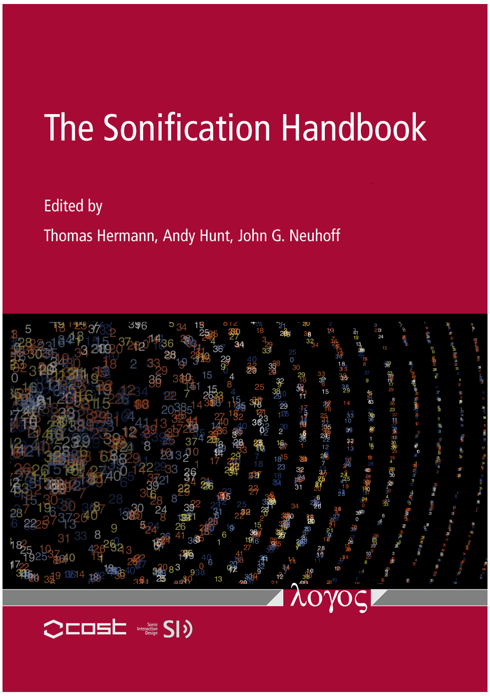

The Sonification Handbook

Edited by
Thomas Hermann, Andy Hunt, John G. Neuhoff

Logos Publishing House, Berlin, Germany
ISBN 978-3-8325-2819-5
2011, 586 pages

Online: http://sonification.de/handbook
Order: http://www.logos-verlag.com

Reference:
Hermann, T., Hunt, A., Neuhoff, J. G., editors (2011). *The Sonification Handbook*. Logos Publishing House, Berlin, Germany.

Chapter 16

Model-Based Sonification

Thomas Hermann

This chapter introduces the sonification technique of Model-Based Sonification. Guidelines for model design are given, and several different example sonification models are explained and shown with sound examples. The advantages and differences to other sonification methods and in particular Parameter-Mapping Sonification are discussed.

Reference:
Hermann, T. (2011). Model-based sonification. In Hermann, T., Hunt, A., Neuhoff, J. G., editors, The Sonification Handbook, chapter 16, pages 399–427. Logos Publishing House, Berlin, Germany.

Media examples: http://sonification.de/handbook/chapters/chapter16

Chapter 16

Model-Based Sonification

Thomas Hermann

16.1 Introduction

Almost every human activity in the world is accompanied with an acoustic response. Interaction in the world typically provides us with rich feedback about the nature of the involved materials, as well as the strength and type of contact. It is stunning that, despite the ubiquity of action-driven informative sounds, we have tended to limit traditional computer interfaces to visual-only displays. *Model-Based Sonification* is a sonification technique that takes a particular look at how acoustic responses are generated in response to the user's actions, and offers a framework to govern how these insights can be carried over to data sonification. As a result, Model-Based Sonification demands the creation of *processes* that involve the data in a systematic way, and that are capable of evolving in time to generate an acoustic signal. A *sonification model* is the set of instructions for the creation of such a "virtual sound-capable system" and for how to interact with it. Sonification models remain typically silent in the absence of excitation, and start to change according to their dynamics only when a user interacts with them. The acoustic response, or sonification, is directly linked to the temporal evolution of the model.

Model-Based Sonification has been introduced by the author [14] and was elaborated in more detail [10]. Several sonification models have been developed since then [21, 27, 4, 23, 18, 3, 16, 12, 15, 10, 11, 13, 14], which give examples for model design, exploration tasks in the context of exploratory data analysis and interaction modes.

This chapter gives a full introduction to Model-Based Sonification (MBS), including its definition, some design guidelines, description of selected sonification models and a discussion of the benefits and problems of MBS in general. Since MBS is a *conceptually* different approach than Audification and Parameter Mapping Sonification, its relation to these will be addressed in detail. Finally, a research agenda for MBS is formulated.

16.1.1 Listening modes

A very helpful experiment to understand how the human auditory system works, is to play a short example sound and ask listeners to describe in as much detail as they can what they have heard. As experiment the reader might try this now with sound example **S16.1**. Please stop reading here until you have listened to the sound, and be as accurate as possible and write down keywords of your description. Done?

Most listeners will now have characterized the sound by a guess of what *source* or *action* might have caused the sound. For instance, you may have described the sound as ‘somebody is coughing’, ‘surely a male, and not a child’, ‘it sounds like bronchitis’, and so on. Such descriptions are very typical and we are not aware of how dominating this source-identification default is. Let us call this listening mode *everyday listening*.

There is, however an alternative way to characterize the example, as ‘a sequence of 7 noise bursts’, ’their roughness and loudness decreases’, ’they form a certain rhythmical pattern’, and so on, characterizing the sound by its acoustic shape, its rhythm, harmony, melody, pattern, structure, etc. Such a description is just as valid as the one given from everyday listening, only the focus is different: rather than focussing on the signified it describes the sign itself. [1]. Let us call this *musical listening*. We can indeed experience our world with ‘other ears’ just by purposefully changing our listening mode.

Obviously, our brain and auditory system is capable of operating in different modes, and ‘everyday listening’ is the dominant or default mode. This is possibly because an accurate and quick sound source identification was evolutionarily advantageous since it enabled quick and correct *reaction*, e.g., to choose to flight or fight [17]. This argumentation would at least explain why our brain is specifically good at interpreting the sound source and source characteristics with a focus on the appropriate reaction rather than on conscious reflection.

There is yet another mode of listening, which we may call *analytical everyday listening, see listening, modes of*: this is the conscious use of all listening skills to distinguish and analyze an object under investigation. To give some examples, think of the task of determining the contents of an opaque box by shaking it, or the task of diagnosing a malfunctioning car engine from the sounds it makes. Such analytical listening is a ‘diagnostic’ use of listening, and thus most inspiring to be used for sonification.

The above list of listening modes is certainly incomplete. For instance the particular modes of listening to language and speech sounds have not been mentioned, or the enjoyment mode when listening to music. A discussion on listening modes can be found in [7] and [17]. Model-Based Sonification addresses our everyday listening and analytical listening skills. In the following section we categorize functions and contexts of sounds in order to better understand how information is encoded into sounds in our physical world.

16.1.2 Sound and Information

The sounds that we have heard in our lives can be categorized in the following classes:

Passive sounds: sounds that come from an external source, not directly caused by our own activity. These sounds give us information about the environment (e.g., a sense of

[1] more on semiotics can be found in chapter 18

where we are), and may direct our attention or even alert us.

Active sounds: sounds that are created in the course of physical activity, which directly accompany the owner's actions. Examples are the rustle of clothes while moving, the clip-clop of footsteps, the soft hiss of breathing, or contact sounds in response to direct or indirect manipulation of physical objects.

There is no strict separation between these classes as, for instance, actions may cause passive sounds. Also, other people's active sounds are indeed passive sounds for us as listeners. Most active sounds are a by-product of the activity and not its goal. As a special case we can identify *intentional active sounds* as active sounds where the subject has performed the (inter-)action intentionally in order to create the sound. Playing musical instruments, shaking an opaque box in order to learn about its content by listening, and clapping the hands to understand the surrounding reverberation characteristics are some examples for such intentional interactions.

Language sounds and musical sounds are highly specific to a cultural tradition, and the relation between the sounds and their meaning are largely learned or memorized bindings. The semantics of sound on the more basic level of environmental sounds and interaction sounds, however, is more universal. Sonification techniques that rely on sounds which the typical human is likely to have encountered are likely to be more culturally independent. For this reason, we now take a closer look at how information is encoded in real-world acoustics or physical sounds.

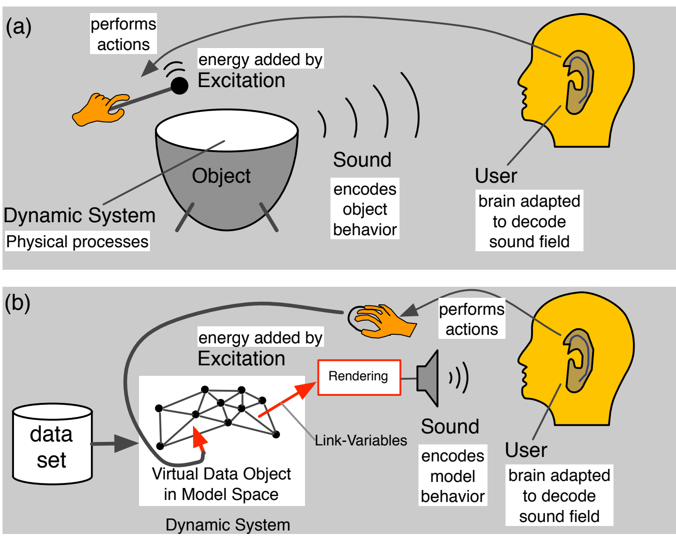

Figure 16.1: (a) Sonic loop in physical interaction: the user is tightly integrated into a closed-loop. The brain is adapted to interpret sonic patterns for source properties and to explore sound dependencies for the given excitation. (b) shows the modifications from real-world sonic loops for a typical Model-Based Sonification, as explained in the text.

Fig. 16.1 (a) illustrates a sonic loop from physical interaction with an object to the interpretation of the sound in the listener's mind. As starting point let us investigate the relation between the sound-capable object and the listener: audible sound is simply a pattern of vibration in a suitable frequency range (20 Hz to 16000 Hz), and we typically perceive sound since it is transported to our ears via sound waves that propagate through air. The detailed pattern of pressure variation, however, is a direct but complex image (or copy) of the vibrating object. The encoding of information into the wave field is neither unique nor invertible: different objects may lead to identical sound signals. Equally, an identical object may under repeated interaction also create slightly different sound responses. How do source properties then relate to the sound signals' There is unfortunately no simple answer to that question. A structural change of the physical object will typically lead to completely different sound signals, so that we may assume that the source properties are holistically encoded into the sound wave field. In addition, the sound will change sensitively with any change of the interaction.

It seems hopeless and overly complex to quickly obtain an inverse estimation of source properties from such distributed information. Bregman compares it to the task of estimating the number of ships, their position and velocity by simply looking at the fluctuations in the water waves in a pair of one meter long channels dug into the beach at the edge of a pond [6]. It would seem impossible to answer these questions from visually observing the water levels going up and down. However, the example is an analogy for human listening with the channels representing our ear canals. With our auditory systems we find that such inverse mappings (which infer source properties via incoming sound signals) are perfectly feasible, and our listening system has even been optimized to infer source-relevant information. Experiments have, for instance, demonstrated that material, size and rolling speed of balls on a surface can be perceived quite accurately [19]. For sonification, we can thus hope to exploit these *inverse mapping skills* to understand systems and in turn the underlying data. Physical systems as shown in Fig. 16.1 typically possess dissipative elements. Internal friction and the radiation of sound waves cause energy loss which makes physical systems converge towards a state of equilibrium. Since in this state there is no more vibration, the sound fades to silence. Often systems are excited and thus perturbed actively from their state of equilibrium by our own interaction. We can think of interaction as actively querying the object, which answers with sounds. Since we can reproduce sounds by repeated interaction, and thereby understand the systematic changes in sound that correlate with our change of excitation, we can gradually build up a mental representation which enables the miraculous inverse mapping from sound to interaction.

In summary, everyday sounds often stem from a closed-loop system where interactions are followed by physical/acoustic reactions which then lead to auditory perceptions and their cognitive interpretation. The human is tightly embedded in the loop and assesses source properties via the indirect *holistic* encoding in action-driven sounds.

16.1.3 Conclusions for Sonification

If we take the abovementioned observations from real-world sonic interactions seriously, there are several consequences for inherently interactive data sonifications:

ubiquity: almost every interaction with data should be accompanied by sound (as almost any interaction with the world causes some sound).

invariance of binding mechanism: the sound-producing laws should be invariant and structurally independent of the actual data – in the same way that the laws of physics and their invariance means that we can understand different objects in the world by attending to their sounds when we interact with them.

immediate response: sonifications should deliver an immediate (real-time) response to the interaction since this is the action-perception pattern we are familiar with from real-world interaction. The brain is tuned to interpret sound in this way; it is even optimized to associate synchronization between different modalities, e.g. our proprioception, visual changes and correlating acoustic patterns.

sonic variability: sonifications should depend on a subtle level on the interaction and data, in the same way that real-world sounds are never strictly identical at the sample level on repeated interaction, but depend very much on the actual dynamic state and the details of excitation.

information richness: sonifications should be 'non-trivial'. In other words they should be complex and rich on different layers of information. This is similar to the way that everyday sounds are complex, due to nonlinearities in the physical systems which produce them. It seems that the human brain expects this 'non-trivialness' and values it highly. If it is missing, the sounds may be perceived as boring, or just may not connect as well as possible with our auditory listening skills.

Model-Based Sonification offers a framework for the creation of sonification models which *automatically* behave according to these requirements, which underly sound generation in the real-world, as depicted in Fig. 16.1. How this is achieved is described in detail in the following section.

16.2 Definition of Model-Based Sonification

Model-Based Sonification (MBS) is defined as the general term for all concrete sonification techniques that make use of dynamic models which mathematically describe the evolution of a system in time, parameterize and configure them during initialization with the available data and offer interaction/excitation modes to the user as the interface to actively query sonic responses which depend systematically upon the temporal evolution model. In this section we will review the different 'ingredients' or elements of this complex and lengthy definition step-by-step. Hopefully this will clarify what is meant and how MBS is generally different from mapping sonification.

Model-Based Sonification (MBS) is the general framework or paradigm for how to define, design and implement specific, task-oriented sonification techniques. A specific design or instance obtained with MBS is called a *sonification model*. Model-Based Sonifications draw inspiration from physics, yet the designer is free to specify otherwise and may even invent non-physical dynamic models. A good procedure for the design of sonification models according to the MBS framework is given by the step-by step definition of the following six components: *setup*, *dynamics*, *excitation*, *initial state*, *link-variables*, and *listener characteristics*, which will be described in turn.

These steps are illustrated by using a simple MBS sonification model called *data sonograms*. In a nutshell, the data sonogram sonification model allows the user to excite a shock wave in data space that slowly propagates spherically. The wave-front in turn excites mass-spring systems attached at locations specified by each data point's coordinates . Fig. 16.3 on page 409 illustrates this setup. Using this sonification users can experience the spatial organization of data and how data density changes relative to the shock wave excitation center. While this sonification model is helpful for a MBS tutorial it should be emphasized that it is only one particular example model – other models can be structurally very different, as will hopefully become clear in section 16.3.

Model-Based Sonification mediates between data and sound by means of a dynamic model. The data neither determine the sound signal (as in audification) nor features of the sound (as in parameter mapping sonification), but instead they determine *the architecture of a 'dynamic' model* which in turn generates sound. Thereby MBS introduces the model space between the data space and the sound space, as depicted in Fig. 16.2.

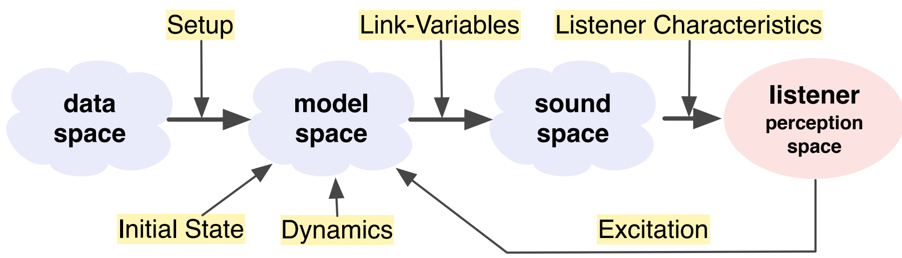

Figure 16.2: Transformations from data space via model space to sound space and to the listener's perception space. The elementary model specification steps are depicted at the location where they provide their specification.

16.2.1 MBS step 1: Model Setup

The model setup determines how data define the configuration of a dynamic system with internal degrees of freedom. The setup bridges the gap between the immaterial, abstract and static world of high-dimensional data and the more tangible world of a dynamic model where elements move in time and thereby cause the sound. It is helpful to distinguish between the data space and the model space. This may become clearer with a concrete example.

For example, assume that a d-dimensional data set with N records is given. The data set can then be represented as a table of N rows with d columns where each column is a feature and each row an instance or sample. In a census data set the columns could for instance be 'income', 'size', or 'sex', and the rows would be different persons. The cell values would then hold information such as 'this particular person's income in euros' etc. A frequently used representation in mathematics is that the data set defines a cloud of N points in the d-dimensional feature vector space, using the feature values as coordinates. So we can imagine the data space as a mathematical vector space.

With this representation in mind, a spatial setup of the model space is tempting. In the example of data sonograms the model setup is defined so that point masses are attached to springs so that they can oscillate or collide with each other. In this case the model space is also spatial, and its dimension may be chosen by the designer. For a d-dimensional data set, however, it is straightforward to create a model space of same dimensionality. Still, it remains the question of how the data vectors should determine these mass-spring elements in model space. Data sonograms use the data vector coordinates as the location vectors of the point to which the spring is attached. Of course, there are manifold different possibilities of how to connect the data space and the model space, examples of which are given in section 16.3.

16.2.2 MBS step 2: Model Dynamics

The ultimate goal is to get a sound signal which is a sonic representation of the data under analysis. Since sound evolves in time it makes sense to introduce a temporal evolution to the model, called *model dynamics*. More precisely, dynamics refers to the equations of motion that describe how the system's state vector changes in time, how the next state $\vec{s}(t + \Delta t)$ is computed from the actual state $\vec{s}(t)$. Since we are dealing with a mathematical description of the model, the equations of motion are usually specified as differential equations, similar to the equations of motion that describe how a mechanical system changes with time. Certainly other laws from electrodynamics, chemistry, or even machine learning are sometimes useful.

For the data sonogram model example where point masses are attached to springs, we need to specify how to update the position and velocities of each mass when the springs exert an actual force to the mass. The dynamics are given by equations $\vec{s}(t + \Delta t) = f(\vec{s}(t))$ which are inspired from physics and the mechanics of spring-mass systems.

Models may need several mechanisms of dynamic behavior. For the data sonogram model, for instance, we need dynamical laws that describe how excitation causes shock waves and how these waves propagate in model space, or how they interact with mass-spring systems. Other mechanisms such as energy flow are presented in section 16.3.

Physical principles such as kinetic and potential energy, and furthermore dissipation mechanisms such as friction, and specifically principles from acoustics provide rich inspiration on how to introduce dynamics that create a specific qualitative behavior. Not only do model developers need to specify the equations of motion, but most dynamical laws demand the inclusion of certain parameters that need to be adjusted. The parameter choice seems to be a source of arbitrariness, yet this is not really a problem if the parameters remain unchanged whatever data set is explored. Then the listener can adapt to the specific sonifications that are implied with the given dynamics and parameter settings.

In addition, the number of parameters is normally much lower than those needed for the specification of a parameter mapping sonification, and furthermore they also have a clear 'physical' meaning with respect to the model, which makes it easy to understand how their change affects the sound. This will be elaborated later in section 16.7.

16.2.3 MBS step 3: Model Excitation

Excitation is a key element in MBS, since it defines how the users interact with the model.

In acoustic systems physical objects (e.g., a bell) eventually come to rest in a state of equilibrium without external excitation. In a similar way the dynamics of sonification models often contain a term which leads to a system state of equilibrium. Excitatory interactions allow the users to feed new energy into the system and in turn users experience the acoustic reaction as a direct response. Firstly this prevents never-ending sound which would be annoying after some time. Secondly it enables the users to bring in rich manual interaction skills to examine a system. Think for instance of how many ways there are to shake, squeeze, tilt, incline, hit, etc. an opaque box to probe its content; such interaction can then be defined for use in interacting with sonification models.

Formally, excitation can be modeled as an external force in the equations of motion, which depends on the state of controllers or input devices. In the data sonogram example, a mouse click triggers a shock wave in model space, but other interactions are possible. For instance, the shaking of the mouse could inject energy into all spring-mass systems within a certain radius simultaneously.

Excitation type can range from elementary triggering (e.g., a mouse click or keystroke) through more detailed punctual interactions such as hitting a surface at specific locations with a certain velocity, to continuous interactions such as squeezing, shaking, rubbing or deforming controllers or tangible objects. Certainly, a mixture of these interactions may occur, depending on the interfaces used.

The better the metaphor binds interaction to the sonification model, the more the users will be capable of developing intuition about model properties, and understanding how these manifest in the resulting sonic response. Therefore, the specification of excitation cannot be done without keeping in mind the bigger picture and the idea of the sonification model.

Besides the mandatory excitation modes, there may be additional interface-to-model couplings that allow users to influence the dynamics. In real life a bottle filled partially with water sounds different when hit at various locations while changing the bottle's orientation. In a similar vein it may make sense, for instance, to allow the user to excite the sonification model at one location while controlling other parameters by rotating or squeezing a controller etc. Such excitation via parameter-rich interfaces brings the users more tightly in touch with the model and allows them to make use of their already available interaction competence from real-world interactions.

16.2.4 MBS step 4: Initial State

The initial state describes the configuration of the sonification model directly after setup. One's first thought might be that this has already happened during the Model Setup phase, yet that merely defines the system and how data are used to determine the architecture of the sonification model. For instance, in the data sonogram sonification model, the data vectors determine the location that the springs are attached to, whereas the initial state would determine the initial location and velocities of the point masses. In other words the Setup phase actually creates the model and then the initialization stage puts it into position ready for the first user interaction.

Normally, the designer knows – from insight into the equations of motion – the equilibrium state and initializes the system accordingly. If this is not possible, that is not a problem since the model will anyway relax from a random initial state to an equilibrium state, assuming that there is built-in dissipation. To prevent disturbingly loud noises, however, it is strongly advisable in this case to mute the audio output until the system has relaxed a bit.

16.2.5 MBS step 5: Model Link-Variables

Link-variables are the 'glue' which connects the model's dynamic processes to sound as shown in Fig. 16.2. In the most straightforward manner, the model's state variables can be used directly as a sound signal, which would be a good and direct analogy to how sound is generated in real-world acoustic systems. Think, for instance, of a drum head whose movement describes more or less one-to-one the sound signal that propagates to the ear. Expressed in terms of sonification techniques such a direct connection of a dynamic state variable with a sound signal could be called *audification* of the model dynamics. Sometimes it is more useful to condense several state variables $x_1(t), \ldots, x_n(t)$ into a single sound signal $s(t)$ by means of a feature function $s(t) = f(x_1(t), \ldots, x_n(t))$. For instance, in the particle trajectory sonification model explained in section 16.3 the kinetic energy of each particle is used as a link variable for the sound signal.

For some sonification models, the designer may consider the linking of state variables in a more complex or indirect way to the sound signal. For instance, the designer might want to map the overall model energy to the sound level. Such explicit parameter mappings can occur in MBS model design, and even help to make sound computation more efficient, yet they introduce a level of arbitrariness and the need for explanation which MBS design principles suggest keeping at a minimum.

One main problem of Model-Based Sonification is that the computation of tens of thousands of update steps necessary to generate even one second of a sonification is complex and time-intensive and even with current computing power in 2011 this is beyond real-time rendition even for moderately large systems. The reason is that the equations of motion may be coupled and demand the computation of the distances to all elements (e.g., masses in the model space) for each single update step of each mass, which leads to an explosion of the number of operations with increasing number of elements. However, real-time computation is crucial for MBS to tightly close the interaction loop. For that reason, implementation shortcuts are often used, which decouple the model update from the sound signal generation to some extent. For the data sonogram example, instead of computing the detailed motion of the mass-spring-system at 44100 steps per second[2], it may suffice to compute the average energy of a mass spring system at 50 Hz and to apply sample-based interpolation between successive amplitude values of an appropriately tuned sine generator. The result may be an acceptable approximation of the *real* model output with a reduced number of operations per second. Similar implementation shortcuts are necessary for many sonification models to reach real-time computability, yet it is most likely that with increasing computing power in a few years they can be minimized or avoided. Actually, while such shortcut procedures may be fine on first sight, they may just cut out subtleties in the sound signals which our ears demand and are tuned to pick up. More examples for implementation shortcuts will be given in section 16.3.

[2]to render CD quality signals at 44100 Hz sampling frequency

16.2.6 MBS step 6: Listener Characteristics

In everyday interaction with sounding objects and environments we either experience an object as a single sound source (e.g., knocking on a melon), or we experience ourselves embedded into a distributed soundscape (e.g., birds in the forest). In the same sense there are sonification models where the suitable metaphor is that the model forms a single sounding object or that the users are located and embedded in a space with the model elements around them. Let us distinguish these types as *microscopic* vs. *macroscopic* sonification models.

Listener Characteristics addresses all issues related to location, orientation or distance between sound sources (link-variables) and the user/listener. Spatial (macroscopic) models usually demand a more complex rendition and sound output, either using multi-channel audio systems or HTRF-convolution [3]. Furthermore they may need head-tracking to achieve a truly convincing spatial model experience. In contrast, the microscopic sonification models are much simpler yet may nonetheless deliver the majority of the information. The metaphor is that the whole model becomes a single sounding object.

For the data sonogram sonification model, the listener is assumed to be located at the shock wave center, so this is a macroscopic sonification model. In a stereo sound setup, it makes sense to play spring-mass sound contributions with stereo panning using the orientation of the spring-mass system relative to the user.

16.3 Sonification Models

The MBS framework is very open, i.e. it enables very different model specifications using very different sources of inspiration. Before providing general guidelines for MBS design in section 16.4, it is helpful to briefly review some existing sonification models. This section gives such an overview, where the model definition steps (setup, dynamics, excitation, etc.) are explained as compact and figuratively as possible. Mathematical details can be found in the referenced articles. However, sound examples are provided and are briefly discussed to bring this section to life.

16.3.1 The Data Sonogram Sonification Model

This model (see Fig. 16.3) has already been used as tutorial example in section 16.2. In summary, the *model setup* is to use one mass-spring system per data vector in a model space of the same dimensionality as the data space, each spring being attached at positions given by the data vector. The user interacts with a scatter plot of the data set and excites shock waves that spherically propagate through the model space. The shock wave speed can be adjusted - typical values for full traversal through the model space are 2 - 5 seconds. The shock wave front, as it passes, displaces mass-spring elements from their equilibrium state and these oscillate with some damping around their position according to the given equations of motion. The resulting sum of all mass-spring displacements constitutes the sonification which is roughly spatialized in stereo around the listener who is imagined to rest at the shock

[3] HRTF = Head-related transfer functions

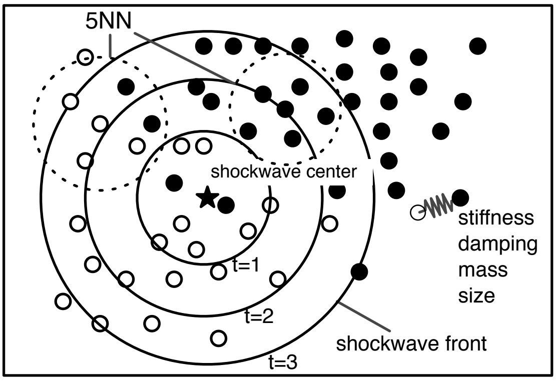

Figure 16.3: Data Sonogram Model Space

wave center. Both mouse clicks and multitouch displays have been used as interfaces to excite the system [14, 27].

Standard Data Sonograms provide information about the data density along a spherical sweep. But task-specific refinements of the model allow specific features such as the class label in data from classification problems to be used to control physical properties of the system, e.g. the stiffness or damping of the individual springs. In general, MBS allows for the definition of individual physical properties at hand of 'local' features. For instance, if the local class mixing entropy[4] among the nearest neighbors of each data point determines the spring stiffness, regions in the data space where different classes overlap will sound higher pitched since the higher local entropy leads to stiffer springs. This may be coined a 'class-border sensitive data sonogram' and it may be useful to quickly assess whether data from classification problems are separable or not. Data sonograms generally support an understanding of the clustering structure of data.

Sound examples **S16.2** are typical data sonograms for clustered data sets. More details on these examples can be found in [10].

16.3.2 Tangible Data Scanning

In Tangible Data Scanning (TDS), data points are represented by localized mass-spring systems just as in the Data Sonogram model as shown in Fig. 16.4). However, now the data are embedded into the 3D-space around the user. Thereby the model is mainly useful for 3D data, or for 3D projections of data. In contrast to data sonograms, interaction is very direct: the user moves a planar object such as a cardboard sheet as an interaction tool which is tracked by a motion capture system. Whenever the surface intersects a mass-spring system in the model space, the latter is excited and oscillates around its position. Even if the sound is played as monophonic audio, the directness allows the user to build up a mental model about the spatial data distribution. It suddenly makes sense to refer to the cluster 'down left around my left knee', or 'in *that* corner of the room'. Similar to Data Sonograms, modified /

[4] which is high when neighboring data points belong to different classes

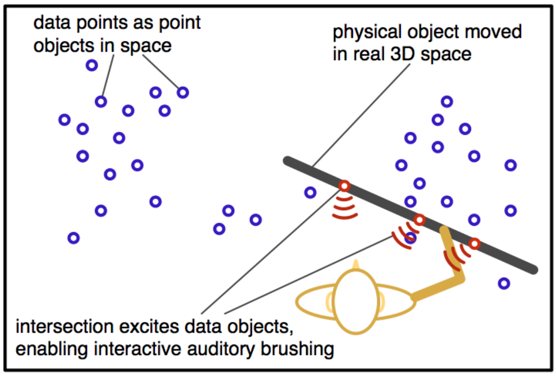

Figure 16.4: Tangible Data Scanning

derived models can use more elaborated definitions of how physical properties depend on local features. Interaction video **S16.3** illustrates a scanning of the space using a clustered data set (Iris data set containing three clusters). More details are reported in [4].

16.3.3 Principal Curve Sonification

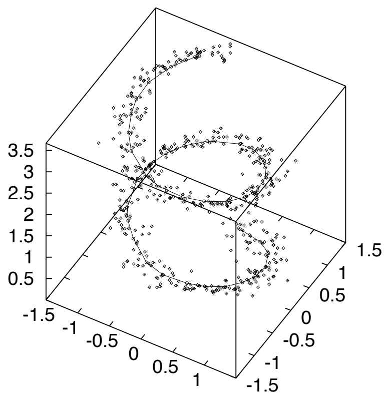

Figure 16.5: PCS for a spiral data set: noise structure along the spiral is difficult to see but easy to hear using PCS

The principal curve (PC) is a machine learning technique to compute a smooth path through a data set which passes nearby all data points [20, 9]. In this sonification model (see Fig. 16.5), each data point in the data space corresponds to a sound source in the model space which may contribute to a continuous overall soundscape, or just be silent. The interaction mode is that the users move along the curve through data space and hear only those data points that project onto their location on the curve. Alternatively, passing along the data points excites the sound sources. As a result, principal curve sonification (PCS) serializes high-dimensional

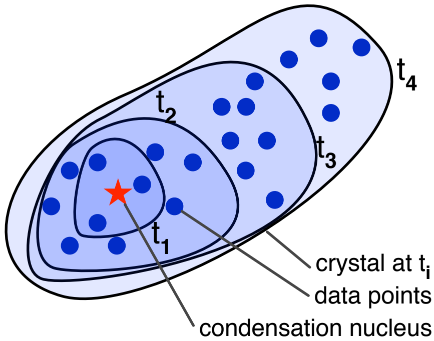

Figure 16.6: DCS crystal growth: crystal hull at various times.

data into a time-organized sequence where movement in space along the curve becomes the main mode of experiencing the data.

This model is very suitable for understanding the clustering structure of data since typically the PC passes once through all clusters. The sound example **S16.4** presents a PCS of a data set where the data are distributed along a noisy spiral: density modulations along the spiral become more easily heard than they can be perceived visually, see [13] for details.

16.3.4 Data Crystallization Sonification

The Data Crystallization Sonification (DCS) is inspired by the chemical process of crystal growth, here applied to the agglomerative inclusion of data points into a growing 'data crystal'. The model is a spatial one: data points specify the locations of 'molecules' in the model space as depicted in Fig. 16.6. These molecules are fixed and never move during the whole procedure. Excitation is done by setting a condensation nucleus, e.g., by clicking the mouse somewhere in the scatter plot. Molecules are then included with increasing distance from this center into a growing 'data crystal'. The metaphor is that the inclusion of a molecule sets free some energy which contributes to the overall vibration energy of the growing data crystal. The crystal's modes of oscillation are not defined in analogy to physics, but instead use the covariance matrix of the data set at each growth step as follows: the eigenvalues determine the harmonic series while the overall variance determines the size and thereby the fundamental frequency of the sound. During growth thereby the pitch drops whereas the brightness signature modulates. Understanding the mathematics helps to better understand the implications of sound changes and to interpret the sound as a fingerprint of the data crystal. Nonetheless, patterns can be discerned, characterized and compared even without this specific knowledge. The technique is suitable for discovering the clustering structure of data and particularly the local dimensionality structure of clusters in data sets. Sound examples **S16.5** illustrate typical sonifications, and more detailed explanations are given in [18].

16.3.5 Particle Trajectory Sonification Model

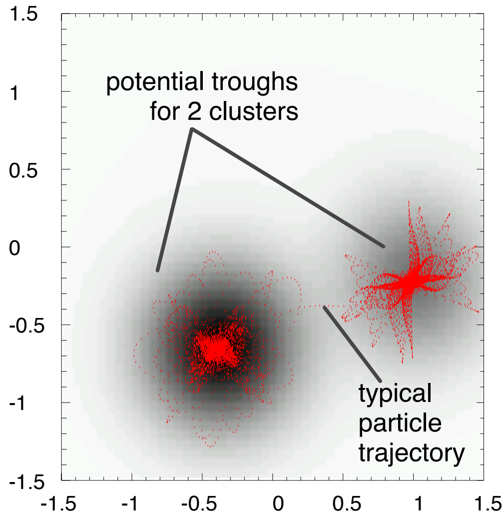

Figure 16.7: PTSM: 2D-potential and a particle for large σ, smooth $V(\vec{x})$

The Particle Trajectory Sonification Model (PTSM) demonstrates how MBS can *holistically* encode information into sound in a way which goes beyond what would be attainable with Parameter Mapping Sonification. For that reason we will discuss the model in more detail. In a nutshell, the model space is a d-dimensional vector space, the same as the data space. Data vectors determine coordinates of point masses in model space, which contribute to an overall 'gravitational' potential function $V(\vec{x})$. There are no dynamic elements connected to these fixed masses so the model remains silent so far.

The model of the universe is a useful metaphor for this, and we can imagine data points as stars that are fixed in space. Additional particles are now introduced to probe the model. They move fast in the 'data universe' according to the laws of mechanics. Staying with the metaphor of the universe, these are like comets as shown in Fig. 16.7.

As potential function, instead of a Coulomb potential, here an inverse Gaussian function $\phi_\alpha(\vec{x}) = -\mathcal{N}\exp(-(\vec{x}-\vec{x}^\alpha)^2/(2\sigma^2))$ is used, where σ controls the width of the potential trough, $\vec{x}^\alpha$ is the position of mass α, and $\mathcal{N}$ is a normalization constant. In the overall potential $V(\vec{x}) = \sum_\alpha \phi_\alpha(\vec{x})$ each particle moves according to Newton's law $m_p\vec{a}(t) = -\nabla_x V(\vec{x}(t)) - R\vec{v}(t)$, where $\vec{a}$ is the acceleration, $\vec{v}$ the velocity, R a friction constant and m_p the particle mass. As a result each particle moves on a deterministic trajectory through the data space. Collisions with other particles or masses are excluded. Finally, due to the friction term, each particle comes to rest at a local minimum of $V(\vec{x})$, which cannot be determined in advance. The link-variables are the instantaneous particle energies $W_i(t) = m_p\vec{v}_i(t)^2$, and their sum represents the overall sonification.

So what can be heard? In the beginning, a particle has enough energy to move freely in the data space, attracted by the data masses, moving on rather chaotic trajectories. This translates to rather noisy sounds. With energy loss, the particle is captured within a cluster (in the metaphor: a galaxy) and finally comes to rest at a minimum of V. The oscillations depend on V and change over time with decaying energy, providing an implicit and partial idea of the data distribution. While a single particle gives limited information, a number of particles create a qualitative sonic image of the data universe structure. Sound examples **S16.6** are single particle sounds.

Excitation in this model means either the injection of a bunch of particles into the model space, or the excitation of existing particles by giving them an impact. Depending on the excitation type, different interfaces can be used, ranging from a mouse click in a plot window for triggering particle injection, to shaking the mouse or other controllers such as an audio-haptic ball interface [12] to inject energy.

An important parameter for understanding the data distribution is the potential width σ: at large values the particles move in a very smooth Gaussian potential; decreasing σ lets more detail appear, first clustering structure, and finally potential troughs around each data point. Thereby the overall sound of the particles depends strongly on σ and this parameter can be offered as control to the user for interactive adjustment. For instance, with an audio-haptic interface [12] it is intuitive to use the squeeze force to control $1/\sigma$. Sound examples **S16.7** are sweeps while decreasing σ. The first example is for a data set consisting of three clusters. Stable pitches occur during decay at middle values of σ corresponding to well-shaped potential troughs at clusters. The second data set is only a single Gaussian distribution without further substructure, and in turn this pitch structure is absent in the sonification.

The primary analysis task of the model is to make perceptible the homogeneity and clustering shape of high-dimensional data. The structure can be understood from stable sonic pitch plateaus and noisy patterns during the transitions between these modes. Timbre complexity is obviously very high and there is no explicit definition of a synthesizer or sound generator. Data points are not explicitly responsible for sound structure. In contrast, data points contribute to the overall potential function and thereby contribute in a complex way to a holistic encoding of information into the sound wave field. Obviously the human auditory system can pick up structural properties, and we are likely to adapt further during sustained use of the model since the sound signal possesses the expected complexity and richness we are familiar with from contact sounds and noises in natural environments.

16.3.6 Growing Neural Gas Sonification Model

Growing Neural Gas (GNG) is a method for computing a topology-preserving graph representation of reduced complexity for a given high-dimensional data set [8]. For the GNG sonification (GNGS) model, the setup consists of the GNG graph trained with the data (see Fig. 16.8). The nodes of the graph are called neurons and can be imagined as points in the data space. For the model setup, an energy level variable is associated with each neuron.

The dynamics of the model operate on two levels: first, via an equation which determines how energy flows along graph edges to neighboring neurons; second, via different equations of motion for neurons to generate sound depending on their local properties (i.e. energy, graph connectivity structure). The model is excited by injecting energy into a neuron, e.g., by touching the location in a visual representation. The equations of motions spread the initially concentrated energy throughout the connected sub-patch of the GNG. Each

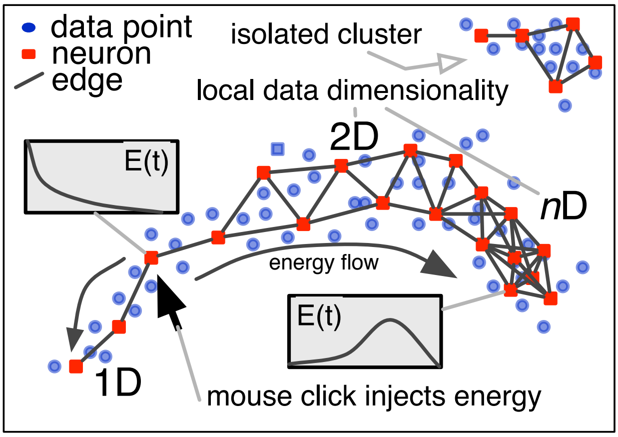

Figure 16.8: GNGS: energy flows through the network.

neuron contributes its sound to the sonification, allowing the perception of graph structure by listening.

What is the sound of a neuron? Assume that edges fix a neuron at its location, so the more edges there are, the higher the net restoring force, and qualitatively, the higher the frequency this neuron oscillates around its position. Following this logic, each neuron generates a sine wave, its energy determining the amplitude, and the number of edges influencing the stiffness and in turn the frequency. As a result, the overall connectivity of the structure becomes audible while energy spreads in the graphs. An important characteristics of GNG graphs is that the edge number at each neuron roughly scales with the local (intrinsic) dimensionality of the data. Thereby the sonification is an *implicit representation* of intrinsic dimensionality, an important feature for modelling and data analysis.

Exciting the GNG at different locations allows the user to perceive, at first, the local properties, then later the average properties of connected GNG patches. As a promising alternative use to excitation, the sonification can be rendered while the GNG grows. This allows the user to perceive the progress of adaptation and even to hear at what point overfitting sets in. Overfitting means that the graph merely describes the randomness of the data instead of the underlying relationship. This shows that MBS is not only useful for active exploration, but is also a suitable technique for process monitoring applications., see chapter 18.

Sonification examples **S16.8** show that clusters of different intrinsic dimension[5] sound differently when energy is injected into one of their neurons: note that higher-dimensional distributions automatically sound more brilliant without this feature having been mapped or computed explicitly during any part of the model construction. Sonification example video **S16.9** shows a sonified GNG growth process. You can hear how the structural hypothesis changes during learning. More details are provided at [16].

[5]degrees of freedom to span the volume, e.g. 1d is a curved line, 2d a twisted plane, 3d a volume, etc.

The reader may also check sonification models omitted here such as Shoogle for shaking text messages [30], the Local Heat exploration model [3], Data bubbles [23], Markov-chain Monte Carlo sonification [11], data solids [12], Multitouch GNGS [28, 21], and for scatter plot exploration for visually impaired people using active tangible objects [24].

16.4 MBS Use and Design Guidelines

How can designers quickly create useful sonification models for a certain task? This section provides guidelines, (a) to decide whether to use MBS at all and if so, (b) how and why to design new models and (c) how best to use MBS.

When to use MBS The motivation is to gain a rapid understanding of what is of interest in the data. If the data are organized in time (e.g., multivariate time series data), it is in most cases straightforward to maintain this dimension and to consider Audification or Parameter-Mapping Sonification. In the latter case it is important to consider if the available features can be meaningfully mapped to acoustic features, allowing the user to experience the temporal evolution in an informative way. MBS is rarely used for time-indexed data.

If, however, there is no time index, nor any other unique continuous feature for temporal organization of the sonification, it may appear unclear how to proceed. To give an example, the *Glass Identification Data Set*[6] contains 10 different physical properties such as refractive index, and chemical analyses such as Na, Mg, Al,... (in weight percents) of different types of glass samples (from buildings, vehicles and containers). The challenge is to identify the glass (or a new unseen glass sample) correctly from its features. In this example there is no time axis! Furthermore, the dimensionality is too high to understand the structure from looking at scatter plots. In this case, mapping all the features to acoustic features would be difficult. 10 meaningful acoustic features would be needed, which is quite challenging. The next problem is that there are infinite possibilities for the mapping, so the question arises of how to map what feature to what acoustic parameter? Any mapping will give an arbitrary sonic image, and it is highly likely that only the features mapped to event onset and pitch will mainly attract the listener's attention.

In such situations, MBS can be very useful. Think of MBS as a kind of tool box, each tool designed for a specific analysis goal. In the same way as you would not use pliers or a screwdriver to hammer a nail into the wall, each sonification model has been (or should be) developed to support a specific analysis task. The task is often so general that it abstracts largely from the concrete data. If, for instance, the task is to detect linear dependencies, a model would be applicable (and ignorant) to whether the data are chemical compound ratios or stock prices or census data features. The sonification model gives structure-specific information, which is good since it may inspire analysts to find new ideas for modelling the data or visualizing them in a way not thought about before. Similar to a motor mechanic who naturally listens to the engine sound before checking part by part for malfunctions, MBS may help data analysts to understand more quickly what's going on and in what direction to proceed with analysis. For example, if you discover linear dependencies you would certainly apply principal component analysis. If you discover clustering, you would proceed with

[6] see http://archive.ics.uci.edu/ml/datasets/Glass+Identification

clustering algorithms etc.

The currently existing sonification models are not strongly optimized for such specific tasks. Moreover, they typically allow the user to perceive additional aspects beyond the main objective. In our glass data example, the data sonogram sonification model using the class-entropy-based stiffness control explained on page 409 provides sounds that allow us to understand how strong the different classes of the glass probes overlap, or whether these classes can be nicely separated. After some interactions, particularly when paired with an interactive data selection to filter out some glass types, you may get a good idea what glass types are easily separable .

The fact that MBS is independent of the concrete data semantics is also a big advantage for another reason: the sound patterns remain stable over many uses with many different data sets. Therefore users can build up knowledge and experience in how specific structures sound.

How to Design Sonification Models If you want to create a new sonification model, the first question should be what is the main analysis task, or what type of pattern or structure should become apparent from listening. Taking a task-centered view helps the designer to focus on the relevant features. For example, assume that the goal was to hear whether the data set contains outliers.

Outliers are data points which are far away from the rest of the distribution, often due to erroneous data. They are sometimes difficult to recognize in multivariate data. Think of a census data set where females provide the information "$x =$age" and "$y =$ number of children". $x = 12$ is not an outlier, nor is $y = 3$. Yet the tuple $(x = 3, y = 12)$ is certainly impossible and must be an outlier. So in order to detect outliers it is not enough to look at single features.

Here is one way to invent a sonification model for outlier detection. We could start by the following observation: outliers typically have few nearest neighbors in data space. So, if we create a dynamic system whose properties depend on this neighborhood emptiness we would obtain sounds where outliers stand out. For instance, we could represent each data point by a mass-spring system and define that the distance to cover the 5 nearest neighbors in data space determines the spring stiffness. After excitation of the masses, the 'outlier candidates' would sound at very high pitch and perceptually stand out. However, data points in a sparse region might also cause similarly high-pitched sounds. Thus, pitch is not necessarily an indication for outliers. Also, if the data space is rescaled, all stiffnesses increase and all oscillations sound higher pitched. One solution is to take the *relative* size of the 5-nearest neighbor sphere, so as to divide the radius by the standard deviation of the complete data set, etc. This should give an idea how the model could evolve further at the next design steps.

However, we could alternatively start from a completely different angle. Assume we connect d guitar strings from each data point in the d-dimensional data set to the points $\vec{x}_i$ that are the center of the k nearest neighbors if we would leave the ith vector component out. We could then send wind through the model space, or hit the whole model and as a result those data points which have long strings will contribute very low-frequency percussive sounds.

It is difficult to imagine what this model would actually sound like, yet certainly it would be possible to iteratively optimize a model to be both satisfying to use and informative. Perhaps, after a series of model inventions and refinements we would arrive at a quite suitable model to perceive outliers. The important point is that the models provide *analog* information and leave the inference and interpretation to the user. This is in contrast to procedures where a detector algorithm simply finds outliers and signals the result since then the user is detached from the analytical and decision-making process. The simpler and easier to understand the dynamic system, the better it will be for users to learn interactively how sound relates to patterns. The hope is that useful sonification models will - by being used - at some point in time become a *standard* tool for a given task, and are then effortlessly understood and routinely used to accelerate data analysis.

However, the designer may lack a concrete idea of what structure the sonification model will work best with, and may start from a random design seed. This probably bears a higher risk of creating useless models, but may eventually offer a higher chance of discovering something really unexpected and new. In the end, it is the utility of the sound to better understand the data which decides if sonification models 'survive' and will be used.

16.4.1 Metaphors for Sonification Model Design

Metaphors are very helpful both for the design process and the user. Some examples are the "shaking objects in a box" metaphor as used with the audio-haptic ball sonification model in [12] or in shoogle [30], or the "moving particles in a data universe" metaphor or the growing data crystal metaphor presented earlier.

To give an example let us start with a metaphor of 'dropping water' for the model design. Going back to the outlier detection sonification model considered in the previous section, we could imagine data points to be little pinholes through which water drips every second. Certainly we need to invent a law to describe in what direction the drops fall (e.g., they could fall towards the plane spanned by the first two principal components of the data distribution) and what sound they make when they touch this plane (i.e. what is the sound rendering process for this virtual water? - will it sound like real water drops?). The metaphor of dripping opens up ideas for new models. It might even inspire new interaction ideas, e.g., squeezing a tube interface to press more drops through the pinholes. If the metaphor works well, we may even consider ideas about how we can shape the dynamics so that the sonification is more similar in perceptual qualities to what we, as the designer, would have expected.

In summary, metaphors are useful both for the design and interpretation of MBS. However, the underlying coherence in a sonification model is usually stronger than just a metaphor (which works in some aspects but fails in others). The model is not a metaphor but has its own logic and consistency – the metaphor, however can be helpful for speeding up design and learning.

16.4.2 Task-oriented templates

Model-Based Sonifications abstract from the application-specific details of the data and are ignorant to the semantics. In other words it does not matter whether data come from chemistry, biology, economy, etc., when used in MBS the focus is on the data's structural properties. This makes MBS a bit more complex to understand and use, but it increases reusability.

The model developer's goal is to have a powerful toolbox of sonification models for whatever structure could potentially be of interest and to quickly explore a new data set with these 'interaction tools' to rapidly understand what's going on. This does *not* replace further investigation, but informs analysts so that their choices regarding their next steps are better rooted in experience. In the same way as there are several types of screwdrivers for similar screws, there may be several sonification models for similar tasks. It will also be a matter of personal preference, taste, or familiarity as to which sonification model works well for whom.

16.4.3 Model optimization

Sonification models are dynamic systems, and these typically contain a number of control variables that determine the detailed dynamics. These parameters need to be adjusted and tuned. However, this tuning is normally only done once by the model designer, so that the model can be applied *without any changes* by the user to arbitrary data sets. Sometimes a few parameters are provided to the users as interactive controls. The data sonogram sonification model for instance allows the user to control the propagation speed of the shock wave. This is useful for moving between very quick scans for rapid comparison of regions and slow spatial scans to attend to spatial density patterns.

Typically the number of parameters is low, compared to the many parameters to be adjusted when working with parameter mapping sonifications of d-dimensional data onto a p-dimensional synthesizer. This reduction of complexity on the side of the parameters goes hand in hand with the additional benefit that the model parameters are meaningful since the users can relate these to their internal imagination of what is going on.

16.5 Interaction in Model-Based Sonification

Interaction is an important part in MBS because MBS is interactive 'by-design' through the necessary excitation of the model. The general motivation for the importance of interaction is given in chapter 11 where some interaction modes are also explained.

The main purpose of interaction is to put energy into the dynamic system. As a result the system develops in time which causes the sound. A strong advantage of this approach is that interaction binds different modalities together. For example, if we excite a sonification model by knocking on some visualized data points using a multitouch display, we obtain a coupled audio-visual-haptic response and media synchronization helps us to relate the different media to each other and to bind them into multimodal units. Importantly, media synchronization does not need to be programmed explicitly, it emerges naturally from the coherence of the model.

Interaction furthermore enables the users to bring in their highly developed manual interaction skills which they have built up since birth: interaction in the real-world is far more complex than our typical interaction with computer interfaces such as mouse and keyboard. Think for example of the richness of interaction while shaking a box to find out what is inside, or while sculpting with clay. Model-Based Sonification aims to connect to such complex interaction abilities.

All sorts of interactions which we perform with real-world objects are candidates for MBS. Examples are scratching, rubbing, hitting, plucking, squeezing, deforming, stretching, bending, touching, etc. Interactions can be organized into the continuum between contact interactions and continuous interactions.

Contact Interactions are interactions where there is a very short energy transfer to the system. If we tap on a melon to hear whether it is matured, or if we knock on a wall to hear whether it is hollow or solid, we use contact interactions. For sonification models the implementation of these interactions can be as simple as using a mouse click in a scatter plot or as complex as using a multitouch surface equipped with contact microphones to sense details of the contact interaction. In objects such as mobile phones, acceleration sensors allow the measurement of contact interactions.

Continuous Interactions are those where the interaction progresses and changes while sound is being generated. Stroking, rubbing or scratching a surface are examples. Practically, they can be sensed by spatially resolved sensors such as touch-sensitive screens or tactile mats of sufficiently high resolution [1]. However, continuous interactions may also be non-excitatory, which means that they only manipulate the system (e.g., rotating an object or squeezing it) without putting energy into it. For example, imagine how a drum head interaction sound changes while the user's other hand moves or changes the pressure at a different position. In this way continuous interactions may control MBS parameters.

16.6 Applications

Model-Based Sonification was introduced as a framework to turn immaterial, non-sounding data sets into something that is sound-capable, so the primary applications are in the area of exploratory data analysis. However, MBS may also be useful in other fields as will be outlined briefly in the following sections.

Exploratory Data Analysis The best data mining 'machine' for the task of discovering and identifying hidden patterns and structures in complex data is the human brain. Our sensory organs and neural networks in the brain are excellent at making sense of the signals we encounter in the world, and allow us to recognize trees, cars, buildings, objects from the signals that come in via our eyes, ears and other sensory channels. However, as highly adapted as the brain is to make sense of structures as they appear in the world, it is bad at finding patterns in huge tables of numbers, which is the most direct representation of data. For this reason, there is the need to bridge the gap between the data spaces (mathematical vector spaces filled with data points) and our brain's preferred perceptual spaces. Model-Based Sonifications offer interaction-based mediators that turn data spaces into model spaces that are capable of creating sound.

The main capability that our brain offers here is *automatic concept formation*: the brain processes the sensory stimuli, automatically discovers patterns and instantiates categories to organize the perceived signals. In machine learning this is called 'symbol grounding', the transition from sub-symbolic signals to symbols. Here is a good opportunity to connect this to Kramer's continuum from analogic to symbolic displays as a means of categorizing auditory displays (see p. 23 in this volume): In exploratory data analysis we do not want to extract symbols (recognized patterns) from the data and represent them by auditory symbols, we rather want to turn the data into complex *analogic* representations which are suitable for the brain to discover patterns and symbols.

The key requirement to enable this learning process is the *invariance* of the binding between the data and the sound. Examples for that have been given in the sonification models discussed in the previous sections. In principle, all sorts of structures can be subject to sonification model design, such as outlier detection, local intrinsic dimensionality, clustering structure, separability of classes, multi-scale structure, and rhythmical patterns (e.g. where data points are aligned on a grid). Furthermore sonification models can also support meta-tasks such as determining how robust a mathematical model is in explaining the data (generalization), or when and how during the training of a machine learning model overfitting sets in.

For cluster analysis, the GNG sonification model, the particle trajectory sonification model, the data sonogram model and the tangible data scanning offer basic tools. For understanding the topology and intrinsic data dimensionality, the GNG sonification model and the data crystallization sonification model can be used. For understanding multi-scale structures, the growth process sonification of the GNG sonification models, and the particle trajectory sonification model (while controlling the bandwidth parameter σ) can be used. For understanding the separability of classes in classification problems, the data sonograms with class-entropy-based spring stiffness may be used. These models are just starting points and hopefully in the future more powerful and optimized sonification models will be developed for specific data exploration tasks.

Augmenting Human Computer Interaction Model-Based Sonification could in future make positive contributions to HCI, for instance, to create more informative, acoustically complex and situation-specific interaction sounds in Computer Desktop interaction. MBS could be used as a principal mechanism to couple any user interaction to acoustic responses, e.g., on the desktop computer or in virtual reality (VR) systems. For instance, a mouse click action could excite the GUI element clicked (buttons, widgets, background, icons, or link) and the resulting sound could help us to be more aware of where we clicked, and what the state of that element is. For instance, a frequently activated link could sound less fresh. There would be a rich, action-dependent informative soundscape while interacting with the computer, similar to the complex and analogous dependencies of real-world interaction sounds,. Furthermore MBS could enhance continuous interaction such as dragging the mouse while holding an object, using a slider, shaking icons with the mouse, or probing objects by knocking on them with a mouse click. Particularly in Virtual and Augmented Reality (VR/AR) where there is often no haptic or tactile feedback when interacting with objects, Model-Based Sonification can create some of the tactile information by sound while adding relevant data-driven information.

Process Monitoring In Model-Based Sonification, the excitation is normally done by the user. If we modify this basic idea so that changes in the data do not only change the model setup, but also provide some excitation, we obtain a sonification model which generates sound without user interaction, and which may be quite useful for process monitoring. Basic ideas for using sonification models for process monitoring have already been given with the

GNGS (section 16.3.6, p. 413) where the adaptation process of a growing neural gas has been used both to excite the sonification model and to configure it.

Auditory Augmentation and Ambient Information Model-Based Sonification also bears the potential for mixed-reality applications that support human activity and provide an ambient information display. Imagine for instance that each time you press a key while typing, in addition to the physical key sound, you also hear an additional sound resulting from the excitation of a sonification model. For example you could hear by a subtle overlapped cue how much space is left in a twitter message or SMS. Sonification models are just the right approach for such action-coupled information displays and would naturally extend the information value of interaction sounds. In [5] we have outlined techniques for augmented acoustics using contact microphones as detectors. Taking such signals as the excitation of a sonification model is the next step.

16.7 Discussion

Model-Based Sonification has been introduced as a mediator between data and sound. Dynamic models bridge the gap between non-sounding numbers and acoustic responses in a different manner to other sonification techniques such as parameter mapping sonification or audification. This section points out the most relevant differences, benefits and drawbacks of this technique compared to other approaches. Much more research in the form of comparative studies is needed to substantiate the claims, which here emerge mainly from long experience and qualitative observations.

Generality of Sonification Models From the brief overview of sonification models in section 16.3 it should have become clear that models are abstract: they are ignorant to the semantics or meaning of the data features, but only demand a certain generic structure. For instance most sonification models can be used independent of the data source, the data dimensionality or the number of data points in the data set and only demand that the data can be represented as a point cloud in a vector space.

Suitability for data that have no time argument Most sonification models have been defined for data sets where there is no time argument in the data, simply because in this case it is most difficult to specify in a canonic way what should be mapped to sonification time. The models also allow us to treat different dimensions equally, without any particular emphasis of one dimension as would happen in parameter mapping sonification due to the different saliency of acoustic parameters.

Dimensionality and Cardinality Independence Model-Based Sonifications can be defined and designed so that they operate on data of any size and dimension. Dimensionality independence is a particularly nice feature since it allows for reusing a model without modification in other contexts. This is in contrast to Parameter Mapping Sonification which requires that for each data set there must be selected a new set of mapping variables onto acoustic features, and a fresh decision about what to do with the remaining unmapped variables.

Learning MBS offers three benefits compared to mapping sonifications concerning learnability and interpretation. Firstly, Model-Based Sonifications address our everyday listening skills which we naturally use to understand everyday interaction sounds when we identify objects and their characteristics. In contrast, the interpretation of mapping sonifications requires more explicit knowledge of the mapping and musical structures to infer meaning from sound. Secondly, MBS sounds are rather stable in structure when using the sonification model with different data sets. This simply gives the user more opportunities to 'tune in' and to learn the 'language of the sound'. In contrast, for mapping sonification, usually you have a new independent mapping and sound structure for different data domains. Thirdly, MBS is interactive by design, naturally allowing the user to connect changes in interaction with changes in sound. Also, users can adapt their exploratory actions immediately as their understanding of the data changes.

Auditory Gestalt Formation Model-Based Sonification aims to provide an analogous auditory data representation according to the continuum definition of Kramer [22]. This analogous representation is particularly useful for auditory gestalt formation since it uses the same mechanisms which encode information into a sound wave as in real-world sound generation. Our listening system is evolutionarily prepared for detecting and conceptualizing gestalts from these kinds of signals.

Ergonomics From the author's experience, the following reasons seem to show that MBS may positively support human well-being and overall system performance. Firstly, since sonification models create sound only after excitation, the sound will be less annoying than sonifications which fill the soundscape decoupled from the user's initiative: they are integrated into a closed-loop (see chapter 11). In addition, interaction sounds accompany the user's actions, so MBS matches their expectations. Secondly, MBS enriches otherwise artificially soundless environments so that the information load is distributed on several perceptual channels. This may reduce fatigue and furthermore engage users into the work process. Thirdly, MBS may increase awareness of the data and actions, thereby helping to avoid misinterpretations or errors. Finally, MBS offers rich and more complex interaction modes such as shaking, scratching, squeezing, hitting a sonification model, for instance by using special interfaces and controllers beyond the mouse and keyboard. This turns data exploration into a much more comprehensive human activity and may also positively impact the healthiness of the work place.

Complexity of sound Sonification models which evolve according to dynamic laws are likely to render sounds which are otherwise intentionally difficult to synthesize. Depending on the model, they may possess a complexity and richness which exceeds the capacity of parameter-mapping sonification sounds. Since the concrete sound depends on the details of the interaction, every sonification will sound slightly different – similar to the way it is impossible to reproduce the signal-identical sound by plucking a real guitar string. However, our ears appreciate this variability and it does not hinder the auditory system to discover the relevant structures behind the 'signal surface' of the sound.

Reusability MBS sonification models are tools, designed to deliver interaction-driven task-specific information. They can be (and often are) defined to operate on a larger class of problems such as 'all data sets which can be represented as point cloud in an Euclidean vector space', or 'all data sets that represent variable distributions on a 2D surface' etc. This makes the sonification highly reusable without the need to adjust any parameters. MBS is a 'design once – use many times' paradigm. Only the developer needs to work hard; it should be simple for the users.

Intuitive Parameters MBS sonification models usually introduce some parameters within the model implementation. Examples are the shock wave velocity of propagation in data sonograms, the energy decay rate in GNG sonification model, etc. These parameters are either specified by the designer, or provided as interactive controls to the users. In the latter case, these parameters will be intuitive controls for users who understand the model. Generally, MBS provides fewer parameters than parameter mapping sonification where both the mapping of data to sound and the parameter ranges are variable. In addition, MBS parameters are often more meaningful since they refer to a physical process that can be imagined.

The Problems of Computational Complexity Sonification models can be extremely demanding in terms of computation. This is especially true for models where the degrees of freedom (e.g. number of moving particles) influence each other so that the number of operations scales quadratically or worse with the number of data points. Since MBS constructs virtual sounding objects from the data, their sound synthesis is as complex as the numeric physical modelling of acoustic instruments, and full-quality rendering of this may exceed the available computer power for many years. There are two alternative ways to address this problem: (i) model simplification, i.e. to invent implementation shortcuts that yield coarsely the expected signals without requiring full numeric simulation, and (ii) model analysis, i.e. using modal analysis from physics or other tricks that enable the efficient computation of the full resolution sound.

16.7.1 Model-Based Sonification vs. Parameter Mapping Sonification

The discussion has pointed out that MBS is quite different from parameter-mapping sonification (PMS). MBS creates dynamic models that are capable of rendering sound themselves whereas PMS maps data values to sound attributes and actively synthesizes the sounds. MBS is interaction-driven whereas interactivity needs to be added artificially in PMS. MBS needs only a few parameters whereas PMS typically needs a more complex mapping specification. MBS addresses everyday listening whereas PMS addresses musical listening. MBS is a 'design-once-use-many' paradigm whereas parameter mapping sonifications need to be set up for each individual data set.

Can we interpret MBS as parameter mapping sonification? On first sight it may appear so in some models. For instance, is the data sonogram model not just a mapping sonification where distance from the shock wave center is mapped to onset? In fact this could be one of the implementation shortcuts to practically implement the model for real-time operation. However, even if mapping is used in MBS for practical reasons such as a more efficient implementation, the model dictates exactly how to map. This may be called model-induced parameter mapping. MBS is also different in character: it can lead to 'holistic' representations, as for instance shown in the particle trajectory sonification model, which parameter mapping cannot create.

Can we understand MBS as audification? On first sight this may appear so as well: For instance, the particle trajectory sonification model is – concerning the rendering – an audification of state variables, specifically the particles' kinetic energies. Yet MBS is not an audification of the data under examination.

Finally, there are two other sources of confusion. Firstly, physical models have become popular for rendering sound signals. If such a physical model is used within a parameter-mapping sonification, this is not a MBS. On the other hand, MBS does not necessarily imply the use of physical modeling synthesis. Secondly, Kramer's *virtual engine* approach, where data are mapped to controllers of a dynamic system [22] is different from MBS despite the fact that a dynamic model is used: again, still the concept of mapping connects data and (in this case a more complex) synthesizer. In MBS, however, the data is not 'playing' the instrument, but the data set itself 'becomes' the instrument and the playing is left to the users. The sonification techniques may appear to lack clear borders, depending on how they are looked at, yet the approaches have their own place. In conclusion MBS is a new category qualitatively different from parameter mapping sonification and audification.

16.7.2 Model-Based Sonification and Physical Modeling

Physical modeling has become a major trend in modern sound synthesis for achieving complex, natural and interesting sounds. The structural vicinity to MBS motivates the question as to how methods from this field can be used for MBS. Few selected examples provide pointers to the relation.

There is a body of research on Sounding Objects [25], which provides assistance for the creation of physics-based models and for controlling their parameters in order to achieve continuous controlled events or interactive systems using these models. These methods are powerful for the generation of parameterized auditory icons (see Ch. 13), yet they can also be used for MBS. A sonification model would be the result if the data set under analysis would determine aspects of the model configuration.

There are also systems developed for music control and synthesis that offer inspiration and useful methods for MBS: for instance Cordis-Anima [26] is a sound synthesis engine, mainly used for music creation, but also capable of visual animation or multimodal simulations. It numerically integrates dynamic processes, e.g., using mechanical interactions, and furthermore it provides the means to excite the physical system via force-feedback gestural controllers. If the mechanical system was determined and set up from the data under analysis (Model Setup) Cordis-Anima would render Model-Based Sonifications.

Scanned Synthesis [2, 29] is a sound synthesis technique which also uses a dynamic system and its temporal evolution to shape sound. Different from simulated acoustics, here the model (e.g., a simulated spring) is scanned cyclically at audio rate to create the audio signal, allowing excitation and interaction to shape dynamic timbre evolutions at a lower control rate. Scanned Synthesis offers an interesting approach to mediate between the model's configuration and the resulting sound, giving inspiration for future sonification models to come.

16.8 Conclusion

This chapter has introduced Model-Based Sonification as a sonification technique that mediates between data spaces and the sound space by means of dynamic interactive models. Starting from an analysis of listening modes, we discovered the potential of human listening to make sense of sound wave fields that represent dynamic processes. This led to the definition of MBS as a paradigm, and *sonification models* as concrete task-centered designs, which need a specification of setup, dynamics, excitation, initial conditions, link-variables and listener characteristics. Various sonification models have been explained and demonstrated. From this background, guidelines for the use and design of MBS sonification models have been formulated. After highlighting interaction and the main application fields, the benefits and problems have been analyzed.

MBS research is still in its infancy. The next step will be to create a toolbox of optimized sonification models for many different tasks, and a good tutorial on how to apply, use, and learn them. For this it will be helpful to have an atlas of reference sonifications for certain structures so that the users can faster assess the structure in the data. Currently existing sonification models are just the first examples and possibly far from optimal. We hope for an evolution where many models will be invented, used, refined or rejected; working towards a set of good standard sonification models tuned to certain tasks. These models will perhaps become as stable and widely understood as pie charts or scatter plots are in visualization. This process will go hand in hand with the evolution of interfaces that allow us to use our skilled manual interactions to manipulate information spaces.

A research agenda for MBS includes, besides the development of the abovementioned MBS toolbox: research into ways of implementing the models so that they can be used for larger data sets with limited computation power; research into how best to interweave MBS with standard visual interfaces and the workflow of data analysts; and finally how to evaluate MBS and how to assess its effects on performance, flow, fatigue, depth of understanding, acceptance, etc. In summary, Model-Based Sonification opens up new opportunities for interactive HCI and multimodal data exploration, and will over time find its way into standard user interfaces.

Bibliography

[1] J. Anlauff, T. Hermann, T. Großhauser, and J. Cooperstock. Modular tactiles for sonic interactions with smart environments. In Altinoy, Jekosch, Brewster, editors, *Proceedings of the 4. International Workshop on Haptic and Auditory Design (HAID 09)*, volume 2 of *Lecture Notes on Computer Science*, pp. 26–27, Dresden, Germany, 09 2009. Springer.

[2] R. Boulanger, P. Smaragdis, and J. ffitch. Scanned synthesis: An introduction and demonstration of a new synthesis and signal processing technique. In *Proc. ICMC 2000*, Berlin, 2000. ICMC.

[3] T. Bovermann, T. Hermann, and H. Ritter. The local heat exploration model for interactive sonification. In E. Brazil (ed.), *Proceedings of the International Conference on Auditory Display (ICAD 2005)*, pp. 85–91, Limerick, Ireland, 07 2005. ICAD, International Community for Auditory Display.

[4] T. Bovermann, T. Hermann, and H. Ritter. Tangible data scanning sonification model. In T. Stockman (ed.), *Proceedings of the International Conference on Auditory Display (ICAD 2006)*, pp. 77–82, London, UK, 06 2006. International Community for Auditory Display (ICAD), Department of Computer Science, Queen Mary, University of London.

[5] T. Bovermann, R. Tünnermann, and T. Hermann. Auditory augmentation. *International Journal on Ambient Computing and Intelligence (IJACI)*, 2(2):27–41, 2010.

[6] A. Bregman. *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambrigde Massachusetts, 1990.

[7] M. Chion. *Audio-Vision: Sound on Screen*. Columbia University press, New York, NY, 1994.

[8] B. Fritzke. A growing neural gas network learns topologies. In G. Tesauro, D. Touretzky, and T. Leen, editors, *Advances in Neural Information Processing Systems*, volume 7, pp. 625–632. The MIT Press, 1995.

[9] T. Hastie and W. Stuetzle. Principal curves. *Journal of the American Statistical Association*, 84:502–516, 1989.

[10] T. Hermann. *Sonification for Exploratory Data Analysis*. PhD thesis, Bielefeld University, Bielefeld, Germany, 02/2002, http://sonification.de/publications/Hermann2002-SFE

[11] T. Hermann, M. H. Hansen, and H. Ritter. Sonification of Markov-chain Monte Carlo simulations. In J. Hiipakka, N. Zacharov, and T. Takala, (eds.), *Proceedings of 7th International Conference on Auditory Display*, pp. 208–216, Helsinki University of Technology, 2001.

[12] T. Hermann, J. Krause, and H. Ritter. Real-time control of sonification models with an audio-haptic interface. In R. Nakatsu and H. Kawahara, editors, *Proceedings of the International Conference on Auditory Display*, pp. 82–86, Kyoto, Japan, 2002. International Community for Auditory Display (ICAD), ICAD.

[13] T. Hermann, P. Meinicke, and H. Ritter. Principal curve sonification. In P. R. Cook, editor, *Proceedings of the International Conference on Auditory Display*, pp. 81–86. ICAD, International Community for Auditory Display (ICAD), 2000.

[14] T. Hermann and H. Ritter. Listen to your data: Model-based sonification for data analysis. In G. E. Lasker, editor, *Advances in intelligent computing and multimedia systems*, pp. 189–194, Baden-Baden, Germany, 1999. Int. Inst. for Advanced Studies in System research and cybernetics.

[15] T. Hermann and H. Ritter. Crystallization sonification of high-dimensional datasets. In R. Nakatsu and H. Kawahara, editors, *Proceedings of the International Conference on Auditory Display*, pp. 76–81, Kyoto, Japan, 2002. International Community for Auditory Display (ICAD), ICAD.

[16] T. Hermann and H. Ritter. Neural gas sonification – growing adaptive interfaces for interacting with data. In E. Banissi and K. Börner, (eds.), *IV '04: Proceedings of the Information Visualisation, Eighth International Conference on (IV'04)*, pp. 871–878, Washington, DC, USA, 2004. IEEE CNF, IEEE Computer Society.

[17] T. Hermann and H. Ritter. Sound and meaning in auditory data display. *Proceedings of the IEEE (Special Issue on Engineering and Music – Supervisory Control and Auditory Communication)*, 92(4):730–741, 2004.

[18] T. Hermann and H. Ritter. Crystallization sonification of high-dimensional datasets. *ACM Trans. Applied Perception*, 2(4):550–558, 2005.

[19] M.J. Houben, A. Kohlrausch, and D. J. Hermes. The contribution of spectral and temporal information to the auditory perception of the size and speed of rolling balls. *Acta Acustica united with Acustica*, 91(6):1007–1015, 2005.

[20] B. Kégl, A. Krzyzak, T. Linder, and K. Zeger. Learning and design of principal curves. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22(3):281–297, 2000.

[21] L. Kolbe, R. Tünnermann, and T. Hermann. Growing neural gas sonification model for interactive surfaces. In Bresin, (ed.), *Proceedings of the 3rd Interactive Sonification Workshop (ISon 2010)*, Stockholm, 2010. ISon, KTH.

[22] G. Kramer. An introduction to auditory display. In G. Kramer, editor, *Auditory Display*. Addison-Wesley,

1994.

[23] M. Milczynski, T. Hermann, T. Bovermann, and H. Ritter. A malleable device with applications to sonification-based data exploration. In T. Stockman (ed.), *Proceedings of the International Conference on Auditory Display (ICAD 2006)*, pp. 69–76, London, UK, 2006. ICAD, Department of Computer Science, Queen Mary, University of London.

[24] E. Riedenklau, T. Hermann, and H. Ritter. Tangible active objects and interactive sonification as a scatter plot alternative for the visually impaired. In *Proc. of the 16th Int. Conference on Auditory Display (ICAD-2010), June 9-15, 2010, Washington, D.C, USA*, pp. 1–7. 2010.

[25] D. Rocchesso, R. Bresin, and M. Fernström. Sounding objects. *IEEE Multimedia*, 10(2):42–52, 2003.

[26] O. Tache and C. Cadoz. Organizing mass-interaction physical models: The cordis-anima musical instrumentarium. In *Proceedings of the international Computer Music Conference (ICMC 2009)*, pp. 411–414, Montreal, 2009. ICMC.

[27] R. Tünnermann and T. Hermann. Multi-touch interactions for model-based sonification. In M. Aramaki, R. Kronland-Martinet, S. Ystad, and K. Jensen, (eds.), *Proceedings of the 15th International Conference on Auditory Display (ICAD2009)*, Copenhagen, Denmark, 2009.

[28] R. Tünnermann, L. Kolbe, T. Bovermann, and T. Hermann. Surface interactions for interactive sonification. In M. Aramaki, R. Kronland-Martinet, S. Ystad, and K. Jensen, (eds.), *ICAD'09/CMMR'09 post proceedings edition*. Copenhagen, Denmark, 2009.

[29] B. Verplank, M. Mathews, and R. Shaw. Scanned synthesis. In *ICMC 2000 Proceedings*, Berlin, 2000. ICMC.

[30] J. Williamson, R. Murray-Smith, and S. Hughes. Shoogle: excitatory multimodal interaction on mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 121–124, San Jose, California, USA, 2007. ACM Press.