Final summary report, Marie Curie IEF: PIEF-GA-2008-220793 (INFERENCEHCI)

Marie Curie Intra-European Fellow: Per Ola Kristensson, Ph.D.
Duration: July 2008 – May 2010 (23 months)
Present address: Per Ola Kristensson, Computer Laboratory, University of Cambridge, United Kingdom
Email: pok21@cam.ac.uk; **website:** http://pokristensson.com

This research project has investigated how to design new fluid and efficient user interfaces that take advantage of machine learning and other uncertain reasoning procedures.

The work that was carried out can broadly be divided into two (partially over-lapping) themes: a) understanding the design space of user interfaces driven by machine learning, and b) design of novel user interfaces based on machine learning. This report presents the results for each theme separately below.

Theme 1: understanding the design space

To design efficient user interfaces based on machine learning it is critical we understand users' capabilities and limitations. In this theme we assessed and analyzed how the design of an intelligent user interface influences users' performance and behaviour.

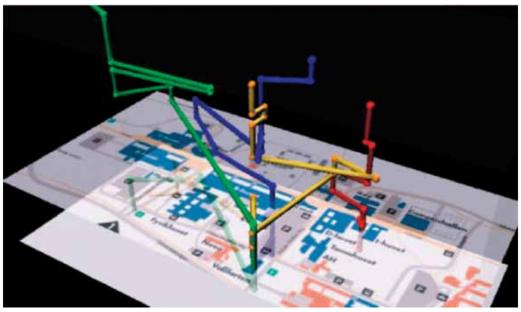

Figure 1. Space time cube visualization

Many intelligent user interfaces are on some level based on users' understanding of spatiotemporal patterns. Examples of such interfaces are handwriting recognition and various forms of gestural interfaces. We conducted a study to assess if so-called space time cube visualization could be used to aid users' understanding of such patterns (Kristensson *et al.* 2009). Space time cube visualization is a method to visualize two-dimensional spatiotemporal data in 3D by visualizing time as the third dimension (figure 1). We found that users were on average twice as fast in understanding complex spatiotemporal patterns when they used space time cube visualization in comparison to a 2D baseline. However, for simple data lookups a 2D visualization is better. Hence, user interfaces that need to convey both simple and complex spatiotemporal data patterns should consider incorporating both display variants. This work was published in *IEEE Transactions on Visualization and Computer Graphics* in 2009.

Another project in this theme was to understand the human performance of state of the art handwriting recognition. Tremendous effort has been invested into developing accurate handwriting recognizers that are capable of recognizing unconstrained handwriting that enables users to mix printed and cursive writing. However, the research literature predicted that users would write much slower using handwriting recognition than a much simpler-to-implement on-screen keyboard (about 16 wpm vs. 25 wpm). Hence it was questionable exactly how much use end-users would benefit from the research invested into creating accurate handwriting recognizers. However, we found that the research literature based its assumption on an old unreliable study that measured handwriting rather than handwriting *recognition.* We conducted an 11-session longitudinal user study with 12 participants to assess the learning curve of state of the art handwriting recognition in comparison to an on-screen keyboard baseline (Kristensson and Denby 2009). We found that, contrary to the prior beliefs in the literature, handwriting recognition performance was on par with an on-screen keyboard. Thus, handwriting recognition performance had been underestimated in the previous literature. This work was published in *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI 2009)*.

The third work in this theme was the authoring of an essay about intelligent text entry methods – text entry methods that enable users to write faster using machine learning and other uncertain reasoning methods. This essay had two primary contributions. First, it put intelligent text entry methods into their historical perspective. Second, it highlighted five challenges in the research field: localization, error correction, editor support, feedback, and context of use. The essay was published in *AI Magazine* in 2009.

Theme 2: design of novel user interfaces

Multi-touch tabletop displays based on FDIR technology was relatively new in 2008. We created a multi-touch table based on this technology to explore how multi-touch could be used in information visualization. Our application enables users to explore tagged images by using bimanual manipulation of an interconnected tag cloud. To make the interaction feel seamless we used a physics model based on viewing related image tags as interconnected springs (figure 2). A formative user study revealed that participants generally liked the system. However, they only used multi-touch about 10% of the time. This indicates multi-touch could be a viable complementary interaction technique for visualization displays. This work was published in *Proceedings of the 5th Nordic Conference on Human-Computer Interaction (NordiCHI 2008)*.

Figure 2. The multi-touch system

Another project was in the area of speech recognition. Speech recognition is now at a stage where it can be used as an alternative mobile text entry method. However, a problem with speech recognition is how to handle recognition errors, particularly in mobile scenarios that have widely varying background noise characteristics. We developed a mobile continuous speech recognition system for mobile Nokia devices. Uniquely, our system used the touch-screen display to reveal not just the best hypothesis of what the user intended to write, but also the next best hypotheses, using a so-called word confusion network. In a formative user study we found that novice users were able to use our system to enter text while simultaneously walking outdoors at a rate of 13 wpm. An expert user managed to reach 45 wpm if we subtracted the relatively long recognition delays. This shows mobile speech recognition coupled with smart touch-screen correction techniques has great potential as an alternative mobile text entry method. This work was published in *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI 2010)* and in *Proceedings of the ACM International Conference on Intelligent User Interfaces (IUI 2009).*

The third system we developed was a programming by example (PBE) application. PBE systems use machine learning to automate users' frequent tasks. We developed a new set of algorithms to be able to infer and automate a wide array of users' interaction patterns. Our set of algorithms uses ensembles of decision trees coupled with decaying confidence functions to decide which decision tree to use in a particular classification task. We were able to show that this technique enables PBE systems to infer many difficult interaction patterns from just a few examples. This work was published in *Proceedings of the 2nd ACM SIGCHI Conference on Engineering Interactive Computing Systems (EICS 2010)*.

Impact on society

The finding that space time cube visualization is beneficial in conveying complex spatiotemporal patterns to users is used by Oculus Software's product *GeoTime* (http://www.oculusinfo.com/SoftwareProducts/GeoTime.html). The work on providing efficient touch-screen correction methods for mobile speech recognition was featured in the leading magazine *The Economist*'s Technology Monitor (http://www.economist.com/node/16577398). In 2008-2009, I was also serving as an executive committee member of Cambridge University Entrepreneurs, a volunteer organization in Cambridge that helps students start their own companies. Two videos of the above research were also uploaded to *youtube*. The first video showcased the multi-touch information visualization display (*Proc. NordiCHI 2008*) and has so far attracted 16,165 views. The second video showcased Parakeet (*Proc. IUI* 2009) and has so far attracted 5,119 views.

Impact in the scientific community

The research results have been presented and included in the proceedings at five competitive ACM and AAAI international conferences with acceptance rates between 25-34%: *NordiCHI 2008*, *IUI 2009*, *CHI 2009*, *AAAI 2010* and *EICS 2010*. One of the systems was also demonstrated at *IUI 2009*. Two journal articles have also been published; the first in *AI Magazine* **30**(4) and the second one in *IEEE Transactions on Visualization and Computer Graphics* **15**(4). According to Google Scholar the above papers have been cited 26 times in total so far. Of these seven papers I am the first author on four of them and the senior author on the rest. During the fellowship I have also been invited as program committee member at several world-leading scientific conferences in my area. I was a Program Committee Member at *IUI 2009, IUI 2010, CHI 2010* (Associate Chair), Faculty Member at the Doctoral Consortium (*CHI 2010*), Judge at the ACM Student Research Competition (*CHI 2009, CHI 2010*, ACM-wide 2010), and a Technical Co-Chair of *HCI 2009* (the British HCI conference).

References

1. Bray, Z. and **Kristensson, P.O**. 2010. Using ensembles of decision trees to automate repetitive tasks in web applications. *Proc. EICS 2010, 35-40.* ACM Press: 35-40. Acceptance rate: 29%.
2. Vertanen, K. and **Kristensson, P.O**. 2010. Intelligently aiding human-guided correction of speech recognition. *Proc. AAAI 2010*. AAAI Press: 1698-1701. Acceptance rate: 25%.
3. **Kristensson, P.O.** 2009. Five challenges for intelligent text entry methods. *AI Magazine* **30**(4): 85-94.
4. **Kristensson, P.O.** *et al.* 2009. An evaluation of space time cube representation of spatiotemporal patterns. *IEEE Transactions on Visualization and Computer Graphics* **15**(4): 696-702.
5. **Kristensson, P.O.** and Denby, L.C. 2009. Text entry performance of state of the art unconstrained handwriting recognition: a longitudinal user study. *Proc. CHI 2009*. ACM Press: 567-570. Acceptance rate: 25%.
6. Vertanen, K. and **Kristensson, P.O.** 2009. Parakeet: a continuous speech recognition system for mobile touch-screen devices. *Proc. IUI 2009*. ACM Press: 237-246. Acceptance rate: 29%.
7. **Kristensson, P.O.** *et al*. 2008. InfoTouch: an explorative multi-touch information visualization interface for tagged photo collections. *Proc. NordiCHI 2008*. ACM Press: 491-494. Acceptance rate: 34%.