IEEK Transactions on Smart Processing and Computing

Real-Time Rotation-Invariant Face Detection Using Combined Depth Estimation and Ellipse Fitting

Daehee Kim, Seungwon Lee, and Dongmin Kim

Department of Image, Chung-Ang University / Seoul 156-756, Korea wangcho100@gmail.com

* Corresponding Author: Daehee Kim

Received June17, 2012; Revised July 14, 2012; Accepted July 31, 2012; Published October 31, 2012

***Abstract*:** This paper reports a combined depth- and model-based face detection and tracking approach. The proposed algorithm consists of four functional modules; i) color-based candidate region extraction, ii) generation of the depth histogram for handling occlusion, iii) rotation-invariant face region detection using ellipse fitting, and iv) face tracking based on motion prediction. This technique solved the occlusion problem under complicated environment by detecting the face candidate region based on the depth-based histogram and skin colors. The angle of rotation was estimated by the ellipse fitting method in the detected candidate regions. The face region was finally determined by inversely rotating the candidate regions by the estimated angle using Haar-like features that were robustly trained robustly by the frontal face.

***Keywords*:** Real-time face detection, Ellipse fitting, Depth histogram

1. Introduction

Face detection is of great importance in the computer vision and pattern recognition fields, and has many applications such as human-computer interaction, intelligent monitoring, face synthesis and object-based coding [1]. Viola et al. [2] proposed a feature-based face detection method using the Adaboost algorithm. Although Viola et al.'s method works well with frontal faces, it often fails to detect slanted faces. Linehart first proposed a method to detect specially rotated face regions, and extended it to multi-view face detection based on Viola et al.'s method [3]. Although color-based methods are robust to variations in rotation and scale, but they cannot deal with complicated background or occlusion [4]. Generally, a face detection method based on skin color is affected by the illumination conditions and camera characteristics [5, 6].

This paper presents a robust face detection algorithm to overcome both occlusion and variable orientation problems with the following four functional modules: i) A candidate of the face region is first detected using skin color, ii) a depth histogram is generated using a stereo camera, iii) the face region is finally determined by fitting the candidate region to an ellipse, and iv) the detected face region is tracked by appropriately predicting the motion of the region. Fig. 1 gives a flow-chart of proposed algorithm.

This research was supported by the Chung-Ang University Research Scholarship Grants in 2010.

2. Face Region Extraction Using Decorrelated Skin Color and Depth Information

The face candidate regions were detected using a decorrelated skin color component and depth information. The chromatic red component in the YCbCr color space, which was first utilized in reference [7], was used to decorrelate the skin color. The chromatic red (Cr) component can extract and compensate for an error signal. The skin regions are detected using the Cr and error components of ranges, R_{ES}=[0.02511,0.1177] and R_{Cr}=[135,173], which are the experimentally obtained error signal and Cr component, respectively [8]. The skin color segmentation process is defined as:

$$f_{skin}(x) = \begin{cases} 1, & \text{if } E(x) \in R_{ES} \text{ and } Cr(x) \in R_{Cr} \\ 0, & \text{otherwise} \end{cases} . \quad (1)$$

A color-based segmentation algorithm, however, is sensitive to occlusion in a complex background. The

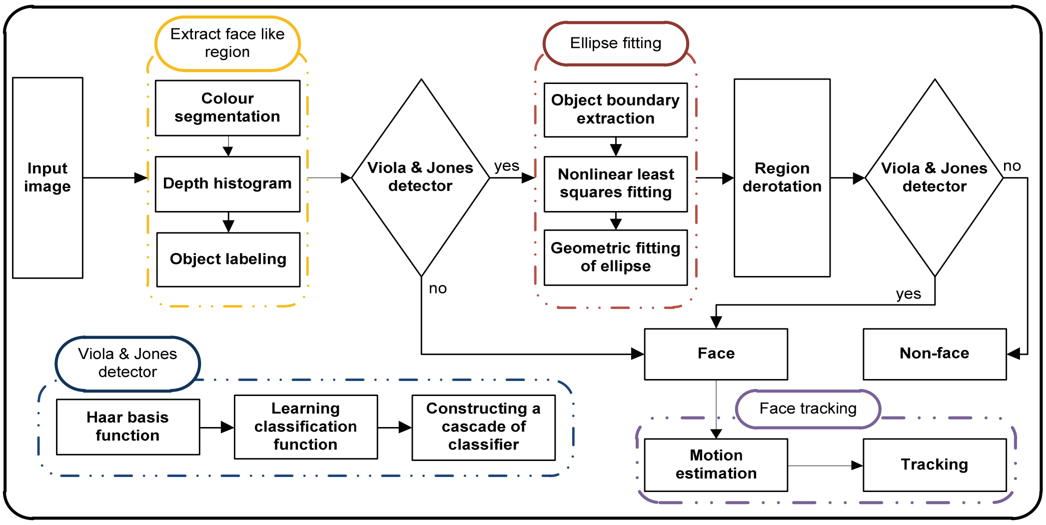


Fig. 1. Flow-chart of proposed algorithm.

following depth histogram of candidate regions was used to overcome this problem:

$$f_{depth}(x) = \begin{cases} 1, & \text{if } th_{front} < d(x) < th_{back} \\ 0, & \text{otherwise} \end{cases}, \tag{2}$$

where d represents the depth map of an image acquired by a stereo camera, and th_{front} and th_{back} are the threshold values of the foreground and background, respectively. The th_{front} and th_{back} values were 30 cm and 80 cm, respectively. Each region is divided according to the histogram. The best candidate region was selected as the final face region as follows:

$$f_{face}(x) = \begin{cases} 1, & \text{if } f_{skin}(x) = 1, \text{ and } f_{depth}(x) = 1 \\ 0, & \text{otherwise} \end{cases}. \tag{3}$$

3. Rotation-Invariant Face Detection and Tracking

Although Viola et al.'s face detection method is fast and accurate; it is sensitive to the training set. In particular, if faces have various poses, a non-face region can be detected erroneously. To overcome this problem, the angle of rotation was estimated and the region was rotated inversely using only front and side faces. The proposed algorithm first detects the front and side candidate regions, and the remaining candidate areas were fitted to the ellipses to estimate the orientation. After inversely rotating the slated face region, Viola et al.'s method was applied to determine the final face.

3.1 Rotation-Invariant Face Detection Method

In this study, the face region was detected using robust features trained by the front and side faces. An ellipse is determined uniquely by the centre of mass $(\bar{x}, \bar{y})$, orientation θ, length of the major axis α, and length of the minor axis β. The centre of mass was derived as follows:

$$(\bar{x}, \bar{y}) = \left(\frac{1}{A}\sum_{(x,y)\in C} x, \frac{1}{A}\sum_{(x,y)\in C} y\right), \tag{4}$$

where A represents the area of the face region. The orientation θ was derived using the centre of mass as follows:

$$\theta = \frac{1}{2}\arctan\left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}\right), \tag{5}$$

where $\mu_{p,q}$ is defined as:

$$\mu_{p,q} = \sum_{(x,y)\in C} (x-\bar{x})^p (y-\bar{y})^q. \tag{6}$$

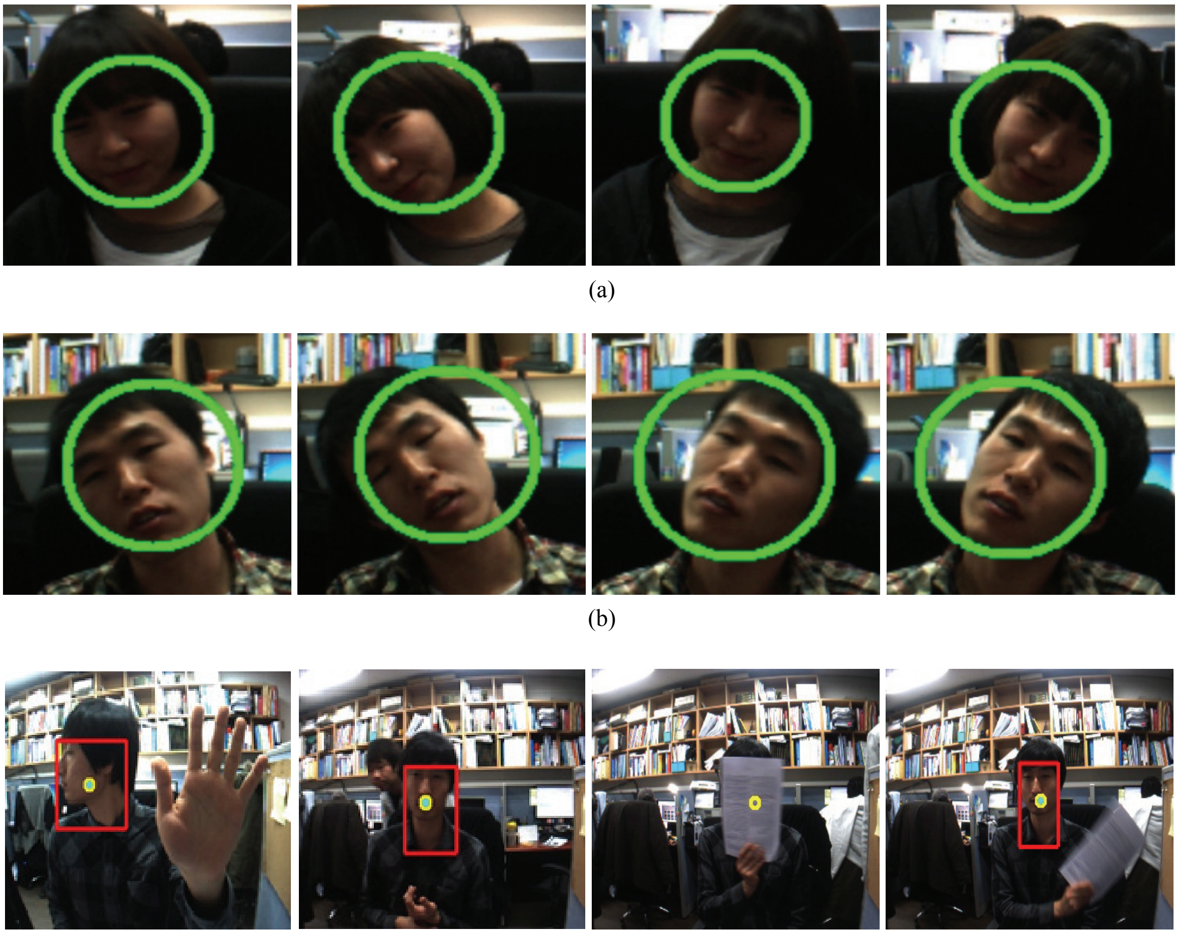
Finally, α and β were respectively calculated using

$$\alpha = \left(\frac{4}{\pi}\right)^{1/4}\left[\frac{I_{\min}^3}{I_{\max}}\right]^{1/8}, \text{ and } \beta = \left(\frac{4}{\pi}\right)^{1/4}\left[\frac{I_{\max}^3}{I_{\min}}\right]^{1/8}, \tag{7}$$

where $I_{\min}$ and $I_{\max}$ are defined as:

$$\begin{aligned} I_{\min} &= \sum_{(x,y)\in C} [(x-\bar{x})\cos\theta - (y-\bar{y})\sin\theta]^2, \text{ and} \\ I_{\max} &= \sum_{(x,y)\in C} [(x-\bar{x})\sin\theta - (y-\bar{y})\cos\theta]^2. \end{aligned} \tag{8}$$

Given the above parameters by ellipse fitting, the rotation matrix was generated using $(\bar{x}, \bar{y})$ with the rotation angle θ. The inversely rotated face was then



(a)

(b)

(c)

Fig. 2. Results of the proposed algorithm; (a) and (b) results of detecting difference faces and angles [15˚, 45˚], and (c) result of tracking an occluded face.

estimated by a geometric transformation using the rotation angle θ. The inversely rotated candidate region was used to detect the correctly rotated faces using Viola et al.'s method.

3.2 Face tracking with Occlusion Handling

This subsection presents a robust face tracking algorithm. Although the face region can be detected accurately in a single frame, the detection might fail in another frame if the face region is partially occluded by an object. If the occlusion occurs in a specific frame, the motion vector of the center of the face regions for robust tracking can be predicted. The face center is then calculated by averaging three points as follows:

$$\left(\frac{x_1+x_2+x_3}{3},\frac{y_1+y_2+y_3}{3}\right). \tag{9}$$

The difference between the current and previous frames is defined as the motion vector, F_{move}, which can be used as a prediction parameter in the event of an occlusion. In other words, if a face region disappears by occlusion, the face location can be predicted by adding F_{move} to the center of the previous frame. The proposed tracking algorithm is based on the linear motion model of the face.

4. Experimental Results

640x480 indoor video sequences were used to evaluate the performance of the proposed algorithm. The test images included both low light level conditions and a complicated background. To detect the face candidate regions, the skin regions were segmented using the skin tone color and depth-based histogram. Fig. 2 shows the results of the proposed algorithm.

For two different faces with four orientations in the range [15˚, 45˚], the proposed algorithm provides correctly detected face regions and the correct trajectories of the centre of the region as a result of the motion-based face tracking algorithm. The current location of the object trajectory provides a migration path for a motion estimation to predict the amount of overlapping in two objects.

The proposed face detection method was compared with Viola et al.'s and LAB color space-based methods.

The experiment used 31 images in the Caltech database. Viola et al.'s face detection method provided a success and error rate of approximately 93.55% and 19.35%, respectively, whereas the LAB color space-based method gave a significantly lower detection rate. The proposed method gave the same performance as Viola et al.'s method for upright faces.

Table 1. Comparison of the face detection rates for 31 upright face images using Viola et al.'s, LAB color space-based, and the proposed methods.

Face detection method	LAB-based face detection method (DETECTION/NUMBER OF IMAGES)	Viola's method [2] (DETECTION/NUMBER OF IMAGES)	PROPOSED ALGORITHM (DETECTION/NUMBER OF IMAGES)
Detected rate (success / number of image)	16/31	29/31	29/31
Error rate (fail / number of image)	19/31	6/31	4/31

Table 1 compares the face detection results of the proposed method with Viola et al.'s method for variously rotated faces. As shown in Table 2, the success rates of Viola et al.'s method became significantly low as the face was rotated by 15˚ or more, whereas the proposed method maintained high success rates regardless of the amount of rotation up to 45˚.

Table 2. Comparison of the success rates of rotated faces up to 45˚.

The rotated angles	Viola's method[2]	Proposed method
0˚	94%	94%
5˚	94%	94%
10˚	90%	94%
15˚	59%	94%
20˚	59%	90%
25˚	6%	94%
30˚	0%	94%
35˚	0%	94%
40˚	16%	90%
45˚	6%	94%

5. Conclusion

A novel face detection and tracking algorithm for solving both rotation and occlusion problems is proposed. The depth histogram of candidate regions generated by a stereo camera for solving the occlusion problem was used. Moreover, ellipse fitting was performed in the detected candidate regions to estimate the angle of rotation. Based on the extended set of experimental results, the proposed algorithm could successfully detect rotated faces, and was robust to rotation, low light level conditions and complicated backgrounds. In future research, a color correction will be incorporated to improve the detection rate in a range of environments.

References

[1] Z. Peng, Y. Zhou, and P. Wang, "Multi-Pose Face Detection Based on Adaptive Skin Color and Structure Model," *Proc. Int. conf. Computational Intelligence and Security*, pp. 325-329, Dec. 2009. Article (CrossRef Link)

[2] P. Viola, and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Feature," *Proc. Int. Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 511-518, 2001. Article (CrossRef Link)

[3] R. Lienhart and J. Maydt, "An Extended Set of Haar-like Features for Rapid Object Detection," *Proc. Int. Conf. International Conference on Image Processing*, vol. 23, pp. 900~903, 2002. Article (CrossRef Link)

[4] R. Hsu, M. Abdel, and A. Jain, "Face Detection in Color Images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 696-706, May 2002. Article (CrossRef Link)

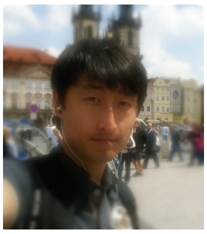
[5] U. Yang, B. Kim, K. Toh, and K. Sohn, "Illumination-Invariant Color Space and Its Application to Skin-Color Detection," *Optical Engineering*, vol. 49, no. 10, 107004-1-10, 2010. Article (CrossRef Link)

[6] M. Jing, and L. Chen, "Novel Face-Detection Method under Various Environments," *Optical Engineering*, vol. 48, no. 6, 067202-1-12, 2009. Article (CrossRef Link)

[7] A Cheddad, J. Condell, K. Curran, and P. Mckevitt, "A New Colour Space for Skin Tone Detection," *Proc. Int. Conf. Image Processing*, pp. 497-500, Nov. 2009. Article (CrossRef Link)

[8] D. Chi, K. Ngan, "Face Segmentation Using Skin-Color Map in Videophone Applications," *IEEE Trans. Circuits, Systems for Video Technology*, vol.9, no.4, pp. 551-564, Jun. 1999. Article (CrossRef Link)

Daehee Kim was born in Buchun, Korea in 1981. He received his B.S. degree in computer engineering from Kangnam University, Korea, in 2005. He received his M.S. degrees in image engineering from Chung-Ang University, Korea, in 2008. Currently, he is pursuing a Ph. D degree in image engineering at Chung-Ang University. His research interests include face detecting, gait recognition, object tracking, object segmentation and gesture recognition.



Seungwon Lee was born in Seoul, Korea in 1983. He received his B.S. degree in Electronic engineering from Seoul National University of Science and Technology, Korea, in 2009. He received his M.S. degree in in image engineering from Chung-Ang University, Seoul, Korea, in 2007. Currently, he is pursuing a Ph. D degree in image engineering at Chung-Ang University. His research interests include object tracking, computational camera based depth estimation, intelligent surveillance system, video processing and human recognition.

Dongmin Kim was born in Taegu, Korea in 1972. He received his B.S. degree in computer engineering from Hannam University, Korea, in 2000. Currently, he is working for the Forensic Science Division, Supreme Prosecutor's Office, Korea, and is pursuing a M.S. degree in image engineering at Chung-Ang University. His research interests include image restoration, enhancement, analysis and filtering.