EMPIRICALLY DERIVED GUIDELINES FOR THE PRESENTATION OF CONCURRENT EARCONS

David K McGookin, Stephen A Brewster
Glasgow Interactive Systems Group
Department of Computing Science
University of Glasgow
Glasgow, G12 8QQ, UK
{mcgookdk, stephen}@dcs.gla.ac.uk
www.dcs.gla.ac.uk/~{mcgookdk, stephen}

ABSTRACT

This paper presents a set of empirically derived guidelines for the presentation of concurrent earcons (short structured audio messages which can be used to effectively communicate information to users [3]). When earcons are presented in such a way, they interfere with each other, making it difficult to determine the information encoded within them. The guidelines presented in this paper cover the impact of varying the number of earcons on their identification as well as how the design and presentation of earcons may be modified to reduce interference when concurrently presented.

Keywords

Earcons, Spatialised Audio, Guidelines, Auditory Display, Evaluation

1. INTRODUCTION

The use of sound in human computer interaction has been an active area of research for the last decade; allowing for the creation of effective user interfaces which do not require the use of a visual display. Recently such advantages of so called auditory displays have been realised in the creation of displays for small form factor mobile devices, where both screen space is limited and it is important for users to be aware of the ever changing environment around them, which necessarily draws attention away from a visual display. Additionally, auditory displays have been used to create usable interfaces for visually impaired people, allowing them to overcome the problems of graphical user interfaces, when one cannot actually see the graphical display [8]. Whilst sound offers such advantages, it also has disadvantages, notably a lack of persistence, making comparisons between multiple sequentially presented data difficult. Additionally, in comparison to the bandwidth of visual displays, auditory displays are more limited, due in part to the reduced spatial acuity of the auditory perceptual system in comparison to the visual [10]. This makes it difficult to create auditory displays which can fully replace visual ones. One way of overcoming such problems is through the use of concurrent auditory presentation. That is, presenting sounds which encode information in a temporally overlapping fashion. Doing so allows the presentation bandwidth of auditory displays to be increased, allowing data to be concurrently presented and therefore allowing comparisons between data to be more easily made [6]. Several systems such as the portable audio based news and diary manager, Nomadic Radio [15], have exploited such advantages by the concurrent presentation of multiple data.

Whilst concurrently presenting sounds is useful, there are serious problems with such presentation that need to be considered. Sounds which are concurrently presented are likely to interfere with each other, making them difficult to segregate and thus making it impossible to extract any data encoded in the sounds, thereby reducing their effectiveness as part of an auditory display. Whilst simple psychoacoustical interference is predictable and can be easily controlled, other interactions are more difficult to predict and are based on the degree of similarity between concurrently presented sounds. Hence, if two sounds are concurrently presented, where each has a different rhythmic structure which encodes data, and each sound is presented with the same timbre (or musical instrument), the sounds may be merged such that the user will hear only one rhythmic structure, which does not correspond to any data mapping. These issues are compounded by a lack of guidelines that are available to designers of auditory displays to design and use sounds which are less likely to interfere with each other when concurrently presented. This is a particular problem for designers who wish to use concurrently presented earcons in an auditory display. Earcons are *"abstract synthetic tones which can be used in structured combinations to create auditory messages"* [3], where different auditory parameters (such as pitch, timbre, rhythm) are mapped to different data attributes (such as file type, file size etc.) according to a set of rules or "grammar''. E.g. a piano timbre might represent a word processing file, a trumpet timbre might represent a paint application file, etc. Rhythm may be mapped to file size, with two or three different rhythms mapping to different categories of file size (e.g. will fit on a floppy disk, will fit on a CD-ROM etc.). Such earcons are easy to learn [4] as only a small set of rules and sounds need to be learned, instead of a completely different sound for each file type and size combination. However, earcons formed from the same grammar are likely to be similar, sharing timbres or rhythms, making them likely to interfere with each other in the way mentioned above.

Auditory Scene Analysis (ASA) [2] is the psychological study of interactions between concurrently presented sounds. It is impossible in the space available to go into a great deal of detail on ASA, (the interested reader is directed towards Bregman [2] and Deutsch [7]). ASA does state however, that the greater the differences between concurrently presented sounds in terms of their timbre, rhythm, pitch etc., the less likely it will be that those sounds will be grouped together, so that they are perceived as only a single sound. However, there is little guidance on which factors dominate the grouping process, or how much of a difference is required between sounds to affect that process. For example, does a small difference in timbre dominate over a large difference in rhythm? In any case, due to the grammar that a particular set of earcons may be formed from, it would not be possible to make arbitrarily large changes between concurrently presented earcons. Whilst such changes may avoid undesirable grouping between the earcons, they may destroy the mapping between the sound and data the earcons useless as part of an auditory display. For example, it is not possible to overcome the problem of interference in the previous example by arbitrarily changing the timbre of one earcon, as this will change the meaning of the earcon, rendering any reduction in interference useless.

This paper presents guidelines derived from a three year research study which has investigated the problems of concurrently presented earcons [12, 13, 11]. It has investigated how concurrent earcons can be redesigned and presented to reduce interference, without destroying the grammatical relationships between earcons. This study has investigated three specific areas of concurrent earcon identification (increasing/decreasing the number of concurrently presented earcons, modifications to the design and presentation of earcons when they are presented from the same location in space, and the impact of each earcon being presented in a different spatial (3D) location around the user's head), within which guidelines for concurrent earcon design and presentation will be outlined.

2. QUALIFYING THE GUIDELINES

In order to make full use of the guidelines presented here, some understanding of the earcons used in the work from which the guidelines are derived is important. The earcons used were of the transformational type [1]. With transformational earcons each auditory parameter, such as timbre, rhythm etc is directly mapped to a data parameter. E.g. in the example from the last section, file type is mapped to timbre, file size to rhythm, with different timbres and rhythms defining different data values for file type and size. Each earcon encoded three data parameters; additionally each data parameter could have three distinct values. This created a set of 27 distinct earcons. In the earcons, timbre (musical instrument), melody (a combination of rhythm and relative pitch spacing between the notes of that rhythm) and register (the set of pitches that the melody is presented in), were used to encode the different data parameters. The guidelines of Brewster, Wright and Edwards [5] (BWE) who investigated the individual presentation of earcons were used to design an initial earcon set, empirical evaluations between that initial set and the same set with some modifications guided by ASA research, being used to derive the guidelines presented in this paper. The guidelines presented here therefore should be used in conjunction to those of Brewster, Wright and Edwards [5] to design earcons which are to be concurrently presented.

3. NUMBER OF EARCONS VS. IDENTIFICATION

Varying the number of concurrently presented earcons has a significant impact on the proportion of those earcons that can be successfully identified by users. In work where participants had to identify earcons and their attributes when sets of one, two, three and four earcons were concurrently presented, identification of earcons fell from 70% (consistent with similar work by Brewster [3] on single earcon identification) when only one earcon was presented, to 30% when four earcons were concurrently presented [12]. Designers should therefore reduce the number of concurrently presented earcons as much as possible. These results are illustrated in Figure 1. Figure 1 also shows how individual earcon attribute identification is affected by the number of concurrently presented earcons. Identification of both timbre and melody encoded data attributes, whilst still following the same trend as earcon identification, is much higher; dropping from 95% to 65% as the number of concurrently presented earcons is increased. This indicates that reducing the density of the earcons (i.e. reducing the number of data attributes encoded in each earcon), may assist in increasing the number of earcons which can be identified, although this would restrict the amount of data which each earcon could encode.

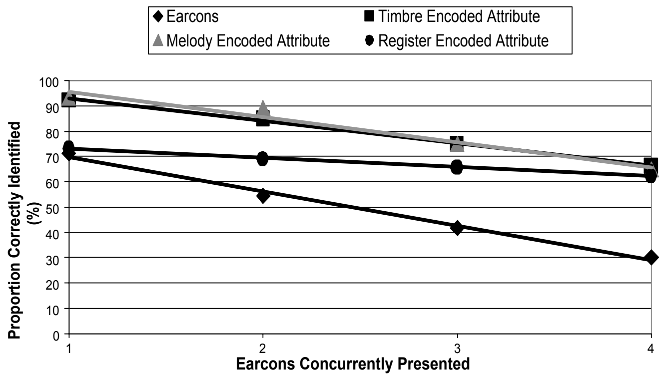

Figure 1: Graph showing how earcon identification is affected by the number of earcons concurrently presented. Adapted from McGookin and Brewster [11].

Identification of the register encoded earcon attribute whilst being much lower than timbre and melody, does fall at a much shallower gradient, due in part to the incorporation of inharmonic musical intervals between the registers used [11]. Designers should therefore use such intervals when register is used to encode a data attribute.

4. ENGINEERING MORE ROBUST EARCONS

There are a limited number of modifications which can be applied to the design of earcons in order to improve their identification when concurrently presented. As explained in the introduction, earcons cannot be arbitrarily modified, as the relationships between sound attributes (such as timbre, melody etc.) and the data which are encoded in such attributes, must be preserved. This constrains the extent to which modification can be applied to improve earcon identification.

Presenting each earcon with a different timbre can significantly improve the identification of any timbre encoded earcon attribute [12]. Whilst the guidelines of BWE already recommend that different data values have different timbres associated with them, they do not take account of two earcons with the same timbre encoded data value, but different melody and register encoded data values being concurrently presented. For situations where the user may be presented with two earcons which encode the same data value, designers should use different timbres for each earcon which are derived from the same instrument group [14]. For example, if a piano timbre is used to encode the data value, one earcon should use a grand piano timbre and another should use an electric piano timbre. By doing so the influence of timbre in the ASA process can be reduced whilst still preserving the earcon grammar. However, when incorporating this guideline a sufficient number of instrument groups must be available (one for each distinct data value encoded by earcon timbre), and each instrument group must contain a sufficient number of instruments (one for each earcon that is to be concurrently presented which encodes the same data value).

It is possible to significantly increase the number of concurrently presented earcons that can be identified by ensuring that all earcons do not start at exactly the same time. Staggering the onsets of concurrently presented earcons by at least 300ms is sufficient to significantly improve their identification, as well as significantly increase the identification of any timbre encoded data attribute [12]. Both timbre and onset modifications can be combined without a significant degradation in identification performance.

Earcons which incorporate the guidelines of this section are likely to still show the same trend as Figure 1 when the number concurrently presented is varied. Other studies indicate that the gradient of the trend can be significantly flattened [9], leading to an increase in the number of earcons which can be sucessfuly identified by users.

5. SPATIALISATION VS. IDENTIFICATION

An important feature of ASA which can be used to improve the identification of concurrently presented earcons is the use of spatialised presentation. Such presentation allows for earcons to be separated in space and thus be less prone to undesirable interference. As Bregman [2] notes *"It would be a good bet that sounds emanating from the same location have been created by the same sound source"*. Spatialisation yields such a significant effect in concurrent earcon identification, that when earcons which do not comply with the guidelines described in Section 4 are presented in spatially distinct locations, their identification is superior to earcons which are non-spatially presented but do incorporate the guidelines from Section 4. When concurrently presenting earcons therefore, spatialisation should be employed whenever possible [13]. The impact of spatialisation however is dependant on the degree of separation between concurrent sounds that can be practically achieved. The closer two earcons are presented, the greater the interference between them will be. This may be a problem when a data parameter is mapped to spatial location, which may make it difficult to ensure concurrently presented earcons are spatially separated enough to reduce interferences between them. The guidelines outlined in Section 4 should therefore also be incorporated with spatial presentation to increase the identification of concurrently presented earcons [13].

6. CONCLUSIONS

This paper has presented a set of guidelines for the concurrent presentation of earcons. We have also shown why such design guidelines are necessary when earcons are concurrently presented. These guidelines, derived from experimental work, have provided guidance as to the impact on identification of varying the number of concurrently presented earcons, as well as design and presentation modifications in both a spatialised and non-spatialised auditory environment. The guidelines presented in this paper will assist designers who wish to use concurrently presented earcons to design interfaces that are more effective at communicating information to users.

7. ACKNOWLEDGEMENTS

This work was supported by EPSRC studentship 00305222.

8. REFERENCES

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

[2] Bregman, A.S. *Auditory Scene Analysis*. MIT Press, Cambridge, Massachusetts, 1994.

[3] Brewster, S.A. *Providing a structured method for integrating non-speech audio into human-computer interfaces*. Department of Computer Science, University of York, 1994.

[4] Brewster, S.A. Using Non-Speech Sounds to Provide Navigation Cues. *ACM Transactions on CHI* 5, 2 (1998), 224-259.

[5] Brewster, S.A., Wright, P.C. and Edwards, A.D.N. Experimentally Derived Guidelines for the Creation of Earcons. *Proceedings of BCS-HCI '95* (Huddersfield, UK) Springer, 1995, pp. 155-159.

[6] Brown, L., Brewster, S., Ramloll, R., Yu, W. and Riedel, B. Browsing Modes for exploring Sonified Line Graphs. *Proceedings of BCS-HCI 2002* (London, UK) BCS, 2002, pp. 6-9.

[7] Deutsch, D. Grouping Mechanisms in Music. *The Psychology of Music*, Deutsch, D. (Ed.), Academic Press, San Diego, 1999, 299-348.

[8] Edwards, K.W. and Mynatt, E.D. An Architecture for Transforming graphical Interfaces. *proceedings of UIST 94* (Marina del Ray, California) ACM Press, 1994, pp. 39-47.

[9] Folds, D.J. and Gerth, J.M. Auditory Monitoring of up to Eight Simultaneous Sources. *Human Factors and Ergonomics Society 38th Annual Meeting*, 1994, pp. 505-509.

[10] Kramer, G. An Introduction to Auditory Display. *Auditory Display: Sonification, Audification and Auditory Interfaces*, Kramer, G. (Ed.), Addison-Wesley, Reading, Massachusetts, 1994, 1-77.

[11] McGookin, D.K. and Brewster, S.A. Understanding Concurrent Earcons: Applying Auditory Scene Analysis Principles to Concurrent Earcon Recognition. *to appear in ACM Transactions on Applied Perception*.

[12] McGookin, D.K. and Brewster, S.A. An Investigation into the Identification of Concurrently Presented Earcons. *Proceedings of ICAD 2003* (Boston, MA) ICAD, 2003, pp. 42-46.

[13] McGookin, D.K. and Brewster, S.A. Space the Final Frontearcon: The Identification of Concurrently Presented Earcons in a Synthetic Spatialised Auditory Environment. *to appear in proceedings of ICAD 2004* (Sydney, Australia) ICAD, 2004.

[14] Rigas, D.I. *Guidelines for Auditory Interface Design: An Empirical Investigation*. Loughborough University, 1996.

[15] Sawhney, N. and Schmandt, C. Nomadic Radio: Speech & Audio Interaction for Contextual Messaging in Nomadic Environments. *ACM Transactions on CHI* 7, 3 (2000), 353-383.