AUDIODB

GET IN TOUCH WITH SOUNDS

Till Bovermann, Thomas Hermann, Helge Ritter

Bielefeld University
Neuroinformatics Group
Bielefeld, Germany
`[tboverma|thermann|helge]@techfak.uni-bielefeld.de`

ABSTRACT

In this paper we present *AudioDB*, a system to collaboratively navigate sound databases via a spatial audio-haptic setup. It provides an environment to sonically sort, group and select sounds which are represented as physical artifacts on a tabletop surface. We give an introduction and insights on implementing interactive overviews for sound databases followed by first impressions of a qualitative analysis of the system.

Figure 1: AudioDB can be used to discuss sounds and their interrelations.

1. INTRODUCTION

As audio researchers, digital audio is ubiquitous in our everyday life. Searching and sorting of sounds collected in extensive databases (e.g. sampling libraries for musical production or seismographical surveys) is difficult and restricted by the standard keyboard/mouse computing paradigm. Also the common technique of tagging sounds and other media files[1] has the drawback that it needs descriptive words, which is often unsuitable for sounds. We therefore created *AudioDB*, an intuitive human computer interface to interactively explore sound by representing them as physical artifacts (*grains*) on a tabletop surface. The system can be used for sonic sorting, grouping and selecting of sounds represented as physical artifacts, and can therefore serve as a basis for collaborative work and to assist the discussion on audio-related tasks. *AudioDB*, however, is not only a solution for problems in a dedicated field of work, but is also designed as an easy-to-use *multi purpose tool* for audio-based

---

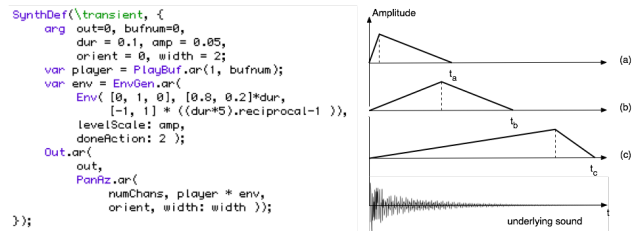[1] as it is common in web2.0-applications such as http://flickr.com

Figure 2: Left: The synthesis definition implementing one sound grain. The underlying sound here is stored in a buffer. Right: Randomly chosen length of the rendered audio grain determines the attack and therefore proportion of the transiental part of the underlying sound.

information. As a side-effect, *AudioDB* can be used for fundamental research on how humans handle digital information that is projected onto physical artifacts.

2. SYSTEM

As a basis for AudioDB we use the tDesk [1], a tabletop tangible computing environment designed and built in our interaction laboratory. By design, the dimensions of the surface allow groups of people to work on tangible applications, providing each member direct access to the physical objects. We capture the position of the objects with a digital camera from below. This prevents possible visual object occlusions by the users. A blob tracking algorithm developed by the authors detects number and position of the used grain objects on the surface and applies a unique ID according to [2].

For the sound display the table is surrounded by an eight-channel audio system arranged in a ring of equidistant loudspeakers. This allows for a natural auditory interface that is directly coupled to the interaction on the tabletop. Every recognized grain object is then linked to a spatially corresponding sound source. If and only if a grain is moved, a part of its sound-file is made audible. In AudioDB the feedback of information about moved grains is realized by granular re-synthesis based on the corresponding sound. Unlike other systems designed to summarize multimedia files (e.g. [3]), AudioDB features highly interactive access to sounds. This fact does not allow us to use standard sound summarization techniques, since their summarization algorithms need to know the exact summary duration beforehand which cannot be given in highly interactive systems as *AudioDB*. Our auditory dis-

play is therefore based on granular re-synthesis. Each rendered audio grain is chosen to be in between a short part of the sound's onset multiplied by a curve with a sharp attack, or a longer part multiplied with a smooth envelope. Transiental respectively decaying parts in the granular sound stream are uniformly distributed. Information on attack and decay of the underlying sound therefore is kept in the resulting steady sound stream. To closely link the auditive display to its corresponding physical object, we render the sound to originate from the same direction as the object is located on the tDesk. For the complete audio rendering we used SuperCollider, a higher-level programming language that is specially suited for realtime sound rendering [4]. The exact Synthesizer definition is shown on the left side of Figure 2.

3. FIRST INSIGHTS

We are currently evaluating *AudioDB* using qualitative user studies, results of which might be already at hand at IDCAD 2008. We provided eight users access to a running *AudioDB* session—one at a time—and gave them the simple task of "arranging the sounds in a meaningful way". We recorded their interaction with a camcorder and, in a subsequent interview, walked them through the video, asking about their thought process while performing specific actions.[2]

One participant first explored the sounds by adding and removing the grains from the active surface of the table. After a while, he arranged all grains in a big circle, while positioning the grain under exploration in its center (Figure 3c).

Another participant first explored the given sounds one by one, and arranged them into on-the-fly-established groups on the table. When settled, she changed her goal, and tried to arrange the sounds of each subgroup into linear sequences (Figure 3a, b).

All eight participants were able to connect the objects with their corresponding sounds without efforts. They were able to bring the objects into an order and could afterwards explain the meaning of the different object groups. We also discovered that our particular setup works best with about 17 or fewer objects; with larger numbers "examined" and "to-be-examined" object groups cannot anymore be differentiated easily.

4. CONCLUSION

In this paper we presented *AudioDB* as a system for interactive and collaborative exploration of sound databases. Though the first insights look promising relating the systems ability to serve as a multi purpose tool for audio-based interactive exploration, additional work has is needed as it is currently not possible to operate on sound databases consisting of more sounds then there are objects on the surface. We, however, plan to build in a mechanism to come over this limitation which nicely integrates into the workflow of the current system.
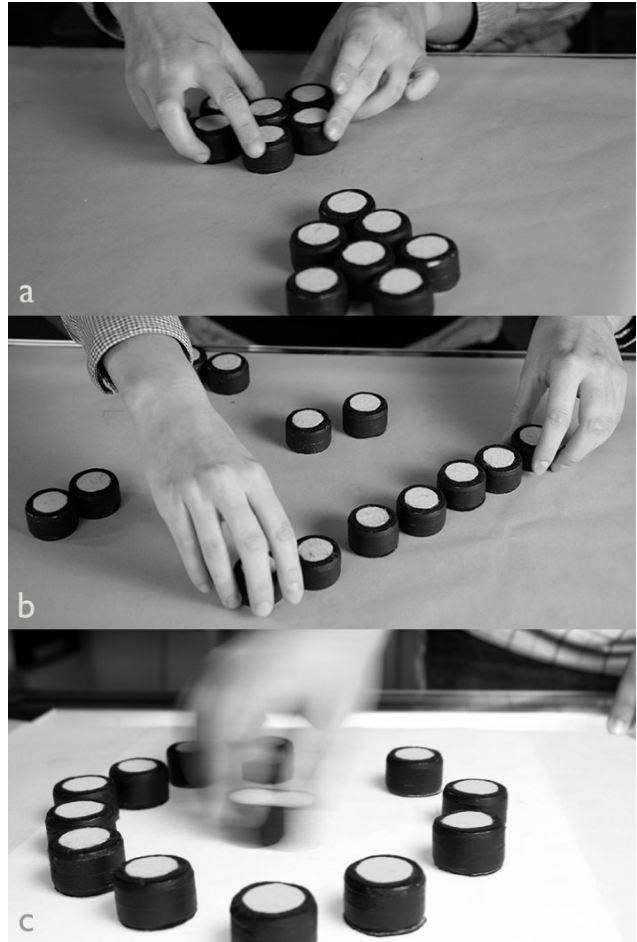

Figure 3: Ordering results of AudioDB-grains in a single user environment: (a) clustering, (b) sequencing, (c) sorting objects in a circle.

5. REFERENCES

[1] T. Bovermann, T. Hermann, and H. Ritter, "A tangible environment for ambient data representation," in *First International Workshop on Haptic and Audio Interaction Design*. multivis, Aug 2006, vol. 2, pp. 26–30.

[2] I. J. Cox and S. L. Hingorani, "An efficient implementation of reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 2, pp. 138–150, 1996.

[3] J. Laaksonen, M. Koskela, M. Sjöberg, V. Viitaniemi, and H. Muurinen, "Video summarization with soms," in *Proceedings of the 6th International Workshop on Self Organizing Maps (WSOM 2007)*, Bielefeld, Germany, 2007, Neuroinformatics Group, Bielefeld University, Germany.

[4] J. McCartney, "Supercollider hub," URL, July 2004, http://supercollider.sourceforge.net/.

---

[2]A video showing *AudioDB* in action can be viewed at *http://seto.LFSaw.de/audioDB.shtml*