TEXTURAL COMPOSITION IN 3D ENVIRONMENT THROUGH SWARM ALGORITHM

Wander Vieira Rodrigues
NICS/Unicamp
Campinas, Brasil
wanvrod@gmail.com

Walber de Macedo Rodrigues
CIn/UFPE
Recife, Brasil
wmr@cin.ufpe.br

ABSTRACT

The presented project uses algorithms based on bird flock behavior to generate sound masses in a 3D environment through "Cave" and sound spatialization in 8.1 Surround. The behavior of animal groups in agglomeration makes us lose the individual's perception in isolation to force us into the whole. Part of this process is identical to listening and composing sound masses in the textural music 20th and 21st centuries. Thus, we want to apply the swarm behavior algorithm to generate sound masses in real time, allowing contemplation, creation of musical material, and interaction through digital immersion.

1. INTRODUCTION

This project deals with creating an environment for projecting images in three dimensions with musical interaction in 8.1 surround. In this environment, the user can act with digital swarms of flocking birds, associating each individual to sounds produced by a human agent immersed in a 3D chamber to generate sound masses in real time.

Textural Composition in 3D Environment through Swarm Algorithm starts from the following steps: 1 - immersion in digital space; 2 - generation of digital birds; 3 - sound capture, interaction, and overlay; 4 - sound and visual results. The developed environment is available at GitHub[1]

To develop this concept, we need to understand that the behavior of agglomeration of animals has called humanity's attention since ancient times, be it herds, fish shoal, swarms, or flocks of birds. From a visual point of view, the behavior pattern of these beings makes us lose focus on the individual animal and start paying attention to a larger body that is the agglomeration. This topic has been the focus of current researchers who strive to model these behaviors due to the complexity generated and efficiency because of the factors resulting from this process, such as low collision rates between individuals, protection from predators, and long-distance travel.

[1] https://github.com/walbermr/TEXTURAL-COMPOSITION-IN-3D-ENVIRONMENT-THROUGH-SWARM-ALGORITHM

[1] argues that a swarm is a large number of homogenous, simple agents interacting locally among themselves and their environment, with no central control to allow a global behavior to emerge.

[2] takes the topic further by focusing on the individual: A swarm particle is a stochastic optimization approach modeled on the social behavior of bird flocks. The PSO (Particle Swarm Optimization):

> Particle swarm optimization (PSO) is a stochastic optimization approach modeled on the social behavior of bird flocks. PSO is a population-based search procedure where the individuals, referred to as particles are grouped into a swarm. Each particle in the swarm represents a candidate solution to the optimization problem. Each particle through the multi-dimensional search space in a PSO system, adjusting its position according to its own experience and neighboring particles. [2, p. 49]

Commonly associated with the concept of Computational Intelligence or Artificial Intelligence, the application of these algorithms, according to [2] again, brings satisfactory results in telecommunications network optimization, graph coloring, quadratic allocation problem, clustering optimization, and structural optimization, or as also advocated by [1]:

> Swarm-based algorithms have recently emerged as a family of nature-inspired, population-based algorithms capable of producing low-cost, fast, and robust solutions to several complex problems. [1, p. 2]

This project is inspired by Multimodal generative installations, and the creation of new Art form based on interactive narratives, the one created and developed by [3]. However, unlike the Generative Multimodal Installation, which uses compositional processes and generative designs linked to a scope of mediated perception, our project uses sound samples produced through improvisation by an immersed human agent feeding the system based on the representation of a bird flock modeled algorithmically.

Practically speaking, the moment the individual is immersed in the Digital Interactive Chamber, he will be guided to produce sounds of his choice in a pre-established or random way, using a musical instrument, voice, body percussion sounds, electronic oscillators. He will be free to choose and not need any formal musical experience. Then the Interactive Chamber system will then start extracting sound samples at a predetermined time and associating or allocating them to each swarm element or virtual bird. After these birds begin to swarm and form a flock, the result will be a mass of sounds that will move around in the sound space in 8.1 surround, projecting the swarm's behavior into the digital environment.

To better understand our goal, we must understand computer modeling and sound mass concepts, Section 2, and Section 3, respectively.

2. SOUND MASSES

Sound Mass is a concept associated with 20th and 21st century Textural Music. [4]. His research dealt with the SM approaching two perspectives: 1 - that linked to orchestral music and 2 - that linked to electronic music.

The Sound Mass is not only a musical resource to compose a work in contemporary style, but it is also a musical aesthetic. From a certain point of view, it is linked to the concept of *canone* in traditional music, resulting in the creation of patterns of sounds that can be called Textural Sound Blocks, as described in [4].

[5] discusses the canon by identifying it as an element of an original musical cell copied through isomorphism or the transformation of preserved information. The author further states that this property reveals itself in a recursive process, in which infinite structures can be defined as being the "*expansion of node after node*".

[6] already starts from the point of view that Musical Texture is based on the idea of sound clouds and argues that the sequence combination of these clouds in the mesostructure induces a process of statistical evolution.

He also states that a synchronous cloud generates metric rhythmic sequences at a grain density rate of 2 to 20 occurrences per second. When the initial density is not the same as the final density, but the variation of this rate is continuously controlled and increasing/decreasing, such a procedure generates an acceleration or deceleration effect. Otherwise, the initial and final density variation produces an irregular acceleration or deceleration for an asynchronous cloud. The cloud must be at a density above 100 grains per second to create continuous sounds.

[7] emphasizes these two distinct ways of using granular synthesis: synchronous and asynchronous, and in granular synthesis, the spacing of the sound grains radically modifies the generated texture. In the synchronous process, the grains are separated in the same time interval involving a particular linear relationship, while in asynchronous granular synthesis, the grains are inserted into the texture without a strictly linear relationship.

> This parameter specifies the temporal distribution of the grains, either synchronous or asynchronous. Synchronous generation means that one grain follows another in series, with the distance between the grain attacks determined by the density parameter. [...] Asynchronous clouds scatter grains at random time points within the specified boundaries of the cloud, at the specified density. Asynchronous clouds produce random and explosive effects. [6]

In a nutshell, we propose the shift of the analytical perspective from the score and orchestra to the sensory and cognitive process of the Sound Mass in the Textural Music *cf.* [8, 9, 10, 11].

Looking at our project through this lens, this means that for each bird that will make up our digital flock, we will have a speck and that the accumulation of birds will generate a cloud. Thus, we can determine a set of sound samples obeying asynchronous patterns. However, we still have another argument that starts from the sounds found in nature and is part of our sound universe. From a perspective that arises from the sounds found in nature that are responsible for our hearing development, [4] also defends:

> We can, then, relate the perception of individuality and sound grouping to the auditory experience when listening to the rain. The isolated sound of a drop of water hitting a surface, floor, or puddle becomes clear and recognizable as a single sound source. In this way, we can identify the sound and pinpoint its direction in space. However, as the number of drops hitting these surfaces increases (randomly), we stop perceiving them as drops and begin to relate them to the sound phenomenon of rain (which we can consider a sound mass). [4, p. 40]

To prove this point, the author [4] created an experiment associated with the phenomenon of rain that involves stochastic processes and gave a clear direction on the principles governing the formation and concept of Sound Masses.

By applying these concepts in our Digital Interactive Chamber, we will create the conditions for these principles and mechanisms linked to the Sound Mass Music, or Textural Music, to be incorporated from the behavior of flocks of birds since it has implicit models of stochastic processes.

We intend our interactive system to allow the individual immersed in the Cave to interact with birds flying through computer simulation and by supplying each bird with sound information using a musical instrument, its body percussion sounds, synthesized sounds, voice, or any other sound source, through prearranged or improvised ideas in real-time. Each bird will then be responsible for capturing a sound, or groups of sounds, with a time frame ranging from 2 to 5 seconds. After the sounds are captured, each individual bird will fly and seek a similar one to form a virtual flock. Consequently, the swarm's movement in space will create a mass of digital birds and several providing Sound Masses that will behave similarly to the spatially and dynamically arranged flock.

3. MATHEMATIC MODEL

Our objective is to create an immersive three-dimensional graphical environment through Unit3D. After the participant is positioned in the center of the Cave, the system will automatically perceive his presence and guide him in the interaction to generate sounds related to the bird swarm. As the system is gradually fed, there will be a waiting time for each sound produced to start the displacement of birds, implying that the environment will be fed by capturing sounds. The captured sounds will be incorporated into a particular bird of the flock, generating movement, and flocking simultaneously and gradually agglomerating based on the mathematical and computational modeling implemented. The established principle expects the captured sounds to have the same behavior as the song or call for each bird in the flock in space, keeping in mind that the captured sound samples will have variable durations and occur at the participant's pleasure. From the musical point of view, which is different from the sound mass compositions in G. Ligeti, W. Lutoslawsky, and others, instead of a static Sound Mass predicted and projected in the score, a dynamic Sound Mass is created in real time.

The user immersed in the Cave will be able to guide the swarm and generate more sounds based on the mass he or she has created or even interact with other people in the same environment. The Cave will work as a form of improvisation for experienced musicians or participants with some musical experience. Whereas, for amateurs, it will work as a playful experience that can provide a musical moment different from the one traditionally offered.

For the algorithm representing the flock of birds, we will have, as described by [12], the following mathematical model.

A *flock B* of birds can be represented by a set containing n birds b_i.

$$B = [b_i, i = 0,1,\dots,n-1] \tag{1}$$

(1) shows the definition of a flock.

Each bird has two randomly generated attributes and three related forces. The attributes are the position p and speed $\vec{v}$. Its forces are separation, cohesion, and alignment.

To visually recognize and discriminate the elements around the bird (isolated element) in the simulation, consider the birds' field of view. In this model, it is assumed that all birds have the same field of vision, and for this reason, the field is not a particular attribute of each individual bird. Since the two attributes of the birds (position and speed) vary during the simulation, this field of view will be represented by a sphere of radius e and the set of birds seen by another bird that is V_i, defined in Equation (2).

$$V_i = \{b_j \in B; \forall b_j: |p_i - p_j| < e, j = 0,1,\dots,n-1; j \neq i\} \tag{2}$$

As Equation (2) shows, we have a set of birds within the bird's field of view.

Three main movements in a flock of birds can be observed: 1- separation, 2 - cohesion, and 3 - alignment. When combined and applied to each bird, these forces should guide them so that the flock does not disperse.

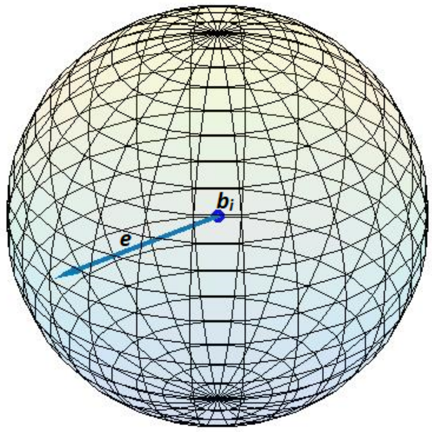

Figure 1: Field of View of the Bird b_i with radius$e$. Designed using the software MATLAB.

The separation force is essential because it prevents collisions between birds in the group. With its field of vision, each bird can identify other birds in the surroundings and estimate the movement needed to avoid a collision. This movement is called the separation movement, defined by the vector $\vec{s_i}$, and is the inverse direction of the sum of the vectors between the bird positions b_i and the birds b_j around them. Equation (3) defines the separation motion vector.

$$\vec{s_i} = -\sum_{\forall b_j \in V_i} p_i - p_j \tag{3}$$

The cohesive force $\vec{k_i}$ is responsible for the non-dispersion of the flock during the separation movement. It is defined as the vector between the bird b_i position, and the center of density c_i of all the other m birds visible by b_i. Equation (4) defines the center of density of visible birds, and Equation (5) shows the vector referring to the cohesion movement.

$$c_i = \sum_{\forall b_j \in V_i} \frac{p_j}{m} \tag{4}$$

$$\vec{k_i} = c_i - p_i \tag{5}$$

The force of alignment $\vec{m_i}$, defined in Equation (6), is what ensures that all the birds in the flock maintain similar speeds. This is observed in nature through the tendency of bird flocks to synchronize their speeds and stay close for extended periods. For a bird b_i, the alignment force is calculated using the average speed of the other birds within its field of view.

$$\vec{m_i} = \sum_{\forall b_j \in V_i} \frac{\vec{v_j}}{m} \tag{6}$$

Combining these three forces ensures that the flock maintains its unity, with close speeds between birds and a low number of collisions between birds. This combination results in the updated speed $\vec{v_i}'$ of a bird. Each force has a related coefficient that represents the share of that force in the final velocity. The updated velocity is defined in the Equation (7), shows here the vector referring to the updated speed of the bird.

$$\vec{v_i}' = \vec{v_i} + S \cdot \vec{s_i} + K \cdot \vec{k_i} + M \cdot \vec{m_i} \tag{7}$$

The next position p'_i that a bird b_i will assume is calculated by adding your current position to your speed multiplied by a time interval Δt. Equation (8) shows the point referring to the updated position of the bird.

$$p'_i = p_i + \Delta t \vec{v_i} \tag{8}$$

Every sound sample on our system must be a periodic signal, represented by a function x in the time domain t, represented by the function $x(t)$. Signal periodicity is defined as any point t in time to have a repetition period T_0, causing $x(t) = x(t + kT_0)$ where k is any integer.

4. MATHEMATIC STATISTIC

The swarm design has three main features, 1) acquisition of a sound sample that is repeated intermittently, 2) several samples are played together, and 3) the samples are arranged in *boids* that behave according to a bird swarm algorithm giving spatialization and providing clutter to the sounds.

Taking a sound sample y. any a period of T_1 seconds, where T_1 is an integer, will project on the swarm a sample repeated for an infinite time or interruption of the environment's execution. Thus, the sound signal will be set for the entire time dimension every time T_1 seconds in which it repeats. This observation fits the definition of a periodic signal, so the sample is periodic at kT_1seconds, where the k is an integer number.

In the theory of Signals and Systems the sum of two periodic signals, with periods T_0 and T_1, is also a periodic signal since the ratio $\frac{T_1}{T_0}$ be a rational number $\frac{M}{N}$ where M N are non-divisible integers.

As new sound samples are inserted into the swarm, the existing periodic sounds are added together to create a different sound signal that is also periodic. This other unreleased signal is periodic since the periods T_1 ad T_0 are integers, and division between two integers does not yield rational numbers. The period of this new sign is defined as $\frac{T_1}{T_0} = \frac{M}{N} \rightarrow MT_1 = NT_0$, that is to say, $T_2 = MT_1 = NT_0$. Each new signal added will always respect this rule, generating new periodic signals.

5. THE LISTENING PERSPECTIVE

The Project in question here deals with sound not as a pure matter dissociated from an artistic context but as a sound construction molded in the function of materials that carry minimalist traits and are allocated, superimposed, and spatialized by the algorithm of agglomeration of animals, specifically birds. At the same time, there is no intention to use or sculpt musical structures as traditionally established. Our idea is built on the concept that it is not the sound that generates the beauty we associate with a landscape of birds in flight but the condition that each bird producing sounds and agglomerating creates a sound and visual phenomenon that has acoustic impact and generates the beauty, or any reaction, that has always caught the attention of humankind.

[13] launches the perspective of sound agglomeration from the point of view of rain and the nuances between the sonorities implicit in it, as well as the different ways we perceive and interpret it (In [4], we took a little far in depth the issue of similarity between sound structures such as rain and sound mass). Furthermore, both rain and other phenomena of nature are responsible for forming our listening. [14] has a similar approach when dealing with the relationship between sound perception, natural phenomena, and the relationship with the development of our listening. He states:

> The first reason comes from the observation and discovery of rare or everyday sound phenomena that nature or society offer us. Thus, for example, during summer in the countryside the song of the cicadas, which rocked humanity and its poets, imposes itself and invites us to join the structure of this event, if only out of childlike curiosity "how it's made" that every adult keeps for the happiness of humanity. Then comes the desire to reconstitute a similar event, no longer with cicadas but with other sound means, with orchestral instruments or with machines. And this desire goes so far as wanting to modulate, according to invention, the sound event inspired by the song of thousands of cicadas. How to get there? By appealing to the internal logic of this natural phenomenon [14].

Furthermore, it seeks a mathematical understanding of the phenomenon:

> This logic is that which also governs the movements of the molecules of a gas and physics has already given its answer in the "kinetic theory of gases", which is a stochastic probabilistic theory. If the observation of the song of cicadas had been made by a mathematician composer before the creation of the "kinetic theory of gases", it is certain that the study of the song of cicadas would have led him to the discovery of similar laws. Once the abstract structure of the mass event made up of thousands of elements has been defined, it can be used to condition a mass of pizzicati or any occasional bowing of the classical string orchestra." [14].

[13] speaks of the lo-fi soundscape due to sound congestion or saturation resulting from the emergence of new sounds. He states:

> The Industrial Revolution introduced a multitude of new sounds with unhappy consequences for many of the natural and human sounds which they tended to obscure; and this development was extended into a second phase when the Electric Revolution added new effects of its own and introduced devices for packaging sounds and

> transmitting them schizophrenically across time and space to live amplified or multiplied existences [13].

And continues:

> Today the world suffers from an overpopulation of sounds; there is so much acoustic information that little of it can emerge with clarity. In the ultimate lo-fi soundscape, the signal-to-noise ratio is one-to-one, and it is no longer possible to know what, if anything, is being listened to. This, in summary, is the transformation of the soundscape [13].

Nowadays we are facing a panorama that offers the possibility of a better understanding of what gave rise to the consolidation of an aesthetic that breaks with traditional concepts of music and focuses greater attention on sonority.

Based on this perspective, we intend to offer the participants of our immersive project a sound experience that will be built based on the interaction with bird flocking algorithms. In our system, the sound material produced will result from the resources available to the human agent, be it an entirely non-musically literate person or even an excellent and experienced musician. The experience derived from this digital medium will allow a different vision of what the immersed human agent is used to observing from his own experience in his daily life and sound experiences, composing sound materials with conceptual support in the masses of sounds coming from the music of the 20th and 21st centuries.

6. IMPLEMENTATION OF THE SYSTEM

The project presented here was initially developed until the implementation phase and the first tests.

Using the Cave set up at the NICS/Unicamp lab at Campinas/São Paulo, we were able to launch the projection of the Boids and perform experiments with sounds produced by agents immersed in the digital environment. In this initial phase, we did not build the birds' bodies. We used spheres as representatives of each bird of the flock, shown in Figure 2:

Figure 2: photo taken from a *boid,* created in the system developed for sound mass composition.

Before deployment, the system was previously executed on a computer using three screens, then projected onto a larger screen to generate the 3-D environment.

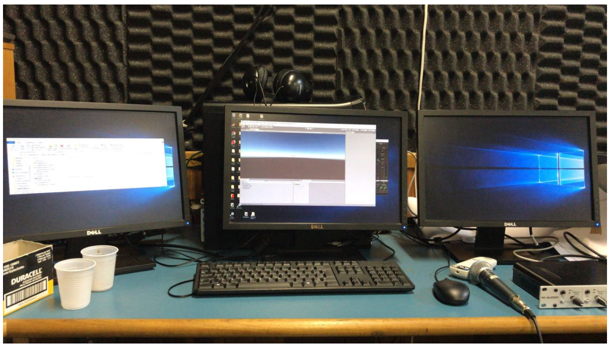

Figure 3: Screen computer where the 3D plot was generated.

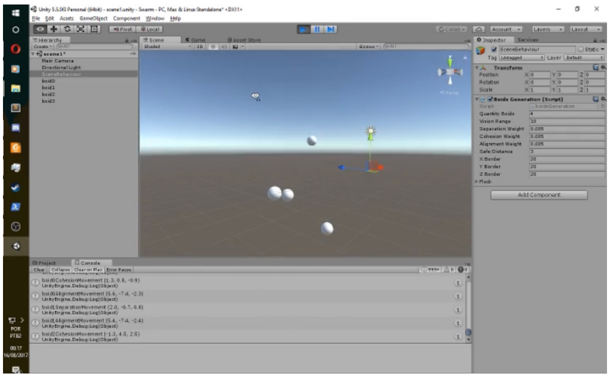

Figure 4: Projection screen in Unity 3D.

Figures 2, 3 and 4 present the system running on the computer screen. Figures 5 and 6 show the Cave.

Figure 5: Projection Screen of the digital Cave.

Next, we have the projection of the swarm on the Cave:

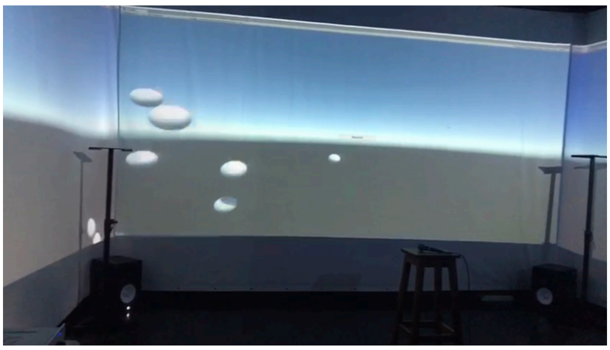

Figure 6: The Cave working with images and sounds with surround 8.1.

In the presented projection we can observe a group of 8 elements that had their sound representation agglomerated into images and sounds within the principles of sound mass.

7. CONCLUSION

The exposed project obtained tangible results from both points of view: visual and sound experience. We worked with sounds generated with body percussion sounds, clapping, hissing, and musical instruments such as the xylophone during this project phase.

The process of sound mass generation, that is, the loss of the perception of the sound unit for the agglomeration in an extrapolation of understanding in a larger dimension, was perceived as more sound elements were generated and composed of the swarm.

The unfolding of this project involves:

- The composition of melodies by collecting melodic, harmonic, and sound masses.
- The experiment may give rise to an ecological approach to birds for students.
- Analysis of the sound extrapolations resulting from the interaction in the digital environment.
- Music/Sound Installation to be presented in artistic presentations.
- The neuroscientific study of the agglomeration process of sound and image. This study is currently in progress but without publications and looking for partnerships.

In this way, we intend to advance with the research and seek partnerships to disseminate and improve the idea.

8. REFERENCES

[1] H. Ahmed and J. Glasgow, *Swarm Intelligence: Concepts, Models and Applications*. School of Computing Queen's University, 2012.

[2] A. P. Engelbrecht, *Computational Intelligence: An Introduction*. John Wiley & Sons, 2007.

[3] J. Manzolli. Multimodal generative installations and the creation of new Art form based on interactivity narratives. *XVIII Generative Art Conference*, [Online] Available: https://www.gentiveart.com/ga2015_WEB/Multimodal-Installation_Manzolli.pdf

[4] W. R. Vieira, Composição Textural Assistida por Computador. PhD Thesis, State University of Campinas (Unicamp), Campinas, 2016.

[5] D. R. Hofstadter, *Gödel, Escher, Bach: an Eternal Golden Braid*. Baic Books, 1999.

[6], C. Roads, *Microsound*. MIT Press, 2001.

[7] T. Opie. Sound in a NutShell: Granular Synthesis, An overview of granular synthesis and the techniques involved. [OnLine] Available: http://www.granularsynthesis.com/hthesis/hthesis.html.

[8] R. Bader, Berechnung Fraktaler Strukturen in den Etüden für Klavier von György Ligeti, *Musikwissenschaftliches Institut Hamburg*. 2003.

[9] M. Chemillier, György Lgeti et la Logique des Textures. *Analyse Musicale*, n. 38, 2001, pp. 75-85.

[10] J. Clendinning, Jean Piper. The Pattern-Meccanico Compositions of György Ligeti. *Perspectives of New Music*. v. 31. 1993, pp.192-234.

[11] A. M. Roig-francoli, Harmonic and Formal Processes in Ligeti's Net-Structure Compositions. *Music Theory Spectrum,* v. 12, n. 2, 1995, pp. 242-257.

[12] W. R. Craig, Flocks, Herds, and Schools: A Distributed Behavioral Model. *Computer Graphics 21 (4)*, 1987. pp. 25-34.

[13] M. Schafer, *Our Sonic Environment and the Soundscape: The Tuning of the World.* Destiny, 1994.

[14] I. Xenakis, *Formalized music: Thought and mathematics in composition*. Pendragon Press, 1992.