

Auditory Augmentation

Till Bovermann, CITEC, Bielefeld University, Germany

René Tünnermann, CITEC, Bielefeld University, Germany

Thomas Hermann, CITEC, Bielefeld University, Germany

ABSTRACT

With auditory augmentation, the authors describe building blocks supporting the design of data representation tools, which unobtrusively alter the auditory characteristics of structure-borne sounds. The system enriches the structure-borne sound of objects with a sonification of (near) real time data streams. The object's auditory gestalt is shaped by data-driven parameters, creating a subtle display for ambient data streams. Auditory augmentation can be easily overlaid to existing sounds, and does not change prominent auditory features of the augmented objects like the sound's timing or its level. In a peripheral monitoring situation, the data stay out of the users' attention, which thereby remains free to focus on a primary task. However, any characteristic sound change will catch the users' attention. This article describes the principles of auditory augmentation, gives an introduction to the Reim Software Toolbox, and presents the first observations made in a preliminary long-term user study.

Keywords: Ambient Computing, Auditory Augmentation, Auditory Display, Interaction Design, Sonification, Tangible Interface

1. INTRODUCTION

The world around us is full of artificially gathered data. Upon that data we draw conclusions and make decisions, which possibly influence the future of our society. The difficulty hereby is not the data acquisition – we already have plenty – but our ability to process it (Goldhaber, 1997). Arising from this circumstance, at least two demands for data preparation can be identified: first, it should gain an appropriate amount of its user's attention depending on both the data domains' nature and the users' needs (Goldhaber, 2006), and second, it should utilise appropriate representations that truly

integrate data and algorithmic functionality into the human life-world. Our awareness of being-in-the-world (Heidegger, 1927) is often caused by the intensiveness of multi-sensory stimuli. The experience of walking through a cavern, feeling a fresh breeze that contrasts with the pure solid rock under the feet, hearing echoes of footsteps and water drops serves as a good example for this: All the simultaneous impressions make us aware of our body and its integration into the cavern. The lack of a single sense or only a misleading impression would change the holistic interpretation of the scene. In traditional computer-related work, however, many of our senses such as hearing, taste or smell are underused. Historically developed paradigms such as the prominent

DOI: 10.4018/jaci.2010040102

Graphical User Interface (GUI) are not able to fully embed the user into the information to be mediated. Possible explanations for their nevertheless widespread use should be searched more in their (historically developed) technical feasibility (Sutherland, 1963), rather than in usability and user-oriented simplicity.

For about the past ten years, though, there has been a shift towards multimodal and tangible representations of computer-based processes and abstract data, which try to close the gap between the users' reality and the abstract environment of data and algorithms. This takes us closer to data representations that benefit from the various aspects of the human's being-in-the-world by incorporating other modalities than vision and general-purpose pointing devices. However, a key prerequisite for an effective and ergonomic interface to digitally stored data is that the interface designer takes care of the common interplay between the human and his environment and integrates the resulting interface into this complex interrelationship.

We argue that haptic feedback, feature-rich control, and the use of many modalities are essential to sufficiently mediate complex information from computers to humans. Tools to achieve this are for example tangible interfaces and auditory displays. While tangible user interfaces (TUI) provide rich and at the same time direct control over digitally stored data (Brave, Ishii, & Dahley, 1998), sound and therefore Auditory Displays (AD) are widely recognised as very direct and flexible in their dynamic allocation of user attention and information conveyance (Bovermann, Hermann, & Ritter, 2006). Tangible auditory interfaces (TAI), a superset of both AD and TUI, has been introduced as paradigm by the authors (Bovermann 2010). They provide valuable guidelines for tangible auditory interface design. We believe that this combination can, after Rohrerhuber (Rohrerhuber, 2008), help to unfold the true potential of ergonomic user interfaces (Bovermann, Groten, de Campo, & Eckel, 2007). TAIs offer an information-rich interface that allows users to select, interpret

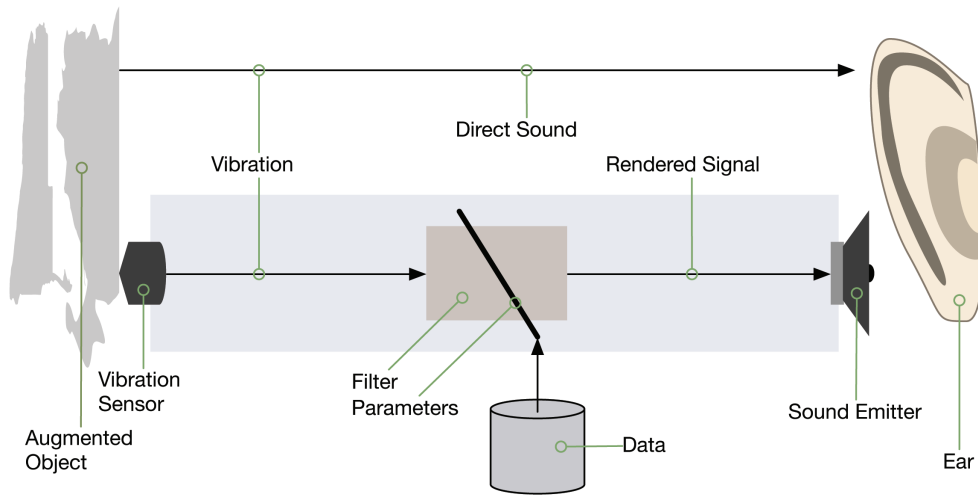
and manipulate presented data such that they particularly profit from their naturally excellent pattern recognition abilities.

One paradigm that evolved from the research in TAI is auditory augmentation. It draws on peoples' knowledge about everyday objects, whether they are simple like stones or more specialised and integrated into our daily work respectively into technology-driven systems as computer interfaces like for instance keyboards or computer mice. To add a data representation to such objects, rather than manipulating their intentional usage, we introduce auditory augmentation as a paradigm to vary the objects' sonic characteristics such that their original sonic response appears as augmented by an artificial sound that encodes information about external data. All this manipulation does not affect the sound's original purpose. The sonic reaction to an excitation of such an enhanced object then does not only reflect its physical structure, but also features the attached data. In other words, the structure-borne sound is artificially altered to render an additional information layer of data-inherent features.

We implemented an auditory augmentation system (Figure 1) called the *Reim* toolbox.¹ It features a lightweight and modular concept that is intended to help users in creating and manipulating custom data-driven auditory augmentations of objects they have ready at hand. *Reim* is currently available as a library for the SuperCollider language.

In the next sections, we will give a detailed overview of data as we understand it in relation to auditory augmentation followed by an overview of the related work and research fields. This is followed by a detailed introduction to the auditory augmentation paradigm and its implementation in the *Reim* toolbox. Various application scenarios are demonstrated with interaction examples, and first insights are reported from a qualitative user study in which we observed people in an unobtrusive data monitoring environment that incorporates an auditory augmentation setup.

Figure 1. General model of Reim-based auditory augmentations



2. DATA: THE NON-MATERIALISTIC MATERIAL

Due to their usage in digital environments, data (e.g., audio, video or text files) are widely viewed as a material such as wood or stone. This implies both a certain materialistic characteristic and a way to treat it that is based on our common experience with reality. This circumstance has its origin in our often subconscious understanding of data. Already the phrases data handling, data processing, or data mining implicate that data is widely recognised as a basic, materialistic resource. The used words originate in crafting or other physical work.

Data, though, is immaterial and disembodied. Its physical shape, the modality it is represented in, does by no means determine or affect its content; even more, data is pure content. There is for example absolutely no difference in a digital recording of Stravinsky's *Sacre du Printemps* whether it is represented as a series of magnetic forces on a rotating plate (i.e., a hard-drive), as states of electronic NAND-gates on computer chips (as it is the representation e.g., in computer memory), or as a series of high- and low-voltages in a copper-cable.

Neglecting this fact, data mining and data analysis, however, suggest its users to handle data as material. They process, analyse, and shape it like other work fields process, analyse and shape material like ore, stone, or wood. Nevertheless, the nature of data being a “non-materialistic material” has some inherent features, marking it different to material in the common sense. One of these features is that a data set is not bound to one phenotype. Its formal information content does not change depending on its actual representation: A change of modality does to no extent change the data itself. The subject matter of a book contains no other information than the same text represented as bits and bytes on a hard disk. A change of representation does, however, change the way people perceive a data set, since we derive our understanding of data from its actual representation. This circumstance makes it essential to look at the influence on the representation on the human perception and interpretation when dealing with data exploration and monitoring tasks.

Technically, however, data is independent of its representation type; nevertheless it has to be represented in some way. If this representa-

tion is well-suited for an algorithmic processing by computers, it is: most of the time – not in a form that supports human perception or structure recognition. The reason for this is not that the machine-oriented representation is too complex to understand. Moreover the pure physical representation (binary values coded as voltages in semiconductors or magnetic forces on hard drives) is completely inappropriate to be sensed or decoded by the human without appropriate tools.

3. RELATED RESEARCH FIELDS AND APPLICATIONS

3.1 Tangible Interfaces

The young research field of tangible interfaces (TI) picks up the concept of physically interfacing users and computers – a circumstance that was not present in the more traditional GUI-based designs (Ullmer & Ishii, 2000). To achieve this, the community around TI introduced physical objects to the virtual world of the digital, fully aware of all their interaction qualities, but also of their limitations caused by their embedding in the physical world. Tangible interfaces exploit real-world objects for the manipulation of digitally stored data, or – from a different point of view – enhance physical objects with data representations (either measured or rendered from artificial algorithms). This on first sight straightforward idea turned out to be a powerful approach to the conscious development of complex yet natural interfaces.

The used physical objects strongly affect the user experience of a tangible interface. Their inherent natural features of which users already have a prototypical concept are valuable for the designer and make it easy to develop interfaces that are naturally capable of collaborative and multi-handed usage (Fitzmaurice, Ishii, & Buxton, 1995). Even further, the usage of tangible objects implicitly incorporates a non-exclusive application such that the system designer does not have to explicitly implement it (Patten & Ishii, 2007).

3.2 Auditory Displays

Not only have research and perception of input technologies changed over the last century, but also the research in display technology has developed by discovering also non-visual modalities. The former focus on primarily visual displays has broadened to cover auditory (Kramer, 1994) and haptic cues (Brave & Dahley, 1997; Massie & Salisbury, 1994). Particularly auditory displays (AD) have seen a strong uplift, since they connect to our human's excellent abilities to perceive auditory structures even in noisy signals. Furthermore, in our auditory perception, we are sensitive to different patterns than those that are pronounced in visual display techniques. Sound rendering provide a way to display a reasonable amount of complexity. Therefore they are suitable to display high-dimensional data. The benefit of sound, compared to other non-visual modalities, is that it can be synthesized in a reasonable quality and spatial resolution.

The human perception of sound differs strongly from visual perception. Humans developed different structure detection and analysis techniques for sound stimuli than those that are used in the visual domain. For instance, timing aspects like rhythm, a spectral signal decomposition and the native support of time-based structures are unique to auditory perception. The combination of visual and auditory displays, however, makes it possible to get a more complete interpretation of the represented data. Thus, the provision of the same data by more than one modality makes it possible to extend the usage of human capabilities in order to reveal the data's structure. Auditory displays also natively support collaborative work (Hermann & Hunt, 2005), and allow for subconscious and ambient data representations (Hermann, Bovermann, Riedenklau, & Ritter, 2007; Kilander & Lonnqvist, 2002).

3.3 Tangible Auditory Interfaces

While both auditory display, as well as tangible interface research are highly promising as in-

dividual research fields, a combination of their techniques and experiences introduces valuable cross-links and synergies beneficial for both. We therefore propose the term tangible auditory interface (TAI) for systems that combine tangible interfaces with auditory displays to mediate information back and forth between abstract data space and user-perceivable reality (Bovermann 2010). The two parts form an integral system for the representation of abstract objects like data or algorithms as physical and graspable artefacts with inherent sonic feedback. The tangible part hereby provides the means for the manipulation of data, algorithms, or their parameterisation, whereas the auditory part serves as the primary medium to display data- and interaction-driven information to the user.

Key features of TAIs are their interfacing richness, directness, capabilities as a multi-person device for ambient augmentation, and their values in ergonomics. The latter is due to the fact that the interplay of sound and tangibility suggests a nature-inspired interface gestalt that can be directly derived from nature. In this regard, audio is a common affiliate to physical objects; most of them already make sound, e.g. when touched or knocked against each other. Furthermore, auditory displays profit from a direct control interface (Hermann & Hunt, 2005). Especially an auditory display that is designed for direct interaction with data profits from a close interaction loop between user and data representation as it can be provided easily by a tangible interface.

3.4 Reality-based Interaction

Reality-based Interaction (RBI) is a framework introduced by Jacob et al. that aims to unify emerging human computer interaction styles such as virtual, mixed and augmented reality, tangible interaction, ubiquitous and pervasive computing (Jacob, Girouard, Hirshfield, Horn, Shaer, Solovey, & Zigelbaum, 2008). Their key statement for unifying these approaches into one field is that all of them – intentionally or unintentionally – utilise at least one of the four principles of RBI that are *Naïve Physics*, *Body*

Awareness and Skills, *Environment Awareness and Skills*, respectively *Social Awareness and Skills*.

As the authors state, these principles – i.e., to base interaction techniques on pre-existing real-world knowledge and skills – can help to reduce the overall mental effort that is required to operate a system because users already possess the needed skills by their being-in-the-world. They claim that this reduction of mental effort may speed up learning, improves performance, and encourages improvisation and exploration, since users do not need to learn interface-specific skills. Designing data monitoring systems according to RBI therefore implies the use of multi-modality in both directions, to and from the user.

RBI forces to think both problem and user centred, rather than tool oriented. As an example, let us consider RBI's answer to the question of what is the typical reality-based approach to handle sounds. Natural sonic events are always connected to objects (re)acting with their environment. A loud bang, for example, always has a cause, be it an explosion or a slamming door. Auditory Displays on the other side grant digital information a physical voice. There is no natural pendant for them, apart from an internal physical model that is completely rendered in the virtual (like it is the case in Model-Based Sonification (Hermann & Ritter, 1999)). Here is where the benefit of RBI comes into play: To be human-understandable and therefore closely linked to RBI themes, not only the sonic outcome of a physical model should be perceivable by the user. Moreover, RBI claims that the overall performance of the system will increase when an interface is part of the user's direct environment, be it integrated either via VR, AR or any other related interfacing technology.

Another feature of RBI is the explicit utilisation of tradeoffs regarding the above-described principles in order to sharpen the designer's awareness in interface design. These tradeoffs are usually caused by the implementation of desired qualities of the system that cannot be implemented without automated algorithmic systems. They further state that

each tradeoff in an RBI-based system should be explicitly made. Tradeoffs, however, are not only optional for RBI-related system design, moreover they deserve a central place: An application that makes use of dynamic/algorithmic data processing (e.g., that has to use a computer) and is designed after the RBI framework has to have parts that result from these tradeoffs. Otherwise, the system could be built better –at least in terms of RBI – without the use of computers (i.e., exclusively in reality). The tradeoff in the design of auditory augmentations, for example, is caused by the need to control the system's sonic appearance by means of externally acquired (i.e., otherwise unconnected) data.

We integrated the tradeoff according to the guideline we derived from the RBI framework: Try to develop the desired application strictly according to the RBI principles, which especially means to avoid the mentioned tradeoffs. When desired features, such as the integration of additional, dynamically changing data, cannot be integrated without breaking these rules, the designer has to introduce tradeoffs. Each compromise has to be accompanied by an explicit discussion of reasons and possible benefits. This approach results in an application that can be located in the Venn diagram exemplified in

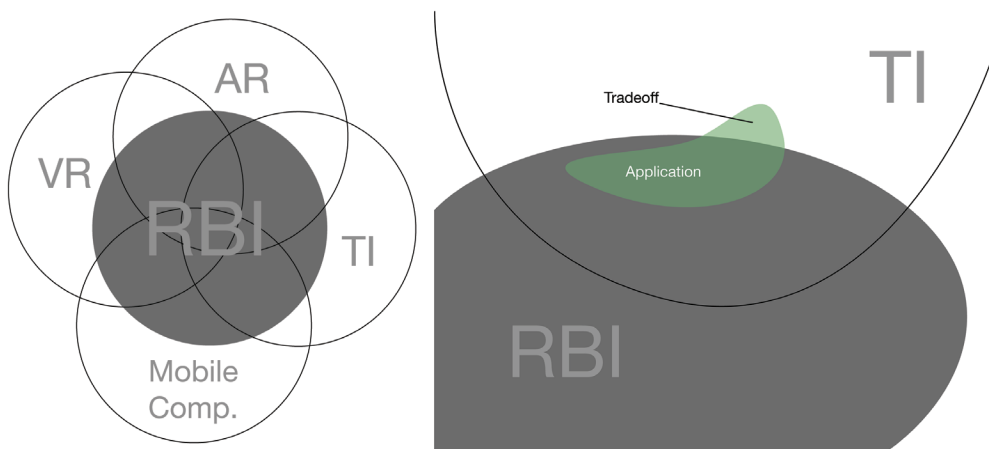
Figure 2. The following sections review several relevant auditory and tangible interfaces.

3.5 Audio-haptic Ball

The audio-haptic ball senses physical interactions such as accelerations and applied pressure, allowing to make use of these interactions as excitations of a Sonification Model (Hermann & Ritter, 1999) resulting in an auditory and dynamic data representation (Hermann, Krause, & Ritter, 2002). By this, the user can experience the model-based sonification as plausible result to interactions such as shaking, rotating or squeezing the ball. Since the auditory output directly corresponds to the users' interaction with the ball, mediated via the sonification model, interaction can be used to explore and interpret data structures. The formal software development process for the audio-haptic ball interface used for Model-Based Sonification can be described as

1. designing a dynamic model, which often borrows from physical principles,
2. parameterizing the model with given data,
3. interacting with the ball (i.e. shake it, etc.),
4. sound is continuously rendered according to the dynamic model.

Figure 2. Venn diagram of RBI and its related research areas (left) and the (hypothetical) location of an RBI-based application



This approach especially requires the re-implementation of basic natural functionality, namely the dynamics of objects in a 3D space. Although this approach makes it literally possible to shake and squeeze data sets of higher dimensionality, it remains difficult to explain and understand what happens in such a space, and how the modelled n-dimensional object can be embedded into 3D reality so that it can be excited with the audio-haptic ball.

3.6 Pebblebox

The *Pebblebox* is another audio-haptic interface for the control of a granular synthesiser which extracts information like onset, amplitude or duration of grain-like sounds captured from physically interacting pebbles in a box (O'Modhrain & Essl, 2004). These high-level features derived from the colliding stones are used to trigger granular sounds of e.g., water drops or wood cracking to simulate rain or fire sounds. The performance of the *Pebblebox* massively relies on the fact that the captured signal has to be a superposition of transient sound events. A change of the sound source such as it is implemented in the Scrubber, another closely related interface also developed by the authors of the *Pebblebox* (Essl & O'Modhrain, 2004), has to extract a completely different feature set from the input signal. It is designed in assuming incoming scrubbing sounds in order to synthesise artificial scrubbing sounds.

Auditory augmentation, however, does not rely on such assumptions: it directly uses the object's sound as the input signal of an audio filter, which is parameterized by given data. The resulting sound is then directly played back to the user. The idea to involve data of the users' interest into the sound filtering process is essential for our approach to auditory augmentation.

4. AUDITORY AUGMENTATION AND THE REIM TOOLBOX

One of the human's natural qualifications is his ability to literally get a grip of almost every physical object easily. Technically speaking, a

human is able to understand the basic features and often also the inner structure of an object by physically exploring it with his various senses and actuators (i.e., ears, nose, skin and eyes, and arms, hands, legs, fingers, etc.). We propose that dealing with data should be as easy as discovering e.g., the current fill-level of a box with sweets. We propose this both for everyday scenarios involving information such as temperature, humidity, stock exchange quotation, etc., but also for technology-oriented measurements like CPU load or network load. Taking this attempt literally motivates a more direct representation of data than it is state of the art. The augmentation of action feedback on everyday objects with appropriate data representations.

The paradigm of auditory augmentation is aimed to help interface designers to represent digitally stored data as auditory features of physical objects. It can be formally described as the process of artificially inducing auditory perceivable characteristics to existing physical objects. The structure-borne sound gestalt hereby is altered according to externally acquired data. However, this process does not change the natural interaction sound's presence or timing. An auditory augmentation system can be used to alter the sonic characteristics of arbitrary objects. Each object can therefore provide a different impression of the data, unveiling a different set of possible structural information of the represented data. Note that, although powerful and built for non-linear analysis and exploration, this paradigm is neither intended nor appropriate to systematically search for specific structure in data, or even to observe exact class labels for a data set. Moreover, it shifts the task of observing structures in possibly unknown data into a naturally perceivable form, where the human ability to find and understand structural information can be utilised.

As shown in Figure 1, an auditory augmentation system consists of the following parts: An audio-transducer (Vibration Sensor) captures structure-borne vibrations of arbitrary objects, which are fed into a parameterised audio filter (Filter). Its parameters are controlled according

to externally acquired data such as the temperature or stock exchange quotations (Data). The filtered signal then is transformed into an audible sound (Sound Emitter), being a superposition of the originating vibration and the data under investigation. The resulting augmentation has negligible latency, and smoothly overlays with the original sound (Direct Sound). The overall auditory character of the complete setup depends on the input's audio characteristic, the filter, the data state, and the sound rendering including possible distortion by the loudspeaker. Note that the resulting sound mixes with the real sound of the interaction.

We introduce the Reim toolbox as an implementation for the auditory augmentation paradigm. Its lightweight and modular concept intends to help people familiar with a basic sound synthesis knowledge in the creation and customisation of such data-driven object augmentations. Systems, build according to Reim, draw on peoples' knowledge about every-day objects, whether they are as simple as pebbles, or more specialised and integrated into daily, technology-driven systems like keyboards or other computer interfaces.

4.1 Usage Scenarios

To show the potential of auditory augmentation as a tool for data exploration and monitoring, this section presents examples on how an everyday usage of such a setup might look like. It especially focuses on an ergonomic interaction design, drawing from familiar manipulation skills.

Let us consider two data sets that share the same characteristics in distribution and local density. There are no obvious differences in their structure. A user wants to investigate if there are other, possibly non-linear structural differences between the data sets. By linking each data set to a Reim augmentation, he investigates into this direction. Around him, the user collected surfaces of various characteristics: one of granite, one made of wooden, etc. He attaches the transducers of the Reim system to small glass objects and scrubs them over the

surfaces. Each combination of surface, glass object/data set and scrubbing technique results in a characteristic sound. Exploring these combinations for differences between the sounds of each object enables the user to find structural differences between the data sets. When he found interesting reactions, he captures and analyses the source vibrations (i.e., the sounds that appear when scrubbing the objects on the surfaces without the data-inherited overlay) for further analysis, because these sounds offer information on the non-linear structures in the data sets under exploration. It can be seen as a classifying discriminant. Instead of using only rigid bodies, it is also possible to attach the transducers to drinking glasses filled with grainy material of different sizes and shapes. The user then sequentially loads the data sets to the glass/tool aggregates and shakes them. This way he can test which of the glasses emit a characteristic sound augmentation that can be used to differentiate between the data sets. Both scenarios become more powerful by Reims feature to record and playback input sounds with different data sets. Also the feature to change the synthesis process as well as the range of the parameter mapping increases the flexibility of the system.

In another scenario, dealing with unobtrusive data monitoring, a person wants to keep track of a slowly changing data stream such as the weather situation around his working place. In order to acquire this information without being disturbed by a constantly sounding auditory display, or having to actively observe e.g. a webpage, he acquires the data automatically from weather sensors and feeds them to his auditory augmentation setup. After this, he attaches the connected transducer to a computer input interface that he is using regularly (e.g., the keyboard, or the mouse), resulting in an auditory augmentation of the artefact's structure-borne sound with the weather data. Every time the attached sensor values change, the auditory character of the augmented device changes, giving the user a hint on the current weather conditions.

Adding auditory augmentation to structure-borne sounds means to insert a thin layer between people's action and an object's auditory re-action. The proposed auditory augmentation can be easily overlaid to existing sounds, and does not change prominent auditory features of the augmented objects like the sound's timing or its volume. In a peripheral monitoring situation, the data gets out of the way for the user if he is not actively concentrating on it. A characteristic change, however, tends to grab the user's attention.

4.2 Level of Abstraction

Reim supports two different abstraction levels: The first level incorporates mostly direct and physical manipulation with direct sonic feedback, whereas the second abstracts from these natural manipulation patterns.

In the first, the user's experience of an augmented object does not differ from handling non-augmented objects, apart from the fact that the object-emitted sounds are also data-driven. Due to his being-in-the-world, the user feels familiar with the objects manipulation feedback. He gets a feel for the process by gaining experience of the data-material compound's reaction over time. Non-linear complexity of material properties and their reactions to e.g., pressure and speed of action can be used intuitively, i.e., without additional cognitive effort. Data easily becomes integrated into everyday life.

The second level allows gaining assessment and increasing repeatability in the explorative process of Reim. It enables the user to capture the vibration of a physical excitation that then can be used to either repeat the data-representation process with the exact same prerequisites or to sonify other data items with it. This demand requires to capture the transducer's input and use it for the representation of several data sets as well as the addition of recording capabilities to the system such that the data's representation can be easily captured and replayed to others. Related to this are the offering of pre-recorded standard excitation sources, or the provision of a standard set of objects to add data-driven

auditory augmentations. This abstraction, or, in terms of RBI, tradeoff allows to programmatically explore and compare data, while still utilizing the sound characteristics of the augmented object.

4.3 Implementation

According to the general model of auditory augmentations (cf., Figure 1), a setup of such a system requires the following hardware: a vibration sensor capable of audio signals (e.g., a dynamic microphone like the AKG C411, or a piezo-based pickup system like the Shadow SH SB1), a computer with an audio interface to capture the sensed signal and to apply the filter model to the signal, and a sound emitter (i.e., either loudspeakers or headphones) for signal playback.

We implemented the Reim toolbox to help with the administration of the data as well as with the filter design. The toolbox makes it easy to apply data based parameters to signal filter chains and to implement, collect, store, and share presets for the synthesis process. Both data processing and sound rendering are realised in the *SuperCollider* language (McCartney, 2002), and are available for free upon request.

5. APPLICATIONS

Auditory augmentation can be used in various usage scenarios. This section describes systems utilising the Reim toolbox for the two, in terms of their usage very different, scenarios of data exploration and unobtrusive monitoring that we described above. All introduced applications are demonstrated in videos on the corresponding website.²

5.1 Exploration

*Schüttelreim*³ is an approach to implement the mentioned use case of active data exploration and comparison. In this setup, the transducers are statically attached to box-shaped objects, which should contain a grainy material such as several buttons or marbles. As shown in

the video example, shaking the box results in an audible reaction that reflects the physical structure as well as the data-inherent parameters. This is realised with the attached transducer that captures the rattling of the box' content and feeds it into a filter. Loudspeakers near the exploration area then play back the augmentation in real-time. When the data attached to the *Schüttelreim* object is substituted by another one, this substantially changes the resulting sound depending on the variation in the attached data item. Since people are trained to listen to manipulation-caused sounds, able to precisely control their handling, *Schüttelreim* allows to turn data into highly controllable sonic data objects. We claim that, by extensive use, people will learn to shake and manipulate the boxes in such ways that they can perceive certain aspects of the data, which possibly leads to a valid differentiation and classification of the structural information of the attached data.

A different example application incorporating auditory augmentation is *Paarreim*. In contrast to *Schüttelreim*, *Paarreim*'s interaction design is not based on the manipulation of self-contained sounding objects. Furthermore, it is focused on the physical interaction between objects and surfaces. It features several independent objects, each attached to one data set. These rigid objects with little natural resonance can be scrubbed over various surfaces that are made of different materials, each with a characteristic haptic texture. It results in substantially different excitations of the data depending on the interplay between their gestalt and the texture, which in turn change the sound of the

auditory augmentation. The user gets detailed insights into the data structures and can learn to use specific material combinations that help him classify data into groups according to their sonic reaction. Having more than one object at hand allows for a comparison of the sounds, and therefore the data items. The actual auditory augmentation is realised by loudspeakers near the exploration area, which play back the sound synthesis. The setup of such a system is shown in Figure 3 and in the corresponding video on the website.

5.2 Unobtrusive Monitoring

Object manipulations result in structure-borne sounds that inherently transport information about the incorporated objects and the accompanying physical reaction. It is packed in a very dense form, yet it is easy to understand.

*Wetterreim*⁴ utilises this feature for a dedicated scenario: the day-to-day work on a computer as it is common at almost any office workplace. As the source for the auditory augmentation, we chose the keyboard, one of the main interfaces for the daily work with computers. Typing on it results in a characteristic sound that is shaped by the design of the keyboard and its interplay with the writer's fingers. A contact microphone captured the keyboard's structure-borne sound, on which we based a Sonification of weather-indicating measurements. When filtering the captured sound by data-driven filter parameters, an audio stream is created, which is close to the characteristics of the original but additionally features charac-

Figure 3. A *Paarreim* exploration session



teristics of the integrated data. The filter output is superimposed to the original sound such that it is perceived as one coherent auditory gestalt. The developed filter parameterisation for the weather data allows people to perceive a drop in pressure or an approaching cloud front as a change in the object's auditory characteristic. An example for the use of *Wetterreim* is given in the corresponding video on the website.

6. WETTERREIM CASE STUDY

To gather feedback on the implemented auditory augmentation system, we conducted a qualitative user study. We asked three people to integrate *Wetterreim* into their day-to-day work for a period of four or more days. After this period, we collected their statements in an unstructured interview. During the setup, the audio transducer was attached to the participant's commonly used keyboard (as shown in Figure 4). Its signal was fed into an external computer that was exclusively used for data acquisition and sound rendering. The data that were augmented to the participant's keyboard were acquired from the nearest publicly available weather station. Its

update rate varied between every half an hour and every hour. We used the filter setup shown in Figure 5. The weather conditions during the study are shown in Table 1. In an initial setup session, filter ranges were adapted for each participant in order to reflect their individual preferences and the sonic character of their keyboard. Overall, our observations based on the unstructured interviews unveiled the following aspects:

Sound design

Participant 2 found the used ringing sound to be natural and pleasant. However, Participant 1 reported that the augmented sound irritated her in the beginning. Participant 1, Participant 2 and Participant 3 stated that they missed the sound when it was absent by accident.

Localization

Participant 2 found it astounding that the sound seemed to originate from the keyboard although the loudspeaker was at a completely different position.

Figure 4. The hardware setup used by Participant 1. The transducer was attached to the external video adapter of her laptop. This made it easy for (dis-)assembly, since she only used Wetterreim at her workplace, but carried her laptop with her



Figure 5. Schematic of the sound synthesis used in the case study

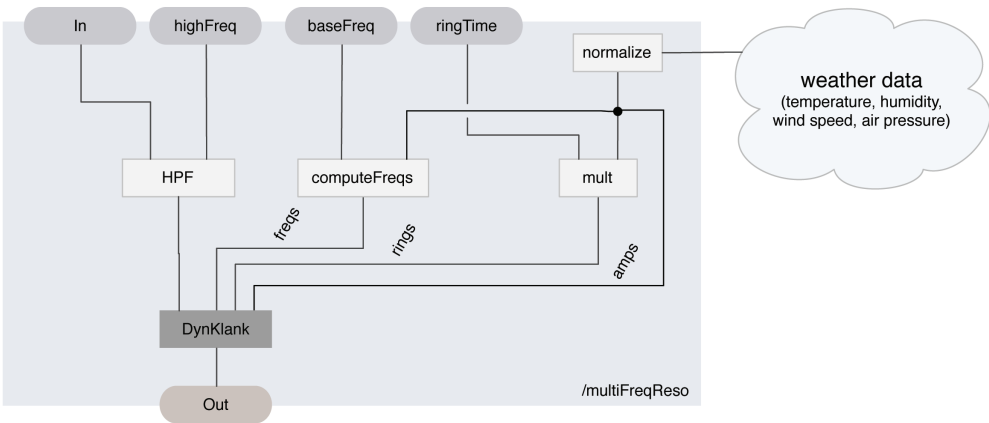


Table 1. The weather conditions for each participant during the Wetterreim study.

User	# of Days	Weather Conditions
Participant 1	4 days	Contrary weather, changes between 35°C, sunny and 20°C with thunderstorm and sometimes heavy rain in the evening.
Participant 2	10 days	Constant over the time, no rain, around 20°C.
Participant 3	8 days	20°C – 25°C, rainy and sunny.

Data-to-sound mapping

The differences in the rendered sound according to the data were considered by Participant 1 and Participant 2 to be reasonably distinguishable, even without direct comparison.

Exploration

All participants reported that they also used the setup playfully; Participant 2 and Participant 3 stated to actively trigger it by purpose to hear the system's actual state.

Attention

Regarding the subconsciousness of the sounds, participants reported mixed feelings. While Participant 1 found it difficult to shift her attention away from the sound, Participant 3 stated that a change in feedback was rising his attention

even when he was concentrating on something different. However, no participant mentioned the system to be bothersome.

Sound Level

The adjustment of the augmentation's volume was experienced by all users to be difficult. Especially Participant 1 reported to usually type relatively weak, making it difficult to properly adjust the amplitude of the augmentation.

In general, the application – unobtrusive monitoring of near real-time data – worked out for the participants. We especially found out that users perceived the auditory augmentation and the original sound as a single natural sound, they were not bothered by the Sonification, and they had difficulties adjusting the volume of the auditory augmentation. For a future setup, we plan to investigate into this issue.

7. CONCLUSION

In this article, we introduced auditory augmentation as a paradigm to represent data as an artificially induced overlay to the common structure-borne sounds of an arbitrary object. With Reim, we presented a toolbox for the design and implementation of such tangible auditory interfaces. It utilises everyday objects and their interrelations to transform abstract data into physically manipulable and auditorily perceivable artefacts. The toolbox has been demonstrated at hand of several design studies featuring different usage scenarios including active data exploration and subconscious monitoring situations. During the setup of the different applications, we experienced that latency plays a prominent role in Reim-based applications. Long delays (more than 20ms) between user action and system reaction broke the illusion of sonic identification and compactness of the object and its augmentation. However, small spatial separations between the structure-borne sound (i.e. transducer location) and the augmentation source (i.e. the loudspeaker) did not affect that illusion.

Because of Reim's simple technical assembly, the participants in the qualitative user study were able to understand the setup without any problems. Additionally, it turned out that the Reim system is well applicable for a long-term case study. During such a study on *Wetterreim*, subjects used an auditory augmentation of weather data in their usual working environment. Local measurements of weather-related data have been augmented to the structure-borne sounds of their computer keyboard. Participants reported that the augmentation worked well, though it turned out that the particular data domain was not of much use. However, the augmentation was perceived as part of the augmented object, a fact that indicates that auditory augmentations can well merge into the everyday soundscape. Participants were also able to differentiate between several weather situations.

Many participants stated that they were not able to separate source sounds from data-driven sounds. Although this is an essential effect regarding the acceptance of the system, it uncovers an inherent issue of Reim-based applications: the sound of the data object combination is perceived as an entity; users are not able to split it into its components to separate the data communicating part from the structure-borne sound. Long-term usage of a Reim-based system, though, should overcome this effect. People will adapt to the auditory specifics of the used objects and develop implicit knowledge on how to separate the physically induced sounds from the data-dependent sounds. This effect is supported by the fact that the physical part of the sound bases in a static set of parameters, reflecting the same object characteristics in all excitations. Changes in the sound therefore always originate in a change of the data-driven augmentation. These observations and considerations suggest that auditory augmentation is a promising approach for tangible auditory interfaces, both for data exploration and subconscious monitoring.

ACKNOWLEDGEMENTS

This work was partly funded by the CRC673-Alignment in Communication and the Excellence Initiative of the German Research Foundation.

REFERENCES

- Bovermann, T. (2010). *Tangible Auditory Interfaces: Combining Auditory Displays and Tangible Interfaces*. PhD thesis, Faculty of Technology, Bielefeld University, Germany.
- Bovermann, T., Groten, J., de Campo, A., & Eckel, G. (2007). *Juggling Sounds*. In *Proceedings of the 2nd International Workshop on Interactive Sonification*, York, UK.

- Bovermann, T., Hermann, T., & Ritter, H. (2006). Tangible Data Scanning Sonification Model. In Proceedings of the International Conference on Auditory Display (ICAD 2006), London, UK (pp. 77-82).
- Brave, S., & Dahley, A. (1997). inTouch: a medium for haptic interpersonal communication. In Proceedings of the Conference on Human Factors in Computing Systems (pp. 363-364).
- Brave, S., Ishii, H., & Dahley, A. (1998). Tangible interfaces for remote collaboration and communication. In Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (pp. 169-178).
- Essl, G., & O'Modhrain, S. (2004). Scrubber: an interface for friction-induced sounds. In Proceedings of the 2005 Conference on New Interfaces for Musical Expression (NIME '05), Singapore, Singapore (pp. 70-75).
- Fitzmaurice, G. W., Ishii, H., & Buxton, W. (1995). Bricks: Laying the Foundations for Graspable User Interfaces. In Proceedings of CHI 1995 (pp. 442-449).
- Goldhaber, M. H. (1997). The Attention and the Net. *First Monday*, 2(4).
- Goldhaber, M. H. (2006). How (Not) to Study the Attention Economy: A Review of The Economics of Attention: Style and Substance in the Age of Information. *First Monday*, 11(11).
- Heidegger, M. (1927). *Sein und Zeit*. Halle A. D. S: Niemeyer.
- Hermann, H., & Hunt, A. (Eds.). (2005). *IEEE Multimedia, Special Issue Interactive Sonification*. Washington, DC: IEEE.
- Hermann, T., Bovermann, T., Riedenklau, E., & Ritter, H. (2007). Tangible Computing for Interactive Sonification of Multivariate Data. In Proceedings of the 2nd Interactive Sonification Workshop.
- Hermann, T., Krause, J., & Ritter, H. (2002). Real-Time Control of Sonification Models with an Audio-Haptic Interface. In Proceedings of the International Conference on Auditory Display 2002 (pp. 82-86).
- Hermann, T., & Ritter, H. (1999). Listen to your Data: Model-Based Sonification for Data Analysis. In Proceedings of the Advances in Intelligent Computing and Multimedia Systems, Baden-Baden, Germany (pp. 189-194).
- Jacob, R. J. K., Girouard, A., Hirshfield, L. M., Horn, M. S., Shaer, O., Solovey, E. T., & Zigelbaum, J. (2008). Reality-based interaction: a framework for post-WIMP interfaces.
- Kilander, F., & Lönnqvist, P. (2002). A Whisper in the Woods: An Ambient Soundscape for Peripheral Awareness of Remote Processes. In Proceedings of the International Conference on Auditory Display 2002.
- Kramer, G. (Ed.). (1994). *Auditory Display*. Reading, MA: Addison-Wesley.
- Massie, T. H., & Salisbury, J. K. (1994). The PHANTOM Haptic Interface: A Device for Probing Virtual Objects. In Proceedings of the ASME Winter Annual Meeting, Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems.
- McCartney, J. (2002). Rethinking the computer music language: SuperCollider. *Computer Music Journal*, 26(4), 61-68. doi:10.1162/014892602320991383
- O'Modhrain, S., & Essl, G. (2004). PebbleBox and CrumbleBag: tactile interfaces for granular synthesis. In Proceedings of the 2004 Conference on New Interfaces for Musical Expression (NIME '04), Singapore, Singapore (pp. 74-79).
- Patten, J., & Ishii, H. (2007). Mechanical constraints as computational constraints in tabletop tangible interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 809-818).
- Rohrhuber, J. (2008). Implications of Unfolding. In *Paradoxes of Interactivity* (pp. 175-189).
- Sutherland, I. E. (1963). Sketchpad, a man-machine graphical communication system. Unpublished doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Ullmer, B., & Ishii, H. (2000). Emerging Frameworks For Tangible User Interfaces. *IBM Systems Journal*, 39(3-4), 915-931. doi:10.1147/sj.393.0915

ENDNOTES

- ¹ The name of the implemented system is motivated by a german saying *sich einen Reim machen auf*, which can be translated best as put two and two together.
- ² Auditory Augmentation Demonstration Media: <http://www.techfak.uni-bielefeld.de/ags/ami/publications/BTH2010-AA/>
- ³ *Schütteln* is German for *to shake*.
- ⁴ *Wetter* is German for *weather*.

Till Bovermann is a research associate at the Ambient Intelligence Group at the Cognitive Interaction Technology Center of Excellence at Bielefeld University (CITEC). He is also involved in the C5 project Alignment in AR-based cooperation of the CRC673-Alignment in Communication. Previously, he worked as a research assistant at the Neuroinformatics Group at Bielefeld University. He received his german diploma in Information Technology and the Natural Science with a focus on robotics in 2004. His current research interests are the integration of auditory displays and tangible interfaces to form an integral system for data emersion into the human life world. His arts-related interests are in media arts, especially interactive performances and just-in-time programming of musical and visual structures. He is a co-founder of Too Many Gadgets, a live-coding group that attempts to capture the relationship of space, sound and vision.

René Tünnermann is a research associate at the Ambient Intelligence Group at the Cognitive Interaction Technology Center of Excellence at Bielefeld University (CITEC). He studied science informatics at Bielefeld University. During his studies he worked as a student worker at the Neuroinformatics Group of Bielefeld University and the Alignment in AR-based cooperation project of the CRC673-Alignment in Communication. His research focus lies with tangible computing and interactive surfaces.

Thomas Hermann studied physics at Bielefeld University. From 1998 to 2001 he was a member of the interdisciplinary Graduate Program "Task-oriented Communication". He started the research on sonification and auditory display in the Neuroinformatics Group and received a Ph.D. in Computer Science in 2002 from Bielefeld University (thesis: Sonification for Exploratory Data Analysis). After research stays at the Bell Labs (NJ, USA, 2000) and GIST (Glasgow University, UK, 2004), he is currently assistant professor and head of the Ambient Intelligence Group within CITEC, the Center of Excellence in Cognitive Interaction Technology, Bielefeld University. His research focus is sonification, datamining, human-computer interaction and cognitive interaction technology.