



HAL
open science

NON-REAL-TIME TRANSDUCTION AND ITS USE AS A COMPOSITIONAL TECHNIQUE FOR THE CREATION OF VISUAL MUSIC

Carlos Lopez

► **To cite this version:**

Carlos Lopez. NON-REAL-TIME TRANSDUCTION AND ITS USE AS A COMPOSITIONAL TECHNIQUE FOR THE CREATION OF VISUAL MUSIC. ICMC - International Computer Music Conference, Aug 2013, Perth, Australia. hal-01665405

HAL Id: hal-01665405

<https://hal.science/hal-01665405>

Submitted on 15 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

NON-REAL-TIME TRANSDUCTION AND ITS USE AS A COMPOSITIONAL TECHNIQUE FOR THE CREATION OF VISUAL MUSIC

Carlos López Charles

Paris VIII University – MSH Paris Nord
CICM – EA 1572

ABSTRACT

The computer allows us to simulate all sorts of behaviours that go beyond the physical, mechanical and optical possibilities imposed by traditional visual music instruments (Levin, 2000, p.33). Nevertheless, the amount of data and the resolution of the sound and image files an artist can use in real-time to create visual music are limited by the processing capabilities of the computer (or computers) that he works with.

Transferring (or, to be more precise, *transducing*) data between digital representations of sounds, gestures and images is one of the main compositional techniques in today's interactive visual music. This article presents different techniques and software tools that can be useful in establishing transductions between digital sounds, images and control gestures in non-real-time. Three examples of the implementation of these ideas in different contexts will be presented in order to analyze the benefits of this approach. The first example, "Elastic Study", is an audiovisual piece in which different psychoacoustic characteristics of an acousmatic tape composition are analyzed and used to control a visual 3D particle system in non-real time. A second example, "The Flowers and the Clouds" is a visual music piece in which user's gestures are recorded and used to control a 3D particle system and an ensemble of synthesis sound processes in non-real time. The third and final example is a video in which a non-real time analysis of an HD video is used to control the parameters of a sound synthesis process.

1. INTRODUCTION

The term "visual music" has been used extensively to describe different forms of expression in which artists create analogies between sound and image through a constellation of ideas: such as the ideas of abstraction, synaesthesia and the aesthetic experiences of time, geometry and gesture. Several works have been created in this spirit by artists such as Kandinsky, Survage, Klee, Kupka, Delaunay, Ruttman, Eggeling, Fischinger, McLaren, Lye and the Whitney Brothers.

More recently, visual music artworks have been made with the use of digital systems in which heterogeneous elements are brought together and stabilised into functional units through the implementation of transduction¹ processes. Regardless of

being used in real-time or non-real-time, transduction processes offer *models* for the creation of structural relationships between different *representational spaces*², where collections of discretely-defined symbols allow us to have a certain control over the forms and behaviour of sounds, images and different objects (patches, spectrograms, control algorithms, etc.) that can form part of a compositional environment. Consequently, within a digital environment, we can consider these transfers of information between different media to be transductions of symbols between different representational modes. The use of transduction as a tool for mediation becomes most effective in this context because it enhances the connection of the most diverse representational spaces. From the moment in which symbols in a digital environment acquire the role of generating new morphologies, they adopt –as Vaggione describes it - a role that is similar to music notation within a score (Vaggione, 2006, p. 14), becoming the elementary units of a language in which different characteristics of a represented object can be described, manipulated and articulated within a network of compositional operations. The use of transduction is an artistically pertinent approach because it sets specific criteria that can be used to endow objects in an audiovisual composition with a multiplicity of forms and behaviours. However, the use of this technique in non-real-time contexts has not been explored in depth.

2. THE USE OF TRANSDUCTION IN THE CONTEXT OF VISUAL MUSIC

In physics, transduction is defined as the conversion of one form of energy into another form of energy (University of Oxford, 2005, p. 672). An electroacoustic transducer, for example, is a device that allows us to convert electric energy into acoustic energy and vice-versa. Condenser microphones and loudspeakers are examples of transducers used for the recording and playback of sounds.

Concerning audiovisual expression, transducers allow us to convert any kind of energy (radiant, mechanic, thermal, magnetic or chemical) into sounds and images³. In this way an input energy, like the

¹ SIMONDON, Gilbert *L'individuation à la lumière des notions de forme et d'information*, Éd. J. Millon, coll. Krisis, Grenoble, 2005, p. 32. explains transduction in the following way: "We understand transduction as a physical, biological, mental, social operation by which an activity is spread within a domain, founding this spreading

process on a structure that operates from place to place: every defined structural region serves the next one as a constitutive foundation, even though a modification extends progressively at the same time as this operation takes place"

² VAGGIONE, H. *Symboles, signaux, opérations musicales*, CICM-Université Paris VIII, unpublished article, 2006, p.14

³ EMMERSON, Simon *Living Electronic Music*, Ed. Ashgate, Anglaterra, 2007, p.117-142 offers a very complete overview of the different transducers used in electroacoustic music.

mechanical energy of a human gesture¹, can be transformed into electricity and, conversely, into magnetic and radiant energies that could control the variations of, for example, the pressure in a loudspeaker and the luminosity variations in a set of lights, respectively. In such cases, the trace of the input energy becomes a controller (or regulator) of the output energies that will be ultimately perceived as sounds or images.

Compared to analogical devices, digital technologies offer a lot of flexibility in the control of audiovisual materials. On the one hand, the use of digital-to-analog and analog-to-digital converters allows us to represent electric energy variations as a discrete (digital) collection of symbols. Conversely, the computer allows us to generate, store and manipulate any collection of symbols and create electric signals from them. This opens up the possibility of operating on different kinds of media not only from at the perceptual level, but also at the level of their code.

In summary, control in computer-made visual music may be achieved *by the use of input transducers, by generating a collection of symbols within the computer that will be eventually transduced into sounds or images* (which has the advantage of allowing for more complex control procedures²), and *by a combination of these techniques*.

3. NON REAL-TIME TRANSDUCTION

Not many tools have been developed to facilitate the circulation of data between different representational spaces (digital representations of sound, image, gesture, etc.) in the creation of visual music in fix video and cinema. As a consequence of this, artists have had to use tools that were originally conceived for real-time interaction. The drawbacks of this situation are considerable: the complexity of possible the transduction processes are limited to very simple operations (such as controlling the colour of the image with the amplitude of a sound, for example). If we wanted to stage a transduction process that involved heavy processing for the average computer (like, for example doing an FFT analysis and using the frequency of each bin as a control of a 100,000 element particle system in which each element took the colour of a high resolution video), there would be considerable latency issues or even unexpected program crashes and, with this, a decrease of the *transformational capacity*³ (the ability of animating the

objects in our compositional network with different forms and behaviours) of the system.

What kind of software solutions could help in the creation of non-real-time visual music works that incorporate the possibility of establishing transduction relationships between different modes of representation?

4. STUDY CASE NO.1: SOUND AS THE CONTROLLER

The software for the piece “Elastic Study” was developed in the Max/MSP/Jitter programming environment. The music for this piece was created after testing how Tristan Jehan’s object *analyzer~4* (which makes real-time estimations of features like the brightness, noisiness, loudness, etc. of a given sound) could be used to control different variables of a particle system on a 3D scene created with the *jit.p.shiva* and *jit.p.vishnu* objects of the Jitter library. The sonic elements of this piece were composed in a way in which they could provide a very flexible control of the elements of this particle system. In order to provide more flexibility to the sound-image relationship of this piece, I integrated a system that would allow to vary the mappings automatically in particular sections of the piece (such as the amount of particles, their variations of colour, duration, inclination, the background colour, transparency and the camera’s position).



Figure 1. Still image from “Elastic Study

Even though the software presented a stable performance in real time, recording the output in high definition quality became impossible due to the amount of CPU consumption that the process demanded. To solve this problem, I created an abstraction called “dataSequencer” (figure 2), which let me 1) record the output values of the sound analysis, 2) sample them at a speed of 30 samples per second and 3) output these values in non-real-time speed so that the individual frames of a 30 frames-per-second movie could be generated individually, without surpassing the processor’s capacity to render them.

Having generated the individual frames of the video in non-real-time, different Final Cut sessions were created to bounce the frames of each section into small

¹ VOLPI, G. et al. *Gesture Analysis dans Enaction and Enactive Interfaces : a Handbook of Terms* édité par Annie Luciani et Claude Cadoz, Enactive System Books, ACROE, Grenoble, France, 2007. define gesture as a human action

² Portillo O. et Polyviani I., “Control, Digital” in *Enaction and Enactive Interfaces : a Handbook of Terms*, édité par Annie Luciani et Claude Cadoz, Enactive System Books, ACROE, Grenoble, France, 2007, p.58.

³ VAGGIONE Horacio, “Vers une approche transformationnelle en CAO Actes des Journées d’Informatique Musicale (JIM) 1996. Les cahiers du CNRS. CNRS-Université de Caen. Electronic version available at <http://jim.afim-asso.org/jim96/actes/vaggione/VaggioneTEXTE.html#Tahil>

⁴ JEHAN, Tristan. Disponible sur le site électronique personnel <http://web.media.mit.edu/~tristan/>

videos that were finally mixed together into a final piece. This strategy made it possible to edit the frames produced by Max/MSP/Jitter and to correct any errors that could have been eventually produced through the use of the Jitter abstractions, which proved to be somehow limited for this kind of approach.

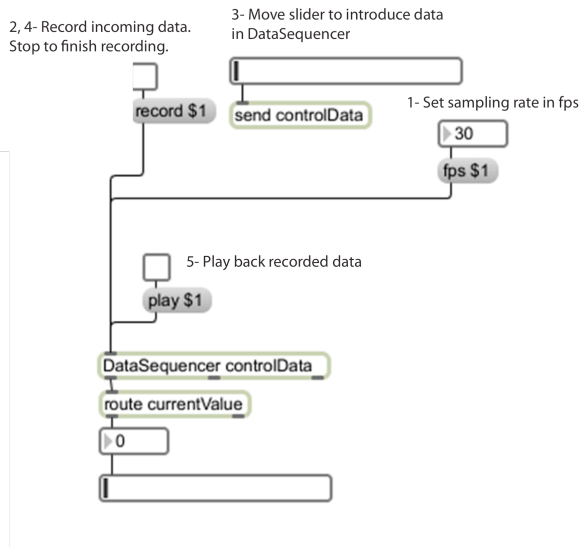


Figure 2. Basic use of the DataSequencer object

5. STUDY CASE NO.2: GESTURE AS THE CONTROLLER

“The Flowers and the Clouds” is a piece made using Processing¹ and Max/MSP in which a visual particle system adopts the colours of a background video in order to create abstract imagery that lies somewhere between the figurative and the non-figurative. The particle system was controlled with Max/MSP through the use of different sliders that allowed to modify different characteristics of their elements, including the number of particles, their speed, the randomness of their trajectories, their transparency, their size, etc. The communication between Max/MSP and Processing was made through the OSC protocol.

The control of the particle system proved to be very efficient within a certain threshold of complexity, but when the number of particles and their transparency increased beyond a certain number (for example, 12,000 particles), a slow down in the response of the system became evident and the level of *performativity*² was affected. To solve this situation, I used a similar solution to the one in “Elastic Study”, but instead of using the analysis’s data of a soundtrack as the control data, I used gestural data that that was 1) recorded from improvisations with a set of virtual sliders, 2) edited in non-real-time so that the undesired accidents in the

¹ www.processing.org

² LEVIN, Golan *Painterly Interfaces for Audiovisual Performance*, Massachusetts Institute of Technology, 2000, p.33. defines performativity as the degree in which a system allows the users to establish a feedback loop in which they can explore their potential as actors in an open work.

improvisations could be corrected and 3) used to generate the individual frames of a movie.

The individual frames for this video were assembled into a HD Quicktime movie by using a custom patch made in Processing. In this case, no synchronisation problems occurred, as it happened with the use of the Jitter objects in the first example. My observation with regard to this is that open-code software like Processing allows for much more precise programming at a low level than that offered by software like Jitter, which translates into more creative flexibility in cases like this one.



Figure 3. Still image from “The Flowers and the Clouds”

The soundtrack for this piece was made with a commercial FM synthesizer (FM8), whose different variables were also controlled in non-real-time through the use of recorded and edited gestural data from improvisations made in sync with the video, which allowed for a very precise writing of the sound materials and facilitated the establishment of synaesthetic relationships between the images and the music during the compositional process.

6. STUDY CASE NO.3:

“Miniature 3” is a short piece in which the movement of a particle system in a pre-recorded video controls the playback speed of two soundfiles (one for each audio channel) to create a synaesthetic effect through the use of synchronisation. Even though this synchronic approach could lead us to a situation in which we create sound-image relationships that seem too direct, it also provides a solid ground to achieve synaesthetic sensations that would be impossible to create otherwise³.

The computational cost of creating the sound for this video in real-time was too high because the image analysis was very demanding. However, the non-real-time solution proved once more to be very efficient within this context.

The audio for this video was created with Processing and Max/MSP through a process in which 1)

³ COLLINS N. et OLOFSSON F., *kipp av : Live Algorithmic Splicing and Audiovisual Event Capture*, Computer Music Journal 30:2., Massachusetts Institute of Technology, 2006, p. 8–18.

Processing divides a high-definition (1280 by 720 frames) video into two regions (left and right), 2) the amount of movement in each region is estimated by Processing (through the use of a frame-difference algorithm) and sent to Max/MSP as a sequence of numbers that Max saves into a text file and 3) the data in the text file is converted in non-real time into an audio signal which is finally used to control the playback speed of the sound files in the left and right audio channels.

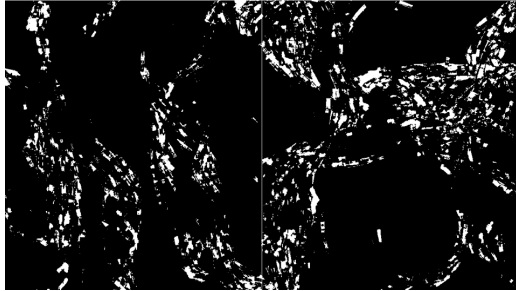


Figure 4. Movement detection on “Miniature 3”. Movement on each side of the screen (white pixels) is used to control the playback speed of sound files in non-real time.

7. CONCLUSIONS

Composing visual music through the use of computers enables the use of digital representations of sounds, images, movement or any kind of data -as well as the analyses that can be made from it- as models for the generation of forms and behaviours in different representational modes. Sounds, images and physical movement may present considerable perceptual differences between them, but represented as symbols, they can attain a certain degree of stability and the open up a space to establish *translations* (or, more precisely, *transductions*) between them. By using a transductive approach to visual music composition, the artist uses a representation of *something* as a model to create *something new*.

As Simon Emmerson points out (Emmerson, 2007, p. 42), the compositional choices that an artist has to make in using models in his creative process are at least four: [I would number these] the identification of the model and its behaviour, the choice of the variables that will correspond to the chosen model and the definition of the scale of their relationships. However, by choosing to compose digital visual music with a transductive approach, another decision has to be made: the nature of the compromise that will take place between the performativity of the system and the quantity and quality of the materials that will be used in the compositional process. In that regard, a non-real time approach seems to be a good option for the creation of visual music.

8. REFERENCES

- [1] Collins N. and Olofsson F., “kipp av: Live Algorithmic Splicing and Audiovisual Event Capture”, *Computer Music Journal* 30 :2, Massachusetts Institute of Technology, 2006.

- [2] Emmerson S., *Living Electronic Music*, Ashgate, England, 2007.
- [3] Levin G., *Painterly Interfaces for Audiovisual Performance*, Massachusetts Institute of Technology, 2000.
- [4] Portillo O. et Polviani I., “Control, Digital” in *Enaction and Enactive Interfaces : a Handbook of Terms*, édité par Annie Luciani et Claude Cadoz, Enactive System Books, ACROE, Grenoble, France, 2007.
- [5] Simondon, Gilbert *L'individuation à la lumière des notions de forme et d'information*, Éd. J. Millon, coll. Krisis, Grenoble, 2005, p. 32.
- [6] University of Oxford, *Oxford Dictionary of Science*, 5th edition, New York, Oxford University Press, 2005.
- [7] Vaggione H., *Objets, représentations, opérations*, CICM-université Paris 8, France, 1995.
- [8] Volpi G. *et al.*, “Gesture Analysis” in *Enaction and Enactive Interfaces: a Handbook of Terms*, édité par Annie Luciani et Claude Cadoz, Enactive System Books, ACROE, Grenoble, France, 2007.

9. WEB RESSOURCES

All the videos from this article are available at www.carloslopezcharles.com