Face Detection using Half-Face Templates

Yi Zhu and Florin Cutu
Computer Science Dept.,
Indiana University, USA
Email: {yiz,florin}@Indiana.edu

Abstract

Face detection is the first important step in many face image processing applications. Although a lot of work has been done on detecting frontal faces much less effort has been put into detecting faces with large image-plane or depth rotations. Most templates used in face detection are whole-face templates. However, such templates are ineffective for faces significantly rotated in depth. We propose to use half-face templates to detect faces with large depth rotations. Our experimental results show that half-face templates significantly outperform whole-face templates in detecting faces having large out-plane rotations and performs as well as whole-face templates in detecting frontal faces.

1. Introduction

Accurately detecting and locating faces in an image is the first step of face image processing applications, such as face recognition and facial expression recognition. It is difficult to detect and locate faces in general background because of insufficient a priori information and the variability of face appearance.

Current face detection methods can be categorized into image-based methods and feature-based methods. Image-based methods use the intensity array as template to matching standardized candidate regions [1][2][3]. All the upright rectangular regions are exhaustively matched with the template after standardizing and neural networks are used to find the best match. Exhaustive search makes image-based methods very time-consuming and it is hard to generalize them to detect faces having large image plane or depth rotations. Feature-based methods first detect facial features using skin color, the specific paterrs of various facial features, and their relative positions. Candidate regions are limited to darker regions and high gradient regions. These methods are not only faster but also can adapt to large in-plane rotation. K.C. Yow and R. Cipolla [7] found facial features by edge detection and then verified candidate regions by the position relations of facial features. C. Lin and K.C. Fan [3] proposed to select candidate pixels by their intensity. Gray scale images are first transformed into binary images by thresholding and then 4-connected components are found as candidates of facial features. They used the ternary spatial relations of two eyes and mouth or eye, mouth and ear to define the candidate face region. The candidate regions are further tested using an evaluation function similar to a neural network.

Although high detection rates have been achieved when detecting frontal faces, the problem of detecting faces having large image-plane and depth rotations has not been satisfactorily addressed. Typically, when a face has a large depth rotation, roughly half of the face is occluded. Even for a frontal face image, if the image is captured even at a slight angle the left half and right half of the face will not by symmetric, affecting the correlation with the template. Thus, a whole-face template will be less effective in detecting faces having significant depth rotations. In this paper we propose using segmentation to detect faces having large image plane rotations and using half-face template to match faces having large depth rotations. The experimental results show that the half-face template is as efficient as whole-face template when detecting frontal faces, and is much more efficient in detecting faces having large depth rotations. Therefore, using half-face templates appears to be a promising method for detecting faces having large depth rotations.

The outline of the proposed method is introduced in Section 2, and in Section 3 we describe it in detail. Experimental results and discussions are given in Section 4, and Section 5 contains the conclusions.

2. Outline of the algorithm

The flowchart of the method is shown in Figure 1. First, the location, orientation and size of the detected face are approximately estimated according to location, orientation, and size of the left eye or right eye. Then, the correlation coefficients of the left half-template and the right half-template with the tested area are calculated. The larger of the two is selected as the final correlation coefficient of the half template with the tested area. Finally, a threshold is used to reject non-face regions in the entire image.

3. Details of the algorithm

3.1 Finding candidate face regions

The process of estimating candidate face regions includes three steps:

Step 1. Transforming the image into binary image by intensity thresholding. In the binary image pixels whose intensities are less than a threshold T are 1 while the other pixels are 0. Since images are captured under different lighting conditions it is not possible to use a unique threshold for all images. Here we use an iterative procedure, i.e., we iteratively increase the threshold T from a *min* value to a *max* value by step ?*t*. For each value of the threshold we repeat steps 2 and 3 below.

Step 2. Find 4-connected components in the thresholded (binary) image and eliminate the components whose sizes are less than 15 pixels from further consideration since the dark area of an eye can not be very small. The remaining components are left or right eye candidates. Obviously, many such regions will not correspond to eyes, and they will be eliminated at Step 3.

Step 3. Find candidate face regions according to the sizes and locations of the 4-connected eye components, as follows. The location $(\bar{x}, \bar{y})$, the orientation *?*, and the major and minor semiaxes *a* and *b* of the ellipse enclosing an image region can be determined from the first and second-order image moments $\bar{x}$, $\bar{y}$, μ of the region, as follows [13]:

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right)$$

$$(a,b) = \sqrt{\frac{2\left(\mu_{20} + \mu_{02} \pm \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\right)}{\mu_{00}^2}}$$

where

$$\mu_{pq} = \sum_{(x,y)} (x-\bar{x})^p (y-\bar{y})^q f(x,y), \bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$$

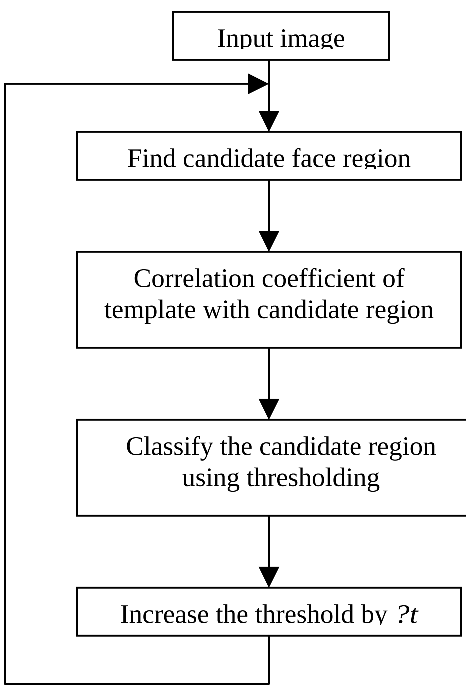


Figure 1: Flowchart of the face-detection algorithm

Elliptical regions with the major axis larger than a 1/3 of the image size are eliminated since the eye regions are unlikely to be very large.
Since in non-frontal face views only part of the face is visible, we will match half-face templates to the candidate regions in order to detect faces having significant depth-rotations.

We use the parameters ($\bar{x}$, $\bar{y}$, *a*,*b*,*?*) of the ellipses best approximating the candidate eye regions to compute the bounding boxes enclosing the candidate half-faces from which the eye regions were extracted. The half-face bounding box is aligned with the major and minor axes of the ellipse, and has dimensions proportional to the semiaxes of the eye ellipse, as indicated in Figure 2.

The parameters (position, length, width, orientation) of the estimated half-face rectangles corresponding to left eye and right eye ellipses are shown in Figure 2. The process of estimating the candidate half-face region is shown in Figure 3. First, the image is thresholded (Figure 3a) and the candidate eye regions are found (Figure 3b). Then, the candidate half-face regions corresponding to the candidate eye regions are computed (Figure 3c). Because it is very difficult to differentiate left and right eyes without the context of the original face each component is regarded as both a left eye and right eye. Therefore, two candidate half-face regions are computed for each eye region.

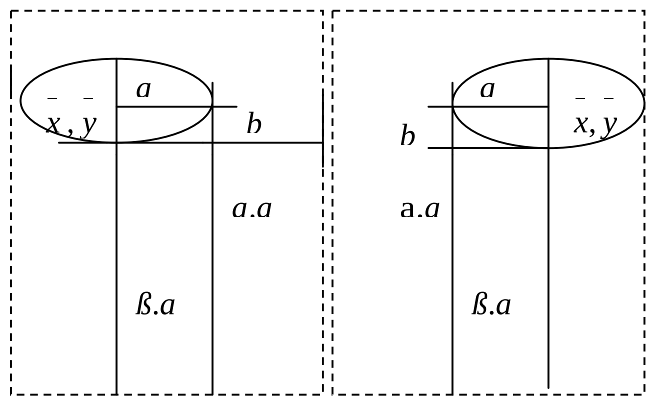
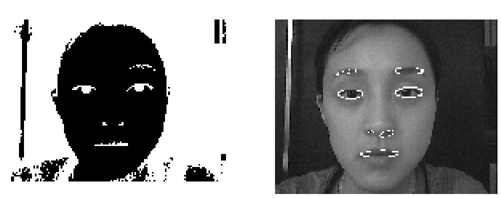
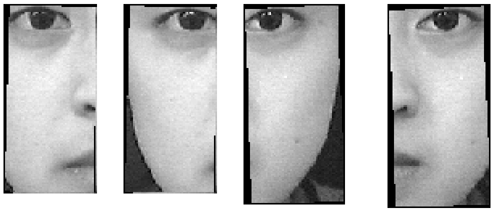


*Figure 2: Candidate half -face region corresponding to left and right eye ellipses. Based on an ellipse fitted to the candidate eye regions the bounding box of the half face is calculated. The length of the half-face is proportional to the minor ellipse semiaxis (alfa*a) and the width of the half-face is proportional to the major ellipse semi-axis.*



(a) *(b)*



(c)

Figure 3: Half-face regions found using candidate eye regions.

3.2 Constructing face templates

We used the ORL face database to build half-face templates. We used faces having large image contrast and neutral expressions. All selected face images were aligned by the upper bound and corners of the eyes.

The ORL image-set mean left half-face and mean right half-face were stored as, respectively, left half-face and right half-face templates. The two half-face templates are shown in Figure 4.

To compare the effectiveness of the half-face template approach to conventional whole-face template matching we used the ORL face database to construct a whole-face template (shown in Figure 4) using the same approach.

(a) *(b)* *(c)*

Figure 4: Face Templates (a) left half-face template (b) right half-face template (c) whole-face template.

4. Experimental results

In several computational experiments we compared the correlation coefficients of the half-face template and of the whole-face template with the detected face regions of faces of various orientations. Our results show (Figure 6) that 99% of the correlation coefficients of the half-template with detected faces regions are larger than those of the whole-face template. The larger the depth rotation the more significant the advantage of the half-face template over the whole-face template becomes. As illustrated in Figure 5, with increasing depth rotation, the correlation coefficient of the half-face template with the detected face region becomes progressively larger than the correlation coefficient of whole-face template with the detected face region.

In order to compare the correlation of the face templates with the detected face region and the correlation of the face template with the non-face regions 6800 landscape images are searched from the internet (see Figure 7). The searched landscape images are divided into subimages of the same size as the face templates. The results are shown that the correlation coefficients of the face templates with the detected faces regions are significantly larger than the correlation coefficients of the face templates with the landscape images (see Figure 8). Thus demonstrates that the face templates are more correlated to the detected face regions.

5. Conclusions

From the experiment results we can see that the half face template is very effective in detecting faces having large depth rotations. It is a prominent method to solve the problem of detecting faces having large depth rotations. In order to sufficiently use the information provided by the images we use a large template instead of small templates. If the tested image is larger than the template it is scale down. Otherwise the template is scaled down to accommodate the tested image. The experiment results show it is better than using small template.

References

[1] H.A. Rowley, S. Baluja, and T. Kanade, “Neural network-based face detection,” IEEE Trans. on Pattern Analysis and Machine Intelligence, 20(1):23--28, 1998. 3

[2] K. Sung and T. Poggio, “Example-based Learning for View-based Human Face Detection,” IEEE Trans. Pattern Recognition and Machine Intelligence, vol. 20, no. 1, Jan. 1998, pp.39-51

[3] C. Lin and K.C. Fan, “Triangle-based approach to the detection of human face,” Pattern Recognition, 34(2001), pp.1271-1284

[4] K.W. Wong, K.M. lam and W.C. Siu, “An efficient algorithm for human face detection and facial feature extraction under different conditions,” Pattern Recognition, 34(2001), pp.1993-2004

[5] E. Osuna, R. Freund, and F. Girosi., “Training support vector machines: An application to face detection,” Proc. of CVPR’97, 1997 pp. 130-13

[6] A. Colmenzrez and T. Huang, “Face detection with information based maximum discrimination,” Proc. IEEE Int'lConf. on Comput. Vis. and Patt. Recog., pages 782--787, 1997.

[7] K.C. Yow and R. Cipolla, “Feature-based human face detection,” Image and Vision Computing, vol. 15, no. 9, pp. 713--735, 1999.

[8] M.H. Yang, D. Roth, and N. Ahuja, “A SNoW-based face detector,” In S. A. Solla, T. K. Leen, and K.-R. Muller editors, Advances in Neural Information Processing Systems 12, pp. 855-861. MIT Press, 2000.

[9] M.H. Yang, N. Ahuja, and D. Kriegman, "Face detection using a mixture of factor analyzers," Proc.

IEEE Int. Conf. on Image Processing, vol. 3, pp. 612-616, 1999.
[10] R. Feraud, O. J. Bernier, J.-E. Viallet, and M. Collobert, "A fast and accurate face detector based on neural networks," IEEE Trans. On Pattern Analysis and Machine Intelligence, 23(1):42--53, January 2001.
[11] S.B. Yacoub, B. Fasel, and J. Luettin, "Fast Face Detection using MLP and FFT," Proc. of Second Int. Conf. Audio and Video-based Biometric Person Authentication, pp. 31-36, 1999.
[12] "On invariants of image," Lecture Notes

[13] S. O. Belkasim, M. Shridhar and M. Ahmadi, "Pattern recognition with moment invariants: a comparative study and new results," Pattern Recognition, vol.24, No.12, 1991, pp.1117-1138

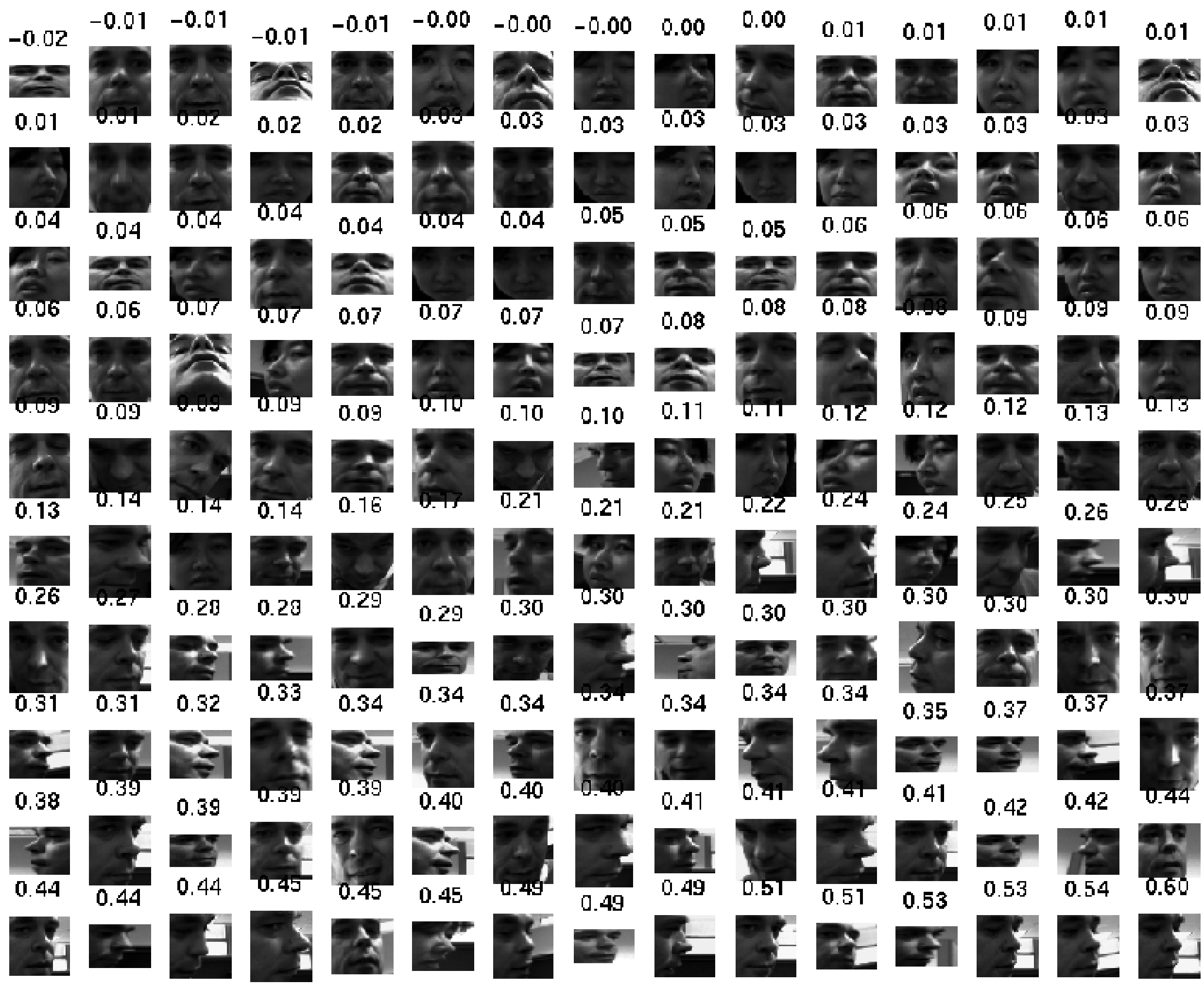


Figure 5: Difference of correlation coefficients of half face template with the detected face regions and those of whole face template with the detected face regions

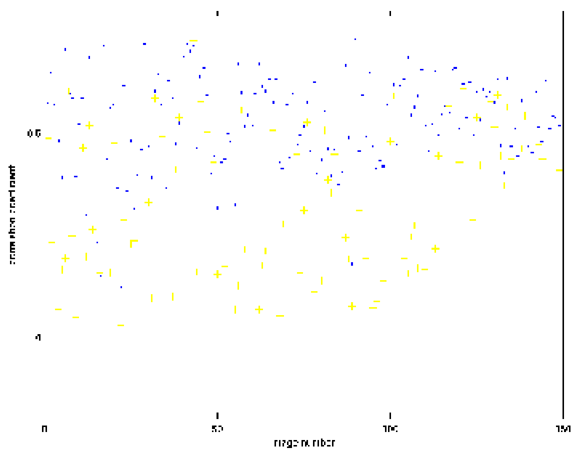


Figure 6: Correlation coefficients of half face template with detected face region and those of the whole face template with the detected face regions

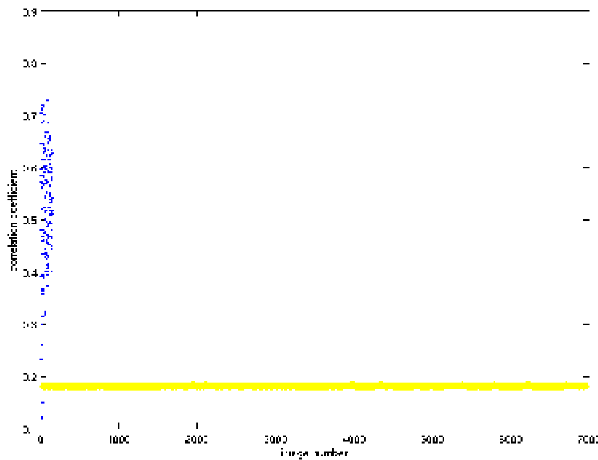


Figure 8: Correlation coefficients of half face template with the detected face regions and those the half face template with the backgrounds

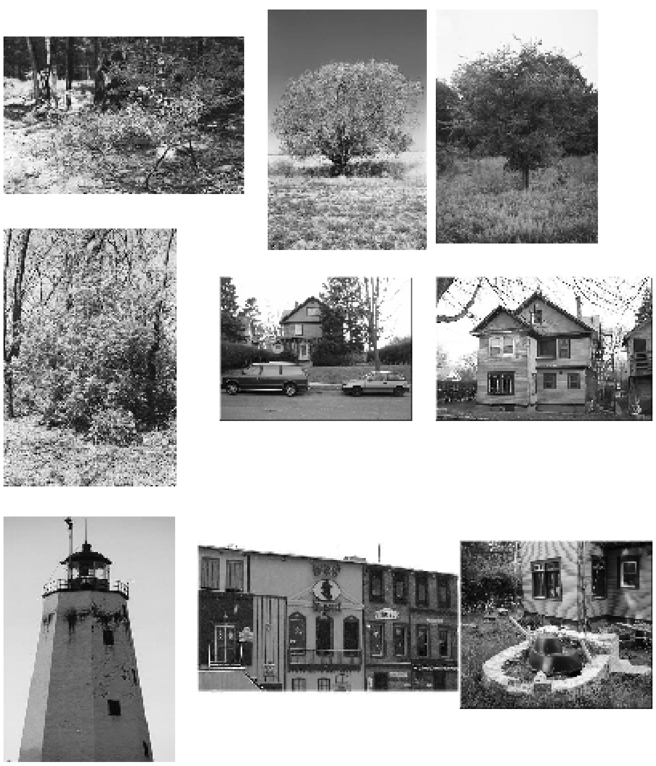


Figure 7: Sample landscape images used in comparing the correlation of the face templates with the detected face regions and the correlation of the face templates with the non-face regions