Eye Contact Camera System for VIDEO Conference
Takuma Funahashi, Takayuki Fujiwara and Hiroyasu Koshimizu
School of Information Science and Technology, Chukyo University
e-mail: takuma@koshi-lab.sist.chukyo-u.ac.jp, {tfuji, hiroyasu}@sist.chukyo-u.ac.jp

Abstract
We took the face, especially eye or eye gaze, into consideration for discussing the non-verbal interface media. We proposed a method for generating eye-contacted facial images by computer image processing for enforcing and improving the quality of facial, nonverbal communication on the net such as VIDEO Conference. We second proposed a passive eye-camera system based on the facial image processing such as Hough transform for iris recognition.
Finally we applied these two proposed methods to the real environment of E-mail daily conference in the production process of the company, and shortly discussed on the prospects of the facial new media on the net.

1. Introduction –Facial Interface Media –

Human Communication on the network environment is becoming primal even in the industrial production processes and the non-verbal media for communication such as face is likely to be neglected or be in fatally low quality.
We introduce a system for supporting VIDEO conference system. Recently, VIDEO conference system could be used easily even in the mobile phone environment with camera, and many people use it in daily life. Since human is likely to look at the face of his partner on monitor not at camera, he will usually fail to send his own eye-contacted facial images to him, and vice versa. The basic idea to improve this fatal communication degradation is to regenerate facial image by changing the direction of the irises in the same original facial image.
However, main current contacting is e-mail in industry. As commonly known, so-called “emoticon (KAO-MOJI in Japanese)” characters have been utilized for compensating the lack of the “non-verbal” information (facial expression, gesture) in computer-mediated communications. However, it is becoming expectable to introduce directly several kinds of facial image media into the communication network infrastructure [1].
We second propose a system for extracting eye gaze information. In the E-mail case, for example, the sender wants to know how well the receiver can understand the content or not. To do this it is usually expected to utilize facial images. But, the face image is too large to be exchanged to each other. So, we paid attention to the digested features of eye gaze pattern on the E-mail letter extracted from the facial images by iris recognition.

2. Fundamentals via Image Processing

2.1 Face Region Extraction
The images with 24bit color and QVGA (320x240) in size were taken in a indoor circumstance. First, extract a skin region by using color image. Here we used HSV color table [2].
To put it concrete, let the color transform system be eq. (1) and generate a hue image, saturation image and value image.

\[
\begin{align*}
H &= \tan^{-1} \left( \frac{C_y}{C_x} \right) \\
S &= \sqrt{C_x^2 + C_y^2} \\
V &= -0.3R - 0.59G + 0.89B \\
\end{align*}
\]

(1)

Next, erosion, dilatation and labeling are applied to defect skin region. Fig.1(b) shows an example of face region extracted from the image given in Fig.1(a).
2.2 Iris Recognition

A method for recognition of irises from the gray images is proposed by using Hough transform for circle detection. Several candidates of a pair of the irises were extracted at first by applying Hough transform for circle detection to the binary image. The binarization method was especially constructed by [3]. The voting ranges of the parameter space \((a,b,r)\) were limited to some extent in order to reduce the computation cost and to enforce the performance. Parameters \(a\) and \(b\) indicate the center of iris. Parameter \(r\) indicates the radius of the iris. The best pair of the irises is detected from the candidates in accordance with the criteria standards given by that the number of votes is the biggest, that the positional relation between left and right irises is horizontal and that the radius of the left equals to the right. Therefore the recognition procedure was prepared as follows:

**Step1.** Set \(K\) as the threshold for the peak detection.

**Step2.** Detect the coordinates of irises whose peak is greater than or equal to \(K\).

**Step3.** Prepare the list of all candidates of the irises.

**Step4.** Choose the candidates pairs whose vertical distance is smaller than the threshold.

**Step5.** Choose the candidates pairs from the list whose radiuses are equal.

**Step6.** Among the rest of the list, extract a pair of right and left irises whose horizontal distance is minimum.

Fig.2 shows the example of the iris selection processing. Fig.2 (a) is a iris candidate extraction result. The one with the shortest segment that connects a right and left iris is adopted for the iris candidate. As shown in Fig.2 (b), it is known that the circle drawn in the extraction result image shows the iris position, and the iris is extracted accurately.

---

(a) Original Image       (b) Skin Region

**Fig.1. Face Region Extraction.**

(a) candidates of irises   (b) select pair of irises in (a)

**Fig.2. Iris Recognition.**

3. Eye Contact Camera System

3.1 Overview

As shown Fig.3, human is likely to look at the face of his partner on monitor not at camera, he will usually fail to send his own eye-contacted facial images to him.

The basic idea to improve this fatal communication degradation is to regenerate facial image by changing the direction of the irises in the same original facial image.

---

(a) Look away from me...

(b) transmitter side

**Fig.3. Problem of glance disagreement on VIDEO conference.**

3.2 Geometric Model of the Eye Contact

In order to model the situation of the VIDEO conference environment, as shown in Fig.3, the parameters \(R\) and \(r\) are specified for modeling the vertical relation between the camera and the monitor, and the parameter \(L\) is specified for the horizontal relation between them. This is depicted in Fig.4. In the beginning, let us imagine the iris moves from the
coordinate \((x_0, y_0)\) extracted before to the new coordinate \((x_1, y_1)\). This new coordinate \((x_1, y_1)\) can be easily calculated by eq.(2) and eq.(3) characterized with the parameters \(\theta\) and \(\theta_2\) indicating the spacial relationship among a person, camera and monitor. In this expression, functions \(\Delta x\) and \(\Delta y\) are designed to convert the parameter \(\theta\) to the number of the pixels in the facial image.

\[
\theta = \tan^{-1} \frac{R}{L} \quad \theta_2 = \tan^{-1} \frac{r}{L}
\]  \hspace{1cm} (2)

\[
x_1 = x_0 + \Delta x(\theta) \quad : \Delta x(\theta) = \frac{\theta}{10}
\]

\[
y_1 = y_0 + \Delta y(\theta_2) \quad : \Delta y(\theta_2) = \frac{\theta_2}{10}
\]  \hspace{1cm} (3)

3.4 Regeneration of the irises
The pixels \((x, y)\) in the region for new iris are painted in black at the region where the distance \(d\) between \((x_1, y_1)\) and \((x, y)\) is less than or equal to the radius \(d_0\) (as shown Fig.6). All pixels \((x, y)\) within the contour of the eyelid are painted in white at the region where the distance is greater than or equal to \(r_1\). The black and white colors are decided as follows:

\[
\text{black} = \min \{F_{ij} | f_{ij} = 1\}, \quad \text{white} = \max \{F_{ij} | f_{ij} = 0\}
\]

After this procedure, the irises are regenerated by smoothing. The procedure is shown in Fig.7.

4. Experiments and Considerations

The system is composed as follows;

1. CPU : Pentium 1.8GHz
2. RAM : 1.0GB
3. OS : WindowsXP
4. Software : Visual C++
5. Camera : 0.3Mpixel CMOS Sensor

4.1 Iris Recognition
Performance evaluation was executed. The experimental result of the iris recognition presented below.

Three kinds of data set are prepared as follows by...
changing three different distances between camera and face, and by changing the lighting without special light;
(a).30cm (looking into the monitor)
(b).50cm (typing closely at key board)
(c).80cm (typing apart from key board)
Fig.8 shows the example of the experiment result in each distance, and 25 frontal facial images (QVGA(320×240) size), and the iris recognition results are shown in Table.1.

<table>
<thead>
<tr>
<th>Distance</th>
<th>30cm</th>
<th>50cm</th>
<th>80cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success</td>
<td>18/25</td>
<td>21/25</td>
<td>16/25</td>
</tr>
<tr>
<td>72%</td>
<td>84%</td>
<td>64%</td>
<td></td>
</tr>
<tr>
<td>Failure</td>
<td>7/25</td>
<td>4/25</td>
<td>9/25</td>
</tr>
<tr>
<td>28%</td>
<td>16%</td>
<td>36%</td>
<td></td>
</tr>
</tbody>
</table>

In the distance 30cm, the edge extraction had failed because of the reflected light of the monitor. In the distance 50cm, the reflected light of the monitor was reduced and there was no problem in the edge extraction. In the distance 80cm, there were a lot of iris recognition failures due to the lack of resolution and glasses. The possibility of real time processing could be obtained because it was clarified that the process speed was about 20fps.

4.2 Iris Regeneration
Fig.9 shows the result of the generation of facial image of which eye gaze is contacted to the camera (to his partner).
In this experiment, since the monitor was set at the right side of the camera, his eye gaze was shifted to the left so that his eye gaze came to the position of the camera.

Since the successive relationship among the image frames was not yet implemented in this algorithm, the robustness of this application could be improved.
5. Passive Eye Camera System

5.1 Overview
Person's eye gaze information can be obtained by using the eye movement-tracking device [4],[5]. However, these devices are necessary the higher expertise for handling and are very expensive. Therefore, it is strongly expected to develop a new eye movement tracking device which is not expensive and is easy to use. We introduce a simple image processing system by using a Laptop PC and a Web camera.

5.2 System Flow
The system flow is shown in Fig.10. This system captures the image from Web camera, the face region extraction is applied to the image, and the iris is recognized in face area. Gaze points are acquired by using the series of the center coordinate values provided from the iris recognition.

5.3 Experiment and consideration
Experiment was executed as follows:
E-mail letter shown in Fig.11(a) with about 15 lines was displayed on the monitor of Laptop PC, Web camera was attached at the top of the monitor, and a person sit in front of this PC.
Fig.11(b) and (c) are an example of the eye gaze mark recorded during several seconds. Image processing for iris recognition was implemented by using VGA (640x480) image in this application.

It is notable that the eye gaze mark let us know how much carefully he watched the letter and consequently know the non-verbal quality of the verbal E-mail communication.

Although the processing speed of this experiment was about 5 fps, this application could be more realistic by introducing the simultaneous procedures to model the non-verbal quality based on the eye gaze mark.
6. Conclusion

In this paper, based mainly on the image processing techniques for the iris recognition, we proposed two kinds of applications for improving the quality of communication: a system of passive eye camera and a system of eye-contacted facial image generation. Through these system developments, we could demonstrate to introduce new facial interface media on the network environment.

As the coming subjects, the proposed systems are now being brushed up so that the real applications could be realized and the recognitions of other facial parts than irises are also being introduced for the complete digital modeling of the face.

Acknowledgment Authors would like to express thanks to all Koshimizu Lab members. A part of the research was supported by IMS HUTOP project, HRC project, JST project and NEDO project.

References


