MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11 - 13, 2002, Nara - ken New Public Hall, Nara, Japan

8—14

Extraction of face regions and eye windows from a set of gray-scale images with different face orientations

Shigeya Sasane*
MMC Computer Research Limited

Hideo Ogawa†
Faculty of Education
Aichi University of Education

Abstract

The enrollment of face templates and eye windows in a detection system of eye blinking or an estimation system of face orientation requires the extraction of face region and eye region in a face image. In order to automatically enroll face templates and eye windows with various face orientations, we present a novel method of extracting face region and eye window by using a set of gray-scale images. Experiments under various conditions(i.e., different illuminations, different backgrounds and so on) show very promising results. These experimental results show the method has possibility of application for enrolling the face templates and the eye windows.

1 Introduction

Recently, in automotive environment, a realization of the blinking detection has been tried in order to detect the driver's drowsiness by using image processing[1, 2, 3, 4]. Since the shortest eyeblink time, from eye close to eye open, is in about 100msec, one calculation period in blinking detection has to be in the video rate(33msec). And illumination condition dramatically changes in the automotive environment. Moreover, it should apply to any face orientations. Therefore, a realization of the blinking detection is not easy.

By the way, by the development of the face image processing, some studies of robust face image processing for the change of the illumination condition are reported[5, 6]. These studies use a face knowledge which is extracted by principal component analysis and so on. We think that face region and eye region of each face orientation must be memorized as prior knowledge in order to realize realtime blink detection.

As usual, in order to make the segmentation easy and accurate for the enrollment of face templates and eye windows, face images are observed in an artificial environment with simple background such as blue back. However, installing such environments needs much labor and it is difficult to install such environments in the automobile. Therefore the extraction method of face region is needed for the enrollment of face templates and eye windows, which is applicable even in various illumination conditions, or even in the complex backgrounds. On the other hand, even if a knowledge of face and facial components is used, it isn't easy to extract face regions and eye regions with various face orientations. Accordingly, not only knowledge of a face and facial components but also other knowledge is needed for the enrollment of face templates and eye windows.

By the way, for the enrollment of face regions and eye windows, we may ask a subject to turn his/her head toward various orientations and obtain a set of face images. Paying attention to the set of face images, we propose a novel method of extracting face regions and eye windows from a set of face images. Our method consists of two parts: extraction of face regions and detection of eye windows. The flow diagram of the procedure is shown in Figure 1.

*Address: Tsuchihara 3-1214, Tenpaku-ku, Nagoya, Aichi 468-0026 Japan. E-mail: sasane@quartz.ocn.ne.jp

†Address: Hirosawa 1, Igaya-cho, Kariya, Aichi 448-8542 Japan. E-mail: hogawa@auecc.aichi-edu.ac.jp

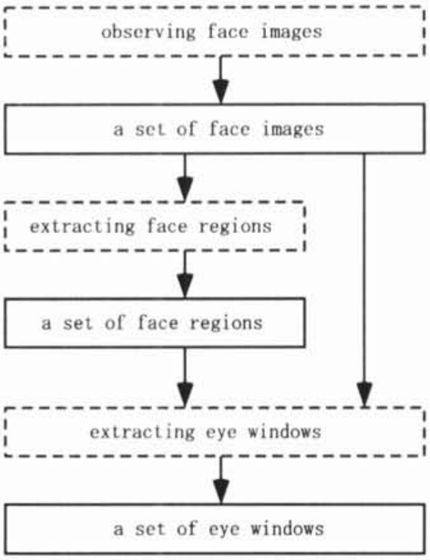

Figure 1: Flow diagram for the extraction of face regions and eye windows.

2 Definition of face orientation vector

In Figure 2, we show the situation of face observation. A subject is facing and looking at the marker $\boldsymbol{V}$. Here we define face orientation vector $\boldsymbol{v}$. We assume a plane $\boldsymbol{P}$ parallel to the view plane (i.e., image plane) of camera, and set the coordinate axes u and v which are parallel to the axes of image plane, because it is easy to estimate the location change of the eye region in image by the face orientation change. The inter-

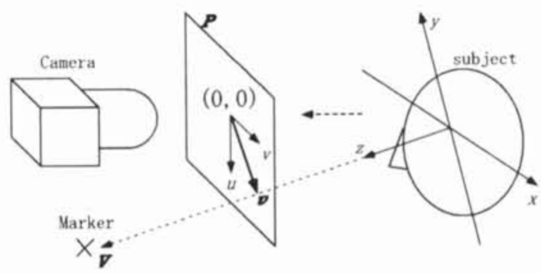

Figure 2: Definition of face orientation vector.

section (u, v) of a vector from the subject's face to the marker $\boldsymbol{V}$ and the plane $\boldsymbol{P}$ is called the face orientation vector denoted as $\boldsymbol{v}$ in Figure 2. The face orientation vector $\boldsymbol{v}$ is especially used in the extraction of eye window. Our method does'nt depend on the scale of the coordinate system. Therefore the distance between the subject's face and the plane $\boldsymbol{P}$ is arbitrary. In other words, our method has the robustness for the change of sheet position of each driver.

3 Extraction of face region

The basic idea of extracting face region in a face image with the specified face orientation is as follows. We ask a subject to turn the head toward the markers, and observe the face images $\{E_i\}$ with corresponding face orientations. Let B_{ij} be a binary image where the pixel value $B_{ij}(\boldsymbol{x})$ is 1 when there exists the difference between the gray values of the corresponding pixels in the images E_i and E_j.

Then B_{ij} contains both face regions in E_i and E_j. Therefore, by calculating $B_{ij}(\boldsymbol{x}) \wedge B_{ik}(\boldsymbol{x})$, where $j \neq k$, the face region in E_i can be enhanced, i.e., the pixels of face region in E_i take the value 1. Furthermore, by suppressing the noise segments and enhancing the face region by employing relaxation labeling, we can accurately obtain the face region in the image E_i.

4 Extraction of eye window

If the positional relation between the subject and the camera is already known, we can estimate the location change of eye region when the subject turns his/her head from a face orientation $\boldsymbol{v}_i$ to a neighboring face orientation $\boldsymbol{v}_j (i \neq j)$. By using this knowledge, we devised the eye window extraction algorithm as follows.

1. The differential image is constructed from the image that the subject opens his/her eyes and the image that the subject closes his/her eyes. There are not only eye regions but also other regions in the differential image.

2. The eye regions are chosen from the differential image by evaluating the correspondence between face orientations as follows.

   2-a. Let a region labeled "a" in the differential image be an eye region candidate $L(\boldsymbol{v}_i, a)$ corresponding to the face orientation $\boldsymbol{v}_i$. If the location of $L(\boldsymbol{v}_i, a)$ is consistent with the location of an eye region candidate $L(\boldsymbol{v}_j, b)$ corresponding to the neighboring face orientation $\boldsymbol{v}_j$, then the compatibility function $f(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b)$ is defined as 1, else the compatibility function $f(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b)$ is defined as 0.

   2-b. If the candidate $L(\boldsymbol{v}_i, a)$ is true eye region, there must be the corresponding, i.e., consistent candidates in all neighboring face orientations. From the compatibility function f, the degree of agreement of the hypothesis that $L(\boldsymbol{v}_i, a)$ is an eye region is defined as follows,

$$F(\boldsymbol{v}_i, a) = \bigwedge_{\boldsymbol{v}_j \in N(\boldsymbol{v}_i)} \bigvee_{b \in G(\boldsymbol{v}_j)} f(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b), \tag{1}$$

   where $N(\boldsymbol{v}_i)$ is a set of the neighboring face orientations of $\boldsymbol{v}_i$ and $G(\boldsymbol{v}_j)$ is a set of the region labels corresponding to the face orientation $\boldsymbol{v}_j$. If $F(\boldsymbol{v}_i, a) = 1$ is satisfied, then the candidate $L(\boldsymbol{v}_i, a)$ is regarded as 'true' eye region.

5 Supplement

Indeed, it is difficult to decide an appropriate threshold required to obtain the binary images B_{ij} mentioned above under various circumstances. In addition, the compatibility function $f(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b)$ in Eq.(1) between the eye region candidates $\{L(\boldsymbol{v}_i, a)\}$ may not be obvious. That is to say, it is difficult to apply *Boolean operators*. Therefore, we conducted various contrivances in our method.

For example, Eq.(1) is improved as follows.

$$Q^{(t)}(\boldsymbol{v}_i, a) = \operatorname*{median}_{\boldsymbol{v}_j \in N(\boldsymbol{v}_i)} \left\{ \max_{b \in G^{(t)}(\boldsymbol{v}_j)} \{q(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b)\} \right\}, \tag{2}$$

where q is the compatibility function. But q is a continuous function and satisfies $0 \leq q \leq 1$. The compatibility function q is defined as follows.

The potential eye region $E(\boldsymbol{v}_j|\boldsymbol{v}_i, a)$ is defined in order to include the eye region candidate $L(\boldsymbol{v}_j, a)$ when the face orientation vector changes from $\boldsymbol{v}_i$ to $\boldsymbol{v}_j$. And the potential eye region $E(\boldsymbol{v}_j|\boldsymbol{v}_i, a)$ is given as a region which surrounds $L(\boldsymbol{v}_i, a)$ and extends rectangle region to the direction of $\boldsymbol{v}_j - \boldsymbol{v}_i$ still the edge of the image. Figure 3 shows an example of the potential eye region $E(\boldsymbol{v}_j|\boldsymbol{v}_i, a)$, where $W_M(\boldsymbol{v}_i)$ is the width of face region corresponding to face orientation $\boldsymbol{v}_i$.

By using the potential eye region $E(\boldsymbol{v}_j|\boldsymbol{v}_i, a)$, the compatibility function q in Eq.(2) is given as follows.

$$q(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b) = \begin{cases} q'(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b) & \text{if } q'(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b) \leq 1 \\ \dfrac{1}{q'(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b)} & \text{otherwise} \end{cases} \tag{3}$$

$$q'(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b) = \frac{\|U(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b)\|}{\|L(\boldsymbol{v}_j, b)\|} \tag{4}$$

$U(\boldsymbol{v}_i, a|\boldsymbol{v}_j, b)$ is a set of pixels which are included in both the eye region candidate $L(\boldsymbol{v}_i, a)$ and the potential eye region $E(\boldsymbol{v}_i|\boldsymbol{v}_j, b)$. And $\|\cdot\|$ is the number of pixels.

The eye region candidate $L(\boldsymbol{v}_i, a)$ which satisfies Eq.(5) is regarded as 'true' eye region corresponding

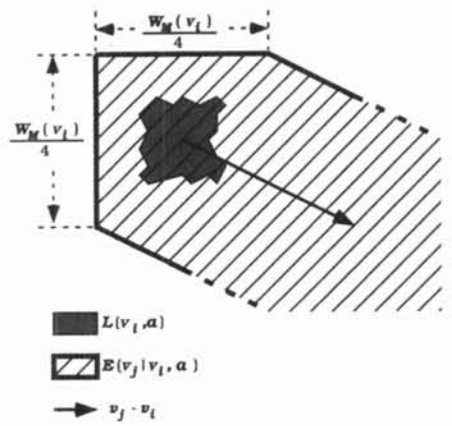

Figure 3: Definition of the potential eye region relative to the region $L(\boldsymbol{v}_i, a)$.

to face orientation $\boldsymbol{v}_i$.

$$a = \underset{b \in G^{(t)}(\boldsymbol{v}_i)}{\operatorname{argmax}} \{Q^{(t)}(\boldsymbol{v}_i, b)\} \tag{5}$$

However, if the set of region labels $G^{(t)}(\boldsymbol{v}_i)$ is large, a region unlike eye region may satisify Eq.(5). Accordingly, if the set of region labels $G^{(t)}(\boldsymbol{v}_i)$ is too large, 'true' eye region isn't determined and the set of region labels $G^{(t)}(\boldsymbol{v}_i)$ becomes smaller step by step as follows,

$$G^{(t+1)}(\boldsymbol{v}_i) = \{a | Q^{(t)}(\boldsymbol{v}_i, a) > \theta^{(t)}(\boldsymbol{v}_i), a \in G^{(t)}(\boldsymbol{v}_i)\}, \tag{6}$$

where $\theta^{(t)}$ is threshold and becomes bigger as follows,

$$\theta^{(t)}(\boldsymbol{v}_i) = \begin{cases} \vartheta^{(t)}(\boldsymbol{v}_i) & \text{if } \vartheta^{(t)}(\boldsymbol{v}_i) > 0.2 \\ 0.2 & \text{otherwise} \end{cases}, \tag{7}$$

$$\vartheta^{(t)}(\boldsymbol{v}_i) = \alpha(t) \times \max_{a \in G^{(t)}(\boldsymbol{v}_i)} \left\{ Q^{(t)}(\boldsymbol{v}_i, a) \right\}, \tag{8}$$

$$\alpha(t) = \begin{cases} 0.1 \times t + 0.1 & \text{if } t \leq 7 \\ 0.8 & \text{otherwise} \end{cases}. \tag{9}$$

Finally, when the set of region labels $G^{(t \to \infty)}(\boldsymbol{v}_i)$ becomes small enough, an eye region candidate $L(\boldsymbol{v}_i, a)$ which satisifies Eq.(5) is regarded as 'true' eye region.

We also conducted various contrivances in the extraction of face region. In section 3, we explained the enhancement of face region using binary images $\{B_{ij}\}$ and Boolean operator *AND*. However, if one image of a set of binary images $\{B_{ij}\}$ is obtained by inappropriate threshold, the accurate face region may not be extracted. But in practice it is difficult to decide an appropriate threshold required to obtain the binary images under various circumstances. Therefore, we apply Arithmetic operator *SUM* instead of Boolean operator *AND*. In order to enhance face region in E_i, we employ a set of binary images$\{B_{ij}\}$ obtained with the images $\{E_j\}$ facing to the neighboring face orientation $\{\boldsymbol{v}_j\}$ of the face orientation $\boldsymbol{v}_i$, and produce enhanced image W_i as

$$W_i(\boldsymbol{x}) = \frac{\sum_{\{j | \boldsymbol{v}_j \in N(\boldsymbol{v}_i)\}} B_{ij}(\boldsymbol{x})}{|N(\boldsymbol{v}_i)|}, \tag{10}$$

where $N(\boldsymbol{v}_i)$ is a set of the neighboring face orientations and $|\cdot|$ is the size of a set. For enhanced image W_i, the relaxation labeling is applied in order to enhance face region more. The detail is written in literature [7].

6 Experimental results and conclusions

In Figure 4, the results of our method are shown for all face orientations of one subject in the indoor. For all face orientations, our method provides a satisfactory result. In other experiments, however in the indoor, our method showed good results for various face orientations and various subjects under various backgrounds and illuminations. Of course, the illumination conditions change more dramatically in the automotive environment than in the indoor. Then, the result of applying our method to automotive environment is shown in Figure 5. For most of the face orientations, our method provides a satisfactory result. This experimental result shows that this method can be employed in automotive environment. Our method didn't extract eye regions when a subject turned the face extremely. But we think this problem can be solved by improving the definition of the potential eye region $E(\boldsymbol{v}_j | \boldsymbol{v}_i, a)$. We will improve our method soon.

This method can extract face regions and eye windows if there is a set of face images. In order to realize the automatic extraction of face regions and eye regions, it must be able to automatically acquire the set of face images. We will soon add the function that acquires a set of face images automatically to our method.

References

[1] Kiichi Tanabe, Makoto Sugiyama, "Automatic Extraction of Eyeblink for Psychological Experiment", IEICE, Vol.J76-D-II, No.5, pp.959-966, 1993.

[2] Tomoaki Nakano, Kazuhiko Sugiyama, Shin Yamamoto, "Measuremnt of driver's blinking durations for detecting drowsiness during vehicle driving", Proceedings of the 1st symposium Sensing via Image Information. pp.59-64, 1995.

[3] Shunji Katahira, Satoko Nara, Masayoshi Aoki, "Driver Drowsiness Detection by Eyelids Movement from Face Image", Proceedings of the 2nd World Congress on ITS, pp.1217-1222, 1995.

[4] Kazuo Okino, Syouiti Tagawa, Masataka Ariga, Ken Ishihara, "Development of Advanced safety system to alarm drivers for sleepy drive by multimedia sensing without adhesive part", Proceeding of the 4th symposium Sensing via Image Information, pp.285-288, 1998.

[5] Masaki Sohma, Kenji Nagao, "Robust Face Recognition under Drastic Changes of Conditions of Image Acquisition", IEICE, Vol.J80-D-II, No.8, pp.2225-2231, 1997.

[6] Kazuhiro Fukui, Osamu Yamaguchi, Kaoru Suzuki, Ken-ichi Maeda, "Face Recognition under Variable Lighting Condition with Constrained Mutual Subspace Method – Learning of Constraint Subspace to Reduce Influence of Lighting Changes –", IEICE, Vol.J82-D-II, No.4, pp.613-620, 1999.

[7] Shigeya Sasane, Hideo Ogawa, "Face Detection Algorithm Using a Set of Gray Image Sequences" IEICE, Vol.J83-D-II, No.3, pp.1064-1069, 2000.

Figure 4: Results of the extraction of face regions and eye windows for one subject in the indoor.

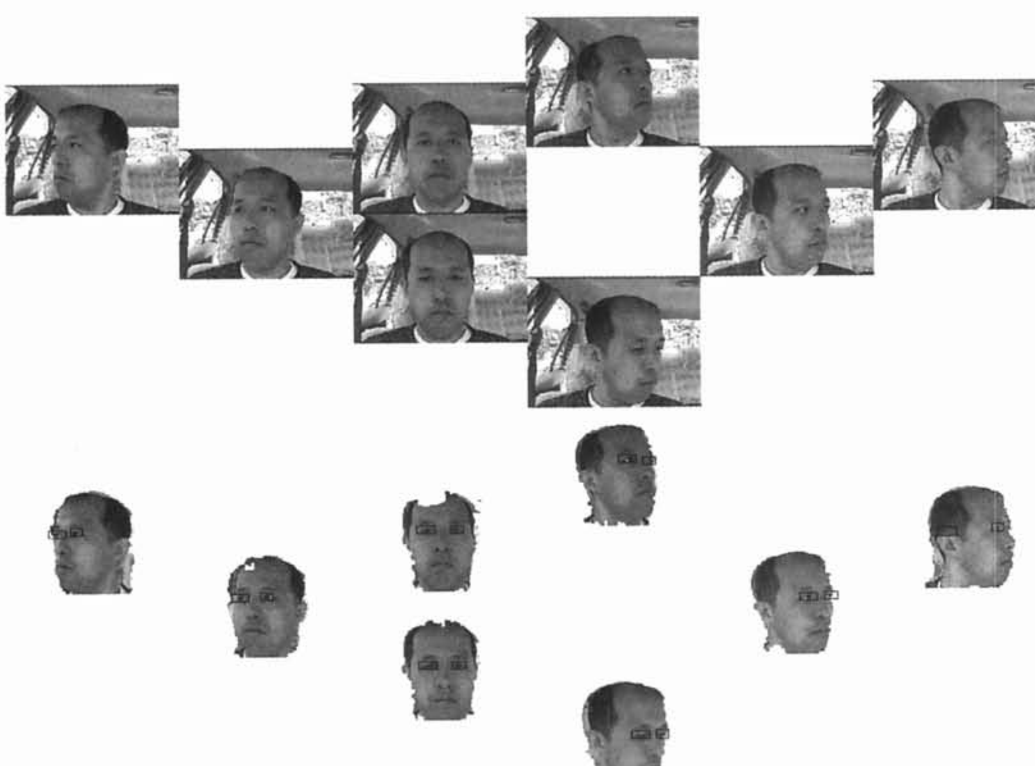

Figure 5: Results of the extraction of face regions and eye windows for one subject in automotive environment.