Gaze Tracking for Mobile Devices Operation

Naoya Takeuchi† Hiroki Takahashi††

Graduate School of Informatics and Engineering
The University of Electro-Communications, JAPAN
††t1230052@edu.cc.uec.ac.jp ††rocky@inf.uec.ac.jp

Abstract

This paper aims for detecting eye gaze with a single camera attached on a tablet PC as an alternative human interface that supports device operations without hands manipulations. A vector from a center of an eye ball to a center of an iris is defined as an eye gaze vector. Feature points around eyes are extracted by ASM(Active Shape Model). The center of the eye ball is obtained by the mean of those feature points. The center of the iris is obtained by the mean of a black region in a binarized image extracted from the eye area. An eye width, that is a bounding box width of the feature points, has small differences among persons. Depth between both centers is estimated by the width and the radius. Angle variations of the vectors are measured. As a result, angle variations in the horizontal direction are relatively measured correctly, while those are not in the vertical direction.

1 Introduction

Currently, many kinds of mobile devices have been developed. Those devices attract attentions as alternative devices that play a part of telephones, the Internet interfaces, e-books and so on. Tablet PCs(Personal Computers) as represented by iPad® are terminals that combine high performance and portability. Those tablet PCs have clear displays. Tablet PCs can be used as alternative paper media for example magazines, newspapers, books, musical scores, scripts and etc. Everyone can easily and intuitively operate the devices with touch screen. It is, however, difficult to use the tablet PCs when both hands are occupied. For example, if you stand in a train, you will hold on to a strap by one hand and the other hand must be used for supporting a tablet PC. In order to solve this problem, this paper aims for detecting eye gaze with a single camera attached on a tablet PC as an alternative human interface that supports device operations.

Eye gaze manipulations are roughly divided into two categories[1], command-based interface and non-command-based interface. The command-based interface is a method that a user uses eye gaze intentionally. The interface employs only the eye gaze. It is a little bit hard to operate because person can not keep to see an exact point. The command-based interface can control at high speed. However, it is very usefull for disabilities.

Non-command-based interface estimates meanings and interests from eye gazes. User can not control non-command-based interface intentionally. Non-command-based interface, however, can be used for marketing applications.

This paper aims for detecting eye gaze with a single camera attached on a tablet PC as an alternative human interface that supports device operations without hands manipulations.

2 Previous work


J.Satake[6] provides an interactive information display that estimates face and gaze directions of a user in real-time. The system finds out user’s interest and obtains feedback from eye gaze with three high resolution cameras and detailed iris model. Users do not need to install any equipments. The system detects eye gaze when a user approaches his/her bodies. Three high resolution cameras capture a face and track feature points. The system allows to change face direction.

Y.Matumoto[7] provides intelligent wheelchair. A face is captured by stereo cameras. The face image included
eyes is a low resolution. The system searches eye gaze vector and facial movement. Wheelchair moves according to the direction of the human face.

These researches use multiple cameras to track eye gazes. A tablet PC has, however, only one camera. K.Arai[8] provides computer input system based on viewing vector estimated from iris center. It uses only one camera. The system, however, cannot detect eye gaze if other faces appears in the image. Because face extraction is realized by flesh color detection.

This paper aims for detecting eye gaze with a single camera, even if some faces exist in an image.

3 Face detection

This paper aims for detecting eye gaze with a single camera attached on a tablet PC. Therefore, this paper employs image processing based method. Overview of proposed algorithm is shown in Figure 1. At first, a face is captured by a camera. Face is detected by using Haar-like features. If some faces are detected, the most largest face is extracted as user’s face. Next, feature points around eyes are extracted. Eye feature points tracking sometimes fails. For example, it tracks a nasal cavity as an eye or left eye as right eye. Upper right and left regions are, therefore, extracted from the face region in advance. Then feature points in right eye are tracked within the right eye region. The left eye feature points tracking is performed in the same way. The extracted eye regions are shown in Figure 2.

4 Eye ball center definition

Feature points around eyes are extracted by ASM (Active Shape Model)[9]. ASM is a robust feature points tracking method to use a distribution of pixels in the contour of the image. Therefore, ASM can extract feature points if orientation of a face is changed as shown in Figure 3. ASM is required training data. This system uses Asian face images from “Face place.org”[10]. 104 eye images of seven men and women are used. Figure 4 shows an example of tracked feature points. The i th feature point coordinate is represented by \((a_i, b_i)\). The center of the eye ball is obtained by the mean of those feature points. It is given by following equation (1).

\[
\begin{align*}
    u_b &= \frac{1}{8} \sum_{k=0}^{7} a_k, \\
    v_b &= \frac{1}{8} \sum_{k=0}^{7} b_k
\end{align*}
\]

where \((u_b, v_b)\) donates the center of the eye ball.

5 Iris center definition

A bounding box of the feature points are shown in Figure 5. The top-left vertex is \((a_2, b_0)\) and the bottom-right vertex is \((a_6, b_4)\). Eye region is converted into gray-scale. It is smoothed by Gaussian filter for denoising. After that eye region is binarized by percentile method. Percentile
6 Eye gaze vector acquisition

A 3D vector from a center of an eye ball to a center of an iris is illustrated in Figure 7. This paper assumes that an image is taken with a camera by parallel projection into camera coordinate system. Deviation of the iris is detected from the center of the eye ball \((u_b, v_b)\) and the center of the iris \((u_i, v_i)\). Depth \(l_0\) between both centers is needed to get an eye gaze vector \(V\). It is given by the following formula (2).

\[
V = (u_i - u_b, v_i - v_b, l_0)^T
\]  
(2)

The depth \(l_0\) is estimated by a width and a radius. The width and the radius of an eye ball have small differences among persons. An eye width is about 30mm, an eye radius is about 13mm. An eye radius \(r_0\) in the image coordinate system from a center of an eye ball to a boundary point in 5 seconds.

Horizontal distance of \(P_{i,j}\) to \(P_{i,j+1}\) is 71.6mm. As a result, an ideal eye gaze angle variation in the horizontal direction is 7.1 degree. Vertical distance of \(P_{i,j}\) to \(P_{i,j+1}\) is 45.0mm. As a result, an ideal eye gaze angle variation in the vertical direction is 8.5 degree. Eye gaze angle variations in the horizontal \(\alpha\) and vertical \(\beta\) directions are measured when the subject is staring at a point in 5 seconds.

Horizontal eye gaze angles in Table 1. Angle variations in the horizontal direction are relatively measured correctly, while those are not in the vertical direction. Because iris is not detected correctly when iris is hidden in the eyelid.

\[
\alpha = \tan^{-1} \frac{u_i - u_b}{l_0}
\]  
(5)

\[
\beta = \tan^{-1} \frac{v_i - v_b}{l_0}
\]  
(6)
Table 1: Eye gaze angle variations in front of a tablet PC (degree)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right eye</td>
<td>4.53</td>
<td>10.6</td>
<td>9.27</td>
<td>11.2</td>
<td>10.7</td>
</tr>
<tr>
<td></td>
<td>-1.87</td>
<td>-1.33</td>
<td>0.42</td>
<td>-3.68</td>
<td>2.11</td>
</tr>
<tr>
<td>Left eye</td>
<td>8.56</td>
<td>8.59</td>
<td>6.67</td>
<td>5.64</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>2.12</td>
<td>1.55</td>
<td>-0.72</td>
<td>-0.79</td>
<td>3.25</td>
</tr>
</tbody>
</table>

7.2 Angle variations for face positions

Eye gaze vectors are measured when user positions are different. Experiment is performed for one subject. A subject is 300mm away from the left of the tablet PC. It shows Table 2. Angle variation in the horizontal direction are relatively measured correctly, while those are not in the vertical direction. The results of “Left”, “Front” and “Right” are similar.

Table 2: Angle variations for face positions (degree)

<table>
<thead>
<tr>
<th></th>
<th>Left</th>
<th>Front</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right eye</td>
<td>5.56</td>
<td>7.58</td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>-0.18</td>
<td>0.39</td>
<td>1.32</td>
</tr>
<tr>
<td>Left eye</td>
<td>5.89</td>
<td>4.24</td>
<td>7.65</td>
</tr>
<tr>
<td></td>
<td>-2.55</td>
<td>-2.46</td>
<td>-4.00</td>
</tr>
</tbody>
</table>

8 Conclusion

This paper aims for detecting eye gaze with a single camera attached on a tablet PC. Angle variations in the horizontal direction are relatively measured correctly but angle variations in the vertical direction are not measured correctly. Because iris is not detected correctly when iris is hidden in an eyelid.

We are planning to extract an iris exactly. Moreover, additional training data are necessary in order to realize feature points extraction. Furthermore, useful effective eye gaze interfaces for tablet PC’s must be discussed.

References


