# SONIFIED AEROBICS – INTERACTIVE SONIFICATION OF COORDINATED BODY MOVEMENTS

*Thomas Hermann and Sebastian Zehe*

Ambient Intelligence Group
Center of Excellence in Cognitive Interaction Technology (CITEC)
Bielefeld University, Bielefeld, Germany
`thermann@techfak.uni-bielefeld.de`

## ABSTRACT

This paper introduces a new hard-/ and software system for the interactive sonification of sports movement involving arm- and leg movements. Two different sonifications are designed to convey rhythmical patterns that become auditory gestalt so that listeners can identify features of the underlying coordinated movement. The Sonification is designed for the application to enable visually impaired users to participate in aerobics exercises, and also to enhance the perception of movements for sighted participants, which is useful for instance if the scene is occluded or the head posture is incompatible with the observation of the instructor or fitness professional who shows the practices in parallel. Furthermore, the system allows to monitor fine couplings in arm/leg coordination while jogging, as auditory feedback may help stabilizing the movement pattern. We present the sensing system, two sonification designs, and interaction examples that lead to coordination-specific sound gestalts. Finally, some qualitative observations are reported from the first uses of the prototype.

## 1. INTRODUCTION

Sonification [1] and particularly Interactive Sonification [2] offers the ability to represent multivariate time series as auditory streams so that coordinated temporal patterns become auditory gestalts. This powerful capability, when applied to sensor data recorded during sports exercises can turn a rhythmic activity into a sonic rhythm so that the listeners can understand the coordination during the movement. Specifically, for sports exercises such as in aerobics[1], fitness gymnastics, and tai chi, to name a few, there are situations where an instructor or tutor presents an exercise to be imitated by the participants. These trainings may be accompanied with music to force a specific rhythm and thus a given execution rate to be followed by the participants. There are, however, problems that participants face in such situations: for instance, the body orientation of the trainer obstructs some details of their movements (e.g. hand movements) so that it is difficult from the view angle to copy the movement pattern correctly. Also, other participants may be in the direct line of sight. Furthermore, some exercises may require the participants to look downwards or to have their heads oriented so that it is impossible to look at the trainer. A particular problem occurs for visually impaired participants who normally do not participate in such sports training since it is impossible for them to copy the (only visually accessible) movement patterns. One alternative would be that trainers explain the movements verbally and also when to change a coordination pattern (e.g. to change from in-phase to alternating arm swings). Yet the verbal channel is not rich enough to explain the details of the exercise sufficiently enough to convey all required information, particularly if the user can not see it at all.

How can this application and problem profit from sonification? The apparent problems motivate to develop a sonification that represents movement details by means of real-time rendered sound streams so that the listeners perceive both the rate (speed) of execution but also details of the coordination between the limbs (arms and legs). The rate is particularly relevant if the aerobics is conducted without accompanying music, but even when performed with music, it remains unclear whether the movements repeat every beat, twice a beat or every two beats.

The goal is to sonify the body movements so that – after some learning time – visually impaired users are enabled to conduct the exercises with much reduced verbal corrections or explanations. A second goal is to support sighted users by offering signals that draw their attention to aspects of the motion or coordination which are difficult to pick up from a mere visual observation.

In this paper we demonstrate a first prototype for the aerobics sonification for the restricted case of knee-and-elbows sensing, which already covers a variety of different bodily coordination patterns.

---

[1] `http://en.wikipedia.org/wiki/Aerobics` (last accessed 2011-02-15)

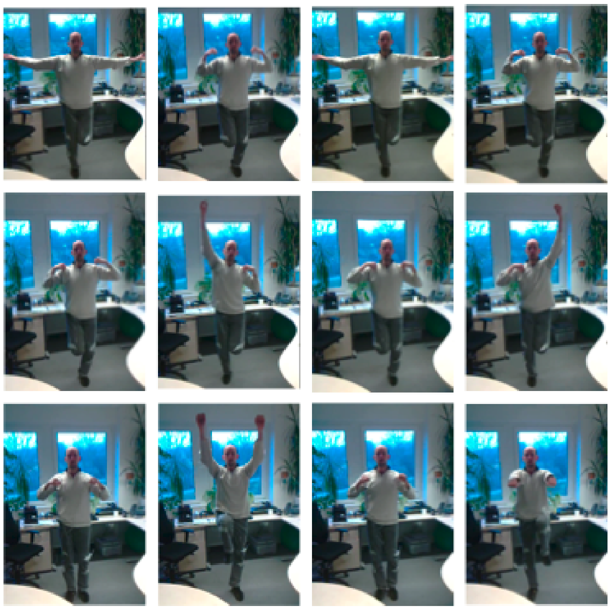


Figure 1: Illustration of 3 coordination exercises: top row: synchronized arm bend/stretch, 2$^{th}$ row: alternating arm-stretch crossed with legs, bottom row: parallel arms with alternating legs.

## 2. MOVEMENT PATTERNS IN COORDINATION EXERCISES

The use of sonification in sports and movement has meanwhile some tradition, e.g. for swimming strokes, jumps, rowing, running etc. [3, 4, 5, 6]. It has been considered as a replacement of visual information for players with visual impairment for ball games in [7]. Movement sonification also proved useful in areas such as support for musical performance [8] or analysis of body movements from their real-time or offline sonification [9, 10].

The sort of exercises we focus on in this paper is arm-leg coordination exercises performed at a fixed location (in contrast to jogging / running exercises where the participant changes continuously the location). Even here, we have a variety of movement patterns: the arms can be stretched upwards, sidewards, downwards, both arms either synchronized or alternating, or any combination (e.g. left arm alternating stretched sidewards and bent while right arm alternating bent and stretched sidewards). Beyond this already large variety of possible rhythmic movement sequences, the combination with the bending of knees, or on-the-spot-walking gives even more possibilities. For instance, when both arms are stretched in front of the body while the left knee is bent, arms bent while both feet are on the floor, arms stretched upwards while right leg is bent and arms bent while both feet touch the floor. Obviously, from reading the past paragraph, a verbal instruction is much too time-consuming to accompany the exercise in real-time.

Besides an exercise effect on the arm and leg muscles these coordination exercises train complex interrelationships and increase the fine-control of the body coordination, so the coordination is not irrelevant to the practitioner. Also the different modes of synchronization can result in subtle training effects such as spinal cord torsion which would otherwise not be obtained.

Figure 1 depicts some frame-sequences of exercises used here for sonification. The shown exercises are the most basic ones which can be measured with few sensors, and the loop size is 4, meaning that every for steps the pattern repeats. More complex patterns may easily have 8 or 16 steps before the pattern loops. As proof-of-principle we start with patterns that can be understood by recording the elbows and the knees' angles.

## 3. WEARABLE WIRELESS SENSING AND DATA PROCESSING

For sensing motion-relevant information we track the arm and leg posture of the subject using goniometers as motion sensors. These instruments measure the angle between two attached sticks around a common axis. In our sensors, a potentiometer is used to convert the measured angle into electrical resistance. A potentiometer is a resistor with three terminals that form an adjustable voltage divider.

We read out these voltage dividers using the analog inputs of an Atmel Atmega168 microcontroller (see Fig. 2). This enables us to finally read the angle of the goniometer with a resolution of 10 bit, leading to a precision of around 3.8 LSB/degree. We stream this motion data to a PC via Bluetooth with a rate around 120 Hz.

The self-made goniometers are surprisingly accurate while being very inexpensive and easy to build. In our setup, we equipped the subject with four goniometers, one on each elbow and knee. The goniometers were attached using adhesive tape. It showed that the sensors do not restrict the subjects' movements.

## 4. SONIFICATION DESIGN

We start the discussion of the sonification design with an analysis of the information that the user needs and usually uses for copying the movement pattern.

Firstly, in the regular, vision-supported situation, the user observes the instructor and can thereby understand directly the movement sequence in all relevant degrees of freedom. However, the visual focus is only on a single location (fovea) at a time and thereby eye-movement will be used to understand different aspects in more detail. Undoubtedly this visual information is the largest source users

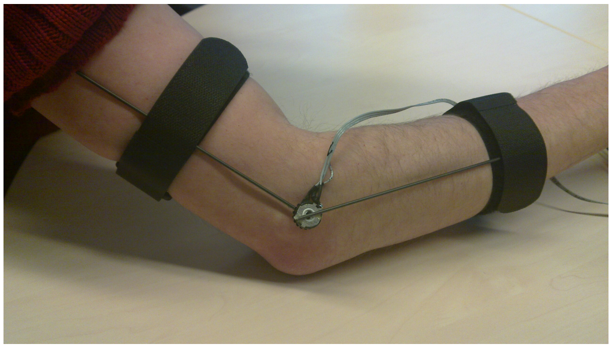


Figure 2: Photo of a goniometer-setup attached to the elbow. The potentiometer acts as a voltage divider, the angle-dependent voltage is sampled and digitized in a microcontroller.

rely upon to understand, imitate or reproduce the exercise. It can furthermore be assumed that from the mere observation neural mirror systems are activated that generate this pattern as motor behavior.

However, some movements may be occluded by the instructor's body so these can not be judged from visual observation. Also, some details may be too subtle or hidden by a larger and fast movement that it is impossible to attend to them, so that there may be information deficits. A pattern that is quite well understood visually is the symmetry of the body during the coordinated movement pattern, i.e. that arms and legs perform movements that show a symmetry with respect to the body axis. Yet subtle deviations between the arms, such as if one arm would return slightly earlier than the other or has a stronger amplitude than the other will be difficult to catch from visual observation. Certain exercises may contain instructions which are invisible such as if the instructor asks to stretch the abdominal muscle tightly.

Interestingly, sound is already a carrier of information in observing the instructor, since both the contact sounds of feet with the floor and other body sounds such as when the arms touch the body deliver helpful information to structure the own movement. In addition, if music accompanies the aerobics this is quite efficient to synchronize the participants in their movement rhythm.

These observations make clear how crucial visual information is for the task and where the limits are. For the design of sonifications the main question will be whether sound shall aim (a) to replace the visual information, for instance to enable visually impaired users to understand the pattern, or (b) to help the users by providing complementary information to what is available from visual observation, for instance to emphasize aspects such as synchronization, movement speed or acceleration. With the two sonification designs presented in this paper, we mainly aimed at (a), to enable a replacement of the visual information, yet we found that the sonifications also provide helpful information concerning (b).

For the substitution of visual information the immediate idea is that at any time point the sound needs to represent a decodable sound that corresponds to the body articulation at that time. This idea leads to stationary sounds that convey as many degrees of freedom (here: joint angles) as needed to cover the whole set of exercises. To include all joints of the human body is an enormous challenge, both for the sensing and the sonification design and this will very likely also overburden the user with sonic complexity. For that reason we limited the task to a smaller subset of exercises, namely arm-leg coordination where furthermore the only required degrees of freedom are elbow and knee bending angles.

An important feature of the relevant exercises is that they are repetitive, with a time period of few seconds. This allows the listeners to attend to changes between subsequent patterns. In consequence, the exercise will turn into a repetitive sound stream which allows the users to employ active listening to better understand the relation of movement parts.

An important part of the performance is to assess whether the own movement is similar or different from the instructor's movement. Without sonification, we here mainly rely on our proprioception, and to a limited degree also on visual observation of our own body. Since birth we are used to relate visually observed patterns to our motor actions. However, in the auditory domain this will be very new and difficult! We identify basically two alternative options to integrate sonification into a practical scenario: (a) the instructor's movements are sonified so that the listener can identify from listening alone the actual movement – an approach which will need much learning, or (b) both the instructor's and the own's movements are sonified so that the listener can attend to the differences. An idea for that would be to present the instructor sonification on the right ear while the own's movement sonification appear on the left channel using headphones. Alternatively, both signals may use different timbres, leaving left/right free for the spatial reference. Probably such a simultaneous sonification will overburden the listener, and the repetitive nature of the exercises opens a third alternative: to allow a A/B-switch between instructor / own movement sonification. This could either occur automatically every two patterns or on the user's initiative.

In the current prototypes we focus on the underlying sonification methods, and these may certainly need optimization according to the actual application context as described above. With the following approaches we introduce two methods to convey limb/action identification and also to turn repetitive patterns into auditory rhythms that become characteristic and recognizable auditory gestalts. Aspects of Gestalt formation by controlling interrelated complex timbre sequences for visually impaired users has been investigated by Grond et al. in [11].

### 4.1. Excitatory Mapping Sonification Design

Our first sonification approach uses the real-time data of elbow and knee angles for a basic continuous parameter-mapping sonification, modified so that the sonification becomes excitatory: without changes in posture, the sound turns to silence. This feature is extremely helpful to avoid largely disturbing sound during physical inactivity in any limb.

For sound synthesis, we create four continuous sound streams using filtered noise with a sharp resonant peaked band-pass filter so that it rings at an adjustable frequency. The parameters are the noise level, the center frequency, the spatial panning and the filter bandwidth. It seemed intuitive to associate left/right limbs to the left/right audio channel to obtain a natural spatial reference. Furthermore we define the center frequencies to identify limbs, representing legs by a significantly lower pitch than arms, in line with the dominant association that vertical position is represented as pitch. The actual angle values, as measured from the goniometers are then mapped to a frequency deviation of the limb sound stream center frequency in certain limits (e.g. three semitones). By the combined pitch and panning the sound both identifies the limb type, the body side and conveys the actual value. The difference between the fully stretched and bent joint is clearly audible. The detailed movement can be followed since frequency is used as a continuous parameter. Sound stream level is not directly controlled by the data. Instead we use a derived feature which measures the activity in the limb: $\bar{\omega}_i = |\phi_i(t) - \phi_i(t-1)|$ is the absolute value of the angle difference between successive measurements $\phi$ for each limb $i$ and we map it linearly to amplitude, clipping small values to zero so that only angular velocities larger than a threshold $\omega_{\text{min}}$ become audible. As a modification we also considered to feed $\bar{\omega}$ into a leaky integrator with adjustable leak rate to obtain a smoother level decay, yet the former more simple mapping showed to be sufficient and also more direct.

### 4.2. Event-based Sonification Design

Our second approach aims to reduce the sound as much as possible so that it becomes more sparse to free the auditory space for more specific and articulated sonic cues that identify critical or significant events. These are those that segment the movement into patterns. As shown in the Fig. 1, we can very well mentally interpolate the continuous movement from few key frames which represent turning (or action) points in the cycle. These key points normally correlate with extrema of the joint angles. This motivates the reduction of the continuous sensor signal to fewer turning point events for each limb. Since these events are few compared to the sensing frame rate, they can easily be represented by more complex sounds, such as an impact sound, or even a transient musical instrument sound.

With some experimentation we found that percussive sounds are particular useful since (a) they offer good identification and precise location in time, (b) listeners can well differentiate a large number of percussive instruments and details therein, (c) they even work to some degree in presence of background music. Furthermore, at least from observing drummers, we have the idea that each limb is associated with a specific drum instrument (e.g. right hand: cymbal, right foot: base drum, etc.). The metaphor of the instructor as a performer at a virtual drum is certainly too narrow, but it is exemplary for other limb-specific bindings.

Technically, we need to identify robustly the turning points of joint angles $\alpha_i$, which we achieve by the following procedure: (index $n$ corresponds to time $t = t_0 + n\Delta t$)

- for each joint we initialize a 'hysteresis' counter $h_i[n] = 0$.
- on each step, we set $h_i[n+1] = h_i[n] + \text{sign}(\alpha_i[n] - \alpha_i[n-1])$
- we clip the value to the range $[-3, 3]$.
- if the $h_i$ changes the sign: trigger a sound event for joint $i$ with situation-specific parameters.

The introduction of the hysteresis offers the advantage to avoid overly frequent events due to signal noise. However, it introduces also a slight delay (about 30 milliseconds at a rate of 120 Hz).

As initial sound design we tried four different bells for the four limbs, and as better working alternative setting a mixed percussion / instrument set.

The sonic details of the sound to play can then depend on different parameters: whether the joint angle reached a minimum or maximum, what the actual value is, what the highest speed was during the past movement segment, etc. We found it interesting to experiment with sound tuples that have some association, such as a 'tic' and 'toc' sound for stretching / bending one joint. For the right knee we used a base drum sound on the stretched (angle maximum) and a snare drum sound for the bent (angle minimum) state. This creates a binary drum machine rhythm on a running-on-the-spot pattern. Likewise we used two differently pitched electric base tones for the other knee, differently pitched cymbals for the right arm and congas and a clavinet sound for the left arm.

Figure 3: Screenshot of the GUI in sc3: The interface allows the playback of data synchronous to the recorded video, allowing playback control (via the rate slider) and real-time adjustment of the mapping functions. This accelerates and facilitates the optimization of the sonification.

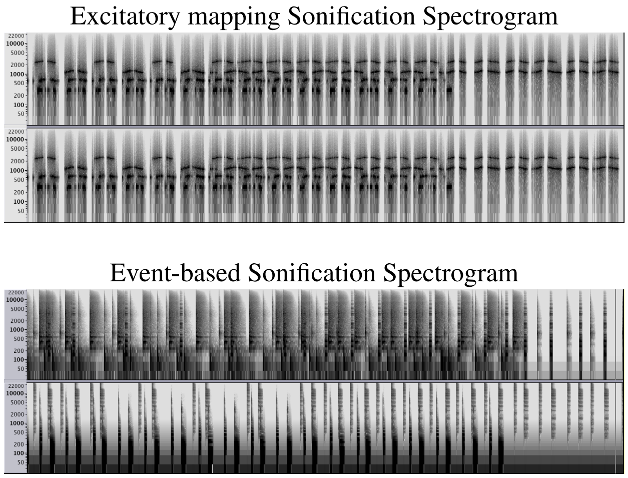
### 4.3. Implementation in SuperCollider

The sonification is fully programmed in SuperCollider. Basically two operation modes are supported: (i) real-time sonification of the sensor data arriving via Bluetooth. Here the program collects the data vectors, adds a time stamp and furthermore allows to save a session to a comma-separated values file. The second mode is (ii) offline playback of recorded data, which allows to replay previously recorded sessions in synchronization with the video playback. This is ideal to work on the sonification design while playing selected movement patterns in a loop. The offline mode furthermore allows slow-motion or speed up of the frame rate which allows to investigate the timing in more detail.

At the core of the player routine is a SuperCollider *Task* which iterates at a certain frame rate (either real-time or file-driven) and processes the incoming data vectors. It sends the sensor vectors as OSC messages, where either the playback data or the real-time sensor data are processed and sonified as they come in. There are furthermore *init* functions for the sonification at the beginning and a *free* function to free the used synths. Different sonification approaches can be implemented and interactively exchanged by overwriting the current definitions during runtime of the system. A code example is provided on our website[2].

## 5. APPLICATION EXAMPLES

We demonstrate the two sonification types discussed before with an interaction video which are provided on our website at `http://www.techfak.uni-bielefeld.de/ags/ami/publications/HZ2011-SAI`. The excitatory mapping sonification videos show typical exercises from the aerobics program. Fig. 4 depicts a spectrogram of the sonification. It is easy to recognize the chirps in pitch. Since the body movement is quite coordinated (arms move together at the beginning) it is first difficult to understand what part of the sound represents what, yet when the users are equipped with the sensors themselves so that they can relate directly their movements to the resulting sound, each movement creates immediately stimuli at hand of which one can – to our experience – quickly learn the association.


Excitatory mapping Sonification Spectrogram

Event-based Sonification Spectrogram

Figure 4: Spectrograms of the two sonifications (first = left channel). The first pattern in time shows alternating arm stretching, then off-phase arm stretching of both arms and in a third episode parallel arm stretching without walking. The gaps show that without activity the sound fades to silence. For the Event-based Sonification, small differences in timing are much more difficult to identify visually than with listening to the sound.

The interaction video for the event-based sonification shows, for comparison, the same movement sequence as before with the percussive sound event selection. Obviously a more musical rhythm emerges and the sonification emphasizes the perception of differences in timing between the turning points of different limbs. The association from instruments to joints is easier learned than with the mapping-based approach. Only from this sonification our attention was drawn to an otherwise neglected pattern: a double base attack occurs on each step. Looking into the video in slow motion we found that indeed the legs are bent after land-

---

[2] http://www.techfak.uni-bielefeld.de/ags/ami/publications/HZ2011-SAI

ing and stretched again to jump off. Before hearing it in the sonification, this feature was almost overseen in visual observation.

## 6. DISCUSSION & CONCLUSION

This paper presented two new sonification approaches for the sonic representation of coordinated body movements for a single performer, as they occur in physical exercises and particularly in aerobics. The main motivation was to render auditory augmentations of body movements as a substitution of visual information for situations where visual observation is not possible, e.g. to enable users with visual impairments to participate to aerobics classes, or where the instructor's body occludes relevant parts of the movement. Furthermore, we expected sonification as a useful augmentation of body movements to support the better understanding of movement features which are difficult to grasp from visual observation alone, specifically concerning aspects such as degree of synchronization, detailed timing.

We introduced a practical method to measure the movements using wearable goniometers, sending digitized data via bluetooth to a computer. We presented a Supercollider program to render data sonifications either in real-time or from pre-recorded data.

The two sonification types demonstrate different approaches to sonify the continuous data stream. The excitatory parameter mapping sonification is closer to the measured signals and thereby contains more of the detailed information available in the actual movement. The Event-based Sonification reduces the information to focus on sonic 'key frames' that allow to perceive coarsely what activity is to be performed with a limb at a time. Being more coarse, the information may not suffice to reproduce the movement accurately, but it is perhaps superior to recognize the pattern, and to emphasize the detailed synchronization between joints. From our current subjective experience, the event-based sonification is more efficient to discover and self-regulate subtle rhythmical irregularities.

Certainly these two prototypes need further evaluation, both concerning the acceptance by users, their value for visually impaired users to pick up and reproduce patterns, and their discrimination power to learn how suitable the sonifications are to differentiate between movement types. While this paper focussed on an initial prototype for the application and on sonification techniques, the studies are subject of our ongoing research.

## 7. ACKNOWLEDGMENT

We thank the German Research Foundation (DFG) and the Center of Excellence 277 Cognitive Interaction Technology (CITEC) that enabled this work within the German Excellence Initiative. We thank Tobias Grosshauser for his initial involvement in building the sensor system. Thanks to Alexandra Barchunova and Diana Grandi for trying the system.

## 8. REFERENCES

[1] T. Hermann, "Taxonomy and definitions for sonification and auditory display," in *Proc. 14th Int. Conf. Auditory Display (ICAD 2008)*, B. Katz, Ed., ICAD. Paris, France: ICAD, 06 2008.

[2] T. Hermann and A. Hunt, "An introduction to interactive sonification (guest editors' introduction)," *IEEE MultiMedia*, vol. 12, no. 2, pp. 20–24, 04 2005.

[3] A. Effenberg, J. Melzer, A. Weber, and A. Zinke, "Motionlab sonify: A framework for the sonification of human motion data," in *The 9th International Conference on Information Visualisation (IV'05)*. IEEE Press, Jul. 2005, pp. 17–23.

[4] A. O. Effenberg, "Movement sonification: Effects on perception and action," *IEEE Multimedia*, vol. 12 (2), pp. 56–69, 2005.

[5] G. Dubus and R. Bresin, "Sonification of sculler movements, developement of preliminary methods," in *Proceedings of ISon 2010 - Interactive Sonification Workshop : Human Interaction with Auditory Displays*, R. Bresin et al., Eds., ISon.

[6] T. B. Stephen Barrass, Nina Schaffert, "Probing preferences between six designs of interactive sonifications for recreational sports, health and fitness," in *Proceedings of ISon 2010 - Interactive Sonification Workshop : Human Interaction with Auditory Displays*, R. Bresin, et al., Eds., ISon.

[7] T. Hermann, O. Höner, and H. Ritter, "Acoumotion - an interactive sonification system for acoustic motion control," in *Proc. 6th Int. Gesture Workshop*, ser. LNCS, S. Gibet, N. Courty, and J.-F. Kamp, Eds., vol. 3881/2006. Berlin, Heidelberg: Springer, 2006, pp. 312–323.

[8] T. Grosshauser and T. Hermann, "Wearable setup for gesture and motion based closed loop audio-haptic interaction," in *ICAD 2010: Proceedings of the 16th International Community for Auditory Display*, Washington, USA, 2010.

[9] F. Grond, A. Bouenard, T. Hermann, and M. M. Wanderley, "Virtual Auditory Myography of Timpani-playing Avatars," in *Proceedings of the 13th Int. Conference on Digita Audio Effects (DAFx-10)*, Graz, Austria, Sept. 2010.

[10] F. Grond, T. Hermann, V. Verfaille, and M. M. Wanderley, "Methods for effective sonification of clarinetists' ancillary gestures," in *Gesture in Embodied Communication and Human Computer Interfaces: Proc. 8th Int. Gesture Workshop*, ser. LNCS, S. Kopp and I. Wachsmuth, Eds. Berlin, Heidelberg: Springer Verlag, 02 2009.

[11] F. Grond, T. Drossard, and T. Hermann, "SonicFunction: Experiments with a function browser for the visually impaired", in *Proc. ICAD 2010*. Washington D.C.: ICAD, 2010.