

Surface Interactions for Interactive Sonification

René Tünnermann, Lukas Kolbe, Till Bovermann, and Thomas Hermann

Ambient Intelligence Group
Cognitive Interaction Technology - Center of Excellence (CITEC)
Bielefeld, Germany
{rtuenner, lkolbe, tboverma, thermann}@techfak.uni-bielefeld.de
<http://www.cit-ec.org/>

Abstract. This paper presents novel interaction modes for Model-Based Sonification (MBS) via interactive surfaces. We first discuss possible interactions for MBS on a multi-touch surface. This is followed by a description of the *Data Sonogram Sonification* and the *Growing Neural Gas Sonification Model* and their implementation for the multi-touch interface. Modifications from the original sonification models such as the *limited space scans* are described and discussed with sonification examples. Videos showing interaction examples are provided. Furthermore, the presented system provides a basis for the implementation of known and novel sonification models. We discuss the available interaction modes with multi-touch surfaces and how these interactions can be profitably used to control spatial and non-spatial sonification models.

Key words: Sonification, Model-Based Sonification, Data Mining, Interactive Surfaces

Introduction

Exploratory Data Analysis aims to develop techniques for users to better *grasp* the hidden structure in complex data. If we take this statement literally, we might not only ask how we could implement techniques to *manually* interact and get our hands on data, but also how it sounds – or should sound – if we interact physically with data. Real-world acoustic responses that we experience when touching (hitting, scratching, tapping, etc.) an object or surface are often very useful and reveal a whole range of information about the object’s properties (material, stiffness, surface properties, etc.). We often underestimate the utility of such direct feedback since it is omnipresent and at the same time effortlessly integrated into our multi-modal perceptions.

The arising questions are how can we inherit the benefits of action-perception loops for a better understanding of complex data and how can we structure surface-based interfaces in a way that users obtain an informative acoustic reaction on arbitrary interactions? Model-Based Sonification takes these aspects of interaction particularly into account [11]. *Sonification models* according to

MBS can be excited by the user. For this excitatory process many different interaction interfaces beyond the mouse, such as the audio-haptic ball interface, or the malleable user interface, have been presented [12, 13]. These are primarily input interfaces and the sonification in many implementations has been the only output modality in the interaction loop.

In this paper, we investigate the above research questions by using *interactive surfaces*. We start by presenting the tDesk system, a device developed within the Ambient Intelligence Group that combines the possibilities of Multi-Touch Interactions and Tangible Interactions in a desk-based system for simultaneous multi-user use. Our multi-touch system allows to create tightly coupled audiovisual interaction loops to represent the temporal evolution of sonification models while at the same time allowing real-time complex manual interaction with a sonification model. The system has been developed to serve as a sound basis to fuse and explore the potential of multi-touch interactions together with tangible interactions, while using truly multi-modal output media. In Sec. 1 we provide categories for flat surface-based interaction and then use these to discuss how interactions can be connected to the excitation for Model-Based Sonifications. We demonstrate the system together with two specific sonification models. The *Data Sonogram Sonification Model* allows the user to use multi-point interactions to set centers of excitation waves that spherically pass through data space. The *Growing Neural Gas Sonification Model* allows the user to listen to the growing neural gas during its adaption process and to visually and auditorily explore the state of the network with the help of multi-touch interaction.

In comparison to the mouse-based interaction used in previous implementations of these two, the multi-touch interaction provokes new interaction styles such as rapid A/B-comparison and simultaneous excitations in different regions. Furthermore the real-time visualization supports a better cross-modal binding.

Beyond the demonstration of new interaction modes for holistic data experiences as exemplified with the use of our system for the interaction with Model-Based Sonifications, we see diverse application fields where sonification can be plugged-in to enhance the experience. For instance, in didactic multi-touch applications such as an interactive visualization of electromagnetic fields, sonification may represent the electric field as sound while the user moves electric charges or touches the surface. In the area of interactive games, sonification could enable games between sighted and visually impaired users where each receives the modalities she could use best.

Model-Based Sonification

Model-Based Sonification (MBS) is a framework for the development of sonification techniques [10]. *MBS* starts from the observation that humans are well trained to interpret the complex acoustic signals in the world with respect to sound source characteristics. To give an example in everyday interaction, imagine to fill a thermos flask with water. By the pitch rise, due to the changing resonance of the bottle, we are aware of the flask's fill level. There is a large

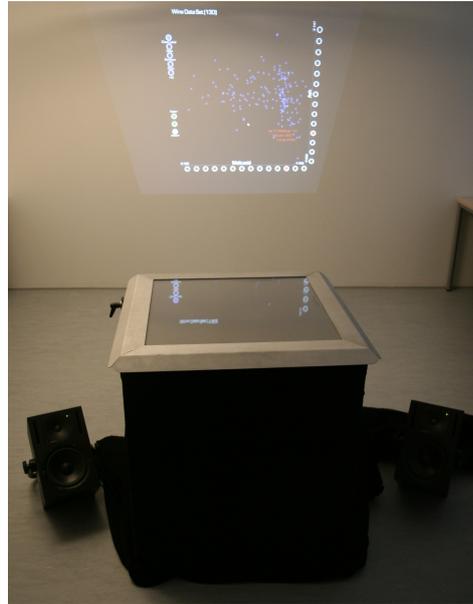


Fig. 1. The multi-touch enhanced tDesk platform with an additional projection on the opposing wall.

variety of situations when we use sound to gain insight into complex systems (e.g. engineers listening to machine sounds or physicians using the stethoscope to support diagnosis) [17, 18].

Most important, sound is connected to the underlying physical system by means of a dynamic (physical) *model*. The model mediates between a system's state and its acoustic response. The human brain is trained to infer source properties from sound that results from physical models. This principle provides the basis for *Model-Based Sonification* which defines in analogy dynamic processes between elements that are typically parameterized by the data. As in physical systems, a *sonification model* is silent without any excitation. Via interaction, the user is able to excite the model which connects MBS to the field of *interactive sonification* [14]. Guidelines and examples for creating interactive sonification models are provided in [9–11].

For the definition of a sonification model according to MBS, six aspects need to be addressed: At first, the *setup* of dynamical elements and the *initial state* of the model have to be determined. This is followed by the *dynamics*, which define how dynamic elements behave in time. The *excitation* defines how the user is able to interact with the model. The *Link Variables* used as transducers between the model and the audible domain have to be chosen and designed before the listener can be positioned in the setting. Finally, depending on the model, the listener needs to be positioned with respect to the data.

Excitation Modes for Sonification Models

The above definition for MBS has already shown that excitation plays a crucial role in the use of sonification models. Designers can take inspiration from all real-world interactions where the world responds acoustically, from 'foot steps' to 'hitting of objects'. If we focus on manual interaction we see that in most situations we either hit, pluck, shake, deform/squeeze or scrub objects. Most of these interactions have several degrees of freedom, e.g. the location of the interaction, the strength, the detailed direction relative to the surface. Depending on the details, interaction with real-world objects provides sonic feedback which includes information about the object. Obviously this richness is far beyond what can be obtained by simple mouse or keyboard interfaces. For those reasons, new interaction devices have been developed to better explore our manual interaction abilities [12, 13].

If we consider the interaction with surfaces in general (e.g. consider to search a wall for hidden holes) we often use tapping, scratching (to examine the surface) and (think of drumheads) bimanual interactions where one hand hits while the other changes an aspect of the surface. Similarly, interactions are natural for interacting with surfaces, and with interactive multi-touch systems we now have the chance to define audiovisual surface reactions, so that an almost as natural utilization of manual interaction procedures may occur by users that explore complex data.

Our long range aim is to implement examples for all available surface-based interactions to explore the potential of MBS to connect manual interactions with exploratory excitations to support the understanding of data under analysis. In this paper we start this investigation with *tapping interactions* as excitations for Sonification Models.

Multi-Touch Technology for the Tangible Desk (tDesk)

As a basis for the development, we started off using the tangible desk (tDesk) [1], a tabletop environment for tangible interaction (see Fig. 2). The tDesk is assembled using aluminium strut profiles. It consists of a 70 cm × 70 cm glass surface resting on four aluminum poles in about 80 cm height (see Fig. 2). The chosen extent of this table allows to conveniently work either alone or collaboratively within a group on and with touchable and tangible applications. Any spot on the surface can be reached with ease regardless of the user's deskside position. Since modularity was a major design issue of the tDesk, the current glass surface is easily exchangeable. We designed a drop-in replacement surface, enabling the tDesk to recognize fingers touching the surface. The used setup consists of the tDesk platform, the constructed acrylic surface with attached aluminium frame, lighting modules covering the pane edges, projector, camera, speakers and a computer system for image processing and multi-modal feedback (see Fig. 1). Basically, the constructed surface is a spatially resolved 2D-sensor recognizing multiple touch inputs. The physical sensor pane is made out of acrylic glass. The display is provided using a screen foil and an inside mounted projector.

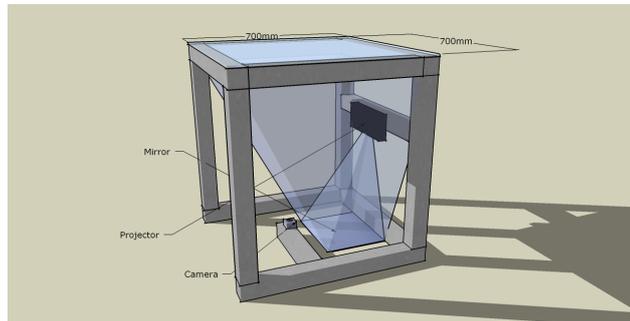


Fig. 2. The tangible desk (tDesk) platform provides the basis for the multi-touch system. It already contains a projector and a firewire camera.

The designed surface allows simultaneous interactions by at least four people in a closed-loop and direct manner. When designing the surface the following aspects were considered:

- Low Latency
Closed-loop interactions require a low latency and therefore a high sampling rate. Reducing latency is of utmost importance, as rising latency can disturb closed-loop human-computer interaction. There has to be immediate feedback while the user interacts with the surface, otherwise the feedback can not be associated with the previous action.
- Input Points
The sensor should support at least 40 input points so that up to four users can use all fingers simultaneously.
- High Resolution
The sensor’s resolution should be close to the display’s resolution to support detection of the characteristics of touch, such as the shape or the orientation of the touched spot on the surface.
- Backprojection
To intensify the degree of immersion, the projection should have its source inside the table. Top projection would lead to disturbing occlusions caused by the user’s hands and fingers.

We used the *FTIR* technique (as proposed by Han [5]) to sense contacts on the surface. By using this technique we were able to sense almost as many input contacts as fit on the surface, achieving a high resolution and sufficient latency. To apply *FTIR* sensing, an acrylic pane is *flooded* with *near-infra-red (NIR)* light. When a finger touches the surface a bright *NIR* blob is caused on the surface. A camera, mounted underneath the surface, is used to capture the resulting blobs. To reduce the latency in the loop, a firewire camera capturing images at a frequency of 60 Hz is used. To improve the image quality, an optical bandpass filter was mounted in front of the camera. Finally, the display is provided by an

underneath mounted projector.

To calibrate the camera and projector coordinate system we have chosen a mapping to resolve:

- camera trapezoid and pincushion distortion
- translation, rotation and scaling
- projector trapezoid and pincushion distortion

Optimal parameter values of the transformation map are determined by minimizing a quadratic error function using a least-squares minimization. Matching point-pairs from both coordinate systems are needed to compute the residuals for each iteration of the optimization process. The Levenberg-Marquardt Algorithm (LMA) [15, 16] was used to train the mapping’s parameters.

Multi-Touch Interactions for MBS Excitation

Since the appearance of publicly available *multi-point* displays, more advanced displays allowing *absolute-spatial*, *multi-point* and *multi-user* interaction by the use of one’s *fingers* are announced frequently, but the characteristics of touch (applied force, angle of approach) are mostly ignored. Whereas orientation and force are commonly exploited in graphics tablets such as Wacom’s devices, these characteristics are sparsely used in other surfaces. An introduction and some framing of that subject by Buxton is available on his website [2]. In general, interactive surfaces can be considered as transducers between the digital and the analog interaction space. The following categories were a useful scope for us to better differentiate and discuss characteristics of surface-based interaction, particularly to excite sonification models. These are similar and partly based on previous work by Buxton [2].

Point vs. Touch: Existing ‘*multi-touch*’ displays often offer *multi-point* instead of *multi-touch* input. The touch of a finger’s tip is only used for a mere pointing, neglecting the details of touch. In addition, touching the surface with the hand or the arm will often lead to undefined behavior.

Single- vs. Multi-Spot: Old fashioned *touchpads*, which are still quite common, support only single-point input. Whereas *single* connotes just one spot input, *multi-spot* refers to devices capable of sensing more than one spot, for example all of the user’s fingers. With *single-* and *multi-spot* as two sides of the continuum, in between there are *n-spot* devices capable of sensing a fixed number *n* of spots.

Collaborative Use: Even though newer notebook computers offer multiple input surface devices, these can hardly be used by more than one person at a time. Even if those pads theoretically could be used by more than one person at a time, in most cases this will lead to odd experiences.

Degrees of Freedom: When using spots on the surface only as pointing input, the surface provides input with two degrees of freedom. The transducer gains

degrees of freedom by adding information about the pressure-vector of touch and direction of approach or other information.

Feedback: Traditional *touchpads* give no direct active feedback at all. *Touchscreens* and multi-spot displays feature visual feedback and thereby create the illusion of being able to manipulate digital items directly. Visual feedback can be enhanced by sound to intensify the degree of immersion. Digital objects with an auditory behavior can create sounds when triggered or when several objects interact with each other. Additionally, vibration motors could be used to create haptic feedback.

Relative vs. Absolute: *Touchpads* are, like mice, relative input devices: If touched, they take the cursors position. The position in the real world is set equal to the position of the cursor in the screen space. Touchscreens on the other hand feature *absolute* input. The user does not have to move the cursor from its current position to the target, but approaches the object *directly* by using a finger or a tool.

Direct vs. Abstract: When an object is moved with a finger or a pen-like tool, the interaction with it can be in a direct manner. If there is a relative transducer in the input chain, for example a mouse, the interaction becomes more abstract. There are a lot of discussions about when interactions are to be seen as abstract or not. We do not doubt that for someone who is familiar with mouse interaction, the relative transducer is ubiquitous and therefore virtually *ready-to-hand* (see dimension *Tools* and [8]). In this work, the term *direct* is used if the input chain of an interaction is free of relative transducers and the application allows the user to touch or move digital items.

Point vs. Gesture: Irrespective of the above-mentioned properties, an application can depend on the actual position, on the trajectory of the input spot, or both. Most common relative *point*-input devices are just using the actual position of the cursor. However, gesture-like input can be used to scale and rotate objects on the surface. Further it can be used to trigger certain actions such as to open a navigation menu.

Discrete vs. Continuous: Discrete interactions can be seen as single actions or events. A continuous interaction can be described by a trajectory of actions or events. Imagine typing on an on-screen-keyboard, or *pressing* displayed buttons, the interaction would be discrete. By moving an object from one position to another, the interaction becomes continuous.

Tools: Surface input devices can be designed to be used with different parts of the body such as a finger or with external tools. A tool can be *ready-to-hand* or *present-at-hand* to the user. A pen for example, when used for sketching or drawing tasks, is *ready-to-hand*. The user does not have to think about how to handle the pen, he just spends time on the drawing task itself [8], [3].

According to the above definition we have implemented direct and *absolute-spatial*, multi-point *Sonification Models*, which allow discrete and continuous use in a collaborative manner. We plan to exploit attributes such as force sensitive input, the use of tools and touch characteristics since the constructed surface already provides these informations.

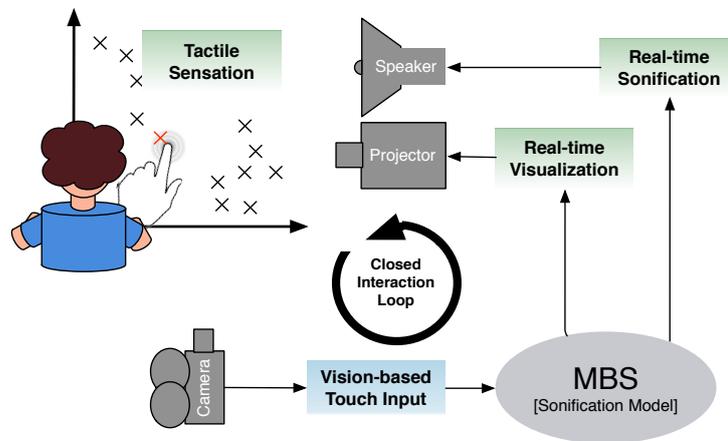


Fig. 3. Multi-Touch Interactions as melting pot for multi-modal coherent binding. The interface causes visual, tactile and auditory perceptions which are bound together via their synchronization to multi-modal perceptual entities.

The Data Sonogram Sonification Model

Overview

The *Data Sonogram Sonification Model* can be described by the following five categories:

Setup: Data points are used as point masses in a model space of the same dimension as the data space. The Data points' coordinates are used as fix point location for a virtual spring connected to the point mass.

Dynamics: (a) oscillation of spring-mass systems modeled by classical mechanics, and (b) the propagation of shock waves in the high-dimensional model space.

Excitation: The user can excite multiple 'shock waves' to emanate at certain locations in the model space.

Link Variables: The kinetic energy of all point masses is used to generate the sound signal, which represents the sonification. Alternatively the elongation from equilibrium can be used as link variable.

Listener: A two-dimensional view of the data and visual controls to navigate the data exploration are provided to the listener on the interactive surface. In the original model, the virtual listener is positioned in model space at the point where the shock wave is initiated. In this implementation, however the listener is centered at the x -axis of the 2D plot in front of the table, since only a stereo panning is used for sound spatialization.

A two-dimensional scatter plot of the data serves as the interaction area for the user to excite data sonograms by selecting a position on the plot. The speed of

the shock wave can be adjusted by the user interactively. The resulting shock wave then passes through the data set within some seconds, so that the user can examine the data step-by-step.

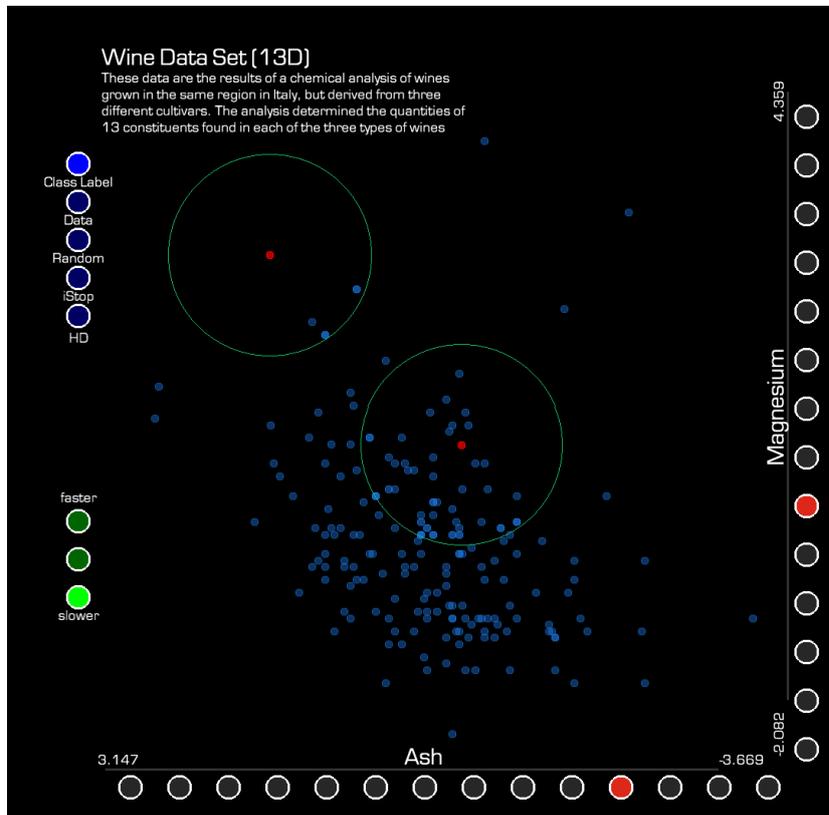


Fig. 4. The graphical user interface of the Data Sonogram application. The 13D *wine data set* is selected. The dimensions *Ash* and *Magnesium* are chosen for the visual feedback. Dimensions are chosen by a tap on a button in the button group for the corresponding axis. In a 13-dimensional data set, 13 buttons are displayed along each axis. Two shock waves are active, propagating through the two-dimensional data space.

Multi-Touch Adaptations for Data Sonograms

We added several features for the multi-touch Data Sonogram implementation. At the point of excitation in a sonogram, a virtual shock wave is initiated. For the sake of usability and in contrast to physical constraints, the speed of the propagating wave can be adapted while the wave is traveling. A *shock wave* has a center and a source point. The distinction has to be made because the

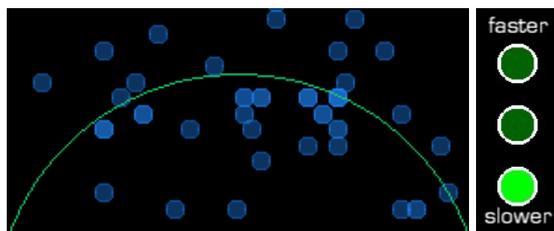


Fig. 5. Left: Point instances in data space with different energy, indicated by the spot's brightness. Energy is induced by the augmented shock wave. Right: An example of a button group used to control the velocity of the traveling wave.

point of excitation is not always the center of the shock wave. In case the user initiates a high-dimensional shock wave, the shock wave center will be located at the coordinates of the data point that is nearest to the excitation point in the two-dimensional display.

In our original implementation [10], no visual feedback was given during the excitation of the data sonogram. Here we have added an interactive visualization of the shock wave front while the sonification is computed in real-time. Thereby the user knows where the sound comes from at the time they are perceived. A visual shock wave front is only meaningful in the case of a two-dimensional shock wave expanding on the table. If a high-dimensional shock wave is propagating data points might be reached by the shock wave in a less predictable pattern. In that case we update the data points' visual blobs by increasing their intensity at the time the high-dimensional shock wave passes through them. Again, these synchronized visual and auditory events help the user to better interpret the sound with respect to the data.

With the increased directness that multi-touch interactions serve, we discovered that users were interested to 'probe' the data sets frequently at different locations. The continuation of the shock wave after it has been triggered is then irritating and thus not helpful for this interaction pattern. We therefore implemented a new interaction style, where a shock wave immediately stops at the time when the user lifts the finger. In result, users can tap on the scatter plot freely and always get and compare their impression about local neighborhood distribution (from the temporal organization of the first sound events at the excitation point). This behavior turned out to be a useful optional feature for a step by step exploration of the data space.

The interface offers several options for the adjustment of parameters and the navigation of data in a coherent surface (see Fig. 4). Parameters such as the wave velocity, the sonic dimensionality and other features of the sonification model can be adjusted by the user interactively. The user is able to navigate in the data and can choose the to be displayed dimensions.

The chosen data dimensions (see Fig. 4) are presented on the interactive surface. Each dimension can be selected by a single touch on a corresponding

button. Each axis is bound to only one dimension at a time. The two-dimensional scatterplot is used as the visual interface by which users are able to excite the sonification model. The system currently supports data sets with up to 16 dimensions, since the size of the interactive surface is limited. The users are able to trigger shock waves in two modes:

Two-dimensional mode (2D): a 2D shock wave is initiated at the touched coordinates on the surface. The mode has its main purpose as a didactic introduction to the system. The traveling shock wave front is visually augmented by a green circle with an increasing radius on the surface. Data points passed by the wave front are excited.

High-dimensional mode (HD): the user triggers a high-dimensional shock wave at the coordinates of the nearest 2D data point in the visual display. In contrast to the 2D mode, a visually spreading wave is not as useful in the high-dimensional mode. Instead of augmenting the propagating wave, passed data points to which energy is transferred are illuminated. The wave can be observed in the visual domain as a sequence of flashing data points (see Fig. 5).

The user can switch between these two modes through a button labeled ‘HD’, located at the left border where all control buttons are placed. At the lower left, a button group is placed consisting of three buttons to control the velocity (slow, normal, fast) of the propagating wave. The graphical user interface was written in *Processing* [7].

Auditory Components for Data Sonograms

The sonification sound signal is the superposition of all instantaneous elongations of masses from their equilibrium position. Since the spring forces the masses to a damped oscillation their representation becomes audible as a decaying sine tone. For the implementation of the spring-mass systems, unit generators for spring-mass systems in *SuperCollider* [19] have been used. This alleviates the problem of numerically integrating the dynamics of all mass-spring-systems, since they are well decoupled.

A stereo speaker setup is aligned towards the listener (as shown in Fig. 1). The virtual listener is centered in front of the table. When the shock wave front passes a data point, a sound event is spawned via the *OSC* protocol. Since the class label is used as spring stiffness, it can be perceived as pitch. The spatial location of the sound source can be estimated via the stereo panning.

Example Data Sets

There are three tutorial data sets available to the user by default:

wine data set: These data are the results of a chemical analysis of wines. They are derived from three different cultivars in the same region in Italy. The analysis measured the quantities of 13 constituents found in each of the three types of wines [4].

iris data set: This is perhaps the best known data set to be found in pattern recognition literature. It contains three classes of iris flowers, with four features each. 50 instances are included for each of the three classes. One class is linearly separable from the other two. The latter are not linearly separable from each other [6].

random data: The random data set contains uniformly distributed random data in four dimensions. It serves here as a benchmark distribution to train audiovisual exploration, i.e. to better learn to associate auditory and visually perceived elements in the multi-modal system.

Every time the data set is changed by the user's demand, a short description of the data set is displayed. This message shows the origin and history of the data set, its dimensionality and cardinality.

Interaction Examples

To discuss the approach, we provide a video showing a user interacting with the application on our website ¹. In the first scene the user demonstrates different functions and explores the data space. The user chooses displayed dimensions and triggers shock waves in the high- and two-dimensional space. Then you can see and hear how the data sonogram evolves over time, starting with one of two pitch levels, depending on where the shock wave is initiated. Thereby the regions of different classes in the data set can be well discerned. Furthermore overlapping classes and class boundaries can be perceived.

The Growing Neural Gas Sonification Model

Growing Neural Gas (GNG), introduced by Fritzke in [21], is an undirected learning algorithm that incrementally 'grows' a network graph into the data distribution. The GNG is a network of neurons and connections between them. During the learning process, the neurons are moved to minimize the error with respect to the original data. New neurons are inserted and connections between them age or are reinforced.

The Growing Neural Gas Sonification Model introduced in [20] is described in brief below. It is categorized in the same way as the Data Sonogram Sonification Model before:

Setup: For the GNG Sonification Model, the connections in the GNG graph are used as transducers that transport energy between the neurons. The frequency of a neurons' tone is determined by the number of connections emanating from it: for each connection, a quint is added to the base frequency.

¹ <http://sonification.de/publications/TuennermannKolbeBovermannHermann2009-SIFSM/>

Dynamics: Using the energy flow equation (1), the energy for each neuron is calculated. It decays over time, depending on parameters g and q (which the user is able to adapt) and the current state of the GNG graph. The energy of each neuron determines the amplitude of the respective tone.

$$\frac{dE_i}{dt} = -gE_i(t) - \sum_{j \in I_N(i)} q(E_i(t) - E_j(t)) \quad (1)$$

The parameter g steers the exponential energy decay, q determines the amount of energy that flows to every neighboring neuron each step. $E_i(t)$ is the energy of the neuron i , $I_N(i)$ is the set of neurons that are connected to neuron i .

Excitation: The user can induce energy into a neuron by tapping near it. This can be done at multiple points simultaneously or subsequently. The energy then propagates through the GNG until equilibrium is reached again.

Link Variables: The sonification is the superimposed sound signal from all existing neurons. This consists of one tone per neuron, with the frequency determined by the number of connections to other neurons and the amplitude determined by the current energy level of the neuron.

Listener: The resulting sonification for all neurons is presented to the user as well as the coupled visual feedback.

Overview

To benefit from the interaction capabilities of the tDesk, the GNG Sonification was reimplemented and simplified with multi-touch capabilities in mind. The goal was to be able to explore the GNG while it was growing, using only the fingers to deliberately excite the sonification. The user should not have to worry about setting up the program and initializing the GNG parameters, but be able to intuitively grasp the structure of the adapting GNG.

To start with, the user is presented with a two-dimensional scatterplot of the data. Five controls are available on the lower right corner, from top to bottom:

- the maximum number of neurons for the GNG
- the maximum age of connections between neurons
- the learning rate parameter
- the energy flow rate (parameter q in eq. (1))
- the energy dissipation rate (parameter g in eq. (1))

The first two parameters control the GNG algorithm itself, the third determines the speed of the learning process. The GNG has more tunables (see [21] for details), but they have well-working default values and most often do not need to be adapted. A future improvement will make it easier for the user to adjust them, without being overwhelmed by too many configuration options to choose from. The last two parameters define the energy decay and transport for the sonification. Finally, in the lower right corner are two buttons: The bottom one starts or pauses the adaptation process of the GNG, the upper one resets the

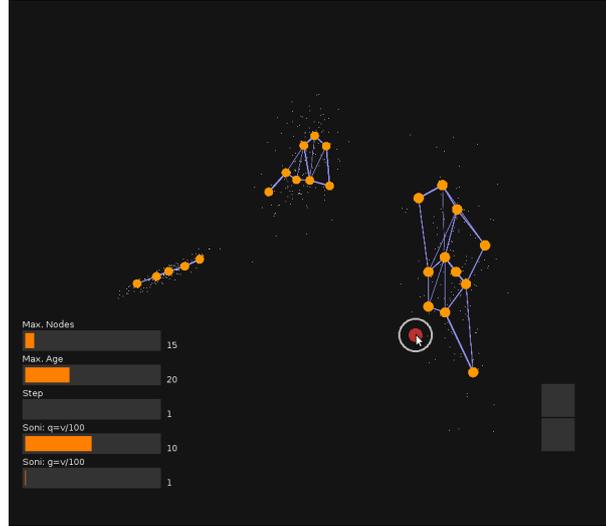


Fig. 6. The user interface of the GNG Sonification application, showing a two-dimensional scatterplot of the *three cluster* dataset overlaid with a GNG during its adaption process. Beneath the right cluster of data, the user induces energy into the network.

GNG to its initial state, taking into account the changes the user made to the three initialization parameters.

Fig. 6 depicts an example of a GNG during the adaption process, using the *three cluster* dataset, which contains three clusters of an intrinsic dimensionality of two, four and eight each. It is a synthetic dataset used to evaluate the GNG Sonification Model, with 449 data points in total.

The bright circles represent the neurons, initially with an energy level of zero. The lines show the connections between them, with their thickness representing their age. In a future implementation of this sonification, we will take the age into account as well to further direct the energy flow to the neighboring neurons.

Implementation Details

The GNG sonification is implemented in python, utilizing the *Python Modular toolkit for Data Processing* [24] for the calculations. For the user interface and multi-touch interaction, *PyMT - A Multi-touch UI Toolkit for Pyglet* [23] is used. The sonification is synthesized in SuperCollider [19], utilizing Stinson's *OSC interface for Python* [25].

Interaction

When tapping near a neuron, the user induces energy into it. This leads to an energy flow within the network of neurons that is immediately sonified. At the

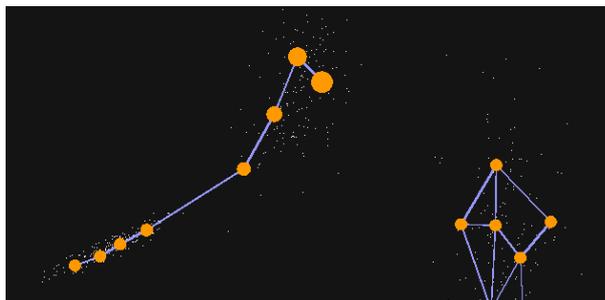


Fig. 7. An earlier state of the adaption process for the *three cluster* dataset. After the user induced energy into the rightmost neuron of the left cluster, the sonification begins to sound as the energy finds its way through the network. We visualize the current energy of a neuron through its size. Note that the right cluster does not emanate any sound, as the energy induced into the left cluster cannot possibly reach it at this point in the adaptation process.

same time, it is visualized – the sizes of the neurons indicate their current energy level (see Fig. 7). The user can influence the speed of energy decay within the network through independently adjusting the g and q parameters of the energy flow equation. For example, setting both to higher values leads to a faster decay and simultaneously reduces the distance the energy travels through the network, resulting in a more localized excitation of the sonification. In choosing lower values, the distance broadens and the decay slows down. In experimenting with different settings during different stages of the adaption process, or repeatedly inducing energy at similar points during the process, the user directly interacts with the sonification model and thereby is able to gain insights in the analyzed data.

For example, Fig. 6 shows an excitation of the GNG sonification on the lower right part. The resulting sound is very bright and slowly fading, as every neuron has between two to six connections to other neurons, resulting in the highest tone being the sixth quint over the base frequency. This indicates a high intrinsic dimensionality and, in fact, the underlying data distribution of this cluster is 8-dimensional. Would the user tap towards the left cluster of neurons, the highest frequency would be the second quint over the base frequency, indicating a less complex network and lower intrinsic dimensionality. The left cluster contains data that have only two independent variance dimensions.

In Fig. 7, the GNG is shown in an earlier state of the adaption process. The user induced energy into the rightmost neuron of the left cluster, so that it flows through the left part of the network only. The pitch of the sound is low, as each neuron has at most two neighbors. When excited while the GNG is adapting, new connections made to other neurons or newly added neurons are clearly audible through a rise in pitch. When connections are deleted, the sound suddenly becomes lower in pitch.

Growing Neural Gas is an undirected learning algorithm, but there exists no established decision criteria as to when it has fully grown into its dataset. After a while, overfitting occurs and ultimately the learned structure becomes diffused again. The user has to make an informed decision as to when to end the learning process, and the GNG Sonification Model provides a multi-modal and highly interactive tool to do just that.

Discussion and Conclusion

In this paper we have presented a multi-touch interface for the excitation of sonification models. We have reimplemented the Data Sonogram and the Growing Neural Gas sonification model and demonstrated multi-point multi-user exploration of scientific data via surface-based interaction.

The main advantage of our approach is that a very natural contact between the user and the surface (as the physical representation of the data) can be established. Interaction modes with typical real-world surfaces such as tapping, hitting, pushing and scratching provide examples of how interactions can be profitably used in the context of sonification models. With the two sonification models we have given first examples that show how spatially-resolved tapping on the surface can be utilized as a tapping into data spaces, using quasi-physical dynamic processes in the space of the sonification model to associate meaningful acoustic responses which then represent the data to the user. In result a qualitative experience is created from the ongoing continuous interaction.

An important aspect is that the interface connects the auditory and visual representation and binds them via the surface to multi-modal data perceptualization units. Synchronization is a key component for the user to be able to connect visual and auditory elements. Since the interaction occurs in the same frame of reference, and tactile sensations complement the experience, a tightly closed interaction loop is created.

In our future work we will particularly focus on sonification models that allow to explore yet *untouched* aspects of continuous interaction with data distributions. Instead of providing a trigger only, we want to enable users to continuously deform data representations in order to perceive the resulting tension by these deformations as informative sound. Sonification can be used in various multi-touch applications. For instance, for didactic applications, the real-time sonification of variables (e.g. stress, magnetic field strength, etc) while interacting with a simulation of a system can deliver complementary information to what is visible on the surface. Also, auditory games where the goal is to competitively or jointly shape sounds via physical interaction with the surface offer a great potential to explore tactile computing in a yet unseen way.

In summary, the presented multi-touch sensitive surface enriches the available modes to interact with complex data and to perceive structure-related features as sound via Model-Based Sonification. The tight coupling of visualization, sonification, tangible interfaces and continuous interaction in one interface contributes to a truly multi-modal experience and shows the potential of an increased level

of understanding of structures in the data. The scope of our ongoing research is to explore and quantify the possibilities in this direction.

References

1. Bovermann, T., Hermann, T., Ritter, H.: A tangible environment for ambient data representation. In *First International Workshop on Haptic and Audio Interaction Design*, volume 2, pp. 26–30, Aug 2006.
2. Buxton, B.: Multi-touch systems that i have known and loved. URL, 2009. <http://www.billbuxton.com/multitouchOverview.html>.
3. Dourish, P.: *Where the action is: The foundation of embodied interaction*. MIT Press, 2001.
4. Forina, M.: Arvus - an extendible package for data exploration, classification and correlation. <http://www.radwin.org/michael/projects/learning/about-wine.html>.
5. Han, J. Y.: Low-cost multi-touch sensing through frustrated total internal reflection. In *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*, pp. 115–118, New York, NY, USA, 2005. ACM.
6. Fisher, R. A.: UCI Repository of Maschine Learning Databases – Iris Data Set. 1999.
7. Fry, B., Reas, C.: Processing programminig environment. URL, 2001. <http://processing.org/>.
8. Heidegger, M.: *Sein und Zeit*. Niemeyer, Halle a. d. S., 1927.
9. Hermann, T.: Taxonomy and definitions for sonification and auditory display. In Katz, B. editor, *Proc. Int. Conf. Auditory Display (ICAD 2008)*, France, 2008. ICAD, ICAD.
10. Hermann, T., Ritter, H.: Listen to your data: Model-based sonification for data analysis. In Lasker, G. E. editor, *Advances in intelligent computing and multimedia systems*, pp. 189–194, Baden-Baden, Germany, 08 1999. Int. Inst. for Advanced Studies in System research and cybernetics.
11. Hermann, T.: Sonification for exploratory data analysis. PhD thesis, Bielefeld University, Febr. 2002 URL: <http://www.techfak.uni-bielefeld.de/ags/ni/publications/media/Hermann2002-SFE.pdf>.
12. Hermann, T., Krause J., Ritter, H.: Real-time control of sonification models with an audio-haptic interface. In Nakatsu, R. and Kawahara, H. editor, *Proc. Int. Conf. Auditory Display (ICAD 2002)*, pp. 82–86, Kyoto, Japan.
13. Hermann, T., Milczynski, M., Ritter, H.: A malleable device with applications to sonification-based data exploration. In Stockman, T. editor, *Proc. Int. Conf. Auditory Display (ICAD 2006)*, pp. 69–76, London, UK. Department of Computer Science, Queen Mary, University of London.
14. Hunt, A., Hermann, T., Pauletto, S.: Interacting with sonification systems: Closing the loop. In Banissi, E. and Börner, K. editors, *IV '04: Proceedings of the Information Visualisation, Eighth International Conference on (IV'04)*, pp. 879–884, Washington, DC, USA, 07 2004. IEEE Computer Society.
15. Levenberg, K.: *A method for the solution of certain non-linear problems in least squares*. Quarterly of Applied Mathematics pp. 164–168, 1944.
16. Marquardt, D.: *An algorithm for least-qquares estimation of nonlinear parameters*. SIAM Journal on Applied Mathematics 11 pp. 431–441, 1963.

17. Kramer, G. (ed.): *Auditory display: sonification, audification, and auditory interfaces*. Addison-Wesley, 1994.
18. Kramer, G., Walker, B., Bonebright, T., Cook, P., Flower, J., Miner, N., Neuhoff, J., Bargar, R., Barrass, S., Berger, J., Evreinov, G., Fitch, W. T., Grohn, M., Handel, S., Kaper, H., Levkowitz, H., Lodha, S., Shinn-Cunningham, B., Simoni, M., Tipei, S.: Sonification report: Status of the field and research agenda, 1997.
19. McCartney, J.: SuperCollider hub. URL, July 2004. <http://supercollider.sourceforge.net/>.
20. Hermann, T., Ritter, H.: Neural gas sonification: Growing adaptive interfaces for interacting with data. In Banissi, E. and Börner, K. editors, *IV '04: Proceedings of the Information Visualisation, Eighth International Conference on (IV'04)*, pp. 871–878, Washington, DC, USA, 07 2004. IEEE Computer Society.
21. Fritzke, B.: A growing neural gas network learns topologies. In Tesauro, G., Touretzky, D. and Leen, T. editors, *Advances in Neural Information Processing Systems*, volume 7, pp. 625–632. The MIT Press, 1995.
22. Kaltenbrunner, M., Bovermann, T., Bencina, R., Costanza, E.: TUIO: A protocol for table based tangible user interfaces. In *Proceedings of the 6th International Workshop on Gesture in Human-Computer Interaction and Simulation (GW 2005)*, Vannes (France).
23. Hansen, T., Hourcade, J. P., Virbel, M., Patali, S., Serra, T.: PyMT: A post-WIMP multi-touch user interface toolkit. In *Proceedings of the International Conference on Interactive Tabletops and Surfaces (ITS 2009)*, Banff (Canada).
24. Zito, T., Wilbert, N., Wiskott, L., Berkes, P.: Modular toolkit for data processing (MDP): A python data processing frame work. *Frontiers in Neuroinformatics* (2008). URL, <http://mdp-toolkit.sourceforge.net>.
25. Stinson, P. K.: SCOSC: SuperCollider OSC interface for python. URL, October 2009. <http://trac2.assembla.com/pkaudio/wiki/SuperCollider>.