REVIEWING THE TRANSFORMATION OF SOUND TO IMAGE IN NEW COMPUTER MUSIC SOFTWARE

Esther Lemi, Anastasia Georgaki

National and Kapodistrian University/ Music Department, Athens, Greece, ester@mail.gr

National and Kapodistrian University/ Music Department, Athens, Greece, georgaki@music.uoa.gr

ABSTRACT

In the following essay we are going to analyse the relationship between sound and image in computer music. We will be examining sound visualisation software, and its evolution over the thirty-five year timespan in which it has existed. How we judge software, is based on aesthetic criteria, the way they were handed down to us from theories of abstract painting (20th century avant-garde) , the theory of montage by Sergei Eisenstein, of neurophysiology (synesthesia, muscular images) and of the successful correspondence of the two media (pixel and music) in the works and theory of James Whitney.

I. INTRO

Since there is always the need to write down the music, the relationship between image and sound is a very important one. An approach of this kind is crucial for composers. Xenakis justifies the composers need to have a first draft by hand using the poetic, but also logical, saying that "the hand is the organ of the body that is closest to the brain" ([1]).

The history of modernism begins with the spiritual relationship between two great men, Kandinsky and Schoenberg. During their long friendship they certainly managed to influence each other, breaking the boundaries of representative art and tonic music respectively, each creating for their art the new era of abstraction. On a deeper level they managed a new approach of art as a constant and alive centre of creativity. Many artists and movements followed. These consisted of musical and visual principles (Duchamp, Mondrian, Matiuschin, Hauer, Fluxus etc), which aided the evolution of art. This evolution led to new art forms (video art, interactive performance, visual music etc).

The booming of the computer age helped the transformation of the relationship between the image as a decoder of the musical language into the language of visual translation with the help of symbolic representations (from the theory of Paul Klee in Bauhaus to the aleatoric scores). The involvement of computer in the production of the common resulting image-sound gave birth to a completely new view that could not have existed up to that point ([2]).

II. FROM NOTATION TO VISUAL MUSIC

The first connection of the relationship between colour and music can be found in the text by Aristotelis 'On Sense and the Sensible' [3]. In this text there is a categorisation of colours according to the ratio between consonance and dissonance, in ratios representative of the harmonic music system of Pythagoras, even though that term will not be invented until the 19th century, Aristototelis talks about the synesthesthetic point of view in art.

On the other hand there are formal methods of approaching a graphic conception of an encoded musical structure. In the generalized typical form (stave), as it was given to us by Guido d' Arezzo, the important information that comes up from it, is the abscissa relationship of the modification of the pitch (vertical axis) to the evolution of time (horizontal axis).

The movement of aleatorism gave a freer description of music composition by using abstract forms, through which the macrostructure of the piece remains clear and (because of the abstract flexible handling) the microstructure of the piece could be controlled with greater detail ([4]).

After the 2nd World War the evolution of musical writing became revolutionary in as far as its focus and analysis options were concerned. So, we can separate the options of representation into the following general categories:

Graphic Representation of:

- the external/ internal structure
- the harmonic and melodic structure (pitch)
- rhythmic structure (time)
- timbral structure
- motive (cellular reconstruction)

The criteria with which these representations were created were always in the context of the synesthetical relationship between two arts, using as a criteria of that, decoding aural symbols, literal and formal congruences (onomatopoeia etc) and kinaesthetic stimuli ([5]). For example, a modern music score can provide information of time, technique, expression; pitch and volume in a visually structured form which will help the performer understand and perform the piece with the accuracy that the composer wishes.
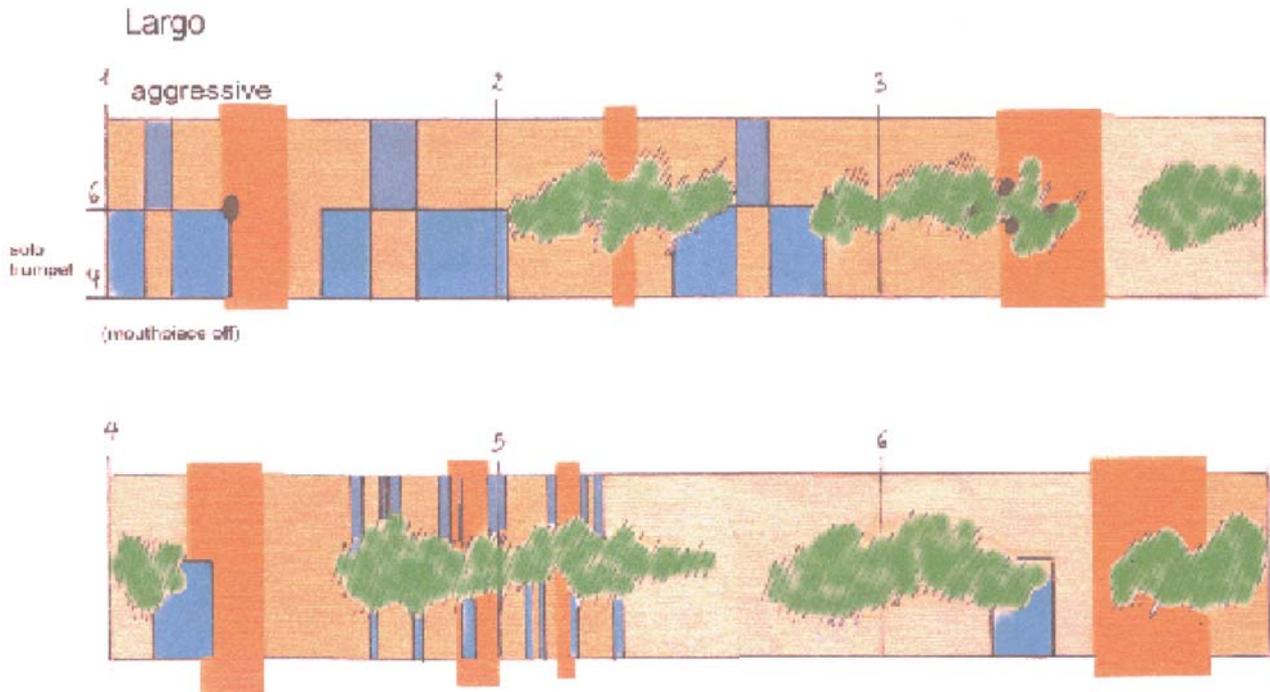
Fig 1 : Fragment from the score "Stamina" Esther Lemi 2007

For example, in the score of Esther Lemi's piece for a trumpet solo "Stamina" (2007) the performer produces sounds of indefinite pitch that cannot be shown with classical notation. In order to present varius timbral mutations on the time domain, symbolic forms and colours were used, which represent different timbers, meanwhile the nuances of luminance in the orange colour define the dynamic evolution between piano and forzatissimo (Fig.1)

A discretionary synesthetical approach provides a vague dimension with neurophysiological and psychological parameters that cannot be perceived in an objective way ([6]). For example, for the time being we can represent a sound as "yellow" as part of a serious correspondence. A basis of this type is vague and doubtful and the correspondence lies within a deeper coding.

A more significant method of correspondence between image and music came with the growth of computer music, particularly the research of James Whitney on the construction of a harmonic relationship, based on the tone as the basic component of music, and the pixel as the basic component of image. During the seventies he founded the term «harmonic pixel phenomena» in his book "Harmony-On the Complementarity of Music and Visual Art". In one of the chapters he includes average programs in Pascal language, which create "differential dynamics", a family of algorithms that activated each pixel point of a cluster differentially. The plasma-like liquidity of such motions permits aggregate architectonic structures to match musical action. James Whitney also gave us the term «differential digital harmony» to express the idea of evolutions of ratios into visual and sound models of harmony ([2]).

The efforts of the artists to combine the two arts created many hybrid forms of art, which consisted of a dynamic further evolution. In this way a form of art became possible, which allows, by the rhythmic change of images, even with the absence of sound, a kind of visual composition. This art was named visual music. Applied on the technical and aesthetic theories at the beginning of the twentieth century, visual composition combines the meanings of consonance and dissonance, as a common language for music and image, because the phrases are built through repetition, contrast and variation of the many visual dimensions including direction, speed, shape, size and colour. ([7])

This is the result of using a theoretical basis of montage and the needed use of a rhythmic continuity as well as the symbolic approaches for time and rhythm in painting and experimental cinematography. The separation of the techniques of the following images, depending on their rhythmic behaviour was given by Sergei Eisenstein who believed that the process of montage is the most important part of the artistic value of film. For him, montage was a multi-tiered construct of tension and release. This is a term that can be also given to the art of music. Furthermore, he listed five basic types of montage, which are applied today in the theory of visual music. These types are: metric, rhythmic, tonal, overtonal and intellectual montage. ([8]) Basically, he values the rhythmic ratios and the meaning of rhythm, as one of the basic elements of narration.

III. A TAXONOMY OF COMPUTER VISUAL MUSIC SOFTWARE

Iannis Xenakis started, in 1972, constructing a music system for computers which specialized in the creation of sounds through graphs. The visual palette of Upic (Unite Polygogique Informatique de CEMAMu) consists of lines, curves and points. Many examples of programs of visualizing sound followed. Most of them are being presented as simplified programs of composition. In the following paragraph we refer only to programs with two dimensional presentation in axis x, y

The programs can be separated according to the visual production into:
• programs with graphical representation
• programs with colour representation

And according to their function, into:
• programs of reproducing sound into image
• programs of reproducing image into sound
• programs of reproducing simultaneous production of image and sound

A general diagram with characteristic samples from programs can be presented as follows:

Programs Name Date of Presentation Inventor	Score	Visual Presentation Process	Pros
Upic 1972 Xenakis CEMAM	The Upic score is a collection of "arcs". An arc is a pitch (frequency) versus time curve.	Black and White From image to sound	Ultimate guidance of the visual process by ear. Detailed macrocompositional/ microcompositional structure analysis.
Phonogramme 1993 Lesbros GAIV	Resembles two-dimensional frequency-domain representations of sound by physical drawing tools	Black and White From image to sound	Option of physical drawing tools (e.g. ink, pencil, watercolour etc) to create different sound results.
Metasynth 1998 Wenger	Translation from static images (PICT files) to sound. A picture is scanned from left to right. Pitch and duration are represented on the vertical and horizontal axis respectively	Colour From image to sound	Uses pictures as filters (by scanning the colors) to produce the desired sound.
Hyperscore 2004 Farbood MIT Lab	The user chooses a colour for each motive and composes a piece versus time curve	Colour From image to sound	Motive presentation of tonic music (Minor and Major Structure).
Soundpaint 2005 Wenger and Spiegel	Definition of sounds according to RGB colours (different colours map to different sounds in order to utilize the colorspace as much as possible)	Colour From image to sound	From **colour** to sound.
IFS music 1989 Bransley	Production of different types of fractals that can define pitch and time	Colour vice-versa	Pieces of modern aesthetic can be produced.
JITER 2003 Cycling 74	Jitter abstracts all data as multidimensional matrices, so objects that process images can also process audio volumetric data.	Colour vice-versa	Further evaluation of three-dimentional and real-time presentations.
SONOS 2005 Thiebaut	It shares the same data between the domains of sound and video in a relationship that makes the biggest possible sense. The abscissa represents frequency and time.	Colour vice versa	Rotation, zoom, blur and saturation from Jitter environment. Due to the corporation with such a complex program as max/msp/jitter, its applications are infinite.

As we can see on the table above, according the categorization, we observe characteristic examples of translation image-sound. The method of categorization here is according to the procedure and the aesthetic result.

In the first category of graphical production we have Upic, in which we are given the opportunity to draw waves and envelopes, to compose a page and to do the mixing. Upic, as a pioneer in the field of visualization changed the creative habits and methods with a simultaneous process of composing in a macro- and micro compositional level through the interaction between man (via hands that design the music) and the machine at the command receiving of the system ([9]).

Phonogramme, a simpler software composition tool, has a visual palette which spreads through all the shades between white and black (silence and maximum volume respectively), but despite the good ideas it presents, this system is not based on empirical studies and does not support design strategies ([10])

In the second category, that of colour presentation, the palette of symbolisms grows, and there is a possibility of choosing further parameters. The chromatic "desymbolism" of sound is an idea that was successfully applied for the first time during the eighteenth century by the French Jesuit Monk Louis Bernard Castel. The monk constructed a visual musical instrument of performing sounds with a parallel showing of colours, and he inspired the later generation. Three centuries later, we still have not found a logical image-sound correspondence, so the correspondences are always used in combination according to what they serve and present.

The first example of production sound from a picture is the one of Metasynth, which provides a range of red- green-yellow colours, which work as filters. As in Phonogramme, but in the opposite direction, the grey shades symbolize the differences of volume (white for the maximum volume and black for silence)([10]).

Constructed for educational purposes, Hyperscore's layout is made in a more playful style. The melodies are composed independently in the motive window and then are presented the way the user wants in the sketch window. So we have a motivic structure. Each motive has its own colour. In a third window, the harmonic line of the piece is defined. The construction follows a simple setting of the motivic elements of the composition. However, the simplistic visual setting sure enables children to comprehend the procedure of a classical form of a composition. Moreover, there is a possibility of copying the visual score in a midi file and its appearance as a stave in one of the special programs (Finale,Sibelius etc) ([11]). Such a play -like procedure (from the abstract to the specific) pleases the imagination and amuses, but its creative possibilities are limited.

Soundpaint is a program that maps from colour to sound by introducing a vector space homomorphism between visual space (in RGB palette) and sound space. The constraints which are used for the process are injectivity, surjectivity, topology preservation and user definable mapping. The program has potential, because its table is regarded in a way that implies its flexible behaviour of movement during a process of painting ([12]).

In the case of fractals, we examine the technique of musical composition through computer that counts the repetitive functions from codes of chaotic procedures ([13]). Through this, the descriptive system of these functions can form the fractals which can be rendered as music scores. Fractal interpolation waveforms are deterministic, so the same melody will be generated repeatedly unless the parameters of the waveform are changed in pitch, dynamic level, behaviour and into a new musical form ([14]). The great interest of a non-linear dynamic system for the use of musical composition is a natural relationship, in behaviour or phenomena, to the real world, which unfolds the mechanical affinity of controlling and contemplating. The chaotic procedures consist of a process of modification. Its internal constancy is verified by the rules that lie coded into its equation ([15]).

Jitter provides a vast field of application, which allows the visual parameters to function as aspirations to the dynamic and non-objective qualities of music ([16]). This is the first time that we speak of the possibility of producing visual music, because the complexity of the program provides a free choice of the circumstances of mapping from a simple sketch to a graph table and the application of video. The positive element of this freedom, as far as the procedure is concerned, but also from the procedure of simultaneous flexible applications in image and sound, produces results of a common relationship image- sound ([17]).

Sonos is a program that complies with Max/MSP/ Jitter, which controls the sound parameters using the abilities of visualization of sound as its own procedure. An important function is graphical transformation using a transfer matrix as control interface. That happens by connecting each plane of a matrix to a transformation, while each pixel simultaneously stores a value. The procedure is activated when the user colours a matrix. Here the abscissa represents frequency in time. A controller, such as a keyboard mouse, or any joystick, may then allow the exploration of a sensitive variation of sound and image, giving it interactive form ([18]).

The general problem of all these programs is that beyond the scientific approach of the representation, such as waveforms or spectrograms, there is no objective representation of sound phenomenon by the computer visualization. There are only subjective, more or less metaphorical representations, in heterogeneous sound or musical context. The subjective relationship between visual and sound results are due to the absence of objective categorization of the colours. According to the neurobiological studies of Margaret Livingstone ([19]), on the process of colour relationships between the eye and the brain, the greater part of what has been written about colors is arbitrary. According to the research, the brain receives the values of shade and luminosity separately. Also, the depth of the field and the separation of the form from the surface is achieved with the absence of colour,

while the crises in the colour range depend on genetic and empirical elements. The generalization that we can make concerns the three-dimensional construction of the colour space without a relation to the kind of colour range (RGB, RYB, CMY). The results of the research show that there is no hierarchy of the color range beyond the value of luminosity and its saturation.

The programming of systems of sound visualization balances, in movement, those subjective-objective levels of hierarchies. Even if most of the criteria are judged by physiological arbitrary approaches, our connection with the systems create a general hierarchy that we can accept as is (e.g. who can throw away Kandinsky's theory on which modernism was based?)

IV. VISUAL COMPUTER MUSIC

The new thing that computer music brought into the world was the way in which music was processed with the focus on sound. The process of sound using mathematical models helped us enter its core. The ability to deconstruct the sound into its basic components, and to alter sensitive parameters shows a new behaviour of easily modifying the material itself, providing us with one more aesthetic advantage of the immediate and precise impression of the sound result and its sculpting reformation, even though the opportunity for a common vocabulary to be built has not been given. Besides that, in all of the programmes there is always a hidden intent of touching a broader public and of discovering a new synesthetic behaviour and a new musical form.

From the practical applications point of view, the programmes of visualization are separated into the categories:

- of musical production and presentation as two different processes, and
- the parallel use of the parameters with the synchronous presentation of the results.

These two different techniques (acousmatic and interactive), because of the use of the computer, are used for the production of more complex applications that without the power of the machine, could not be produced [(4)]. In the computer music repertoire we meet inspirational works in some visualization programs, while successful creation is usually due to its ability to compose with its talent.

Xenakis' Mycenae Alpha, is the first work entirely realized on the UPIC (1979). Others followed, composers such as Julio Estrada with eua'on (1980), but the big bloom came during the nineties with pieces such as "Saxatile" by Jean-Clause Risset's and "L'Autel de la Perte et de la Transformation" ("the Altar of Loss and Transformation") by Brigitte Robindore (1993).

Beyond the researchers and the modern composers, pop artists worked with many of the programs. One of the most known examples is the case of Aphex Twin who hid an image

of himself in a spectrogram using Metasynth, an image of a spiral shape in his first track from "Windowlicker".

On the other hand, the programs such as Hyperscore,are used mainly for educational purposes. The pieces that have been written are cute, with a respective optic and musical lightness such as "Creepy Raindrops" by Chelsea O'Hara (2002) and "My Very Happy Hyperscore" by Garry Hughes (2002).

Nowadays, when the technological revolution allows more delicate manipulation in sound and image, it is now possible to combine more complex techniques that bring a more fundamental result.

Even though we have examples of pieces of visualization by some software, the flexible new forms of programming allows us to include, in the prehistory of the modern bases, visionary steps that have been applied by artists that worked with sound and experimental cinema. Norman MacLaren, in his film "Synchrony" (1971) composed music and drew directly on the optical soundtrack of 35-mm film blocks of different vertical and horizontal sizes, which are audible as square waves of different frequencies and amplitudes, respectively. The visual component of the film was created by manipulating the soundtrack on an optical printer, to create multiple copies in different foreground and background colours. In this way, Mac Laren used the technology of film to associate sound to image millisecond by millisecond" ([16]). The gathering and comparison of such pieces can bring results for a new fundamental theory of reversibility, transformation and interaction.

V. TOWARDS A VISUAL MUSIC THEORY

With the term interactive, we generally refer to any real-time adaptation in relation to another action. However, we accept that this relation does not only concern the two senses, but also exists for the viewer in a simultaneous use of the rest of the senses to get the stimuli.

This belief also works in the case of acousmatic music, where, even though the stimuli are less or more constant, the rest of the senses are always open to the stimuli, and the combination of receiving creates parallel activations that have to be measured because they affect the audience. An interesting point of view of this phenomenon is that of Antonio Damasio [20].

Antonio Damasio claims that the complex aesthetic stimuli of visual and auditory incentives, creates a neurological model inside the brain, which transforms into what can be described as an "image". As the "image" brings information from the physical aesthetic action field of hearing and touching, the term "image", according to Antonio Damasio, does not refer to the visual correspondence but to something more complex. Albert Einstein was one of the first who studied that phenomenon and named those models of brain images, "muscular images". From all this general information we gather that the creation of the system of the visualization of sound includes parameters from the fields of neurophysiology and

history of art, and if these fields are taken into consideration, the results allow the user a more friendly relationship with the machine -(the detest of a large part of the population towards technology because of the incomprehensible programs is known). There is a stable basis in all those facts. We need to give our faith to a new kind of reading of music, which does not need the time consuming familiarity of classical notation. Something of that sort would help the audience to experience music creation more consciously, and give them levels of sentiment beyond amusement, levels that, until today, only musicians shared.

Charles Rosen[(21)] has written convincingly about the important role of the written score in the Western classical tradition. The score was used to circulate new compositions, and preserve them for future generations. The audience's ability to read the score, at a time when most of the bourgeoisie learned musical notation, was critical to the reception of new works. Until the end of the 19th century, music was in large part a private experience. Most people would first encounter a Beethoven sonata alone at a piano, paging through the score. This private dialogue of discovery, between amateur pianist and composer, suggests that the Western relationship to music had once been closer to our contemporary relationship to poetry -- engaged with a page, searching, meditative.

Today, conversely, we think of music as belonging mostly to the public sphere. Rosen writes: "Our assumption today, made unconsciously, that almost all music is basically public is a radical distortion of Western tradition. We no longer have a public that largely understands how the visual experience of a musical score is transformed into an experience of sound, and to what extent this transformation is not a simple matter but is capable of individual inflections."

The transformation of sound into image enables an audience to better comprehend the musical structure of a work, by presenting them with another level of sensory involvement. Yet, in order to facilitate an engagement with a visual aesthetic, it is necessary to re-evaluate the criteria upon which the visual representation is based, and re-locate them from the technical arena to an aesthetic one. This could be done on three levels

- Firstly, by reviewing the technical software interface, encouraging composers and computer scientists to collaborate to improve software usability.
- Secondly, by instigating a co-ordinated effort to compile theories of music/sound reversibility, in order to consistently improve on the aesthetic criteria upon which the software is based.
-Finally, by examining new possibilities of acheiving reversibility of image and sound through mapping.

The parallel action of image and sound, offers to our comprehension of music a more drastic, vivid feeling about sound, the same as the one of the description of music, sci-fi books like "Brave New World" by Aldous Huxley [(22)] back in the thirties. This reminds us that technology is here today to serve the creative fantasy, and our wishful imagery, unless we turn it into an unnerving state of consuming ideas.

VI. CONCLUSION

Unforeseen events of auditory imagery in combination with visual representation offers a way of presenting and communicating in art, that emulates the way we get the stimuli in our everyday life. "This use of auditory information should be consistent with normal experience even if the phenomenon being represented is far outside normal experience"[(23)]. Human perception works in that way and in a way it helps a deeper understanding of musical composition, but the results so far in music are only experimental works of a repetition of all trivial musical systems. The chaos is a result of a combination of the absence of applied theory of structure and the denial of the musician computer-scientists to observe with greater attention and sensitivity the evolution of musical history.

As far as the visual part of the presentation is concerned, the absence of common parameters in the programmes apart from the visual depiction of the abscissa pitch in time, does not involve any new ideas but the refreshment of an old one, and presents a truth that in this time of evolution we tend to forget. In the general rush of evolution and presentation of new media, the new remains on the surface, because getting into a deeper sector demands the slowness and the introduction to codes that lie hidden in the history of art and not technology.

A formal language of representation is needed in order to give composers the opportunity to fully explore the possibilities offered by computer music-sound reversibility. If some consistent mathematical rules, and aesthetic values, were applied to create a formal environment, a collaborative and dialectic engagement with this new medium would become the source of many interesting approaches to composition.

VII. REFERENCES

[1] H. Lohner, "Interview with Iannis Xenakis", Computer Music Journal 1986, pp50-55

[2] J.Whitney, "To Paint on Water: The Audiovisual Duet of Complementarity", Computer Music Journal, 1994, pp.45-52

[3] Aristotelis, "On Sense and the Sensible",http://etext. library.adelaide.edu.au/mirror/classics.mit.edu/ Aristotle/sense.html

[4] T.E.Janzen, "Aesthetic Appeal in Computer Music", Computer Music Journal 1986, pp83-88

[5] V.S.Ramachandran and E.M.Hubbard, "Synesthesia: A Window into Perception, Thought and Language", Journal of Conaciousness Studies, 2001, pp34

[6] B. Evans, "Foundations of a Visual Music", Computer Music Journal 2005, pp11-24

[7] R.B.Dannenberg, "Interactive Visual Music: A Personal Perspective", Computer Music Journal 2005, pp25-35

[8]   S.M. Eisenstein,"Film Form" ,Harvest Books 1949, pp.3[9]   H.Lohner, "The Upic System: A User's Report", Computer Music Journal 1986, pp42-49

[10] K. Giannakis and M.Smith, "Auditory-Visual Associations for Music Compositional Process: A Survey"pp1-4

[11] K. Jennings.,"Hyperscore: A Graphical Sketchpad for Novice Composers", IEEE Computer Society 2004, pp50-54

[12] J. Reuter, "Soundpaint-Painting Music" http://lac. zkm.de/2005/papers/juergen_reuter.pdf pp1-7

[13] M.F.Barnsley, "Fractals Everywhere" 1989, Springer-Verlag

[14] G.Monro, "Fractal Interpolation Waveforms", Computer Music Journal, 1995, pp88-98

[15] M. Gogins, "Iterated Functions System Music", Computer Music Journal 1991, pp40-48

[16] O. Mattis, "Visual Music: Synaesthesia in Art and Music since 1900 ", Thames and Hudson Ltd, 2005, pp271

[17] R. Jones and B. Nevile, "Creating Visual Music in Jitter: Approaches and Techniques", Computer Music Journal 2005, pp55-70

[18] A. Sedes, B. Courribet, J. Thiebaut "Visualization of Sound as a Control Interface", Proc. of the 7th Int. Conference on Digital Audio Effects (DAFX-04), 2004 pp1-6

[19] M.S. Livingstone, "Vision and Art, the Biology of Seeing", Harry N Abrams, New York, pp.208

[20] A.R. Damasio, "The feeling of what happens, Body and Emotion in the Making of Consciousnes, Harcourt Brace and Co, New York, pp400

[21] C.Rosen, "Critical Entertainments, Music Old and New, Cmbridge: Harvard University Press, 2001, pp.36

[22] A. Huxley, "Brave New World", Bantam 1953, pp.266

[23] G.S.Kendall, "Visualization by Ear: Auditory Imagery for Scientific Visualization and Virtual Reality" Computer Music Journal 1991, pp70-74