

Concentration Detection by Eye Movements: Towards Supporting a Human*

Yoshisuke Tateyama[†], Yoshio Matsumoto[‡], Satoshi Kagami[†]

[†] Digital Human Research Center, National Institute of Advanced Industrial Science and Technology

[‡] Nara Institute of Science and Technology

y.tateyama@aist.go.jp, yoshio@is.naist.jp, s.kagami@aist.go.jp

Abstract – *Machines like humanoids or personal computers are needed to assist humans. The tasks of humans become so complicated that machines need to act without detailed orders from humans. Eye movements provide very important information about a human. Many researchers have tried to measure accurate fixation points using an eye tracking system, from which they attempted to extract a human's orders. With limited resources, it is impossible to calculate accurate fixation points while allowing natural head movement. The authors have attempted to separately analyze the change in face positions and change in direction of gaze without detecting fixation points. This eye tracking system requires 3 features: (1) detection of a face position change and eye movements, (2) wide measurement area to allow for natural head movement, (3) high sampling rate over 20 Hz for detection of saccades. The chosen eye tracking system is based on a stereo camera. To verify the ability to detect whether a human is concentrating or not, experimental measurement is made while a subject is doing several tasks. The results confirm the possibility of detecting a human's state of attention.*

Keywords: Face detection, iris detection, human interaction.

1 Introduction

Attempts were made to obtain information from a human by observing his/her face, especially from an eye. The requirements of the measurement system were:

- to avoid any unusual requirement for measurement.
- to avoid head restraint during measurement.
- To observe natural human movements

Additionally required were:

- A large measurable range
- appropriate precision

*0-7803-8566-7/04/\$20.00 © 2004 IEEE.

Sophisticated eye tracking systems have become available. But essentially their measurement range is obtained with the sacrifice of accuracy. They detect the line of gaze. When one human sees another, he/she does not appreciate target's precise view. Devising methods to detect what he/she is seeing is a challenging and important research area, but requires the extraction of precise 3D positions. Extracting gaze direction is a delicate technique. If one wish to know what a user is seeing, one must know exact eye position and gaze direction. If one wish to know only a user's status, one may not need to know these parameters precisely. Using a normal camera, it is difficult to get precise information. With a face tracking system based on a stereo camera an indication is sought of what a human is doing.

Eye movements can be roughly classified as followings:

- fixation
- saccade
- small eye movements (e.g. tremor)

Detection is required of spontaneous and intentional changes, related to what the subject is doing. Observation of a subject's eye may give very useful information about their state. However the subject cannot intentionally control his/her gaze point. So if the system reacts each time the gaze point moves, the subject may be distracted by extraneous noise. For a system to support a human, detection of the target user's state is important. If a user concentrates on a task, the system may provide little or no support. If a user is ready to perform a task and waits for a system response, the system can then offer some help. If the user is tired, the system can advice rest, as it is not a good idea for the system to offer information continuously.

2 Experiments

The following need to be known:

- The level of eye movements that can be observed with this measurement equipment.
- The accuracy with which this measurement equipment observes during saccadic movements.

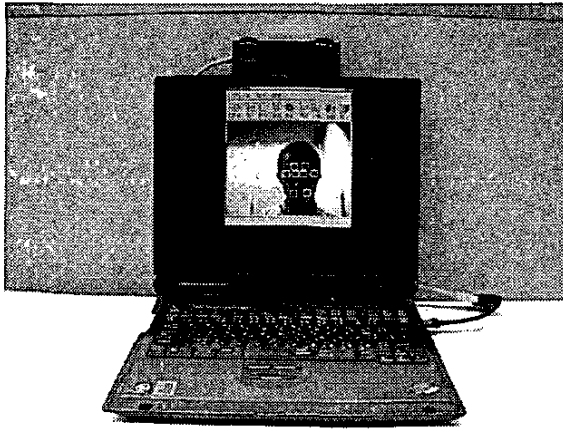


Figure 1: An external view of the face tracking system used.

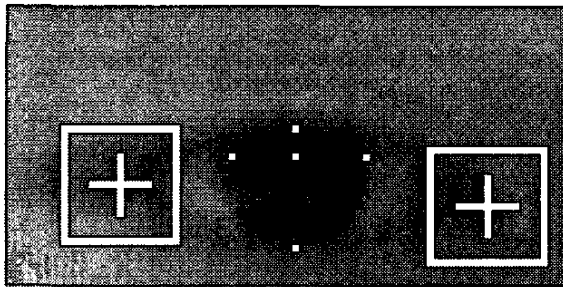


Figure 2: An extracted image of a right eye. The face tracking system detects both edges of the eye by template matching.

- Can the system differentiate between tasks.

2.1 Apparatus and methods

It is possible to extract information from a human face at 30 Hz [1]. Figure 1 shows an external view of the face tracking system. Though the original [1] system uses two NTSC cameras, an IEEE 1394 stereo camera (VidereDesign Mega-DCS) is now used. This captures two 640×480 pixel images simultaneously at a 30 Hz rate. These images are processed in a PC (CPU: Pentium III 450 MHz). 640×480 [pixel] $\times$ 30 [Hz] binocular images were taken at a time. A modification of the system of [1] was used to find a position of a face. This system is stable and robust to find a face and to track it.

Particular attention was paid to two features — the right edge of the right eye and the left edge of the left eye. Basically these are found by template matching. Although the face tracking system outputs 3D positions of a face, intermediate values are used — 2D feature positions on an image. The reason for this is to recognize sources of errors. The face tracking system also gives also 3D positions of each feature, such as the right edge of the right eye, the left edge of

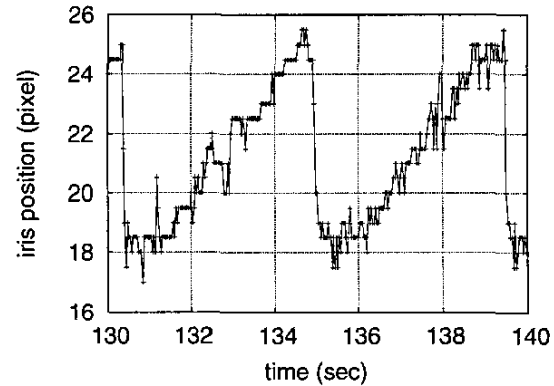


Figure 3: Eye movement when a subject has the task of reading a document.

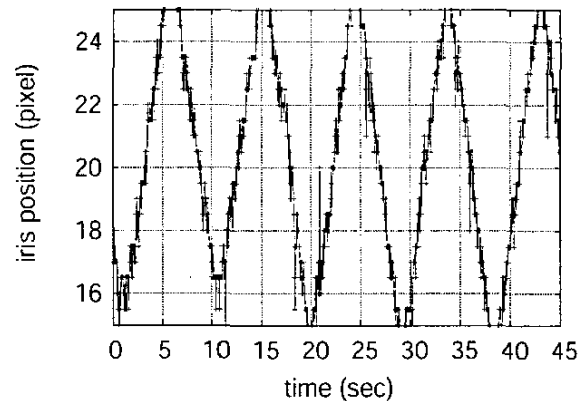


Figure 4: Eye movement when a subject is observing a moving object.

the right eye, etc. However, these are calculated from 2D images employed by singular value decomposition. If a phenomenon is observed using these 3D positions, the measurements contain nonlinear calculation and template matching error, and the source of which cannot be specified. Only the change of positions require to be known, so it was decided to examine the 2D positions of features on an image.

The left edge of the right eye was regarded as the face position. The template images of the left edge of the right eye and of the right edge of the right eye were taken when a subject's eyes were open. While a subject was blinking, the correlation value between the template and captured image of both edges of the right eye falls, so the data from these frames were excluded.

2.2 Single task experiments

Experiments were conducted to examine the precision of the system in detecting eye movements. Eye movements are basically combination of fixations and saccades, but when following a continuously moving object, the eyes move

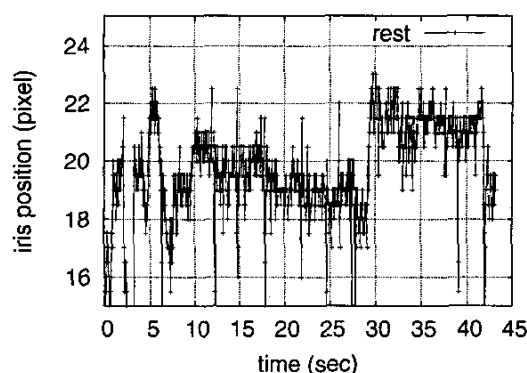


Figure 5: Eye movement when a subject is resting.

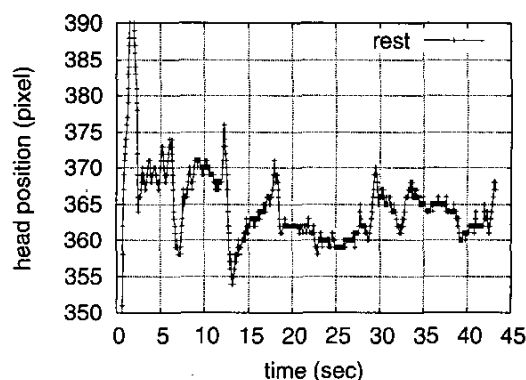


Figure 7: Head movement when a subject is resting.

smoothly [3].

Figure 3 shows eye movements of a subject reading a document. Theoretically it may contain fixations and saccades. In this measurement, though, template matching errors became a nuisance, some saccades were observed.

Figure 4 shows the result of a task in which a subject watches a moving circle. The moving circle was displayed on the LCD display, placed in front of a subject. The circle moved repeatedly horizontally from left to right, then right to left. Theoretically, the change is continuous, but the result is constrained to the pixel units. This system detects spontaneous eye movement in pixel units.

Figure 5 shows eye movements when the display is blanked and the subject is at rest. The eye movement was larger than in the previous situations.

Head movements in the single task experiments are shown in Figures 6 and Figure 7. Figure 6 shows that when a user concentrates on reading, the head movement is quite small compared to resting (Figure 7).

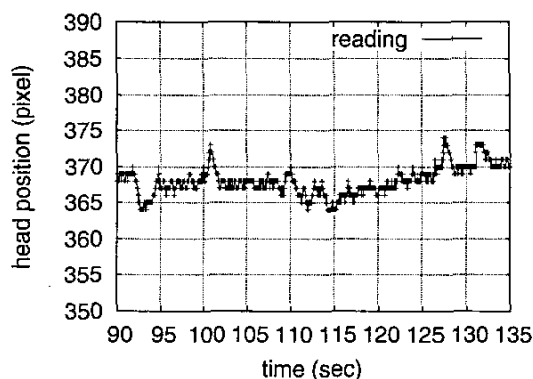


Figure 6: Head movement when a subject has the task of reading a document task.

2.3 Combination of concentration and rest experiment

An experiment was conducted to examine whether the tracking system could detect any changes whether a user is concentrating on a task or is at rest. In this experiment, to avoid any fixation phenomena when looking at an object, the subject looks at a blank LCD display. Iteration consists of the following procedures:

1. The experimental system generates an arithmetic question which requires adding two two-digit numbers.
2. The system gives the question to a subject using a synthesized voice.
3. The subject calculates the answer, and responds verbally.
4. After about 10 seconds, the system begins next iteration.

After ten iterations, a ten second rest was inserted. In the experiment, images of the subject were taken by the stereo camera, and recorded. The face tracking system extracts face information from these images.

Figure 8 shows vertical positions of the face of a subject in the experiment. After 6 minutes elapsed from the start of the experiment, a subject appeared to tire. Between iterations the head of the subject was moving. But when he was calculating, the face positions and eye movement observed by the tracking system were seen to stop.

Figure 9 shows the squared velocity of the face position of a subject in the experiment. The system can detect if a user is moving or not.

It was found the system could detect the cessation of movement of a user's head and eye when he/she concentrated on a task.

Figure 10 shows the eye position in the experiment. As with the face position, while the subject seemed to concentrate on the task, the eye position stabilised.

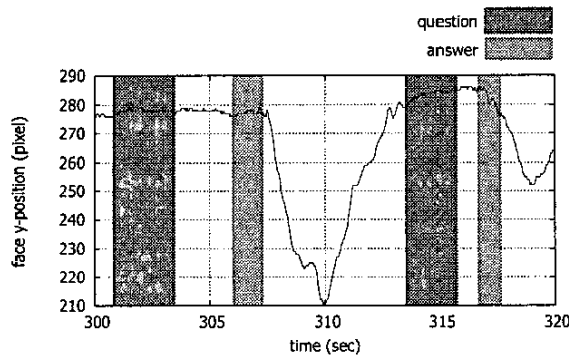


Figure 8: Face y-position in the voice based arithmetic task.

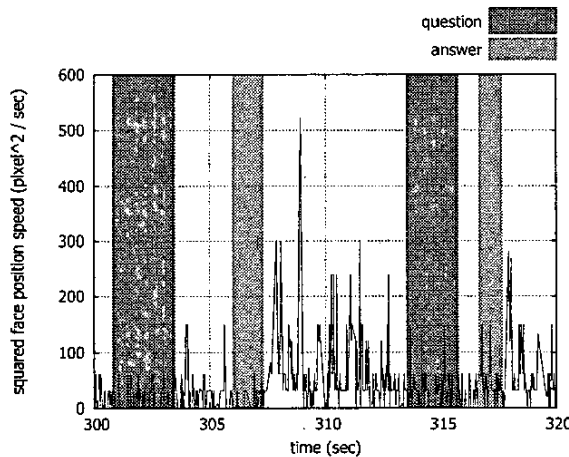


Figure 9: Squared values of face position speed in the voice based arithmetic task.

At this stage, it is now possible to detect changes in a face position and eye movement. A simple system was implemented to detect these changes. Figure 11 shows a sample of this implementation. Figure 11 shows a line indicating the time at which the system is giving a question, and a line indicating the time that the subject is speaking an answer. The points show the system detecting eye or face changes at these times. At the start of the system question, the subject seems not to react for about some tenths of seconds. After the subject answers, he/she seems to be relaxed and the face and eyes move. At the 275 second, the system does not give a question but declares a rest time, and the subject seems to be relaxed. This system recognized a user's major changes of face and eye. For example, if the system decides that the user is tired, it could advise him/her to rest.

3 Discussion

Several sophisticated ideas such as attentive user interfaces [2] have been proposed. For example, consider a display system in a public space. If it can detect changes in

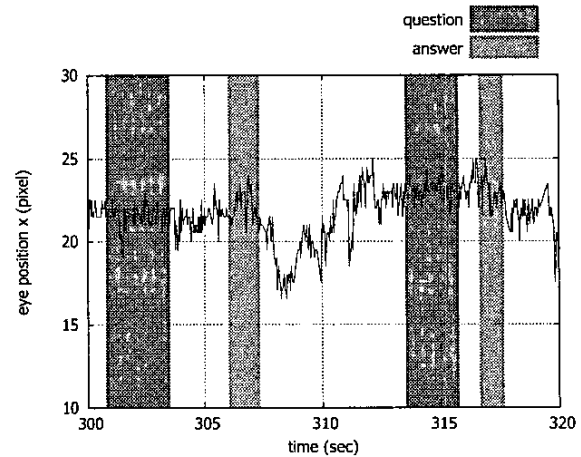


Figure 10: Eye x-position in the arithmetic task.

facial position and eye movements of an individual, it can detect their reaction and decide whether the content is attractive to him/her. It can then provide new contents in which he/she might be more interested.

Consider now small eye movements at fixation. According to [4], they can be very small. For example, tremors are very rapid and fine movements (20 to 40 arc second, 30 to 90 Hertz). If the eye ball radius is 13 mm, changes might be 1.3×10^{-3} mm. Thus a normal vision based system measuring over a wide area is unlikely to detect them.

Some psychologists have tried to find relationships between eye movements and mental workload[5]. Their contributions are very informative, but they want to observe phenomena and with to understand their physiological basis. If conditions change, a phenomenon will not reappear consistently. To implement the system that watches a human, it is important not only to apply psychological knowledge, but also to exercise the system in the assumed scene and to see whether phenomena are reproducible.

4 Conclusion

To support a human, it is important to detect whether he/she is concentrating or resting. To this end, a computer vision based face tracking system was applied to the problem. The system does not calculate "gaze direction" nor the point at which a subject is looking. It was confirmed that the system can detect the stability of eye movements and of head positing that may have a strong relationship to whether a subject is concentrating or not.

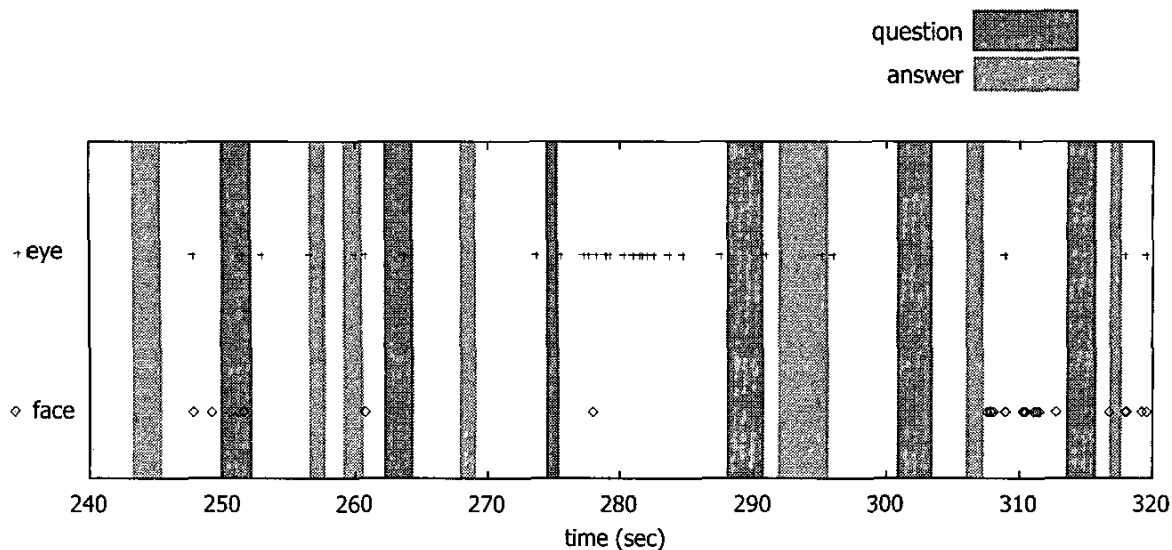


Figure 11: Result of an implementation of the position change detection system.

Acknowledgment

This research was supported by the Core Research for Evolutional Science and Technology (CREST Type), Japan Science and Technology (JST).

References

- [1] Matsumoto, Y. and Zelinsky, A. "An Algorithm for Real-time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement", Proceedings of IEEE Fourth International Conference on Face and Gesture Recognition (FG'2000), pp.499–505, 2000.
- [2] Vertegaal, R. "Attentive User Interfaces", Communications of the ACM, Vol. 46, No. 3, pp. 30–33, 2003.
- [3] Pashler, H. "Attention", Psychology Press, 1998.
- [4] Yarbus, A. L. "Eye movements and Vision", Plenum Press, New York, 1967.
- [5] May, J.G., Kennedy, R.S., Williams, M.C., Dunlap, W.P. & Brannan, J.R. "Eye movement indices of mental workload." Acta Psychologica, 75, pp. 75–89, 1990.