SoundVision

Advanced Auditory Displays for the Blind and Visually Impaired

Associated Researchers: Christopher Frauenberger [frauenberger@iem.at]
Veronika Putz [corrado@sbox.tugraz.at]
Robert Höldrich [hoeldrich@iem.at]

SoundVision research work aims at the development of a human-computer interface based on the auditory mode, exploiting the extensive capabilities of the human auditory system.
As the first application of such an advanced auditory display SoundVision will focus on a system providing efficient access to digital information resources for the visually impaired and the blind. Having the focus on this specific target group, SoundVision will develop its concepts and techniques to be applicable in different fields like mobile computing, personal digital assistants (PDAs) or wherever the visual mode is not available, restricted by form factors or insufficient for the amount of information to be presented. An increasing amount of information is available in digital form making computers the main information source in our information society nowadays. Conventional human-computer interfaces are mainly based on the visual mode, providing minor or no access for the visually impaired and the blind and establishing the danger of creating a "digital divide".
The major objectives of SoundVision can be identified as:

1) Development of Auditory Information Display techniques considering their reusability for pushing the performance of the auditory interaction mode generating solutions in a wide range of applications.
2) Determination of usability engineering methodologies for an Auditory Information Display in order to ensure a user centred design.
3) Development of human-computer interfaces for the visually impaired and the blind providing similar efficiency than common visual interfaces for non-handicapped users.
4) Making assistive technology more affordable to provide the people concerned with access by similar costs like for the non handicapped user.
5) Strengthening the role assistive technology plays in the common software market by supporting existing standards and a wide range of applications.

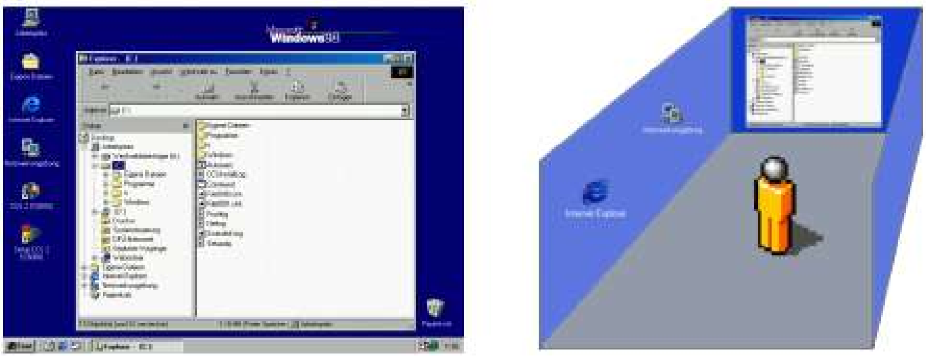

Figure 1

Exploiting the most recent developments in audio engineering the project vision is to introduce spatial dimensions into auditory human-computer interaction in order to make it an equivalent replacement for visual displays or being a powerful extension to the visual mode.

Figure 1 illustrates a possible approach for a common desktop replacement and is intended to provide an idea of how such displays could "sound" in a virtual audio environment.

The major advantage of this concept is the independence of the actual device size because of its virtual nature and, in consequence, its large interface area (making complex and rich displays possible). It is expected to reach a similar standard of performance as conventional interfaces by employing high definition audio rendering techniques, intelligent information modelling, advanced mapping methods and usability engineering.

State-of-the-art

The currently available access for blind and visually impaired people to computers is neither satisfactory in terms of efficiency nor in terms of cost. The two most commonly used technologies are Braille-lines and screen-readers – both are sequential techniques "reading" computer screen texts out to the user. On top of the obvious disadvantage of lower performance compared to the parallel working visual interfaces, these techniques impose a significant additional learning effort on users. In the case of Braille-lines 20% of all blind people were not taught Braille and cannot use this interface at all. In order to achieve better performance with screen-readers blind people are trained to increase the speaker's speed to a level up to 3.5 times faster than normal making it not understandable for untrained persons [1].

An early attempt to build advanced audio interfaces for the visually impaired and the blind was made by the project Mercator [2]. It was begun in 1991 by the Multimedia Computing Group at the Graphics, Visualisation & Usability Centre at Georgia Tech with the aim to provide access to X Windows applications. Due to the complexity of the X client-server structure it became clear that the aim to provide access to most of the commonly used X applications is hard to achieve and finally, Mercator was stopped. More recently, windowing systems provide accessibility interfaces for exposing information about the user interface, making it possible to build alternative displays. Microsoft Active Accessibility is a technology which provides accessibility to applications running under the Microsoft Windows operating system. The AT-SPI (Assistive Technology Service Provider Interface) was developed by the Gnome Accessibility Project [3] and is implemented in the Gnome Desktop – a windowing system running on Linux, IRIX (SGI) and other Unix operating systems – and will also be integrated into the KDE desktop. Both accessibility interfaces are mainly used for text to speech user interfaces (Gnopernicus), but provide more information suitable for more advanced user interfaces.

The development of audio interfaces has seen very different approaches and lead to solutions very specific to the problem domain. Research was conducted in how specific parts of a user interface should be transformed [4, 5], but the scientific field lacks a generic approach in which the scope is not restricted to a specific problem [6, 7].

With investigations in information modelling and classification schemes SoundVision will contribute to find more generic solutions in order to make its concepts and techniques reusable in other applications. Virtual audio reality and the simulation of real sound environments are techniques available for some time in theory, but only recent developments in signal processing methods and the increasing computational power made the first real time software solutions possible [8]. Here, improvements in definition, flexibility and efficiency are essential for their use in audio interfaces.

Enhancement of the state-of-the-art

With three leading audio research institutes combining their expertise in different sound reproduction techniques, SoundVision establishes a network of expertise with the aim to find solutions incorporating the existing advantages and developing new solutions [6,9]. In collaboration with new multi-media standards like MPEG-4 we have sophisticated techniques at our hands that are able to improve the auditory interaction mode significantly.

Sound Vision aims to lift the state-of-the-art in computer interfaces to a next, advanced auditory level. In order to reach this ambitious goal significant enhancements have to be achieved in the following domains:

Virtual Audio Engineering:

To make virtual audio environments (VAE) available as interactive interfaces for any kind of content it is indispensable to have high definition audio rendering. Rendering techniques like Higher Order Ambisonics (HOA), Wave Field Synthesis (WFS) and Vector Based Amplitude Panning (VBAP) will be investigated and a solution matching the requirements of SoundVision will be developed. To achieve reusability the audio renderer will be driven by an MPEG-4 stream. The MPEG-4 standard supports scene description, interaction and 3D audio rendering and promises to be the up-coming multimedia standard.

Besides the VAE system, audio engineering is also needed for the process of auditory depiction. Finding efficient representations of information in the auditory domain requires a deep understanding of how the human auditory perception functions. Psychoacoustic research will show which effects may contribute to the optimisation of auditory display concepts.

Information Modelling:

SoundVision will identify the information entities to be depicted for representing a user interface, and investigates whether certain properties of this information may be candidates for classification schemes considering the depiction on an auditory display. Such classifications will be the basis for generic frameworks defining rules for the transformation of information or interaction entities into the auditory domain. This methodology ensures that concepts developed in SoundVision will be suitable for alternative problem solutions in other fields. Whenever the visual mode is not available or insufficient to present the amount of information these concepts and techniques can be applied to design efficient audio interfaces. Possible applications include miniaturised mobile devices, the observation of complex processes or the sonification of multidimensional datasets. SoundVision concepts might, for example, significantly increase safety and reduce strain for employees in air traffic control tasks.

Human-computer interfaces create a mental model that should present the system task model intuitive and efficient to the user. Investigations in the field of mind (or mental) models and interaction models will provide insight into how interfaces work, also concerning the visually impaired and the blind.

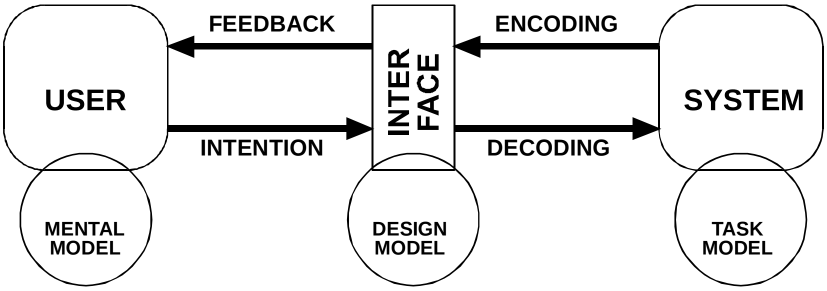

Figure 2

By understanding possible mental models and having a pre-defined task model given by the system the design model of an interface can then be developed. This methodology basically features similar design principles as traditional text or graphics based interfaces, but applied in the auditory domain.

Usability Engineering:

Knowing the user is the first rule for designing a good user interface. The user requirements will determine the goals of development ensuring user centred design. Usability engineering also includes the evaluation of user interfaces. Because of the new approach of the proposed interface also the methodology of evaluation must be developed.

The sensitivity on the auditory system of the blind and the visually impaired, and their experience of mapping an outside world to the audio domain will support research on a generic usage model from which also non-handicapped users will benefit.

With its social partners SoundVision has strong expertise for the determination of the user requirements as well as for conducting large-scale user evaluation. As an important result SoundVision will develop guidelines for interface designers on the basis of the research and evaluation results.

Accessible user interfaces:

SoundVision will combine all research results and will develop a prototype for a human-computer interface for the Windows operating system and the two most important Linux desktop flavours, the Gnome and KDE, to verify the research results on standard hardware platforms with a large user group, building a basis for further product development. Figure 3 illustrates the concept.

All information about the display content is received through the AT-SPI / MS Active Accessibility interface and will be transformed into the auditory domain by the Audio Information Display framework which is a generic framework for any kind of auditory interface. This transformation will result in a scene description of a virtual audio environment (VAE). The scene represents the display content and implements all necessary interface interactions. It will be rendered in the MPEG-4 driven VAE system by sophisticated audio techniques producing output in various formats (binaural, loudspeaker arrays).

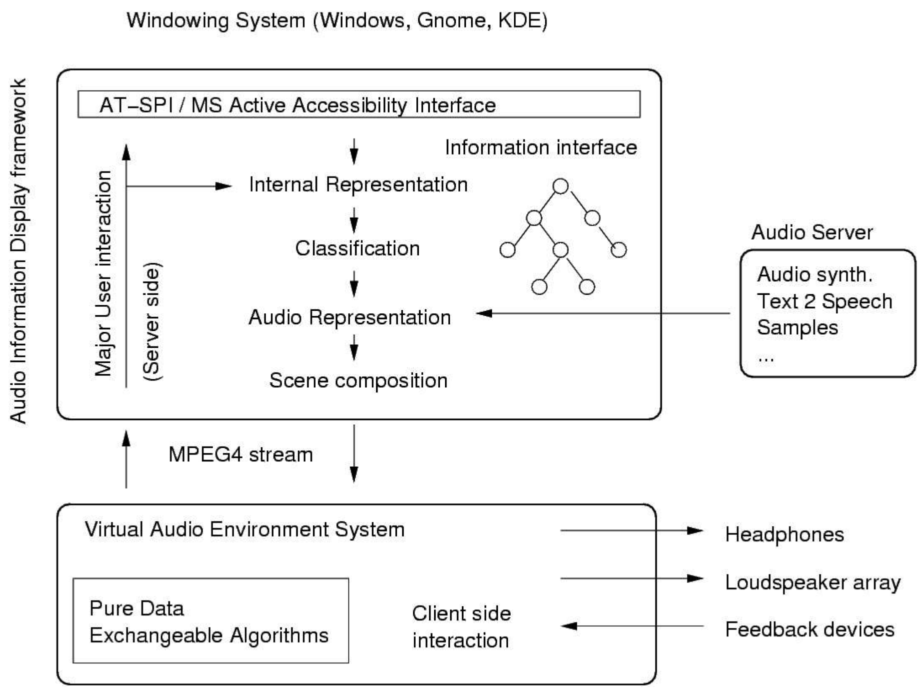

Figure 3

The design will support the reusability of the concept by stressing modularity, so that modules are applicable in different problem domains also.

Social Perspective and Dissemination:

Research in the field of assistive technology has to deal with a minor market and limited financial power within the supporting industry. Thus a major objective is the dissemination of the project results with a high coverage of the target user group and the relevant industry on a

European level.

[1] C. Asakawa, H. Takagi, S. Ino, and T. Ifukube, “Maximum listening speeds for the blind”, International Conference on Auditory Displays, 2003, Boston, USA

[2] Mercator Homepage: http://www.cc.gatech.edu/gvu/multimedia/Mercator.html

[3] Gnome Accessibility Project: http://developer.gnome.org/projects/gap/

[4] R. M. Greenberg M. M. Blattner, D. A. Sumikawa, “Earcons and icons: Their structure and common design principles,” Human-Computer Interaction, vol. 4, no. 1, pp. 11–44, 1989.

[5] G. Weber E. Mynatt, “Nonvisual presentation of graphical user interfaces: Contrasting two approaches,” in Proc. ACM Conference of Computer Human Interactions, Boston, MA, USA, 1994, ACM Conference on Human Factors in Computing Systems (CHI’94), pp. 166–172, ACM press.

[6] C. Frauenberger M. Noisternig, “3d audio interfaces for the blind,” in ICAD Proceedings, Boston, USA, July 6–9 2003, Internation Conference on Auditory Display.

[7] S. Brewster A. Walker, “Spatial audio in small screen device displays,” in IHM-HCI Proceedings, Lille, France, September 10–14 2001, Interaction Homme-Machine – Human Computer Interaction, IHM-HCI, pp. 531–540.

[8] M. Noisternig et.al., “A 3d real time rendering engine for binaural sound reproduction,” in ICAD Proceedings, Boston, USA, July 6–9 2003, Internation Conference on Auditory Display.

[9] T. Lokki et.al., “Creating interactive virtual auditory environments” IEEE Computer Graphics and Applications, special issue Virtual Worlds, Real