ACCV2002: The 5th Asian Conference on Computer Vision, 23--25 January 2002, Melbourne, Australia.

An Iris Detection Method Using the Hough Transform and Its Evaluation for Facial and Eye Movement

Hideki Kashima* Hitoshi Hongo** Kunihito Kato* Kazuhiko Yamamoto*
**Faculty of Engineering, Gifu University **HOIP, Softopia Japan/JST*
kashima@yam.info.gifu-u.ac.jp

Abstract

Our research aims at applying information with regard to a user's gaze to computer interaction. In this paper, we propose an iris detection method that is adaptable to various face and eye movements taken from a facial image. First, our method determines the standard skin color from previous frame. The face region is determined using the color difference from the standard skin color. Secondly, the eye and mouth regions are extracted from the face region by hybrid template matching, which consists of edge and color features. Using our method, facial parts can be extracted against individual variation of faces and head movements. Finally, the iris regions are detected by using saturation and brightness from the extracted eye regions. Then, the iris positions are detected by the Hough transform. We evaluated our method using facial images containing various positions of faces and eyes. Our method achieved an accuracy rate of approximately 96% for eye region detection, and approximately 92% for iris detection.

1. Introduction

In recent years, various methods that to detect irises for gaze detection, applied to man-machine interfaces, have been proposed [1][2][3]. However, most research did not report the relationship between the accuracy rate of the iris detection and the directions of the face and the eyes in detail. Actually, the shape of facial parts are not only different depending on the person, but also dynamically change by movement of the head, facial expressions, and so on. In particular, when a person faces down, the eye regions tend to be dark. In this case, it is difficult to detect irises in the eye region. Therefore, we propose an iris detection method that is adaptable to individual variation and directions of faces and eyes. Our method extracts the eye regions before detecting irises. Extraction of facial parts requires adaptation to great change in shape. We have proposed a facial part extraction method, which utilizes the four directional feature fields and the color distance feature [4]. First, to detect irises, iris regions are extracted by saturation and brightness. Next, the position of the iris is detected by the Hough transform.

In this paper, we will describe the eye-region extraction method in section 2. We will explain about the iris detection method in section 3 and in section 4 we will present the results of the iris detection experiment. Finally, in section 5 we will conclude our experiment.

2. Extraction method of facial parts

2.1. Face detection

Before detecting the irises, we think it is necessary to identify face regions from the input image. Some methods that extract face regions by color information have been proposed [5]. Popular methods often use hue and saturation values in the HSV color system. However these methods do not work well if the target objects are not bright enough. We decided to use U V values in the CIE-LUV color system from the preliminary experiment. The CIE-LUV color system consists of lightness(L) and color values(U and V). Theoretically, the CIE-LUV color system is just as capable as humans when distinguishing between any two colors of the same brightness. Our method that extracts face regions is described as follows.

First, a two-dimensional histogram on the UV plane is constructed using the facial image from the previous frame. In the two dimensional histogram, our method determines the standard skin color, which indicates the maximum number of pixels within the range of skin UV values. The range of skin UV values has been decided from a sample of several Japanese people. Figure 1 shows an example of the UV value distribution. As shown in figure1, the rectangle shows the range of skin UV values, and the position of the cross indicates the standard skin color. The distances from the standard skin color to each pixel of UV values in the input image are calculated. Figure 2 is an original image.

Figure 3 shows the UV value distance from the standard skin color. In figure3, the level of brightness indicates the distance from the standard skin color. Thus, dark areas indicate a close match to standard skin color. Next, we make a UV value difference histogram from the above results and then extract the skin regions by the discriminant of Otsu. Figure 4 shows an example of the detected face region.

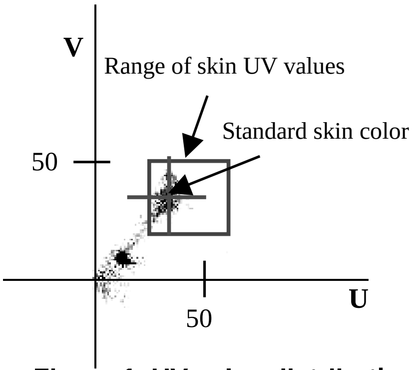

Figure 1. UV value distribution

Figure 2. Original image

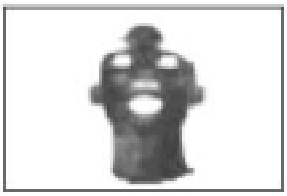

Figure 3. Distance image Figure 4. Detected face region

2.2 Extraction of facial parts

Many methods that extract facial parts have been proposed. For example, a method exists that detects eye regions by edge projection analysis [6]. Another method detects facial parts in Eigenspace [7]. However, these methods could not obtain the same accuracy rate when subjects tilted or moved their heads. Therefore, we proposed a hybrid template-matching method that uses the four directional feature fields [8] and the color distance feature. This method allows for an improved accuracy rate, even if a person tilts or moves their head.

Figure 5 is an example of the four directional feature fields and the color distance feature template for facial parts. The four directional feature fields consists of four images which are divided by edges into the horizontal, vertical, upper right, and upper left. They are reduced to low resolution through the Gussian filter. Even in low resolution, this feature keeps the information of the edge directions in comparison with the feature that does not divide the direction of edges. Therefore, the four directional feature fields are adaptable to change in shape. The color distance feature finds the mean of color distance from the standard skin UV values of skin to each pixel of UV values. Although the shape of facial parts dynamically change when the head is moved, this feature is useful. Because there is little change of color compared with the change in shape. Before extracting facial parts, the right eye, left eye, and mouth images are prepared in order to construct the initial template. Then the four directional feature fields and the color distance feature are extracted from each facial part image.

Next, similarity is calculated between the templates and the input image, which was analyzed with the same processing as templates. Similarity is decided according to the criteria (1), (2) and (3). The similarity of the four directional feature fields '*E*', is calculated by equation (1), and the similarity of color distance feature '*C*', is calculated by equation (2). '$I_{(x,y)}(i, j)$' represents a feature of the input image, '$t(i, j)$' represents a template feature. Let the size of the template be '*mn*'. '*d*' indicates a direction of an edge. '*U*' and '*V*' indicate the distance from the standard skin color for the U and V components, respectively. Total similarity '*S*' is evaluated by equation (3), and '*w*' is the weight associated with color similarity. In this paper, we let *w* equal 1. Regions whose total similarity marks higher than the threshold are registered as candidates. After matching, our method tests all combinations that satisfy certain restrictions regarding the relative positions between the eyes and mouth.

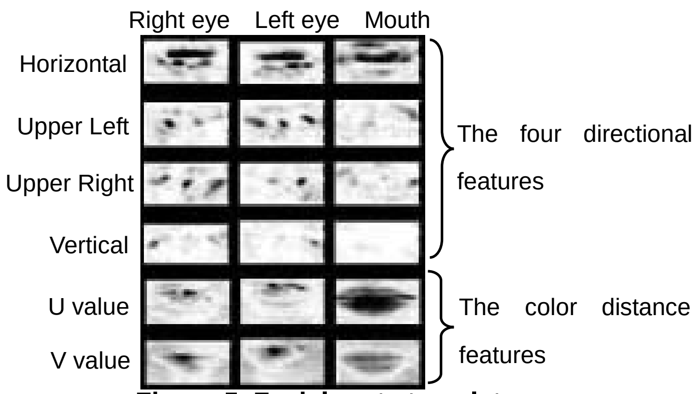

Figure 5. Facial parts template

$$E=\frac{\left\{\sum_{d}^{4}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{d(x,y)}(i,j)\times t_d(i,j)\right\}^2}{\sum_{d}^{4}\sum_{i=1}^{m}\sum_{j=1}^{n} I_{d(x,y)}(i,j)^2\times\sum_{d}^{4}\sum_{i=1}^{m}\sum_{j=1}^{n} t_d(i,j)^2} \quad (1)$$

$$C=1-\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(I_{u(x,y)}(i,j)-t_u(i,j)\right)^2+\sum_{i=1}^{m}\sum_{j=1}^{n}\left(I_{v(x,y)}(i,j)-t_v(i,j)\right)^2}{m\times n\times((2\times U_{max})^2+(2\times V_{max})^2)} \quad (2)$$

$$S=E+wC \quad (3)$$

3. Iris detection method

In this section, we describe iris detection from the extracted eye regions. First, since skin regions tend to be higher in saturation values than eye regions, the extracted eye region is divided into the eye and skin regions by a saturation histogram using the discriminant of Otsu. Then, the eye region is divided into the iris and the white of the eye from the brightness histogram also using the discriminant of Otsu. Thus, the dark region is the iris. Employing the discriminant of Otsu, we can separate the iris and the white of the eye even if the eye regions are dark. Figure 6 is an extracted eye region image, and the result of the segmented skin and eye is shown in figure 7. The iris region is calculated from the extracted dark region. Next, the edge of the iris is detected from the gray-scale image in the iris region by Prewitt’s operator. Figure 8 shows the detected edges of the iris in the iris region. The size of the frames in figure 6, 7 and 8 are 72×48 pixels. After the edges of the iris are detected, the position of the iris is determined by the Hough transform. Our method is shown as follows.

First, our method searches for an outline of the iris edge. The possibility of a circle projects on a three-dimensional parameter space, which consists of the position of the center of the iris and the radius of the circle. The Hough transform is robust against noise, and detects the iris even if only a part of the outline of the iris appears. However, the possibility of many iris candidates have appeared by using the Hough transform. In order to reduce these candidates, our method restricts the edge to within the iris region, which is extracted in the previous process. In addition, to improve the speed of the process, the center of the circle is limited within the iris region. In this experiment, the radius is coordinated in a range of 7 to 12 pixels, which have been determined from previous experiments.

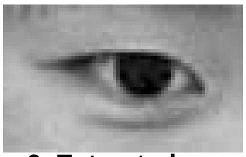

Figure 6. Extracted eye region

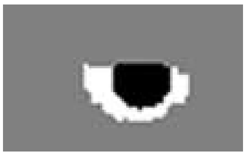

Figure 7. Segmented eye and skin

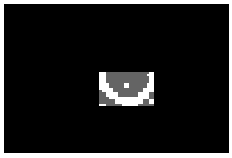

Figure 8. Detected edge of iris

4. Iris detection experiment

We experimented with iris detection by using our method. In this experiment, we experimented with many directions of facial and eye images. Experimental data is explained in 4.1, and then the result of iris detection is discussed in 4.2.

4.1. Experimental data

We have established a multiple camera system for the collection of the experimental data. Using the multiple camera system, we gathered facial images containing various positions of the head and eyes. Figure 9 shows the environment where experimental data is collected. The multiple camera system consists of 15 cameras, and it can take images simultaneously. The subject sits on the chair and faces index 13, 28, and 43 in order making it possible to take pictures of the face in 45 directions. These faces vary by 40 degrees in the horizontal and 20 degrees in the vertical at intervals of five degrees. The subject then gazes at 25 indexes when facing the index of 13, 28, or 43. Thus, it is possible to take pictures of the subject’s gaze in 25 directions. These gazes vary by 10 degrees in the horizontal and vertical at intervals of five degrees.

In this experiment, we used experimental data which was collected by the multiple camera system. Experimental data consists of a total of 1,125 different facial and gazing directions, 25 gaze directions for every 45 facial directions. Ten subjects, who did not wear glasses, were used in the experiment. The number of images of experimental data totals 11,250 images. The image data consists of input with dimensions of 640×480 pixels in 24-bit. In these images, the size of the face is about 220×260 pixels. The background is simple, because we are aiming to evaluate the extraction of facial parts in this experiment.

Figure 9. Multiple camera system

4.2. Results of iris detection

We experimented with eye extraction and iris detection from the successful, extracted eye regions taken from the experimental data. We checked the results of facial part extraction shown in table 1. Figure 10 shows the results of each facial part extraction. In figure 10, each rectangle indicates the extracted eye and mouth regions. The circle in the rectangle indicates detected irises. Hereafter, the results of eye region extraction and the results of iris detection are evaluated. Figure 11 (a) and (b) show the accuracy rate of eye region extraction for each direction faced. Figure 11 (a) shows the rate of extraction when the direction of face changes from the front horizontally. Figure 11 (b) shows the rate of extraction when the direction of the face changes vertically. In figure 11, the rate decreased when subjects faced away horizontally than vertically. However, our method was more than 90% accurate when evaluating subjects facing 20 degrees away from center. An example of a failure in this experiment was when eyebrows were incorrectly, extracted. However, since the relative position relation between an eye and a mouth is unusual when eyebrows were incorrectly, extracted, we considered a method for reducing incorrect detection from this position relationship. Figure 12 (a) and (b) show the accuracy rate of iris detection for the each facial direction. Figure 12 (a) shows the rate of extraction when facial direction changing from the front horizontally. Figure 12 (b) shows the rate of extraction when facial direction changing vertically. In figure 12, it turns out that the rate of detection decreased when the direction of faces changed horizontally. It is thought that this is because an iris becomes difficult to detect when the direction of a face changes horizontally. In generally, it is thought that when a person looks directly ahead an iris can be detected more clearly than a person who is looking away. Therefore, we made a figure that shows the relation of iris detection between the camera position and gazing point. The rate of iris detection at each gazing point is shown in figure 13 (a) and (b). Figure 13 (a) shows the rate of detection when the gazing point changes from directly in front moving away horizontally. Figure 13 (b) shows the rate of detection with the gazing point changes vertically. Figure 14, 15, and 16 show the results of the extracted eye region. Each image on the left is an eye region image with the result of the detected iris. Each image on the right is the edge around an iris. In figure 13, the accuracy rate of detection decreased as the subjects faced away. The accuracy rate decrease particularly when subjects looked downward. When a subject looks downward, the eye region becomes very small and a shadow is cast making the iris almost undetectable. Therefore, one of the problems of detecting gaze points is thought to be when subjects look in a downward direction. On the other hand, light is reflected off an iris when a subject looks upward. Figure 14 shows a subject looking upward. In this case the reflected light creates a false edge. This makes it possible for our method to misinterpret the direction of the iris. However, since we have used the Hough transform our method automatically adapts to interference of this level. Figure 15 shows a subject looking sideways. When looking in a horizontal direction, the outline of an iris may not even appear. Even though it is thought that the Hough transform would detect this, only part of the outline appears. Figure 16 shows a subject looking downward. Figure 16 was detectable, although the outline of the iris did not appear well under the lighting conditions.

Table 1. Results of facial part extraction

	Iris detection	Eye region extraction
Accuracy rate (%)	91.7	95.8

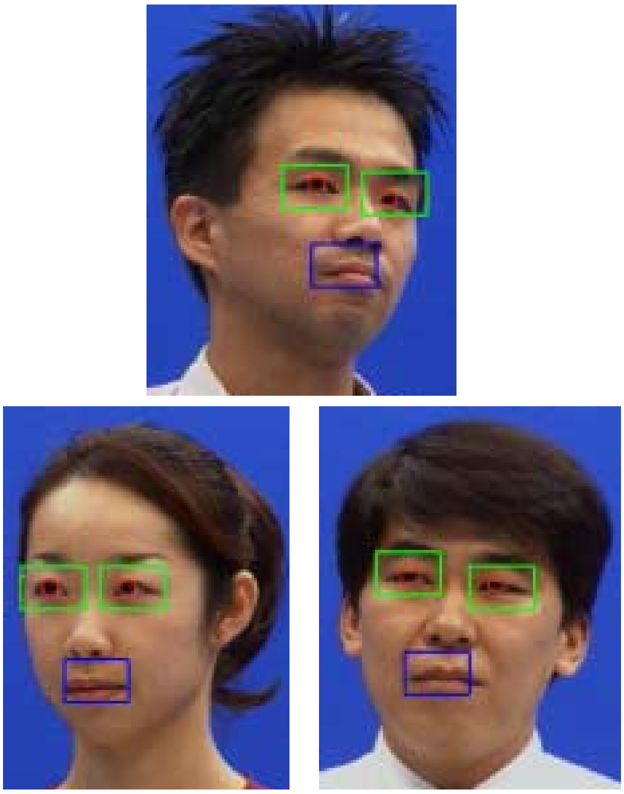

Figure 10. Results of each facial part extraction

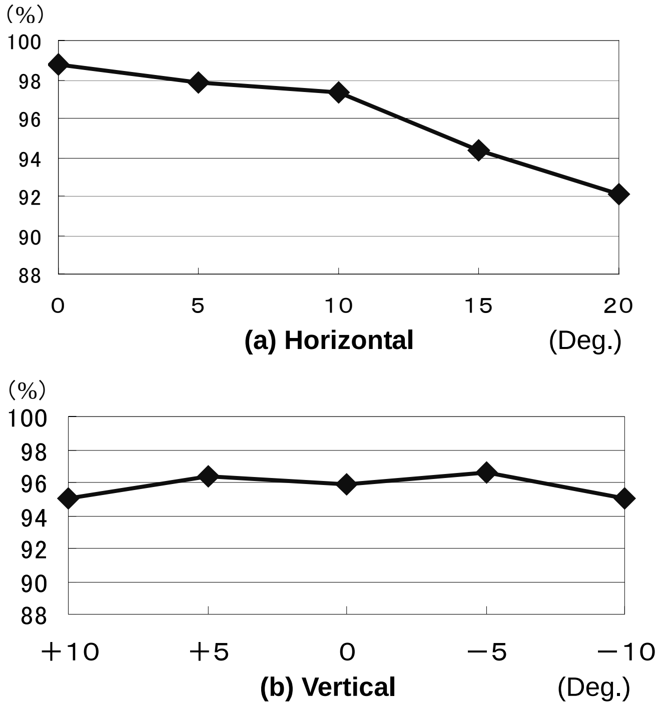

Figure 11. Accuracy rate of eye region extraction

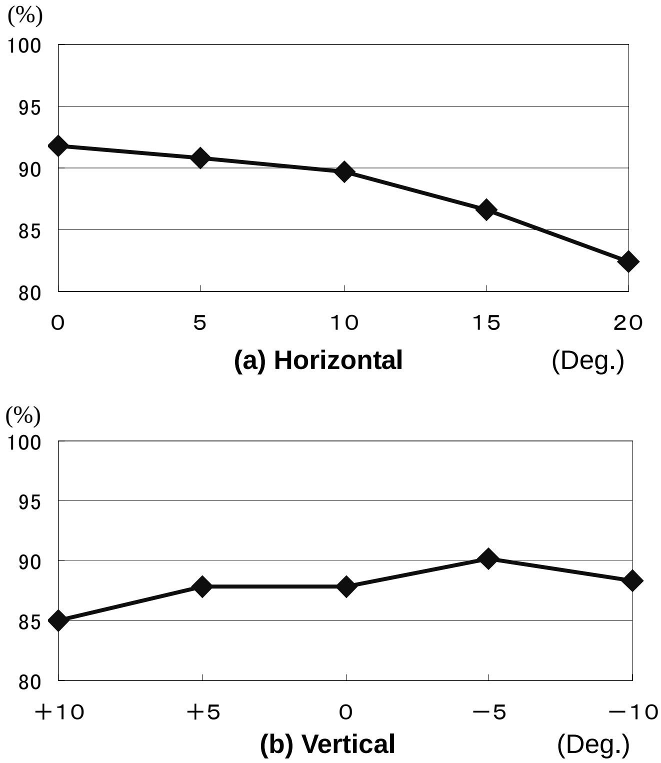

Figure 12. Accuracy rate of iris detection

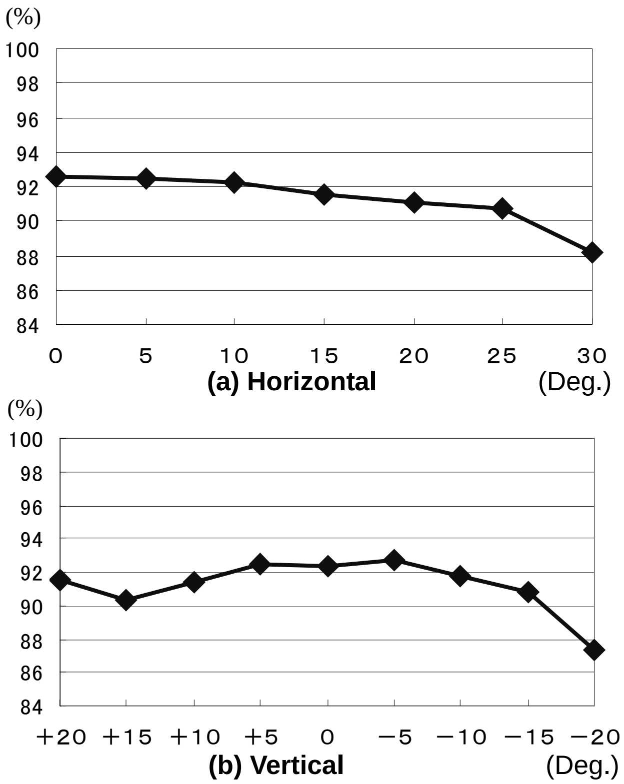

Figure 13. Accuracy rate of iris detection

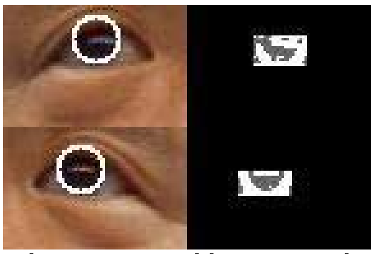

Figure 14. Looking upward

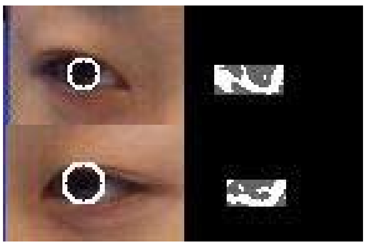

Figure 15. Looking sideways

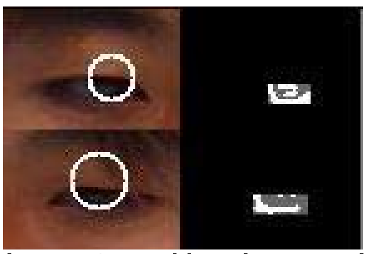

Figure 16. Looking downward

5. Conclusions

We proposed an iris detection method using the Hough transform that adapts to various face and eye positions. First of all template matching, which consists of the four directional features field and the color distance features, was used in the extraction of the eye regions. Then the positions of irises were detected by the Hough transform.

We experimented with facial parts extraction from facial images with various face and eye positions. Experiments showed that our detection method was stable, but a few problems were found when experimenting with various gazing points. Taking this problem into consideration we aim to further stabilize our method in future works.

Acknowledgments

This research was partially supported by the Gifu Prefecture Regional Intensive Research Project and JST.

The authors would also like to thank Alec Hansen for comments on this paper.

References

[1] Y.Segawa, H.Sakai, T.Endoh, K.Murakami, T.Toriu , and H.Koshimizu. "Face Recognition Through Hough Transform for Irises Extraction and Projection Procedures for Parts Localization -for Facial Caricaturing System PICASSO-," Proc. PRICAI'96, pp.625-636, 1996

[2] Y.Matsumoto, and A.Zelinsky. "An Algorithm for Real-Time Stereo Vision Implementation of Head Pose and Gaze Direction Measurement," Proc. FG2000, pp.499-504, 2000

[3] O.Yamaguchi, K.Fukui, and K.Maeda. "Face recognition using temporal image sequence," FG98, pp.318-323, 1998

[4] H.Hongo and K.Yamamoto. "Detection of Face and Facial Features using Directional Features and Color Information ," Proc. of 4th, Joint Conference on Information Sciences(JCIS'98) Vol.4, pp.350-353, 1998

[5] A.R.Smith. "Color Gamut Transform Paris," Proc. of ACMSIGGRAPH'78, pp.12-18, 1978

[6] R.Brunelli and T.Poggio. "Face Recognition: Feature versus Templates," IEEE, Transactions on PAMI, Vol. 15, No.10, pp.1042-1052, 1993

[7] A.Pentland, B.Mohaddam and T.Starner. "View-Based and Modular Eigenspaces for Face Recognition," CVPR, pp.1-7, 1994

[8] K.Yamamoto, "Present State of Recognition Method on Consideration of Neighbor Points and Its Ability in Common Database," IECE Transactions, Vol.E79-D, No.5, pp.417-422, 1996

[9] N.Hagai, H.Hongo, K.Kato, and K.Yamamoto, "Detection of Facial Parts Using Edge and Color Features," Proc. of 4th International Conference on Virtual Systems and MultiMedia, Vol.2, pp.380-385, 1998