Self-Space: Interactive, Self Organized, Robotics Mediated, Cross-Media Platform for Music and Multimedia Performance

Giannoukakis Marinos, Zannos Iannis
"Sound Arts and Technologies" Graduate Program, Music Department, Ionian University, Corfu, Greece, gaidaro2@gmail.com
Audiovisual Arts Department, Ionian University, Corfu, Greece, zannos@gmail.com

***Abstract* — In this paper, we present a platform for experimental music and multimedia performance that uses swarm theory, a subfield of complexity theory, to model of complex behaviors in intelligent environments. The platform is a kinetic structure equipped with sensors and mobile speakers that functions as a dynamically changing space for music performance. Game development software and Internet technologies have been used for the creation of a 3D virtual environment for remote control and data exchange, thereby adding an augmented reality component. The paper describes the hardware and software architecture of the system and discusses implementation issues concerning multilayered control architecture. This research combines approaches and techniques from the fields of kinetic architecture and intelligent environments on the one hand and of generative music algorithms on the other hand.**

I. Introduction

The notion of "self" as a phenomenon of human cognition and consciousness is intimately linked to the notion of space both at the cognitive and at the (neurophysiological) perceptual level, due to evolutionary factors. The mechanisms of spatial perception depend on the perceptual processing of light, sound and sensomotoric stimuli. The work described here aims at creating a robotics mediated, interactive, self organized, cross media platform providing combined stimuli related to the sense of self as created or influenced by the perception of a dynamically changing interactive physical space. The design of the platform is inspired by several historical and contemporary movements and theories from diverse artistic and scientific areas such as kinetic architecture, swarm theory, complexity theory, and emergence [1] intelligent environments, interactive multimedia performance, and interactive digital games.

Motion was introduced in sculpted or architectural structures by artists at beginning of the 20th century. Static form elements interpreted by the viewer as representing motion and it's characteristics (speed, directionality etc), were replaced by actual moving parts, thereby allowing the artists to add an extra dimension to plastic art forms, that of their temporal evolution. This was the beginning of a new art form, kinetic art, represented since ca 1910 in works by Marcel Duchamp, Naum Gabo, Alexander Calder and other artists [2].

More particularly, the idea of evolving forms and structures in architecture appeared around 1920, an early example being the Schroder-Scrader House built in Utrecht in 1923-1924 by the Dutch architect Gerrit Rietveld. In this building moving parts could form dynamically changing spatial formations. These early insinuations of kinetic architecture gave birth to a new area of interest for architects. In the 60's these ideas were further developed under the influence of modern technology by avant-garde architectural research groups. A characteristic example is the idea of "The Walking City", proposed by Ron Herron and Archigram (1964), consisting of a large number of mobile robotic structures, with their own intelligence, traveling the world according to their 'needs'. Walking cities could interconnect into forming 'walking metropolises' where individual structures are able to detach themselves from the main city cluster and move freely.

In the field of sound, the use of loudspeakers and tapes in electronic music transposed the focal point from physical gestures and positions of live musicians to the movement of real or virtual sound sources in space. As early as 1950 three-dimensional attributes were attached to sonic objects in works by Stockhausen, Xenakis and other pioneers. As the spatial characteristics of sound attained the status of compositional parameters, fundamentally novel ideas emerged for the presentation of sound compositions in performance spaces. Moving sound sources and other types of sonic installations in architectural structures such as the Phillips Pavilion 1958 and the West German Pavilion Osaka World Fair 1970 created intimate links between music and architecture, thereby showing the potential for previously unimagined developments in both fields. In a way, sound acquired its own position in space and spatial formation its own sonic characteristics or dimensions.

Prompted by the ideas and developments outlined above, the question arose in the course of the present work, of how to explore the intimate relationship between sound and other media and architectural space in an environment which can change its own form in time during the course of a performance. This exploration should have at least semi-experimental character, in the sense of providing data about the actions of performers and/or other participants in the space, measured via all kinds of sensor devices. This has twofold implications: On the one hand it creates an interactive performance space, on the other hand it creates a measurable cybernetic loop between a controlled environment and human agents acting in it. Neuroscience is recently investigating brain functions taking place whilst actively perceiving. It has been proven that a unique and separate area of the human brain is committed to analyzing and processing movement and its characteristics. Brain cells analyzing movement are different from those analyzing visual stimuli, such as color or shape. Thus, when motion is perceived, essentially different processes take place than when perceiving shape alone, which also implies that the nature of the evoked sensations are different [2]. Furthermore, the areas perceiving motion, are also responsible for estimating the quality of the motion, possibly playing an important role when judging kinetic art works [2].

If one regards that one of the functions of art is to explore and expand the processes of human perception, then the present work can be viewed as an experimental art platform with applications in performance arts, architecture, cognitive science research and related interdisciplinary research [2]. Scientific and technological breakthroughs in areas such as artificial intelligence and neural networks combined with the ability of real-time computation of complex algorithms have prompted a convergence of science, architecture and music into a new form of *information art*. Here we intend to use technological advancements in microprocessors, robotics, human-machine interface systems to create a dynamic combined physical and virtual environment, as a tool for research projects in information art and its above mentioned related fields.

The second major field of application for this project is that of generative music systems. By that is meant the area of algorithmic composition and real-time "automatic" computer improvisation that developed mainly since the 1970ies, with influences from research in fractals, chaos theory, cellular automata, artificial life, genetic algorithms, data sonification and others. Research combining aspects from both the convergent sound-space fields and the generative algorithm fields is being conducted by many research groups and individuals, including: Hyperbody Research Group and ONL in the Netherlands, Media Lab in MIT, La Fondation Daniel Langlois, companies like ART+COM and Hexagram (made up of designers, scientists, artists, technicians) and individual artists such as Bill Vorn and Simon Penny working with robotics in interactive installations.

II. Hardware Setup

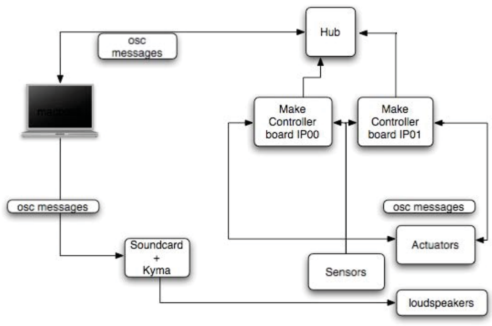

Fig.1 System setup schematic

The system presented here is a robotic installation cross media platform that functions as an intelligent interactive virtual and physical environment. Interaction between the virtual and actual installation was achieved using the Make Controller Kit as a "portal" between the virtual and the physical model and a setup of sensors and software components extending the system's input to accept user information. Output relies on the continuous interaction between virtual and physical agents and is controlled by behavioral models implemented in the system's software.

The components of the system are:

A. Sensors

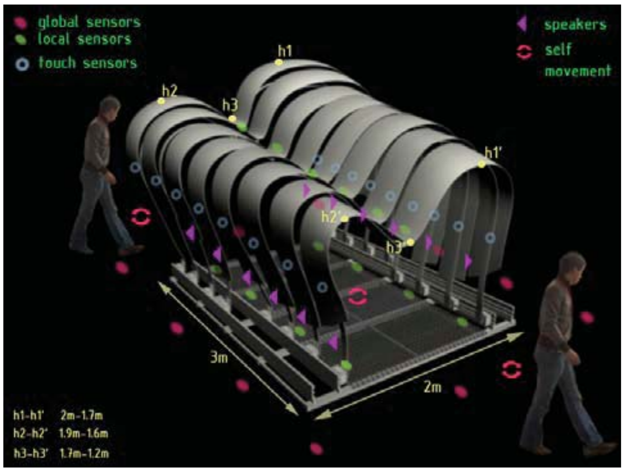

Fig.2 Virtual model of sensors setup

Movement and position control of the installation is controlled based on information gathered by sensors that measure the distance and positions of participants relative to the installation. Sensors are placed to read information both from the outside and the inside part of the frame, allowing users to control the qualities of sound, the source positioning of sound sources and the form of the physical structure of the installation itself.

B. *Control/communication*

Fig.3 Make Controller

Communication between the computer and all the different sensors and actuators was implemented using two Make Controller boards. The Make Controller (http://www.makingthings.com/products/KIT-MAKE-CTRL/) is built around an Atmel AVR family architecture RISC Flash memory microcontroller, the AT91SAM7X256. The controller's board provides control, computation, along with CAN and Ethernet communication.

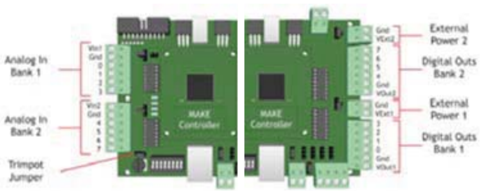

Fig.4 Analog inputs and Digital outputs

The application board supports connectivity through USB, CAN interface and Ethernet. It provides 8 analog inputs, 8 digital outputs and 4 servo outputs. These input and output ports are used to read values from the sensors and control all the actuating devices.

Two Make Controller boards are required to provide a minimum sufficient number of inputs and outputs. The boards are connected on a local Ethernet network with a 2 GHz Intel double core MacBook, running Mac OS X version 10.4.8.

C. *Actuators*

DC and stepper motors are used as actuators controlling 3 different types of movements of structural components of the installation, which affect shape of the enveloping the space of the installation and and the positioning of the loudspeakers within it.

Types of motors:

- One 180 V, 1800 rpm Leeson DC motor.

Fig.5 180V DC motor

This motor is used to provide the displacement of the structure in the mechanical slides.

- Two 24 V, 2000 rpm DC motors.

Fig.6 24V DC motor

- Three Sanyo Denki bipolar stepper motors

Fig.7 Stepper motors

The Sanyo Denki bipolar stepper motors are used to drive the conveyor belts, which control the position of the sound source loudspeakers.

D. *Sound sources*

As sound sources we employ 6 passive loudspeakers arranged in 3 stereo pairs. The physical orientation and position of each speaker is controlled by to small dc and stepper motors.

E. *Electrical Panel*

An electrical panel is used to convert PWM signals from the board's digital outputs to analog signal driving the DC motors. On the whole, the installation's control panel is a 'multi-converter' of OSC and MIDI messages to voltages capable of operating different electrical, pneumatic, mechanical components (motors, solenoids, pneumatic pistons etc.) and assuring electrical safety during the process.

F. *Steel Structure*

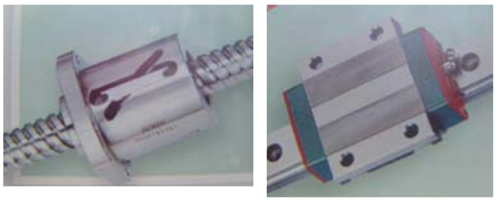

Fig.8 Screw guides and linear guideways and wagons

Mechanical slides and wagons were used in order to position the arcs along the horizontal and vertical axis.

A 3000x2000 mm stainless steel frame has been constructed to support the structure and to create the walking path through the steel frame.

A number of steel arcs of similar shape arranged in order of decreasing in height are attached on the wagons and are capable of sliding along the frame. Thus, the overall structure can expand and contract in a telescopic pole manner, attaining a maximum variation of 17-20% in height and 60% in length. Several models of the enveloping structure at different contraction and expansion stages are shown in figures 9-13.

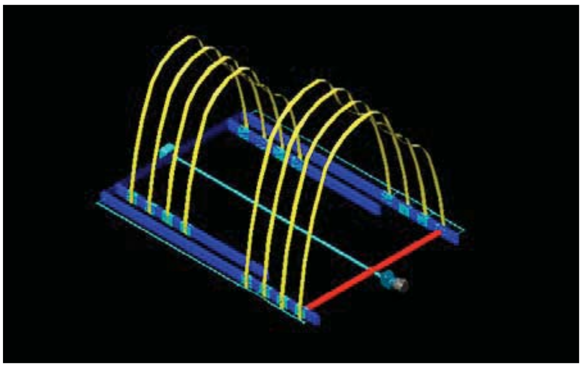

Fig.9 Horizontal motion of the arcs

Fig.10 Folded view

Fig.11 Expanded view

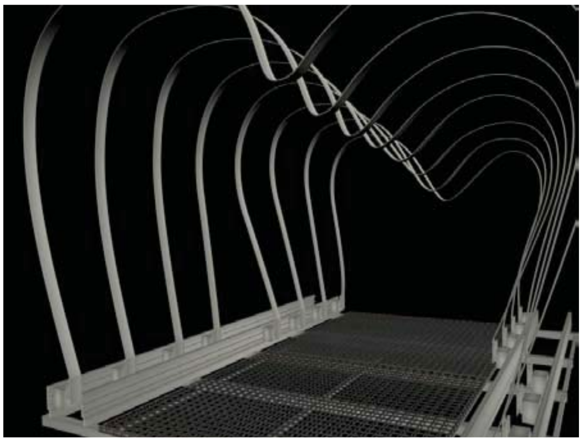

Fig.12 Internal view

Fig.13 View of structure lighted with reflective steel surfaces

III. SOFTWARE ARCHITECTURE

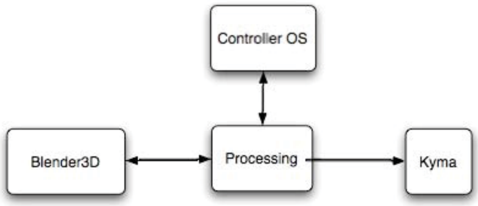

Fig.14 Software setup schematic

A. *Firmware*

Interfaces between the sensors and actuators and the microcontroller interface device (Making Controller) have been developed for the following systems:

1. *Processing*. Java based programming environment for real time animated graphics synthesis and processing. http://www.processing.org
2. Kyma: Graphic programming environment for real-time sound synthesis and control.
3. SuperCollider 3: Real time object oriented programming environment for sound synthesis and processing.

4. Blender3D: Development environment for 3d animation and interactive games with Python API.

The interfaces of these software units to the hardware platform are implemented on a RTOS (Real Time Operating System). This activates all components for computation and communication. The firmware is programmed in C programming language and includes files that provide the following services:

- Network functionality
- Basic operating system functions
- OSC and USB functionality

Protocols and functions supported on the firmware are:

- TCP/IP, used to provide network operations over the internet. Both UDP and TCP are supported
- OSC, is used as an option to connect all the subsystems of the board
- Controller library functions, giving access to the controllers subsystems, for example the CAN interface or PWM devices
- Application Board library functions, providing access to additional subsystems for hardware control. For example the AnalogIn() functions help read any of the 8 inputs or the Motor() functions provide DC motor control.

The RTOS uploaded on the microcontroller communicates with *Processing*. *Processing* performs parameter computations and adjustments and controls the flow of information between the software and hardware components.

The Blender3D game engine is used to create an interactive virtual space. A 3d model of the structure is presented in a game like environment, allowing interaction with the user through the MacBook keyboard. OSC messages are sent to Processing from the Blender game engine as additional parameters to the computations controlling the state of the installation in a continuous mode. Messages are also sent to the Kyma sound engine, via the Osculator application, allowing control in real time of sound parameters.

A uniform protocol was developed on the basis of the OSC (Open Sound Control) protocol, that serves as uniform transparency to the different software and hardware platforms.

Additional software used includes 3D modeling software like AutoCAD and 3DStudio Max.

IV. Interaction Model Prototype

The purpose of the development was the creation of a three dimensional environment designed to respond to agent input using sensing devices, real time data processing and control of actuating devices. The response gives feedback to the agents through spatial configuration and sound synthesis. To amplify the effect of the sonic environment, stainless steel sheets are used as reflective surfaces placed on top of the steel arcs. This gives the ability to control, along with source placement, the diffusion of sound in the installation. The whole system is also implemented in a 3D game like environment providing control over the sound placement between speakers and affecting the sensor data. The constant information input from the sensing devices and the virtual environment, provides data for the parameters controlled by Processing generating in real time ‘next state’ computations. Control statements in Processing condition the behavior of the installation and decide the way to respond to user input. Simple behavioral structures are used to determine the ‘character’ of the machine entity and the response towards the user input information. The term ‘character’ is used here to describe the perceived qualities of interaction of the machine with the users, in other words, the perceived qualities of its behavior.

A. Behavior:

The assigned character of the system is an illusory attribute of the machine-entity implemented through the connection between the different subsystems and user interaction with their elements. Software design issues were based upon conceptual framework derived from evolutionary organic behavior and allowed characters to immerge on a unified machine-organism. This artificial organism, endowed with machine motion and producing sound and light while functioning, creates the illusive notion of behavioral structures. These structures are closely mapped to sound morphologies creating fictitious behavioral responses connected to the actual machine responses. The sonic aspect is vital in enhancing the characteristics of the different behavioral patterns through appropriately chosen synthesis strategies.

Sound morphologies combined with the notion of movement evoke a dual effect on user’s perceptual process. On the one hand movement perception simulated or enhanced in sound space and on the other sound morphologies affected by the three dimensional positioning and displacement of the sound source. This combination of stimuli creates a fictitious perceptual space for the user to translate, analyze and characterize. This space is a mental playground where characters, behavior and language or a machine-mediated counterpart of these elements is created. The parallel control of the subsystems through a central ‘brain’ in the computer, allows global and local behaviors to emerge changing the level of user immersion with the installation.

The resultant behaviors vary according to the mentioned elements forming them, from peaceful to aggressive, linear to chaotic, attractive to repulsive, individual to flocking. These behaviors effectively create fictitious characters carrying these qualities in the different installation space expressions.

B. User interaction:

The A.I. and evolutionary algorithms that simulate the system's behavior are based on the fact that the game engine/sensor sub-network uses events to represent different types of user interaction. Physical events (sensors or keyboard input) are evaluated as low-level events, interpreted by the computational engine as high-level actions presented to the user. For example a simple collision event, taking place in the virtual domain, triggers a sequence of predefined behaviors in the system's event chain translating to movement or generation of sonic event. The direct reaction on low-level events is affecting the user's perceptual processes, closing a feedback loop, suggesting acting again (or not) accordingly. This continuous interpretation is based on the behavioral model system, where high-level actions are recognized and executed. Starting with a default set of high level events as response to discrete/deterministic low-level actions, the intelligent system in Processing generates altered sets of behaviors by evaluating the fitness factor of the particular low level actions. User input determines the degree of abrupt change (surprise element) affecting overall system behavior. With this simple mechanism a large number of different behavioral combinations are generated from discrete low-level activities.

Fig.(15) demonstrates how deterministic behavior defined by the artist can be used to alter behaviorally complex responses to continuous user input by through exchanging data between the different subsystems.

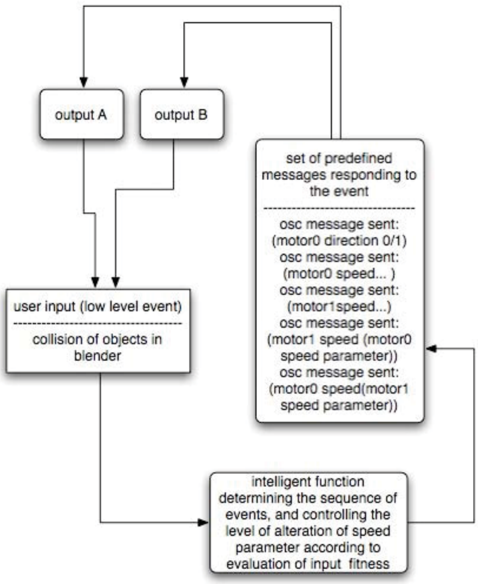

Fig.15 Diagram showing the event chain triggered from a simple low-level action

C. Sonic Events:

Sonic events generated follow a similar architecture in their structure with the rest of the system. This was considered essential in order to accomplish conditioned unpredictable responses to the rest event chain designed in the software platform. The key point in the design strategy involved creating relations and dependencies in the infrastructure of the system's responses. Effectively, there appeared sonic morphologies in accordance with the movement of speakers and the structure's temporal space formation.

To take this issue into account during the design process, the sound sub-system had to be dealt as a swarm in its own accord. Thus the conditioned behavioral model responding to the overall network is subject to the global rules of the system.

Sound synthesis techniques available in the Kyma engine were modified with scripting to acquire sensitivity to the data from the rest of the network. The global genetic code running on Processing was used to initialize and process the *parameter space* of the sound system. Dealing with different modules and synthesis techniques in Kyma, a global *parameter space* was necessary to be defined. These global parameters were directly related to data coming from evolutionary computations. Further scripting was required to map the global parameters to local ones, controlling the properties of each module. Conditioning of parameter data locally was applied to achieve more control over the output.

The desired behavioral model of the sound contained strategies for presence and source placement (applied to all the sounds) and control of the specified parameters for the different synthesis techniques. Basic techniques used included granulation techniques, cross synthesis, and processing sampled material. Each different process was modified to evolve according to the global computations, mapped to properties of each sound synthesis module.

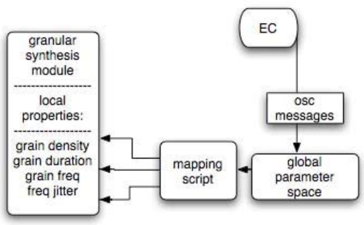

Fig.16 Global parameters constantly changing from network input mapped to local synthesis properties

The user's interaction with sound is implemented on a nonlinear global level, supplying the necessary input for the system to decide its state. User input combined with the state of the installation itself affects the initial state of the modules in the sound engine, the selection of parent parameters and control the fitness, crossover points (determining the children state) and mutation (altering the children state) of the parameters.

Presenting the system in such a way allows the exploration of each module's sound-space in relation with the system's general behavior. Direction, movement or not on the horizontal axis and rotation of the speakers, combined with dynamic sound object-source placement and the changing spatial configuration, generates sonic morphologies.

As a result these morphologies are affected from the global scenario instead of the technique's local properties.

V. DISCUSSION

Appropriate design and architecture, even in small scale interaction systems that are able to allow open flow of information between different systems on a predefined level, can produce complex, structured behavioral strategies. This permits agents to operate in unison so as to generate an integrated output. In his paper on swarm architecture, Kas Oosterhuis, describes key points in the morphology of such systems in architecture, setting the design basics of multi-agent networks (swarm) interaction. The focus while designing these systems is not on their material appearance, but on the 'membranes' allowing the flow of information. Agents carrying input and output information are processed by higher level agents, which in the case of Oosterhuis' arguments is space [3]. Space is not simply accepting and outputting that information, but digests the agent in order to process the information carried (strangely enough the same suggestions were found in Le Corbusier's-Xenakis' project the Phillips Pavilion). Thus a common network allowing the input-process-output of information is created, based on the permeability degree of the membrane. Under the perspective of this general approach, different qualities can be parameterized and conditioned to produce a swarming output. The generated result seems complex enough to be unpredictable, but on discrete recognized *quanta* level simple in its interactions. Merging the agent with the network can be extended in a vast degree increasing the level of complexity but still retaining control over the system's output.

It was of concern in the design process of the installation to build a system with a certain level of transparency in its 'membranes' (gateways controlling data flow). Allowing that, enabled the system to exchange parameters and data between different subsystems and to apply a common processing and behavioral strategy towards all the different agents. The interaction between the agents is organized hierarchically by employing dynamically altered degrees of predictability in the events generated by low-level actions. As a result the agents were stripped from their 'external' characteristics keeping intrinsic basic qualities as simple as possible.

The main challenge of this work is to combine architectural and general environmental system designs principles with those of sound/music generation and systems. The resulting system is perceived as a musical instrument with unique acoustic properties that permits the exploration of new sonic morphologies. At an abstract level, the existence of a unified network of interconnected sub-systems provides multiple stimuli to the 'human brain' subsystem, thereby closing the cybernetic feedback loop and harmonizing the installation with its dual function as environment and performance instrument.

VI. FUTURE WORK

Additional software development and investigation of communication protocols, allowing the remote exchange of data (music, video) is required. One example is the DMX protocol for control of stepper devices, in order to achieve a greater number of control channels with less wiring.

This project presented is the first in a series of installations employing the same structural design elements able to form a larger scale swarm, following the same properties and principles. This larger network is going to enable us exploring swarm networking between remote places, connected through the internet, exchanging cultural and environmental elements from different locations/cities/countries.

For the completion of different installations the utilization of various OSC to voltage converters is necessary, providing flexibility during the design process.

It is also considered essential the development of different sensor devices, for the implementation of different interaction scenarios. Sensors varying from noticeable personal, wearables, to invisible global sensors for crowd behavior and environmental measurements.

ACKNOWLEDGMENT

Many thanks to Constrinox S.A., D. Giannoukakis and A. Giannoukakis for providing the means and materials used in the installation, D. Tomaras and G. Tsotsos, helping in the realization of the installation with their experience in mechanical engineering and electronics, P. Giannakou and Z. Michael helping with the virtual models and in deciding and design the form of the installation and providing material for architectural research.

This research is part of a theses project for the "Sound Arts and Technologies" Graduate Program of the Music Department of Ionian University. This program is funded by the Operational Programme for Education and Initial Vocational Training of the Hellenic Ministry of National Education and Religious Affairs.

REFERENCES

[1] S. Johnson, "*Emergence: The Connected Lives of Ants, Brains, Cities, and Software*." New York: Scribner, 2001.

[2] E.Zeki, *"Inner vision, an exploration of art and brain"* University Edition of Crete, pp.179–205, Heraclio 2002.

[3] K. Oosterhuis, L. Feireiss, "*Game Set and Match II."* Rotterdam: Episode Publishers, 2006.

[4] D. Ashlock, *"Evolutionary Computation for Modeling and Optimization"*, Springer Science+Business Media, Inc, pp. 1-60, New York 2006.

[5] M. Maybury, O. Stock, W. Wahlster *"Intelligent Technologies for Interactive Entertainment"*, Springer Science+Business Media, Inc, pp. 64-74 , New York 2006.

[6] G. L. Nelson, *"An application of genetic algorithms to the growth and development of musical organisms"*, unpublished.

[7] T. Blackwell, *"Swarming and music"*, unpublished.

[8] Bill Vorn,, "*Machine-Mediated Communication: Agents of Representation*". *In human cognition and social agent technology"Advances in Conciousness Research"*, , pp. 377-393, Amsterdam 2000.

[9] Bill Vorn, Louis-Phillipe Demers, "*Real Artificial Life as an Immersive Media*", Convergence: 5th Biennal Symposium for Arts and Technology (New London, Conn.), pp. 190-203, London 1995.