1 2 9 0

UNIVERSIDADE D COIMBRA

João Carvalho Oliveira das Neves

Acoustic Mirrors
From Sound to Tangibility

Dissertation in the context of the Master Degree in Design and Multimedia, advised by Professor Artur Luís Gonçalves de Azevedo Rebelo Alves, Professor Pedro José Mendes Martins and by Professor Fernando Amílcar Bandeira Cardoso and presented to the Department of Informatics Engineering of the Faculty of Sciences and Technology of the University of Coimbra.

September 2024

Acoustic Mirrors

From Sound to Tangibility

Written by
João Neves

Supervised by
Artur Rebelo,
Pedro Martins and
Amílcar Cardoso

Master Degree Dissertation in
Design and Multimedia
Faculty of Science and Technology
Department of Informatics Engineering
University of Coimbra

2023
2024

Festina lente.

Resumo

A relação entre a computação e a arte tem-se intensificado tanto em contextos performativos como expositivos. A criação e acessibilidade de novas ferramentas e métodos computacionais facilitam a disseminação e popularização da arte computacional e de novos meios por artistas com formações mais ecléticas. Paralelamente a este desenvolvimento, a velocidade, imediata e efemeridade do mundo moderno estabeleceu uma dominância visual, negligenciando a atenção ao som e ao silêncio.

Nesta dissertação, exploramos a relação entre computação, som, imagem e estética através do desenvolvimento de uma instalação audiovisual interativa: S—I (Som, depois Imagem). A instalação utiliza composições sonoras generativas e oferece dois momentos: A) um momento de exploração (de natureza artística); B) um desafio de espacialização, utilizando a memória auditiva experimentada no primeiro momento (de valor científico). O nosso objetivo é integrar o público no processo algorítmico de criação sonora, enquanto utiliza a sua percepção auditiva no ambiente criado como uma ferramenta para a espacialização. A instalação é conceptualmente inspirada pelos Espelhos Acústicos, dispositivos monolíticos de betão britânicos desenhados por William Sansome Tucker nas décadas de 1920 e 1930, destinados a detetar aeronaves inimigas através do uso do som como medidor de distância e localização. A abordagem baseia-se na recolha de material de pesquisa relevante, no desenvolvimento do objeto final, na criação de artefatos visuais estáticos resultantes da interação com a instalação e na reflexão à luz de uma avaliação fenomenológica.

Palavras Chave

Arte Sonora
Arte Computacional
Arte de Novos Meios
Síntese Sonora
Composição Generativa
Espacialização Sonora
Design Gráfico

Abstract

The relationship between computing and art has been intensifying in both performative and exhibition contexts. The creation and accessibility of new computational tools and methods facilitate the dissemination and popularisation of computational and new media art by artists with more eclectic backgrounds. Alongside this development, the speed, immediacy, and ephemerality of the modern world have established visual dominance, neglecting attention to sound and silence.

In this dissertation, we explore the relationship between computation, sound, image and aesthetics through the development of an interactive audiovisual installation: S—I (Sound, then Image). The installation employs generative sound compositions and provides two moments: A) an exploration moment (artistic in nature); B) a spatialization challenge, using the auditory memory experienced in the first moment (of scientific value). Our goal is to integrate the audience into the algorithmic sound creation process while using their auditory perception in the created environment as a tool for spatialisation. The installation is conceptually inspired by Acoustic Mirrors, British monolithic concrete devices designed by William Sansome Tucker in the 1920s and early 1930s, intended to detect enemy air force through the use of sound as a measure of distance and location. The approach is based on gathering relevant research material, developing the final object, creating static visual artefacts resulting from interaction with the installation, and discussing it in light of a phenomenological evaluation.

Keywords
Sound Art
Computational Art
New Media Art
Sound Synthesis
Generative Composition
Sound Spatialisation
Graphic Design

Index

List of Figures

List of Tables

Chapter I
Introduction

> *"… art, science and the creative process. Experimentation and ultimately failure are an intrinsic commonality of all three."*
> — PETTET-SMITH (BBC NEWS, 2019)

Over the course of the last decades, technology has expanded the possibilities of how computers can be used and the roles they can adopt. From the most mundane to the most complex tasks, their ever-increasing presence in our daily lives is an established reality. The boundaries of the artistic field are constantly being expanded and redefined, as the available methods and creative processes evolve, allowing for new approaches with innovative possibilities. The growing access to computational systems and development tools has allowed artists from more diverse backgrounds to explore all sorts of expression materials, from sound to visuals, giving rise to the proliferated and popular new media of today.

This dissertation presents an installation that integrates the audience into the algorithmic process of sound creation. The co-creative system in question employs generative sound synthesis and, interactively, provides two moments: A) an exploration moment (artistic in nature); B) a spatialization challenge, using the auditory memory experienced in the first moment (of scientific value).

The aim is to relate sound, aesthetics, and the use of computation with the participation and interaction of the audience, evaluating their capacity for spatialisation and perception in the developed environment. By exploring ways to expose and materialise interaction, we create static computational artefacts that help us to analyse the results.

We chose Acoustic Mirrors (AM) as a metaphorical foundation for the installation, drawing inspiration from these British devices used in the 1920s and early 1930s to detect enemy aircraft through sound. The physical characteristics, purpose, and historical and geographical context of AM serve as a guiding influence in the creation and operation of the installation.

Motivation

This research is motivated by my background as a designer and a jazz musician. Throughout my artistic practice as a composer and pianist, the connection between stimuli and creativity has always puzzled me. I often find myself recurring to standalone pictures, random movie clips, or other types of visual media when trying to come up with new melodies or harmonies. The process goes both ways, as I usually mentally wonder how certain visual content could be sonified.

To some extent, I personally consider computation and music improvisation to share similar paradigms. A) They should both have rules that can potentially take us to unknown places — improvisers are influenced by tempo, key, genre, and other theoretical music parameters (apart from non-idiomatic/free improvisation contexts), while algorithms dictate the behaviour of every system. B) They can both converge or diverge from their context — when we compare renowned jazz groups we may stumble upon completely opposite musical approaches. Keith Jarrett Trio and Brad Mehldau Trio are a clear example of brilliant yet distant musical aesthetics. While Jarret's rhythmic section fills the moments where the piano is silent (Keith Jarrett Trio, 1997), Mehldau's is constantly searching for mutual and synchronous climaxes (Brad Mehldau Trio, 1997). As for computation, the rules of the system define it as more or less deterministic. C) They usually pretend to be what they are not — when sharing artistic contexts, computation tends to mimic nature-inspired behaviours while musical improvisation is often viewed as an extension of the theme's melody.

From a designer standpoint, I am also inspired by the exploration of multidisciplinary universes that bring together distinct subjects such as digital media and editorial design, image and typography, mathematics and poetry, and so on. I believe that exploring distinct fields and their intersections, whether the result is successful or not, contributes to the development of a more well-rounded designer, as new options and resources as well as conceptual and empirical awareness arise. These guidelines have been the key throughout my academic and freelance journey.

Goals

The abundance of visual stimuli in the modern world leads to a growing deterioration of our auditory capacity (Ikoniadou, 2014). One of the main objectives of this dissertation is to create a space that provides moments that stimulate the auditory apparatus and its use for tasks that were archaically fundamental to humans, such as spatialisation through sound. Subsequently, we aim to evaluate the experience by combining a phenomenological approach with a comparison of results from experiments conducted with audiences having diverse academic backgrounds and varying levels of knowledge about the installation.

With the purpose of developing a sonically rich and computationally robust installation, we studied work and developed skills particularly in the areas of sound synthesis, algorithmic composition, body detection, and sound design. Concurrently, the framework and contextualisation of this dissertation are based on a survey of works of artistic nature, sound art, new media art, musical notation, sonification, visualisation, and relevant models and algorithms.

Document Structure

The work presented herein is based both on the study and analysis of related work and the development and discussion of computational experiments that investigate the synergy between the various fields of focus. The present Chapter I — *Introduction* introduces the theme of this dissertation, its motivation and its goals. In Chapter II — *Work Plan* we present the process agenda and methodology. In Chapter III — *Prelude* covers key concepts for contextualising the developed work. Chapter IV — *State of the art* contains a selection of relevant research from various field studies that inspired or guided the practical work. Chapter V — *Preliminary Work* presents a compendium of audiovisual static explorations that guide the installation. In Chapter VI — *Installation* we provide a detailed description of the process that led to the final installation. We present the system, its evaluation and all the associated graphic material. We end with Chapter VII — *Conclusion* reflecting on the developed work.

Chapter II
Work Plan

This chapter displays both the methodology adopted and an overview of the plan for the development of this project. The plan is divided into major tasks with deadlines that keep the progress accountable and ensure its overall fulfilment.

Methodology

We adopted Cross (2000, as cited in Dubberly, 2005) methodology "Four stage design process" (figure 1) for its simplicity and pragmatism. This model is based on the essential activities that a designer performs — Exploration, Generation, Evaluation, and Communication. After exploring the problem and its context in the first phase, the process enters an iterative loop where various proposals are conceived and evaluated. The last stage starts when a proposal is found adequate. The construction of the initial chapters (*Prelude*, *State of the Art*, and *Preliminary Work*) corresponds to the Exploration phase as does the process of building synthesisers to be selected for the installation and the drafts of the visual computational outputs resulting from the experiments. The generation and evaluation pertain to the experimentation with the system and the constant adjustments made after and between tests. The communication phase involves the presentation and exhibition of the final system.

Tasks

The project is partitioned into five distinct tasks. In *Bibliographic Review*, we research and collect bibliographic references about the dissertation-related subjects so we are properly informed and technologically and culturally up to date. *Writing* refers to the conception of the intermediate report and the final submission of the essay. *Development* details S—I development process, which is divided into categories necessary for the materialisation of the installation — conceptualization of the object, experimentation and preliminary tests, implementation of the system, familiarisation with required and uncharted technologies, and design of the object's graphic material. In *Evaluation*, we assess the developed work. In *Dissemination*, we planned to promote the project by preparing a research paper for submission and formally exhibiting the installation.

We developed a *Gantt* diagram to properly schedule all the processes (figure 2.A). Throughout the process, other professional projects overlapped with the development of this dissertation, delaying its completion by one year. Figure 2.B addresses that change.

Exploration
↓
Generation
↑↓
Evaluation
↓
Communication

Figure 1
Four stage design process diagram (Cross, 2000, as cited in Dubberly, 2005).

Furthermore, despite our intentions to formally and publicly exhibit S—I, we are still awaiting the results of two open calls (Semibreve 2024 and MATE 2024). We decided to postpone writing the research paper to align with the specifications of a future conference, which will take place after the submission of this dissertation.

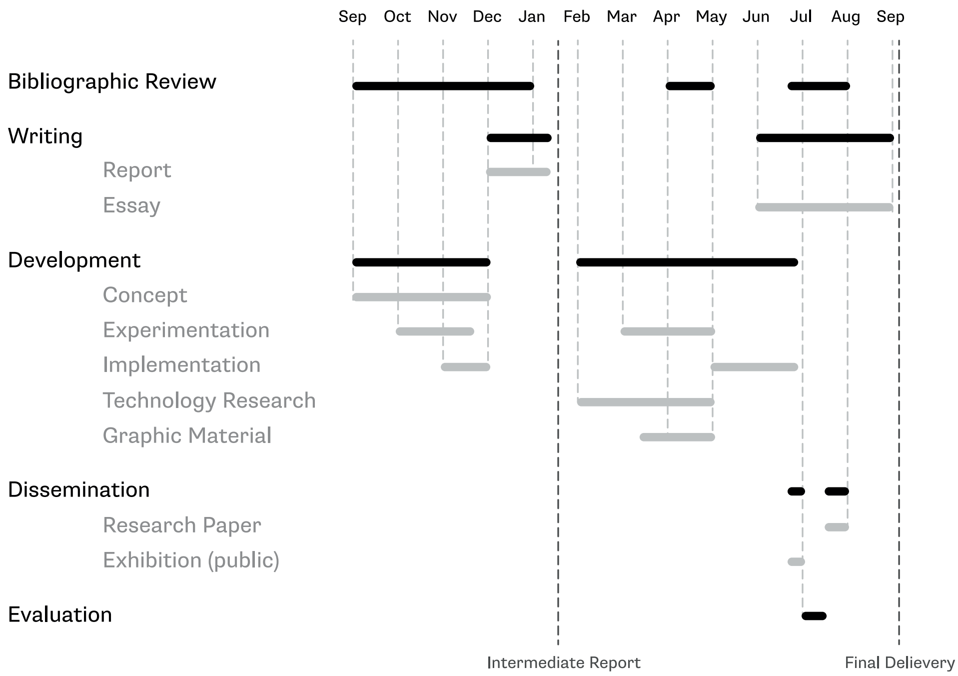


Figure 2.A.
Initial *Gantt* diagram with predicted tasks.

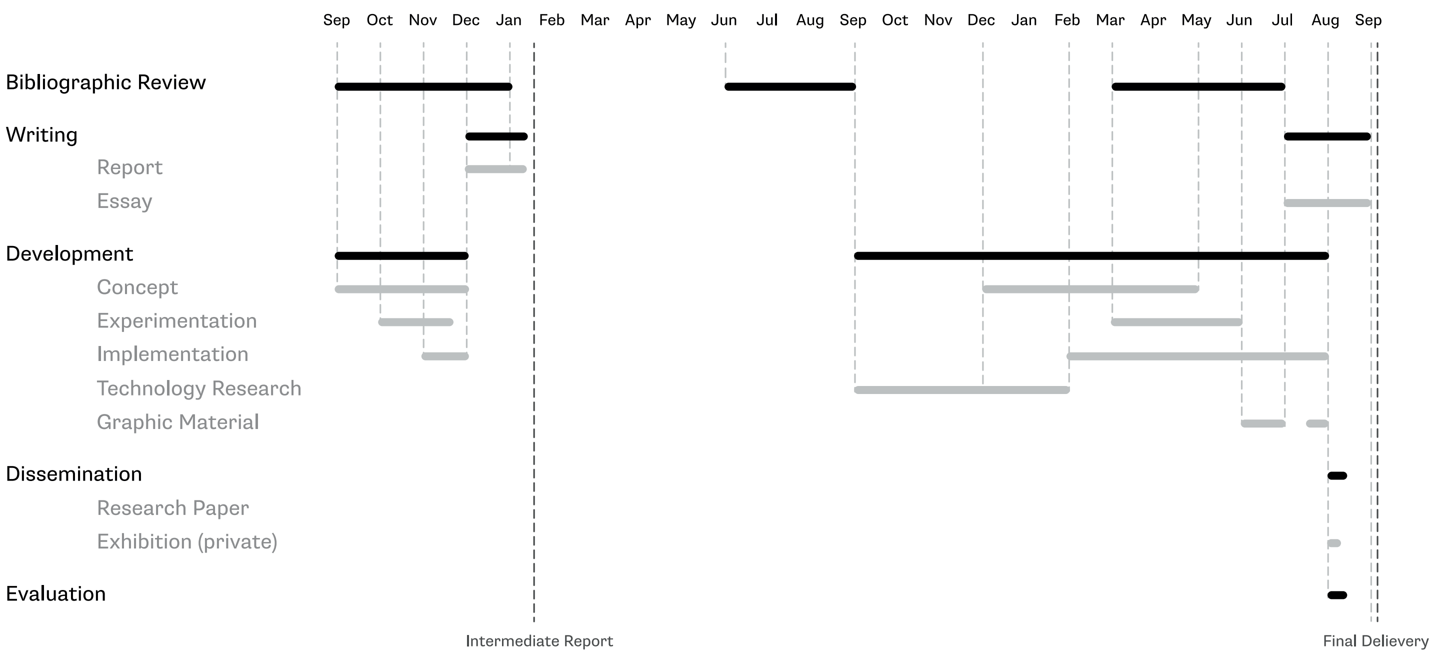


Figure 2.B.
Realistic *Gantt* diagram.

Chapter III
Prelude

This chapter introduces basic and necessary concepts for the dissertation comprehension, aiming to contextualise it historically from a psychological and anthropological standpoint. It also presents the source and conceptual foundation for the development of our work.

An Overview on Perception

Hearing is continuous and does not require activation (Chion, 1994). Moreover, the hearing experience is almost mandatory as sound cannot be touched and we cannot shut our ears (Dyson, 2009). Unlike vision, whose primary characteristic is analytic, hearing's primary characteristic is synthetic (Solomos, 2018). In constant recurrence, it is responsible for warning us when our body is at rest. In ancient times, it was essential for detecting environmental disturbances that could compromise survival and preservation of the species.

The first registered beings with binocular vision are estimated to have existed in the *Carboniferous* period, around 300 million years ago (Ted Institute, 2013). This heritage allows us to move around and interact with our surroundings with no room for error in an uncomplicated fashion. It is responsible for the perception of depth in a three-dimensional space, which we call binocular vision stereopsis (Bhola, 2013). There are, however, other animals with different and equally effective visual systems, some of which inspired the development of computer vision due to their unique characteristics. The jumping spider, with its eight eyes, two of which allow it to analyse patterns (Jackson, 1982); Cubozoans, with their twenty-four eyes, which provide them with a detailed view of the constantly dangerous aquatic environment (Nilsson et al., 2005) are just two examples among others.

> *" ...perception is a process of controlled hallucination..."*
> — Bruno Olshausen (2020)

Perception is the ability to absorb collected information, interpret it through the analysis of hypotheses and subsequently react to it. The brain manages the process of transforming these stimuli by filtering noise and assigning them a meaning. According to many researchers (e.g., Snyder, 2001; Knudsen 2007), incoming information is stored in three different places depending on the stage of the process. It changes from sensory memory to short-term memory and thento long-term memory, from which it can be accessed permanently.

Modern world brought a clear dominance of visual stimuli over sound's (Ikoniadou, 2014). In the practical component of this dissertation we take into account this reality by manipulating the experiences so that vision does not dominate over audition (Sinnet et al, 2007). It cannot be denied that sound has a haptic dimension, which can be quite literal if we consider sound in terms of vibration (Solomos, 2018). We are interested in creating an experience that also explores sound as a sensorial and emotional trigger. We convey Carvalhais' (2010) intention of trying to make the audience's experience less obvious, and "explore grey areas (...) to produce a greater variety of perceptual effects".

Lombard and Ditton (1997) present six notions of presence: presence in social richness, presence as a social actor within the medium, presence as a social actor, presence as realism, presence as transportation, and presence as immersion. The latter is what defines and represents the immersiveness we describe. They define this concept as a "perceptual and psychological immersion," where the senses are deeply engaged in a virtual environment, leading to a distorted perception of reality. This immersion can be achieved by limiting sensory inputs (Kim, 1996), which requires the interpretation of virtual scenarios. Additionally, presence as immersion has a psychological component, as it can involve (Palmer, 1995) and absorb (Quarrick, 1989) individuals within the experience.

Sound Physics Synopsis

In physics, "sound (or noise) is the result of pressure variations, or oscillations, in an elastic medium (e.g., air, water, solids), generated by a vibrating surface or a turbulent fluid flow." (Hansen, 1951). These oscillations propagate in the form of longitudinal waves, creating disturbances in the medium. According to Cipriani and Giri (2013), there are three fundamental sound parameters — frequency, amplitude and waveform. Each one of them influences how we perceive sound and translates into ordinary qualities of our sonic perceptual sensation. Frequency variation reflects on pitch, amplitude translates into intensity and waveform influences our ability to distinguish timbres (table 1).

Characteristic	Parameter	Perceptual Sensation
Frequency	Pitch	High - Low
Amplitude	Intensity	Forte - Piano
Waveform	Timbre	Sound colour

Table 1.
Correspondences between sound characteristics, musical parameters and perceived sonority (Cipriani & Giri, 2013).

Frequency is measured in hertz and allows us to distinguish between high and low-pitched sounds. When sound (or a mechanical perturbation) is registered, the medium is affected at a molecular level. This disturbance is called pressure. When analysing pressure over time, we observe a compression and a rarefaction of the medium (figure 3). This phenomenon is called a cycle. While a wave period is the time required to complete a full cycle, frequency includes the number of cycles completed per second. The range of frequencies that is audible to humans extends from approximately 20 to about 20.000 hertz (Cipriani & Giri 2013).

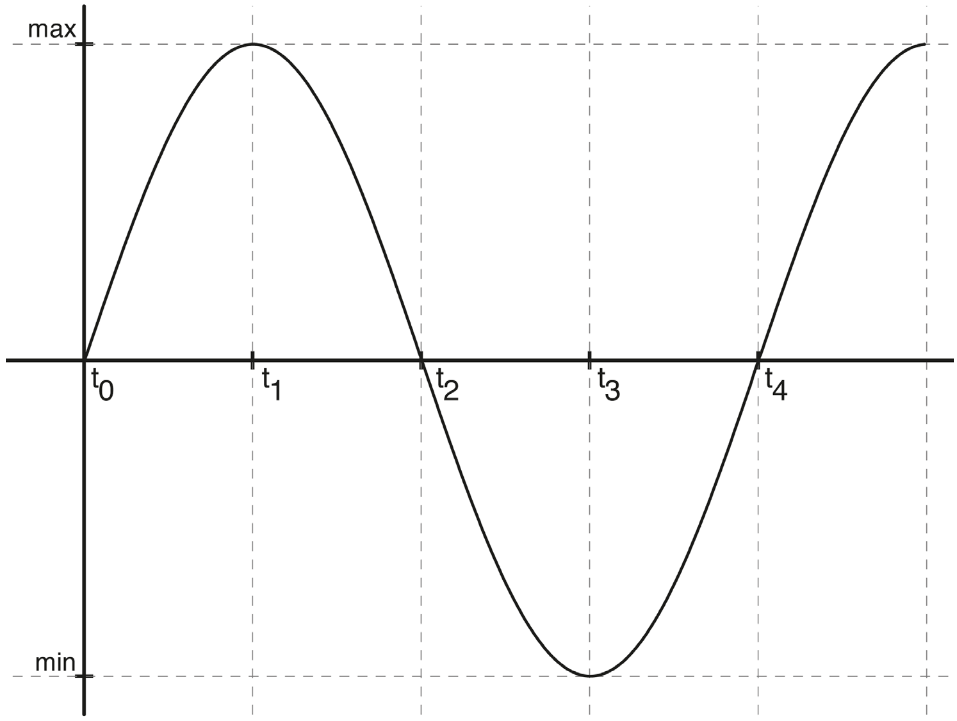


Figure 3.
A graphical representation of compression and rarefaction (Cipriani & Giri, 2013).

While waveforms are directly associated with timbre and with our ability to distinguish instruments (figure 4), amplitude expresses the pressure of the medium — how much it compresses and expands, impacting on intensity (or volume).

We can manipulate these waveforms by applying certain filters. Filters are essential tools in audio processing, used to shape and modify sound by allowing certain frequencies to pass through while blocking others. A band-pass filter allows frequencies within a specific range to pass through and attenuates frequencies outside that range. A high-pass filter allows frequencies above a certain threshold to pass while blocking those below it, while a low-pass filter does the opposite, allowing only lower frequencies to pass and softening higher ones. There are also low-shelf and high-shelf filters that reduce the volume of sound signals after the set frequency by the desired amount rather than entirely removing them (figure 5).

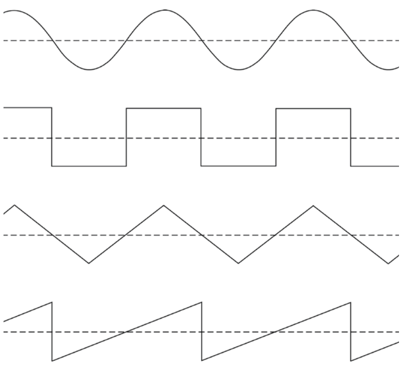


Figure 4.
Different types of waveforms. Adapted from Cipriani and Giri (2013).

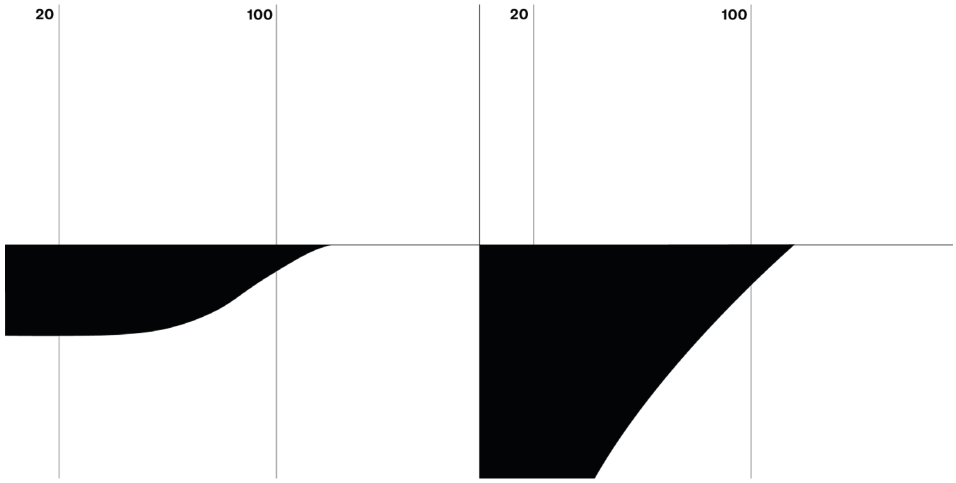


Figure 5.
Examples of low-shelf filter (left) and low-pass filter (right) shapes. Adapted from Cipriani and Giri (2013).

The reciprocal quality (or Q factor) of a filter, particularly relevant for band-pass filters, determines how narrow or broad the range of frequencies that it affects. A high Q factor results in a narrow, selective frequency range, while a low Q factor means a wider, less selective range. By adjusting the reciprocal quality, we can control the sharpness and precision of the filter's effect on the sound.

For the context of this thesis, we add panning as a fundamental concept to be clarified. It translates to the distribution of sound across a spatial axis and is a powerful factor to manipulate when conceiving immersive sound experiences, as it helps simulate distance and positioning. Panning involves placing audio signals at various points within the stereo or multi-channel sound field, creating the illusion of sound originating from different directions.

Sound Synthesis Basics

Sound synthesis is a technique in the field of audio engineering and music production, encompassing the creation of sound through artificial means rather than through traditional acoustic sources. It involves the generation and manipulation of sound waves using various algorithms and electronic components. The primary methods of sound synthesis include additive synthesis, in which complex sounds are built by summing waves; subtractive synthesis, which involves shaping sound by filtering out frequencies from a harmonically rich signal; and frequency modulation (FM) synthesis, where the frequency of one waveform modulates the frequency of another, creating a wide range of timbres (Cipriani & Giri, 2013). These methods allow for the creation of a diverse array of sounds, from realistic instrument emulations to entirely novel audio textures. Additionally, granular synthesis and wavetable synthesis are important methods that offer unique approaches to sound design. Granular synthesis breaks sound into small grains and reassembles them to create new textures, while wavetable synthesis uses pre-recorded waveforms to create evolving timbres. In addition to these fundamental techniques, sound synthesis often employs various modulation and effects processes to further shape and enhance sound. Modulation involves varying parameters such as pitch, amplitude, or filter cutoff over time to introduce movement within a sound. Common modulation sources include envelopes, which shape how sound evolves over time. Effects such as reverb, which adds echo-like reflections; delay, which repeats sound at specified intervals; and distortion, which introduces harmonics to alter the sound, can also be applied in synthesis to add depth and complexity.

As technology advances, sound synthesis tools are becoming more efficient and accessible (Schwarz, 2011). Modern software and hardware now offer enhanced features for precise sound design, including real-time processing and modular systems. These developments enable more detailed and practical applications of sound synthesis in music production and audio engineering. Improvements in algorithms and interfaces further enhance these tools, streamlining the process for artists, sound designers, and engineers.

Acoustic Mirrors

Until the 20th century, a variety of devices were developed for spatial orientation using sound, such as the large-scale ear trumpet and the topophone (Mayer, 1880), as well as smaller devices like the miniature locator (van der Voort & Aarts, 2009; figure 7). The topophone (figure 6), which was based on the principle of sound reflection, helped sailors identify the precise position of other ships around them (Helfer, 2009). With characteristics in common, *Acoustic Mirrors* (AM) are large monolithic concrete structures designed by William Sansome Tucker meant to detect approaching enemy aircraft through sound (BBC News, 2019). They were built during the 1920s and were used during and after World War I but were only viable for a short time, as the appearance of the first radars made them obsolete. Over the course of the last few decades, most of them have been demolished and vandalised, but a few can still be found scattered across the shores of Great Britain. *Acoustic Mirrors* vary in size and structural characteristics depending on their intended frequency range and the distance of the targets they were designed to detect. Nevertheless, they all share a key feature: a curved surface known as a paraboloid, which is engineered to reflect and focus sound waves. Incoming sound waves hit this curved surface and are concentrated at specific points in front of the structure, known as *focal points* (figure 8). At these *focal points*, microphones were installed to amplify the incoming sound waves, allowing operators to determine the direction of the sound source.

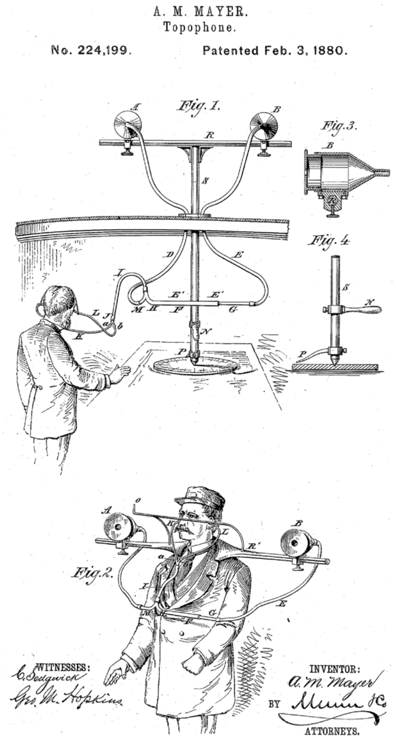


Figure 6.
Topophone (Mayer, 1880).

Figure 7.
Miniature locator
(van der Voort, 2009).

We found AM to have plenty of characteristics that could be explored. From their impactful structure (figure 9) to themetaphor around reflection and how everything converges to one point; from their historical context to their surface whose shape is perceived as either convex or concave depending on the perspective; from their passive functioning and reactivity to active and hostile agents. From their poetic proximity to the coast to the relation between sound and image, since there is detection before visual notice. In Chapter VI — *Installation*, we investigate how AM properties inspire and translate into the developed installation.

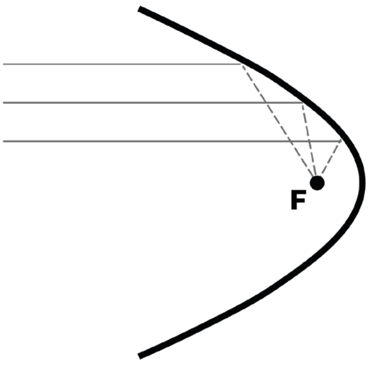


Figure 8.
Function scheme of a paraboloid acoustic mirror. F denotes the focal point (Helfer, 2009).

Figure 9.
Acoustic Mirror at Abbot's Cliff, built in 1928 (BBC News, 2019).

Chapter IV
State of the Art

This chapter aims to establish theoretical grounds by presenting relevant research about the themes of this dissertation. The selection of references takes into account historical relevance, personal interest and novelty.

The first section is dedicated to Sound Art (SA) and the early pioneering work in Sound Synthesis. In the second part of this chapter we explore sonification, visualisation and music notation, as these subjects intersect with our work. The third section introduces computational art and primarily showcases contemporary installations. We conclude by presenting systems that integrate algorithmic processes with creative elements.

Sound Art

The term SA arose from the need to name the sonic activities that take place outside auditoriums and music rooms (Castro, 2022). It is difficult to define the boundaries with other artistic areas since it often incorporates image, experimental music or conceptual art from the *Fluxus*[1] movement. Its first manifestations, in the 1910s and 1920s, were largely performative and are often presented as electro-acoustic explorations.

Over the decades, SA has been ambiguously associated with all sorts of sonic material. The term has been referenced in the work of different artists such as minimalist composers, classic experimentalists, *free jazz* musicians, experimental pop and rock groups, and so on. In this dissertation, we do not intend to catalogue and restrict SA to a specific context, but we believe that it distances itself from music on a conceptual level.

To start with, we can compare both subjects: A) While music searches for a linguistic connection, SA does not (Aldrich, 2003); B) SA tends to involve the audience with its sense of space whereas music does not; C) SA seeks a listener-listener relationship with the audience unlike music which seeks a player-listener relationship. We highlight the last topic as a mutual listening attitude that promotes reactive and interactive artefacts and installations.

Sound Art is not tied to a specific time, geographic location or group of artists. Luigi Russolo's early interest and investment in noise music paved the way for further conceptual experiments with sound. Despite Annea Lockwood's and Max Neuhaus' explorations in the early 60s, the term SA was first used in a MoMA exhibition in 1979. In the 90s, more festivals and exhibitions began to emerge, such as SoundArt '95 Internationale Klangkunst, Sounding Islands Sound Art Festival

1. An avant-garde art movement that emerged in the 1960s, *Fluxus* sought to break down traditional boundaries between art forms and everyday life. Emphasizing experimental and participatory practices, it challenged established norms through performances, happenings, and unconventional art objects.

and Six Exquisites International Sound Art Festival, among others. Over the years, its roots spread and invited more artists with diverse backgrounds to participate in the movement.

But how can we define a movement like SA, given that it encompasses a wide range of formats and presentations? Licht (2019) proposes two definitions of SA: A) "An installed sound environment that is defined by the physical and/or acoustic space it occupies rather than time and can be exhibited as any visual artwork would be"; B) "A visual artwork that has a sound-producing function, such as sound sculpture." Christian Marclay states that in SA diffusion, "Everything is so portable and so easy to share that you don't need an art institution to tell people what to listen to." (Schütze, 2013). We believe sound embodiment to be greatly influenced by physical context. In fact, why would the experience not be influenced by space if every other art genre shares a similar analogy? — Is photography, painting or cinema perceived equally if exhibited in a small or a large format? Nevertheless, when exhibiting SA works a few problems arise. Unlike primarily visual installations, where insight about scale and material positioning is more straightforward, sound installations are more challenging to set up. The visitors' circuit and movement must be specially planned (Schaeffer, 1966). According to Licht (2019), there is generally less knowledge about sound and museums and galleries often lack ideal acoustic settings, as seen in works like *Fans Corridor* (1970) and *Get Out of My Mind, Get Out of This Room* (1968). Finding the optimal placement for speakers, the right equalisation and volume (to prevent it from impacting other areas), and identifying frequencies that might compromise the experience are all regular concerns when planning a SA exhibition or installation. The sound installation artist must take responsibility for all of these features, those that are controlled and those left to chance, such as the behaviour of the auditor (Revill, 2016).

An ongoing criticism of SA installations is that they usually fail to connect emotionally with the audience. This criticism may also explain SA's perennial outsider status in the art world. For the purpose and context of this dissertation, we try to address this issue carefully to create enriching experiences. We proceed by presenting projects, research, and historically significant artists who challenge preconceived notions and attitudes toward sound and auditory experiences.

> *"Sound installations merge the eye and the ear, the haptic, the spatial and the temporal, dissolving the boundaries articulated by all other fine and musical arts."*
> — ROS BANDT (2006)

As a visionary composer, music theorist, philosopher and graphic artist John Cage was one of the leading figures of post-war avant-garde. A pioneer of electroacoustic music, he explored randomness, unconventional use of objects for sonic purposes, and non-standard approaches to musical instruments. One of his most important contributions is his “silent” piece *4’33”* (figure 10), composed after visiting an anechoic chamber and realising the impossibility of total sound abstraction. The composition consists of three parts of apparent silence, interrupted only by the opening and closing of the piano lid. It was deliberately premiered at the Maverick Concert Hall, an open room surrounded by nature, in Woodstock, New York. Cage sought to create an intense experience fostered by incidental and discreet sounds of the immediate surroundings that usually go unnoticed, thus transcending standard musical composition. With this composition he established “a separation between music and what would eventually become the environmental quality of sound installation” (Licht, 2019).

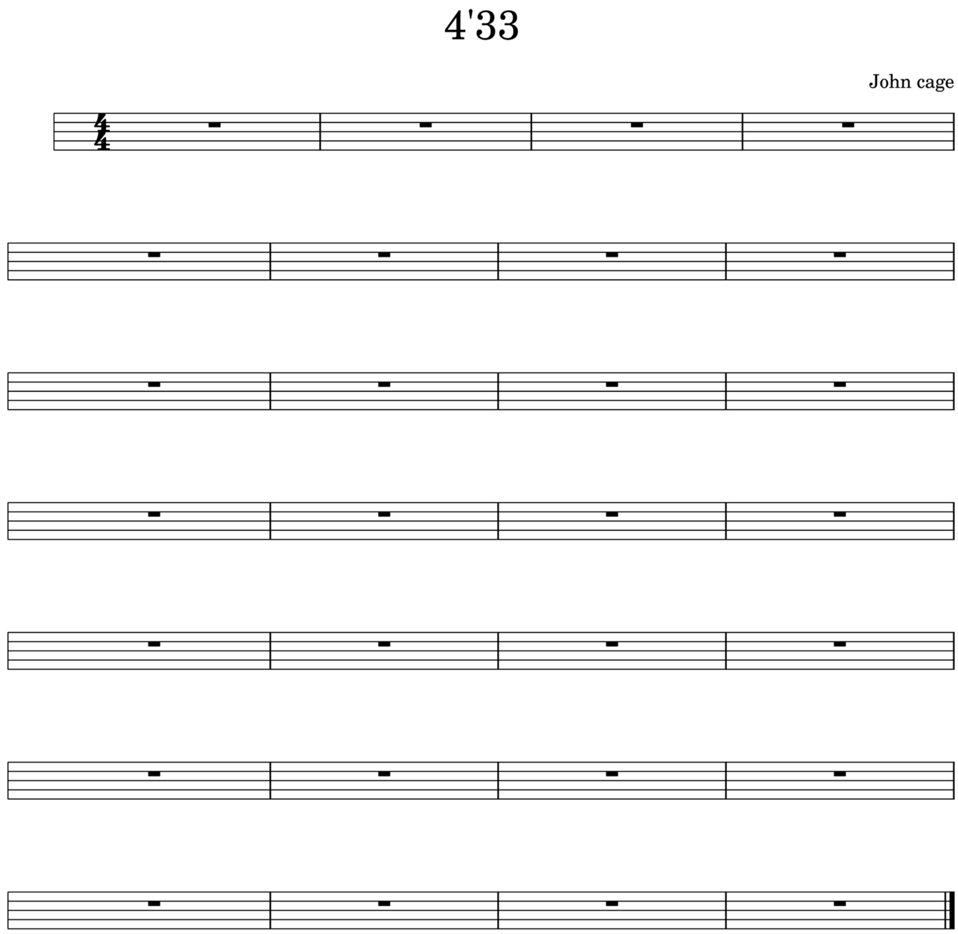


Figure 10.
4’33” musical sheet (Cage, 1952).

In his most famous work, *I Am Sitting in a Room* (1968), Alvin Lucier, an explorer of acoustic resonance and sound perception, records himself speaking a short text and then repeatedly plays it back in the same room. Each iteration amplifies certain frequencies while diminishing others due to the room's acoustics, resulting in a complex, resonant soundscape made up of pure tones that reflect the room's unique acoustic characteristics (Minkin, 2018). Similarly, the installation *Respiro* employs a detailed exploration of sound frequencies. In its first stage, *Respiro* used cymbals stimulated by transducers[2] to produce close-range sounds, while the second stage involved amplifying these vibrations with speakers and contact microphones (Tudela & Carvalhais, 2022). Although we initially envisioned a two-room partition for S—I, similar to *Respiro*, logistical constraints prevented this setup. Both projects illustrate a profound interest in how acoustic environments shape and transform sound, highlighting a exploration of frequencies and their perceptual effects across different periods and technological approaches.

Luigi Russolo was an Italian Futurist artist who worked as a painter, composer, and builder of instruments. In his manifesto from 1913, *The Art of Noise*, Russolo (1967) he argued that the traditional orchestra, with its reliance on conventional musical instruments, was limited and outdated. He proposed that the modern world needed a new palette of sounds, including those produced by machines and everyday life. To achieve this, he invented new instruments called *Intonarumori* (figure 11), which could produce a variety of noises. These instruments were designed to mimic industrial sounds, such as hisses, buzzes, and roars.

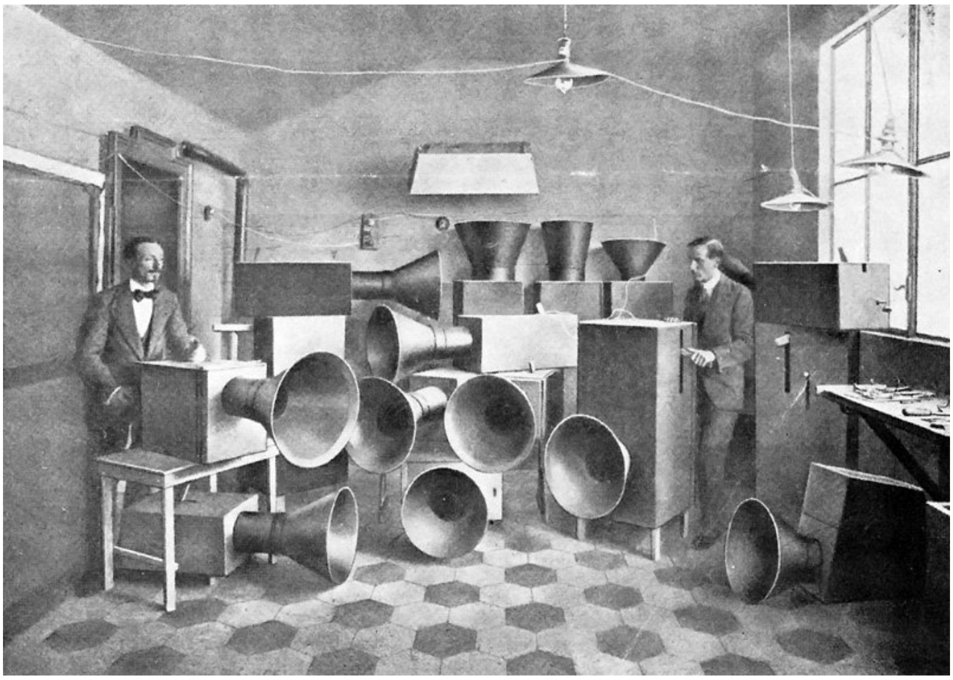


Figure 11.
Intonarumori, Luigi Russolo (left) and his assistant Ugo Piatti (Russolo, 1967).

The artist Tarek Atoui brings the same intentionality to contemporary times by seeking new sounds and instruments. He explores how sounds can circulate through various materials (such as stone, wood, metal) and by various means (through water, through air).

2. Devices that convert one form of energy into another, commonly used in sound equipment to convert sound waves into electrical signals or vice versa.

The exhibition context after the creation of the artefacts allows for both *The Whisperers* (Galerie Chantal Crousel, 2021), a project in which Atoui exhibits part of his work (figure 12), and *Intonarumori* to establish a connection with the audience. These exhibitions create spaces for experimentation and exploration of unique sounds, influenced by the interactions and choices of the participants. In fact, Kester (2011) writes about the increasingly common characteristic of SA: emphasising direct engagement rather than reflective interpretation.

> *"Sound installations invite the auditor to move around and through the space to consider the different behaviours and perspectives of the sound installed."*
> — ROS BANDT (2006)

Figure 12.
The Whisperers, Atoui
(Galerie Chantal Crousel, 2021).

While Luigi Russolo was developing his ideas about noise music, a group of Russian innovators, including Evgeny Sholpo and Arseny Avraamov, were contemporaneously exploring the frontiers of sound synthesis and music theory. Driven by a desire to revolutionise the musical landscape, these figures emphasised the need for a scientific analysis of music phenomena with regard to creation, performance, and perception, integrating scientific principles into their work (Sholpo, 2001 as cited in Smirnov, 2011).

Sholpo envisioned a *mechanical orchestra*, a machine designed to synthesise complex sound spectra and automatically transcribe them onto a music score. This early concept of a sound synthesiser was notable for its novelty in using sine wave oscillators[3], which formed a discrete scale covering the entire audible range. This unprecedented control over sound at a spectral level allowed for distinct interval pitches that were imperceptible to the human ear. In 1927, Sholpo built a device called the *Melograph*, designed to capture and record the rhythmic nuances of musical performances, enabling more detailed analysis and reproduction of musical expression. Meanwhile, Avraamov was working on blending traditional musical elements with the sounds of the modern world by manipulating pianos and harmoniums through special retuning and incorporating noise instruments and symphony orchestras (Smirnov, 2013).

This pioneering research by Sholpo and his contemporaries laid the groundwork for the development of later sound synthesis technologies like Murzin's ANS, in 1957 (figure 13). This instrument used an optic disc to generate 720 simultaneous sinusoidal sounds, representing a significant advance in sound synthesis, allowing composers to manipulate the entire sound spectrum rather than just individual sound waves, thereby setting the stage for future innovations in the field (Smirnov, 2011).

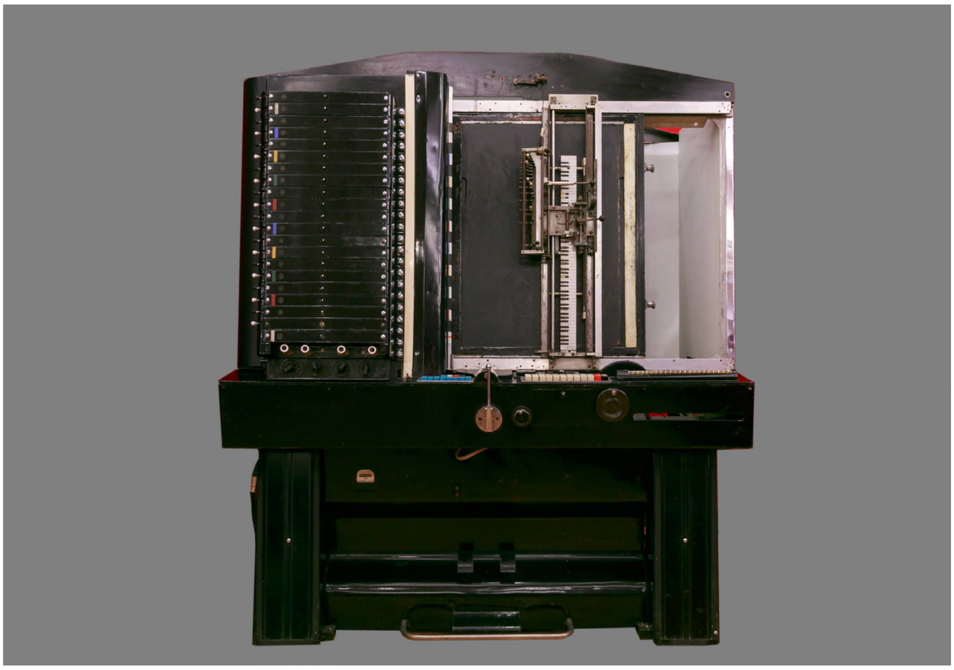


Figure 13.
ANS Synthesiser, Murzin, 1957
(Russian Cultural Heritage, n.d.).

Building on these innovations, the concept of algorithmic composition began to be more concrete, exploring how mathematical principles could drive the creation and evolution of sound. This transition from Sholpo's early mechanical orchestras to contemporary algorithmic practices illustrates a broader shift in how sound artists approached the generation and manipulation of musical material.

3. Oscillators are electronic circuits or devices that generate periodic waveforms, such as sine, square, or triangle waves. They are fundamental components in sound synthesis, used to produce continuous tones by creating vibrations at specific frequencies.

Young's and Zazeela's *Dream House* (1969; figures 14.A & 14.B), originally envisioned as a continuous performance by numerous musicians, instead used electronic equipment to create an immersive sound environment (Young, 2018). By employing sine wave oscillators to produce a continuous field of precisely tuned tones, Young's work concretely applied algorithmic principles such as: A) systematic variation, which involves altering components like pitch and rhythm based on mathematical criteria or sequences; and B) harmonic structuring, the use of mathematical methods to generate chord progressions. This approach reflects the foundational techniques in sound synthesis developed by Sholpo and his contemporaries.

Figure 14.A.
Dream House, New York (Zazeela, 1993).

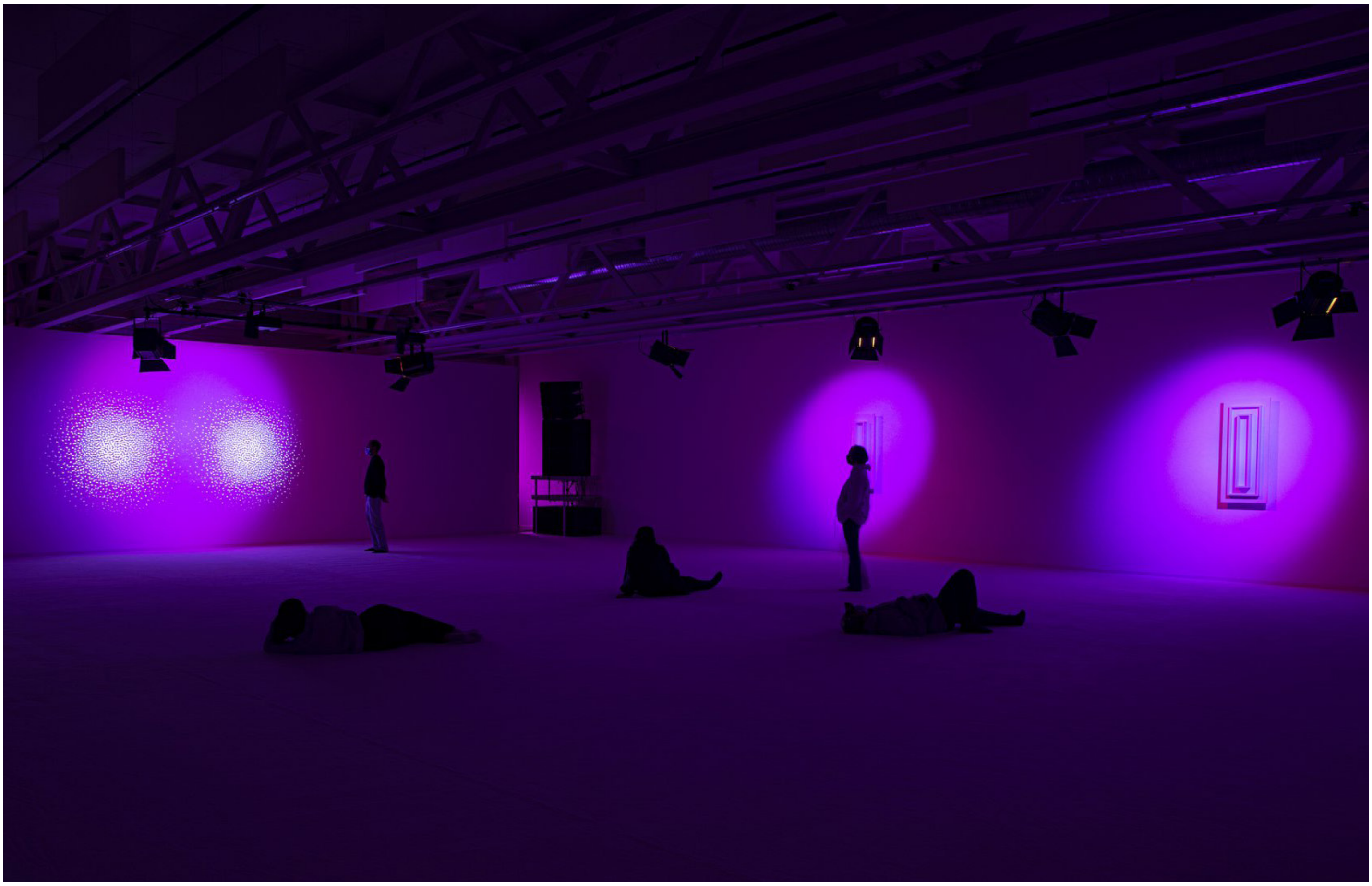


Figure 14.B.
Dream House: Sound and Light Environment, Lausanne, Suisse (Choi, 2021).

Similarly, Christina Kubisch's *Clocktower Project* (1996) harnessed environmental data to algorithmically influence sound, recontextualizing the bells in the clocktower beyond their original timekeeping function. This project exemplifies how algorithmic composition can integrate external variables, like weather conditions, to shape and transform auditory experiences, further expanding the scope of how mathematical principles can interact with sound (Cox, 2021).

We proceed by presenting Dick Higgins, an artist who inspired our work in various ways. A proponent of using computers as a tool for art making, he embraced randomness as an artistic potential. Through sound, painting and print artworks (primarily using lithography and silkscreen) he reflected on politics, nature, and music notation (figure 15).

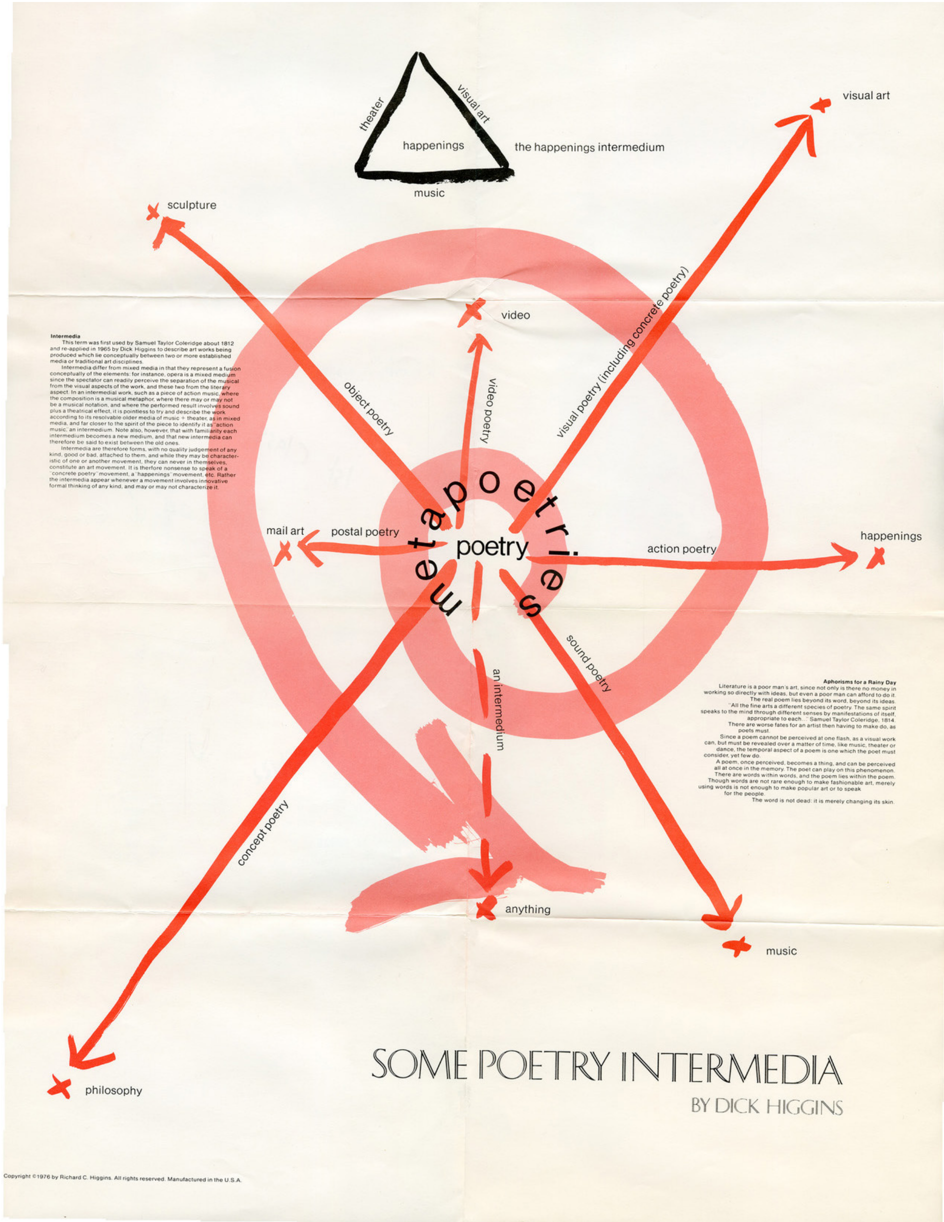


Figure 15.
Some Poetry Intermedia (1976) (Higgins, 2018).

Following a line of visual inspiration and outside the scope of SA, we were also influenced by the artworks of three fine artists during the development process of the visual outputs for S—I and also in our preliminary work. *Rythme / Rhythm* (1939, figure 16.A) by Nicolaas Warb, the collages by Olga Rozanova (1916, figure 16.B), and *Homage to the Square* (1968, figure 16.C) by Josef Albers are abstract compositions that share a focus on depicting non-representational forms as they explore colour and geometry.

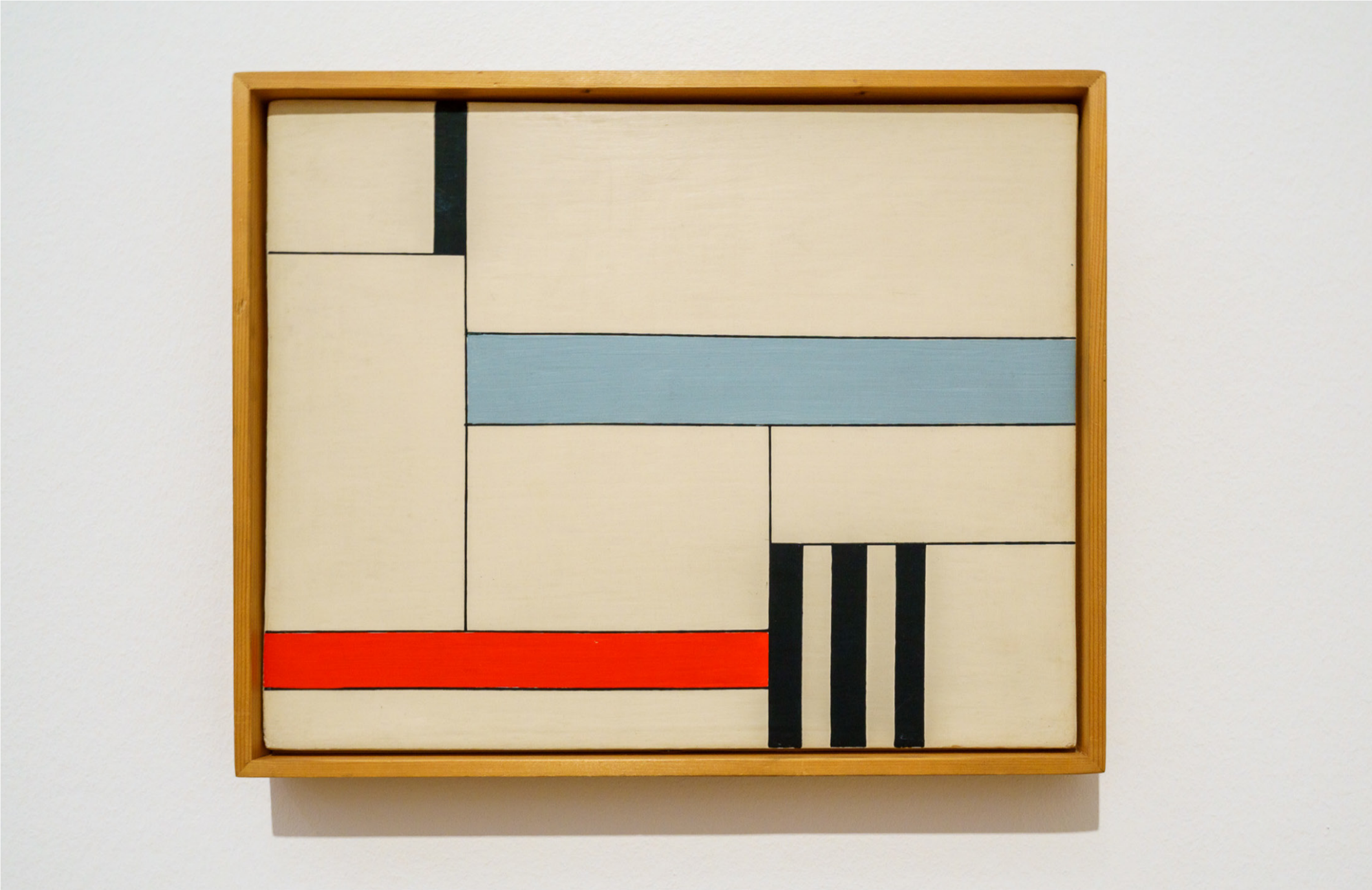


Figure 16.a.
Rythme / Rhytmn (Warb, 1939).

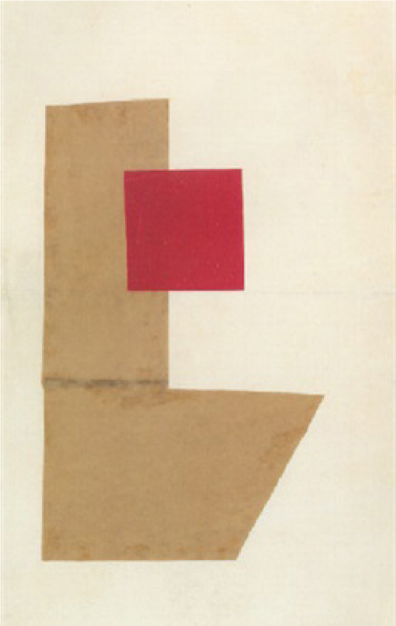


Figure 16.b.
Composition, 1918–1918 (Rozanova, 1918).

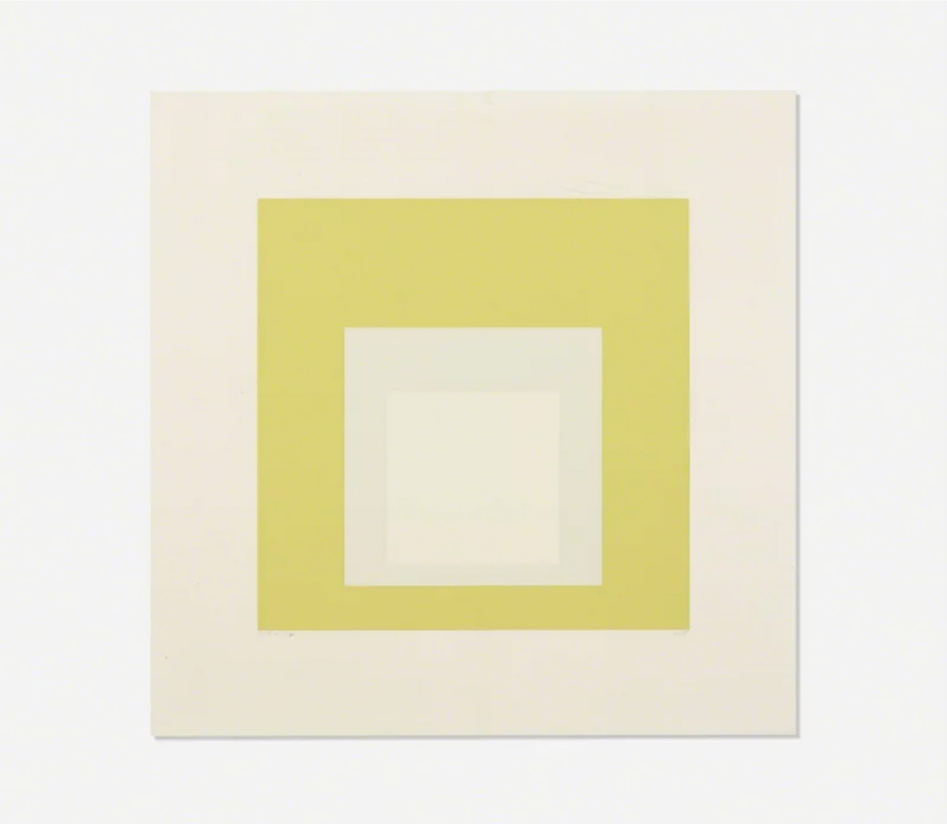


Figure 16.c.
Homage to the Square (Albers, 1968).

Cross-Media Portrayals

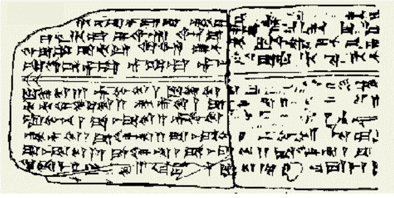


Figure 17.
Hymn to Nikkal inscription, 1300 B.C. (Bucellati, 2003).

Since the Hurrian *Hymn to Nikkal* (figure 17), the earliest music inscription registered, dating from the 13th century B.C. (Buccellati, 2003), music notation has adopted many forms and interpretations throughout musical eras and tendencies. For the purpose of this dissertation, we focus on abstract, experimental, and visually differentiated specimens rather than concrete and classic examples.

Iánnis Xenákis was an avant-garde composer, mathematician, engineer, light sculptor, and architect. His military experience during World War II is reflected in his compositions which exploit dynamics and a broad range of frequencies. He was a pioneer in granular synthesis and stochastic music, and he had a deep interest in creating self-made sounds. His first stochastic composition, *Pithoprakta* (1955), is inspired by crowd behaviour where one can perceive the aggregate but not follow individual movements (Ghez, 2022). The music notation in *Pithoprakta* suggests sound outcomes and results from thoughtful, arbitrary encodings (figure 18).

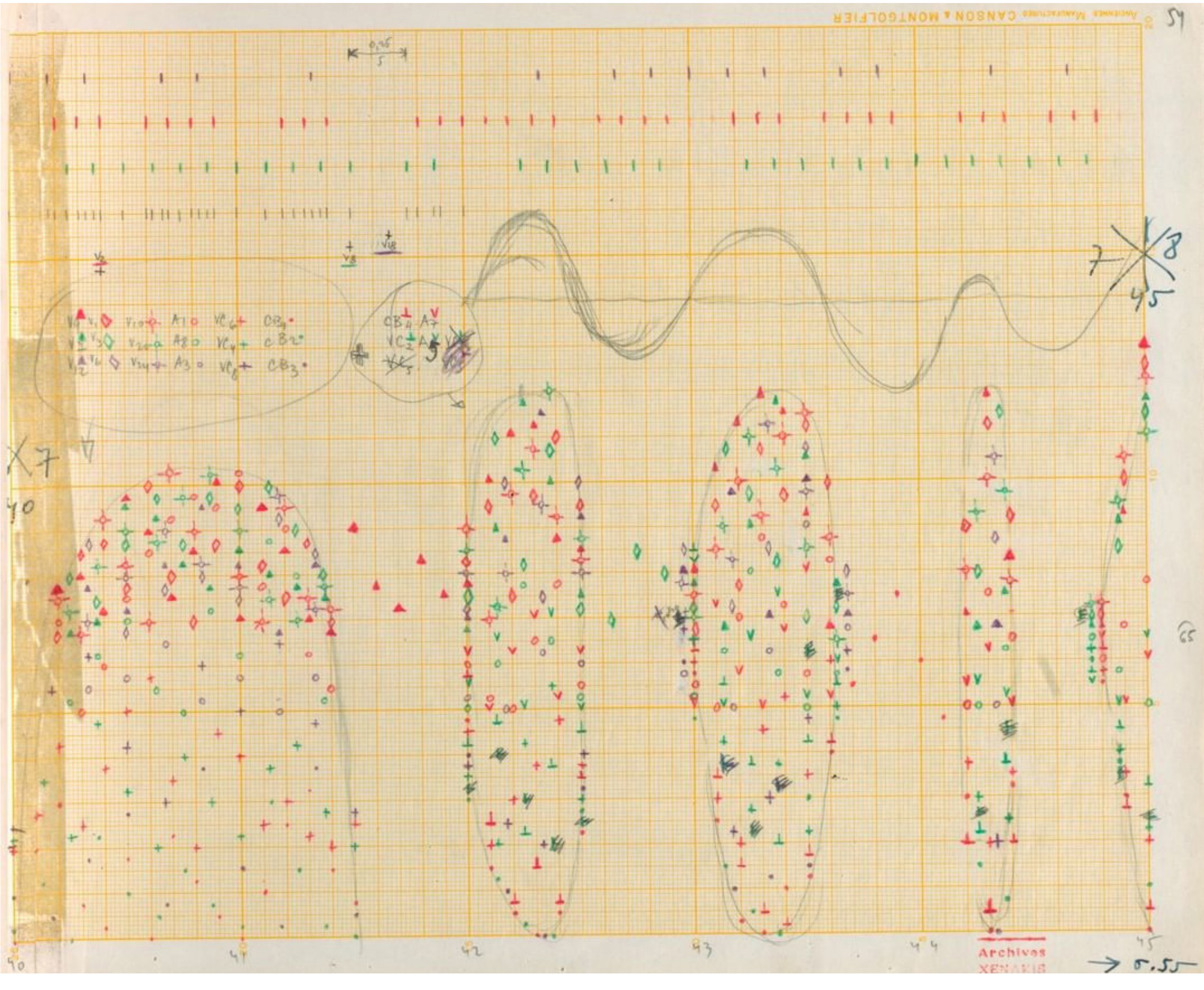


Figure 18.
Studies on *Pithoprakta* (Xenákis, 1955).

> *"It's not just about making music and sound but also about transforming Mankind."*
> — iánnis xenákis (Ghez, 2022)

Xenákis looked at music as architecture in motion and he believed that they can feed from each other as he often resorted to the intersection of both areas. For example, many of the sound solutions for *Metastasis* (1954) come from architectural ideas conceived for *Sainte Marie de La Tourette* (figure 19.a), and parts of the monument are influenced by some of the music sections (Ghez, 2022). He created a space where the application of physics, statistics, mathematics, and random processes plays an autonomous role in musical arrangement and composition. However, "the theory of probabilities does not exist to explain his work but to contaminate it" (Manzolli, 2022). In 1956 he was hired by Le Corbusier, a painter, sculptor, and influential modernist architect, to create the revolutionary *Philips Pavilion* (figure 19.b). The pavilion was envisioned to be a structure that combined art, technology, and architecture beyond its pragmatic purpose of displaying artwork (Spatial Music Forum, 2021). Nowadays, it stands as a landmark of architecture. Corbusier's admiration for music led him to search for mathematical rationing that could bring forth the beauty found in sound into the realm of vision. For both him and Buckminster Fuller, architect and philosopher, music and poetry were used to justify not only the systems they proposed but also to justify themselves "beyond the sphere of simple principles of architecture and design" (Calovski, 2022). We highlight Buckminster Fuller's museum *Biosphere* (figure 19.c) as its futuristic nature relates to am.

Figure 19.a.
Sainte Marie de La Tourette, Xenákis & Corbusier, 1960 (Ghez, 2022).

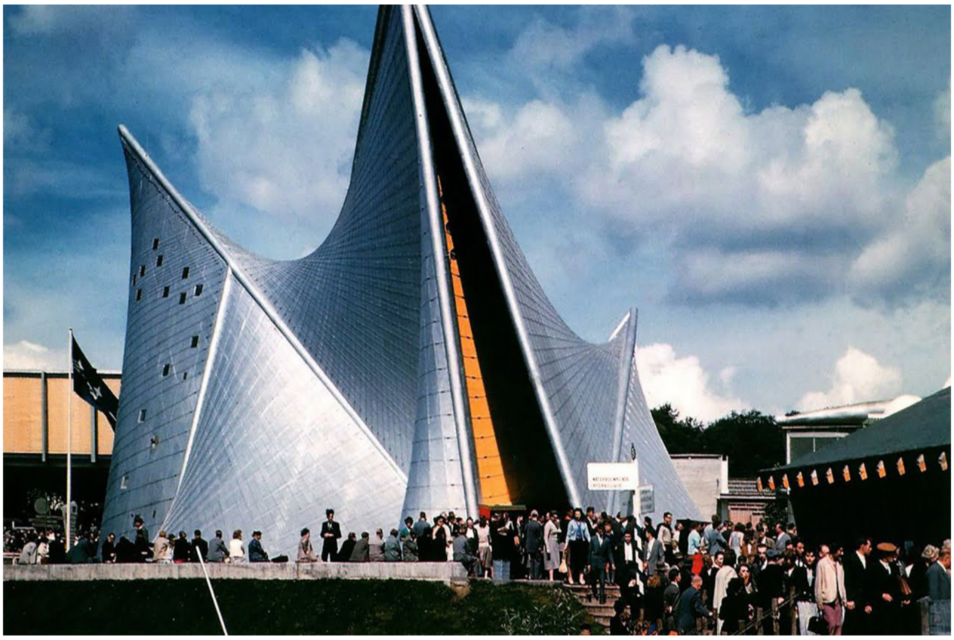


Figure 19.b.
Philips Pavilion, Xénaquis & Corbusier, 1958 (Spatial Music Forum, 2021).

Figure 19.c.
Biosphere, Fuller, 1967 (Calovski, 2022).

John Cage explored music composition from two distinct approaches. He created music scores that cannot be straightforwardly played, and he used unconventional graphic material as music scores to create new compositions. To compose *Atlas Eclipticalis* (1961; figure 20.A), he used an atlas of the stars made by Antonín Becvár (John Cage Trust, 2016) and superimposed musical staves on the star charts. On the other hand, *Fontana Mix* (1958, figure 20.B) is a piece composed of ten pages of paper and twelve transparencies of original graphical notation of music. The score compresses letters from various languages beyond all sorts of shapes. When reading *Fontana Mix*, the performer is not only a musician who makes musical elements sound but also a composer who selects and arranges these musical elements in the first place, based on predefined rules created by Cage. *Fontana Mix*, featuring electronic sounds, is often paired with *Aria* (Sauer, 2009), one of Cage's vocal works, resulting in a complex and layered presentation. The graphic score of *Aria* (figure 20.C) comprises pages, each one intended to last 30 seconds thus providing guidance to the vocalist. In this score, lines indicate pitch, shapes denote intensity, and colours signify timbre. This attitude withdraws the performer from the actual performance, creating a freer expression environment with space for wider interpretation possibilities.

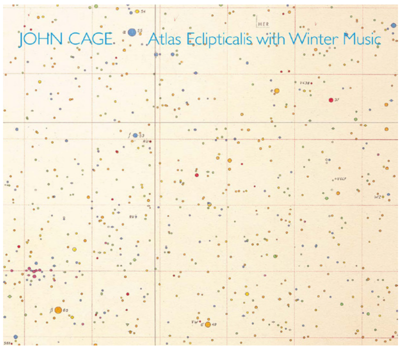


Figure 20.A.
Atlas Eclipticalis CD cover (Cage, 1961).

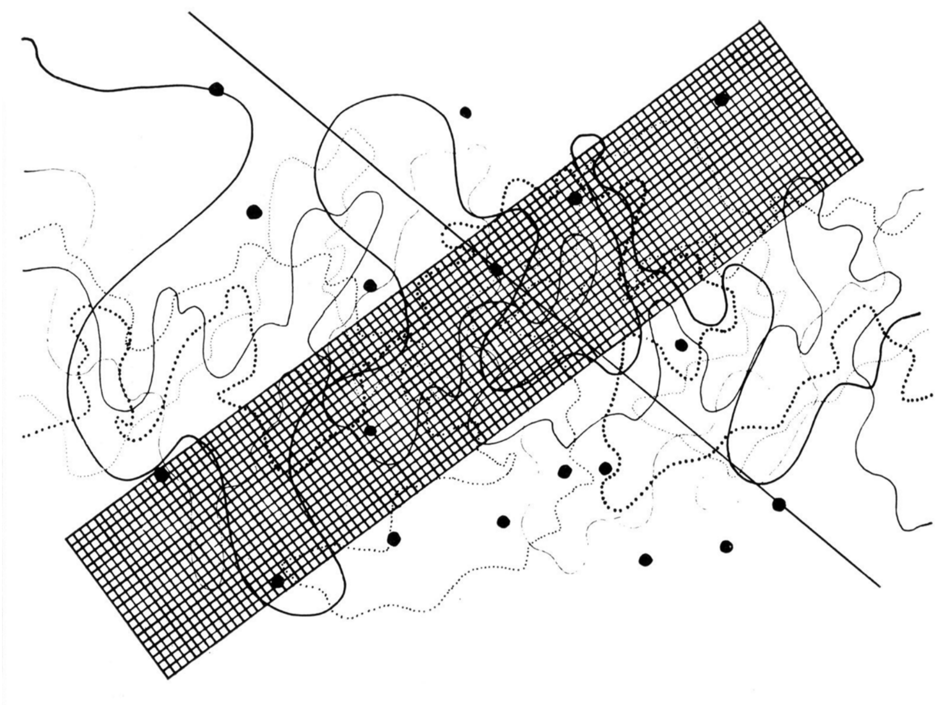


Figure 20.B.
Fontana Mix, Cage, 1958 (Reichardt, 1968).

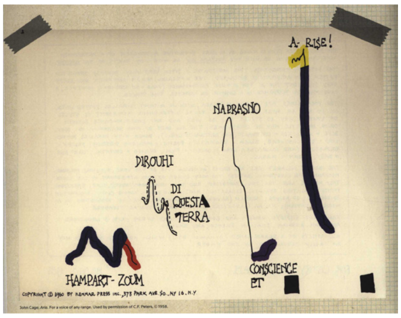


Figure 20.C.
Aria, Cage, 1958 (Sauer, 2009).

Treatise (1967) is Cornelius Cardew's most renowned work. It is a document with 193 pages of lines, symbols, geometric glyphs, abstract shapes, and no explicit instructions on how to musically interpret it (Dennis, 1991). On the bottom of every page, there are two musical staves that contextualise the graphic elements as a music score and not only purely artistic in character (figure 21.A). His contemporary, Raymond Murray Schafer's vast works also extend to music notation. We highlight *Snowforms* (1986; figure 21.B), an 8-minute hummed chorale for children's voices visually composed by the overuse of glissandos, which arguably gives a sense of peaceful quietness. It began as a series of sketches of snowdrifts and turned into a musical score when Schafer traced a pentagram over them (Vancouver Chamber Choir, 2013).

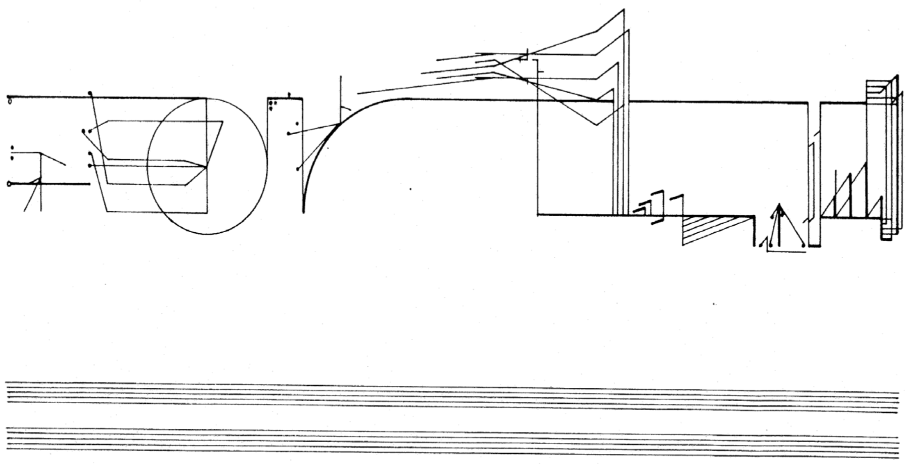


Figure 21.A.
Page 49 of *Treatise* (Cardew, 1967).

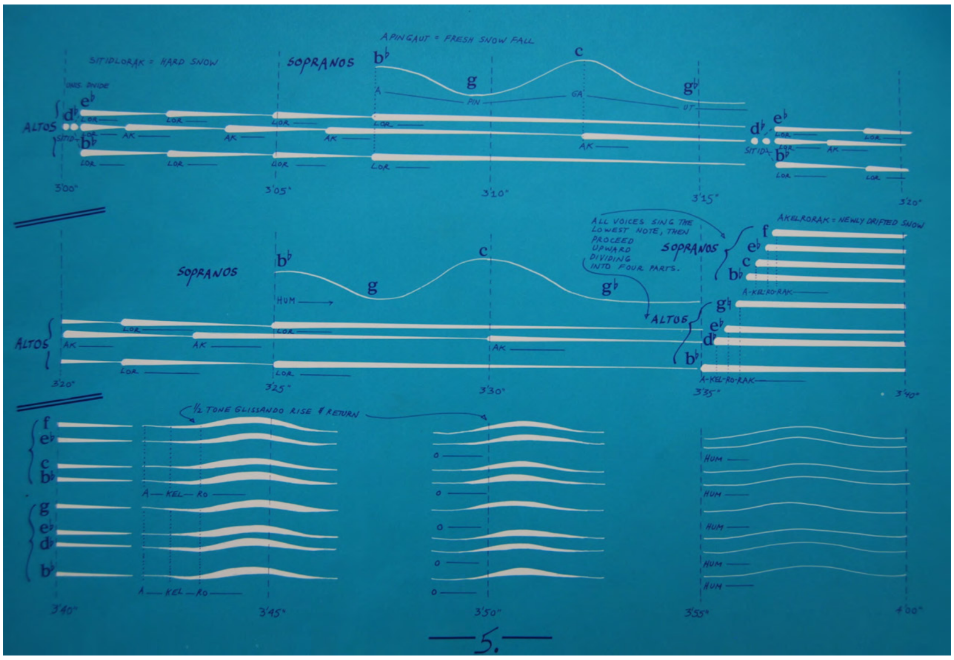


Figure 21.B.
Snowforms (Schaefer, 1986).

In the final examples of music notation we present worksby Karlheinz Stockhausen and Brian Eno. Figure 22 shows Stockhausen's method of music notation in which each note is precisely detailed, limiting the performer's freedom and allowing the composer to control random elements. Eno states that the nature of his notation enables the exploration of sound textures, unlike classic notation, since it was invented when sound texture possibilities were limited (Hope, n.d.).

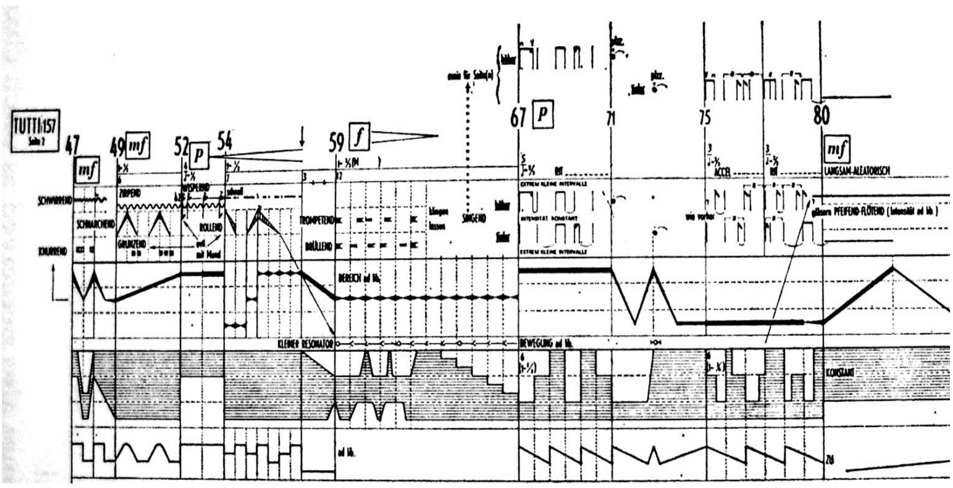


Figure 22.
Tutti 157 (Stockhausen, 1959).

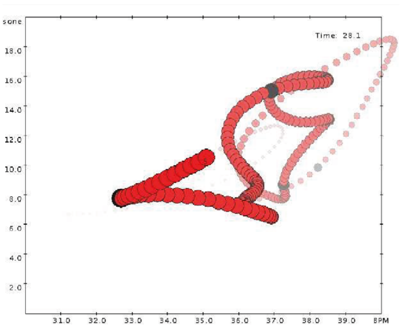


Figure 23.
Snapshot of the *Performance Worm* when analysing four bars. Horizontal axis: tempo in bpm; vertical axis: loudness; fainter points are older registers (Widmer & Zanon, 2004).

Approaching from the standpoint of visualisation studies, Widmer & Zanon (2004) aims to distinguish pianists by analysing their performances using machine learning techniques. His system generates analytic representations of recordings (figure 23), where the performative characteristics of the artist are interpreted based on fundamental sound components such as tempo, loudness and crescendo.

From a less analytical perspective, Sá (2017) developed a 3D system that processes sound and image based on frequency analysis from a string instrument. The visual results arise from the superimposition of several graphical scores created for a set of musical pieces. Consequently, they no longer aim to provide instructions. Instead, their use "evokes a personal creative continuum which traverses different approaches and processes" (Carvalhais, 2022). Sá (2016) contrasts two major opposing audiovisual approaches, which she categorises as the "one-to-one approach" and the "structural approach." The former method involves audiovisual relationships where all sonic and visual events are synchronised, while in the latter, synchronisation does not occur for at least part of the sounds and images. We began our preliminary work by exploring these two approaches in a static environment.

To conclude this section, we will review some examples of sonification, a relevant topic for the encoding component in S—I. Sonification, defined as "the use of nonspeech audio to convey information" (Hermann, Hunt, & Neuhoff, 2011), is a powerful method for data representation that maps data attributes to sound parameters for improved user comprehension. Its benefits include leveraging inherent auditory pattern recognition, facilitating rapid data interpretation and assisting visually impaired individuals.

Two Trains (Foo, 2014) is a music composition that emulates a ride on the New York Subway through three boroughs: the Bronx, Brooklyn and Manhattan. The number of instruments and the dynamics of the piece correspond to the median household income in each location, revealing economic inequality across the city while capturing the energy and chaos of the subway system. Brian House interprets a year's worth of location-tracking data to create a recording that sonifies every moment of his daily routine. The 11-minute piece, *Quotidian Record* (House, 2012), suggests that habitual patterns possess inherent musical qualities that may form an "emergent portrait of an individual". While the first two examples use data from fixed datasets, the system described in *Sonic Kayaks* (Griffiths et al., 2017) employs real-time data to inform kayak paddlers about water temperature and underwater surroundings, mapping marine world data to a generative live composition as they navigate. As demonstrated in the previous example, sonification works can also take on a physical format. Both *Rain Beats* (Garcia, 2022; figure 24.A) and *Soundscape4DEI* (Neves et al., 2023; figure 24.B) are physical devices that reveal and hint at their purpose. The former is a musical machine that uses meteorological data from across the globe, while the latter involves the sonification of data from the daily routine of CISUC[4].

4. Centre for Informatics and Systems of the University of Coimbra

Figure 24.a.
Rain Beats (Garcia, 2022).

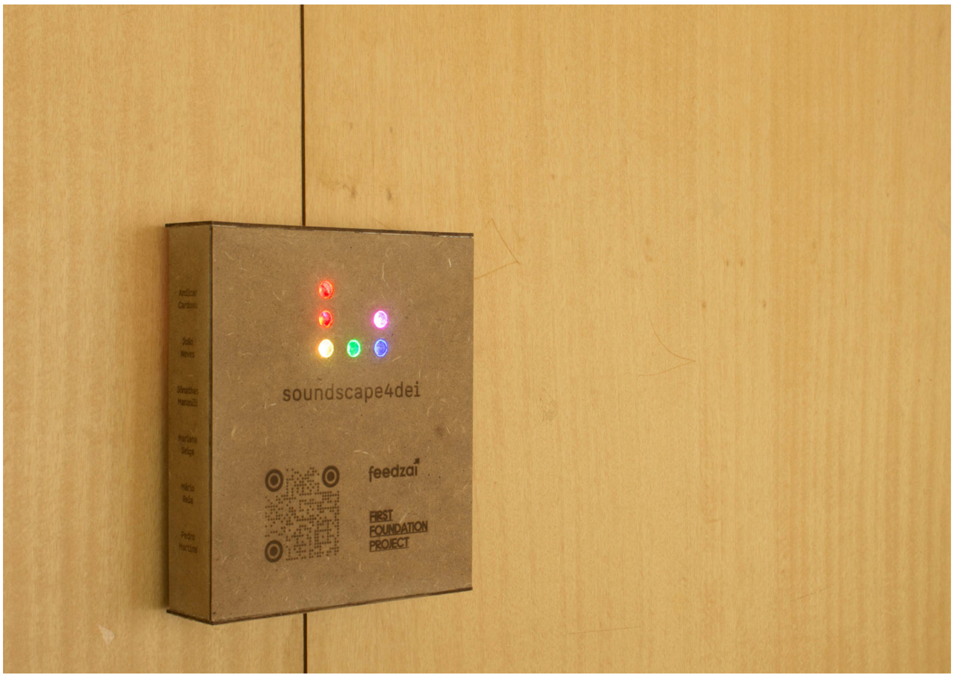


Figure 24.b.
Soundscape4DEI (Neves et al., 2023).

Computational and New Media Art

One of the earliest outcomes of the relationship between sound and image began to be digitally explored in a cinematographic context. Mary Elle Bute (Basquin, 2020), Jordan Belson (Balme, 2020) and Oskar Fischinger (Gardiner, 2007) are some of the animators whose work is centred on abstract films that connected both media. Since John Whitney experiments in motion and generative graphics (figure 25), technology has come a long way on skewing these relationship barriers (Moritz, 2019). Nowadays, as sound and image are able to influence each other in real time, it is possible for the experience to be not merely expository but also interactive.

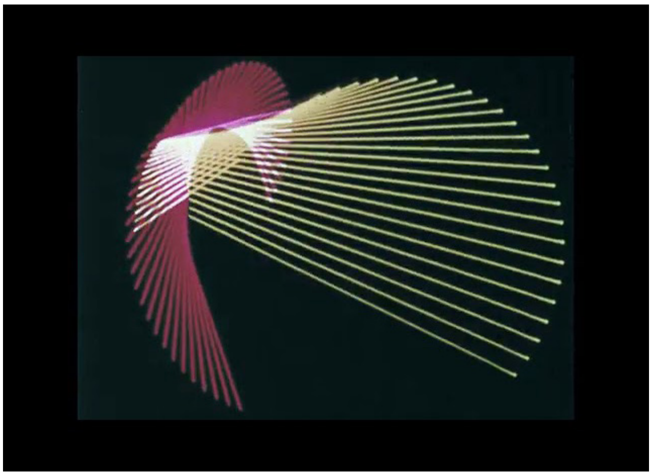


Figure 25.
Experiments in motion graphics (Whitney, 1968).

According to Manovich (2002), all new media objects possess the following characteristics: A) numerical representation, meaning they are composed of digital code and their functions can be described mathematically; B) modularity, as they are represented as collections of discrete elements (such as pixels, polygons, voxels, characters); C) a degree of automation, which enables them to operate independently to some extent; D) variability, allowing them to be presented in multiple versions; and E) transcoding ability, enabling them to be transformed into different formats. When referring to New Media (NM) we are not alluding to the traditional forms of media such as television, radio, or early digital platforms. Instead, the focus is on contemporary NM, characterised by its integration into installations and modern artworks that explore immersive and sensorial experiences. These works often push the boundaries of sensory perception and challenge conventional notions of art, using cutting-edge technology to create dynamic, multi-sensory experiences that redefine how we interact with and interpret media and go beyond the passive consumption of information. Nevertheless, they share key characteristics with their predecessors.

Arguably the most significant cross-media artist for the context of this dissertation, Ryoji Ikeda has consistently engaged with sound and image in both performative and exhibition contexts throughout his career. His work (figure 26) is characterised by a distinct aesthetic, predominantly using a black-and-white palette to create chaotic yet meticulously crafted minimalist audiovisual landscapes rich in patterns.

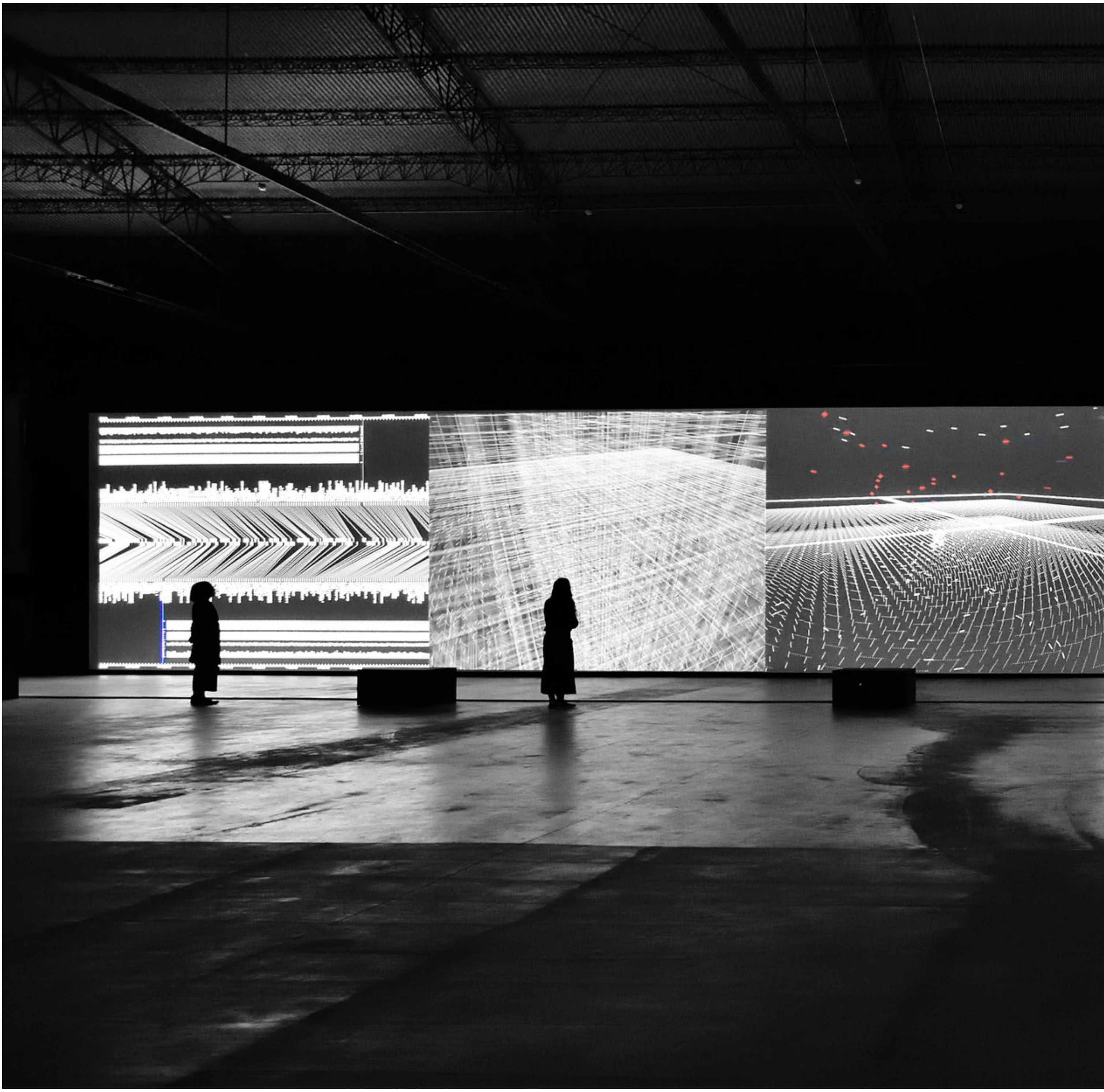


Figure 26.
Data-flux (Ikeda, 2006).

Ikeda has had a significant influence on the NM domain. The following projects present similar approaches, showcasing a minimalist, and sensory approach. In *Aurora Cerebralis* (Rush, 2019), generative images of cerebral activity and neural connections form the basis of an immersive audiovisual experience (figure 27.A). *VAST* (Synthestruct, 2022) uses the mathematics of sound to drive its visual elements (figure 27.B). *Pladis: Data Universe* (Anadol, 2018) manipulates light to alter the viewer's perception and awareness (figure 27.C). *Wayfarer* by Weidi (2023) employs AI to generate graphics that depict Tiangou, a Chinese mythical creature (figure 27.D). Despite the unique approaches of each project, they share notable similarities. All are housed in expansive black box environments, allowing for extensive movement and placing the audience in an immersive and vulnerable position due to the overwhelming scale of the projections and the intensity of the sound. In the development of S—I, we focused on addressing the challenge of achieving immersiveness despite the inaccessibility of such physical conditions and technology.

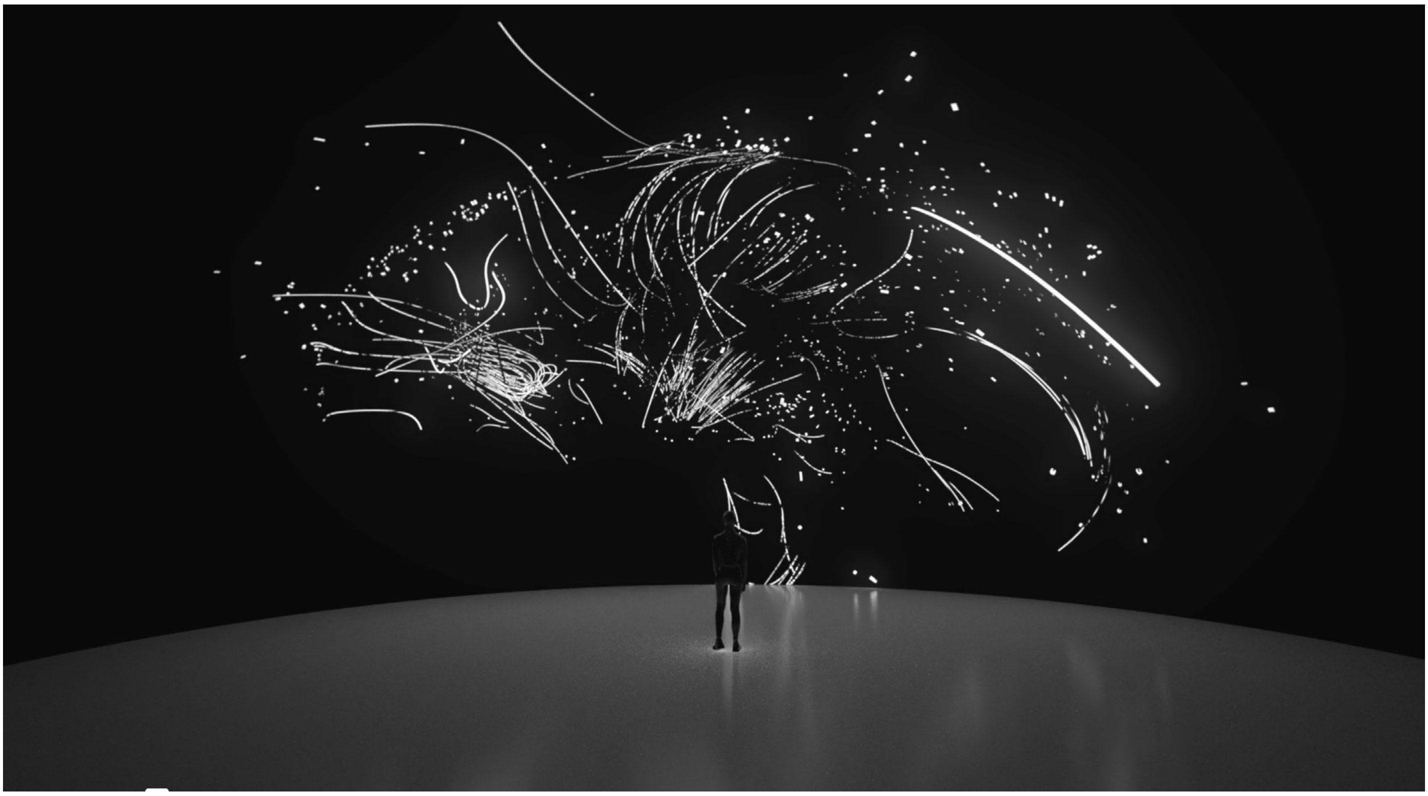


Figure 27.A.
Aurora Cerebralis (Rush, 2019).

Figure 27.B.
VAST (Synthestruct, 2022).

Figure 27.c.
Pladis: Data Universe (Anadol, 2018).

Figure 27.d.
Wayfarer (Weidi, 2023).

Along with the previously presented artists whose artistic language influenced our work, we continue by presenting additional installations that further demonstrate these influences. In *Found in Translation* (TheGreenEylr, 2021), the system generates visuals and sound based on the audience's speech prompts and its corresponding machine translation (figure 28.a). These minimalist visualisations, composed of overlapping lines and typography, are displayed asynchronously across 24 screen panels, each representing a different language family. *Living Room* (Random International, 2023) uses spotlights and their absence to draw paths of the audience (figure 28.b). In a separate room those results are displayed in small screens along with notes of the development of the project (figure 28.c).

Figure 28.A.
Found in Translation (TheGreenEylr, 2021).

Figure 28.C.
Living Room process notes
(Random International, 2023).

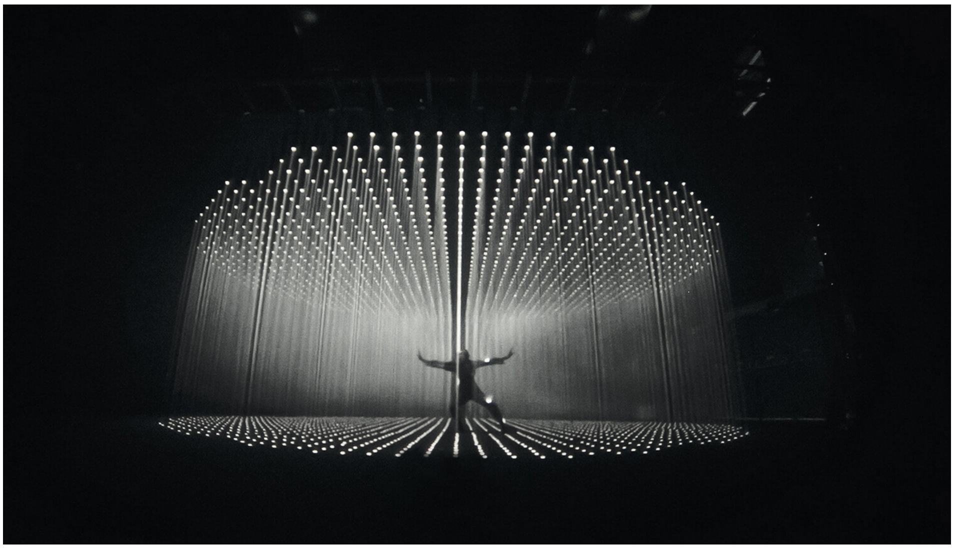


Figure 28.B.
Living Room (Random International, 2023).

To conclude this section, we showcase two more installations that promote a meditative engagement and that employ physical concave walls as their surroundings, resembling some of the AM variants. *HavObservatoriet* (Autogena, 2023) uses a panoramic screen to display real time simulation of ocean waves (figure 29.A). It is a permanent installation situated in Vejle Klimapark, Denmark, differing from the previous examples as it is set in an outdoor environment. *Static Balance* (Nicolai, 2007) invites the audience to explore spatiality by having parabolic mirrors to reflect sound, creating areas with higher and lower concentration of acoustic information (figure 29.B). It is a work from Carsten Nicolai, also known as Alva Noto, a German artist and musician who works in the transitional area between music, art and science. He describes his work as an attempt to "overcome the separation of human sensory perceptions by making scientific phenomena like sound and light frequencies perceivable to both eyes and ears", with several other artworks that are relevant to this dissertation.

Figure 29.a.
HavObservatoriet (Autogena, 2023).

Figure 29.b.
Static Balance (Nicolai, 2007).

Algorithmic Processes and Creativity

Throughout the last decades we have been engaging with increasingly more capable computational systems. From the simplest to the most complex ones, the act of computation remains the same. *Cent Mille Milliards de Poèmes* (Queneau, 1961) is an unconventional book that contains ten poems with fourteen verses each. Every page is cut in strips, which enables the verses to be reorganised (figure 30). When the reader creates new poems within this search space by adopting either thoughtful or random approaches, they are performing a task that not only relates to but represents the computational process.

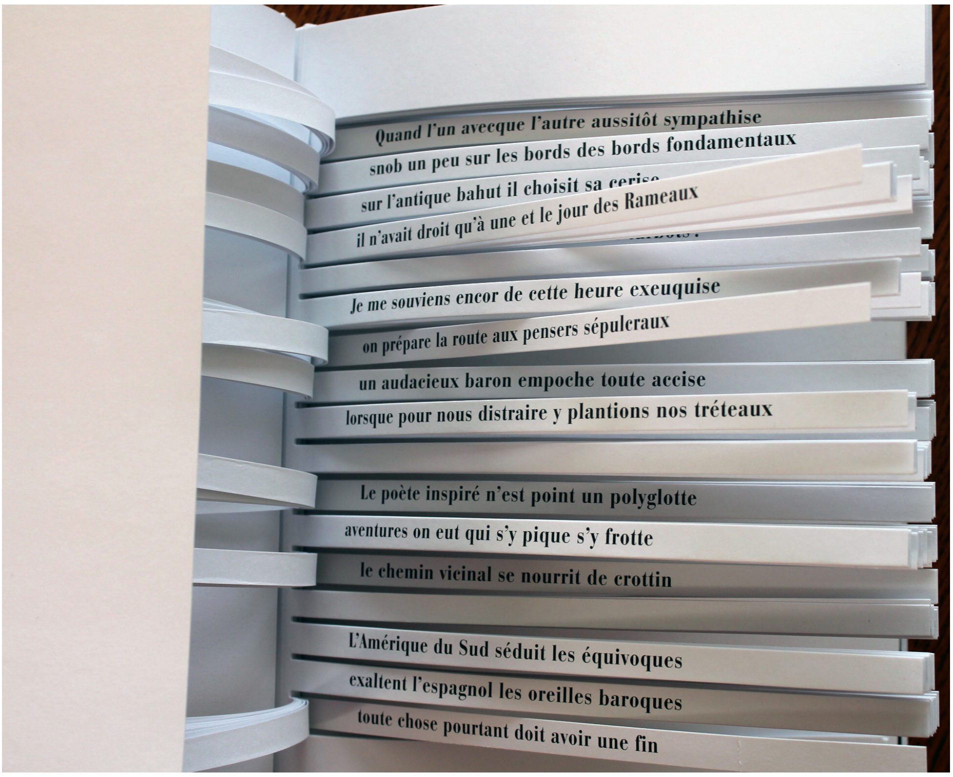


Figure 30.
Cent Mille Milliards de Poèmes (Queneau, 1961).

We proceed by sharing more examples of works that integrate the human element into non-computational algorithmic processes. Duchamp anticipates for nearly forty years Cage's and Xenáki's thinking of how the laws of chance might apply to music. In *Erratum Musical* (1913), Duchamp wrote a musical for three voices, one for each of his two sisters and one for himself. Each part was written separately and was inscribed with their names: Magdelaine, Yvonne, and Marcel. He then made three sets of 25 cards—one card per note, one set per voice. He shuffled each set of cards in a hat and withdrew the cards one by one. By writing down the outcome, he created a new composition (Ya-Ling, 2019). "Duchamp created a musical *jeu d'esprit*[4] and Cage explored chance as a compositional principle (Vergo, 2010, as cited in Romero & Machado, 2007).

In Making Pictures Today (1964; figure 31), Gerstner (2007) introduces a systematic approach to image creation. Its core idea is the usage of algorithms or pre-defined systems to generate visual outcomes, rather than relying solely on intuition or subjective artistic decisions. This method is inspired by the belief that art and design can be enhanced by incorporating mathematical and logical frameworks, resulting in a more objective and reproducible creative process.

Figure 31.
Making Pictures of Today (Gerstner, 2007).

Following the relationship between systems and the human algorithmic role, Ambient Machine (Suzuki, 2022) is a wooden sound device that allows users to create and mix soundscapes by flicking 32 switches (figure 32). Here again, the decisions are constrained by the design-defined parameters and are essential for generating the outputs.

5. Small childish and ingenious work or exercise of the mind

Figure 32.
Ambient Machine (Suzuki, 2022).

We conclude this section by referencing Reynolds' (2002) work on steering behaviours. Reynolds introduced algorithmic steering behaviours that enable autonomous agents to navigate their environment with limited perception, simulating actions such as fleeing, pursuing, and evading. We will discuss how and when this technique was applied in Chapter VI — *Installation*.

Summary

This chapter has traversed a broad spectrum of research and historical context, focusing on SA, cross-media portrayals, computational and new media art, and algorithmic processes and creativity. The review aimed to provide a comprehensive theoretical foundation relevant to the themes of this dissertation.

Our review of SA reveals its multidisciplinary roots and divergence from conventional music. Emerging from early 20th-century performative experiments, SA has evolved to encompass diverse sonic practices beyond traditional music settings. Foundational figures like Luigi Russolo and John Cage established key concepts that continue to influence contemporary artists such as Tarek Atoui. Despite its potential for emotional engagement, SA frequently encounters challenges with audience engagement, emphasising the need for careful installation practices to address spatial and perceptual issues.

The multiple forms of music notation demonstrates a complex interplay between sound and image. From the ancient *Hurrian Hymn* to abstract notations by Xenákis and Cage, these cross-media approaches offer ways to encode and interpret different musical ideas. The exploration of graphical scores and their implications for performance underscores a dynamic relationship between visual art and musical composition.

New Media Art has redefined the relationship between sound and image through the use of digital technology, enabling the creation of innovative and immersive experiences. Pioneers such as John Whitney, known for his groundbreaking work in digital animation, laid the foundation for this evolution, while contemporary artists like Ryoji Ikeda use real-time data and interactive elements to explore new dimensions of artistic expression. These advancements highlight a broader trend towards the development of multi-sensory experiences that challenge traditional art forms and actively engage audiences, transforming them from passive observers into participants in the artistic process.

Algorithmic processes have increasingly influenced creative practices, blending computational methods with artistic intuition. Systems like Queneau's structured poetry or Duchamp's chance-based compositions highlight the potential for algorithms to shape artistic outcomes.

Overall, the research provided a thorough understanding of the diverse fields interconnected with this dissertation, bridging historical context with contemporary practices. This comprehensive overview not only clarified the current landscape but also offered critical guidance for navigating the various dimensions of our work.

Chapter v
Preliminary Work

Before working with motion, dynamism, and interactivity, we explored the visual representation of sound. To achieve more visual variety, we focused on interpreting a single composition using different approaches. For that purpose, from February to May 2023, we developed a compendium of eclectic computational visual essays of *data.matrix* (Ikeda, 2005), a recording and composition by the audiovisual artist Ryoji Ikeda. Its whole purpose is to provide visual references and guidelines for the practical component of this dissertation. This chapter is dedicated to its presentation.

Approach

Sá (2016) confronts two major opposite audiovisual approaches and catalogues them as the "one-to-one approach" and the "structural approach", as previously discussed in Chapter IV — *State of the Art*. Building on her framework, we applied these concepts to static environments, envisioning multiple interpretations from two standpoints: a sequential approach, where audio and its representation are recorded synchronously, and a global approach, where visuals are derived from an overall audio analysis. There is space for either a more pragmatic and linear methodology or a more symbolic and interpretably less obvious way of using the audio data for each of the presented approaches. In every stage, we took advantage of computational methods and the "unexpected outcomes driven by data that sometimes arise" (Shim, 2020). Despite using simple analysis of the sound components like amplitude and frequency, in some cases, we dissected the composition in order to understand its structure and distinguish and highlight its different sections. The generated visual essays were inspired by the computational and new media artworks presented in Chapter IV — *State of the Art*, as well as by algorithms that produced visually interesting outputs. We envisioned a minimalist and a black-and-white environment that could be applied both at this stage of the project and at later stages.

We started by writing a list of possible systems and sketches to be developed. Before starting to implement them, we had to standardise the artwork's proportion, analyse the mentioned list, and sort it out. We used the following criteria:

1. Type (sequential or global approaches)
2. Similarity with its precedent
3. Time-consuming level

The goal was to have printable outputs on a short term so they could guide the remaining work. This order also ended up by defining the compendium structure, as we organised the artworks by different labels or types, consequently respecting and displaying the chronological procedures.

Specifications

In every output generated while using a sequential approach, we focused on finding patterns originated by the composition structure itself (figure 33). To make them visible, different grids composed of distinct glyphs were tested. By analysing the sound more or less regularly, it was possible to adjust the number of elements in the artwork. A sequential approach also has the ability to show the flow of the composition — how the frequency values vary from start to finish and where the volume peaks are. Some of the generated and selected results display this more effectively than others.

The first series of artworks are sinistrodextral — read from left to right. We found the outputs more engaging when the behaviour of the sound is more clearly defined, as it provides insight into how the composition unfolds.

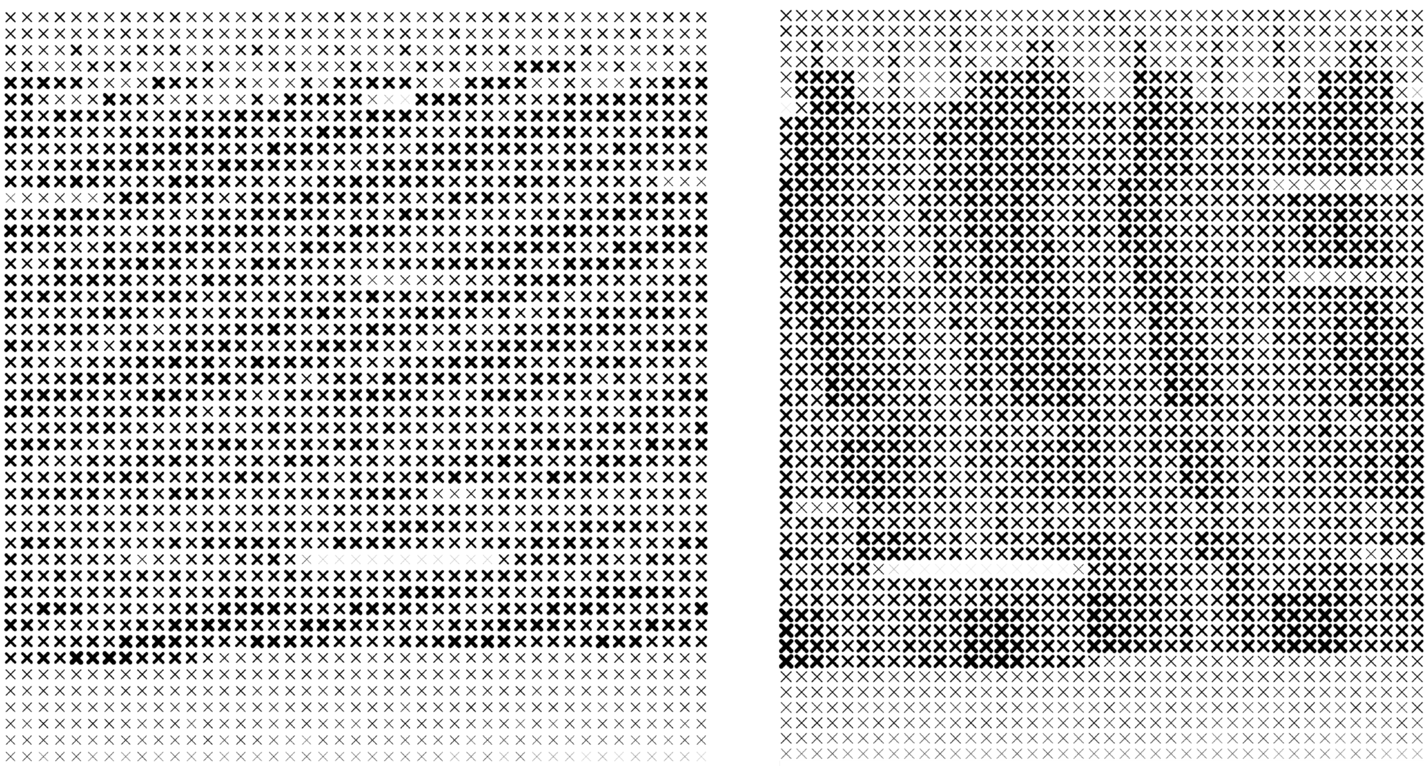


Figure 33.
How different grids affected the patterns found.

Following the same principles, we then developed top-to-bottom patterns. We highlight multiple results for different reasons. Figure 34.A displays a minimalist artwork that portrays overall amplitude analysis and figure 34.B shows six-band frequency analysis. The output in figure 34.C uses chaos and a glitchy appearance to its advantage.

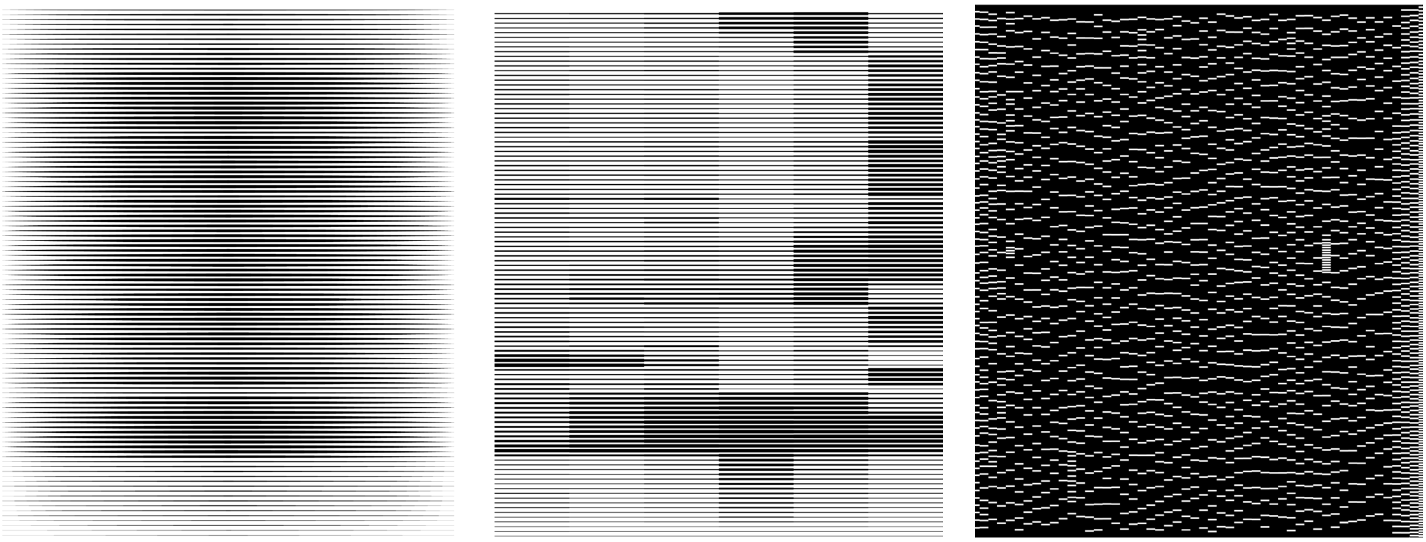


Figure 34.
A. Amplitude analysis on line thickness (left);
B. Six band frequency analysis (centre);
C. Amplitude analysis on rectangle height (right).

We highlight the outputs on figure 35 as they embody less pragmatic and straightforward methods. Upon finding the different sections of the composition based on its dynamic, we emphasised them by confining them to a specific space. This originated seven individual rectangles. For each of the composition sections, the system divides and subdivides the specific area by drawing an horizontal line if a high pitch beat is registered, or a vertical line if a low pitch beat is identified. By slightly changing the threshold of the beat identifier, the outputs generated changed greatly without losing their informative capability.

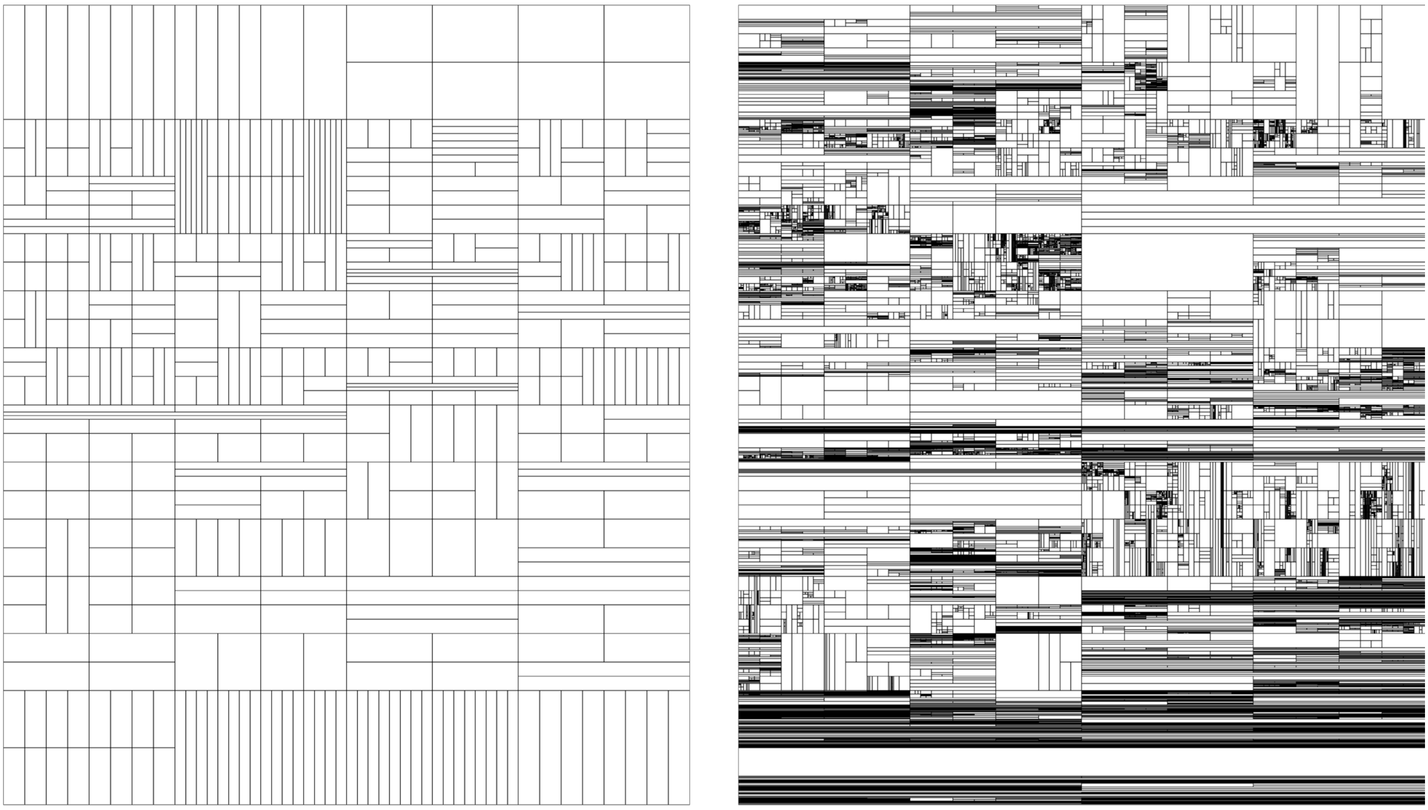


Figure 35.
Outputs for manipulation of threshold for beat detection and respective section dividing.

To end the sequential approach methods, we mention circular patterns where a clockwise reading is implied. In figure 36, the system drew a line every two degrees for each of the frequencies used, with thickness translating into value. This means that the composition is analysed 180 times (0 to 2, 2 to 4, and so on). For legibility purposes, the low segment of the spectrum is drawn on the outer ring of the circle, as its density tends to be higher. We chose not to dig further into this kind of interpretation, as the outcome of this type of system is very common in general in ordinary music visualisation (Benadon, 2007; Mrugala, n.d.).

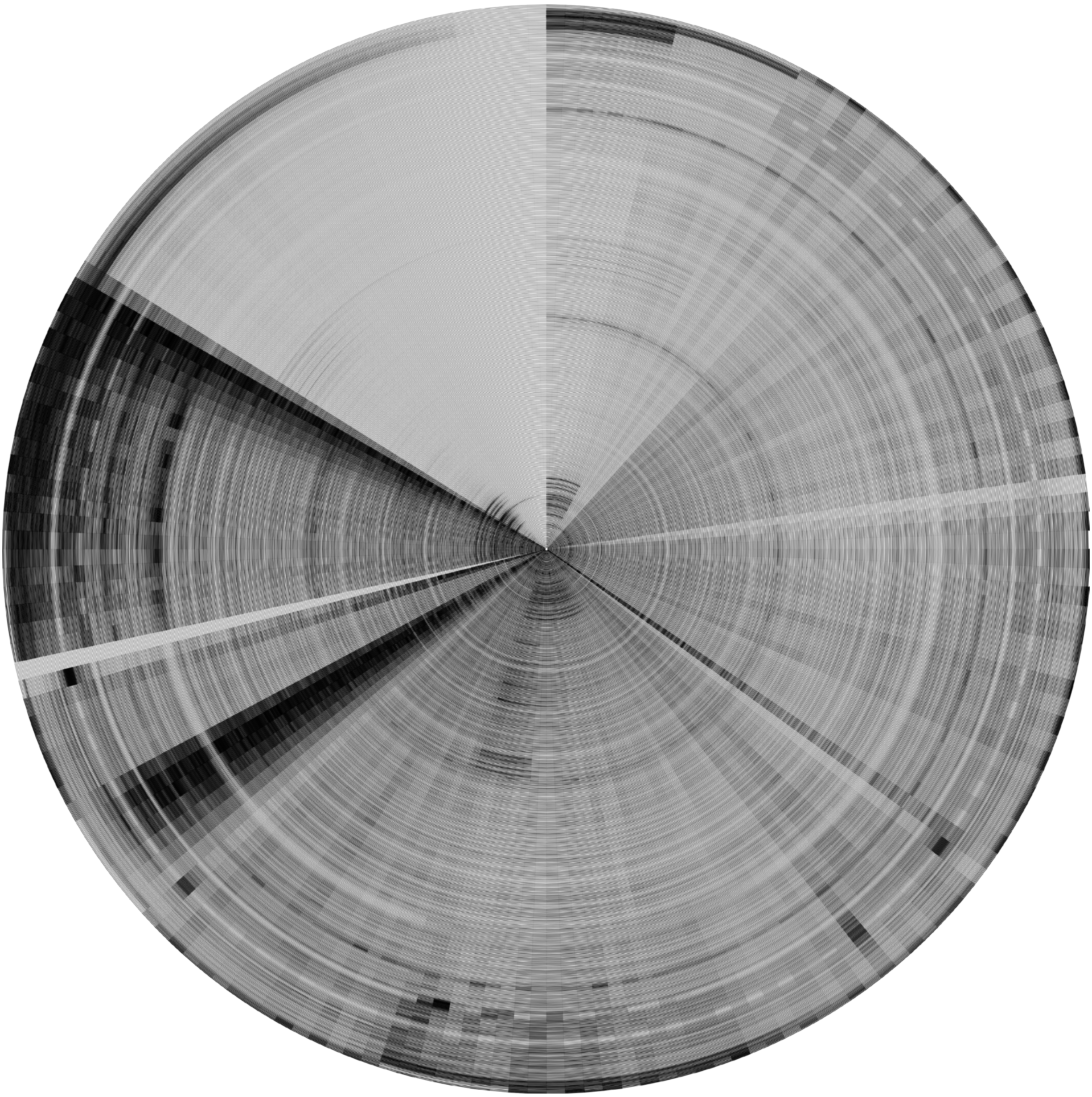


Figure 36.
Circular frequency analysis.

When we looked into the composition from a general standpoint, we envisioned new-generation systems. We began by representing the most common frequency partitions by drawing areas that correspond to their relevance (the amount in %). In figure 37.A, the second graphic partition which represents the frequencies from 250.00 to 500.00 HZ, reveals itself as the most present. Figure 37.B shows a circular version of the same idea, where this distribution distorts the shape of the ellipse, transforming it into a blob with a higher radius in more audible frequency areas. Figure 37.C uses beat detection to make shapes fall from the top to the bottom. The low and high pitch beats are distinguished by a binary fill colour of the shapes so the reader can get hints from the visual density. We considered this type of method and level of abstraction to be our boundary since we did not intend for the outputs to verge on total ambiguity.

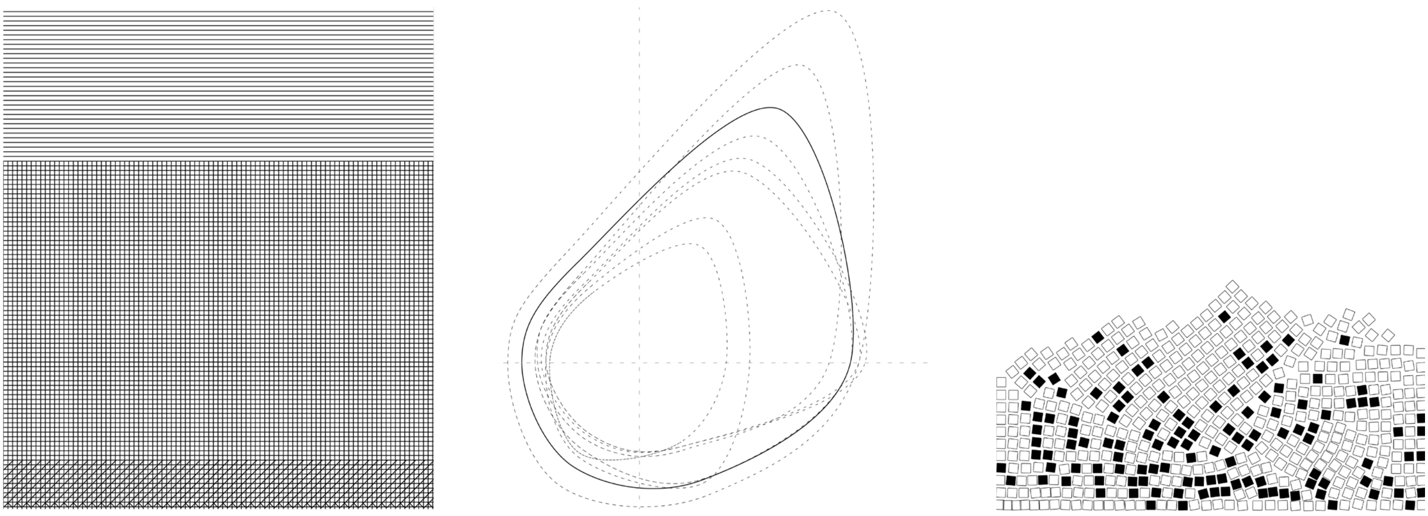


Figure 37.
A. Pattern with the dominant frequencies (left);
B. Blob with the dominant frequencies (centre);
C. Beat detection contrast (right).

Notes

From all the generated outputs, we discarded the less effective ones (primarily the 3D variants) and selected the most successful results for printing, along with some coding errors that were used as section dividers. We realised that, despite following the initial list, some of the best artworks emerged from analysing and redesigning the work as it developed.

The created compendium has helped to define the project's motto and has inspired the next phase.

Chapter VI
Installation

Figure 38.
S—I at Círculo de Artes Plásticas de Coimbra.

S—I[6] is an interactive installation that generates sound compositions and visual artefacts. The experience unfolds in two stages and is designed for individual participation, with one person engaging at a time. The first stage, the experimentation stage, allows the participant to explore and learn the mechanics of the installation. The second stage, the execution stage, challenges their auditory perception by requiring them to respond to the sound based on their memory and insights gained during the first stage. The experience takes place in a black box with minimal light, a camera (*Kinect V2*[7]) and four speakers concavely positioned, mimicking the appearance of an AM (figure 38).

The main objective of this installation is to create an environment for sound exploration, enhancing auditory memory, perception, and capabilities. This is achieved by using sound synthesis as the sole audio source, building each synthesiser from scratch, and manipulating them computationally, allowing for greater flexibility in generating the composition. Finally, we evaluate the experience phenomenologically by comparing interactions from different experiments.

6. Video demonstration at joaoneves.art/dissertation#showcase

7. A motion-sensing device that detects body movements and gestures. It is equipped with a colour camera, an infrared camera, a depth sensor and four microphones.

In this section, we discuss S—I at both a conceptual and a technical level. Conceptually, we delve into how am properties are translated into the object. Technically, we provide a comprehensive overview of the system by explaining the decisions made during its development, detailing the architecture and describing the encodings used.

> *"I believe that what our mind does best is to be sensitive to new things, rather than to reiterate successful habits."*
> — JÔNATAS MANZOLLI (2022)

Approach

The chain of events that depicts the functioning of AM translates into a straightforward process that begins with movement from an aircraft. The next factor is sound, involving a source and a detection agent, which is the focal point of the structure. As the sound source approaches, it becomes optically identified, leading to a visual stimulus. We approach S—I by exploring the following principles, concepts and metaphors:

1. Co-creation — the audience plays a fundamental role in the algorithmic process by simulating an enemy aircraft in the first stage, facing an AM (the system), and acting as a detection device in the second stage. Audience interaction also feeds the system, as we record their paths, which are used in future interactions. We embrace the choices of the audience as they decide to act against or along the algorithms, making a choice between augmentation and erosion (Pošćić & Kreković, 2020).
2. Immersiveness — combining rich and powerful sounds with a black box setup and minimal light helps the audience immerse themselves in the experience.
3. Sound before image — the audience acts solely based on their sonic perception. The visual artefact generated from each experience is only presented after its end. This reflects the asynchronous audiovisual relationship identified in AM.
4. Sound memory — like Burt (2005), exploring what kind of mapping might be easier to hear and also what kind of mappings the audience is able to learn to hear. We aim to use data mapping as a resource for ear training.
5. Delayed response — the gap between auditory perception and subsequent action relates to the non-immediate and sometimes faulty detection characteristic of AM.
6. Detection as composition — the audience's detection and behaviour translate into both sound and visual compositions, depending on the stage.

7. Movement and positioning as triggers — by exploring spatiality, the audience triggers and manipulates the sound composition according to pre-established encodings.
8. Proximity as danger — the system responds to audience proximity based on the principle that closeness equates danger.
9. Mirroring — we use panoramics to explore this concept;
10. Randomness as a tool — when applied carefully, it contributes to the variety and richness of the compositions without impairing the sonification process.

Architecture

The system supporting S—I is subdivided into four main sectors: a) detection of bodies and their movement in the room; b) sound synthesis, composition and reproduction; c) encodings; and d) creation of visual artefacts. Figure 39 demonstrates the system pipeline.

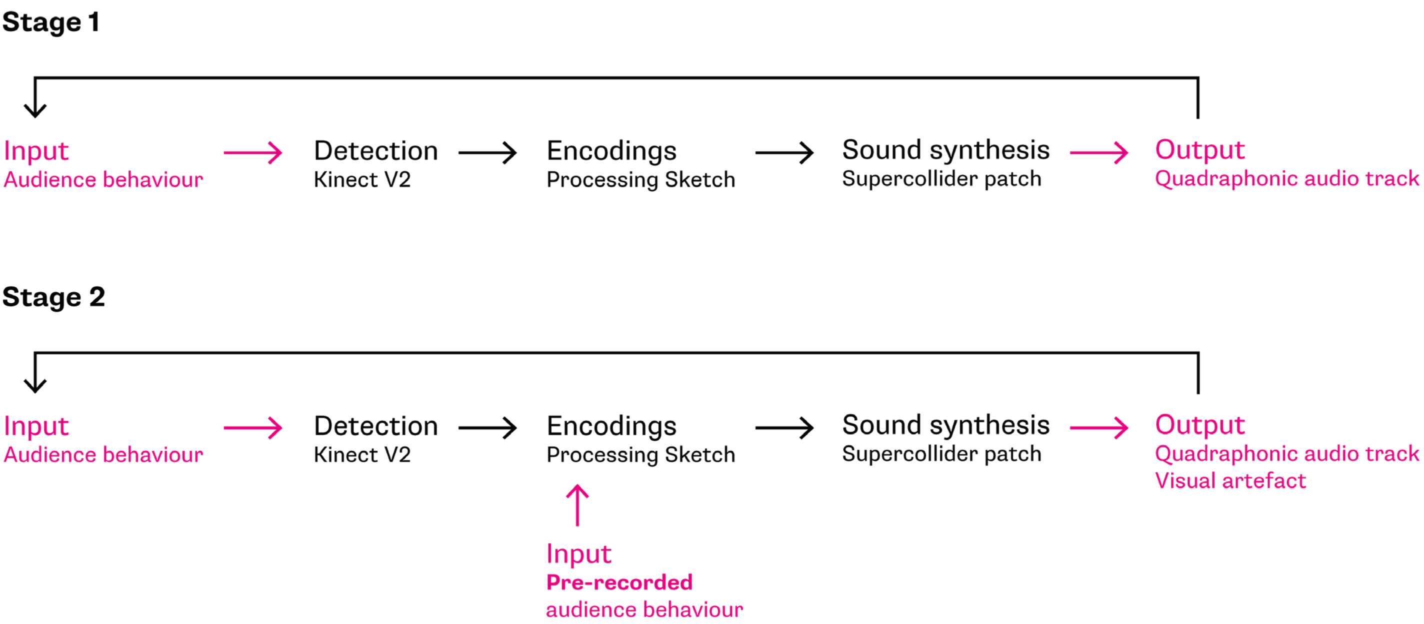


Figure 39.
S—I architecture.

The audience's behaviour is detected by a *Kinect V2* and processed by a *Processing*[8] sketch. In the first stage, their movement and velocity directly impact the sound generated, according to the encodings in the *Processing* sketch. The sketch sends OSC[9] messages to a *SuperCollider*[10] patch, triggering each synthesiser which is then outputted as a quadraphonic audio track. In the second stage, the audience no longer impacts the sound. Instead, the system reads a previously recorded path, analyses it, and sends OSC messages to the *SuperCollider* patch accordingly. In this context, the audience's movement and velocity are recorded for two purposes: a) to provide a visual comparison between their path and the reproduced path, the visual artefact, and b) to feed the system with a new path that becomes available for future interactions.

8. An open-source software and language for visual arts, often used for creative coding and interactive installations.

9. Open Sound Control is a protocol used to send and receive messages between computers, musical instruments, and other multimedia devices over a network.

10. An open-source platform for audio synthesis and algorithmic composition.

We use a *Kinect V2* for its depth camera capabilities, *Processing* for its efficiency as an open-source language suited to visual debugging and encoding, and *SuperCollider* for its strength as an open-source, object-oriented language for audio synthesis and algorithmic composition.

Detection

The computer vision of this system is achieved using a *Kinect V2*, whose primary purpose is to track the audience's head movements, specifically their horizontal (x) and depth (z) positioning. The code that interprets the *Kinect V2* data is encapsulated in a function called `kinect()`, which orchestrates the workflow by capturing skeleton data, processing joint positions, and mapping them to a 2D space. The function also handles scenarios where the user is no longer detectable by the *Kinect V2* device, ensuring that the system can reset and prepare for new interactions.

Mapping these variables into a 2D space presents certain challenges. Due to the *Kinect V2*'s field of view, the minimum and maximum x positions of a detected body vary at different z values. This means that for each room where S—I is installed, we must register these variances and update the function accordingly. Figure 40 illustrates how we addressed this issue for themost recent installation of S—I in the black box at the Círculo de Artes Plásticas de Coimbra facilities, which has an irregular polygonal shape.

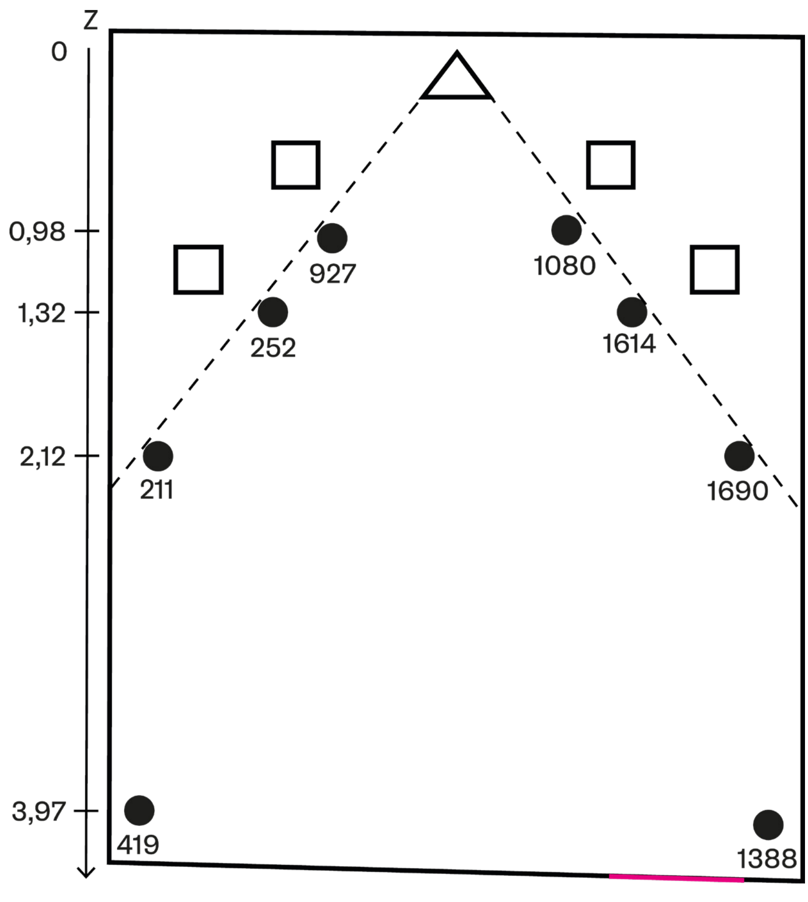


Figure 40.
Mappings for the black box at the Círculo de Artes Plásticas de Coimbra. The black circles represent the minimum and maximum x positions registered along the z-axis as defined by the *Kinect V2* (represented by a triangle). The squares indicate the locations of the speakers, and the magenta line marks the entrance to the room.

We then use linear regression and constrained mapping algorithms to achieve a highly reliable prediction of the audience's position within the installation. The linear regression function calculates the best-fit line for a set of data points. The function first computes the necessary summations. Using them, it then calculates the slope of the line which represents how much the `x` value changes for each unit change in `z`. The function returns an array containing the slope and intercept which are then used to predict the `x` position for any given `z` value.

For development, encoding and debugging purposes, we mapped these outputs to a small window that displays the positioning. To prevent position values from going out of bounds, we constrained their range to a minimum and maximum limit.

Sound Synthesis

Developing all the synthesisers involved a lengthy process of experimentation, frequently characterised by trial and error. Before addressing the composition, we first review some of the synthesisers we built, including those used in the final system: `synth1`, `synth2`, `reverb`, `click`, `drone`, `glitch`, `pinknoise`, `brownoise`, `kick`, `hithat`, `highpitch`, `highpitch_sin`, and `ping`, as well as some that were saved but ultimately not selected such as `dial` and `glitchsin`.

`synth1` is composed of a sawtooth wave signal that passes through a band-pass filter and then a low-shelf filter. An envelope that ensures gradual fade-in and fade-out is applied to the signal (figure 41), along with an array of four different amplitude values — one for each speaker. The use of random modulation in frequency and filter parameters ensures that the sound remains lively and constantly evolving. By setting the resonance (Q factor) to very low values, the resulting frequency is perceived as rhythm, and its pitch can be controlled by the lower range of the band-pass filter. This value is randomised within the Lydian dominant scale (1 2 3 ♯4 5 6 ♭7), the fourth mode of the minor major scale, which has no avoid notes[11], ensuring that all output sounds remain consonant.

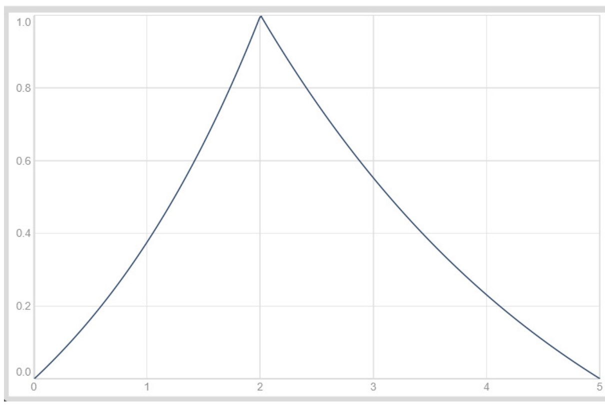


Figure 41.
Envelope used for most of the synthesisers. Created with *SuperCollider*.

11. Notes within a given scale or chord that are considered dissonant or clashing when played over certain harmonies.

synth2 is a synthesiser that generates a sawtooth wave signal that undergoes several stages of processing. Initially, the it is detuned using low-frequency noise modulation, which introduces subtle pitch variations, making the sound more dynamic. This detuned signal is then passed through a short delay line with randomly modulated delay times, creating a chorus-like effect that enhances the spatial depth of the sound. Following this, the signal is filtered using a low-pass filter to smooth out the high frequencies, and then further shaped by a high-shelf filter that attenuates the high-frequency content, resulting in a darker tone. The same envelope applied to synth1 is applied to control the amplitude. The output is then distributed across four speakers, with panning managed by an azimuth panner (figure 42), allowing the sound to be perceived from different directions.

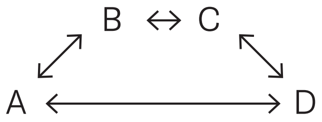


Figure 42.
How the azimuth panning moves the sound between speakers A, B, C & D.

reverb is designed to add reverberation effects to audio signals. The synth begins by processing the input signal with a pre-delay of 0.1 seconds, which creates a slight initial gap before the reverberation effect begins. The core of the reverb effect is achieved through a series of 16 Allpass filters, each with randomly modulated delay times and a decay time of 1.8 seconds. The signal is then passed through a low-pass filter. The final step blends the original dry signal with the processed wet reverb signal using an xFade UGen[12], which adjusts the mix level between the dry and wet signals based on the mix parameter. The result is an output that combines the clarity of the original signal with the reverberated sound, with amplitude control applied to the final output.

Unlike the subtractive synthesis used in synth1 and synth2, the click synthesiser employs wavetable synthesis. The synthesis begins with the generation of a wavetable using an envelope which defines the waveform shape over time with a sine wave function. This wavetable is then allocated and loaded into a buffer, preparing it for use in the synthesis process. The primary signal is produced by reading from this wavetable via an oscillator that is modulated by a specific frequency. The signal is subsequently shaped by the already presented envelope generator and submitted to the same panning UGen.

All the other synthesisers share similar characteristics despite producing very different results. Although the initial signals and synthesis methods vary, each applies the same envelope and is processed through an azimuth panner UGen. This setup allows for precise control over the placement of the sound source within an array of speakers. Table 2 describes the type of synthesis used for each one of them, as well as the UGens involved in their creation. Each one can be listened to on the shared webpage[13].

12. A unit generator is a building block of synth definitions on the server, used for generating or processing both audio and control signals.

13. Synthesisers samples at joaoneves.art/dissertation#synths

Table 2.
Type of synthesis used for each developed synthesiser and respective UGens.

Synth Name	Type of Synthesis	UGens
\brownnoise	Noise-based	WhiteNoise.ar, EnvGen.kr, PanAz.ar, Out.ar
\click	Additive	SinOsc.ar, EnvGen.ar, PanAz.ar, Out.ar
\dial	Subtractive	SinOsc.ar, TIRand.ar, Impulse.ar, Splay.ar, Out.ar
\drone	Additive, Subtractive, Noise-based	Pulse.ar, SinOsc.ar, Saw.ar, BrownNoise.ar, EnvGen.kr, PanAz.ar, Out.ar
\glitch	Additive	Impulse.ar, BPF.ar, EnvGen.ar, PanAz.ar, Out.ar
\glitchsin	Additive	SinOsc.ar, EnvGen.kr, Out.ar
\highpitch	Additive	Pulse.ar, EnvGen.ar, PanAz.ar, Out.ar
\highpitchsin	Additive	SinOsc.ar, EnvGen.ar, PanAz.ar, Out.ar
\hithat	Additive, Noise-based	SinOsc.ar, WhiteNoise.ar, EnvGen.ar, PanAz.ar, LFNoise1.kr, Out.ar
\kick	Subtractive	LFNoise1.ar, LPF.ar, EnvGen.ar, PanAz.ar, Out.ar
\ping	Additive	SinOsc.ar, EnvGen.ar, LFNoise2.ar, PanAz.ar, Select.kr, TIRand.kr, Impulse.kr, Out.ar
\pinknoise	Noise-based	PinkNoise.ar, EnvGen.kr, PanAz.ar, Out.ar
\reverb	Effects (Reverb)	In.ar, DelayN.ar, AllpassN.ar, LPF.ar, XFade2.ar, Out.ar
\synth1	Subtractive	Saw.ar, BPF.ar, BLowShelf.ar, LFNoise1.kr, EnvGen.kr, Out.ar
\synth2	Subtractive	Saw.ar, DelayC.ar, LPF.ar, BHiShelf.ar, LFNoise2.kr, PanAz.ar, EnvGen.kr, Out.ar

SuperCollider allows for the asynchronous plotting of separate functions. In this system, we pre-run a patch that enables *Processing* to trigger each pre-built synth separately, according to the encodings. The patch is organised in the following structure: A) server configuration; B) global variable initialisation; C) piece-specific function definitions; D) registration of the functions; E) server boot; F) remaining processes that require a booted server. This is an advised structure for piece compositions (Fieldsteel, 2024) that ensures that every line of code is executed in the correct order. We now provide a detailed overview of the patch.

SuperCollider allows for the asynchronous plotting of separate functions. In this system, we pre-run a patch that enables *Processing* to trigger each pre-built synth separately, according to the encodings. The patch is organised in the following structure: A) server configuration; B) global variable initialisation; C) piece-specific function definitions; D) registration of the functions; E) server boot; F) remaining processes that require a booted server. This is an advised structure for piece compositions (Fieldsteel, 2024) that ensures that every line of code is executed in the correct order. We now provide a detailed overview of the patch.

When configuring the server we define the output device, the number of bus channels (matching the number of speakers — 4), the sample rate format (which should be the same as the device output we're using, in this case 44100 HZ) and the memory allocated for the server (1GB is sufficient for the system encodings). Next, we create global variables to ensure they are accessible throughout the code. These include variables for the synths played with patterns, which allow for easy activation and deactivation, a variable for the output bus to simplify output management and a variable defining the IP[14] address for communication with *Processing*.

The next step involves defining functions. `~makeBusses` initialises a dictionary for the buses and adds the reverb synth. `~cleanup` deallocates any buses and unregisters functions that may have been left registered from previous executions of the patch. Finally, `~makeNodes` is responsible for setting up the configuration of groups and synths on the server, ensuring the reverb group is placed after the main group so that the desired synths can pass through that bus.

We then register these functions for server initialization and shutdown and proceed to boot the server. To prevent adding synths and busses before the server is fully booted (as this is not an immediate process), we asynchronously add them to allow it to receive messages from *Processing*. We proceed by describing these encodings.

Encodings

The *Processing* sketch is organised into three stages: 0, 1, and 2. Stage 0 can be referred to as the *rest stage* as its sole purpose is to check if a body is detected. When a body is detected, the timer for the first stage, the experimentation stage, starts. During our tests we set this timer to 5 minutes, which we predict to be sufficient for the audience to apprehend and memorise the sonic cues. When there are 10 seconds left on the timer, the sound output stops, signalling the start of the second stage — the execution stage. Following the execution stage, the sketch saves the path taken by the audience as text coordinates, creates

14. A set of rules governing the format of data sent over the internet.

visual artefacts, saves them, and then resets. At any point, if the audience leaves the room for more than 5 seconds, the system also resets. Everytime the system resets, it checks for every path inside a specific folder and loads them, making them available for being selected on the execution stage.

The two main classes responsible for content generation are the `Output` class, which is used exclusively for generating visual artefacts and will be discussed in the next section, and the `Human` class. The `Human` class has different functionalities depending on the stage of the system. When a body is detected a `Human` element is added and stage one begins. From this point on, the body mapping is displayed in a small window for debugging purposes. To achieve a smoother and continuous path without bumps, we created an autonomous agent that seeks the body's real position using steering behaviours described by Reynolds (2002). This was implemented following methods presented by Shiffman (2012). The distance gap between the real position and the autonomous agent's position contributes to the system's delayed responsiveness. Higher velocity (unquietness) increases the time it takes for the autonomous agent to reach the correct position, conceptually linking velocity to reading difficulties (figure 43).

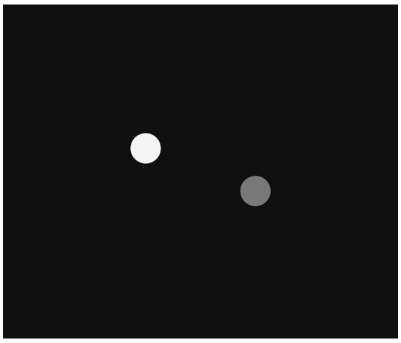


Figure 43.
Position mapping on a given room. The grey circle — the autonomous agent — is always chasing the white circle — the real position of the body detected.

Within the `Human` class, the system continually updates two values: the position of the autonomous agent and the agent's velocity, calculated as the difference in position between the current and previous update. This data is used differently depending on the stage in progress. Figure 44 illustrates how the system operates and varies across these stages. In the first stage, the system updates the position and velocity of the autonomous agent, which tracks the body's position to generate a sound composition. This process enables experimentation with the composition's encodings. In the second stage, the sound composition is influenced solely by playing a recorded path from a preloaded list of paths. This creates the illusion of another entity in the room, which the audience is encouraged to follow using their sound memory and knowledge acquired during the first stage. The execution of this challenge generates the visual artefact and feeds the system for future iterations.

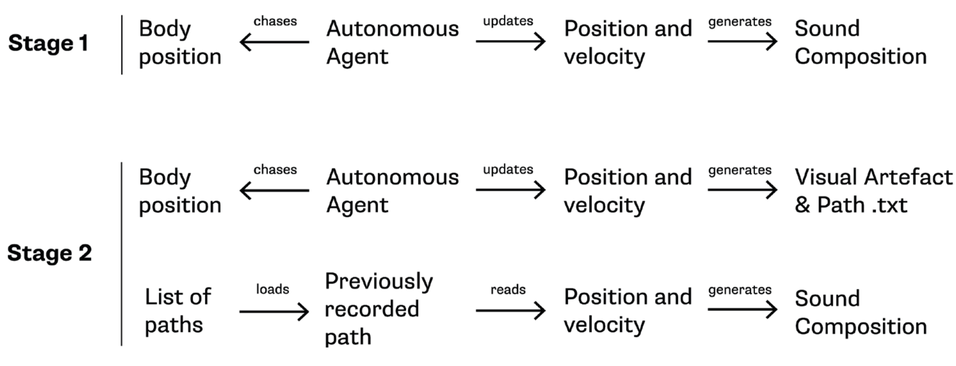


Figure 44.
S—I pipeline variation between stages.

Despite the stage, the composition[15] and its responsiveness adhere to the same encodings. The composition comprises three distinct layers: a) harmony; b) melody; and c) noise. We proceed by describing them and how they react to the position and velocity that is being read (table 3).

Table 3.
Reactiveness and layers affected by each factor.

Axis	Factor	Genre	Synth	Affected Layer
y	Proximity	Intensity, Rhythm	\synth1, \synth2	Harmony, Melody
x	Horizontality	Panoramics	\synth1	Melody
x and y	Velocity	Noise	\synth1, \brownnoise, \click, \drone, \glitch, \highpitch, \hihat, \kick, \ping	Noise

The harmony, generated in real time by `synth2`, consists of a chord with seven voices. This system alters the chord in two main ways:

1. When a new body is detected, the Human element generates a random number that is used to adjust the chord's tonality.
2. The chord's fifth voice has a low probability (2%) of being changed every frame to a random note from an array (figure 45).

This stochastic behaviour enriches the composition by preventing stagnation and predictability, introducing movement, adding harmonic richness and creating an element of surprise, thereby immersing the audience.

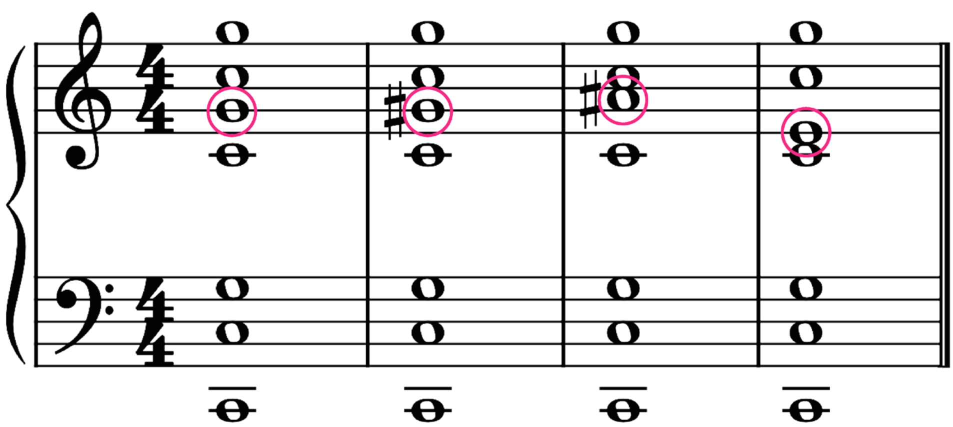


Figure 45.
Variances in the chord's fifth voice result in different: C5 (first measure), Cmaj7♭5 (second measure, introducing dissonance), E♭7 (third measure), and E♭ (fourth measure).

The proximity to the detection device, positioned centrally near the front wall of the room, directly impacts this layer by detuning `synth2` creating tension and chaos. This effect is achieved by mapping `synth2`'s detune parameter (explained in the previous subchapter) to the `y` position from the 2D mapping of the path.

15. Generative composition example at joaoneves.art/dissertation#composition

The melody is generated by triggering synth1 which selects random notes from a pre-established scale. The rhythm of these notes is also determined by the distance to the detection device. The notes_rhytmn variable contains the possible rhythmic values: 1/8, 1/4, 1/2, 2/3, 1, 4/3, 2, 5/2, 3, 4, 6, 8, 16, 32. These values were chosen for their arithmetic relationships, creating polyrhythms when overlapped, and representing the number of percussive articulations per second. We map the distance to the number of elements in the array and select the rhythm accordingly. Furthermore, this melody also reacts to the horizontal position detected. The volumes of the speakers mirror this position. The leftmost speaker gradually increases in volume as the position moves from the left edge toward the centre, while the rightmost speaker mirrors this behaviour, decreasing in volume as the position moves toward the right edge. The two central speakers adjust their volumes accordingly, with each speaker's volume increasing or decreasing based on whether the position is on the left or right third of the space, reflecting a mirrored response across the setup (figure 46). The decisions of this mapping were inspired by the way AM reflected the sound waves towards the focal point.

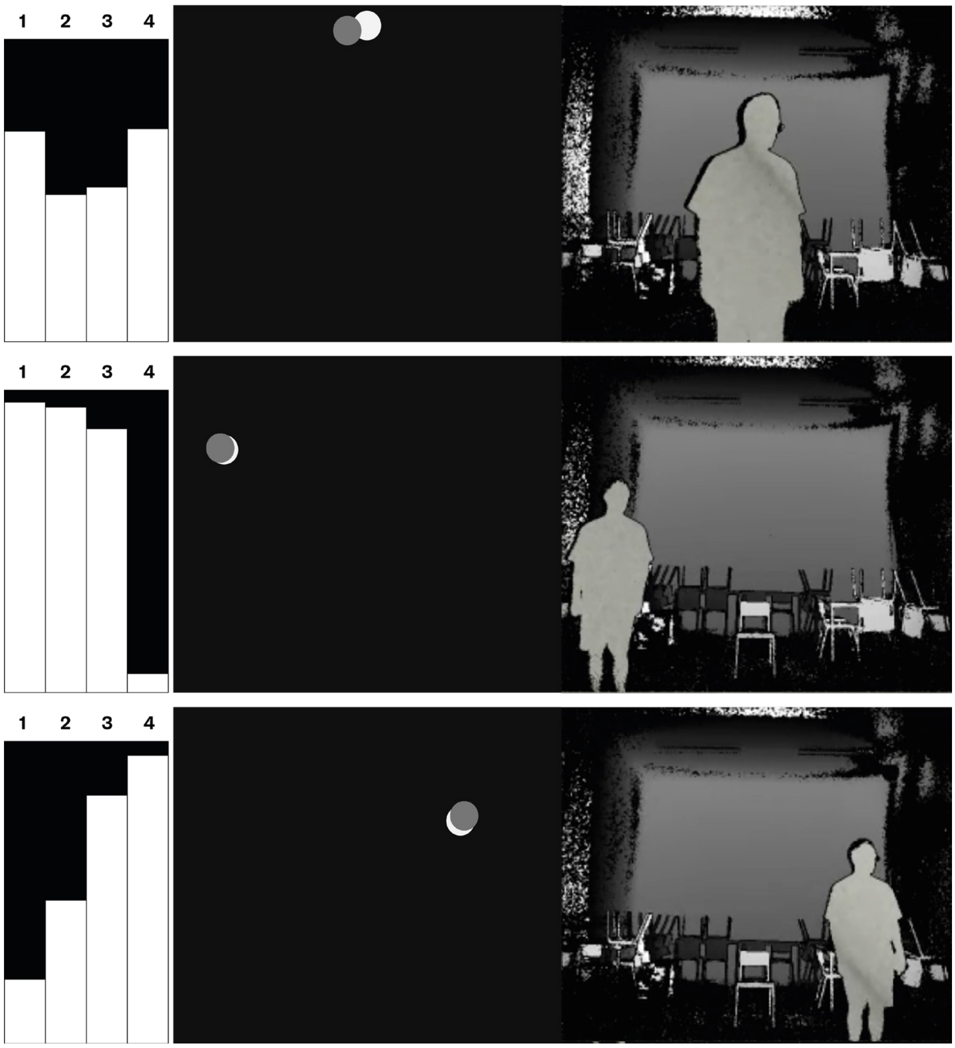


Figure 46.
Tests for mapping the horizontal position to the speakers, with speaker 1 as the leftmost, followed by speakers 2 and 3, and speaker 4 as the rightmost.

Finally, while some noise synths are triggered based on a probability that increases with higher registered velocity, others have their volume adjusted according to this velocity. These choices were made through experimentation with various values and were finalised based on personal preference (table 4). In addition to alerting the audience to the impact of sudden movement, the noise synths are designed to create a more challenging auditory experience, encouraging greater focus on the critical sound cues — the harmony and the melody.

Synth Name	Every x frame	Probability to be played	Volume change
\click	0 (every frame)	100%	Yes
\kick	12	15%	Yes
\highpitch	14	0.7%	No
\hihat	12	3%	Yes
\highpitchsin	16	2%	No
\glitch	8	15%	No
\ping	18	0.9%	No
\pinknoise	48	10%	Yes
\brownnoise	24	2%	Yes

Table 4.
Every noise synth probability to be played for every x frame.

Visuals

From the beginning of this project, we wavered on the idea of including a visual component. Since visual stimuli often overshadow sound stimuli, we sought to balance their contributions. In the end, we decided to include a visual component only after the installation experience, aligning with our interest of materialising the interaction — the tangibility. In this section, we present the visual creation process.

> *"...often, the two goals of information access and psychological engagement compete within the same new media object."*
> — LEV MANOVICH (2002)

Before deciding to use the audience as the error factor in the detection process, we explored ways to simulate this lack of precision using computational tools. We began by examining the 2D representation of the body within the installation room and implemented a genetic algorithm (Sims, 1994) that evolves curved lines. These lines are formed by the intersection of several points, with the fitness function rewarding closeness to the true path[16]. However, we found the results to be too informative, potentially hindering a more immersive and engaging experience (figures 47.A & 47.B).

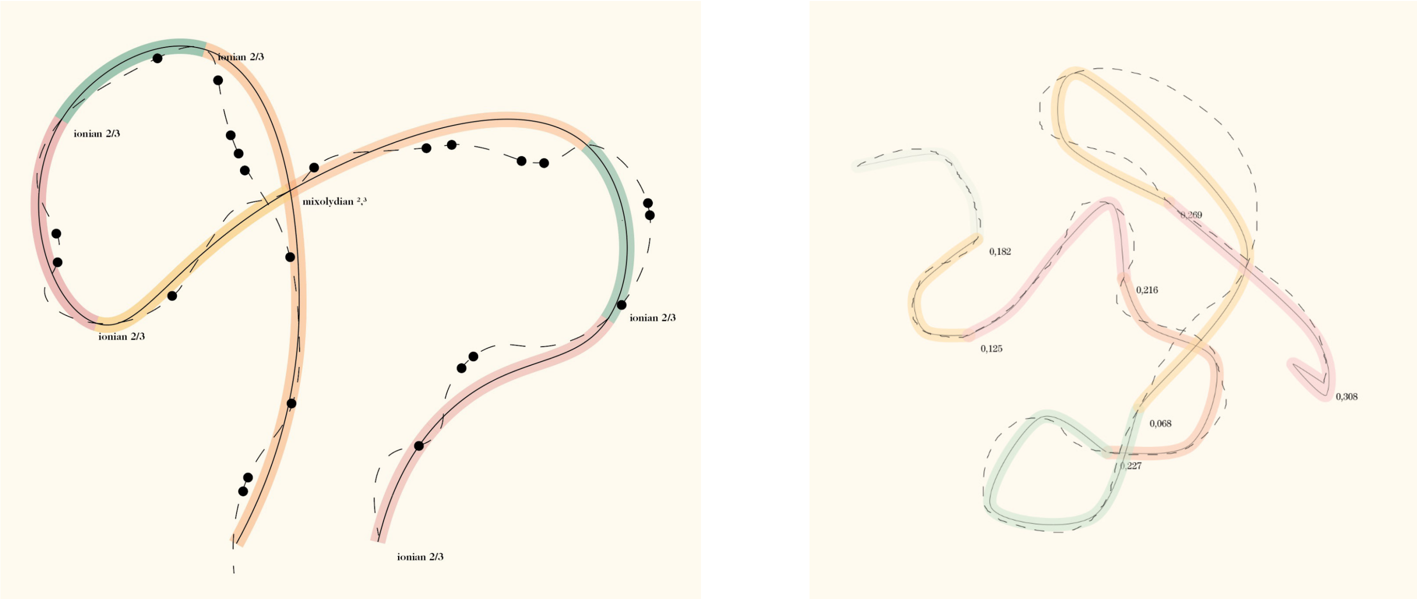


Figure 47.
A. Output digital sketch (left).
B. Output implemented version (right).

To illustrate the concept of delay, we implemented a sketch that captures and stores frames from the *Kinect V2* device. The buffer is set to store 300 frames (equivalent to five seconds). Once the buffer reaches capacity, it begins removing the oldest frame as new ones are added. The oldest frame is then displayed on screen, reflecting the audience's actions from five seconds earlier[17] (figure 48). The transparency of the frames is directly linked to the velocity of movement, creating an additive effect that enhances the noise metaphor. We also experimented with various colour palettes for the displayed frames. However, informal tests showed that these visuals were confusing, as the audience relied heavily on their vision to guide themselves. As a result, we decided to present the visual artefact only after the S—I experience was concluded.

Figure 48. Delayed visual response visuals by displaying the oldest frame of a five second buffer. Real-time position displayed in the top left.

16. Video of visual artefacts tests with genetic algorithms at joaoneves.art/dissertation#genetic

17. Video of detection and visual artefacts tests at joaoneves.art/dissertation#tests

When designing the visual artefact we initially aimed to create a frame that would represent stage 2 — execution — experience, by displaying and overlapping the audience's path with the pre-recorded path. We believed this approach would help in comparing and evaluating the experiences, as time-based processes like this can be frozen to make them more accessible for human perception.

The lines drawn in the initial outputs lacked smoothness and fluidity, as sudden changes in direction disrupted their continuity (figure 49.A). To improve the results, we implemented a workaround to address *Processing*'s limitations in drawing curved lines. By drawing a line connecting the first four points, and then each subsequent set of four points (e.g., points 2 to 5, then 3 to 6, and so on), we created a series of overlapping lines that smoothly connect the points (figures 49.B & 49.C). This approach also allowed us to implement line thickness variations according to the velocity recorded at each point.

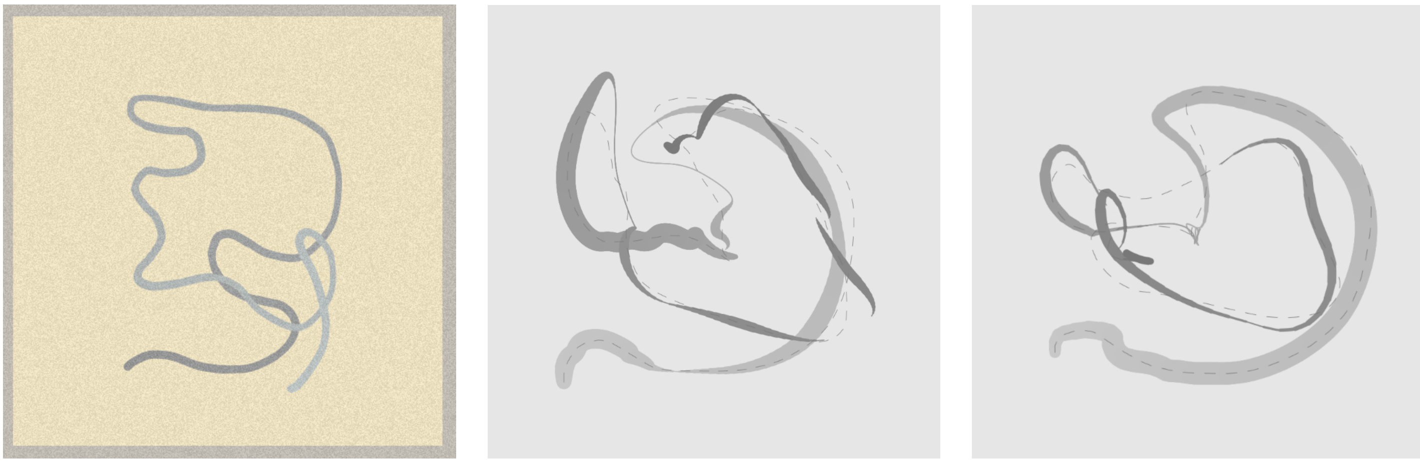
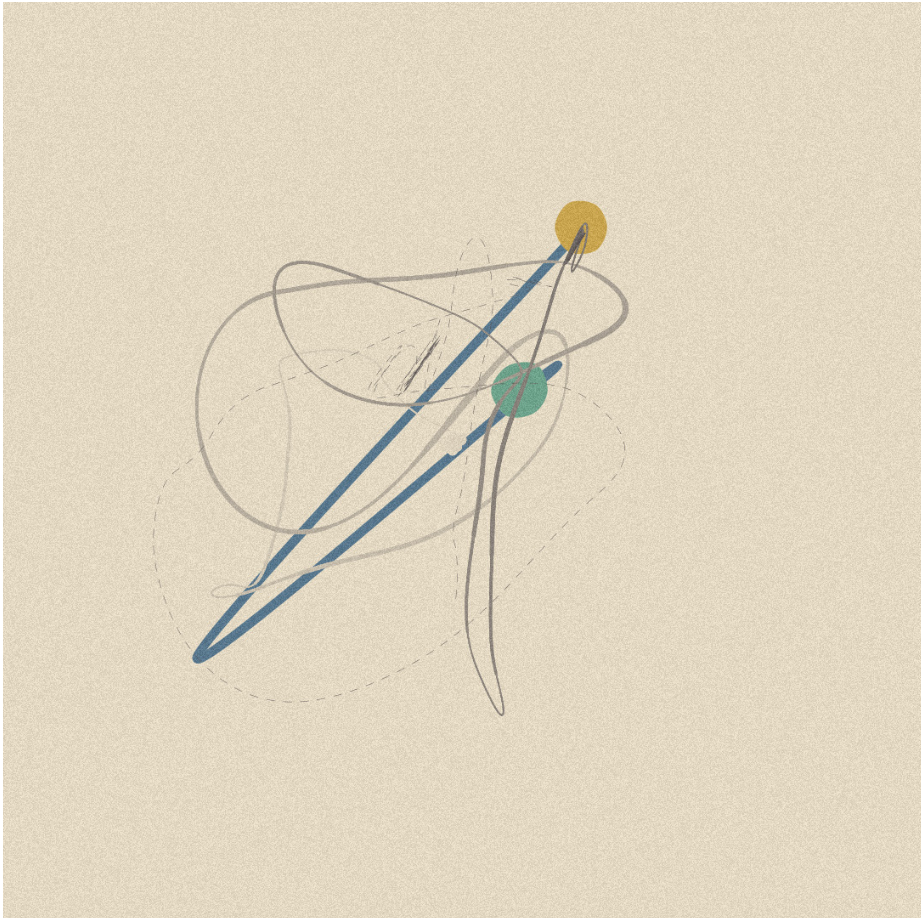
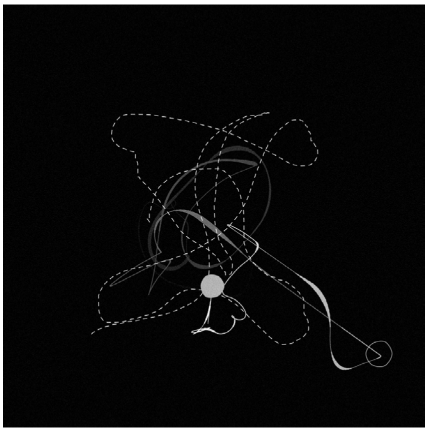


Figure 49.
A. First visual artefact output.
B. & C. Two outputs of a series of overlapping lines with thickness variations according to velocity. The dashed line represents the pre-recorded path.

The final version of this sequence of outputs adds some nuances (figures 50.A, 50.B & 50.C). The green and yellow blobs represent the points where the body is closer to and farther from the pre-recorded path, respectively, while the blue line connects them. Computationally, it adds grain and transparency to the path: the older it is, the more transparent it becomes. In the end, we realised that despite the enticing visuals, this approach would fail to represent lengthy experiences, as the constant overlapping of lines makes it very difficult to read and draw conclusions.

To enhance readability, we looked to improve contrast and we chose to divide the information to be evaluated into three separate outputs that together form the final visual artefact. Each output consists of two lines: one solid and one dashed. The dashed line represents different aspects of the pre-recorded path, while the solid line displays data related to the participant's activity.





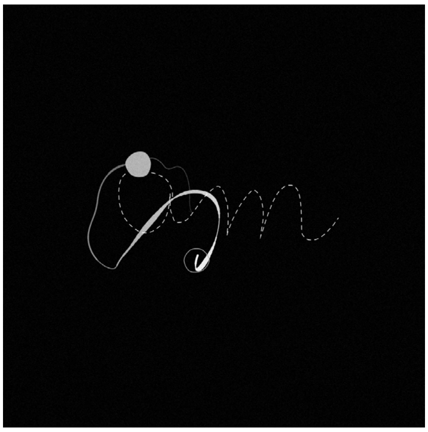


Figure 50.
A. Final version of this concept development: real Stage 2 example (left).
B. A greyscale version with higher contrast (corner right).
C. A playful interaction with the system, walking along a pre-recorded path in the shape of the acronym AM (bottom right).

The first output aims to show how closely the participant followed the path using sound, by highlighting the segments where they succeeded. This proximity is calculated by measuring the distance between each coordinate of the pre-recorded path and the participant's corresponding position (figure 51).

In the second output, the dashed line indicates the proximity of the pre-recorded path to the detection device over time, shown on a horizontal axis. The solid line reflects the participant's proximity in comparison with the former. The success percentage is calculated by comparing these values throughout the interaction (figure 52.A). The third output tracks the horizontal position over time, with the y-axis (from top to bottom) charting the participant's horizontal movements, using a similar calculation method as the previous one (figure 52.B).

When these three visual outputs are combined, they create a unique and intelligible visual artefact that varies with each participant and interaction, offering a personalised reflection of their experience (figure 53). This visual artefact can also be described as an informal musical score, as reproducing it exactly would result in a similar sound output.

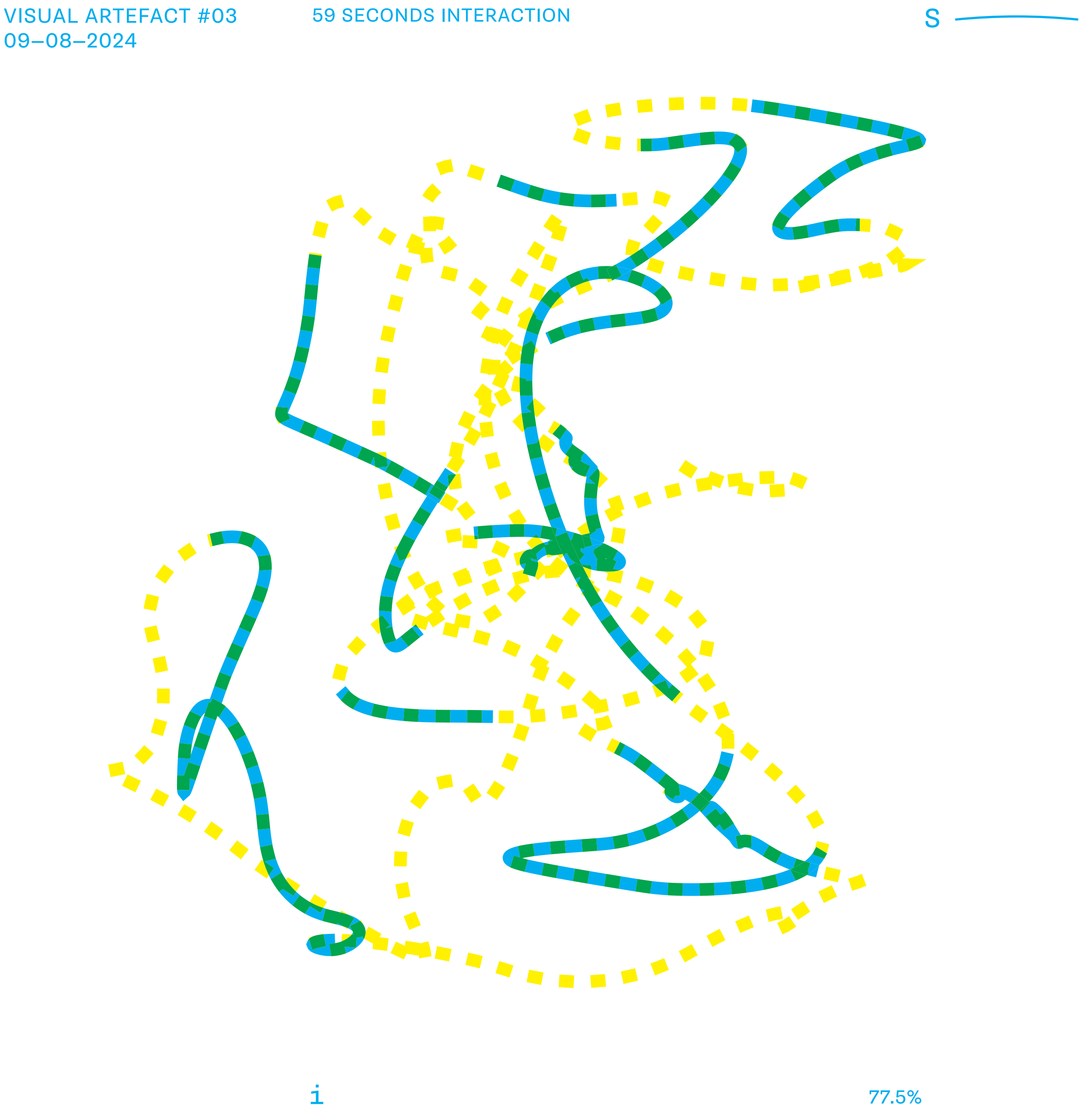
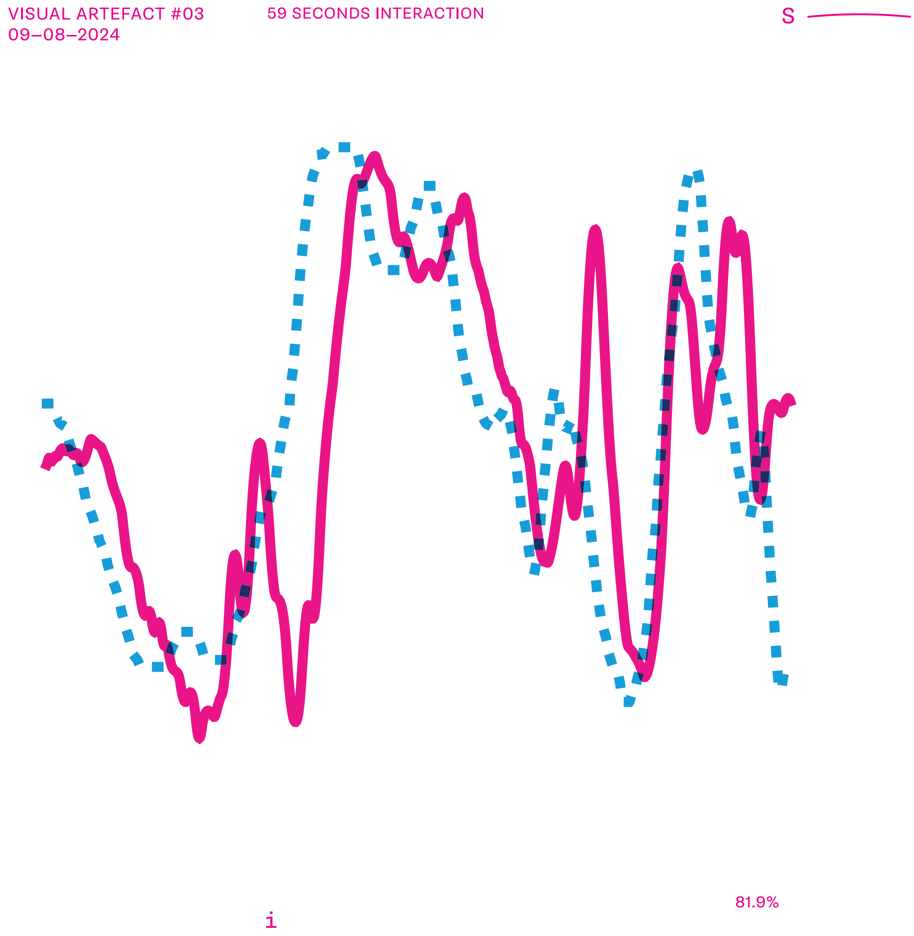


Figure 51.
Final output with pre-recorded path highlighting overlaps with the participant's path.



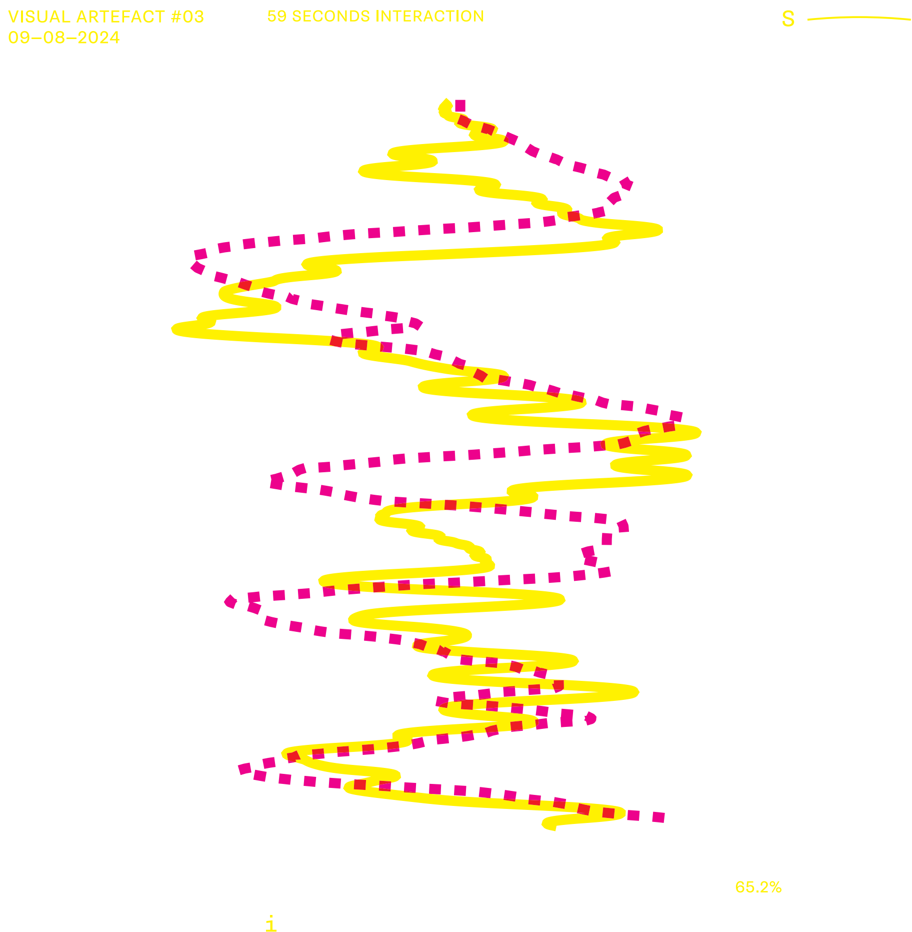


Figure 52.
A. Participant proximity overlapping with pre-recorded positions over the x time axis (left to right).
B. Participant horizontal position overlapping with pre-recorded horizontal positions over the y time axis (top to bottom).

Figure 53.
S—I final artefact example: the overlap of the three outputs.

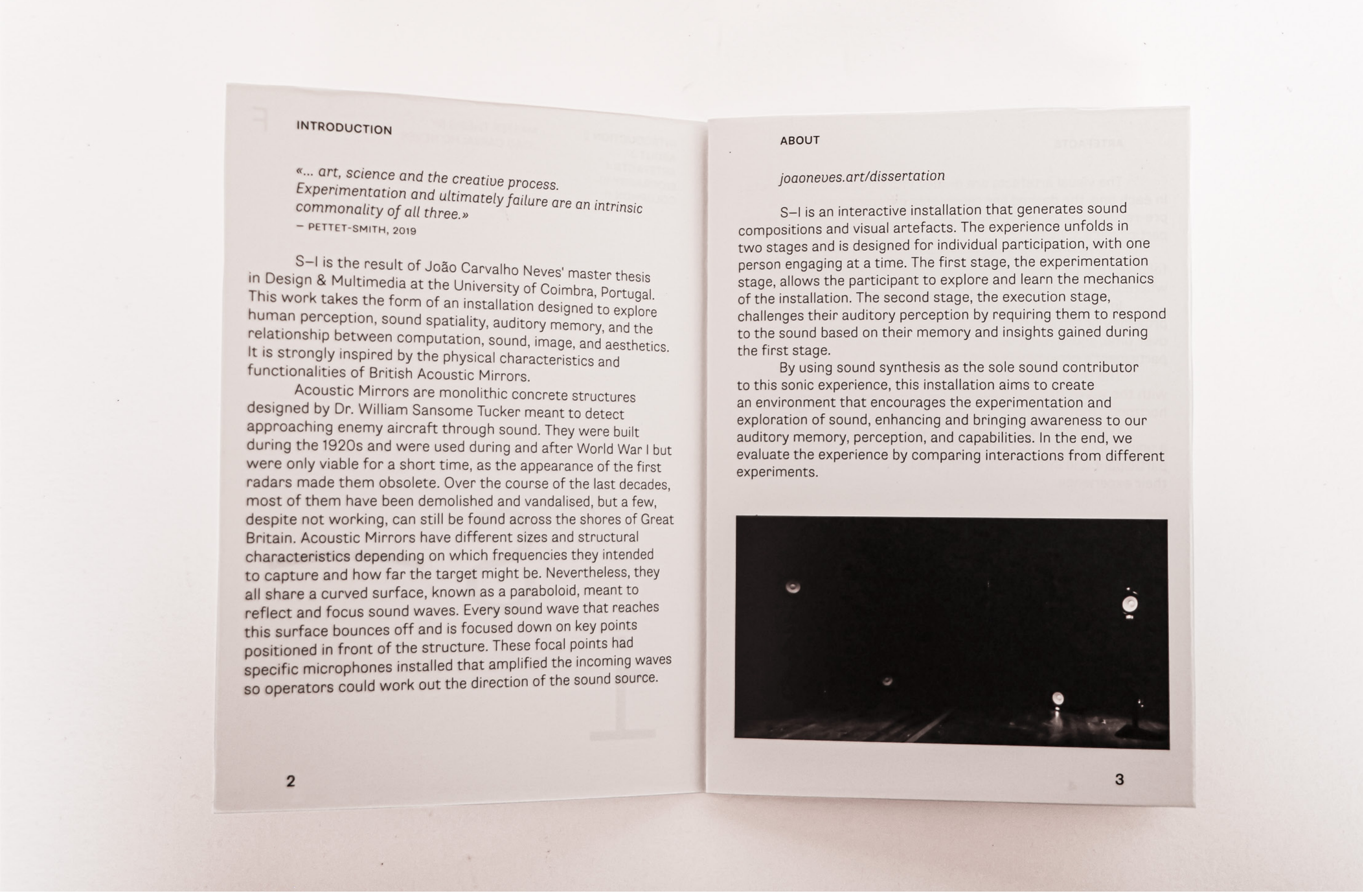
Graphic Material

The visual identity of S—I (Sound, then Image) features a distinct separation between the initials of the acronym, symbolising the temporal gap between the two phenomena. The dash that separates them takes on a curvilinear form, evoking the concavity of AM, which served as inspiration for this design translation. The minimalist approach to the visual identity, as observed in similar Chapter IV — *State of the Art* projects, reflects the nature of such work, establishing a sense of neutrality and a *zero state* before interaction, thereby giving prominence to the experience and its absorption.

The graphic materials[18] are designed to support the installation across various contexts. The program booklet (figures 54.A, 54.B, 54.C & 54.D), presented in booklet format, provides context for the project and assists the audience in its interpretation. The design accommodates the integration of variable information depending on the event and exhibition context.

Figure 54.A.
S—I program booklet's cover.

18. Documents may be consulted at joaoneves.art/dissertation#documents



Figures 54.b. & 54. c.
S—I program booklet's interior.

Figure 54.D.
S—I program booklet's outputs example.

To present the artefacts, we developed a tangible proof of concept for display. By using a light table and printing the different outputs on acetate, it is possible to assemble the final artefact without sacrificing its informative character while maintaining the ability to reference each performance (figures 55.A, 55.B & 55.C). The cyan, magenta, and yellow palette of the outputs allows for the creation of a variety of colours when they are overlapped, adding a dynamic visual element to the audience's interaction. This setup invites the audience to physically and freely engage with the component, enhancing the overall experience.

For wider audiences and larger installation spaces, we envision a more expansive light table that accommodates multiple participants simultaneously. Additionally, we plan to print the artefacts in real time, immediately after participants complete their experience with the sound interaction stages.

Figures 55.a. & 55.b.
S—I visual artifacts: acetate prints displayed on a light table.

Figure 55.c.
S—I visual artifacts: acetate prints displayed on a light table.

Pilot

The installation pilot took place at the facilities of Círculo de Artes Plásticas de Coimbra over a 3 day period. During this time, we addressed and clarified issues related to lighting, speaker layout, overall volume and required hardware.

We created a technical rider[19] that consists of the following items (figure 56):

— Four speakers.
— Four speaker stands.
— One audio interface with 4 channels.
— Four 6-metre XLR-jack cables.
— One computer and peripherals.
— One *Kinect v2*.
— Two extension cords with at least 4 outlets each.
— Five c13 power cables.

After assembling and connecting all the devices, we covered the entrance of the black box with black fabric to prevent any light from coming in and disturbing the experience. We then proceeded to configure the detection system to match the room's layout as mentioned above and adjusted the output volumes. Next to the black box (figure 57), there is a room that we used as a waiting area for testers to wait and be introduced to the installation before entering.

19. The formal document may be consulted at the Appendix section and may be consulted at joaoneves.art/dissertation#documents

Figure 56.
S—I technical rider's items.

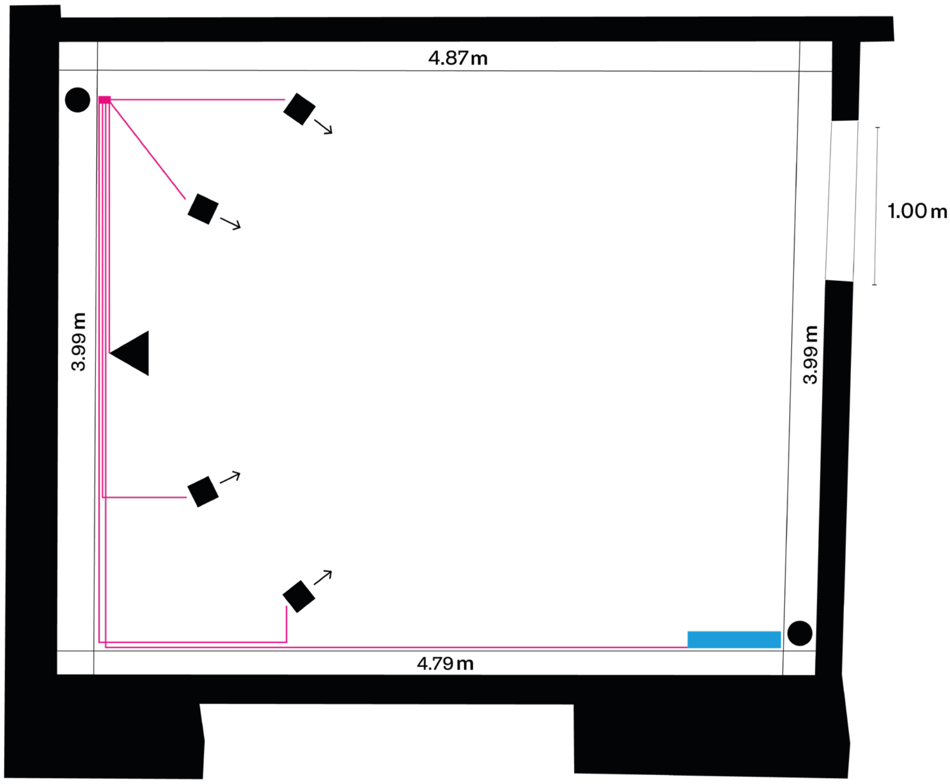


Figure 57.
Installation and waiting room blueprints: the cyan rectangle indicates the location of the computer, the circles mark the power sockets, the squares represent the speakers, and the triangle denotes the detection device.

Evaluation

S—I evaluation comprises a phenomenological approach (Bevan, 2014) combined with an analysis of various interactions through its visual outputs. This evaluation method focuses on describing phenomena as experienced by individuals. It involves the systematic study of perception through guided conversations with participants about their interpretation of the experience, rather than attempting to explain why the phenomena occur. We applied Bevan's (2014) structured procedure by allowing participants to experience the installation in two states: initially without any prior information about the system, and then after being debriefed on its behaviour and reactivity. Similar to the approach of Seiça et al. (2019), in this part of the evaluation our focus is on understanding how the system is experienced as a phenomenon, "uncovering meanings and common themes that users associate while listening to it." Additionally, by comparing and analysing different outputs, we aim to study participants' ability to decode auditory information and recognise patterns and changes. During this trial period, we gathered 15 participants from diverse backgrounds (table 5) to test the installation. Each participant interaction followed a strict procedure consisting of five stages: A) a brief contextualization of the installation; B) the first interaction; C) the first part of the interview, followed by an explanation of the system encodings; D) the second interaction; E) the second and final part of the interview.

Participant	Age	Background	Musical Knowledge
P1	23	Photographer	3
P2	24	Designer	2
P3	24	Designer	1
P3	24	Front-End Developer	1
P5	30	Economist	2
P6	31	Designer	4
P7	32	Educational Psychologist	1
P8	32	Designer	2
P9	32	Photographer	4
P10	32	Developer	3
P11	32	Civil Engineer	2
P12	33	Designer	1
P13	35	Architect	4
P14	40	Musician	5
P15	63	Communication and Visual Culture Educator	2

Table 5.
Participants' age, backgrounds, and musical knowledge assessed using a Likert scale (Likert, 1932).

To begin with, we provided minimal contextualisation of the installation, disclosing only that it was a sonic experience conceptually inspired by AM. The first interaction involved a 5-minute experimentation stage before Stage 2 began. Both the first and second interactions used the same pre-recorded path for all participants to ensure consistency. Subsequently, we started the first part of the interview, establishing a trustworthy and receptive environment that facilitated flowing conversations. The collected information was later transcribed and organised into different categories (table 6, 7 & 8). After the second interaction we concluded the test by conducting a follow-up conversation, which was shorter and was mainly to collect suggestions.

Sensation	Ocurrences	Percentage
Anxiety	4	26.7
Constrained	3	20.0
Discomfort	4	26.7
Foreboding	2	13.3
Relaxed	3	20.0
Sensory	6	40.0
Tension	8	53.3

Table 6.
Total number of occurrences (across all participants) that each sensation was mentioned when talking about the experience.

Experience	Ocurrences	Percentage
Challenging	6	40.0
Confusing	9	60.0
Engaging	4	26.7
Free	2	13.3
Fun	3	20.0
Immersive	12	80.0
Meditative	2	13.3
Mysterious	4	26.7
Surprising	5	33.3

Table 7.
Total number of occurrences (across all participants) that each adjective was used to characterise the experience.

Composition Aesthetics	Ocurrences	Percentage
Volume	12	80.0
Panoramics	2	13.3
Rhythm	4	26.7
Glitches	6	40.0

Table 8.
Total occurrences (across all participants) of each composition component mentioned when describing the interaction.

The initial reactions of all participants after the first interaction were focused on qualifying the experience. Twelve interactions (80%) were described as "immersive". One participant mentioned hearing a motorcycle passing on the street and thought it was happening inside the room, while another participant commented on closing their eyes to follow the auditory stimulus. At this level, we can draw a distinction between participants who sought to understand how the system worked and those who found different types of value in the interaction. The first group includes participants who, on one hand, described the experience as “confusing” (8), and on the other, those who found it “fun” (3) and “engaging” (4). In this context, the ability to recognise the Composition Aesthetics proved significant, as understanding the system demystified the interaction and made it a more straightforward experience. Among the fifteen participants, three felt lost, unable to justify their role despite trying and feeling the need to do so. In this regard, while two of them described the discovery aspect as unintelligible and noted that the presence of a timer, a guide or an explanation was essential to the interaction, the other suggested that some concepts of reactivity should be introduced beforehand so that the experience could unfold more informatively. These opinions support our decision not to reveal the encodings from the start, which might have disoriented the participants. However, this approach was not always well received.

On the other side of the spectrum six participants did not seek to understand how the system worked and instead chose to explore different sounds. In these cases, they described the experience as “mysterious” (4), “meditative” (2), and “free” (2). Throughout the interviews, small associations with the cinematic universe were mentioned, with participants saying that the installation could belong to the world of *Take Shelter* (Nichols, 2011), *Dune* (Villeneuve, 2021), *2001: A Space Odyssey* (Kubrick, 1968) and *Alien* (Scott, 1979). In line with this type of scenario, the most predominant sensations recorded were “tension” (8), “sensory” (6), “anxiety” (4), and “discomfort” (4). Due to the intensity, timbres, vibrations, and lack of light, four participants mentioned that this is not an experience they could endure for long. One participant even mentioned constantly searching for a “safe place” because the sound “penetrated her body.”

Regarding the Composition Aesthetics, we can identify a discrepancy in the mention of volume (12) compared to other components, which is not surprising. Volume, in the form of chords detuning, was always associated with proximity to the system, and four people were able to discover a second layer related to this proximity: the change in the rhythm of the melody, which one participant compared to "a frightened creature we want to calm down," and another to “little bubbles.” The difficulty in decoding the system is confirmed by the experience being described as “challenging” (6). One participant mentioned a challenge on two levels: sensory, in the sense that it requires physical effort, and mental, due to the constant attempt to understand what is happening. A second participant, despite understanding some of the encodings, found it difficult to distinguish and follow them simultaneously. The glitches, which evoked a “danger zone” (6), and the panoramics (2) were rarely mentioned and never associated with specific behaviour.

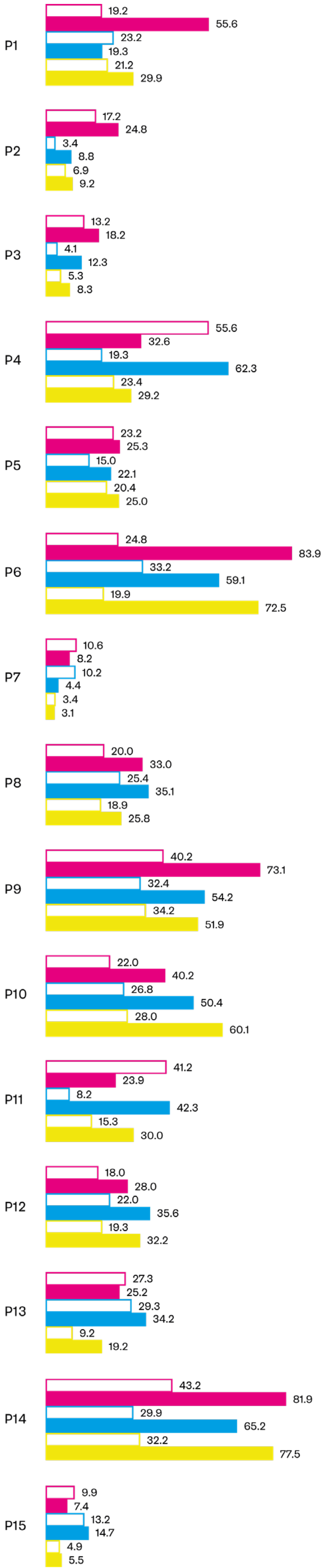


Figure 58.
Participants' performance in percentage. Pink, blue, and yellow represent distance, horizontal position, and overall path success, respectively. Blank bars correspond to the first iteration, and filled bars correspond to the second iteration.

Three participants also explored the system in non-parameterised ways, performing gestures (2), jumping (3), crouching (1), and touching the walls (1). Another participant mentioned that there could be better contextualization regarding what should not be done, such as not touching the speakers or getting too close to the camera, explaining that this self-questioning acts as a barrier to immersion.

Figure 58 displays the percentage of visual artefacts produced by each participant. In a preliminary analysis, we can observe that for participants P2, P3, P7, and P15, the explanation of the system provided little or no help, as they were unable to substantially improve their initial results. With the exception of these cases, sharing the encodings improved performance, although not absolutely. The data reveal a clear difficulty for participants in simultaneously tracking multiple auditory stimuli. Participants P4 and P11, despite improving their horizontal accuracy, neglected the distance component. On the other hand, participant P1 experienced a decrease in horizontal accuracy from the first to the second iteration.

The participants with the most musical knowledge (P6, P9, and P14) were the most successful at following the sound cues. In contrast, P5, who admitted during the interview to recognising the reactivity to distance early on, was unable to apply this knowledge either due to issues with perception or interpretation.

This collection reveals that, in this system, distance was the easiest factor to understand. Figure 59 are two examples that illustrate learning ability during the second iterations, as they show closer alignment with the intended path in the final segment of the interaction. For future work, we are interested in exploring how additional interactions affect performance for each participant, in order to understand the value of the installation as an auditory training tool.

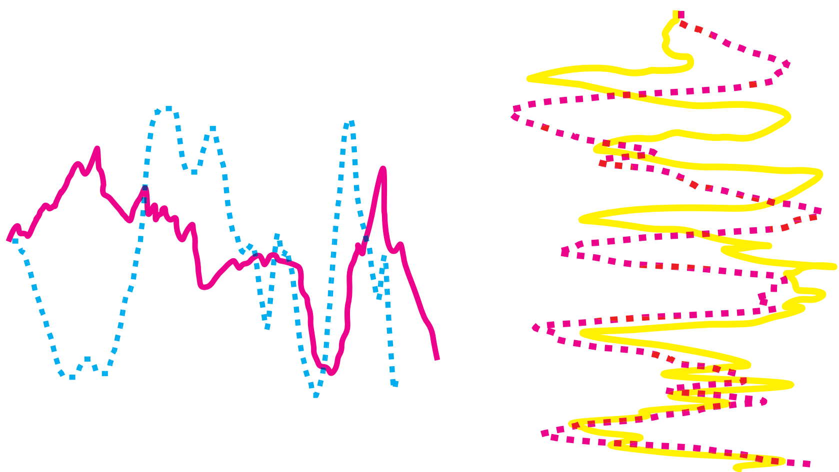


Figure 59. Tendency to learn during the interaction: proximity (left) horizontal position (right)

Notes

The development of this installation went through several stages, including an extended period of conceptualization where the type of interaction and reactivity were redefined. This process involved investing in and deepening knowledge in areas such as sound design, sound programming and object-oriented programming, HCI (Human-Computer Interaction), computer vision, and graphic design. Once realised, the project also provided some hands-on experience with the physical assembly of such installations. Documentation included photography, front-end development, and book binding techniques.

In future iterations of the project, we aim to refine some of the issues identified, such as providing clearer distinctions between different phases (e.g., using visual markers or slightly adjusting the lighting in the room to indicate transitions) and clarifying what should be avoided (e.g., getting too close to the detection device, which may cause interference). We also plan to improve the nuances of the composition, making the horizontal alignment more perceptive and rewarding (e.g., by adjusting the spatial positioning or adding visual cues to enhance the user's spatial awareness).

We intend for the installation to operate entirely autonomously, self-resolving potential bugs and interaction issues that occasionally arise due to simultaneous detection of multiple bodies. To achieve this, we will implement more robust error-handling mechanisms and ensure that the system can dynamically adjust to varying user interactions. Regarding spatialization, (if we gain access to the necessary technology and resources) we will explore testing the system with a more complex and extensive sound setup (e.g., expanding to an 8-speaker surround system) once the quadraphonic system is more thoroughly used and studied.

Chapter VII
Conclusion

In this dissertation, we explored the relationship between sound, computing, image and aesthetics through the study of Acoustic Mirrors and the subsequent development of S—I, an interactive audiovisual installation. The conceptualisation of the presented installation arises alongside a literature review and research on authors and projects within various fields such as Sound Art, Music Notation, Visualisation, Sonification, Computational Creativity, and New Media Art. The preliminary work, fundamentally inspired by the work of Ryoji Ikeda, contributed to establishing the universe in which S—I acts and seeks to enhance through human-computer interaction.

The implemented system offers an auditory experience that unfolds in two stages, designed for individual participation. Initially, the audience discovers the encodings through experimentation, and then, relying on their memory, they follow the sound and the paths it suggests, testing their abilities and perception. This interaction results in the generation of a visual artefact composed of three distinct visual outputs that serve both as evaluation objects and, when overlapped as a unique and informal musical score.

Our aim was to create an environment for sound exploration and to investigate and enhance auditory perception and the audience's capacity for spatialization through memory and sensory three-dimensional experiences. We tested the installation over a three-day period at the facilities of Círculo de Artes Plásticas de Coimbra, where 15 participants individually experienced both stages and provided feedback on their feelings and findings, guided by Bevan's (2014) phenomenological structured procedure.

Although our findings cannot be generalised, as they are based on experiences within a specific system and scenario, we observed that it is easier for participants to recognise and respond to sound intensity rather than to the sound's origin when multiple sound sources are present. Additionally, we noted difficulties in interpreting simultaneous auditory stimuli. This dissertation contributes to this field of research as a case study, presenting a method and approach that can be adapted and reinvented with numerous other possibilities in terms of interaction, sound composition, and visual feedback. It also serves as a contribution to the normalisation of computing in artistic contexts, creating a novel space for audience discovery that is capable of provoking sensory and emotional responses, as well as providing tactile visual artefacts for an experience initially conceived in one's mind.

The evaluation of the system is limited by the number of participants who tested it. A larger sample size could yield new insights. We plan to expand this sample size and address the issues identified thus far, improving feedback on the different stages of interaction and making the system more autonomous and capable of resolving sporadic anomalies. This may be achieved by adding visual markers or slightly adjusting the room's lighting to indicate transitions between stages and clarify what should be avoided, such as getting too close to the detection device.

In future iterations, the generative compositions may also be further explored. Beyond using more complex sound systems, the use of panoramic sound could be enhanced, making the relative component (horizontality) more apparent, thereby improving spatial awareness.

By refining and expanding upon the approaches developed here, we hope to inspire new explorations that continue to push the boundaries of sensory and interactive art, offering audiences richer and more engaging experiences.

References

Albers, J. (1968). *Homage to the Square.*

Aldrich, N. B. (2003). *What is sound art?*. Unpublished manuscript, University of Maine, Brooks, ME, USA.

Anadol, R. (2018). *Pladis: Data Universe.* Retrieved September 1, 2024 from refikanadol.com/works/pladis-data-universe/

Autogena, L. (2023). *Havobservatoriet.* Retrieved September 1, 2024 from autogena.org/work/havobservatoriet

Balme, H. (2020). *Audiovisual hallucinations in the synesthetic films of Jordan Belson and James Whitney.* In Music and Synesthesia (p. 150).

Bandt, R. (2006). *Sound installation: Blurring the boundaries of the eye, the ear, space, and time.* Contemporary Music Review, 25(4), 353-365.

Basquin, K. (2020). *Mary Ellen Bute: Pioneer animator.* In Personal and professional life of Mary Ellen Bute.

BBC News. (2019). *In pictures: The year in wildlife photos.* Retrieved September 1, 2024 from www.bbc.com/news/in-pictures-46348917

Benadon, F. (2007). *A circular plot for rhythm visualization and analysis.* Music Theory Online, 13(3).

Bevan, M. T. (2014). *A method of phenomenological interviewing. Qualitative Health Research*, 24(1), 136–144.

Bhola, R. (2013). *Binocular vision.* Retrieved September 1, 2024 from webeye.ophth.uiowa.edu/eyeforum/tutorials/bhola-binocularvision.htm

Brad Mehldau Trio. (1997). *The art of the trio volume 3 [CD].* Warner Bros. Records.

Buccellati, G. (2003). *Hurrian music.* Retrieved September 1, 2024 from www.urkesh.org/urkeshpublic/music.htm

Cage, J. (1952). *4'33''.* C.F. Peters.

Cage, J. (1961). *Atlas eclipticalis.* C.F. Peters.

Calovski, Y. (2022). *Le Corbusier - Buckminster Fuller: Architecture between music et poetry.* La Première Rue.

Cardew, C. (1967). *Treatise.* The Gallery Upstairs Press.

Carvalhais, M. (2010). *Towards a model for artificial aesthetics: Contributions to the study of creative practices in procedural and computational systems.* Universidade do Porto.

Carvalhais, M. (Ed.). (2022). *The Book of X: 10 years of computation, communication, aesthetics and X.* i2ADS: Research Institute in Art, Design and Society.

Castro, R. (2022). In M. Carvalhais & P. Tudela (Eds.), Installations.

Chion, M. (1994). *Audio-vision: Sound on screen.* Columbia University Press.

Choi, J. H. (2021). *Lausanne Dream House: Sound and Light Environment.* ECAL. Retrieved September 1, 2024 from ecal.ch/en/feed/events/974/dream-house-exhibition-by-la-monte-young-marian-zazeela-jung-hee-choi/

Cipriani, A., & Giri, M. (2013). *Introduction to sound synthesis.* In A. Cipriani & M. Giri (Eds.), Electronic music and sound design: Theory and practice with Max and MSP (2nd ed., Vol. 1, pp. 1-47). Rome, Italy.

Cox, C. (2021). *Sound art and time.* In J. Grant, J. Matthias, & D. Prior (Eds.), The Oxford handbook of sound art. Oxford University Press.

Dennis, B. (1991). *Cardew's 'Treatise' (mainly the visual aspects).* Tempo, (177).

Dubberly, H. (2005). *How do you design?*. San Francisco: Dubberly Design Office.

Dyson, F. (2009). *Sounding new media: Immersion and embodiment in the arts and culture.* University of California Press.

Fieldsteel, E. (2024). *SuperCollider for the creative musician.* Oxford University Press.

Foo, B. (2014). *Two trains: Sonification of income inequality on the NYC subway.* Data Driven DJ. Retrieved September 1, 2024 datadrivendj.com/tracks/subway/

Galerie Chantal Crousel. (2021). *Tarek Atoui, The Whisperers, 2021* [Video]. YouTube. Retrieved September 1, 2024 from youtube.com/watch?v=MwI1RPpB4OU

Garcia, J. F. (2022). *Rain Beats.* In Artworks: ISEA2022 Barcelona: 27th International Symposium on Electronic Art (Catalogue).

Gardiner, I. (2007). *Oskar Fischinger. Music, Sound, and the Moving Image*, 1(1), 117–118.

Gerstner, K. (2007). *Designing programmes* (1st ed.). Lars Müller Publishers. (Original work published 1964)

Ghez, S. (Director). (2022). *Xenakis revolução: O construtor de sons*. RTP Play.

Griffiths, A. G. F., Kemp, K. M., Matthews, K., Garrett, J. K., & Griffiths, D. J. (2017). *Sonic kayaks: Environmental monitoring and experimental music by citizens*. PLOS Biology, 15(11), Article e2004044.

Hansen, C. H. (1951). *Fundamentals of acoustics*. University of Adelaide.

Helfer, M. (2009). *Sound source localization with acoustic mirrors*. Paper presented at NAG / DAGA 2009, Rotterdam, Netherlands.

Hermann, T., Hunt, A., & Neuhoff, J. (Eds.). (2011). *The sonification handbook*.

Higgins, D. (2018). *Intermedia, Fluxus, and The Something Else Press: Selected writings by Dick Higgins* (S. Clay & K. Friedman, Eds.). Siglio.

Hope, C. (n.d.). *The art of visualizing music: A brief history of graphic notation*. David Hall. Retrieved September 1, 2024 from davidhall.io/visualising-music-graphic-scores

House, B. (2012). *Quotidian record*. Retrieved September 1, 2024 brianhouse.net/works/quotidian_record/

Ikeda, R. (2005). *Dataplex* [Album]. Raster-Noton.

Ikeda, R. (2006). *Datamatics*. Retrieved September 1, 2024, from ryojiikeda.com/project/datamatics/

Ikoniadou, E. (2014). *The rhythmic event: Art, media, and the sonic*. The MIT Press.

Jackson, R. R. (1982). *The behavior of communicating in jumping spiders (Salticidae).* In P. Witt & J. Rovner (Eds.), Spider communication: Mechanisms and ecological significance (pp. 213-247). Princeton University Press.

John Cage Trust. (2016). *John Cage*. Retrieved September 1, 2024 from johncage.org/

Keith Jarrett Trio. (1997). *Tokyo '96* [CD]. ECM Records.

Kester, G. H. (2011). *The One and the Many: Contemporary Collaborative Art in a Global Context*. Durham, NC: Duke University Press.

Kim, T. (1996). *Effects of presence on memory and persuasion* (Unpublished doctoral dissertation). University of North Carolina, Chapel Hill.

Knudsen, E. I. (2007). *Fundamental components of attention*. Annual Review of Neuroscience, 30(1), 57-78.

Kubrick, S. (Director). (1968). *2001: A space odyssey*. MGM.

Licht, A. (2019). *Sound art revisited*. Bloomsbury Publishing.

Likert, R. (1932). *A technique for the measurement of attitudes*. Archives of Psychology, 22 140, 55.

Lombard, M., & Ditton, T. (1997). *At the heart of it all: The concept of presence*. Journal of Computer-Mediated Communication, 3(2).

Manovich, L. (2002). *The language of new media*. The MIT Press.

Manzolli, J. (2022). *A auto-organização como agente do processo criativo e laboratório de significação musical*. Universidade Estadual de Campinas, Campinas, São Paulo, Brazil.

Mayer, A. M. (1880). *Topophone (U.S. Patent No. 224,199)*. United States Patent and Trademark Office.

Minkin, S. (2018). *Alvin Lucier -- I Am Sitting In A Room* [Video]. YouTube. Retrieved September 1, 2024 from youtu.be/peahfxDx8m8

Moritz, W. (2019). *John Whitney: Technological visionary, animation pioneer.* In J. Coorey, A. Murnieks, H. Shaw, & R. Tegtmeyer (Eds.)

Mrugala, D. (n.d.). *David Mrugala's profile*. Fine Art America. Retrieved September 1, 2024 from https://fineartamerica.com/profiles/david-mrugala

Neves, J., Martins, P., Manzolli, J., Cardoso, F. A., Seiça, M., & Rela, M. Z. (2023). *Soundscape4DEI as a model for multilayered sonifications*. In Proceedings of the 16th International Symposium on CMMR (pp. 221-232). University of Coimbra, CISUC, Department of Informatics Engineering. Tokyo, Japan.

Nichols, J. (Director). (2011). *Take shelter*. Sony Pictures Classics.

Nicolai, C. (2007). *Static balance*. Carsten Nicolai. Retrieved September 1, 2024 from carstennicolai.de/?c=works&w=static_balance

Nilsson, D. E., Gislén, L., Coates, M. M., Skogh, C., & Garm, A. (2005). *Advanced optics in a jellyfish eye.* Nature, 435(7039), 201-205.

Olshausen, B. (2020). *Visual perception and the brain* [Video]. YouTube. In P. Scaruffi (Producer). UC Berkeley. Retrieved September 1, 2024 from youtu.be/x8srRJhKw2A

Palmer, M. T. (1995). I*nterpersonal communication and virtual reality: Mediating interpersonal relationships.* In F. Biocca & M. R. Levy (Eds.), Communication in the age of virtual reality (pp. 277–302). Lawrence Erlbaum.

Pošćić, A., & Kreković, G. (2020). *On the human role in generative art: A case study of AI-driven live coding.* Journal of Science and Technology of the Arts, 12, 45–62.

Quarrick, G. (1989). *Our sweetest hours: Recreation and the mental state of absorption.* McFarland.

Queneau, R. (1961). *Cent mille milliards de poèmes.* Gallimard.

Random International. (2023). *Living Room Variation 1.* Retrieved September 1, 2024 from random-international.com/living-room-variation-1/

Reichardt, J. (Ed.). (1968). *Cybernetic serendipity: The computer and the arts.* Studio International.

Revill, G. (2016). *How is space made in sound? Spatial mediation, critical phenomenology and the political agency of sound.* Progress in Human Geography, 40(2), 240-256.

Reynolds, C. (2002). *Steering behaviors for autonomous characters.* Journal of Graphics Tools.

Romero, J., & Machado, P. (Eds.). (2007). *The art of artificial evolution: A handbook on evolutionary art and music.* Springer.

Rosanova, O. (1918). *Composition, 1918–1918.*

Rush, D. (2019). *Dasha Rush presents: Aurora cerebralis* [Video]. Vimeo. Retrieved September 1, 2024 from https://vimeo.com/357400767

Russian Cultural Heritage. (n.d.). *Sintezator-ans. Culture.ru.* Retrieved September 1, 2024 from ar.culture.ru/en/subject/sintezator-ans#

Russolo, L. (1967). *The art of noise: Futurist manifesto, 1913.* New York: Something Else Press.

Sá, A. (2016). *A perceptual approach to audio-visual instrument design, composition, and performance.* Goldsmiths, University of London.

Sá, A. (2017). *Designing musical expression.* In Proceedings of xCoAx 2017, Lisbon, Portugal.

Sauer, T. (Ed.). (2009). *Notations 21.* Mark Batty Publisher.

Schafer, R. M. (1986). *Snowforms* [Musical score]. Arcana Editions.

Schäffer, P. (1966). *Traité des objets musicaux: Essai interdisciplines.* Éditions du Seuil.

Schütze, P. (2013). *Audio visual.* Retrieved September 1, 2024 from frieze.com/article/audio-visual

Schwarz, D. (2011). *State of the art in sound texture synthesis.* In Proceedings of the 14th International Conference on Digital Audio Effects (DAFx 2011).

Scott, R. (Director). (1979). *Alien.* 20th Century Fox.

Seiça, M., Martins, P., Roque, L., & Cardoso, A. (2019). *A sonification experience to portray the sounds of Portuguese consumption habits.* In Proceedings of the 25th International Conference on Auditory Display (ICAD 2019). Northumbria University.

Shiffman, D. (2012). *The nature of code.* The Nature of Code. Retrieved September 1, 2024 from natureofcode.com

Shim, K. (2020). *Computational approach to graphic design.* The International Journal of Visual Design.

Sims, K. (1994). *Evolving virtual creatures.* Thinking Machines Corporation.

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). *Visual dominance and attention: The Colavita effect revisited.* Perception & Psychophysics, 69(5), 673-686.

Smirnov, A., & Pchelkina, L. (2011). *Russian pioneers of sound art in the 1920s.* In Catalogue of the exhibition 'Red Cavalry: Creation and Power in Soviet Russia between 1917 and 1945' (La Casa Encendida, Madrid).

Smirnov, A. (2013). *Sound in Z: Experiments in sound and electronic music in early 20th century Russia.* Koenig Books in partnership with Sound and Music.

Snyder, B. (2001). *Music and memory: An introduction.* MIT Press.

Solomos, M. (2018). *From sound to sound space, sound environment, soundscape, sound milieu or ambiance…* Paragraph, 41(1), 95-109.

Spatial Music Forum. (2021). *Le Corbusier / Varèse / Xenakis: Philips Pavilion* [Video]. YouTube. Retrieved September 1, 2024 from youtube.com/watch?v=P-a5UUiyyWY

Stockhausen, K. (1959). *Tutti* [Musical score]. Stockhausen-Verlag.

Suzuki, Y. (2022). *Ambient Machine*. Dezeen. Retrieved September 1, 2024 from dezeen.com/2022/05/23/yuri-suzuki-ambient-machine-sound-conditioner/

Synthestruct. (2022). *VAST | Audiovisual experience [U-108] Taiwan [clips]* [Video]. Vimeo. Retrieved September 1, 2024 from vimeo.com/676709864

TED Institute. (2013). *Achin Bhowmik: Interactive experiences with perceptual computing* [Video]. YouTube. Retrieved September 1, 2024 from youtu.be/wrSE8TcYlHU

TheGreenEyl. (2021). *Found in Translation*. The Green Eyl. Retrieved September 1, 2024 from www.thegreeneyl.com/found-in-translation

Tudela, P., & Carvalhais, M. (2022). *Installations / Instalações*. i2ADS.

Vancouver Chamber Choir. (2013). *Snowforms - R. Murray Schafer* [Video]. YouTube. Retrieved September 1, 2024 from youtu.be/GiOhtgR1T0k

van der Voort, A. W. M., & Aarts, R. M. (2009). *Development of Dutch sound locators to detect airplanes (1927–1940)*. In Proceedings of NAG/DAGA 2009. Rotterdam, The Netherlands.

Villeneuve, D. (Director). (2021). *Dune*. Warner Bros. Pictures.

Warb, N. (1939). *Rythme / Rhythm.*

Whitney, J. (1968). *Experiments in motion graphics.* IBM Corporation / IBM Direct / IBM K-12 Assist Group / IBM Publications.

Widmer, G., & Zanon, P. (2004). *Automatic recognition of famous artists by machine: Recognising famous musicians (concert pianists), based on their style of playing.*

Xenakis, I. (1955). *Study for Pithoprakta* [Photograph]. Iannis Xenakis Archives, Bibliothèque nationale de France, Paris.

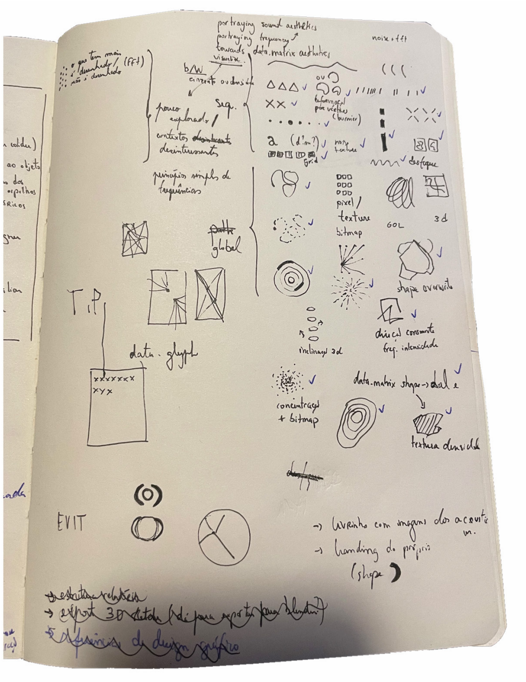
Ya-Ling, C. (2019). *Erratum musical, 1913*. Retrieved September 1, 2024 from toutfait.com/erratum-musical-1913

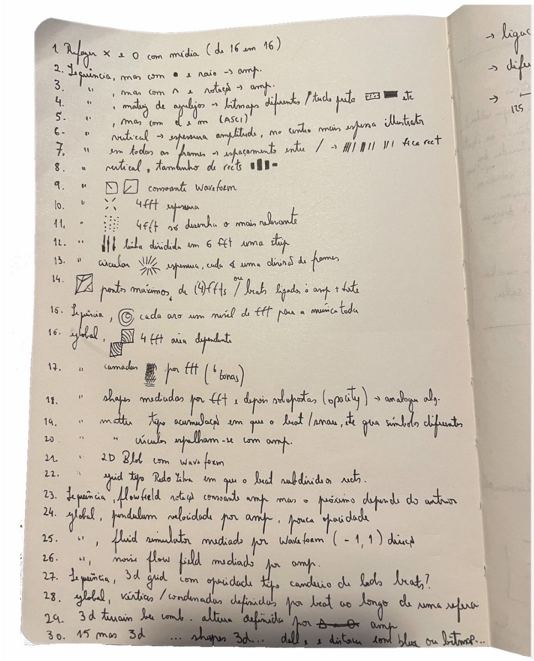
Young, L. M. (2018). *Lecture at Wesleyan University, November 5, 1996*. In A. Lucier (Ed.), Eight lectures on experimental music (pp. 59–79). Wesleyan University Press.

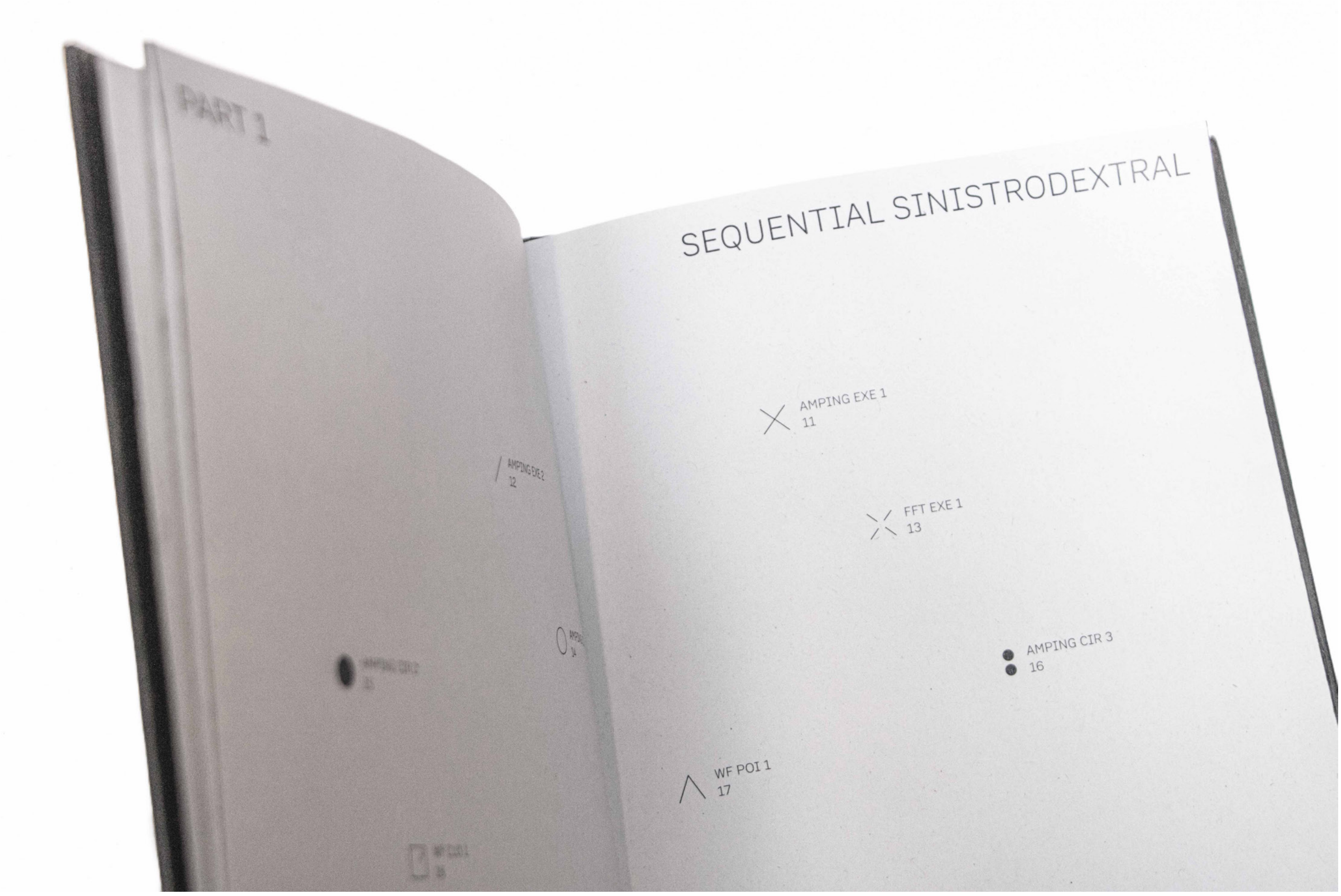
Zazeela, M. (1993). *Dream House: Sound and Light Environment*. MELA Foundation. Retrieved September 1, 2024 from melafoundation.org/dream02.htm

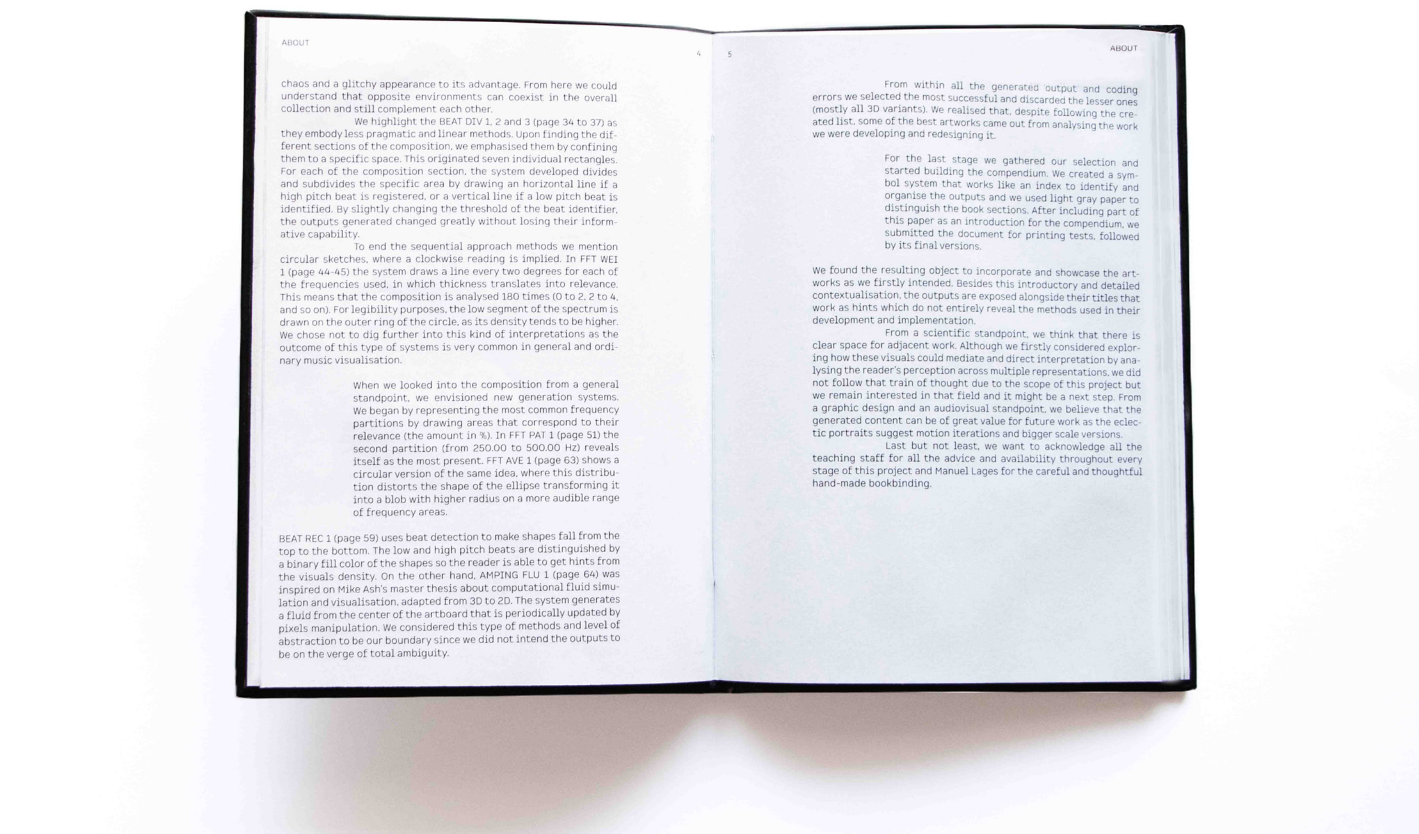
Zhang, W. (2023). *Wayfarer Series. Zhang Weidi*. Retrieved September 1, 2024 from zhangweidi.com/wayfarer

Appendix








ABOUT

4

chaos and a glitchy appearance to its advantage. From here we could understand that opposite environments can coexist in the overall collection and still complement each other.

We highlight the BEAT DIV 1, 2 and 3 (page 34 to 37) as they embody less pragmatic and linear methods. Upon finding the different sections of the composition, we emphasised them by confining them to a specific space. This originated seven individual rectangles. For each of the composition section, the system developed divides and subdivides the specific area by drawing an horizontal line if a high pitch beat is registered, or a vertical line if a low pitch beat is identified. By slightly changing the threshold of the beat identifier, the outputs generated changed greatly without losing their informative capability.

To end the sequential approach methods we mention circular sketches, where a clockwise reading is implied. In FFT WEI 1 (page 44-45) the system draws a line every two degrees for each of the frequencies used, in which thickness translates into relevance. This means that the composition is analysed 180 times (0 to 2, 2 to 4, and so on). For legibility purposes, the low segment of the spectrum is drawn on the outer ring of the circle, as its density tends to be higher. We chose not to dig further into this kind of interpretations as the outcome of this type of systems is very common in general and ordinary music visualisation.

When we looked into the composition from a general standpoint, we envisioned new generation systems. We began by representing the most common frequency partitions by drawing areas that correspond to their relevance (the amount in %). In FFT PAT 1 (page 51) the second partition (from 250.00 to 500.00 Hz) reveals itself as the most present. FFT AVE 1 (page 63) shows a circular version of the same idea, where this distribution distorts the shape of the ellipse transforming it into a blob with higher radius on a more audible range of frequency areas.

BEAT REC 1 (page 59) uses beat detection to make shapes fall from the top to the bottom. The low and high pitch beats are distinguished by a binary fill color of the shapes so the reader is able to get hints from the visuals density. On the other hand, AMPING FLU 1 (page 64) was inspired on Mike Ash's master thesis about computational fluid simulation and visualisation, adapted from 3D to 2D. The system generates a fluid from the center of the artboard that is periodically updated by pixels manipulation. We considered this type of methods and level of abstraction to be our boundary since we did not intend the outputs to be on the verge of total ambiguity.

5

ABOUT

From within all the generated output and coding errors we selected the most successful and discarded the lesser ones (mostly all 3D variants). We realised that, despite following the created list, some of the best artworks came out from analysing the work we were developing and redesigning it.

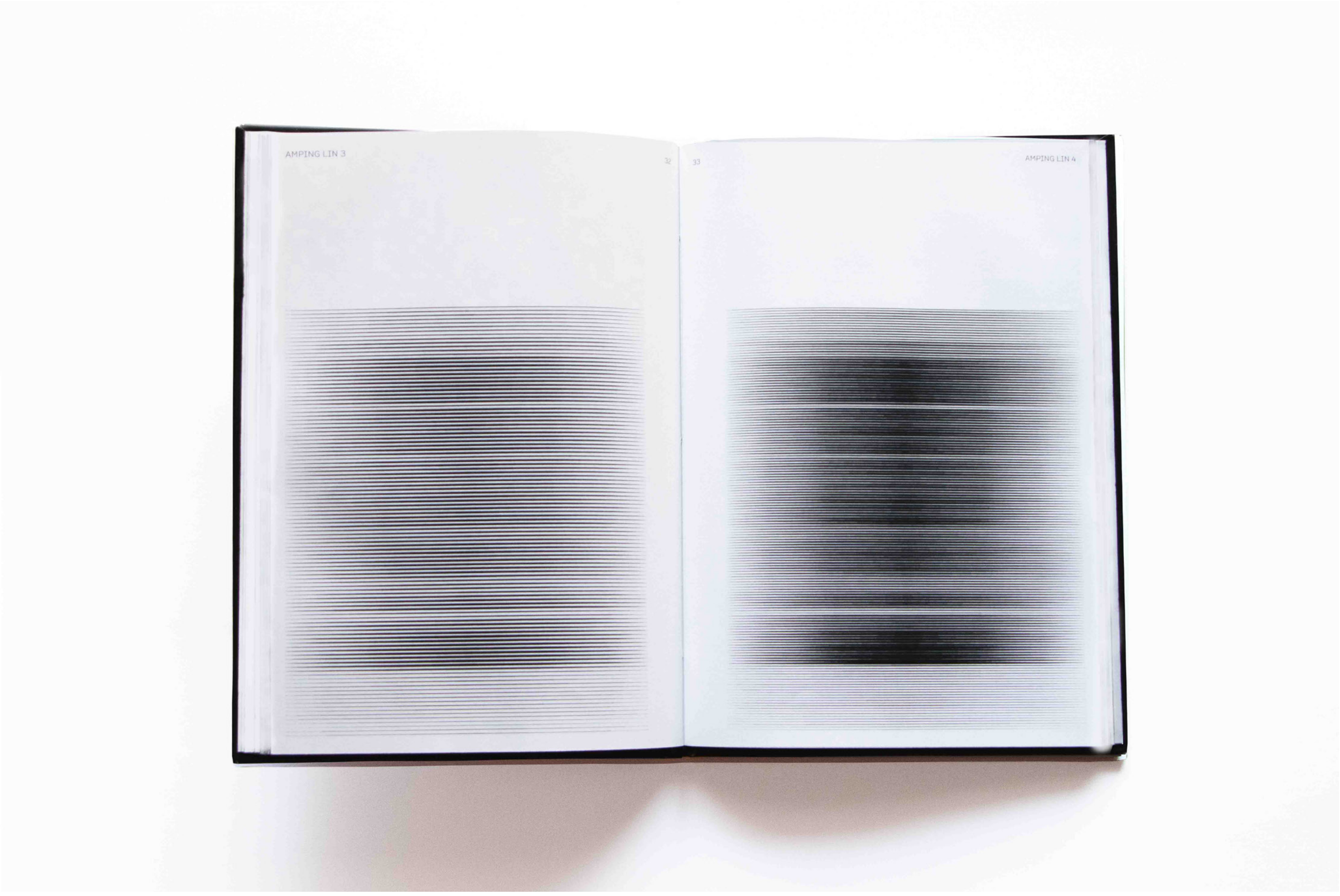
For the last stage we gathered our selection and started building the compendium. We created a symbol system that works like an index to identify and organise the outputs and we used light gray paper to distinguish the book sections. After including part of this paper as an introduction for the compendium, we submitted the document for printing tests, followed by its final versions.

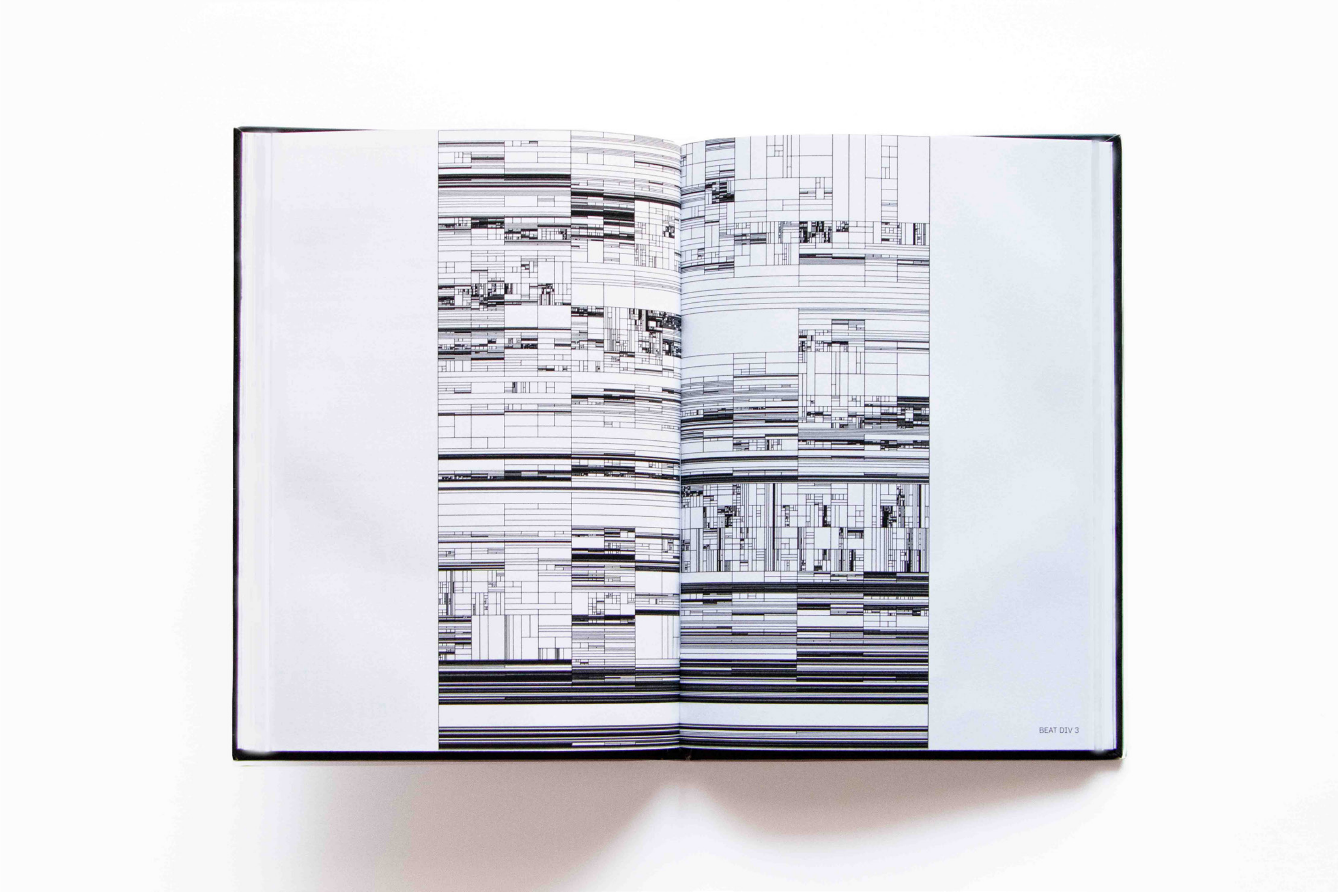
We found the resulting object to incorporate and showcase the artworks as we firstly intended. Besides this introductory and detailed contextualisation, the outputs are exposed alongside their titles that work as hints which do not entirely reveal the methods used in their development and implementation.

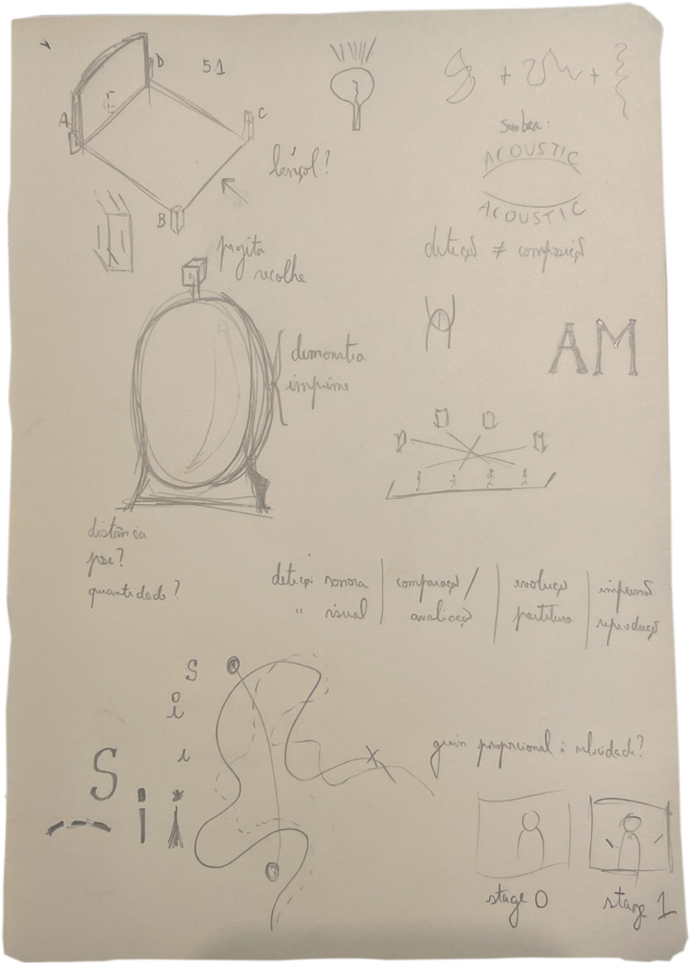
From a scientific standpoint, we think that there is clear space for adjacent work. Although we firstly considered exploring how these visuals could mediate and direct interpretation by analysing the reader's perception across multiple representations, we did not follow that train of thought due to the scope of this project but we remain interested in that field and it might be a next step. From a graphic design and an audiovisual standpoint, we believe that the generated content can be of great value for future work as the eclectic portraits suggest motion iterations and bigger scale versions.

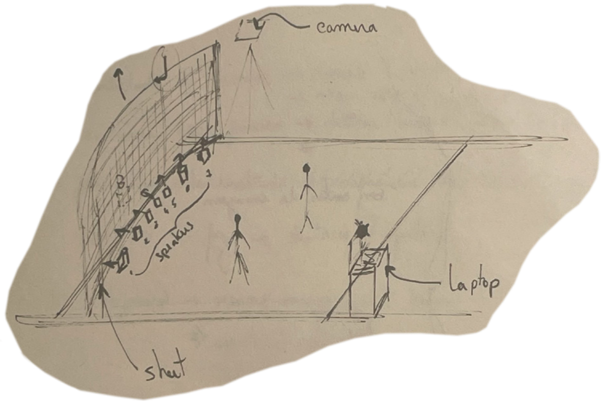
Last but not least, we want to acknowledge all the teaching staff for all the advice and availability throughout every stage of this project and Manuel Lages for the careful and thoughtful hand-made bookbinding.

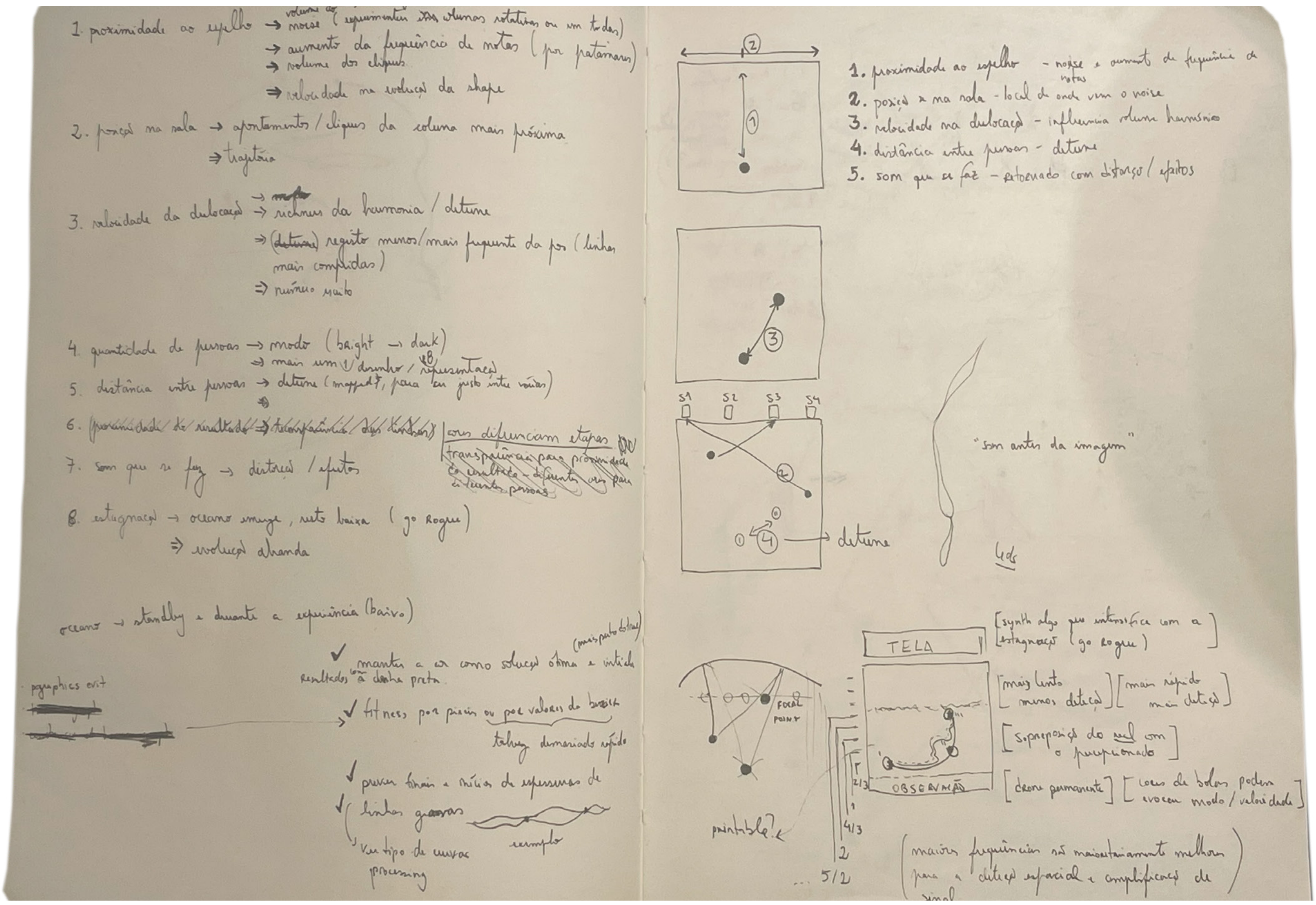

AMPING REC 2


AMPING LIN 3
AMPING LIN 4

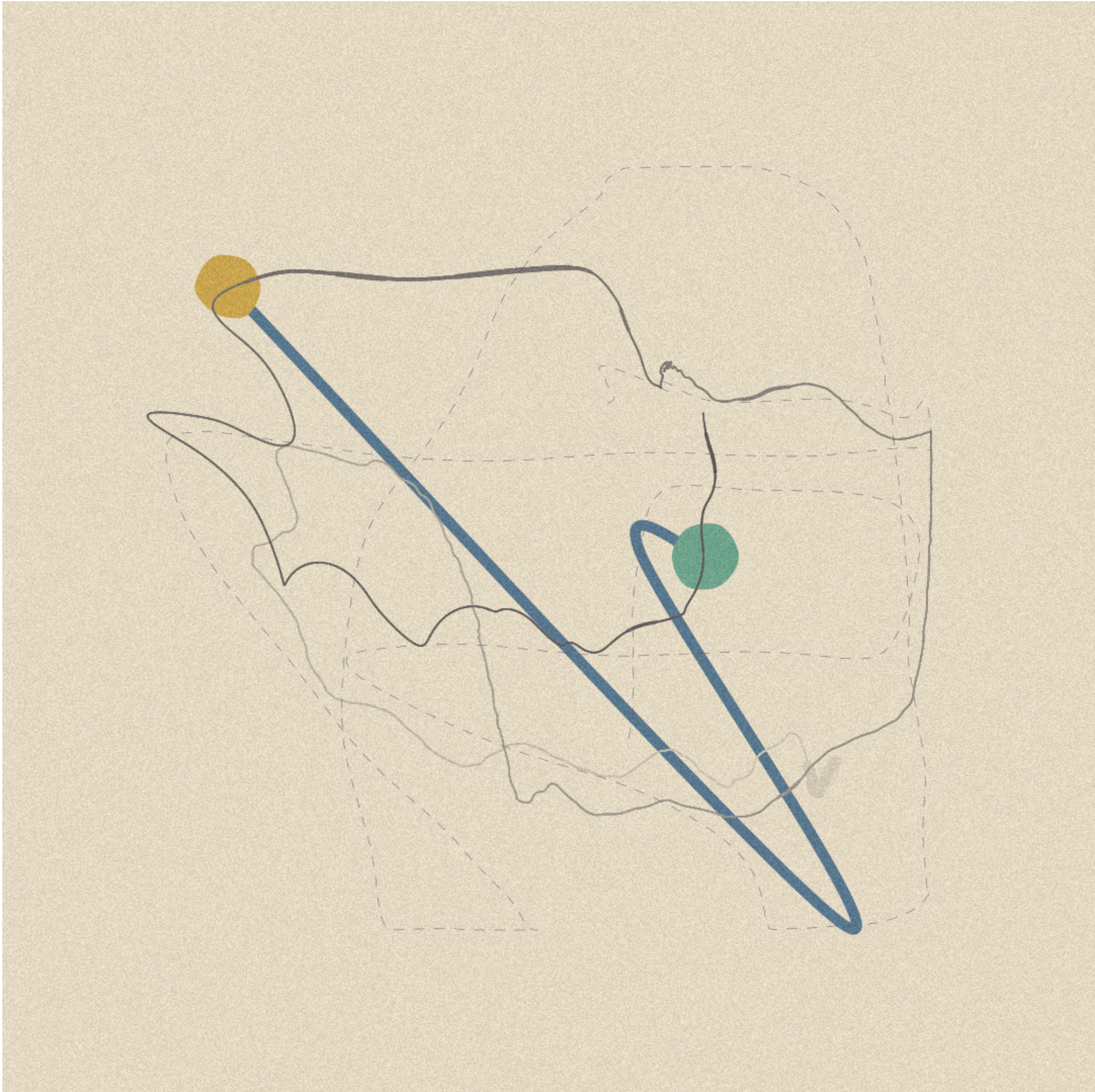

BEAT DIV 3



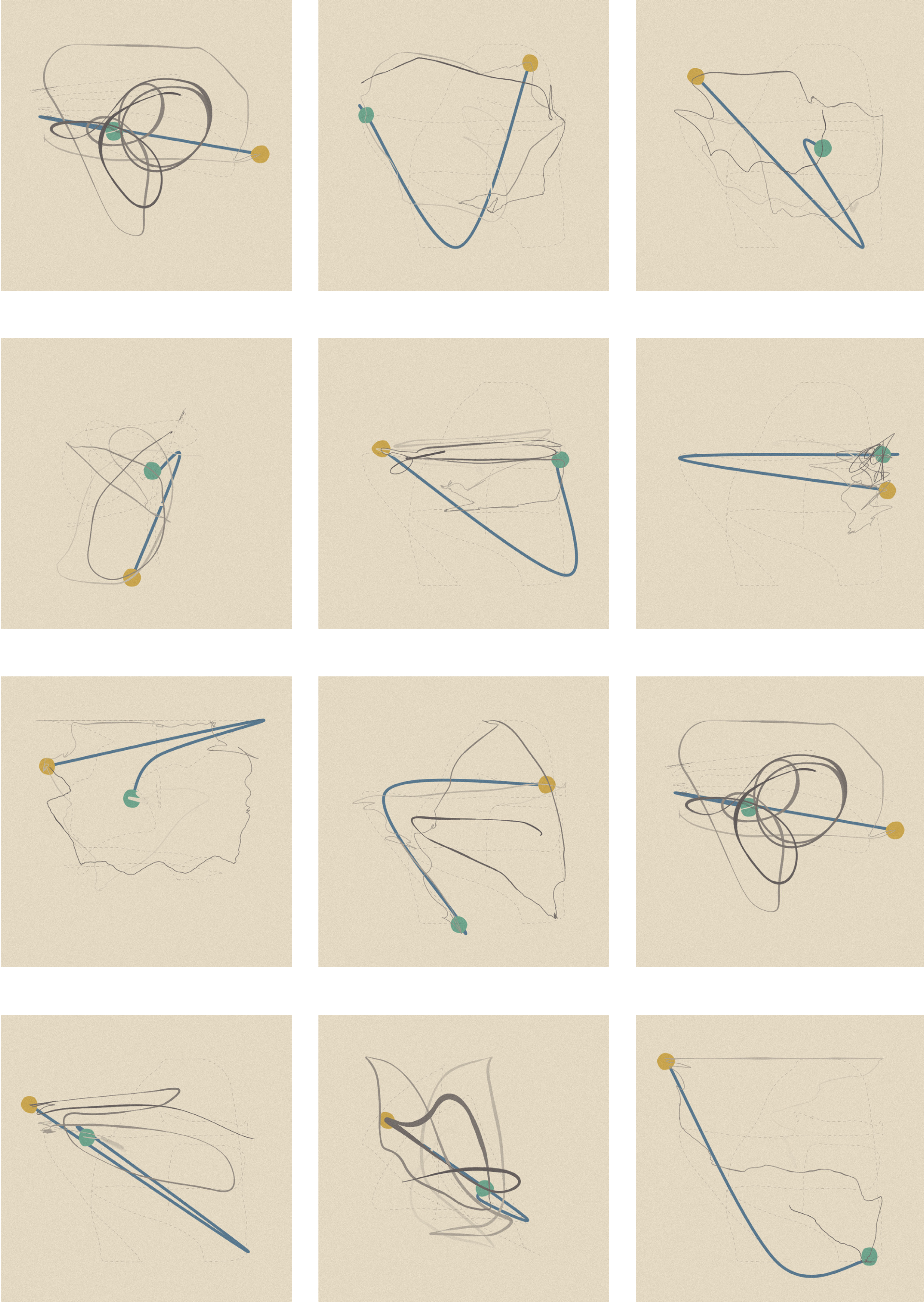




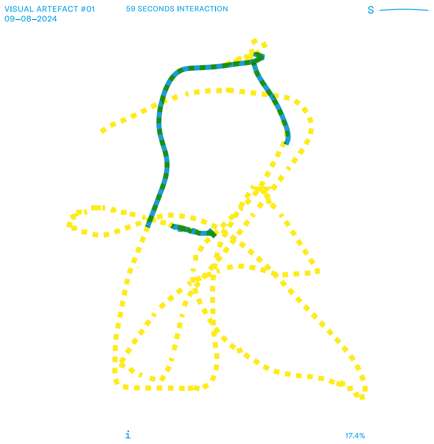
S—I
Process notes.



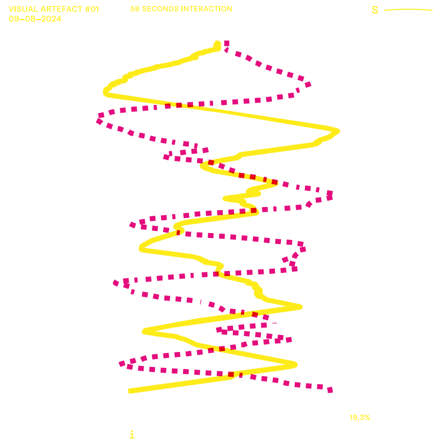
S—I
Visual artefacts development process.



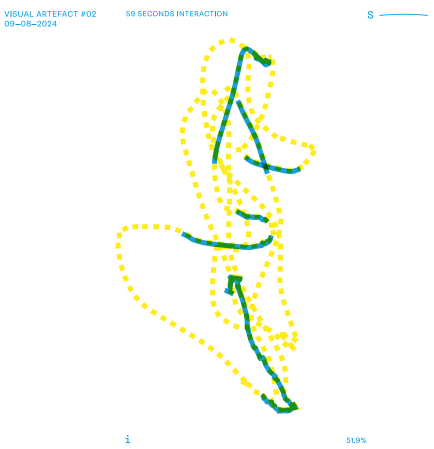
S—I
Visual artefacts development process.

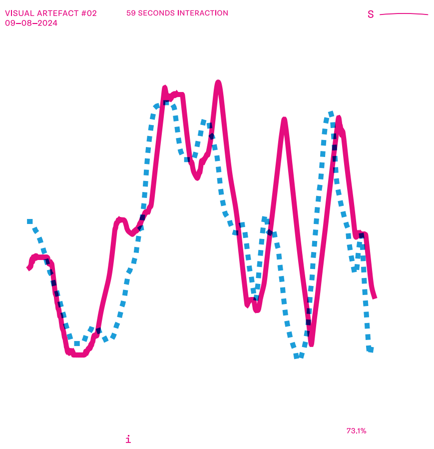

VISUAL ARTEFACT #01
09–08–2024
59 SECONDS INTERACTION
S
i
17.4%

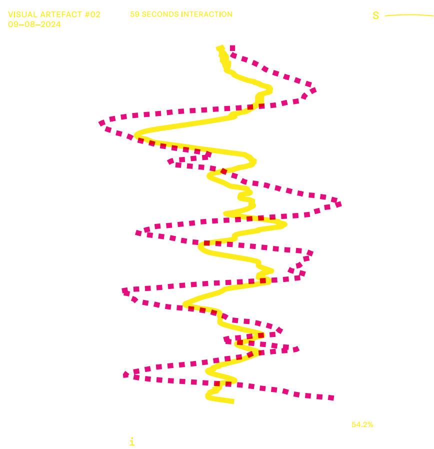

VISUAL ARTEFACT #01
09–08–2024
59 SECONDS INTERACTION
S
i
55.6%

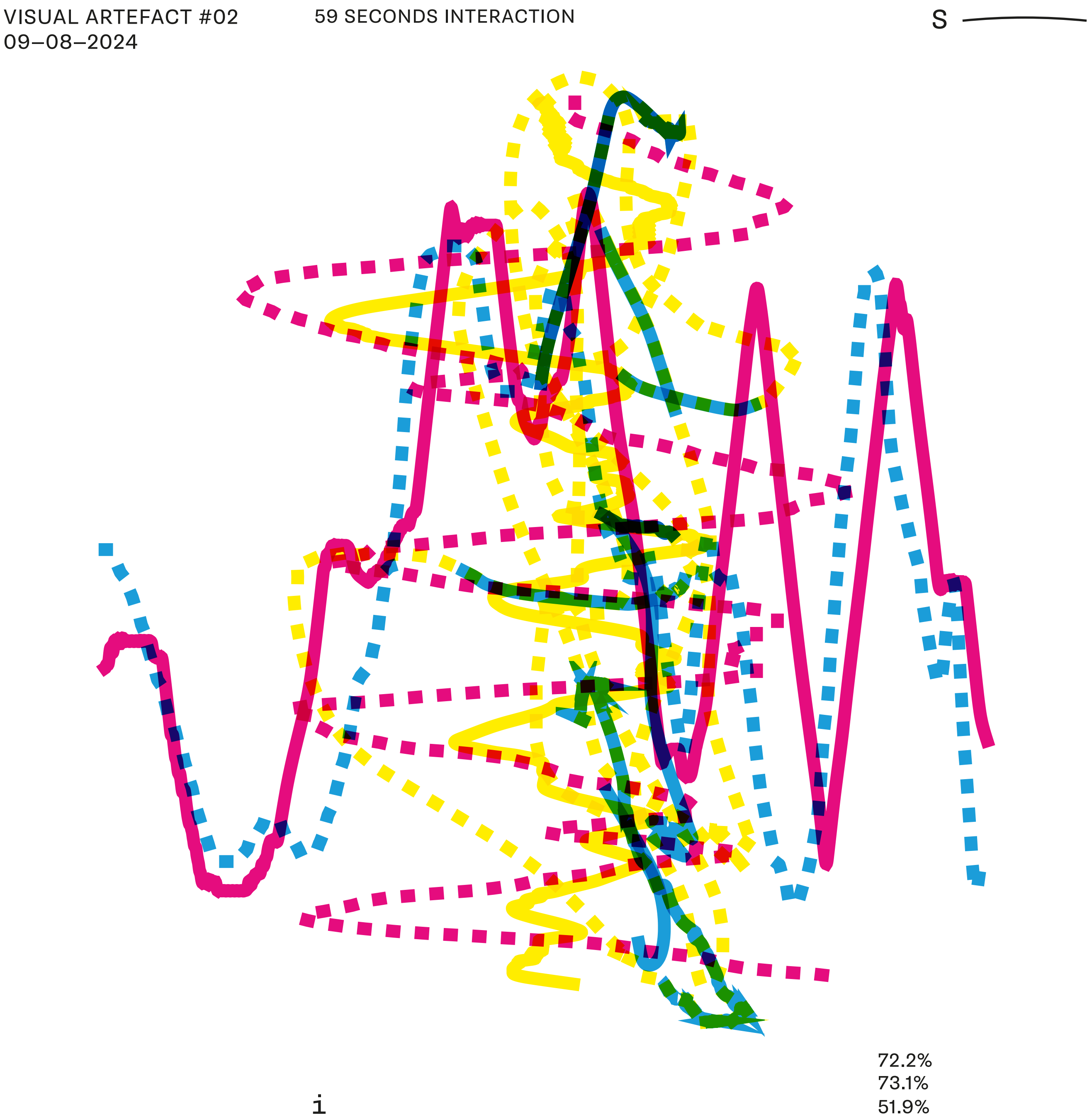

VISUAL ARTEFACT #01
09–08–2024
59 SECONDS INTERACTION
S
i
19.3%

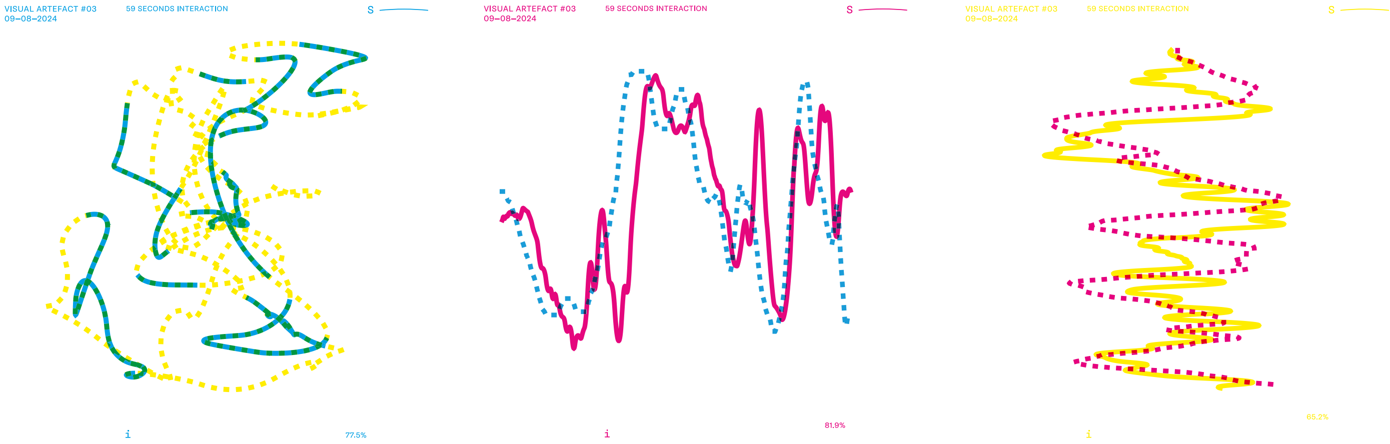

VISUAL ARTEFACT #01
09–08–2024
59 SECONDS INTERACTION
S
i
19.3%
55.6%
17.4%


VISUAL ARTEFACT #02
09–08–2024
59 SECONDS INTERACTION
S
i
51.9%


VISUAL ARTEFACT #02
09–08–2024
59 SECONDS INTERACTION
S
i
73.1%


VISUAL ARTEFACT #02
09–08–2024
59 SECONDS INTERACTION
S
i
54.2%


VISUAL ARTEFACT #02
09–08–2024
59 SECONDS INTERACTION
S
i
72.2%
73.1%
51.9%



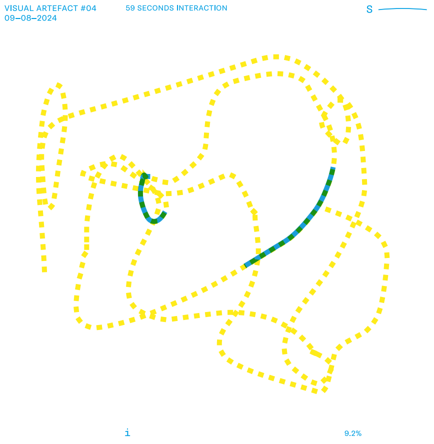
VISUAL ARTEFACT #03
09–08–2024

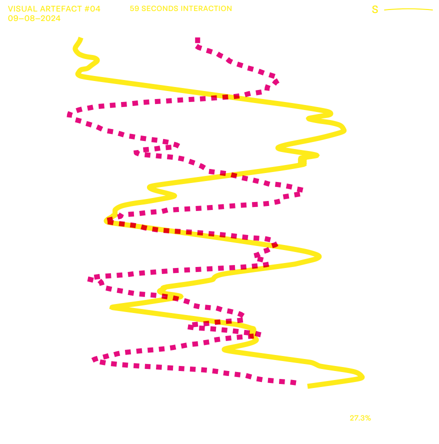
59 SECONDS INTERACTION

S

i

65.2%
81.9%
77.5%

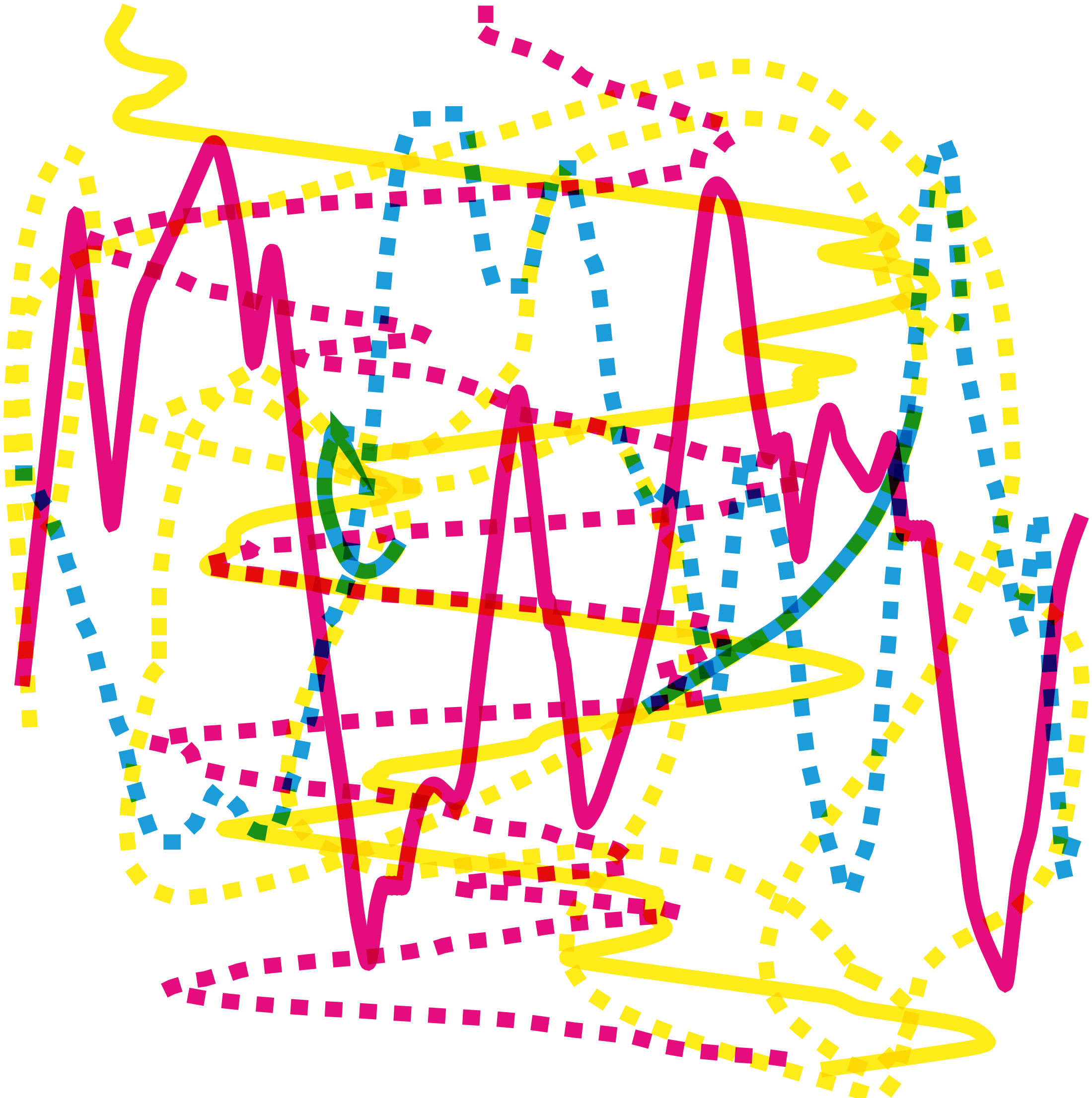




VISUAL ARTEFACT #04
09–08–2024

59 SECONDS INTERACTION

S



i

27.3%
29.3%
9.2%

MASTER THESIS BY
JOÃO CARVALHO NEVES

F

S—i

12

1

INTRODUCTION

«... art, science and the creative process. Experimentation and ultimately failure are an intrinsic commonality of all three.»
— PETTET-SMITH, 2019

S–I is the result of João Carvalho Neves' master thesis in Design & Multimedia at the University of Coimbra, Portugal. This work takes the form of an installation designed to explore human perception, sound spatiality, auditory memory, and the relationship between computation, sound, image, and aesthetics. It is strongly inspired by the physical characteristics and functionalities of British Acoustic Mirrors.

Acoustic Mirrors are monolithic concrete structures designed by Dr. William Sansome Tucker meant to detect approaching enemy aircraft through sound. They were built during the 1920s and were used during and after World War I but were only viable for a short time, as the appearance of the first radars made them obsolete. Over the course of the last decades, most of them have been demolished and vandalised, but a few, despite not working, can still be found across the shores of Great Britain. Acoustic Mirrors have different sizes and structural characteristics depending on which frequencies they intended to capture and how far the target might be. Nevertheless, they all share a curved surface, known as a paraboloid, meant to reflect and focus sound waves. Every sound wave that reaches this surface bounces off and is focused down on key points positioned in front of the structure. These focal points had specific microphones installed that amplified the incoming waves so operators could work out the direction of the sound source.

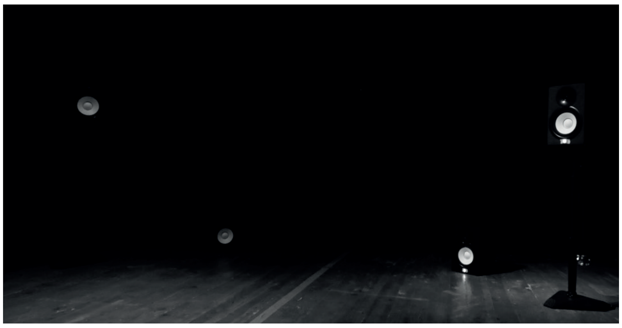
2

ABOUT

joaoneves.art/dissertation

S–I is an interactive installation that generates sound compositions and visual artefacts. The experience unfolds in two stages and is designed for individual participation, with one person engaging at a time. The first stage, the experimentation stage, allows the participant to explore and learn the mechanics of the installation. The second stage, the execution stage, challenges their auditory perception by requiring them to respond to the sound based on their memory and insights gained during the first stage.

By using sound synthesis as the sole sound contributor to this sonic experience, this installation aims to create an environment that encourages the experimentation and exploration of sound, enhancing and bringing awareness to our auditory memory, perception, and capabilities. In the end, we evaluate the experience by comparing interactions from different experiments.



3

ARTEFACTS

The visual artefacts are divided into three distinct outputs. In each one, the dashed line represents different aspects of a pre-recorded path, while the solid line displays data related to the participant's activity.

The first output aims to show how closely the participant followed the path using sound by highlighting the segments where they succeeded.

In the second output, the dashed line indicates the proximity of the pre-recorded path to the detection device over time, shown on a horizontal axis. The solid line reflects the participant's proximity in comparison.

The third output tracks the horizontal position over time, with the y-axis (from top to bottom) charting the participant's horizontal movements.

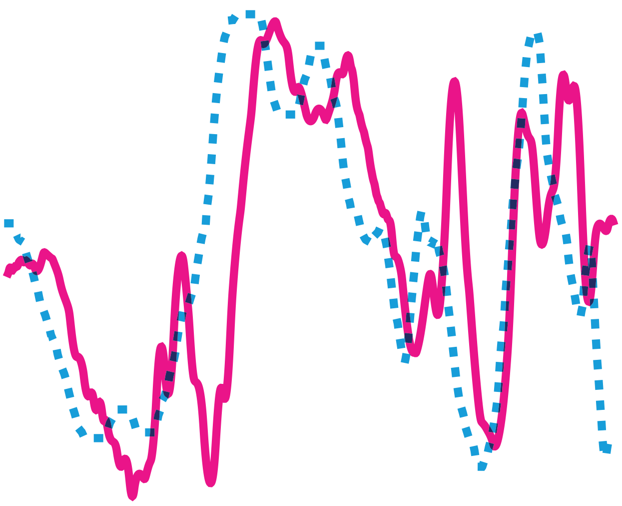
When these three visual outputs are combined, they create a unique and intelligible visual artefact that varies with each participant and interaction, offering a personalised reflection of their experience.

4

ARTEFACTS
PATH OUTPUT

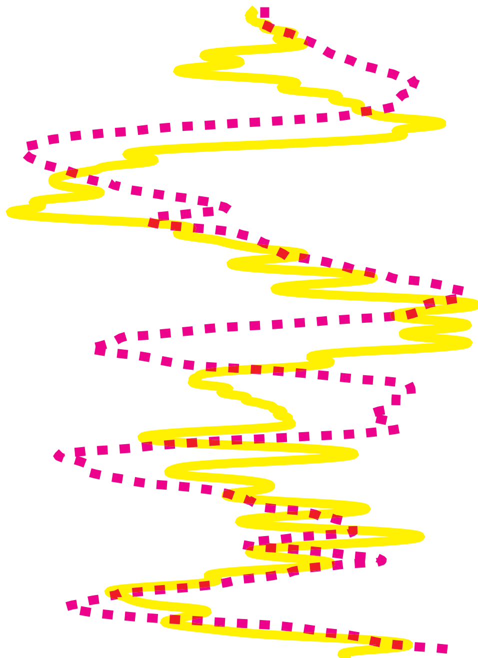
5

ARTEFACTS
DISTANCE OUTPUT



6

ARTEFACTS
HORIZONTAL OUTPUT



7

8

9

BIOGRAPHY

João Neves was born in Coimbra, Portugal, in 1991. He works as a computational and graphic designer and as a musician. He holds a Bachelor's degree in Design and Multimedia from the University of Coimbra (PT) and a Bachelor's degree in Music Performance, specializing in Jazz Piano, from the Lisbon School of Music (PT).

Throughout his career, he has composed several soundtracks, including those for *A Aldeia Adormece* and the TV series *Sempre*. He is currently working as a freelancer designer and performing with the bands *Duques do Precariado* and *Fred Menos Mais Amigos*.

COLOPHON

10

11

João Neves
www.joaoneves.art

+351 916 378 062
joao91neves@gmail.com

TECHNICAL RIDER
S—I

Speakers	4
Stands	4
Audio Interface (4 channels)	1
6-metre XLR-jack cable	4
Extension cords (4 outlets)	2
C13 power cable	5
*Computer (8gb ram) and peripherals	1
*Kinect V2	1

Notes
* We can provide these items if they are not available.
The installation must take place in a dark room that is at least 4x4 meters in size.

July 2024