WebSound: a generic Web sonification tool allowing HCI researchers to dynamically create new access modalities

Lori Stefano Petrucci[1], Eric Harth[1], Patrick Roth[1], Assimacopoulos André[2], Thierry Pun[1]

(1) Computer Science Department
CUI, University of Geneva
CH – 1211 Geneva 4, Switzerland
+41 22 705 76 32
Lori.Petrucci@cui.unige.ch

(2) UCBA
Swiss Central Union of and for the Blind
Schützengasse 4, CH- 900 St-Gallen,
Switzerland

ABSTRACT

The World Wide Web (WWW) has recently become the main source of digital information accessible everywhere and by everyone. Nevertheless, the inherent visual nature of Internet browsers makes the Web inaccessible to the visually impaired. To solve this problem, non-visual browsers have been developed. One of the new problems, however, with those non-visual browsers is that they often transform the visual content of HTML documents into textual information only, that can be restituted by a text-to-speech converter or a Braille device. The loss of spatial layout, and textual attributes such as boldface, italic, underline, color or even size should be avoided since they often bear visually important information. Moreover, typical non-visual Internet browsers do not allow visually impaired and sighted individuals to easily work together using the same environment. These new problems have to be solved with new alternative non-visual display techniques. This paper presents *WebSound*, a new generic Web sonification tool and its application to a 3D audio augmented Internet browser (Internet Explorer 5.0) developed at the University of Geneva.

Keywords

Non-sound speech, 3D virtual audio environment, earcons, visual impairment, haptic sense

USING AUDIO-HAPTIC INTERFACE TO RESTITUTE VISUAL AND SPATIAL INFORMATION

The visual channel has a tremendous capacity for information transfer. As a result, it is heavily relied upon in normal human-computer interfaces. Internet browsers (for example Internet Explorer or Netscape Navigator) do not escape from this consideration. Most of the information that is presented, however, is not visual in nature, and could easily be presented in an alternate verbal form. This consideration is certainly the basis of non-visual Internet browsers that typically transform the two-dimensional visual representation of HTML documents into a one-dimensional representation, that is simple raw text which can be presented by a text-to-speech converter.

The need to use standard visual Web browsers such as Internet Explorer even for visually impaired leads us to develop new alternative display techniques to restitute the spatially and visually related information that are present in an HTML document. Since the new challenge is to present the spatial relationship between HTML elements (tags) and their visual attributes, it immediately suggests the use of techniques and senses, which are commonly used by blind and visually impaired people to find locations and to perceive their environment. These include both auditory and tactile approaches. One technique is to provide different auditory attributes (earcons / auditory icons) to each different tag. Moreover, the use of a 3D immersive audio environment, which permits to make a sound appear from a given position, may give blind users the sense of object location. This hypothesis which has been validated in our preliminary studies [1,2] have shown that the use of non-speech sound in human-computer interfaces can considerably increase the “bandwidth” of computer output. This is also shown in the work of Stephen A. Brewster [3] or in the studies of Frankie James[4].

To enhance the mental representation of the virtual sonic environment, we propose to add to the hearing sense the haptic modality, that is the sense of where one’s hand or arm is in space. Using a haptic tablet or a touch-sensitive screen would allow an individual to move his/her finger about while at the same time keeping track of his/her position on the device. The system then would respond with auditory feedback while the user moves his/her finger (device pointer) around the virtual screen. This approach, as shown by M. Lumbreras et al [5] validates the hypothesis that a 3D immersive virtual sound environment combined with haptic manipulation of the audio environment can enable blind users to construct a mental representation of the spatial environment (in our case, the spatial layout of HTML documents).

WEBSOUND, ARCHITECTURE AND IMPLEMENTATION

One of the goals of *WebSound* is to provide researchers of HCI working on non-visual Web interfaces, with a tool that allows them to easily add new accessible functionalities to a standard visual browser. To reach this goal, the design and implementation of a generic Web sonification tool (*WebSound*) has been done on the basis of two fundamental ideas:

- use of the same augmented visual browser, based on the standard Internet Explorer 5.0 which will be used both by "Internet surfers" (visually impaired and sighted users) and by programmers of new behavioral models;
- creation of an add-on tool (the *Workspace*) allowing to dynamically and visually create new sonic models.

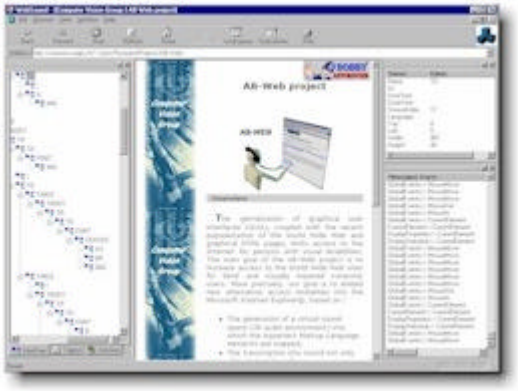
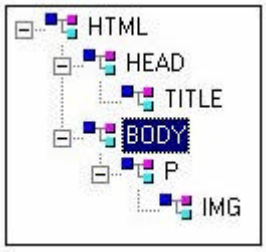


Figure 1. Snapshot of the WebSound interface




```
<HTML>
<HEAD>
<TITLE>home page</TITLE>
</HEAD>
<BODY>
<P>
<IMG SRC="images/logo.gif">
</P>
</BODY>
</HTML>
```

Figure 2. hierarchical (a) and textual representation (b)

Two of the most difficult problems to solve have been firstly to find a way to obtain all the internal events that occur in Internet Explorer, in order for example to determine the device pointer's position while moving on the browser and secondly to determine which HTML tag is pointed on. The first problem has been solved using some possibilities that Microsoft offers to get informed of the internal events that occurs in a Windows based software or in an ActiveX such as Internet Explorer. The technical details of this solution are omitted here. To solve the second problem, we need to access the hierarchical structure of HTML documents (Fig. 2) in order to create an off-screen structure, called *HtmlNode*, of each HTML element that can receive and send messages. For example, if the user moves his/her finger across an image present in the HTML document, the *HtmlNode* associated with the image tag will send messages of type *MouseIn*, *MouseMove* (repeated while the device is inside the image border) and *MouseOut* to the WebSound application.

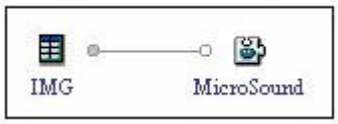


Figure 3. An example of a dynamic sonic model

Those messages have no behavior until they are connected, via the *Workspace*, to entities called *Services* (Fig. 3). A *Service* is a piece of code that performs a particular task regarding of the message type it receive. For example, if the programmer wants to play a sound when the device pointer moves over an *HtmlNode*, he needs to create a *Service* (*MicroSound* in the fig. 3) which can take a sound and play it when he receives a *MouseMove* message from the *HtmlNode* associated with the image.

This mechanism can be extended to any accessibility modalities such as reading the HTML document using keyboard, navigating inside HTML tables, associating force-feedback behavior to elements, etc.

CONCLUSION

In this article, we describe a new tool (*WebSound*) that allows researchers to dynamically create alternative access modalities to a standard Web browser (Internet Explorer 5.0). It also describes a possible way to permit visually impaired individuals to explorer spatial information.

ACKNOWLEDGMENTS

This project is financed by the Swiss Priority Program in Information and Communication Structure and by the Swiss Central Union Of and For the Blind. The authors are grateful to A. Barrillier, V. Hintermeister and V. Rossier for their help in the design and evaluation of the prototype.

REFERENCES

1. Petrucci L., Roth P., Assimacopoulos A., Pun T. *An Audio Browser for Increasing Access to World Wide Web Sites for Blind and Visually Impaired Computer Users.* HCI'99, Munich, pp. 995-998.
2. Harth E., *Plate-forme de Sonification Dynamique de Pages Web Pour Aveugles et Malvyants.* Diploma thesis, University of Geneva, 1999.
3. Brewster S.A. *Using Non-Speech Sounds to provide Navigation Cues.* ACM transaction on Computer-Human Interaction 5(3), pp. 224-259.
4. James F. *Presenting HTML Structure in Audio: User Satisfaction with Audio Hypertext.* ICAD'96 Proceedings. Xerox PARC, 4-6 Nov. 1996, pp. 97-103.
5. Lumbreras M., J. Sanchez. *Interactive 3D Sound Hyperstories for Blind Children*. CHI'99 Pittsburgh PA, USA.