MULTIPLE-RESOLUTION EDGE-BASED FEATURE REPRESENTATIONS FOR ROBUST FACE SEGMENTATION AND VERIFICATION

Yasufumi Suzuki and Tadashi Shibata

Department of Frontier Informatics, School of Frontier Sciences, The University of Tokyo
701 Kiban Bldg. 5-1-5 Kashiwanoha, Kashiwa, Chiba, 277-8561, Japan
yasufumi@else.k.u-tokyo.ac.jp, shibata@ee.t.u-tokyo.ac.jp

ABSTRACT

Robust face segmentation and verification algorithms have been developed employing edge-based feature representations in the multiple resolution scheme. The face segmentation algorithm in our previous work [1] has shown the robustness against illumination, focus and scale variations. In addition to these features, a rotation-invariant scheme has been introduced in the present work. As a result, a false-negative-free face segmentation robust against circumstance variations has been developed. In order to eliminate false positives occurring in the segmentation stage, a face verification algorithm has been also developed, in which edge information in a finer resolution is utilized to confirm the existence of facial parts. Since multiple facial parts are utilized as clues for verification, the system is capable of detecting partially occluded faces. As a result, a robust face detection algorithm has been successfully developed in which the occurrence of both false positives and false negatives has been drastically reduced.

1. INTRODUCTION

The development of robust image recognition systems is quite essential in a variety of applications such as intelligent human-computer interfaces, robot visions, security systems, and so forth. Real-time face recognition, in particular, plays an important role in establishing user-friendly interfaces between humans and computers. An automatic face recognition system consists of two stages, the face detection and the face identification [2]. In the face detection, facial images must be correctly localized in an input image without prior information about illumination, scales, numbers of faces, and so on. A lot of face detection algorithms, using principal component analysis [3] and neural networks [4], for instance, have been developed. In these algorithms, however, a large amount of numerical computation is required, making the processing extremely time-consuming. It seems not feasible to build real-time-responding low-power embedded systems by solely relying upon software programs running on general-purpose computers. In this regard, the development of hardware-friendly algorithms compatible to dedicated VLSI chips is quite essential.

For this purpose, we have developed search engine VLSI chips in both CMOS digital [5] and analog [6] technologies. The projected principal-edge distribution (PPED) [7] and cell edge distribution [8] have been developed as image representation schemes compatible to the vector-generation VLSI engine we developed [9]. These algorithms have been applied successfully to hand-written pattern recognition, medical X-ray analysis [10][11], and face segmentation [1][8].

In our previous work [1], a false-negative-free face segmentation system (never missing true faces while allowing some false positives) and its robust nature against variations in illumination, focus, and scales have been demonstrated. Although the segmentation system detected faces very robustly, tilted faces could not be detected. In addition, a number of false positives were left after segmentation.

The purpose of this paper is firstly to make the system robust against facial image rotation. Secondly, a face verification algorithm is developed by introducing the concept of multiple resolution in the edge-based feature representations. Finer edge information is utilized to validate facial parts for face candidates selected in the segmentation stage. Multiple clues derived from the results of confirmation of each facial part are employed to reject false positives. Using such multiple clues has made it possible to recognize partially occluded faces. The enhanced robustness of the face detection algorithm developed in this work has been demonstrated for a number of sample images.

2. FACE SEGMENTATION ALGORITHM

2.1 Edge-Based Feature Maps

Edge-based feature maps are employed as the bases of our image representation algorithm [7]. The feature maps represent the distribution of four-directional edges extracted from a 64×64-pixel input image as illustrated in Fig. 1. The input image is firstly subjected to pixel-by-pixel spatial filtering operations using kernels of 5×5-pixel size to extract edges in four directions, i.e. horizontal, +45 degree, vertical, or -45 degree. The threshold for edge detection is determined taking the local variance of luminance data into account. Namely, the median of the 40 values of neighboring pixel intensity differences in a 5×5-pixel kernel is adopted as the threshold. This is quite important to retain all essential features in an input image in the feature maps. All the pertinent parameters have been optimized in the medical radiograph recognition as a test vehicle. But the robust nature has been proven in more general applications [7]. Therefore, the parameters determind in Ref. [7] were utilized as they are in the present work. Fig. 1 shows an example of feature maps generated from the same person under different illumination conditions. The edge information is very well extracted from both bright and dark images and this has enabled illumination-invariant detection.

2.2 Feature Representations

In this work, the two general-purpose feature vectors developed in Ref. [8] were utilized for face segmentation. The Eyes-and-Mouth extraction scheme was not employed because it is not easy to accommodate to image rotation. Generation scheme of these two feature vectors from the same

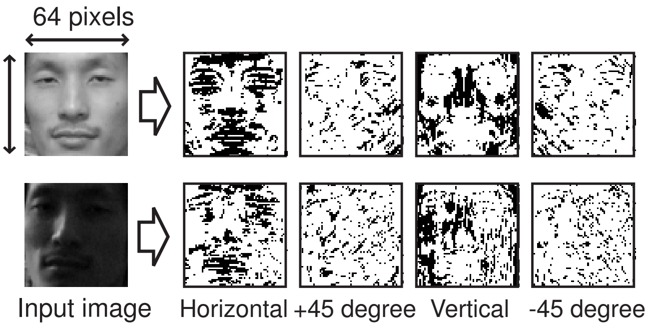


Fig. 1. Feature maps generated from bright and dark face images.

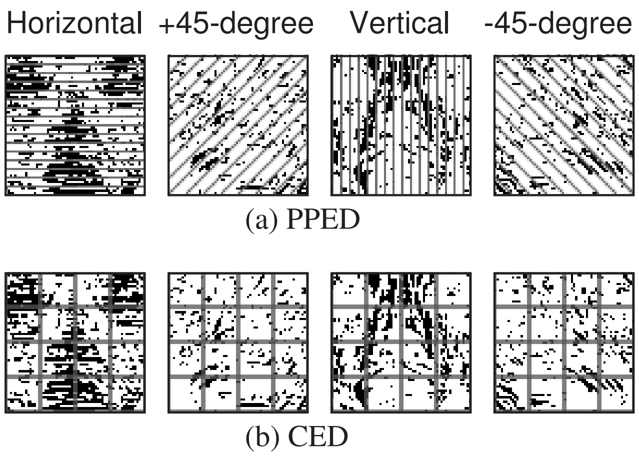


Fig. 2. Partitions of feature maps for vector generation based on projected principal-edge distribution (PPED) (a) and cell edge distribution (CED) (b).

set of feature maps is briefly reviewed in the following.

Each feature map is divided into 16 blocks as illustrated in Fig. 2, thus 64 blocks are generated from four directional edge maps. In the projected principal-edge distribution (PPED) [7], feature maps are divided in the same direction of the edge as shown in Fig. 2 (a). In the horizontal edge map, for example, it is cut in every four rows. Fig. 2 (b) illustrates another feature vector called the cell edge distribution (CED) [8], which divides feature maps into square cells. Each of the four feature maps is cut into 4×4 cells. Namely, each cell contains 16×16 pixels.

Each element in the 64-dimension feature vector represents the number of edges within the corresponding block in the feature maps. During the counting of edge flags within each block, the spatial resolution of the original image is moderately reduced while retaining the essential features in the image as a spatial distribution histogram. This is the origin of the robustness of edge-based feature representations in template matching.

2.3 Face Segmentation by Template Matching

Face segmentation is carried out as follows. A 64×64-pixel area is taken from an input image and feature vectors are generated. Then, the feature vectors are matched with all

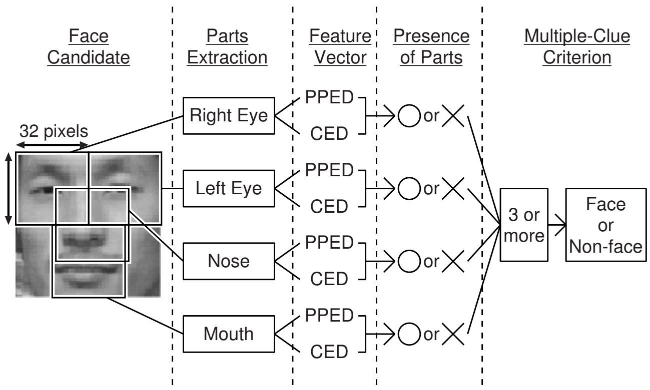


Fig. 3. Face verification using finer edge information of eyes, nose and mouth.

template vectors of face samples and non-face samples stored in the system and classified as a face or a non-face according to the category of the best-matched template vectors. The matching is carried out using the Manhattan distance as the dissimilarity measure.

As the criterion in the face segmentation, the multiple-clue method [8] is utilized. Template matching is performed using both PPED and CED vectors. If a local image is classified as a face by both feature vectors, then it is adopted as a face candidate. This classification is carried out by pixel-by-pixel scanning of the 64×64-pixel window over the entire image. The same operation is repeated for the input image whose size is enlarged or reduced by a factor 4, 2, 1/2, 1/4, or 1/8 in order to accommodate scale variations.

3. ROTATION-INVARIANT SEGMENTATION

In order to accommodate our system to rotated facial images, tilted face images are included in the set of face template vectors. Due to the symmetry of vector generation algorithm for 90-degree rotations in PPED and CED vectors, the feature vector for 90, 180, and 270 degree-rotated images can be generated by just swapping elements in the original vector. Therefore, only optional samples we need to prepare to cover the entire 360 degrees rotation are tilted faces within the angles between -45 and $+45$ degrees. In this work, 0, ± 18, and ± 36-degree tilted face images are included as face templates for rotation invariant face segmentation.

4. FACE VERIFICATION ALGORITHM

The face verification algorithm employing the feature vectors generated from local blocks with finer resolutions has been developed. In the verification algorithm, the presence of facial parts is confirmed using finer edge information. If facial parts cannot be found, the candidate will be discarded. Among four directional feature maps, as shown in Fig. 1, a lot of edge flags are detected around the areas of eyes and a mouth in the horizontal edge map and the area of a nose in the vertical edge map. In this work, these four facial parts are utilized in the face verification.

The procedure of the verification is as follows. Four 32×32-pixel windows are taken from the left-top, right-top, center, and bottom of the candidate window as illustrated in Fig. 3. They correspond to the locations of the right eye, the

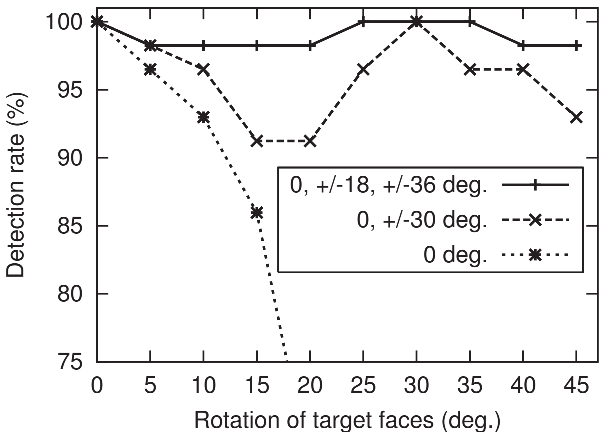


Fig. 4. Detection rate from rotated target images with different combinations of tilted template vectors.

(a) (b)

Fig. 5. Results of rotation-invariant face segmentation using image in Ref. [13]. Face verification was not conducted in this experiment.

left eye, the nose and the mouth, respectively. Each 32×32-pixel window is converted to the PPED and CED feature vectors. In the face verification, finer edge information is utilized as compared to the original PPED and CED vectors employed in the segmentation. Namely, each element of the PPED vector represents the number of edges within two columns or rows instead of four columns or rows in the face segmentation. And each element of the CED vector indicates the number of edges within the 8×8-pixel cell instead of 16×16-pixel cell. Template matching is performed using the eight feature vectors generated from four facial parts with two options of PPED and CED representations. The presence of each facial part is determined when both PPED and CED vectors match the respective templates. If three out of four parts are confirmed, the candidate is determined as a true face. Since we require only three parts as clues for verification, the system is capable of detecting partially occluded faces.

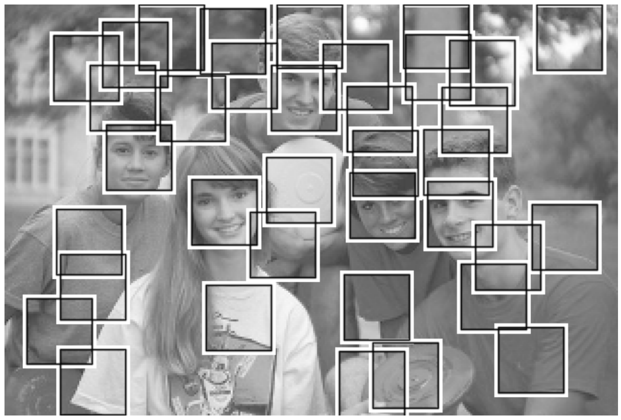


Fig. 6. Face segmentation result using picture in Ref. [4].

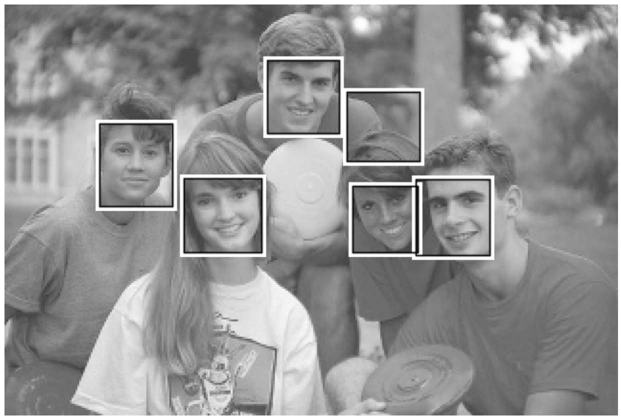


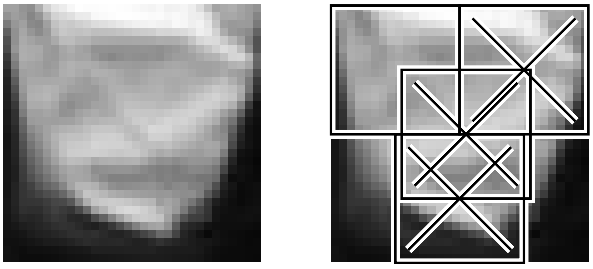
Fig. 7. Result of face verification using result of segmentation in Fig. 6.

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this work, 900 face images and 2000 non-face images are utilized as templates. Template images of faces are all frontal faces of 300 people [12]. It should be noted that the face samples are all from Japanese people. As scale variations of face templates, 80% and 60% sized face images of them are used in additional to the original images. Template images of non-faces were chosen randomly from the background scenery of pictures which were not used in the test.

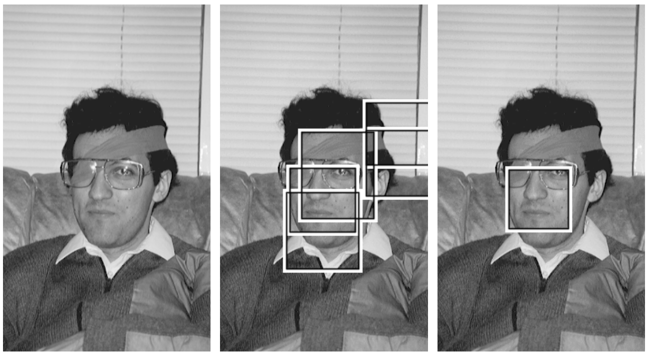
Segmentation of rotated faces between -45 and $+45$ degrees was carried out using three sets of tilted template vectors. The results are presented in Fig. 4. The results obtained with the template set employed in our system show the detection rate over 95% all over the range. Using only upright templates, namely only 0-degree template vectors, the detection rate falls rapidly at 10 degrees of tilting. Although the set of 0 and ± 30-degree template vectors covers the range, the detection rate fluctuates depending on the tilt angle.

Fig. 5 shows the results of face segmentation using the images that contains different size and angle of faces in Ref. [13]. Some false positives were eliminated by density selection criterion [1] which retains only high density clusters of face candidates. As a result, all faces are detected correctly with some false positives. Fig. 6 demonstrates the result of face segmentation using a picture in CMU test-



(a) (b)

Fig. 8. An example of false positives in face segmentation (a); this was rejected in verification since only a right eye passed facial-parts validation (b).



(a) (b) (c)

Fig. 9. Results of face segmentation (b) and verification (c) from original face image (a) whose right eye is occluded.

set [4]. All faces in the picture detected, however, a lot of false positives are also detected. These detected faces were screened using the face verification algorithm and the results are demonstrated in Fig. 7. All false positives except one were rejected by the verification without the occurrence of false negatives.

Fig. 8 (a) shows an example of false positives in our previous work [8]. This false positive were rejected by the verification algorithm since a left eye, a nose and a mouth failed to be detected as illustrated in Fig. 8 (b).

Face verification of a partially occluded face is demonstrated in Fig. 9. All false positives in Fig. 9 (b) were eliminated by the face verification and the only true face was remained. In this case, the true face was certified although the right eye is covered since the algorithm confirmed all three parts except the right eye.

6. CONCLUSION

Robust face segmentation and verification algorithms based on the edge-based feature representations have been developed. The rotation-invariant scheme has been introduced to the face segmentation by adding tilted face samples to the templates. The face verification is performed by validating facial parts within each face candidate. The multiple-resolution feature representations have been employed for the confirmation of the presence of facial parts. As a result, the occurrence of both false negatives and false positives has been greatly reduced. And the robustness of the algorithm has been demonstrated for sample images.

REFERENCES

[1] Y. Suzuki and T. Shibata, "An edge-based face detection algorithm robust against illumination, focus, and scale variations," in *Proc. EUSIPCO 2004*, Vienna, Austria, September 6-10. 2004, pp. 2279–2282.

[2] T. Kondo and H. Yan, "Automatic human face detection and recognition under non-uniform illumination," *Pattern Recognition*, vol. 32, pp. 1707–1718, 1999.

[3] J. S. Jang, K. H. Han, and J. H. Kim, "Evolutionary algorithm-based face verification," *Pattern Recognition Letters*, vol. 25, pp. 1857–1865, 2004.

[4] H. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.20, pp. 23–38, January 1998.

[5] M. Ogawa, K. Ito, and T. Shibata, "A general-purpose vector-quantization processor employing two-dimensional bit-propagating winner-take-all," in *IEEE Symp. on VLSI Circuits Dig. Tech. Papers*, Honolulu, United States, June 11-15. 2002, pp. 244–247.

[6] T. Yamasaki and T. Shibata, "Analog soft-pattern-matching classifier using floating-gate MOS technology," *IEEE Trans. Neural Networks*, vol. 14, pp. 1257–1265, September 2003.

[7] M. Yagi and T. Shibata, "An image representation algorithm compatible with neural-associative-processor-based hardware recognition systems," *IEEE Trans. Neural Networks*, vol. 14, pp. 1144–1161, September 2003.

[8] Y. Suzuki and T. Shibata, "Multiple-clue face detection algorithm using edge-based feature vectors," in *Proc. ICASSP 2004*, Montreal, Canada, May 17-21. 2004, vol. 5, pp. 737–740.

[9] M. Yagi, H. Yamasaki, and T. Shibata, "A mixed-signal VLSI for real-time generation of edge-based image vectors," in *Proc. NIPS 2003*, Vancouver, Canada, December 8-13. 2003, pp. 1035-1042.

[10] M. Yagi, M. Adachi, and T. Shibata, "A hardware-friendly soft-computing algorithm for image recognition," in *Proc. EUSIPCO 2000*, Tampere, Finland, September 5-8. 2000, pp. 729–732.

[11] M. Yagi and T. Shibata, "Human-perception-like image recognition system based on the associative processor architecture," in *Proc. EUSIPCO 2002*, Toulouse, France, September 3-6. 2002, pp. 103–106.

[12] Softpia Japan, Research and Development Division, HOIP Laboratory, "HOIP facial image database," http://www.hoip.jp/web_catalog/top.html

[13] H. Rowley, S. Baluja, and T. Kanade, "Rotation invariant neural network-based face detection", in *Proc. Computer Vision and Pattern Recognition 1998*, Santa Barbara, United States, June 23-25. 1998, pp. 38–44.