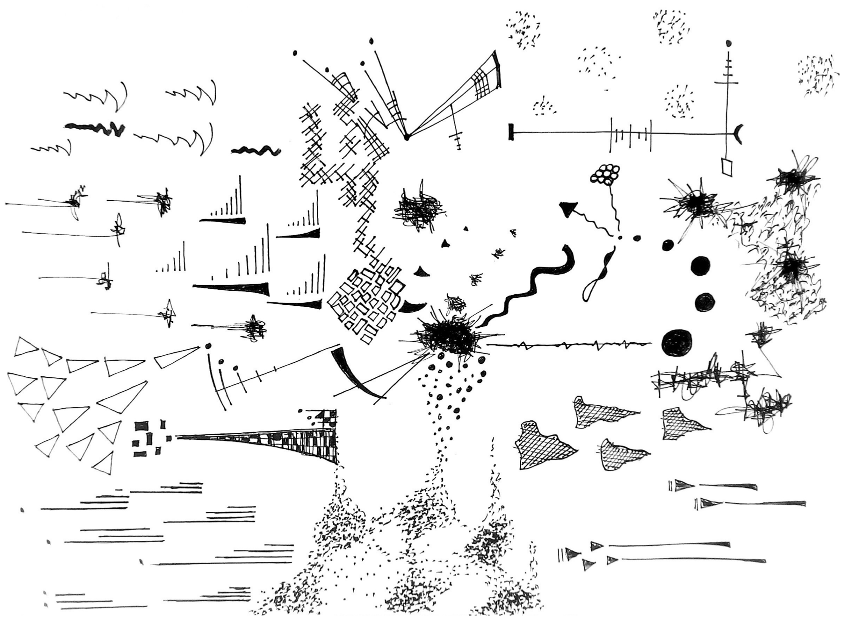
Royal Conservatoire in The Hague

Master Thesis

Sailing Through the Score-Map: Scanning Graphic Notation for Composing and Performing Electronic Music



Riccardo Marogna

Institute of Sonology

May 25, 2018

"We're cooking quinoa" David Lynch

Royal Conservatoire in The Hague

Abstract

Institute of Sonology

Master of Music in Sonology

Sailing Through the Score-Map: Scanning Graphic Notation for Composing and Performing Electronic Music

by Riccardo Marogna

Graphic notation has always been fascinating to me. As an improvising musician and composer, I tend to think about music and sound in a visual way. In my experience within the free improvisation context, I have found that graphic notation can be a useful tool for guiding the musicians, defining macro-structures but leaving them with enough freedom to let them express their ideas and personalities. Starting from these explorations, I came to the idea to develop a similar system for electronic music composition. We can identify two threads in this thesis. The first focuses on graphic notation and the definition of a graphic vocabulary for representing musical material, and how certain aesthetic choices are conveyed by this kind of representation. This topic is covered in chapters 2 and 3, as well as in chapter 5. In particular, chapter 2 presents some inspiring sources for the development of the graphic notation and some previous works of mine on the topic. Chapter 3 describes the graphic notation developed for the system, as well as the musical ideas that I want to express through this graphic vocabulary. The second thread is about optical sound and the concept of *raw scanning* as opposed to *symbolic* representation. Chapter 4 presents a historical survey on optical sound and, more generally, graphic-based systems for synthesizing sound, both in the analogue and in the digital domain. Chapter 5 goes into the detailed description of the main outcome of this research, an instrument/interface for live electronics called CABOTO, which implements both kind of ideas int the form of a *symbolic classifier* and a *raw scanner*. Chapter 6 presents some final remarks and future developments. The first chapter serves as an introduction.

Acknowledgements

I think that any kind of human achievement, especially in art and science, cannot be considered as the result of the effort of a single human being. It is the outcome of a community, a *scene*. Brian Eno calls it the "scenius", as opposed to the concept of "genius". This thesis is no exception. Studying, working, living in the Sonology community has shaped my research as well as - hopefully - my work has influenced my colleagues and friends. Thus, I want to thank all the teachers and colleagues for all the discussions we had together and all the advice, critiques, suggestions that I have received, but also for all the great creative work they are doing, because it inspired and influenced me a lot.

First of all, I would like to thank my mentor, Prof. Richard Barrett, for encouraging me from the very beginning, for his constant and kind support, and for his patience in reading the first draft of this thesis. I would like to thank Prof. Gabriel Paiuk for his advice and suggestions regarding Chapter 2, Prof. Bjarni Gunnarsson for teaching me Supercollider and for our fruitful discussions on the sound synthesis modules of the instrument, Prof. Raviv Ganchrow for some illuminating hints regarding the score-map metaphor. Thanks to Prof. Kees Tazelaar for his support and for his frank and open way of expressing opinions and thoughts. Thanks also for having supported my funding request, thus giving me the opportunity to present my work at the NIME Conference. Thanks to Prof. Johan Van Kreij for his suggestions on the live performance with the CABOTO instrument, for sharing the stage with me in some great improv sessions and for giving me the opportunity to assist him during an improvisation workshop. Thanks to Prof. Joel Ryan for his inspiring lectures and discussions. Finally, thanks to Prof. Peter Pabon for a conversation we had which has been very precious to me, and he knows why.

I would like also to thank some colleagues and friends that helped me: Giammarco Gaudenzi for sharing his knowledge in graphic design and fine art, Rob Schuttz for proofreading some chapters and checking my "broken English", the musicians that performed with CABOTO: Lauge Dideriksen, Volkan Turgut, Annick Odom, Irene Ruiperez, Dalton Danks. Thanks to Julio Molina, friend, housemate, colleague, for cooking together a very funny "Mighty Sup".

I feel blessed for the opportunity I had to spend these two years at the Institute of Sonology. What I have found here is not only an excellent artistic and cultural environment which encourages freedom *and* responsibility, but also a stimulating community of creative human beings who share their gifts with kindness and respect, and this is the most important lesson that I have - hopefully - learned.

Contents

Chapter 1

Introduction

1.1 Motivations

In my experience within the free improvisation context, I have found that graphic notation can be a useful tool for guiding the improvising musicians, defining macro-structures but leaving them with enough freedom to let them express their ideas and personalities. This led me to the development of a graphic notation system for improvisation, called *Graphograms* (see 2.5). In that system, a graphic vocabulary which defines some *sonic states* is organized in a graph-like structure. The sonic states are connected with each other, and the musician can then move from one state to another, whether on cue or freely. With this system, I can keep overall control of the macro-structure, defining a framework in which musicians could “move”, but at the same time there is a certain degree of stochastic behavior due to the concurrent choices that are made available to the musicians. From these experiments in improvisation came the idea to explore a similar graphic-based approach for composing and performing electronic sounds. Both improvisation and electronic music composition share a similar issue: within these scenarios, traditional notation systems are maybe not the most useful tools for representing the sonic material and the musical gestures. In a deeper sense, as noted by Trevor Wishart (Wishart and Emmerson, 1996), traditional western notation is based on a time/pitch lattice logic, which strongly influences the way music is composed. Nowadays, a lot of music is produced without the aid of traditional notation. However, the pitch/time lattice logic, to use Wishart’s terminology, is largely employed in composing and performing electronic music (consider, for example, the grid-based logic of digital sequencers). The idea was then to design a graphic-based notation system which could be defined in a continuous two-dimensional domain, as opposed to the lattice.

1.1.1 The Disappearing Computer is still There on Stage

In 2005 a paper with an intriguing title appeared in the ACM Communication Journal. In "The Disappearing Computer" (Streitz and Nixon, 2005), Norbert Streitz and Paddy Nixon theorized that in the future computers would have been embedded in everyday objects and would have thus disappeared from our view, and we would never had to deal with keyboards and screens. The disappearing computer paradigm implies that the machine has to be adapted to human beings and their behavior, and not vice versa. Ten years later we realize that, although this prophecy has become true for certain applications (think for instance about voice recognition software), in electronic music it is still common to see laptops on stage in front of musicians. If we consider that a laptop is an object that was designed for typewriting documents, we should be quite disappointed in seeing people making music out of it. In a sense, it is like seeing DJs with typewriters instead of turntables.

1.1.2 Gestural power of Drawing

Another important motivation behind this research deals with the role of control, playability and gesture in electronic composition and performance. In dealing with electronic music composition, I have realized that I need some kind of gestural and/or visual control over the sound design, and this need could not be satisfied by using traditional MIDI controllers, GUIs or sequencers. This need probably derives from my background as an improvising musician. The quest for new interfaces and controllers in electronic music is a vast topic which goes way off the topic of this research. In my work, I have focused on a specific scenario where the drawing gesture on a canvas is captured by a camera, scanned by a computer vision system and some information is then collected and mapped for synthesizing sound. Nowadays there are many devices available for digitizing drawing gestures (e.g. Wacom tablets). Nevertheless, I have decided to stick with real drawing with pencils and brushes. This have two main objectives: 1) to keep the richness of the real drawing experience 2) to have an interface that is based on a intuitive and *enactive* action (Varela, Rosch, and Thompson, 1992). The drawing gesture has become a fundamental element in performing with the system, and one of the main original outcomes of this research. The performance develops through the constant tension between the score, the real-time semi-autonomous scanning and synthesis of sound material, and the performative gesture which can, at any moment, modify the score itself, even to the point of causing a dramatic and sudden change. The need for this gestural component reflects an aesthetic vision. In performing electronic music I do not want to be the sailor who adjusts the sails from time to time, and let the boat follow the wind. I want to have a dialectical relationship with the sonic material produced by the machine, to *trade* and *fight* - even in a physical way. Scanning hand-made drawings is also a way of *challenging* the machine (Goodiepal, 2009). We can look at the computer as a digital *black box*. If we input an analogue stream through some kind of sensor, we are in a sense injecting a *perturbation* into the digital system. I find this way of looking at the human-machine interaction quite intriguing. With a camera it is possible to capture the complexity and richness of hand drawing and this richness is translated by the machine into unexpected behaviors, glitches that give life to the sonic output.

1.2 Aims and Objectives

In the beginning of this research, the idea was to develop an offline tool for composing, a computer-aided composition environment. Therefore, the graphic sketch was intended as a way of organizing sonic events in time on a canvas. The first experiments were actually inspired by some historical examples, such as the realization score of *Studie II* by Karlheinz Stockhausen (Figure 1.1). In 2014 I developed the first prototype of the system, called *KarlHeinz Score Scanner* (*KarlHeinz Score Scanner (2014)*). The underlying paradigm was quite straightforward: the canvas was interpreted as a time/frequency Cartesian space, where rectangular shapes defined band-limited and time-limited signals, and a set of envelopes defined amplitude and panning (see Fig. 1.2). Later on in the research I started questioning this framework, in particular the assumption that the score had to be time-oriented or time-based. I started thinking about other ways of looking at the graphic score, in which time was not represented on the canvas. In a way, I felt that having to deal with the time axis became a limitation instead of a resource or a stimulus for developing new ideas. This led to the shift from the time-based score to the *cartographic map metaphor*. This shift naturally evolved towards a performative scenario, where time is *now* and where the graphic score becomes an interface for real-time synthesis of electronic sounds. In the map metaphor, the score is traversed by a set of semi-autonomous *navigators* or *explorers*. The way we traverse (or explore) the score becomes the musical output or rendering of the composition/performance.

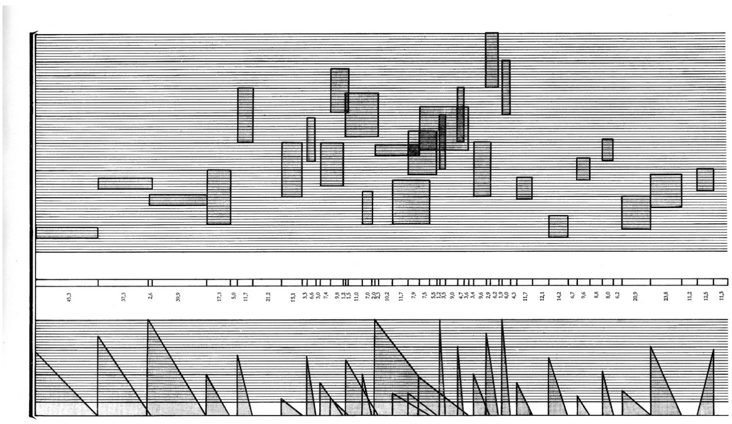


FIGURE 1.1: Karlheinz Stockhausen: an excerpt from the realization score for Studie II (1954).

1.3 Blurred Edges: the Score as an Instrument/Interface

The shift towards the performance scenario has represented a crucial one in my research, and opened up a whole set of new questions and threads. If we look at the system as a live electronics setup used for the real-time synthesis of sounds, the

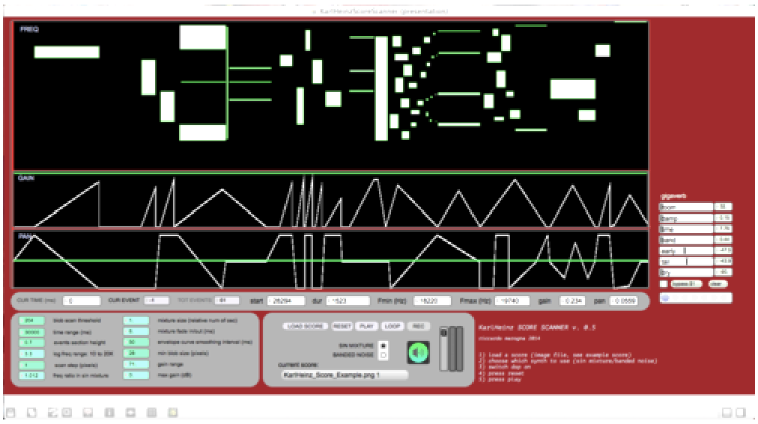


FIGURE 1.2: The first prototype: KarlHeinz Score Scanner.

graphic score becomes (part of) an interface or an instrument. In this sense, the system tries to challenge the boundaries between the concepts of composition, score, instrument, performance. All these terms can be put in relation with the others, in a sort of inter-connected graph. In particular, we can identify four relationships:

- Composing <-> Performing
- Drawing <-> Playing
- Composition <-> Score <-> Instrument

The first and second relationships are somewhat related to the juxtaposition of *allographic* vs *autographic* arts (Goodman, 1968). Allograph literally means "other grapheme" [1] and in phonetics it refers to the set of letters or letters combinations that represent the same phoneme, like the */f/* in *fish* and *phase*. Nelson Goodman uses the term to refer to the arts that need to be *put into action* in order to exist. Music and theater are typical examples. A composed score is just a collection of symbols on a piece of paper, until an orchestra starts playing it. Moreover, each time that composition is played, will result in a different *realization* of the same composition. In this sense, the written score, as in the case of the allograph, is an abstract symbolic entity that stands for its different realizations in the physical realm of sounds. On the other hand, autographic refers typically to fine art: a drawing by Picasso has been realized by himself in person, and the piece of art can be identified with a physical object that is always there, always expressing its content as far as someone is staring at it. You cannot have different realizations of a Picasso: you can only have reproductions of it. [2] Regarding the third relationship, the system can be defined as an *inherent score-based system* (Maestri and Antoniadis, 2015). According to Eric Maestri, when we create a notation system which is inherent to the developed instrument, it *becomes* part of the instrument itself. The composer blends with the instrument maker,

[1] In linguistics, a grapheme is "the smallest unit of a writing system of any given language" (Coulmas, 1999)

[2] We can note, however, that the introduction of recording techniques on one side and the advent of Andy Warhol's work on the other has challenged this dichotomy.

and the boundary between the score and the instrument disappears: the score *is* the instrument. In this sense, inherent scores "are [...] an expansion of what an instrument normally is: these instruments expand and reinforce their affordances, turning into objects acting in the sense of musical composition". An example are the tangible scores developed by Enrique Tomás and Martin Kaltenbrunner (Tomás and Kaltenbrunner, 2014).

In CABOTO, the act of composing and performing can be combined and merged in different ways. For instance, we can have a two-step scenario in which the score is pre-composed offline, and then played during the performance. Note that here the term *play* is used in the sense of instrumental playing, since the way we traverse the score will affect the sonic result. Therefore, the "scenario" or "setting" of the piece is defined by the offline composing activity, while the time development of the composition/performance is decided during the performance. Another possibility is available when we make use of a live camera for capturing the canvas in real time. In this case, the composing act can be performed directly on stage, as a sort of live painting. The score will be completed during the performance, becoming almost a sort of trace of what has already happened instead of the plan of something that is yet to come.

1.3.1 Morphophoric Features

According to Marie-Elisabeth Duchez (Duchez, 1989):

> The notion herein referred to as morphophoric - or form-bearing - element, has always and unfailingly guided musical action, that is to say strategies of production (inspiration, invention, representation, execution) and reception (listening, memorization). But this essential guidance is first of all only a more or less conscious, empirical practice based on immediate perception. Its efficiency, therefore, though direct and reliable, is limited, and it corresponds to what are generally called "primitive", orally-transmitted musics”.

We can see that CABOTO contains this morphophoric feature. First, we have the drawing gesture, which becomes a composing act or an improvisational act (in the sense of improvisation as real-time composition). This gesture results in a graphic sketch which has intrinsic morphophoric characteristics. This becomes an immediate and intuitive set of cues for the performer. The non-mediated perception of the different shapes, gradients, spaces, densities, which are represented on the canvas, becomes a powerful guide in developing the live performance. In this sense, the performer can look at the score as a palette of *sonic elements* available for the live performance.

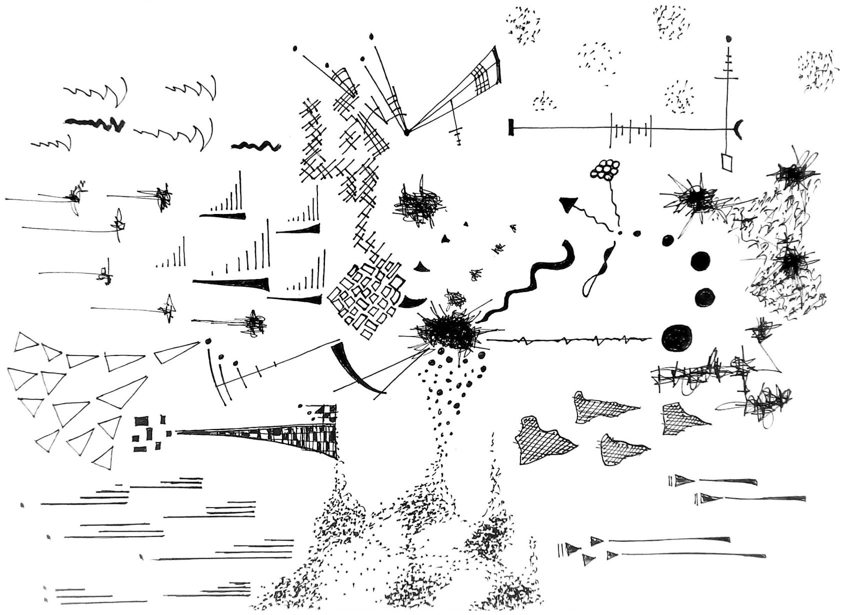


FIGURE 1.3: An example of a graphic score composed for CABOTO.

1.4 Background

The idea of synthesizing sound from graphics has a long history, which can be traced back to the early experiments by pioneers in Soviet Russia during the early decades of the twentieth century (Smirnov, 2013), the experiments by Oskar Fischinger in Germany (1932), Norman McLaren (1940), Daphne Oram (1959), and arrives at computer-based interfaces such as the UPIC system conceived by Xenakis in 1977 (Lohner, 1986). More recent studies have been dealing with augmented paper tools for traditional composers (Tsandilas, Letondal, and Mackay, 2009), projects inspired by the UPIC system (Thiebaut, Healey, and Bryan-Kinns, 2008), audiovisual interfaces (Levin, 2000) and more educational and game-oriented app (Farbood and Pasztor, 2004).

In early developments, graphic-to-sound synthesizers were based on analogue technologies, such as optical sound track. According to this technique, the graphic elements are directly translated into sound signals, without any *interpretation* or *mediated translation* (other than a photocell-based circuit). Later on, with the advent of the ANS synthesizer and the Oramics instrument, the sketched shapes began to be interpreted as voltage control signals, and this is quite important since it marks the passage from a *non-mediated sonification technique* to a more complex interpretation of the graphic sketch, which becomes then more similar to a score. This will then evolve further with the advent of computers into a symbolic interpretation of the graphic score. In developing CABOTO, I tried to combine these different approaches, creating a sort of *hybrid* instrument which makes use of a symbolic classifier along with analogue-like scanning techniques for extracting optical signals and waveforms.

Chapter 2

Graphic Notation: Inspiring Sources

As noted by Trevor Wishart (Wishart and Emmerson, 1996), traditional Western notation is based on what can be defined as a *time/pitch lattice*. According to this logic, the sonic material is organized in a two-dimensional Cartesian plane, where the two axes represent time and pitch, and these parameters are quantized according to certain scales. The use of this two-dimensional representation in music precedes the ideas of Descartes: in one of his writings Iannis Xenakis (Xenakis I., 1994) cites the

> [...] Invention of the bi-dimensional representation of pitches versus time by the use of staves and points (Guido d'Arezzo), three centuries before the coordinates by Oresme and seven centuries before (1635–1637) the superb analytic geometry from Fermat and Descartes.

This way of representing the musical content strongly influences the way music is composed. Graphic notation appeared around the middle of the twentieth century as a way of tackling this lattice logic and finding new ways of representing musical structures. Earle Brown, one of the first composers to make use of graphic notation, wrote:

> Time is the actual dimension in which music exists when performed and is by nature an infinitely divisible continuum. No metric system or notation based on metrics is able to indicate all of the possible points in the continuum, yet sound may begin or end anywhere along this dimension. (*Program note to Music for Cello and Piano, 1955*)

Each graphic notation system is unique, since it is defined by the composer, and sometimes it is specifically conceived for a single composition. It conveys the composer's aesthetic and thus it becomes part of the composition itself [1]. In the following sections, some inspiring sources in defining a graphic notation system will be presented. First, a note about early music examples. Then, a look at some examples

[1] We can express this in relation to the score/instrument/interface topic: the graphic vocabulary, the system of signs and rules used for composing the score is also part of the equation, and can be considered as an *instrument* for generating the score

from contemporary composers of the twentieth century. In Sections 2.3.1 and 2.3.2 the focus will move towards graphic systems in improvised music. Section 2.5 an original notation system for improvisation called *Graphograms*, is presented. The vocabulary introduced in that system forms a basis for the graphic notation developed for CABOTO. Finally, a previous work for bass clarinet and tape, *Ossi di Seppia*, is presented, which also made use of graphic notation.

2.1 Early Examples

The use of curvilinear graphic shapes for representing musical gestures can be found in different cultures and geographical areas. The first documented use of neumatic notation is in the Paleobyzantine chant, in the ninth century. This kind of writing was a purely mnemonic aid in the context of a well-established oral tradition: curvilinear shapes above the text represented the pitch contour. A similar technique was developed in the Japanese *shomyo* tradition (Figure 2.2) as well as in the Buddhist Tibetan chant (Figure 2.1). It is interesting to note how in the case of Tibetan neumatic notation, the graphic shapes could represent pitch contours as well as other parameters, such as durations (Kaufmann, 1967).

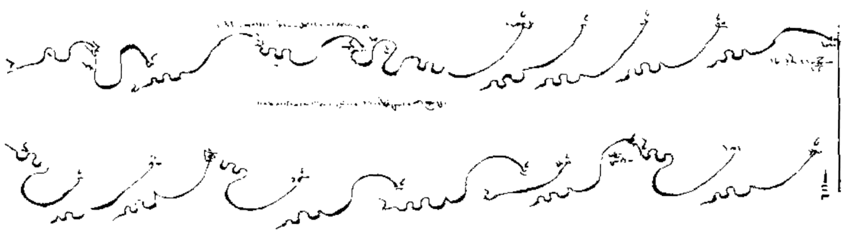


FIGURE 2.1: Tibetan neumes

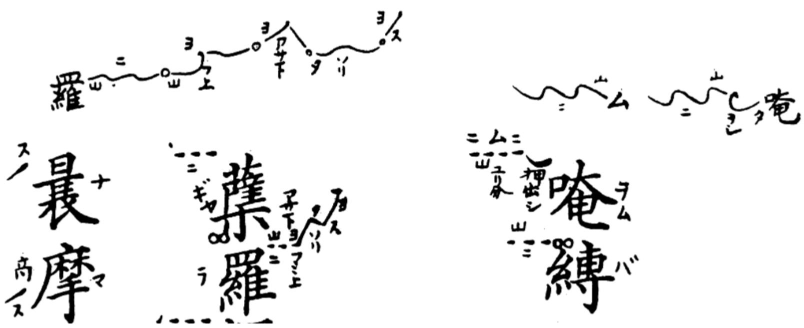


FIGURE 2.2: An Excerpt from Japanese shomyo notation (from (Nelson, 2008))

2.2 Graphic Notation in the Postwar Era

2.2.1 Morton Feldman, Earle Brown

The first examples of graphic notation appeared in the circle of American composers who orbited around John Cage, in particular, Morton Feldman and Earle Brown. Feldman's *Projection* series includes what is considered to be the first example of graphically scored work of the postwar era (Boytwell, 2012). In *Projection 1* for solo cello (1950-51, Figure 2.3), rectangular shapes organized in a grid-based layout provide the performer with instructions about the style, pitch range and durations of the notes. The graphic score is based on a vocabulary and the composer provides instructions on how to interpret it. A striking different approach can be found in Earle Brown's *December 1952* (1954), a piece composed around the same years. Looking at the score (Fig. 2.4), we can point out two main differences with respect to Feldman's approach: a) the grid-based layout has disappeared, leaving the graphic elements floating in an empty space; b) the composer does not provide any set of instructions nor a legend for interpreting the score.

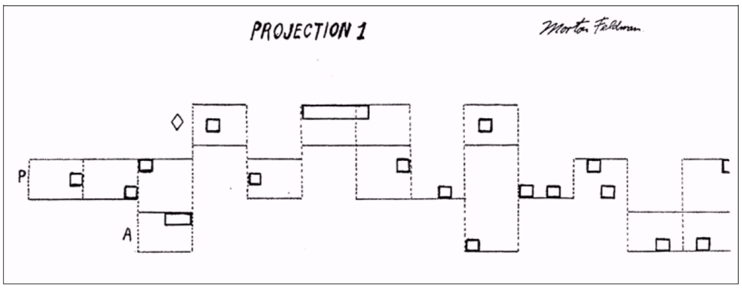


FIGURE 2.3: Morton Feldman, Projection 1 for solo cello (1951)

Brown was aware that graphic notation can be a tool for exploring the *continuum* in the sound parametric space, as well as a way for conveying a from of knowledge that goes beyond language. In the prefatory note to *Folio and 4 Systems* (1954) he writes:

> Similarly, all of the other characteristics of a sound — frequency, intensity, timbre, modes of attack-continuation-decay — are infinitely divisible continua and unmeasurable. The imposition of approximate scalar-systems is obviously possible and efficacious, but to deal directly with the experience of a continuum on its own unknown terms seems to imply that the unmeasuring eye and ear are their own terms and experiential justification and compatible with unmeasured experience. An ambiguous but implicitly inclusive graphic "notation" and alert, sympathetic performers, are conceivable catalysts for activating this "process" within continua.

The approaches that Brown decided to explore were two: the first led to the definition of "mobile scores", that is, scores defined by a set of modules of notated

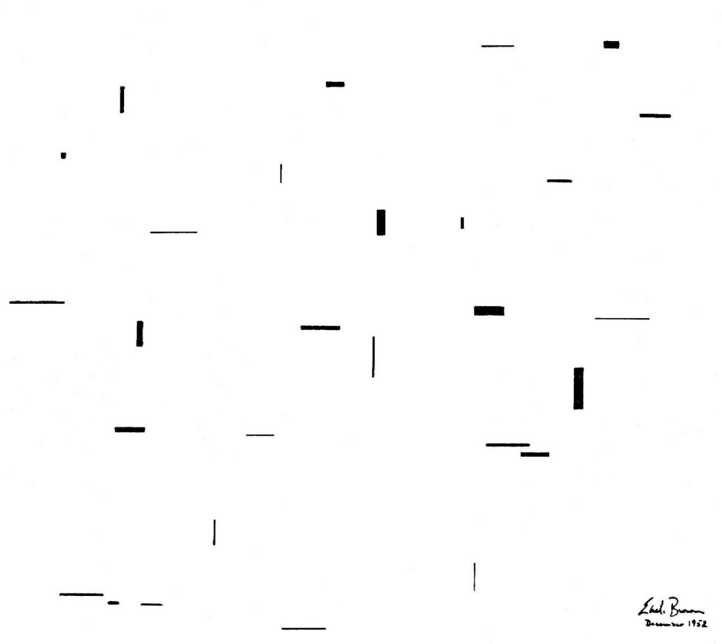


FIGURE 2.4: Earl Brown, *December 1952* (1954)

material that could be organized by the performer. The second approach led to the introduction of graphic elements:

> [...] a conceptually "mobile" approach to basically fixed graphic elements; subject to an infinite number of performance realizations through the involvement of the performer's immediate responses to the intentionally ambiguous graphic stimuli relative to the conditions of performance involvement.

2.2.2 Cornelius Cardew

This approach was developed to its extreme consequences by Cornelius Cardew (1936-1981). His *Treatise* (1963-67) has become iconic, a graphic score of huge proportions (193 pages), without any instruction for interpreting it. The score becomes a visual artifact which fascinates the reader with its aesthetic beauty (Fig. 2.5). In later years, Cardew strongly criticized his work (Cardew, 1974). This criticism has to be framed into the wider discourse of his political involvement into the Marxist-Leninist movement in the early 1970s, which led him to abandon avant-garde music altogether. In 174 he published a collection of essays titled "Stockhausen Serves Imperialism and Other Articles", in which he not only criticized the works of Cage and Stockhausen, but also his own. Although the strong political bias which informs this essay has to be taken into account in reading the Maoist-like self-criticism he applies to his works, it is worth discussing here the two main criticisms he raises about the

Treatise. The first and most important has to do with the arbitrary link which pure [2] graphic notation establishes between the visual and the sonic realm:

> Let's start with the idea very widespread in the avant garde and implicit in the score of Treatise that anything can be transformed into anything else. Now everybody knows (not only Marxists and farmers) that a stone, no matter how much heat you apply to it, will never hatch into a chicken. [...] And yet in Cage's work Atlas Eclipticalis patterns of stars in a star atlas are transformed into a jumble of electronic squeals and groans. This transformation is carried out through a system of notation (a logic) that has no connection with astronomy and only a very sketchy connection with music. (Cardew, 1974)

The main issue with which Cardew is concerned here is that graphic notation has the character of arbitrariness. He points out Cage as the first promoter of this tendency, and that arbitrariness is *ethically wrong* since it is a *lie*, it does not reflect the reality of the world (e.g. scientific knowledge) nor the reality of the human condition (which is defined by social and political struggle). The other criticism points out the relationship between the interpreter, the score and the audience:

> In performance, the score of Treatise is in fact an obstacle between the musicians and the audience. Behind that obstacle the musicians improvise, but instead of improvising on the basis of objective reality and communicating something of this to the audience, they preoccupy themselves with that contradictory artefact: the score of Treatise. So not only is Treatise an embodiment of (not only irrelevant but also) incorrect ideas, it also effectively prevents the establishment of communication between the musicians and the audience. (*ibid.*)

2.2.3 Breaking the linear score: Open Forms

Another aspect that emerged along with the development of graphic notation, as already mentioned in Section 2.2.1, deals with the introduction of *open form*. Both Feldman and Brown began exploring open form around the same period: as a matter of fact, the first modular piece by Earle Brown, *Twenty-Five Pages* and Morton Feldman's *Intermission 6* were both realized in 1953. In his piece, Brown presents to the interpreter a set of twenty-five pages that can be performed in any order and either side up. Moreover, the duration of each fragment may vary according to the performer's will. In *Intermission 6*, Feldman organizes a set of fragments on a single page, along with the zen-flavoured, fascinating instruction: *"Composition begins with any sound and proceeds to any other"*. Three years later, Karlheinz Stockhausen

[2] The term pure graphic notation is used here in order to refer to a notation system that does not provide any cue on how to interpret the score.

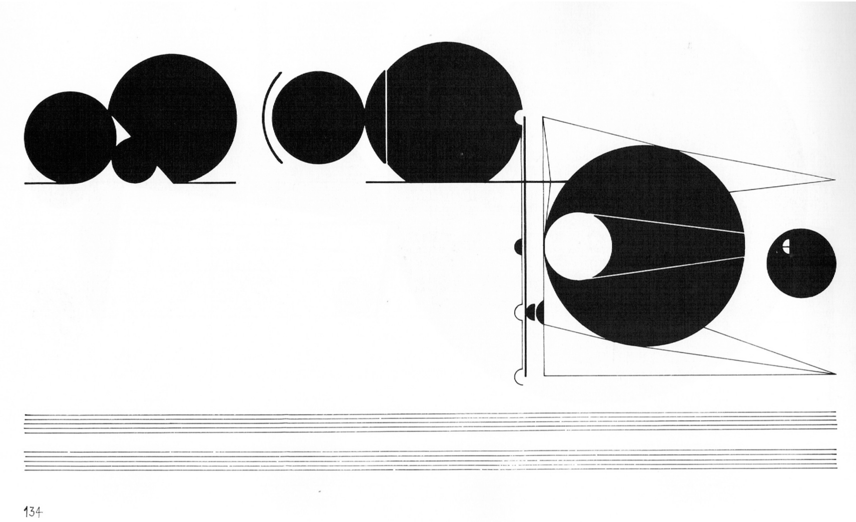


FIGURE 2.5: Cornelius Cardew, Treatise, p.134

explored a similar modular structure for composing one of his piano pieces, *Klavierstük XI* (1956). Here, a set of 19 fragments are presented on a single page, and the instructions tell the performer to make their own path through them. The piece ends when each fragment has been played three times. Although this apparently sounds very similar to Feldman's approach, Stockhausen introduced an interesting detail: the markings from a previous fragment (tempo, dynamics, articulations) are to be applied to the next one. In this way, Stockhausen achieves two interesting results: a) the combinatory power of the piece grows to a new level of complexity, since each fragment is not defined only by itself, but also by the one preceding it and b) the composition is no longer a mere stochastic juxtaposition of modules, a collage, but a graph of interrelated fragments: the whole is more than the sum of its parts.

2.2.4 Anestis Logothetis

Another interesting example is represented by the work of the Greek composer Anestis Logothetis (1921-1994). He started developing his original graphic notation around 1958, and since then he composed a corpus of more than one hundred graphic scores. The main feature of his notation, which is an inspiring source for my research, is the concept of *polymorphism* (Magnus, 2015), which for Logothetis has to do with the spatial dimension of the score and the way it is interpreted (Logothetis, 1973, as translated in *Anestis Logothetis Music Portal*):

> What fundamentally differentiates graphic notation from traditional notation is the aforementioned polymorphism, which clearly enables all performers to retain their subjective reaction times. The composer takes into consideration the divergences of the different performers in composing and expects a certain degree of surprise through the new formalization of musical form in every performance.

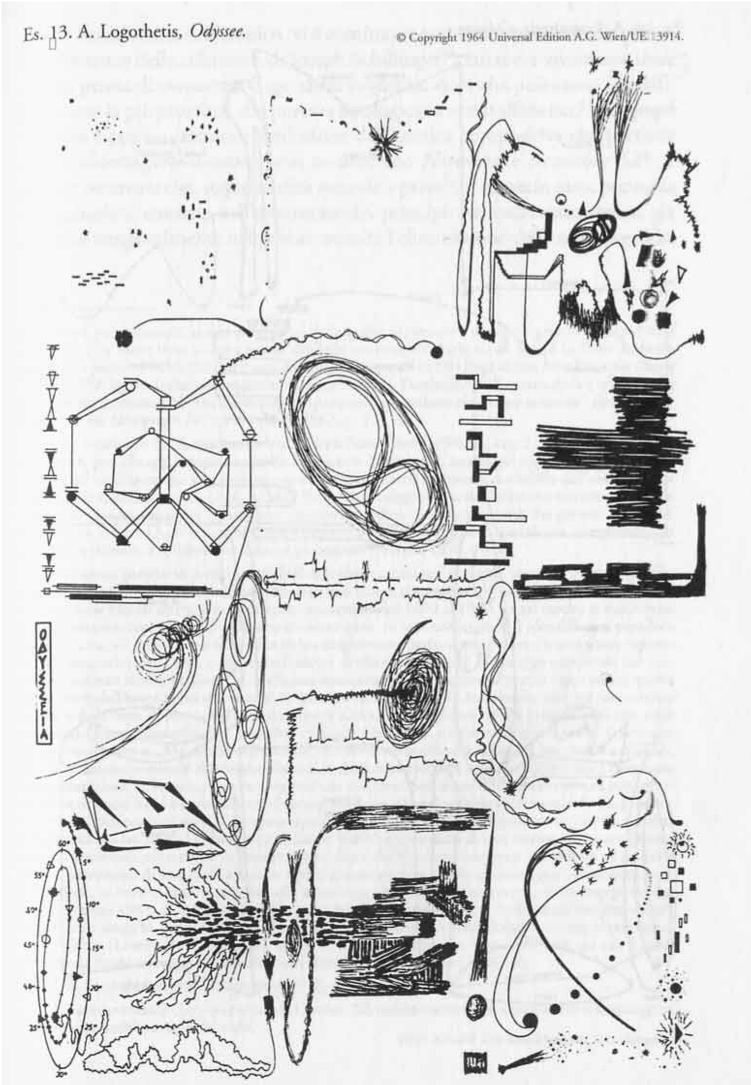


FIGURE 2.6: Anestis Logothetis: *Odyssee* (1964)

This polymorphism is achieved through two strategies: 1) the composer does not provide any instruction on how to interpret the graphic elements, thus the interpretation is left to the performer; 2) the graphic elements are distributed in the score in a non-linear way, thus the score does not provide a preferential orientation for reading it. As stated by Logothetis (Logothetis, 1973):

> Traditional notation is divided into systems and is read from left to right, like books. But since sound does not behave in the way written word does, we could think about using pictorial notation to represent musical events. (...)because musical time doesn't follow any direction, let alone the conventional left to right writing found in literary forms.

If we look for example at the score of *Odyssee* (1964) - and the title itself reveals the underlying concept - we can note how the distribution of the various elements on the canvas gives no clear indication about which route to follow in this voyage:

the interpreter has to make her/his own path through the score, as if traversing a cartographic map (Figure 2.6).

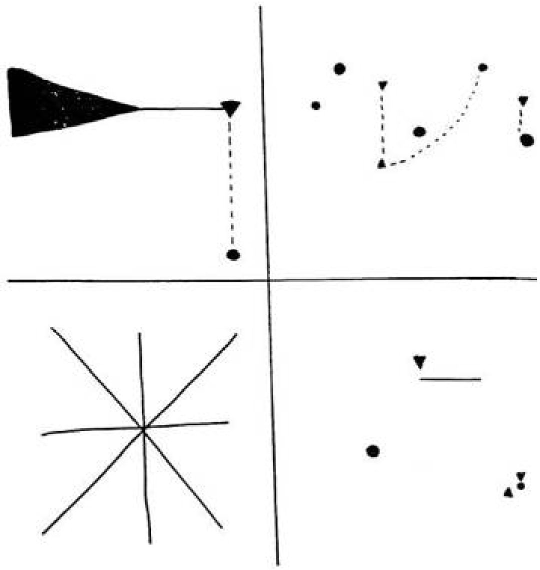


FIGURE 2.7: An excerpt from Composition 10 for solo piano (1969)

2.3 Graphic Notation in the Context of Improvised Music

Finding a way of organizing or guiding an improvisation process is an issue that has been explored by jazz musicians and composers since the Sixties. These explorations reflected some aesthetic choices but also some historical, sociopolitical ones (Lewis, 2008). In the flourishing scene of Afro-American experimental music centered around the Chicago-based Association for the Advancement of Creative Musicians (AACM), composers were taking inspiration from European composers such as Schönberg, Stockhausen, Boulez in order to get away from a certain idea of jazz as pure entertainment, and to present it as an artistic practice that deserved more attention. This meant, for instance, presenting the music in a proper venue (a theater instead of a restaurant), to get rid of jazz idioms and cliches (or to manipulate and reinvent them), to create a discourse and the ability to conceptualize an aesthetic, and to explore new ways of organizing the musical material through original notating systems. One of the prominent figures in this revolution is Anthony Braxton. His complex, articulated and multi-layered work spans many decades and a proper analysis of it is out of the scope of this research. The next section will focus on some original notation systems developed by Braxton during the early years of his career, when he started his solo saxophone explorations in 1968. In the subsequent sections, other interesting examples of improvising musicians are presented. In the last section the *Graphograms* system is presented. This is an original graphic vocabulary for improvisation developed since 2012 as a tool for guiding an ensemble of improvising musicians. It is worth presenting here this system, since that graphic vocabulary form the basis for the evolution of the graphic notation developed in this research.

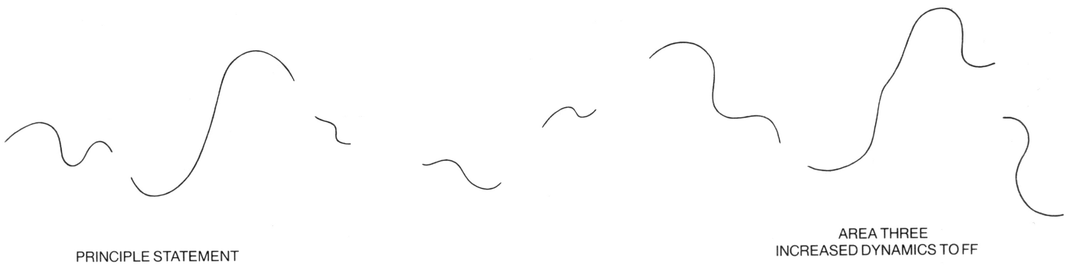


FIGURE 2.8: An example of Braxton's graphic score for *Composition 77F*.

1. SUBTONE PHRASE LINE FORMINGS
2. MELODIC LIKE PHRASE REGION
3. LONG FORMINGS (length of a total forming)
4. REGISTER FOCUS (low register is principle)
5. DARKER SOUND TIMBRE (on one sound)
6. LONG SOUND LINE FORMINGS (on one sound)
7. MEDIUM SOUND TIMBRE
8. A SLOW AND SOLEMN SOUND STATE

FIGURE 2.9: Anthony Braxton's text score for *Composition 77F*.

2.3.1 Anthony Braxton

Braxton started using graphic notation in 1969 at least, when he finished his *Composition no. 10* for solo piano (1969) (Fig.2.7). At that time, he had already written some pieces under the influence of some European composers like Schönberg and Stockhausen. But while these were composed using traditional notation, *Composition no.10* contains for the first time what he defined as "symbolic notation". Each graphic symbol here refers to a certain playing technique, such as cluster, staccato, trills. As noted by Radano (Radano, 2009), the influence of John Cage, Morton Feldman and Earle Brown is quite clear, and this is confirmed by the fact that the instructions for performing the piece stated that the pages could be arranged in any order and the durations were also left to the interpreter. Although these first attempts were strongly inspired by the so-called "concert music" or "modern music" of the time, we can argue that the introduction of graphic notation was for Braxton much more than an aesthetic fascination: it was also a natural bridge between the practice of composing and that of improvising. Giving the performer (a significant) part of the responsibility of the musical material organization and/or its interpretation, was a natural shift for a composer who was *also* or maybe we should say *first of all* a musician and improviser. Braxton himself contextualize the researches of that period in the dynamic activity that was taking place in the AACM:

> [...] It was in this time period that the organization began to profoundly investigate the reality of creative music - especially as this subject pertained to the post-Coleman/Sun Ra extension that was radically changing the music. In this period the musicians of the AACM would move to investigate every area of creative music - from ragtime music to African ritual music. *(Anthony Braxton, Composition Notes A (Frog Peak, 1988: 47-50))*

Around the same period, Braxton started performing solo, and he soon realized that he needed some kind of organization of the sonic material in order to structure the improvised set beforehand:

> I didn't want to give a concert and just play saxophone for five hundred hours; I wanted to find a way of presenting materials and separating things so I wouldn't repeat myself during the whole evening. So I began to create musical ideas beforehand so that I would have different problems to deal with in each composition. *(Interview by Ronald Radano, February 17 1983, reported in (Radano, 2009))*

Braxton developed two kinds of notation for organizing this material: we could define the first one as "text scores" (Fig. 2.9), the latter as graphic notation. In the text notation we can already note the use of a vocabulary that is related to a shape-like or spatial-like way of thinking about musical ideas. Take for instance the term "formings" which Braxton used extensively in defining different musical entities. Other terms like "lines", "shape", "stream", "flow" are used as well. Some pieces contains graphic indications (see Figure 2.8). He eventually came to the definition of a vocabulary of "formings", reported in his *Composition Notes* (Braxton, 1988). In this taxonomy of sounds, each item was denoted by a descriptive name and a graphical counterpart (Figure 2.10).

2.3.2 Barry Guy and the London Jazz Composers Orchestra

Barry Guy has been exploring graphic notation since 1992, but he recalls (Guy, 2012) that already in 1968 he presented a piece entitled *Dada Requiem* which contained some sort of graphic notation. If we take a close look at the scores he composed for the Jazz Composers Orchestra, we note a sort of hybrid approach: we have the

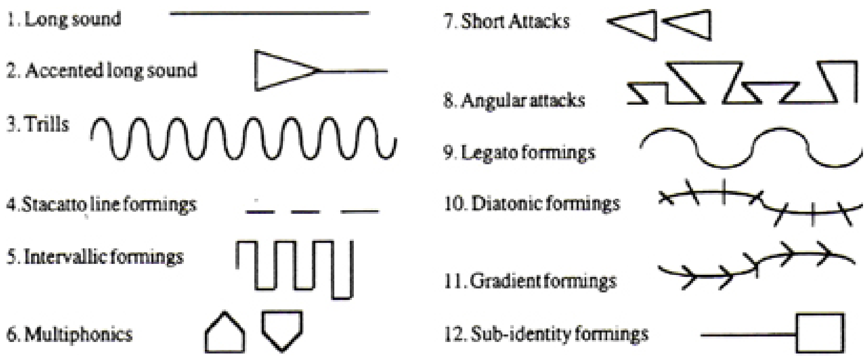


FIGURE 2.10: Excerpt from Anthony Braxton's graphic vocabulary (Braxton, 1988).

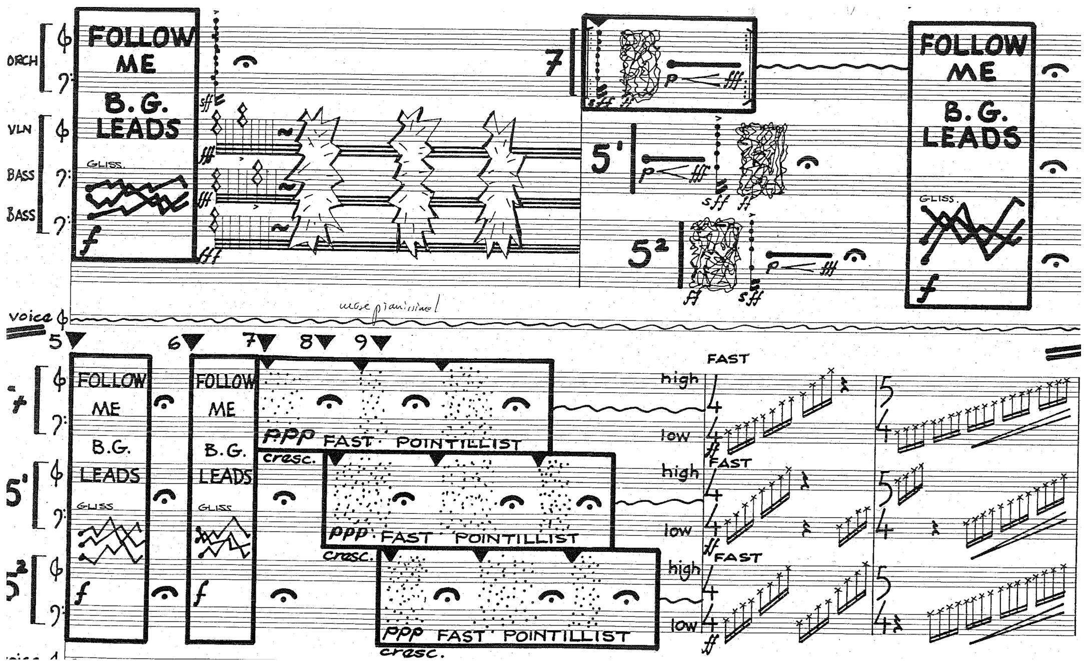


FIGURE 2.11: Excerpt from *Strange Loops* by Barry Guy

co-existence of traditional notation and graphic elements which sometimes appear in a frame-like environment which disrupts the continuum of the traditional staves (Figure 2.11). As he stated:

> [...] the scores written for the London Jazz Composers Orchestra (LJCO) and latter for the Barry Guy New Orchestra (BGNO) [...] were characterized by my utilizing the creative improvisational voice within defined (and sometimes not so defined) structures that allowed freedom as well as finite through-composed sonorities. The important element was that improvising musicians formed the ensemble and gave it its specific sound. Simple graphic representations of sound areas informed the players of sonic expectations, but I also used more conventional symbols which continuously changed over the years in an attempt to refine not

only the clarity of intention but also the ease with which they could be interpreted (Guy, 2012).

The kind of graphic vocabulary used by Barry Guy has an immediate, *intuitive* character, or at least the composer probably thinks so, since no vocabulary definition nor list of symbols were provided along with the score. Nevertheless, for an improvising musician it is quite clear what the composer suggests using the provided graphic hints. Moreover, sometimes written indications accompanied the graphics, such as the note "pointillistic" and "gliss" in 2.11.

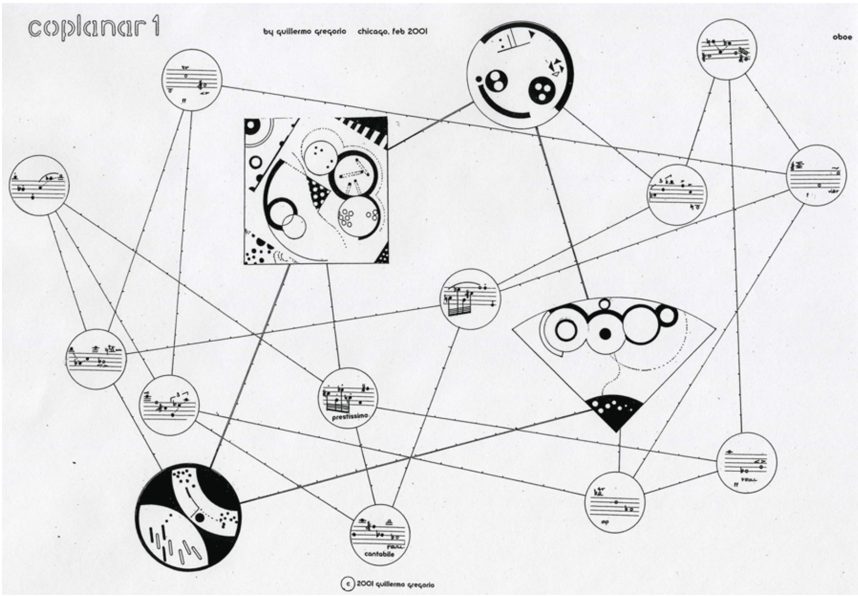


FIGURE 2.12: Guillermo Gregorio: Coplanar 1 (2001)

2.3.3 Guillermo Gregorio

It is worth citing here also the work of Argentine composer and improviser Guillermo Gregorio (b. 1941), since in his graphic scores we can find another interesting way of organizing the material in the score-space. In works like *Moholy 2* and *Coplanar 1* (2.12), graphic elements are organized in fragments enclosed in circular frames. In particular, in Coplanar 1 these fragments are connected together in a graph structure.

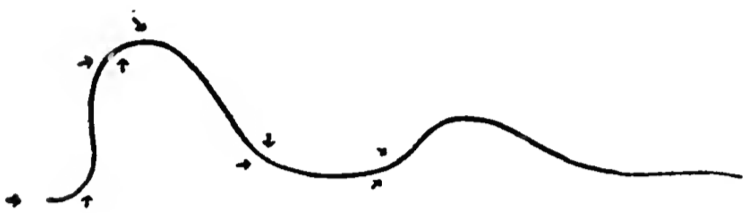


FIGURE 2.13: Vassily Kandinsky: a curved line shaped by forces (from *Point and Line to Plane*, 1947)

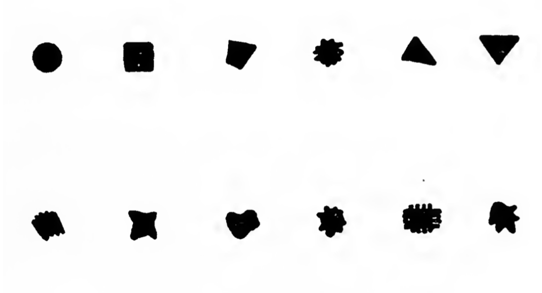


FIGURE 2.14: Vassily Kandinsky: point forms (from *Point and Line to Plane*, 1947)

2.4 Visual Arts: Kandinsky

Other influences in developing the graphic notation are - not surprisingly - derived from visual arts and in particular from abstract painting. These influences are sometimes difficult to point out, since they are sometimes a sort of subconscious inspiration from various sources. Nevertheless, one strong influence has to be pointed out: the work of Vassily Kandinsky. Apart from his paintings, an inspiring source is represented by his book *Point and Line to Plane* (Kandinsky, 1947), in which he elaborates a theory of abstract geometric forms and their composition, as well as a sort of synaesthetic theory of shapes and colors. In the first part of his essay, a *taxonomy* of basic shapes (point, straight line, curved lines, plane...) is given, and the composition of these elements is discussed. An interesting part of this work deals with the concept of forms shaped by dynamic relations: according to Kandinsky, a curved line is, for instance, the result of forces acting on the line from different directions (Figure 2.13). This relates with the concepts I have introduced earlier: the graphic composition is the result of gestures and thus of physical relationships (force distributions, mass densities...). Another interesting topic in Kandinsky's discourse on shapes is the intuition that a shape appearance depends on the point of view. A point seen from a distance has no measurable size nor shape, but, if you look closer, its shape becomes perceivable 2.14:

> [...] Abstractly or imaginatively, the point is thought of as ideally small, ideally round. In actuality, it is an ideally small circle. Nevertheless, just as in the case of its size, its limits are equally relative. In its material form, the point can assume an unlimited number of shapes: it can become jagged, it can move in the direction of other geometric forms, and finally develop into entirely free shapes. [...]

This idea will be discussed in detail in section 3.3 in the next Chapter, where I will discuss the cartographic map metaphor and the change of perspective introduced by zooming in/out of the score.

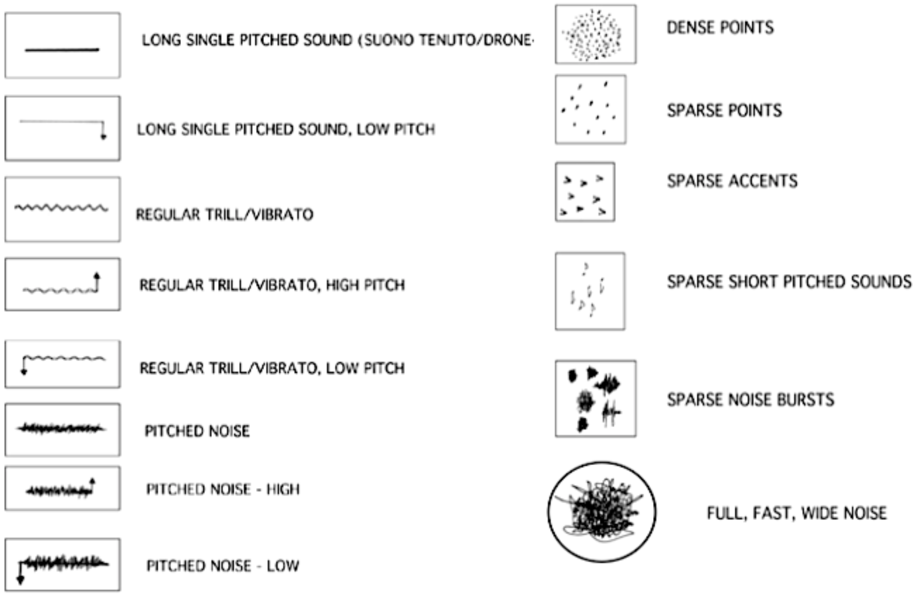


FIGURE 2.15: An excerpt from the Graphograms vocabulary

2.5 Graphograms

My first experiments with graphic notation for improvising ensemble started in 2012, when I was involved in some free improvisation projects and was looking for a way to guide the improvisation, to organize the development by defining some sort of *bendable* structure. Apart for the inspiring example of Anthony Braxton, I had in mind also the *Conduction* technique developed by Butch Morris and the strategies of John Zorn (*Cobra* and game pieces in general). The aim was to find a similar way of defining macro-structures, rules, organizational principles, and to suggest some kind of sonic materials, while at the same time giving the musicians enough freedom to let them express their ideas and personalities. I was also interested in defining a *nonlinear* kind of notation, that is, a scoring system that is not structured along a unique time-line. This led to the definition of a graphic notation system for improvisation, called *Graphograms* (*Graphograms*). This notation is based on the graph theory and a parametric vocabulary, which guides the musicians through the improvisation process (Figure 2.15). The score-graph defines some *sonic states* in which the musician can stay, and the chance to make some transitions to other states. The sonic states contain graphic elements defined in a vocabulary. Transitions to other states can be free or conducted, that is, triggered on cue by one or more leader(s). The choice of the path through the graph can be also cued, or left to the improviser. In this case, each musician has the possibility to make its own path through the graph, colliding and/or collaborating with the others. At each transition, a new sonic scenario suddenly emerges, forcing the improvisers to deal with the situation and to act rapidly and in a very focused way. This results in a "quasi-stochastic" music, which can rapidly move from quiet, soft scenarios to full noisy bursts of energy and then turn again to some sort of *obbligato* passage. The graphic vocabulary can be classified

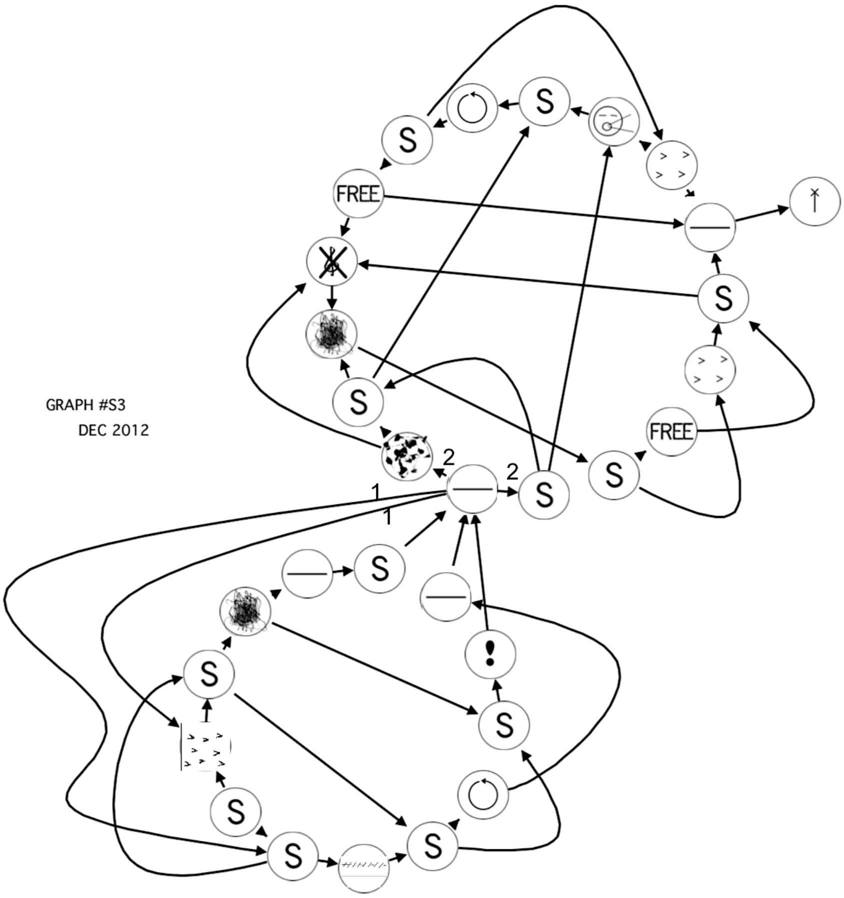


FIGURE 2.16: Graphograms example: Graph No.S3 (2012)

in three different categories of symbols/shapes:

- *Sonic material formings*: graphic shapes that suggest an area in the multi-dimensional parametric space (silence, staccato sounds, accents, long sustained drones, noise bursts...)
- *Rules*: states that contain some organizational rules or strategies (Copy, Develop, Focus, Repeat...)
- *Unconventional States*: a set of symbols that force the improviser to do something surprising, theatrical and/or unconventional

It is worth noting how in the Graphograms we already have the idea of the score as a non-linear structure which can be traversed like a cartographic map. This metaphor then became part of the CABOTO system. The Graphograms notation has been developed further while working with an octet, Oktopus Connection in 2013 (*Oktopus Connection*). A new edition of this graphic-based improvising ensemble, *Oktopus Connection Redutch*, has been recently put together for further explorations.

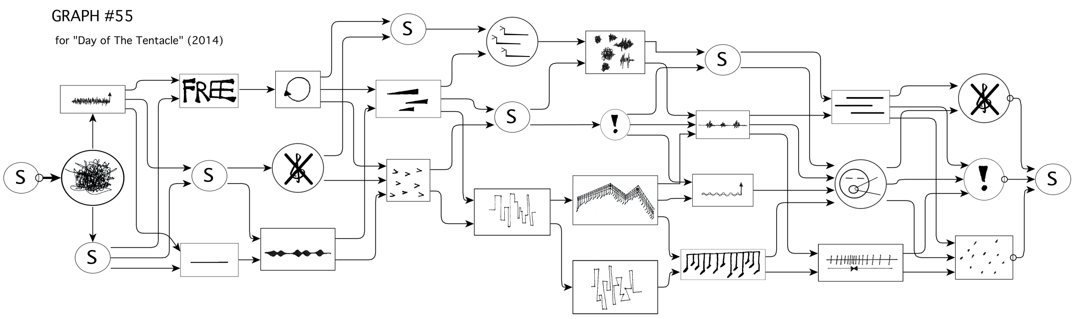


FIGURE 2.17: Graph No.55 (2014)

2.6 Ossi di Seppia (2016)

In 1925 Eugenio Montale published his most famous collection of poems, Ossi di seppia. The title means "Cuttlefish bones" and has always been strongly evocative for me: the inner essence of things, of reality itself, is presented here as a residual, a relic which survives the sea and arrives to the beach in the form of a white, polished cuttlefish bone. Montale's poems are full of sounds, like the "snakes rustle", the "blackbirds catcall", the "cicadas' wavering screaks ... a whole soundscape is presented to our ears through concrete, onomatopoeic words. For the electronic part, I decided to start from this poetic idea – Ossi di Seppia – and to work on prerecorded material based on bass clarinet sounds, processing them through spectral decomposition techniques, in order to extract the bones, the residuals, the relics of the original sounds. Starting from this concrete material, the piece was developed into five sections, each inspired by a sonic image taken from Montale's poems. For the bass clarinet score, I decided to use graphic notation, with a set of instructions for interpreting it (Figures 2.18, 2.19). It is worth presenting this work here, since the graphic vocabulary used in this score contains many elements that are found in the graphic notation developed for CABOTO. In particular, we can note the definition of some sonic formings such as points clouds, masses, lines, for suggesting some particular playing techniques. These graphic elements are described in the introductory instructions of the score (see Fig. 2.19).

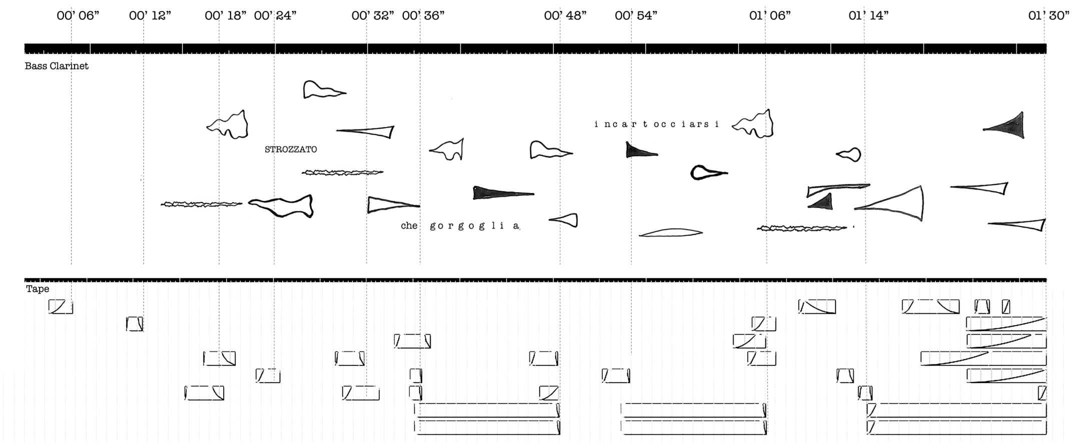


FIGURE 2.18: An excerpt from *Ossi di Seppia* (2016)

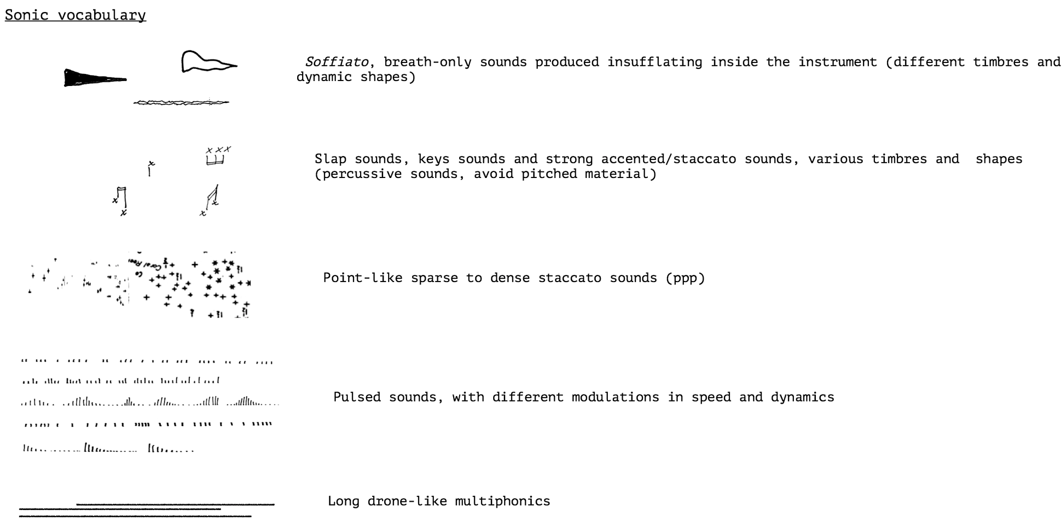


FIGURE 2.19: Graphic vocabulary for *Ossi di Seppia* (2016)

Chapter 3

Defining a Graphic Vocabulary

In this Chapter, the graphic vocabulary developed for the system is presented. This graphic notation is the result of personal aesthetic choices, and it draws inspiration from the various sources and previous works presented in Chapter 2. This vocabulary is not only a collection or *taxonomy* of graphic abstract elements. It also reflects musical choices, and it reveals a personal view on how to organize electronic sounds, so as to compose with them. In this sense, formalizing a graphic notation for a machine becomes a way to elaborate, to reflect on the composing process and to formalize some underlying aesthetic choices.

Once a vocabulary of elements is defined, we can look at the visual composition on the canvas and its relationship with musical composition. The way these graphic elements are combined on the score-canvas can be viewed as another stage in the composing process, although it cannot be separated: we have to look at the process as a whole. In fact, drawing the sketch is not merely putting *graphemes* together, it is a more integrated process, the result of a seamless movement from one element to another in the continuum of the drawing gesture.

In the next section, a detailed description of the graphic vocabulary is given. The vocabulary items are organized according to some geometric features, and classified into a set of *classes of shapes*. In Section 3.2 the composition of these graphic elements is addressed, as well as the relationship between spatial/visual composition and sound composition. Sections 3.2.1 and 3.3 focus on the issue of time representation and the cartographic map metaphor. Finally, in Section 3.4, some considerations about possible synaesthetic relationships are presented.

3.1 Points, Lines, Masses

In this abstract graphic vocabulary, geometry plays a leading role. Geometric elements such as points, lines and planar shapes, form the basis for the development of the graphic score. These elements are combined according to relations that can be expressed in terms that are used in fine arts and graphic design as well as in music and physics: *repetition*, *rhythm*, *density*, *rarefaction*, *tension*, *release*, *unity*, *variety*. A similar approach has been followed by Cornelius Cardew for composing *Treatise*. In the *Treatise Handbook*, which collects the notes he made while composing the score, he writes:

> The way the elements act on each other it is like chemical processes: acid bites, circles roll and drag, and bend the stave lines of "musical space"

And again in (Cardew, 1974):

> Treatise arbitrarily combines images of transformations that occur in the real world: images of mathematical or logical transformations (multiplication of elements, relations between pairs of dissimilar elements, presence and absence of elements), and of physical transformations (by fragmentation, exploding, squashing, bending, melting, interpenetrating, etc.)

All the elements of this notation are to be considered as *gestural shapes*, that is, graphic elements that embed the drawing gesture, the energy that created them. This results from the fact that the score is sketched by hand, and therefore the hand gesture is clearly recognizable in the sketch (e.g. in the non-uniformity of the line strokes). This has been taken into consideration also in designing the sound synthesis algorithms: in this way, the resulting sound contains a glimpse of the gesture that generated the sketch. Apart from its intrinsic shape, every geometric element can be associated to a set of parameters that describes its relative position on the canvas and its *bounding box*, that is, the smallest rectangle within which all the points of the shape lie. These parameters have been used to characterize the resulting sound, mapping them to certain sound synthesis parameters. This topic will be discussed in detail in 5.

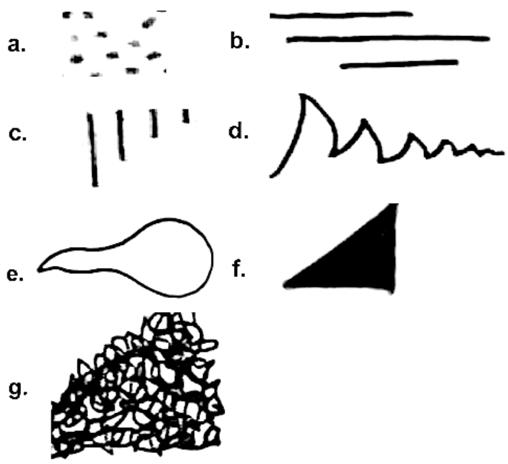


FIGURE 3.1: Graphic Vocabulary: seven macro-classes of shapes.

3.1.1 Points

The point is the result of the fastest, smallest gesture, and thus it is interpreted as the shortest, grain-like sonic event in the palette. These sound points are scattered in time as a result of a stochastic process, with a variable density controlled by the optical scanner value. Except for the relative position of the element, we can not gather any other piece of information from the point. Its size is too small to analyze its contour or its filling. Nevertheless, if we have the option to zoom in on the score, our point of view on this element could change drastically. This topic is discussed in detail in Section 3.3.

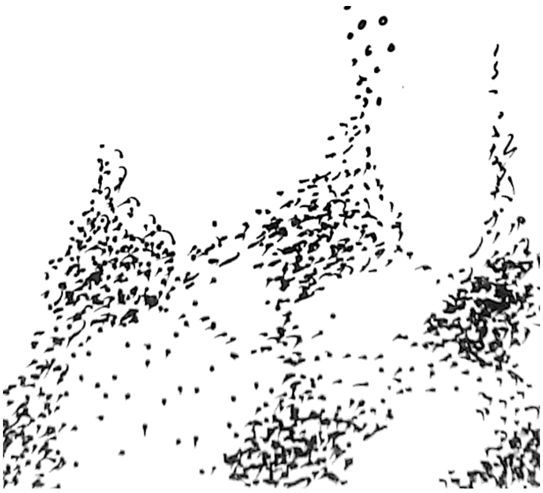


FIGURE 3.2: Graphic vocabulary: points formings

3.1.2 Lines

We may have different kinds of lines, but we can define all of them as gestural lines: lines that embed the energy of the performer's drawing gesture. A line is in fact the result of a single gesture, a seamless movement across the canvas. In interpreting the lines, the score is seen as an oriented two-dimensional plane with a top and a bottom. According to their shape, the lines are classified into *horizontal*, *vertical* and *noise* lines respectively.

Horizontal Lines

The horizontal line is associated with stability, stasis, balance, quietness. This element represent a long, sustained sound (Fig. 3.3). The uniformity of the line is translated into the uniformity of the sound through time, that is, a sound with minimal or no modulation of its parameters in time. In Western cultures, a longer horizontal line may suggest a longer time duration, since time is usually represented on the x axis. Here, a certain size won't be directly translated into a time duration. This will depend on the scanning rate of the waveform scanner, which is a separate mechanism. This will be discussed in detail in Chapter 5.

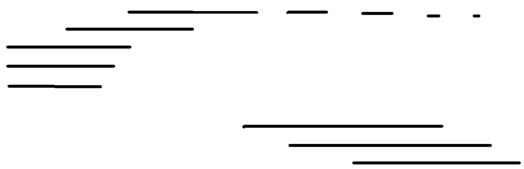


FIGURE 3.3: Graphic vocabulary: horizontal lines formings

Vertical Lines

A vertical line calls for an alert, attention, a strike or a beat. It is interpreted as a sonic event which has no time extension, but is entirely defined by its happening in a certain, precise moment. A vertical forming [1] is a *sign in time*, a pulse. These pulses are emitted in time, according to a certain rate controlled by the waveform scanner. This can result in a regular or irregular sequence in time. The only information conveyed, apart from the position, is the size along the Y axis. This is mapped to certain timbral characteristics of the sound.

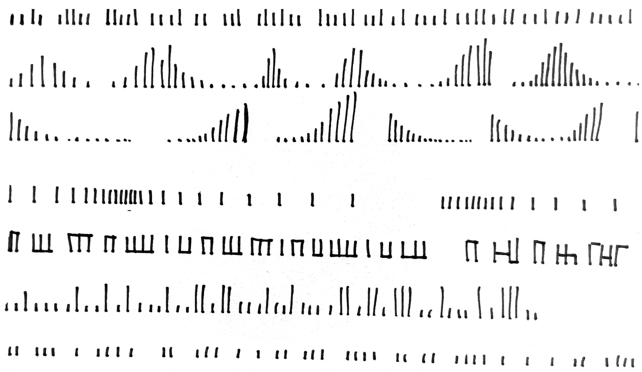


FIGURE 3.4: Graphic vocabulary: vertical lines formings

Line Variations: noise lines

When the gesture becomes more complex, line variations arise. If the hand of the performer starts deviating from an ideal average y position of the horizontal line, we move from quietness to agitation, disturbance, eventually turbulence and noise. Let's call these formings *noise lines* (Fig. 3.5). The contour of the line is now more complex and rich (curved, sharp, jagged, noisy...) and the wave scanning mechanism becomes the primary source of information for the sound synthesis.

3.1.3 Masses and Clusters

When a curved line gesture goes back to its point of departure, a two dimensional shape emerges. The planar region included then becomes part of the graphic element. I refer to this a *mass*. The mass can be made of a contour line only, and is then

[1] I use the term *forming* here and in the following sections to define a certain abstract audio-visual shape, as suggested by Braxton (see 2.3.1)

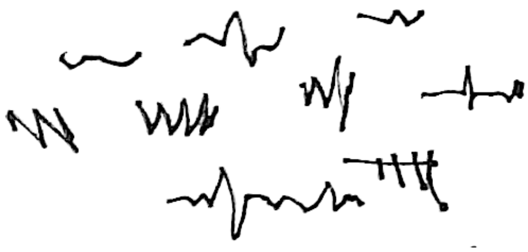


FIGURE 3.5: A set of *noise lines*

defined as a *void mass*. If the internal space is filled with an uniform color field, it is called a *filled mass*. If the internal space is filled with a texture or a curved line, the mass becomes more dense, and complex formings emerge. The resulting shape may have a strong geometric character (e.g. triangular, rectangular shapes) or a more organic character. Thus, the contour shape has to be taken into account in designing the sound.

Void Masses and Filled Masses

If the internal space of the mass is empty, or filled with an uniform color, the mass is defined only by its contour line. No other information is conveyed, except for its size and relative position on the canvas. For both of them, the chosen sound material is dense and uniform, with a timbral character that can be defined by its spectral width, density and quasi-static behavior. The contour line drives a modulator that moves the masses in time-space. The shape's character (regularity, irregularity, level of complexity) will then emerge.

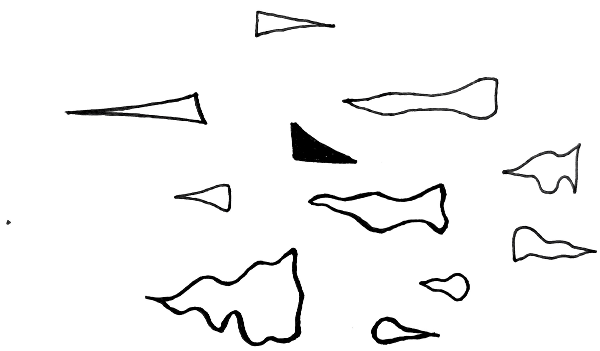


FIGURE 3.6: Graphic vocabulary: mass-like formings.

Noise Clusters

If the internal space of the mass is more complex, the mass may lose its compactness and become a cluster. I call this a *noise cluster* (Fig. 3.7). The sound is still dense, but more granular, and the timbre is stochastically evolving in time at a high rate,

FIGURE 3.7: Graphic vocabulary: noise clusters and textured masses.

according to the optical scanner signal. As for the other masses, the contour line drives an amplitude modulator that moves the grains in time, that is, it will shape the amplitude contour of the grains' population in time.

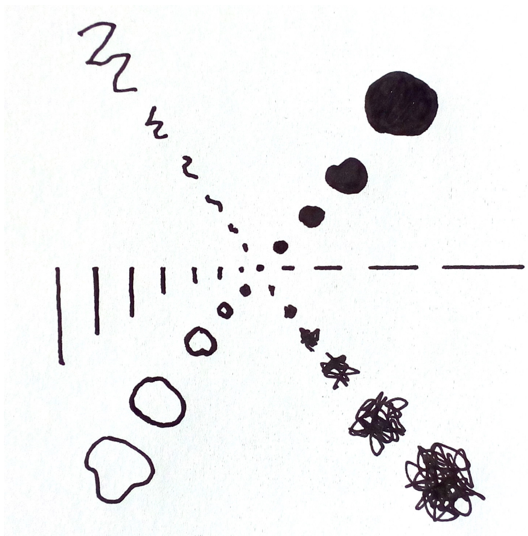


FIGURE 3.8: Composing graphic elements: the "morphing star diagram".

3.2 Composing the score: Establishing Relationships Between the Graphic Elements in the Compositional Space

The score is not made up only by the graphic elements that we see. It is defined also by what we don't see, by the void, the space in which these elements are combined. This two dimensional space is the canvas, and when we look at it as a whole, the concept of spatial composition arises, as defined in fine art. The problem here is how this concept is related to the sound composition. This correspondence (if present) will depend on how the score is interpreted or, more precisely in my case, how it is scanned and interpreted. We could for instance look at the score as a whole, and realize (in the sound realm) what we see, as a sort of persistent sonic image. This would result in a (more or less) complex *sonic state* that is *now*, and continuously perpetuating itself until we decide to stop the scanning process. This does not mean that the resulting composition would be static or quasi-static. The sonic processes that realize the graphic elements could be designed in a way to generate ever-changing material. Nevertheless, the development would not be dictated by information contained in the score itself. If, on the other hand, as in this case, the system scans the score cropping a view out of it and then moving along a certain path or trajectory in the two-dimensional space, a development in time arises. If this trajectory is continuous, for a given point in the score the resulting sound composition will develop according to the surrounding elements. Drawing two graphic elements close to each other will put them in relation both graphically and sonically. In organizing the elements on the canvas, I have focused on three different composing techniques that I define as: a) *contrast*, b) *morphing*, c) *force relationships*. *Contrast* means combining elements of striking different characters, such as a tiny horizontal line and a noise mass (Figure 3.9). On the other hand, *morphing* looks for continuity, moving seamlessly from one shape to another (see Figures 3.10, 3.11). With the term *force relationships* I refer to a kind of composition which relates to physics, and thus looks at the elements on the canvas as physical bodies with certain properties (e.g. stiffness, mass) subjected to forces of various kinds. For instance, in Figure 3.12, a mass becomes a center of gravity, dictating the composition of the surrounding elements.

The scanning trajectory can be oriented according to a score dimension, as in traditional Western notation and literacy, or we can break this linearity and look at the score according to a different metaphor, as described in detail in the next Section.

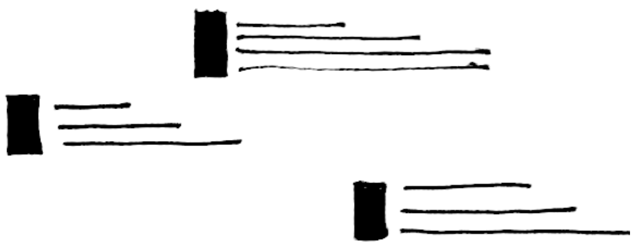


FIGURE 3.9: Composing graphic elements through contrast: masses and horizontal lines

3.2.1 From the Time/Pitch Space to The Carthographic Map Metaphor

The concept of time expressed as space is of crucial importance in representation of music. As noted by Emilios Cambouropoulos (Cambouropoulos, 2015):

> Many researchers have studied the correspondence between space and

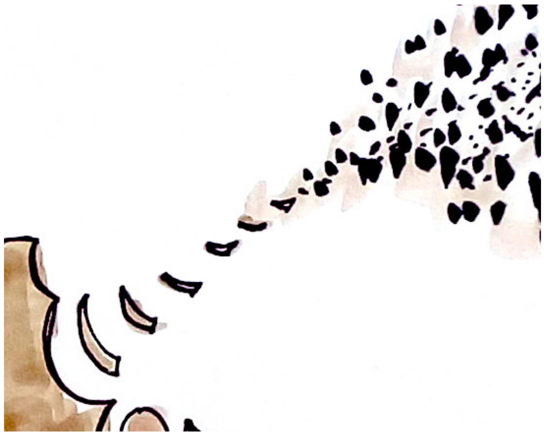


FIGURE 3.10: Composing graphic elements through morphing: mass to points

FIGURE 3.11: Composing graphic elements through morphing: vertical lines to points.

> time in human conceptualisation, and more specifically the use of spatial metaphors in temporal reasoning, [..] and, more specifically, the employment of two basic spatialisation metaphors of time in the conceptualization of musical time and motion. Frequency/pitch is also commonly conceptualised in terms of spatial metaphors, namely along the high/low spatial axis [..] The combination of the x-axis spatial representation of time with an orthogonal high/low pitch y-axis gives rise to the most common 2D representation of music [..] The correspondence between time and space is so strong that it is difficult to conceptualise music without thinking about shape or pattern.

George Athanasopoulos et al. (Athanasopoulos, Tan, and Moran, 2016) have recently published a comparative study on the visual representation of sound in different cultural environments. An interesting result of this study is that the Cartesian representation of sound events, where time is represented on the *X* axis, is a cultural influence probably derived from literacy. In developing CABOTO, the issue of time has been a crucial one. In the first prototypes, in which the system was intended as a composing tool rather than an instrument for live performance, time was represented on the *X* axis, as in traditional Western notation. This led to a conventional representation of the composition, which was quite intuitive in some ways, while on the other hand it led to predictability, and forced one to think about the composition process in a very strict and time-oriented way. These considerations led

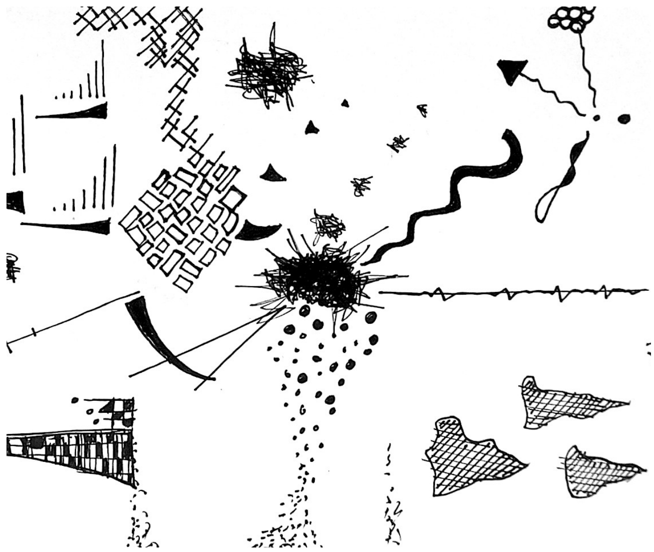


FIGURE 3.12: Composing graphic elements through forces relationships: a mass that creates a "gravity field".

me to the shift from a time-based score to the concept of the score as a cartographic map (Miller, 2017). According to the map metaphor, the two-dimensional canvas is viewed as the representation of some kind of *terra incognita* which is explored by navigators (Figure 3.15). There are different kind of maps, and different ways of read them. The way we traverse the score and the path we choose, affects the resulting information we gather from the score itself. One or more paths (or algorithms that generate paths) can be defined in order to explore the score. Moreover, if time is removed from the x axis, this Cartesian dimension is now available to be used for controlling other parameters.

3.3 Through the Magnifying Glass: Zooming in on the Score

A cartographic map is designed according to a certain scale. This is quite important for the user, since a blue stain on the map could represent a pond or an ocean. In the past, a traveler had to bring maps with different scaling factors, and use them accordingly. Nowadays, we have interactive online maps that allow us to change the scale, zooming in and out of a certain area. A similar zooming action can be applied to the score-map. In the first stages of development of CABOTO, this zoom feature was introduced as a standard preprocessing module for correcting the score image view. Nevertheless, it became a powerful tool for modifying the score in real-time. The most interesting result of zooming is that the scale of the geometric elements changes: points become lines, lines become masses. Zooming into a noise cluster may result in a compact filled mass. As we zoom deeper, we eventually reach the pixel level, where everything is quantized into squares. The way we look at the score and the perspective we choose changes the score itself.

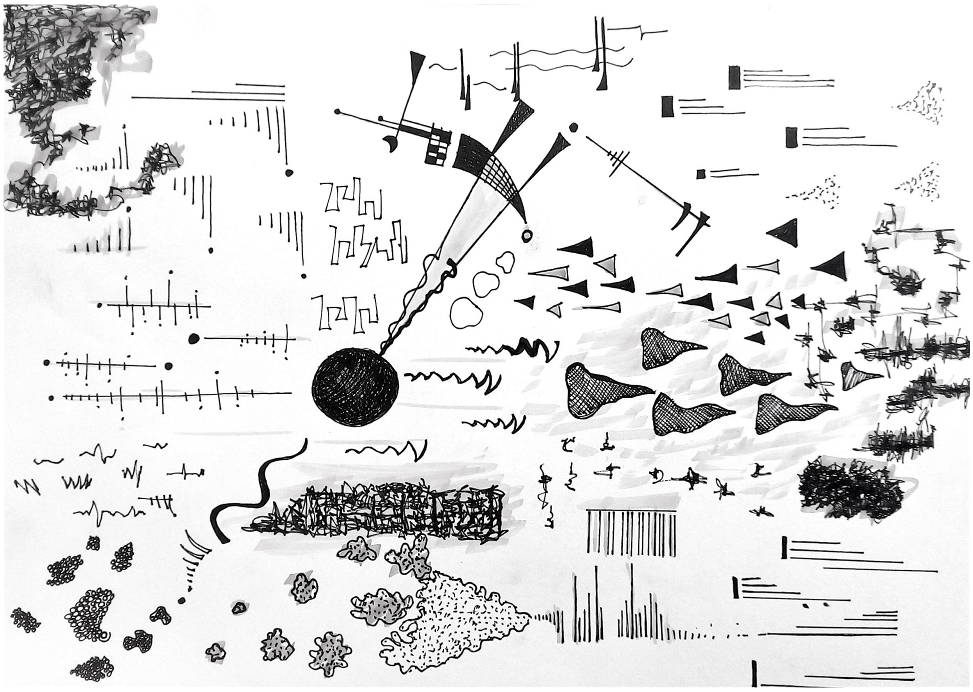


FIGURE 3.13: An example of a score composed for the system.

3.4 Mapping: The Synaesthesia Connection

In a famous experiment by Vilayanur S. Ramachandran and Edward Hubbard (Ramachandran and Hubbard, 2001), derived from Wolfgang Köhler (Köhler, 1929), people were asked to assign names to two different geometric shapes. The provided names were "*Bouba*" and "*Kiki*", and the shapes were a curved, smooth shape and a more sharp-angled one. The results of this experiment suggested that the association between shapes and sounds is not connected to cultural biases, but to a human brain feature. Other studies in phonaesthetics (Crystal, 1995) and sound symbolism (Hinton, Nichols, and Ohala, 2006) suggested that the link between the sound of certain words and their meaning is also not completely arbitrary. In developing the graphic vocabulary and its mapping into the sonic domain, I have decided to avoid exploring the issue of synaesthesia, since this would have implied a kind of research more related to phonetics and cognitivism, which would have taken me off topic and away from the focus of this research. Although the synaesthetic approach has been taken into account in designing the mapping strategy and sound synthesis processes, the graphic system developed here is based on personal choices and the mapping is still arbitrary and justified only by personal aesthetic choices. Nevertheless, we can observe an interesting link between the results of these studies in linguistics and psychology and the mathematical properties of waveforms. Consider the graphical representation of a sound pressure wave, that is, the pressure vs time Cartesian plot (or voltage vs time). If we listen to the synthesized sound corresponding to that shape by reading the wave as a wavetable (in the digital domain)

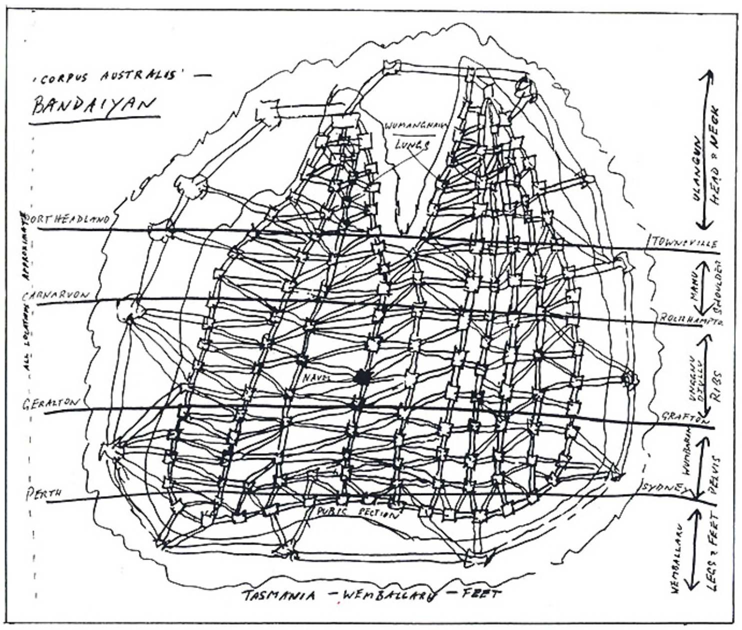


FIGURE 3.14: The cartographic map metaphor: Songlines in the Australian bush. The songlines are paths across the land, but also mythological chants.

or playing it with an optical device similar to the ones used in analogue film technique, we can verify that a "sharper" kind of waveform will sound harsher, since its spectrum will contain more components, more partials. On the other hand, a sinusoidal shape will have few or even just one spectral component (the fundamental), resulting in a "smoother" sound output. Since the system includes a waveform scanner, this synaesthetic link is embedded in the system. For instance, a noise line will be recognized by the symbolic classifier, and the waveform scanner will extract the shape of the line contour translating the shape into an audio signal with certain spectral characteristics. This "raw" signal will be used into the synthesis process,

FIGURE 3.15: The traces of four explorers scanning the graphic score.

both as a waveform and as a modulator and/or control signal. In this way, the visual quality of the shape will be reflected in the resulting sound. This integrated symbolic/raw scanning technique is described in detail in chapter 5.

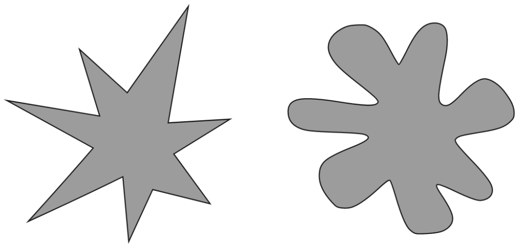


FIGURE 3.16: Bouba and Kiki (from Wikimedia)

Chapter 4

Background: Scanning Graphics

4.1 Introduction

We can divide this historical survey into two main phases. The turning point between the two is - not surprisingly - represented by the introduction of digital machines in the electronic music realm, thus marking the beginning of sound and music computing. It is worth noting how the interest in making sound out of graphics was more extensively explored in the pre-computer decades than after the introduction of digital computing. In a time where the available technology was greatly more "primitive" with respect to today's computers, scientists and composers were struggling with celluloid, photocells and valve oscillators, obsessed by the dream of sketching sounds and thus hearing to new, unexplored music. This is not so surprising if we consider the historical context. In the years preceding the Second World War, cinematic art was the *new sensation* and a powerful instrument for propaganda. In 1929 the optical soundtrack was introduced and for the first time audiences could experience a synchronized audio/video stream in the cinema. In Soviet Russia, in the aftermath of the October Revolution, a vivid artistic scene flourished, which was utopian, anarchist, materialist and inspired by science and technology (Smirnov, 2013). This chapter moves from the very first experiments in visualizing sound waves, through the introduction of the optical sound recording in film industry, to the explorations in "hand-drawn sound" in Soviet Russia and Germany during the 1930s. Then we focus on other applications of these analogue techniques by various artists and researches during the second half of the twentieth century. Finally, a survey of graphic-based systems in the digital domain brings us to the present day and the state of the art.

4.2 From Ur-Images to Tonschreibekunst: Early Phonographic Studies in the Twentieth Century

The first visual patterns that can be considered as a graphic transcription of sound appeared in 1787 in a text entitled "Entdeckungen "uber die Theorie des Klanges" by German physicist and musician Ernst Florens Friedrich Chladni. He called these images *Klangfiguren* and he discovered that these were related in a distinctive and

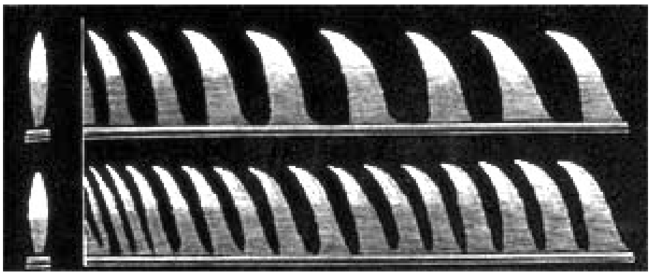


FIGURE 4.1: Illustration of two manometric traces. From Koenig's Acoustic Catalogue, 1865.

precise manner to specific tones. For the first time, we had a non-arbitrary visual representation of sound which was not a product of human conventions such as in musical notation:

> What was so exciting about these acoustic "ur-images" (as a contemporary of Chladni called them) was that they seemed to arise from the sounds themselves, requiring for their intelligibility not the hermeneutics appropriate to all other forms of musical notation but instead something more akin to an acoustic physics.

In 1862, German physicist Rudolph Koenig realized a particular manometric capsule in which the changes in pressure due to sound vibration produced a variation in the opening of a gas valve and thus a change in the intensity of a burning flame. Edward L. Nichols and Ernst George Merritt captured these flickering images on photographs, thus realizing what was believed to be the first *visual* recording of an acoustic phenomenon. Nevertheless, in 2008, the audio historian David Giovannoni, along with a pool of researchers from Berkeley lab, (Dearen, March 28, 2008) discovered some recordings from the French Academy of Sciences that dated back to 1853-54. These were created using a "phon-autograph" or *phonautograph*, a device developed by Edouard-Leon Scott, which is now considered to be the first sound recording device. It is interesting to note that the recording mechanism was inspired by human anatomy: Scott, a printer and bookseller, became interested in a book about auditory anatomy. He then had the idea of replicating the timpani-ossicles system through the means of a membrane, some levers and a bristle that scratched a surface covered in lampblack. Scott collaborated with Koenig in building different devices of this kind, but never succeeded in making any kind of profit from his invention. The output of the phonoautograph was a visual image of sound, and it was not possible to playback the recording. In fact, the purpose of the device was to investigate the properties of sound waves, and in particular of vocal sounds.

It is worth noting how scientists and researchers involved in these early experiments were principally interested in speech sounds and language (Levin, 2003). In fact, around 1830 a branch of linguistics known as *phonography* or *vibrography* was taking its first steps, and the main focus of these researches was the relationship between speech and inscription. Physiologists and phoneticians such as Franciscus

Cornelius Donders used such devices for recording speech and then try to relate the different sound wave shapes to letters. These - of course - mostly unsuccessful experiments were also attempted by Thomas Alva Edison himself, after he invented the first acoustic read/write device: after many transcriptions of the letter 'a', he realized - we could imagine with great disappointment - that if he compared two different transcriptions, they "were absolutely dissimilar". (Gitelman, 1999). Scott himself was not interested in sound reproduction, but in "writing speech, which is what the word phonograph means" (as stated in a self-published memoir in 1878). The possibility of visualizing sound opened up many scientific and philosophical questions about the relation between sound, image and meaning. As observed by Jonathan Sterne in his book *"The Audible Past: Cultural Origins of Sound Reproduction"* (Sterne, 2003):

> "Through modern physics and acoustics, and through the new relation between science and instrumentation, auditory and visual phenomena could be first isolated and then mixed or made to stand in for one another. Scott's discourse on the phonautograph and its successors suggests that this kind of synesthesia—of mixing codes and perceptible material—is a constitutive feature of technological reproduction of sound and image."

In Rainer Maria Rilke's "Ur-Geräusch" (Primal Sound) we have the first glimpse of an intuition about using the new technology for reproducing sounds that have never been recorded. In his essay, Rilke shares a memory of his youth, when he was studying anatomy and became fascinated with the skull, to the point of keeping one on his desk and observing it from time to time at the light of a candle. He then noted the suture traces on the skull surface and these reminded him of phonographic cylinder engravings:

> The coronal suture of the skull (this would first have to be investigated) has – let us assume – a certain similarity to the closely wavy line which the needle of a phonograph engraves on the receiving, rotating cylinder of the apparatus. What if one changed the needle and directed it on its return journey along a tracing which was not derived from the graphic translation of sound, but existed of itself naturally–well, to put it plainly, along the coronal suture, for example. What would happen? A sound would necessarily result, a series of sounds, music... *(Rainer Maria Rilke, Primal Sound (1919))*

As noted by Thomas Levin, in this desire to listen to these *unheard signs* there is not much of a scientific interest or a pure aesthetic exercise, what we could nowadays call a data *sonification*. There is also the implicit assumption that in those suture lines a message of some sort is hidden, and could be revealed. Rilke was looking for a secret meaning, a message from the unknown.

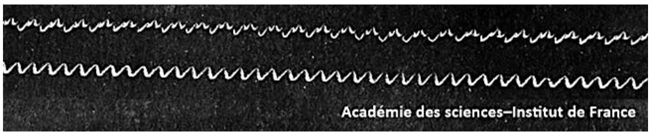


FIGURE 4.2: A sound inscription made by Scott's Phon-autograph

4.3 Tri-Ergon, Lee De Forest and the Rush for Optical Soundtrack in film Industry

In the years following the end of the First World War, a number of patents related to sound-on-film started to appear both in Europe and in the United States. Most of these inventions were an attempt of give voice to movies. At that time, cinema was the new sensation, but movies were still silent: the main issue being how to synchronize the soundtrack with the scrolling film. The Finnish inventor Eric Tigerstedt, the Tri-Ergon company in Germany and the American inventors Lee De Forest and Theodore Case independently developed similar techniques for recording sound on film. This eventually resulted in some controversy about patents attributions (De Forest was accused by his collaborators of taking full credit for the work, while Tri-Ergon engaged in a long legal battle in the US). The battle for gaining the role of film industry standard was finally won by RCA with its Photophone system.

Sound-on-film techniques can be classified according to the recording method, that is, how the waveform is represented on the film. The "Variable Density Method", implemented in De Forest's Phonofilm and in the Tri-Ergon device, modulated the intensity of a light bulb for transferring the signal on the photosensitive film. The resulting optical track is a series of vertical lines of different brightness, according to the sound pressure level (Figure 4.4. It was then possible to reconstruct the audio signal by means of an electronic circuit driven by a photocell which picked up the light signal obtained while the film was rolling in front of a lamp (see Figure 4.3. On the other hand, the "Variable Area Method", implemented by the RCA Photophone, produces a filled waveform which will be broader or narrower depending on the magnitude of the signal. This method has two variations: *Unilateral Variable Area* and the *Bilateral Variable Area*. In the latter, the resulting waveform is mirrored and thus symmetric with respect to an axis parallel to the film direction.[1]

4.4 Laszlo Moholy-Nagy, prophet of the optical sound synthesis

The artist, filmmaker, designer, professor Làszlò Moholy-Nagy has been a crucial figure of the avant-garde since the early 1920s and a strong supporter of the integration of technology in the art. In an era of new art-related technological wonders

[1]Interestingly, this is the kind of waveform representation with which we all became familiar and that is widely used, for instance, in digital audio workstations and audio software in general.

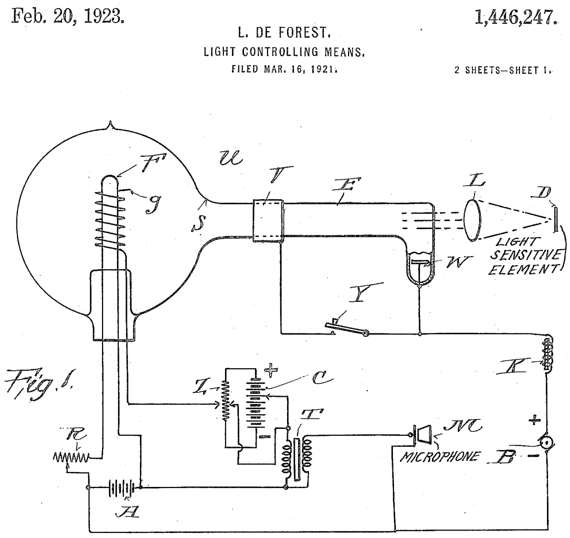


FIGURE 4.3: An illustration from the original Patent No. US1446247 by Lee De Forest, which shows the sound-controlled mechanism for recording the optical soundtrack.

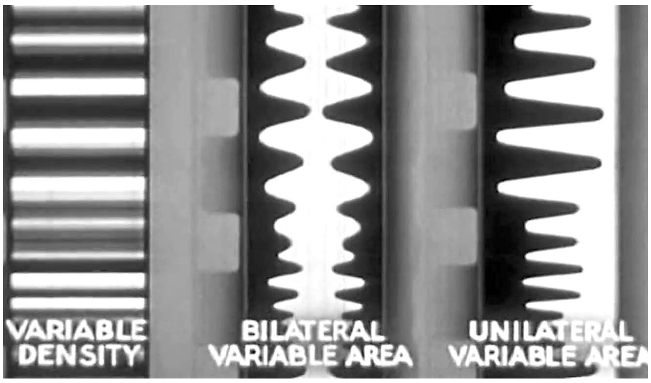


FIGURE 4.4: (Three different methods for sound-on-film recording. From "Sound Recording and Reproduction", a 1943 educational documentary by EB Films-ERPI)

such as the Trautonium, the Theremin, and sound-on-film technique, Moholy-Nagy encouraged and challenged artists to use these new instruments as means for exploring the unknown [2]. In his essays he presented his almost prophetic vision about technology and art, depicting a future where composers would master the phonographic vocabulary to the point of being able of compose new, unheard sounds and music directly in the “groove-script alphabet”:

> the grooves are incised by human agency into the wax plate, without any external mechanical means, which then produce sound effects that

[2]Many of these instruments were, in fact, primarily used in the most conventional and familiar way, and we may suppose that this should have quite disappointed Moholy-Nagy. For instance, the Theremin was being used as a weird imitation of the violin or the soprano voice for performing classical arias. Levin suggests also another explanation:[...] radically new technology for the generation of sound attempting to legitimate itself not by foregrounding its own unprecedented sonic capacities but by slavishly simulating well-known classical pieces”

> would signify—without new instruments and without an orchestra—a fundamental innovation in sound production (of new, hitherto unknown sounds and tonal relations) both in composition and in musical performance. The primary condition for such work is laboratory experiments: precise examination of the kind of grooves (as regards length, width, depth, etc.) brought about by the different sounds; examination of the man-made grooves; and finally mechanical-technical experiments for perfecting the groove-manuscript score. (Or perhaps the mechanical reduction of large groove-script records.)

Moholy-Nagy himself conducted some experiments along with George Antheil, and Paul Hindemith was also interested in the matter (he premiered a piece for gramophones in Berlin in 1930). Nevertheless, they were not able to achieve any significant results, due to technical limitations: the grooves were actually too small to be controlled by hand. As described in the previous section, this is the point when the new optical sound devices arrived, and Moholy-Nagy's dream could finally become true. He also immediately recognized that optical sound was the technology he was waiting for. In his essay "Problems of the Modern Film", first published in 1928 (Moholy-Nagy, 1964) he called for a new experimental "abstract sound film" art. The optical soundtrack could finally overcome the technical limitations of the gramophone, since the waveform was now visible to the human eye and could then be drawn by hand directly on the film by means of traditional painting tools such as ink pen and brushes. It can be noted that Moholy-Nagy assumption was still related to the idea that the waveform trace is the expression of a kind of language which can be codified and classified in an alphabet in order to be fully understood and mastered:

> It will not be possible to develop the creative possibilities of the talking film to the full until the acoustic alphabet of sound writing will have been mastered. Or, in other words, until we can write acoustic sequences on the sound track without having to record any real sound. Once this is achieved the sound-film composer will be able to create music from a counterpoint of unheard or even nonexistent sound values, merely by means of opto-acoustic notation. *("Problems of the Modern Film", 1928)*

The focus here is still on the language, an universal, kind of naturally-given code that these devices are able to bring to human perception. The modulated light signal impressed on the film, as previously the vibrating stylus engraving the groove, are means that reveal the same *truth*: a sort of new, foreign, never heard language. There is a shift here from a physical phenomenon which can be ascribed to acoustics and physics of vibrations, to an anthropocentric view. We could argue whether this perspective has been influenced by how these devices were designed: the fact that the sound was translated into a trace which develops linearly across a surface (a cylinder, a film roll), had probably encouraged an immediate connection with writing and from this metaphor a whole semantic universe exploded, bringing in the concepts of language, alphabet, signs, and so on. This resulted also in breaking up the continuum of the sonic phenomenon into *quanta* of sounds, into phonemes, following the assumption that we are dealing with the transcription of a language. As presented in the next Section, the ideas of Moholy-Nagy became reality thanks to the work of optical sound pioneers such as Rudolph Pfenninger, who was praised by Moholy-Nagy himself as the inventor of such technique. In that period many artists and scientists were working independently on the topic, thus is probably not possible to identify a unique, first inventor of "hand-drawn sound".

4.5 Optical sound experiments in Soviet Russia

Starting from 1930, different groups of researchers were exploring the new possibilities made available by the optical soundtrack, both in Russia and in Germany. As far as we know, these groups were working independently from one another, and none of them knew of the advancements of the others. Moreover, it appears that researchers were quite jealous about their discoveries and they tried to be secretive about them. For instance, it is known that the sheets shown by Oskar Fischinger in some pictures published at the time were actually a fake [3].

In Soviet Russia, the film industry was highly developed and encouraged. The Communist Party envisioned in the new art a powerful means for propaganda and for teaching the illiterate population. Lenin himself, in the 1919 decree on film industry nationalization, proclaimed (Taylor and Christie, 1994):

> 'The art of cinema is the most important of all arts for us today!'

As noted by the English critic John Berger (Berger and Neizvestny, 1969), a curious short-circuit took place between avant-garde artists' ambitions and the needs of the Revolution: artists wanted to be free, to experiment, to be anarchist, to break with the past and the tradition, to engage with science and technology. On the other hand, the regime could accept some of these desires, in as far as they were framed in or *dressed as* revolutionary concepts. For instance, the fascination with technology could perfectly work for both, since the Revolution needed technological advancement for industrialization, for modernizing Russian society and creating the "new humanity", while artists wanted to experiment with new devices for creating a new art. The film production industry appeared to be the perfect realm for condensing all these aspirations. Colossal projects were produced, and in this vivid scenario, in which different personalities such as directors, animators, engineers and composers were clashing together, optical sound appeared. Andrei Smirov Smirnov, 2013 states that during the realization of the movie *Piatiletka. Plan velikih rabot* by Abram Room,

[3] This could have been also due to practical reasons, since the actual optical soundtrack was too tiny to be displayed

within the working group was the illustrator Tsekhanovsky, who, talking with the composer Arseny Avraamov, said: "what if we take some Egyptian or Greek ornaments and use them as the soundtrack? Perhaps we will hear some unknown archaic music?". The day after, the group was already working on this new concept, and the graphical sound technique was born. In 1930, Avraamov produced the first hand-drawn motion picture soundtracks, realized by means of shooting still images of sound waves sketched by hand 4.5. In his experiments, he focused on geometrical patterns that were called "ornaments" at the time. If examined today, an electronic musician would easily recognize many familiar shapes that became standard waveforms in analogue synthesizers: sawtooth, triangular, square waves. Those shapes were not designed according to any knowledge of the sound spectra, but were inspired by geometrical and/or aesthetic considerations. In some cases, Avraamov took inspiration from mathematics and chemistry, but he also introduced some decorative-like patterns or even figurative elements (4.6). Unfortunately, the archive of Avraamov works (which included more than 2000 meters of film) went lost during the 1930s (apparently as a result of a domestic accident provoked by his sons).

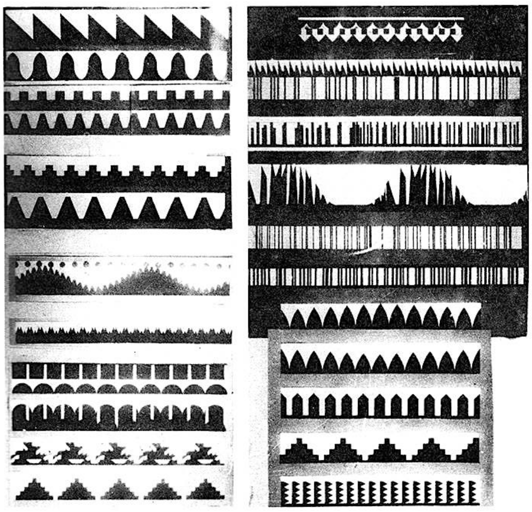


FIGURE 4.5: Hand-drawn 'ornamental sounds' by Arseny Avraamov (1930-31).

In 1931 Evgeny Scholpo, in collaboration with Georgy Rimsky-Korsakov (grandson of the famous composer) built the first version of the *Variophone*. The method implemented in Scholpo's device was quite unique and opened up new possibilities: instead of drawing or using fixed silhouettes photographed on the optical soundtrack, Scholpo made use of rotating discs with different shapes, that were moving in sync with the film strip. Controlling the speed of the discs he could then obtain different pitches and glissando effects.

FIGURE 4.6: Profiles of Yankovsky and Avraamov used as optical soundtracks, 1931 (from Smirnov, (Smirnov, 2013)).

4.5.1 The ANS Synthesizer

Evgeny Murzin started working on a new photo-electric musical instrument around 1939, and the device was finally patented in 1957, and called the ANS synthesizer. This device is quite innovative with respect to the previous optical sound devices. In fact, here the sound was synthesized by adding a number of sinusoidal signals, thus implementing a sort of additive synthesis technique. These sinusoidal components, up to 576, were generated through optical oscillators, tuned to a microtonal scale, and controlled in amplitude by a graphic score which represented amplitude trajectories in the time-frequency plane, etched on a black-covered glass surface (Figure 4.7). This allowed the composer to directly design the spectrum of the sound, or what nowadays is called a *sonogram*.

FIGURE 4.7: Graphic scores composed for the ANS Synthesizer.

4.6 Rudolf Pfenninger, Allan Humphriss and Oskar Fischinger

In the meanwhile, similar researches were conducted in England by Eric A. Humphriss and in Germany by Rudolf Pfenninger and Oskar Fischinger. Humphriss made quite a sensation in 1931 when he presented what was then called the first "robot voice". He painstakingly studied the waveforms of the different syllables of the English language, to the point that he claimed to be able to read and draw a soundtrack. To demonstrate that, he synthesized the sound of a voice, sketched by hand, saying the sentence "All of a tremble". His ability was then employed by the RKO-Pahté film company for changing the name of a character in the soundtrack of the movie "Born To Love" (1932). While Humphriss was interested in the resynthesis of existing sounds, Rudolf Emil Pfenninger was intrigued by the possibility of creating new sounds. Pfenninger was an animator working at the Kulturfilmabteilung of EMELKA, a large film production company in Munich. It appears that he arrived at optical sound experiments for economic reasons: he could not afford the musicians nor the recording studio for creating the music for his animations, thus he decided to draw the soundtrack by himself. He sketched the waveforms on strips of paper, that were then photographed and impressed on the optical soundtrack. EMELKA produced a series of his films with "handwritten sound", entitled "Die tönende Handschrift: Eine Serie gezeichneter Tonfilme eingeleitet durch ein Film-Interview", which was premiered on October 19, 1932.

The work of Oskar Fischinger in Berlin, developed around the same years (1932-33), although technically very similar to the one of Pfenninger, had some original aspects related to the kind of sonic material Fischinger was interested in. In fact, his research focused on non-musical sounds and more precisely on iconic sounds: he was trying to relate sounds and shapes in order to find a sort of common audio-visual vocabulary. We can relate this fascination for "sound shapes" with successive explorations in psychology like the famous "bouba-kiki" experiment 4. As a final remark, we can note however as in the cases of both Fischinger and Pfenninger, the focus was more on exploring a phono-graphic vocabulary or an audio-visual semantic, than on making something artistically interesting with the new technique.

4.7 Norman MCLaren and the experimental cinema scene

After the first pioneering experiments described in the previous sections, optical sound techniques were used in different scenarios, but especially in avant-garde cinema and animation. Norman McLaren started exploring synthetic animated sound while he was still an art student at the Glasgow School of Art , and he recalls himself how he was strongly influenced by Pfenninger's system (Levin, 2003). He then realized a number of hand-made animations, among others *Synchromy* (1971), produced by the National Film Board of Canada. McLaren used two different kind of techniques for his animations. In the first part of his career he experimented with free-hand drawing on the sound track, in an improvised and gestural manner. Later on, following he example of Pfenninger, he developed a library of film strips, each containing a specific waveform at different frequencies. Later on, other animators and experimental film makers made use of optical sound techniques from the 1950s on. Among the others, it is worth citing the work of the Whitney Brothers, Barry Spinello and Guy Sherwing.

4.8 Daphne Oram and The Oramics

With the work of Daphne Oram at the BBC Workshop during the 1960s Hutton, 2003 we shift our attention on a branch of the topic which can be considered as an evolution of the work of Murzin and his ANS synthesizer, and thus bring us closer to the topic of our research, that is, making use of a graphic system for composing electronic music. Such an instrument was first imagined by Edgar Varése in 1939, when he wrote:

> . . . (My machine) . . . will work something like this: after a composer has set down his score on paper by means of a new graphic notation, he will then [...] transfer the score directly to this electric machine...The advantages of such a machine [...] (are) [...] liberation from the arbitrary, paralyzing tempered system; [...] the formation of any desired scale [...] the possibility of obtaining any differentiation of timbre, of sound- combinations; new dynamics far beyond the present human-powered orchestra [...] (Varése 1939 in Schwartz et al. 1998: 200)

Oram had a clear vision of the instrument she wanted to build, and wrote about it when she started working on its developed:

> I visualize the composer learning an alphabet of symbols with which he will be able to indicate all the parameters needed to build up the sound he requires. These symbols, drawn . . . freehand on an ordinary piece of paper, will be fed to the equipment and the resultant sound will be recorded onto magnetic tape. (Oram 1962: 9)

The Oramics instrument's main component was a set of ten 35mm film loops on which the composer could sketch waveforms and envelopes. These graphic elements were then scanned by photocells which controlled a bank of oscillators and filters. The resulting sound was then recorded on tape. Although Oram received some grants for developing the project, the instrument was never really completed in all its details, and the device was perhaps too complex and not really practical enough to become a commercial product that could be used by other composers, although Oram did carry out some collaborations with the composers Thea Musgrave, Hugh Davies and Tristram Cary.

Nevertheless, Oramics can be considered a predecessor of other graphic-based tools for composing electronic sounds. The evolution of such systems, however, did not appear in the analogue domain: In fact, the research in the analogue domain evolved towards building instruments for performing, such as the Moog and the VCS3. But at that point, the promising new world of digital machines appeared at the horizon, and electronic composers perceived in this new realm, whose language consisted of routines, schedulers, tasks, a promising land for machine-assisted composition.

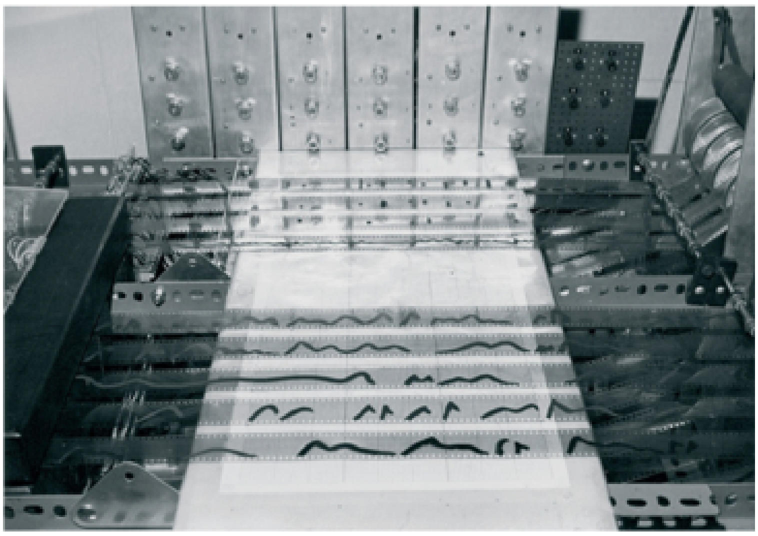


FIGURE 4.8: Daphne Oram's Oramics instrument.

4.9 The Computer Era: from Optical Sound to "Drawing Sound"

With the advent of computers these explorations moved into the digital domain. Even if this might seem a well-known topic or even a trivial one, I think it is worthwhile to point out some reflections about this issue. As in many other fields of application, the shift to the digital *paradigm* has been not only a technological but also an ontological one. The digital machine is based on the concept of symbolic representation: in order to be processed, an input has to be *translated* into a sequence of numbers, and becomes *data*. Analogue signals (like voltage levels, sound) then become pieces of *information*. Although this process is, for some tasks, *transparent* to the user [4], as in conversion between analogue and digital signals, for some other musical activities it became a fingerprint, a marker that influenced the music creation process at its deepest level. Think, for instance, of the MIDI protocol, which is a message-based communication protocol with a discrete number of steps (128) which is still widely used today, despite its poor scale resolution. Another example is the grid-based logic of all DAWs and sequencers. The controls available to the performer in these environments are quantized, discrete and icon-based.

[4] In the software engineering sense

Music software developers have certainly realized that they had to deal with this issue. One of the solutions they found was to build a software emulation of an analogue counterpart of the system. One well-known example is the Max/MSP visual programming language. Here, the code is written by visually connecting different objects, thus resembling the manual patch-bay connections of an analogue synthesizer.

4.9.1 Xenakis and the UPIC System

Iannis Xenakis had the intuition of developing a graphic-based tool for composing since the 1950s, and maybe this intuition came from the ideas of Varése (Matossian, 2005). As noted by Peter Nelson (Nelson, 1997), the fact that Xenakis' compositional work embedded some graphic design can be related with his background in architecture and engineering. In his preliminary work, he used to compose graphic scores that had then to be translated into traditional notation, and seems like the UPIC system was conceived also as a way of getting rid of this passage and have a direct translation of graphics into sound. Nevertheless, Xenakis had to wait until 1977 in order to see the realization of a first prototype of the system. The device was called UPIC (Unité Polyagogique Informatique du CEMAMu). The core of the system was an electromagnetic drawing board for digitizing the graphic elements sketched by the user, while a video monitor showed relevant information about the point coordinates and other parameters. This allowed, if requested, an accurate control over the hand drawing gesture. The compositional workflow can be divided into two phases: a *microcompositional* and a *macrocompositional level* (Lohner, 1986). At the micro level, the user could draw waveforms and amplitude envelopes, combine them together and store them in the system memory. At the macro level, the canvas represented a pitch/time plane, where the user could draw lines, curves, shapes and then assign the previously composed sonic elements to these graphic elements. An interesting feature of the system was that the time axis did not have a fixed duration: this parameter could be set by the user. Therefore, a page could represent a piece of music of 0.2 seconds up to approximately 15 or 30 minutes, depending on the memory size. Moreover, the time axis could be *bended* or *warped*: in this way, the score could be scanned not only from left to right, but also in a non-linear way.

4.9.2 Other works: Music Sketcher, HighC, Iannix...

In more recent years there have been a number of projects inspired by the UPIC system. One is *HighC*, a software developed by Thomas Baudel, released in its first version in 2007 (*HighC, Draw your Music*). The idea was to create a software version of the UPIC that could run on a typical desktop machine, and it was intended to be an educational and entertaining tool as well as a instrument for professional composers. Although many features of the original UPIC system have been implemented in the software, the drawing tablet is - in most common cases - replaced

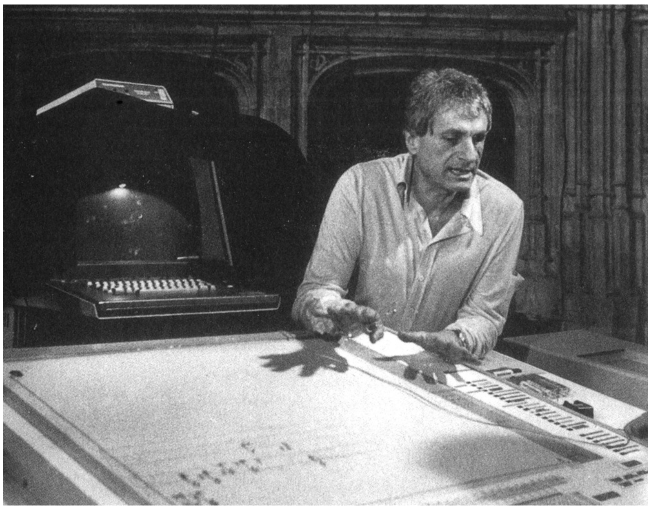


FIGURE 4.9: Yannis Xenakis with the UPIC system.

by the mouse and this results in a quite different kind of interaction. In fact, the drawing gesture was one of the crucial aspects in Xenakis' instrument.

Music Sketcher is a system developed by Jean-Baptiste Thiebaut et al. (Thiebaut, Healey, and Bryan-Kinns, 2008). An interesting feature of this software is that the two-dimensional space where the user can sketch curves and shapes is highly customizable: the x and y axes can be mapped to any couple of parameters in the sound synthesis. The issue of time representation have also been addressed in developing the instrument: time can be represented on one axis, but it can be also associated to the velocity of the hand-drawn stroke.

Another interesting system has been presented in 2006 by Mark Zadel and Gary Scavone. The software, designed as an instrument/interface for live performance, allows the user to sketch some free-hand curves that are then used to control sample-based sound generators. These lines represents trajectories for a set of moving cursors, or *particles*, and their motion controls the sound generation.

This idea is inspired by the work of Toshio Iwai (Iwai, 2001), who developed a musical software called*Music Insects* in 1991. This was an interactive sequencer where the notes, represented by colored pixels, were triggered by "insects" moving on the screen. A similar concept has been implemented in Iannix, a graphical sequencer that is a clear hommage to Xenakis' work. Here, hand-drawn curves in the three-dimensional space become paths for cursors whose motion can be programmed in different ways. A set of triggers can be defined along these paths, and an osc or midi message can be associated to each trigger. This concept of multiple cursors or *agents* scanning the score has been inspiring for CABOTO.

Golan Levin's (Levin, 2000) work focuses on an audio/video interface, which was intended to allow the user to express audiovisual ideas in "a free-form, non--diagrammatic context". The user could paint on a virtual canvas using an electronic drawing device. The resulting sketches were then interpreted by both a visual and an audio synthesizer. Though the sound output designed by Levin was intended to be more a kind of *sonification* than the output of an instrument or a composition, the author has discussed several interesting design issues, such as the representation of time and the quest for an intuitive but rich expressive interface.

Other projects have focused on more educational/entertaining tools. Two examples are *Hyperscore* (Farbood and Pasztor, 2004) and *PhonoPaper*, a software emulation of the ANS synthesizer developed by Aleksander Zolotov. PhonoPaper is available as an application for smartphone and tablet, and thus allows the user to scan a sonogram sketched or printed on paper by using the camera and synthesize the resulting sound in real time.

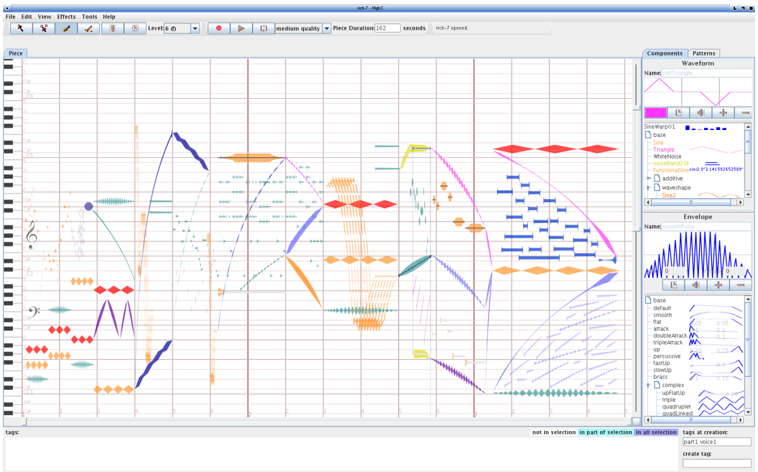


FIGURE 4.10: HighC: A software inspired by UPIC.

Tools for composing in traditional notation: MusInk, InkSplorer, pOM

Other authors have explored the possibilities of scanning hand-made sketches for composing in traditional notation, for example *MusInk* by Theophanis Tsandilas et al. Tsandilas, Letondal, and Mackay, 2009. This tool allows composers to sketch graphic gestures on paper and import this material into the OpenMusic environment, in order to create scores that combine traditional and graphic notation. Jérémie Garcia et al. proposed an evolution of this idea, *InkSplorer* (Garcia et al., 2011), in which curves sketched on paper are scanned by an Anoto pen (*Anoto digital pen*) and translated into traditional notation. The project *pOM* by the same author, created in collaboration with composer Philippe Leroux, presents a variation of this technique, in which the calligraphic strokes of a manuscript by Guillaume de Machaut are translated into polyphonic sequences (Figure 4.11).

4.9.3 Tangible Scores: Tomas, Adams

Another kind of projects deal with the concept of the score as an instrument or interface, as already discussed in 1.3. This concept of a score which is inherent in the instrument has been introduced with the use of electronic circuits in performing live music: Alvin Lucier, while referring to the works produced by the Sonic Arts Union,

FIGURE 4.11: pOM: a paper-based tool for translating hand-drawn shapes into traditional notation.

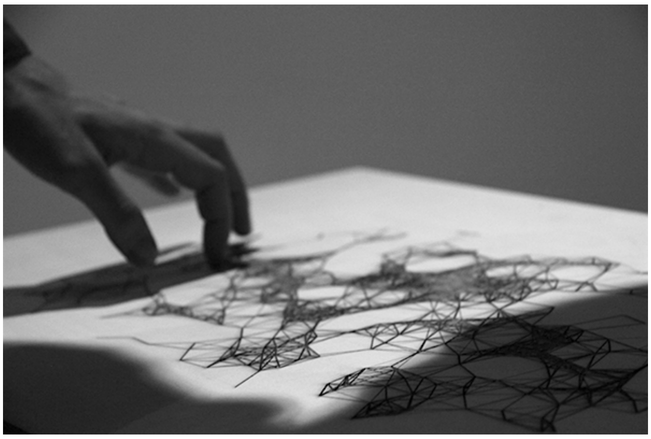


FIGURE 4.12: The score as an instrument/interface: Tangible Scores.

writes: "there were no scores to follow; the scores were inherent in the circuitry" (Lucier, 1998) (another example is *Rainforest IV* by David Tudor). In the computer music realm we have some well-known examples of devices that can be seen as instruments, interfaces or scores: for instance, a Max/MSP patch. An intriguing example is represented by the *tangible scores* created by Enrique Tomás and Martin Kaltenbrunnen (Tomás and Kaltenbrunner, 2014). Starting from the concept of *composed instruments* proposed by Schnell and Battier (Schnell and Battier, 2002), the authors developed the idea of a graphic score that can be played as an instrument, by touching it.

Chapter 5

CABOTO: A Graphic-Based Interactive System for Live Electronics

5.1 Description of the System

This chapter focuses on the detailed description of the main outcome of this Research, CABOTO. CABOTO is an instrument/interface designed for the live performance of electronic music. A graphic score sketched on a canvas is scanned by a computer vision system. The graphic elements are then recognized following a symbolic/raw hybrid approach, that is, they are interpreted by a symbolic classifier, but also as waveforms and optical raw signals. The graphic score is a map, traversed by an set of navigators or explorers. The performer tries to keep control over these agents, but they are designed to exhibit a semi-autonomous behavior. The graphic score and the scanners' paths are projected on the screen, creating a visual feedback of the ongoing sonic process. As previously stated, the instrument has evolved from an offline composition-oriented tool to a real-time scenario. This was a natural evolution, due to the fact that - as a musician and improviser - I felt the need to work with those feedback-loop dynamics which can originate only in the realm of live performance with an instrument (Maestri and Antoniadis, 2015) This instrument embeds different hardware and software modules:

- The Canvas, where the performer can sketch in real time using traditional drawing and painting tools.
- A Computer vision system, which scans the graphic sketch by means of a camera
- An image processing module, for image recognition, classification and features extraction
- The sound synthesis engine
- The video output for rendering the visual feedback

The different components of the system are described in details in the following paragraphs.

5.2 Design

The possibility of sketching the graphic score with pencils and brushes on a canvas in real time is one of the key aspects of CABOTO, and one of the main design constraints that I decided to incorporate from the very beginning of the project. Nowadays, many interfaces are available for digitizing hand drawing (e.g. Wacom tablets). Nevertheless, I decided to stick with paper-based drawing for two reasons: (a) the gestural control and the feeling that I can achieve with traditional painting tools still cannot be emulated in all its nuances by a digital system; (b) as previously stated in Chapter 1, I think that having some sort of analogue input injected into a digital system is an interesting way for *challenging* the system (and our expectations). For example when imperfections in the paper becomes glitches in the digital domain.

Another key aspect in designing the system has to do with playability and control. Since the sound output depends on the explorers' trajectories, and we have a number moving on the score at the same time, we need a way of controlling them on a macro-level. This means designing the system in a way that it can develop in time according to certain algorithms, but at the same time having the option of taking control and adjust the overall course of one or more explorers at the same time. I had then to select and/or define a limited number of (macro)parameters for the live control of the instrument through a MIDI controller.

5.3 Score Image acquisition and preprocessing

The first step in the image processing section of CABOTO requires the acquisition of a top view of the score through a live camera. The system is designed to grab the view in two modes: "one-shot" or "streaming". The one shot option is useful for keeping the CPU load within a certain limit, since the continuous video stream from a camera in real time is quite demanding. However, when the performer wants to interact with the canvas, it is possible to activate the streaming mode, giving the possibility of modifying the score while it is being played by the system.

The acquired image is preprocessed by a set of standard image processors. Brightness, contrast and saturation levels can be modified in real time, as well as the zooming factor. It is worth noting how these processing modules, introduced in the first place in order to correct the quality of the input image, have become an interesting extra feature of the instrument. In fact, the score can then be modified in real time just by modulating these parameters. For instance, an erosion effect on the scanned shapes can be obtained by simply adjusting the contrast and brightness levels, thus modifying the graphic score itself.

The zooming factor opened up another interesting way of exploring the score-map. As described in more detail in Chapter 4, the map concept serves as a powerful metaphor which suggests a certain way of approaching the graphic score. As stated before, one interesting feature of maps is that they define a *scale*. When we look at a map we need to know the scaling factor in order to figure out the actual size of the mapped territory. In more recent years, we have become accustomed to interactive online maps which give us the possibility to change the scaling factor, that is, to zoom our view in or out. In a score-map, changing the zooming factor results in a radical modification of the sonic result. For instance, a mass or cluster-like shape becomes a point if looked at from a distance, and vice versa. This gives the possibility of moving from a hyper-zoomed out view where the image is folded and multiplied as in a kaleidoscope, to a pixel-level view. The point of observation changes the score itself, and the way we interpret it.

5.4 Cropping the view and blob recognition

At this point, the image is converted to the greyspace (that is, a single channel is extracted) and the processing chain branches into four identical modules, each connected to a navigator. For each one, a view is cropped from the score, centered at the navigator [x,y] position. The size or scope of the cropped view can be set in real time by the performer. This operation is similar to applying a "magnifying glass" tool to the map and then results in an effect similar to the overall zooming described in the previous section. In the next step, a threshold is applied to the image, which then becomes a binary b/n matrix. A blob recognition algorithm is applied, that is, all the connected components in the cropped view are labeled and their boundaries (that is, the bounding box of the shape) are extracted 5.1. In this way the system recognizes the shapes and store them in a database for being used by the following classification algorithm.

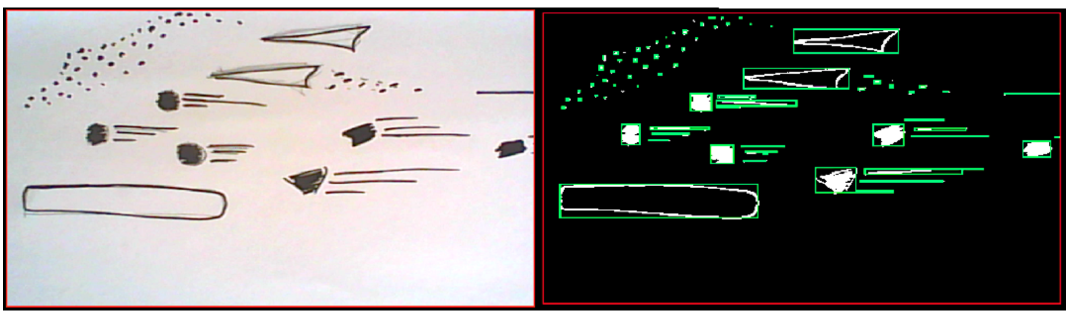


FIGURE 5.1: Blob recognition algorithm applied to a score.

5.5 The Symbolic Classifier

For each connected component, a set of geometric features are extracted:

- Centroid (x,y) position relative to the score
- Blob bounds size along x and y axis, δX and δY
- Shape ratio
- Filling
- Fatness
- Noisiness

The Shape ratio R is the ratio between the blob dimensions:

$$R = \frac{\Delta Y}{\Delta X} \tag{5.1}$$

This factor is useful for telling horizontal and vertical straight lines from mass-like shapes. The filling F is a measure of the total luminance with respect to the blob area, that is:

$$F = \frac{\sum_0^N val(x, y)}{\Delta X \cdot \Delta Y} \tag{5.2}$$

The filling can then be used to tell empty masses from filled or noise masses.

The *Fatness* FAT is a parameter that measures the average thickness of the shape along its main orientation, in order to tell curved lines from plane-like shapes. If we consider x as the main orientation axis, this results in:

$$FAT = \frac{\sum_0^N dy(n)}{N} \tag{5.3}$$

where:

$$dy_n = y_{max}(n) - y_{min}(n) \tag{5.4}$$

$y_{max}(n)$ being the highest edge of the shape, and N is the number of sampled values along the x dimension: this subsampling is introduced for optimization, since the computation is performed in real time and may result in a high CPU load. The *noisiness* of the blob is defined by the average number of zero-crossings of the first derivative along a set of paths that traverse the shape. Thus, a compact blob which is mostly filled or mostly empty will have a very low noisiness value, while a complex line will exhibit high noisiness:

$$NOI = \frac{\sum_0^N \frac{\delta val(x,y)}{\delta x}}{N} \tag{5.5}$$

According to these features and a set of thresholds, each blob is classified into 7 categories (Figure 3.1):

- point

- horizontal straight line
- vertical straight line
- complex line
- empty mass
- filled mass
- noise mass

.

The diagram of the classification algorithm is presented in Figure 5.2. It can be noted that this kind of classification algorithm is an untrained one, therefore it could be objected that is a quite naive kind of classifier. Nevertheless, this choice is deliberate. In a previous version, a more sophisticated classifier was developed, which made use of a trained pattern recognition algorithm. This led to an over-classification of shapes, which tend to become a sort of dictionary or a *taxonomy* of graphic elements. A symbolic mapping implies an interpretation. In this sense, classifying is a way of quantizing the collected data, and thus, in a certain sense, it is an operation which leads to a reduction of information and is not able to catch the richness of the hand-drawn sketch. This is the reason why the classifier has been kept quite simple and general.

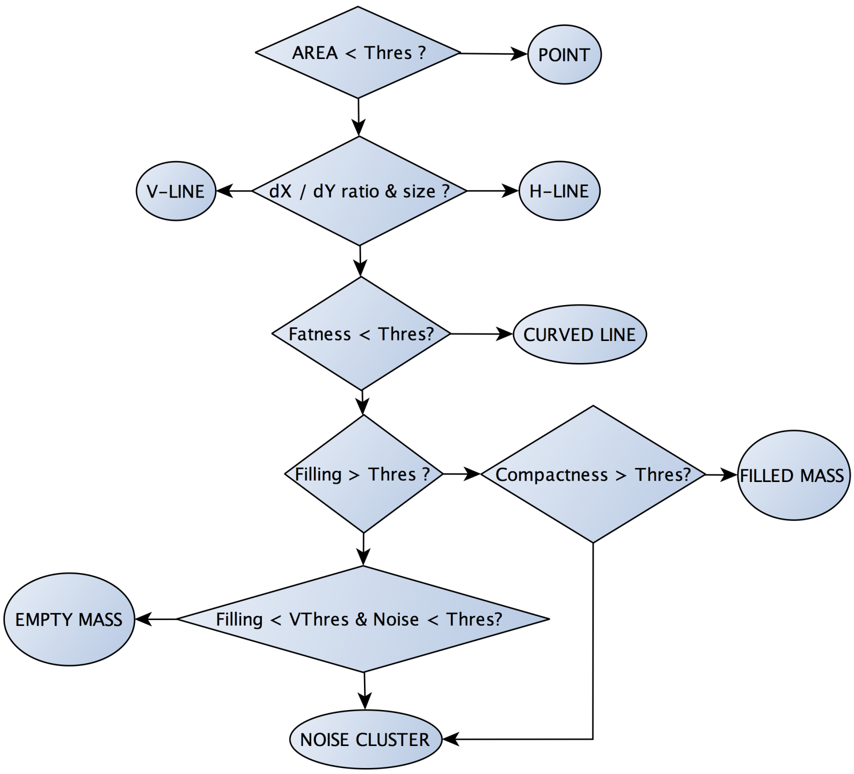


FIGURE 5.2: Diagram showing the classification procedure.

5.6 The Optical Scanner

The optical scanning of the image is inspired by the traditional optical sound techniques. As previously described in Chapter 4, there are different formats for sound recording on film. The one implemented here is the *Variable Density* method. According to this technique, the scanner looks for variations in the density of optical trace. In the digital domain, this is equivalent to look at the image matrix in the greyscale domain. An optical scanner is associated with each explorer traversing the score. The scanner crops a view in a chosen color channel and extract the overall mass (as a measure of the luminance):

$$OPT(x, y) = \sum_{Xmin}^{Xmax} \sum_{Ymin}^{Ymax} val(x, y) \tag{5.6}$$

The area covered by the scanner can be controlled in real time, thus varying the resolution and gain of the resulting signal. The output of the optical scanner is a raw signal, that is, it is not derived from some sort of interpretation according to a vocabulary, but from a scanning operation upon the values stored in the image matrix, and it brings richness and unpredictability to the sound synthesis. Moreover, since it depends strongly on the instantaneous position of the explorer, it has an immediate correlation with the visual feedback that can be seen on the visualized score. This allows the performer to have a high degree of control on the optical signal output. An interesting outcome of the optical scanning is that, since it can act on a pixel resolution level, it is highly affected by the imperfections of the hand drawn sketch and the canvas.

5.7 The Waveform Scanner

The other scanning mechanism implemented in CABOTO is called the *waveform scanner*, and it is also based on an optical sound technique called *modified unilateral variable area*. The observed blob is cropped and its edges are scanned along its main orientation axis. The optical signal is extracted as the distance between the outer edge of the shape and the median line with respect to the blob size (see Figure 5.3). Once the scanner reaches the bound of the blob (with respect of its main axis), it wraps around the shape and goes backward scanning the opposite edge. The output signal is sent to the synthesis engine as an audio stream, and can then be used as an envelope, modulator, control signal or directly as an audio signal. One of the main issues when dealing with such a technique is that the resulting signal suffers from the scanning periodicity, that is, at audio rate a fundamental frequency corresponding to the scan rate will clearly appear and dominate. In order to cope with this problem, the waveform scanner speed is controlled and modulated in real time. In particular, a stochastic scan rate modulator has been introduced. In this way, the

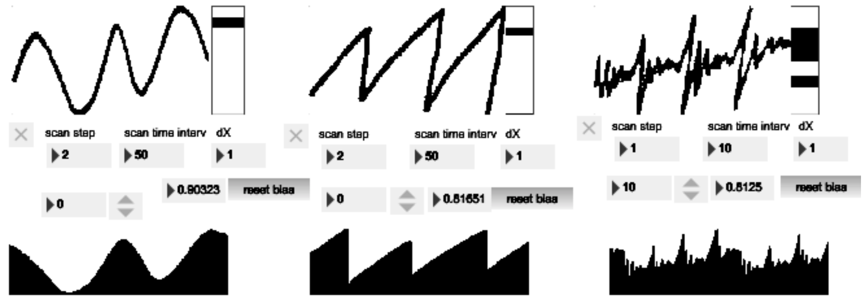


FIGURE 5.3: Scanning the waveforms.

performer can choose in real-time whether to set the scanning rate at a fixed value or to modulate it, thus moving from regularity to irregularity.

5.8 Mixing Different Approaches

The main original feature of CABOTO is that the three scanning mechanisms described in the previous sections have been combined together in the system. This synthesis can be described with the following procedure:

- The explorer crops a view out of the score-map
- The symbolic classifier looks for connected elements (blobs) in the view, and classify them according to the dictionary of basic elements.
- At the same time, the waveform scanner scans the contour of the graphic elements found in the view, and streams the resulting signal to the synthesis engine
- the optical scanner associated with the explorer scans the density value in an area centered at the explorer's position, and streams the resulting signal to the synthesis engine
- the symbolic classifier sends the blob(s) data to the synthesis engine. According to the detected shape, the engine starts a certain kind of process, and makes use of the wave and optical signals for modulating certain parameters and/or makes use of the waveform directly as a wavetable for generating the sound

5.9 Adjusting the Sails

The navigators' trajectories are generated in real time according to four different modes: *forced*, *random*, *jar of flies*, *loop*. The *forced* mode allows the performer to send a navigator to a certain position in the score, using a cursor on the score view interface (Fig. 5.6). In *random* mode, the navigators move autonomously, performing a random walk. A more interesting motion is defined by the *jar of flies* algorithm. This is a random walk in which the step increment is inversely proportional to the optical signal value detected at the current position. This means that a navigator will move slowly when it is in a densely populated area of the score (that is, with more elements), and faster when nothing is detected. This simple technique results in a sort of "organic" motion, which has some interesting effects on the development of the sound output. Finally, a *loop* mode is available, which generates a trajectory modulating the X and Y coordinates of the navigator with periodical signals. Since the rate and amplitude of these signals can be set independently for the two axis, it is then possible to have different kind of motions, from simple loops along one axis to more complex trajectories. Some of these modes can be mixed or superimposed. For example, a navigator can perform a random walk while looping in a certain interval across the X axis.

5.10 Sound Design

The main aesthetic guidelines in designing the sound have been presented in Chapter 3. Here, some details about the sound synthesis algorithms are given, introducing the concept of *sonic realization*. Each sound synthesis produced by the system is a possible realization of the detected class. Different realizations have been designed, and the engine will output different sonic results for the same given shape, according to a stochastic process. The same line will not sound the same, but the result will be consistent with the graphic/sonic vocabulary described in Chapter 3. This idea is inspired by the *Graphograms* experience (see Section 2.5): in that system, the graphic elements are interpreted by improvisers, who play a *realization* of that graphic hint. The *point*, *horizontal line* and all the *mass* classes are implemented with granular synthesis techniques, applied to both sampled material and audio input collected by the waveform scanner. A library of samples, classified according to the different items in the symbolic dictionary, has been created. The synthesis engine chooses randomly from this collection. For the *noise line*, the waveform signal is the main source for the synthesis, since it is been used both as an audio signal and as a modulator.

5.10.1 Mapping: The Polymorphic Parametric Space

For rendering the different shape classes, different processes and synthesizers have been designed, each one characterized by a set of control parameters. This results in a *polymorphic mapping*, that is, different mapping strategies for different kind of sonic events. For instance, the relative position of the navigator with respect to the sound object boundaries is mapped and used for the *noise cluster*, but is ignored in the case of the *point* class. Part of the mapping is presented in Figure 5.4.

	points (PBind)	h line (Synth)	v line (PBind)	noise mass 1 (PBind)	noise mass 2 (PBind)	noise line (Synth)
Ymin,max	grain rate	sin freq, max cutoff freq	wave rate / sin freq	max grain rate	max grain rate	wave scan rate
dY	-	freq rand deviaton range	impulse duration	rate range	rate range	rate range
Xmin, max	grain position, rate deviation	pulse duration	-	grain pos	grain pos max	
dX	-	pulse trigger freq	-		grain pos range	pulse duration
opti signal	-	cutoff freq	LP cutoff	gain envelope	gain envelope	LP cutoff
waveform	-	audio signal 2, mod signal	audio signal 1	mod signal	audio signal	audio/mod signal
Xe	-	-	-	grain pos	grain pos	
Ye	-	-	-		min grain	max freq
noise	-	-	pulse speed	grain pos mod speed	grain pos mod speed	grain pos mod speed

FIGURE 5.4: Parameters mapping for different classes of shapes.

5.11 Spatialization

The sound is projected in the performance space through a 4-channel audio system. For spatializing the sound, two different approaches have been explored. According to the first, each navigator is linked to a certain position in the acoustic space, that is, to a different channel. This is a quite effective choice both musically and from the practical point of view, since the explorers are treated as a sort of polyphonic choir distributed in the space of the performance. In this way, it is easier to keep track of what each explorer is actually producing. Another approach is to map different areas of the score to different positions in the virtual sound space. Although this idea is appealing and creates an interesting link between the visual feedback and the sonic result, it results in an unbalanced spatial distribution of the sonic events.

5.12 Implementation

The system is designed according to a modular logic, with different pieces of software integrated through Open Sound Control and a virtual audio driver. The image acquisition and processing and the navigators logic have been developed in the Max/MSP programming environment. For blob recognition and some feature extraction the cv.jit library developed by Jean-Marc Pelletier (*CV-JIT*) has been used, which provides an effective set of objects for computer vision. The image processing is quite CPU intensive; therefore, some routines have been written in C++, such as

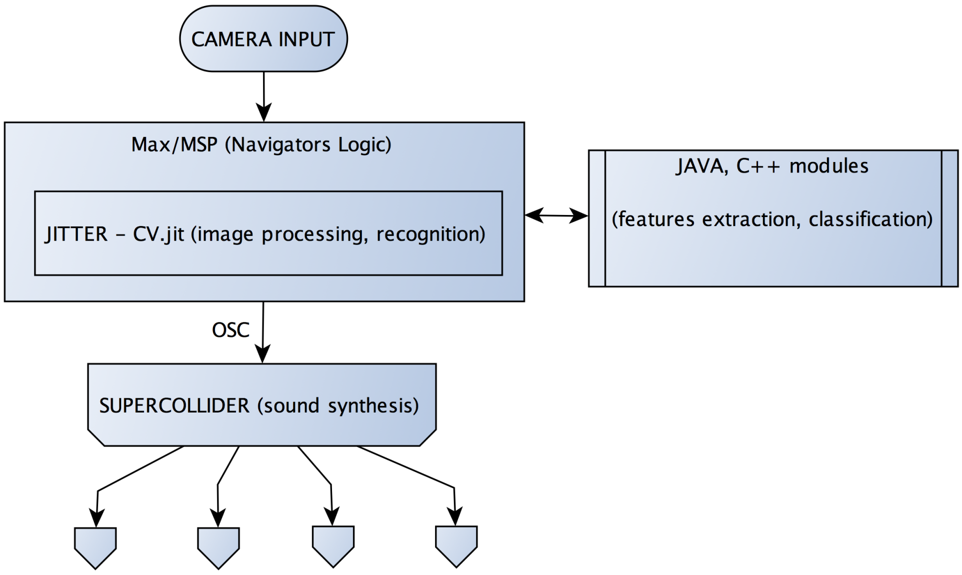


FIGURE 5.5: Implementation diagram of the CABOTO system.

some feature extraction, the symbolic classifier and the waveform scanner. The data from navigator position, features extraction and blob classification is formatted as OSC messages and sent to the synthesis engine. There are four independent streams of messages for the four navigators. A message is sent when a navigator finds a blob in its cropped view, or each time the view itself changes for any reason, for instance if the navigator changes its position relative to the blob, or if the score has been modified (by hand or in the digital preprocessing).

The signals from the waveform scanners are sent as audio streams through the virtual audio driver, and these signals are then acquired as audio inputs from the synthesizer. The sound synthesis engine was developed in the Supercollider language, which provides a powerful framework for generating complex sound events in the form of processes controlled by a set of macro-parameters. A general diagram of the whole software architecture is depicted in Figure 5.5.

5.13 Live Performance with CABOTO

As previously noted in Chapter 1, The system was originally conceived as a tool for composing. However, the project evolved towards the design of an instrument for live performance (Figure 5.7). This evolution is connected to the fact that, as a musician and improviser, I felt the need for a system for live performance and improvisation. During the performance it is possible to sketch or modify the score: in order to avoid the hand interference, the image can be grabbed with a one-shot button, once the drawing gesture has been completed. Another option is to disable the video streaming according to a motion detection algorithm. Nevertheless, I found it more interesting to keep the video streaming on and let the drawing action interfere with the score scanning, thus resulting in glitches, noise and unexpected sonic

FIGURE 5.6: The CABOTO consolle.

output. In designing the live setup, some decisions had to be made regarding the parameters to be controlled. Since I am dealing with multiple navigators as well as the drawing action, I decided to retain control over a few macro-parameters, such as the output gain of each navigator (which also enables/disables the navigator itself), the trajectories generation mode and speed, and the score image settings (brightness, contrast, saturation, zoom, blob recognition thresholds). A video documentation of a live performance with the instrument can be found in *CABOTO - Live at KonCon, March 21 2018*.

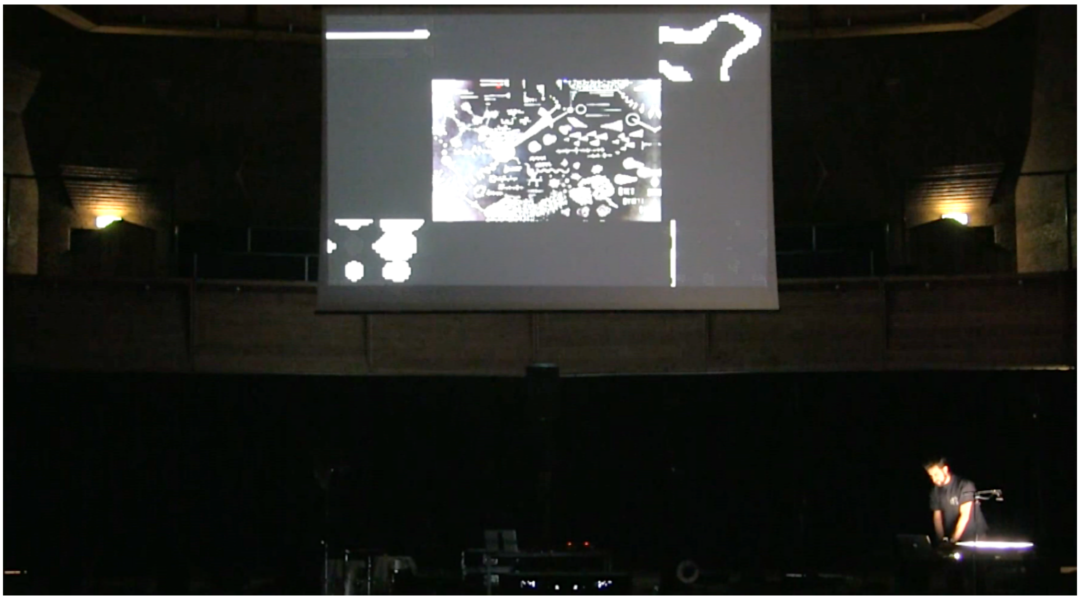


FIGURE 5.7: Live performance with CABOTO.

Chapter 6

Conclusions and Future Developments

6.1 Conclusions

This dissertation has followed two main threads. The first one deals with the development of a graphic notation for composing electronic music, and how this vocabulary is related to certain aesthetic choices and how it leads to certain musical results. The second thread is related to the development of the system for scanning the graphic material, and the mapping strategies that have been adopted. These two threads are strongly interrelated. In programming the software tool, I had to take into account the kind of graphic material I wanted to deal with. On the other side, in designing the graphic vocabulary, I had to consider the constraints imposed by the machine. Formalizing in a visual form some concepts and ideas that have been implicit in my creative process led to an interesting side effect: a new level of awareness about my musical aesthetics.

There are other two parallel tracks in this research, which can be related to the previous ones. These can be defined by the *symbolic/raw* dichotomy. As explained in detail in Chapter 5, the term *symbolic* refers to a kind of graphic notation that requires a certain degree of interpretation, which can be achieved only through the use of a dictionary of basic elements. The symbolic approach can typically be found in language. The *raw* approach, on the other side, is derived from the audiovisual technologies of the early twentieth century described in Chapter 4. Here, the graphic element is viewed as a pure optical signal which is translated into a voltage signal and then into sound. We could object that this is an interpretation as well. The fact that we choose to read a certain signal *as a sound* is *per se* an arbitrary choice, an interpretation. Nevertheless, we can argue that this is a very low-level kind of interpretation of a graphic element.

In developing CABOTO I have been looking for a synthesis of these two approaches. A symbolic classifier has been developed, as well as a waveform scanner and an optical scanner. These three devices are used in combination for synthesizing the sonic material. This results in an instrument that can provide a certain degree of control, since I know in advance how certain graphic elements will be interpreted, but at the same time it can provide some rich and unexpected sonic material, thanks to the raw scanners.

Along this process of synthesis of the different approaches, another crucial issue has been explored: the one of time representation (see 3.2). This led to an evolution of the project from the initial idea of a computer-aided composition tool to the concept of the score-map and the score as an instrument.

6.2 Future Developments

CABOTO must be considered a work in progress and many improvements are currently under development. The sound synthesis algorithms are constantly being updated, and new sonic material is being added to the palette. Further explorations will focus on developing the visual feedback which is presented to the audience during the live performance. Another important aspect is optimization: the image processing algorithms are quite CPU demanding, thus some code optimization is still required in order to achieve a better quality for the visual rendering.

CABOTO has been presented in some live performances, both in solo and in collaboration with an ensemble of improvisers. This collaborative scenario with musicians is still under development. In this scenario, the musicians play along with CABOTO, while reading from the same graphic score that is scanned by the system. A set of instructions is given to the musicians in advance on how to interpret the graphic vocabulary. During the performance, the group of improvisers is divided into four sub-groups, and a different cropped view of the score is presented to each sub-group. Each cropped view is linked to the trajectory of one of the explorers, thus becoming a sort of live notation system.

Nevertheless, the main focus at the moment is on developing the live performance with the system. Like every other instrument, CABOTO has to be explored and practiced in order to achieve a certain degree of expertise. Through this process, it is also possible to understand the pros and cons of certain design choices, and try to explore other solutions, if necessary. As already mentioned in the introductory chapter, in practicing with the instrument the gestural part of the performance is coming more and more into focus. In a continuous dialectical confrontation (and sometimes a clash) between the machine and the human gesture. My explorations are now pointing in this direction, after having finally realized that the quality of the performance, its strength and intention are all connected to the gestural power of drawing.

Appendix A

Additional Material

A download link to additional material is provided, which contains:

- A video of a performance with CABOTO, recorded live at the Sonology Discussion Concert, Arnold Schoenbergzaal, March 21st 2018.
- A folder containing the CABOTO code, which includes: a Max/MSP patch, a Supercollider script, two Max/MSP externals written in C/C++.
- A collection of selected scores composed for the system
- A pdf version of the article: *"CABOTO: A Graphic-Based Interactive System for Composing and Performing Electronic Music"*, published in *Proceedings of the 2018 NIME Conference*, June 3-6 2018, Blacksburg, VA.

Bibliography

Anestis Logothetis Music Portal. accessed on November, 29th 2017. URL: http://anestislogothetis.musicportal.gr.

ANOTO. *Anoto digital pen*. accessed on May, 22th 2018. URL: http://www.anoto.com/creative/technology/the-digital-pen.

Athanasopoulos, George, Siu-Lan Tan, and Nikki Moran (2016). "Influence of literacy on representation of time in musical stimuli: An exploratory cross-cultural study in the UK, Japan, and Papua New Guinea". In: *Psychology of Music* 44.5, pp. 1126–1144.

Baudel, Thomas. *HighC, Draw your Music*. accessed on November, 29th 2017. URL: https://highc.org.

Berger, J. and E. Neizvestny (1969). *Art and Revolution*. Pantheon Books.

Boytwell, Brett (2012). "Morton Feldman's Graphic Notation: Projections and Trajectories". In: *Journal of the Society for American Music* 6.

Braxton, A. (1988). *Composition notes*. Composition Notes 1. Synthesis Music.

Cambouropoulos, E. (2015). "The Shape of Music - The Music of Shape". In: *Proceedings of the Third Interdisciplinary Workshop SHAPES 3.0 - The Shape of Things (SHAPES 2015), Larnaca, Cyprus, November 2, 2015*, pp. 19–21.

Cardew, C. (1974). *Stockhausen serves imperialism, and other articles: with commentary and notes*. Latimer New Dimensions.

Coulmas, F. (1999). *The Blackwell Encyclopedia of Writing Systems*. Ill., graph. Darst. Wiley.

Crystal, David (1995). "Phonaesthetically speaking". In: *English today* 11.2, pp. 8–12.

Dearen, J. (March 28, 2008). "Newly Found Clip Might Be World's Oldest Recording". In: *The Washington Post*.

Duchez, Marie-Elisabeth (1989). "An historical and epistemological approach to the musical notion of "form-bearing" element". In: *Contemporary Music Review* 4.1, pp. 199–212.

Farbood, Morwaread and Egon Pasztor (2004). "Hyperscore: a graphical sketchpad for novice composers". In: *IEEE Computer Graphics and Applications* 24.1, pp. 50–54.

Garcia, Jérémie et al. (2011). "InkSplorer: Exploring Musical Ideas on Paper and Computer". In: *New Interfaces for Musical Expression (NIME 2011)*. Oslo, Norway.

Gitelman, Lisa (1999). *Scripts, Grooves, and Writing Machines: Representing Technology in the Edison Era*. Stanford University Press.

Goodiepal (2009). *Radical Computer Music & Fantastisk Mediemanipulation: A Corrected and Illustrated Transcript of The Official Mort Aux Vaches Ekstra Extra Walkthrough*. Pork Salad Press.

Goodman, N. (1968). *Languages of Art: An Approach to a Theory of Symbols*. Bobbs-Merrill.

Guy, Barry (2012). "Graphic Scores". In: *Point of Departure, Online journal* 38. URL: http://www.pointofdeparture.org/PoD38/PoD38Guy.html.

Hinton, Leanne, Johanna Nichols, and John J Ohala (2006). *Sound symbolism*. Cambridge University Press.

Hutton, Jo (2003). "Daphne Oram: innovator, writer and composer". In: *Organised Sound* 8.1, pp. 49–56.

Iwai, Toshio (2001). "Piano–As Image Media". In: *Leonardo* 34, pp. 183–183.

Kandinsky, Wassily (1947). *Point and Line to Plane*. Dover Publications.

Kaufmann, W. (1967). *Musical notations of the Orient: notational systems of Continental, East, South and Central Asia*. Indiana University humanities series. Indiana University Press.

Köhler, Wolfgang (1929). *Gestalt psychology*. H. Liverights, New York.

Levin, G. (2000). *"Painterly Interfaces for Audiovisual Performance"*. M.S. Thesis, MIT Media Laboratory.

Levin, T. (2003). "Tones from out of Nowhere: Rudolf Pfenninger and the Archaeology of Synthetic Sound". In: *Grey Room* 12, pp. 32–79.

Lewis, G.E. (2008). *A Power Stronger Than Itself: The AACM and American Experimental Music*. University of Chicago Press.

Logothetis, Anestis (1973). "Über die Darstellung des Klanges im Schriftbild". In: *Impluse: für Spielmusikgruppen, p. 3-9*.

Lohner, Henning (1986). "The UPIC System: A Users Report". In: *Computer Music Journal* 10.4, p. 42.

Lucier, Alvin (1998). "Origins of a Form: Acoustical Exploration, Science and Incessancy". In: *Leonardo Music Journal* 8, pp. 5–11.

Maestri, Eric and Pavlos Antoniadis (2015). "Notation as instrument: from representation to enaction". In: *First International Conference on Technologies for Music Notation and Representation TENOR 2015*. Proceedings First International Conference on Technologies for Music Notation and Representation TENOR 2015. IRCAM - IReMus. Paris, France.

Magnus, David (2015). "Polymorphy and Adequacy. The Role of Perception in the Pictorial Notation of Anestis Logothetis". In: *Georgaki, A. Topología*.

Marogna, Riccardo. *CABOTO - Live at KonCon, March 21 2018*. accessed on April, 28th 2018. URL: http://riccardomarogna.com/caboto.

— *Graphograms*. Accessed on November, 29th 2017. URL: http://riccardomarogna.com/graphograms.

— *KarlHeinz Score Scanner (2014)*. A graphic-based system for composing electronic music. Accessed on April, 28th 2018. URL: http://riccardomarogna.com/karl.

Matossian, N. (2005). *Xenakis*. Moufflon Publications.

Miller, Daniel (2017). "Are scores maps? A cartographic response to Goodman". In: *Proceedings of the International Conference on Technologies for Music Notation and Representation - TENOR2017*. A Coruña, Spain: Universidade da Coruña, pp. 57–67.

Moholy-Nagy, L. (1964). *The New Vision, 1928*. Documents of modern art: director, Robert Motherwells. Wittenborn.

Nelson, Peter (1997). "The UPIC system as an instrument of learning". In: *Organised Sound* 2.1, 35–42.

Nelson, Steven G. (2008). "Court and religious music (2): music of gagaku and shōmyō". In: *The Ashgate Research Companion to Japanese Music*. Routledge. Chap. 3.

Oktopus Connection. Accessed on April, 29th 2018. URL: http://www.setoladimaiale.net/catalogue/view/SM2680.

Pelletier, Jean-Marc. *CV-JIT*. Computer Vision Library for Jitter. Accessed on December, 11th 2017. URL: http://jmpelletier.com/cvjit.

Radano, R.M. (2009). *New Musical Figurations: Anthony Braxton's Cultural Critique*. University of Chicago Press.

Ramachandran, V S and Edward M Hubbard (2001). "Synaesthesia — A window into perception, thought and language". In: *Journal of Consciousness Studies* 8.12, pp. 3–34.

Schnell, Norbert and Marc Battier (2002). "Introducing Composed Instruments, Technical and Musicological Implications". In: *Proceedings of the 2002 Conference on New Interfaces for Musical Expression*. NIME '02. Dublin, Ireland, pp. 1–5.

Smirnov, Andrey (2013). *Sound in Z experiments in sound and electronic music in early 20th century Russia*. König.

Sterne, J. (2003). *The Audible Past: Cultural Origins of Sound Reproduction*. Duke University Press.

Streitz, N. and P. Nixon (2005). "The Disappearing Computer". In: *Communications - ACM* 48.3, pp. 32–35.

Taylor, R. and I. Christie (1994). *The Film Factory: Russian and Soviet Cinema in Documents*. Soviet cinema. Routledge.

Thiebaut, Jean-Baptiste, Patrick GT Healey, and Nick Bryan-Kinns (2008). "Drawing Electroacoustic Music." In: *ICMC*.

Tomás, Enrique and Martin Kaltenbrunner (2014). "Tangible Scores: Shaping the Inherent Instrument Score". In: *Proceedings of the International Conference on New Interfaces for Musical Expression*. London, United Kingdom: Goldsmiths, University of London, pp. 609–614.

Tsandilas, Theophanis, Catherine Letondal, and Wendy E Mackay (2009). “Mus ink: composing music through augmented drawing”. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pp. 819–828.

Varela, F.J., E. Rosch, and E. Thompson (1992). *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press. MIT Press.

Wishart, Trevor and Simon Emmerson (1996). *On sonic art*. Vol. 12. Psychology Press.

Xenakis I., Gibson B. (1994). *Kéleütha: Écrits*. Paris: L’Arche.