Shengdong Zhao

December 2007 Research Statement

Interactive Media Lab (IML)
Dynamic Graphic Project Lab (DGP)
Bahen Center for Information Technology
40 St. George St.
Toronto, Ontario, Canada, M5S 3E4

Phone: 1 (416) 978-7581
Fax: 1 (416) 978-4765
Email: sszhao@dgp.toronto.edu
WWW: http://www.dgp.toronto.edu/~sszhao
Citizenship: Canadian

RESEARCH INTERESTS AND GOALS

My research interest is in the area of Human Computer Interaction, with an emphasis on mobile HCI and information visualization. I am particularly interested in enhancing mobile human computer interaction and mobile information access with effective and novel interactive techniques and interfaces, leveraging alternative methods of interaction such as gesture, audio, and haptics. I systematically study how these new approaches can be seamlessly integrated with existing solutions across mobile scenarios with varying degrees of distraction. Finally, I am interested in the design and evaluation of *highly interactive* visualization systems that allow rapid manipulation of views and rearrangement of data.

PREVIOUS WORK

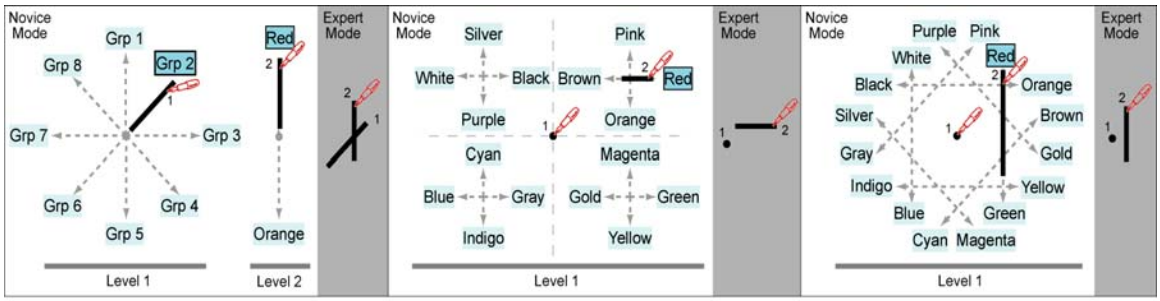
From offices to airports, from subway trains to cars, from walking on the street to running on a treadmill, mobile devices are increasingly used in a wide variety of different contexts and settings. Mobile interface design has become increasingly challenging due to the different cognitive demands, physical constraints, and social constraints imposed by such diverse usage scenarios. Effective mobile interfaces in the future need to incorporate alternative methods of interaction, and should be designed while envisaging various usage scenarios.

Figure 1: earPod and its various mobile eyes-free usage scenarios

Eyes-free interfaces, that do not require visual attention for its operation, are particularly attractive in mobile HCI since dedicated visual attention is often not available. However, due to the innate difficulty in designing auditory interfaces, previous attempts suffered from performance and scalability problems that prevented serious use. In my dissertation work, I have made important contributions to solving this problem. I made a number of design breakthroughs and invented an eyes-free menu technique called *earPod* using touch input and reactive audio feedback (Fig. 1). *earPod* is the first scalable eyes-free menu interface with a comparable performance with a popular visual menu selection technique (iPod) [2]. The interaction techniques embodied in *earPod* moved the state of the art a step ahead toward building realistic eyes-free interfaces in the future.

One of the compelling scenarios for eyes-free interfaces involves in-vehicle usage. To investigate in-vehicle eyes-free interaction, I systematically investigated earPod and a set of related menu techniques with a driving simulator. My evaluation results showed that eyes-free interfaces have significant advantages over visual interfaces while driving. In comparing menu selection while driving with desktop menu selection, I found that different usage scenarios can have strong effects on performance, and that the best solution for one scenario may be the worst solution for another. Findings such as these are essential for guiding mobile interface designers so that they can improve interface design across multiple usage scenarios.

In addition to eyes-free interfaces, I have the benefit of a strong research base in using gesture input for pen-based computing [1, 3, 4, 6]. Pen-based interfaces are often ideal for casual creative tasks such as sketch design, reflecting on a topic and extending creative ideas. Optimal experience for such tasks demands concentration without interruption; therefore, interfaces supporting such activities need to be efficient and to carry a light load for cognitive attention.



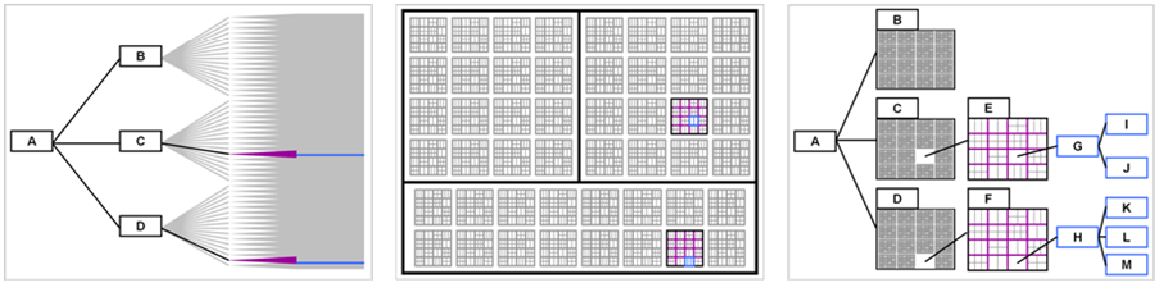
(a) Multi-stroke Marking Menu (b) Zone Menu (c) Polygon Menu

Figure 2: the various Marking Menus invented by me

Compound-stroke Marking Menus are an efficient gesture-based menu technique that can make menu selection up to 10 times faster than the corresponding linear menu selection performance when used in expert mode. However, the original design [8] suffered from a number of flaws (e.g.: ambiguity in item recognition in expert mode, and limited breadth and depth for the menu hierarchy) that prevent it from being used by applications with larger numbers of commands. I designed and implemented Multi-stroke Marking Menus (Fig. 2 a) to solve the ambiguity problem and to extend the menu depth (increase the number of items by 5 times) [6], and further developed Zone and Polygon Menus (Fig. 2 b,c) with colleagues at Microsoft Research to improve the menu layout and increase the breadth limit by 4 times to allow natural allocation of menu items [4]. My work on Zone and Polygon Menu has also led to a patent application (*app# 20070168890, pending*).

Figure 3: various components and features of the InkSeine in-situ search tool

My interest in interaction extends to the use of new input methods that optimize user experience for creative tasks. The *Active note taking task* (e.g.: using a notebook to sketch designs, reflect on a topic, or capture and extend creative ideas) is one type of such tasks. Together with Ken Hinckley and his team from Microsoft Research, I designed and implemented an in-situ search application called InkSeine (Fig. 3), which integrates a number of key concepts including: leveraging pre-existing ink to initiate a search; providing tight coupling of search queries with application content; persistent search queries as first class objects that can be commingled with ink notes; and enabling a quick and flexible workflow where the user may freely interleave inking, searching, and gathering content [1]. InkSeine has been released as an internal application at Microsoft Corporation, and will be available to the public in February, 2008. A patent application for InkSeine was submitted (*serial# 11/733,133*) in April, 2007.



(a) Node-Link Diagram (b) Treemap (c) Elastic Hierarchy

Figure 4: an illustration of a tree with ~3500 nodes using 3 different visual representations

I also have a strong interest in using information visualization to reveal abstract data visually and to allow scientists to gain quick insights into complex phenomena. Visualization for hierarchical data has always been a difficult problem. Node-Link diagrams show topology clearly, but distribute nodes unevenly, leaving upper

level nodes separated by white space, and lower nodes densely packed (Fig. 4 a). Treemaps use space efficiently, but are less familiar and can be difficult to interpret (Fig. 4 b). My work in Elastic Hierarchies [5] is an example of using a hybrid combination of node-link diagram and Treemap to interactively manipulate views and flexibly rearrange data. By combining the two complementary techniques, Elastic Hierarchies allow chosen structures and content to be emphasized and presented in a flexible and space-efficient manner (Fig. 4 c).

CURRENT AND FUTURE WORK

I aim to become a leading researcher in the area of mobile HCI and information visualization. Building on my past research, my research plan is organized into two stages.

In the next 1 or 2 years, my focus will be on mobile eyes-free interfaces, and on building *highly interactive* information visualization systems. Unlike graphical user interfaces (GUIs), where general interaction models (e.g., WIMP) have been extensively studied and used, there is not yet a similar "standard" for eyes-free interfaces. Due to the significant differences between the operating modalities, a straightforward translation from visual interfaces often fails to leverage the full potential of eyes-free interfaces. Thus, systematic and independent investigation is needed to identify general eyes-free interactive tasks, and to establish eyes-free interaction as a new paradigm that can be incorporated into our everyday lives. There is also a huge opportunity to study related issues in safety and privacy. Projects to be worked on include:

- Extending the earPod approach of eyes-free menu selection to support long and dynamic lists. In addition to menu selections, identify and establish other fundamental computing tasks that are suitable for eyes-free interfaces and design the corresponding eyes-free techniques for those tasks (such as dragging and dropping of auditory objects)
- Investigating eyes-free techniques within a broader context by comparing them with related approaches across multiple scenarios to understand the interplay of eyes-free techniques with vision-based techniques, leading to guidelines for designing multimodal interfaces.
- Pursuing case studies of real world usage of these techniques in applications to further investigate their usability, safety, and other related issues. In an initial case study, I am currently designing and evaluating an eyes-free version of a voice conference application called Vocal Village (a product of *Vocalage Inc*.) in mobile environments.
- Continuing with an ongoing collaboration with the National Federation of the Blind (NFB) to study on how to improve interface techniques for the blind.

In the area of Information Visualization, following the trend of my Elastic Hierarchy work, my focus will be on proposing interaction techniques that allow a user to quickly traverse sequences of many states of visualizations, with lightweight input, and with smooth, visually continuous feedback interactive visualization. I am also interested in using audio and haptic feedback to augment visual displays. This approach can be particularly interesting for small-screen devices, in which the audio and haptic feedback can offer additional means to overcome the screen real estate problem.

Over a longer term (next 3 years and beyond), I intend to extend the scope of my research into the following exciting areas.

Mobile Multi-touch Gesture and Eye-tracking: Multi-touch gesture and eye-tracking techniques have started to demonstrate their effectiveness in domains such as tabletop displays and desktop systems. I have started the investigation of multi-touch gesture in the tabletop environment. With rapid advance in hardware technology, it is expected that such capabilities will be incorporated into mobile and wearable devices, and mobile use of these technologies will become an exciting avenue of research in the future.

Theoretical Framework of Sensory Channels: While it's beneficial to investigate novel interaction techniques leveraging individual sensory channels, more importantly, we need to develop a theoretical framework that can account for interactions between the non-visual sensory channels and the visual channel. I plan to further investigate their relationship and establish a framework which can ideally serve as a foundation for future multi-input multimodal interfaces.

Evaluation Methods for Mobile Applications: Effective evaluation of mobile applications remains challenging. Mobile applications are often used in a variety of scenarios that include complex phenomena not easily reproducible in conventional laboratory experiment methods. It is important to design simulated environments that capture the essence of many real world phenomena. Constructing a lab environment that is capable of providing realistic simulations of various mobile contexts will be a priority in my research agenda.

Evaluation Methods for Information Visualization Systems: The evaluation of information visualization systems is a challenging problem due to the exploratory nature of information visualization tasks, and the lack of standard metrics to evaluate the utility of such systems. However, developing plausible and effective evaluation methods is an essential task that will promote the scientific standing of the field of information visualization. I want to make contributions to solving this important problem.

Mobile Social and Gaming Applications: Finally, social computing and gaming are two exciting research areas that can effectively bridge academia and industry. I am interested in investigating the possibilities of eyes-free social and entertainment applications for people to use on the go. Research in these areas will allow me to bring state of art technologies and ideas to consumers. I expect that this topic will attract substantial research funding from both governmental and industrial sources.

REFERENCES

1. Ken Hinckley, Shengdong Zhao, Raman Sarin, Patrick Baudisch, Edward Cutrell, Michael Shilman, Desney Tan (2007). InkSeine: In Situ Search for Active Note Taking. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).* pp. 251-260.
2. Shengdong Zhao, Pierre Dragicevic, Mark H. Chignell, Ravin Balakrishnan, Patrick Baudisch (2007). earPod: Eyes-free Menu Selection with Touch Input and Reactive Audio Feedback. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).* pp. 1395-1404.
3. Xiangshi Ren, Jinbin Ying, Shengdong Zhao, Yang Li (2007). The Adaptive Hybrid Cursor: A Pressure-based Target Selection Technique for Pen-based Interfaces. *Proceedings of the Interact 2007 Conference (INTERACT).* pp. 310-323.
4. Shengdong Zhao, Maneesh Agrawala, Ken Hinckley (2006). Zone and Polygon Menus: Using Relative Position to Increase the Breadth of Multi-stroke Marking Menus. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI).* pp. 1077-1086.
5. Shengdong Zhao, Michael J. McGuffin, Mark H. Chignell (2005). Elastic Hierarchies: Combining Treemaps and Node-link Diagrams. *Proceedings of IEEE Symposium on Information Visualization (InfoVis).* pp. 57-64.
6. Shengdong Zhao, Ravin Balakrishnan. (2004). Simple vs. Compound Mark Hierarchical Marking Menus. *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST).* pp. 33-42.
7. Monica C. Schraefel, Yuxiang Zhu, David Modjeska, Daniel Wigdor, Shengdong Zhao. (2002) Hunter Gatherer: Interaction Support for the Creation and Management of Within-web-page Collections. *Proceedings of the Eleventh International Conference on World Wide Web (WWW).* pp. 172-181.
8. Gordon Kurtenbach (1993). The design and evaluation of marking menus. *Ph.D. Thesis*, University of Toronto.