Mobile Crossmodal Auditory and Tactile Displays

Eve Hoggan and Stephen Brewster
Multimodal Interaction Group | GIST
Department of Computing Science
University of Glasgow
Glasgow, UK, G12 8QQ
{Eve,Stephen}@dcs.gla.ac.uk

Abstract
This paper describes an alternative form of interaction for mobile devices using crossmodal output. These crossmodal displays allow alternative senses such as hearing and touch to be used to perceive information. The aim of our work is to investigate the equivalence of audio and tactile displays so that the same messages can be presented in one form or another. Initial experiments show that rhythm, texture, and spatial location can be perceived as equivalent in both the auditory and tactile modalities.

Keywords
Tactons, Earcons, crossmodal interaction, mobility

Introduction
Unlike desktop users, mobile device users are usually in motion whilst using their device. This means that their visual focus is on their primary task (e.g. walking or exercising) and they cannot dedicate visual attention to interacting with the mobile device. It is difficult to design entirely visual interfaces that work well under these mobile circumstances. Despite this however the interfaces used by most mobile devices are based on desktop designs. With the amount of information handled by mobile devices increasing everyday most mobile interfaces have problems displaying such a vast quantity of data due to their reliance on small visual displays. If the information available from these devices is to be accessible in an effective and safe way to all users on the move, it will be necessary to address the mobility restrictions enforced by such visually demanding interface designs.

It can be unnatural to be forced to interact with the environment around us using only our vision. For example, in a dark environment we may choose to use touch for navigation as opposed to vision. Similarly, people with sensory disabilities are often forced to use alternative senses. A common technique used to help the sensory impaired is sensory substitution where one sensory modality is used to supply environmental information normally gathered by another sense. The replacement of a sense by another one could be employed by mobile devices too. It could be argued that mobile device users are 'situationally impaired' for example wearing gloves, being in a very noisy environment, or driving. Using notions drawn from sensory substitution, mobile device users could simply use the appropriate modalities as desired.

The manufacturers of mobile devices like PDAs and phones commonly include audio and vibrotactile feedback in their products. This research will build on these features by using auditory/tactile crossmodal output to overcome problems experienced by situationally impaired users of mobile devices.

Crossmodal Interaction

It is important to outline the definition of crossmodal interaction in the context of this research. Crossmodal interaction relates to both synesthesia and sensory substitution. Unlike multimodal interaction where each modality is used to transmit a different type of information e.g. audio for alerts and vision for graphical data, crossmodal interaction uses the different modalities to present the same data. Crossmodal use of the different senses allows the characteristics of one sensory modality to be transformed into stimuli for another sensory modality [1]

The crossmodal interactions of different sensory inputs have advantages because they can help to overcome a specific sensory deprivation or situational impairment (e.g., auditory and tactile senses in darkness) and they can also reduce perceptual ambiguity.

Audio and tactile displays are ideal candidates for crossmodal use because our senses of hearing and touch share several important similarities, in particular their temporal characteristics and their ability to perceive vibrations. An attribute that can communicate comparable information across modalities is considered to be amodal. Intensity, spatial location, rate, texture, and rhythmic structure are common types of amodal attributes [2]. The auditory/tactile crossmodal interaction design described here is based on the amodal attributes available in the auditory and tactile domains.

Crossmodal Parameters

If we wish to provide effective crossmodal output to mobile displays, it is first necessary to investigate the different parameters available for manipulation in the auditory and tactile modalities because sensory substitution requires the same information to be encoded and presented interchangeably via both modalities.

There is a significant amount of research on individual modalities. Earcons are a type of non-speech auditory display, which Blattner defines as "non-verbal audio messages that are used in the computer/user interface to provide information to the user" [3]. Similarly, in the tactile domain, Brewster and Brown have developed Tactons [4] for structured vibrotactile messages which can be used to communicate information non-visually.

In order to explore the possibilities of crossmodal auditory/tactile output, we have created several crossmodal icons. Crossmodal icons [5] are abstract icons which can be automatically instantiated as either an Earcon or Tacton, such that the resultant Earcons or Tactons are intuitively equivalent and can be compared as such. Crossmodal icons allow the same information to be accessible interchangeably via several different modalities.

The current parameters or amodal attributes under investigation for auditory/tactile crossmodal icons are:

Rhythm
This is an extremely important parameter in both Earcons and Tactons [4,6]. Rhythm is an amodal property in the audio and tactile domain because it can be directly transferred between modalities. The MICOLE project [7] has conducted an experiment in crossmodal equivalence in rhythm recognition and reproduction using loudspeakers for the audio and a Logitech WingMan mouse for the tactile. The results of the experiments showed that there were more correctly reproduced rhythms in tactile than visual modality. In that sense, the tactile modality settles between the audio and visual modalities.

Given that these experiments have shown that rhythm can be perceived as equivalent in the audio and tactile modalities, rhythm is a potential parameter for crossmodal interaction.

Texture
Modulating the amplitude of a tactile pulse creates differing levels of roughness [4]. An experiment was conducted to determine which version of audio roughness (dissonance, flutter-tonguing, amplitude modulation, or timbre) can be perceived as equivalent and maps most effectively to tactile roughness. It has been shown that using timbre or amplitude modulation produces better results than flutter tonguing and dissonance. This suggests that participants found it easier to match audio and tactile cues when the audio cues used amplitude modulation or timbre. Initial results also show that subjects preferred the use of differing timbres in audio. However, the results also show no significant difference in performance between timbre and audio amplitude modulation.

These results suggest that crossmodal roughness in the auditory domain should be created using either amplitude modulation or differing timbres.

Spatial Location
This is another amodal attribute found in both the auditory and tactile domains.

Stationary Spatial Location Experiment
Experiments were conducted to investigate which body location can be mapped most effectively to locations in a 3D audio soundscape. The experiment involved a computer-controlled belt/wrist band/ankle band with four embedded vibrotactile transducers: each of the small transducers are evenly spaced around the circumference of the body area (waist, wrist or ankle) and mapped to spatial audio played through a pair of headphones.

Results show that participants are able to map the presented 3D audio positions to tactile body positions on the waist most effectively and that there are significantly more errors when using the ankle.

Mobile Spatial Location Experiment
Users of mobile devices are often in motion when they use their devices (e.g. receiving calls, sending text messages, etc.). Interfaces must be designed to work well under these circumstances too, not just when the user is stationary.

Given the promising results of the stationary spatial location experiment, the same experiment was conducted again in a mobile situation to see if motion affects the results. There are many ways in which motion could affect perception of crossmodal output: mobile environments tend to change frequently, the user’s main attention may be on safety whilst crossing a road instead of the mobile device, a user can become physically tired, and during natural motion such as walking, a user’s hands are likely to be moving.

This mobile experiment used a treadmill in a usability lab to simulate mobility because the tactile actuators used were not wireless and were controlled from a PC and therefore inappropriate for use in a real mobile environment. Furthermore, using a treadmill permitted the experimenter to set a standard speed for all participants (in this case, all walked at a speed of 6km per hour).

Results show that participants are able to map the presented 3D audio positions to tactile body positions on the waist most effectively when mobile and that there are significantly more errors made when using the ankle or wrist. Unlike the previous experiment, a greater number of participants preferred the waist to the wrist and ankle. However significantly more participants still preferred the wrist to the ankle.

Combining Parameters

As discussed above, initial research into crossmodal icons has shown parameters such as rhythm, texture, and spatial location to be easy to map between the audio and tactile domains. There is, however, no complete set of parameters so the next steps in this research will be to develop a complete set of cues by combining parameters to identify what works well across the two modalities

The current experiment being conducted within this research incorporates three parameters: rhythm, texture, and spatial location as these have been shown to be successful in the earlier experiments. A complete set of crossmodal cues will be created to represent, for example, alerts that may be generated when an important appointment in the mobile device user's calendar. Given that the parameters can be perceived as equivalent in both parameters, the appointment information may be presented via audio or tactile or both.

The aim of this experiment will be to investigate absolute identification of both the audio and tactile cues and also absolute matching between the modalities. As with the spatial location experiment, this experiment will be conducted in both a stationary and mobile environment in order to ensure that the results are applicable to the various different situations experienced by mobile device users.

It will also be necessary to investigate parameter interactions. For example, using a certain spatial location may force a particular level of roughness. It will also be necessary to establish the resolution of these parameters so that the number of different distinguishable roughness levels and distinguishable spatial locations can be outlined. This will help to determine how much information can be encoded in these crossmodal displays

Conclusions

This paper has described the features of crossmodal auditory/tactile interaction and has investigated some of the potential parameters that could be used to create crossmodal audio and tactile icons. Initial experiments have shown that both roughness and spatial location can be perceived as equivalent in both the auditory and tactile domain. The current experiment involving a complete set of crossmodal cues using multiple parameters will help to inform designers as to the best ways to include crossmodal icons in various mobile applications.

References

[1] Lenay, C., Canu, C., Villon, S.P. Technology and perception: the contribution of sensory substitution systems. In 2nd International Conference on Cognitive Technology (CT '97), (1997).

[2] Lewkowicz, D.J. The development of intersensory temporal perception: an epigenetic systems/limitations view. Psychological Bulletin 126, (2000), 281 -308.

[3] Blattner, M.M., D.A. Sumikawa, and R.M. Greenberg, Earcons and Icons: Their Structure and Common Design Principles. In Human Computer Interaction, 1989. 4(1): p. 11-44.

[4] Brown, L.M., Brewster, S.A. and Purchase, H.C. A First Investigation into the Effectiveness of Tactons. In Proceedings of WorldHaptics 2005, IEEE Press, (2005), 167-176.

[5] Hoggan, E. and Brewster, S. Crossmodal icons for information display. In Proceedings of CHI 2006, ACM Press, (2006), vol. II, 857-862.

[6] Brewster, S.A. Providing a structured method for integrating non-speech audio into human-computer interfaces. PhD Thesis, University of York, UK, (1994).

[7] Kosonen, K. and Raisamo, R. Rhythm perception through different modalities. In Proceedings of Eurohaptics, France, (2006).