# Evolutionary Sound Synthesis Controlled by Gestural Data

José Fornari (UNICAMP)
Mariana Shellard (UNICAMP)
Jônatas Manzolli (UNICAMP)

**Abstract:** This article focuses on the interdisciplinary research involving Computer Music and Generative Visual Art. We describe the implementation of two interactive artistic systems based on principles of Gestural Data (WILSON, 2002) retrieval and self-organization (MORONI, 2003), to control an Evolutionary Sound Synthesis method (ESSynth). The first implementation uses, as gestural data, image mapping of handmade drawings. The second one uses gestural data from dynamic body movements of dance. The resulting computer output is generated by an interactive system implemented in Pure Data (PD). This system uses principles of Evolutionary Computation (EC), which yields the generation of a synthetic adaptive population of sound objects. Considering that music could be seen as "organized sound" the contribution of our study is to develop a system that aims to generate "self-organized sound" – a method that uses evolutionary computation to bridge between gesture, sound and music.

**Keywords**: Generative art. Gestural data. Evolutionary computation. Self-organization. Sound design.

**Título: Síntese Evolutiva de Sons Controlada por Dados Gestuais**

**Resumo:** Este artigo trata de pesquisa interdisciplinar que envolve Música Computacional e Arte Visual Generativa. Descrevemos aqui a implementação de duas instalações artísticas interativas baseadas nos princípios de coleta de dados gestuais (WILSON, 2002) e na auto-organização (MORONI, 2003) para controlar um Modelo de Síntese Evolutiva Sonora (ESSynth). A primeira implementação utiliza como dados gestuais o mapeamento por visão computacional de imagens de desenhos feitos a mão. A segunda usa dados estuais de movimentos dinâmicos de coreografia. O resultado computacional é formado por dados gerados por um sistema interativo implementado em Pure Data (PD). Este sistema utiliza princípios de Computação Evolutiva (EC), que permitem a geração de uma população sintética de objetos sonoros dinâmicos. Considerando que música pode ser definida como "som organizado", a contribuição deste estudo consiste em desenvolver um sistema computacional que gera "som auto-organizado" – um método de computação evolutiva que intersecciona gestos, som e música.

**Palavras-chave**: Arte generativa. Computação evolutiva. Auto-organização. Projeto sonoro.

FORNARI, José; SHELLARD, Mariana; MANZOLLI, Jônatas. Evolutionary Sound Synthesis Controlled by Gestural Data. *Opus*, Porto Alegre, v. 17, n. 1, p. 133-153, jun. 2011.

Computer music research studies digital sound synthesis methods described by discrete mathematical models, from one of the following categories: Linear vs. Nonlinear; Deterministic vs. Stochastic; Static vs. Dynamic (BENDER, 1978). Among Dynamic models, there are the Adaptive ones, such as Evolutionary Computation and Genetic Algorithms. They resemble, in a computational environment, the adaptive characteristics of problem-solving, observed in biologic individuals and populations. Such models are capable of changing their search strategies to find the best possible solution for a problem that is not well-defined (HOLLAND, 1996).

For such problems, there is a class of interactive genetic algorithms in which human-machine interactions are required, in order to find their best solutions. Evolutionary environments that use body movements to control the evolution are one example within a family of interactive adaptive computation systems introduced by Stephen Todd and William Latham (1992) and described by Alexander Kosorukoff (2001) as Human-Based Genetic Algorithm (HBGA).

Gesture is here defined as a set of body movements bound by one specific intention. Lately, the retrieval of gestural data has been a key issue in the study of adaptive systems applied to computer music (WANDERLEY, 1999). The body continuously corrects its movements to perform an intended task, which – in spite of its complexity – can be trained by repetition, and reproduced – in most cases – at ease.

This article describes a computer musical model using evolutionary computational methodology to synthesize sound objects generated through human-machine interaction based on gestures. This model can automatically generate sound synthesis originated from gestural data retrieved from a collection of drawings and dance movements. As further explained, this method leads to a computer model that generates "self-organized sound". This can be used in interactive sonic installations, such as the one presented by the group of Georg Fleischmann (1994). We understand our concept as an extension of the well-known Edgard Varèse's viewpoint about music, as being "organized sound". In the next sections we present a review of the research background, which is followed by the description of this computer evolutionary model and a discussion of the achieved results.

## Theoretical Background

In this section we describe the theoretical basis of our approach. We start with a brief review of the contemporary concepts of generative art, converging to computer generated music. It follows a brief review on Evolutionary Sound Synthesis, the methodology here used to generate the sound output of this musical application.

## Contemporary Generative Art

Generative Art are broadly defined as rule-based artwork systems, which are "set into motion with some degree of autonomy" (GALENTER, 2003). Algorithmic and Generative processes had already been extensively explored in arts and music composition. Athanasius Kircher – believing that musical harmony should reflect universal proportions – wrote in the 1650's the book entitled: *Musurgia Universalis*, where he described the design of a music generating machine (CRAMER, 2005). In 1793, Hummel published a system to generate musical notation, whose creation was attributed to Mozart, in which music scores were assembled by predefined bars, chosen through a dice tossing game. Mozart's system embeds most of generative contemporary methodologies in which musicians might create – throughout simple rules and building blocks (i.e. the predefined musical bars) – an astonishing amount of new compositions. Later, this algorithm was known as the Mozart's dice game and influenced many composers, such as John Cage and Lejaren Hiller, in the creation of their musical piece entitled HPSCHD (HUSARIK, 1983).

The main proposition of Generative Art is to release the artwork from the artist's total control, thus letting it, as an independent entity, free to explore subtle variations through reiterative generative processes. Reiteration and loosening of control where well explored during the 1960's, by musicians, visual artists and dancers, sometimes also exploring partnerships with scientists and engineers. In 1968, Billy Kluver, John Cage and Robert Rauschenberg organized the *Project 9 Evenings*, where artists from different fields and engineers from the Bell Laboratories worked together on developing human-machine interactions, during 9 evenings of artistic performances. Some of the choreographies created were mainly based on simple repetitive tasks that interacted humans and machines (programmable electronic devices). This project resulted in the development of new technologies and artistic approaches over new tools and led the engineers Billy Kluver and Fred Waldhauer, with the artists Robert Rauschenberg and Robert Whitman to create the E.A.T. (Experiments in Art and Technology), a non-profit organization that gathered researchers and artists to work together, developing art and technology.

An artwork autonomously created by a generative process is not restricted within any specific field, but in a multitude of different areas of knowledge, even beyond visual arts and music. Adaptive methods, such as Artificial Intelligence, Neural Networks and Evolutionary Computation, are recent areas of interfacing Science and Art. This fruitful dialogue fits well the principles of Generative Artwork, due to their potential of creating a process that is dynamic and immersive.

## Evolutionary Art and Interactive Genetic Algorithm

Evolutionary Art – a recent form of creation – is almost exclusively created with computational aid. Genetic Algorithms (GAs) and their creative use in Arts were introduced by Todd and Latham (1992) when they evolved visual forms based on GAs. Peter Bentley and David Corne (2002) introduced the concept of evolutionary design and there is a collection of creative applications of evolutionary systems in the book *Creative Evolutionary Systems*. From this pioneering work, several computer models for Evolutionary Art were designed. This is based on a Selection process, inspired by the artist viewpoint, guiding the artwork development – that is somehow similar to the Darwinian natural selection guided by environmental conditions. In Music, one of the first works of this type is GenJam (BILES, 1994), a system designed to generate jazz solos in synchronism with a predefined improvisation harmony line. Another pioneering GAs musical application, developed by Peter Todd and Gregory Werner (1999), is a system for Evolutionary Music Composition.

The approach described in this article relates to a family of interactive evolutionary computation systems, referred as Interactive Genetic Algorithm (IGA). This is a GA subclass that uses human judgment as the fitness evaluation. In computational terms, IGAs are successfully applied when it is not possible to find an analytical definition as fitness function. This is also a form of incorporating human arbitrariness in the flow of a computer generated synthetic evolution. Kosorukoff (2001) also describes this class of interaction as the previously referred Human-Based Genetic Algorithm (HBGA), that he defines as a genetic algorithm that allows humans to contribute with solutions, suggestions and guidance to the evolutionary process. Hee-Su Kim and Sung-Bae Cho (2000) studied computer-aided adaptable systems using IGAs for fashion design. Dunwei Gong, Xin Yao and Jie Yuan (2009) studied a hybrid approach, combining IGA with fitness functions not assigned by human. This type of semi-automatic decision process is a way of solving the problem of human fatigue in taking decisions over a large amount of possible solutions data. In this context, we created the concept of ArTbitrariness (MORONI; MANZOLLI; VON ZUBEN, 2002a) that mediates the application of computer creativity in art production (MORONI; VON ZUBEN; MANZOLLI, 2002b).

**Evolutionary Sound Synthesis**

It is well-known that adaptive computing methodologies can produce emergent, self-organized complex systems (HOLLAND, 1992 and 1996). Among them, there is the Evolutionary Computation (EC), an adaptive method inspired in the problem-solving approach observed in biological species. This method seeks out – in evolutionary steps – for the best solution, within a landscape of evolving candidates (possible solutions).

Evolutionary systems are also described by the theory of Complex Adaptive Systems (CAS). As described by John Holland (1992), CAS consists of a large number of agents with interconnected parameters that, altogether, exhibits coherent emergent properties. It is also known that CAS can generate emergent properties by means of its agents' competition and/or cooperation (HOLLAND, 1996). Its systemic behavior is the result of interactions of a large number of agents, leading to the process of self-organization, in which, a CAS may pass through several distinct organizational states (FOERSTER, 1960), which is the fundamental principle of the methodology presented here.

Since 2001, the research group of NICS (Interdisciplinary Nucleus of Sound Communication, at UNICAMP) has worked with ECs, in sound design and music composition (MORONI; MANZOLLI; VON ZUBEN, 2002a). Some of these techniques created highly textural environment (FELS; MANZOLLI, 2001). Then, we developed the Evolutionary Sound Synthesis (ESSynth) in which a Population set of waveforms evolves in time, in generation steps, by the action of two genetic operators: crossover and mutation, and a predefined fitness functions (MANZOLLI, 2001), (FORNARI et al., 2001). Later, the genotype of each individual in the population set was defined by psychoacoustic arrays of acoustic aspects that aimed to define how the waveform was perceived and understood. Finally we incorporated sonic spatial localization parameters in this evolutionary method (FORNARI; MAIA JÚNIOR; MANZOLLI, 2006. FORNARI; MAIA JÚNIOR; MANZOLLI, 2007).

This article describes a new musical development of EESynth for dynamic sound design in which there is no Selection through the usage of fitness functions, but through the definition of a lifespan for each individual in the variable-size population set. As said before, this implementation is based on the concept of Self-Organization and our hypothesis is that the whole population is evolving towards an organizational attractor. The usage of gestural data retrieved from drawings and dance movements constitute a type of disturbance input on this computer music model that, in reaction, adapts to the stimuli. The resulting sound output describes emergent paths of adaptation, in the sound domain.

## Methodology

This section describes the method here used to retrieve gestural data, from drawings and dance, used to create a Population of sound objects, through Evolutionary Sound Synthesis.

## From Gestures to Sounds

Our development started from the computational analysis of digital images taken from drawings. First, we developed a model to find and retrieve information from graphic objects and map them into sound objects. These were then constituted as genotypes of individuals, in a population set. Later, this method was expanded to also handle kinetic data – the gestures retrieved from body movements of dance. These two cases will be further described.

Both data types (drawings and dance) share the common characteristic of being similar yet variant, which means that all drawings from the collection, as well as all dance movements, are similar but never identical. This resembles an important feature found in biological populations: Individuals of the same species are similar, but there is no natural occurrence of clones (even for twins, there is a slight genotypic difference, due to mutation). These data constituted the Genotype, representing acoustical characteristics of each Individual: the sound object.

## Computer Code as Individuals

One of the aims of this work is to use an ESSynth model to generate emergent sonic properties, based on external data retrieved from gestures. In order to implement this new implementation of ESSynth, we defined the Individual as the sound synthesis algorithm itself. Here, this is implemented as an instantiation of a PD abstraction – an independent sub-patch, similar to the C language subroutine that can be replicated in independent instantiations, having their own arguments and handling their internal variables as local variables. In a nutshell, each Individual is a process that synthesizes a sound object. Different from other studies, instead of evolving synthesis parameters, here we aim to evolve the Code itself.

Resembling biology, Individuals have their own genotype, defined here by acoustic descriptors related to features of their corresponding sound objects. Genotype data is initially created from the mapping of gestural data, as described in section 3. There are also interactive parameters to control the evolutionary sound process of ESSynth, in real-time.

They are: (1) Proliferation, which describes the rate of reproduction in the Population set; (2) Lifespan, that sets the average of how long each individual, in the Population set, can be alive; (3) Crossover, that sets the amount of crossover, or recombination, during the reproduction process; and (4) Mutation rate, which sets the amount of mutation in the new individuals' genotype.

The Population set, in this version of ESSynth, starts with only two individuals (two instantiations of a PD abstraction). Its size grows as the reproduction process creates new individuals, and reduces when the individuals' lifespan is over. During the evolutionary process, new individuals will be born, thus creating sound objects, and old individuals will die, when the correspondent sound object also disappears. The ESSynth sound output is given by all individuals alive, which creates an emergent sound realm by all active individuals within the Population set. Individuals will also independently move around. The location of each sound object is given by simple Interaural Time Difference (KELLY; PHILLIPS, 1991) and Interaural Loudness Difference (BIRCHFIELD; GANGISHETTY, 2005) sound location algorithms.

Here, individuals don't have gender differentiation, but will – in pairs – generate offsprings. Genetic operators: Crossover and Mutation assemble the offsprings new genotypes from the genotypes data of their two parents. Each genotype has six chromosomes, given by arrays of floating-points. They are separated in two sets of three arrays, controlling sound synthesis characteristics: Tonal and Stochastic. Tonal set synthesizes sound with clear pitch. Stochastic one synthesizes sounds without clear pitch. Each group of 3 arrays control synthesis characteristics related to Intensity, Frequency and Distortion. Intensity controls the sound intensity of each set. Frequency is associated with pitch (for the Tonal set), which is the center frequency filter (for the Stochastic one). Distortion controls the amount of partials in the Tonal set, and the filter bandwidth, in the Stochastic set. It is interesting to realize that the Distortion parameter works as a bridge between Tonal and Stochastic features. Without distortion, Tonal set generates a sine-wave sound and the Stochastic one, a white-noise. As the Distortion rate increases, Tonal output increases the number of partials, by clipping the sine-wave, thus making it more similar to a square-wave; and the Stochastic output by narrowing the filter center-band, generates a sound similar to a whistle.

As shown in Figure 1, each array is a time series that describes the progression of one parameter along of the individual lifespan. Each array has a fixed length of 100 elements, that are real numbers, normalized between [-1,+1]. These 6 arrays compound the individual genotype. In the system, they are given as a text file which contains a matrix $M_{6,100}$ – the Gestural Matrix. From top to bottom, the 6 lines of the gestural matrix text file are: Tonal

Intensity, Tonal Frequency, Tonal Distortion, Stochastic Intensity, Stochastic Frequency, Stochastic Distortion.

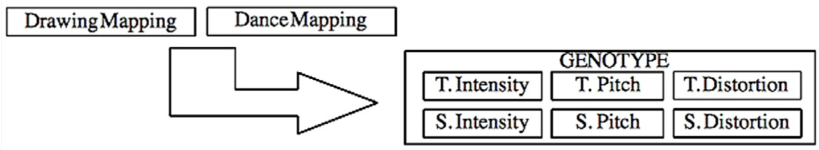


**Fig. 1:** Diagram of the mapping from gestural data to the six arrays of the Genotype.

## Retrieving Gestural Data

Gesture, as an artistic expression, is here seen as movements and actions embodying artistic intention. Our hypothesis that is possible to access information that resembles the generative artistic process by retrieving gestural data conveyed by the final artwork (drawings) or performance (dance). This can be expressed in the sonic domain, by using the retrieved gestural data, to generate dynamic sound synthesis through an evolutionary system. The retrieval of such data is described in the next subsections.

## Gestures from Drawings

We started our work with gestural data retrieved from drawings. They belong to a large collection of over 200 conceptual drawings, all sharing the same characteristic of being very similar, but never identical. They were created by the artist through the repetition of a similar (back-and-forth) gesture that created a collection of drawings alike, somehow resembling a biological population of individuals belonging to the same species.

Some of these drawings were scanned and their digital images were mapped by a computer algorithm. These mappings were used with the ESSynth to create an artistic installation where graphic characteristics from each drawing could be mapped into sonic characteristics of sound objects, in the form of individuals' genotypes. These individuals belong to the Population set of ESSynth, where its dynamic evolution creates a landscape of sound objects. Figure 2 shows a detail of one of these drawings. On the left side is the digital image of one of the drawings. On the right side is its mapping. The large blobs are the accumulation of dripped paint, and the quasi-parallel lines, were created by the back-and-forth paintbrush movement.

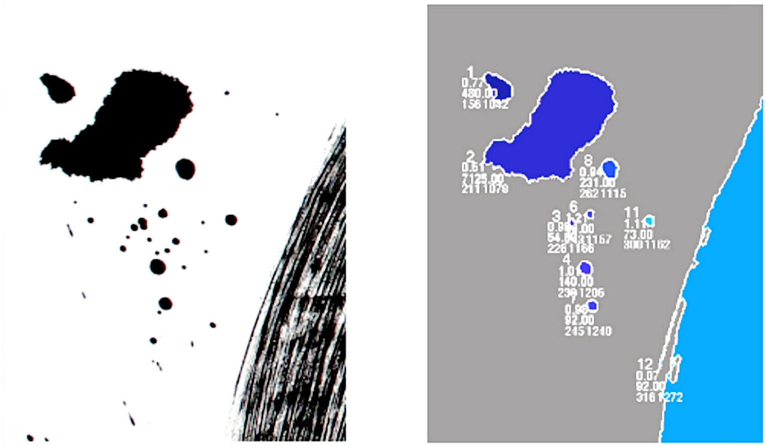


**Fig. 2:** Mapping a processual drawing collection. Detail of the digital image of one drawing (left) and its mapping, into several graphic objects (right).

The first step was to create a convention to map drawing features into sonic features. This mapping was developed in MATLAB. On the right side of Figure 2, it is shown the mapping of this image, done by this algorithm that is able to retrieve several graphic objects for each image. Each object has also several features associated with it. Some of these features are shown in this figure, imprinted in white, at the left side of each graphic object.

For this work, we established a convention that these objects belong to three categories. Altogether, they characterize the identity of each drawing. Such graphical elements are also found in any drawings from this collection. They are: (1) Accumulation, (2) Repetitions and (3) Fragments. Accumulation is the biggest object found in each drawing. As such, there is only one accumulation per drawing. This is usually the concentration of paint found at the bottom of the image, where the artist's gesture normally initiated. Repetitions are objects stretched in shape, similar to fragments of lines. They are normally the quasi-parallel traces found at the middle of each drawing, generated by the back-and-forth gesture of the artist's paintbrush. Fragments are the small, detached and circular spots of paint that dripped at the outlying areas of each drawing paper, spilled by the gesture, while increasing speed and intensity. Following that, we related each graphical element with any unique sonic aspects that we believe to synestesically represent – in the acoustic domain – graphical aspects of the drawings. We related the object Accumulation to the stochastic constant and low-frequency sonic features, that are steady and with long duration

(10 seconds, or more). Repetitions were related to sine-wave sounds, with middle-ranged frequencies and middle duration (about 3 seconds). Fragments were related to short-time (less than 1 second) pulses, similar to sparks of either stochastic or tonal sounds. Each image mapped generates several graphic objects; one is Accumulation, and all the others, either fall into the category of Fragments or Repetitions. Each object has several features associated with it. They are mapped into a matrix conveying the genotypes of individuals in the Population set that started the evolutionary process.

The mapping is done through the projection of the bi-dimensional shape of each object, into the horizontal and vertical coordinates of the whole image. This projection is sampled into a time series of 100 elements each. For tonal sounds, the horizontal projection is used. For stochastic sounds, the vertical projection is used. Distortion curves are calculated by the difference between horizontal and vertical projections. To express the influence of the objects position, each projection was circularly shifted, according to the distance between its objects and the whole image origin.

The tonal intensity curve is calculated by the blend of all horizontal projections, modulated by each object eccentricity parameter. Tonal frequency curve is given by the blend of all vertical projections, modulated by the orientation angles of their respective objects. Tonal distortion is given by the blend of the difference between horizontal and vertical projections, modulated by the inverse of its eccentricity.

Stochastic intensity is given by the blend of all vertical projections, modulated by the square root of each object normalized area. Stochastic frequency is the blend of all vertical projections, modulated by the objects normalized orientation angle. Stochastic distortion is the blend of the projections difference, modulated by its eccentricity.

## Gestures from Dance

One of the first to study the movements of dance was Rudolf Laban, a famous choreographer and movement theoretician. He gathered his findings in the title Laban Movement Analysis that explains eight categories of basic movements that are the combination of three independent types of effort actions: (1) Space, (2) Weight and (3) Time. According to Laban, the eight basic movements are: (1) Float, (2) Punch, (3) Glide, (4) Slash, (5) Dab, (6) Wring, (7) Flick, and (8) Press. They have been used by several schools of dance and acting, as movements embodying specific emotions (PFORSICH, 1977). Data from these gestures can be computationally retrieved and used in several manners. For instance, the InfoMus Lab has developed the software EyesWeb, an interactive system for the real-time analysis of movements and acquisition of expressive gesture (MANCINI;

CASTELLANO, 2007). Our research group, in collaboration with GITD – the Interdisciplinary Group of Theatre and Dance at Unicamp – developed a performance called Elementaridades. It was inspired in the physical movement of particles of matter, and its application in Laban´s principles of movement in dance (MAIA JÚNIOR et al., 2001).

In this work, similar gestures were collected as movement data, using as gestural interface two remote-controls of a Nintendo video-game – wii-remote (wiimote) – with it accessory, the nunchuck. Each wiimote and nunchuck has embedded a tridimensional accelerometer that retrieves acceleration parameters for the three coordinates of space, in real-time and transmits these data wirelessly – via bluetooth – in the form of seven motion parameters. Three of them are rotational angles of the space dimension. They are named, in aviation terms: Yaw, Pitch and Roll (LAVALLE, 2006). The next three parameters are the raw accelerations over the coordinate axes: X, Y, and Z (for each space-dimension). The seventh parameter is: accel, and expresses raw general acceleration, despite its direction.

The equations below show the rotation matrixes describing the correlation between Yaw, Pitch and Roll with its rotation around the orthogonal axes, related to each respective space-dimension x, y and z.

$$yaw(x)=\begin{pmatrix}\cos(x) & -\sin(x) & 0\\ \sin(x) & \cos(x) & 0\\ 0 & 0 & 1\end{pmatrix} \quad pitch(y)=\begin{pmatrix}\cos(y) & 0 & \sin(y)\\ 0 & 1 & 0\\ -\sin(y) & 0 & \cos(y)\end{pmatrix} \quad roll(z)=\begin{pmatrix}1 & 0 & 0\\ 0 & \cos(z) & -\sin(z)\\ 0 & \sin(z) & \cos(z)\end{pmatrix}$$

The data was collected by a computational model, developed as a PD patch that recorded each movement in synchronism with the seven parameters transmitted by each accelerometer (given a total of 28 time-series) at 8 bits of resolution and sampling rate of 100 Hz. As shown in Figure 3, the accelerometers (the "sensors") were attached (taped) to the dancer's knees and elbows. The resulting data was saved in text files, automatically created by the patch, every time it was running.

The dancer performed eight short improvisations (of about 1 minute each) for each entry described in the rows of Table 1. It shows these eight body actions, as described by Rudolf Laban, and its formant aspects of: Space, Weight and Time. The movements performed by the dancer where improvised according to these premises.

| **The Body Actions described by Rudolf Laban.** | | | |
|---|---|---|---|
| **Action** | **Space** | **Weight** | **Time** |
| **Sliding** | Direct | Light | Slow |
| **Fluctuating** | Flexible | Light | Slow |
| **Punctuating** | Direct | Light | Rapid |
| **Shaking** | Flexible | Light | Rapid |
| **Pressing** | Direct | Firm | Slow |
| **Twisting** | Flexible | Firm | Slow |
| **Punching** | Direct | Firm | Rapid |
| **Whipping** | Flexible | Firm | Rapid |

**Tab. 1:** The Body Actions described by Rudolf Laban.

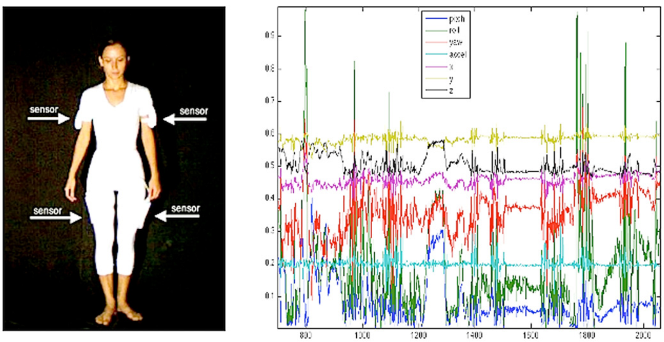


**Fig. 3:** Sensors placements on the dancer's body (left). Excerpt of gestural time-series retrieved from one performance (right).

The PD model that recorded the body movements received, through OSC protocol, seven streams of data, at the rate of one floating numbers each 10ms, for each stream. We attached the nunchucks at the dancer's elbows – once they are lighter than the wimote – and the wiimotes at her knees. In Figure 3, they denoted by the generic term "sensor." At the right side of this figure is depicted an excerpt of the real-time data generated by one of these sensors, for one of the dancer's performances: an improvisation over one of the body action movements described in Table 1.

Using these sensors, it was possible to collect gestural data out of the dancer's movements, wirelessly and in real-time. The dataset excerpt shown in Figure 3 (right side) represents only one of the four sensors, compounded of 7 synchronous streams of data: three rotation angles: pitch, roll, yaw; three raw accelerations: x, y and z; and one general acceleration activity. They were mapped to create six arrays, corresponding to the lines of the gestural matrix M, which corresponds to the first individual genotypes in the Population set, at the beginning of its evolutionary process. Compounding the genotype, these gestural matrixes were fed with data collected from one specific body action recording. In this implementation, the mapping translated the mean variation of each sensor's parameters – the ones attached to the dancer's arms – for the Tonal Intensity and Tonal Frequency curves; and the ones attached to the dancer's legs, were translated to Stochastic Intensity and Stochastic Frequency. Tonal Distortion was created by the difference between Tonal Intensity and Tonal Frequency. Stochastic Distortion was given by the difference between Stochastic Intensity and Stochastic Frequency.

## Computational Implementation

We implemented an ESSynth model in PD because this is an open-source, multi-platform, visual programming environment, specially designed for the implementation of real-time multimedia data processing (www.puredata.info). A program developed in PD is a patch. This resembles a data-flow structure, made with interconnected objects, message boxes, number boxes, symbol boxes, and so forth. An interesting feature of PD is the possibility of developing patches that can create and control other patches. This touches the paradigm of meta-programming, in which programming code can automatically write new code, without human supervision. There have been recent efforts to develop objects better acquainted for the meta-programming, such as the iemguts library, currently under development and written by IOhannes Zmölnig that aims to computationally emulate the behavior of self-aware agents in a system (ZMÖLNIG, 2008). Nonetheless, PD is already capable of, introductorily, exploring the automatic generation of patches by other patches, which is particularly useful for the implementation of adaptive systems, such as ESSynth.

The Evolutionary Sound Synthesis, originally introduced by the group of Jônatas Manzolli (2001) is a computational method for sound synthesis based on principles of the Evolutionary Computation. In this work, we developed a new version of ESSynth with some new features. Instead of being fixed-size, here the Population set is variable-size. The evolutionary process can start with few Individuals, whose Genotypes were mapped from a select group of drawings. This eliminates the evolution is Generation steps. In turn, the Individuals here must have a lifespan. In this current implementation, offprings are begotten by pairs of genderless predecessors. Individuals that reach their lifespan time are eliminated from the Population set. The system erases their Genotypes, so it will never be repeated (which would create a clone). While alive, Individuals will generate sound objects whose perceptual identity is defined by their Genotypes. The sound output of this version of ESSynth is given by all sound objects generated by the coexistence of all Individuals alive, within the Population set. In this implementation, there is no Selection process. Individuals are only eliminated when their lifespan is over. New individuals are created by the Reproduction process. This behavior emphasizes the system capacity of generating self-organized sound structures.

Figure 4 shows the implementation of the Reproduction process in PD. There are twelve tables depicted on the top and another six on the bottom. The ones on the top represent the parents' Genotypes. The ones on the top left are the genotype of the first individual. The one on the top right side corresponds to the second individual. These genotypes are the mappings from the gestural data (drawings or dance), that are stored in the gestural matrix. Both gestural data types have the same genotypical organization of 6 tables for each Individual Genotype, as described in section 3.2. There are two sliders on the right side of this figure. They represent the rates for crossover and mutation. They determine how much recombination (crossover) and variance (mutation) the offspring genotype, shown in the 6 tables at the bottom, will have.

In its first version, ESSynth output was given by a queue of best Individuals (the waveforms selected by the Selection process) given by the Hausdorff distance between Population and Target set. In this current implementation the sound output of ESSynth is synthesized by all Individuals in the Population set. They act as agents in a self-organized system. The sound output is compounded by all unique Individuals coexisting within this variable-size set. There is a chance that the Population set reduces in size until all Individuals are extinct. On the other hand, there is also the chance that the Population size grows so much that the ESSynth computer model crashes. To avoid that, the system has minimum and maximum thresholds for the Population size, where it can vary in between.

Reproduction process creates a new Individual through the Genetic Operators: Crossover and Mutation. Crossover creates a new offspring Genotype from the recombination of a copy of the Genotypes of its parents. Mutation inserts variability, through small bits of random information, in the offspring Genotype. These genetic operators are implemented in the PD subpatch called "pd genetic operators," as seen also in Figure 4. This sub-patch receives two coefficients, given by the pair of vertical sliders, at the right side of this figure. These are variables normalized from zero to one, where 0 means "no operation," and value 1 means "full" operation. They can be changed on-the-fly, during the ESSynth processing.

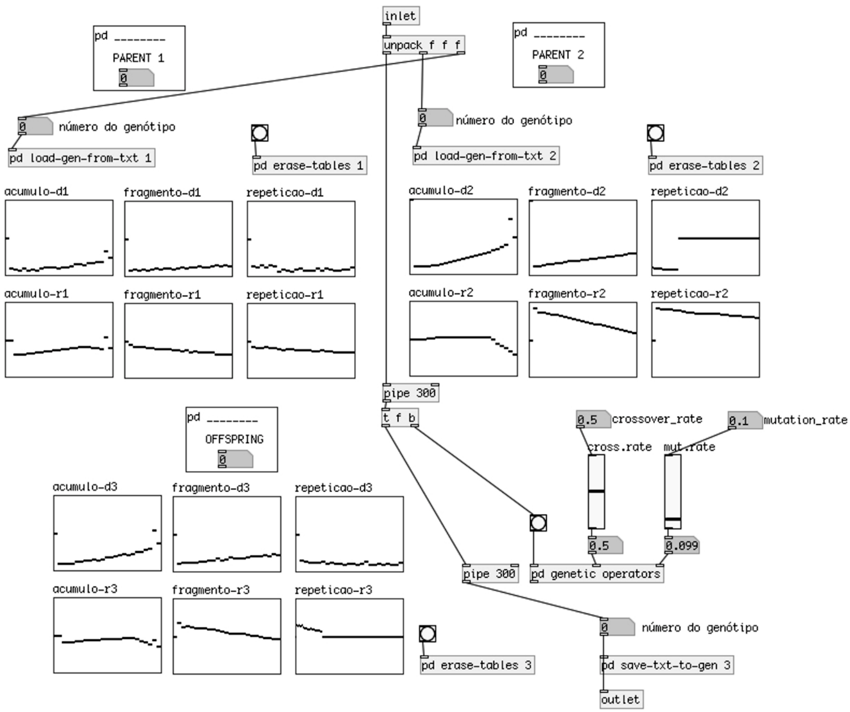


**Fig. 4:** The Reproduction process, as implemented in PD.

## Results and Discussion

The sonic results produced by the Evolutionary Sound Synthesis (ESSynth) implementation, here presented, are based on two mappings using different types of gestural data: processual drawings and body movements of dance. An important aspect to consider here is the degree of similarity between the drawings from the collection, in comparison with the dancer improvisations on Laban's body actions. The drawing collection is noticeably more self-similar than the body actions. The purpose of the artist, during the creation of the drawing collection, was to reach a "perfect gesture"; the one that would be registered in a final drawing and express all the evolutionary path of gestures that composed the whole collection. Given that it is impossible to draw by hand two identical sketches, it follows that there would never be two identical mappings, which implies a Population set without clones; although these individuals are strongly related. In contrast to that, dance body actions did not aim for similarity. To deal with such context, Rudolf Laban developed a system to represent distinct movements of dance. Movements, supposed to be unique that were defined as dimensions, would represent any dance movement, somehow working like the orthogonal coordinates of the Cartesian space.

Here, these dimensions were mapped into sonic genotypes of the Individuals Population set (see Table 1). That is why the Individuals whose Genotypes came from Dance Gestures were more heterogeneous than the Individuals whose Genotype came from Drawing Gestures. This is observed in Figure 5, which presents two sonograms taken from audio excerpts, recorded from the sound output created by these two types of gestural data. It is possible to observe that the left sonogram generated by dance movements produce Tonal sounds, in contrast with the right sonogram, were the gestural drawings engender Stochastic sounds. These examples can be heard at the following link: http://www.4shared.com/dir/GXH0h6_G/JNMR2011.html.

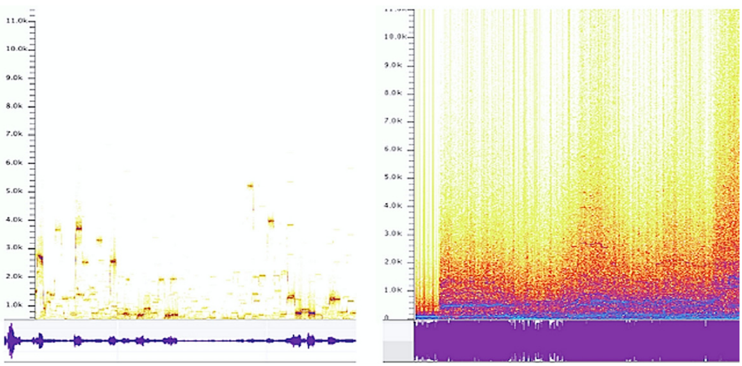

**Fig. 5:** Sonogram (top) and waveform (bottom) of two audio recordings from sound outputs generated by dance movements (left) and drawing mappings (right).

It is interesting to observe in Figure 5 that the audio excerpt generated by the ESSynth fed with gestural data retrieved from dance movements presents more variation of intensity (dynamics) than the one generated by the drawings, which almost seems as having been compressed. Dance movements produced a sound texture with cumulative frequency concentration, while the drawings created dense and spread sonograms. For each drawing image, the mapping algorithm automatically found about thirty different graphic objects: one Cumulation and many Repetitions and Fragments. The mapping process took all of them into consideration, to create the Genotype. This seems to have enlarged the sonogram's density, as seen in the sonogram at the right side of Figure 5. Dance gestural data, on the other hand, was constituted by 24 time-series, generated by 4 movement sensors (accelerometers). They described near-to-periodic movements, which resemble sinusoidal waveforms. Thus, the movements tended to be acoustically represented as quasi-periodic sounds. That corresponds sound objects more deterministic (with pitch) than stochastic (without pitch). This is observed in the sonogram at the left side of Figure 5.

When comparing these two distinct results, the drawings and the dancer's movements emerged as two distinct sonorities, which can be seen as different CAS attractors (HOLLAND, 1992). These two aspects show different forms of representing adaptive gestural processes, in the Evolutionary Sound Synthesis context. Therefore, these two types of sound objects converged to sound realms with distinct sonic features. According to Varèse's viewpoint, the concept of organized sound is about arranging together timbre and rhythm in musical structures (BERNARD, 1987). We understand the self-organized sonic results of our system as transcending Varèse's concept, where those arrangements are here extended by the emergent behavior of attractors.

## Conclusions

A new generation of artists has applied Evolutionary Computation to emulate creative processes in computers (BENTLEY, 1999). The Evolutionary Sound Synthesis method presented here can be seen as a process of this sort. It integrates a Processual Artwork and a real-time Computer Music system, by means of Gestural Data. This method can be used to indefinitely extend, in the sound domain, the process of artworks already finished, such as a collection of drawing or records of dance movements. This is especially useful for Multimedia Art Installations. The gestures of the artist, imprinted in a physical medium (such as drawings on paper), can now be retrieved as Genotypes of a Population of Individuals, altogether forming a new, never-repeating, never-ending evolutionary sonic process, that converges to a self-similarity, thus providing reminiscent traces of the previous processual artwork. This is like bringing back to life, in sonic media, a processual artwork already over.

This current implementation does not have Target set, or Selection process. However, it does have a variable-size Population set. Here, the evolution process is adrift, with Individuals having limited lifespan, forming dynamic sound objects that altogether could self-organize a sonic environment. In this implementation, we used gestural interfaces to collect data mapped – in offline – to Genotypes, for the Individuals in the Population set. Further developments may experiment with interfaces to control, in real-time, parameters from the evolutionary synthesis model, such as the genetic operator's rates, individual lifespan average, the population proliferation rate, or even the parameters within a Selection process fitness function. Another extension to be implemented is the process of energy intake, also known as "synthetic forager", where Individuals may seek out and compete for food. The concept of Individual multi-gender is also an interesting subject to be explored in further works. In this current version, there is no gender distinction between Individuals, although it takes two Individuals to create a new offspring. Further implementations may also have a childhood period, for the Individuals in the Population set. During such period, Individuals shall not be able to go through the Reproduction process, but may exchange and acquire information from other active ones, somehow resembling the childhood learning period. As seen, there is a myriad of interesting possibilities using this evolutionary method to create new computing models, for sound synthesis, processual artworks and installations – exploring multi-modality and interactivity, to reach together new immersive and adaptive sonic and musical experiences. [1]

## References

BENDER, E. A. *An Introduction to Mathematical Modelling*. Nova Jersey: John Wiley & Sons, 1978.

BENTLEY, P (Ed.). *Evolutionary Design by Computers*. San Francisco: Morgan Kaufmann, 1999.

BENTLEY, P.; CORNE, D. (Ed.). *Creative Evolutionary Systems*. San Francisco: Morgan Kaufmann, 2002.

BERNARD, J. W. *The Music of Edgard Varèse*. New Haven: Yale University Press, 1987.

BILES, J. GenJam: A Genetic Algorithm for Generating Jazz Solos. In: *Proceedings of Computer Music Conference*. San Francisco: [s.n.],1994.

[1] Acknowledgements: we would like to thank Tatiana Benone, for performing the body movements; the financial support of FAEPEX/ UNICAMP, that sponsored this project. Manzolli is supported by CNPq, Shellard and Fornari are supported by FAPESP and CNPq.

BIRCHFIELD, S.T.; GANGISHETTY, R. Acoustic Localization by Interaural Level Difference. In: *37th International Conference on Acoustics, Speech, and Signal Processing*. Philadelphia, March 2005, p. 1109-1112.

CRAMER, F. *Words Made Flesh*: Code Culture Imagination. Rotterdam: Piet Zwart Institute, 2005.

GONG, D., YAO, X.; YUAN, J. Interactive Genetic Algorithms with Individual Fitness not Assigned by Human. *Journal of Universal Computer Science*, v. 15, n. 13, p. 2446-2462, 2009.

FELS, S.; MANZOLLI, J. Interactive Evolutionary Textured Sound Composition. In: *6th Eurographics Workshop on Multimedia*, Sept. 2001, p. 153-164.

FLEISCHMANN, G. et al. SMDK: An Interactive Self- Organizing Sound Environment. *AAAI-94 Workshop, Artificial Intelligence, Artificial Life, & Entertainment, Working Notes*. Seattle, Washington, 1994.

FOERSTER, H. On Self-Organizing Systems and Their Environments. In: YOVITS, M. C.; CAMERON, S. (Org.). *Self-Organizing Systems*. London: Pergamon Press, 1960, p. 31-50.

FORNARI, J.; MAIA JÚNIOR, A.; MANZOLLI, J. Creating Soundscapes using Evolutionary Spatial Control. In: *Proceedings of the EvoMusarts*. Valencia, España: Springer-Verlag, 2007.

FORNARI, J.; MAIA JÚNIOR, A.; MANZOLLI, J. A Síntese Evolutiva Guiada pela Espacialização Sonora. In: *Anais do XVI Congresso da Associação Nacional de Pesquisa e Pós-graduação em Música*. Brasília: UnB, 2006, p. 274-279.

FORNARI, J. et al. The Evolutionary Sound Synthesis Method. In: *Proceedings of the ninth ACM international conference on Multimedia*. NY: ACM, 2001, p. 585-587.

GALENTER, P. What is Generative Art? Complexity Theory as a Context for Art Theory. In: *6th Generative Art Conference*. Milan, 2003.

HOLLAND, J. *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. Cambridge: MIT Press/Bradford Books, 1992.

HOLLAND, J. *Hidden Order: How Adaptation Builds Complexity.* New York: Helix Press Books, 1996.

HUSARIK, S. John Cage and Lejaren Hiller: HPSCHD, 1969. *American Music*, University of Illinois Press, v. 1, n. 2, p. 1-21, Summer 1983.

KELLY, J.; PHILLIPS, D. Coding of Interaural Time Differences of Transients in Auditory Cortex of *Rattus norvegicus*: Implications for the Evolution of Mammalian Sound Localization. *Hearing Research*, v. 55, n. 1, p. 39-44, 1991.

KIM, H.; CHO, S. Application of interactive genetic algorithm to fashion design. *Engineering Applications of Artificial Intelligence*, v. 13, p. 635-644, 2000.

KOSORUKOFF, A. Human-based genetic algorithm. In: *IEEE Transactions on Systems, Man, and Cybernetics*, 2001, p. 3464-3469.

LAVALLE, S. *Planning Algorithms*. Cambridge: Cambridge University Press, 2006.

MAIA JÚNIOR, A. et al. *ELEMENTARIDADES: Algorithmic Composition and Gestural Art performance*. Based on the elementar movements of subatomic particles. Elaborated by NICS (Núcleo Interdisciplinar de Comunicação Sonora) and GITD (Grupo Interdisciplinar de Teatro e Dança). Campinas, Brazil: UNICAMP, 2001.

MANCINI, M.; CASTELLANO, G. Real-time analysis and synthesis of emotional gesture expressivity. *2nd International Conference on Affective Computing and Intelligent Interaction*. Lisbon, September 2007.

MANZOLLI, J. et al. The Evolutionary Sound Synthesis Method. Short-paper. In: *9th ACM International Conference on Multimedia*. Ottawa, 2001, p. 585-587.

MORONI, A. *ArTEbitrariedade: uma reflexão sobre a natureza da criatividade e sua possível realização em ambiente computacionais*. Tese (Doutorado). UNICAMP, Campinas, 2003.

MORONI, A.; MANZOLLI, J.; VON ZUBEN, F. ArTbitrariness: Putting Computer Creativity to Work in Aesthetic Domains. *Advances in Logic, Artificial Intelligence and Robotics, Frontiers in Artificial Intelligence and Applications*, IOS Press, v. 85, p. 46- 54, 2002a.

MORONI, A.; VON ZUBEN, F.; MANZOLLI, J. ArTbitration: Human-Machine Interaction in Artistic Domain. *Leonardo Music Journal*, MIT Press, v. 35, n. 2, p. 185-188, 2002b.

PFORSICH J. *Handbook for Laban Movement Analysis.* New York: Janis Pforsich, 1977.

TODD, P. M.; WERNER, G. M. Frankensteinian Methods for Evolutionary Music Composition. In: GRIFFITH, N.; TODD, P. M. (Ed.). Musical Networks: Parallel Distributed Perception and Performance. Cambridge: The MIT Press, 1999, p. 313-339.

TODD, S.; LATHAM, W. *Evolutionary Art and Computers*. London: Academic Press, 1992.

WANDERLEY, M. et al. Non-obvious Performer Gestures in Instrumental Music. In: In: BRAFFORT, A. et al. (Eds.). *Gesture-Based Communication in Human-Computer Interaction*. Berlin: Springer-Verlag, 1999, p. 37-48.

WILSON, S. *Information Arts: interactions of art, science and technology*. Cambridge: MIT Press, 2002.

ZMÖLNIG, I. Pure Agents: Augmenting Live Patching. Piksel 08 Presentations. In: *Talks & Presentations at StudioUSF*. December 4-7. 2008.

**Jose Fornari** é Bacharel em Engenharia Elétrica (1991), Bacharel em Música Popular (1994), Mestre em Engenharia Elétrica (1994) e Doutor em Engenharia Elétrica (2003) pela Universidade Estadual de Campinas, UNICAMP. Possui Pós-Doutorado em Síntese Sonora Evolutiva (2006) e Cognição Musical (2007). Sua pesquisa está voltada a interface gestual, arte multimodal, psicoacústica, computação musical e computação evolutiva. tutifornari@gmail.com

**Mariana Shellard** é Mestre em Artes Visuais pela UNICAMP, com a pesquisa intitulada *A plasticidade do sonoro* (2010), tendo contado com apoio da FAPESP. Em parceria com José Fornari foi contemplada pelo programa de incentivo *Rumos arte cibernética* do Itaú Cultural (2009), com o projeto *RePartitura*. Participa do grupo de estudos *Arte & meios tecnológicos* (FASM/CNPq). marianashellard@gmail.com

**Jônatas Manzolli** é Bacharel em Matemática Aplicada Computacional (1983), Bacharel em Composição e Regência (1987) e Mestre em Matemática Aplicada (1988) pela UNICAMP; é Doutor em Filosofia pela University of Nottingham (1993). Atualmente é Livre Docente do Instituto de Artes da UNICAMP e Coordenador do Núcleo Interdisciplinar de Comunicação Sonora (NICS). Pesquisa a interação entre arte e tecnologia, entrelaçando criação musical, computação musical e ciências cognitivas. jotamanzo@hotmail.com