Research Proposal
Eye Tracking using Infrared

David Wild
Supervisor: James Connan
Computer Science Department, Rhodes University

2nd March 2012

1 Principle Investigator

David Jonathan Wild
Warden Flat, Victoria Mxenge Residence
Rhodes University, Grahamstown
082 495 8689
g09w0474@campus.ru.ac.za
Supervised by: James Connan

2 Project Title

This project is proposed under the title of:
Gaze Tracking using a Regular Webcamera

3 Statement of Problem

Eye tracking is a valuable technological utility to have, whether it is used for research of the way people observe their surroundings, or in specific tasks such as reading or driving, or whether we are trying to further human interfacing with computers. Sibert and Jacob say that "Eye gaze interaction can provide a convenient and natural addition to user-computer dialogues" [4]. The requirements of the two different systems differ.

If we wish to create a new human interface then "A reasonable definition of performing well is if eye gaze interaction does not slow down interaction and can "break even" with the mouse" [4]. So we need to create an easily accessible system that is equivalent to currently used interfacing devices. The second system is usually known as Gaze tracking, where you take the data down for future analysis rather than for immediate interaction. The problem for gaze tracking is to recognise the difference between "saccades and fixations" [2].

4 Objective of Research

The objective of the proposed research is to firstly get an eye-tracking system working using a regular web camera, and then to test the limits of the efficiency using a self-designed and implemented program, and using an open source software program that utilises the same hardware framework. If the efficiency of the system is good enough then the next step will be to create a user interface that is equivalent to mouse usage. If the efficiency of the system is not good enough for an interface, then other suitable applications will be examined.

5 Background

There has been a lot of work in this field already, so none of the proposed research is likely to be original. There are already products available, both software and hardware [1]. Rather than simply duplicating what others have already done, this section shall describe some general concepts that underlie this project.

When we read a static visual display there is no actual smooth progression of eye movements, rather there are "saccades and fixations" [2], as mentioned earlier. Fixations are periods of time between 100 to 500ms where the eye comes to rest, and saccades are rapid, ballistic eye movements interspersed between the fixations. When we are reading, the fixation times drop to 200-300ms and the saccades usually jump about 5 to 9 characters ahead. The eye picks up virtually no information during these saccades. There are many different types of saccades including return sweeps that return the reader to the start of the next line. These sweeps are usually one longer saccade but can include a shorter, corrective saccade to get the position correct. For a gaze tracker, taking into account which recorded movement is which is very important in the analysis of the data found. [3]

To see where the eye is pointing uses different indicators. For the human eye to focus on an object it is necessary to move the eye so that the object appears on the fovea. The fovea is a small area in the center of the retina of the eye. So being able to track the position of the fovea is useful for figuring out exactly where the focus of the gaze is. The problem with tracking the fovea is that the eye makes jittery movements rather than a smooth motion the user would be used to from a mouse. Another problem, is if you consider the eye tracker as a complete replacement of and to work exactly the same as a mouse then you end up with the problem known as the "Midas Touch Problem" [1]. If you make the eye tracker issue a command everytime the user's eyes focus on an object then the user will be unable to browse the desktop without opening every object that is focused on. So instead of forcing a command Jacob streamed the interaction of objects as tokens similar to the raw input from the keyboard. This meant

[1] See this website for an example of available products: www.tobii.com/en/gaze-interaction/global/lp/Gaze-the-Future-of-Computers/

that the data was only acted on under special circumstances, like the pressing of a button. [1]

To conclude this section; it is easy to find information on how do this project but the implementation will still not be easy.

6 Approach

The first part of this research will be to find and read more literature on this topic. Understanding the problems and aspects of the eye movement, algorithms for focus calculations, image processing and image capture of the eye will allow the research to progress and should allow for the creation of a step-by-step methodology.

The next phase of the research will be to set up the image capture. Once we can capture the image in a raw format or using a program, then the next step is the image processing. Processing the image in order to recognise the pupil, and eventually the fovea, is useful to finding the focus point of the user's vision, but if this is not possible with the hardware then just recognising the pupil will suffice and the deficit in accuracy will be dealt with in the algorithms.

The third phase of the research is to use clean and efficient algorithms to find the focus point of the eye field according to where the pupil or fovea is looking. Once this stage is completed the efficiency can be tested and raised to the best it can be.

The final stage will be to find an application for the system but, as previously stated, the final decision on what application is dependent on the efficiency of the program.

7 Requirements

The hardware requirements for this project are:

- A standard web camera

The plan is to use only freely available software for the programming aspects of this research, such as OpenCV.

8 Time-line

- By May all of the major research into this field should be done and the image input should be nearly ready to receive raw input video feed.
- The image processing to recognise the pupil or fovea needs to be done by the end of July.
- From July till the end of September will be dedicated to the algorithm for finding focus points and maximising the efficiency of the system

- By the end October the application for the system should be complete and running with the whole system available for use on similar hardware systems.

References

[1] Jacob, R. J. K. The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems 9* (1991), 152–169.

[2] Rayner, K. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* (1998), 372–422.

[3] Reichle, E. D., Pollatsek, A., Fisher, D. L., and Rayner, K. Toward a model of eye movement control in reading. *PSYCHOLOGICAL REVIEW 105*, 1 (1998), 125–157.

[4] Sibert, L. E., and Jacob, R. J. K. Evaluation of eye gaze interaction. pp. 281–288.