

Facial Features for Template Matching Based Face Recognition

¹Chai Tong Yuen, ²M. Rizon, ¹Woo San San and ¹Tan Ching Seong

¹Department of Mechatronic and Biomedical Engineering, University Tunku Abdul Rahman,
Jalan Genting Kelang, Setapak 53300, Kuala Lumpur, Malaysia

²Department of Electrical Engineering, King Saud University, P.O. Box 800,
Riyadh 11421, Kingdom of Saudi Arabia

Abstract: Problem statement: Template matching had been a conventional method for object detection especially facial features detection at the early stage of face recognition research. The appearance of moustache and beard had affected the performance of features detection and face recognition system since ages ago. **Approach:** The proposed algorithm aimed to reduce the effect of beard and moustache for facial features detection and introduce facial features based template matching as the classification method. An automated algorithm for face recognition system based on detected facial features, iris and mouth had been developed. First, the face region was located using skin color information. Next, the algorithm computed the costs for each pair of iris candidates from intensity valleys as references for iris selection. As for mouth detection, color space method was used to allocate lips region, image processing methods to eliminate unwanted noises and corner detection technique to refine the exact location of mouth. Finally, template matching was used to classify faces based on the extracted features. **Results:** The proposed method had shown a better features detection rate (iris = 93.06%, mouth = 95.83%) than conventional method. Template matching had achieved a recognition rate of 86.11% with acceptable processing time (0.36 sec). **Conclusion:** The results indicate that the elimination of moustache and beard has not affected the performance of facial features detection. The proposed features based template matching has significantly improved the processing time of this method in face recognition research.

Key words: Facial features, face detection, iris detection, mouth detection, face recognition

INTRODUCTION

In past decade, face recognition has made great strides. It has attracted global attention and rapidly maturing growth in researches. Nowadays, face recognition has been widely applied in security system, surveillance system, credit card verification and other applications. Automatic face and facial features detection have become important for face recognition. Automated face recognition has been applied in holistic and feature-based methods. The holistic approach^[4,5] treats a face as two dimensional pattern of intensity variation. Eyes and mouth are important facial landmarks which significantly affect the performance of face recognition. The feature-based approach^[3,12] recognizes a face using the geometrical measurements taken among facial features.

Many researches have proposed methods to find the eyes^[9], mouth regions^[5,6] and locate the face region^[3,4,8] in an image or video^[11]. These methods have shown the

popularity of using information such as template matching, geometrical and intensity features. Brunelli and Poggio^[3] and Beymer^[4] located eyes using template matching. Template matching and eigenspace methods require normalization of the face for variation of size and orientation. A large number of templates are needed for template matching to accommodate varying poses.

The vertical and horizontal projections are one of the popular methods used in feature based approaches. The projections can be applied on the first derivative of the image^[3] or directly on the intensities values. Projection based methods^[7,9] have been used particularly to find coarsely the position of the facial features. The projection method is being combined with other methods such as genetic algorithm^[8] and gabor transformation^[18]. The amount of work that gave qualitative results about the performance of features localization is limited as most of the previous studies^[7,9,12,13] excluded the effects of beard or moustache.

Corresponding Author: Chai Tong Yuen, Department of Mechatronic and Biomedical Engineering,
University Tunku Abdul Rahman, Jalan Genting Kelang, Setapak 53300, Kuala Lumpur, Malaysia

In the most recent researches, a fast facial features detection algorithm has been developed based on generative frameworks^[19]. Adaboost is used to search for features through the entire face to get all possible combination of facial features. The final combination is determined by maximum likelihood. The detection rate is arguable as a detection of more than three features is considered as successful detection rather than calculate the rates separately. There is a need to manually zoom the image before running detection for features which are too small. Besides, a new facial features localization technique^[20] has been proposed based on multi-stream Gaussian Mixture Modal (GMM) to locate regions of eyes, nose and mouth corners. PCA is used for feature extraction but the detection of eyes at the initial stage was not fully automated.

As for classification method, template matching^[3] and neural networks^[15] have been widely implemented. Template matching is sensitive to variation of poses while neural networks consume time in training when dealing with bigger database. Fuzzy clustering has been combined with parallel neural networks^[14] to classify human faces. In this study, an algorithm has been proposed to reduce the effects of thick beard, moustache and facial expressions on facial features detection. The proposed algorithm will first locate the face region using skin-color. Iris candidates are extracted and selected from valleys based on computed costs. Then, the algorithm detects the lips region using the RGB color space. Image processing and connected component labeling have been used to eliminate unwanted noise and extract the lips region. SUSAN algorithm is applied to detect the mouth corners in order to obtain the accurate position of mouth. Finally, template matching is used to recognize faces in database based on extracted eye and mouth templates.

MATERIALS AND METHODS

Face region extraction: In this experiment, color images for males and females with beards, moustaches, long hair and facial expressions are specifically chosen from the online face database^[16]. The images are assumed to be head-shoulder images with plain background and head rotation on y-axis within the interval $\pm 30^\circ$. The proposed algorithm first creates a skin-color model from color images:

$$r = \frac{R}{R+G+B}, g = \frac{g}{R+G+B} \quad (1)$$

Using the r,g color space as in Eq. 1, a skin-color model is created. First, select skin-color pixels whose

color value $v = (r,g)$ satisfies the set threshold. Second, estimate the mean and covariance matrix of the pixels in the selected regions^[2]. Then, the Gaussian distribution model is computed using the probability density function. Mathematical morphology: opening and closing are applied to the region of skin-color pixels. Finally, the largest area of connected component of the region is defined as face region.

Detection and selection of iris: The proposed algorithm first converts the color image to grayscale image and applies grayscale closing^[11] to the face region to extract intensity valleys:

$$v(x,y) = g(x,y) - i(x,y) \quad (2)$$

where, $g(x,y)$ and $i(x,y)$ denote the value obtained by applying grayscale closing and its intensity value. Then, region consisting of pixels which is greater than or equal to a threshold value are determined to be valleys. Preprocessing steps such as illumination normalization^[17] and light spot deletion are performed to enhance the quality of image and reduce illumination effect in this algorithm.

This algorithm performs the similar method^[2] for iris selection. For iris candidates' selection, the costs for all pixels in the valleys are computed and select m pixels according to non-increasing order that give the local maxima. First, the eye template is being placed at each candidate location and measures the separability^[2] between outer and inner regions. Second, Canny edge detector^[10] is applied to the face region. The fitness and maximum vote ratio of iris candidates are measured using Hough transform. Third, the fitness of the iris candidates to the intensity image is measured by placing two different templates^[11]. Fourth, the average intensity of each iris candidate over the average intensities for all candidates has to be calculated. The sum of all the four calculated costs for each iris candidate as mentioned is being used for final determination as in Eq. 3:

$$F(i,j) = t\{C(i) + C(j)\} + (1-t)/R(i,j) \quad (3)$$

Where:

- C(i) and C(j) = The sum of the computed costs for each pair of iris candidates
- R(i,j) = The normalized cross-correlation value computed by using eye template which is manually cut from a face image
- t = The weight to adjust two terms of the cost



Fig. 1: The detected mouth corners after applying the image processing steps and SUSAN corner detector

The lowest cost $F(i,j)$ is selected as the pair of irises.

Mouth detection and refinement: First, the detection process starts by using the detected irises location as the left X_{l-5} and right X_{r+15} boundaries in estimating the mouth region. Then, the vertical positions of both irises are used to calculate a mean value, Y_{min} :

$$\begin{aligned} Y_1 &= Y_{min} + (X_l - X_r) \times R \\ Y_2 &= Y_1 + H \end{aligned} \quad (4)$$

Where:

- Y_1 and Y_2 = The upper and lower boundaries
- R = The height ratio
- H = The preference height of the region

The region of interest with the calculated width and height from Eq. 4 is being clipped out of the original color image.

The proposed algorithm applies RGB color space on estimated mouth region. RGB color space is being applied to the clipped image to allocate only the red lips region. So, the red component is compared with the green and blue in RGB color space as in Eq. 5 with a threshold being set:

$$\begin{aligned} I1 &= \left(\begin{array}{l} (R/R + G + B) > \\ \text{threshold} * (G/R + G + B) = 1 \end{array} \right) \\ I2 &= \left(\begin{array}{l} (R/R + G + B) > \\ \text{threshold} * (B/R + G + B) = 1 \end{array} \right) \\ I &= I1 * I2 \end{aligned} \quad (5)$$

where, I is the detected lips which appeared in white region.

Among all the image processing steps, connected component labeling plays an important role to refine the mouth location. In this case, it is being used not only to eliminate unwanted regions but also detected the lips region as the biggest area in the estimated region. Opening, closing and dilation of the mathematical

morphologies are applied according to the thresholds being set for various areas of connected components to have a more complete lips region. Median filter is also being used to eliminate noises and small clusters. After applying all the image processing methods, the largest member of the connected component is being chosen as lips.

SUSAN corner detector is then applied on the processed image. Corners will be detected at the boundary of the white region. The two corners located at the left most and right most are determined as mouth corners as shown in Fig. 1. When there is more than one corner with the same x-position at each side, the point with the nearest y-position to the middle of the image will be considered as a corner. Finally, the average values of y-coordinates from both sides can help in minimizing the error rate and adjusting the mouth corners' positions. This can be very useful when there is only one failure corner detection with small deviation at either side. The result of mouth corners detection is used to refine and center the position of mouth.

RESULTS

There are 72 color images from AR face database which have been chosen to evaluate the performance of the proposed algorithm. The size of each image is 768x576 and the database of 12 subjects for male and female are used respectively. These images chosen are covered variety of poses, head orientations and different face contours. Most of these images are with beard, moustache, long hair, facial expressions, mouth closed and opened conditions.

DISCUSSION

The size of the face database used in this experiment is not large but the variety of faces and subjects are more than adequate to test the robustness of this algorithm. From Table 1, the overall performance of this algorithm in iris detection has achieved 93.06%. The unsuccessful cases are caused by the color of female's irises. When the intensity difference between iris and outer region is lower, failure detection might happen. Besides this, when the subject frowns too hard, the eyebrows will come closer to iris. If the iris is almost covered by the eyelid, a failure might also occur. However, the algorithm has shown good accuracy in iris detection dealing with these types of situations. The iris detection rate can still be improved by using an eye template with more similarity compared with most of the images in the face database.

Table 1: Successful detection rate for iris and mouth corners

Facial features	Detection rate (%)	Processing time (sec image ⁻¹)
Iris	93.06	2.04
Mouth	95.83	0.46

Table 2: Successful face recognition rate using template matching

Template types	Face recognition rate (%)	Processing time (sec)
Eye	79.17	0.25
Mouth	51.39	0.10
Eye + Mouth	86.11	0.36

For mouth detection, the proposed algorithm achieved 95.83% in detecting the mouth corners for refinement purpose. This algorithm combining RGB color space and SUSAN corner detection techniques has shown significant detection rate over the images with beards, moustache, facial expressions and orientation angles. The performance of this method has been proven better than template matching (83.33%) and integral projection (81.94%) by testing with the same face database. From this experiment, template matching has generated bad detections when dealing with certain head orientation angles and facial expressions. It is found not effective in refining the mouth location. Projection method has failed in most of the images with thick beard, moustache and 'mouth-opened' conditions. Both of the methods will be tested again with a bigger database to perform under external lighting from various directions.

The failure cases from the proposed method are mostly caused by 'mouth-opened' but there are also successful detections in this condition. The examples of successful and failed mouth detections are shown in Fig. 2 and 3. The image processing steps carried out in the proposed method to eliminate beard, moustache and unwanted region work effectively. The processing time for mouth detection is 0.46 sec image⁻¹ while 2.04 sec for iris detection as shown. The simplicity of the mouth detection's algorithm has an advantage of processing speed over algorithm for iris detection which is much more complicated. There are successful iris detections from incomplete detected face region and successful mouth detection from failure iris detection even the geometric information of these irises is being used to estimate the mouth region.

From Table 2, the combination of eye and mouth templates which is used for face recognition has recorded a success rate of 86.11%. Vividly, eye template contains more useful information than mouth template with 79.17% successful face recognition. However, mouth template can be used as additional information to further enhance the success rate of face recognition due to its low processing time, 0.10 sec.



Fig. 2: The examples of successful mouth detection from self-selected AR face database



Fig. 3: The examples of failed mouth detection from self-selected AR face database

A total of 0.36 sec has clocked by using eye and mouth templates for face recognition in this experiment. The success rate of face recognition still can be improved by adding more images per person rather than three. In order to have a better recognition rate, the mouth corners need to be detected precisely as it is very useful in calculating correlation values for template matching method.

CONCLUSION

Referring to previous researches, there are lack of mouth detection results specifically dealing with thick beard, moustache and open mouth. The aim of this proposed algorithm is to effectively detect the facial features from different faces with head tilts, poses, facial expressions, beard and moustache. Face recognition using template matching can still be improved by obtaining more precise location of mouth corners as discussed. In future, this proposed algorithm will be tested with a larger face database, additional images with external lighting from different directions and more facial expressions. Artificial neural networks will be implemented for face recognition as a comparison to template matching technique and to improve the recognition rate.

REFERENCES

1. Kawaguchi, T. and M. Rizon, 2003. Iris detection using intensity and edge information. *Patt. Recog.* 36: 549-562. DOI: 10.1016/S0031-3203(02)00066-3

2. Kawaguchi, T., D. Hidaka and M. Rizon, 2000. Detection of eyes from human faces by Hough transform and separability filter. Proceedings of the IEEE 3rd International Conference Image Processing, Sept. 10-13, IEEE Xplore Press, Vancouver, BC, Canada, pp: 49-52. DOI: 10.1002/ecjb.20178
3. Brunelli, R. and T. Poggio, 1993. Face recognition: Features versus templates. IEEE Trans. Patt. Anal. Mach. Intel., 15: 1042-1052. DOI: 10.1109/34.254061
4. Beymer, D.J., 1994. Face recognition under varying pose. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 21-23, IEEE Xplore Press, Seattle, WA., USA., pp: 756-761. DOI: 10.1109/CVPR.1994.323893
5. Li, C.M., Y.S. Li, Q.D. Zhuang and Z.Z. Xiao, 2004. The face localization and regional features extraction. Proceedings of the 3rd International Conference on Machine Learning and Cybernetics, Aug. 2004, IEEE Xplore Press, USA., pp: 26-29. DOI: 10.1109/ICMLC.2004.1380504
6. Sobottka, K. and I. Pitas, 1996. Face localization and facial feature extraction based on shape and color information. Proceedings of the International Conference on Image Processing, Sept. 16-19, IEEE Xplore Press, Lausanne, Switzerland, pp: 483-486. DOI: 10.1109/ICIP.1996.560536
7. Baskan, S., M.M. Bulut and V. Atalay, 2002. Projection based method for segmentation of human face and its evaluation. Patt. Recog. Lett., 23: 1623-1629. DOI: 10.1016/S0167-8655(02)00037-5
8. Wong, K.W., K.M. Lam and W.C. Siu, 2001. An efficient algorithm for human face detection and facial feature extraction under different conditions. Patt. Recog., 34: 1993-2004. DOI: 10.1016/S0031-3203(00)00134-5
9. Ryu, Y.S. and S.Y. Oh, 2002. Automatic extraction of eye and mouth fields from a face image using eigenfeatures and ensemble networks. Applied Intel., 17: 171-185. DOI: 10.1023/A:1016160814604
10. Canny, J., Nov. 1986. A computational approach to edge detection. IEEE Trans. Patt. Anal. Mach. Intel., 8: 679-698. DOI: 10.1109/TPAMI.1986.4767851
11. Balasubramanium, M., S. Palanivel and V. Ramalingam, 2009. Real time face and mouth recognition using radial basis function neural networks. Int. J. Expert Syst. Appli., 36: 1-8. DOI: 10.1016/j.eswa.2008.08.001
12. Song, Y., K. He, J. Zhou, Z. Liu and K. Li, 2004. Multi-resolution feature extraction in human face. Proceedings of International Conference on Information Acquisition, June 21-25, IEEE Xplore Press, USA., pp: 417-421. DOI: 10.1109/ICIA.2004.1373403
13. Chew, W.J., L.M. Ang and K.P. Seng, 2008. Automatic model based face feature detection system. Proceedings of the International Symposium on Information Technology, Aug. 26-28, IEEE Xplore Press, Kuala Lumpur, Malaysia, pp: 1-6. DOI: 10.1109/ITSIM.2008.4631696
14. Jianming Lu, Xue Yuan and Takashi Yahagi, 2006. A method of face recognition based on fuzzy clustering and parallel neural networks. Sign. Process., 86: 2026-2039. DOI: 10.1016/j.sigpro.2005.09.023
15. Bae, H. and S. Kim, 2005. Real-time face detection and recognition using hybrid-information extracted from face space and facial features. Image Vis. Comput., 23: 1181-1191. DOI: 10.1016/j.imavis.2005.07.017
16. Martinez, A.M. and C.K. Avinash, 2001. PCA versus LDA. IEEE Trans. Patt. Anal. Mach. Intel., 23: 228-233. DOI: 10.1109/34.908974
17. Tan, X. and B. Triggs, 2007. Enhanced local texture feature sets for face recognition under difficult lighting conditions. Lecture Notes Comput. Sci., 4778: 168-182. DOI: 10.1007/978-3-540-75690-3_13
18. Zhao, Y., X. Shen and N.D. Georganas, 2008. Combining integral projection and Gabor transformation for automatic facial feature detection and extraction. Proceedings of the IEEE International Workshop on Haptic Audio Visual Environments and Games, Oct. 18-19, IEEE Xplore Press, Ottawa, Ont., pp: 103-107. DOI: 10.1109/HAVE.2008.4685307
19. Erukhimov, V. and K.C. Lee, 2008. A bottom-up framework for robust facial feature detection. Proceedings of the 8th IEEE International Conference on Automatic Face and Gesture Recognition, Sept. 2008, IEEE Xplore Press, pp: 1-6. http://vision.ucsd.edu/~leekc/papers/facedetect_fg08.pdf
20. Kumatani, K., H.K. Ekenel, H. Gao, S. Rainer and E. Aytul, 2008. Multi-stream Gaussian mixture model based facial feature localization. Proceedings of the 16th IEEE International Conference on Signal Processing, Communications and Applications Conference, Apr. 20-22, IEEE Xplore Press, Aydin, Turkey, pp: 1-4. DOI: 10.1109/SIU.2008.4632752