di nuovi campi di ricerca e di sperimentazione nell'Arte, nel Design e anche nelle tecnologie per la Musica. Il Physical Computing e l'Interaction Design sono due settori di ricerca che testimoniano questa evoluzione. Infatti, in passato, se si esclude da tali considerazioni lo sviluppo storico della Musica Elettronica e dell'Informatica Musicale in cui, sin dalle prime sperimentazioni pionieristiche, i musicisti e i compositori hanno sempre condotto un confronto diretto e, spesso, in prima persona con le problematiche tecnologiche, l'idea di avere artisti o designer dediti, nella propria pratica creativa, a scrivere codice o a progettare sistemi hardware è sempre stato qualcosa di piuttosto raro e, in qualche modo, inconsueto. Oggi, al contrario, tutto ciò è ormai diventato non soltanto un fenomeno piuttosto comune e diffuso, ma, anche, ha determinato la nascita di un importante campo di sperimentazione e di esplorazione per nuove pratiche in tutti i diversi linguaggi artistici. Così oggi questo profondo legame tra creazione artistica e ricerca tecnologica è divenuto un vitale e vibrante fenomeno che influenza in maniera sostanziale sia lo sviluppo delle nuove forme d'Arte, sia la ricerca tecnologica.

Copyright: ©2024 Costantino Rizzuti et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

L'affermarsi di queste tecnologie ha portato a interessanti conseguenze anche nel campo dell'Informatica Musicale e della Musica Elettronica. In particolare, è importante osservare la convergenza positiva con altri fenomeni già esistenti quali: lo sviluppo di nuove interfacce e sistemi d'interazione innovativi per la Musica, le pratiche di autocostruzione di sintetizzatori e la creazione di strumenti musicali elettronici sperimentali. Queste ultime pratiche possono essere chiamate, utilizzando la definizione inglese, come: *analog synthesizer do it yourself* o, nella più nota forma abbreviata: *synth DIY* [2]. Negli ultimi anni l'autocostruzione di sintetizzatori e di dispositivi analogici per la creazione ed elaborazione del suono ha avuto una crescente diffusione soprattutto in relazione all'affermazione dei sintetizzatori modulari Eurorack e al consistente ritorno all'uso di tecnologie hardware rispetto a quanto accadeva nei venti anni precedenti dominati dall'impiego di soluzioni software eseguite su calcolatori general purpose.

Oggi, infatti, si osserva una strana e promettente commistione tra l'utilizzo e la diffusione di soluzioni hardware piuttosto che software: moduli hardware basati su microcontrollori e single-board computer, da un lato, in cui l'elaborazione interna dei segnali è di tipo digitale e si realizza mediante l'uso di opportuni software; lo sviluppo di software per dispositivi touchscreen e per calcolatori tradizionali, dall'altro lato, volto a simulare il comportamento dei sintetizzatori modulari hardware.¹ Il senso profondo di simili possibilità è, probabilmente, ancora oggi tutto da investigare. Sembra però certo che il rapporto hardware-software sia dominato da processi di feedback in cui lo sviluppo e il prevalere di un aspetto, inevitabilmente, sia controbilanciato da fenomeni che spingono all'evoluzione simbiotica anche dell'altro.

Il presente lavoro è volto a illustrare una estensione delle funzionalità di controllo del prototipo *SynthBerry Pi*: un sintetizzatore autonomo basato sul single-board computer (SBC) Raspberry Pi e sull'ambiente di sviluppo Pure Data (Pd). In un articolo precedente sono stati presentati in dettaglio gli aspetti tecnici legati alla costruzione del prototipo e il suo funzionamento [3]. Nel seguito, si discuterà lo sviluppo di un sistema per il controllo da remoto delle patch di Pure Data basato sull'invio di messaggi *Open Sound Control (OSC)* attraverso una connessione di rete WiFi. In tal modo dispositivi quali, per esempio, tablet o smartphone possono essere utilizzati per controllare i processi di generazione del suono attraverso interfacce touchscreen.

¹ A tal proposito si veda, per esempio, *VCVRack* un programma open source che consente di simulare il comportamento di un sintetizzatore modulare: <https://vcvrack.com/>.

Il prototipo utilizza delle patch di Pure Data per generare ed elaborare i suoni. La collezione di patch utilizzate per questo progetto è stata chiamata *PDSynth* e costituisce un insieme di moduli (un *toolkit*) finalizzato alla creazione di sintetizzatori digitali ispirati al funzionamento dei sistemi analogici. Raspberry Pi² è stato scelto per eseguire in tempo reale le architetture di *PDSynth* al fine di creare un compatto sintetizzatore autonomo. SynthBerry Pi è dotato di una interfaccia di controllo hardware composta da otto potenziometri che consentono di modificare, in tempo reale e in maniera tra loro indipendente, altrettanti parametri delle patch di Pd.

Raspberry Pi è un single-board computer ormai piuttosto noto e utilizzato in diversi campi di applicazione. Il progetto è basato su dei *system-on-a-chip* della Broadcom che incorporano un processore ARM, una GPU VideoCore IV e della memoria RAM. La scheda si avvale di una scheda SD esterna per il caricamento del sistema operativo (*boot*) e per la gestione della memoria non volatile.

Raspberry Pi, così come anche altre schede di sviluppo e prototipazione simili, sono state utilizzate negli ultimi anni per realizzare diversi progetti nel campo dell'audio e dei sintetizzatori modulari e dei *synth DIY* [4, 5, 6, 7]. Infatti, sono disponibili sul mercato dei moduli Eurorack basati sull'uso di questo single-board computer come per esempio il progetto *Terminal Tedium* e il *Nebulae 2* di Qu-Bit Electronix. Al pari di *SynthBerry Pi*, oggetto del presente articolo, l'uso del Raspberry Pi in questi dispositivi rende possibile ottenere una *macchina aperta*, ossia riprogrammabile, che consente di implementare anche sistemi di sintesi ed elaborazione del suono che possono essere personalizzati e modificati dall'utente attraverso l'uso di diversi linguaggi e ambienti di sviluppo dedicati all'audio.

2. PDSYNTH

*PDSynth*³ è una collezione di patch sviluppata nella versione Vanilla di Pure Data per sperimentare la creazione di sintetizzatori digitali programmabili. Il nome deriva, da un lato, dall'acronimo: *Programmable Digital Synthesizer*. Dall'altro lato, la sigla *PDSynth* serve anche a designare un synth realizzato con Pure Data (Pd). In un articolo precedente sono stati presentati in dettaglio gli aspetti tecnici legati all'implementazione di *PDSynth* [3].

L'idea iniziale che ha spinto a sviluppare il progetto *PDSynth* deriva dal bisogno di disporre di una serie di patches, facilmente interfacciabili tra loro, capaci di simulare il comportamento dei moduli essenziali di un sintetizzatore analogico. I moduli di *PDSynth* implementano diversi sistemi di generazione ed elaborazione del suono e sono tutti controllabili tramite il protocollo Open Sound Control. Gli utenti possono facilmente creare e interconnettere i moduli tra loro per costruire delle architetture di alto livello per la sintesi e l'elaborazione del suono in tempo reale.

² Per maggiori informazioni sul progetto si veda il sito della Raspberry Pi Foundation: <https://www.raspberrypi.org>.

³ Per maggiori informazioni e per scaricare il toolkit si veda la pagina ufficiale del progetto: <https://www.artislab.it/it/programmable-digital-synthesizer-pdsynth>.

2.1 Architettura di PDSynth

Il toolkit offre attualmente cinque moduli per la generazione del suono che emulano il comportamento dei classici oscillatori dei synth analogici. I moduli offrono la possibilità di generare le seguenti forme d'onda:

GENPULSE — generatore di treni d'impulso;

GENSAWTOOTH — generatore di onda a dente di sega;

GENSIN — generatore di onda sinusoidale;

GENSQUARE — generatore di onda quadra;

GENTRIANGLE — generatore di onda triangolare.

Il modulo **GENSIN** utilizza l'oggetto nativo di Pure Data `osc~` per la generazione dell'onda sinusoidale. Tutti gli altri moduli producono dei segnali limitati in banda e sono basati sulla lettura di dati salvati in tabelle (*Sintesi Wave Table*) [8].

Oltre gli oggetti per la generazione del suono, sono stati realizzati anche dei moduli che implementano filtri e sistemi di elaborazione del suono. Finora, sono stati creati tre differenti filtri del quarto ordine:

FLTBandPass — filtro passa banda;

FLTHighPass — filtro passa alto;

FLTLowPass — filtro passa basso.

I filtri del quarto ordine sono stati realizzati ponendo in cascata gli oggetti nativi di Pd per la creazione dei filtri. Il filtro passa banda è basato sull'oggetto `vcf~`, mentre per il passa alto si usa l'oggetto `hip~` e, infine, l'oggetto `lop~` per la realizzazione del filtro passa basso.

Sono stati sviluppati anche due moduli utili per generare degli involucri:

ENVTable — generatore di involucri basato su tabella;

ENVADSR4 — generatore di involucri di tipo ADSR con interpolazione polinomiale del quarto ordine.

2.2 Open Sound Control

Il protocollo MIDI, nonostante la sua diffusione e le grandissime potenzialità offerte all'epoca della sua creazione, è caratterizzato da alcune forti limitazioni, specialmente imputabili alla sua rigidità e al limitato sviluppo che ha subito negli anni. Per superare tali limitazioni, a partire dalla fine degli anni '90 dello scorso secolo, è stato sviluppato, presso il CNMAT dell'Università di Berkeley, da Adrian Freed e Matt Wright [9, 10] l'ormai ben noto protocollo di comunicazione chiamato *Open Sound Control (OSC)* che rende possibile anche lo scambio di dati via rete.⁴ OSC nasce, nelle intenzioni degli autori come un formato per la condivisione di dati relativi all'esecuzione musicale (parametri, sequenze di note, gesti) tra strumenti musicali (principalmente sintetizzatori e strumenti elettronici), calcolatori e altri dispositivi multimediali. Questo protocollo nasce come aperto, flessibile ed estendibile e si è affermato

⁴ Per approfondimenti si veda il sito ufficiale del progetto: <http://opensoundcontrol.org>.

negli anni come una valida alternativa al MIDI soprattutto in quelle applicazioni in cui è necessario disporre di livelli di risoluzione e di accuratezza maggiore, oppure di uno spazio di parametri maggiore, rispetto a quanto offerto dai sistemi precedenti.

OSC è un protocollo di comunicazione appositamente ottimizzato per sfruttare le moderne tecnologie di rete. I messaggi OSC, infatti, sono comunemente trasmessi, tramite opportuni protocolli (UDP/IP oppure Ethernet), in reti locali o all'interno di uno stesso calcolatore — *in locale* — oppure attraverso Internet — *in remoto*. I vantaggi offerti da OSC rispetto all'uso del MIDI consistono nella possibilità di utilizzare la connettività offerta dalla rete, dalla maggiore risoluzione dei tipi di dati che è possibile trasmettere, dalla semplice definizione degli indirizzi dei messaggi, definiti mediante una sequenza alfanumerica, che offre la possibilità di creare delle strutture gerarchiche degli indirizzi. Da ciò deriva che il protocollo riesce a fornire a musicisti e sviluppatori una maggiore flessibilità sia nella tipologia di dati supportati, sia nell'indirizzamento dei messaggi al fine di realizzare sistemi di comunicazione e di scambio di dati anche di elevata complessità.

Ovviamente, il protocollo OSC è caratterizzato anche da alcuni svantaggi che è importante considerare nella fase di sviluppo di applicazioni basate sul suo utilizzo. I due principali svantaggi consistono nella maggiore complessità e onerosità della trasmissione, analisi e distribuzione dei messaggi i cui indirizzi sono codificati in formato stringa rispetto a protocolli, come il MIDI, che adottano sistemi binari e la possibilità che si manifestino perdite di pacchetti durante il trasporto dei messaggi su UDP, specie su reti WiFi. Questo è un problema piuttosto serio soprattutto nel caso di messaggi come gli equivalenti di Note On e Off del MIDI in cui è necessaria la ricezione di entrambi i messaggi nella giusta sequenza temporale per evitare che l'emissione del suono rimanga sempre attiva per la mancata trasmissione del messaggio di interruzione.

I messaggi OSC devono conformarsi con una precisa sintassi che sta alla base stessa del protocollo di comunicazione. I messaggi consistono in una successione alfanumerica che definisce l'**Indirizzo**, una stringa che definisce il **Tipo** dei dati inviati, l'**Argomento** ossia i dati del messaggio e, infine, un'indicazione di tempo opzionale (*timestamps*) che, se necessario, consente la temporizzazione della lettura nella fase di ricezione.

L'**Indirizzo** è costituito da una sequenza alfanumerica che comincia con il carattere / (barra obliqua — *slash* in inglese) composta secondo una sintassi simile ai percorsi nei file system di tipo Unix oppure alle URL comunemente utilizzate in Internet. L'indirizzo definisce l'indicativo del messaggio (definito come *Address Space* in inglese) secondo una sequenza gerarchica aperta e modificabile a piacimento a seconda dei casi e delle applicazioni. Di seguito sono riportati alcuni indirizzi di messaggi OSC:

```
/frequenza
/synth1/modulante
/synth1/portante
/synth2/filter/cutoff
/synth2/sin/frequenza
```

Gli esempi proposti mettono in luce la possibilità di sfruttare la struttura gerarchica degli indirizzi. L'esempio sulla prima riga mostra un indirizzo denominato *frequenza* relativo a un messaggio con un *address space* a un solo livello. Sulla seconda e terza riga si trovano due indirizzi caratterizzati da due livelli. In questo caso il primo livello potrebbe essere utilizzato per definire l'indirizzo dello strumento da controllare (*synth1*) e il secondo livello per definire il nome del parametro da controllare (*modulante* oppure *portante*). Le restanti due righe mostrano un esempio di *address space* a tre livelli. In questo caso: il primo livello è simile agli esempi precedenti, mentre il secondo potrebbe individuare delle specifiche aree funzionali dello strumento (filtro, generatore sinusoidale) e, infine, il terzo definisce il nome dei parametri da controllare.

La costruzione degli *address space* è assolutamente libera e non è vincolata dalla sintassi del protocollo quindi nel caso di sviluppo di applicazioni proprie può essere realizzata secondo le esigenze del caso. La libera scelta degli indirizzi dei messaggi assieme alla possibilità di creare una struttura gerarchica degli indirizzi sono due caratteristiche molto utilizzate nello sviluppo delle architetture di PDSynth. Ciò, infatti, consente di lasciare agli utenti la possibilità di modificare i protocolli di comunicazione esistenti o, anche, di crearne di propri.

L'unità di trasmissione del protocollo è un **Pacchetto OSC**. Qualsiasi applicazione capace di inviare pacchetti OSC viene definita come **OSC Client**, mentre qualsiasi applicazione che riceve pacchetti viene chiamata come **OSC Server**. Alla luce di questa distinzione, è possibile affermare che i moduli di PDSynth sono degli **OSC Server** capaci di modificare i parametri degli algoritmi di generazione ed elaborazione del suono in risposta ai messaggi OSC che ricevono dall'esterno. L'invio di un messaggio attiva nel server l'invocazione di un metodo OSC tramite una chiamata a procedura del metodo corrispondente al messaggio. Il metodo viene invocato con una serie di argomenti forniti dai dati ottenuti dalla ricezione del messaggio. I metodi di un server OSC, in maniera analoga a quanto accade per gli indirizzi dei messaggi, sono organizzati in una struttura ad albero chiamata spazio degli indirizzi (**Address Space**) in cui le foglie sono i metodi OSC e i rami sono chiamati **contenitori OSC (OSC Containers)**.

Uno dei grandi vantaggi offerti da OSC consiste nella opportunità di disporre di un *Address Space* dinamico che può essere creato e modificato nel tempo durante l'esecuzione delle applicazioni. Quando un server OSC riceve un messaggio OSC deve invocare l'appropriato metodo OSC che viene selezionato in base all'*address space* a sua disposizione, questa procedura viene chiamata consegna del messaggio OSC al metodo OSC corrispondente.

2.3 Implementazione in Pd del controllo OSC

Per realizzare le funzionalità di controllo delle patch di PDSynth attraverso il protocollo OSC è stata utilizzata una libreria di oggetti (*External* nel gergo di Pd) che fornisce il supporto per la gestione dei messaggi OSC.⁵ Gli oggetti di

⁵ Si veda la pagina dedicata alla libreria sul sito ufficiale di Pure Data: <http://puredata.info/downloads/osc>.

questa libreria non consentono di realizzare il trasferimento dei pacchetti OSC attraverso la rete, a tal fine è necessario utilizzare delle altre apposite librerie. Tra le diverse soluzioni possibili si è scelto di impiegare *IEMnet*⁶ che fornisce oggetti utili a tale scopo. In particolare è possibile utilizzare `udpreceive` per implementare all'interno di Pd un server in ascolto su una data porta per la ricezione dei messaggi OSC. L'immagine in Figura 1 mostra la patch realizzata in Pd per la ricezione e lo smistamento dei messaggi OSC. L'oggetto `udpreceive` posto in alto realizza un server per la ricezione dei messaggi in ascolto sulla porta definita mediante il suo argomento, *10000* nel caso dell'immagine in Figura 1.

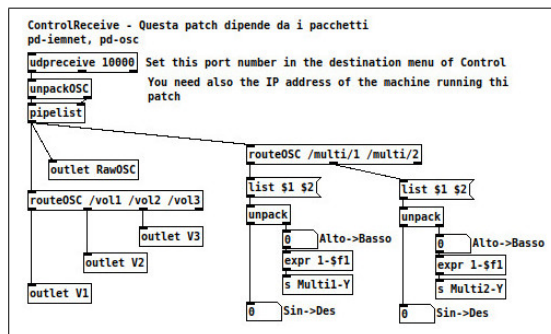


Figure 1. Patch per la ricezione e lo smistamento dei messaggi OSC.

L'elaborazione dei messaggi ricevuti dal server implementato dall'oggetto `udpreceive` può essere realizzata tramite gli oggetti forniti dalla libreria OSC. L'oggetto `unpackOSC` è utile a convertire i pacchetti OSC, composti da dati binari, in messaggi compatibili con il sistema di messaggistica interno a Pure Data. In seguito si inserisce l'oggetto `pipelist` per ottenere una scansione dei messaggi temporalmente coerente nel caso in cui vengano ricevuti messaggi con un dato *timestamps*. Infine, la libreria OSC fornisce un oggetto, chiamato `routeOSC`, che consente di realizzare l'indirizzamento dei messaggi secondo una struttura gerarchica definita mediante l'address space. Gli argomenti forniti all'oggetto definiscono un insieme di indirizzi a cui possono essere inviati i messaggi corrispondenti. I messaggi corrispondenti al primo indirizzo `/vol1` nell'oggetto `routeOSC` posto a sinistra in Figura 1 verranno inviati sulla prima uscita a sinistra, quelli corrispondenti al secondo (`/vol2`) sulla seconda e così via. L'ultima uscita a destra viene utilizzata per inviare tutti i messaggi ricevuti che non trovano una corrispondenza con gli argomenti forniti all'oggetto. Ponendo in cascata due o più oggetti è possibile ricevere e indirizzare opportunamente messaggi che hanno indirizzi con più livelli.

La scelta di utilizzare il protocollo di comunicazione Open Sound Control è sembrata essere la via più semplice per realizzare un sistema di comunicazione affidabile e robusto tra i moduli stessi e da e verso l'esterno. I messaggi OSC possono essere, infatti, facilmente gestiti, all'interno

⁶ Il codice sorgente della libreria può essere scaricato dal seguente link: <http://puredata.info/downloads/iemnet>.

di Pure Data, attraverso il sistema di scambio di messaggi nativo dell'ambiente di sviluppo. Tutto ciò semplifica la creazione dei sistemi di controllo dei moduli e rende i sintetizzatori facilmente controllabili in remoto attraverso dispositivi connessi in rete oppure da controller e superfici di controllo. Nel caso in cui si usino dispositivi che non supportano in maniera nativa il protocollo OSC è possibile realizzare delle patch di Pd per convertire i dati provenienti dagli strumenti esterni (MIDI o connessione seriale nel caso di schede Arduino o simili) in corrispondenti messaggi OSC opportunamente indirizzati a modificare i valori dei parametri dei moduli.

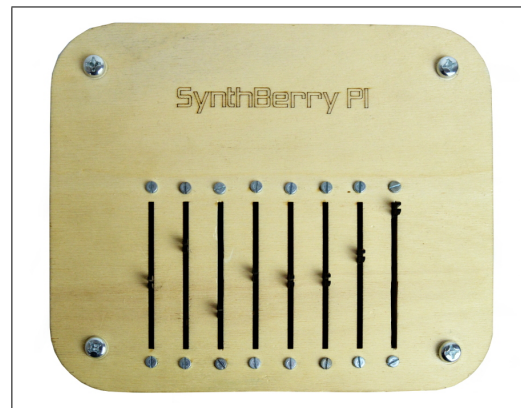


Figure 2. La superficie di controllo di SynthBerry Pi.

3. SYNTHBERRY PI

SynthBerry Pi è un compatto sintetizzatore capace di eseguire patch create attraverso Pure Data. È il primo prototipo di sintetizzatore autonomo basato sul toolkit PDSynth; per eseguire le architetture di sintesi di PDSynth si è scelto di utilizzare un Raspberry Pi 3, sperimentando l'impiego di entrambe le versioni disponibili (Model A+ e B+). Otto potenziometri a slitta sono connessi al computer per realizzare un'essenziale superficie di controllo che rende possibile controllare in tempo reale le patches di Pure Data.

Il prototipo è stato realizzato assemblando due pannelli di multistrato sagomati mediante taglio laser. Sul pannello superiore sono stati disposti gli otto potenziometri a slitta che compongono l'interfaccia di controllo del prototipo. Sul pannello inferiore sono stati montati il computer e la circuiteria elettronica necessaria a interfacciare i potenziometri con il Raspberry Pi. L'assemblaggio tra i due pannelli del prototipo è stato realizzato attraverso opportuni distanziali esagonali in acciaio. L'immagine in Figura 2 mostra la superficie di controllo del prototipo.

Nell'autunno del 2019 è stata realizzata una seconda versione del prototipo sotto forma di un modulo Eurorack da 18 hp che a differenza del precedente è dotato di otto potenziometri rotativi. La Figura 3 mostra una vista laterale del prototipo che evidenzia la struttura interna del modulo. Una scheda PCB progettata e realizzata a mano è connessa al connettore GPIO del Raspberry Pi. Il computer non è dotato di convertitori analogico digitale, per questo motivo

per effettuare la lettura delle tensioni degli otto potenziometri è stato utilizzato in entrambi i prototipi un apposito circuito integrato che realizza un ADC (Analog to Digital Converter). Il circuito integrato utilizzato come ADC, è denominato *MCP3008*. Questo IC fornisce otto canali analogici d'ingresso e una risoluzione digitale pari a 10 bit. I potenziometri sono collegati attraverso dei cavi alla scheda PCB su cui è collocato il convertitore analogico digitale. Questo secondo prototipo realizzato come modulo Eurorack è del tutto analogo al primo dal punto di vista della struttura hardware. Per quanto riguarda il software è ovviamente possibile differenziare il comportamento sonoro dei due prototipi modificando opportunamente le patch di Pd che devono essere eseguite e i parametri di controllo associati ai diversi potenziometri.⁷

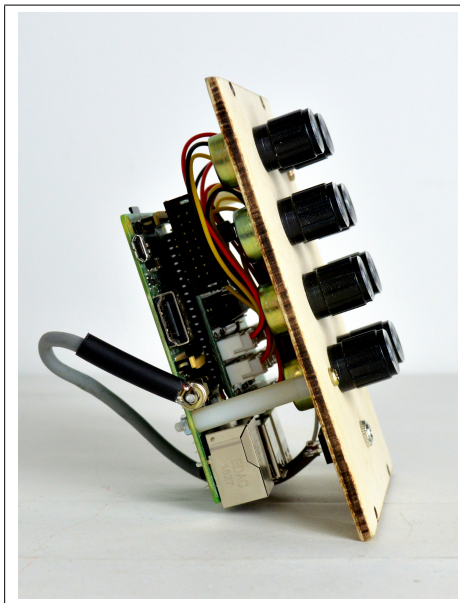


Figure 3. La struttura interna del prototipo di modulo Eurorack.

In seguito per semplificare il processo di costruzione del modulo Eurorack è stata progettata una PCB che può essere prodotta industrialmente per sostituire la prima versione realizzata a mano. La scheda è dotata del connettore GPIO a cui connettere il Raspberry Pi. Inoltre su di essa sono montati gli otto potenziometri, il chip ADC, due connettori audio mono per l'uscita stereo del computer e cinque pulsanti tattili. La Figura 4 mostra sulla destra la scheda PCB completa di tutti i componenti necessari, il pannello frontale, le manopole e i componenti meccanici necessari per l'assemblaggio del modulo.

3.1 Architettura del prototipo

La comunicazione tra Raspberry e l'ADC avviene attraverso il protocollo di comunicazione seriale *SPI* (*Serial Peripheral Interface*). SPI è un sistema di comunicazione tra

⁷ Un video che mostra il funzionamento del prototipo è disponibile al seguente link: https://youtu.be/dQ261949_ng.

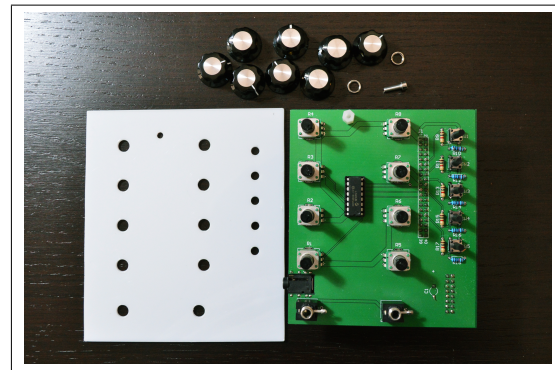


Figure 4. La scheda PCB.

un microcontrollore e altri circuiti integrati o tra più microcontrollori. È uno standard di comunicazione, ideato dalla Motorola, in cui la trasmissione avviene tra un dispositivo di controllo (detto *master*) e uno o più dispositivi controllati (detti *slave*). Il dispositivo master controlla il bus di comunicazione, emette il segnale di *clock*, decide quando iniziare e terminare la comunicazione. La Figura 5 mostra lo schema elettrico delle connessioni tra il convertitore analogico digitale e il Raspberry Pi.

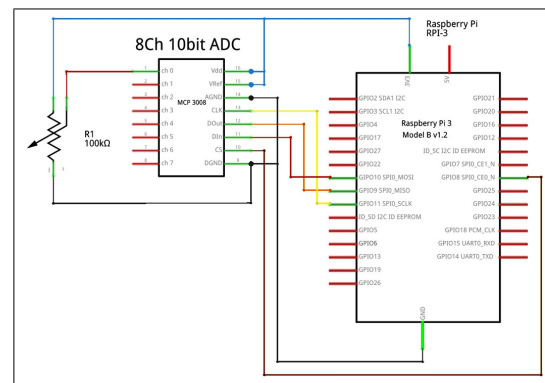


Figure 5. Schema elettrico delle connessioni.

Il sistema di comunicazione SPI è comunemente definito a quattro fili, poiché per la trasmissione dei dati vengono utilizzati 4 segnali distinti in genere definiti come: Serial Clock (emesso dal master), MISO (Master Input Slave Output), MOSI (Master Output Slave Input), CS (Chip Select). Di questi quattro il Chip Select, in alcune applicazioni può essere omesso poiché la sua presenza è necessaria nel caso in cui si debbano gestire contemporaneamente più dispositivi slave. A queste quattro connessioni deve essere aggiunta un collegamento che definisce il livello di riferimento della tensione spesso indicato come **GND**.

La lettura dei dati acquisiti dall'ADC si realizza attraverso uno script di Python che impiega la libreria *SPIDEV* per la gestione dei dispositivi SPI. Lo script esegue la lettura dei dati dall'ADC, il loro filtraggio (calcolo del valore medio ogni dieci letture) e il trasferimento dei dati così elaborati verso Pure Data attraverso l'invio di messaggi di rete

locali su una specifica porta.⁸ Per questo scopo si utilizza il programma `pdsend`, fornito nel pacchetto standard di Pd, che viene utilizzato come sotto processo (*subprocess*) all'interno dello script di Python. Un semplice protocollo consente di inviare i messaggi da Python a Pure Data mantenendo separati e facilmente differenziabili i dati dei diversi potenziometri. Si invia una lista di due numeri: il primo è un'etichetta (da 0 a 7) utile per identificare il potenziometro, il secondo numero è il dato numerico ottenuto dalla lettura effettuata dal convertitore. Come mostrato nella Figura 6, la ricezione dei dati in Pure Data si realizza impiegando l'oggetto nativo `netreceive` che apre un server in ascolto sulla porta corrispondente a quella utilizzata da `pdsend`. L'espediente utilizzato nella costruzione del messaggio inviato da Python semplifica lo smistamento dei dati che può essere facilmente compiuto attraverso l'oggetto nativo di Pd `route`.

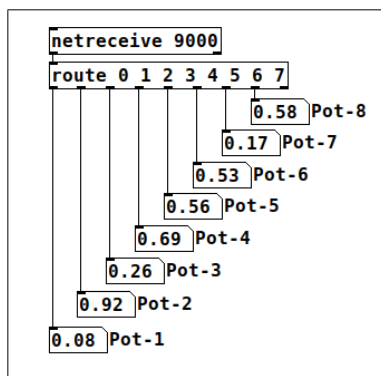


Figure 6. Patch per la ricezione dei messaggi.

Uno script appositamente creato viene eseguito all'avvio del sistema operativo al fine di avviare l'esecuzione di Pure Data con la patch di PDSynth prescelta e lanciare il programma realizzato in Python per la gestione dell'ADC. Poiché si vuole utilizzare il prototipo come un comune strumento elettronico attraverso l'uso della sola superficie di controllo si è scelto di utilizzare il Raspberry Pi in modalità *headless*.⁹ Per compiere l'avvio automatico dello script durante la fase di avvio del Raspberry Pi è stato creato un apposito servizio, avviato in automatico durante le fasi iniziali di esecuzione del sistema operativo, che invoca lo script di lancio dei processi audio e del programma di Python per la gestione dell'ADC.

3.2 Il motore audio

Sono state sperimentate, finora, due differenti patches che possono essere usate come motore audio nei due diversi

⁸ Per maggiori dettagli sugli aspetti tecnici legati alla costruzione e al funzionamento del prototipo si veda l'articolo presentato alla Linux Audio Conference 2020 citato in precedenza.

⁹ Nel gergo degli utilizzatori di Raspberry ciò significa senza la connessione di schermo, mouse e tastiera. È importante sottolineare che in questa modalità di funzionamento le prestazioni del calcolatore aumentano in maniera significativa poiché la gestione dell'interfaccia grafica e dei dispositivi d'interazione in genere assorbono molte delle risorse di calcolo del computer. Per risparmiare risorse in modalità *headless* è utile anche configurare il sistema operativo per avviarsi in modalità console evitando l'avvio dell'ambiente che gestisce l'interfaccia grafica.

prototipi. La prima offre la possibilità di generare suoni mediante tre oscillatori con forme d'onda differenti, mentre la seconda realizza un sistema di sintesi sottrattiva.

3.2.1 Sintesi mediante tre oscillatori

La prima patch offre la possibilità di controllare separatamente l'ampiezza e la frequenza di tre oscillatori con forme d'onda differenti (onda quadra, impulsi e onda a dente di sega).¹⁰ Inoltre, fornisce una linea di ritardo dotata di un percorso di retroazione del segnale. La linea di ritardo è controllata mediante due parametri che possono essere modificati in tempo reale tramite i potenziometri della superficie di controllo del prototipo: il tempo di ritardo e il coefficiente di feedback. Gli otto potenziometri del prototipo sono stati associati ad altrettanti parametri di controllo della patch secondo il seguente schema:

- A0 - frequenza dell'oscillatore onda quadra;
- A1 - ampiezza dell'oscillatore onda quadra;
- A2 - frequenza del generatore d'impulsi;
- A3 - ampiezza del generatore d'impulsi;
- A4 - frequenza dell'oscillatore dente di sega;
- A5 - ampiezza dell'oscillatore dente di sega;
- A6 - tempo di ritardo del delay;
- A7 - coefficiente di feedback.

3.2.2 Sintesi sottrattiva

La patch che realizza la sintesi sottrattiva è costituita da un generatore di rumore bianco e da quattro filtri passa banda del quarto ordine.¹¹ Gli otto potenziometri della superficie di controllo consentono di controllare, per ogni filtro, la frequenza centrale e il fattore di merito (Q). I filtri sono stati accordati per operare in regioni di frequenza diverse dal più grave, la coppia di potenziometri a sinistra, fino al più acuto a destra. Gli otto potenziometri del prototipo sono stati associati ad altrettanti parametri di controllo della patch secondo il seguente schema:

- A0 - frequenza del primo filtro;
- A1 - Q del primo filtro;
- A2 - frequenza del secondo filtro;
- A3 - Q del secondo filtro;
- ...
- A7 - Q del quarto filtro.

¹⁰ Al seguente link è disponibile un video che presenta il comportamento del prototipo mediante l'uso della patch: <https://youtu.be/Sec7Sf4q34s>.

¹¹ Un video che mostra il funzionamento del prototipo mediante l'uso di questa patch è disponibile al seguente link: <https://youtu.be/1PdLRfr3xwM>.

4. CONTROLLO REMOTO DA SMARTPHONE

Il Raspberry Pi 3 utilizzato per creare il prototipo è dotato di una interfaccia WiFi. Questa possibilità consente, quindi, di controllare Synthberry Pi anche attraverso un tablet, uno smartphone o qualsiasi altro dispositivo dotato di connettività wireless. Sul telefono è necessario installare un'applicazione capace di inviare messaggi di controllo basati sul protocollo OSC. Tra le tante possibilità attualmente disponibili si è scelto di utilizzare un'applicazione chiamata *Control* [11, 12] che fornisce delle interfacce grafiche dotate di oggetti che, in risposta alle azioni dell'utente, inviano messaggi OSC attraverso la connessione WiFi.

4.1 Control

Control¹² consente di creare le proprie interfacce touchscreen per il controllo di parametri musicali. Le interfacce possono essere configurate attraverso la creazione di script JSON utili per definire il comportamento e la disposizione sullo schermo di widget. Le interfacce grafiche possono essere create componendo insieme i seguenti oggetti: cursori, pulsanti e altri oggetti GUI (Graphic User Interface). Tali oggetti sono tutti capaci di inviare dati MIDI e OSC verso dispositivi connessi attraverso la connessione WiFi.

La possibilità di configurazione dei widget di Control consente di definire i namespace OSC utilizzati per l'invio dei messaggi dai diversi oggetti che compongono le interfacce. Nel caso in cui, invece, si vogliono usare le interfacce già fornite dall'applicazione bisogna creare in Pure Data un sistema di ricezione dei messaggi OSC coerente con l'architettura dei namespace utilizzati in Control. Poiché non esiste un metodo standard per definire il namespace OSC, bisogna prestare molta attenzione alla esatta corrispondenza dei namespace utilizzati sul controllore (l'app Control in questo caso) e il sintetizzatore per evitare che errori umani, anche banali come quelli di digitazione, possano impedire la corretta trasmissione e ricezione dei messaggi. Nell'immagine in Figura 1, infatti, si può notare l'oggetto `routeOSC` posto a sinistra impiegato per ricevere i messaggi con un namespace definito mediante un unico livello gerarchico caratterizzato da indirizzi del tipo: `/vol1`, `/vol2`, ... (il numero che segue l'etichetta dell'indirizzo serve a identificare uno specifico potenziometro virtuale dell'interfaccia grafica). Questi messaggi sono inviati da una interfaccia di Control, tra quelle fornite come esempio insieme all'applicazione, che simula il comportamento di un mixer. L'oggetto `routeOSC` posto, invece, a destra consente di ricevere i messaggi con un namespace definito `/multi`. Questi messaggi sono inviati da una interfaccia di Control bidimensionale che supporta il controllo multitouch. In questo caso è da notare che viene utilizzato un namespace con due livelli gerarchici (`/multi/1`, `/multi/2`, ...) per identificare i diversi oggetti grafici forniti dell'interfaccia multitouch. Il tipo di dato inviato in questo caso contiene due numeri che definiscono le coordinate orizzontali e verticali del punto

¹² Per maggiori informazioni sul progetto Control si veda: <http://charlie-roberts.com/Control/>.

in cui l'utente colloca l'oggetto all'interno del widget multitouch. Per tale motivo l'argomento del messaggio OSC viene convertito in una lista di Pd attraverso il messaggio `list $1 $2`. Successivamente, mediante l'oggetto `unpack`, si scompone la lista nei due numeri che consentono d'individuare le coordinate orizzontale e verticale del punto sullo schermo.

Sono state sviluppate due interfacce per il controllo della patch di Pd che realizza il sistema di sintesi sottrattiva. La prima interfaccia ha un funzionamento assimilabile a un mixer, poiché consente di modificare separatamente il volume e la posizione del panning stereo di ognuno dei quattro filtri attraverso l'uso del widget che implementa la superficie multitouch. La dimensione verticale dell'oggetto grafico è stata associata al controllo del volume di ogni filtro, mentre la dimensione orizzontale consente di controllare la distribuzione del suono tra canale sinistro e canale destro.

La seconda è una interfaccia dotata di quattro bottoni che consentono di scegliere quattro diverse impostazioni per il campo di variazione delle frequenze dei filtri. Il valore della frequenza dei quattro filtri è controllato dalla posizione del potenziometro corrispondente mediante un processo di mappatura in un ben definito ambito di frequenze definito mediante un valore massimo e minimo. I quattro filtri sono stati accordati in maniera tale da essere disposti in regioni di frequenza differenti dal grave (sinistra) verso l'acuto (potenziometri di destra). Ogni pulsante dell'interfaccia di Control consente di richiamare nella patch una configurazione differente per i valori estremi (massimi e minimi) delle regioni di frequenza in cui si realizza l'interpolazione. Inoltre, posizionando i potenziometri tutti in alto o tutti in basso è possibile ricreare in maniera precisa la successione di frequenze imposte mediante le diverse configurazioni.

4.2 Connessione dei dispositivi

Per realizzare l'invio dei messaggi OSC da Control verso Pure Data è necessario seguire i seguenti passi:

1. attivare l'hotspot WiFi sul telefono o sul tablet;
2. connettere il Raspberry Pi alla rete WiFi del dispositivo mobile;
3. dopo aver avviato la connessione copiare l'indirizzo IP del Raspberry Pi dal menù di configurazione dell'hotspot del telefono o del tablet;
4. aprire Control e nella voce *destinations* del menù aggiungere l'indirizzo IP del synth, incollando l'indirizzo copiato in precedenza; inserire il numero della porta in maniera coerente a quanto fatto per l'oggetto `udpreceive` di Pd;
5. in Pure Data verificare che l'oggetto `udpreceive` sia configurato con il numero di porta coerente a quello usato in Control.

Questa procedura deve essere realizzata soltanto una volta nella fase iniziale di configurazione del prototipo me-

dante l'utilizzo dell'interfaccia grafica del sistema operativo. In tal modo è possibile configurare opportunamente la gestione delle connessioni di rete e verificare la corretta ricezione ed elaborazione dei messaggi OSC provenienti dal dispositivo mobile. In seguito, invece, è possibile ritornare all'utilizzo del computer in modalità headless, ossia senza interfaccia utente, e avviando il sistema operativo soltanto in modalità console per risparmiare risorse di calcolo. In questa condizione, infatti, è sufficiente attivare l'hotspot sul dispositivo mobile e aspettare che Raspberry Pi riconosca la rete WiFi e automaticamente si connetta per riuscire a ricevere i messaggi OSC.

5. CONCLUSIONI

SynthBerry Pi è un sintetizzatore autonomo basato sull'uso del single-board computer Raspberry Pi e sull'ambiente di sviluppo Pure Data. SynthBerry Pi è stato utilizzato in diverse occasioni per realizzare performance dal vivo e, in studio, per realizzare alcune composizioni. In questo ambito di utilizzo non sono emerse problematiche in merito alla eccessiva latenza o alla perdita di messaggi nella trasmissione dei dati attraverso il protocollo OSC tra dispositivo mobile e i due prototipi di synth. L'utilizzo combinato dell'interfaccia fisica dei prototipi e delle interfacce touchscreen sul telefono o sul tablet sembra essere un'interessante strada per estendere le possibilità di controllo e di interazione delle architetture di generazione ed elaborazione del suono eseguite dal sintetizzatore. Sarebbe molto utile in futuro stabilire delle collaborazioni con musicisti e compositori interessati a sperimentare l'utilizzo di tale approccio nella loro quotidiana pratica artistica al fine di raccogliere un maggiore numero di opinioni e impressioni in merito.

Nel presente lavoro si è presentata l'architettura generale del prototipo e il sistema di comunicazione che rende possibile effettuare il controllo remoto delle patch di Pure Data mediante l'invio di messaggi Open Sound Control. L'interfaccia WiFi di cui è dotato il Raspberry Pi 3, infatti, consente di controllare in remoto le architetture di sintesi mediante dispositivi quali tablet o smartphone.

Per il futuro si pensa di aggiungere un secondo convertitore analogico digitale per disporre di ulteriori otto canali di ingresso analogici. Ciò può consentire di connettere altri otto potenziometri da inserire nella superficie di controllo, oppure, per implementare degli ingressi Control Voltage (CV) utili per controllare dall'esterno i parametri del Synth interfacciandolo con altri dispositivi. Inoltre si pensa di sviluppare ulteriormente la superficie di controllo mediante l'aggiunta di pulsanti, LED, encoder rotativi e di uno schermo LCD.

6. REFERENCES

- [1] J. Noble, *Programming Interactivity. A Designer's Guide to Processing, Arduino and openFrameworks*. Sebastopol, CA: O'Reilly, 2009.
- [2] R. Wilson, *Make: Analog Synthesizers*. Sebastopol, CA: Maker Media, 2013.
- [3] C. Rizzuti and F. Rizzuti, "Synthberry Pi: an autonomous synthesizer based on Raspberry Pi," in *Proceedings of the 18th Linux Audio Conference (LAC-20)*, (Bordeaux), 2020.
- [4] J. Reuter, "Case study: Building an out of the box Raspberry Pi modular synthesizer," in *Proceedings of Linux Audio Conference (LAC14)*, 2014.
- [5] F. Meier, M. Fink, and U. Zölzer, "The JamBerry - a stand-alone device for networked music performance based on the Raspberry Pi," in *Proceedings of Linux Audio Conference (LAC14)*, 2014.
- [6] V. Lazzarini, J. Timoney, and S. Byrne, "Embedded sound synthesis," in *Proceedings of Linux Audio Conference (LAC15)*, 2015.
- [7] H. von Coler and D. Runge, "Teaching sound synthesis in C/C++ on the Raspberry Pi," in *Proceedings of Linux Audio Conference (LAC17)*, 2017.
- [8] M. Puckette, *The Theory and Technique of Electronic Music*. World Scientific, 2007.
- [9] M. Wright, A. Freed, and A. Momeni, "Open Sound Control: State of the art 2003," in *Proc. of the 2003 Conference on New Interfaces for Musical Expression (NIME-03)*, pp. 153–159, 2003.
- [10] M. Wright, "Open Sound Control: an enabling technology for musical networking," *Organised Sound*, vol. 10, no. 3, pp. 193–200, 2005.
- [11] C. Roberts, "Control: Software for end-user interface programming and interactive performance," in *Proceedings of the International Computer Music Conference (ICMC)*, pp. 425–428, 2011.
- [12] C. Roberts, A. Forbes, and T. Höllerer, "Enabling multimodal mobile interfaces for musical performance," in *Proceedings of the New Interfaces for Musical Expression conference (NIME)*, 2013.

UNA PROSPETTIVA CRITICA SUI DATASET PER LA SINTESI TEXT-TO-AUDIO

Riccardo Ancona

Conservatorio Licinio Refice

anconariccardo@gmail.com

ABSTRACT

La sintesi text-to-audio (TTA) promette generazione sonora mediata esclusivamente da linguaggio naturale. Come tutte le tecniche di deep learning generativo, essa deriva la sua memoria del mondo e i suoi confini semantici dal dataset con cui viene addestrata. Seguendo una tradizione umanistica dell'analisi critica degli algoritmi, questo articolo prende in esame i dataset impiegati da diversi modelli di sintesi text-to-audio e il modo in cui essi sono concepiti e manipolati dai ricercatori. I dataset non sono strutture informazionali inerti, bensì oggetti socio-tecnici che implicano specifiche simbolizzazioni mediate da prospettive culturali, politiche e industriali. Uno studio comparato di labeled dataset, caption-based dataset e dataset aumentati algoritmicamente rivela i limiti tecnici ed etici legati a un approccio quantitativo alla raccolta dati. A fronte di un paradigma di valutazione dei dataset di matrice puramente computazionale, si propone una metodologia di analisi qualitativa volta a interpretare la capacità di connotazione dei modelli e suggerire pratiche alternative nella collezione delle informazioni.

1. CONTESTO

Il rapido avanzare delle tecnologie di generazione di immagini da prompt testuali (text-to-image) e dei modelli linguistici di grandi dimensioni (LLM) ha catalizzato le ricerche nel campo della sintesi da testo a suono. Gli investimenti industriali per il perfezionamento di tali tecniche di deep learning hanno permesso a numerosi ricercatori, afferenti principalmente ad aziende tecnologiche, ma anche a centri di ricerca universitari, a tentare di sviluppare algoritmi e piattaforme che permettano di generare suono mediante prompt scritti in linguaggio naturale. Se già dal 2019 erano stati compiuti alcuni esperimenti di sintesi del suono mediante *generative adversarial networks* (GAN [1, 2, 3]), il successo di nuovi paradigmi di architettura per le reti neurali, quali i *modelli di diffusione* nella generazione text-to-image [4] e i *modelli autoregressivi* nei LLM [5], ha portato allo sviluppo di numerose implementazioni adatte al contesto della sintesi text-to-audio. Dalla fine del 2022 nuovi modelli TTA si susseguono a distanza di pochi

mesi o giorni, proponendo perfezionamenti tecnici e differenti architetture, che puntualmente superano i risultati nei test benchmark dei modelli precedenti. Oggetto di analisi per questo studio sono stati sia modelli di diffusione, quali DiffSound [6], Noise2Music [7], AudioLDM [8], MoSai [9], Tango [10], sia modelli autoregressivi, come AudioGen [11], MusicLM [12] e MusicGen [13]. Mediante una lettura comparata degli articoli di ricerca che presentano questi modelli – intesi, secondo gli insegnamenti di Bruno Latour [14], come veri e propri documenti antropologici della ricerca ingegneristica – si osservano comportamenti, presupposizioni e scelte di design caratteristiche del paradigma statistico-inferenziale dominante¹. Nell'ambito di questo articolo si è voluto studiare nello specifico un aspetto cruciale, eppure ancora sottostimato, quello dei dataset impiegati per l'addestramento dei modelli. Si è osservato che gli ingegneri autori di tali ricerche sembrano essere quasi esclusivamente interessati all'architettura delle reti neurali, esaminando molto più sinteticamente (o sorvolando completamente) la questione dei dati impiegati². I dataset sono manufatti tecnici ma anche storici e sociali, il cui contenuto non può mai essere neutrale; al contrario, informa sulla visione del mondo dei suoi creatori³. L'orizzonte conoscitivo dei modelli di deep learning, la loro finestra sul mondo reale, si limita a ciò che i dataset contengono e descrivono. Nel caso della sintesi TTA, dove si tenta di relazionare il dominio acustico con quello testuale, sono di estrema rilevanza non soltanto i dati in sé, ma anche il modo in cui sono interconnessi tra loro. Ci si è proposti dunque di analizzare i dataset impiegati per la realizzazione di modelli TTA, il modo in cui essi sono interpretati e manipolati, nonché l'effetto che producono sui risultati sonori, a partire da una già nutrita letteratura critica impegnata nello studio dei dataset [20, 18, 21]. Melanie Feinberg [22] ha suggerito che una lettura lenta e diretta del contenuto e della struttura dei dataset, contrapposta a tecniche analitiche automatizzate, possa far emergere criticità altrimenti nascoste; Lindsay Poirier [19] ha identificato tre modalità analitiche dei dataset, denotazione, connotazione e decostruzione; il *ground truth tracing* (GTT), proposto da Edward B Kang [23], è un procedimento analitico di tipo umanistico per riportare il contenuto di un dataset alla supposta caratteristica oggettiva a cui si riferisce, in-

¹ Una prospettiva etnografica sugli algoritmi è stata sviluppata in [15] e in [16]. Si confronti anche con l'attenta lettura sociologica di [17].

² Come osservato da Sambasivan et al [18]: "Paradoxically, data is the most under-valued and de-glamorised aspect of AI."

³ Poirier [19] suggerisce che i dataset non vadano intesi come "merely instrumental artifacts tarnished by politics, but as always already iterating cultural artifacts privileging certain symbolic orders over others".

dividuandone il grado di dissonanza ontologica. Un'interpretazione dei dataset per la sintesi text-to-audio informata da questi principi metodologici ha lo scopo di rilevare eventuali criticità e limiti insiti nei dati raccolti e nella loro capacità rappresentativa.

2. DATASET

Gli ingegneri occupati nella realizzazione di algoritmi di deep learning che analizzano testi o immagini (o una combinazione dei due) possono contare su collezioni di dati pressoché illimitate, estratte dalla rete attraverso specifici algoritmi di *web crawling*, che ogni mese setacciano internet e raccolgono trilioni di byte [24]. Di difficile reperibilità sono invece gli accoppiamenti tra suono e descrizioni testuali. Causa di tale scarsità è la struttura del web contemporaneo che incoraggia maggiormente la comunicazione visiva rispetto a quella uditiva, nonché la natura stessa delle informazioni sensoriali, essendo la vista figurativa per definizione, mentre l'udito presenta un quadro epistemologico più sfumato e imprevedibile⁴. È ben più semplice osservare un'immagine e descriverne il contenuto testualmente, che non fare la stessa cosa col suono, i cui contorni restano ambigui. Contribuisce probabilmente anche un'educazione orientata a fornire una terminologia adeguata per gli stimoli visivi, ma spesso carente per ciò che concerne le informazioni uditive. Per sopperire a tali difficoltà, gli ingegneri che si occupano di sintesi TTA devono spesso ricorrere a dati scarsamente connotativi⁵, oppure a tecniche algoritmiche per l'aumento artificiale dei dati [27]. Entrambi i compromessi sollevano dubbi sulla qualità dei dataset impiegati.

2.1 Labeled dataset

Fino a pochi anni fa l'approccio prevalente nel campo della descrizione testuale del suono non contemplava l'utilizzo di intere frasi, bensì di etichette (in inglese *label* o *tag*). La ricerca nel campo del *music information retrieval* prevede l'assegnazione di etichette a grandi corpora di suoni, al fine di addestrare algoritmi di classificazione in grado di distinguere identità sonore, eventi, parole, stati d'animo dei parlanti, generi musicali e stili [28, 29, 30]. La classificazione del suono per etichette è un processo fruttuoso anche per la gestione e navigazione di grandi database, come quelli che raccolgono i suoni stock forniti in licenza per applicazioni pubblicitarie e multimediali. Librerie di suoni come Audiostock [31] ed Epidemic Sound [32], entrambe tutt'ora usate come dataset per l'addestramento di modelli TTA, raccolgono decine di migliaia di file audio classificati secondo categorie funzionali, suddivise in luoghi, tipi d'uso, forme e strumenti musicali, tipologie d'immaginario. La destinazione economica di queste librerie ha un influsso determinante sui loro principi tassonomici, che dedicano ampio spazio a tipologie sonore di frequente utilizzo, come i suoni per videogiochi, i sottofondi per i video d'intrattenimento e i loghi sonori. La

tassonomia di Audiostock dedica un'intera sezione a suoni di armi e combattimenti, suddivisi in ventisei tag, mentre la categoria dedicata alle emozioni consta soltanto di sei descrittori: *feel down*, *fun*, *funny*, *sad*, *scary/fear*, *surprise*. Per quanto questa categorizzazione del suono possa assolvere efficacemente alle funzioni per cui è stata progettata, risulta evidente che ereditare una tale struttura del mondo acustico in un modello TTA sia tutt'altro che auspicabile. Per tentare di descrivere il reale con delle etichette, è necessaria una chiara tassonomia, derivata da un'ontologia del suono che non è mai neutrale, ma sempre frutto di scelte ingegneristiche o industriali precise, le quali a loro volta riflettono prospettive culturali implicite ma non irrilevanti. Nel tentativo di costruire un dataset di tag fondativo e quanto più neutrale, Google ha sviluppato Audioset [33], un corpus di 1789621 segmenti audio di dieci secondi, estrapolati da video Youtube ed etichettati da ascoltatori umani seguendo una tassonomia predefinita di 527 parole chiave. Il gruppo di ricerca ha impiegato tecniche statistiche per inferire l'ontologia sonora di Audioset, in primo luogo collezionando gli eponimi del suono più ricorrenti sul web, dunque eliminando le ridondanze e costruendo una gerarchia tassonomica, ulteriormente filtrata verificando che ciascuna tag fosse statisticamente rappresentativa e intelligibile dagli operatori umani. La tassonomia ottenuta non è pertanto un ordinamento ontologico universale, bensì il riflesso statistico di ciò che in un dato momento temporale risulta essere quantitativamente più rappresentato nel discorso digitale in lingua inglese. Ciononostante Audioset rappresenta al momento il dataset più ricco e significativo per numerose applicazioni che riguardano la classificazione automatica e il riconoscimento dei suoni⁶. Ai fini dell'addestramento di un modello TTA, l'uso di etichette appare invece limitante. Lo scopo dichiarato dagli sviluppatori dei modelli più recenti è l'interazione per mezzo di frasi strutturate scritte in linguaggio naturale, che definiscono *caption*, come "qualcuno suona virtuosamente arpeggi discendenti su un arciliuto, all'improvviso passa un treno e migliaia di tortore volano via". Ammesso che il modello sia effettivamente in grado di interpretare il linguaggio naturale, ridurre questa frase alle tag Audioset "plucked string instrument, train, flapping wings" non sortirebbe lo stesso risultato, perché una tale semplificazione perde tutte le specificità della connotazione testuale: il grado di competenza dello strumentista, il tipo di figura melodica prodotta, il timbro specifico dello strumento impiegato, la consequenzialità temporale degli eventi, il numero e la specie di uccelli – mancano, in sostanza, tutte le relazioni logiche, causali, temporali, contestuali, come di una lingua fatta solo di sostantivi e aggettivi, ma priva di avverbi e complementi. Inoltre un labeled dataset può constare di qualche centinaia o migliaia di tag, un numero insufficiente a etichettare la totalità del mondo acustico; incrementi di scala nel numero di tag sono fortemente impraticabili. Per questo è necessario che nell'addestramento di modelli TTA si impieghino dataset di *caption*.

⁴ È ciò che sostiene Salomé Voegelin in [25]: "Listening's focus on the dynamic nature of things renders the perceptual object unstable, fluid and ephemeral."

⁵ In [26] vengono esplicitamente definiti "weakly-associated".

⁶ Alternative, seppur in scala ridotta, sono i dataset FSD-50K [34] e MagnaTagATune [35].

2.2 Caption-based dataset

Il più ampio dataset open source contenente coppie di caption testuali e suoni è opera di LAION, un gruppo no-profit intento a fornire dataset accessibili e trasparenti⁷. LAION-Audio-630K [37], formato da 633526 caption per un totale di oltre 4000 ore di audio, è un aggregato di alcune delle più grandi librerie sonore gratuite, quali Freesound [38], BBC Sound Effects [39] e i già citati Audiostock ed Epidemic Sound. La maggioranza delle caption, circa mezzo milione, provengono da Freesound, un database collaborativo attivo dal 2005, nel quale chiunque può condividere i propri suoni con licenza Creative Commons. Al suo interno si trovano field recordings, frammenti vocali, musica, suoni oggettuali, effetti sonori, timbri elettronici. Si tratta di un importantissimo manufatto di condivisione acustica che meriterebbe una ricerca antropologica approfondita. Bisogna domandarsi, tuttavia, se i dati in esso contenuti siano opportuni allo scopo di addestrare algoritmi di connotazione testuale del suono. Sfogliando il dataset, ci si può imbattere in caption contenenti link promozionali o precisazioni sui diritti d'autore ("For commercial purposes, please visit [http:\[ecc.\]](http://[ecc.])"), dati di natura diaristica o personale ("This one was interesting to create"), informazioni tecniche eccessivamente accurate ("Edited with TDR Kotelnikov and Waves Loudness Meter"), contestualizzazioni troppo solerti ("Date/time: June 21th 2022, 7:34 PM. Weather: 20c, clouds 67%, wind E - NE 2 bft with gusts to 13km/h, 1014 hPa, humidity 50 - 60%"), commenti agli ascoltatori ("Hey guys new track", "I would normally explain how I recorded it, but I want to see if anyone can guess how I did it"). Questo genere di annotazioni abbondano nel dataset e sono usate per l'addestramento di modelli text-to-audio⁸, introducendo un gran numero di informazioni scarsamente significative in quanto poco o per nulla connotative del contenuto sonoro associato. Unico sottoinsieme di caption in LAION-Audio-630K raccolte con una metodologia formalizzata è la libreria BBC Sound Effects, che tuttavia ne rappresenta soltanto una porzione estremamente esigua e dunque poco rappresentativa dal punto di vista statistico⁹. Un'alternativa a LAION è Clotho [40], un dataset ottenuto selezionando 4981 suoni da Freesound e chiedendo a esseri umani di descriverli con annotazioni testuali non vincolate da specifiche prescrizioni. Le persone preposte a tale compito, tutte parlanti la lingua inglese, sono state ingaggiate mediante la piattaforma di crowdsourcing Amazon Mechanical Turk¹⁰. Il dataset consta di 24905 caption, un numero considerato generalmente insufficiente per la sintesi text-to-audio – difatti, nessuno dei modelli studiati lo usa in maniera esclusiva¹¹. In mancanza di dataset gratuiti sufficientemente estesi, per

il suo modello TTA incentrato sulla generazione di musica strumentale Meta ha acquisito licenze da Shutterstock Music [43] e Pond5 [44], due servizi di musica stock generalmente impiegati per l'intrattenimento multimediale. Osservando i dati forniti dal paper [13], risulta che il dataset sia costituito prevalentemente da suoni etichettati come "electronic/edm" e "corporate", il che riflette la destinazione per cui sono stati originariamente realizzati. Un quadro complessivo dei dataset a disposizione per l'addestramento di modelli TTA mostra dunque un'evidente carenza di dati disponibili e una scarsa diversità al di fuori di categorie sonore fortemente funzionalizzate. La quantità di tempo e risorse necessarie per annotare con pertinenza terminologica centinaia di migliaia o milioni di file audio eccede le possibilità dei gruppi di ricerca coinvolti, costretti pertanto a fare uso di dataset di bassa qualità o a dover trovare metodologie algoritmiche non supervisionate per estrarre moli di dati che non si ha il tempo di far ascoltare a orecchie umane.

2.3 Dataset aumentati

Una tecnica di generazione artificiale dei dataset consiste nell'annotazione automatica delle caption mediante algoritmi di classificazione. VGG-Sound [45] è un dataset audiovisivo costituito da 200000 caption prodotte artificialmente da un algoritmo di riconoscimento visivo. Esso analizza il contenuto di clip YouTube raccolte mediante web scraping e verifica la presenza di entità nelle immagini, restringendo il campo di etichette possibili a oggetti e azioni concrete aventi un chiaro corrispettivo visivo. Si tratta di una forma di classificazione intermodale in grado di denotare eventi ed enti predefiniti, ma non di connotarli linguisticamente: le annotazioni ottenute somigliano più a delle tag che a delle caption in linguaggio naturale. Inoltre, la mera presenza visiva di un oggetto non garantisce che esso stia producendo suono durante la clip, né esclude che nella clip siano presenti suoni di altri oggetti o processi non riscontrati dall'immagine. Per sopperire a tale problematica, un algoritmo di classificazione del suono tenta di escludere i falsi positivi. Al netto di quanto osservato, sembrerebbe che chiedere a un algoritmo sordo ma con un'ottima vista di riconoscere i suoni non sia la migliore soluzione. Una metodologia alternativa consiste nel trasformare un labeled dataset come Audioset in un caption-based dataset attraverso l'utilizzo di un algoritmo di processamento testuale, come il modello linguistico T5 [46]. Il modello TTA AudioLDM ha sfruttato questa procedura *key-to-text* per trasformare Audioset in un dataset di caption [8]. Non è chiaro tuttavia in che modo un algoritmo di processamento del testo possa giungere a una maggiore connotazione del suono, non avendo esso alcuna esperienza del contenuto audio relazionato. Un tale procedimento sarebbe forse utile a sintetizzare un testo lungo in uno più breve e puntuale, ma non può di certo aggiungere informazione che non esiste nelle tag fornite, tutt'al più può diluirla in un maggior numero di parole. L'esempio mostrato sulla pagina Github dell'algoritmo di estensione testuale indica come tag "India, Capital, New Delhi" e dà in output il testo "The capital of India is New Delhi" [47]. Un test

⁷ Si noti che LAION è stato altrove criticato per la metodologia di raccolta dati, ad es. in [36].

⁸ Lo impiega ad esempio AudioLDM [8] come dataset primario.

⁹ La libreria BBC contiene caption del tipo: "Church Atmosphere - Church congregation assembling before service with bells in background".

¹⁰ Data la paga media infima [41], questo tipo di crowdsourcing è stato talora considerato una forma di colonialismo mediante intelligenze artificiali [42].

¹¹ Lo impiegano Audiogen [11] e AudioLDM [8] in aggiunta ad altri dataset.

più approfondito usando come input delle tag arbitrarie di Audioset ne mette in evidenza i limiti: le etichette “cash register, sobbing, boiling” danno come risultato “A man is sobbing and boiling cash at a cash register”, con “harp, whistling, jingle” si ottiene “A man is playing a jingle and whistling on the harp”, con “bassline, didgeridoo, singing bowl, tender music” si ha “A young boy is playing the bassline of a didgeridoo while a young girl is singing”. In luce di quanto ottenuto dai test empirici, questo tipo di tecnica sembra essere più un modo di inquinare i dati che di aumentarli. Risulta evidente che impiegare il testo o l’immagine per espandere i dataset sia poco fruttuoso, poiché si perde la relazione diretta con l’informazione acustica. Una soluzione più efficace consiste nell’utilizzo di metodologie di codifica che proiettino sia il testo sia il suono in un unico spazio latente, vale a dire in una rappresentazione numerica condivisa. Una volta definite le proprietà dello spazio latente multimodale sulla base di un addestramento preliminare – a sua volta compiuto con informazioni provenienti da un caption-based dataset, ma di dimensioni inferiori – è possibile usare questo modello (detto modello di *pre-training*) per generare automaticamente caption avendo in input un qualsiasi suono. Questo tipo di procedura è stata inventata per la relazione tra immagini e linguaggio naturale ed è una forma di apprendimento contrastivo auto-supervisionato [48]. MuLAN [26] e CLAP [49] sono due implementazioni per il suono, rispettivamente impiegate per addestrare i modelli text-to-audio MusicLM [12] e AudioLDM [8]. Entrambi gli algoritmi di pre-training dapprima comprimono le informazioni usando un encoder per i testi (BERT o suoi derivati, [50]) e uno per i suoni (variabile a seconda delle diverse implementazioni), ottenendo i cosiddetti *embedding*, cioè dei vettori multidimensionali contenuti nello spazio latente, dunque impiegano l’apprendimento contrastivo per trovare delle relazioni tra due insiemi numerici. Una volta addestrati, è sufficiente fornire l’embedding di un nuovo suono per ottenere l’inferenza di un possibile embedding testuale; quest’ultimo non prende mai la forma di un testo intelligibile da un essere umano, ma viene immediatamente immagazzinato per essere usato come dato per l’addestramento del modello text-to-audio. Tale procedura riduce la problematica delle caption a una questione totalmente numerica, sostituendo le annotazioni testuali con rappresentazioni latenti difficili da interpretare da un essere umano, eliminando così la possibilità di verificare fenomenologicamente i risultati ottenuti durante la fase di pre-training. I dataset impiegati non sono più studiabili con un metodo qualitativo, ma dipendono da altri algoritmi di basso livello, a loro volta addestrati usando altri dataset, la cui provenienza non è sempre specificata in maniera chiara. I ricercatori che hanno sviluppato MuLAN dichiarano di aver usato 44 milioni di caption scarsamente connotate, raccolte in maniera automatizzata da video musicali trovati su internet. Dal momento che non esistono dataset pubblici così ampi, è verosimile che siano state usate come caption le descrizioni di video YouTube; le quali, notoriamente, non sono sempre connotative del contenuto sonoro, poiché spesso rimandano a link esterni, promozioni, descrizioni non pertinenti direttamente al suo-

no e riflessioni a latere. Il presupposto logico implicito è che anche se le annotazioni potrebbero contenere molte informazioni non rilevanti o errate, il mero fatto di estrarre una tale mole di dati tenderà a diluire statisticamente gli errori e a espandere il campo semantico rappresentato: una soluzione puramente quantitativa a un problema sia di natura quantitativa sia qualitativa. Un numero così grande di video raccolti sembra anche aprire all’eventualità che alcuni dei materiali presenti potrebbero essere coperti da copyright; tuttavia un algoritmo come MuLAN dissolve ogni traccia del suo dataset di origine, perché lo riduce a vettori matematici usati in fasi intermedie della produzione di un modello TTA. Non c’è modo di stabilire a posteriori con certezza assoluta che un dato video musicale sia stato usato per l’addestramento, in quanto esso ha soltanto la funzione di informare il meccanismo di generazione automatica delle caption, ma non viene usato direttamente come esempio per l’apprendimento della generazione sonora. Eppure le nozioni linguistiche che verranno usate per descrivere il dataset di suoni sono ereditate dalle informazioni originariamente impiegate per l’addestramento del modello di pre-training: non sopravvivono i dati specifici, ma se ne deriva una rappresentazione numerico-semantica del mondo. Usando questi modelli di pre-training, si può addestrare un modello TTA senza fare ricorso a dataset di caption, sfruttando invece grandi raccolte di musica e suoni non annotate. Qualunque file audio diventa materiale potenziale per l’addestramento, senza alcun vincolo di annotazione, una notevole conquista tecnica non priva di criticità. MusicLM [12] segue questa metodologia e dichiara di aver usato cinque milioni di file musicali, senza specificarne la provenienza¹². L’opacità di questa procedura fa in modo che non sia possibile stabilire a posteriori la presenza di una traccia musicale nel dataset, eccetto in casi di memorizzazione da parte dell’algoritmo, che gli autori del paper hanno misurato essere statisticamente rarissima¹³. La segretezza industriale circa i dataset impiegati, verosimilmente imposta per evitare problematiche di copyright e per non informare la concorrenza, provoca anche l’impossibilità di compiere uno studio approfondito della natura dei dataset, valutando qualitativamente la varietà dei contenuti, il grado di connotazione delle caption, nonché la presenza di eventuali bias.

Table 1 mostra i dataset impiegati dai modelli TTA presi in esame; sono stati esclusi dallo studio i modelli a uso commerciale che non forniscono alcuna informazione sulla provenienza dei dati, come Suno [51] e Udio [52].

3. CONNOTAZIONE

Se la frequente omissione di informazioni sulla provenienza dei dati e l’artificialità delle caption generate dagli algo-

¹² Ho chiesto chiarimenti sull’origine dei dati agli autori dell’articolo, ma non ho ricevuto risposta.

¹³ Ho compiuto dei test nella versione distribuita da Google di MusicLM, MusicFX, dove ho riscontrato che la piattaforma analizza i prompt e li blocca se individua stringhe di caratteri corrispondenti a nomi di musicisti noti. Il prompt “Meshuggah”, nome di un gruppo djent metal svedese, nonché sequenza di caratteri altrimenti poco ricorrente, non viene bloccato e produce generazioni aventi sonorità estremamente simili alla musica del gruppo. Questo non dimostra con certezza assoluta che la loro musica sia stata impiegata nell’addestramento, poiché non si ottengono corrispondenze 1:1.

| Modello | Dataset |
|-------------|--|
| DiffSound | AudioCaps, AudioSet |
| AudioGen | AudioCaps, AudioSet, Clotho, VGG-Sound, FSD-50k e altre librerie di sound effects |
| MusicLM | Embedding testuali generati da MuLan usando più di 280000 ore di musica |
| AudioLDM | CLAP addestrato con LAION-Audio-630K, AudioSet aumentato da T5, AudioCaps, Clotho |
| Noise2Music | Caption generate da un LLM e assegnate automaticamente a suoni mediante MuLan |
| Moûsai | 7000 “Spotify top recommendations” con caption estratte automaticamente dai metadati YouTube |
| MusicGen | Shutterstock, Pond5 |
| Tango | AudioCaps |

Table 1. Dataset impiegati dai modelli text-to-audio.

ritmi di pre-training rendono talora impossibile una valutazione qualitativa diretta del contenuto dei dataset, l'unico metodo per indagarne le proprietà è osservarne i risultati a posteriori, studiando le generazioni sonore ottenute dai modelli TTA addestrati. Come valutare i risultati di un modello di sintesi text-to-audio? La questione è tutt'altro che banale e richiede una chiara definizione di quali sono gli obiettivi perseguiti nell'addestramento del modello. Comparando le diverse implementazioni prese in esame, sembrerebbe che uno degli scopi prevalenti sia la generazione di suono per mezzo del linguaggio ordinario, senza necessità di ulteriori mediazioni né competenze a carico dell'utente, cercando quindi di operare una traduzione diretta dal pensiero verbale al suono. Si presume quindi che un modello propriamente programmato e addestrato sia capace, per il solo mezzo dell'inferenza statistica, di esibire un comportamento di tipo interpretativo nei confronti del prompt scritto dall'utente e che di conseguenza produca un risultato sonoro quanto più affine alla richiesta. I paper di alcuni modelli forniscono esempi di caption impiegate per l'addestramento, supponendo che tali modelli ne abbiano assimilato le informazioni e siano in grado di rappresentarle efficacemente. Una caption come “enchanting jazz song with a memorable saxophone solo and a solo singer” (menzionata in [12]), aggettivando con “enchanting” e “memorable” sembra intendere che lo scopo del modello TTA non sia soltanto di tipo denotativo, ma anche connotativo. Per stabilire l'efficacia della generazione è necessario dunque valutare quanto essa sia propriamente connotativa del prompt dato. Il che, naturalmente, è tutt'altro che semplice da stabilire in termini obiettivi, perché non sempre esiste una *ground truth* a cui fare riferimento: se domando “musica ruvida ma spiritosa per saltare la corda”, è evidente che sto chiedendo all'algoritmo qualcosa di puramente

qualitativo, per il quale io stesso non ho che un'idea approssimativa. Il linguaggio del prompt in questo caso è un'allusione, non ha valore puramente indicale. A meno che non si forniscano nel prompt indicazioni quantitative inequivocabili, come “un'onda quadra a 726.3 Hz”, le richieste espresse in linguaggio naturale chiamano in causa referenti dalla natura ambigua e polivalente; a tali richieste, l'algoritmo tenterà di generare un contenuto coerente sulla base della mera rappresentazione statistica dei dati forniti, ragion per cui la capacità connotativa di un modello TTA, al netto delle diverse architetture e implementazioni, dipende dal modo in cui è connotante il dataset di addestramento. Se però il dataset è ignoto oppure è a sua volta il risultato di un processo computazionale, non rimane che compiere il percorso inverso, valutando quanto il risultato sonoro risulti essere connotativo del prompt fornito. Nei modelli di deep learning generativo la forma del prompt ha un impatto sostanziale sui risultati possibili, al punto che esiste una disciplina preposta allo studio delle formalizzazioni linguistiche, il *prompt engineering* [53]. Da studi relativi ai modelli di diffusione text-to-image risulta – e lo si può verificare empiricamente anche in tutti i modelli text-to-audio disponibili – che questi algoritmi non mostrano di saper interpretare né la grammatica né la sintassi del prompt, ignorando o confondendo negazioni, relazioni e consequenzialità [54]. Prendendo atto di tale limite intrinseco, forse valicabile con miglioramenti tecnici delle architetture dei modelli o con un mero incremento di scala dei dati impiegati per l'addestramento, si può comunque tentare di testare l'efficacia connotativa dei risultati sonori specificando nei prompt soggetti, eventi e aggettivazioni. Una valutazione empirica approfondita dell'efficacia connotativa dei vari modelli a seconda del tipo di prompt esula dalle possibilità e dagli scopi di questo articolo, tuttavia è opportuno affermare che la valutazione delle generazioni text-to-audio intesa in questo senso richiede una capacità interpretativa di tipo fenomenologico e qualitativo, per sua natura soggettiva e personale. Quando si domanda “un suono vetroso”, si richiama una categoria esperienziale derivante da una complessa stratificazione di memoria e percezione, ottenuta da una pluralità di esperienze private, soggettive e talvolta contraddittorie tra di loro, che insieme formano qualità uditive di complessa definizione¹⁴. Tali identità sonore divergono tra gli individui e non è detto che una loro media statistica sia sufficiente a coglierne la complessità. La valutazione dell'efficacia connotativa di un modello dovrebbe perciò basarsi sulla negoziazione tra la percezione soggettiva, le proprie aspettative e le possibilità tecniche del modello, mettendo al centro l'esperienza fenomenologica di chi compie l'analisi. Confrontando i risultati ottenuti con uno studio sia qualitativo sia statistico dei dataset, è possibile trovare delle correlazioni tra risultati inaspettati e aporie nei dati, verificando l'efficacia del modello a seconda dello scopo prefissato e individuando soluzioni specifiche per la redazione di un dataset maggiormente rappresentativo. Se la connotatività di un modello dipende in larga parte dal suo dataset, le osservazioni raccolte durante l'analisi fenomenologica dei risultati dovrebbero infor-

¹⁴ Ho affrontato questo argomento in [55].

mare la creazione di nuovi dataset, innescando un circuito di feedback tra design algoritmico e percezione. Quanto osservato sinora nella ricerca sulla sintesi text-to-audio sembra tenere molto poco in considerazione questo genere di metodologia analitica qualitativa, affidando la valutazione dell'efficacia dei modelli a tecniche quantitative, come l'uso di *benchmark test*. Alcuni modelli, come MusicLM e AudioLDM, hanno compiuto test con esseri umani, ma sempre di natura quantitativa, domandando agli ascoltatori quale suono preferissero tra due opzioni e comparando il numero di preferenze, oppure chiedendo di indicare il grado di "qualità" di un suono su una scala numerica data¹⁵. Tra le metriche di benchmark più impiegate per la sintesi text-to-audio si menzionano la Fréchet Audio Distance [56] e la divergenza di Kullback-Leibler [57]. Se queste metriche possono fornire risultati utili in termini statistici per il miglioramento delle architetture, sembra tuttavia paradossale delegare il compito della valutazione delle capacità connotative e interpretative del linguaggio naturale di un algoritmo a un altro algoritmo. Misurare l'efficacia di un modello con un benchmark permette agli ingegneri di stabilire delle metriche quantitative da superare, deviando l'obiettivo della ricerca dal miglioramento effettivo dei risultati al semplice superamento di soglie numeriche precedentemente raggiunte da altri algoritmi. Seguendo tale criterio si rischia di addestrare modelli che soddisfano le richieste degli algoritmi di benchmark, ma non quelle degli utenti. L'idea che l'efficacia nella connotazione linguistica di un modello possa essere misurata quantitativamente presenta limiti evidenti se si considera la natura incommensurabile del linguaggio. Soltanto una valutazione di tipo esperienziale può fornire un'indicazione significativa delle caratteristiche qualitative dei risultati sonori; se la denotazione è tutto sommato computabile come rappresentazione media di un referente, la connotazione è per sua natura allusiva, ragion per cui solo un'entità in grado di cogliere le allusioni può esaminarne la qualità.

4. CONCLUSIONI

Uno studio comparato dei dataset e dei modelli di sintesi text-to-audio ha rilevato una scarsa curatela dei dati, derivante da un'obiettivo insufficiente di informazioni significative a disposizione e dall'uso di tecniche di aumento artificiale dei dati di dubbia efficacia. Si sono riscontrate possibili problematiche etiche legate alla raccolta dei dati, il cui processo è spesso opaco, nonché l'impiego di sistemi di valutazione dell'efficacia dei modelli di tipo puramente quantitativo, i cui risultati non rappresentano necessariamente le loro capacità di connotazione. In luce di quanto osservato, si è proposto un possibile cambio di paradigma nella raccolta dei dati e nella valutazione dei modelli, orientato all'analisi qualitativa e fenomenologica delle informazioni. Si suggerisce inoltre l'inclusione di musicisti, musicologi e semiologi all'interno dei gruppi di ricerca, al fine di orientare il design dei modelli e i pro-

¹⁵ I test soggettivi quantitativi potrebbero costituire una metrica significativa per l'analisi dei modelli se la loro valutazione si basasse su un protocollo formalizzato, che dovrebbe tener conto delle teorie cognitive contemporanee e delle perturbazioni indotte nelle misurazioni dal contesto e dal formato dei test.

cessi decisionali a partire da punti di vista e competenze differenziate.

5. REFERENCES

- [1] K. Kumar, R. Kumar, T. de Boissiere, L. Gestin, W. Z. Teoh, J. Sotelo, A. de Brebisson, Y. Bengio, and A. Courville, "MelGAN: Generative Adversarial Networks for Conditional Waveform Synthesis," Dec. 2019. arXiv preprint.
- [2] J. Nistal, S. Lattner, and G. Richard, "DrumGAN: Synthesis of Drum Sounds With Timbral Feature Conditioning Using Generative Adversarial Networks," June 2022. arXiv preprint.
- [3] J. Kong, J. Kim, and J. Bae, "HiFi-GAN: Generative Adversarial Networks for Efficient and High Fidelity Speech Synthesis," Oct. 2020. arXiv preprint.
- [4] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," Apr. 2022. arXiv preprint.
- [5] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language Models are Few-Shot Learners," July 2020. arXiv preprint.
- [6] D. Yang, J. Yu, H. Wang, W. Wang, C. Weng, Y. Zou, and D. Yu, "DiffSound: Discrete Diffusion Model for Text-to-sound Generation," 2022. arXiv preprint.
- [7] Q. Huang, D. S. Park, T. Wang, T. I. Denk, A. Ly, N. Chen, Z. Zhang, Z. Zhang, J. Yu, C. Frank, J. Engel, Q. V. Le, W. Chan, Z. Chen, and W. Han, "Noise2Music: Text-conditioned Music Generation with Diffusion Models," Mar. 2023. arXiv preprint.
- [8] H. Liu, Z. Chen, Y. Yuan, X. Mei, X. Liu, D. Mandic, W. Wang, and M. D. Plumbley, "AudioLDM: Text-to-Audio Generation with Latent Diffusion Models," Sept. 2023. arXiv preprint.
- [9] F. Schneider, O. Kamal, Z. Jin, and B. Schölkopf, "Moûsai: Text-to-Music Generation with Long-Context Latent Diffusion," Oct. 2023. arXiv preprint.
- [10] D. Ghosal, N. Majumder, A. Mehrish, and S. Poria, "Text-to-Audio Generation using Instruction-Tuned LLM and Latent Diffusion Model," May 2023. arXiv preprint.
- [11] F. Kreuk, G. Synnaeve, A. Polyak, U. Singer, A. Défossez, J. Copet, D. Parikh, Y. Taigman, and Y. Adi, "AudioGen: Textually Guided Audio Generation," Mar. 2023. arXiv preprint.

- [12] A. Agostinelli, T. I. Denk, Z. Borsos, J. Engel, M. Verzetti, A. Caillon, Q. Huang, A. Jansen, A. Roberts, M. Tagliasacchi, M. Sharifi, N. Zeghidour, and C. Frank, “MusicLM: Generating Music From Text,” Jan. 2023. arXiv preprint.
- [13] J. Copet, F. Kreuk, I. Gat, T. Remez, D. Kant, G. Synnaeve, Y. Adi, and A. Défossez, “Simple and Controllable Music Generation,” Jan. 2024. arXiv preprint.
- [14] B. Latour, *Laboratory life: the construction of scientific facts*. Princeton, N.J: Princeton University Press, 1986.
- [15] N. Seaver, “What Should an Anthropology of Algorithms Do?,” *Cultural Anthropology*, vol. 33, pp. 375–385, Aug. 2018.
- [16] A. Christin, “The ethnographer and the algorithm: beyond the black box,” *Theory and Society*, vol. 49, pp. 897–918, Oct. 2020.
- [17] M. Airoidi, *Machine Habitus. Sociologia degli algoritmi*. LUISS, 2024.
- [18] N. Sambasivan, S. Kapania, H. Highfill, D. Akrong, P. Paritosh, and L. M. Aroyo, ““Everyone wants to do the model work, not the data work”: Data Cascades in High-Stakes AI,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, (Yokohama Japan), pp. 1–15, ACM, May 2021.
- [19] L. Poirier, “Reading datasets: Strategies for interpreting the politics of data signification,” *Big Data & Society*, vol. 8, July 2021.
- [20] F. Jatón, *The constitution of algorithms: ground-truthing, programming, formulating*. Inside technology, Cambridge, Massachusetts: The MIT Press, 2020.
- [21] W. Orr and K. Crawford, “The social construction of datasets: On the practices, processes and challenges of dataset creation for machine learning,” Nov. 2023. knowingmachines.org.
- [22] M. Feinberg, “Reading databases: slow information interactions beyond the retrieval paradigm,” *Journal of Documentation*, vol. 73, pp. 336–356, Mar. 2017.
- [23] E. B. Kang, “Ground truth tracings (GTT): On the epistemic limits of machine learning,” *Big Data & Society*, vol. 10, p. 205395172211461, Jan. 2023.
- [24] “Common Crawl.” <https://commoncrawl.org/>.
- [25] S. Voegelin, *Listening to Noise and Silence*. Continuum, 2010.
- [26] Q. Huang, A. Jansen, J. Lee, R. Ganti, J. Y. Li, and D. P. W. Ellis, “MuLan: A Joint Embedding of Music Audio and Natural Language,” Aug. 2022. arXiv preprint.
- [27] K. Drossos, S. Adavanne, and T. Virtanen, “Automated Audio Captioning with Recurrent Neural Networks,” Oct. 2017. arXiv preprint.
- [28] K. Choi, G. Fazekas, and M. Sandler, “Automatic tagging using deep convolutional neural networks,” June 2016. arXiv preprint.
- [29] Y. Xu, Q. Kong, W. Wang, and M. D. Plumbley, “Large-scale weakly supervised audio classification using gated convolutional neural network,” Oct. 2017. arXiv preprint.
- [30] S. Vishnupriya and K. Meenakshi, “Automatic Music Genre Classification using Convolution Neural Network,” in *2018 International Conference on Computer Communication and Informatics (ICCCI)*, (Coimbatore), pp. 1–4, IEEE, Jan. 2018.
- [31] “Audiostock.” <https://audiostock.net/>.
- [32] “Epidemic Sound.” <https://www.epidemicsound.com/>.
- [33] J. F. Gemmeke, D. P. W. Ellis, D. Freedman, A. Jansen, W. Lawrence, R. C. Moore, M. Plakal, and M. Ritter, “Audio Set: An ontology and human-labeled dataset for audio events,” in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (New Orleans, LA), pp. 776–780, IEEE, Mar. 2017.
- [34] E. Fonseca, X. Favory, J. Pons, F. Font, and X. Serra, “FSD50K: An Open Dataset of Human-Labeled Sound Events,” 2020. arXiv preprint.
- [35] E. Law, K. West, M. I. Mandel, M. Bay, and J. Stephen Downie, “Evaluation of Algorithms Using Games: The Case of Music Tagging,” *10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Oct. 2009.
- [36] C. Buschek and J. Thorp, “Models All the Way Down,” 2024. <https://knowingmachines.org/models-all-the-way>.
- [37] “Laion-audio-630k,” 2022. <https://github.com/LAION-AI/audio-dataset>.
- [38] F. Font, G. Roma, and X. Serra, “Freesound technical demo,” in *Proceedings of the 21st ACM international conference on Multimedia*, (Barcelona Spain), pp. 411–412, ACM, Oct. 2013.
- [39] “Bbc sound effects.” <https://sound-effects.bbcrewind.co.uk/>.
- [40] K. Drossos, S. Lipping, and T. Virtanen, “Clotho: an Audio Captioning Dataset,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Barcelona, Spain), pp. 736–740, IEEE, May 2020.
- [41] K. Hara, A. Adams, K. Milland, S. Savage, C. Callison-Burch, and J. Bigham, “A Data-Driven Analysis of Workers’ Earnings on Amazon Mechanical Turk,” Dec. 2017. arXiv preprint.

- [42] K. Hao and A. P. Hernández, “How the AI industry profits from catastrophe,” *MIT Technology Review*, Apr. 2022.
- [43] “Shutterstock music.”
<https://www.shutterstock.com/music>.
- [44] “Pond5.” www.pond5.com.
- [45] H. Chen, W. Xie, A. Vedaldi, and A. Zisserman, “VGGSound: A Large-scale Audio-Visual Dataset,” Sept. 2020. arXiv preprint.
- [46] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, “Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer,” 2019. arXiv preprint.
- [47] G. Bhatia, “keytotext.”
<https://github.com/gagan3012/keytotext>.
- [48] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, “Learning Transferable Visual Models From Natural Language Supervision,” Feb. 2021. arXiv preprint.
- [49] B. Elizalde, S. Deshmukh, M. A. Ismail, and H. Wang, “CLAP: Learning Audio Concepts From Natural Language Supervision,” June 2022. arXiv preprint.
- [50] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” May 2019. arXiv preprint.
- [51] “Suno.” <https://suno.com/>.
- [52] “Udio.” <https://www.udio.com/>.
- [53] M. Diab, J. Herrera, M. Sleep, B. Chernow, and C. Mao, “Stable Diffusion Prompt Book,” tech. rep., OpenArt, Oct. 2022.
- [54] H.-H. Wu, O. Nieto, J. P. Bello, and J. Salamon, “Audio-text models do not yet leverage natural language,” in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Rhodes Island, Greece), IEEE, 2023.
- [55] R. Ancona, “Material identities in corpus-based algorithmic improvisation,” Master’s thesis, Institute of Sonology, 2023.
- [56] K. Kilgour, M. Zuluaga, D. Roblek, and M. Sharifi, “Fréchet Audio Distance: A Metric for Evaluating Music Enhancement Algorithms,” Jan. 2019. arXiv preprint.
- [57] S. Kullback and R. A. Leibler, “On Information and Sufficiency,” *The Annals of Mathematical Statistics*, vol. 22, pp. 79–86, Mar. 1951.

Considerazioni su VCV Rack come piattaforma didattica per l'ingegnere e il musicista

Leonardo Gabrielli

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

ABSTRACT

Le istituzioni accademiche che insegnano materie come l'elaborazione numerica del suono sono piuttosto ridotte e adottano tagli diversi in funzione del tipo di formazione richiesta: da quella dell'ingegnere a quella del compositore di musica elettroacustica. Nonostante ciò, la quantità di linguaggi e piattaforme software utili alla didattica è estremamente varia, rendendo difficile la scelta per il docente che stia preparando nuovi materiali didattici. In questo articolo viene discussa la potenziale adozione di un software relativamente recente, VCV Rack, come strumento che può servire un buon numero di diversi contesti didattici. Rack è un software open-source sviluppato da VCV che emula un sistema modulare Eurorack dove gli utenti possono caricare plugin scritti in C++ oltre a quelli forniti dalla community. La programmazione di tali plugin è ideale per l'ingegnere che voglia cimentarsi con la programmazione di moduli DSP, mentre la presenza di un ambiente di lavoro modulare rende il software adatto all'improvvisazione musicale, alla sintesi e alla composizione aleatoria. Alcune esperienze dell'autore vengono riportate insieme a suggerimenti e ad una panoramica dei vantaggi e degli svantaggi.

1. INTRODUZIONE

L'Elaborazione Digitale del Segnale (nel seguito utilizzeremo il comune acronimo inglese DSP) è una materia consolidata all'interno dei corsi di laurea in ingegneria elettronica e dell'informazione, che viene insegnata anche in altri ambiti formativi, come i conservatori o le scuole private che formano tecnici per il suono. Essa comprende scoperte fondamentali del ventesimo secolo come il teorema di Shannon-Nyquist e fornisce strumenti fondamentali nell'elaborazione dei segnali come i filtri, che evolvono dalla teoria dei circuiti elettrici lineari e passando per la teoria dei filtri ottimi e adattativi permettono di comprendere appieno le reti neurali, che ne rappresentano l'estensione al caso non-lineare, oggi tanto utilizzate quanto poco studiate in maniera approfondita. Questa mole di conoscenze, è stata in parte funzionale all'evoluzione del pensiero compositivo e musicale. Si pensi ad esempio al dualismo

tempo-frequenza e alla sua analogia matematica con la fisica quantistica, che erano alla base degli articoli di Denis Gabor, poi riletti in chiave compositiva da Iannis Xenakis tra gli altri [1], oppure all'estetica glitch [2], che nasce dall'errore nel trattamento del segnale digitale.

In sostanza, il corpus di conoscenze correlate al DSP è vasto e spazia - per quanto concerne questo articolo - dalla teoria matematica alla pratica compositiva o strumentale, rendendone l'insegnamento un compito complesso. Sebbene la teoria di base sia ben coperta da testi storici come [3, 4, 5], rimane una materia piuttosto complessa, specie se si desidera raggiungere una comprensione piena della teoria. D'altra parte, l'uso di tali testi e approcci è sconsigliato al di fuori dei corsi di laurea in ingegneria dell'informazione¹, perciò altri approcci educativi più "funzionali" sono necessari². Qui si intendono come tali quegli approcci per cui la funzione e la manipolazione pratica dell'oggetto diventano primarie per l'apprendimento. Ciò non è da intendersi in senso negativo, perlomeno nell'ambito musicale dove a mediare la conoscenza è necessario anche l'organo di senso dell'udito.

Questi approcci, con l'avvento dei personal computer e di Internet, sono diventati popolari perché permettono di mediare tra teoria e pratica rendendo la materia accessibile a più studenti. La loro diffusione è tanto ampia che è prassi consolidata che un esame universitario di DSP oggi si sostenga attraverso una prova al computer³.

1.1 Scopo del lavoro

Sebbene l'introduzione di nuove piattaforme e linguaggi, avviene senza sosta, rendendone complessa la scrematura e la scelta, si vuole qui aggiungere una ulteriore recente piattaforma alla rima di possibili software per la didattica in un ambito piuttosto vario che abbraccia i corsi di DSP, quelli di ingegneria del suono e quelli di musica elettronica.

Il software, di recente introduzione, verrà prima introdotto, evidenziando le sue caratteristiche chiave, e verrà

¹ Va detto che anche all'interno della classe di Laurea L-8 esistono molti corsi di studio che non includono un insegnamento di teoria dei segnali a tempo continuo, necessario fondamento per trattare i segnali a tempo discreto con la giusta consapevolezza.

² Nell'opinione dell'autore, un apprendimento limitato delle basi teoriche rappresenta un rischio collettivo per due motivi: da un lato l'uso improprio della tecnica che ne può scaturire, dall'altro il mancato tramandarsi di branche della conoscenza che la società considera desueti o meno attrattivi. D'altra parte l'ipertrofia della tecnica impone una selezione degli aspetti focali su cui concentrare la didattica a parità di tempi, rispetto ad alcuni decenni o sono.

³ Sarebbe interessante stabilire se c'è una diretta relazione tra l'abilità di risolvere problemi computazionali e il padroneggiare la teoria, ma non sembrano esserci studi a riguardo.

messo in relazione ad altri più noti software per l'elaborazione di segnali in ambito musicale. Sebbene il suo scopo originariamente fosse piuttosto definito, ha la flessibilità di adattarsi a vari contesti didattici che verranno esposti. Verranno evidenziate le caratteristiche desiderabili per varie tipologie di discenti, fornendo suggerimenti ed esempi dalla esperienza personale dell'autore, raccolta in corsi accademici, workshop organizzati indipendentemente e dall'interazione con i lettori di un libro di testo dedicato che l'autore ha pubblicato nel 2020 [6].

La struttura dell'articolo segue. Al capitolo 2 vengono introdotte le difficoltà nella scelta di una piattaforma software per la didattica e le possibilità attualmente disponibili. Il capitolo 3 introduce VCV Rack, mentre nei capitoli 4 e 5 viene introdotto il suo utilizzo lato utente e poi lo sviluppo di moduli al suo interno. Al capitolo 6 vengono forniti esempi didattici e spunti per altri docenti. Infine il capitolo 7 discute alcuni vantaggi e svantaggi. Il capitolo 8 chiude l'articolo.

2. PIATTAFORME DIDATTICHE

Come già detto, in funzione del tipo di settore in cui viene svolta la didattica, ad un software vengono richieste caratteristiche diverse. Cerchiamo qui di seguito di elencare alcuni dei parametri per la scelta e alcune delle opzioni possibili.

2.1 Parametri di scelta

Innanzitutto, alcuni corsi sono prettamente professionalizzanti o calati verticalmente su un determinato settore professionale, mentre altri sono di respiro più ampio e sono propedeutici a numerosi tipi di applicazioni. Il docente dovrà, dunque, valutare per ogni software preso in considerazione la "distanza" dal settore industriale di riferimento. Un corso professionalizzante per la produzione musicale dovrà, presumibilmente, adottare una Digital Audio Workstation (DAW) commerciale e dei plugin, mentre un corso base di DSP per futuri ingegneri potrà adottare un qualsiasi tipo di linguaggio di programmazione capace di effettuare elaborazione numerica di segnali. E' anche importante definire il tipo di piattaforma di calcolo: PC, tablet, sistemi embedded e così via.

Un altro requisito da valutare è il tipo di elaborazione: se in tempo differito o in tempo reale, e con che modalità le due possibilità possono essere mescolate. Un software di editing non-lineare come una DAW, ad esempio, permette entrambe le modalità a seconda delle necessità, mentre una piattaforma come Pure Data o Max/MSP lavora esclusivamente in tempo reale. Nell'ambito dell'ingegneria, è molto importante poter elaborare alcuni brani di segnale offline per poi studiare le loro caratteristiche e la loro correttezza rispetto ad un compito assegnato, ma nell'ambito di un corso votato all'elaborazione di segnali audio è molto probabile che si voglia anche lavorare in tempo reale e possibilmente modificare dei parametri degli algoritmi DSP in tempo reale. Purtroppo le due cose vanno difficilmente di pari passo, giacché il debug degli algoritmi o l'analisi di grafici relativi ai segnali sono operazioni che richiedono

l'interruzione dell'elaborazione, cosa spesso non permessa dalla piattaforma stessa.

Nell'ambito della sintesi, si deve rendere possibile sviluppare soluzioni personalizzate conformi al tipo di didattica. La piattaforma deve permettere di fare coding o scripting in un linguaggio consono alla future necessità professionali degli studenti (nonché, presumibilmente, già in loro possesso, o in fase di acquisizione), oppure fare patching o routing utilizzando metafore grafiche. La necessità di utilizzare strumenti grafici per la misura dei segnali deve essere adeguata alle finalità del corso.

Se la piattaforma permette di elaborare dei propri algoritmi a basso livello, e se il corso richiede di farlo, il docente deve valutare il linguaggio e la toolchain eventualmente necessaria. In alcuni casi, si implementano oggetti base (chiamati in vario modo, ad esempio opcode, external o plugin a seconda della piattaforma) con un linguaggio imperativo compilato come il C/C++, che rende necessario una suite di compilatore, linker, debugger, etc. E' anche compito del docente valutare la complessità realizzativa di un'interfaccia grafica, se disponibile, e le sue potenzialità di interazione con l'utente.

Tra i dettagli tecnici vanno considerati anche la compatibilità con i sistemi operativi più utilizzati ed eventuali dispositivi hardware necessari (interfacce audio, schermi aggiuntivi e controller), la compatibilità con altri software utili e il supporto di protocolli di comunicazione utilizzati nel settore (MIDI, OSC, DMX, protocolli di networked music performance, etc.).

In ultimo, la licenza e gli eventuali costi di acquisizione del software rappresentano un importante scoglio all'adozione da parte degli studenti e vanno, dunque, ben considerati. In presenza di costi importanti, l'acquisizione del software è opportuno che venga fatta dall'istituzione formativa all'interno di un laboratorio dedicato, il che però limita la fruizione e lo studio personale degli studenti alle ore di accesso alla struttura. Va inoltre considerata l'importanza di educare all'uso e alla comprensione del paradigma free/open software, se possibile.

2.2 Panoramica

Allo stato attuale, la maggior parte degli insegnamenti nell'ambito dell'elaborazione di dati e segnali fa uso, almeno in parte di Matlab⁴. Tra le motivazioni di questa scelta ci sono sicuramente la sua flessibilità nel trattare tipi di dati diversi, la semplicità del linguaggio di scripting e l'ampia disponibilità di librerie che man mano sono state aggiunte. Il suo equivalente open source, GNU Octave, è anch'esso utilizzato, seppur con minore frequenza. Simulink è un pacchetto aggiuntivo per Matlab che offre un ambiente di progettazione grafica tramite blocchi funzionali.

Sebbene ampiamente adottati nell'ambito dell'ingegneria, Matlab e Octave presentano alcune limitazioni didattiche. Non sono di semplice utilizzo, non sono adatti all'elaborazione real-time e non hanno una vera valenza al di fuori dell'ingegneria. Anche per gli ingegneri possono presentare delle lacune, come anche evidenziato da altri

⁴ Matlab è un marchio registrato di The MathWorks, Inc.

autori [7], perché non aiutano a comprendere problemi di precisione finita e di implementazione real-time.

Gli altri linguaggi e piattaforme che esaminerò nascono in seno all'informatica musicale e nessuno di questi è perfetto sotto tutti gli aspetti come già evidenziato in [8]. Pure Data, Max/MSP, Chuck e Faust sono particolarmente indicati per la sintesi e rendono il flusso del segnale molto chiaro. Nel caso di Pure Data e Max/MSP, questo è dovuto all'ambiente grafico di patching, che peraltro distingue chiaramente segnali audio e segnali di controllo. Faust segue un paradigma funzionale costruito attorno a operazioni da effettuare sui segnali, mentre Chuck, attraverso il suo operatore => effettua connessioni tra oggetti che rende il flusso del segnale piuttosto evidente seppure la sintassi sia piuttosto unica.

Linguaggi che forniscono un supporto più esplicito per la composizione come CSound e Supercollider, sono anch'essi adatti allo studio della sintesi perché dotati di funzionalità DSP che possono essere utilizzate efficacemente in tempo reale. Tuttavia, il loro apprendimento passa per l'uso di sintassi e opcode specifici, per cui ne è consigliata l'adozione solo se si intende utilizzarli a lungo.

Passando a sistemi compilati, invece, esistono SDK di livello industriale come VST di Steinberg e AU di Apple che sono normalmente utilizzati per implementare algoritmi in software commerciali e possono altresì essere impiegati per l'insegnamento, anche se presentano molte problematiche tecniche tediose e poco rilevanti ai fini della didattica. Per superare queste difficoltà, Mathworks ha sviluppato una catena di strumenti che aiuta a creare un plugin compatibile con VST da codice Matlab [9]. Un altro strumento popolare per lo sviluppo di applicazioni audio e plugin VST/AU è JUCE. Anche questo può essere visto come un utile strumento didattico e in qualche misura può rendere più semplici i progetti degli studenti per lo sviluppo di plugin VST utilizzabili, sebbene sia ancora moderatamente complesso per la creazione di un progetto. Un autore che ha dedicato molto tempo all'insegnamento di DSP per audio e sintesi, Will Pirkle, ha deciso di mantenere due percorsi paralleli nelle sue risorse didattiche: plugin VST/AU/AAX e la sua piattaforma (prima RackAFX, poi ASPiK) [10, 11], attualmente open-source. Vale la pena menzionare ASPiK perché fornisce un buon compromesso tra la necessità didattica di lavorare sugli aspetti più teorici del DSP, rendendo semplice l'implementazione di algoritmi, ma al contempo permette di sviluppare plugin avanzati ed esportarli in formato VST/AU/AAX. Purtroppo la scarsa adozione della piattaforma e l'assenza di una estesa libreria di oggetti già disponibili, così come l'assenza di un ambiente di patching, rende difficile la prototipazione rapida di sistemi complessi.

Un'ultima menzione va a Python, che sta diventando sempre più il linguaggio di scelta per l'insegnamento, dalle prime esperienze di coding nelle scuole dell'obbligo ai corsi universitari avanzati in data science, informatica e ingegneria. Al momento però, le librerie più comuni sono dedicate all'elaborazione audio offline a fini di analisi e ricerca, come Librosa⁵ e non sembra esserci una libreria

sufficientemente universale da essere adatta ad un uso didattico stabile negli anni a venire. D'altra parte, la creazione di un'interfaccia grafica dedicata all'uso in real-time non è ancora semplice, e dovrebbe essere sviluppata ad hoc, a seconda dello scopo.

3. VCV RACK

VCV Rack è un'applicativo software multi-piattaforma progettato principalmente come un sintetizzatore modulare virtuale Eurorack. Sviluppato da VCV, il software è implementato in C++ ed è basato su un audio engine multithread per una efficiente elaborazione in tempo reale. Per un uso da semplice utente è disponibile la versione gratuita standalone che permette di caricare numerosi plugin forniti gratuitamente da VCV o da terze parti, rendendo immediato l'utilizzo e l'improvvisazione o lo studio. Per chi volesse utilizzare Rack come piattaforma per lo sviluppo dei propri plugin, viene messa a disposizione una toolchain basata sul compilatore GNU che può essere installata su uno qualsiasi dei sistemi operativi supportati con pochi comandi da terminale.

A differenza dei sistemi di elaborazione convenzionali che sfruttano buffer, VCV Rack opera campione per campione. Questa scelta peculiare comporta un aumento del carico della CPU per via delle più frequenti chiamate a funzione, ma d'altra parte permette di aumentare l'accuratezza nella simulazione dei sistemi analogici quando sono collegati in feedback. Un utile effetto collaterale di questa scelta è una significativa semplificazione nella scrittura delle routine DSP, come vedremo in seguito.

L'interfaccia grafica utente (GUI) di VCV Rack è costruita utilizzando nanoVG⁶, una libreria di rendering grafico vettoriale 2D per OpenGL. La GUI è molto semplice, fornendo agli utenti un rack virtuale che l'utente può riempire con i moduli disponibili.

In termini di modello commerciale, VCV Rack offre una versione standalone gratuita così come una versione a pagamento per l'integrazione nelle workstation audio digitali (DAW) come plugin. Il software fornisce solo una piccola libreria di moduli di base, e si appoggia sulla disponibilità di moduli di terze parti per espandere le sue possibilità. Poiché promuove un modello di sviluppo open source, un gran numero di appassionati si sono uniti al progetto creando i propri moduli e rendendoli disponibili in una delle seguenti forme: gratuito e open source, gratuito ma closed source, a pagamento. La distribuzione dei moduli è semplificata dalla disponibilità di un repository centrale, la libreria dei moduli, ospitata da VCV, dove un utente può trovare moduli e aggiungerli a Rack. Qualsiasi sviluppatore può aggiungere i propri moduli alla libreria a condizione che segua alcune semplici regole. La procedura per aggiungere i moduli alla libreria è snella e viene facilitata da VCV.

Rack viene considerato alla stregua di una DAW, poiché ospita plugin caricati dinamicamente che possono essere collegati a ingressi e uscite audio e MIDI. Tuttavia, la versione a pagamento è disponibile come plugin VST/AU e,

⁵ <https://github.com/librosa/librosa>

⁶ <https://github.com/memononen/nanovg>

quindi, essere a sua volta ospitata sulle più comuni DAW. Limitate funzionalità di registrazione sono disponibili attraverso un modulo dedicato. L'interfaccia MIDI è compatibile con l'estensione MIDI Polyphonic Expression (MPE), consentendo agli utenti di controllare i parametri di sintesi con migliori capacità di espressione e modulazione.

Alcune ramificazioni del progetto sono state create da altri sviluppatori, come Mi Rack⁷ e Cardinal⁸ in un tentativo di mirare a scenari di applicazione diversi. Cardinal, ad esempio, è rilasciato anche come applicazione online, che può essere interessante per demo e lezioni rapide tramite semplice accesso da browser.

4. UTILIZZO

VCV Rack opera attraverso la creazione di *patch*, collegando moduli attraverso le loro porte e configurando i loro parametri, espressi nella maggior parte dei casi attraverso switch e potenziometri rotativi o lineari come se ne trovano normalmente nei moduli reali.

L'interfaccia è composta principalmente da un pannello principale, dove vengono posizionati i moduli e connessi tra loro tramite cavi virtuali. Un aspetto distintivo di VCV Rack risiede nella sua gestione dei segnali, dove tutti i cavi trasportano valori di tensione virtuali. A differenza di altri software per la computer music non esiste una distinzione tra segnali di controllo e segnali audio in VCV Rack, bensì tutti i segnali hanno la stessa frequenza di campionamento. Questo favorisce una maggiore flessibilità creativa, anche se può ridurre la chiarezza di una patch. Inoltre, i cavi possono essere di tipo polifonico, ovvero atti a condurre segnali multipli affasciati insieme. Questo rimuove una grossa limitazione di tutti i sistemi modulari reali, che per avere voci multiple richiederebbero di avere più istanze dello stesso modulo, cosa costosa, dispendiosa di spazio, e tediosa vista la necessità di ripetere le impostazioni di ogni modulo negli analoghi per ogni voce.

Una volta installato il software, gli utenti possono accedere ad una piccola gamma di moduli basilari che coprono tutte le principali categorie della sintesi modulare, come oscillatori, filtri, generatori di involuppo, effetti e altro ancora. Questa piccola palette di moduli è molto istruttiva e nella sua limitatezza permette di mantenere semplice lo studio delle basi. In Figura 1 viene mostrata una semplice patch per la sintesi sottrattiva. Tra i moduli sono mostrati un'interfaccia MIDI, che può anche sfruttare la tastiera QWERTY del computer (per una semplice demo), un'interfaccia audio e un modulo per scrivere alcuni appunti, utile per fornire delle patch ai propri studenti con indicazioni operative o semplici spiegazioni.

L'utilizzo del sistema modulare richiede di comprendere il paradigma di controllo in tensione. La convenzione in tal senso è la stessa dello standard Eurorack: pitch e frequenze di taglio utilizzano la convenzione V/OCT (ad un aumento 1V consegue un aumento di 1 ottava nel pitch), i segnali di controllo unipolari spaziano tra 0 e 10V, mentre i segnali bipolari oscillano tra -5 e 5V, mantenendo il

⁷ <https://mirack.app/>

⁸ <https://cardinal.kx.studio/>

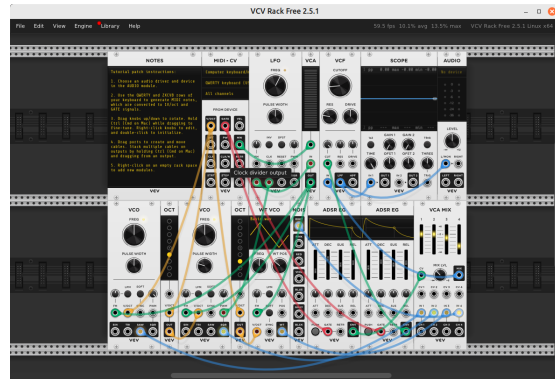


Figure 1: Una schermata di esempio di VCV Rack dove sono allocati i moduli per avere un semplice sintetizzatore a sintesi sottrattiva.

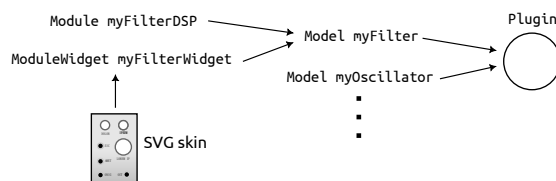


Figure 2: Struttura degli oggetti di un ipotetico plugin, contenente più moduli (Model). Ogni Model aggrega il Module (che contiene il codice DSP) e il ModuleWidget (che descrive l'interfaccia utente).

range di 10V. Trigger e gate raggiungono il valore massimo di 10V. Di recente Rack ha reso possibile trarre più cavi da una uscita e di sommare più segnali ad un ingresso, rendendo il sistema più flessibile rispetto ad un analogo hardware.

5. IMPLEMENTAZIONE DI MODULI

Nel contesto di VCV Rack, i termini "plugin" e "modulo" si riferiscono a componenti distinti all'interno dell'architettura del software. Un plugin in VCV Rack è essenzialmente una libreria dinamica (.dll, .dylib, o .so a seconda del sistema operativo) che viene caricata dinamicamente dall'applicazione e contiene uno o più moduli. I moduli con cui l'utente interagisce, d'altra parte, sono istanze della classe Model all'interno del plugin, ognuna rappresentante un'unità distinta di funzionalità di elaborazione del segnale. Ogni modulo è composto da due componenti principali: una GUI definita ereditando la struttura ModuleWidget, che permette l'interazione dell'utente e l'aggiustamento dei parametri, e un oggetto che eredita la struttura Module, definendo il comportamento di elaborazione del segnale del modulo. La struttura gerarchica viene mostrata in Figura 2.

Ad esempio, il codice sorgente di un ipotetico plugin HelloWorld consiste in un file sorgente HelloModule.cpp che contiene la definizione del Model e l'implementazione del Module e del ModuleWidget. Due file sorgenti aggiuntivi, HelloWorld.cpp e HelloWorld.hpp,

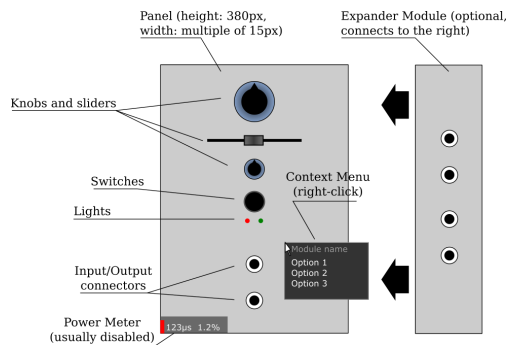


Figure 3: Un modulo in Rack e i suoi componenti.

contengono semplicemente la definizione e l’istanziamento di tutti gli oggetti `Model`, uno per modulo. Risorse come i file grafici vettoriali, eventuali file audio (ad esempio, risposte all’impulso o campioni) possono essere collocate in una apposita cartella allo stesso livello dell’oggetto compilato. Un manifesto `plugin.json` contiene alcune informazioni specifiche per il plugin e i suoi moduli.

La creazione di un plugin è semplificata dalla disponibilità di oggetti già presenti che ne standardizzano l’interazione con l’utente, e traggono spunto dal mondo dei sintetizzatori modulari: connettori di ingresso o uscita, parametri (potenziometri rotativi o lineari, pulsanti), e luci. Questi vanno enumerati quando si definisce il `Module` e si può attingere a oggetti grafici già presenti. In Figura 3 si mostra un esempio. Gli oggetti grafici già esistenti possono essere modificati dallo sviluppatore nell’apparenza o nel comportamento in maniera semplice. Inoltre, la GUI può essere dotata di widget progettati dallo sviluppatore. Questi devono essere disegnati utilizzando la libreria `nanoVG` che consente un’ampia varietà di interfacce basate su primitive di disegno come punti, linee, ellissi, poligoni e così via. Pertanto è abbastanza facile progettare, ad esempio, un plot FFT in tempo reale.

5.1 Esempio: un filtro low-pass

Per poter realizzare un modulo è necessario implementare le classi suddette. Per moduli con un’interfaccia standard il processo è molto rapido e meccanico, ad eccezione per la funzione di elaborazione del segnale `process`.

Come esempio, in questa sezione mostreremo come realizzare un semplice filtro. Il modulo richiesto deve avere un ingresso, un’uscita e una manopola per controllare la frequenza di taglio. Una sezione passa basso del secondo ordine sarà realizzata utilizzando il tipo `BiquadFilter` fornito da Rack.

Il modulo è dichiarato come una struttura che eredita il tipo `Module` e deve implementare un costruttore e il metodo `process` per il processing in tempo reale. Inoltre, è richiesto un insieme di `enum` per indicare il numero di ingressi, uscite, parametri e luci. Il costruttore deve indicare il loro numero tramite una funzione specifica `config`, mentre l’intervallo dei parametri deve essere impostato utilizzando `configParams`.

La `process` per il modulo di esempio è molto semplice e viene mostrata nel Listato 1. Com’è possibile vedere non è necessario iterare attraverso un buffer perché l’elaborazione avviene per un singolo campione. Nella stessa funzione vengono letti gli ingressi e i parametri, effettuato l’eventuale ricalcolo di parametri ed effettuata l’elaborazione. Il codice sorgente di questo e di tutti i moduli *ABC* presentati sul libro [6] è disponibile online⁹.

```
void ASimpleFilterModule::process(const ProcessArgs &
args) {
    float cutoff = params[PARAM_CUTOFF].getValue();
    float x = inputs[MAIN_IN].getVoltage();

    LPF.setParameters(BiquadFilter::LOWPASS, cutoff, Q,
0);
    float y = LPF.process(x);

    outputs[LPF_OUT].setVoltage(y);
}
```

Listing 1: Esempio di metodo `process` che implementa un filtro passa-basso utilizzando la primitiva `BiquadFilter`.

Per completare il modulo va definito il costruttore di un `ModuleWidget`, dove vanno semplicemente inserite le posizioni di ingressi, uscite e knob, oltre alla grafica di sfondo. Infine, deve essere definito un `Model`, che in una riga indica quali `Module` e `ModuleWidget` vanno collegati a formare un modulo. Nel nostro esempio, un modulo filtro semplice che utilizza il tipo `BiquadFilter` integrato richiede meno di 60 righe di codice, che sono in gran parte simili tra ogni modulo e possono essere rapidamente copiate e adattate.

La GUI richiede una grafica vettoriale SVG, che può essere disegnata, ad esempio, con un software open source come `Inkscape`¹⁰.

6. POTENZIALITÀ PER LA DIDATTICA

Rack può essere utilizzato come piattaforma per esercitazioni di laboratorio, compiti per gli studenti, tesine e progetti di ricerca.

6.1 Studio di paradigmi di sintesi non convenzionali

Rispetto al paradigma di sintesi East-coast, in cui le tre variabili del suono (intensità, pitch e timbro) vengono “ortogonalizzate” affidandone la cura a moduli (o gruppi di moduli) separati, la sintesi modulare permette di seguire approcci molto diversificati. I sistemi modulari Buchla dapprima, e poi Serge e Tchperenin, hanno stimolato paradigmi di sintesi più complessi, in cui uno stesso modulo può avere effetti su tutti e tre i piani del suono. Ne è un esempio eccelso il modulo `Lopass-Gate` [12] (es. Buchla 292C), che può lavorare come leggero filtro, VCA o attuare entrambe le funzioni contemporaneamente.

Tuttavia, lo studio di tecniche di sintesi non convenzionali, come quelle disponibili con i moduli Buchla [13, 14], non sono alla portata di tutti, visto il loro costo elevato.

⁹ <https://github.com/LOGUNIVPM/VCVBook/>

¹⁰ <https://inkscape.org/>

Con VCV Rack si possono sperimentare una varietà di tecniche note nell'ambito della sintesi modulare, come: la tecnica di *pinging*, ovvero la stimolazione di filtri attraverso impulsi rapidi; l'utilizzo dei suddetti *lopass-gate* (il plugin Nysthi mette a disposizione una replica del Buchla 208 senza codice sorgente ma in forma gratuita); il *wavefolding* e il *waveshaping* (es. Surge XT Waveshaper, porting in Rack dello storico synth open source Surge); la creazione di forme d'onda periodiche o pseudoperiodiche attraverso generatori di inviluppo rapidi in retroazione (es. Befaco Rampage, disponibile come plugin open source). In aggiunta i moduli virtuali possono implementare avanzate tecniche digitali come la granolazione (es. Audible Instruments Texture Synthesizer, porting del software open source del popolare modulo Eurorack Mutable Instruments Clouds), la modellazione fisica di corde e membrane (es. Audible Instruments Modal Synthesizer e Resonator, a loro volta porting dei moduli open source Mutable Instruments Elements e Rings) e così via.

Naturalmente, i sintetizzatori modulari permettono di sperimentare tecniche di composizione aleatoria e generativa. L'uso di LFO, sequencer basati su cellular automata, sequencer poliritmici ed Euclidei, generatori di tensioni random, *sample&hold*, assieme a moduli pensati per implementare algoritmi caotici, possono certamente arricchire il lavoro compositivo, se utilizzati con cognizione di causa.

6.2 Progetti d'esame bilanciati

Per i corsi avanzati di ingegneria, il docente di solito richiede agli studenti lo svolgimento di un progetto finale sugli argomenti del corso. Tuttavia, è spesso difficile trovare un compromesso tra la necessità di diversificare i loro compiti ed equilibrare lo sforzo di ogni studente. Inoltre, se i progetti individuali sono troppo diversi, è anche difficile valutare tutti i risultati in modo equo e calibrare correttamente i voti finali. Tanto più che per un progetto di DSP ci sono aspetti extradisciplinari che influiscono sulla valutazione, come il design dell'UI/UX, i risultati estetici e l'utilità e usabilità dell'algoritmo sviluppato. Pertanto, un progetto che tratta la sonificazione può richiedere poco lavoro a livello matematico e più a livello creativo, che può influenzare il docente se i risultati sono sonicamente interessanti, mentre un compito eccezionale di modellazione analogica virtuale potrebbe non essere altrettanto divertente da valutare, penalizzando uno studente brillante che ha fatto un ottimo lavoro con la teoria appresa in classe.

Per evitare questo dilemma, nei primi due anni dell'insegnamento di laurea magistrale "Digital Circuits for Music Processing and Sound Synthesis" [W001240] presso l'Università Politecnica delle Marche, l'autore ha optato per una soluzione piuttosto semplice: il progetto finale consisteva nell'implementare diverse parti di una drum machine analogica vintage, la Eko Computerhythm, la prima drum machine programmabile conosciuta. Il compito consisteva nell'implementare un modulo VCV Rack, uno per studente, che sintetizzasse il suono di una delle 12 parti della drum machine. L'importanza del compito è chiara per lo studente dal momento che esistono solo pochi esem-

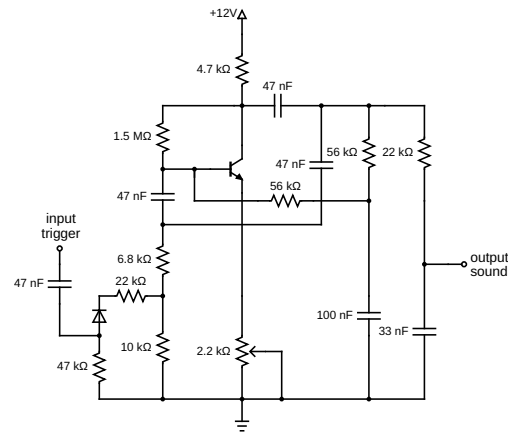


Figure 4: Schematico del circuito Bass Drum della Eko Computerhythm, uno dei circuiti assegnati agli studenti per un'implementazione virtual analog.

plari di questo strumento in tutto il mondo, pertanto una modellazione analogica virtuale dei circuiti è un mezzo per preservarne l'eredità. La parte dello schematico che riguarda la sintesi del suono è costituita da 12 sottocircuiti simili a quello mostrato nella Figura 4, che genera il suono del bass drum. Il circuito risponde ad un trigger in ingresso e genera un tono con decadimento esponenziale. Da questa base, gli studenti possono studiare il circuito e il suo output e implementare una simulazione adeguata in VCV Rack. Tutti gli studenti devono lavorare sullo stesso repository di codice condiviso (per imparare a utilizzare un sistema di versionamento come Git) che permette al docente di monitorare il loro lavoro, suggerire correzioni e condurre la valutazione finale compilando ed eseguendo il codice.

6.3 Controllo via hardware

Rack può essere controllato mediante hardware MIDI. Un progetto di esempio che richiede alcune competenze di elettronica potrebbe essere la costruzione di un'interfaccia di controllo. Ad esempio, diversi gruppi di studenti potrebbero costruire un controller fisico per un modulo specifico nella libreria Rack. Anche in questo caso, ciò consente di assegnare compiti coerenti agli studenti, di bilanciare il loro impegno e di rendere la valutazione finale equa, senza consentire loro di copiare facilmente dagli altri. A ciascun gruppo potrebbe essere dato un microcontrollore, come un Arduino, e, dato un insieme di potenziometri e pulsanti, potrebbero realizzare un circuito stampato e un pannello con le dimensioni esatte del modulo virtuale. Infatti, i moduli virtuali in Rack sono progettati utilizzando unità metriche standard, non pixel, quindi quando si progetta un pannello virtuale, lo stesso può essere utilizzato per un servizio di incisione laser per ottenere il corrispettivo del mondo reale. Alcune piccole aziende che vendono moduli Eurorack reali hanno anche progettato un gemello digitale all'interno di VCV Rack (es. Befaco, Grayscale, etc.), mentre alcuni sviluppatori di VCV Rack hanno prodotto moduli hardware dopo aver verificato l'interesse dei propri moduli in Rack (es. Vult).

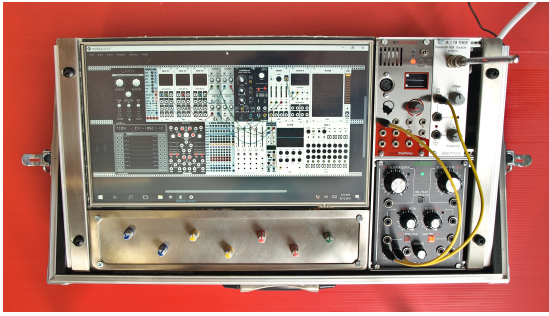


Figure 5: Un sistema ibrido basato su PC e alcuni moduli Eurorack.

Va aggiunto che è possibile interfacciare anche veri modulari attraverso una apposita scheda senza DC-blocker. Infatti, esistono in commercio moduli Eurorack con ingressi e uscite CV, che dialogano via USB con il PC e possono dunque trasferire valori di tensione provenienti da o indirizzati a moduli Eurorack reali. In Figura 5 si vede un prototipo che include una scheda PC al suo interno su cui viene eseguito Rack, uno schermo touch, un set di potenziometri letti da una scheda di sviluppo Teensy (simile ad Arduino, ma con supporto USB-MIDI) ed alcuni moduli Eurorack. Tra questi moduli troviamo un'antenna capacitiva (Doepfer A-178, analoga a quella che si trova in un Theremin), un'interfaccia per onde cerebrali (Soundmachines B11, richiede un caschetto apposito dotato di sensore ECG) e un'interfaccia CV-USB. Quest'ultima può inviare la tensione di controllo proveniente dai suddetti moduli a VCV Rack, integrandolo nel workflow.

6.4 Progetti multimediali

Nella sua versione a pagamento, VCV Rack può essere lanciato su qualsiasi VST/AU host, come le tradizionali DAW ma non solo. Ad esempio, sistemi per la progettazione di live visual come TouchDesigner possono ospitare VCV Rack. In Figura 6 viene mostrata una schermata del software che ospita una patch di musica generativa, il cui output agisce sulla dinamica di una nuvola di punti che costituisce parte dei live visuals, realizzata durante un workshop del 2023 dall'autore in occasione di Acusmatic Festival.

7. CONSIDERAZIONI FINALI

Nell'ambito dello sviluppo di algoritmi DSP, Rack offre un buon compromesso tra la complessità di SDK commerciali (come JUCE o VST/AU) e sistemi di patching modulare come Pure Data o Simulink. Lo sviluppo di moduli è molto più agile rispetto allo sviluppo di plugin VST/AU o external di Pure Data, mentre la sua architettura modulare consente di separare le funzionalità in singoli moduli.

Rack consente una rapida prototipazione di idee e concetti di elaborazione del segnale. Gli oscilloscopi disponibili permettono di monitorare i segnali in tempo reale e fare parte del debug. I moduli esistenti possono essere utilizzati per creare una patch e, quando necessario, gli stu-

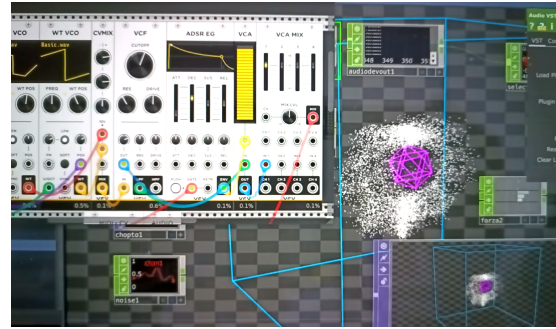


Figure 6: Una schermata di TouchDesigner che ospita VCV Rack come plugin VST.

denti possono costruire i propri in C++ oppure eseguendo patch in Faust Pure Data. L'architettura modulare facilita la prototipazione, poiché molte funzioni di base sono già implementate. Va notato, tuttavia, che i moduli sono spesso progettati per il musicista e, diversamente da sistemi come Simulink, che implementano funzionalità più adatte all'ingegneria, i moduli Eurorack possono nascondere le loro funzionalità sotto nomi bizzarri. L'utilizzo di tag e il rapido accesso al manuale del modulo (tramite click con tasto destro) ne facilita la comprensione. Va tuttavia notato che la qualità della documentazione è molto eterogenea e dipende dalla volontà e l'impegno di ciascuno sviluppatore.

In Rack non c'è distinzione tra segnali a controllo e segnali audio, semplificando l'approccio e consentendo di sperimentare l'elaborazione del segnale con qualsiasi segnale. I controlli MIDI vengono quindi convertiti in segnali analogici virtuali. La mancanza di dati ad alto livello, come i messaggi Open Sound Control (OSC), può rappresentare uno svantaggio nella costruzione di strumenti complessi o nella connessione ad interfacce per l'espressione musicale.

Il software ha superato i sette anni di vita, il che promette una continuità anche per gli anni a venire. Anche se è mantenuto da un'azienda privata, se quest'ultima interrompesse il supporto per il software, Rack è open source ed è possibile che alcuni volontari eseguano un fork, prolungando così la sua vita, il che può essere valutato positivamente in un'ottica di durability e sostenibilità [15].

Come tutti gli altri software basati su C/C++, gli studenti devono imparare a compilare utilizzando una standard toolchain GNU. Questo potrebbe essere troppo complesso per alcuni corsi di studio. Tuttavia, la configurazione della toolchain è estremamente facile grazie a una semplice guida passo-passo per ciascuno dei sistemi operativi supportati, che lo ha reso popolare anche tra neofiti che, anzi, hanno imparato a sviluppare in C/C++ proprio scegliendo questa piattaforma (tra i primi contributor della library ce ne sono alcuni).

Rack è scritto secondo le recenti linee guida del C++ (C++11), e adotta moderne politiche di ingegneria del software, che lo rendono ideale per studenti in ingegneria, rispetto a sistemi come Pure Data, in quanto preparano ad

una programmazione più moderna. Il software è molto leggibile e fornisce molto codice open source da cui gli studenti possono apprendere.

Diversamente da altri linguaggi, non c'è uno strumento per esportare i moduli VCV Rack in altri formati, come plugin VST/AU o LV2, ma Rack stesso (e ciascuno dei suoi moduli) può essere ospitato all'interno di una DAW. Pertanto non è possibile creare un plugin commerciale finito per DAW da un modulo Rack, ma è possibile utilizzare qualsiasi patch Rack su un DAW (è necessaria la versione a pagamento). Al contrario ci sono strumenti per esportare strumenti Faust su Rack, e il modulo VCV Prototype può ospitare script da Lua, Javascript, Faust, Pure Data e SuperCollider, rendendolo molto interessante per chi abbia già esperienza consolidata in quegli ambienti di sviluppo.

Il debugging potrebbe essere più semplice se fosse disponibile un'interfaccia più avanzata, ad esempio per monitorare segnali che non sono esposti attraverso una porta di output. Un misuratore di carico CPU è disponibile così come oscilloscopi e un registratore di segnali. Gli oscilloscopi disponibili consentono solo segnali nell'intervallo $(-10 : 10)V$, che è adeguato per segnali di tensione analogica. Tuttavia, se le variabili assumono valori al di fuori di tale intervallo (o sono molto piccole), è più difficile analizzarle. Ad esempio, la simulazione di fenomeni fisici, come un'onda di pressione all'interno di un tubo acustico che assuma valori di alcuni kPa, richiede la scalatura per adattarsi all'asse y di un oscilloscopio.

Infine, l'attuale implementazione del multi-touch è problematica. Questo è legato a limitazioni nella libreria nanoVG e VCV non ha piani per contribuire a nanoVG per migliorare l'usabilità su touchscreen. Questo rende Rack inutilizzabile su touchscreen, cosa che sarebbe desiderabile in strumenti didattici come le lavagne interattive che hanno il supporto nativo per il multitouch.

8. CONCLUSIONI

In questo articolo è stata discussa una recente piattaforma software chiamata VCV Rack come strumento didattico per corsi di DSP per l'audio e la musica. I suoi vantaggi e svantaggi sono stati esaminati e confrontati con quelli di altri strumenti software ben noti. Nel complesso, la semplicità dell'architettura del software, l'elaborazione a livello di campione e l'interfaccia semplice di Rack sembrano promettenti per corsi di ingegneria e non solo. In particolare, chiunque sia interessato a realizzare tecniche di sintesi e strumenti elettronici può apprezzare la piattaforma sia a livello di patching che di sviluppo in C++.

9. NOTE

L'autore non è affiliato in alcun modo con VCV o qualsiasi altro brand riportato nell'articolo. L'autore vuole ringraziare tutti i suoi studenti della laurea magistrale in Ingegneria Elettronica per il loro entusiasmo e impegno.

10. REFERENCES

- [1] A. Di Scipio, "I quanta acustici di Gabor nelle tecnologie del suono e della musica," *Musica/Tecnologia*, pp. 17–42, 2016.
- [2] J. Vanhanen, "Virtual sound: examining glitch and production," *Contemporary Music Review*, vol. 22, no. 4, pp. 45–52, 2003.
- [3] A. V. Oppenheim and R. W. Schaffer, *Discrete-Time Signal Processing*. Upper Saddle River, NJ, USA: Prentice Hall Press, 3rd ed., 2009.
- [4] S. Haykin, "Adaptive filter theory," *Prentice Hall google schola*, vol. 2, pp. 67–94, 2002.
- [5] J. G. Proakis and D. K. Manolakis, *Digital Signal Processing (4th Edition)*. Prentice Hall, 2006.
- [6] L. Gabrielli, *Developing Virtual Synthesizers with VCV Rack*. Focal Press, 2020.
- [7] F. Duignan, J. Kidd, and A. Medina, "DSPPAL: An open source library to support real-time implementation of dsp and control algorithms on a range of hardware and software platforms," in *INTED2009 Proceedings*, 3rd International Technology, Education and Development Conference, pp. 2555–2563, IATED, 9-11 March, 2009 2009.
- [8] M. Robinson and J. Bullock, "A comparison of audio frameworks for teaching, research, and development," in *Audio Engineering Society Convention 132*, Apr 2012.
- [9] C. DeVane and G. Bunkheila, "Automatically generating VST plugins from matlab code," in *Audio Engineering Society Convention 140*, Audio Engineering Society, 2016.
- [10] W. Pirkle, *Designing audio effect plugins in C++: for AAX, AU, and VST3 with DSP theory*. Routledge, 2019.
- [11] W. C. Pirkle, *Designing Software Synthesizer Plugins in C++: For RackAFX, VST3, and Audio Units*. Routledge, 2014.
- [12] J. Parker and S. D'Angelo, "A digital model of the Buchla lowpass-gate," in *Proc. Int. Conf. Digital Audio Effects (DAFx-13)*, Maynooth, Ireland, pp. 278–285, 2013.
- [13] M. Subotnick, "The use of the Buchla synthesizer in musical composition," in *Audio Engineering Society Convention 38*, Audio Engineering Society, 1970.
- [14] T. Gordon, "The Buchla Music Easel: From cyberculture to market culture," in *Modular Synthesis*, pp. 49–80, Focal Press, 2024.
- [15] L. Gabrielli and L. Turchet, "Towards a sustainable internet of sounds," in *Proceedings of the 17th International Audio Mostly Conference*, AM '22, (New York, NY, USA), p. 231–238, Association for Computing Machinery, 2022.

A mobile game app for adaptive assessment of pitch discrimination in children with different hearing ability

Andrea Gulli, Federico Fontana

Department of Mathematics,
Computer Science and Physics
Università di Udine, Udine, Italy
gulli.andrea@spes.uniud.it
federico.fontana@uniud.it

Hanna Järveläinen

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

Michele Geronazzo

Department of Engineering
and Management
Università di Padova, Padua, Italy
michele.geronazzo@unipd.it

ABSTRACT

Recent advancements in audiological testing and rehabilitation prioritize utmost personalization and minimal stress in patients, especially the young ones. Hearing health assessments integrate innovative approaches, emphasizing ecological listening scenarios and patient engagement. Driven by these principles, a mobile application tailored for pitch discrimination in children has been designed. The app implements intuitive game mechanics on a captivating graphical interface and harnesses machine learning algorithms to adapt the sound pressure levels to individual comfort levels. It utilizes simple yet effective acoustic stimuli obtained from second-order digital resonators, ensuring a more ecological approach. The pitch discrimination threshold is obtained with adaptive psychometric techniques to guarantee reliable and faster measurements. Preliminary qualitative evaluations involving normal hearing and single-sided deaf with cochlear implant children yield promising outcomes. The resulting perceptual thresholds align with established literature, envisioning the app’s efficacy in delivering accurate assessments. The presented tool paves the way for the use of gameplay in young hearing-impaired individuals’ rehabilitation after treatment with cochlear implants.

1. INTRODUCTION

In everyday life, the ability to discriminate pitch is important for music perception [1] and verbal communication [2]; sequences of pitches form melodies, and prosodic information in spoken languages carry important information, especially in tonal languages [3]. Pitch is also an essential perceptual cue for segregating sound sources [4]. Reduced access to information about changes in speech fundamental frequency (F_0 , an acoustic correlate of pitch) degrades speech-in-noise recognition ability [5].

Although many cochlear implant (CI) users have good word recognition capabilities in quiet conditions, they struggle to detect differences in pitch and timbre [6]. CI algorithms analyze the envelope of an incoming sound and

allocate spectral energy to electrodes positioned along the cochlea’s basilar membrane, depending on their central frequency [7]. Pitch, as a perceptual category largely based on the sound’s frequency content, can be encoded with different cues, such as place-pitch (selecting a cochlear region of excitation) and temporal-pitch (selecting a period of excitation for the electrodes) [8]. The place-pitch encoding mapping is based on the well-known tonotopic map of the cochlea [9]. Nevertheless, maps can vary across individuals and even within the same individual over time [10]. Moreover, neural plasticity deviations linked to hearing loss result in tonotopic map changes [11].

Among patients rehabilitated for hearing loss, the largest interaural mismatch exists in cases with one CI and acoustic hearing of whatever degree—from normal to severely impaired—in the contralateral ear [12]. If supported by a hearing aid (HA), acoustic hearing is referred to as a “bimodal” strategy. In the case of normal hearing, it is indicated as a “single-sided deafness and CI” (SSD-CI) strategy. The degree to which asymmetrically impaired patients can benefit from binaural processing in the brainstem largely depends on the fitting of their devices, which involves the critical dimensions of level, latency, and frequency. A mismatch in any of these dimensions can interfere with the benefits of binaural processing. Moreover, the interaural difference in each dimension is frequency band-specific. An ideal fit for the level dimension needs to be more than simply adjusting an amplitude-scaling factor.

In electrical stimulation, the perceived pitch depends on rate, level, waveform, and the place of stimulation, namely the position of the electrode [13]. To measure the pitch discrimination ability of an SSD-CI listener, we considered all these parameters except for the positioning of the electrodes. We treated the CI as a “black box” because we do not know enough about the electrically elicited representation of speech at the periphery and in the cortex to generate an appropriate model. Thus, we adopted an ecological approach that drove our stimulus design and delivery choices.

The stimuli used in pitch discrimination experiments are usually pure tones [14, 15] or 1/3-octave band noise [16]. Optimal stimuli have been designed to simulate the broad spread of excitation, with an adjustable pulse rate to minimize intrinsic modulations after auditory filtering [17, 18]. In discrimination experiments with complex tones, just-noticeable differences ranging from less than one semi-

Copyright: ©2024 Andrea Gulli, Federico Fontana et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

tone [19] to 5.7 semitones [4] were found. Another study reported large individual variability in pitch-ranking thresholds, ranging from 1 to 8 semitones [20]. Studies using real musical instruments have estimated pitch discrimination thresholds for CI listeners of around 80%–90% [21, 22]. In all of these studies, the variance across participants was considerable, with some participants only being able to discriminate pitch with a frequency difference slightly less than 100% (one octave) and best individual thresholds around 3% (0.5 semitones). The best CI users performance is similar to pitch discrimination thresholds with musical instruments for normal-hearing listeners [21].

In the current study, an exponentially decaying sinusoidal stimulus was used because of its simplicity and ecological characteristics. Such a stimulus is widely adopted by the sound and music computing community [23]. Compared to the use of steady tones, models based on exponentially decaying sinusoids are more suitable for the simulation of transients such as those found in speech [24] or collision events [25]. The physical representation of a vibrating object under free oscillations through this elementary model convinced us to use it as a viable probe to evaluate pitch discrimination in everyday contexts.

Another important step toward an ecological hearing evaluation was to move the pitch discrimination test outside the clinical setting through the use of a mobile application. A wide variety of apps for audiometric tests have been made available and evaluated [26]; an audio app provided reliable hearing thresholds with test-retest reliability comparable to clinical audiometry [27], and users preferred to perform the test at home instead of the hospital. We developed a mobile app on iOS platforms to measure the individual pitch discrimination threshold of SSD-CI children and checked with preliminary testing that the results are aligned with the literature.

2. DESIGN PRINCIPLES

2.1 Stimuli

Acoustic stimuli were exponentially decaying sinusoids obtained as impulse responses of a resonant second-order infinite impulse response (IIR) filter [28]:

$$H(z) = G \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}}, \quad (1)$$

with

$$\begin{aligned} r &= 10^{-\log_{10} \frac{\text{dB}}{20} / \tau F_s}, \\ a_1 &= -2r \cos \frac{2\pi f}{F_s}, \quad a_2 = r^2, \\ G &= (1 - r) \sqrt{1 + a_2 - 2r \cos \frac{4\pi f}{F_s}}, \end{aligned}$$

where r is the magnitude of the poles, F_s is the sampling frequency, a_1 and a_2 are the filter coefficients. The stimuli were designed by setting the parameters τ , determining timbre and duration, and f setting the central frequency. With $F_s = 44.1$ kHz, τ was constantly kept to 0.1, and f was varied accordingly to the adaptive procedure described later.

The audio during the test was synthesized at runtime with Pure Data running on an Apple iPad Pro 11-inch 3rd generation 256 GB via libpd [29]. Stimuli were delivered via Sennheiser HD 600 open-back headphones. In view of working with children with higher audiometric thresholds than normal population, and to provide sufficient loudness at the listening point without introducing distortion, the digital output from the iPad was sent to an Antelope Zen Go Synergy Core USB sound interface via a USB-C 3.5 mm jack adapter, whose analog output in its turn was amplified by an RME Babyface Pro sound card. The resulting system was calibrated with a tolerance of 2.5 dB SPL and a peak amplitude of 104 dB SPL.

2.2 Method

The pitch discrimination session consisted of five stages:

1. a preliminary task illustration with a set of variable numbers of training trials performed collaboratively by the participant, the experimenter, and possibly a parent and the speech therapist;
2. a loudness balancing stage in which the participant adjusted the stimuli to the most comfortable loudness (MCL), and defined a contralateral equal-loudness contour for each one;
3. a two-interval, two-alternative forced choice (2I-2AFC) adaptive pitch discrimination test;
4. a catch trial, interrupting the pitch discrimination as a distractor task, where the listener has to answer the question “Do you hear one or two different sounds?” after a pair of pure tones is presented with the same frequencies as the last pair of stimuli;
5. an evaluation of the quality of the experience through short questionnaires on task load and usability.

2.2.1 Loudness balancing

Sound pressure levels in the tested ear were adjusted to a comfortable level, starting from 50 dB SPL. Participants could adjust the level with three animated buttons increasing SPL by 1 or 4 dB steps and decreasing it by 1 dB steps. The equal loudness of pure tones is a function of frequency. Fletcher and Munson [30] determined the first equal-loudness contours (ELCs) for normal-hearing (NH) listeners, and these still form the basis for the A-weighted SPL. These contours are the starting point of the ELC estimations. For each test frequency, we found the curve whose level is nearest to the corresponding frequency’s MCL. Then an adaptation of the procedure described in [31] starts, using a Bayesian machine-learning approach [32]. A Gaussian Process (GP) regression [33] was utilized to estimate the ELC by interpolating data from previous trials. The estimated curve was represented as a Gaussian distribution. Frequencies were selected on a grid with a spacing of 0.1 Cam on the Equivalent Rectangular Bandwidth of the Auditory Filter (ERBN)-number scale [34], ranging from 250 Hz (6.8 Cam) to 8000 Hz (33.2 Cam).

For each subsequent trial, the GP regression predicted responses on a grid with a resolution of 1 dB in level and 0.1 Cam in ERBN-number. The GP employed a constant mean function (μ) and a Matérn covariance function with a ν parameter. The Matérn kernel can be seen as a generalization of the Gaussian radial basis function [35]. It furthermore contains the absolute exponential kernel, and it is better in capturing the physical processes due to its finite differentiability (for finite ν). The hyperparameters μ and ν were determined through likelihood maximization at each iterative step. The μ parameter drives the offset of the overall curve, while the ν parameter controls the smoothness of the function. For the next trial, the frequency whose level required adjustment by the participant was randomly selected from all points where mutual information [36] exceeded 90% of that for the point with maximum mutual information. During each iteration, previously defined frequency levels were utilized to determine the Fletcher and Munson curve that minimized the root-mean-square error (RMSE) at those frequencies. The points in this curve were then integrated into the user-defined frequency - level points for subsequent interpolation. During testing, each participant heard the test frequency set at the MCL in one ear, while a different frequency stimulus was presented to the contralateral ear with a 500 ms stimulus onset asynchrony. A layout similar to the one used for MCL determination allowed level adjustments, with buttons enabling increments of 1 or 4 dB SPL and decrements of 1 dB SPL.

2.2.2 Pitch Discrimination

In the 2I-2AFC pitch discrimination adaptive procedure two stimuli are presented on each trial with slightly different frequencies, and the listener must decide whether the latter has a higher pitch than the former. We used a weighted up-down method [37] similar to the simple up-down rule. The magnitude of the frequency step between the tones to be compared was determined to track the 75% correct point on the psychometric function. After every wrong answer, the relative frequency difference $|\Delta f/f|$ was multiplied by 2; otherwise, it was divided by $2^{1/3}$. The test frequency, heard for first, is kept constant throughout the whole pitch discrimination sequence. For each sequence, the step size was initially set to 10% and adapted with a single adaptive track for rapid convergence for at least 16 trials. This value was empirically set to let the experiment last approximately 30 minutes. After 16 trials the mean and the confidence interval at 95% are computed. If the last frequency step is found within that interval, the resulting threshold is considered sufficiently stable, the pitch discrimination sequence terminates, and the participant starts with the next sequence; otherwise, another sequence of 16 trials starts. Between the two sequences, a catch trial is inserted. These catch trials aimed to convey relaxation to participants by providing a lapse of pause through a relatively easier task with a different acoustic stimulation (pure tones) and graphic interface interrupting the horizontal scrolling of the screen.

Three frequencies were tested: 750 Hz, 1 kHz, and 2 kHz. Besides tuning well with the “speech banana” [38],

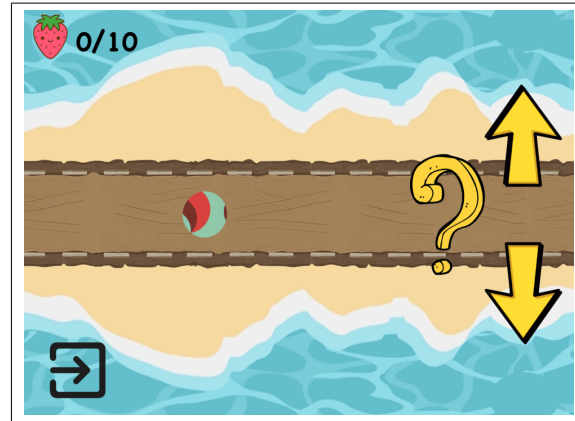


Figure 1. GUI for a pitch discrimination trial.

they lie in the frequency range in which CIs are effective. Since speech quality is quite satisfactory in the 300–3400 Hz frequency range, CI companies peak their devices’ performances around that band [39]. The pitch discrimination session lasted approximately 30 minutes. Data was collected and saved on the device.

2.2.3 Graphical User Interface

A screenshot of the pitch discrimination GUI is illustrated in Figure 1. The GUI and the interaction dynamics are designed to create an engaging experience, providing elements of serious game mechanics [40]. Two bouncing “marbles” are positioned on the sides of the screen. The two marbles run on a track, and the screen scrolls horizontally without noise. When they bounce, corresponding stimuli are presented. The runaway marble bounces when the test stimulus is delivered. After it jumps, it rolls out of the screen. Contextually, a question mark and two arrows pointing up and down appear. During the task illustration, the task to answer whether the second stimulus is higher or lower than the first is explained to the participants as “Where did the marble go, up or down?” This visual design choice was based on consolidated space-pitch associations made by infants across cultures [41]. Feedback was provided with an animation that showed the two marbles traveling the same route and the chaser approaching the runaway when the answer was correct, and two different routes and more distanced marbles otherwise. A strawberry was collected as a reward for providing the correct answer and enhancing engagement [42], while a score counter was visible to track performance. We note that during experimentation, it occurred several times that the participant compared the score made in previous runs with the current score.

2.2.4 Quality of the Experience

Sessions included assessments using a modified version of the NASA Raw Task Load Index (RTLX) [43] and a Short User Experience Questionnaire (UEQ-S) [44] to gauge workload and subjective user experience. Experimenters and parents aided participants in completing the questionnaires,

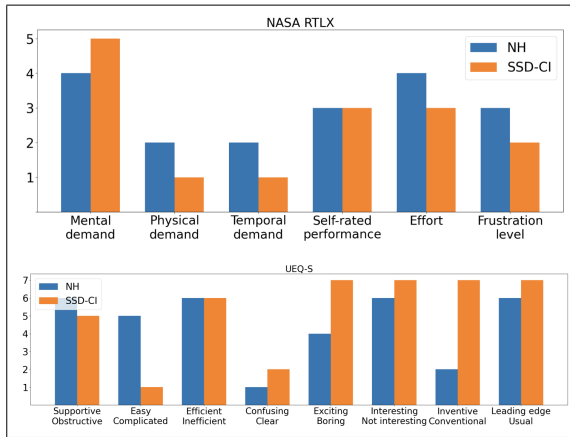


Figure 2. NASA RTLX and UEQ-S results.

followed by discussions of the results.

3. PRELIMINARY RESULTS

Two participants, including an 18-year-old male with single-sided deafness and cochlear implant (SSD-CI) and a 9-year-old male with normal hearing (NH) took part in the study. Ethical standards set by the Department of Mathematics, Computer Science, and Physics at Udine University were followed, with approval from the Review Board obtained on August 31, 2023. Both participants provided written informed consent, adhering to the Declaration of Helsinki guidelines, facilitated by their parents during their minor status. Recruitment occurred through online presentations to users and clients of Studio Riabilitare Insieme, a rehabilitation facility in Maerne (Venice), Italy.

NH child exhibited thresholds of 1.98% for all tested frequencies, while the SSD-CI listener showed thresholds of 66.14%, 69.45%, and 93.75% for 750 Hz, 1 kHz, and 2 kHz, respectively. These thresholds align with the literature, indicating the NH child's proficient discrimination ability and the SSD-CI listener's challenges, consistent with prior studies [21, 22]. We note that the SSD-CI listener's thresholds increase monotonically with the tested frequency.

The results of the NASA RTLX and UEQ-S are visually displayed in Figure 2. NASA RTLX results indicate high mental demand for both participants, but higher demand for the SSD-CI listener. The NH child faced higher physical and temporal loads, probably due to his young age. Both participants rated their performance the same, even if the SSD-CI listener's thresholds were largely exceeding the NH's. The NH listener reported higher effort and frustration than the SSD-CI participant. Likely this evidence is related to the fact that the SSD-CI listener could detect only large pitch differences between the stimuli. Accordingly, the SSD-CI should be involved in a training protocol to appreciate more subtle changes and to be challenged with greater commitment as a result. With a cochlear implant simulator [45], we can see that processed stimuli display broader spectral energy (see Figure 3).

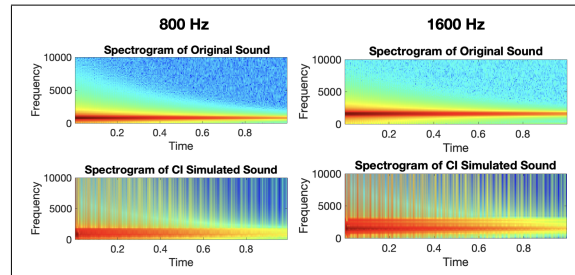


Figure 3. Spectrograms of two decaying sinusoidal stimuli with central frequencies of 800 Hz and 1.6 kHz and their CI-processed respective.

Moreover, UEQ-S outcomes suggest the NH child's familiarity with mobile games, as evidenced by lower excitement levels than the SSD-CI listener. The latter found the app less conventional and more confusing than the NH child, likely due to the difficulty in understanding what task to perform in pitch discrimination. We report that during the pitch discrimination test, participants were generally engaged by the game, feeling frustrated by a wrong answer and enthusiastic when the correct one was hit instead.

4. CONCLUSION

Our study presented a mobile application tailored for pitch discrimination in children, integrating innovative approaches to enhance ecological listening scenarios and patient engagement while prioritizing personalization and minimizing stress. Through intuitive game mechanics, Gaussian contours, and adaptive psychometric techniques, the app offers a user-friendly and effective means of assessing hearing health. Preliminary qualitative evaluations involving participants with normal hearing and a monolateral cochlear implant have yielded promising outcomes. The perceptual thresholds align with established literature, suggesting potential for this app to deliver accurate assessments. The primary limitation of the current study lies in the limited pool of participants. Future endeavors will prioritize a devoted examination of the loudness balancing procedure and incorporate additional stimuli, such as the pulse-spreading harmonic complex [17], into a pitch discrimination test tailored for CI users. Furthermore, integrating our experimental approach with preliminary audiometric assessments and subsequent rehabilitation therapies may offer a more comprehensive evaluation scenario. Future work will address different user groups' varying needs and experiences, as evidenced by differences in subjective workload and user experience between participants. Further research is warranted to refine the app's design and functionality, considering users' feedback and the complexities of auditory processing in individuals with cochlear implants. In conclusion, the presented app is a step toward more ecological and engaging audiological assessment and rehabilitation.

5. ACKNOWLEDGMENT

Roberta Rebesco's support to participants' recruitment and testing was funded by the Institute for Maternal and Child Health IRCCS "Burlo Garofolo" (Trieste, Italy) under the project "Ricerca Corrente 42/22".

This work was funded by NextGenerationEU PNRR M4-C2.1.1 PRIN 2022 project "S-TWIN" - 2022F9FWZ8_003 - CUP G53D23002840006.

6. REFERENCES

- [1] B. C. Moore, *An introduction to the psychology of hearing*. Brill, 2012.
- [2] J. S. Laures and G. Weismer, "The effects of a flattened fundamental frequency on intelligibility at the sentence level," *Journal of Speech, Language, and Hearing Research*, vol. 42, no. 5, pp. 1148–1156, 1999.
- [3] M. L. Deroche, H.-P. Lu, A. M. Kulkarni, M. Caldwell, K. C. Barrett, S.-C. Peng, C. J. Limb, Y.-S. Lin, and M. Chatterjee, "A tonal-language benefit for pitch in normally-hearing and cochlear-implanted children," *Scientific Reports*, vol. 9, no. 1, p. 109, 2019.
- [4] W. Wang, N. Zhou, and L. Xu, "Musical pitch and lexical tone perception with cochlear implants," *International journal of audiology*, vol. 50, no. 4, pp. 270–278, 2011.
- [5] C. Binns and J. F. Culling, "The role of fundamental frequency contours in the perception of speech against interfering speech," *The Journal of the Acoustical Society of America*, vol. 122, no. 3, pp. 1765–1776, 2007.
- [6] C. J. Limb and A. T. Roy, "Technological, biological, and acoustical constraints to music perception in cochlear implant users," *Hearing research*, vol. 308, pp. 13–26, 2014.
- [7] K. Vermeire, D. M. Landsberger, P. H. Van de Heyning, M. Voormolen, A. K. Punte, R. Schatzer, and C. Zierhofer, "Frequency-place map for electrical stimulation in cochlear implants: Change over time," *Hearing Research*, vol. 326, pp. 8–14, 2015.
- [8] V. Marimuthu, B. A. Swanson, and R. Mannell, "Cochlear implant rate pitch and melody perception as a function of place and number of electrodes," *Trends in Hearing*, vol. 20, p. 2331216516643085, 2016.
- [9] G. Von Békésy, "Experiments in hearing.," 1960.
- [10] M. Saenz and D. R. Langers, "Tonotopic mapping of human auditory cortex," *Hearing research*, vol. 307, pp. 42–52, 2014.
- [11] E. A. Koops, R. J. Renken, C. P. Lanting, and P. van Dijk, "Cortical tonotopic map changes in humans are larger in hearing loss than in additional tinnitus," *Journal of Neuroscience*, vol. 40, no. 16, pp. 3178–3185, 2020.
- [12] S. H. Pieper, N. Hamze, S. Brill, S. Hochmuth, M. Exter, M. Polak, A. Radeloff, M. Buschermöhle, and M. Dietz, "Considerations for fitting cochlear implants bimodally and to the single-sided deaf," *Trends in hearing*, vol. 26, p. 23312165221108259, 2022.
- [13] B. Townshend, N. Cotter, D. Van Compernelle, and R. White, "Pitch perception by cochlear implant subjects," *The Journal of the Acoustical Society of America*, vol. 82, no. 1, pp. 106–115, 1987.
- [14] D. Arzounian, M. de Kerangal, and A. de Cheveigné, "A sliding two-alternative forced-choice paradigm for pitch discrimination," *The Journal of the Acoustical Society of America*, vol. 142, no. 1, pp. 167–172, 2017.
- [15] L. Wagner, R. Altindal, S. K. Plontke, and T. Rahne, "Pure tone discrimination with cochlear implants and filter-band spread," *Scientific Reports*, vol. 11, no. 1, p. 20236, 2021.
- [16] Y. Adel, S. Nagel, T. Weissgerber, U. Baumann, and O. Macherey, "Pitch matching in cochlear implant users with single-sided deafness: Effects of electrode position and acoustic stimulus type," *Frontiers in neuroscience*, vol. 13, p. 470018, 2019.
- [17] G. Hilkhuisen and O. Macherey, "Optimizing pulse-spreading harmonic complexes to minimize intrinsic modulations after auditory filtering," *The Journal of the Acoustical Society of America*, vol. 136, no. 3, pp. 1281–1294, 2014.
- [18] Q. Mesnildrey, G. Hilkhuisen, and O. Macherey, "Pulse-spreading harmonic complex as an alternative carrier for vocoder simulations of cochlear implants," *The Journal of the Acoustical Society of America*, vol. 139, no. 2, pp. 986–991, 2016.
- [19] X. Luo, M. E. Masterson, and C.-C. Wu, "Melodic interval perception by normal-hearing listeners and cochlear implant users," *The Journal of the Acoustical Society of America*, vol. 136, no. 4, pp. 1831–1844, 2014.
- [20] R. Kang, G. L. Nimmons, W. Drennan, J. Longnion, C. Ruffin, K. Nie, J. H. Won, T. Worman, B. Yueh, and J. Rubinstein, "Development and validation of the University of Washington clinical assessment of music perception test," *Ear and hearing*, vol. 30, no. 4, pp. 411–418, 2009.
- [21] S. Brockmeier, D. Fitzgerald, O. Searle, H. Fitzgerald, M. Grasmeyer, S. Hilbig, K. Vermiere, M. Peterreins, S. Heydner, and W. Arnold, "The MuSIC perception test: a novel battery for testing music perception of cochlear implant users," *Cochlear implants international*, vol. 12, no. 1, pp. 10–20, 2011.
- [22] L. Bruns, D. Mürbe, and A. Hahne, "Understanding music with cochlear implants," *Scientific reports*, vol. 6, no. 1, p. 32026, 2016.

- [23] R. Boyer and K. Abed-Meraim, “Audio modeling based on delayed sinusoids,” *IEEE Transactions on speech and audio processing*, vol. 12, no. 2, pp. 110–120, 2004.
- [24] K. Hermus, W. Verhelst, P. Lemmerling, P. Wambacq, and S. Van Huffel, “Perceptual audio modeling with exponentially damped sinusoids,” *Signal processing*, vol. 85, no. 1, pp. 163–176, 2005.
- [25] M. Aramaki, M. Besson, R. Kronland-Martinet, and S. Ystad, “Controlling the perceived material in an impact sound synthesizer,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 2, pp. 301–314, 2010.
- [26] J. Saliba, M. Al-Reefi, J. S. Carriere, N. Verma, C. Provencal, and J. M. Rappaport, “Accuracy of mobile-based audiometry in the evaluation of hearing loss in quiet and noisy environments,” *Otolaryngology–Head and Neck Surgery*, vol. 156, no. 4, pp. 706–711, 2017.
- [27] A. Hazan, J. Luberadzka, J. Rivilla, A. Snik, B. Albers, N. Méndez, N. Wack, O. Paytuvi, A. Zarowski, E. Offeciers, *et al.*, “Home-based audiometry with a smartphone app: Reliable results?,” *American Journal of Audiology*, vol. 31, no. 3S, pp. 914–922, 2022.
- [28] G. De Poli and F. Avanzini, “Sound modeling: signal-based approaches,” *Algorithms for Sound and Music Computing*, pp. 2005–2012, 2005.
- [29] P. Brinkmann, D. Wilcox, T. Kirshboim, R. Eakin, and R. Alexander, “Libpd: Past, present, and future of embedding pure data,” in *Proceedings of the 5th Pure Data Convention (Pd Con)*, New York City, 2016.
- [30] H. Fletcher and W. A. Munson, “Loudness, its definition, measurement and calculation,” *Bell System Technical Journal*, vol. 12, no. 4, pp. 377–430, 1933.
- [31] J. Schlittenlacher and B. C. Moore, “Fast estimation of equal-loudness contours using Bayesian active learning and direct scaling,” *Acoustical Science and Technology*, vol. 41, no. 1, pp. 358–360, 2020.
- [32] A. Gulli, F. Fontana, S. Serafin, M. Geronazzo, *et al.*, “An active learning procedure for the interaural time difference discrimination threshold,” in *Proceedings of the International Conference on Digital Audio Effects, DAFx*, pp. 273–280, DAFx, 2023.
- [33] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. Adaptive computation and machine learning, MIT Press, 2006.
- [34] B. R. Glasberg and B. C. Moore, “Derivation of auditory filter shapes from notched-noise data,” *Hearing research*, vol. 47, no. 1-2, pp. 103–138, 1990.
- [35] M. L. Stein, *Interpolation of spatial data: some theory for kriging*. Springer Science & Business Media, 2012.
- [36] C. E. Shannon, “A mathematical theory of communication,” *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [37] C. Kaernbach *et al.*, “Simple adaptive testing with the weighted up-down method,” *Perception & psychophysics*, vol. 49, no. 3, pp. 227–229, 1991.
- [38] N. Nguyen, L. Curry, and M. J. Goupell, “Bionic hearing: When is it time to get a cochlear implant?,” 2023.
- [39] P. SM, “What frequency range of the cochlear implant is the best for speech perception?,” Jan. 2022.
- [40] P. K. R. S. Bhopal and N. Senan, “Hearing screening test mobile games application for kids,” *Applied Information Technology And Computer Science*, vol. 3, no. 1, pp. 33–46, 2022.
- [41] S. Dolscheid, S. Hunnius, D. Casasanto, and A. Majid, “Prelinguistic infants are sensitive to space-pitch associations found across cultures,” *Psychological Science*, vol. 25, no. 6, pp. 1256–1261, 2014.
- [42] S. Nicholson, “A recipe for meaningful gamification,” *Gamification in education and business*, pp. 1–20, 2015.
- [43] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” in *Advances in psychology*, vol. 52, pp. 139–183, Elsevier, 1988.
- [44] M. Schrepp, A. Hinderks, and J. Thomaschewski, “Design and evaluation of a short version of the user experience questionnaire (UEQ-S),” *International Journal of Interactive Multimedia and Artificial Intelligence*, 4 (6), 103-108., 2017.
- [45] H. J. McDermott, C. M. McKay, and A. E. Vandali, “A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant,” *The Journal of the Acoustical Society of America*, vol. 91, no. 6, pp. 3367–3371, 1992.

WAVEPILOT: FRAMEWORK MULTIDIMENSIONALE PER L'ESPLORAZIONE DELLO SPAZIO PARAMETRICO DI STRUMENTI DIGITALI

Alessandro Anatrini

Hochschule für Musik und Theater Hamburg
Conservatorio Statale di Musica J. Tomadini - Udine
alessandro@anatrini.com

ABSTRACT

La crescente diffusione di tecniche di deep learning per la generazione di contenuti multimediali ha aperto nuove possibilità nell'esplorazione sonora e visiva e nell'interazione uomo-macchina (HCI). In questo contesto presento *WavePilot* un framework progettato per agevolare l'esplorazione e la manipolazione dei parametri di Digital Multimedia Instrument (DMmI). *WavePilot* utilizza un Variational Autoencoder (VAE) per tradurre i valori di uno o più DMmI, in una rappresentazione multidimensionale del suo spazio parametrico. L'obiettivo principale è facilitare l'interazione, offrendo all'utente un'interfaccia grafica (GUI) di alto livello sotto forma di spazio virtuale navigabile che semplifichi la programmazione dei DMmI. A differenza di approcci più convenzionali e limitati al campo della rappresentazione multidimensionale dello spazio timbrico di un Digital Music Instrument (DMI), che si basano sull'analisi delle caratteristiche percettive dell'audio, qui ci si concentra esclusivamente sui valori parametrici al fine di fornire un contesto di rappresentazione visiva immediato, agnostico verso la tipologia di dispositivi cui viene applicato, siano essi generatori audio, video o effetti.

Questo approccio mira a stimolare una riflessione critica tra i diversi attori del campo musicale e tecnologico quali musicisti, ricercatori, sviluppatori e designer. L'attenzione è posta sulle modalità di interazione con il materiale artistico e su come queste possono essere ripensate alla luce delle nuove possibilità offerte dai dispositivi basati su IA. Questo articolo fornisce le motivazioni che hanno portato alla realizzazione di *WavePilot*, insieme a un'analisi della sua architettura e delle modalità di utilizzo.

1. MOTIVAZIONI

WavePilot è frutto dell'affinamento della tecnica che l'autore ha elaborato per mappare automaticamente DMmI nella propria pratica artistica. A partire dal 2021 sono stati presentati tre lavori che, pur avvalendosi di metodologie differenti utilizzano tecniche di riduzione dimensionale dello spazio parametrico della sintesi audiovisiva: *Wun-*

derkammer (2021), *Here Be Dragons* (2021), *Sandy Island* (2022)¹. Nonostante le differenti esigenze tecniche, questi lavori condividono la necessità di gestire i parametri dei DMmI tramite una GUI che offra un'interazione di alto livello, una volta delimitato lo spazio di ricerca parametrico. Originariamente, *WavePilot* è stato concepito come uno strumento specificamente dedicato alla gestione dei parametri della sintesi sonora, successivamente, la sua applicabilità è stata estesa per gestire anche contenuti multimediali. La definizione preliminare dello spazio di ricerca è operazione di fondamentale importanza poiché indipendentemente dal contesto, di norma solo una ristretta gamma di risultati sonori ad esempio, sarà appropriata alla natura dell'intervento sia esso di natura installativa, live performance o fixed media.

L'interesse verso questa tipologia di approccio alla composizione e la consapevolezza dell'importanza di circoscrivere in partenza lo spazio della ricerca sonora, sono maturati nell'ambito del progetto di ricerca *Healing Soundscape* sviluppatosi tra il 2016 e il 2018 come collaborazione tra i dipartimenti di multimedia e musicoterapia della Hochschule für Musik und Theater (HfMT) e la clinica universitaria di Amburgo (UKE) al quale l'autore ha preso parte come compositore e sviluppatore. Il progetto era rivolto allo studio della riduzione dei fattori di stress negli utenti della struttura ospedaliera tramite l'esposizione ad ambienti sonori generativi appositamente progettati [1]. Laddove solo un set estremamente ristretto di stimoli sonori risulta appropriato al contesto, si pone ancor più stringente la necessità di:

- **Identificare i parametri della sintesi** capaci di fornire risultati in linea con determinate linee di intervento;
- Fornire ad utenti non esperti modalità di **interazione di alto livello** con generatori audio che diversamente presenterebbero una ripida curva di apprendimento.

2. CONTESTO

WavePilot, pur evolvendosi nel tempo in uno strumento agnostico capace di gestire una varietà di contenuti, affonda le sue radici nelle ricerche nel campo della programmazione automatica dei DMI (ASP). Questo strumento si svi-

¹ Per una descrizione dettagliata dei lavori citati si faccia riferimento a www.anatrini.com.

luppa nel solco di alcune idee relative alla notazione musicale: l'autore avverte infatti l'esigenza di superare, nella sua pratica artistica, la tradizionale separazione tra il processo di programmazione dei parametri della sintesi sonora e l'attività compositiva in sé. In questa sezione, si farà riferimento alla letteratura pertinente al fine di posizionare questo dispositivo nel contesto attuale della ricerca nel campo della programmazione automatica dei DMI.

2.1 Meta-GUI come partitura

Negli ultimi decenni il campo dei DMI ha visto un fiorire di approcci interdisciplinari che sempre più spesso mettono insieme differenti professionalità: musicisti, sviluppatori, designer. Dal punto di vista operativo ciò si traduce in una tendenza verso il superamento dei confini che delimitano i tradizionali ruoli di compositore-interprete-ascoltatore nel contesto musicale. Già nel 2002 si faceva notare come i sistemi tecnologici utilizzati nelle performance musicali contenessero in sé sia la nozione di strumento che quella di partitura [2]. La strada delineata negli ultimi anni sembra andare nella direzione di un "approccio ecologico al fare musica" che si concentra, tra le altre cose, sul contesto di azione tra elementi umani e tecnologici [3] e sulla mediazione operata dagli strumenti della tecnica [4]. Da qui si traccia un percorso che va nella direzione di una progressiva integrazione di elementi simbolici e iscrizioni nello strumento stesso [5]. È proprio questa riflessione che apre nuove prospettive nell'*artistic research* dove la tradizionale tripartizione tra "research-on -for -in-the arts" [6] pare oramai superata dalle attuali dinamiche compositive e artistiche *tout court*.

La pratica compositiva dell'autore, sensibile a questa famiglia di approcci, mira a integrare in modo organico i diversi aspetti del sound design, della generazione di materiale e della gestione formale dell'atto compositivo attraverso l'uso di una meta-GUI, come quella qui proposta. Il termine meta-GUI è utilizzato in questo contesto per indicare una GUI di alto livello che viene applicata a uno o più dispositivi. Essa offre un controllo immediato dello strumento fornendo un livello di astrazione che semplifica l'interazione. Un elemento distintivo di questo approccio è l'implementazione della Meta-GUI su browser. Questa scelta offre numerosi vantaggi, tra i quali compatibilità tra differenti sistemi operativi, portabilità, facilità di manutenzione e accessibilità universale [7]. Inoltre le applicazioni scritte in JavaScript e progettate per funzionare nel browser offrono portabilità tra piattaforme mobili e desktop, e longevità grazie alla standardizzazione. Data l'ubiquità del browser sia su desktop che su dispositivi mobili, esso è una delle classi di applicazioni più diffuse e rappresenta una soluzione "write once, run anywhere" per le interfacce musicali [7]. Questo approccio è inoltre facilmente estendibile a contesti multimediali, e mira a contribuire alle metodologie compositive multimediali basate su traiettorie spaziali in uno spazio virtuale multidimensionale.

L'idea di mappare caratteristiche musicali ad elementi grafici non è certamente nuova e trova antecedenti illustri nel campo dell'informatica musicale. Dal *Graphic I* di Matthews [8] che per la prima volta permetteva di speci-

care una partitura come un gruppo di grafi, al sistema *UPIC* di Xenakis nel quale il livello macro-compositivo consentiva di specificare una partitura tempo-frequenza fatta di linee, curve e punti, fino ad approcci più vicini a noi nel tempo dalle più marcate finalità didattiche [9] [10]. La novità di questo approccio consiste nel basarsi su una meta-GUI agnostica degli strumenti cui viene applicata e che pertanto consente una gamma di operazioni di crescente complessità: dal fornire un'interfaccia intuitiva di alto livello per affrontare le problematiche relative alla programmazione dei DMI, alla possibilità di definire aspetti quegli macroformali dell'evoluzione di un'opera e configurandosi pertanto come affine a una *iperpartitura*².

2.2 Perdere il controllo per acquisire complessità

La motivazione della programmazione automatica dei DMI risiede nella riduzione della curva di apprendimento necessaria per poter utilizzare uno strumento musicale digitale, di norma un sintetizzatore, in maniera consapevole ed espressiva [12]. La difficoltà di lavorare coi sintetizzatori digitali è stata individuata nella distanza concettuale che si crea tra l'interfaccia di controllo (spazio dei parametri), il suono effettivamente prodotto dallo strumento (spazio percettivo) e come viene interpretato qualitativamente dall'utente (spazio semantico) [13]. In questo senso, le procedure algoritmiche che consentono all'utente di automatizzare la sua interazione con lo strumento servono a colmare questa lacuna concettuale fornendo strumenti di controllo a livello più elevato e/o più intuitivi.

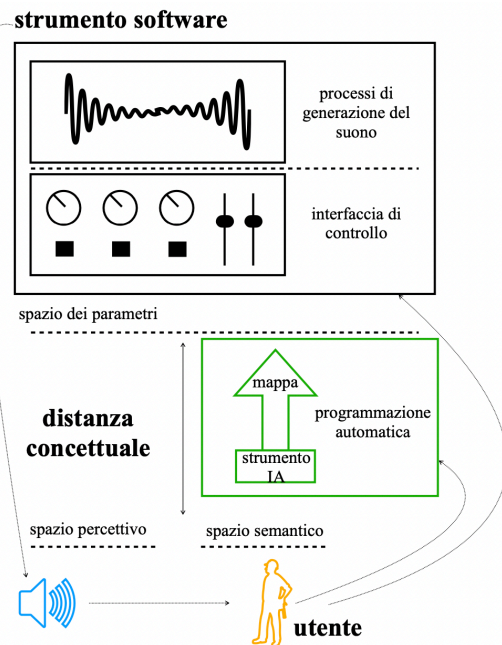


Figure 1. Rappresentazione della distanza concettuale tra utente, GUI e suono prodotto in uno strumento software.

² Per una definizione del concetto di iperpartitura si rimanda a [11].

All'interno del dominio di ricerca dell'ASP, si possono individuare due categorie distinte di approcci:

- **Esplorazione dei parametri:** il sistema genera nuovi set di parametri basandosi sulle valutazioni degli utenti, spesso utilizzando algoritmi genetici [14] [15], strategie di mapping [16] [17], o apprendimento non supervisionato [18]. Questi approcci si focalizzano sull'individuazione e la manipolazione dei parametri senza richiedere una conoscenza preventiva dell'output del sistema;
- **Sound matching:** questo approccio mira a produrre parametri di sintesi che minimizzino la differenza percettiva tra il suono generato e un suono target. Sebbene non operi direttamente sui parametri della generazione sonora, si basa sull'analisi delle caratteristiche spettrali e percettive del target. Utilizzando algoritmi evolutivi [19] o tecniche di deep learning [20] [21], questo approccio fornisce una buona approssimazione del target, ma richiede una conoscenza preventiva dell'output desiderato e un precedente addestramento dello strumento.

Come si può constatare, la programmazione automatica dei DMI si basa tradizionalmente sull'analisi delle caratteristiche percettive e spettrali del suono per determinare i parametri in grado di produrre risultati sonori specifici. Questo approccio è motivato dal fatto che, indipendentemente dal metodo di sintesi, il legame tra il controllo e le caratteristiche percettive del suono risultante è generalmente debole e difficile da stabilire con precisione [22]. Lo strumento presentato adotta un approccio basato sull'esplorazione dei parametri. Non richiede una conoscenza preliminare dell'output del sistema e mira a mappare automaticamente lo spazio parametrico di DMmi ad uno spazio virtuale a bassa dimensionalità. A differenza di altri contributi in questo campo [18], la costruzione dello spazio virtuale non si basa sull'impiego di descrittori audio per correlare le qualità timbriche di un suono al set specifico di parametri che lo generano. *WavePilot* non si propone di rappresentare lo spazio timbrico di uno strumento, né di indagare le complesse relazioni non lineari tra le caratteristiche timbriche di un suono e i parametri che lo generano. Piuttosto, l'obiettivo di *WavePilot* è fornire un ambiente agnostico in termini di numero e tipologia di dispositivi cui può essere applicato, estendendo il suo campo di azione oltre i singoli generatori audio. La finalità di questo strumento è promuovere un'esplorazione immediata dello spazio parametrico, ciò si traduce in una complessità d'interazione resa possibile grazie alla meta-GUI. Poiché l'architettura basata su VAE produce una mappatura non lineare dei parametri all'interno di uno spazio multidimensionale, è possibile modificare simultaneamente tutti i parametri muovendo il cursore che definisce il punto di lettura all'interno dello spazio virtuale. Questa interazione non lineare tra lo spazio virtuale e i parametri del sistema offre un controllo immediato, facilitando l'esplorazione e la scoperta di nuove configurazioni sonore. La rappresentazione a bassa dimensionalità dello spazio parametrico assume di conseguenza una valenza percettiva ambigua.

Il carattere agnostico del tool implica che la valenza semantica delle dimensioni ridotte sia fortemente influenzata dai preset scelti in partenza. Da un lato, questa ambiguità deriva dall'assenza di una fase di analisi che correli i set di parametri con i timbri che essi producono. Dall'altro un gradiente di ambiguità è insito nell'adozione di tecniche di riduzione dimensionale. Nonostante esse siano ampiamente utilizzate per la rappresentazione di set di dati ad alta dimensionalità, le dimensioni ridotte sono "entangled", ovvero rappresentano combinazioni non lineari delle dimensioni originali e non hanno necessariamente un'interpretazione diretta nel contesto dei dati originali. In ogni caso, l'impiego di tecniche di riduzione dimensionale comporta inevitabilmente una certa perdita di informazione: questo è uno dei motivi per i quali l'identificazione dei valori percettivi e semantici nelle dimensioni di uno spazio dimensionale ridotto rappresenta una delle aree di ricerca più attualmente rilevanti nel campo del machine learning [23] [24].

La conseguente ridotta capacità di controllo deterministico dello strumento si colloca nel più ampio contesto della HCI, in particolare nel rapporto tra utente e sistemi che implementano forme di IA. Laddove non vi sia la codifica esplicita di regole che garantiscano un determinato spettro di risultati, emerge la necessità di creare una rappresentazione della conoscenza attraverso dataset statisticamente significativi, dai quali inferire le regole operative. A parità di input, i risultati possono variare significativamente a seconda del dataset e degli iperparametri utilizzati. Questi strumenti possono essere considerati **qualitativi**, poiché rappresentano una transizione dalla manipolazione quantitativa di dati a una definizione di identità che può essere influenzata anche da piccole variazioni. In questo senso, non siamo più di fronte a un'interazione tra utente e un sistema deterministico con degli obiettivi predefiniti, ma a un **dialogo uomo-macchina che coinvolge diverse modalità di rappresentazione della conoscenza**. L'apparente intelligenza di questi sistemi non è solo un prodotto della tecnologia, ma riflette anche la capacità dell'utente di fornire dati pertinenti durante la fase di addestramento. In altre parole, l'espressione artistica attraverso tali strumenti non è predeterminata dalla tecnologia, ma è modellata dalle informazioni che scegliamo di incorporare nel processo di addestramento.

3. WAVEPILOT

WavePilot è un framework open source³ disponibile come command-line tool e modulo Python. *WavePilot* impiega tecniche di riduzione dimensionale dello spazio parametrico e di interpolazione non lineare, esso è quindi in grado di affrontare contemporaneamente: l'esplorazione delle differenti configurazioni di DMmi, l'apprendimento automatico del macro-controllo, la capacità di interpolare tra i diversi stati dello strumento che sono stati definiti come punto di partenza. In altre parole il compito della programmazione di un DMmi è ridotto ad un problema

³ github.com/anatrini/WavePilot.

di mapping non lineare all'interno di uno spazio virtuale multidimensionale come illustrato in Figura 2.

Il tool utilizza la DAW Reaper come host per i plug-in audio cui viene applicato. Tuttavia, la sua versatilità permette di utilizzarlo in qualsiasi ambiente con supporto OSC che possa ospitare plug-in o che li integri nativamente, come TouchDesigner, Max o qualsiasi Standalone FAUST App. Il contributo alla conoscenza è rappresentato da:

1. La possibilità di costruire spazi virtuali con un numero arbitrario di dimensioni. Sebbene sia stato perlopiù utilizzato per generare uno spazio tridimensionale, grazie alla sua intuitiva navigabilità tramite superfici di controllo, impiego di sensori o motion tracking, il tool è agnostico riguardo al numero di dimensioni dello spazio ridotto;
2. La possibilità di navigare lo spazio virtuale in maniera procedurale utilizzando traiettorie temporizzate tra i punti $[p_1, p_2, \dots, p_n]$ o comandi OSC da dispositivi quali smartphone, tablet, sensori indossabili o altro. Questa caratteristica permette di modulare sia aspetti microformali a livello di sound design che strutture articolate a livello macroformale, delineate dall'esplorazione programmata dello spazio;
3. La capacità, grazie alla sua architettura, di minimizzare fino a valori trascurabili la distanza tra i parametri ricostruiti P_r a partire dai parametri ridotti p , e i parametri originali P_o ;
4. La capacità di produrre set $P_{r_{int}}$ tramite interpolazione operata contemporaneamente tra i differenti set $[p_1, p_2, \dots, p_n]$ nello spazio virtuale;
5. La caratteristica di essere agnostico riguardo la tipologia e il numero di DMmI cui può essere applicato simultaneamente, senza vi sia un numero massimo predefinito di parametri per ogni set di P_o ;
6. Funzionare come tool per il rendering automatico di set di parametri generati casualmente o meno, per lo studio di un determinato DMmI. Le coppie composte da un set P_o con associato il relativo snapshot audio possono essere impiegate per ulteriori studi sull'interazione tra i parametri dello strumento e i valori di analisi dell'audio⁴.

3.1 Riduzione dimensionale

L'architettura che consente la riduzione dimensionale è un VAE, applicato mediante transfer learning non supervisionato. Per ottimizzare l'accuratezza della ricostruzione di ogni set P_r , è necessario costituire un dataset D di un numero statisticamente rilevante di P_o . Questo può essere fatto utilizzando lo script `plugin_renderer.py`, iterando attraverso i preset disponibili nel DMmI o, se questi non

sono presenti o se si sta lavorando con più DMmI, generando preset con valori casuali⁵. Una volta addestrato il modello su D , con gli stessi valori degli iperparametri il modello può essere addestrato su un dataset di dimensioni ridotte d , in modo che solo i set P_o in esso contenuti siano trasformati in altrettanti punti p visualizzati sul browser.

È importante notare che d non necessariamente deve essere un sottoinsieme di D , in quanto costituisce solamente l'insieme di set P_o che si è scelto di rappresentare nello spazio virtuale, dai quali poi ricavare sia tutti i P_r che i $P_{r_{int}}$. In altre parole d rappresenta lo spazio parametrico di esplorazione arbitrariamente definito dall'utente. Ci sono tre metodi per selezionare le entry che costituiscono d :

1. selezione dai preset già disponibili nel DMmI;
2. generazione casuale con `plugin_renderer.py`;
3. programmazione manuale.

Il numero di entry in d parte da un minimo di 2 ed è stato testato fino a raggiungerne 250, si veda Figura 3.

La capacità del modello di ricostruire fedelmente P_{r_n} a partire da p_n è misurata dalla funzione di perdita del VAE composta di due parti: la **perdita di ricostruzione** e la **divergenza KL**. La **perdita di ricostruzione** è data dalla differenza quadratica media (MSE) tra l'output del modello e i dati originali. Questa parte della funzione di perdita cerca di far sì che il modello ricostruisca i dati il più fedelmente possibile. In termini matematici essa è espressa come:

$$L_r = \frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2 \quad (1)$$

dove x_i è il dato originale, \hat{x}_i è l'output del modello, e n è il numero totale di dati.

La **divergenza KL** è una misura della differenza tra due distribuzioni di probabilità. Nel VAE, viene utilizzata per far sì che la distribuzione latente appresa dal modello sia simile a una distribuzione normale standard. Si ha pertanto:

$$D_{KL} = -\frac{1}{2} \sum_{i=1}^n (1 + \log(\sigma_i^2) - \mu_i^2 - \sigma_i^2) \quad (2)$$

dove μ_i e σ_i sono rispettivamente la media e la varianza della distribuzione latente appresa per il dato i . La funzione di perdita totale del VAE è quindi la somma della perdita di ricostruzione e della divergenza KL, pesata da un fattore β :

$$L = L_r + \beta \cdot D_{KL} \quad (3)$$

Il modello dispone inoltre di una routine di ottimizzazione tramite grid search contenuta in `optim_grid_vae.py`

⁴ Lo sviluppo di questa funzionalità si è resa necessaria a causa del limitato supporto offerto da tool analoghi, si veda ad esempio github.com/DBraun/DawDreamer. Questa funzionalità è attualmente disponibile solo per i generatori audio usando Reaper come host

⁵ Nello script in questione è implementato un meccanismo per la stima dell'energia totale contenuta negli snapshot al fine di ignorare quelle configurazioni di parametri che generano solamente silenzio. Mediante lo stesso script è anche possibile applicare una maschera per escludere determinati parametri dello strumento dall'apprendimento.

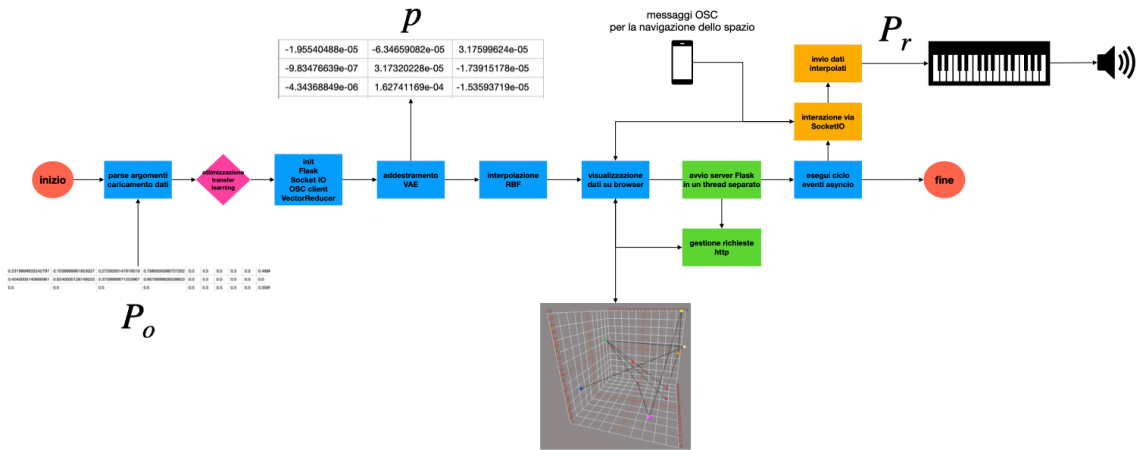


Figure 2. Schema di funzionamento di *WavePilot*. In blu le operazioni di elaborazione dati, in verde le operazioni relative al server Flask, in arancione operazioni di comunicazione in tempo reale con SocketIO, in magenta operazioni che comportano decisioni.

per determinare il valore ottimale degli iperparametri del VAE e della routine di interpolazione descritta di seguito ⁶.

3.2 Interpolazione

Al fine di calcolare i possibili $P_{r_{int}}$ originati da un qualsiasi p_{int} si utilizza una tecnica di interpolazione non lineare con funzioni radiali di base (RBF). La pertinenza di questa tecnica risiede nelle sue caratteristiche distintive: le RBF sono versatili nell’analisi di dati multivariati, indipendentemente dalla complessità delle variabili coinvolte. Le differenti RBF permettono di catturare relazioni complesse e non lineari, anche in contesti ad alta dimensionalità. La loro capacità di ricostruire esattamente i P_r partendo dai p , insieme alla generazione di nuovi $P_{r_{int}}$ realistici su superfici lisce, le rende particolarmente efficaci. Inoltre, la loro flessibilità nell’adattarsi a una vasta gamma di contesti senza la necessità di una specifica RBF predefinita è già supportata dalla letteratura in applicazioni analoghe [25].

Una RBF generica è definita come:

$$f(x) = \sum_{i=1}^N \omega_i \phi(\|x - x_i\|, \epsilon) \quad (4)$$

dove ω_i sono i pesi che vengono calcolati durante il processo di interpolazione, $(\|x - x_i\|)$ è la distanza euclidea tra il punto di dati x_i e il punto in cui si desidera valutare la funzione interpolante e ϕ è la funzione RBF o kernel che, ad esempio nel caso gaussiano, è definito come:

$$\phi(r, \epsilon) = e^{-(\epsilon r)^2} \quad (5)$$

dove r rappresenta la distanza dall’origine (o dal centro della funzione) e ϵ è il parametro che controlla la larghezza della RBF. Al crescere di ϵ l’influenza di un punto di dati si estenderà su una distanza maggiore, risultando in una

⁶ Se usato in modalità transfer learning è necessario escludere l’ottimizzazione dell’interpolazione poiché i suoi valori ottimali sarebbero determinati sull’intero D e non su d .

superficie di interpolazione più regolare o “liscia”. Valori minori di ϵ porteranno a una superficie di interpolazione più irregolare, dove si potrebbero registrare cambiamenti repentini nei $P_{r_{int}}$ restituiti dalla RBF.

La performance della RBF viene valutata in `optim_grid_vae.py` attraverso l’ottimizzazione degli iperparametri ϕ , ϵ e dello smoothing, misurando la distanza euclidea come la differenza media tra tutti i P_o e i corrispondenti P_r con l’obiettivo di minimizzare questo valore al fine di ottenere un’interpolazione ottimale. Un valore più basso indicherà un allineamento migliore tra i P_o e i P_r , riflettendo una maggiore accuratezza.

4. APPLICAZIONI PRATICHE

L’ideazione di *WavePilot* è iniziata nel 2021 come patch Max, motivata dalla necessità riscontrata nella pratica artistica dell’autore di gestire ampi set di parametri su più DMmI contemporaneamente. I DMmI sono organizzati in catene e la transizione fluida da un preset all’altro è fondamentale. Nelle installazioni multimediali basate su feedback, cui la pratica dell’autore fa riferimento, l’esplorazione dello spazio parametrico dei DMmI va oltre il semplice concetto di interpolazione. È necessario gestire l’evoluzione delle azioni su lunghi archi temporali, in cui diverse configurazioni di parametri, con caratteristiche timbriche e visive contrastanti, si susseguono e si sovrappongono senza soluzione di continuità. Inizialmente, *WavePilot* è stato impiegato in *Wunderkammer* con una GUI bidimensionale con capacità limitate di ricostruzione dei parametri. Tuttavia, nell’ambito della creazione dell’ambiente generativo *Sandy Island* nel 2022⁷, il tool è stato estesamente utilizzato per definire gli stati delle catene di effetti audio e video attraverso sovrapposizioni di traiettorie deformabili da input audio esterni. In quel contesto sono stati definiti manualmente due d distinti composti rispettivamente da

⁷ www.youtube.com/watch?v=nzZBNiAqRdA&t=28s.

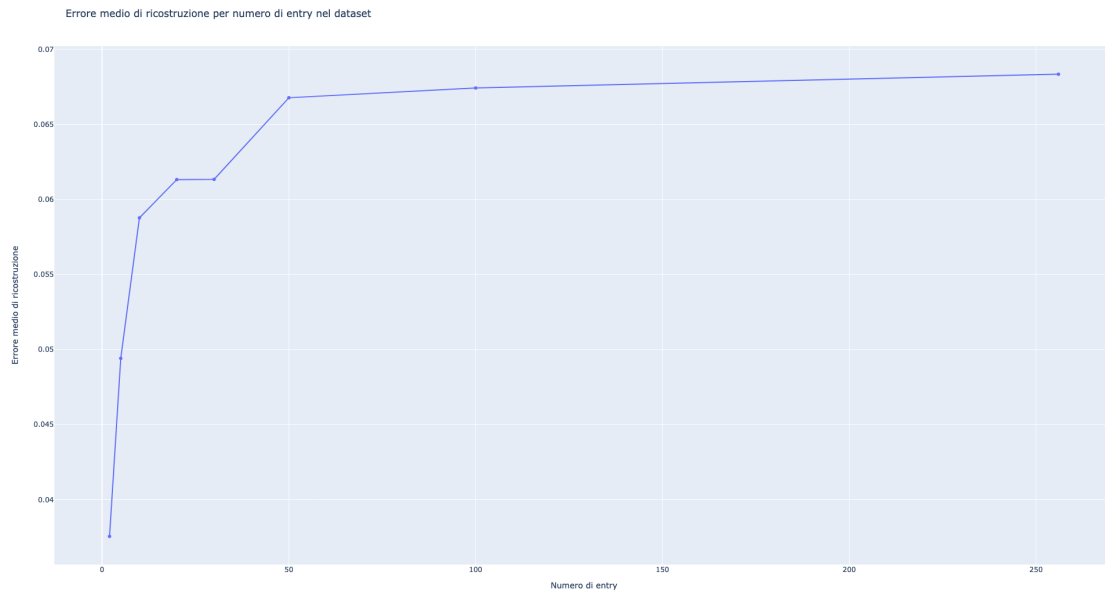


Figure 3. Il grafico rappresenta l’evoluzione dell’errore medio di ricostruzione in relazione al numero di entry in d . L’errore medio di ricostruzione è determinato calcolando la media delle differenze quadratiche tra P_o e P_r . Si osserva che l’errore medio di ricostruzione tende a crescere all’aumentare del numero di P_o in d . Tuttavia, l’incremento dell’errore si attenua per un numero di entry superiore a 50. Ciò suggerisce che l’aggiunta di nuovi P_o in d oltre questo valore porterà sì a un peggioramento dell’errore di ricostruzione, ma con un incremento progressivamente minore. L’osservazione si basa su un dataset contenente 1167 preset di 83 parametri ciascuno di un Ob-Xd, plug-in emulazione di un Oberheim OB-X.

82 P_o di 120 parametri ciascuno per l’audio e 26 P_o da 42 parametri per il video. Nonostante l’elevato numero di entry del d , che ha influenzato la capacità di ricostruzione dei P_o , un certo grado di approssimazione era una caratteristica deliberatamente ricercata per ragioni artistiche. I P_r venivano poi inviati a plugin appositamente sviluppati dall’autore con supporto OSC usando Live come host per l’audio e TouchDesigner per il video.

Successivamente, al fine di migliorare le capacità di ricostruzione dei P_o , è stata sviluppata la versione attuale di *WavePilot* come modulo Python. L’obiettivo è quello di integrarlo in contesti di produzione più tradizionali, consentendo anche ad utenti meno esperti di interagire in modo espressivo con DMmI commerciali di una certa complessità attraverso la meta-GUI.

5. CONCLUSIONE E SVILUPPI FUTURI

La comunicazione con i DMmI ai quali il tool viene applicato è sempre stata problematica, principalmente a causa del limitato supporto OSC offerto dai DMmI commerciali e delle limitate capacità di mapping dei parametri offerte dalle DAW quando si parla di generazione audio. Nella prossima fase di sviluppo, l’attenzione sarà rivolta alla gestione più efficiente del masking per i parametri categorici, così come per quelli che influenzano dinamicamente altri parametri nel tempo, come ad esempio gli LFO, e per quelli il cui impatto sul risultato sonoro non è immediatamente percepibile. Inoltre, si prevede di integrare la funziona-

lità di creazione delle traiettorie direttamente nel tool, eliminando la necessità di ricorrere ad una patch Max come avviene attualmente. Si prevede anche un ulteriore perfezionamento delle tecniche di riduzione, con l’obiettivo di decodificare, almeno in parte, il valore percettivo dello spazio virtuale in funzione del suo contenuto. Infine, si valuterà la possibilità di includere il tool, nella sua forma attuale o in una diversa, in scenari di extended reality.

6. REFERENCES

- [1] G. Hajdu, B. Carey, G. Lazarević, and E. Weymann, “From atmosphere to intervention: The circular dynamic of installations in hospital waiting areas,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, (Copenhagen, Denmark), pp. 364–369, 2017.
- [2] M. Battier and N. Schnell, “Introducing composed instruments, technical and musicological implications,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, (Dublin, Ireland), pp. 156–160, 2002.
- [3] M. Gurevich, “A repertoire-driven approach to design new interfaces for musical expression,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, (Copenhagen, Denmark), pp. 163–168, 2017.

- [4] A. DiScipio, *Pensare le tecnologie del suono e della musica*. Napoli, Italia: Editoriale Scientifica, 2013.
- [5] E. Tomás, *The interface-score: electronic musical interface design as embodiment of performance and composition*. PhD thesis, Kunstuniversität Linz, 2018.
- [6] H. Borgdorff, *The conflict of the faculties. Perspectives on artistic research and academia*. Leiden, Nederland: Leiden University Press, 2012.
- [7] C. Roberts, G. Wakefield, M. Wright, and J. Kuchera-Morin, “Designing musical instruments for the browser,” *Computer Music Journal*, vol. 39, no. 1, pp. 27–40, 2015.
- [8] M. Matthews and L. Rosler, “Graphical language for the scores of computer-generated sounds,” *Perspectives of New Music*, vol. 6, no. 2, pp. 92–118, 1968.
- [9] M. Farbood, H. Kaufman, and K. Jennings, “Composing with hyperscore: an intuitive interface for visualizing musical structure,” in *Proceedings of the International Computer Music Conference*, (Copenhagen, Denmark), pp. 111–117, 2007.
- [10] S. Pearse, L. Landy, D. Chapman, D. Holland, and M. Eniu, “Composing with sounds: designing an object oriented daw for the teaching of sound-based composition,” in *Proceedings of the 16th Sound and Music Computing Conference*, (Malaga, Spain), pp. 310–315, 2019.
- [11] M. Farbood, *Hyperscore: a new approach to interactive, computer generated music*. PhD thesis, Massachusetts Institute of Technology, 2001.
- [12] G. Kreković, “Insights in habits and attitudes regarding programming sound synthesizers: a quantitative study,” in *Proceedings of the 16th Sound and Music Computing Conference*, (Malaga, Spain), pp. 316–323, 2019.
- [13] B. Pardo, M. Cartwright, P. Seetharam, and K. Bongjun, “Learning to build natural audio production interfaces,” *Arts*, vol. 8, no. 3, pp. 110–130, 2019.
- [14] N. Collins, “Experiments with a new customisable interactive evolution framework,” *Organised Sounds*, vol. 7, no. 3, pp. 267–273, 2002.
- [15] P. Dahlstedt, “Creating and exploring huge parameter spaces: interactive evolution as a tool for sound generation,” in *Proceedings of the International Computer Music Conference*, (La Habana, Cuba), 2001.
- [16] R. Bencina, “The metasurface: applying natural neighbour interpolation to two-to-many mapping,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, (Vancouver, Canada), pp. 101–104, 2005.
- [17] M. Marier, “Designing mappings for musical interfaces using preset interpolation,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, (Ann Arbor, Michigan, USA), 2012.
- [18] S. Fasciani, “Tsam: A tool for analyzing, modeling, and mapping the timbre of sound synthesizers,” in *Proceedings of the 13th Sound and Music Computing Conference*, (Hamburg, Deutschland), pp. 129–136, 2016.
- [19] M. J. Yee-King, L. Fedden, and M. D’Inverno, “Automatic programming of vst sound synthesizers,” *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 2, pp. 150–159, 2018.
- [20] G. L. Vaillant, T. Dutoit, and S. Dekeyser, “Improving synthesizer programming from variational autoencoders latent space,” in *Proceedings of the 24th International Conference on Digital Audio Effects*, (Wien, Austria), pp. 276–283, 2021.
- [21] P. Esling, N. Masuda, A. Bardet, R. Despres, and A. Chemla-Romeu-Santos, “Flow synthesizer: universal audio synthesizer control with normalizing flows,” *Applied Sciences*, vol. 10, no. 1, pp. 302–317, 2020.
- [22] T. Wishart, *On Sonic Art*. Reading, United Kingdom: Harwood Academic Publishers, 1996.
- [23] S. N. S. Hasan, N. W. Jamil, and H. Ahmat, “A review of dimension reduction techniques for classification on high-dimensional data,” in *Proceedings of the 7th International Conference on Quantitative Sciences and its Applications*, (Syntok, Malaysia), 2023.
- [24] A. Bardos, I. Mollas, N. Bassiliades, and G. Tsoumakkas, “Local explanation of dimensionality reduction,” *arXiv:2204.14012*, 2022.
- [25] A. Freed, J. MacCallum, A. Schmeder, and D. Wessel, “Visualization and interaction strategies for hybridization interfaces,” in *Proceedings of the International Conference on New Interfaces for Musical Expression*, (Sidney, Australia), 2010.

Il potere dell'intermedialità nella creazione della memoria artistica: il caso di *enigma 33* di Lucia Romualdi

Angelica Speroni

Ecole Doctorale Esthétique, Sciences et Technologies des
Arts, Université Paris 8
angelica.speroni@etud.univ-paris8.fr

Alessandro Bile

Department of Fundamental and Applied Sciences for
Engineering, Sapienza Università di Roma
Department of Electronic Music, Conservatorio di Musica
Santa Cecilia, Roma
alessandro.bile@uniroma1.it

ABSTRACT

Nel panorama delle arti contemporanee, l'intermedialità è spesso elemento costitutivo delle opere stesse e l'orizzonte entro cui queste si muovono. Inoltre, l'intermedialità, intesa come metodologia artistica che coinvolge dialetticamente differenti linguaggi espressivi, non solo dissolve i confini tradizionali tra le discipline, ma agisce anche come un potente catalizzatore per la memoria. In questo contesto, l'opera di Lucia Romualdi si distingue per la sua capacità di tessere insieme elementi musicali, visivi e concettuali, dando vita ad opere che oltrepassano i confini convenzionali dell'arte creando una solida memoria negli spettatori.

Questo manoscritto mira a fornire una breve analisi del potere dell'intermedialità nella costruzione della memoria artistica, focalizzandosi sull'esempio eloquente di *enigma 33*, l'ultima opera dell'artista.

Questa si presenta come un'epica contemporanea, una favola incantata, scritta nella prigionia nei campi nazisti che si dipana attraverso i meandri della memoria storica e dell'immaginazione, un'installazione a cui concorrono le immagini proiettate, la fisarmonica spazializzata di Claudio Jacomucci e la voce narrante di Franco Mazzi.

Il fatto stesso che sia concepita come un'opera intermediale ne determina la strutturazione come memoria. Infatti, i processi mnemonici biologici costruiscono i ricordi mettendo in relazione distretti di memoria diversi e ove ne viene impedita la comunicazione si perde il ricordo creato.

1. INTRODUZIONE

Nelle ultime derive delle arti contemporanee, l'intermedialità sembra emergere come un concetto fondamentale che attraversa e sfida le tradizionali definizioni artistiche. Questo fenomeno si manifesta nel tessuto stesso delle opere d'arte creando un luogo intermedio, liminale tra differenti forme espressive, che si configura come un terreno fertile per l'emergere di nuove narrazioni e significati.

Copyright: © 2024 Speroni et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0 International](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

In questo manoscritto si tratta dell'intermedialità, nella originale accezione proposta dall'artista Dick Higgins, per valutarne la potenzialità come catalizzatore della costruzione di memoria. In particolare verrà preso in esame il caso di *enigma 33*, l'ultima opera dell'artista Lucia Romualdi, che si presenta come opera puramente intermediale in cui la memoria è un tema e una metodologia di lavoro e che, proprio per la sua strutturazione, favorisce a sua volta la creazione di una memoria nella mente dello spettatore.

2. LA NASCITA DELL'INTERMEDIALITÀ

Nello spirito dell'estetica inclusiva di fluxus, è Dick Higgins che conia il termine intermedia [1]. Con questo termine egli indica una strategia sperimentale di convergenza tra diversi linguaggi artistici, che avviene sia a livello tecnico di linguaggio sia a livello concettuale. Questa prima teorizzazione sottolinea la dimensione liminale delle pratiche artistiche, in opposizione a quella di mixed-media, più popolare negli anni Sessanta, denotando una attitudine di pensiero volta alla fluidità piuttosto che alla categorizzazione.

Higgins riconosce nella sua pratica artistica intermediale uno dei tratti fondamentali della società moderna, infatti nei decenni successivi, le pratiche artistiche hanno abolito i confini – dove non li hanno deliberatamente ignorati - tra le diverse categorie critiche, e così il concetto di intermedialità è diventato sempre più presente e il termine sempre più utilizzato e abusato, soprattutto dopo la rivoluzione digitale, costituendosi come un concetto fondamentale per comprendere l'estetica e i meccanismi comunicativi dell'epoca contemporanea.

I prodromi storici dell'intermedialità si hanno già all'inizio del Novecento, provocati dalla ribellione degli artisti alle metodologie delle arti istituzionalizzate, le "belle arti" teorizzate dall'estetica tardo settecentesca. Questa rivolta ha dato adito ad un campo di esperienze ampio e variegato.

Nella prima metà del secolo avviene quella che, riprendendo la concezione del suono in campo allargato di Rosalind Krauss, Giacomo Albert [2] definisce la rottura del piedistallo, ovvero della forma da concerto, ma in generale, anche per le arti visive, la sperimentazione di nuovi modi di diffondere l'arte e quindi trovare

nuove forme di spettatorialità. Protagonista di questa rivoluzione è ovviamente Marcel Duchamp, e il suo lavoro di scardinamento di ogni categoria critica. Per non citare il trito esempio del ready made, si pensi a *Bruit secret* (1916), forse la prima scultura sonora, che determina la nascita della sound art, o *Étant donnés* (concepita già nel 1946) che rende esperienziale e partecipativa l'arte visiva, o ancora a come già nel 1936 Duchamp e Man Ray avevano costruito la prima forma di installazione nella nota esposizione surrealista.

Sempre risalente alla prima metà del secolo è un rinnovato interesse per la sinestesia, si pensi all'opera del teatro espressionista *Die glückliche Hand*, in cui Schönberg cerca un equilibrio tra le arti, in particolare tra musica e luce, e procede nel tentativo di organizzare unitariamente musica e libretto (di cui egli stesso è autore) nonché ogni minimo dettaglio della messa in scena [3], o ancora il *Prometeo* per cui Skrjabin progettò uno strumento che associava ad ogni nota un fascio di luce colorata che avrebbe dovuto inondare la sala. Tutte queste sperimentazioni non danno adito alla creazione di un genere ma ad opere che hanno una propria forma mediale, e quindi non ad un fenomeno unitario ma ad una molteplicità di esperienze diverse che si muovono in altrettante direzioni, basate su problemi estetici diversi.

Dopo queste prime sperimentazioni, già dalla metà degli anni Cinquanta ma soprattutto a partire dal decennio successivo, si assiste al culmine di tutte queste esperienze intermediali; ad esempio con le installazioni sonore sperimentate da artisti di prima levatura come Atsuko Tanaka, Allan Kaprow, Nam June Paik, Stockhausen, Xenakis. Negli stessi anni nasce la teorizzazione della musica ambientale di Walter Olmo, tutta la serie di esperimenti partecipativi che culminano in fluxus e la videoarte, e le già ampie possibilità di media da utilizzare si arricchiscono ulteriormente con il digitale.

In Italia, in particolare, c'è una situazione artistica molto prolifica ma allo stesso tempo una assenza di istituzioni e mercati specifici e mancanza di grandi commissioni, per questo si rimane in una situazione di sperimentalismo diffratto, senza un vero e proprio centro. Forse anche per questo prima di Romualdi, a parte alcune sperimentazioni legate soprattutto al movimento fluxus e nonostante l'apertura alla sperimentazione e la curiosità per quello che in altri paesi già era quasi istituzionalizzato, nessun artista visivo ha posto l'intera sua ricerca in questa direzione intermediale, soprattutto tra musica e arti visive, mentre da parte dei compositori molti sono gli esempi di ricerca in tal direzione – si pensi anche solo a Silvano Bussotti e alla sua pittura musicale che era generatrice della musica e viceversa. Infatti, la ricerca di Romualdi viene riferita proprio a queste sperimentazioni, soprattutto della Nuova Musica post-weberniana [4].

3. L'INTERMEDIALITÀ COME INTER-MEDIA

3.1 La proposta di Higgins

Analizzando il termine *intermedia*, si dovrà riflettere sul significato delle sue due parti. La prima è *inter*, la preposizione latina che indica proprio il *tra*, ovvero la condizione e posizione intermedia nello spazio fra due (o più) oggetti, fra limiti di spazio o di tempo, che indica anche comunanza e collegamento, o addirittura reciprocità. Questa dunque non indica banalmente l'essere posto spazialmente fra due luoghi. La seconda parte della parola è il sostantivo *medium*, letteralmente 'mezzo', è il termine ripreso dal latino per indicare qualunque mezzo di comunicazione e di informazione, che sia veicolo di un messaggio.

Higgins è il primo a parlare di intermedia nelle arti visive e nel teatro (prendendo come esempio l'happening e la partecipazione del pubblico agli esperimenti teatrali). Egli cita anche il campo della musica che sconfinava nella filosofia o nella scultura, cita in particolare il suo maestro, John Cage, e in generale scrive che questa si sta facendo strada in tutte le arti proprio perché la nuova mentalità non è più categorizzante [1]. Ancora di più, oggi, si può dunque affermare l'attualità di questo concetto, considerando, proprio dal punto di vista socio-politico a cui fa riferimento Higgins, la tendenza della società odierna - si pensi al diritto alla fluidità sessuale, di genere, sociale oggetto dei più recenti dibattiti.

L'intermedialità non coinvolge solamente le tecniche artistiche ma accoglie l'irruzione di opere non classificabili nella sfera dell'arte, che si inseriscono tra i media consolidati. Queste opere dunque possono trovarsi all'incrocio di diversi mezzi artistici, o anche all'incrocio tra i mezzi artistici e i nostri mezzi mondani. Oltre al testo originale Higgins pubblicherà anche *Statement on intermedia* in cui introduce l'aggettivo 'intermediale' e, sottolineando la posta politica in gioco del suo tempo, vede nell'arte, e in particolare proprio nell'intermedia, il mezzo per gli artisti di parteciparvi. E questo perché l'intermedia è secondo lui l'unica possibilità di ricongiungersi direttamente alla vita, per scuotere le vecchie categorie e per produrre un'arte impegnata.

Il sostantivo 'intermedialità' compare invece in *Synesthesia and Intersenses: Intermedia*, un articolo doppio che riprende quello del 1966 con una importante aggiunta. Egli evidentemente sentiva di dover precisare ulteriormente un concetto che potesse essere utile al pubblico per comprendere e apprezzare la sua arte d'avanguardia. Interessante è che egli come artista abbia sentito la necessità di farsi comprendere, utilizzando un linguaggio che non è propriamente quello dell'arte, ma quello di parola. Diventando comunicatore del proprio lavoro e della propria poetica, dando precisazioni addirittura a più riprese. Egli distingue chiaramente l'intermedia dalla mixed media: quest'ultima categoria dispone i media uno accanto all'altro, mentre la prima enfa-

tizza la dialettica tra di loro. In questo articolo Higgins utilizza il singolare *intermedium*, individuando finalmente questo luogo indefinibile e illocalizzabile, sospeso tra l'arte e la vita, di cui porta l'happening come esempio, scrivendo: «Thus the happening developed as an intermedium, an uncharted land that lies between collage, music and the theater. It is not governed by rules; each work determines its own medium and form according to its needs. The concept itself is better understood by what it is not, rather than what it is» [5].

E questo è l'aspetto veramente interessante, che, come si vedrà, è stato travisato nella letteratura successiva: questa prima teorizzazione propone il termine *intermedium* come un sostantivo che designa questo non-luogo liminale come dimensione della pratica artistica stessa, senza che questa debba definirsi come multimediale [6], ma in opposizione a qualsiasi categorizzazione forzata; propone l'*entre-deux* come luogo dell'arte.

Un'altra importante precisazione che Higgins tiene a dare è che il termine *intermedia* non è prescrittivo, ma descrittivo. Infine, l'artista insiste sul carattere non storico dell'arte intermediale, che è in realtà una tendenza costante nell'arte: «Intermediality has always been a possibility» [5]. Secondo lui, l'*intermedia* non può diventare un movimento artistico databile: la nozione deve quindi essere usata e poi abbandonata, e non rispettata come un dogma. Essa attraversa la storia dell'arte e le molte arti, per quanto diverse esse siano. Può mantenere la sua vitalità solo rifiutando di diventare un nuovo accademismo. È questa la concezione di intermedialità che spinge Lucia Romualdi, senza che – e questo è forse il punto più interessante – ella abbia mai utilizzato il termine.

3.2 L'intermedia nella letteratura critica

Ovviamente, il successo dell'*intermedialità*, qui sempre intesa in questo senso generico, è determinato in particolare dal fatto che la società contemporanea stessa è caratterizzata da interattività e multitasking, oltre che dalla fluidità di cui già parlava Higgins. Il tentativo della letteratura critica di definire gli esperimenti intermediali in senso lato che si sono prodotti a partire dagli anni Sessanta ha portato, come si accennava, ad una proliferazione di etichette iper-specifiche (e spesso anche ripetitive). Indubbiamente il termine *intermedialità* si è rivelato uno dei più efficaci, ma non come termine specifico, bensì come termine generico, proprio per la sua versatilità e mancanza di univocità e quindi ancora non è stato sostituito da altri termini forse più largamente diffusi [7].

Ad oggi moltissime sono le categorizzazioni e le diverse sfumature del concetto di incontro tra più media. Basta pensare che, grazie ad una residua specificità contestuale o storico-geografica, ogni proposta terminologica ha trovato successo in un diverso contesto e si è affermata come genere preciso e connotato culturalmente: multimedia, unimedia [8], ipermedia [9]... fino alla transmedialità [7], ma nessuno rende conto della dimensione originaria di questo concetto.

3.3 L'importanza di riconnotare il termine oggi

È evidente invece che nel panorama contemporaneo l'intermedialità in senso higginsiano sia caratterizzante. Sarà dunque necessario proporre una teoria dell'intermedialità specifica, come dialettica di convergenza tra diversi linguaggi in cui non avviene fusione, come terreno conflittuale che si impone come soglia, inter-mezzo, e ne riscopre il valore anche etico. Infatti, proprio in questo suo senso si ritiene che l'intermedialità abbia un grande valore per le arti e la società contemporanee e sia quindi di grande attualità.

Inoltre, la consapevolezza che anche la comunicazione e documentazione sia parte dell'opera d'arte contemporanea porta a dover considerare anche queste come determinanti l'intermedialità di qualsiasi opera d'arte, anche che sia concepita e realizzata dall'artista con un medium puro, rendendo il discorso intermediale definitivamente e sempre attuale e, come voleva Higgins, trasversale.

Inoltre, se lo spazio artistico intermediale è un intermedium, una linea di confine che attraversa il reale, questo *in-between*, questo spazio liminale non definisce solo un'ibridazione tecnica tra codici, ma, per la sua dimensione di confluenza, confronto e cooperazione dell'eterogeneo, assume un valore etico: nelle tensioni che i codici disciplinari mantengono, si estende la tensione tra arte e vita, tra forma e comportamento, tra proiezione estetica e azione politica. In questo senso l'intermedialità si configura come una scelta non solo estetica ma anche etica dell'artista.

3.4 L'intermedialità come catalizzatore di memoria

È indiscussa la funzione mediale dell'arte, ma è proprio la pluralità dei sensi e dei linguaggi coinvolti a determinarne il potenziale mediatico.

Nel 1967 Marshall McLuhan apre il suo *Understanding media* dichiarando che la nostra società è «da tempo abituata a separare e dividere tutte le cose come mezzo di controllo» [10]; anche qui si può notare una tendenza a considerare e prediligere un "totale" rispetto a segmentazioni e frammentazioni – non è un caso che il testo sia praticamente coevo all'articolo di Higgins – e questo è diventato possibile secondo il mediologo grazie all'avanzamento di certe tecnologie e alla crescente meccanizzazione che porta ad estremi i comportamenti usuali della materia, ma anche a quelle sperimentazioni artistiche, come la scomposizione cubista, che mostrano le forme nella simultaneità permettendo una consapevolezza istantanea del tutto. In particolare, egli individua un esempio perfetto nella luce elettrica in quanto essa normalmente sfugge all'attenzione come mezzo di comunicazione proprio perché non sembra avere contenuto e quindi non viene percepita come medium. Stabilito però che il medium è il messaggio, la luce diventa pura informazione, un mezzo di comunicazione senza messaggio, il medium perfetto. Sarebbe qui interessante approfondire il discorso considerando le ultime derive degli studi di fisica ottica, in particolare di

quanto attraverso i fotoni si possano addirittura trasportare informazioni complesse come fanno i neuroni umani [11].

Alla luce di questa nuova consapevolezza, che attribuisce all'estetica una dimensione "scientifica", emancipandola al contempo dal soggettivismo romantico e dalla normatività modernista, l'intermedia indica pratiche di incrocio e ibridazione sperimentale non solo tra stimoli sensoriali eterogenei, ma anche tra codici sedimentati da tradizioni culturali. Anche McLuhan scrive un intero capitolo, intitolato *Energia ibrida. Les liaisons dangereuses*, in cui descrive l'incontro fra i media come capace di produrre nuova energia.

In particolare, è noto che la memoria viene catalizzata dalla cooperazione di elementi sensoriali ed emozionali [12]. Si dovrà concludere allora che proprio la multi-medialità artistica, la compartecipazione di fattori della vita, l'immediatezza e il senso di fluidità dinamica con cui si presenta la complessità dell'opera intermediale, tutti questi fattori sono ciò che permette la costruzione di una memoria che sia persistente nella mente umana.

La natura intermediale di un'opera, come il qui preso ad esempio *enigma 33*, favorisce infatti una forte stimolazione sinaptica che si traduce in una sua memorizzazione profonda all'interno delle mappe neurali. L'utilizzo contemporaneo di media diversi comporta una maggiore eccitazione di distretti neurali che "lavorano" insieme, in accordo ai postulati Hebbiani [12]. I collegamenti inter-neuronali si costruiscono così più velocemente ed in modo più intenso generando ricordi e memorie più difficilmente debili dal tempo. Quando il messaggio che viene elaborato dal cervello è il prodotto di una stimolazione sensoriale multipla, infatti, vengono attivate contemporaneamente più connessioni sinaptiche [13]. Ciò rende la conoscenza appresa più complessa e meglio radicata nel circuito neurale [11, 13-14]. Tuttavia, gli stimoli multisensoriali devono essere latori di un unico messaggio. Quando infatti agiscono stimoli che lavorano trasportando messaggi indipendenti tra loro, si ottiene il processo inverso alla memorizzazione, la distrazione [15].

In particolare, nel caso della fruizione dell'opera intermediale, il ricordo nella mente dello spettatore si crea attraverso l'esperienza plurisensoriale che compie e questo stesso fatto determina la formazione di un ricordo molto solido e complesso, e dunque duraturo. Di conseguenza, proprio per come è strutturata la funzionalità dei neuroni umani, perché il ricordo sopravviva si dovrà preservarne la qualità intermediale.

4. IL CASO DI ENIGMA 33

4.1 Lucia Romualdi

Attiva dalla fine degli anni Settanta in ambiti pertinenti alla ricerca concettuale, dopo un'interessante e coerente sperimentazione pittorica, è dal 1989 che Lucia Romualdi opera in stretta simbiosi con i linguaggi musica-

li. Nel 1990 è iniziata la collaborazione con Franco Donatoni che in dialogo con il suo lavoro ha composto i due noti brani *cloches II* per due pianoforti (1990) e *feria IV* due brani per accordéon (1997). Ha collaborato con altri compositori contemporanei: Fausto Sebastiani (*codice c* per chitarra elettrica e nastro magnetico, nel 1995 – *cromatico* per fisarmonica, nel 1996 – *melodie I*, diagramma sonoro per nastro magnetico, nel 1988 – *melodie II*, diagramma sonoro per nastro magnetico e computer, nel 1999 – *op. KM33* per interventi digitali e suoni meccanici, nel 2007); Ivan Fedele (*canone infinito* per nastro magnetico, nel 2002 – *due notturni* con figura per pianoforte e elettronica, nel 2010); Francesco De Gregori (*cardiologia* per voce e cinematica, nuova trascrizione nel 2014); Claudio Jacomucci (*incantesimi* per fisarmonica spazializzata, nel 2023).

Ad una accurata analisi della sua opera, fin dagli esordi, si riconosce subito che l'intermedialità è per lei una metodologia di lavoro. Le sue installazioni non si situano solo sulla soglia tra musica e arti visive, ma precisamente in quel luogo intermedio che sorge nell'incontro tra il linguaggio delle arti visive, il linguaggio di parola e altri media artistici: procedimenti tipici del montaggio cinematografico, del teatro, e ovviamente della musica, la cui composizione è affidata a grandi compositori contemporanei.

L'intesa con i compositori è sempre stata per lei molto importante, solo a loro ha concesso di operare all'interno di un lavoro che, per come lei stessa lo concepisce, non è solo il frutto del suo genio individuale, ma anche il risultato di una collaborazione sinergica. Utilizzando cifre, diagrammi numerici e tabulati, l'artista crea partiture di luce che si trasformano in installazioni ambientali di grande suggestione. Questo approccio conferisce infatti alle opere una profondità e una complessità che va oltre il racconto narrativo, plasmando una memoria strutturata in modo unico e coinvolgente [16].

4.2 Enigma 33

Caso di studio è stato in particolare l'opera *site-specific enigma 33*, l'ultimo lavoro di Romualdi, presentata nella Sala Assoli di Napoli nel maggio 2023. Si tratta di una favola inedita che narra infatti le avventure di viaggio di un principe favolistico, figlio del Re Gontrano, che le Fate Sapienza Bellona e Poesia per un equivoco triplicano in tre fratelli gemelli, ognuno dei quali eccellerà in un Ideale: la Cultura, il Coraggio, l'Arte. L'incantesimo porterà i tre Principi ad affrontare molte prove: l'incontro con l'Ombra, la lotta con l'Orco, il segreto dell'Acqua, fino a quando Merlino, stanco dei pasticci creati dalle Fate, deciderà di sciogliere la magia. Scritta durante la prigionia nei campi nazisti di Chölm e Wietendorf fra il 1943 e il 1945, la favola traspone su un piano poetico e fantastico memorie di guerra e affronta in forma di metafore temi quali il potere, le leggi razziali, la prigionia, il Nazismo, l'armistizio, la Liberazione. L'artista lo definisce come un «lavoro politico nomade e musicale sulla lontananza,

una partitura di luce numerale e cinematografica in variazione, un'opera sul concetto di intervallo che si apre alla memoria collettiva, raccontata a bambini e adulti dai pesci/marionette e dalla barca "vagabondo" che se ne va per mare, accompagnata dalla fisarmonica spazializzata di Claudio Jacomucci e dalla voce narrante di Franco Mazzi che segnano il tempo e il ritmo del mare».



Figura 1. Lucia Romualdi, *enigma 33*, per voce e marionette; musica di Claudio Jacomucci (*Incantesimi 2023*, per fisarmonica spazializzata 2023), Claudio Jacomucci, fisarmonica spazializzata, Franco Mazzi, voce narrante - Sala Assoli, Napoli 2023. Foto: Claudio Fabian de Rosa. Courtesy dell'artista.

Il lavoro giace sulla soglia tra teatro e arti visive, musica, cinema, vivo per una settimana e accessibile per sempre solo grazie allo streaming in diretta di cui rimane traccia, che ne crea una memoria anche tecnologica. Il titolo del lavoro è una citazione di "enigma" un dispositivo elettromeccanico per cifrare e decifrare messaggi che mascherato da macchina da scrivere fu ampiamente impiegato dalle forze armate tedesche durante la Seconda Guerra Mondiale. Modificata più volte Enigma riesce comunque a essere decifrata e decodificata e la decrittazione dei messaggi cifrati fornisce per quasi tutta durata della guerra importanti informazioni alle forze alleate. 33 è un numero che emerge dal testo, è la soluzione dell'enigma che il Nano ad un certo punto propone ai tre principi per accedere al Regno di Fantasia, quel non luogo-ideale intermediale che è poi il luogo dell'opera, il luogo creato da Romualdi per accogliere lo spettatore.

Il pubblico che entra in Sala Assoli, seminterrata rispetto al piano stradale, viene fatto sedere di lato, dimesso lungo le pareti; altri rimangono su, nel foyer, seduti lungo le ringhiere, come se fossero alla finestra, affacciati su una piazza di Napoli. Franco Mazzi, Claudio Jacomucci e un mezza coda nero sono immobili, in scena. Tutto è nero, o bianco, i colori di Romualdi, sulle pareti solo numeri proiettati, immobili. Sono maree e grate. Le maschere fanno sedere il pubblico alla luce delle torce elettriche, un pubblico quasi intorpidito da un lavoro che è già lì, esposto, senza il filtro di nessun si-

pario, e insieme immobile, come in attesa. In attesa di partire per mare. Ed è con il suono del mare che si apre l'opera, e l'invito della voce narrante, calda e calma: "Ascoltate". La barca "vagabondo" inizia il suo viaggio sul monitor scorrendo nell'intervallo tra le linee orizzontali proiettate sul fondo in un equilibrio perfetto.

L'installazione è stata prodotta da Studio Trisorio e Casa del Contemporaneo, con il sostegno di Seda, e sotto il patronato della Fondazione Donnaregina per le arti contemporanee. La musica è stata composta da Claudio Jacomucci, con il supporto del Fonds Podiumkunsten, Performing Arts Fund NL, si intitola *Incantesimi 2023* ed è ispirato al romanticissimo *Time on my hands*, una canzone americana degli anni '30 proposta come spunto da Romualdi, qui irriconoscibile, scomposta, lontana come tutti i riferimenti dell'artista. D'altronde, tutto *enigma 33* è diventato una favola sospesa, non più per bambini, ma per l'uomo contemporaneo, divisa in un prologo, tre atti, quattro variazioni, e varie altre parti non esattamente definibili che affrontano temi complessi con la leggerezza e la sintesi narrativa propria dell'arte. Questo lavoro, estremamente complesso quanto esteticamente e concettualmente minimale può essere considerato la *summa* di tutta la ricerca di Romualdi ed è perciò il caso di studio più consistente anche per lo specifico nesso tra l'intermedialità e la memoria. L'artista lavora infatti su alcuni appunti organizzati per un'opera rimasta inedita nel 2017, di cui esiste solo il progetto, intitolata *nell'intervallo op.K42°06*, una "partitura di luce" – così chiama i suoi lavori – in cinque movimenti in cui affronta direttamente il concetto di intervallo, quello che qui si definisce *intermedium*.

Romualdi infatti non si limita a collaborare con compositori e musicisti, ma utilizza lei stessa graficamente, visivamente e concettualmente determinati elementi propri del linguaggio musicale – come il pentagramma, il ritmo, il tempo. Alla luce di questa riflessione possiamo riconoscerla come una compositrice che scrive non opere oggettuali, ma vere e proprie partiture che possono essere montate e smontate, riprese e portate nuovamente in esecuzione.

Questa terminologia, anch'essa proveniente dall'ambito musicale, viene adottata da Romualdi nel 1993 e mantenuta per tutte le installazioni successive. La partitura come è noto è la notazione grafica complessiva delle parti di un brano musicale, l'artista sceglie questo termine per indicare quella che a tutti gli effetti è la raccolta complessiva delle parti di un'opera le cui unità fondamentali però, in luogo delle note, sono le diapositive che vengono proiettate. Oltre all'impiego di questo tipo di notazione nel campo delle arti visive, un'altra particolarità è che in questo caso non si tratta solo di una notazione grafica ma delle diapositive che fisicamente sono conservate in contenitori trasparenti (i "fogli") poi ordinati in scatole nere divise per movimenti, per proiettori, per opera. In realtà, come chiarisce l'artista in un talk tenuto al Museo MAXXI di Roma nel 2016, non si tratta di normali diapositive: «è un lavoro

in camera oscura cioè un lavoro fotografico, sono dei frame che io taglio manualmente e metto nel telaietto quattro per quattro, non sono slide, è un lavoro cinematografico, una tecnica che rimanda a quella cinematografica, immagini in sequenza, un procedere temporale, una serie di operazioni che si aprono, come avviene nel cinema, a linguaggi diversi, musica, movimento, suono, una sequenza di segni che formano poi un tutt'uno» [17]. Questa importante precisazione sulla metodologia con cui prepara la “diapositiva” che poi viene proiettata dal diaproiettore, ovvero il *frame* che, appunto, rimanda all'immaginario e alla tecnica del cinema, chiarisce che nel processo stesso del suo lavoro, anche nella più banale operazione tecnica, è considerata l'appartenenza di quel frame ad una sequenza cinematica, che quindi sottende e non può disgiungersi dal procedere temporale e dal ritmo, entrambi fattori essenziali della sua ricerca. Proprio questa considerazione che l'artista ha del suo lavoro pone quest'ultimo - in quanto opera - ma soprattutto la sua stessa metodologia compositiva a tutti gli effetti nel regno dell'intermedialità. E così la stessa presenza del suono - quel clic meccanico del diaproiettore e il rumore delle ventole - che è nel lavoro di Romualdi prima ancora di qualsiasi musica, è già invernata nella composizione dei frame.

Inoltre, in tutte le sue opere a partire dal 1990, c'è la presenza della musica, realmente udibile e concettualmente protagonista, in *enigma 33* oltre alla fisarmonica c'è anche un altro strumento, potremmo dire *suonato* da Romualdi: il pianoforte a mezza coda che troneggiava silenzioso al centro della scena. Uno strumento muto che si ascolta con gli occhi e nella memoria di chi è stato testimone dei lavori passati dell'artista.

Tutti questi elementi determinano l'intermedialità dell'opera di Lucia Romualdi, di cui *enigma 33* si ritiene essere miglior esempio, ne determinano le modalità di percezione e la strutturazione come ricordo.



Figura 2. Lucia Romualdi, *enigma 33*, per voce e marionette; musica di Claudio Jacomucci (Incantesimi 2023, per fisarmonica spazializzata 2023), Claudio Jacomucci, fisarmonica spazializzata, Franco Mazzi, voce narrante - Sala Assoli, Napoli 2023. Foto: Claudio Fabian de Rosa. Courtesy dell'artista.

5. CONCLUSIONI

In sintesi, l'intermedialità nell'arte dimostra di favorire la comprensione di ciò che viene rappresentato e di conseguenza ne agevola la memorizzazione. Questi processi sono dovuti alla natura dell'apprendimento del cervello umano: esso viene maggiormente attivato in presenza di una moltitudine di stimoli. In particolare, sia in ambito scientifico che in ambito artistico, è stato mostrato che quando i diversi media concorrono nel trasmettere un messaggio preciso e conciso, si attivano regioni neurali diverse che iniziano a lavorare insieme, collegandosi tra loro attraverso i neuroni che le costituiscono. In questo contesto si inserisce il lavoro di Lucia Romualdi che fa dell'intermedialità il fondamento della sua arte. La luce dialoga in stretta connessione con la musica e con l'ambiente che la ospita. Si genera così una sensazione sinestetica altamente neuro-generatrice.

In *enigma 33*, le proiezioni di luce, la musica della fisarmonica e quella meccanica prodotta dai proiettori, la voce umana e gli elementi fisici coinvolti nel lavoro - tra cui lo stesso spettatore -, costituiscono stimoli diversi ma allo stesso tempo comunicano un messaggio specifico che stimola, direziona e governa la formazione della memoria.

6. RIFERIMENTI

- [1] D. Higgins: “Intermedia”, *Something Else Newsletter*, Vol. 1, n. 1, 1966.
- [2] G. Albert: “La dimensione sonora nelle opere di Nam June Paik e Bill”, ProQuest, Ann Arbor 2019.
- [3] P. Gelli (a cura di): “Dizionario dell'Opera 2008”, Baldini Castoldi Dalai editore, Firenze 2008.
- [4] E. Torelli Landini: “Interferenze del sessantotto”, “Il Trovaroma”, “La Repubblica”, Roma 6 Maggio 1993.
- [5] D. Higgins: “Horizons: The Poetics and Theory of the Intermedia”, Chicago, Southern Illinois University Press, 1984.
- [6] F. Menna: “La soglia. L'opera d'arte tra riduzione e costruzione”, Edizioni concordia sette Pordenone, Pordenone 1985.
- [7] I. Rajewsky: “Percorsi transmediali. Appunti sul potenziale euristico della transmedialità nel campo delle letterature comparate”, *Schermi. Rappresentazioni, immagini, transmedialità*, Eds. F. Agamennoni, M. Rima, S. Tani, Between, VIII.16, 2018.
- [8] P. Lévy: “Cyberculture. Rapport au Conseil de l'Europe dans le cadre du projet «Nouvelles technologie: coopération culturelle et communication»”, Odile Jacob, Paris 1997, trad. it. “Cybercultura. Gli usi sociali delle nuove tecnologie”, Feltrinelli, Milano 1999, 2013.

- [9] G. Landow: “Hypertext: The Convergence of Contemporary Literary Theory and Technology”, Johns Hopkins University Press, Baltimore 1992.
- [10] M. McLuhan: “Understanding Media”, McGraw-Hill, New York 1964.
- [11] A. Bile: “Solitonic Neural Networks: An Innovative Photonic Neural Network Based on Solitonic Interconnections”, In (Ed.), *AI in Materials*, Springer, 2023.
- [12] C. M. Tyng et al.: “The Influences of Emotion on Learning and Memory”, *Front. Psychol.*, 8, 2017.
- [13] T. J. Sejnowski, G. Tesauro: "The Hebb Rule for Synaptic Plasticity: Algorithms and Implementations". Academic Press, 1989.
- [14] Pusch, R., Packheiser, J., Azizi, A.H. et al. Working memory performance is tied to stimulus complexity. *Commun Biol* 6, 2023.
- [15] Harris, A.M., Eayrs, J.O. & Lavie, N. Establishing gaze markers of perceptual load during multi-target visual search. *Cogn. Research* 8, 56, 2023.
- [16] A. Speroni: “Raggiungere l’orizzonte. Significato, valore e funzione della ricerca artistica”, *L’Erma di Bretschneider, Roma-Bristol (CT)*, pp. 158, 2024.
- [17] <https://www.studiotrisorio.com/press-1/achille-bonito-oliva-con-lucia-romualdi-e-francesco-de-gregori-la-lontananza-lestraneita-il-nomadismo-la-misurazione>

A Pierre. Dell'azzurro silenzio, inquietum (1985) Analisi Formale

Claudia Rinaldi

CNIT

Consorzio Nazionale Inter-Universitario per le Telecomunicazioni

claudia.rinaldi@univaq.it

Marco Pennese

Istituto Comprensivo Giovanni Pascoli, Rieti

marcopennese1981@gmail.com

ABSTRACT

Questo studio offre un'analisi formale del brano "A Pierre. Dell'azzurro silenzio, inquietum" composto da Luigi Nono nel 1985, un'opera emblematica della tradizione elettroacustica. Utilizzando un approccio interdisciplinare che integra l'ascolto attento con l'analisi della partitura, esploriamo le stratificazioni sonore e le strutture formali che caratterizzano questa composizione. Inoltre, contestualizziamo il lavoro all'interno della storia e dell'evoluzione estetica del compositore, concentrandoci sulle sue ultime opere per comprendere il contesto creativo e le influenze che hanno plasmato questo capolavoro.

1. INTRODUZIONE

L'opera di Luigi Nono di cui si presenta l'analisi in questo lavoro appartiene all'ultimo periodo di attività del compositore (1985), vale a dire dopo l'incredibile fatica del Prometeo, quando ormai il concetto di utilizzo del live electronics e dello spazio come elementi compositivi era ben sviluppato e radicato nell'estetica dell'autore.

La prima opera elettronica del compositore risale al 1960, Omaggio a Emilio Vedova, ed è l'unica in cui Nono utilizza esclusivamente suoni di sintesi. Affermerà infatti che comporre con materiale concreto insieme al suono elettronico puro consente di «superare l'impronta naturalistica del primo e quella freddamente meccanica del secondo», [1], e di creare una maggiore flessibilità nel gioco fusione/contrasto con il suono strumentale o corale dal vivo.

Anche a livello stilistico, sull'utilizzo dell'elettronica da parte di Luigi Nono si possono riscontrare due periodi ben distinti: le composizioni nate all'interno dello Studio di Fonologia Musicale della RAI di Milano, e quelle che hanno visto la luce in seguito all'inizio della collaborazione tra il compositore e lo Studio di Friburgo (a partire dal 1980), con conseguente abbandono delle attrezzature ormai desuete dello Studio di Milano. Dal 1960 al 1980 Luigi Nono scrive 27 composizioni di cui 17 sono per nastro magnetico, con o senza strumenti o coro. Nel 1979 il compositore operò per la prima volta con l'elettronica dal vivo nel brano *Con Luigi Dallapiccola*, per percussioni e live electronics (4 pickup, 3 modulatori ad anello, 3 generatori

di frequenza). Questo brano, assieme al successivo quartetto per archi *Fragmente-Stille, an Diotima* (1979-1980), aprirono un nuovo periodo compositivo in seguito alla crisi creativa, e l'inizio della collaborazione con Friburgo fu una spinta fondamentale. Il primo brano che vide la luce in seguito a questo importante inizio fu *Das atemde Klarsein* (1980-1983), per flauto basso, piccolo coro, nastro magnetico e live electronics. La ricerca espressiva del compositore, a partire dagli anni '80 diventa più radicale poiché la sperimentazione sul timbro diviene una costante di tutte le sue opere. Il concetto di "comporre con i suoni" si trasforma nella volontà di "comporre il suono". Anche il concetto di tempo ne risente, il suono infatti non è più inscritto nel tempo, ma contiene il tempo e diventa "parametro portatore di forma" [2].

Altro elemento importante nell'evoluzione stilistica del compositore sta nella presenza di lunghe fasi di sperimentazione in stretta collaborazione con gli esecutori. Questo permetteva agli esecutori da un lato di proporre soluzioni di prassi esecutiva originali, dall'altro di acquisire familiarità con il linguaggio musicale del compositore. Il carattere aperto della scrittura di Nono garantiva agli interpreti una certa libertà di esecuzione, seppure limitata a livello della microstruttura del brano e comunque coerentemente con le intenzioni del compositore. Sulla base di questo la stesura di una partitura sintetica risulta necessaria e sufficiente, in quel particolare frangente.

2. A PIERRE. DELL'AZZURRO SILENZIO, INQUIETUM (1985)

Luigi Nono scrisse questo brano "a più cori, per flauto contrabbasso, clarinetto contrabbasso e live electronics" in occasione dei 60 anni di Pierre Boulez, al quale lo dedicò. La prima esecuzione ebbe luogo a Baden-Baden il 31 Marzo 1985 (il compleanno di Boulez era stato il 26 marzo) e vide la presenza degli esecutori con cui il compositore aveva lavorato per la stesura del brano stesso, cioè *Ciro Scarponi* (clarinetto) e *Roberto Fabbricani* (flauto). Il live electronics venne realizzato dall'Experimental Studio di Friburgo, presso il quale come già detto Nono compose tutte le sue opere con elettronica a partire dal 1980.

Come nel caso del brano *Post-Prae-Ludium No.1 Per Donau* (1987), anche qui il contributo degli strumenti e del live electronics è difficilmente distinguibile, come del resto era intenzione del compositore, infatti lo scopo del brano sembra proprio quello di voler fondere i suoni dal vivo e quelli elaborati elettronicamente. La fascia sonora

risultante, che procede ininterrotta fino alla fine del brano, è protagonista, l'attenzione dell'ascoltatore deve essere rivolta alle sue variazioni interne che comprendono anche quelle introdotte dallo spazio in cui l'opera è eseguita.

Sono appunto i termini "mobilità" e "spazialità" che giustificano l'organico indicato da Luigi Nono: "a più cori", riprendendo la terminologia veneziana del secolo dei Gabrieli, da lui usata in molte altre occasioni (ad esempio chiamò cori i sette gruppi strumentali di 2° *No hay caminos, hay que caminar...* Andrej Tarkowskij), [8].

Inoltre, la volontà di rendere indistinguibili le sorgenti sonore è evidente anche dall'indicazione data ai due esecutori, di non staccare mai le labbra dall'imboccatura, anche durante le lunghe pause segnate in partitura e dunque anche l'aspetto visivo concorre a rinforzare il paradigma del labirinto acustico in cui il lavoro sembra evolversi alle soglie del silenzio. Chiaramente la motivazione alla base di tale scelta, che Nono chiama spesso "suono mobile", non sta nella volontà di confondere l'ascoltatore in sé, ma nel desiderio di rendere l'ascoltatore partecipe e attento nella ricerca dell'oggetto estetico, e.g. l'ascoltatore deve trovare la sua strada per districarsi dal labirinto acustico che è stato costruito.

L'opera è stata pubblicata in due edizioni dalla casa editrice Ricordi, la prima nel 1995 e la seconda nel 1996, che è quella utilizzata in questa analisi. In entrambe le versioni le parti di flauto contrabbasso e clarinetto contrabbasso sono riportate in due pentagrammi distinti tra i quali sono riportate una serie di linee indicanti i livelli dinamici (*fff*, *mp*, *p*, silenzio) che il regista del suono deve seguire nel controllo delle due coppie di altoparlanti coinvolte nella diffusione del live electronics. In realtà il compositore non ha mai scritto l'intera partitura che è stata pubblicata, ma sembra che essa derivi da 6 fonti diverse: (i) il manoscritto autografo; (ii) il manoscritto della partitura preparato da Roland Brieitfenfel, che lavorò con Nono allo Studio di Friburgo; (iii) dati relativi al live electronics conservati presso lo studio di Friburgo; (iv) i racconti scritti dei collaboratori di Nono; (v) le parti autografe per flauto contrabbasso e clarinetto contrabbasso con le annotazioni di Fabbriani e Scarponi; (vi) le registrazioni del brano. Il processo creativo che ha dato forma all'opera, ha avuto inizio nel Settembre del 1984, durante le prove per la prima esecuzione del Prometeo (Venezia, 25-29 Settembre 1984), [5]. Negli archivi della Fondazione Nono sono conservate ben 104 pagine di appunti relative a quest'opera, delle quali ricostruire l'ordine cronologico appare lavoro abbastanza arduo dato che la maggior parte non sono datate. Ad ogni modo, sembra che la prima traccia di *A Pierre* sia segnata nel pentagramma mostrato in figura 1, in cui si accenna alla melodia del clarinetto contrabbasso. Le prime sei note sulla pagina di appunti corrispondono a quelle indicate nelle prime 19 battute della partitura pubblicata di *A Pierre*.

2.1 LE NOTE DI SALA

Il testo di presentazione scritto da Nono è riportato di seguito:

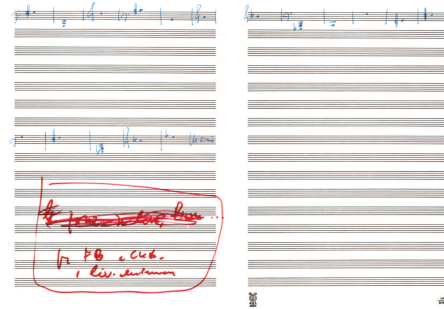


Figure 1. Prima idea del brano.

Più cori continuamente cangianti
per formanti di voci - timbri -
spazi interdinamizzati
e alcune possibilità di
trasformazione
del live electronics

a PIERRE.
"LE MUR MURANT PARIS REND PARIS MURMURANT"
CLAUDE-NICOLAS LEDOUX.
GRADINI SCALE RIPIDISSIME STRETTISSIME A ANGOLI
ACUTI CIRCOLARI PER ARRIVARE (1952) ALLO
SPAZIO DI PIERRE SOPRA I TETTI NEL CIELO AZZURRO (PURO CA-
SO? MI RICORDO BENE) DI PARIGI.
(LES DEUX STRUCTURES POUR DEUX PIANOS)
LABIRINTI TECNOLOGICI SILENZIOSI
("IRCAMNIANI"?)
PER ARRIVARE-SALIRE ALLA SPAZIANTE LUCE DI PIERRE SOPRA
LA VISIONARIETÀ REALE
UTOPICA
DI PARIGI (1984?), ALTRA "MEMORIA" DI
VISIONARIA LUMINOSITÀ (NON PROPRIAMENTE AZZURRA)
DI FOTOGRAFMI DAL FILM DI ELEM KLIMOV ADDIO A MATJO-
RA.
NON SOLO MEMORIE, NON SOLO ECHI LONTANI,
"NON DIRE DELL'IERI"(W. BENJAMIN)
OGGI IL CONTINUO INNOVANTE POSSIBILE.
SAPER ASCOLTARE ANCHE IL SILENZIO.
NON IN UNA POSSIBILITÀ DI ASCOLTO.
ASCOLTARE NEL SILENZIO GLI ALTRI L'ALTRO.
GLI INFINITI
RESPIRI-SENTIMENTI-PENSIERI-TRAGEDIE
"MURMURANTES" DI PIERRE
DEL MANIFESTO DEI 121 DI J.P. SARTRE
DEI VAGABONDAGGI INQUIETANTI BRAHMSIANI
DI BADEN BADEN
DELLA VISIONARIETÀ UTOPICA PENETRANTE
LUMINOSISSIMA
NEL
FUTURO
NELL'OGGI
PROFONDA DI MEMORIE DI SCOPERTE DI
ILLUMINAZIONI
DELL'IERI DEGLI OCCHI-SPAZI di "AFFINITÀ SENTIMENTALE" CON
A. ARTAUD, VIVACIZZATI DAL SOLE, ANCHE COME,
SECONDO SAPIENZE ASTRO-FISICHE, NELLA CATTEDRALE DI CHAR-
TRES
Luigi Nono, Berlino 10 - 5 - 87

In prima analisi lo scritto di Nono riflette la frammentarietà di molti dei brani dell'ultimo decennio, come se fosse la traccia, lo schema, della sua composizione. Se poi si vuole sottolineare il fatto che il compositore ritorni sulla necessità dell'ascolto, si potrebbe collegare lo scritto e la dedica del brano, alla volontà del compositore di "ascoltare" i ricordi dell'epoca di Darmstadt, mantenendone le distanze, ma risolvendo il conflitto. Boulez era infatti uno dei compositori da cui Nono volle prendere le distanze ne 1959. Boulez per Nono rimane l'altro da ascoltare in silenzio, senza più spaccature, senza più chiusure, ragion per cui ne cita anche un'opera, di cui egli stesso presentò un'analisi nel 1957, *Les deux structures pour deux pianos*. In *A Pierre*, Nono cerca un silenzio "azzurro", colore che per Mallarmé, poeta caro a Boulez, è simbolo di purezza, [9].

In questa chiave si può forse rintracciare una delle peculiarità di quest'opera, in cui i due strumenti dal vivo, flauto e clarinetto, usano prevalentemente i toni del piano. La prevalenza dei toni del piano non impedisce a Nono di realizzare un lentissimo "crescendo" fino al mezzoforte per poi tornare verso il pianissimo alla fine della composizione.

Nel testo si fa più volte riferimento alla storia di Parigi e della Francia, come nel seguente passaggio: DEL MANIFESTO DEI 121 DI J.P. SARTRE. Qui Nono si riferisce ad una dichiarazione che nel 1960 Sartre, insieme ad un gruppo di intellettuali francesi, tra cui Boulez, firmò contro la guerra di Algeria e sembra così voler ricucire un'altra frattura che, negli anni '60 e '70, divise i due compositori: ci riferiamo alla diversa posizione relativa al rapporto tra musica e politica. Soprattutto in quel periodo infatti, per Nono l'arte era necessariamente funzionale all'impegno; per Boulez, al contrario, essa non aveva alcun potere di agire sulla società. Inoltre, a completamento del titolo di *A Pierre* le parole *A più cori*, indicano come i due strumenti, spazializzati nella sala attraverso il Live electronics, rappresentino quella policoralità tipica della musica veneziana tra '500 e '600. In tal modo flauto e clarinetto acquistano un carattere collettivo che Nono attribuiva al coro fin dagli anni '50. Ne è un esempio *Il canto sospeso*, in cui le voci dei condannati a morte della Resistenza europea divengono voci della memoria ed assumono un valore universale. Nell'ultima parte della produzione noniana esse divengono più rarefatte, lontane, come si vede proprio in *A Pierre*, dove la dinamica si mantiene appunto prevalentemente sui toni del piano. Altri mezzi per giungere a questa smaterializzazione del suono sono le modalità di emissione: soffio, aria, fischio, suoni "eolien" (parziali so-pracute isolate e rinforzate), armonici, "suoni ombra" (la cui fondamentale indeterminata appare e scompare con discontinuità). Si tratta di tecniche messe a punto soprattutto con Roberto Fabbriciani. Lo stesso Nono fa riferimento a questo particolare uso del termine "cori" in *A Pierre* nel primo dei due testi sull'opera:

Più cori continuamente cangianti
per formanti di voci - timbri -

Poche sono infatti le note "piene" presenti nella composizione, che in questo modo non trova mai certezze definitive. Essa esprime dunque una perpetua condizione di ricerca ed è indicata dal termine "inquietum", contenuto nel titolo e suggerito da quel

CONTINUO INNOVANTE POSSIBILE.

di cui parla il secondo testo. La trasformazione del suono è ottenuta anche col Live electronics che viene espressamente citato nelle note di sala, come elemento compositivo:

alcune possibilità di
trasformazione
del live electronics.

Esso potenzia la dimensione corale dell'opera come testimonia l'uso dell'Harmonizer che agisce sulle frequenze trasportando le altezze di intervalli, in questo caso, dissonanti come semitoni, tritoni e settime minori. Si ottiene

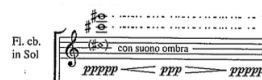


Figure 2. Incipit del brano: il flauto e il suono ombra.

così un insieme di voci che mantengono la loro identità: i loro conflitti non vengono completamente annullati, ma inseriti in un autentico contesto di ascolto.

2.2 SCRIVERE PER FLAUTO CONTRABBASSO E PER CLARINETTO CONTRABBASSO

Luigi Nono utilizza flauto e clarinetto contrabbasso in maniere del tutto non convenzionale, infatti gli esecutori sono chiamati ad eseguire note agli estremi dell'estensione con dinamiche quasi sempre intorno al piano. Il primo effetto di tali richieste è quello di rendere il suono risultante fortemente instabile.

Il brano richiede di utilizzare tre fondamentali metodi di emissione per entrambi gli strumenti: (i) prevalenza del rumore di aria, con poca o nessuna presenza di suono ad altezza determinata; (ii) mescolanza di suono ad aplettezza determinata e rumore d'aria, dosati in proporzioni differenti a seconda della prossimità all'emissione (verso l'aria o verso il tono); (iii) suono in emissione ordinaria.

Vengono inoltre specificati altri tipi di effetti. Ad esempio, per entrambi gli strumenti c'è il suono *Eolien*, inteso come "whistle" nella prassi esecutiva del flauto. Il risultato deve essere un suono composto da armoniche molto alte.

Per il flauto viene anche descritto il "suono ombra" come costituito da diadi armoniche in cui la tonale ha una presenza intermittente. Questo tipo di suono viene indicato da una nota romboidale fra parentesi tonde e la partitura indica che deve apparire e sparire con la discontinuità di un'onda, alternando le sue due componenti ma senza separarle. Il brano inizia appunto con un "suono ombra" come riportato in figura 2. Un'altra modalità di emissione del flauto è il "fischio", si tratta di replicare, con il fischio appunto, una delle due note che costituiscono un bicordo oppure di eseguire note brevi con il fischio, indicate in partitura non trasposte. Si noti infatti che in partitura non sono indicate le note reali poiché si tratta in entrambi i casi di strumenti traspositori. Infine per il flauto c'è anche la modalità "lasciar sospeso", in cui l'esecutore smette di suonare lasciando alla sola elettronica la trasformazione del suono.

Per quanto riguarda il clarinetto una differente modalità di emissione richiesta è quella del "cluster", tale termine denota l'aggregazione di armonici, con differenti gradi di densità, sviluppati a partire da un suono fondamentale. In sostanza questi aggregati sono prodotti nella forma di sortite irregolari e frastagliate sui registri acuti (cluster acuto), medi (cluster medio), o bassi (cluster basso), in funzione della pressione delle labbra e della concomitante azione dei denti.

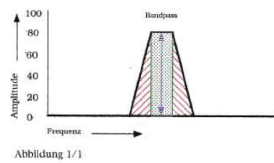


Figure 3. Appunto in cui Haller suggerisce l'utilizzo di un filtro passa banda.

2.3 Utilizzo dell'Elettronica e dello spazio come elemento compositivo

Come si è già detto, anche per la stesura di questa opera Nono si basa molto sulla collaborazione con gli strumentisti ed i tecnici dello studio di Friburgo ed a tal proposito Haller racconta un aneddoto molto particolare che ha portato ad alcune importanti scelte compositive per *A Pierre*, [10]. Il responsabile dello studio di Friburgo ricorda che in occasione di uno dei primi incontri per le sperimentazioni che avrebbero portato alla composizione dell'opera, i due musicisti che erano stati coinvolti, Fabbricani e Scarponi, arrivarono prima di Nono in studio ed il compositore comunicò poco dopo che sarebbe giunto solo nel primo pomeriggio. Haller chiese allora la disponibilità dei due musicisti ad effettuare alcuni esperimenti sul limite temporale del ritornello, ossia sulla lunghezza di una forma canonica musicale, oltre la quale, un'eventuale sua ripetizione, non venga percepita come tale, perché si è persa la memoria musicale di troppi elementi. Imponendo che tra una ripetizione e l'altra di una successione di note non ci fosse silenzio acustico, il risultato dell'esperimento portò ad un limite temporale pari a 24 secondi. Ulteriori esperimenti portarono alla deduzione che l'estensione di questo intervallo temporale dipendeva anche dal timbro dei suoni, dal loro ritmo e dalla posizione della sorgente. Così Haller inserì alcuni filtri sui suoni, ne cambiò la sequenza e vi aggiunse del riverbero, il risultato che ottenne fu che il limite temporale si riduceva a 12 secondi. Suonando la sequenza originale su quella modificata che veniva riprodotta da due altoparlanti, era inevitabile ri-registrare alcuni tratti del segnale così costituito (originale+elaborato) dando luogo ad un leggero feedback ai limiti dell'udibile. In quel momento Haller racconta che arrivò Nono, inaspettatamente in anticipo, ed il risultato fu che uno degli elementi principali di *A Pierre* è il canone a tre voci basato su un ritardo di 12 e 24 secondi rispetto al suono originale prodotto dal duo, cui vengono aggiunte delle trasformazioni molto gentili, come quella indicata nell'appunto di Haller in figura 3 che consiste in un semplice filtro passa banda.

Per il canone viene usato il principio di eco radiale, come indicato in un altro appunto di Haller (figura 4). L'eco radiale si basa sul semplice fenomeno psicoacustico in base al quale, se il suono diretto diminuisce in intensità, mentre quello riverberato rimane costante, si percepisce il suono diretto come proveniente da un punto sempre più distante. Tale fenomeno può diventare più evidente attraverso opportune modifiche del suono diretto.

A riguardo dei suoni prodotti dai due strumenti tradizio-

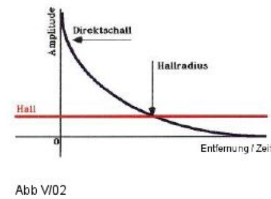


Figure 4. Appunto in cui Haller suggerisce l'utilizzo della eco radiale.

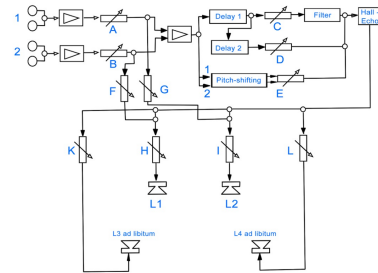


Figure 5. Schema di diagramma tecnico.

nali, Nono richiede un range vasto di emissioni e di pitch allo scopo di produrre suoni che vadano dal pacifico all'aggressivo tramite appunto la combinazione dei due parametri di cui sopra. Inoltre viene applicato un pitch shift in due altezze al suono diretto, cui viene poi aggiunto un vibrato molto lento, il che implica un controllo del pitch variabile. Di conseguenza per evitare l'effetto sirena, la trasposizione deve essere controllata in maniera molto delicata. La ricchezza dei suoni emessi dagli strumenti, assieme alla selezione di porzioni dei suoni tramite i filtri passa basso e alla trasformazione in altezza, genera dei suoni molto ricchi spettralmente e variabili in modo cangiante. I singoli valori per i controlli dinamici e dei tempi di riverbero vengono indicati come percentuali e richiedono quindi di essere adattati alla specifica sala da concerto. Nella maggior parte dei casi i suoni emergono dal silenzio, quindi a partire da un pianissimo indicato con 5 p ed in tutto il brano vi è un solo forte. Nono sviluppò ulteriormente la ricerca di Haller allo scopo di evitare ripetizioni udibili del canone. Il risultato è dato da pause di circa 6 secondi che dissolvono completamente la rigida sequenza di ripetizioni da 12 e 24 secondi. Chiaramente tutto questo procedimento rafforza ulteriormente il concetto di "suono mobile" precedentemente discusso.

Il diagramma tecnico fornito in partitura è mostrato in figura 5. Ciascuno strumento è amplificato con una coppia di microfoni, date le dimensioni fisiche. In figura 6 lo stesso schema di cui sopra, ma con gli effetti applicati esplicitati: La catena elettroacustica è dunque costituita da: harmonizer, banco di filtri passa banda, riverbero. Il diagramma in figura 7 mostra invece come il compositore richiede che vengano variate le dinamiche dei 4 altoparlanti coinvolti. Nonostante questa apparente semplicità, il performer al live electronics è chiamato ad occuparsi di due importanti possibili feedback. Come già detto la diffusio-

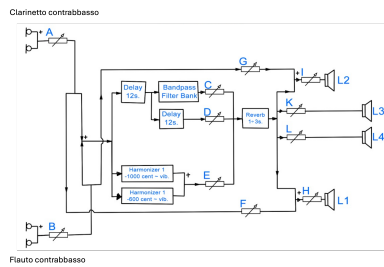


Figure 6. Schema di diagramma tecnico con effetti.

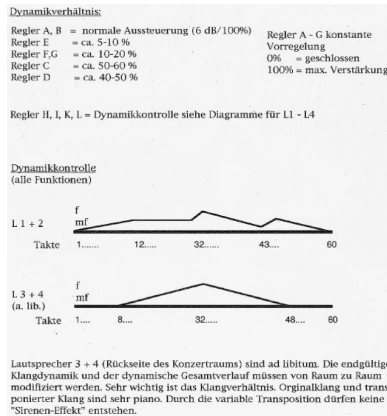


Figure 7. Gestione delle dinamiche per il performer al live electronics.

ne del suono è affidata a 4 altoparlanti, di cui 2 si trovano dietro gli esecutori e gli altri in fondo alla sala. Poiché il suono degli strumenti viene rilevato con 2 microfoni ciascuno, è chiaro che la prima sorgente di feedback è dovuta a questa vicinanza tra microfoni ed altoparlanti. La seconda causa di feedback proviene dall'utilizzo di filtri passa banda per riprodurre la prima ripetizione dei suoni dopo i 12 secondi. Le istruzioni del brano richiedono che il segnale diretto prelevato dai microfoni di ciascuno strumento, venga riprodotto sull'altoparlante opposto. Tutte queste indicazioni rendono il brano particolarmente delicato, non solo per gli esecutori agli strumenti tradizionali ma anche dal punto di vista dell'elettronica. Interessante è la raccomandazione di un tecnico-compositore che descrive la sua esperienza con il live di *A Pierre*, di non utilizzare strumenti di controllo del feedback o compressori disponibili con le tecnologie attuali, ma di correre il rischio con il setup indicato da Nono, proprio per mantenere perfetta coerenza con gli strumenti che si avevano a disposizione all'epoca, e che il compositore deve aver certamente tenuto in considerazione, [6].

3. ANALISI FORMALE

L'analisi formale qui discussa è effettuata sulla base di tre diversi elementi: l'analisi della partitura, la ricerca di una suddivisione formale, l'analisi dell'ascolto nella versione indicata [7]. Ad alto livello si osserva che il brano inizia

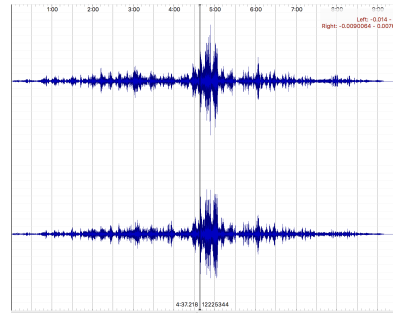


Figure 8. Forma d'onda compressiva di *A Pierre*.

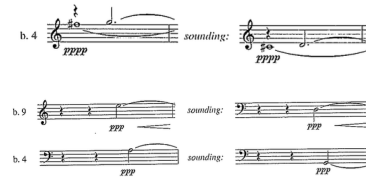


Figure 9. Indicazione sulla scrittura per il flauto contrabbasso e clarinetto contrabbasso ed il risultato acustico.

dal nulla e finisce nel nulla, tratto stilistico del compositore e che il massimo della dinamica è posto esattamente a metà del brano, come si evince chiaramente da una visione d'insieme della forma d'onda in figura 8. Tutto è caratterizzato da un crescendo e diminuendo molto fluido, senza particolari avvenimenti "impulsivi" nel mezzo. L'elettronica si dimostra chiaramente elemento fondamentale del brano, poiché con la gestione delle sue dinamiche tale andamento linearmente fluido è in qualche modo garantito. A proposito della dinamica c'è da osservare che tutto il brano tende a crescere fino alla sua metà, per poi diminuire fino al nulla del finale, ma questo aumento e decremento di volume caratterizza ogni "frase" del brano, come una sorta di microstruttura che viene poi replicata sulla macrostruttura totale.

Partendo dall'analisi della partitura prima di tutto si osserva che essa è costituita da 60 battute, presumibilmente una per ogni anno di vita di Boulez, e dunque, sulla base dell'indicazione metronomica (semiminima a 30), ci si aspetta che il brano duri circa 8 minuti. Una prima difficoltà che si incontra quando si consulta la partitura è legato al fatto che le altezze indicate in partitura non corrispondono a quelle effettive, in quanto entrambi gli strumenti utilizzati sono traspositori e dunque si dovrà tener conto del fatto che: (i) per il flauto contrabbasso in SOL il suono effettivo si trova una undicesima sotto, (ii) per il clarinetto contrabbasso in SI bemolle il suono effettivo si trova una sedicesima maggiore sotto del suono indicato in chiave di violino e una nona maggiore sotto del suono indicato in chiave di basso, si veda figura 9. Inoltre, la consultazione della partitura avulsa al confronto con l'ascolto è piuttosto complicata in quanto il risultato sonoro è difficilmente prevedibile, dato il fatto che la maggior parte delle note suonate vengono emesse con modalità non tradizionali e dunque

il risultato è un mix "inatteso" di suoni ombra, suoni al confine con il respiro, multifonici, toni whistle, il tutto impastato di elettronica che come prima prerogativa utilizza due lunghi ritardi (uno da 12 e uno da 24 secondi). D'altronde, stando a quanto afferma Fabbricani, Nono disse a proposito della prima performance del brano: "Non si deve capire niente!", [11]. Obiettivo perfettamente raggiunto.

Inoltre, sebbene nelle note introduttive della partitura pubblicata siano indicate le dinamiche da assegnare ai potenziometri, i tempi di ritardo, la durata del riverbero, cercare di capire cosa succede all'elettronica, al di là dell'aspetto dinamico sui 4 altoparlanti durante l'evoluzione del brano è molto difficile. Ad ogni modo le indicazioni per l'esecutore al live sono che le dinamiche degli altoparlanti L1 e L2, quelli fronte pubblico, sono regolate da una linea continua, mentre quelle degli L3 e L4 (retro, emettono solo elettronica e non suono diretto dagli strumenti) da una tratteggiata. A sottolineare la volontà di confondere l'ascoltatore si noti il fatto che L1, vicino al flautista riproduce il suono del clarinetto sommato a quello dell'elettronica, mentre L2, altoparlante vicino al clarinetista, riproduce flauto ed elettronica.

La difficoltà che si incontra nell'estrapolare informazioni significative dalla partitura è perfettamente coerente con l'idea che il compositore aveva di scrittura come "work in progress", senza alcun interesse particolare di tramandare ai posteri informazioni significative sull'esecuzione dei suoi lavori: "quando non ci sarò più, altri compositori scriveranno nuove cose".

Sulla base di queste osservazioni, per individuare una prima suddivisione formale si è deciso di partire dalla forma d'onda e dall'ascolto, accompagnata dalla consultazione della partitura. Guardando alla durata del brano rispetto a quella valutata moltiplicando la durata di ciascuna battuta per il numero di battute, il brano risulta più lungo di 1' e 10", ma questo ha senso perché la partitura è piena di corone, su alcune delle quali è indicata anche la durata. Sommando tutte le durate indicate sulle corone (in numero pari a 8) si arriva a 34" in più, mentre il numero di corone prive di indicazione temporale è pari a 4, comprensiva dell'ultima che è la quadrata quindi di durata più lunga possibile.

Per quanto detto in precedenza, una prima idea potrebbe essere quella di ragionare in termini di micro-macrostruttura, cioè cercare di effettuare una suddivisione in funzione delle sequenze crescendo-diminuendo che si riescono ad individuare. In questo modo però le microstrutture sono talmente tante, molte della durata di un paio di battute, che elencandole tutte non si riuscirebbe poi ad ottenere alcuna visione d'insieme.

Basandosi invece principalmente sull'ascolto e sulla forma d'onda del brano, si nota una specularità in termini di dinamiche proposte, nel mezzo della macrostruttura discussa ci sono dei momenti in cui, pur osservando le microstrutture di crescendo e diminuendo, il brano si aggira attorno ad una specifica dinamica. Sulla base di ciò è possibile grossolanamente individuare 4 sezioni, indicate direttamente sulla forma d'onda in figura 10.

Si può dunque cercare di dettagliare ciò che succede al-

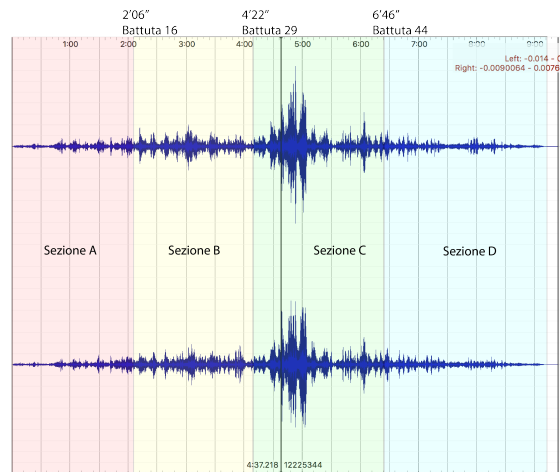


Figure 10. Proposta nuova suddivisione.

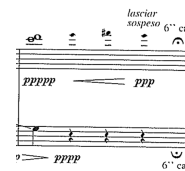


Figure 11. Indicazione in partitura sulla durata della corona.

l'interno di ciascuna sezione individuando eventualmente altri elementi che costituiscano delle sottosezioni.

SEZIONE A. Questa prima sezione è caratterizzata da dinamiche degli strumenti che si mantengono tra il *ppppp* e il *p*. L'elettronica sugli altoparlanti vicino agli strumentisti (L1 e L2) inizia il suo fade alla prima battuta e si assesta su un *mf* a battuta 12, mentre quella su L3 e L4 inizia in fade in battuta 8 e arriva a battuta 16, continuando a crescere, su *p*. Rispetto al continuum generato dagli strumenti con i loro effetti cui si sovrappone in maniera crescente l'elettronica, emergono delle note brevi suonate dal flauto, le prime due a battuta 8, seguite da quelle replicate dall'elettronica in lontananza e poi da altre tre suonate alla battuta 11, che si chiude con la prima delle corone che hanno una durata indicata, si veda figura 11. Queste note evidentemente senza elaborazione possono segnare una cesura per un'eventuale sottosezione. La corrispondente forma d'onda è riportata in figura 12 da cui si nota uno svuotamento abbastanza improvviso rispetto alla lenta accumulazione appena avvenuta. Il medesimo elemento di cesura, cioè le note rapide suonate dal flauto, si determina alla battuta 16, che si chiude con una corona ma senza indicazione sulla sua durata, figura 13.

SEZIONE B. Questa seconda sezione, come detto, è caratterizzata da un generico aumento delle dinamiche per entrambi gli strumenti, che questa volta si muovono tra il *ppp* e il *mf*. L'elettronica su L3 ed L4 continua a crescere mentre quella su L1 e L2 rimane fissa sul mezzo forte. Anche in questo caso si riescono ad individuare, sostanzial-

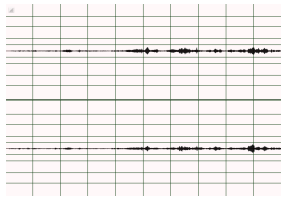


Figure 12. Svuotamento di ampiezza che segna una cesura.

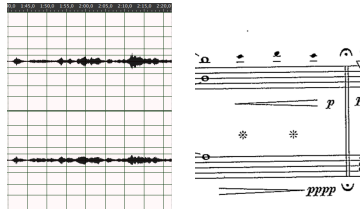


Figure 13. Cesura con corona senza durata.

mente tramite l'ascolto, delle sottosezioni, ma questa volta l'elemento caratterizzante non è più dato dalle note brevi suonate dal flauto, bensì da una variazione del contenuto spettrale del risultato sonoro complessivo. In particolare si determina uno stacco rilevante intorno ai 3'14", corrispondenti alla battuta 25, quando il fischio richiesto al flauto introduce una variazione nello spettro acustico presentato fino ad ora, generando suoni al limite del sibilo. Questo è evidente nell'ultima parte dello spettro riportato in figura 14. Da notare che anche questa sezione si conclude con due note brevi suonate dal flauto, in un registro più acuto rispetto alle precedenti.

SEZIONE C. La sezione C vede gli strumenti arrivare fino al forte e l'elettronica su L1 ed L2 ricominciare ad aumentare a partire dalla battuta 31. Le note acutissime suonate dal flauto a battuta 29 sono un forte elemento caratterizzante il finale della sezione precedente e preludono ad un incremento dinamico importante che caratterizza questa nuova sezione dovuto sia all'aumento delle dinamiche dell'elettronica, sia al cluster acuto suonato dal clarinetto a partire dalla battuta 34 (4'52").

Per la sezione C si possono individuare 3 sottosezioni. La prima è una sorta di introduzione all'importante aumento dinamico e va dalla battuta 30 alla battuta 33, in termini di minutaggio sulla registrazione [7] si tratta dell'intervallo tra 4'22" e 4'52", si veda figura 15. Le dinamiche degli strumenti non vanno sopra al *mf*, ma osservando la forma

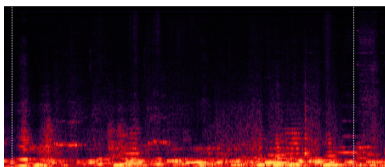


Figure 14. Caratterizzazione di una sottosezione tramite osservazione dello spettro.

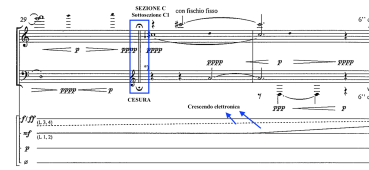


Figure 15. Inizio della sottosezione C1.

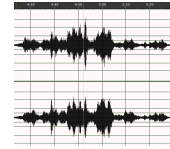


Figure 16. Forma d'onda relativa alla porzione di partitura in figura 15.

d'onda e sulla base dell'ascolto il crescendo è evidente, come del resto indicato in partitura.

Dalla battuta 34 inizia quindi la sottosezione C2 che vede gli strumenti toccare la loro massima dinamica sul forte e anche l'elettronica arrivare al suo picco *ffff*, sulle battute 35 e 36, per poi iniziare un lento e lunghissimo fade out.

Si può dunque individuare il successivo momento di cesura in relazione al diminuendo che arriva a battuta 37, circa 5'30". Questo perché il momento clou del crescendo è determinato da un cluster "acuto, massima pressione labbra", la cui eco ritorna più lontana ("cluster basso, minima pressione labbra") e ad un'altezza diversa alla battuta 38, da cui si stabilisce di far partire la terza sottosezione C3. Quello che si ascolta in questa sottosezione è dunque una sorta di eco lontana a quanto è stato appena ascoltato, chiaramente ad altezze diverse che preludono ad una lunghissima coda conclusiva. Tale eco va indicativamente da battuta 38 a battuta 44 e vede complessivamente una diminuzione delle dinamiche (sempre considerando però la microstruttura crescendo-diminuendo). Si osserva che a partire dalla battuta 43, mentre le dinamiche per L3 e L4 continuano a diminuire, L1 ed L2 subiscono un innalzamento dinamico abbastanza rapido rispetto al resto del pezzo, per poi ricominciare la graduale discesa fino alla fine, come a non ritardare il lungo fade out della coda. Per concludere il discorso sulla sezione C, il crescendo-diminuendo complessivo si nota chiaramente sia dall'ascolto che dalla forma d'onda che dalle dinamiche degli strumenti indicati in partitura, ma si vuole osservare la presenza di numerose microstrutture dinamiche (crescendo diminuendo) che

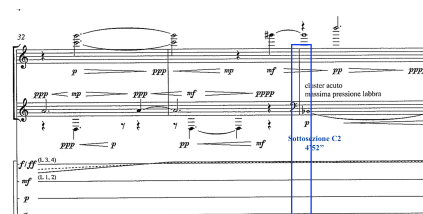


Figure 17. Sottosezione C2.

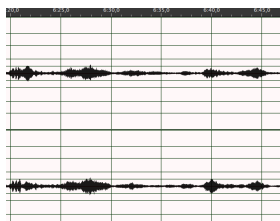


Figure 18. Forma d'onda della porzione di partitura in figura 17.

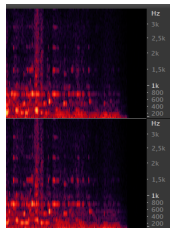


Figure 19. Spettrogramma del finale del brano.

lo costituiscono e che hanno tutte pressappoco una durata compresa tra 8 e 10 secondi. Dunque il compositore procede alla realizzazione della dinamica complessiva per microstrutture.

SEZIONE D: CODA. Ciò che si ascolta a partire dalla battuta 45 è in realtà una lunghissima coda che porta il brano a terminare praticamente nel nulla. Da questo punto in poi si nota uno svuotamento, sia dell'elettronica sia delle note sulla partitura. Ad esempio per quanto riguarda il flauto la frequenza dei bicordi tende a diminuire e c'è in generale una prevalenza di soffio e tecniche di emissione che rendono il risultato acustico se possibile ancora più aleatorio. Il clarinetto si sposta in chiave di violino e a battuta 57 ad esempio con la realizzazione degli armonici si muove anche sopra al flauto. La presenza di note ad alta frequenza si vede chiaramente dalla porzione finale dello spettro riportato in figura 19. Nonostante ciò la diminuzione dinamica è evidente ed anche in questo caso si possono individuare tante microstrutture dinamiche che complessivamente vanno in diminuendo. L'elettronica su L3 ed L4 si azzerava a partire dalla battuta 49, mentre quella su L1 e L2 prosegue sino alla fine, in una modalità un po' vaga. Infatti proprio alla fine della partitura, l'indicazione per l'esecutore al live è "fade out poco a poco", mentre per gli strumentisti "lasciando sospeso". Ora, il sospeso degli strumenti risulta abbastanza chiaro, finito il fiato, rimangono con gli strumenti sulle labbra fino a quando non termina anche l'elettronica. La questione sull'elettronica non è invece esplicita poiché non è chiaro se l'interprete al live debba attendere o meno che venga a concludersi il contributo del secondo delay, e anche l'ascolto in questo senso non è di aiuto. Il risultato complessivo è comunque incredibilmente affascinante.

Di seguito un sunto dell'analisi esposta:

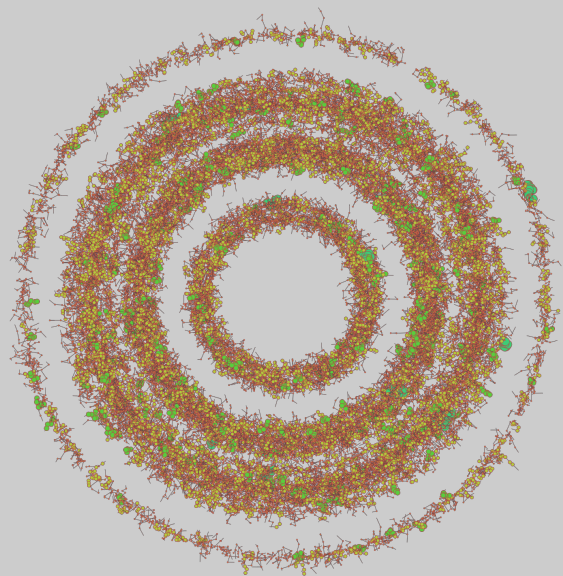
| Sezione | Sottosezione | Tempo | Battute | Dinamiche |
|---------|--------------|---------------|---------|------------|
| A | | fino a 2'22" | 1 - 16 | ppppp - p |
| | A1 | fino a 1'40" | 1 - 11 | |
| | A2 | 1'40" - 2'22" | 12 - 16 | |
| B | | 2'22" - 4'22" | 17 - 29 | ppp - mf |
| | B1 | 2'22" - 3'14" | 17 - 24 | |
| | B2 | 3'14" - 4'22" | 25 - 29 | |
| C | | 4'22" - 6'20" | 30 - 44 | ppppp - f |
| | C1 | 4'22" - 4'52" | 30 - 33 | ppppp - mf |
| | C2 | 4'52" - 5'32" | 34 - 37 | ppppp - f |
| | C3 | 5'32" - 6'20" | 38 - 44 | ppppp - mf |
| D | | 6'20" - fine | 45 - 60 | ppppp - p |

4. CONCLUSIONI

In questo articolo è stata presentata una proposta di analisi formale del brano composto da Luigi Nono nel 1985 "A Pierre. Dell'azzurro silenzio, inquietum". Il processo che ha portato a questo lavoro ha dovuto necessariamente comprendere l'analisi della partitura, dell'elettronica, delle esecuzioni e dell'evoluzione estetica dell'autore, tenendo in massima considerazione la procedura compositiva che ne ha caratterizzato gli ultimi brani e basata sulla stretta collaborazione con gli esecutori, che se da un lato hanno aperto alla possibilità di utilizzare tecniche esecutive innovative, da un lato rendono le opere fragili alla conservazione nel tempo.

5. BIBLIOGRAFIA

- [1] Da uno scritto di Luigi Nono su La fabbrica illuminata (1964) (Luigi Nono, Scritti e colloqui, a cura di A. I. De Benedictis, V. Rizzardi, ed. Lim., Ricordi, Milano 2001).
- [2] Agostino Di Scipio, La musica di due culture. Tracce di una mutazione. In AA. VV., Musica e Scienza. Il margine sottile, a cura di Carlo Boschi, 1991, ISMEZ.
- [3] Paolo Pinamonti - tratto dal catalogo "Con Luigi Nono. Festival internazionale di musica contemporanea, La Biennale di Venezia, 1992-93", Ricordi, Milano 1993, p. 246.
- [4] Luigi Nono. Scritti e colloqui, a cura di A.I. De Benedictis e V. Rizzardi, Ricordi-LIM (Le Sfere, 35), Milano 2001, vol. I, p. 505.
- [5] Laura Zattra, Ian Burleigh and Friedemann Sallis, "Studying Luigi Nono's A Pierre. Dell'azzurro silenzio, inquietum (1985) as a Performance Event", 2011.
- [6] Parra Cancino, Juan Arturo, "Multiple paths : towards a performance practice in computer music", 2014, Leiden University Repository.
- [7] Roberto Fabbriciani, Ernesto Molinari & South West German Radio Experimental Studio, 2013 NEOS Music.
- [8] <http://www.luiginono.it/opere/a-pierre-dellazzurro-silenzio-inquietum/#tab-id-2>
- [9] "Luigi Nono e Pierre Boulez. Due cori che ascoltano la memoria in A Pierre, dell'azzurro silenzio, inquietum", Orfeo nella rete, di Stefania Navacchia, <http://www.orfeonellarete.it/rubriche/articoli.php?idart=00092>
- [10] A PIERRE OMAGGIO A GYÖRGY KURTÁG, Lecture given at Venice by Hans Peter Haller, 1999, <http://www.hp-haller.homepage.t-online.de/venice.html>
- [11] https://medias.ircam.fr/x84b595_ seeking-virtual-voices-in-luigi-nonos-a-p



Memorie proiettive/Projecting Memories
Atti del XXIV Colloquio di Informatica Musicale
Proceedings of the XXIV Colloquium on Music Informatics

