Developing Multi-modal Interfaces for Visually Impaired People to Access the Internet

Wai Yu, Graham McAllister, Emma Murphy, Ravi Kuber, Philip Strain

Queen's University of Belfast
Northern Ireland
{W.Yu, G.Mcallister, E.Murphy, R.Kuber, P.Strain}@qub.ac.uk

Abstract

This paper describes the work being carried out at the Queen's University of Belfast (QUB) on improving visually impaired people's access to information on the Internet. In particular, the project is focused on problems that visually impaired people have on navigating and reading information from Web pages. These problems will be addressed by using a multi-modal approach of combining visual, audio and haptic technologies. The first prototype of these interfaces has been developed based on the results of the user requirements capture conducted with visually impaired people. This paper will present a review of the current technology to assist visually impaired people to access the Internet and also the users' comments on this technology. As a result of the user feedback, the paper will present a prototype multi-modal interface and discuss the issues that should be considered when designing such interfaces.

1 Introduction

The Internet has become increasingly prominent due to the vast amounts of information and array of services that is available. However, its benefits to visually impaired people, especially blind people, are not as much as to sighted people. This is mainly due to two factors: (1) the poor accessibility of the Web content, and (2) the limitations of the currently existing assistive technology.

Most Web pages are graphically oriented with different combinations of fonts, colour, frames and tables which can cause severe difficulties for visually impaired users to access this information. Although the W3C's Web Content Accessibility Guidelines provide information about how to make Web content accessible for disabled people, research has shown that they are insufficient to ensure Web pages to be truly accessible even though they have met the guidelines requirements [Di Blas 04] [Brajnik 04]. In a recently study conducted by Helen Petrie and her colleagues [DRC 04], some shocking findings have been produced on the current status of Web accessibility. In this comprehensive survey, 1000 websites in the UK were tested by using both automatic checking and manual checking. The results show that no home pages passed combined Priority 1, 2 and 3 checks specified in the W3C guidelines. Only 0.6% of home pages passed both Priority 1 & 2 automatic checks and 0.2% passed both automatic and manual checks at Priority 1 & 2 levels. Only 19% of home pages passed the automatic Priority check 1. Furthermore, in the in-depth testing of 100 websites, besides the automated testing, 50 blind and partially sighted people were asked to perform a series of tasks on the websites. In the results, the successful rate of blind people is only at 53% and the successful rate of partially sighted people is at 76%. In the test, 585 different problems were encountered by the participants. Only 55% of the problems are related to the checkpoints specified in the guidelines and the other 45% are not. This shows that the guidelines are insufficient to ensure Web pages to be really usable by visually impaired people. The good news from the study is that accessibility also means usability. A Website that is accessible is also usable.

In addition to the problems caused by the inaccessibility of Web content, the assistive technologies that visually impaired people use for accessing the Internet is also placing hindering limitations. To access information on Web pages, visually impaired people mainly use screen readers, screen magnifiers, and keyboards with or without Braille display. First of all, screen readers cannot handle the graphical information on Web pages if no alternative text message is attached to the image. Secondly, screen readers handle textual information on Web pages by reading them out in a sequential order. It could take a long time for the screen readers to reach the information required by the user. Web pages developed according the W3C's guidelines would help the content to be usable by the screen

readers. However, the problems associated with screen readers would remain. Therefore an enhanced tool/interface is needed to address these problems and improve visually impaired people's access.

WebSpeak [Hakkinen 96] and BrookesTalk [Zajicek 98] have been developed to output text-to-speech as well as providing an overview of Web content for the user. BrookesTalk also summarises and abridges text on a site in an attempt to recreate the way that a sighted user would scan a Web page. Besides synthesized speech, non-speech audio has been used to improve Web accessibility. Goose et al. [Goose 99] and Donker [Donker 02] have created audio browsers based on 3D sound specialization to convey information on Web pages. Susini et al. have attempted to create Sonified Hyperlinks analogous to visual hyperlinks in HTML [Susini 02]. Audio Enriched Links software has been developed to provide previews of linked Web pages to users with visual impairments [Parente 03]. WebSound [Roth 00, Pentrucci 00] is a sonification tool where sonic objects are associated with HTML tags, which are then projected into a virtual 3D sound space according to the finger position on a tactile touch screen.

Research has shown that advantage can be gained when the haptic modality is used in conjunction with the visual and auditory channels [Yu 02]. The IFeelPixel multi-modal application [IFeelPixel] enables the user to mediate structures such as edges, lines and textures, via a tactile/force feedback mouse. Multi-modal solutions have the capacity to extend visual displays and thus making objects more realistic, useful and engaging [Brewster 01]. Multi-modal interfaces also provide assistance in the mental mapping process, allowing the user to develop a greater awareness of objects contained within the environment. A clearer spatial representation created by multi-modal feedback, enhances the user's ability to orientate and navigate within their environment [Lahav 02, Caffrey 04]. It is apparent that a multi-modal assistive interface can provide one solution to reducing barriers that are faced by the blind and partially sighted community.

2 QUB's Multimodal Approach

At QUB, research is being conducted to address the problems of blind peoples' access to the Internet in three main areas: (1) navigation on the Web, (2) representation of information, and (3) accessing graphical content. To achieve the objectives of the project, visual, audio and haptic technology are used to develop the multi-modal interface. Issues of multi-modal interaction and defining the role of each modality according to the task are being researched with an emphasis on the assistive technology. An overview of the QUB approach can be seen in Figure 1.

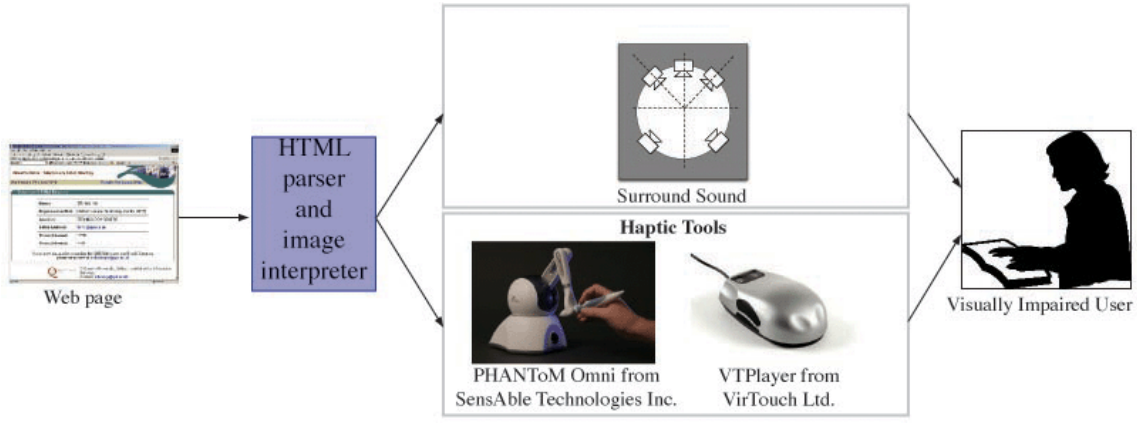

Figure 1. Overview of the multi-modal approach at QUB.

3 User Requirements

In order to develop truly usable interfaces for visually impaired people, a user-centric design approach is adopted. Current problems, users' requirements and expectations have to be taken into account in the design and development.

A series of user interviews has been conducted in order to find out the current situation of visually impaired people's use of the Internet and problems in accessing the Web. A questionnaire has been designed for the user interviews and consists of questions concerning with the following issues:

- Technical Experience
- Internet Usage
- Navigation
- Perception of graphics and images
- Strategies and reading patterns and Information Content
- Accessibility & Usability Issues
- Perceptions of a New System
- Reasons for not accessing the Internet/ Owning a Computer

Besides identifying the current problems faced by visually impaired people, the questionnaire was also designed to survey their technical competence and experience in using the Internet. Therefore, the section on Technical Experience contains questions on the level of familiarity that people have on computers. Then in the section on Internet Usage, people were asked to state how often they use the Internet and on what purpose. If people being interviewed did not use the Internet, they would be led to the last section to find out the reasons in an attempt to identify the barriers that prevent people from using the Internet. Otherwise, they would go through the other five sections in which questions would be asked to determine the issues of accessing various forms of information on the Web. In the section on Perception of graphics and images, different sets of questions were designed to accommodate blind people and partially sighted people. The user interview will usually take one hour to complete for the Internet users.

In the current stage of the user interviews, 28 blind and visually impaired people have been interviewed. 18 of them are male and 10 of them are female. The age range is between 21 and 59 years old. Table 1 shows their condition of visual impairment, whereas Table 2 shows their experience in using the Internet.

Table 1. Distribution of Sight Loss.

Status of Sight Loss	Number of People
Congenitally Blind	12
Adventitiously Blind	15
Partially Sighted	1

Table 2. Distribution of Internet Experience.

Experience in using Internet	Number of People
No Experience	1
Beginner	2
Intermediate	8
Expert	17

From the data collected from the questionnaire, significant differences can be found between expert users and beginner and intermediate users. Expert users can do most of the tasks that sighted users would normally do on the Web such as shopping, banking, chatting and emailing. They are familiar with the use of assistive technologies and able to modify the settings of screen readers to suit their needs and preferences. They can access information faster than beginners and identify problems as they occur. They have developed strategies to form an overview of a page and scanning the page for useful information. The common problems encountered by expert users include:

- Flash
- Javascript
- Graphics & Images
- Embedded files
- Web forms

Some Javascripts i.e. on a news site would automatically refresh the page which as a result bringing the screen reader's current reading postion back to the top of the screen and thus confuse the users. Graphics and Images would become a problem when alternative text is not informative enough. Expert users would prefer short and succinct alternative text. Some files such as Word and PDF that embedded in the page are not always obvious to the users. PDF are not accessible using assistive devices due to poor tagging. Web forms could become a problem if they were poorly labelled. Expert users would need to move screen readers' Forms mode back to Non-Forms mode to view each label.

As expected, beginner users encounter more problems than expert users. There are also differences between these two groups of users in terms of their usage of internet and attitude towards the technology. Beginner users have very clear objectives in using the Internet such as sending emails, getting the bus timetables, reading news from etc. It is very task specific and they do not really grasp the meaning of "surfing the Web". They tend not to purchase products or tickets online due to three reasons: (1) Difficulty in interacting with Web forms using assistive technologies; (2) Fear that they might do something wrong; (3) Security concerns about giving out the credit card number.

Navigation turned out as a frustrating experience due to the lack of feedback. They usually explore the Web in a linear fashion from top to bottom therefore all Web pages appear to be in a single column to them. They have worries about getting lost in navigation between pages. If they become lost within a page or over a number of pages, they tend to close the application. In the worst case scenario, they may shut down the computer. Therefore, they would keep only one Window open at one time to reduce confusion. They also have concerns over downloading files due to (1) Security, may download a virus by mistake; (2) Lack of training to download and interpret the file; (3) Subtle fear that the system settings may be altered, which prevents them from using the system as usual.

Common comments from all groups of users show that they appreciate more feedback on the spatial layout of a page. Assistive devices that can provide additional guidance and assistance would be welcome. They are interested in the technologies that allow them to work with sighted users so that they can show where they are on a page when asking for help or for work purpose. Training is another issue needs to be addressed. In our interviews, visually impaired people pointed out the lack of IT training in their local area. Shortage of trainers who are familiar with the assistive technology and expensive cost in attending trainings seems to be the major causes.

4 Design Considerations

Based on the information gathered so far, the following points have been identified as crucial features for the multi-modal interface. There are 2 major requirements and 3 desirable features.

Major requirements:

- Ability to accommodate different types of visual impairments; different modalities can be selected to suit different user needs, such as audio and haptics for blind people and visual and haptics for partially sighted.
- Compatibility with current screen readers which are familiar to visually impaired users. This will ensure that the new interface should be able to work alongside with the existing screen readers instead of replacing them.

Desirable features:

- To provide an overview of Web pages, this is a main feature missing from screen readers. Some existing work can provide an abstract of the Web page [Zajicek 04], however, an overview on the layout Web pages is also required.
- To enhance users' short-term memory; this would be useful on a busy Web page. This feature allows users to record some important information and its location on the Web page so that users can retrieve this information and return to the same place later.
- To provide active guidance using haptics; this is extremely useful when users need to go to a particular place on a Web page quickly. The force feedback & tactile device used in the project can direct the user to the desired location.

5 First Prototype

The first prototype of the multi-modal interfaces is currently being developed. Multi-modal approaches are recognized as the most effective way to overcome the problems that visually impaired people have in accessing information. This is due to the ability to maximize the benefits of each modality and then further enhance the accessibility by facilitating interaction between the modalities.

5.1 CONTENT-AWARE PLUG-IN

An assistive tool which consists of a multi-modal interface with haptic and audio feedback and a content-aware Web browser plug-in, which allows users to actively navigate on a Web page has been developed (Figure 2).

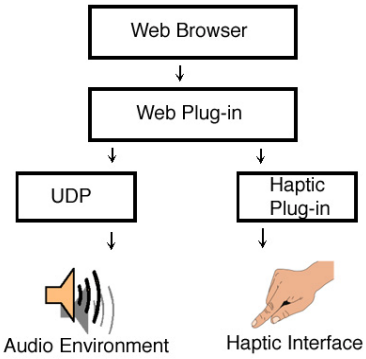

Figure 2. System Architecture.

The main purpose of this approach is to give users an opportunity to explore a Web page's layout through active haptic interaction. A force feedback mouse is used and its cursor position is constantly monitored by the plug-in which detects the surrounding objects. If the plug-in finds an object nearby then it will inform the user by enabling the haptic and audio features. Depending on the user's intention and the context of the task, appropriate prompts can be given, such as giving users guidance to the desired destination or informing users about nearby objects, such as images and hyperlinks.
The open-source Firefox browser has been selected to work with. The browser is based on Mozilla, which fully implements the W3C standards in the Web Content Accessibility Guidelines 1.0. Extensions can be developed that can be utilised to work in conjunction with the browser.

5.2 Haptics

The Logitech Wingman Force-Feedback mouse (Figure 3) has been chosen to facilitate on-screen navigation, due to its force feedback capability and compatibility with the Firefox browser. The Immersion Web plug-in has been linked to the content-aware Web plug-in. Supporting software can model a small array of haptic effects including stiffness, damping, and various textures, which can then be called through the Javascript. This has facilitated the exploration of objects with additional force-feedback.

Figure 3. Logitech Wingman Force-Feedback Mouse

The following haptic primitives have been employed; the ‘enclosure effect’ has been coupled with clipping effects bordering the image. This has given the illusion of a ridge, which needs to be mounted. Cursor clipping motion increases a user's psychological perception of the wall's stiffness. Upon rolling over the image, a ‘buzz effect’ has been produced along with force-feedback. The dual effect of audio coupled with force-feedback, will heighten the sense of awareness that the user is directly on the image. To provide location awareness of the cursor when directly hovering over a hyperlink, the ‘periodic effect’ has been used. This effect produces a wave that varies over time, promoting a locked sensation, when directly hovering over the link. It is intended that this will promote a sense of orientation within a page for the visually-impaired user.

5.3 Real-time Audio

The audio feedback for this system gives the user a sense of navigation in relation to an image or a link on the page. The audio is designed and played back in Max/MSP, a real-time audio programming environment. Netsend, an MSP external object is used to receive x and y location co-ordinates sent via UDP from the Web plug-in. Figure 4 illustrates how the element is divided up, and the range of coordinates that are associated with each section. As the user rolls over an image or a link with the force-feedback mouse, a sound is played to reinforce the haptic response. These sounds can be classified as auditory icons as defined by Gaver [ ] as sounds sampled from real world environments based on the attributes of source events such as size, material, and force. For this system the image sound icon is a short descriptive auditory clip of a camera shutter clicking, suggesting a photograph or image. The auditory icon used to depict a link is a short “metallic clinking” sound suggesting the sound of one link in a chain hitting off another.

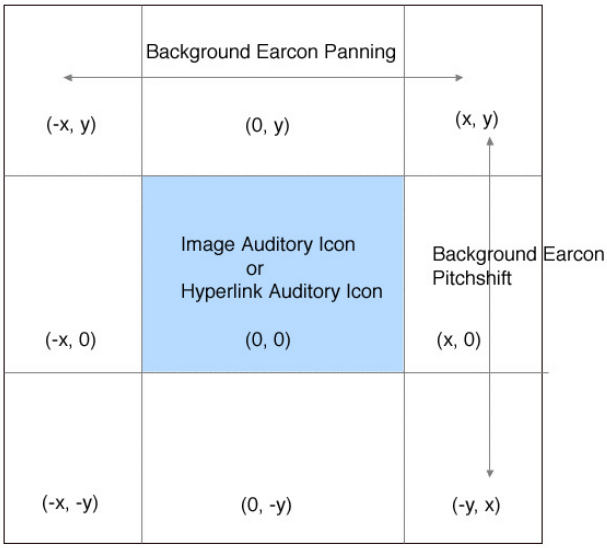

Figure 4. Object Co-ordinates and Audio Feedback

Outside the image or link space the cursor location is mapped to panning and pitch-shift parameters of a continuous background sound. The x-value co-ordinates are mapped to a panning patch in Max/MSP so that as the user moves the cursor along the x-axis the audio is panned to that position. Similarly the pitch varies according to the position on the y-axis; as the user moves the cursor upwards, the background sound is pitch-shifted upwards to represent this movement. This background location sound is based on musical parameters rather than a real world sound source and is therefore classified as an earcon [Brewster 95].

6 Conclusion & Future Work

This paper has given a review of the current situation of blind and visually impaired people’s access to the information on the Internet. Data collected from the user interviews have shown the problems encountered by visually impaired people. Based on such information, a multi-modal interface has been proposed with the design

consideration taken into account. The first prototype with a novel approach which can determine when a user's cursor is in close proximity to a region of interest on a Web page, e.g., a hyperlink or image has been described. By rendering the spatial visual information using multi-modal interfaces, users are not only informed of these regions of interest, but are also guided to them using audio and haptics. This technique also promotes the use of novel interactive devices to increase spatial awareness, allowing users to construct a mental spatial map of a Web page.

The content aware Web plug-in will be extended to include more features to assist visually impaired Internet users. The system will be developed with a view to reducing cognitive workload associated with the need to remember structural and contextual information for Web pages. The content-aware system will be adapted so that it will function with other browsers. The type of information that can be sent to haptic and audio environments will be enhanced to allow improved feedback to convey more complex information. A component will be designed that will summarise and interpret sections of a web page.

In order to overcome the problem of screen-readers outputting extraneous information from a web page, a haptic feature will be developed to automatically direct users to certain sections of the content. The path taken through a page could be customised by the user, allowing time to be saved by skipping over information that is not useful to the individual.

One of the most important considerations in designing audio for a non-visual interface is that the auditory space overloads much faster than the visual space. Initial user testing has highlighted the fact that too many sounds presented simultaneously will confuse the user. Short auditory clips are preferable to convey simple information quickly. Continuous sounds should be used sparingly and a hierarchy of significance of each sound should be determined before it is used.

The audio and haptic interface designed for this plug-in is an example of the ways that multi-modality can be used to enhance a visually impaired user's awareness of navigational position and location of objects on a page. Following a further survey and analysis which is specifically targeting different needs of expert and beginner users, a more comprehensive system will be developed.

7 Acknowledgements

The work described in the paper is supported by Eduserv.

References

[Brajnik 04] Brajnik G., Achieving Universal Web Access through Specialized User Interfaces. 8th ERCIM Workshop "User Interfaces for All", 28-29 June 2004, Vienna, Austria.

[Brewster 95] Brewster S., Wright P. and Edwards A., Experimentally Derived Guidelines for the Creation of Earcons, Adjunct Proceedings of the British Computer Society Conference on Human Computer Interaction, 1995, 155-159.

[Brewster 01] Brewster S. A., The impact of haptic 'touching' technology on cultural applications. In *proceedings of EVA'01*, (2001), 1–14.

[Caffrey 04] Caffrey A. & McCrindle R. Developing a multi-modal web application. *In proceedings of ICDVRAT'04* (2004)

[Di Blas 04] Di Blas N., Paolini P., and Speront M., Usable Accessibility to the Web for blind users. 8th ERCIM Workshop "User Interfaces for All", 28-29 June 2004, Vienna, Austria.

[Donker 02] Donker, H., Klante, P. & Gorny, P., "The design of auditory user interfaces for blind users", Proceedings of the second Nordic conference on Human-computer interaction, 2002, pp. 149-156

[DRC 04] The Web Access and Inclusion for Disabled People, A Formal Investigation conducted by the Disability Rights Commission. http://www.drc-gb.org/publicationsandreports/report.asp

[Gaver] Gaver, W. W., Smith, R., & O'Shea, T., Effective sounds in complex systems: the ARKola simulation. Proceedings of the SIGCHI conference on Human factors in computing systems, New Orleans, Louisiana, USA, 1991, pp.85-90

[Goose 99] Goose, S. & Moller, C., "A 3D audio only interactive Web browser: using spatialization to convey hypermedia document structure", Proceedings of the seventh ACM international conference on Multimedia, 1999, pp. 363-371

[Hakkinen 96] Hakkinen, M & Dewitt J., WebSpeak:User Interface Design of an Accessible Web Browser. *White Paper, The Productivity Works, Inc., Trenton, New Jersey, USA*, (1996)

[IFeelPixel] IFeelPixel Features, http://www.ifeelpixel.com/faq/#whatitwill

[Lahav 02] Lahav O. & Mioduser D. Multisensory virtual environment for supporting blind persons' acquisition of spatial cognitive mapping, orientation and mobility skills. *In proceedings of ICDVRAT'02,* (2002)

[Parente 03] Parente, P. & Bishop, G., "BATS: The Blind Audio Tactile Mapping System", 41st Annual ACM Southeast Conference, 2003

[Pentrucci 00] Petrucci, L.S., Harth, E., Roth, P. & Assimacopoulos, André , Pun, Thierry, "WebSound: a Generic Web Sonification Tool, and its Application to an Auditory Web Browser for Blind and Visually Impaired Users ", Proceedings of the 2000 International Conference on Auditory Display

[Roth 00] Roth, P., Lori Petrucci, Assimacopoulos, André & Pun, T., "Audio-Haptic Internet Browser and Associated Tools for Blind Users and Visually Impaired Computer Users", COST 254 Intelligent Terminals, Workshop on Friendly Exchanging Through the Net, 2000, pp. 57-62

[Susini 02] Susini, P., Vieillard, S., Deruty, E., Smith, B. & Marin, C, "Sound Navigation: Sonified Hyperlinks", Proceedings of the 2002 International Conference on Auditory Display

[Yu 02] Yu W., Reid D. & Brewster S.A. Web-based Multi-modal Graphs for Visually Impaired People. In *Proceedings of CWUAAT'02*, (2002), 97-108

[Zajicek 04] Zajicek M., Powell C. and Reeves C., A Web Navigation Tool for the Blind, 3rd ACM/SIGAPH on Assistive Technologies, 15-17 April 1998, California, USA.

[Zajicek 98] Zajicek, M., Powell C., and Reeves C. A web navigation tool for the blind, *In Proceedings of ASSETS '98*, (1998), 204-206