Project description / Martin Tall / info@martintall.com / +46 730 611677 / Varmdogatan 72, 257 33 Rydeback, Sweden

Noise tolerant gaze interaction with computers

Introduction

Since the introduction of graphical user interfaces for human computer interaction the main input device for the general population has been the keyboard and a pointing devices such as the mouse, trackballs, touch-pads etc. These pointing devices have evolved to support the common two dimensional desktop interfaces and are used to select/manipulate objects and activate functions, while the keyboards are in general used for text entry. Most users have spent an substantial amount of time to master these devices and there has been few real alternatives available. This type of interaction excludes a range of users who do not posses the required precision of arm and hand movements needed to operate these devices. Additionally, for the general population the repetitive strain injuries caused by a static prolonged everyday usage could be reduced if computer interfaces supported other modalities, such as eye tracking for gaze based interaction. The range of modalities available today creates an opportunity for a new era of Human Computer Interaction where computers adopt to human behavior, narrowing the gap between the machine and man.

Human Cognition

Central to all of this is knowledge about human behavior, our mind, brain and the ability for cognition. Traditional interaction with the mouse and keyboard in a silent two dimensional environment leaves a large part of our cognitive capabilities behind for the great majority of users while at the same time excluding those who cannot use those limbs. To narrow the gap between man and machine the interface needs to support and understand natural human behavior. Ideally a computer interface is capable of sensing a wide range of human behavior for establish a rich two-way interaction. However, the interface needs to be capable to cater for the capabilities of the individual, which differs greatly in between users and also over time for each single person. Today computers are totally oblivious of their users without any sense of tact. The advances in technology have produced modalities capable of sensing a wider range of human behavior than ever before. This have created an opportunity to track eyes, facial expressions, bodies, limbs, hands, interpretate and produce speech, collect crude brainwave patters (EEG), register subtle muscle contractions (EMG) as well as arousal levels (GSR). This opens up the possibility for developing new forms of interaction that enables the use of technology of those who previously has been excluded. Moreover, it has the opportunity to revolutionize the man-machine interaction as a whole.

Eye Trackers and Gaze Interaction

It seems appalling that eye trackers which provides real time data on eye movements and coordinates of where the user is looking at the screen could come to use in Human-Computer Interaction (HCI). After all there is a clear connection on where we look and what we are interested in. Several studies for gaze based human-computer interaction has been performed during the last two decades. Many of which have explored how eye tracking devices could be used to interact in a two dimensional Graphical User Interfaces (GUI) such as Microsoft Windows. A great deal of attention has been directed towards developing solutions which compensate for the differences between the mouse and gaze. The discrimination of current eye tracking equipment lies around 0.5-1 degree of visual angle and is prone to “jitter”. This makes is difficult to use gaze for operating traditional desktop interfaces which have been developed for use with a precise pointing device such as the mouse.

Using eye trackers to gather real time gaze data for the purpose of interacting with a computer interface poses several challenges and requires novel interaction techniques. For example the issuing of a command has to be identified as something that differs compared to normal gazing to view the scene, a common issue usually referred to as the "Midas-touch problem" (Jacobs et al., 1993) The human visual system is ideally constructed for surveying and observing the environment. Overloading the visual channel with motor commands has been found to be undesirable in several studies (Zhai et al 2003, Kumar et al 2007) This poses a problem for interaction, how do you build a system that distinguish between a user just looking around and gazing with the intent to perform an action? Several methods have been developed to work around this problem. A common solution is to apply dwell-times where the user fixates on a point for a prolonged period of time which is interpreted as an intention to activate or execute commands (Hansen 2003, Majaranta 2004). In general this poses two problems. First, the user is stressed because where ever he looks an command seems to be executed which causes him to constantly move his eyes around. Second, the interaction is delayed since the user has to sit through the dwell time and fixate on a point for the specified period of time before the command is activated. The dwell-time can be adjusted and tuned but it still poses a delay. Thus, many projects have come to conclusion that dwell-time activation is only preferred when the user cannot use any other mean of activation/modality (buttons, voice etc) Other systems rely on speech (Miniotas, 2005) or series of keyboard hot-keys (Kumar, 2007) to perform activation, selection etc.

Within the field of assistive technology eye tracking and gaze interaction has successfully been implemented to improve the quality of life and the ability to communicate for users diagnosed with ALS or similar conditions. The GazeTalk (Hansen, 2001) and Dasher (Ward, 2000) software is today used on a daily basis for text input and communication utilizing gaze alone. The rate of input is ranging from 6-15 words per minute with the GazeTalk to 25 words per minute with the Dasher. A normal chat room conversation typically goes at 40 WPM while speech easily reaches above 100 WPM (Hansen et al, 2004). In other words, there is room for improvement by new interfaces and better probabilistic models to reduce the number of fixations needed to write each word. The StarGazer (Skovsgaard, 2008) interface is one example of a novel approach with its 3D approach that utilizes pan and zoom for efficiently handling the noisy data associated with eye tracking.

Establishing an effective interface for text entry is crucial for empowering the disabled with a tool for communication, such as text to speech or email. Going further there are few open alternatives for everyday computer activities. Unfortunately, the few available solutions are highly vendor specific and comes in a package with a price tag matching a modern automobile. As in any domain, alternatives would improve the field as a whole while also lowering the costs associated with the technology.

Combining gaze with other modalities

The usage of eye trackers could provide beneficial advantages which can be exploited in two ways, to enhance the direct interaction by gaze (overt) or to provide the computer with information on where our attention is directed (covert). This covert attentional awareness could be used for creating interfaces that are aware of the humans that operate them and adapt according to their behavior. In terms of covert usage eye trackers in combination with other modalities creates interesting opportunities. For example, the German firm Eye-Square have developed a tool for recording several physiological measurements (such as skin conductivity) in combination with eye tracking. When used for studying reactions to TV commercials the software visually illustrates what object the subject was fixating and how this attention modulated his/her arousal state. Besides this measurements of galvanic skin response (GSR) for arousal levels, electroencephalogram (EEG) could be used as a crude sensor of cortical wave patterns as a measurement of engagement. The electromyography (EMG) detects electric potentials generated by muscle cells and could be attached to the forehead to enable interaction by subtle facial gestures. Recently, the combination of eye tracking and EMG has been incorporated in a system developed by J.C Mateo, J.S Agustin and J.P Hansen (2008) This solution seems promising as it allows for effective hands free interaction. Personally I see a great potential for a wide range of users and usages, especially since the initial studies indicate that the performance in terms of speed is comparable or above that of a mouse. The learning effect has yet to be studied but at first glance it seems capable of beating the mouse.

Specific gaze based interfaces

Most gaze-based interfaces that have been developed so far are more or less solutions for using gaze with existing operating systems such as Microsoft Windows. This poses a problem since these systems are designed from the start to support the mouse as the main input device, which is a rather static mechanical device compared to the human visual system. In my opinion, gaze based interfaces requires novel interaction methods and other types of representations which can be easily identified and fixated to navigate and manage information. Gaze-based interfaces have to be developed from an initial intention to support gaze based interaction and not just substitute a mouse for the traditional desktop operating system. Special consideration for supporting the various capabilities of the end users is in place.

At the time of writing I am currently developing gaze based interface components that take several factors related to eye tracking and eye movements into account. The components such as dwell buttons, saccade selection menus and expanding canvases have been incorporated into a prototype application for a photo gallery and a music library/player. The components have been designed with reuse in mind so that they can be incorporated into future application simply by “drag and drop” or a few lines of code. This work will be presented in my master thesis which is finalized in late May 2008.

Research topics

The domain of gaze interaction struggle with a set of challenges that needs attention. Since my aim is to advance the field of gaze interaction some of the more important ones, in my opinion, are listed below. My main focus of interest lies in the areas of interaction methods and interfaces.

- Interaction methods
  Eye movements are noisy. In fact, our eyes are never still. This is due to the small eye movements such as drifts, tremors and involuntary micro-saccades (Yardbus, 1967) which essentially provides the ability to observe objects in focus, as a natural method to counterbalance retinal fatigue. The knowledge on the nature of eye movements, such as fixations, saccades, smooth pursuit etc. should be utilized when designing gaze interfaces. Improved algorithms for filtering and stabilizing these factors in combination with interaction methods that take these into account would improve the usability and experience of these interfaces. For example, a wide range of research findings indicate the zooming/expanding interfaces are suitable for gaze interaction since it has the ability to counterbalance the inaccuracy. This and other methods needs to be further investigated and developed.

- Interface components
  Solving the midas touch problem is key issue in gaze interaction. Added modalities or applying dwell times are the most common solutions, however a more flexible solution would allow multiple interaction methods. During the progression of ALS the muscle control degenerates over time. Starting with two modalities for a fast interaction and the switching to gaze-only while still using the same interface would be an ideally flexible solution. The development of open and accessible interface components that at a bare minimum can be driven by gaze alone while catering for additional modalities have the possibility to accelerate the field of gaze interaction as a whole.

- Accuracy of eye trackers
  From my own experience I have learned the value of a good calibration, and maintaining it. In most eye trackers the quality of tracking deteriorates over time due to changes in body posture and physical properties of the eye changes (eye becomes dry, changes in ambient light etc.) Unnoticeable or simple calibration as well as dynamic/adaptive re-calibration over time is crucial for the next generation of gaze interfaces and deserves further investigation. Likewise, interaction methods that counterbalance the inaccuracy as discussed above is crucial.

- Accessibility of eye trackers
  Most systems today contains a firewire camera, a reverse InfraRed filter, a set of IR-LED´s and software for image processing. The camera for popular remote systems have a resolutions in the 1-2 MP range capable of 50 images per second. In my own opinion, it is a unreasonable that these systems come at the same price as a new modern car. Open source software such as openEyes (Li&Parkhurst, 2006) in combination with high quality consumer cameras could assist in spreading awareness and raising the interest on the field of gaze interaction.

Development process

The purposed structure of the process relies heavily on a iterative process applying User Centered Design methodology (Norman&Draper 1986) as well as the Contextual Design (Beyer&Holzblatt, 1998) paradigm. This ensures a focus on the end users and their needs instead of developing from the perspective of what is technically possible to achieve. Incorporating human factors experiments (such as Shneiderman, 1986) to drive the development is especially important when developing for users with specific needs and a high variance of capabilities. Many of the proposed activities are highly iterative and are at this early stage not in any way set in stone.

- Step 1. Understand users needs
  a) Evaluate and establish appropriate method for conducting a pre-study to understand the variance of users capabilities and needs. The gathered data will drive the development process and ensure that the interface components cater for a wide rage of users and their needs. B) Evaluate the state of technology and identify possible combinations of modalities used in conjunction with eye tracking. These actions will set the initial base plate of requirements for the project.

- Step 2. Research suitable interaction methods
  Evaluate current state of research. Focus on end users needs. Consider use with or without additional modalities. Specify interaction methods and define a set of mockups/prototypes.

- Step 3. Develop flexible components
  Implement prototypes to support the vast majority of users and conditions as specified by previous research.

- Step 4. Evaluation
  Test the interface. Gather raw data on selection times, error rates etc. as well as subjective user experiences. Evaluate the data and perform iteration of step 2 until a reasonable solution has been established.

- Step 3. Collaborate with the research community (ongoing)
  Participate on conferences, discussions, seminaries and online communities. Publish papers and reports. Communicate and coordinate research. Evaluate and specify guidelines for further research. Iterate from step 2.

- Step 4. Incorporate the components into applications for evaluation
  Base these on the real needs and requests from the user community (extend the COGAIN software application section) These applications can be used to evaluate the individual components together in a real-world everyday task, this is to ensure that they work as intended and meets expectations of end users.

- Step 5. Release
  Document and distribute software components and applications, making it accessible to the public and the end users.

Anticipated contribution

The expected contribution of my research will further develop the interaction methods associated with gaze interaction. Besides contributing to the collected knowledge the project will result in a range of reusable software components becoming available. My mission is to raise the awareness of gaze based interaction, working towards making the technology more accessible while improving the usability for a better user experience. In the field of technology and eye tracking I make a few assumptions a) the performance of hardware such as cameras and processors will continue to increase at a constant or increased rate b) accessibility to hardware will increase due to lower costs c) performance and accuracy of image processing algorithms for eye tracking will improve with a higher level of availability. The development path for eye tracking as a technology is rather certain compared to that of gaze based interaction which is a minor subset within the eye tracking field. My assumption is that as this technology becomes more accessible the limiting factor for gaze interaction will be the availability of applications. The development of these are dependent on openly accessible and standardized interface components with an established method for interaction. This is what I wish to contribute to with a determination for driving the evolution of gaze interaction forward. It has the potential to empower the millions of people who are unable communicate or participate in society as well as revolutionizing the field of Human Computer Interaction as a whole.

References

Beyer, H., Holzblatt, K. (1998), Contextual Design: Defining Customer-Centered Systems. Morgan Kaufmann, USA.

Chin, J. P., Diehl, V. A, Norman, K. (Sept. 1987) Development of an instrument measuring user satisfaction of the human-computer interface, Proc. ACM CHI '88 (Washington, DC) 213-218. CS-TR-1926, CAR-TR-328

Duchowski A (2007). Eye Tracking Methodology. Theory and Practice. Springer-Verlag London Ltd.

Duchowski A, Michael Ashmore, Garth Shoemaker (2004) Efficient Eye Pointing with a Fisheye Lens

Hansen, J. P. et al. Command Without a Click: Dwell Time Typing by Mouse and Gaze Selections. In Proceedings of Ninth IFIP TC13 International Conference on Human-Computer Interaction (Interact '03) (Zurich, Switzerland, Sept 1-5, 2003)

Kumar, M. et al. EyePoint: Practical Pointing and Selection Using Gaze and Keyboard. In Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '07) (San Jose, CA, 2007).

Li, D., and Parkhurst, D. J. (2006). Open-source software for real-time visible-spectrum eye tracking. Proceedings of the COGAIN Conference, pgs. 18 - 20.

Majaranta, P. et al. Effects of Feedback on Eye Typing with a Short Dwell Time. In Proceedings of the Eye Tracking Research & Application Symposium (ETRA 2004) (San Antonio, TX, 2004)

Miniotas, D. et al. Extending the Limits for Gaze Pointing through the Use of Speech. In Information Technology and Control, 2005, Vol. 34, No. 3, 225-230.

Ohno, T. (1998). Features of Eye Gaze Interface for Selections Tasks. Proceedings of The Third Asia Pacific Computer Human Interaction – APCHI '98. IEEE Computer Society. 1– 6

Shneiderman, Ben (1982): The Future of Interactive Systems and the Emergence of Direct Manipulation. In Behaviour and Information Technology, 1 (3) pp. 237-256

Shneiderman, Ben (1986) Human factors of computer programming.
In Handbook of Human Factors (G. Salvendy, Ed.), John Wiley & Sons, Inc., (1986), 1398-1415,
Reprinted in Tutorial on Software Restructuring, (R. Arnold, Editor), IEEE EH0244-4 (1986), 67-81.

Sutherland, Ivan E. (1963): Sketchpad: A man-machine graphical communication system. In: Proceedings of the AFIPS Spring Joint Computer Conference 1963. pp. 329-346.

Yarbus, A. L., Eye Movements and Vision. New York: Plenum Press pp. 1967.