Sound for enhanced experiences in mobile applications

Mats Liljedahl and Katarina Delsing

Interactive Institute

Background

When visiting new places you want information about restaurants, shopping, places of historic interest etc. Smartphones are perfect tools for delivering such location-based information, but the risk is that users get absorbed by texts, maps, videos etc. on the device screen and get a second-hand experience of the environment they are visiting rather than the sought-after first-hand experience.

One problem is that the users' eyes often are directed to the device screen, rather than to the surrounding environment. Another problem is that interpreting more or less abstract information on maps, texts, images etc. may take up significant shares of the users' overall cognitive resources.

The work presented here tried to overcome these two problems by studying design for human-computer interaction based on the users' everyday abilities such as directional hearing and point and sweep gestures. Today's smartphones know where you are, in what direction you are pointing the device and they have systems for rendering spatial audio. These readily available technologies hold the potential to make information more easy to interpret and use, demand less cognitive resources and free the users from having to look more or less constantly on a device screen.

Description of the activities

Two prototypic smartphone applications were developed. The first app was a tourist guide and the second a gaming application. Both applications had multimodal user interfaces based on GPS positioning and electronic compass for input from the users, and a balanced mix of spatial and non-spatial audio, text and graphics as output to the users.

The guide app allowed users to search for points of interest (POIs) and to get guidance to selected locations. To perform a search, the users swept the device horizontally in front of them. As the users swept the device in different directions, the app played a short sound for each POI the users swept by. In this way the users could get an overview of the density of POIs in different directions. When an interesting direction was identified, the users could get more detailed information about the POIs in that direction by tapping an on-screen button. Detailed information was then presented using text and images.

The app could also guide users to selected locations through spatial audio, a graphic arrow and text. The GPS accuracy of smartphones is often not high enough to reliably guide users through turn-by-turn navigation. Instead a method based on the direction to the end target location was developed. The app put a virtual sound source on the target location. Depending on how the user pointed her device in relation to the target, the sound from the virtual sound source was moved between left and right ear. The effect resembled hearing in real-life. The more to the right of the target the user was pointing, the more to the left ear the virtual sound source was moved and vice versa. The users' everyday ability to locate sound sources was used to guide towards the target without the need to interpret more or less abstract information. The direction to the target was also shown with a graphic arrow on the device screen. The distance to the target was encoded into the sound and displayed in digits. Users were free to choose their personal mix of spatial audio and on-screen information for guidance.

The game application used GPS location and pointing direction as input and mainly spatial and non-spatial audio as output. The users' movements in the physical world were reflected in the game's virtual world. Information from the virtual game world about directions and distances to opponents was conveyed through audio. Opponents were found by pointing the device and listening. Attacks were made by pointing and activating an on-screen "fire-button". The damage made by an attack was conveyed through audio to both the attacked player and the attacking player. Current health status was shown graphically on the device screen.

To test the guide application, 24 test users aged 14 to 50 years where enrolled, 15 females and 9 males. The test users were asked to use the app to search for three specified targets and to use the guide function and walk to these. 11 users aged 15 to 30 years, three females and eight males, tested the game application. The test users were simply asked to play the game. After the tests, all test users filled out a questionnaire and took part in focus group interviews.

Results

All users of the guide application managed to use the multimodal search function to get information about the designated targets without problems. All users could also effectively and without problems navigate to these target locations using the multimodal guide function. 75% of the test users stated that they wanted to explore other cities using the application. A NASA Task Load Index showed low overall mental and physical demands and the users reported being successful and unstressed in fulfilling the navigation task. In the focus group interviews, some users reported using the audio feedback to a large extent, both when searching for targets and while getting guidance there. These users appreciated being able to use to application more or less “eyes free”. Other users reported relying almost exclusively on the graphical information. Some users asked for greater diversity between the different sounds in order to more easily discriminate between them and their different meanings. One user confused the audio feedback from the application with the ring signal from his mobile phone. At some occasions the sounds from the application was drowned by background noise from traffic or machines. To some users the application did not convey enough information through audio about direction (left/right) or distance to target. Using speech to give the information “turn left” and “turn right” was suggested as a solution.

The test of the game application showed that all players managed to locate opponents by pointing and listening for audio cues. There was also a strong agreement that the players enjoyed playing the game.

All in all the results show that the multimodal interfaces were both effective and fun to use. The spatial and non-spatial audio cues contributed significantly to the overall user experience.

Conclusions

Applications featuring multimodal user interfaces hold the potential to unburden users from cognitive loads and interpretative tasks. Guide applications with such interfaces can effectively help users find points of interest and to guide the way, while at the same time leaving users more free to experience and explore the environment compared to using for example traditional maps. In game applications multimodal interfaces can support rich pervasive gaming experiences.

For a system to give a trustworthy user experience, its information must match the users real-life experience. Low-accuracy GPS data A positive side effect with using virtual sound sources is that they can help overcoming problems with low accuracy in GPS-data. We are used to handle the fact that the eyes can more precisely determine direction and distance to targets compared to the ears. Therefore, auditory guide information based on low-accuracy GPS and compass data can be less disruptive for the user experience compared to visual information based on the same data.

Keywords

multimodal, interface, experience, spatial, audio

References

1. M. Anastassova, C. Magnusson, M. Pielot, G. Randall, and G. B. Claassen. Using audio and haptics for delivering spatial information via mobile devices. In Proceedings of the 12th international conference on Human computer interaction with mobile devices and services (MobileHCI '10). ACM, New York, NY, USA, 525-526, 2010. DOI=10.1145/1851600.1851734 http://doi.acm.org/10.1145/1851600.1851734

2. S. Holland, D. R. Morse and H. Gedenryd. AudioGPS: spatial audio navigation with a minimal attention interface. Personal Ubiquitous Computing 6(4), 253–259, 2002.

3. M. Jones, S. Jones, G. Bradley, N. Warren, D. Bainbridge and Holmes, G. Ontrack: dynamically adapting music playback to support navigation. In Personal and Ubiquitous Computing, 12. Springer-Verlag, 513-525, 2008.

4. M. Liljedahl and S. Lindberg. Sound Parameters for Expressing Geographic Distance in a Mobile Navigation Application. In Proceeedings of the 6th Audio Mostly, A Conference on Interaction With Sound. September 2011, Coimbra, Portugal. ACM, pp. 1 – 7, 2011.

5. C. Magnusson, B. Breidegard, K. Rassmus-Gröhn. "Soundcrumbs – Hansel and Gretel in the 21st century", In Proceedings of the 4th international workshop on Haptic and Audio Interaction Design (HAID '09), 2009.

6. C. Magnusson, M. Molina, K. Rassmus-Gröhn and D. Szymczak. Pointing for non-visual orientation and navigation. In Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10). ACM, New York, NY, USA, 735-738, 2010. DOI=10.1145/1868914.1869017 http://doi.acm.org/10.1145/1868914.1869017

7. C. Magnusson, K. Rassmus-Gröhn and D. Szymczak. Angle sizes for pointing gestures. In Proceedings of Workshop on Multimodal Location Based Techniques for Extreme Navigation, Pervasive 2010 (Helsinki, Finland), 2010. Retrieved October 27, 2011, from http://www.english.certec.lth.se/haptics/HaptiMap/proceedings_combined_final_with_frontmatter.pdf

8. D. McGookin, S. Brewster and P. Priego. Audio Bubbles: Employing Non-speech Audio to Support Tourist Wayfinding. In Proc. HAID 2009, Springer-Verlag, 41 – 50, 2009.

9. M. Pielot, S. Boll. "In Fifty Meters Turn Left": Why Turn-by-turn Instructions Fail Pedestrians. In Proceedings from the Workshop Using Audio and Haptics for Delivering Spatial Information via Mobile Devices. MobileCHI 2010, Lisbon, Portugal, 2010. Retrieved October 27, 2011, from http://www.english.certec.lth.se/haptics/HaptiMap/MobileHCI2010workshop_proceedings.pdf

10. B. Poppinga, C. Magnusson, W. Heuten, D. McGookin, N. Henze, G. B. Claasen, M. Pielot, H. Eftring and J. Peters. Observing the Mobile user Experience, 2010. Retrieved October 27, 2011, from http://omue10.offis.de/files/OMUE10-Proceedings.pdf

11. S. Robinson, M. Jones, P. Eslambolchilar, R. Murray-Smith and M. Lindborg. "I did it my way": moving away from the tyranny of turn-by-turn pedestrian navigation. In Proc. of Mobile-HCI '10. ACM, 341-344, 2010.

12. D. Spath, M. Peissner, L. Hagenmeyer and B. Ringbauer. New Approaches to Intuitive Auditory User Interfaces. In M.J. Smith, G. Salvendy (Eds.): Human Interface, Part I, HCII 2007, LNCS 4557, pp. 975–984, 2007.

13. S. Strachan, P. Eslambolchilar and R. Murray-Smith. gpsTunes - controlling navigation via audio feedback. In Proc. MobileHCI 2005, vol. 1. ACM, 275–278, 2005.