REAL-TIME CONTROL OF SONIFICATION MODELS WITH A HAPTIC INTERFACE

Thomas Hermann, Jan Krause and Helge Ritter

Faculty of Technology
Bielefeld University, Germany
{thermann|jkrause|helge}@techfak.uni-bielefeld.de

ABSTRACT

This paper presents a new interface for controlling sonification models. A haptic controller interface is developed which allows both to manipulate a sonification model, e.g. by interacting with it and to provide a haptic data representation. A variety of input types are supported with a hand-sized interface, including shaking, squeezing, hammering, moving, rotating and accelerating. The paper presents details on the interface under development and demonstrates application of the device for controlling a sonification model. For this purpose, the Data-Solid Sonification Model is introduced, which provides an acoustic representation of the local neighborhood relations in high-dimensional datasets for binary classification problems. The model is parameterized by a reduced data representation obtained from a growing neural gas network. Sound examples are given to demonstrate the device and the sonification model.

1. INTRODUCTION

Within the research field of *data mining*, the aim is to detect structures in data, to uncover hidden regularities or patterns [1]. For the data mining process, the exploratory analysis of data constitutes an important part. Prior to validating to what degree a pattern is evident in given data, one needs to have a hypothesis [2] what pattern there might be at all. Exploratory Data Analysis (EDA) techniques help researchers to better understand their data (and thus to enable them to formulate hypotheses) by allowing them to "experience" their data in various ways. While most of the applied EDA techniques focus on data visualization, a rather new branch is concerned with developing *auditory representations* of data for this purpose [3].

Auditory perception has many things in common with visual perception but shows also several differences. One important difference is that we are able to focus on those parts of a visualization we are interested in, thus we may interactively control exploration. Many sonifications lack this possibility, since in most cases the sonification is rendered offline and only presented afterwards, similar to listening to a piece of music. Listening to such sounds, the only way for interaction is to focus our attention on certain parts or auditory streams within the sound. Certainly, our brain is very developed to perform this task, as a mixture of sound signals is a typical auditory input in environmental sounds.

However, there exists also an important class of sounds where we have greater control on their generation – interaction sounds. Most of our own actions (every footstep, every keystroke on the keyboard, etc.) causes an acoustic feedback and we are able to examine objects and learn about their material properties from their acoustic reaction on our interaction with them. For example, the sound while shaking a bottle filled with rice provides a lot of information about its content.

To exploit such interactions for the field of exploratory data analysis is one aim of the framework of Model-Based Sonification [4]. A sonification model describes a virtual sounding object and how it is parameterized from the data, and furthermore it introduces model-specific modes of interaction with the virtual sound object. In previous sonification models plucking and point excitations (hammering) were investigated due to the limitation that mouse clicks as excitation were used and the sonification was rendered offline. This paper presents a much more flexible and versatile interface for controlling and interacting with sonification models, a hand-sized audio-haptic ball. Besides providing a rich means of manipulation, the audio-haptic ball also offers the chance to be used as an output interface that provides a tactile data representation. Integration of the tactile components is subject of future work while the setup of the sensory part is currently proceeding and the working interface will be demonstrated at the ICAD 2002 conference.

The paper is structured as follows: Section 2 regards acoustical and haptic interaction in the real world and derives requirements for the interface. Section 3 gives an overview of Model-Based Sonification and introduces the Data-Solid Sonification Model to be used as a first application of the audio-haptic ball. In Section 4, the design of the interface and its technical specification is given. Section 5 will present some sonifications rendered by interaction with the audio-haptic ball during exploration of datasets from binary classification problems, and issues like ergonomics and usability will be addressed. The paper ends with a conclusion and prospects for future work.

2. AUDIO-HAPTIC INTERFACES

When we explore our environment or an object of interest, we make extensive use of our tactile sense. The notion of grasping in the sense of understanding and insight indicates how important our hands were and are to get comprehension. Basic manipulations which can be applied to real-world objects are moving, rotating, shaking, squeezing, rubbing, torquing, scratching on its surface and hammering on it. In combination with other objects, we can further position an object or attach one to another. As a result of such manipulations we usually get both haptic and acoustic feedback.

With our interface we aim at taking over as much as possible of these interactions to the control of virtual data-driven objects. Therefore it is necessary to record the user's actions by various sensors and to give tactile and auditory feedback to the user.

For the moving, shaking and rotating interaction the position and inclination of the interface device must be tracked at a high temporal resolution. The rotational degrees of freedom can be measured by inclination sensors directly. For measuring the position there exist different possibilities: Visual tracking of the ball device is not suited here due to likely occlusions by the hands, whereas other direct methods for instance by using pulsed magnetic fields and magnetic sensors [5] are difficult to calibrate. Therefore we decided to use an indirect measurement technique which integrates the signals of a 3D acceleration sensor to obtain relative position information.

Squeezing interaction is realized by six force sensitive resistors (FSR), one for each finger – for practical reasons, there are two FSRs for the thumb, one for right-handed and another for left-handed persons.

Hammering, scratching and rubbing interaction is recognized in a restricted area on top of the device only, because it was not feasible to cover the whole surface with sensors.The sensors being used are four piezo-electric elements which respond to mechanical excitation with a corresponding induced analog voltage. In addition to that we plan to integrate an industry standard mouse-pad which supplies 2D-position and finger-pressure output values on top of this area. With this combination it will be possible to sense spatially resolved hammering as well as using the pad for navigation in terms of the standard computer mouse.

Vibrations are able to inform the user about dynamic features of an object like surface texture, slip and impact. Vibrations within the frequency range of 2 – 200 Hz are suited for tactile output. We realize this with a small DC motor with an unbalanced weight attached to its axis.

For that our actions and the auditory and tactile feedback form a coherent entity, it is required that the output is rendered in real-time. A time lag in the order of magnitude of the sound wave run-time is of course permitted and provides a coarse orientation about the acceptable latency (sound waves travel 1 meter in about 3 msec).

Summarizing, we arrive at the following list of requirements for a suited audio-haptic ball device

- Size: Should be hand-sized with an ergonomic shape
- Sound Generation: Real-Time with short latency of about 5 msec
- Input Modalities:
  - 3D acceleration measurement (for high-resolution 3D-position tracking, shaking, moving)
  - 2D inclination measurement (for rotating the device)
  - Surface force measurement (for squeezing, hammering, scratching, rubbing interactions)
- Output Modalities: (so far only)
  - Vibration

Most of the existing haptic interfaces provide just one type of interaction. For example the “Squeezables” [6] from the MIT Media Lab have an interesting surface mounted force-sensor in form of an embroidery but do not offer interactions like shaking or position-tracking. Another class of haptic interfaces includes devices like the Pantograph developed by Hayward et. al. [7] or the Phantom from SensAble [8], a 6 DOF tactile input/output controller that allows the tactile exploration of a virtual surface or object. DiFilippo and Pai used the Pantograph to build an audio and haptic interface for contact interaction [9]. Compared to those interfaces our device contains more different types of sensors to map more types of interaction modalities allowing the user to manipulate a “virtual object” much more like it is possible with real objects.

3. MODEL-BASED SONIFICATION

The prevailing type of sonifications in literature are Parameter Mapping Sonifications [10], where data attributes are mapped onto sound attributes and sound events are superimposed to yield the sonification. The major disadvantage of Parameter Mapping is that a complicated mapping needs to be specified and be referred to in order to interprete the sound correctly. Besides this, Parameter Mapping lacks intuitive types of control and interaction. The framework of Model-Based Sonification supplies a solution to these problems by introducing a sonification model. The data here is not anymore mapped to sound attributes but used to parameterize a virtual sounding object. Exploration of the dataset is done by interacting with the model. Suited interaction types and the (acoustic) model reaction on it is part of the definition of a sonification model. Model-Based Sonification may be easier to “edit” at higher semantical levels since the model allows to “ground the semantics”. The model provides the key for interpreting the data with respect to the sound.

3.1. The Data-Solid Sonification Model

To demonstrate the application of the audio-haptic ball interface, we apply the *Data-Solid Sonification Model* which is also introduced in this paper as a new means to render acoustic representations for high-dimensional datasets.

In this model, a growing neural gas (GNG)[11] is used to obtain a reduced representation of the dataset. The neuron weight vectors can be interpreted as prototype positions in the high-dimensional data space. A sonification model using a dynamic growth process of a neural gas was already introduced and discussed in [12]. The neural gas is characterized by a set of neuron weight vectors and a list of pairwise undirected connections (edges) between the neurons. Both the neuron weights and the edges are adapted by a learning algorithm. In the Data-Solid Sonification Model, a virtual sounding object is constructed from the network graph as follows: The projection of all neuron vectors onto the first three principal components is used to determine the positions of corresponding point masses in a 3d data-solid model. In equilibrium, so without external activation, the model masses remain at their original position. However, the masses are able to move around their position and this motion is governed by dynamical laws given by Newton’s law of motion in model space assuming that each prototype object has a fixed point mass m attached to its original coordinate vector by a spring whose stiffness k is determined from local properties like the number of edges to topological neighbors. In addition the motion is damped with a friction coefficient γ. Let $\mathbf{x}$ be the coordinate of the mass, then the motion is determined by

$$m\ddot{\mathbf{x}}(t) + \gamma\dot{\mathbf{x}}(t) + k\mathbf{x}(t) = \mathbf{F}_{ext}(t) \ , \qquad (1)$$

where $\mathbf{F}_{ext}(t)$ is determined by interaction with the haptic interface. Figure 1 illustrates the model elements for a GNG with 9 neurons and 11 edges. Furthermore, each object is given a size (radius) determined from the data. Practically, this is achieved by mapping the average variance per dimension of all data being within the

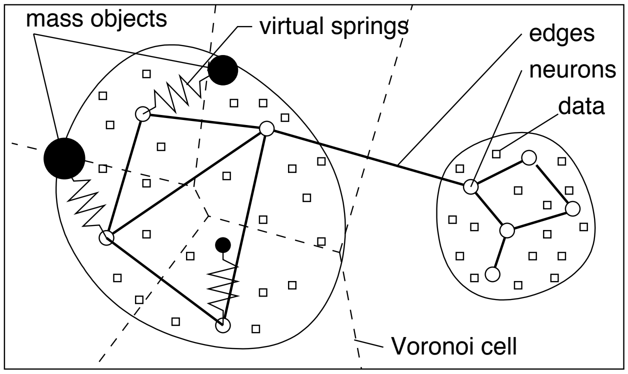

Figure 1: Illustration of Data-Solid Sonification Model.

Voronoi cell given by the neurons to the radius range $[0, 1]$. The stiffness k is directly proportional to the number of edges emanating from each neuron. The edges may thus be thought of fixing the mass tighter to its position. To keep the model simple, the masses as well as the friction constants are taken constant for all model objects. They are once adjusted manually so that the interface reacts in a reasonable manner.

Without any interaction, all objects will remain at rest at their center coordinates. As modes of interaction, the data-solid can be shaken which displaces all elements from their position.

An important event within the model dynamics is the collision of objects: these events cause the objects to emit sound events that provide information about the collision and the object itself. The events sound attributes are associated to the object properties in a rather intuitive way: the event level depends on the relative kinetic energy of the masses, the resonances (or the pitch) scales with the number of data points within the Voronoi cell of a model mass. A very important role for the current model is taken by the timbre. We assign an object's "material" as either plastics or wood, depending on the data within the voronoi cell. For the sounds presented within this paper, data from binary classification problems are used. The material type is wood if most data points within the Voronoi cell belong to the first class, else plastics. Thus collisions of objects of different types can be detected from the sound. As we are particulary interested in class boundaries, we use a noticeable timbre for collisions between model objects of different material: a glass-like sound with brilliant timbre[1]. There are many ways to integrate further information about the data into the sound events. For instance, we low-pass filter the object sounds so that collisions of particles close to their equilibrium position sound dull and collisions far away sound more brilliant. In contrast to Parameter Mapping sonifications, these assignments and adjustments have only to be done once and then the model can be used without any parameter tuning with all kinds of datasets of arbitrary dimensionality.

Besides shaking the data-solid, the user can currently squeeze the ball. This reduces the mean distance between the prototype objects, which results in more frequent collisions. The squeezing force thus controls the spatial resolution for system inspection. In many techniques for exploratory data analysis like principal curves or density estimation, smoothness parameters are needed to control the resolution. The squeezing interaction can for instance be used for such a purpose.

4. THE AUDIO-HAPTIC BALL

Figure 2 illustrates the design of the audio-haptic ball device. The device consists of various sensors which are integrated in a ball-shaped housing which fits into a human hand. Some sensors are inside the housing others are surface-mounted.

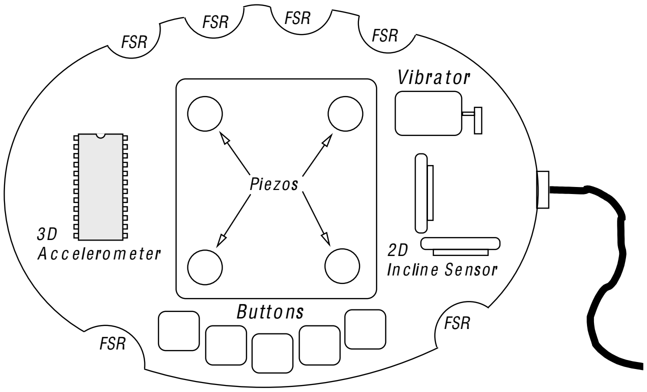

Figure 2: Scheme of the audio-haptic ball device. Shown are the various sensors and the coarse setup of the system. (FSR=force sensitive resistor)

The used sensors/actuators are:

Force Sensors: Six force sensitive resistors (FSR) are mounted in depressions on the surface so that the fingers fit naturally on top of them. To enable the use of the interface for both right-handed and left-handed persons, there are two FSRs for the thumb.

Acceleration Sensors: A 3D-on chip-accelerometer from Analog Devices (two 2d sensors, orthogonally mounted) is used to measure the forces that act upon the whole device. This is needed to realize position-tracking and to recognize shaking of the device.

Piezo Sensors: Four piezo-electric elements are used to build a plane pad that is able to recognize spatially resolved hits and scratches. The hitting position is inferred from the relative measured amplitudes at the four sensors. This element is currently implemented and will be shown at ICAD conference.

Incline Sensors: Inclination can be inferred from the 3d accelerometer sensors in the low-frequency limit.

Buttons/Switches: Four free programmable buttons are included to quickly access application-specific functions.

Vibrational Actuator: a small DC motor with an unbalanced weight is used to generate vibrations with adjustable amplitude and frequency.

Further Actuators: Further actuators are planned in a future version of the device.

The sensor data is preprocessed by a 16 bit RISC Microcontroller of the MSP430 Family from Texas Instruments which is hosted in an external housing together with the analog parts of the preprocessing electronics. Preprocessing includes filtering of the raw sensor data and combination of atomic data-events to macro-events like a rubbing-event with length, amplitude and frequency parameters.

---

[1] All those isolated sounds are accessible on our web site [13]

The MSP430 processors are ultra-low power devices so that a future version of the Audio-Haptic Ball possibly works wireless without the need of a cable connecting the device to a computer. For the data interfacing with the computer a special protocol was defined which allows to transmit the various sensor data in a time-multiplexed manner over a standard serial or parallel port. Special care was taken to minimize the latency to about 1 ms. So every sensor signal can be recorded at 1 kHz sampling rate.

A prototype of the interface devive housing was formed from *Efaplast* [14], an air-drying Plasticine. It was spread around a Styrofoam egg and shaped to show slots for the finger/thumb FSRs. The Plasticine egg was then cut into halves and the circuit board holding the sensors and some electronics were integrated. Figure 3 shows photographs of the current prototype.

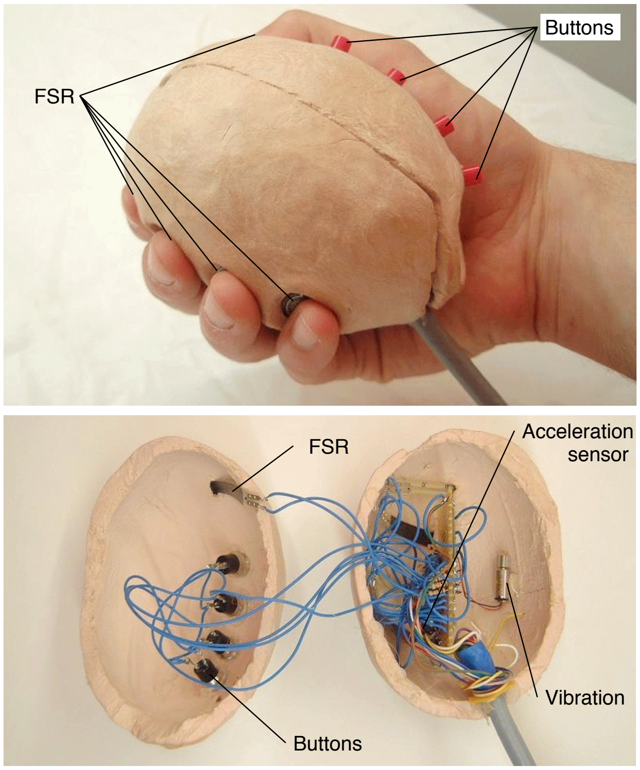

Figure 3: Photograph of the prototype of the audio-haptic ball device.

For interfacing with the data, a graphical programming and visualization environment called NEO [15] is applied. For sound synthesis, the "Real-Time Audio Toolkit for Sonification" (rats) is being developed. It is a C++-class library which is adapted to the specific needs for implementing sonification models. Parts of the architecture have been inspired by STK [16], the Synthesis Toolkit from Perry Cook. Rats allows for instance to model each mass as a C++ object that offers the required kinds of interactions and manipulations.

Figure 4 shows the measured sensor data from interacting with the audio-haptic ball. Some shaking actions along different directions can be seen; the third shaking was done while squeezing the ball with all fingers. At the end, all fingers are bend and a button is pressed.

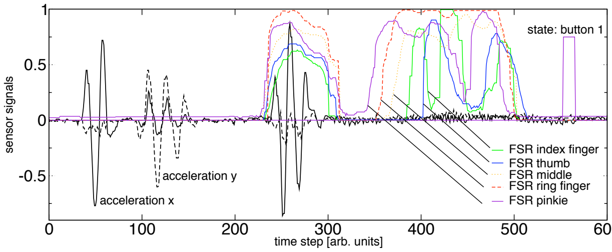

Figure 4: Sensor data during typical interaction with the ball.

5. APPLICATIONS

To demonstrate the new device, we present its application to controlling the Data-Solid Sonification Model introduced in Section 3.1. For testing we first used synthetic datasets for a binary classification problem. The data were drawn from spherical Gaussian distributions with different mean and covariance matrix. Figure 5 shows a scatter plot of a typical dataset used for the model. Besides the data points, the plot shows the network graph obtained from the Growing Neural Gas algorithm. The Data-Solid Sonification Model is used and may be excited by specifying the external force. This force is specified in a very intuitive way be shaking, moving or rotating the haptic ball. As a consequence of the user's actions, the ball reacts with the sound events in real-time. Besides the excitation force, the FSRs are currently used to specify a scale factor: the stronger the ball is squeezed, the smaller the "virtual container" is. This causes the collisions to become more frequent.

The following sonifications are obtained by shaking the interface. Unfortunately, the naturalness and experience of using the device is difficult to illustrate. Figure 6 shows a sensor plot of the acceleration sensors and a spectrogram of the resulting sonification. As the sound is directly related to the user's actions, it is

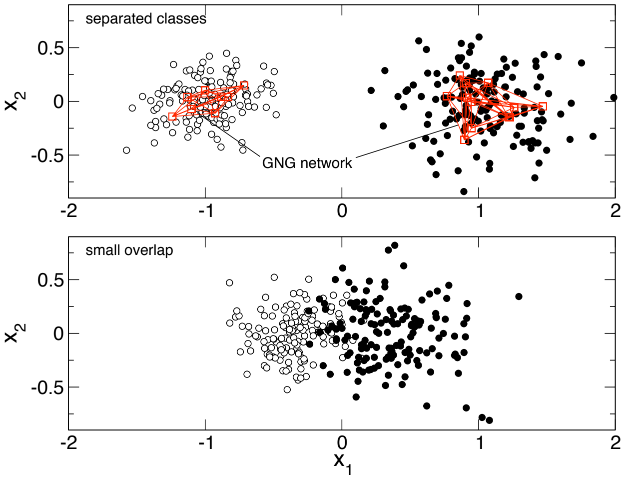

Figure 5: Datasets used to demonstrate sonifications for the Data-solid sonification model. Both plots show a projection onto the first two Principal components of the 6d data. The classes are separated in the upper plot and mixed in the second plot.

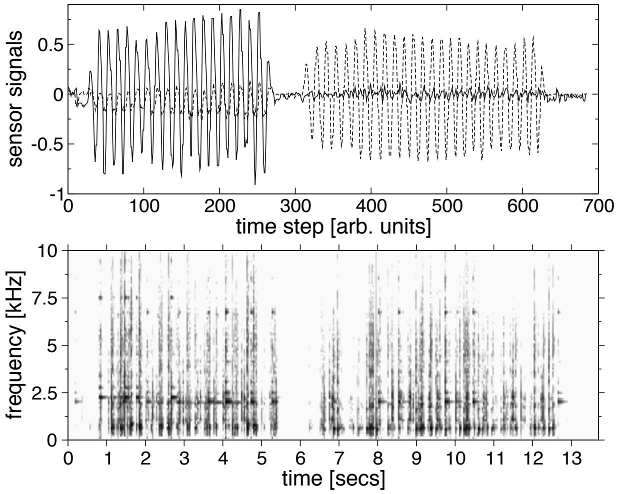

Figure 6: Sensor data and model sonification for the dataset where two classes show a small overlap. Shaking along the x axis causes much more collisions between neurons that belong to different a class.

easily understood. In addition it is much less disturbing according our experiences than sounds that are rendered offline and played without real-time control.

The following sound examples present sonifications for three datasets (all examples are provided on our web site [13]):

- two classes that are well separated: no glass-like sound events occur during the shaking interaction. This indicates that the classes are well separated.
- two classes that have a small overlap. The glass-like sound events occur more frequently while shaking along the x-axis than shaking along the y-axis. This indicates that the classes may be separated perpendicular to this shaking axis.
- two classes which are overlapping (gaussians with same mean but different variance): many glass-like sounds are heard. The classes obviously mix.

Another example presented on the web site [13] shows how sound properties of the data-solid change with increasing model complexity, resp. with the number of neurons. Since less and less data are found within a neuron's voronoi cell, the pitch decreases. Thus the interface might also be used to query and trace the progress of machine learning processes.

6. CONCLUSION

The audio-haptic ball device provides a rich interface for the manipulation of virtual objects. It implements basic interaction types that humans routinely apply to examine objects. Using this interface to control sonification models offers the chance to improve the experience of data, and facilitate manipulation of computer simulated objects. To demonstrate the audio-haptic ball interface, the Data-Solid Sonification Model was introduced. It allows the user to probe the data in a dynamical way, by shaking, striking and squeezing it.

With this design, the sound and the actions of the user are brought together. However, the visual interface is still static. Integrating a dynamical visualization and dynamical controls both on the graphical and acoustically presented information by using the audio-haptic ball interface seems to be very promising and will be followed in future work.

7. REFERENCES

[1] U. M. Fayyad et al., Ed., *Advances in Knowledge Discovery and Data Mining*, MIT Press, 1996.

[2] J. W. Tukey, *Exploratory Data Analysis*, Addison-Wesley, 1977.

[3] G. Kramer, Ed., *Auditory Display - Sonification, Audification, and Auditory Interfaces*. Addison-Wesley, 1994.

[4] T. Hermann and H. Ritter, "Listen to your Data: Model-Based Sonification for Data Analysis," in *Advances in intelligent computing and mulimedia systems*, M. R. Syed, Ed. 1999, Int. Inst. for Advanced Studies in System Research and Cybernetics.

[5] Ascension Technology Corporation, "Flock of birds," `http://www.ascension-tech.com/products/flockofbirds/`.

[6] G. Weinberg, M. Orth, and P. Russo, "The embroidered musical ball: A squeezable instrument for expressive performance," in *Proceedings of CHI 2000*, The Hague, 2000, ACM Press.

[7] V. Hayward and C. Ramstein, "The pantograph: a large workspace haptic device for a muti.modal human-computer interaction," in *Conference on Human Factors in Computing Systems ACM/SIGCHI'94*, 1994, `http://www.cim.mcgill.ca/@haptic/pub/CR-VH-CHI-94.pdf`.

[8] SensAble, "Phantom," `http://www.sensable.com/haptics/products/phantom.html`.

[9] D. DiFilippo and D. K. Pai, "The ahi: An audio and haptic interface for contact interactions," in *Proceedings of UIST'00*, 11 2000, `http://www.cs.ubc.ca/spider/pai/interfaces.htm`.

[10] C. Scaletti, "Sound synthesis algorithms for auditory data representations," in *Auditory Display*, G. Kramer, Ed. 1994, Addison-Wesley.

[11] Bernd Fritzke, "A growing neural gas network learns topologies," in *Advances in Neural Information Processing Systems*, G. Tesauro, D. Touretzky, and T. Leen, Eds. 1995, vol. 7, pp. 625–632, The MIT Press.

[12] Thomas Hermann, *Sonification for Exploratory Data Analysis*, Ph.D. thesis, Bielefeld University, Bielefeld, 2 2002.

[13] Thomas Hermann, "Sonification for exploratory data analysis – demonstrations and sound examples," `http://www.techfak.uni-bielefeld.de/~thermann/projects/index.html`, 2002.

[14] Eberhardt Faber, "Efaplast modelling material," `http://www.eberhardfaber.de/efaplast_en.htm`.

[15] H. Ritter, "The graphical simulation toolkit neo/nst," `http://www.techfak.uni-bielefeld.de/ags/ni/projects/simulation_and_visu%al/neo/neo_e.html`, 2000.

[16] P. R. Cook, "Synthesis toolkit in c++," SIGGRAPH 1996, `http://www.cs.princeton.edu/~prc/STKPaper.ps`, 1996.