London: Chapman & Hall, pp. 38-42

6

DESIGNING ENVIRONMENTAL SOUNDS BASED ON THE RESULTS OF INTERACTION BETWEEN OBJECTS IN THE REAL WORLD

A. DARVISHI, E. MUNTEANU, V. GUGGIANA, H. SCHAUER
Department of Computer Science (IfI),University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland,
tel. +41-(1)-632 7098, e-mail: darvishi@ifi.unizh.ch

M. MOTAVALLI
Swiss Federal Labs for Material Testing and Research (EMPA),Uberlandstr. 129, CH-8600 Dubendorf, Switzerland

M. RAUTERBERG
Usability Laboratory, Work and Organizational Psychology Unit
Swiss Federal Institute of Technology (ETH), Nelkenstr. 11, CH-8092 Zurich, Switzerland

KEY WORDS: human computer interaction, auditory interfaces, sound synthesis, sound models, auditive feedback, usability engineering, multimedia, virtual reality, visual impairment.

ABSTRACT This paper presents an object oriented layered software architecture for describing and designing environmental (everyday) sounds in user interfaces based on a new sound model (audio framework). This new architecture is defined by different layers including the physical layer, system sound software layer, sound analyser/synthesiser layer and the interface layer. The sound model described can be used as the basic design of environmental sounds in user interfaces. This paper describes the different components: physical modelling, interaction, context sensitivity, and metaphorical description. In the paper the term *audio framework* is ultimately used for the sound model presented. This paper first provides an overview of existing approaches for modelling environmental sounds, then presents the new audio framework, a comparison between real and model generated sounds and finally discusses potential applications.

1. INTRODUCTION

This section describes various computer applications in which sounds are currently being used and discusses distinguishing dimensions for describing sounds. Some examples of such applications are:

- Data sonification/Scientific audiolization [Kramer, 1994], [Blattner, 1992]

- User interfaces [Gaver,1986] for the following tasks:

  • status and monitoring messages;
  • alarms and warning messages; [Momtahan,1993].
  • sounds as redundancy information to the visual displays to strengthen their semantics.
  • sound in collaborative work [Gaver,1991]
  • multimedia application [Blattner,1993]
  • visually impaired and blind computer users [Edwards, 1994].

Similary to light, sound has many different dimensions. Visual perception distinguishes dimensions such as colour, saturation, luminescence, and texture. Audition has an equally rich space. In the physical dimension (physical level) human can perceive differences in sound like pitch, timbre, and amplitude. There are also more complex so-called higher level dimensions (semantic description), e.g., the differentiation of interacting objects concerning their physical condition (state of aggregation (solid, liquid and gaseous), reverberance, locality, phase modulation, and others. Humans have a remarkable ability to detect and process minute changes in a sound along anyone of these dimensions [Rossing,1990]. The hearing of sounds in everyday life is based on the perception of events and not on the perception of sounds as such. For this reason, everyday sounds are often described by the events they are based on. Thus, the presented model for describing sounds (environmental sounds) in this paper offers a framework for a semantic description of environmental sounds.

The next section describes existing approaches for describing environmental sounds. The new audio framework, its components, the object oriented architecture of layered software system for its implementation and some potential applications are then introduced.

2. DIFFERENT APPROACHES FOR DESCRIBING ENVIRONMENTAL SOUNDS

The two figures below illustrate schematically two different approaches for designing environmental sounds:

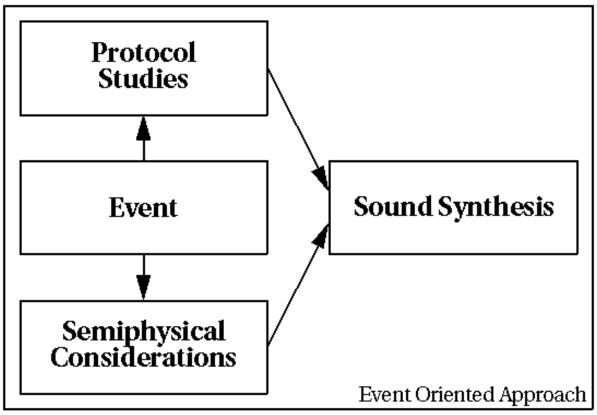

fig.1 Gaver's approach for modelling environmental sounds

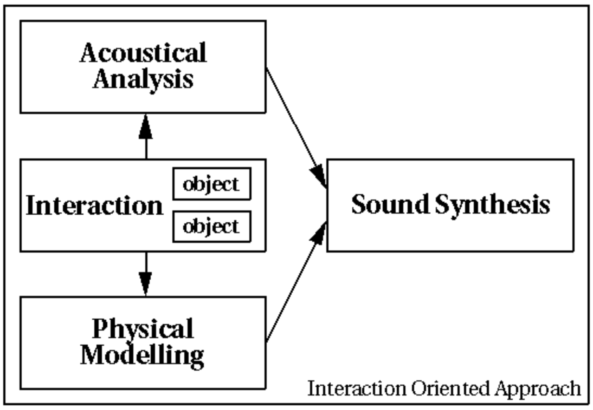

fig.2 Our approach for modelling environmental sounds

The first approach (event oriented) is used in the work of William Gaver [Gaver,1993]. It is based on perception of events in the real world and uses results of protocol studies and semi-physical considerations of objects (wood, metal etc.) to derive parameters for sound synthesis.

The second figure illustrates our (interaction oriented) approach [Rauterberg,1994]. It is based on the perception of interaction of objects in the real world and analyses real impact sounds and concurrently encounters physical modelling of interacting objects in order to derive appropriate parameters for sound synthesis. The recorded impact sounds were analysed via spectral analysis.

3. THE SUGGESTED MODEL (AUDIO FRAMEWORK) FOR DESCRIBING ENVIRONMENTAL SOUNDS

In this section and the following subsections are introduced different components of the new model. We have focused first on sound generation of *impact sounds* in particular the interaction of different spheres and beams. The interaction of these objects are analysed (sound analysis) and implemented (sound synthesis) on a SIG-Indigo workstation in the object oriented programming language OBERON.

3.1 Physical Modelling

General considerations. We start off by describing the physical models of simple interactions, e.g., the collision between an homogeneous and isotropic sphere and an homogeneous and isotropic plate/beam. For more information and detailed physical description of these objects and their interactions see [Rauterberg,1994]. On the one hand, the finite elements method can help us to simulate impacts for complex structures between more complicated objects. On the other hand, a second category of complex impact sounds (like bouncing, scraping, rolling, breaking, etc.) can be simulated by repeatedly reproducing simple impact sounds, by applying adequate time modulations.

The physical description of the behaviour of the plate or beam oscillations following the impact with the sphere provides variations in air pressure that we are able to hear. The sphere hits the plate or beam and stimulates vibrations with the natural frequencies. The natural frequencies of the small spheres are usually not in the audible range, therefore we disregard their contribution in the sound generating process. However, we are concerned about including the essential influence of the interaction on the impact sound in our simulations. This influence is especially important in the case of short impact sounds, where the vibrations are quickly attenuated. Therefore, it is necessary to take into consideration the transitory effect from the beginning of the sound generating process when the structure has a non-zero loading.

The following table illustrates the combination of different structures which are physically modelled .

Material	**Material shape**	**Mallet material**	**Mallet shape**
steel	- beam - plate	- steel - glass	sphere
aluminium	- beam - plate	- steel - glass	sphere
glass	- beam - plate	- steel - glass	sphere
Plexiglas	- beam - plate	- steel - glass	sphere
PVC	- beam - plate	- steel - glass	sphere
wood	- beam - plate	- steel - glass	sphere

Through physical modelling we are able to calculate the natural frequencies of the interacting objects, the shapes and initial amplitudes of the natural frequencies of the objects being hit.

3.2. Interaction

Every sound could be described as the result of one or several interactions between one or several objects at a specific place and in a specific environment. Each interaction has attributes, which influence the sound generated. Also, the participating objects, which take part in the sound generating process, can consist of different physical conditions (states of aggregation), various materials as well as having different configurations. Additionaly the materials themselves have attributes, which influence the generated sound.

An example of an interaction specific parameter is the height from which the sphere falls. Another example is the radius of the sphere used to hit the beam. The bigger the sphere the louder the perceived sound. These two examples emphasise the importance of interaction parameters in events producing sound. Both interaction parameters are implemented in our system.

3.3 Context Sensitivity

Sounds are context sensitive, i.e., the generated sounds differ depending on the environment in which the interaction of objects takes place and the combination of interacting objects. (The same sphere hitting a wood beam sounds different to hitting a steel beam.). Other examples of context sensitivity parameter are the co-ordinates of the place of impact. If the plate/beam is hit in the middle it sounds different to being hit on the edge.

3.4 Sound Metaphors

Basically, there are three ways of describing everyday sounds on a metaphorical level: (1) linguistic description of everyday sounds which are often ambiguous; (2) technical description of environmental sounds in terms of frequency, duration, timber etc.; (3) semantic description of environmental sounds in terms of interacting objects, their interaction and the environment. Sounds generated by our model fall into this last category. Further user tests are necessary to insure adequate identification of the sounds.

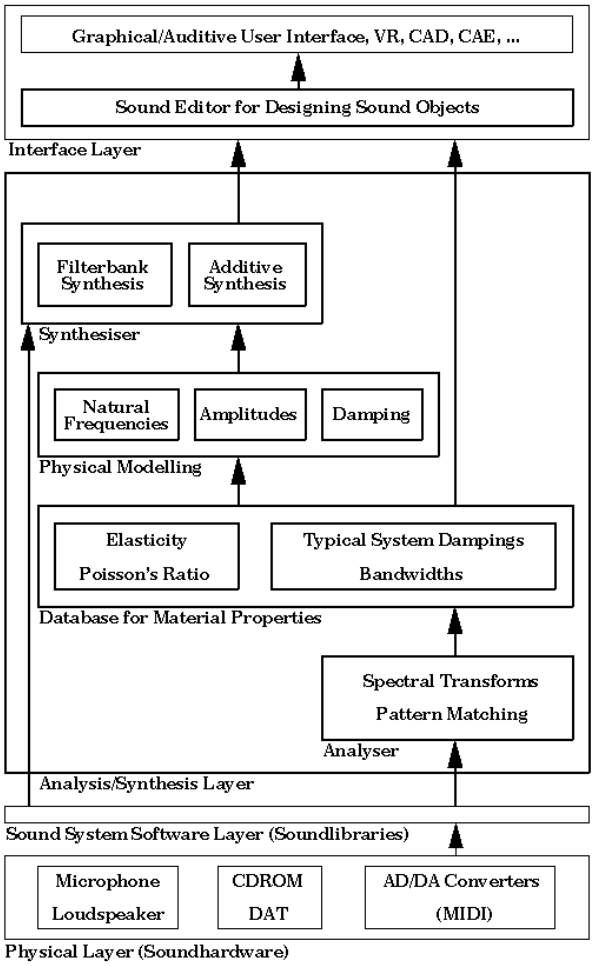

fig.3 The object oriented architecture of a layered software system for implementing the new model

4. THE OBJECT ORIENTED ARCHITECTURE OF A LAYERED SOFTWARE SYSTEM FOR IMPLEMENTING THE NEW MODEL

Physical Layer

This layer builds the required hardware for sound recording (Microphone, DAT, AD-converter etc.), sound processing (storage, CD-driver etc.) and sound generating devices (DA-converter, loudspeaker etc.).

Sound System Software Layer

This layer builds the interface to different sound hardware and supplies basic procedures for sound processing in the form of a software library.

Sound Analyser / Synthesiser Layer

This layer consists of several units (software tools) which fulfil different tasks. These software tools are:

Wave form editor. Represents the set of samples for each channel (for stereo signals,) on the screen and has all the standard functions of an editor, like *copy*, *paste*, *cut*, *save*, etc., as well as functions specific for our purposes (zoom, play, etc.)

Spectral analysis. The software makes the Fourier transformation of natural sounds through means of fast algorithms. One can specify the desired sound fragment to analyse in terms of the number of samples or in terms of time interval. In order to observe the spectrum's time evolution it was implemented a procedure that builds and draws a so-called spectrogram of the signal.

Parameter extraction. Fast algorithms were implemented which sweep the time axis and/or frequency axis to find the frequencies corresponding to the natural modes of vibration, the initial amplitudes for each of these waves, and the damping coefficients that describe the spectrum's evolution in time domain.

Data base system. This data base system stores two types of information: (1) different material properties; (2) formal description of system dampings of different materials which are derived from an analysis of real sound of designated materials.

Sound synthesis objects. Different objects are provided for sound synthesis. Currently we have implemented additive synthesis and filter bank algorithms for synthesising impact sounds. However new objects (other synthesis algorithms,) can be added to this layer. The synthesis algorithms receive natural frequencies of the vibrating objects, their initial amplitudes and the damping function for each frequency. For more information about synthesis algorithms see [Kramer,1994].

Physical modelling. The unit for the physical modelling of interacting objects takes as input the object definition from a graphic editor (see interface layer) and generates as output different natural frequencies of the vibrating objects and their initial amplitudes.

Interface Layer
This layer defines the interface to the user (software-developer) and offers an interactive editor for the definition of interacting objects, e.g., object type (wood, metal etc.), object shape (beams, plate etc.), and the environment (room etc.). The editor generates as output a meta-description of sounds which can be linked to a programming environment or which can be executed directly (real-time sound generation). The output of this layer is used as input for the underlying layer (sound synthesiser layer).

5. COMPARISON BETWEEN REAL SOUNDS AND SYNTHESIZED SOUNDS

As an example the following spectrograms illustrate the real (fig. 4) and model generated sounds (fig. 5) of the same interacting objects (steel sphere with steel beam). This proves that (at least) simple impact sounds can be generated through modelling as presented in this paper. Further investigation will be carried out to explore other possibilities and define the constrains of the introduced audio framework.

6. POTENTIAL APPLICATIONS OF THE NEW MODEL

The model generated sounds can be applied for different applications, e.g.:

Virtual Reality:
Most of the sounds used in current virtual reality applications are sampled sounds. Model generated sounds (everyday sounds) offer new possibilities in virtual reality applications.

Visually Impaired Computer Users:
Blind or visually impaired people make use of everyday sounds for orientation. The integration of everyday sounds in user interfaces introduces new ways for this community to work more effectively with computers. The use of software systems and applications, (for instance learning tools for training,) supplemented with everyday sounds become easier to use and more intuitive because these sounds are close to their mental model.

7. CONCLUSION

A new audio framework for describing and designing environmental sounds in user interfaces has been discussed in this paper. An object oriented, layered software architecture based on the proposed model was introduced. This layered architecture can easily be extended for describing new types of environmental sounds. The implementation of environmental sounds based on the audio framework offers new possibilities for the design of sound in virtual reality applications and special interfaces for blind or visually impaired computer users.

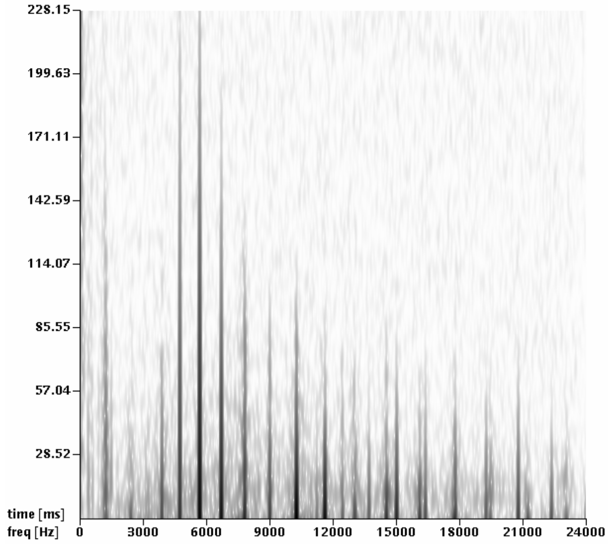

fig.4 Spectrum of the real sound

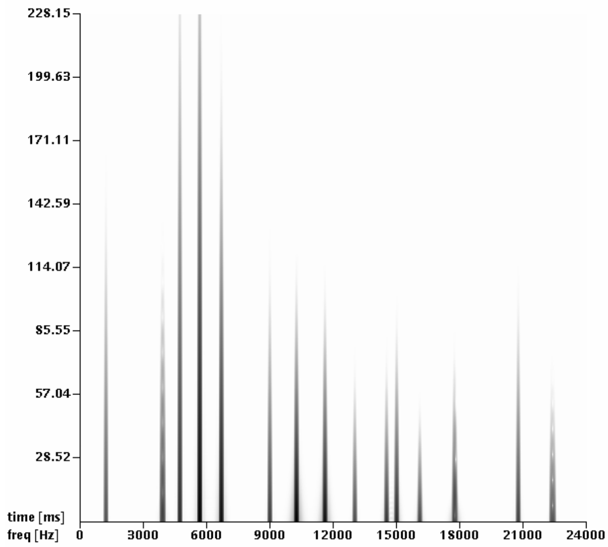

fig.5 Spectrum of the synthesized sound

REFERENCES

Astheimer, P.: What you see is what you hear - Acoustics applied to Virtual Worlds, *IEEE Symposium on Virtual Reality*, San Jose, USA, October 1993.

Blattner, M. M., Greenberg, R. M. and Kamegai, M. (1992) Listening to Turbulence: An Example of Scientific Audialization. *In: Multimedia Interface Design*, ACM Press/ Addison-Wesley, pp 87-102.

Blattner, M. M., G. Kramer, J. Smith, and E. Wenzel (1993) Effective Uses of Nonspeech Audio in Virtual Reality. *In: Proceedings of the IEEE Symposium on Research Frontiers in Virtual Reality*, San Jose, CA. (In conjunction with IEEE Visualization '93) October 25-26, 1993.

Bloyd, L.H., Boyd, W.L., and Vanderheiden, G.C. (1990) The graphical user interface: Crisis, danger and opportunity. *Journal of Visual Impairment and Blindness*, p.496-502.

Crispien, K. (1994) Graphische Benutzer-schnittstellen für blinde Rechnerbenutzer. Unpublished manuscript

Darvishi, A., Guggiana, V., Munteanu, E. Motavalli, M., Rauterberg, M., Schauer, H. (1994) "Synthesising Non-Speech Sounds to Support Blind and Visually Impaired Computer Users" *In: 4th Int. Conference on Computers for Handicapped Persons*, Vienna, pp. 385-393.

Darvishi, A., Munteanu, E., Guggiana, V., Motavalli, M., Rauterberg, M., Schauer, H. (1994) "Automatic Impact Sound Generation for use in Nonvisual Interfaces" *In Proceedings of 1st Annual International ACM/SIGCAPH Conference on Assistive Technologies*,, pp. 100-106.

W.K. Edwards, E.D. Mynatt, K. Stockton (1994), Providing Access to Graphical Users Interfaces- Not Graphical Screens. In: ASSETS`94, Marina del Rey, CA, USA, 1994, ACM ISBN 0-89791-649-2, pp: 47-54.

Edwards, W.K., Mynatt, E.D. and Rodriguez, T (1993) The Mercator Project: a non-visual interface to the X Window system. *The X Re-source*,4:1-20.(ftp multimedia.cc.gatech.edu /papers/Mercator /xresource).

Gaver, W. W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction.* 2, 167-177.

Gaver, W. W. (1988). Everyday listening and auditory icons. Doctoral Dissertation, University of California, San Diego.

Gaver, W. (1989) The SonicFinder: an interface that uses auditory icons. *Human Computer Interaction* 4:67-94.

Gaver, W. & Smith R. (1990) Auditory icons in large-scale collaborative environments. In: D. Diaper, D. Gilmore, G. Cockton & B. Shackel (eds.) Human-Computer Interaction - INTERACT'90. (pp. 735-740), Amsterdam: North-Holland.

Gaver, W., Smith, R. & O'Shea, T. (1991) Effective sounds in complex systems: the ARKola simulation. in S. Robertson, G. Olson & J. Olson (eds.), *Reaching through technology CHI'91*. (pp. 85-90), Reading MA: Addison-Wesley.

Gaver, W. (1993) What in the World do We Hear? An Ecological Approach to Auditory Event Perception. *Ecological Psychology*, 5(1): 1-29.

Kramer, G. Auditory Display - Sonification, Audification and Auditory Interfaces. *Proceedings Volume XVIII in the Santa Fe Institute Studies in the Sciences of Complexity*, Addison-Wesley, April 1994.

Momtahan, K., Hetu, R. and Tansley, B. (1993) Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics* 36(10): 1159-1176.

Mynatt, E.D. and Edwards, W.K. (1992) The Mercator Environment: A Nonvisual Interface to XWindows and Workstations.*GVU Tech Report GIT-GVU-92-05*

Mynatt, E.D. and Edwards, W.K. (1992) Mapping GUIs to auditory interfaces. In: *Proceedings of the ACM Symposium on User Interface Software and Technology UIST'92*.: 61-70

Rauterberg,M., Motavalli, M., Darvishi, A. & Schauer, H. (1994) Automatic sound generation for spherical objects hitting straight beams. *In: Proceedings of "World Conference on Educational Multimedia and Hypermedia"* ED-Media'94 held in Vancouver (C), June 25-29, 1994, pp: 468-473.

Rossing, T.D (1990) The Science of Sound 2nd Edition, Addison Wesley Publishing Company

Sumikawa, D. A., M. M. Blattner, K. I. Joy and R. M. Greenberg (1986) Guidelines for the Syntactic Design of Audio Cues in Computer Interfaces. In: *19th Annual Hawaii International Conference on System Sciences*.

Takala, T.& Hahn, J. (1992) Sound Rendering Computer Graphics, 26, 2,: 211-219.