Abstract: This paper presents the software part of an inexpensive hands-free eye tracking system. The system works using infrared illumination like most of the available eye trackers. Two methods allowing estimation the gaze point on computer screen are compared. Research on effectiveness of these methods is discussed and the better one is indicated.

1 Introduction

Eye tracking is a technique allowing to determine gaze point on the computer screen. The information of analysing the eyeball movements (fixations and saccades) and characteristic parts of the eye image are used to determine the gaze point. Eye tracking devices are based on various hardware and software approaches. The most of hardware approaches are head-mounted eye trackers (Fig. 1a) or hands-free eye tracking systems (Fig. 1b). This paper concerns hands-free systems, characterised by the lack of additional physical equipment directly associated with the user. An image of the user’s face is registered by a video camera located under the display user looks at. Both head-mounted and hands-free eye trackers are based on the analysis and processing of video signals. This process is hardware supported by infrared diodes (IR LEDs). As a result, pairs of images are processed – with an illuminated and with a dark pupil. Various software approaches of eye trackers rely on using different analysis methods of the pairs of images and different methods of extraction of information from characteristic parts of the eye image. This paper describes two software approaches researched at the Multimedia Systems Department, GUT.

2 Research range

As mentioned in the Introduction, image processing is supported by diodes emitting infrared light. Initially, two images, with bright and dark pupil associated with the IR LEDs synchronization, allow obtaining a difference image (DI). The DI is the result of the subtraction of relevant pixel values between both images. A simplified scheme of obtaining the DI is shown in Fig. 2.

Fig. 2. Obtaining the difference image

Images are processed in gray scale with 8-bit depth. The determination of the pupil area on the DI is relatively simple. Afterwards, a characteristic corneal reflection (CR), the first image of Purkinje, technically known as “glint”, is visible in the dark pupil image [1] [4]. The glint remains almost a stable point despite eyeball movements. Therefore, it is regarded as a reference point for the center of the pupil.

The Multimedia Systems Department researched an eye tracking system focusing on two methods of estimating the gaze point based on an analysis of the relations between the glint and the pupil center. A common part of these two methods is the necessity of determining the pupil center (discussed in Part 3).

The first method uses four groups of IR LEDs located in the four corners of the display and one group of diodes located on the camera axis (around the lens). The LEDs on the corners generate four glints on the eyeball surface. The relative position of this set of glints depends mainly on the relations between head, camera axis and display. This method guaranties a relatively high tolerance against head movements.

The second method uses LEDs placed on the camera axis and diodes placed outside of the axis. These two groups of diodes are mounted in one device. This method is based on processing an eye image with one glint. The glint is created by the diodes lighting outside of the camera axis. This method is simpler than the first one but is not as resistant against head movements. The system can be calibrated only in one head position.
An assumption we have made before the development of the system was the discrimination of nine areas on the computer screen as shown in Fig. 7. These requirements are not too much restrictive, therefore both methods should be researched and assessed regarding their effectiveness, speed of working and hardware requirements, in order to choose the optimal one.

3 Determination of the pupil center

There exist many methods of determining the pupil center and the position of the glints. The Multimedia Systems Department researched some approaches individually and specified some conclusions. The first stage of the DI processing is binarization of the images with a threshold adopted because the glint characterises with a very high intensity pixels value compared to the rest of image. Then, the region of interest (ROI) is determined to reduce the possibility of detecting wrong glints. Some characteristic points of the eye image (pupil centre and glints) are determined within the ROI. Experiments on image samples certify that the ROI could be determined according to the brightest point on the DI. The brightest point is a reflection of the infrared light beam emitted by LEDs on the camera axis and it is always located close to the bottom side of the pupil. This point is not a glint in the meaning described in Part 2, but it could be qualified as a "quasi-glint" because it is visible as a characteristic light pattern. This quasi-glinit is directly associated with the bright eye image (Fig. 3). The proper glints are detected instead on the dark eye image.

Fig. 3. Quasi-glinit on a bright eye image

Thus, the quasi-glinit is the brightest point on the DI. Therefore, the location of the ROI is determined based on the detection of this point and defining the arbitrary horizontal and vertical coordinates of the range in which the characteristic points are located.

Specifying the pupil center (PuCe) is the most important stage in the gaze estimation process, because the system accuracy relies directly on the accuracy of the pupil center determination. Both researched methods use the information about the computed pupil center and there are several approaches of specifying it.

At the Multimedia Systems Department four approaches were investigated. Initially, a simple method based on the geometric center was implemented. After the first tests, it was sure that the precision of this approach is not enough because of some real variable conditions like irregular illumination or spherical eyeball. The second researched method was more complicated and based on the center of mass. This one is able to compute an accurate center of the pupil even though the pupil is incorrectly illuminated. This method also does not ensure the best accuracy of PuCe determination but is enough for the single glint method. If user does not move his head, the pupil is regularly illuminated. The next method was based on a horizontal and vertical projection as it is described in literature [1]. Unfortunately, also this approach does not guarantee enough accuracy. The last researched method is based on ellipse approximation. Some characteristic points on the edge of the pupil region are obtained and then the contour of an ellipse is determined. The PuCe is the point of intersection of both two ellipse axes [1] [3]. It is experimentally proved that this last approach allows obtaining the required accuracy and this is the one implemented in the final version of the four-glints method algorithm.

4 Four-glints method

As mentioned in Part 2, an eye tracking system based on the analysis of the PuCe and 4 glints uses a video camera and 5 groups of IR LEDs. The scheme in Fig. 4 shows the configuration of the researched system.

Fig. 4. Hardware configuration of eye tracking system with 5 groups of IR LEDs

After determining the PuCe, one can successively find the four glints in the ROI as the brightest points. An image sample with dark eye and four glints is presented in Fig. 5.

Fig. 5. Dark eye effect with 4 glints

Some pixels around each of the obtained points are discarded because the reflections of the LEDs produce the image bigger than 1 pixel. Moreover, sometimes some artefacts may appear nearby the pupil. Therefore, in order to avoid these artefacts, the four glints are searched only in a region determined as union of the ROI discussed in Part 3 and the pupil area. In Fig. 6, the bright points are glints on the cornea formed by the IR LEDs attached to the monitor.
If some of the glints could not be obtained, the results of the detection are not reliable.

**Gaze point estimation with characteristic point method**

The position of the gaze point can be approximated to one of nine areas based on the information of the PuCe, glints location and the size of the monitor screen. At the beginning, nine characteristic points (p1 to p8 and E – center of the polygon ABCD) are determined. The locations of the characteristic points are presented in Fig. 7. The system determines the observed by the user screen area based on the distances of these characteristic points to the PuCe.

![Fig. 6. Detecting four glints – concept illustration](image)

The points are determined with a weight ratio dependent on the glints positions. There are two iterations performed in order to improve accuracy of the estimation. Firstly, the algorithm determines the direction of the eyesight if user looks at the top, centre or bottom, left or right side of the computer screen. If the PuCe is close to one of the corners A, B, C, D then it is already known that user looks at area 1, 3, 7 or 9 without a second iteration performed, yet. If the shortest distance is between the PuCe and some of the p-points, then the another iteration is performed. The second shortest distance is computed to complete the approximation of the gaze point. This step of the algorithm allows estimating one of the nine areas.

![Fig. 7. Estimation of the eye gaze point by characteristic points method](image)

**5 Single glint method**

In order to research the single glint method, a special hardware device was designed by the Multimedia Systems Department. The device is shown in Fig. 8. The single glint in the cornea is generated by two sections of IR LEDs located out of the camera axis.

![Fig. 8. Hardware device designed for the single glint method](image)

Also, the first stage of this method requires determining the PuCe coordinates, as described in Par. 3. Then, one glint on the dark pupil image is detected. An image sample of the dark pupil with a single glint is presented in Fig. 9.

![Fig. 9. Dark eye with visible single glint](image)

The principle of the glint receiving is the same as in the four-glints method and depends on determining the brightest point in the ROI. The single glint method is based on the analysis of the relation between the PuCe and the glint coordinates. No complex mathematical operations or transformations are used. Unfortunately, the simplicity of this approach determinates some requirements of user’s head setup and head movements restrictions. In the first version of the application, it was assumed that user head is located on camera axis and larger head movements are forbidden. The area on which user looks is determined by the specific localization of the PuCe in relation to the glint. The representation of the relations between the characteristic points and the display areas is shown in Fig. 10.

![Fig. 10. Relations of characteristic points: „+” – PuCe, „.” – glint](image)

The main advantage of the single glint method is the tolerance to little changes of the eye position. The glint usually moves itself a little, but the relation between it and the PuCe is constant for small eye movements.
6 Methods comparison

The research was carried out on ca. 500 pairs of images. The methods were implemented in the MATLAB environment in order to simulate them and compare their effectiveness. At the beginning, different conditions of natural illumination were checked. Sunlight contains a lot of IR radiation which disturbs the proper work of eye tracker. For that reason system works better after the sunset, during cloudy days or in rooms with exclusively artificial illumination. The test image samples were recorded in the afternoon (normal light condition). The numbering of the screen areas consistent with order of looking is shown in Figs 7 and 10. During the tests the user was sitting in front of the display at a distance of 60cm. The used lens had a focal length of 12mm and the chosen image resolution was 800x600 pixels. Figures 12 and 13 present plots of eye gaze tracking course, respectively for the four-glint and the single glint method. The user was looking from region No. 1 to No. 9 in sequence. The vertical axis represents the parts of the computer screen. When the user runs his/her eyes over the monitor surface, the ideal plot of gaze tracking should be stepped-shaped plot.

![Fig. 12. Eye gaze tracking course – four-glint method](image1)

![Fig. 13. Eye gaze tracking course – single glint method](image2)

It is easy to notice that the second method (with a single glint) is more sensitive to transitions between looking areas. The four-glint method does not track all areas correctly but its plot tends to be stepped-shaped. Large deviations, seen in Fig. 12, between areas No.: 5 and 2, 4 and 1, 6 and 3, 9 and 6 do not cause large mistakes, because these areas are adjacent.

7 Conclusions

The described methods of eye gaze tracking present different approaches and solutions. There are many methods to determine the coordinates of the characteristic points and interpreting them. The results of the tests indicate, however, that the four-glint method is more effective in gaze estimation. Nevertheless, this method is still imperfect, so that it should be developed in the nearest future.

The system works in special conditions, mainly with stable user’s head and does not use any movements compensations methods. A permanent stable head positioning is not a natural situation, therefore the algorithm should cope with this kind of disturbances. Also, an important reason of inaccuracy of both methods is the lack of any calibration procedure. Consequently, it is planned in the next step of the research to apply calibration for several of the most important head positions. Another problem is associated with the actual size of user’s eyes. Some people have smaller eyes than others. The four-glints method could not work correctly in this situation, because the four glints would not situated within the pupil region in this case.

Also, the hardware equipment influences the accuracy of the system. The resolution of the tested images was 800x600 pixels with 12mm lens. In such a configuration, the diameter of the pupil region is represented by approximately 20-22 pixels. In the next version of the system a 16mm focal length and a video camera should be used with a better resolution, in order to provide a larger size of the pupil region for a more accurate determination of the PuCe regions.

It is worth mentioning also, that the system with the four-glints method, implemented in C++ programming language, works in real time on a standard PC.

References