Pre-print: Strain, P., McAllister, G., Murphy, E., Kuber, R. & Yu, W. 2007: A Grid-Based Extension to an Assistive Multimodal Interface. In Extended Abstracts of on Human Factors in Computing Systems - CHI'07, San Jose, USA, 2675-2680

A Grid-Based Extension to an Assistive Multimodal Interface

Philip Strain
Sonic Arts Research Centre
Queens University, Belfast
Belfast, BT7 1NN, UK
p.strain@qub.ac.uk

Graham McAllister
Sonic Arts Research Centre
Queens University, Belfast
Belfast, BT7 1NN, UK
g.mcallister@qub.ac.uk

Emma Murphy
Sonic Arts Research Centre
Queens University, Belfast
Belfast, BT7 1NN, UK
e.murphy@qub.ac.uk

Ravi Kuber
Sonic Arts Research Centre
Queens University, Belfast
Belfast, BT7 1NN, UK
r.kuber@qub.ac.uk

Wai Yu
Virtual Engineering Centre
Queens University, Belfast
Belfast, BT7 1NN, UK
w.yu@qub.ac.uk

CHI 2007, April 28 – May 3, 2007, San Jose, USA
ACM 978-1-59593-642-4/07/0004.

Abstract
This paper describes an extension to a multimodal system designed to improve Internet accessibility for the visually impaired. Here we discuss the novel application of a grid (patent pending) to our assistive web interface. Findings from our evaluation have shown that the grid enhances interaction by improving the user's positional awareness when exploring a web page.

Keywords
Assistive technology, accessibility, navigation, auditory & haptic interfaces.

ACM Classification Keywords
H.5.2 User Interfaces: auditory feedback & haptic I/O; H.5.4 Hypertext/Hypermedia: navigation.

Introduction
The Assistive Technology Act of 1998 [1] defines an assistive technology device as "any item, piece of equipment, or product system, whether acquired commercially, modified or customized, that is used to increase, maintain, or improve functional capabilities of individuals with disabilities". Most visually impaired persons use screen readers to access the Internet. However, Di Blas [2] has shown that the linear model employed by current screen readers has inherent

limitations. Mynatt [7] has identified four models of screen reader implementation – linear, spatial, hierarchical and conversational. Linear, hierarchical and conversational models are limited, as the lack of spatial information provided makes communication between visually impaired and sighted users difficult. As the Web is primarily a visual medium, developers make use of spatial layout to enhance the attractiveness and usability of websites. In addition, Rourke et al [11] have pointed out that some web pages refer to spatial directions in their help pages, for example "go to the link at the top right hand corner of the page".

A requirements capture with thirty blind and partially sighted people, was undertaken to investigate the problems that visually impaired Internet users face [6]. Participants indicated that they would be like to be made more aware of the position of images and links on a page, to develop a mental model similar to the visual depiction of the page. This would make collaboration with sighted users easier and also enhance their browsing experience. Additionally, visually impaired web developers stated that they would find it useful to be able to conceptualise their designs as perceived by sighted users.

Previous Work

An initial prototype of a multimodal system to utilize the spatial model when accessing the Internet has been developed [12]. The system enables the user to gain a mental model of the overall spatial layout of the web page via free exploration. When the page loads, the absolute position of each HTML element is calculated, along with relevant information, such as the element height and width. A semantic value is also determined for each element. For example, the alt attribute value is used to communicate the information contained within an image. As the user moves the mouse, the relative position of each element in relation to the mouse cursor is calculated. This enables the HTML element nearest to the mouse cursor to be determined. If the mouse cursor is in the vicinity of the element (defined as being within 50 pixels), the user is given multimodal feedback as to the position of the element in relation to the cursor. As the user moves over the element, additional multimodal feedback is presented.

The system consists of four main components which interact to form the system – a web plug-in, haptics, speech audio and non-speech audio. A force-feedback mouse has been used, providing the user with a variety of haptic effects when accessing the interface. Non-speech audio feedback has been designed and rendered in Pure Data, a real-time audio programming environment. Speech is provided via the development of a plug-in utilizing the Microsoft Speech SDK. All of the components are co-ordinated via a Firefox extension which parses and analyzes the webpage, then presents pertinent information to the relevant modalities.

An extensive user evaluation has revealed that users were able to develop a spatial representation of a web page when using the system. However, some users were confused due to locational awareness issues. The physical distance moved by the mouse did not correspond to the distance moved by the mouse cursor on the screen. As a result, users were unaware of their position in the context of the page. One participant suggested additional feedback to provide awareness of the mouse cursor on the web page would be beneficial.

It is hypothesised that the addition of a new layer to the system to convey positional information to the user through the use of a grid would enhance the usability of the multimodal system.

Current Grid Systems

The use of a grid to enhance locational awareness for the visually impaired has been explored in both academic and non-academic contexts.

In 1947 a BBC commentator employed a grid representation to convey the current location of the ball on the pitch in a game of soccer – the pitch was segmented into 9 "cells", each cell being labelled with a number. Current real-world spatial mobility training methods for visually impaired adults and children are often based on grid-based game metaphors. For example, real world spatial routes and concepts are often described in terms of chessboard movements because tactile chess is an accessible and familiar game for visually impaired users.[1]

Systems that make use of a grid for navigation have been researched by [3], [4] and [13]. iSonic [13] makes use of a grid to enhance the navigation of an interactive sonification of large datasets, to enable access to Geographic Information Systems. The use of a grid to enhance exploration of auditory pattern perception has been investigated via the use of the GUESS system [4]. The IC2D [3] system enables visually impaired users to "draw" and interpret shapes and diagrams. Both the iSonic and IC2D systems use a 3x3 "recursive" grid– the grid is recursive in that each grid cell can contain another grid cell, enabling the user to "zoom" into the interface. Also each system enables a one-to-one mapping between the numeric keypad and the grid structure, and in both cases evaluations shown that this was easy for visually impaired users to conceptualise.

Some speech recognition software (e.g. Dragon Naturally Speaking) enables sighted users to zoom into any part of an interface by employing a recursive grid.

Extended Grid System

It was hypothesised that by dynamically overlaying a grid over the web page, the user would be able to develop a better sense of positional awareness. A grid created via DHTML, consisting of 9 rectangular squares proportional to the viewable area of the browser, is overlaid on the web page and rendered with multimodal feedback. The layered approach [5] enables existing multimodal feedback (e.g. representing images and hyperlinks) to be accessed by the user.

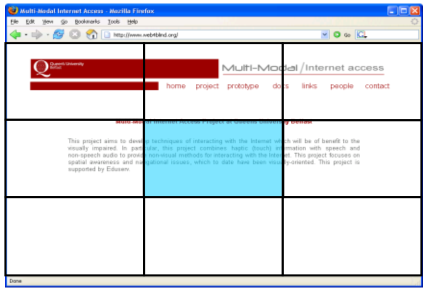

Figure 1. Grid overlaid on web interface

1. Interview with trainer at St. Joseph's School for the Blind, Dublin, Ireland. 28th March, 2006

Audio for grid mode

For this prototype, speech feedback was implemented as a means for conveying cell position. Although it is intended that this information will be conveyed through non-speech sound in future versions of the interface, it was considered that non-speech feedback could distort initial testing of the effectiveness of the grid layer. Currently, the user's mouse position on the grid is conveyed via speech in a direct mapping between the square and its position in the context of the entire grid, for example "top left", "bottom right" or "middle right". The user's position in the grid is conveyed each time he/she moves from one grid cell to another.

Haptics for grid mode

As the user explores the grid structure, a distinctive periodic wave coupled with a pulse effect is played to communicate the location of the borders surrounding each cell.

Evaluation

An evaluation of the system with the grid has been carried out in a collaborative context with fourteen visually impaired users. An experiment was designed to evaluate whether blind and sighted peers were able to work together to reach a common goal.

A help-desk scenario was introduced, and users were asked to perform two web-based tasks. Both tasks required each participant to explore a webpage using the multimodal interface discussed to search and locate a specific element, suggested by a sighted user who was playing the role of a telephone based customer services advisor. One task was performed with the use of the grid, the other without the grid.

Results from the evaluation have shown that the grid was thought to be an effective method of conveying location within the interface by the majority of participants. Thirteen out of fourteen participants found the concept of the grid easy to learn, with ten participants strongly agreeing that the system was more effective to use with the grid in place. This was thought to be because the grid cells could provide a spatial reference for the user, facilitating movement to the "top left" or "bottom right" of the page. One participant stated that the grid would help to reduce search time, as the user could initially locate a cell of interest, and then search within it for the target element.

However, a few participants stated that although the cell information conveyed via speech was easy to comprehend and useful for the evaluation task, it may become irritating and potentially overloading. Alternative suggestions from participants included the substitution of non-speech sounds, alternative labelling of grid cells (e.g. A1, A2 etc), or to have the grid layer as a customised "mode" that can be toggled on or off.

A description of the evaluation methodology and results will be described in a full paper for future publication.

Future Work

Several adaptations of this prototype have been developed, and will be fully evaluated.

Recursive "zooming"

The evaluated grid has a resolution of 3 by 3 cells. This is sufficient to give users an indication of their current position on the page. However, it would be desirable for a finer resolution to be conveyed so the user can explore the page in more detail. Thus the option to "zoom" into a grid cell has been developed. This is achieved via the provision of two keystrokes which enable the user to "zoom" into the grid. When the user navigates to a cell of interest, they can press a keystroke to move further into the cell.

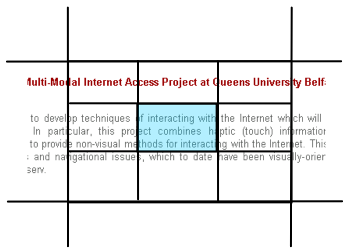

Figure 2. Recursive Grid

Overview

Currently, the system conveys the user's position within the page as he/she moves over a grid cell. It is hypothesised that by additionally providing an indicator of the complexity of a cell, it would be possible for an overview of the page to be conveyed to the user. This would relate to the macro and micro analysis stages described by [10]. The user could explore the page while a general description overview of each cell was conveyed ("macro viewpoint"). When the user discovers a cell of interest, they could then "zoom" into the cell to further explore its contents in finer detail ("micro viewpoint"). This also relates to the "overview and detail" interaction technique described by [9].

Complexity Algorithm

Determining the complexity of each cell is an interesting research question. A set of heuristics is to be developed that will enable a complexity rating to be determined. An initial analysis based on previous user studies suggests that a rating will be based on the number of elements within a cell, the distance between elements and the presence of an appropriate value for each element. An experiment is currently being designed to determine the effectiveness of non-speech audio against speech in conveying complexity.

Horizontal exploration

Research by various research groups employing eye tracking has shown that for some websites, users scan the page using an "F" pattern [8]. By providing a horizontal grid, it is hypothesised that users may be able use the mouse to move horizontally left and right to quickly scan the page. This is possible using the previous system, but the addition of the grid provides a structured framework to guide the user in the scanning process.

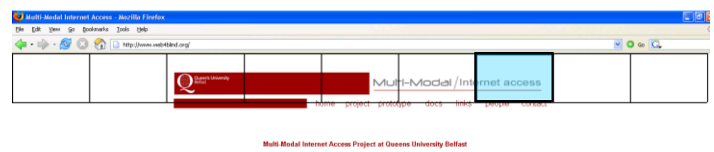

Figure 3. Horizontal Grid

Multimodal Feedback

In addition to investigating the effectiveness of conveying cell complexity and overview information through non-speech sound, it is intended that cell location information will also be conveyed through non-speech sound rather than speech. In terms of haptic output, additional awareness will be provided by applying textural effects to the body of each of the nine cells contained within the grid. This will be done using the metaphor of a chessboard (i.e. varying levels of roughness are applied to adjacent squares)

Keyboard-only navigation
The potential for the implemented grid system to enhance accessibility for sighted keyboard-only users, (e.g. those with motor impairments who are unable to use a mouse), will also be explored. It is hypothesised that by providing keyboard access to the grid, these users will be able to access sections of the page quicker than having to tab through multiple links.

Scrolling
The existing system does not allow for scrolling, which is a limitation for longer web pages. Therefore, a novel method for scrolling will be integrated with the current system and evaluated.

Screen reader integration
Currently, speech on the interface is provided by a Firefox plug-in. However, we hope to integrate the grid with current screen reader applications.

Conclusion

This paper has outlined an approach to enhancing the positional awareness of visually impaired users within the context of interactive systems conveying spatial layout via the application of a recursive grid. An extensive user evaluation has shown that the grid is a useful addition to the system. However, further research is needed to fully develop and evaluate all aspects of the system.

References

[1] Assistive Technology Act of 1998. http://www.section508.gov/docs/AT1998.html

[2] Di Blas, N., Paolini, P., Speroni M. and Capodieci, A. Enhancing Accessibility for Visually Impaired Users: the Munch's Exhibition. *Museums and the Web,* (2004).

[3] Hesham M., Kamel and Landay, J. Sketching Images Eyes-free: A Grid-Based Dynamic Drawing Tool for the Blind. *Proc. ASSETS 2002,* ACM Press (2002).

[4] Kamel, H., Roth, P. and Sinha, R, "Graphics and User's Exploration via Simple Sonics (GUESS)." *Proc. ICAD (2001),* HUT (2001), 261-265.

[5] McAllister, G., Yu, W., Strain, P., Kuber, R. and Murphy, E. Improvements Relating to the Access of Image Information. Patent No GB0624035.2 (Patent Applied For)

[6] Murphy, E., Kuber, R., Strain, P., Yu, W. and McAllister, G. An Empirical Investigation into the Difficulties Experienced by Visually Impaired Internet Users, To appear in *UAIS Journal*, (2007).

[7] Mynatt, E. Transforming Graphical Interfaces into Auditory Interfaces for Blind Users. *Human-Computer Interaction, 12,* (1997), 7-45.

[8] Nielsen, J. F-Shaped Pattern For Reading Web Content. http://www.useit.com/alertbox/reading_pattern.html

[9] Plaisant, C., Carr, D. and Shneiderman, B. Image-Browser Taxonomy and Guidelines for Designers. *IEEE Software, 12*, 2 (1995), 21–32.

[10] Roth, P. Petrucci, L., Pun, T. and Assimacopoulos, A. Auditory Browser for Blind and Visually Impaired Users*. Ext. Abstracts CHI 1999*, ACM Press (1999), 218–219.

[11] Rourke, C., Sloan, D. and O'Loughlin, M. Accessibility beyond the Guidelines. *Proc. HCI 2005*, Springer-Verlag (2005), 179-186.

[12] Yu, W., Kuber, R., Murphy, E., Strain, P. and McAllister, G. A Novel Multimodal Interface for Improving Visually Impaired People's Web Accessibility. *Virtual Reality, 9,* 2 (2006), 133-148.

[13] Zhao, H. iSonic: Interactive Sonification of Spatial Data for Auditory Information Seeking. PhD Thesis, 2006.