Using Sonification and Haptics to Represent Overlapping Spatial Objects: Effects on Accuracy

Junlei Yu, Kris Lohmann, and Christopher Habel

Department of Informatics, University of Hamburg, Germany
{jyu,lohmann,habel}@informatik.uni-hamburg.de

Abstract. For blind and visually impaired people, the access to spatial information is crucial. Therefore, the development of non-visual interfaces to spatial representations, e.g. maps and floor plans, are important tasks. In earlier empirical work [19], we investigated virtual haptic floor plans, accessible through a Phantom force feedback device, which allows haptic exploration of virtual objects (walls), in combination with sonification for representing overlapping objects (windows). In the following we present an empirical study on multiple-overlap constellations (in the room-plan scenario: walls, windows and radiators). We reduced the complexity of the environments from complete apartments to only one wall and overlapping subsections, to focus on the spatial accuracy of acquired knowledge. This one-wall experiment has two goals: to compare the accuracy of size and position estimation with the experiment with complete apartments and to investigate the usability of sonification to represent overlapping entities at walls. Qualitative measures on the correctness of overlap-relationship identification and quantitative investigation on the accuracy of size and position estimations are discussed. The results can be extended to the use of sonification to represent overlapping entities in general.

Keywords: Spatial Knowledge Acquisition, Virtual Haptics, Sonification, Representational Multimodality.

1 Haptic-Audio Representations of Indoor Environments

The use of spatial knowledge is ubiquitous in our daily life. But, for blind and visually impaired people, the access to spatial information is limited to non-visual perception. Therefore, in recent years, modern human computer interfaces that rely on other perceptual channels to interact with users have been implemented. These are intended to help blind and visually impaired people to overcome problems in acquiring spatial knowledge caused by the absence of vision. For example, virtual haptic floor plans can provide spatial knowledge of indoor environment [9, 17, 19]. In a prior study with blindfolded participants, which explored palm-structured apartment using a Phantom Omni force feedback device (http://www.sensable.com), we have shown that humans—even when the sense of vision is absent—acquire spatial knowledge with very good performance [19]. After exploring the virtual haptic floor plan—using mostly a wall following strategy—their knowledge about the explored apartment was tested with respect to layout topology and size estimations. A second

C. Stephanidis and M. Antona (Eds.): UAHCI/HCII 2013, Part I, LNCS 8009, pp. 602–611, 2013.

group of participants explored the same apartment layouts containing supplementary information about the size and the position of windows provided by sonification. The participants of this study showed good knowledge with respect to the position of windows. Besides, we discovered that when the sonification can be invoked during wall following, higher quality of shape knowledge is achieved in some cases, since beyond space perception common sense knowledge about apartments can be exploited.

However, the results also showed that the accuracy on the size estimation of the windows was not satisfying [19]. Analyzing the exploration behavior of the participants as well as knowledge they acquired, there seems to be primary candidates for bringing about these inaccuracies, (1) they are caused by perceptual limits of using sonification to re-present windows in virtual haptic floor plans, and, (2) they are effects of intentional and attentional patterns of behavior, namely that users intending to acquire a floor plan may focus their attention mainly on the layout of rooms and the apartment, and only with minor priority on the windows.

To solve this problem, we foreground in the present paper the interplay of perceptual and conceptual processing of overlap constellations. In particular, to inspect how precise the size estimation of linear spatial entities can be achieved by using a sonified haptic interface, we conducted an experiment considering explorations of individual walls with varieties of overlap constellations (Fig. 1). By this experimental design, the cause candidate (2) is applicable. The investigated wall-window configurations were the same as that in the virtual haptic apartments. The accuracy of both size and position estimation of the windows were compared.

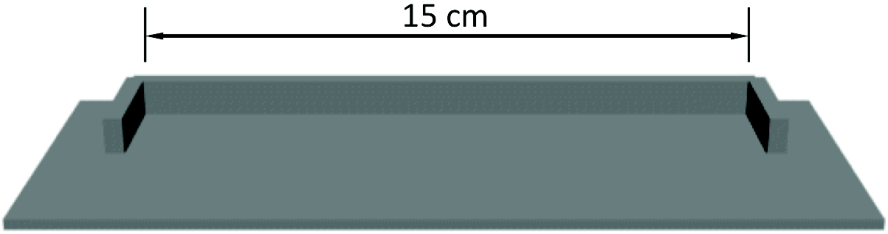


Fig. 1. The one-wall virtual haptic model employed in the empirical experiment

On the other hand, wall following with concurrent ‘window following’ is not the only case of spatial overlapping in room plan exploration. In modern residence apartments, not only windows, but also radiators are commonly regarded as integral elements of the wall constitution on the internal surface. It is also worth noticing that a radiator is often located below a window in the vertical dimension. These facts can easily arouse the overlapping between ‘window following’ and ‘radiator following’ to emerge during wall following in the exploration of 2.5-D virtual haptic floor plans. An appropriate re-presentation solution of overlapping spatial entities explorable without the sense of vision is still lacking. So the second study described in this paper investigates the usability of re-presenting spatial overlapping among three different types of linear entities in sonified haptic environments. More specifically, we focus (i) on a qualitative investigation on the recognition of overlaps and (ii) on a quantitative investigation on size and position estimations of the linear entities. In the following sections of this paper, we discuss our designing of the human computer interface, the conducting of the empirical experiments as well as the results.

2 Haptic-Audio Representation of Spatial Overlaps

2.1 Designing a Multimodal Interface

Multimodal designs are playing more and more important roles in modern human computer interfaces. In some HCI applications where the sense of vision is absent, other perceptual channels are used as substitution. The haptic channel might be one of the most popular perceptual substitutions for graphical representations for people without the sense of vision [8, 17]. But as pointed out by Loomis, Klatzky and Lederman there is a drawback brought by the localized characteristics of the haptic sense [12]. Verbal assistance may further improve the knowledge acquisition of haptic graphics [11]. But as the complexity of graphics increases, more mental effort will be needed to process larger amount of verbal assistance, especially for (geo-)metric or size information, e.g. length, breadth or width of spatial entities, which, in addition, is challenging to describe in textual forms. Non-speech and sonification has been widely employed in HCI interfaces. But so far sonification has been mainly used to address the information overload of verbal assistance [4, 13, 14]. Recent work has shown that sonification offers an effective help in representing 'spatial-overlaps' involving two different types of entities in graphical representation by representing the two types of entities separately via haptic and audio channel in an analog fashion [19]; in visual graphics color-coding is an easy and highly adequate solution. However it is common for graphical representations to bear spatial overlaps involving more than two types of entities. A typical example could be that in the floor plan view (2D), the location of the wall, the window and the radiator usually overlap one another. Fortunately human beings are able to form concurrent audio input into separate sound streams depending on several particular properties (e.g. frequency, rhythm, timbre, and the location of sound origins) [2, 16, 18]. This implies the potential of mapping of concurrent haptic following to concurrent audio stimuli. Yet the plausibility of using representational multimodality of sonification to support the acquisition of analog knowledge involving 'spatial-overlap' in haptic graphics has not been explored. In this paper, we realize the representational multimodality of sonification by assigning different frequency and rhythm to audio stimuli.

2.2 Realization of the Interface

We used a Phantom Omni haptic device for virtual haptic exploration, and the software was implemented based on the open source SDK released by CHAI 3D.

In order to keep the virtual haptic exploration free from complexities induced by complex virtual apartment-environment, and to focus only on perception and recognition of size and position, as well as of linear spatial overlaps, we used a virtual haptic model that only re-presents an individual wall section (Fig. 1). The one-wall model was 15 cm in length, with a boundary of 2 cm at both end of the wall section, and was placed on a solid horizontal plane. The wall and the two boundaries are 1 cm in height. So wall following exploration strategy can be perfectly supported.

For the windows and radiators in this virtual haptic environment, users are supposed to suffer from the "lost in haptic space" phenomenon when these entities are rendered as raised blocks [7]. If these entities are represented by means of different haptic features (such as different friction, magnetic properties and etc.), users could suffer from instability of virtual haptic texture [6]. And different haptic features may also leads to instability in exploration speed, so as to introduce 'experimental artifacts' in length and proportion estimation during wall following. In order to maintain steady and smooth haptic wall following, we decided to re-present the windows and radiators with audio stimuli.

For a window, we used a stable harmonic tone of 261.63 Hz (C4 as musical note). For a radiator, we used a harmonic tone of 277.18 Hz ($C^{\#}4$ as musical note) but with salient and regular rhythm, so as to be distinguishable from the windows. But as we are not experts in synthesizing audio stimuli with scientific approaches, we use salience within the rhythm to support the taxonomy for the stimuli employed in our research.

The moment the haptic interact point (see the grey point as a depiction in Fig. 2) enters the defined sonification field, the corresponding sonification will fire out. The spatial field that re-presents either a windows or a radiator is defined as haptically non-perceivable cuboid, whose intersection is a square with sides of 1 cm. All the sonification fields were placed sticking to the walls seamlessly. So the invoking of the sonification can be done during the wall following, with tolerance of reasonable noise of the hand movements.

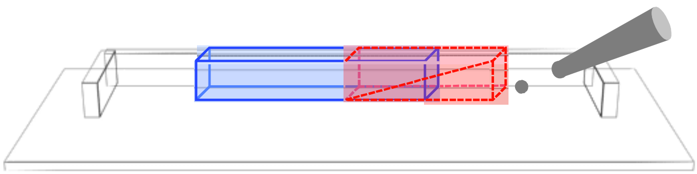


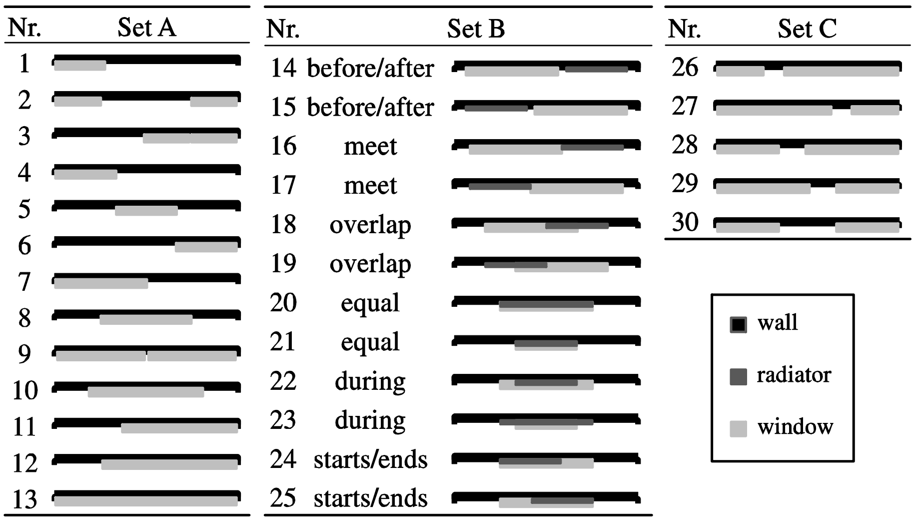
Fig. 2. The depiction of the one-wall virtual haptic model with other two different types of objects, where the blue block (left) represents a window and the red one (right) represents a radiator

3 Experiment

3.1 Materials

We conducted a repeated-measures experiment with 30 different virtual wall models. The virtual walls numbered from 1 to 13 were named as "Set A". Virtual walls of number 17 to 25 were "Set B". Number 26 to 30 were "Set C". In Set A, the same window configurations that occurred in the complete apartment experiment were repeated. A quantitative between group comparison of the performance with respect to the size and position reproduction of the windows was investigated. In Set B, all theoretically possible spatial relationships between a window and a radiator were simulated. The window-radiator configurations in Set B were a genuine sonified haptic adaption of *Allen's interval algebra* for temporal [1, 15]. Set C was covered for comparison and balancing reasons. With these experimental materials, we investigated the usability of re-presenting window and radiator by representational multimodality within auditory stimuli (harmonic tones with and without rhythm), and we contribute to discover the influence on the accuracy of perception brought by the overlapping of the stimuli.

Table 1. Depiction of virtual wall models employed in the experiment (the black color was used to represent the wall, the dark grey for the radiators, and the light grey for the windows)



3.2 Procedure

25 participants (15 female, 10 male, mean age: 24.9 years, *SD*: 3.5 years) contributed to this empirical study. They were all sighted and right-handed university students having no or little experience with haptic force feedback devices. The participants were blindfolded when they were doing the haptic exploration. We selected sighted university students as participants mainly for two reasons. First, blind and visually impaired people may not be familiar with the force feedback device we use. Thus, the time and efforts for training section could be difficult to estimate. Second, based on a big number of relevant research results collected by Cattaneo and Vecchi [5], sighted people and blind people are not significantly different in elementary audio and haptic sensitivity. With this experiment, our purpose is to investigate the usability of representational multimodality of audio stimuli. And the complexity of the virtual haptic model and the sonification stimuli we used in the experiment was also seriously limited. With these facts taken into consideration, the reported experiment was a pilot Human-Computer Interaction study based on a haptic-audio perception task. Thus the blind, the visually impaired, as well as the sighted were all appropriate experiment participants.

All the participants were first trained how to operate the force feedback device and how to deal with the experimental tasks for about half an hour. Then they were asked to explore 30 different virtual wall models in randomized order. There was no time limit for the exploration. After every exploration trial, the participants were given an answer sheet, on which a depiction of the wall of 10 cm in length was printed. The experimental task was to sketch all the windows with the blue color and the radiators with the red color as accurate as they could (Fig. 3). The participants were well informed that the width of the sonification filed was not of interest through out this experiment. The entire experiment of one participant lasted from 60 to 150 minutes. There was a short break of 10 to 15 minutes when the experiment was half done. At the end of the entire experiment, the participants were asked to fill in a questionnaire about their attitude towards the usability of this sonified haptic interface.

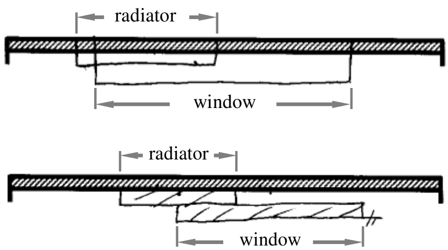


Fig. 3. Two examples of sketches produced by different participants after exploring wall configuration number 19, where the shorter object was sketched in red re-presenting a radiator, and the longer one was in blue re-presenting a window

3.3 Evaluation

As sketch is considered as a reliable data source of spatial cognition [3, 10], the evaluation was based on the analysis on the sketches.

The qualitative evaluation on the recognition of spatial relationship between the window and the radiator was done by analysis of the sketches of Set B. Referred to the original models (Table 1), if the corrected color was used to sketch the corresponding entity, and the correct spatial relationship could be recognized by the researchers, the spatial relationship of the wall configuration was regarded as correctly recognized by the participant.

Quantitative evaluation was done on the accuracy of size and position estimation of the entities. As exemplified by Fig. 4, for the size estimation analysis of an entity AB, we measured the length of the sketched entity, which is the length of segment A_1B_1. The value taken in statistic analysis for size estimation was in percentage. It was yield by dividing the absolute value of the difference between the measured length A_1B_1 and the correct length A_0B_0, by the length of the printed wall (10 cm). So the size estimation error of an entity could be computed by equation (1):

$$\Delta\text{Size} = | A_1B_1 - A_0B_0| / (10\text{ cm}) * 100\% \quad (1)$$

Similarly, the position estimation error could be computed by equation (2):

$$\Delta\text{Position} = |M_0M_1| / (10\text{ cm}) * 100\% \quad (2)$$

where M_0 is the midpoint of the entity in the original model, and M_1 is the midpoint of the sketched entity.

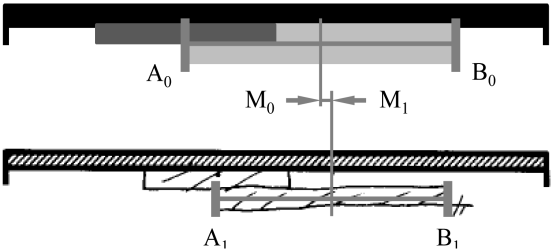


Fig. 4. Example of sketch analysis

4 Results

By going through the qualitative investigation, we tried to discovery whether it is possible to re-present spatial knowledge with overlapping between different types of linear entities with representational multimodalities of auditory stimuli in sonified virtual haptic exploration.

On the other hand, with the quantitative investigation, we were able to address two issues. First, we were able to know whether the accuracy of size and position estimation is different when the entities are explored in an apartment and only long an individual wall. Second, we focused on the influence of complexity on the accuracy of size and position estimation. In particular, we focused on the influence on the accuracy of estimations brought by the overlapping of the entities and by having different number of entities at the wall.

4.1 Qualitative Investigation

The usability investigation was constrained upon the qualitative correctness of the spatial relationship between one window and one radiator. According to the result of the analysis, all participants were able to reproduce all the re-presented spatial relationship correctly. This result strongly indicates that, in the qualitative level, spatial overlapping of linear entities can be re-presented by using representational multimodality. Again, we would like to argue that depending on having various rhythms is only one appropriate modality in auditory stimuli that could support spatial overlapping in virtual haptic exploration, and we do not exclude the existence of other (even better) representational modality for overlapping re-presentation purpose.

In the questionnaire, the participants reported that it was easy for them to understand and interact with the virtual force feedback device and the sonification. They were able to acquire the desired information by individual wish. In general, the experiment task was not challenging for them. However the wall configurations where an overlap between a window and a radiator takes place were considered as the most challenging situation by almost the participants.

4.2 Quantitative Investigation

Apartment Exploration vs. One-wall Exploration. We compared the data of Set A with the data collected in the (apartment) experiment [19]. As illustrated by Fig. 5, in the one-wall scenario experiment, the error made by the participants was significantly decreased with respect to both size estimation ($t(19.54) = 9.50$, $p < .001$) and positioning of the windows in comparison to the apartment experiment reported in [19] ($t(21.08) = 8.38$, $p < .001$[1]). This result implies that the low performance in the complex apartment scenario results from the complexity of the environment in the apartment experiment. The weak performance of the estimation in apartment scenario was very likely due to the cognition focus during the exploration was on other perspective, such as the layout of the apartment or the sizes of every room.

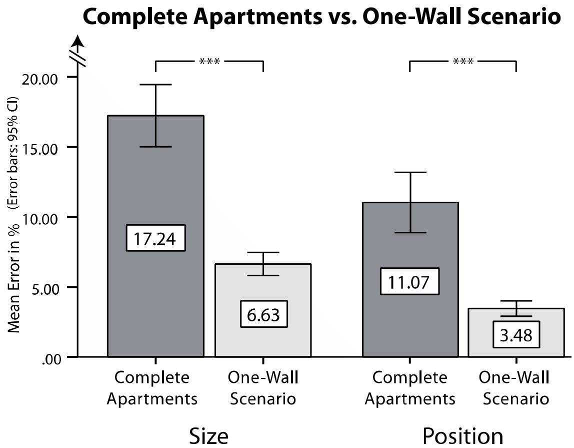


Fig. 5. Comparison of size estimation and positioning between exploration of complete apartment and one-wall scenario

Other Factors on Estimation Accuracy. In this part, we made quantitative analysis on the data of Set B to control for influences on accuracy of both size and position estimation.

First, in configurations with overlap between the window and the radiator, the error in size estimation was significantly larger than in configurations without overlap ($t(24) = -3.49$, $p < .01$; see Fig. 6 (a)). The accuracy of positioning the objects was not significantly different ($t(24) = -.77$, $p > .05$).

Second, when there were two objects on the wall, the accuracy of size estimations was not significantly different ($t(24) = -.22$, $p > .05$), but positioning accuracy was lower ($t(24) = -2.43$, $p < .05$; see Fig. 6 (b)).

---

[1] To correct for deviations from normality, the position estimation data was square-rooted. The reported p-values for the comparison are corrected for significant deviation from homogeneity of variances.

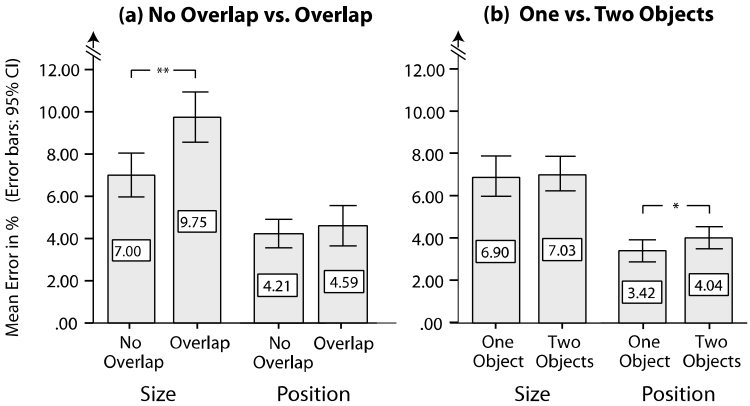


Fig. 6. Within-group comparison over size estimation and positioning accuracy

5 Discussion and Conclusion

With a combination of virtual haptics and sonification, overlapping entities can be represented as (geo-)metric spatial knowledge entities. Size and position of the represented objects can be perceived fairly good with a small amount of error (under 10% for size estimation, and about 4% for positioning). But it is more challenging to perceive precisely when overlap between two types of sonification takes place or when there are more than one object at the wall. The results of this study seem to be extendable to other types of overlaps, in particular 2D-overlap.

With respect to the design of human-computer interfaces, the results point towards the necessity to respect the interrelationship of representational complexity and accuracy of information presented by sonification early in the design process.

Acknowledgement. The research reported in this paper has been partially supported by DFG (German Science Foundation) in the International Research Training Group "Cross-modal Interaction in Natural and Artificial Cognitive Systems" (CINACS, IRTG 1247). We thank the anonymous reviewers for their valuable commentaries.

References

1. Allen, J.F.: Maintaining knowledge about temporal intervals. Communications of the ACM 26, 832–843 (1983)
2. Bernsen, N.O.: Foundations of Multimodal Representations: a Taxonomy of Representational Modalities. Interacting with Computers 6, 347–371 (1994)
3. Blades, M.: The Reliability of Data Collected from Sketch Maps. Journal of Environmental Psychology 10, 327–339 (1990)
4. Brewster, S.: Using Non-speech Sound to Overcome Information Overload. Displays 17, 179–189 (1997)

5. Cattaneo, Z., Vecchi, T.: Blind Vision: The Neuroscience of Visual Impairment. MIT Press, Cambridge (2011)
6. Choi, S., Tan, H.: Perceived Instability of Virtual Haptic Texture. I. Experimental Studies. Presence 13, 395–415 (2004)
7. Colwell, C., Petrie, H., Kornbrot, D., Hardwick, A., Furner, S.: Use of a haptic device by blind and sighted people: Perception of virtual textures and objects. In: Placencia, I., Porrero, E. (eds.) Improving the Quality of Life for the European Citizen: Technology for Inclusive Design and Equality, pp. 243–247. IOS Press, Amsterdam (1998)
8. King, A.: Re-presenting Visual Content for Blind People. PhD thesis, University of Manchester (2007)
9. Lahav, O., Mioduser, D.: Haptic-feedback Support for Cognitive Mapping of Unknown Spaces by People who are Blind. Intern. J. Human-Computer Studies 66, 23–35 (2008)
10. Lohmann, K.: The Use of Sketch Maps as Measures for Spatial Knowledge. In: Wang, J., Broelemann, K., Chipofya, M., Schwering, A., Wallgrün, J.O. (eds.) Proceedings of the COSIT 2011 Workshop "An Interdisciplinary Approach to Understanding and Processing Sketch Maps", pp. 45–54 (2011)
11. Lohmann, K., Eschenbach, C., Habel, C.: Linking Spatial Haptic Perception to Linguistic Representations: Assisting Utterances for Tactile-Map Explorations. In: Egenhofer, M., Giudice, N., Moratz, R., Worboys, M. (eds.) COSIT 2011. LNCS, vol. 6899, pp. 328–349. Springer, Heidelberg (2011)
12. Loomis, J., Klatzky, R., Lederman, S.: Similarity of Tactual and Visual Picture Recognition with Limited Field of View. Perception 20, 167–177 (1991)
13. Nees, M.A., Walker, B.N.: Auditory Interfaces and Sonification. In: Stephanidis, C. (ed.) The Universal Access Handbook, pp. 507–521. CRC Press, New York (2009)
14. Rassmus-Gröhn, K.: User Centered Design of Non-Visual Audio-Haptics. Doctoral Thesis, Certec, Lund University 2:2008 (2008)
15. Rauh, R., Kulik, L.: The Influence of Linear Shapes on Solving Interval-Based Configuration Problems. In: Habel, C., Brauer, W., Freksa, C., Wender, K.F. (eds.) Spatial Cognition 2000. LNCS (LNAI), vol. 1849, p. 239. Springer, Heidelberg (2000)
16. Shamma, S.A., Elhilali, M., Michey, C.: Temporal Coherence and Attention in Auditory Scene Analysis. Trends in Neurosciences 34, 114–123 (2011)
17. Sjöström, C., Danielsson, H., Magnusson, C., Rassmus-Gröhn, K.: Phantom-based Haptic Line Graphics for Blind Persons. Visual Impairment Research 5, 13–32 (2003)
18. Song, H.J., Beilharz, K., Cabrera, D.: Evaluation of Spatial Presentation in Sonification for Identifying Concurrent Audio Streams. In: Proceedings of the 13th International Conference on Auditory Display, Montreal, Canada, June 26-29 (2007)
19. Yu, J., Habel, C.: A Haptic-Audio Interface for Acquiring Spatial Knowledge about Apartments. In: Magnusson, C., Szymczak, D., Brewster, S. (eds.) HAID 2012. LNCS, vol. 7468, pp. 21–30. Springer, Heidelberg (2012)