Self-Organizing Bio-Inspired Sound Transformation

Marcelo Caetano[1], Jônatas Manzolli[2], Fernando Von Zuben[3]

[1] IRCAM-CNRS-STMS 1place Igor Stravinsky – Paris, France F-75004
[2] NICS/DM/IA - University of Campinas – Brazil PO Box 6166
[3] LBiC/DCA/FEEC – University of Campinas – Brazil PO Box 6101
caetano@ircam.fr, jonatas@nics.unicamp.br, vonzuben@dca.fee.unicamp.br

Abstract. We present a time domain approach to explore a sound transformation paradigm for musical performance. Given a set of sounds containing *a priori* desired qualities and a population of agents interacting locally, the method generates both musical form and matter resulting from sonic trajectories. This proposal involves the use of bio-inspired algorithms, which possess intrinsic features of adaptive, self-organizing systems, as definers of generating and structuring processes of sound elements. Self-organization makes viable the temporal emergence of stable structures without an external organizing element. Regarding musical performance as a creative process that can be described using trajectories through the compositional space, and having the simultaneous emergence of musical matter and form resulting from the process itself as the final objective, the conception of a generative paradigm in computer music that does not contemplate *a priori* external organizing elements is the main focus of this proposal.

Introduction

The digital computer allows great flexibility in sound processing. As a consequence, the spectrum of possibilities is so vast that the exploration of the full musical potential of the digital computer has become a major problem in computer music. Many different approaches have been proposed in order to create aesthetically interesting music for composition and performance, with results that vary from the unexpected to the undesired, depending upon a vast number of factors and on the methodology itself [7]. Musical sounds are complex and hard to synthesize, since they generally have a dynamic spectrum, such that each partial frequency has a unique temporal evolution envelope. Our ears are highly selective and frequently reject mathematically perfect and stable sounds [7]. Traditional sound synthesis techniques are limited, especially because they do not take the dynamic and/or subjective nature of music into consideration, using deterministic processes that were not specifically designed for sound manipulation [9].

In the perspective of this work, music composition and performance are creative processes that can be described using trajectories through soundspace that create musical matter and form by means of a self-organizing generative and organizational process. Self-organizing systems exhibit emergent properties and promote complex patterns from simple rules. Here, self-organization is the principle used to generate, organize and structure sound material during musical performance, guiding the composer through the compositional soundspace.

The use of bio-inspired or AI based systems for musical composition, performance, and free improvisation has already been suggested independently by several researchers as a way to permit more flexibility in the aesthetical exploration of the

resulting compositional space. The applications of bio-inspiration in composition involve artificial neural networks [11], cellular automata [6], artificial immune systems [7], particle swarms [2], ant algorithms [14], and evolutionary computation [1]. Nevertheless, most of these systems use MIDI codification (i.e. representation of musical events rather than musical sounds), consequently, they generate only form using bio-inspired algorithms. On the other hand, Miranda's CAMUS [17] and Blackwell's Swarm Granulator [3] extract parameters from the agents allowing the generation of matter by means of granular synthesis, as well as form, represented as the dynamic state of the system. Independent works have proposed the use of genetic algorithms [8], artificial neural networks [9], and artificial immune systems [7] in a sound synthesis technique in time domain, highlighting the profitable aspects of each. The user is enabled to find candidate solutions that meet certain musical requirements by using a set of waveforms (attractors) as examples of the desired sound qualities, instead of describing the sounds using numerical parameters or any other linguistic tool. Developing this idea one step further, we are proposing here that in self-organizing systems with emergence of sound patterns, the synthesis process is integrated into the performance process so that as the system generates emergent patterns, this material constitutes sonorities for musical performance. By regarding self-organizing bio-inspired algorithms under the general umbrella of self-organization, the method allows the composer to express a certain degree of subjectivity by simply choosing the algorithm and setting the parameters adequately, according to aesthetical preferences that can be reflected in the particularities of each candidate paradigm.

The next section explains briefly the conceptual foundations of our method, followed by an introduction to the bio-inspired paradigms chosen to explore the compositional space in this study, trying to emphasize that although they all share the same high level organizational principle, that is, self-organization, the application of different approaches results, in the case of this work, in different trajectories through the space, and therefore, in different sound transformation procedures. This feature of the proposal is exemplified in the following section, where an example is used to illustrate the differences and similarities that are considered most relevant. Finally, the conclusions and future perspectives are considered.

Conceptual Foundations

In this section we present general ideas behind our model that lay the foundations of the populational search through soundspace in time domain. Here we show how time-domain codification results in a search space that approximates the sonic continuum. The aim of this section is to show that trajectories in the search space correspond to musical gestures, or manipulations of the sonic continuum.

Sound Transformations as Trajectories

At first, in traditional western music, the usual musical parameters are: frequency, time, intensity, and timbre. Often, these parameters are traditionally described with discrete values. Following this idea, musical composition can be understood as the organization of a finite set of notes, rhythms, timbres and different intensities using a summative notation (compass) and fixed durations [21]. Our approach is an attempt to escape the traditional musical lattice. The use of the computer as a musical instrument

permits to go beyond the discrete, not being limited by its finite set of possibilities, reaching the sonic continuum and the transformation concept. A trajectory in this space represents a gradual transformation in one or more musical dimensions, that is, a gesture [21]. In this space, frequency is a real variable, as is time and intensity. Traditionally, timbre is described by a discrete set of values, since each different acoustical instrument characterizes a specific timbre and there is no middle point between two or more instruments. However, in the continuum it is possible to obtain sounds with characteristics originated from multiple instruments by just setting a position between two or more instruments in the dimensions representative of their timbral qualities.

Self-Organization as Compositional Paradigm in Computer Music

Self-organization can be summarized as a process where a global pattern emerges from multiple interactions of simple system components. In this sense, the rules that specify components interactions only use local information, without any reference to a global pattern. It is in the emergence of structures and organization that are not imposed from outside of the system and that were not pre-established that lays the essence of self-organization [10]. In self-organized systems, properties that cannot be understood by examining the properties of the system's components arise from interactions among them [20]. The decision-making and emergent pattern processes of these systems are often non-intuitive due to the large number of nonlinear interactions involved. For this reason, mathematical models and computational simulations provide useful techniques for the study of self-organizing systems and for the exploration of the consequences of the myriad interactions among component subunits. Many of these simulations are agent-based, i.e., each individual subunit is monitored during the simulation, and its behavior over time is determined through local interactions with other subunits and local cues from the environment. This is precisely the case of this work, where each sound wave characterizes an agent from a population searching through soundspace. Agents interact locally and it is the resultant set rather than isolated individuals that represents a particular region of interest of soundspace, characterized by the attractors.

Codification and Resultant Search Space

Figure 1 illustrates the codification used in this work and the associated transformational mappings involved. Each individual (agents and attractors) is codified as N samples of a waveform sampled at a rate SR, as shown in part a). Each of these samples is interpreted as a vector component. Therefore, each individual is a vector with N dimensions. Part b) represents the two-dimensional Euclidean space with a pictorial example of the distribution of a set of sounds interpreted as vectors in this space. This is only an illustrative example in which N=2. The sounds are manipulated by the algorithms in this space. There is a corresponding mapping to soundspace since each sound has duration, pitch (for harmonic sounds), timbre and dynamics. The generative process is understood as manipulations of this space, i.e., trajectories in the Euclidean search space correspond to articulations in one or more dimensions in the correspondent continuous sound space.
Smalley [19] declared that the information contained in the frequency spectrum cannot be separated from the time domain since "*spectrum is perceived though time*

and time is perceived as spectral movement". So, once the user specifies the waveforms (attractors), he is also specifying the spectral content and sound qualities of the tones. Our objective is the attainment of results that correspond to gestures through temporal representation and manipulation of tones in order to allow musical performance. Each bio-inspired approach selected to guide the search through Euclidean space does so in a different fashion. They all share some basic characteristics, though, as follows. The agents of the system and the attractors characterizing the space belong to the same vector space, i.e., they have the same number of samples (dimensions), which corresponds to the same maximum duration.

   The agents are initialized in points in the space that do not correspond to the attractors, meaning they are different sounds. Upon starting the simulation, the agents are drawn towards the regions containing the attractors. When one follows the trajectory described by each individual agent in pursuing the attractors, it can be perceived as a gradual transformation from the starting sound to the attractor (also a sound), drifting according to the bio-inspired approach, thus characterizing a unique gesture. Also, upon convergence, each approach tends to place the agents in different positions relative to the attractors when attempting to preserve certain features of the original space, i.e., the attractors themselves, resulting in variations of the sounds represented by the attractors.
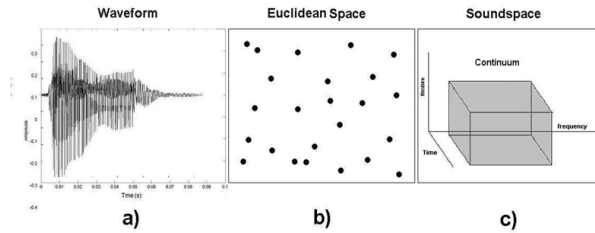

Fig. 1. Codification, pictorial resulting space and mapping to soundspace. In part a) there are N samples of a sound wave. Each waveform can be interpreted as a vector with N dimensions. Part b) represents the two-dimensional Euclidean space and the distribution of a set of sounds interpreted as vectors in this space (N here is taken to be 2, but the sound vector may contain thousands of elements), to be manipulated by the algorithm. There is a mapping to soundspace (part c) since each waveform has duration, a definite pitch (harmonic sounds), timbre (represented as one-dimensional for simplicity) and intensity. The compositional process happens in this space.

Bio-Inspired Algorithms

Artificial Immune Systems

The immune system is a complex of cells, molecules and organs with the primary role of limiting damage to the host organism by antigens. One type of response is the secretion of antibodies, receptor molecules with the primary role of recognizing and binding, through a complementary match, with an antigen. Antigens can be recognized by several different antibodies. The antibody can alter its shape to achieve a better match (complementarity) with a given antigen [12].

Artificial Immune Systems (AISs) are adaptive procedures inspired by the biological immune system for solving several different problems. The model used here, aiNet [13], follows some ideas from the immune network theory, the clonal selection [5], and affinity maturation principles. The resulting self-organizing system is an antibody network that recognizes antigens (input data set, in this case, the attractors) with certain (and adjustable) generality.

The antibodies generated by aiNet will serve as *internal images* (*mirrors*) responsible for mapping existing clusters in the data set (Figure 2a) into network clusters (Figure 2b). The resultant memory cells represent common features present in the data set that were extracted by aiNet. Let us picture a set of sounds as antigens and its internal (mirror) image as variants. Inspired by Risset's sound variants idea [18], it is possible to imagine, for example, variants as a type of immune-inspired transformation applied to the sound population. Wishart's [21] ideas on the manipulation of the sound continuum as a compositional procedure induce one to regard the convergence process of the algorithm in time as a gesture (sound transformation procedure). In this sense, the waveforms can be regarded as the repertoire to which the system is exposed, and the associated sound qualities may be linked to the specific response it elicits. It is of critical importance to notice that, upon system convergence, when an antibody-sound is representing more than one antigen-sound, it is placed in such a spot in soundspace that allows it to present features that are common to all the sounds it is representing. Figure 2c depicts the intersection of characteristics shared by three different sounds.
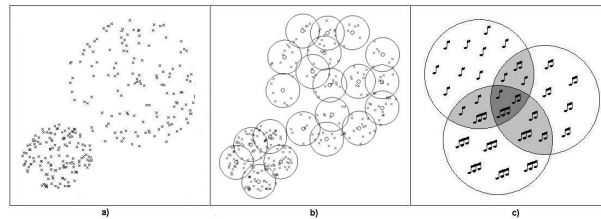

Fig. 2. Depiction of the representation of aiNet. Part a) shows the original dataset (attractors) in soundspace, part b) shows the resultant antibodies representing the attractors, and part c) illustrates the common timbral features of three classes of sounds.

Artificial Neural Networks

One important organizing principle of sensory pathways in the brain is that the placement of neurons is orderly and often reflects some physical characteristic of the external stimulus being sensed. Although much of the low-level organization is genetically pre-determined, it is likely that some of the organization at higher levels is created during learning by algorithms which promote self-organization. Kohonen [16] presents one such algorithm which produces what he calls self-organizing feature maps (SOMs) similar to those that occur in the brain.

Kohonen's algorithm creates a mapping of high-dimensional input data into output nodes arranged in a low-dimensional grid. The artificial neurons (output nodes) are extensively interconnected by many local connections with associated weights. The weights will be organized such that topologically close nodes (neurons) are sensitive

to inputs (attractors) that are physically similar. The artificial neurons will thus be ordered in a natural manner [16].

There is no consensus ordering or classification for soundspace (Figure 3a). Due to the self-organizing feature of SOM, it is possible to propose arrangements that respect the original topology (Figure 3b). The key feature of SOM that allows this process is that attractors with similar characteristics trigger neurons in close regions of the one-dimensional mapping that represents the topological neighborhood in the original soundspace. In our application, self-organization gives rise to a musically profitable phenomenon. The result of the training may cause the neurons to represent more than one attractor (zoomed in areas in Figure 3b). The expected result is a merger of their qualities (Figure 3c). The concept of musical performance emerges from the possibility of following the dynamic convergence process of the neuronal sounds from the initialization to the final result. This process would reveal the neurologically induced transformation resulting from the path followed by each neuron during the self-organizing process. Moreover, the orderly self-organizing cyclic path provided by the method could also be advantageous.
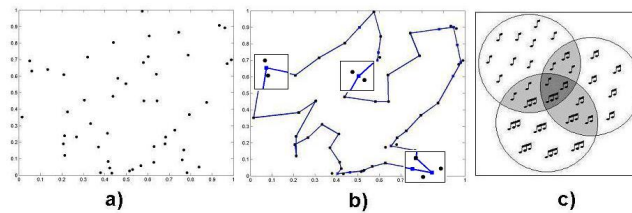

Fig. 3. Depiction of the representation of the one-dimensional SOM used. Part (a) shows the original topology of the space as represented by the waveforms (black dots) – again, an illustrative two-dimensional space is considered for visualization purposes; part (b) shows the resultant one-dimensional SOM representing the original space topologically arranged; and part (c) has the same meaning of Figure 2c.

Swarm Intelligence

Bonabeau et al. [4] define swarm intelligence as "*the emergent collective intelligence of groups of simple agents*." While intelligence is usually considered to be a trait possessed by humans alone, it can be shown to arise from interactions among individuals. Most species interact for adaptation purposes. Research in swarm intelligence is based on the grounds that there is a relationship between adaptability and intelligence, and that social behavior increases the ability of organisms to adapt.

Particle swarm optimization (PSO), the swarm intelligence paradigm used in this work, is a bio-inspired computer paradigm based on human social influence and cognition [15] whereby the particles interact to find the attractors, flying through the search space. The power of the particle swarm comes from the interactions of the individuals [15]. Individuals in the swarm have memory, they can remember the closest to the attractor they have been. Also, the particles are connected to other particles in a kind of social network. The particles it is connected to are called its neighbors. Each particle chooses the next point to visit by referring to its own previous best success and the best success of its best neighbor. In this approach, it is also influenced by the position of the nearest attractor. Thus the topological

positioning of the individual particles relative to one another in the sociometric space has a profound effect on the swarm's ability to successfully find the attractors.

In this kind of representation the multidimensional psychometric model of human sociocognition, represented by our perception of the sounds, is nested in a kind of topological sociometric space (Figure 4a) that is comprised of the waveforms. Therefore, the self-organizing search for the attractors performed by the swarm corresponds to their approximation to the attractors in Euclidean space (Figure 4b), and the generation of sounds that share perceptual qualities in the cognitive evaluative space (Figure 4c). The way the particles interact affects the dynamics of the process, which, in turn, affects the trajectories (sound transformation procedures).
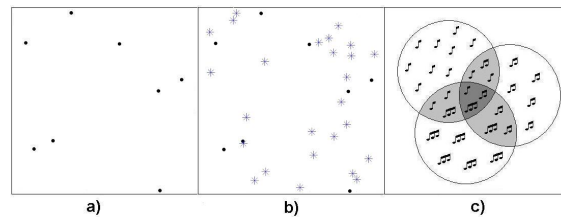

Fig. 4. Depiction of the capability of representing different regions of the environment owing to the cognitive exchange of information among individuals. Part (a) shows the original attractors as black dots, part (b) shows the particles as asterisks searching the space to best represent the original data, and part (c) has the same meaning of Figure 2c.

Results and Discussion

This section aims at illustrating the similarities and differences among the bio-inspired algorithms, trying to emphasize the musical applications of each. Each algorithm was run under very similar conditions, except for parameters that find no match in the other approaches, in order to highlight essentially the differences that are due to the distinct approaches. The experiment was performed as follows: seven attractors (different musical sounds) were used; each algorithm was initialized at random (white noise) with 14 agents and they were run for 50 iterations (discrete steps). Apart from aiNet, whose population varies dynamically, the others maintain the same number of agents over the iterations.

All algorithms converged before 50 iterations with at least one agent representing each attractor. When more than one agent moves toward one attractor, they do so through different paths, representing distinct sound transformations, specially because they were initialized in separate locations in Euclidean space, thus being different starting sounds.

At the top row of Figure 5, the trajectory of one of the agents for each algorithm is presented over all the iterations, as well as the final result (bottom row) achieved by each algorithm. All the agents chosen to be shown pursued the same attractor, which can be confirmed by visual inspection of the final result for each paradigm. The trajectory (top row) is represented as the agent at each iteration, whereas the agent is represented as its dynamic spectrum. Since each algorithm features diverse bio-inspired mechanisms to seek the attractors, different trajectories are expected to arise from the same conditions. Nevertheless, all algorithms are performing the same
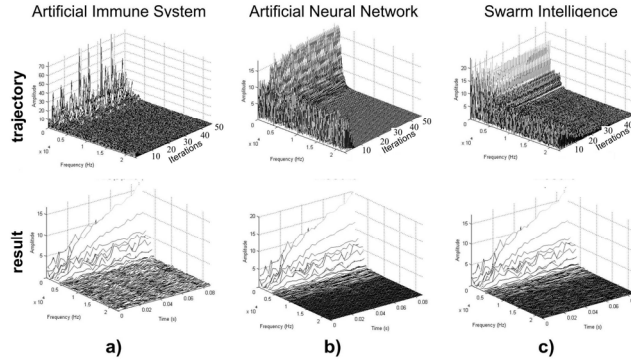
Fig. 5. Depiction of the different trajectories followed by the diverse bio-inspired algorithms over 50 iterations from white noise (random) initialization at the top row, and the resultant sounds at the bottom row. Part (a) shows one antibody-sound resulting from the application of aiNet, part (b) shows one neuronal-sound resulting from the application of Kohonen's self-organizing map, and part (c) shows one particle-sound resulting from the application of particle swarm intelligence. Each individual in the trajectory is represented just as the result, the only difference being the number of agents (50) shown at the top of the figure, to represent the entire trajectory followed by it.

musical task: find the attractor and represent it following pre-specified criteria that vary accordingly. A different trajectory can be visually identified at the top of Figure 5 as a different array of dynamic spectra from the white noise initialization to the final result achieved, whereas visual inspection of the final result at the bottom confirms that the agent shown for each algorithm converged to the same attractor, representing a variant. Close examination of Figure 5 reveals that the different approaches pursued dissimilar trajectories upon performing the task of representing the same set of attractors. Notice in Figure 5 that, along the iterations, the trajectory resulting from the application of the AIS presents a very irregular pattern. It is heard as sounds differing a great deal from one another. Also, the noisy contents disappeared very fast, in the first five iterations. SOM and PSO present similar features when visually examined. The noisy contents of the spectrum persists for more iterations and dies away more gradually. They both appear to be more homogeneous upon convergence, although PSO retained a little noisiness.

The abovementioned characteristics result in varied musical gestures emerging from the separate approaches applied to the same scenario. That is, a unique type of sound transformation is obtained with the application of each bio-inspired algorithm, summarized in Table 1. AIS searches the space in a jerky, lumpy way, resulting in discontinuous transitions between sounds, causing the impression of a bumpy transformation. SOM is the smoothest of all, gradually winding over the iterations in a snaky fashion until it converges. Finally, PSO is shaky and twitchy, with a spiraling gesture that resembles the swarming of insects around flowers. Table 1 also shows the total elapsed time each algorithm took to run on the same machine, for comparison purposes.

Also, it can be inferred from Figure 5 that, although different, the three algorithms successfully achieved their final objective of representing the attractors with different degrees of accuracy; resulting, in this case, in representations that correspond to

variations of the original attractors. Although visually they seem to differ only in high frequency contents, that is, noisiness, perceptually they differ in sound qualities.

Moreover, this is true not only for the final result, but also for each iteration, where each agent occupies a distinct position in soundspace depending upon its interaction with its neighbors, the total number of agents searching the space, the strategy used, among other factors. That is, one can see that the emergence of stable, complex patterns (each individual agent representing each waveform) results from the populational strategy allied to the self-organizing features present in all algorithms. It is very important to notice that the proposal generates both form, represented by the trajectory, and matter, the individual agent in each iteration.

Table 1. Self-organizing features and sonic correspondence

	AIS	SOM	PSO
Trajectory	Jerky	Smooth	Shaky
Gesture	Discontinuous transitions	Gradual winding	Spiraling
Transformation	Bumpy	Snaky	Swarmy
Time	46.75 s	33.02 s	38.64 s

Conclusions

In this work we have proposed a populational approach to exploring soundspace in time domain with self-organization as generative and structuring paradigm by means of bio-inspired algorithms. Self-organizing systems feature the emergence of complex patterns through local interactions among simple individuals. Many physical and biological systems have been found to present patterns that appear to be self-organizing. A large number of bio-inspired algorithms have been proposed to solve a wide range of problems in engineering as well as other areas, where traditional methods have failed. Here, we have shown the emergence of both musical form and matter by means of self-organizing temporal manipulation of sounds. In the system described here, the process of synthesis is integrated in the performance process so that the system can be used in musical performance. We have shown that the application of different bio-inspired paradigms leads to different paths followed by the agents in pursuing the attractors, which, in turn, results in different sound transformations and distinct variations of the attractors along the trajectory.

Future perspectives of this work include allowing the user to control the transformation in one or more musical dimensions (frequency, for example), reducing the self-organizing transformation to the remaining sound qualities. The adoption of moving attractors can greatly enhance the musical potential of the method, for the attractors themselves would be sound transformations. We also consider relevant to investigate other bio-inspired paradigms that could result in different transformations.

Acknowledgements

References

1. Biles, J. A. (1994) GenJam: A Genetic Algorithm for Generating Jazz Solos, Proceedings of the 1994 International Computer Music Conference, (ICMC'94), pp. 131-137.
2. Blackwell, T. M., Bentley, P. Improvised Music with Swarms. Proceedings of IEEE Congress on Evolutionary Computation, 2002.
3. Blackwell, T. and Young, M. (2004) Swarm Granulator. In G. R. Raidl et al. (Eds): EvoWorkshops, Lecture Notes in Computer Science 3005, pp 399-408.
4. Bonabeau, E., Dorigo, M., Theraulaz, G. Swarm Intelligence: From natural to artificial syatems. New York: OxfordUniversity Press, 1999.
5. Burnet, F.M. The Clonal Selection Theory of Acquired Immunity, Cambridge University Press, 1959.
6. Burraston, D., Edmonds, E. A., Livingstone, D. and Miranda, E. (2004) Cellular Automata in MIDI based Computer Music. Proceedings of the International Computer Music Conference, pp. 71-78.
7. Caetano, M., Manzolli, J. and Von Zuben, F. J. (2005 a) Application of an Artificial Immune System in a Compositional Timbre Design Technique. In C. Jacob et al. (Eds): ICARIS 2005, Lecture Notes in Computer Science 3627, pp 389-403.
8. Caetano, M., Manzolli, J. and Von Zuben, F. J. (2005 b) Interactive Control of Evolution Applied to Sound Synthesis. in Markov, Z., Russel, I. (eds.) Proceedings of the 18th International Florida Artificial Intelligence Research Society (FLAIRS), Clearwater Beach, Florida, EUA, pp. 51-56.
9. Caetano, M., Manzolli, J. and Von Zuben, F. J. (2005 c) "Topological Self-Organizing Timbre Design Methodology Using a Kohonen's Neural Network". 10 Simpósio Brasileiro de Computação e Música, Belo Horizonte, Brazil.
10. Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., Bonabeau, E. (2001) Self-Organization in Biological Systems. Princeton University Press.
11. Chen, C. J. and Miikkulainen, R. (2001) Creating Melodies with Evolving Recurrent Neural Networks, Proceedings of the International Joint Conference on Neural Networks (IJCNN-01), 2241-2246.
12. de Castro, L. N. & Timmis, J. I. Artificial Immune Systems: A New Computational Intelligence Approach, Springer-Verlag, London, 2002.
13. de Castro, L.N and Von Zuben, F. aiNET: An Artificial Immune Network for Data Analysis, in Data Mining: A Heuristic Approach. Abbas, H, Sarker, R and Newton, C (Eds). Idea Group Publishing, 2001.
14. Guéret, C., Monmarché, M., Slimane, M. Ants can play music. Fourth International Workshop on Ant Colony Optimization and Swarm Intelligence (ANTS 2004), Université Libre de Bruxelles, Belgique.
15. Kennedy, J. Particle swarms: Optimization based on sociocognition. In: Recent Developments in Biologically Inspired Computing, De Castro, L. Von Zuben, F. J. (Eds), Idea Group Publishing, ISBN:159140312X, 2004.
16. Kohonen, T. (2000) Self-Organizing Maps. Springer.
17. Miranda, E. R. Granular Synthesis of Sound by means of a Cellular Automaton. Leonardo, 28 (4), pp. 297-300, 1995.
18. Risset, J. C. Computer Study of Trumpet Tones. Murray Hill, N.J.: Bell Telephone Laboratories, 1966.
19. Smalley, D. 1990. Spectro-morphology and Structuring Processes. In The Language of Electroacoustic Music, 61-93.London: Macmillan.
20. Von Foerster, H. , "On Self-Organizing Systems and Their Environments." In: Self-Organizing Systems, M. C. Yovits und S. Cameron (Hg.), Pergamon Press, London, pp. 31–50, 1960.
21. Wishart, T. "On Sonic Art". Simon Emerson: Harwood Academic Publishers, ISBN 3-7186-5847-X, 1998.