smdk: An Interactive Self-Organizing Sound Environment

Georg Fleischmann, Michael Hoch, Detlev Schwabe,
Christian Hübler, Alexander Tuchaçek, Yvonne Wilhelm

Academy of Media Arts
Dept. of Computer Science/Audio-Visual Media
Peter-Welter-Platz 2
50676 Cologne, Germany
georg@khm.uni-koeln.de

Abstract

smdk cross-disciplinary project by Knowbotic Research that results from an exchange of working techniques between media artists, computer musicians and computer scientists. The interactive installation *smdk* consists of a sound data base that forms a self-organizing system by means of simple artificial life rules. A visitor can interactively explore the data base in an action space and will organize sound by manipulating the duration, volume and direction of sound elements, which in turn depend on the speed and type of his or her movements. In this paper we describe the interactive environment of the installation as well as the concepts of the artificial life system, including self-organization of the data sounds and the real-time composition performed by a visitor in the action space.

Key words: interactive simulated environments, artificial life, self-organization, particle systems, virtual reality, digital sound analysis

1. Introduction

smdk (simulation space mosaic of mobile sound data) is an interactive installation by Knowbotic Research (a free grouping of media artists and scientists) that has been presented at the Mediale '93 in Hamburg, Germany, and the Ars Electronica '93 in Linz, Austria, where it received the Golden Nica Award for interactive art. The installation consists of three parts (see Fig. 1): firstly, sound samples are collected, analyzed and put into a data base that forms a virtual sound room by means of self-organization of the data sounds that are part of an artificial life system; part 2, the action space is a physical room where the data sounds are made audible to the visitor who can navigate and interact within the self-organizing system; in part 3 these events are visualized in real-time on a large screen to the audience. The sound samples are collected from all over the world and represent personal attitudes of their creators. By entering the self-organizing system, new interrelations between the different attitudes are established while the sound samples move in the virtual sound room.

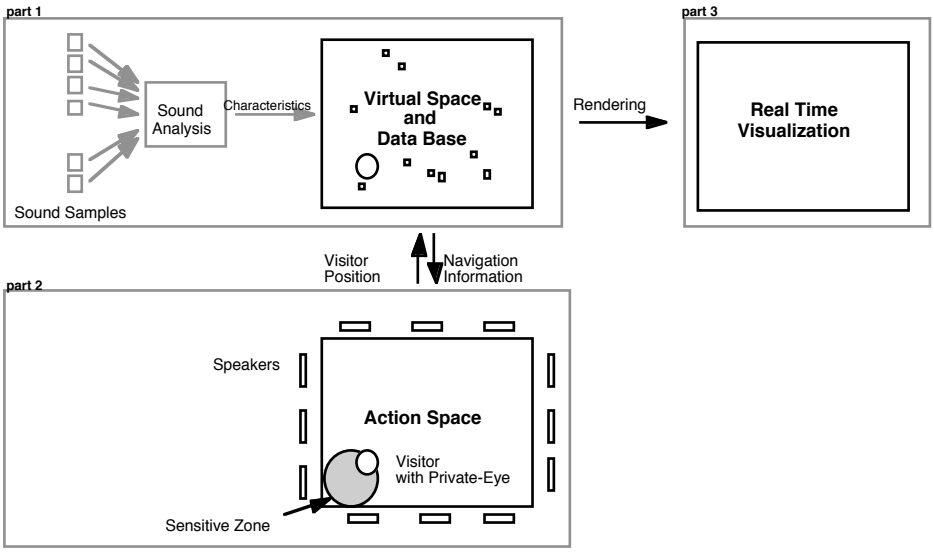

Fig. 1: Block diagram of installation parts

The term "self-organization" stems from intensive efforts in various scientific disciplines to clarify the spontaneous emergence of order and structure and their evolution into increasingly complex systems [Haken and Wunderlin, 1991]. The study of macroscopic components, which initially emerged in the form of several nearly unrelated efforts, has today been consolidated under this concept. This cross-disciplinary research approach increasingly erodes the conventional clear-cut distinctions between simple systems as studied in physics and chemistry and complex systems as examined in biology and human science. Processes that tend to generate spontaneous ordered structures can be found in all branches of science and all areas of animated and non-animated nature. Examples include the emergence of frost-flowers on a freezing window, life in an anthill or a bee colony, thermodynamic convection flows in fluids, the inherent dynamics of flocks, herds or other animal communities [Reynolds, 1989] (e.g., schools of fish, aggregate behaviour of insects or birds), the emergence of social or industrial structures (e.g., corporations, public authorities, families), and the self-regulation of traffic.

In this project, the concept of self-organization is used to form an interactive artificial life system that can be explored while moving in the action space. The artificial life system is implemented as a particle system based on simple rules of Newtonian physics. By using a space tracking device to determine the current visitor position, the visitor becomes part of the system and is able to explore the interrelations between the sound samples and influence the emerging structures. A sensitive zone centered at one hand of the visitor determines which sound is to be played. Depending on the direction and velocity of movement, either small fragments of sounds or the entire statements are played. Hence, the visitor performs a real-time composition and enters a virtual world of sound and information. In contrast to common virtual reality systems, which imitate real world aspects, *smdk* creates a new, different reality that is determined by the spatial configuration of the sound groups and the order in which sounds are played.

2. Installation Setup

For this project, Knowbotic Research sent out an international call for participation about two months before the first installation. The sound samples, i.e. personal, acoustically formulated statements with a maximum duration of 6 sec, sent in via computer networks like Internet and CompuServe or on ordinary audio and DAT tapes, are analyzed both automatically and manually and given up to 10 different characteristics reflecting acoustically distinguishable features. Next, the sound samples are placed in a data base and become mobile elements (agents) in real space and their equivalent in virtual space. On account of their characteristics, self-similar groups form and by means of a simple artificial life system, the behaviour of the entire system is determined. There are 10 different groups that are based on the characteristics given to the sound data. Equipped with basic degrees of freedom, the agents organize spatially in sound groups that represent one characteristic each, forming an organism that is continously restructuring itself, e.g. agents may form a group or join a group if they have the appropriate characteristic or may freely float in space as free agents.

This dynamic system is made accessable for the user in the action space. The action space is part of a larger, mostly darkened room outlined by fiber optics light cables that show the base boundaries of the space (e.g. at the Ars Electronica the action space was located in the center of one floor of a parking garage, the 7 by 7 m action space was outlined by fiber optics light cables). Any visitor entering the action space is spotted by an ultrasonic tracking system. His or her position is transmitted to the virtual space. The visitor is equipped with the ultrasonic sensor attached to one hand and a small monitor mounted in front of one eye that is linked to the data space via radio transmission and provides textual information. This "private eye" shows directions for entering into contact with the three nearest sound groups and information about the characteristic and quantity of groups located within a certain range of action (Fig. 2). A sensitive zone in the virtual space surrounds the position of the visitor's hand, allowing him or her to activate the original sound information. In this way the visitor will 'organize' sound by manipulating the duration, volume and direction of each sound element, which in turn depend on the speed and type of his movements. The visitor becomes part of the system and may influence its complex behaviour and build new connections.

Fig. 2: Private-Eye display: a) while navigating in free space, b) while being in group 5

The large screen which is located outside of the action space shows a rendered, real-time sketch of the virtual organism which, in turn, is described in the action space in acoustical and textual terms only. The computer visualizes the overall system and the dialog between the visitor and the data base (Fig. 3). Each of the 10 possible groups has its own geometric representation and once an agent joins a group it will adopt the geometric representation of this specific group. The current position of the visitor in the action space is represented by a red cursor. The perspective of the virtual camera follows the movements of the red cursor, i.e. the movements of the visitor, and sounds or agents that are being activated will light up.

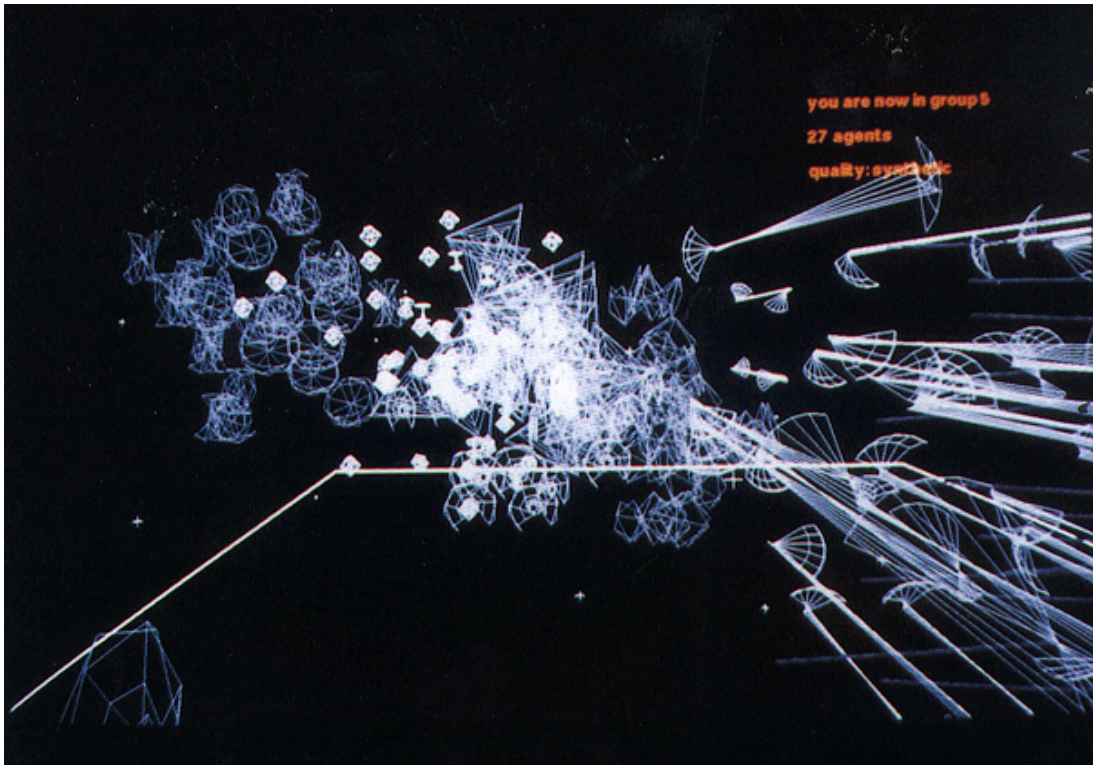

Fig. 3: Real-time visualisation of the virtual organism

3. Concepts

3.1 Spatial behaviour of the agents

The set of sound data used within this project forms a complex, self-organizing system. While the basic elements of Newtonian particle systems are used, each agent (particle) follows its own local rules and possesses its own limited world view. As a result, the overall system exhibits an orderly, complex behaviour. Its operation, which appears to be directed at the viewer level and moulded by external forces, is derived entirely from a lower-order set of laws. The audience watching the real-time visualization of the system can perceive this process by observing the creation and dissolution of groups and the migration of sounds between these groups. In the action space, the visitor can become part of the system himself and is given the opportunity to influence its complex behaviour and build his own connections.

In computer graphics, particle systems have been used to model visually complex phenomena. The complex behaviour of particle systems results from large numbers of independent particles reacting to forces. Particle systems can be classified according to the interactions between the particles. The simplest systems are systems of *independent particles*, where forces on each particle are independent of other particles. More complex behaviour results, when particles interact. The interaction between particles can be fixed, as in spring systems, or it can be dynamic. In *dynamically coupled* particle systems, the interactions are spatially defined and evolve over time.

Our animation system can be viewed as a dynamically coupled particle system in which each agent is modeled as a particle. Each agent has its own individual behaviour which depends on the position of the agents belonging to the same group and some other characteristics that influence the migration of agents between groups. The current state of an agent is characterized by its acceleration, position and velocity. At any time, an agent can decide to change its current movement by applying forces. Knowing the forces, the state of each agent can then be updated using Euler's difference approximation for discrete time intervals [Goldstein, 1981]. Using the notation

a = acceleration

m = mass

v = velocity

x = position

t = current time

Δt = time step

the equations of motion for an agent i in R^3 are:

$$a_i = F_i / m_i$$
$$v_i(t + \Delta t) = v_i(t) + a_i \Delta t$$
$$x_i(t + \Delta t) = x_i(t) + v_i(t)\Delta t$$

At each time step Δt and for each agent i, the total force F_i, which depends on the present spatial position of the agents, is computed. The vector F_i and the mass m_i determine the acceleration a_i. A change of acceleration results in a change of velocity. The position x_i is updated based on the current velocity v_i. To get a real-time system the state of all agents has to be updated at least 12 times per second.

The forces F_i the above equations are triggered by a behavioural model. Whether an agent changes its current velocity and direction depends on several simple rules. These spatial rules are:

(1) try to avoid collisions with the boundary of the action space
(2) try to avoid collisions with other agents of your group
(3) try to stay close to your own group
(4) try to emphasize the current movement of the whole group
(5) try not to penetrate into other groups
(6) try to stay inside the sensitive zone of the vistor's hand once there

Rule (1) keeps the agents within a predefined space. Rules (2) and (3) cause the dynamic movement of agents within a group. Rule (4) is responsible for the migration of the whole group and rule (5) avoids the penetration of groups. Rule (6) is an interaction rule between the agents and the visitor. Although this rule is very simple, the visitor is able to influence the behaviour of the whole system. All rules, 1-6, are simultanously applied, i.e. a force vector for each rule is calculated and all these vectors are added up to one force vector F_i. Priorities for some rules can easily be realized by intensifying the corresponding force vector.

There is another set of rules that is responsible for the formation of groups and the migration of agents between groups. After starting the system, all agents are free agents, which means that they do not belong to any group. Their initial direction and velocity in space is randomly chosen and then the following rules are permanently applied:

(7) two meeting free agents with matching characteristic form a new group, if a group of this characteristic does not already exist
(8) a free agent joins a nearby group if it matches the characteristic of this group
(9) an agent leaves its group if an 'indicator' has reached a certain threshold. If there is a nearby group with matching characteristic, it changes to that group, otherwise it becomes a free agent

Rules 7 and 8 cause groups to form and grow and rule 9 diminishes the size of groups. The indicator in rule 9 is implemented by a numerical value that reflects the endeavour of the agent to leave its current group. If this value exceeds a certain threshold, the agent tries to change to another nearby group or becomes a free agent. This tendency to change is permanently updated and depends on the time the agent has already spent in the group, an individual fluctuation parameter, the size of the group and the sound characteristics of the agent. Rules 7-9 are responsible for the permanent reorganization of the sound groups. Thereby each sound can again and again be experienced in different context to other sounds.

Rules 1-9 are checked at each time step. Certainly, not all rules are applicable at each step. Rules 2-5 for example have no influence on free agents. But as soon as an agent joins a group, its movement immediately depends on all the other members of this group. When an agent leaves a group it just keeps its current direction and velocity. Rules other than those used within our system are possible and could lead to a different global behaviour of the system. For example, we did some experiments with rules for the formation and destruction of groups depending on the movement of the visitor.

3.3 Sound Analysis

The incoming data sounds were analyzed and given several acoustically distinguishable characteristics. Five out of ten overall characteristics, i.e. noise, environmental, instrumental, speech, and synthetic where determined manually. The other five characteristics, i.e. rough, high, low, floating, and pulsing where determined in an automatic way by using simple signal processing techniques. Based on the Fast Fourier Transform (FFT), or simple low pass filtering techniques, some empirical features where selected to classify the sound samples.

3.4 Real-Time Composition

The transformation of sounds into digitized form allows them to be manipulated by a computer, i.e. a sound becomes a freely usable component detached from its physical context. This allows the fragmentation of sound entities into small acoustic units and to dissolve given time sequences. By establishing a dialogue with the system, the visitor acting inside the action space will evoke time fragments of these sound data. The direction of activation determines which of the 12 speakers is used to play the sound. The relative velocity between an agent and the visitor controls the volume of the sound fragment. All sound samples are played in a loop as long as they are activated by the visitor. Depending on the specific movements of the visitor, either the entire sound statement is played or many fragments are

played subsequently altering speaker direction and volume. The transitions between successively activated sound samples are performed by cross-dissolving the sound data and volume as well as the speaker direction.

4. Technical Description

In this section we will describe some of the technical details of our environment. Fig. 4 shows schematicaly the technical setup of our system as it was installed at the Ars Electronica '93. The entire system can be subdivided into four subsystems which are connected to a Graphics & Control System via ethernet and serial lines.

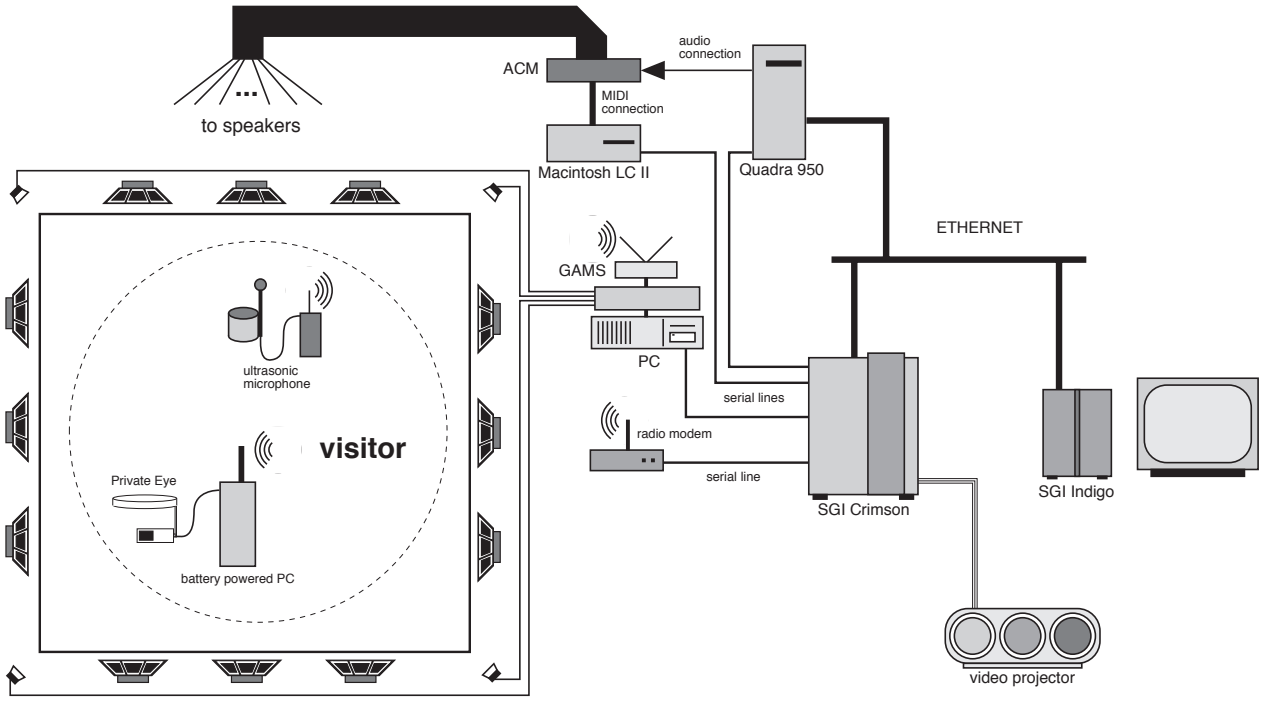

Fig. 4: Schematic diagram of the technical setup

4.1 Graphics & Control System

The Graphics & Control System, which performs the calculations for the particle system, is also responsible for displaying the wireframe models of approx. 400 to 500 agents in real-time. The heart of the system is a Silicon Graphics Crimson workstation with a fast RealityEngine graphics board. With this configuration we are able to calculate and display approx. 12 to 18 frames per second. This real-time animation is then displayed by video projection on a large screen outside the action space. The Crimson gets the position data of the visitor from the position tracking device and sends information about the number of the sound to play, its volume and the number of the speaker to play through, to the audio subsystem. It also transmits data to the wireless PC that controls the *PrivateEye* to update the information displayed on the *PrivateEye* monitor.

4.2 Sound Database Maintenance

The database maintenance and automatic and manual classification of incoming sound data is done on a Silicon Graphics Indigo workstation, equipped with an integrated digital audio device and a DAT drive capable of playing and recording audio data. At the Ars Electronica '93 the Indigo was connected to the Internet through a telephone line and a SLIP[1] connection. New sound data contributions where transfered from the *smdk* mailbox in Cologne and analyzed once a day. With the DAT drive we are able to make use of sounds which were not sent via the Internet. The resulting feature list of all sounds is then read by the main program on the Crimson.

4.3 Position Tracking System

For detecting the position of the visitor, we use an ultrasonic system called GAMS[2]. The GAMS was developed by Canadian engineer Will Bauer and is based on measuring the travel time of ultrasonic signals [Bauer and Foss, 1992]. Four ultrasonic speakers, located in the four corners of the action space at a level of approx. 180 cm send multiplexed ultrasonic signals, which are then detected by a microphone attached to the hand of the visitor. By using a small battery-powered radio transmitter carried by the visitor,

[1] SLIP: Serial Line Internet Protocol
[2] GAMS: Gesture and Media System

these signals are sent back to a controlling PC where the best three measurements are used to calculate the position.

4.4 Audio System

Two Apple Macintosh computers are the core of the audio system. One is connected via a MIDI[3] interface to an NICH Audio Control Module (ACM) and controls the speaker system, which consists of 12 active speakers located around the action space. The Macintosh gets the control data from the main system and activates the speaker corresponding to the direction of the sound. The second Macintosh is a Quadra 950, equipped with an AudioMedia II audio device board and a large harddisk, which contains all the sounds. The actual sound to be played is triggered by data received over a serial line from the Crimson.

4.5 *PrivateEye* System

The PrivateEye system consists of a battery-powered PC, which is carried by the visitor in a special jacket and a small LED monitor which is attached to the head similar to a headphone. The portable PC receives its data via a radio modem from the Crimson and displays the information on the connected LED monitor.

5. Conclusion

This cross-disciplinary project offers a non-traditional immersive environment were diverse ideas from different contexts are integrated. In comparison with the usual meaning of the term 'virtual reality', *smdk* follows a completely different course. Rather than just producing a computer generated copy of a real world, the visitor can experience a new reality, using minimal abstract navigational clues to find his way in an abstract world of sounds. The essential element is the visitor's freedom of moving in a relatively large action space.

The behaviour of our system would be even more interesting if new arriving sound data were fed directly into the system. This would require that all ten sound features be determined automatically. Then, to prevent the system from getting overloaded with agents, some sort of aging and a maximum lifetime could be added to the rules.

References

[Haken and Wunderlin, 1991] Haken H. and Wunderlin A., "Die Selbststrukturierung der Materie", Vieweg & Sohn, Braunschweig 1991

[Reynolds, 1989] Reynolds Craig W., "Flocks, Herds, and Schools: A Distributed Behavioral Model", Computer Graphics, vol. 21, No 4, 1989

[Goldstein, 1981] Goldstein H., "Classical Mechanics", 2nd edition, Addison Wesley, Reading, Massachusetts, 1981

[Bauer and Foss, 1992] Bauer W. and Foss B., "GAMS: An Integrated Media Controller System", Computer Musical Journal, vol. 16, No. 1, 1992

[3] MIDI: Musical Instruments Digital Interface