



## ACCESSIBILITY AND COMPUTING

A Regular Publication of the ACM  
Special Interest Group on  
Accessible Computing

September 2006

Number 86

### Contents:

	From the Editor	1
1	ASSETS 2006 Doctoral Consortium	2
2	Personalized and Adaptive Navigation based on Multimodal Annotation	4
3	Prototyping and Evaluation of Landcons: Auditory Objects that Support Wayfinding for Blind Travelers	8
4	A Visio-Haptic Wearable System for Assisting Individuals Who Are Blind	12
5	Developing An Assistive Haptic Framework For Improving Non-Visual Access To The Web	16
6	Multi-Layer Dialog Generation for Non-Visual Web Access	20
7	ViCRAM: Visual Complexity Rankings and Accessibility Metrics	24
8	Designing Cognitive Supports for Dementia	28
9	Understanding and Supporting the Use of Accommodating Technologies by Adult Learners with Reading Disabilities	32
10	An Evolutionary System for the Sc@ut Platform	36
11	Virtual Sign Animated Pedagogic Agents To Support Computer Education For Deaf Learners	40

## SIGACCESS Newsletter

A regular publication of the ACM Special Interest Group on  
Accessible Computing

### SIGACCESS Officers & Information

#### *Chairperson*

Vicki Hanson  
IBM T.J. Watson Research Center  
19 Skyline Drive  
Hawthorne, NY 10532, USA.  
+1-914-784-6603 ( work )  
chair\_SIGACCESS@acm.org

#### *Vice-chairperson*

Andrew Sears  
Information Systems Dept.  
UMBC  
1000 Hilltop Circle  
Baltimore, MD 21250, USA.  
+1-410-455-3883 ( work )  
vc\_SIGACCESS@acm.org

#### *Secretary/Treasurer*

Noelle Carbonell  
LORIA,  
Campus Scientifique  
F54506 Vandœuvre-lès-Nancy  
Cedex  
France  
treasurer\_SIGACCESS@acm.org

#### *Newsletter Editor*

Simeon Keates  
ITA Software  
141 Portland Street,  
Cambridge, MA 02139.  
editors\_SIGACCESS@acm.org

#### *Who we are*

SIGACCESS is a special interest group of ACM. The SIGACCESS Newsletter is published regularly in January, June, and September. We encourage a wide variety of contributions, such as: letters to the editor, technical papers, short reports, reviews of papers of products, abstracts, book reviews, conference reports and/or announcements, interesting web page URLs, local activity reports, etc. Actually, we solicit almost anything of interest to our readers.

Material may be reproduced from the Newsletter for non-commercial use with credit to the author and SIGACCESS. Deadlines are one month before publication dates. Submissions may be sent as hard copy (paper), but machine-readable files are preferred. Postscript or PDF files may be used if layout is important, but word-processor files, text files, or e-mail are also acceptable. Ask the editor if in doubt.

Finally, you may publish your work here before submitting it elsewhere. We are a very informal forum for sharing ideas with others who have common interests.

Anyone interested in editing a special issue on an appropriate topic should contact the editor.

# Accessibility and Computing



## The Newsletter of ACM SIGACCESS

NUMBER 86, September 2006

---

### **A note from the Editor**

Dear SIGACCESS member:

Welcome to another issue of the SIGACCESS Newsletter.

This is a special issue focused on the Doctoral Consortium held at ASSETS 2006 and sponsored by NSF. As Yeliz Yesilada and Andrew Sears (the Doctoral Consortium co-chairs) describe in their introduction to the papers and authors, the Consortium participants were initially selected based upon an extended abstract that was submitted through the on-line conference management system. Those authors whose papers were selected were then invited to participate in the Consortium and to prepare a brief presentation to be given to a select group of established academics. That panel of experts gave their feedback and the authors invited to revise and extend their papers. This issue of this newsletter contains those final versions of the papers.

### **And finally...**

After mention of ASSETS 2006, it seems appropriate to remind readers that the call for ASSETS 2007 is now available on-line at:

<http://www.acm.org/sigaccess/assets07/>

*Simeon Keates*

Newsletter Editor

## ASSETS 2006 Doctoral Consortium

*Yeliz Yesilada*<sup>1</sup>, *Andrew Sears*<sup>2</sup>

<sup>1</sup>School of Computer Science, University of Manchester Manchester, UK

yeliz.yesilada@manchester.ac.uk

<sup>2</sup>Information Systems Department, Interactive Systems Research Center, UMBC,  
Baltimore, USA

asears@umbc.edu

### Introduction

The Eight International ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2006) was held between 22nd and 25th of October 2006 in Portland, USA. As part of this conference, a Doctoral Consortium was held on Sunday the 22nd of October. Ten students from five different countries and five faculty advisors were welcomed. Overall the Doctoral Consortium was a great success and was organised as follows: each student prepared and presented a brief summary of their dissertations followed by open discussion among the faculty and students. This helped us to ensure that each student got valuable feedback from all advisors. Presented dissertations covered a broad range of research topics focusing on advancing state of the art for disabled people. The following four topics were organised into the four sessions:

#### ***Wayfinding, Mobility and Wearable Systems***

In this first session, three doctoral candidates presented their dissertations. First Thorsten Volkel (University of Kiel) presented his dissertation titled “*Personalised and Adaptive Navigation based on Multimodal Annotation*”. His presentation was followed by Robert Lutz (New Jersey Institute of Technology) whose dissertation titled “*Prototyping and Evaluation of LandCons: Auditory Objects that Support Wayfinding for Blind Travelers*”. The last presentation in this session was given by Troy McDaniel (Arizona State University) whose dissertation was titled “*Visio-Haptic Wearable System for Assisting Individuals Who Are Blind*”.

#### ***Visually Impaired Users and the Web***

In this second session, three doctoral candidates presented their work on enhancing Web accessibility for visually impaired users. Ravi Kuber (Queen’s University) was the first presenter whose dissertation was titled “*Developing An Assistive Haptic Framework For Improving Non-Visual Access To The Web*”. Kuber’s presentation was followed by Yevgen Borodin’s (Stony Brook University) talk whose dissertation titled “*Multi-Layer Dialog Generation for Non-Visual Web Access*”. The last presentation in this session was given by Eleni Michailidou (The University of Manchester) whose dissertation was titled “*ViCRAM: Visual Complexity Rankings and Accessibility Metrics*”.

#### ***Cognitive and Learning Disability***

In this third session, there were two presentations. Joe Wherton (University of York) presented his dissertation titled “*Designing Cognitive Supports for Dementia*” and

Katherine Deibel (University of Washington) gave the second presentation whose dissertation titled “*Understanding and Supporting the Use of Accommodating Technologies by Adult Learners with Reading Disabilities*”.

### ***Education and Disability***

In this last session, there were two presentations. Maria Dolores Paredes-Garrido (University of Granada) gave the first presentation whose dissertation titled “*An Evolutionary System for the Sc@ut Platform*”. The last presentation was given by Benjaporn Saksiri (Clemson University) and was titled “*Virtual Sign Animated Pedagogic Agents To Support Computer Education For Deaf Learners*”.

Our faculty advisors included Dr.Armando Barreto (Florida International University), Dr.Stephen Brewster (University of Glasgow), Dr.Jinjuan Feng (Towson University), Dr.Arthur Karshmer (University of San Francisco) and Dr.Richard Ladner (University of Washington). It was important for us to ensure that students with such a broad range of dissertation topics would get valuable feedback from different perspectives and they could benefit from most of this event therefore our faculty advisors were specialised in different areas of accessibility.

Students were accepted to participate in the Doctoral Consortium based on two-page long descriptions of their dissertations and a letter of recommendation from their PhD supervisors and/or advisors. In this abstract, they were asked to clearly describe their research problem, the motivation behind their research, the proposed solution, the progress of their research and finally discuss envisioned contributions. This year each submission was peer-reviewed by our faculty advisors. Students were evaluated based on metrics related to the originality of work, importance and potential impacts to the accessibility field, and the soundness and correctness of the proposed approach to address the problem. This review process helped us to ensure that we have chosen students whose dissertations focus on advancing the state of the art of accessibility, which is the major theme of the conference. Updated versions of these two-page long abstracts are included in this newsletter to give the SIGACCESS community an idea of the research topics covered at the Doctoral Consortium.

The Doctoral Consortium was sponsored by National Science Foundation. Expenses of all doctoral consortium candidates, faculty advisors and doctoral consortium organisers were covered in part by this NSF funding. This year part of this funding was also used to invite students from last year’s Doctoral Consortium to meet students from this year and share their experiences.

Many people contributed to the success of the program. First we would like to thank the students for their participation. Without their dedication this event would have not been possible. We would also like to thank the faculty members for their exceptional work and dedication in the review process and giving exceptional feedback to students and making this event very successful. Finally, we express our deepest appreciation to the National Science Foundation for its support.

# Personalized and Adaptive Navigation based on Multimodal Annotation

*Thorsten Völkel*

Human-Centered Interfaces Research Group, University of Kiel

t.voelkel@hci-research.de

## Abstract

Although much effort is spent in developing navigation systems for pedestrians, many users with special needs are mostly excluded due to a lack of appropriate geographical data such as landmarks, waypoints, or obstacles. Such data is necessary for computing suitable routes because the best route might not be the shortest or fastest one. In this paper, the concept of multimodal annotation of geographical data for adaptive and personalized navigation is described. Direct input by the user is combined with data derived from the observation of the user's LOM-Modality (Location, Orientation, Movement) to annotate geographical data. Based on this data and data derived from other users of the same user group, suitable routes even in unknown territory can be calculated.

## Introduction

Carrying out an independent and autonomous life seems usual for most people around us. However, many disabled people and elderly people face immense problems in overcoming difficulties imposed by our common environment. For example, for people restricted in their mobility due to an impairment, a task including covering a specific distance for buying daily life products may become a heavy burden.

Travel aids and assistive technologies have been developed to guide elderly people and blind people in unknown environments. For example, the MOBIC travel aid [6] is based on geographical information systems (GIS) and the Global Positioning System (GPS) for guidance during macro-navigation which mainly covers distances greater than 50 meters. The system also provides a component for pre-journey planning, on which the later navigation relies. The MOBIC system proved successful for macro-navigation although distances between micro-navigation (up to 10 meters) and macro-navigation are not supported. Other approaches also used for indoor navigation incorporate beacons which provide additional information about the environment. Although promising results have been reported [5], such systems require immense investments regarding necessary infrastructure and may therefore only available at special places.

Using current GPS systems, promising results have been reported for navigational support for instance for visually impaired people [3]. However, one of the most important problems – as was also reported for the MOBIC system – is the acquisition of specific data including obstacles, specific waypoints, and landmarks [2,4]. The additional information is necessary to adaptively calculate the best route for people with special needs and provide accurate route descriptions. Although the direct and shortest route might seem the best choice for people without impairments, for instance blind people try

to avoid crowded cross-ways. Consequently, a route avoiding such cross-ways – though it might be of longer distance – becomes more suitable.

### **The concept of multimodal annotation**

As current navigation systems do not fulfil most of the requirements of disabled users, the annotation of the underlying geographical data such as the position of obstacles, landmarks or waypoints as well as information about the adequacy of routes for specific users offers new potentials for macro-navigation and micro-navigation as well as for distances in between.

The acquirement of annotation data is based on two integral parts, namely direct input by the user via different modes such as pen or speech input combined with input derived by analysing the users' LOM-Modality (Location – Orientation – Movement). The LOM-Modality, in contrast to solely considering the location of the user as context, combines information of location, orientation and movement and the corresponding histories into one modality. Annotations can be obtained either by using the users' direct input or the derived LOM data isolated or the combination of both respectively. For example, the user might provide the location of a mailbox by using speech input. Using the LOM-Modality, specific routes can be assigned with weights indicating the suitability for the user. If the user needs significantly longer than the average for covering a specific distance, the route might therefore not be suitable for him. Deriving isolated annotation data by analysing the users movement, stops, locations and orientations might lead to improper conclusions regarding the weighting of routes as the system is not able to determine the actual reason for any difference which varies significantly from the average. Therefore, additional information derived by direct input can help to rate the initial results leading to multimodal input for the annotation of the underlying geographical data.

### **Personalized and stereo-type based navigation**

Most geographical information systems allow for adding the location of individual waypoints and landmarks of the user. As described in the introduction, this data is not sufficient for calculating the best route for people with special needs. Consequently, the navigation problem is extended beyond the problem of finding the shortest or fastest route to a given destination. Additionally, not only the profile of a user must be considered when calculating routes but also the actual situation of the user. For example, it becomes crucial whether a blind user only uses a white cane – and consequently is able to use escalators – or is accompanied by a guide dog which in contrast is not able to use escalators.

As described in the previous section, geographical data is annotated in a multimodal way leading to data directly related to the individual user. This approach works for environments of the user's daily life, however, in unknown territory only well-known geographical data will be available. One possible solution is the concept of stereo-type based navigation. By using data derived from users of the same user group, suitable routes can also be calculated in unknown environment. Therefore, the problem of giving the user required navigation information must be solved, as the system must decide whether directions for micro-navigation, macro-navigation or for distances in between are most important for the user's current situation. In detail, it becomes crucial whether to

provide the user with a description of the next 20 meters or the overall direction. For example, for partially sighted users obstacles within the next 20 meters may be of higher importance compared to a change of direction 100 meters later. Although for example the MOBIC system incorporates a pre-journey component for planning a route, the guidance during the actual journey is not adaptive, that is, neither the user's actual behaviour (except his current location) nor the switching between information concerning micro-navigation and macro-navigation is considered.

## Conclusion and Outlook

As yet, first results have been reached on accessibility issues concerning small devices such as mobile phones and smart phones and their potential applications[7]. Concerning location-based services – besides an accessible presentation of the corresponding information – the personalized and adaptive generation of appropriate directions is one of the main problems due to inaccurate data provided by state-of-the art navigation systems. A personalization of information and services is one of the key requirements for navigation systems to be successful for a wide range of different users [1]. We thus believe that an application of the concept of multimodal annotation for both the acquisition of data as well as for the calculation of adaptive route descriptions has the potential to aid in particular users with special needs in carrying out their life more independently regarding their grade of mobility.

As this research is still in an early stage, only basic concepts and a general architecture for the system have been developed. We are currently gathering user requirements for the system. Based on these requirements a refinement of the concepts as well as a basic implementation allowing for first evaluations is currently in progress. Concerning the development of a prototype, the extension of a commercial navigation system with the above described functionality is intended allowing an easy use for many people without the need of buying special and potentially expensive hardware.

## References

1. Bradley, N.A., Dunlop, M.D. An Experimental Investigation into Wayfinding Directions for Visually Impaired People. In: *Personal and Ubiquitous Computing* 9, 6 (November 2005), 395-403.
2. Golledge, R.G., Marston, J.R., Loomis, J.M., Klatzky, R.L. Stated Preferences for Components of a Personal Guidance System for Nonvisual Navigation. In: *Journal of Visual Impairment and Blindness*, 99, 4 (April 2005), 219-232.
3. Loomis, J.M., Marston, J.R., Golledge, R.G., Klatzky, R.L. Personal Guidance System for People with Visual Impairment: A Comparison of Spatial Displays for Route Guidance. In: *Journal of Visual Impairment and Blindness*, 98, 3 (March 2004), 135-147.
4. May, A.J., Ross, T., Bayer, S.H., Tarkiainen, M.J. Pedestrian Navigation aids: information requirements and design implications. In: *Personal and Ubiquitous Computing* 7, 6 (December 2003), 331-338.
5. Ross, D.A., Lightman, A. Talking Braille: A Wireless Ubiquitous Computing Network for Orientation and Wayfinding. In *Proceedings of the 7th international ACM conference on Assistive technologies (ASSETS'05)* (Baltimore, Maryland, 2005), ACM Press, New York, NY, 2000, 98-105.

6. Strothotte, T., Fritz, S., Michel, R., Raab, A., Petrie, H., Johnson, V., Reicher, L., Schalt, A. Development of Dialogue Systems for a Mobility Aid for Blind People: Initial Design and Usability Testing. In Proceedings of the second annual ACM conference on Assistive technologies (ASSETS'96) (Vancouver, Canada, 1996), ACM Press, New York, NY, 1996, 139-144.
7. Völkel, T., Weber, G. Scenarios for Accessible Multimedia Messaging Services. Proceedings of the 9th ERCIM Workshop "User Interfaces for All" (UI4All) (Königswinter, Germany, 2006), Springer, Berlin, in print.

# Prototyping and Evaluation of Landcons: Auditory Objects that Support Wayfinding for Blind Travelers

*Robert J. Lutz*

New Jersey Institute of Technology

[lutz@acm.org](mailto:lutz@acm.org)

## Introduction

Sighted people use vision to quickly access a rich amount of information about their environment. Blind users are deprived of this information, which compromises their abilities both to understand their surroundings and to navigate within them. Alternative sensory information, such as sound or touch, is substituted for the missing stimulus [7]. This information supports navigation in the immediate vicinity, e.g., the location of obstacles and hazards, but does not provide the location of distant unique objects which sighted people often use as navigating landmarks. The opportunity exists to electronically augment a blind user's environment with information about distant landmarks, e.g., information that would allow a blind person to turn in a circle, listen to distant landmarks, and then proceed in a direction guided by a chosen landmark.

Technological solutions in several key areas permit access to contextual information about environments. These include: location sensing systems [3]; open and rich geographic databases; high performance portable computing devices; ubiquitous high speed networking; and high quality computer-based sound rendering. While these solutions are not new, past and continued development permits affordable access that was previously prohibitive for all but very wealthy consumers of technology and the research community.

Previous work, referenced subsequently, suggests that the current technology portfolio can be applied in an affordable way to augment a blind traveler's environment with relevant navigation information. This research will focus on developing and evaluating such a navigation aid.

The rest of this document is organized as follows. First, the background literature supporting this work is presented followed by a description of the proposed navigation aid. This is followed by a presentation of the current status of the research and then by a discussion of this work's contribution to the field. Finally, a brief description of the author / consortium applicant is presented.

## Background

A review of related work suggests several important features and concepts to consider in designing the landmark navigation aid.

- Navigational tasks decompose into two main areas: (1) nearby hazard / obstacle avoidance and (2) navigating to remote destinations [6]. Type 1 is the purview of the Orientation and Mobility (or O&M, previously known as foot travel) professional practice of occupational instruction, e.g., canes and seeing-eye dogs. [2]. Type 2 represents the perceived opportunity space for this research.

- Significant research has been undertaken to evaluate methods for auditory information display. Earcons [1] and sonification [5], provide rich information detail via terse audio representations. Speech can provide information with minimal ambiguity, but imposes higher cognitive load. Several navigation systems for destination planning and wayfinding have been developed for both academic and commercial applications [6], [8], [11], [13], <http://www.humanware.ca>, <http://www.senderogroup.com>. Feedback is provided via sound, touch or a combination of the two. Users of systems providing tactile feedback outperform those using systems with audio feedback. [12].
- Audio feedback that provides location information through variations in musical quality or pitch has been shown to be a good mechanism for providing information on relative differences and situational positioning. [9].
- Using silence during audio display has been found to be an effective mechanism for helping users to maintain direction in the user studies performed in our labs at the New Jersey Institute of Technology.
- Recent research has investigated the manipulation of user-selected music for navigational feedback for sighted users with spatial audio being used to guide travelers through the simulation of the music origin as the navigation target. [4]. It is unclear whether this passive type of support will work effectively for blind users because of its interference with other useful audio navigation sources.
- Researchers have demonstrated display of continuous values to subjects via an audio abacus[14]. During blind-user navigation, this mechanism could be used to display real, continuous values such as distance.
- A software-based real-time virtual acoustic environment (VAE) rendering system has been developed at NASA Ames Research Center [10]. This system offers developers a software-based experimental platform with low-level control of a variety of signal processing parameters for conducting psychoacoustic studies. This system has characteristics that are similar to the test apparatus being constructed.

## Proposed solution

This work will leverage the research described in the previous section to assemble a new design that will give the blind navigator information about both distant and near objects in his or her surroundings. Since haptic interfaces are not uniformly found on commodity computing platforms, the problem is further constrained to eliminate haptic interfaces from consideration. The goal of this research is thus: to develop audio interfaces which will give a blind navigator wayfinding information captured from the environment to support local and distant navigation that will approach the effectiveness of the touch-based systems.

Two audio feedback approaches will be developed and tested. These have already been prototyped and found to show promise among our blind test user population.

The first approach uses an egocentric coordinate system (the navigator is at the zero point and all paths lead from the navigator), and provides feedback of a desired landmark's distance and direction through sonification. This is done by providing sound that falls off in pitch and timbre when the user orients away from the chosen landmark. Distance from

the landmark is given either as a variation in clicks per time or in generated speech. Landmarks, themselves, are identified by speech, but the plan is to also explore ecologically-based sounds.

The second approach will use silence. When a landmark is chosen as the navigation destination, silence will be provided as long as the traveler is within a desired threshold of the optimum path to the landmark.

Further iteration in the use of continuous audio feedback to provide navigation orientation and guidance is expected, based on our evaluations conducted with blind users.

### **Status of research**

A test environment has been constructed using a laptop computer and a Wacom tablet, with both cartesian and angular reporting capability. A user navigates within a virtual space that is mapped onto the tablet surface. The user is given a navigational objective to find 10 distinct landmarks. As they explore, appropriate feedback is provided via binaural headphones. Feedback includes both sonified audio as well as generated text-to-speech. Performance of the navigation task within the environment is captured via several commonly used measures: elapsed time, distance traveled, and path error.

It is realized that this miniature environment is a proxy for the real environment, and that evaluation “in the large” is optimal and may be required. If not, justification needs to be provided to support the generalizability of results to life scale. The benefit of the small environment is that many design iterations can be tested. Because the current landmarks are virtual and sonified, they have been called Landcons, akin to the labeling of virtual sounds as earcons. Eventually a landcon will represent a real landmark.

Pilot tests have been conducted with the apparatus in its current state. Over the next 3 months, the test apparatus will be finalized and a user study will be conducted. Results from this phase will be available at ASSETS 2006. A full demonstration of the test apparatus will also be available.

### **Contribution to the field**

The most significant contribution that would come from this work is the generation of a more comprehensive auditory feedback mechanism for real-time wayfinding. Such a system can work in the small and guide a blind user along a curved sidewalk but also work in the large guiding a blind user to a key but distant location, e.g., the subway station.

A secondary contribution would be the potential application of this research to stationary settings in which a blind user could explore an environment before traversing said environment.

### **About the author**

I am currently a part-time Ph.D. student at the New Jersey Institute of Technology in the Information Systems Department. I am also a full-time employee of Sun Microsystems, Inc., currently working in product engineering on a Utility Computing offering At

ASSETS, I hope to meet others with similar research interests and to benefit from the invaluable feedback I'll receive if accepted to the Doctoral Consortium.

## References

1. Blattner, M. M., Sunikawa, D. A. and Greenberg R.M. Earcons and Icons, Their structure and common design principles. *Human Computer Interaction*, 4, 1, 1989, 11-44.
2. Blasch, B., Wiener, W., and Welsh R. *Foundations of orientation and mobility* (2nd Edition). New York, NY: American Foundation for the Blind, 1997.
3. Hightower J. & Borriello G. Location Systems for Ubiquitous Computing, *Computer*, 34, 8, August 2001, 57-66.
4. Jones, M. and Jones S. The Music is in the Message. *ACM Interactions XIII*, 4, July + August 2006, 24-27.
5. Kramer, G., Walker B., Bonebright T., Cook P., Flowers J., Miner N. and Neuhoff J. Sonification report: Status of the field and research agenda, <http://www.icad.org/websiteV2.0/References/nsf.html>. 1999.
6. Loomis, J. M., Golledge, R.G., Klatzky, R.L., Speigle, J., and Tietz, J. Personal Guidance System for the Visually Impaired. *Proceedings of the First Annual ACM /SIGCAPH Conference on Assistive Technologies*, Marina Del Ray, California, October 31-November 1, 1994, 85-90.
7. Loomis, J. M. Sensory replacement and sensory substitution: Overview and prospects for the future. In M. C. Roco & W. S. Bainbridge (Eds.), *Converging Technologies for Improving Human Performance: Nanotechnology, Biotechnology, Information Technology and Cognitive Science*, Boston: Kluwer Academic Publishers, 2003.
8. The MoBIC website. Retrieved Nov.11, 2005 from <http://phoenix.herts.ac.uk/SDRU/MOBIC/mobic.html>.
9. McQueen, J. C. *Auditory Pie Menus*. Masters Thesis, University of Toronto, Toronto, CA, 1995.
10. Miller, J.D. and Wenzel, E.M., Recent Developments in SLAB, A Software Based System for Interactive Spatial Sound Synthesis, *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto Japan, 2002, ICAD02-1 through ICAD02-6.
11. Ogata, T., Makino, H., Ishii I., and Nakashizuka, M. Location guidance system for the visually impaired using an invisible bar code. *Transactions of IEICE*, J80/D-2, 1997, 3101-3107 (in Japanese).
12. Ross, D. A. and Blasch B. B. *Wearable Interfaces for Orientation and Wayfinding*, *Proceedings of the Fourth International ACM Conference on Assistive Technologies*, Arlington Virginia, 2000, 193-200.
13. Ross, D. A. and Lightman, A. *Talking Braille: A Wireless Ubiquitous Computing Network for Orientation and Wayfinding*. *Proceedings of the 7th International ACM SIGACCESS Conference on Computers and Accessibility*, Baltimore, Maryland, 2005, 98-105.
14. Walker, B., Lindsay, J., and Godfrey, J., *The Audio Abacus, Representing Numerical Values with Nonspeech Sound for the Visually Impaired*, *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility*, Atlanta, GA, 2003, 9-15.

## **A Visio-Haptic Wearable System for Assisting Individuals Who Are Blind**

*Troy L. McDaniel, Sethuraman Panchanathan*

Center for Cognitive Ubiquitous Computing, Arizona State University

699 South Mill Ave., Tempe, AZ, 85281

Email: Troy.Mcdaniel@asu.edu

### **Abstract**

Computer vision algorithms for visio-haptic information analysis, i.e., the conversion of visual data into haptic (tangible) features, can be utilized in wearable assistive devices for individuals who are blind. Touch is an important modality for individuals who are blind, but it is limited to the extent of one's reach. By estimating how an object feels from its visual image, we are able to overcome this limitation. This paper proposes a wearable assistive device to estimate haptic features from visual data to enable users to feel objects from a distance.

### **Introduction**

Individuals who are blind rely on their sense of touch for perceiving their proximal environment. Unfortunately, this modality is limited to the extent of one's reach, thereby limiting haptic perception of objects in the distal environment. This paper proposes a wearable system to estimate haptic (tangible) features from visual data (i.e., visio-haptic information analysis or simply visio-haptics) to enable users to feel objects from a distance.

Haptic features are perceived through touch, but humans can visually estimate haptic features through intermodal processing and transfer [2]. Similarly, computer vision algorithms can be utilized to estimate an object's haptic features including shape, size, texture and material. Once haptic features have been estimated, they may be presented to the user through a haptic user interface.

We propose the integration of Radio-Frequency Identification (RFID) technology and computer vision for visio-haptics. RFID tags may be placed on objects in the environment, and their haptic and visual features stored in a central database. As tagged objects are encountered, a memory map is built using Multidimensional Scaling (MDS) where objects are clustered based on salient, haptic features. Utilizing this ground truth, we may tune a learning system for intermodal transfer, i.e., the prediction of haptic features from visual input. Essentially, the learning system provides a mapping between visual input and haptic concepts stored in memory.

Computer vision algorithms often work fine in controlled settings, but good performance in uncontrolled environments can become challenging. Rather than have the system entirely handle the task of analysis, it is important to utilize human capacity and involve the user in this process, creating a human-in-the-loop system. The system design we present here is unique in that it incorporates confidence measures to communicate to the user how confident it is about its decisions regarding how an object feels. For example, if

confidence is poor because of a dimly lit scene, the user is informed and may find a way to improve lighting to assist the system in analysis.

## Related work

There have been several attempts at developing algorithms for estimating haptic features from visual data. One wearable system for individuals who are blind is the Tactile Vision Sensory Substitution System (TVSS) [1]. An image is first divided into blocks, and the average intensity value of each block is computed. These values are then converted directly into vibro-tactile stimulations delivered to the user through a grid. However, this approach merely presents visual features, i.e., illumination values, to the haptic modality, which is a major design flaw, as the haptic modality should receive haptic features.

Kahol [3] developed a wearable assistive device for individuals who are blind that utilizes visio-haptic algorithms to enable users to feel from a distance. The conceptual framework outlines two types of algorithms: those that operate at a perceptual level, and those at a physical level. The perceptual algorithms classify features into predetermined classes, and subsequently, these features are conveyed to the user through a haptic glove in the form of tactile cues, which Kahol shows are capable of invoking concepts in a user's mind. While the algorithms achieve high accuracy in controlled environments, uncontrolled environments present a problem. A novel approach to visio-haptics is presented here that utilizes a unique framework for wearable computing that integrates computer vision and RFID. Further, the use of confidence measures is proposed to enhance the system's use in uncontrolled environments.

## Conceptual framework

The goal of my research is to design and develop algorithms for estimating haptic features from visual images, which will be utilized in a wearable system for individuals who are blind, depicted in Fig. 1.

Computer vision algorithms for visio-haptic information analysis may operate at either a perceptual or physical level [3]. At a perceptual level, haptic features are classified into predetermined classes, e.g., the texture of an object may be classified as rough, medium or smooth. At a physical level, haptic features such as texture or shape are reconstructed from an object's image. Haptic user interfaces are limited for haptic presentation of physical information, and hence our wearable system will utilize algorithms that work at a perceptual level. Haptic information will be delivered to the user using the tactile cueing methodology of [3], which has been shown to invoke haptic concepts. A concept is an average representation of a category; for example, the object category dog invokes an image in our minds that is an average of all the different types of dogs we've encountered.

We propose a novel framework, depicted in Fig. 2, to integrate RFID and visio-haptics. RFID provides a method to easily learn about objects in the environment to create a memory map through Multidimensional Scaling of haptic features. Using computer vision, we may estimate haptic features from visual data through intermodal transfer, and recall from memory haptic concepts. A learning system may be trained for intermodal transfer using ground truth provided by RFID and the visual input of tagged objects.

Perceptual visual and haptic ground truth for RFID may be collected using human participants.

A human-in-the-loop paradigm is crucial as computer vision is much more challenging in uncontrolled environments. Scale changes, pose changes, illumination changes, blur, occlusion and noise are often encountered when using a vision-based wearable computer. Rather than have the system entirely handle the task of analysis while ignoring the user, it is important to involve the user in the process, creating a human-in-the-loop system to take advantage of human capacity. The system communicates to the user how confident it is about its decisions regarding how objects feel. Confidence will depend on a) environmental cues, e.g., illumination, motion blur, etc.; b) context-dependent cues, e.g., the object was recognized as a bowl, but this conflicts with its material, which was recognized as cloth; and/or c) algorithm-dependent cues that rely on training data.

Lastly, the usability of the wearable system will be extensively tested using a group of participants who are blind. The system will be judged on its ease-of-use, comfort and usefulness. Changes will be made to system design based on user comments.



Figure 1: Wearable system

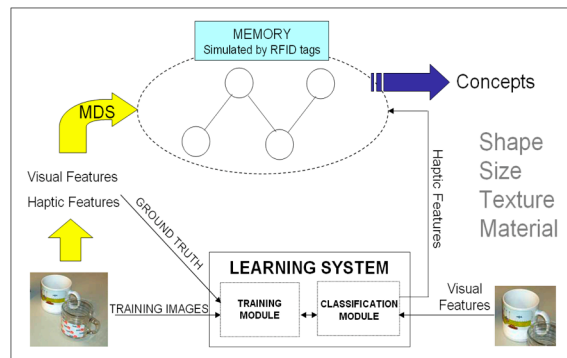


Figure 2: Framework for visio-haptics and RFID

### Current status of research

A methodology to collect ground truth for computer vision algorithms for estimating haptic features from visual data was developed [4]. Perceptual ground truth in both the visual and haptic modality was collected from participants. Physical ground truth was collected utilizing stereo analysis and 3D tracking equipment. Confidence measures based on illumination and motion blur were developed [4]. Illumination is classified as

poor, good or great by computing the mean pixel value of a grayscale image, then classifying the sample mean using a Bayesian classifier. Accuracy on novel test images was 96%. For motion blur classification, a metric based on average line width computed from a difference image is classified as no motion blur, small motion blur, large motion blur or extreme motion blur using thresholds found through experimentation. Accuracy on novel test images was 95%.

A wearable system capable of video processing and RFID sensing has been built. Video is captured using a small, hidden camera built into a pair of ordinary sunglasses, and is sent to a wearable computer for processing. Confidence measures based on illumination and motion blur may be accessed by the user at any time. RFID sensing is performed using an RFID reader contained in a lightweight carrying case with shoulder strap. Once a tag is read, it is sent to the wearable computer through 802.11. An object's haptic features are accessed through a database, and presented to the user through audio output or vibro-tactile stimulations using a haptic glove and the cueing methodology of [3].

## Conclusion and future work

This paper proposes a wearable system to estimate haptic features from visual data to enable users to feel objects from a distance. This research will make several important contributions: 1) Visio-haptic algorithms for estimating haptic features from visual data that are robust enough to use in uncontrolled settings. 2) Confidence measures for conveying to the user how confident the system is about its decisions. Further, the measures will inform the user how he or she may increase confidence, further improving the reliability of the system. 3) The final contribution is a framework for combining computer vision with RFID technology where RFID builds a memory map and computer vision invokes concepts from memory. Future work will involve 1) the development of a learning system for intermodal transfer of visual shape, size, texture and material to their haptic counterparts contained in memory; 2) development of confidence measures based on video noise and out-of-focus blur; and 3) usability testing.

## References

1. Bach-y-Rita, P., Kaczmarek, K. A., Tyler, M. E., and Garcia-Lara, J. Form perception with a 49-point electro-tactile stimulus array on the tongue: a technical note. *Journal of Rehabilitation Research and Development*, 35, 4 (Oct. 1998), 427-430.
2. Hatwell, Y., Streri, A., and Gentaz, E. *Touching for knowing: cognitive psychology of haptic manual perception*. John Benjamins Pub., Amsterdam, Philadelphia, 2003.
3. Kahol, K. *Distal object perception through haptic user interfaces*. Ph.D. Thesis, Arizona State University, Tempe, AZ, 2006.
4. McDaniel, T.L. <http://www.public.asu.edu/~tmcdani>, 2006.

## Developing An Assistive Haptic Framework For Improving Non-Visual Access To The Web

*Ravi Kuber*

Sonic Arts Research Centre  
Queen's University Belfast  
Belfast, BT7 1NN, UK  
r.kuber@qub.ac.uk

### Abstract

With the growing focus on user-centered design by HCI researchers and practitioners, it is perhaps ironic that the visually impaired community continues to encounter challenges when interacting with the Web. Access is believed to be improved through the use of haptic technologies. However, due to the limited availability of targeted haptic design guidelines, developers could arbitrarily select and map sensations, with little consideration given as to how the sense of touch could provide assistance to a blind user. Research aims to address the limitations associated with existing assistive technologies, by developing haptic feedback which provides both spatial and navigational assistance when accessing the Internet. This research aims to culminate in a haptic framework, for designers to reference when developing accessible solutions for the Web.

### Background

The Internet offers considerable potential to the visually impaired community. However, recent evidence suggests that needs of blind and partially sighted web users are often overlooked [6]. Barriers to access are attributed to the graphically-oriented presentation of information on a page, coupled with limitations imposed by assistive devices. The process of navigating through a busy page is known to be tedious, largely due to the linear output of screen reading technologies [8]. Gaining an overview of the spatial layout of page elements can also pose a challenge. Improved non-visual presentation of information and additional navigational cues would not only assist visually impaired users when exploring unfamiliar web pages, but would also help when targeting particular sections of each page. Considerable benefit would be gained by those users who collaborate on web-based tasks with sighted peers, as discussed in further detail by Mynatt & Weber [4].

In contrast with other modalities, the haptic channel presents the advantage of bi-directional information transfer between the user and the virtual environment, and when utilised, is known to reduce the burden on vision and audition, freeing these senses for other tasks [5]. O'Modhrain & Gillespie [5] have stated that in certain instances, it is possible to substitute haptics for other sensory modalities, such as graphical user interface (GUI) exploration for visually impaired individuals. As haptic technologies develop and devices reduce in cost, using the sense of touch becomes a more viable method for creating a universally accessible solution to browsing the Web. However, existing web accessibility guidelines do not extend to the use of haptic feedback. Similarly, haptic

design guidelines do not provide the focused assistance needed for use on the Web. Limited haptic interface design guidance may encourage designers to select arbitrary sensations to represent elements on a web page, rather than develop stimuli which provide assistive benefit to the user. Choice may be influenced from the limited range of effects currently offered by haptic technologies. Poorly designed or inappropriate sensations may reduce the quality of the subjective experience. Lack of design conventions could lead to the reuse of the same haptic sensation to represent multiple page elements, which could result in confusion for the visually impaired user. Mappings could also vary from site-to-site, meaning that the user would need to invest time and effort explicitly learning the meaning associated with each sensation. A haptic framework housing a library of assistive feedback would allow designers to select and replicate standardized sensations, which would provide assistance to a visually impaired user when exploring a web page. Accompanying recommendations would aim to discuss the application of assistive haptic icons within the context of a page.

### **Design challenges**

The main design challenge would be translating an interactive, spatially presented, visually dense interface, into an efficient, intuitive and non-intrusive non-visual interface [4]. A direct graphical-to-haptic conversion would not provide benefit to users who have never experienced sight in earlier life. Haptic feedback would need to be specifically designed to assist the exploration and targeting processes. Assistive cues should not only enable visually impaired users to browse pages independently, but should also allow blind and sighted parties to work together to reach a common goal. To support joint operation, the respective presentations of both visually and non-visually-orientated technologies should be synchronized [4]. This could be achieved by providing visually impaired users with an assistive haptic layer for use with current browsing technologies (e.g. Internet Explorer or Firefox).

In order to develop an assistive haptic layer, research would need to assess the information that needs to be perceived by the user to provide assistance, determine how to represent this through the use of haptic technologies, and investigate how these haptic sensations can be perceived effectively by the user. Developing assistive feedback that communicates an underlying sense of meaning to the user would offer considerable advantage.

### **Related work**

Haptic technologies have been used to promote spatial awareness within a virtual environment. Mynatt & Weber [4] refer to the development of the GUIB system, whose design allowed visual information to be transformed into a tactile representation based on the spatial representation of a GUI. Icons, buttons, controls and menus on the desktop have been tactually perceived using the Betacom Screen Rover [1] and the Moose [5]. However, it is thought that haptic-based solutions have yet to target the issues pertaining to communicating the semantic content associated with HTML markup, presented on a web interface.

In order to assist navigation, Yu and Brewster [7] have enabled visually impaired users to haptically follow the grooved lines in graphs, to promote awareness of the trends in

complex numerical datasets. Following a groove is thought to reduce time spent locating reference points, minimizing confusion within a virtual environment. Non-visual feedback has been found by Jacko et al. [2] to assist with complex mouse-based tasks, such as the ‘dragging and dropping’ of files on an interface. The users were able to experience a haptic sensation upon the correct positioning of the file using the Logitech Wingman force-feedback mouse.

With regards to haptic interface design, MacLean [3] has discussed the use of a “top-down approach”, allowing the designer to consider the need for providing a non-visual interface, and finding a solution from a suite of technologies and methods. The developer should ideally have knowledge of tools and affordances, when designing feedback. However, this is not always the case. An assistive framework would benefit both experienced and non-experienced developers, allowing them to adhere to a standardized set of mappings.

## Research goals & contribution

- To design and refine assistive haptic feedback to inform positional awareness and to provide navigational feedback when exploring a web page. The aim is to develop haptic sensations which improve understanding of the information presented and heighten levels of subjective experience.
- To develop targeted haptic feedback for visually impaired web users to assist them when working on collaborative tasks in conjunction with their sighted colleagues. The aim is to synchronize the presentation of both non-visual and visual-oriented browsing technologies.
- Establish an assistive vocabulary, with accompanying recommendations covering issues of haptic salience, aesthetics, device and perceptual constraints. The framework aims to contribute to the limited body of knowledge on haptic interface design for the Web, and provide a valuable reference tool to support developers in both the design and evaluation processes.

## Methodology

An exploratory approach is required to develop assistive sensations, due to the limited level of research in haptic interface design for an essentially two-dimensional interface. Scenario-based techniques are utilized as a tool for inspiring and evaluating design ideas for assistive feedback. The scenarios themselves are based upon the observation and task analysis of visually impaired screen-reader users accessing the Web. This process is thought to identify interactions made during the exploration and collaboration process, which could be improved through the use of assistive haptic feedback.

The scenarios are presented within the context of an assistive web interface, allowing small design teams to consider the unique qualities offered by touch [3], and how these can be used to inform design. Affordances and metaphors appropriate to the haptic domain will be considered, to produce sensations which convey an underlying sense of meaning to the user. A participatory approach is applied to acquire the perspectives from visually impaired screen reader users and haptic interface designers, both of whom are thought to be vital in the process of targeted assistive haptic feedback design. The method enables teams to consider how haptic sensations can work in conjunction with

other haptic stimuli suggested for use on a page. Interacting with an input device other than a keyboard is often a new experience for a screen reader user, so design should also consider the additional support required when using a mouse. Prototypes can be refined to a sufficient standard through follow-up sessions, where designs are evaluated against the same scenario, and against additional scenarios exploring other contexts of use (e.g. when gaining general or in-depth knowledge from a page).

## Current status & open issues

One comparison study profiling the use of a screen reader against using assistive force-feedback has been conducted [8]. Scenario-based design techniques have been used to develop targeted assistive feedback for the purposes of assisting a visually impaired user to perform a web search. Three iterations of design have been completed, with a fourth in progress. Lower-fidelity prototypes of commonly occurring page elements have been strengthened and implemented dynamically on web pages using the content-aware plug-in discussed by Yu et al. [8]. Findings from the research will form the basis of the assistive vocabulary. The next logical step would be to design sensations which support visually impaired users when collaborating on web-based tasks with sighted users. A method for evaluating the resulting feedback would also need to be formulated.

## References

1. Betacom Screen Rover. (<http://www.betacom.com>)
2. Jacko, J.A. et al. Older Adults and Visual Impairment: What Do Exposure Times and Accuracy Tell Us About Performance Gains Associated with Multimodal Feedback? In Proceedings of HFCS'03 (Fort Lauderdale, FL, 2003), 33-40.
3. MacLean, K.E. Designing with Haptic Feedback. In Proceedings of ICRA (San Francisco, CA, 2000), 783-788.
4. Mynatt, E.D. & Weber, G. Nonvisual Presentation of Graphical User Interfaces. In Proceedings of CHI'94 (Boston, MA, 1994), 166-172.
5. O'Modhrain, M.S. & Gillespie, B. The Moose: A Haptic User Interface for Blind Persons. In Proceedings of Third WWW6 Conference (Santa Clara, CA, 1997).
6. Petrie, H. & Hamilton, F. The Disability Rights Commission Formal Investigation into Website Accessibility. In Proceedings of HCI'04 (Leeds, UK, 2004), 157-158.
7. Yu, W. & Brewster, S.A. Evaluation of Multimodal Graphs for Blind People. *Universal Access in the Information Society*, 2, 2 (Jun. 2003), 105-124.
8. Yu, W. et al. A Novel Multimodal Interface for Improving Visually Impaired People's Web Accessibility. *Virtual Reality*, 9, 2 (Jan. 2006), 133-148.

## Acknowledgements

Thanks to Dr. Wai Yu and Dr. Graham McAllister at Queen's University Belfast, for their support and advice. The research is funded by Eduserv - <http://www.eduserv.org.uk>

# Multi-Layer Dialog Generation for Non-Visual Web Access

*Yevgen Borodin*

Department of Computer Science

Stony Brook University

Stony Brook, NY 11794, USA

borodin@cs.sunysb.edu

## Abstract

People with visual disabilities use screen-readers to browse the Web. The existing screen-readers have limited ways of presenting Web page content. I propose to turn non-visual Web browsing into a multi-layer mixed-initiative dialog-based interaction between users and computers. The suggested layers of dialog navigation are: basic screen-reading, DFS or BFS, and domain-specific. The support of adaptive dialogs is also planned. This research is aimed at improving and accelerating non-visual Web browsing for blind users.

## Introduction

The Web has become an indispensable aspect of our lives. In our daily activities we turn more and more to it for information: from checking the weather to searching for air tickets or getting directions. We use it for reference, news, shopping, etc. The Web was designed for graphical modes of interaction, making all these activities simple for sighted users. Users with visual disabilities, however, have to use screen readers, which process Web pages sequentially making the process of Web browsing complicated and time-consuming.

A typical screen reader simply speaks out the content of a page to the user. The only way to interact with a screen reader is by means of shortcut keys, which allow skipping paragraphs, searching, pausing, changing the rate of speech, etc. Some screen readers also provide their users with extended features, such as summarization [11], lists of current and visited links [3], etc. All these features greatly improve Web-browsing. However, they provide little control over the way the information is presented to the user. Also, the interaction between the user and the screen reader is one-sided and requires improvement.

In this paper I propose to study dialog generation for non-visual Web access to provide blind people with better ways of exploring Web content. This research will improve non-visual Web browsing by offering multi-layer dialog-based interaction with screen readers. The approach will build on features of the best screen-readers, extending the levels of interaction and information presentation.

## Related work

To provide blind users with multi-layer Web access, the system will have to implement non-visual Web browsing; use a collection of methods for Web-page segmentation and analysis; and implement dialog generation and processing.

**Non-Visual Web Access** has been addressed by a number of research projects [9, 1]. Among the well-known screen readers there are JAWS [3] and IBM's Home Page Reader [8]. BrookesTalk [11] summarizes Web pages. Braille-Surf [2] converts HTML into Braille. These systems give their users reasonable facilities to browse the Web. However, most screen readers present information sequentially and provide little control over the form of presentation. In contrast, I propose to give visually disabled users access to Web-page content at multiple levels, keeping all of the standard features of the best screen readers.

**Web Content Analysis** uses segmentation of Web pages into blocks of data, which are then analyzed, grouped, labeled, etc. Most of the techniques used for segmentation are either domain specific [8] or rely on sets of manually specified rules [10]. Some of these approaches are not suitable for dynamically changing Web sites. I build on the previous work on structural and semantic analysis described in [7, 5]. The geometric partitioning method I am using is fully automated and scalable over domains and does not depend on manually specified rules or domain knowledge.

**Dialog Generation** is a large area of research in NLP and substantial research has been done on various aspects of dialog generation, including VoiceXML dialogs. Mixed initiative dialogs and generation of cooperative responses using VoiceXML are described in [4]. An example of use of VoiceXML dialogs in Web browsing is described in [6]. However, dialog generation for non-visual Web Browsing is not a well-studied subject. My work is based on and extends the research described in [7].

## Approach

Non-visual Web navigation can be viewed as a mixed-initiative dialog between a human and a Web browser, in which the person asks questions and the browser gives answers or asks for clarification. The browser can make suggestions and prompt the user for input, e.g. if a Web form has to be filled out. User input can be in text or speech. The new approach will allow to access Web pages at three different layers of dialogs. The system will generate all three layers and will allow the user to switch between these layers and browse the Web at the level, which is most convenient for a given task.

**Basic Screen-Reader** functionality is the first layer of dialog generation. At this level the interaction between the user and the screen-reader is limited to an event-driven dialog, in which the page is simply read out to the user. The user can press standard screen-reader shortcut keys that control the flow of the dialog, e.g. skip, pause, repeat, change voice settings, etc. All these shortcuts are implemented within the dialog interpreter, while the dialog itself can be a simple narration, in which the screen-reader has the initiative and the user only selects the dialog direction.

**DFS/BFS** navigation is the second level of content presentation, requiring more sophistication in generating dialogs. The user should be able to freely navigate on the Web page using breadth-first, depth-first, and mixed approaches. The user will be able to choose which part of the page to listen to, at which level of detail, as well as get brief summaries of the page parts. The implementation of this level will require a mix of simple partial-document summarization techniques to be able to abstract, classify, and label different types of data.

**Domain-Specific** level of dialog generation will require specialized templates, e.g. news, shopping, etc. The dialog generator will analyze the Web page content and try to fill the corresponding template with information. The resulting dialog will give the user a more structured view of the page. The implementation of this layer may also require the use of classifiers to determine to which domain a page belongs.

Dialogs will be enriched with earcons, special sounds that could help distinguish between plain text, links, visited links, taxonomies, etc. Dialogs will be made adaptable to the user's choice of vocabulary, verbosity, navigation style, etc. This will require the use of sophisticated speech grammars to be able to interpret a variety of user commands.

## **Project status**

### ***Infrastructure***

This work is a part of the HearSay [7] project (supported by NSF grants CCR-0311512 and IIS-0534419). The HearSay system has basic non-visual browsing facilities and can perform simple structural and semantic analysis of Web page content. The system provides a frame tree representation of Web pages, which can be further analyzed, grouped, and partitioned. The frame tree content is used to generate VoiceXML dialogs that will then be processed by a voiceXML interpreter. To be platform independent, the system is being developed entirely in JAVA.

### ***VoiceXML interpreter***

Since no suitable open-source VXML interpreter was available, I am implementing a custom VXML interpreter that will comply with the VoiceXML specification and allow both voice and keyboard input. However, the interpreter is designed to go beyond the specifications and provide more control over the dialogs, speech properties, and event-handling.

### ***Dialog generator***

The HearSay system already implements the first layer of dialog interaction, supporting basic screen-reading and extended speech controls. The implementation of the second layer of dialog generation will start as soon as reasonable partial-document summaries can be obtained. Dialog templates have to be developed for the third layer of domain-specific dialogs. And, finally, the dialog generation system and the VoiceXML interpreter have to be extended to support adaptive dialogs.

### ***Summarization***

The second layer of dialog generation will require the use of summarization. Most of the research on Web page and document summarization concentrated on full-and multi-document summarization. I am currently investigating summarization techniques that will be able to provide intelligent labels, summaries, and abstracts of parts of a Web page. The choice of techniques will depend on the type of information in any given part of the Web page.

## **Evaluation**

The research goals of this project have been formulated through collaboration with Helen Keller School for the Blind (HKSB) at Hempstead, NY. The design of the project is guided by individuals with visual disabilities who are teachers at HKSB. The ideas have been obtained and clarified in meetings with instructors and students of the school. A series of progressive evaluations will be performed by the students of the school as the system takes shape.

## **5 Conclusion**

In this paper I proposed to research dialog generation techniques as part of the ongoing HearSay project [7] for non-visual Web browsing. The goal of my research is to provide better ways of presenting Web page content by means of audio. A multi-layer dialog-based interaction has the potential to make the Web more friendly and accessible for people with visual disabilities. The system can be also adapted to provide Web access by phone.

## **References**

1. C. Earl and J. Leventhal. A survey of windows screen reader users: Recent improvements in accessibility. In *Journal of Visual Impairment and Blindness*, 1999.
2. D. Hadjadj and D. Burger. Braillesurf: An html browser for visually handicapped people. In *Proceedings of Tech. and Persons with Disabilities Conf.*, 1999.
3. <http://www.freedomscientific.com>.
4. K. Komatani, F. Adachi, S. Ueno, T. Kawahara, and H. G. Okuno. Flexible spoken dialogue system based on user models and dynamic generation of voicexml scripts.
5. S. Mukherjee, I. Ramakrishnan, and A. Singh. Bootstrapping semantic annotation for content-rich html documents. In *Intl. Conf. on Data Engineering (ICDE)*, 2005.
6. <http://www.internetspeech.com>.
7. I. Ramakrishnan, A. Stent, and G. Yang. Hearsay: Enabling audio browsing on hypertext content. In *Intl. World Wide Web Conf. (WWW)*, 2004.
8. H. Takagi, C. Asakawa, K. Fukuda, and J. Maeda. Site-wide annotation: Reconstructing existing pages to be accessible. In *ACM Intl. Conf. on Assistive Technologies (ASSETS)*, 2002.
9. G. Weber. Programming for usability in non-visual user interfaces. In *ACM Intl. Conf. on Assistive Technologies (ASSETS)*, 1998.
10. S. Yu, D. Cai, J.-R. Wen, and W.-Y. Ma. Improving pseudo-relevance feedback in web information retrieval using web page segmentation. In *Intl. World Wide Web Conf. (WWW)*, 2003.
11. M. Zajicek, C. Powell, and C. Reeves. Web search and orientation with brookestalk. In *Proceedings of Tech. and Persons with Disabilities Conf.*, 1999.

## **ViCRAM: Visual Complexity Rankings and Accessibility Metrics**

*Eleni Michailidou*

School of Computer Science

The University of Manchester

Manchester, M13 9PL, UK

michailidou@cs.man.ac.uk

### **Abstract**

The World Wide Web (Web) has become the means of distribution and use of information by individuals around the world. However, access to this information by visually impaired people is limited due to the Web's visual complexity. ViCRAM is a project that will relate user's implicit understanding of Web page visual complexity with its layout. Eye tracking methods and knowledge acquisition techniques will be used to elicit sighted people's visual perception. During this project we will also determine whether pages that sighted users identify as visually complex are complex for visually impaired users as well, from a Web accessibility perspective. We aim to develop a heuristic framework that will be used for describing Web page's visual complexity and as guidelines for transcoding a Web page into a less visually complex and more accessible one.

### **Introduction**

Most Web pages focus on visual presentation to implicitly help users understand and interact with the content. When sighted users reach a Web page, they can scan the page and get a comprehension of it in an average of 5 seconds. This view helps them to decide if the page is relevant to their task and move towards the part of the page that interests them. On the other hand, if visually impaired users want to get an idea of how the page looks they spend much more time because they have to listen to the entire page being read from the top left corner of the screen to the bottom right. This is because assistive technologies, such as screen readers, render the source code of the page. In addition, they might spend even more time because the source code is not always accessible. This happens when designers do not always follow the correct guidelines for accessibility and use different coding conventions to represent page elements such as headings and links [1].

ViCRAM is a project that will mainly contribute to the improvement of accessible Web design. We will study sighted users' behaviour and eye movement while interacting with a Web page to elicit their implicit knowledge of visual perception. In addition, we will investigate the coping strategies visually impaired users develop to interact with complex Web pages. A relationship will then be distinguished between Web page visual complexity, sighted users' cognition and visually impaired users' coping strategies. The project's objective is to create a framework that will be used to identify Web page visual complexity for two purposes: to give feedback to the user regarding the presentation of

the page; and to help reduce the visual clutter of the page by using it as a guide for the Web page transcoding process.

## **Related work**

Within the literature, visual complexity of an HTML document is described through Web site accessibility and usability research. Studies try to identify Web page design metrics that determine whether a site is complex. These studies relate Website design guidelines with complexity explaining that the way a Website is presented depends on the way its page is designed and what elements are used. For example, Ivory et al. [4] performed a quantitative analysis on Web page attributes (e.g. number of images) using a large collection of Web sites. In this case, page composition metrics, such as word count, could distinguish between good and bad pages with respect to design and usability.

Moreover, eye movement behaviour while interacting with a Web page can give further information regarding Web page design issues. Eye tracking studies are mostly performed for Website usability evaluations. Poynter [5] is one such study with large collections of Web pages in their evaluation set. The authors extensively tested eye movements across several news homepage designs and noticed a common pattern between a user's fixation order on the page (scanpath) and the page layout. In [6], the authors show that Website viewing behaviour is driven by gender, the order of Web pages being viewed and the interaction between Website types. Their study also revealed a possible relationship between scanpath variability among individuals and the structural/visual complexity of the Web page.

Throughout the literature, studies show that user's cognition, Website content and the way information is grouped and presented affects Website usability and accessibility. We want to determine how these factors correlate and how we can define Web page visual complexity.

## **Research goals**

ViCRAM aims to relate user's implicit understanding of a visually complex Web page with its layout. In this way, Web page designs can be associated with a common Web behaviour and visual complexity that will give further insight into accessible Web page design and facilitate visually impaired users. In order to achieve this we will answer three important questions: "What is a visually complex Web page?", "How does a sighted user interact on a Web page?" and "What is the relationship between a sighted user's Web page interaction and its visual complexity?".

A variety of usability and accessibility studies show that complicated Web page structure depends on the page composition and layout [2, 3, 4]. During our research we will identify how Web page elements, such as links, text, and images, interact with each other to produce a visually complex page. Sighted users' Web behaviour using eye movement tracking methods can give better perception of page presentation. This is because users' glancing habits can give important information on where they look first when they reach a Web page, where they pay more attention and for how long they concentrate on specific parts of the page. A related literature survey, a study using knowledge elicitation methodologies, such as card sorting, and users' eye movement behaviour on Web pages will be our research methods for learning more about human visual perception.

Our objective is to develop a framework to distinguish the visual complexity of a Web page. This framework will provide information about the level of visual complexity to visually impaired users so less time will be needed to get an overview of the page. In addition, the framework will be used as a guide for transcoding a page into a less visually complex one.

### **Current research status**

To date, we were able to perform an empirical study to extract more information on how sighted users describe a visually simple or complex Web page. By using card sorting and triadic elicitation techniques we were able to encourage participants to articulate their opinion and extract their implicit knowledge with respect to page presentation. During this study we concluded that the main factors that affect the visual complexity of a Web page are the diversity (how many different elements are used), density (how many of each element are used) and position of the page elements (e.g. links, words, images). That is, a visually simple page has a significantly lower number and variety of page elements than a visually complex page. The most important difference between visually simple and complex pages is the number of different subjects that each present. For example, visually complex Web pages like MSN (<http://www.msn.co.uk>) provide a variety of information to a reader (news, entertainment etc.) where visually simple pages like the Mint Group (<http://mint.cs.man.ac.uk>) focus on one subject. Hence, fewer page composition elements are needed to organize a visually simple page than a complex one.

In addition, we were able to perform a pilot study-using eye tracking methods. We wanted to investigate how sighted users perceive the visual presentation of Web pages, where they look first and for how long. A descriptive analysis of this study's results revealed some interesting patterns. For example, Web page salient elements, such as big logos, pictures and animations, attracted the subjects' attention first but the main content of the page attracted them for the longest. Also, menus were not completely read and no specific reading order between right and left menus or columns was determined. In addition, this pilot study gave implicit information on user's common interaction scanpaths. After gazing at dominant graphics the participants scanned through the main content of the page. They did not fully read the text, in the main content area, and they paid more attention to the links and the first sentence of each block or paragraph. Then, the participants tended to look and read the first three links on the menu, either on the left or right hand side of the screen. A quantitative analysis could not reveal solid conclusions due to experiment design problems, such as participants' page familiarity and imprecise task assignment. However, we did notice a possible relationship between visually complex designs identified by our empirical study and participants' Web behaviour. For example, participants spent more time interacting with a visually complex page than a simple page. In addition, the more visually complex the page was, the more scattered and disordered the participants' scanpath was.

### **Future research plan**

The next stage of our research is to perform an empirical investigative study. Sighted users will compare and rank Web pages from a visual perspective. We will investigate how font size, image location and menu size affects sighted users' perception. Then, we

will define heuristics that describe the visual complexity of a Web page. These heuristics will be based on the users' feedback of gazing behaviour and visual perception taken throughout our research.

We plan on evaluating our framework to determine whether it gives sufficient results using both sighted and visually impaired users. Sighted participants will be used for an eye tracking study. We will determine the significance between the eye tracking data and our framework feedback. Visually impaired participants will evaluate the visual complexity feedback and the transcoded pages with respect to the time needed to understand the layout of the page with and without the feedback.

## Conclusions

ViCRAM's main contribution is to define a framework that describes visual complexity of Web pages. The framework will be defined by the relationship between a Web page's visual presentation and structure, along with a sighted user's interaction and browsing behaviour. It will be used for both giving feedback directly to designers regarding visual clutter in a page and as guidance for transcoding a page to result in a simpler and more accessible one. Visually impaired users will then be able to access Web pages faster and easier than they used to.

## References

1. C. Asakawa. What's the web like if you can't see it? In W4A '05: Proceedings of the 2005 International Cross-Disciplinary Workshop on Web Accessibility (W4A), pages 1-8, New York, NY, USA, 2005. ACM Press.
2. M. Germonprez and I. Zigurs. Causal factors for web site complexity. *Working Papers on Information Environments, Systems and Organizations*, (3), 2003.
3. M. Y. Ivory, R. R. Sinha, and M. A. Hearst. Preliminary findings on quantitative measures for distinguishing highly rated information-centric web pages. In *Proceedings of 6th Conference on Human Factors and the Web*, 2000.
4. M. Y. Ivory, R. R. Sinha, and M. A. Hearst. Empirically validated web page design metrics. In *CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 53-60, New York, NY, USA, 2001. ACM Press.
5. S. Outing and L. Ruel. *Eyetrack III: What We Saw Through Their Eyes*, 2006. A Project of the Poynter Institute, <http://www.poynterextra.org/eyetrack2004/index.htm> (Accessed: January 2006).
6. B. Pan, H. A. Hembrooke, G. K. Gay, L. A. Granka, M. K. Feusner, and J. K. Newman. The determinants of web page viewing behavior: an eye-tracking study. In *ETRA'2004: Proceedings of the Eye tracking research & applications symposium on Eye tracking research & applications*, pages 147-154, New York, NY, USA, 2004. ACM Press.

## Designing Cognitive Supports for Dementia

*Joe Wherton and Andrew Monk*

Centre for Usable Home Technology

Department of Psychology, University of York

York, YO10 5DD, UK

[j.wherton@psych.yrk.ac.uk](mailto:j.wherton@psych.yrk.ac.uk)

[a.monk@psych.york.ac.uk](mailto:a.monk@psych.york.ac.uk)

### Abstract

This paper gives an overview of a PhD programme focusing on the design of cognitive prostheses to support people with mild-moderate dementia in the home. The approach taken involves two main phases. The first phase explores the problems that need to be addressed, from the perspective of professional carers, patients, and informal caregivers. The second phase focuses on designing a cognitive support that can assist patients through a specific activity. This paper reports the findings from the first phase (exploring problems of dementia in the home), and describes plans for the subsequent designing phase.

### Introduction

During the early to middle stages of dementia, cognitive deficits are most profound in episodic memory (memories of experienced events) and executive function processes (required for planning, sequencing, and attentional control) [3]. Executive function processes are essential for goal-directed behaviour, and so impairment in this domain severely disrupts patients' ability to perform everyday tasks [1]. Existing concepts aimed at supporting dementia patients through daily activities involve systems that monitor task progress and provide prompts when necessary. Mihaildis et al. [4] devised the Cognitive Orthosis for Assisting Activities in the Home (COACH), which is designed to support hospitalised dementia patients when washing their hands. The system tracks progress during the activity and provides audio cues when the patient fails to initiate an action or performs an action out of sequence. Similarly, Dishman [2] devised a prototype that uses radio frequency identification (RFID) tags to monitor progress when making a cup of tea, and provides video clips of to-be-performed actions on a monitor located on the worktop.

Developments in motion tracking and artificial intelligence demonstrate a real potential in implementing technological supports across a broad range of tasks. However, little work has been done to explore the requirements of such systems so that they meet the needs of the users. The project aims to address this issue by designing a cognitive prosthesis that is grounded in an understanding of the problems patients experience during daily activities. The first phase includes two interview studies exploring the problems of dementia in the home [7,8]. The second phase focuses on the designing and evaluation of a Wizard of Oz prototype. This section will summarise the findings from the interview studies, before going on to describe plans for the designing phase.

## Problems of dementia in the home: Professional carer and patient-caregiver perspectives

Two interview studies were conducted to identify aspects of daily living that are important to support. In the first study, nine interviews and one focus group were conducted with 20 care assistants and Occupational Therapists (OTs). In the second study five mild-moderate dementia patients and seven informal caregivers were interviewed in their own home. All transcripts were analysed using Grounded Theory Analysis [6].

### Study 1: Professional carer perspective

The analysis revealed three themes that summarise the problems of dementia in the home, from the perspective of the professional carers. These were ‘Problems in the Home’, ‘Underlying Deficits’, and ‘Consequences for Patients and Caregiver’. The themes consisted of ten main categories, and 29 sub-categories. Figure 1 depicts the three themes and their constituent categories.

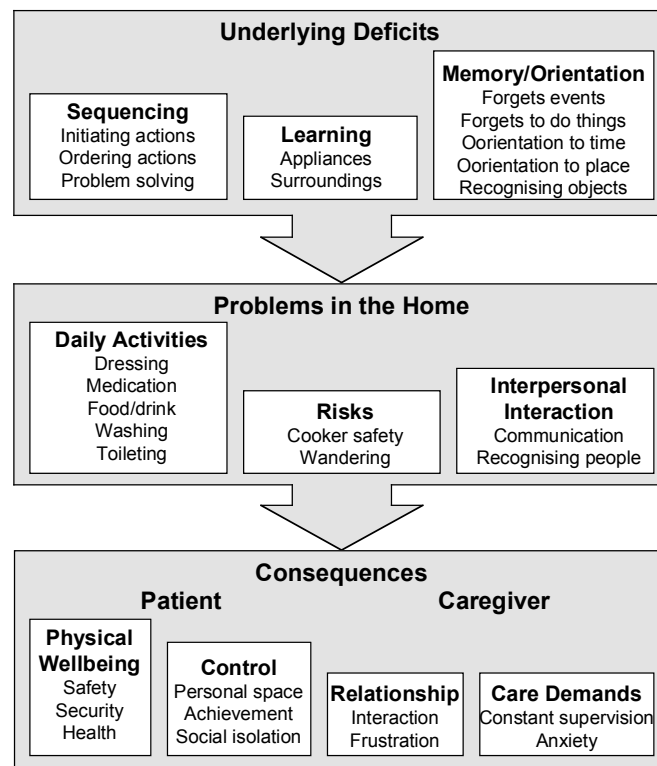
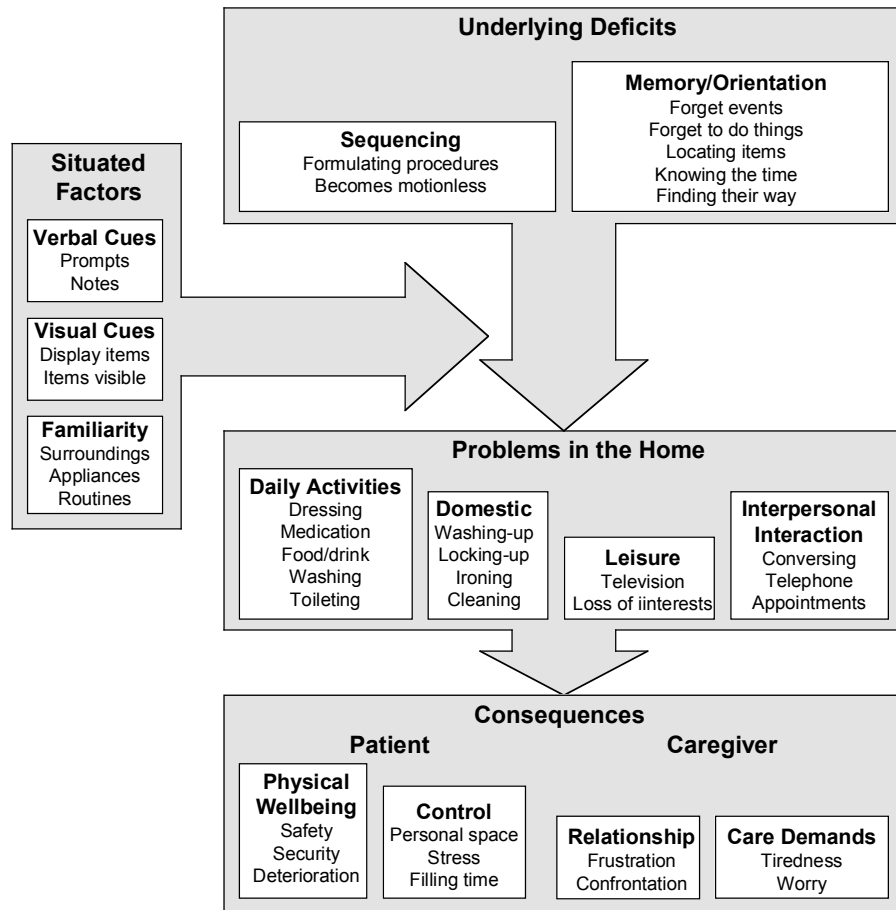


Figure 1. A summary of the problems in the home, as perceived by professional carers

### Study 2: Patient-caregiver perspective

Interviews with patients and informal caregivers revealed four themes. Three themes covered similar issues raised in Study 1 (Problems in the Home, Underlying Deficits, and Consequences for Patients and Caregivers). The fourth theme covers ‘Situational Factors’ that influence the impact of the underlying deficits on patients’ functional status. The

themes composed of 13 main categories and 38 sub-categories. Figure 2 depicts the four themes and their categories.



**Figure 2.** A summary of the problems in the home, as perceived by patients and informal caregivers.

***Video analysis of kitchen tasks***

The next phase of the project will explore how technology can be designed to support patients through a daily task. Problems in cooking were reported frequently in the interview studies, and so this activity will be the focus for the design. In order to further understand the types of problems that occur during kitchen tasks patients are being video recorded performing activities in their own kitchens. Actions will be transcribed using the Action Coding System, devised by Schwartz [5], and action errors will be classified based on behavioral characteristics. In conjunction with cognitive theory the error classifications provide a basis to address specific problems and design strategies that are compatible with patients’ cognitive capabilities. Preliminary analysis has shown that problems occur with regards to sequencing (intrusion, omission, and repetition), orientation (locating items and identifying items), operation of appliances, and incoherence (toying and inactivity).

## Conclusion

The interview studies provide an account of dementia in the context of the home. This helps direct research efforts towards aspects of daily living that are important to support. Slight discrepancies between the two perspectives illustrate the benefits in using different elicitation methods and sources of information to gain an insight into the various problems. The naturalistic observations highlight aspects of the task that a technological support should address. Understanding how these errors relate to executive function deficits is an essential part in designing appropriate methods of support. The next phase will test different strategies of support using a single case experimental design.

## Acknowledgements

The work is part of the Dependability Interdisciplinary Research Collaboration (DIRC) funded by EPSRC (GR/N13999).

## References

1. Bell-McGinty, S., Podell, K., Franzen, M., Baird, A. D., and Williams, M. J. Standard measures of executive function in predicting instrumental activities of daily living in older adults. *International Journal of Geriatric Psychiatry*, 17 (9), 2002, 828-834.
2. Dishman, E. Inventing wellness systems for aging in place. *Computer*, 37 (5), 2004, 34-41.
3. Graham, N. L., Emery, T., and Hodges, J. R. Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia. *Journal of Neurology, Neurosurgery, and Psychiatry*, 75 (1), 2004, 61-71.
4. Mihaildis, A., Barbenel, J. C., and Fernie, G. The efficacy of an intelligent cognitive orthosis to facilitate handwashing by persons with moderate to severe dementia. *Neuropsychological Rehabilitation*, 14 (1/2), 2004, 135-171
5. Schwartz, M. F., Reed, E, Montgomery, M., Palmer, C., & Mayer, N. H. The quantitative description of action disorganisation after brain damage: A case study. *Cognitive Neuropsychology*, 8 (5), 1991, 381-414.
6. Strauss, A. L., and Corbin, J. *Basics in qualitative research: Grounded theory procedures and techniques*. Sage Publications: Newbury Park, California, 1990.
7. Wherton, J. Exploring the requirements of cognitive support for dementia. In *Proceedings of the home-oriented informatics and telematics (HOIT) conference*, Volume 2, 45-50, (York, UK, April 13-15, 2005).
8. Wherton, J., & Monk, A. Cognitive support for dementia. *CHI'06 Workshop: Designing technology for people with cognitive disabilities* (Montreal, Canada, April 22-23, (2006).

# Understanding and Supporting the Use of Accommodating Technologies by Adult Learners with Reading Disabilities

*Katherine Deibel*

Dept. of Computer Science and Engineering

University of Washington

Seattle, WA, USA 98195

deibel@cs.washington.edu

## Abstract

Developing computer accommodations for users with reading disabilities involves several challenges: diversity of needs, stigma risks, and self-advocacy issues. This paper proposes a two-fold approach to address these issues. First, participatory design with reading-disabled users will inform necessary directions for technology development. Second, to help individual users identify what accommodations can benefit them, intelligent software will be developed. This software will also aid in the configuration of the accommodations.

## Introduction

The importance of reading in today's information society is without debate. This is particularly true in the later years of education (high school, college, etc.) when students move from learning to read to reading to learn [11]. However, for those with reading disabilities, the task of reading can become challenging and stressful. Developing digital technology, like electronic documents and computer-based reading, to support and accommodate this population involves several inherent challenges.

One challenge is due to the size and diversity of this user population. Dyslexia, one of the more common reading disabilities, is estimated to affect 7–15% of the population [12]. Moreover, nearly half of all university students in the U.S. who register as having a disability experience significant difficulty with reading [10]. With such large numbers, there is an unsurprising diversity of abilities, needs, and degrees of severity [7]. How one attends to these differences influences the universal helpfulness of an accommodation.

Another challenge is related to the notion of self-advocacy. Students with invisible disabilities tend to actively avoid seeking out help until a crisis motivates them [1]. This hesitancy has been traced to numerous causes, including low self-confidence due to previous bad experiences asking for help [2] and wanting to avoid the stigma of having a disability [1]. This fear of stigma can interfere not only with the adoption of technology but also in recruiting study participants [4]. Considering these issues of stigma and self-advocacy could potentially prevent a beneficial technology from collecting dust on a shelf unused.

This paper proposes a project to create a computer-based accommodation that will address these challenges. The key intention is that embracing the diversity of needs and

the notion of self-advocacy will help identify how to design technologies that not only benefit the reading process but are also usable and adopted into regular use.

## **Previous work**

These notions of being beneficial, usable, and adopted provide a means for analyzing previous work on computer accommodations for reading disabilities. In terms of being beneficial, most of the previous research in using computers to accommodate reading disabilities has focused on text-to-speech technology (TTS) in systems like Bookwise [4] and Kurzweil 3000R[3]. Hearing the text read aloud improves word identification and can even be used for reading remediation for some people [4, 12]. TTS does require strong auditory skills to be effective, resulting in a significant subpopulation that remains unhelped by TTS.

The usability and learnability of TTS systems appear to have been mostly unstudied. The SeeWord project [7] studied interface design for people with reading disabilities. Some of their key findings include an emphasis on direct manipulation, the need for fine-tuning of visual features, and consistent design. An inspection by the author found that most TTS systems do not use these design recommendations.

Finally, the reasons why an accommodation is adopted or not has also received little research. Elkind et al. studied factors that influenced the use of Bookwise in the workplace by adults; positive factors included an environment supportive of disabilities and motivation to adapt their work practices accordingly [4]. At the college level, however, only the motivations behind students seeking help have been studied extensively [1]. Little is known about how students with reading disabilities adopt digital technologies (accessible or not) to support their reading process.

## **Research plan**

Working from a philosophy that design should be influenced both by the needs and wants of users and the potentials of technology, this project involves two research threads.

### ***Understanding the needs of the users***

The research question in this thread is to understand the users' relationship with technology used for reading to inform new technologies and/or improvements to existing solutions. This process began with informal interviews among students with disabilities. These casual conversations were exploratory and meant to improve understanding of the problem space and to identify potential research questions. Two of the key findings were the importance of privacy to the students in regards to their use of AT and the difficulty in learning and configuring accessible technology.

Future plans are to formalize these results through several approaches. To gain perspective on how digital technologies are used by this population, a general survey on technology use is being developed that will potentially be distributed at multiple universities. Participatory design sessions using the future workshops method [9] will be conducted to identify directions on how to improve reading technology.

### ***Technology, diversity, and self-advocacy***

Given the diversity of needs among those with reading disabilities, a single accommodation is unlikely to be effective. Instead, one can provide a collection of interoperable tools for the user to choose among. This toolkit would be a component of basic computer-reading application and would contain tools like text-to-speech, text magnification, etc. Findings from the previous thread will likely suggest other accommodations to include. Each tool is then configured as needed. Importantly, though, by letting a user select only what he or she needs, this can address both the challenges of diversity and self-advocacy.

Letting users customize a base application through extensions is not a new idea; the Firefox web browser [6] does exactly that. However, Firefox provides little support in helping users determine what would benefit them. With a large number of tools, finding the few useful ones becomes difficult. Moreover, the identified tools still need to be configured. Given that fine-tuning is often required for accommodating reading disabilities [7] and that informal interviews identified configuration as a known user issue, the toolkit is not yet a sufficient solution.

Thus, a software wizard will be integrated with the toolkit to guide each user to appropriate accommodations and walk them through the configuration process. The wizard will use assessments of varying complexity: from simple preference questions to formal performance measurements. The key idea is that the wizard will be an expert system that applies the knowledge and procedures similar to those that a human disability expert would utilize when helping someone. Since the user will not have to approach others for help, the stigma risk is lessened. Furthermore, the toolkit plus the wizard gives the users a chance to advocate for their own accommodations.

To demonstrate this approach, I will develop software demonstrating the key aspects of the expert system. In particular, I will implement a software version of an established diagnosis protocol for an accommodation with a large configuration space. The accommodation addresses visual stress syndrome, a condition involving excessive eyestrain during reading that affects 15–20% of people with dyslexia [5, 13]. The standard treatment consists of an optometric examination and reading test that identifies an optimal color for the text's background [13]. This procedure, including the oral reading test, will be automated and tested for validity and reliability. This system will likely be extended to other typographic parameters in light of recent research on their potential as an accommodation [8].

### **Contributions**

This research has several contributions. First, a better understanding of the user population, their needs, and their views on technology will help inform future work. Second, emphasizing self-advocacy and diversity forms a new design approach for accessible technology that not only addresses disabilities but also usability and adoptability. Finally, the suggested approach of an intelligent toolkit of accommodations has further applications. The same technology could aid disability experts and caretakers in their jobs or even benefit the general user in customizing complex software applications.

## References

1. R. C. Cory. Identity, Support And Disclosure: Issues Facing University Students With Invisible Disabilities. PhD thesis, Syracuse University, Syracuse, NY, USA, 2005.
2. J. H. Edwards. *The Scars of Dyslexia: eight case studies in emotional reactions*. Cassell, London, 1994.
3. J. Elkind. *Computer Reading Machines for Poor Readers*. Lexia Institute, Portola Valley, CA, USA, 2001.
4. J. Elkind, M. S. Black, and C. Murray. Computer-based compensation of adult-reading disabilities. *Annals of Dyslexia*, 46:159–186, 1996.
5. B. J. W. Evans. *Dyslexia and Vision*. Whurr Publishers, London, 2001.
6. Firefox web browser, 2006. <http://www.mozilla.com/firefox>.
7. P. Gregor, A. Dickinson, A. Macaffer, and P. Andreasen. SeeWord— a personal word processing environment for dyslexic computer users. *British Journal of Educational Technology*, 34(3):341–355, June 2003.
8. A. R. Grounds and A. J. Wilkins. Separating the visual and semantic factors that affect reading speed in dyslexic children. In preparation.
9. F. Kensing and K. H. Madsen. Generating visions: Future workshops and metaphorical design. In J. Greenbaum and M. Kyng, editors, *Design at Work: Cooperative design of computer systems*, chapter 8, pages 155–168. Erlbaum, Hillsdale, NJ, USA, 1991.
10. L. Lewis, E. Farris, and B. Greene. An institutional perspective on students with disabilities in postsecondary education. *Statistical Analysis Report NCES 1999-046*,
11. U.S. Department of Education, Washington, DC, 1999.
12. L. Peer and G. Reid, editors. *Dyslexia—Successful Inclusion in the Secondary School*. David Fulton Publishers, London, 2001.
13. S. Sands and E. S. Buchholz. The underutilization of computers to assist in the remediation of dyslexia. *International Journal of Instructional Media*, 24(2):153–175, 1997.
14. A. J. Wilkins, R. J. Jeanes, P. D. Pumfrey, and M. Laskier. Rate of reading test: Its reliability, and its validity in the assessment of the effects of coloured overlays. *Ophthalmic and Physiological Optics*, 16(6):491–497, 1996.

## **An Evolutionary System for the Sc@ut Platform**

*Maria Dolores Paredes-Garrido*

Dpt. Lenguajes y Sistemas Informaticos

Computer Science Engineering High School

University of Granaca, Spain

lolapg@ugr.es

### **Abstract**

The Sc@ut platform is a series of tools for creating augmentative and alternative communication (AAC) systems. This platform is based on a two-level architecture: the communicator and the meta-communicator. The meta-communicator is used by designers (usually the child's educators or parents) to create and adapt the communicator. The communicator is used by a person with disabilities (normally a child) who interacts with it in order to communicate. In this paper, we will propose the addition of a new system to help designers and which provides the platform with automatic evolution and adaptation.

### **Introduction**

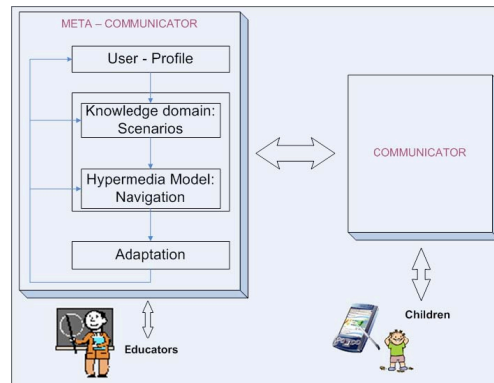
Augmentative and alternative communication systems (AAC) are a growing field of study. These are mainly concerned with providing devices and techniques to increase the communicative ability of people with disabilities which impede speech or which make their communication difficult to understand [3].

The Sc@ut project is a platform that allows adaptive communicators [4] to be created for people with certain communication needs due to disability. The communicator runs on a Pocket PC device, which has the advantages of being cheap, versatile and portable. We are currently focusing on autism, but due to the numerous settings that Sc@ut offers, we have designed communicators that can be adapted for other people with disabilities such as cerebral paralysis or dysphasia.

Next section of this paper presents Sc@ut's current architecture which comprises two levels and which we will see in the subsections. Following section describes the thesis to be developed over the coming years. Finally, we outline future work arising from this thesis.

### **Sc@ut architecture**

As we are currently working on autism, we will consider autistic children to be the communicator's end user. Figure 1 shows the Sc@ut architecture, which is organized, into two levels: the communicator and the meta-communicator. The first of these is used by the child, and the second is used by the child's educators, parents or teachers. In this figure, we can also see the complete process for designing and creating a communicator which has been adapted to one particular child. We will now briefly examine *who* the final users of each level are and what each level is for.



**Figure 1:** Sc@ut Process

### ***Communicator***

The communicator is a hypermedia system which enables the child to navigate pages with different scenarios (templates) and to select components from each page to express what he/she wants. Each child has their own communicator because of differences in their ability to understand language, their capabilities and their skills[4]. In order to collect information about the child's navigation, log files are created.

### ***Meta-Communicator***

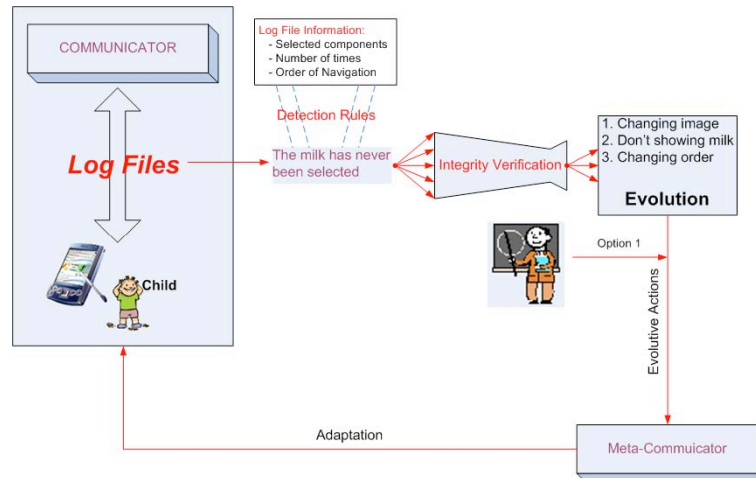
The meta-communicator is the application used by educators to create the communicator for each particular child. Figure 1 shows the steps taken to create the communicator.

## **3 Towards an evolutionary system**

At present, the communicator must be evolved and adapted manually. Our aim, however, is to semi-automate this process by designing a new system (the Evolutionary System) at the same level of the communicator system in order to carry out changes at runtime with the educators' intervention.

To adapt the communicator, the following steps must be followed (see Figure 2):

1. Educator defines detection rules informally to detect anomalous behaviours in the child
2. Evolutionary system translates detection rules into a formal language (colored net petri, set theory, ...).
3. System analyzes log files with detection rules.
4. System detects alerts of anomalous behaviours and searches the set of evolutionary actions (previously prepared) associated with them.
5. System checks communicator integrity with the Integrity Rules for each evolutionary action.
6. System shows evolution and change alternatives to educator. Every action shown is coherent with communicator.
7. Educator chooses an option available.
8. System applies evolutionary actions in order to make the changes automatically.



**Figure 2:** Example of evolution and adaptation

The following information can be obtained about the child's interaction from the log file: which templates, components or links between templates have been visited or selected, and when; the number of times that each one has been visited; navigation order followed; the time taken to observe a component, etc.

Educators must intervene to control the child's evolution. It is in this area that we should focus our effort with the help of psychology and pedagogy experts, and educators and parents who know the children and know whether their profile or environment has changed, making decisions about the most suitable changes at each moment.

Evolution and Adaptation rules must be formally specified and verified in order to allow the automation of changes in the communicator. Our ideas are based on previous work by members of our research group [2] (her thesis is supported by Brusilovsky's work which adaptation rules are being considered in this research within other authors), [5], [1], who uses formalisms as temporary logic and occurrence graphs of Petri nets. These rules must be specified with the help of psychologists and pedagogues of Sc@ut to ensure the right change.

## Conclusions and future lines of work

In this paper we have presented a platform for creating ACC systems. We have proposed the inclusion of a new system to manage the evolution and adaptation process so that it may better suit each user's characteristics and environment.

Our first line of work is to select and use a formal method or language for writing the evolution and adapting rules used during the process. Secondly, we aim to create a tool which uses the rules and which may be used by developers to help with the creation and evolution of the communicator. We are also interested in collecting information from the educators while the child is using the communicator and which can then be used to better understand the log files and to make decisions about changes. With this proposal, we are researching the exchange of files between the computer and the Pocket PC device. Finally, information about the environment (such as where the user is -ubiquity) can also be used to make decisions about user adaptation. We are interested in studying how this may be obtained and used.

## Acknowledgements

This research is supported by the Sc@ut Project (Consejera de Educacion de la Junta de Andalucia, Regional Government of Spain) and the ADACO project TIN2004-08000-C03-02 (CICYT, Spanish Government).

## References

1. L. Garcia Cabrera, M. Rodriguez Fortiz, and J. Parets Llorca. Evolving hypermedia systems: a layered software architecture. *Journal of Software Maintenance and Evolution: Research and Practice*, 14:389-406, 2002.
2. N. Medina Medina. Un Modelo de Adaptacion Integral y Evolutivo para Sistemas Hipermedia. El Sistema de Aprendizaje de SEMHP. Doctoral Thesis in Spanish. Granada, Spain, 2004.
3. P. Mirenda. Toward functional augmentative and alternative communication for students with autism: Manual signs, graphic symbols, and voice output communication aids. *Language, Speech and Hearing Services in Schools*, 34:203-216, July 2003.
4. M.D. ParedesGarrido, O. PinoMorillas, M. Rodr'iguez-F'ortiz, et al. A platform for creating adaptive communicators. *Lecture Notes in Computer Science*, 4061:847-854, July 2006.
5. M. Rodr'iguezFortiz, P. PaderewskiRodr'iguez, M. Rodr'iguezAlmendros, and M. Gea-Meg'ias. An approach for evolving an adaptive and unanticipated system: a communicator for children with autism. *Proceeding of the Workshop on Foundations of Unanticipated Software Evolution*, pages 40-47, February 2004.

## Virtual Sign Animated Pedagogic Agents To Support Computer Education For Deaf Learners

*Benjaporn Saksiri,<sup>1,2</sup> William G. Ferrell,<sup>1</sup> Pintip Ruenwongsa<sup>2</sup>*

<sup>1</sup>Dept. of Industrial Engineering, Clemson University, USA

<sup>2</sup>Institute for Innovation and Development of Learning Process, Mahidol University, Thailand

E-mail: bsaksir@clemson.edu, rsbss@mahidol.ac.th

### Abstract

In Thailand, there is a desperate need to improve the educational opportunities for deaf and hearing-impaired university-aged students. The research described herein is targeted at two aspects of this issue: 1) design and test a virtual sign animated instructional tool for the Thai sign language, and 2) investigate how instructors of university classes can most effectively teach deaf students and assess their performance, particularly by using this tool. Currently, I have completed a thorough review of the literature, constructed a 3-D human model that includes rudimentary facial expressions, and devised a framework to investigate the instructional process.

### Introduction

During my studies program at Mahidol University in Thailand, I became aware of the severe problems encountered by deaf and hearing-impaired university-aged students. (I will henceforth refer to all such students as “deaf.”) A little investigation uncovered two important facts. First, the opportunity for deaf students to learn Thai sign language is very limited. Second, deaf students primarily learn through their sense of vision unfortunately, very few instructors understand this or receive any training in the most effective way to educate deaf student. This research is aimed addressing these two problems. The first research focus is to construct and test a virtual sign animated instructional tool that will enable students to learn Thai sign language. The second focus is on investigating the efficacy of different techniques in the classroom relative to deaf students’ learning and assessment, particularly the possibility of using this tool, with the goal of providing instructors with insights into how they can more effectively teach the deaf.

At the heart of this research is computer animation which offers great potential for better supporting science and technology education for the deaf. To investigate this potential, I am developing the Virtual Sign Animated Pedagogic Agents which uses computer generated 3-D images to teach Thai sign language as well as provides a platform for studying instructional techniques. This approach has had a great impact on the technological development of animated pedagogical agents, especially with the emphasis on lifelike and controllable agent behaviors. This is particularly important in this research because it supports essential features of the Thai sign language such as hand gestures, facial expressions, and gaze. In addition, the research on animated interface agent provides a new way of human computer interaction by applying features of face-to-

face human communication to human-computer interaction. The result is that animated pedagogical agents are believed to enhance computer-based learning and are also effective at improving the motivational aspects of learning experience [4].

Many sign language researches use a concatenated synthesis system based on data collected with a "DataGlove" such as a machine interpreting system that translates written Japanese to Japanese Sign Language and vice versa [6]. The English company Televirtual produced a prototype captioning system known as "Simon the Signer"[1] and a successor, ViSiCAST[8]. For computer animation usability, VCom3D [7] has a system that produces signed English as the input and produces H-Anim-compliant Web3D [9] as the output. The first part of my research, on the other hand, is focused on development of 3-D virtual human and sign language animation with facial expression techniques in the context of education technology. We call this Virtual Sign Animated Pedagogic Agents (VSAPA). This concept is applied sequentially to the various steps of exploiting a virtual human approach: hand animation, then facial expression, and finally cognitive skills evaluation. The aim of this research is to enhance computer literacy and provide greater opportunities for deaf learners by developing and assessing the effectiveness of a real learning environment and interactive learning to teach deaf students. My first application area will be computer science, but the research results should provide a foundation for exploitation in other disciplines as well. The second aspect of this research centers on defining and assessing the effectiveness of different techniques in the classroom. I selected computer science as my first application area to reduce the external complexity of the initial work. Since VSAPA is a computer-based system and the material we will use in the prototype applications is also computer-based, many of the complications associated with environmental factors and human instructors are eliminated. I intend to start this part of my research by adapting currently taught material to different learning domains that are better suited for deaf college students and assessing the effectiveness relative to several measures. This process of interactive learning, especially for deaf students, is complex and affected by many factors. Clearly, the instructional design component is critical because it is important for users to have the flexibility for building knowledge by manipulating different types of media at their own pace and according to their own preferences. Equally important is the user interface which, in so many cases, is the difference between a tool being useful and used or ignored.

The final aspect of my research is to use VSAPA and the application to computer science material to provide instructors with some insights into techniques that will improve the educational experience and learning of deaf students. For example, it has been suggested that some students are more readily engaged when an instructor uses sign language rather than teaching by voice with a sign language interpreter translating the information[2]. While this is interesting, it is unlikely that instructors at universities not exclusively for deaf students will learn sign language, so it is important to investigate other techniques that can help.

The overall motivation for this work is my belief that deaf people should have full access to resources for education and employment – and this research is targeted at fulfilling this belief.

## **Research methodology**

The VSAPA will be based on 3-D technology (e.g., 3dsmax[10], LIFESTUDIO:HEAD [3]) dedicated to sign language features such as location, movement, hand-shape and orientation combining facial expression, Web technologies, innovative user interfaces, and pedagogical strategies. Each of these is now briefly addressed.

### ***Virtual Sign Animated Pedagogic Agents***

#### **Design of VSAPA**

Develop the virtual human with sign language and facial expression by scripting animation languages and visual modeling tools.

#### **Design computer content**

The content currently under investigation includes text, sign language with the agent, animations, and adjunct questions. To effectively learn the material, deaf learners typically first read the text screen, then view the sign animated pedagogic agent, and finally watch the graphic animations.

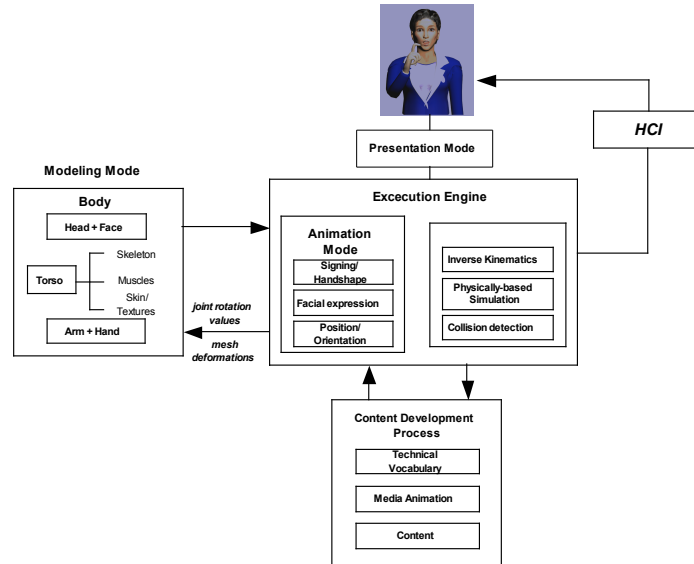
#### **Design User Interface**

System links for graphic user interface design and usability testing. The focus for GUIs is on optimizing the static displays that interactive systems are based on. There are elegance but simplicity, perceptual organization and visual structure. It is particularly investigating to identified the attributes that contribute to usability such as learn ability, efficiency, memory ability, errors and satisfaction[5].

#### ***Assessment process***

Measures will be constructed that assess the effectiveness of the tool as well as how these measures are obtained. For example, to understand how much content the students retain, one might use written or oral exams. It is not known if one is more effective than the other relative to assessing deaf students' knowledge of content or if the way a written exam is constructed has an impact. These are the types of questions that we wish to address using VSAPA.

The research scheme is illustrated by Figure 1.



**Figure 1.** Research Scheme

## Research status

Currently I have performed three tasks associated with this research: (i) analyzed the literature related to the construction of 3-D human models; (ii) constructed a prototype that includes hand sign animation and rudimentary facial expression; and (iii) begun to develop a framework of instructional design.

## Contribution

We believe that performing applied research into development of VSAPA and integrating this into a learning system to assist deaf college student is important work. With advances of computing and software, there is much promise that this approach can significantly improve the opportunities for deaf people in Thailand and around the world.

## Acknowledgments

I wish to thank Mahidol University for providing me the opportunity to study for a year in the Department of Industrial Engineering at Clemson University and the faculty there for their support. Thanks also to the deaf students who have provided valuable suggestions.

## References

1. Bangham, J. A., Cox, S., Elliott, R., Glauert, J., Marshall, I., Rankov, S., Wells, M.: Virtual Signing: Capture, Animation, Storage and Transmission - An Overview of the ViSiCAST Project.
2. Lang, H.G. Higher education for deaf students: Research priorities in the new millennium. *Journal of Deaf Studies and Deaf Education*, 7, 2002, 267-280.
3. LIFESTUDIO:HEAD Import-Export Plug-in for 3dsmax, Lifemode Interactive Ltd.

4. Mary Margaret Merrill, The Role of Animated Pedagogical Agents in Multimedia Learning Environments, Mississippi State University, December 2003.
5. Nielsen, J.: Usability Engineering, Academic Press, San Francisco, 1993.
6. Ohki, M., Sagawa, H., Sakiyama, T., Oohira, E., Ikeda, H., Fujisawa, H.: Pattern Recognition and synthesis for sign Language Translation System. ASSETS 10 1994, 1-8
7. VCom3D. <http://www.vcom3d.com/>
8. Verlinden, M., Tuijsseling, C., Frowein, H.: A Signing Avatar on the WWW. Presentation to Gesture Workshop 2001.
9. <http://h-anim.org/Specifications/H-Anim1.1/>
10. <http://www.3dmax.com>