



PERGAMON

Pattern Recognition 34 (2001) 1993–2004

**PATTERN
RECOGNITION**

THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

An efficient algorithm for human face detection and facial feature extraction under different conditions[☆]

Kwok-Wai Wong, Kin-Man Lam*, Wan-Chi Siu

*Centre for Multimedia Signal Processing, Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hung Hom, Hong Kong*

Received 28 January 2000; received in revised form 25 August 2000; accepted 25 August 2000

Abstract

In this paper, an efficient algorithm for human face detection and facial feature extraction is devised. Firstly, the location of the face regions is detected using the genetic algorithm and the eigenface technique. The genetic algorithm is applied to search for possible face regions in an image, while the eigenface technique is used to determine the fitness of the regions. As the genetic algorithm is computationally intensive, the searching space is reduced and limited to the eye regions so that the required timing is greatly reduced. Possible face candidates are then further verified by measuring their symmetries and determining the existence of the different facial features. Furthermore, in order to improve the level of detection reliability in our approach, the lighting effect and orientation of the faces are considered and solved. © 2001 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Face detection; Facial feature extraction; Genetic algorithm; Eigenface technique

1. Introduction

Digital images and video are becoming more and more important in the multimedia information era. The human face is one of the most important objects in an image or video. Detecting the location of human faces and then extracting the facial features in an image is an important ability with a wide range of applications, such as human face recognition, surveillance systems, human–computer interfacing, video-conferencing, etc. In an automatic face recognition system [1], the first step is to segment the face in an image or video irrespective of whether the background is simple or cluttered. For model-based video coding [2], the synthesis performance is quite dependent on the accuracy of the facial feature extraction

process. In other words, a reliable method for detecting the face regions and locating the facial features is indispensable to such applications. This paper presents an efficient method for face detection and facial feature extraction in a cluttered image.

1.1. Problems of face detection and facial feature extraction

In fact, detecting human faces and extracting the facial features in an unconstrained image is a challenging process. It is very difficult to locate the positions of faces in an image accurately. There are several variables that affect the detection performance, including wearing of glasses, different skin coloring, gender, facial hair, and facial expressions. Furthermore, the human face is a three-dimensional (3-D) object, and might be under a distorted perspective and uneven illumination. As a result, a true face may not be detected. Moreover, facial feature extraction is a time-consuming process due to the lack of constraint on the number, location, size, and orientation of faces in an image or a video scene.

[☆]This work was supported by HKPolyU research grant G-V709.

* Corresponding author. Tel.: + 852-276-66207; fax: + 852-236-28439.

E-mail address: enkmlam@polyu.edu.hk (K.-M. Lam).

1.2. Existing work on face detection and facial feature extraction

Recently, human face detection algorithms based on color information have been reported [3–5]. The face regions are initially segmented based on the characteristic of skin tone colors. The color signal is usually separated into its luminance and chrominance components in an image or video. Experimental results show that the skin-like regions can be segmented by considering the chrominance components only. Although skin colors differ from person to person, and race to race, they are distributed over a very small area on the chrominance plane. The major difference between the skin tones is intensity. However, human face detection and facial feature extraction in gray-level images may be more difficult because the characteristics of skin tone color are not available. Sung et al. [6] proposed an example-based learning approach for locating vertical frontal views of human faces in complex scenes. A decision-making procedure is trained based on a sequence of “face” and “non-face” examples. Six “face” clusters and six “non-face” clusters are obtained according to 4150 normalized frontal face patterns. The face regions are located by matching the window patterns at different image locations and scales against the distribution-based face model. Yang et al. [7] proposed a hierarchical knowledge-based method consisting of three levels for detecting the face region and then locating facial components in an unknown picture. Mosaic images of different resolutions are used in the two higher levels. Two sets of rules based on the characteristics of a human face region are applied to the mosaic images. At third level, the edges of facial components are extracted for the verification of face candidates. However, the computational requirements of these methods may be too high for some applications, which may be unable to detect and locate a tilted human face reliably.

Extraction of facial features by evaluating the topographic gray-level relief has been introduced [3,8,9]. Since the intensity is low for the facial components, the position of the facial features can be determined by checking the mean gray-level in each row and then in each column. In [9–11], facial feature detection based on the geometrical face model was proposed. The model is constructed based on the relationships among facial organs such as nose, eyes, and mouth. However, these methods can work properly only under well-lit conditions. Therefore, the pre-processing step for reducing the lighting effect is very important for the methods.

In our previous work [12,13], possible face candidates in a gray-level image with a complex background were identified by means of valley features on the human eyes. In this paper, we propose an efficient method for locating the face region and facial features based on the characteristics of eye regions. The face regions are segmented

based on a pair of possible eye candidates. The facial features are then extracted from the detected face regions. In order to improve the level of detection reliability, the lighting effect is also considered and alleviated for the possible face regions. This method is tested with the MIT face database and some other complex images. Experimental results show that faces can be detected more reliably and efficiently now compared with our previous work. The details of our approach for face and facial feature detection will be described in the following sections.

2. Human face detection using the genetic algorithm and eigenface technique

Our method for detecting and extracting the facial features in a gray-level image is divided into two stages. Firstly, the possible human eye regions are detected by testing all the valley regions in an image. A pair of eye candidates are selected by means of the genetic algorithm [14] to form a possible face candidate. The fitness value of each candidate is measured based on its projection on the eigenfaces [15]. In order to improve the level of detection reliability, each possible face region is normalized for illumination; the shading effect, when the head is tilted, is also considered as well. After a number of iterations, all the face candidates with a high fitness value are selected for further verification. At this stage, the face symmetry is measured and the existence of the different facial features is verified for each face candidate. The facial features are determined by evaluating the topographic relief of the normalized face regions. The facial features extracted include the eyebrow, the iris, the nostril, and the mouth corner.

Genetic algorithm is an optimization technique that operates on a population of individual solutions. It has been successfully applied for many purposes, such as object recognition [16], human face detection [17,18], facial feature extraction [19], and motion estimation for video coding [20]. In our approach, genetic algorithm is also applied to search for possible facial regions in an image. The first step in locating the face regions in our approach is to select a pair of eye candidates using genetic algorithms. The fitness value for each face candidate is calculated by projecting it onto the eigenfaces space. Since eigenfaces [15] is a successful approach for face recognition, we therefore adopt it as a fitness function.

2.1. Possible eye candidates detection

In our approach, the possible eye regions are located by detecting the valley points in an image. Since the human iris in a gray-level image is of low intensity, a valley exists at an eye region. The valley field Φ_v can be

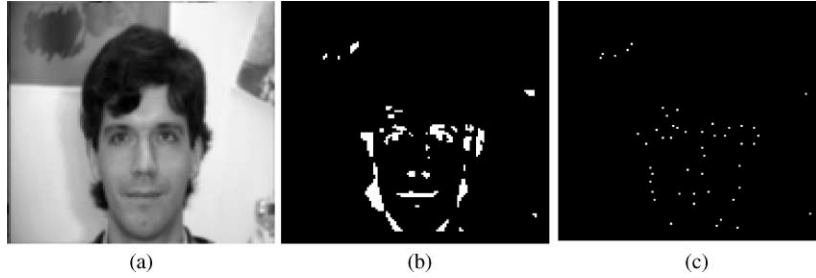


Fig. 1. Eye candidates detection: (a) original image, (b) possible eye regions, and (c) the good candidates for the detected eye regions.

extracted using morphological operators [21]. The equation for valley field extraction is

$$\Phi_v = f(x, y) \cdot B - f(x, y), \quad (1)$$

where $f(x, y)$ is the image and B is the structuring element. The valley image is obtained by performing a closing operation, which is then subtracted by the original image. A pixel at (x, y) is considered as a possible eye candidate if the following criteria are satisfied:

$$f(x, y) < t_l \quad \text{and} \quad \Phi_v(x, y) > t_v, \quad (2)$$

where t_l and t_v are thresholds. Fig. 1(a) and (b) shows the original image and its corresponding possible eye regions, respectively. The segmented possible eye regions are then reduced to a point or a number of points by choosing the good candidates in each region. The good eye candidates are those which have large values in the functions, $F1(x, y)$ and $F2(x, y)$. The two functions are defined as follows:

$$F1(x, y) = W_{1,1} \left(\frac{f(x-2, y) + f(x+2, y)}{2} - S_{1,1}(x, y) \right) + W_{1,2} \Phi_{1,1}(x, y), \quad (3)$$

$$F2(x, y) = W_{2,1} \left(\frac{f(x-3, y) + f(x+3, y)}{2} - S_{2,1}(x, y) \right) + W_{2,2} \Phi_{2,1}(x, y),$$

where W 's are the weighting factors, $S_{1,1}(x, y)$ and $S_{2,1}(x, y)$ are the average gray-level intensities of the region under 3×3 and 5×5 windows, respectively, and $\Phi_{1,1}(x, y)$ and $\Phi_{2,1}(x, y)$ are the average value of the valley field under the 3×3 and 5×5 windows, respectively. This arrangement allows us to detect the eyes according to different scales. Fig. 1(c) illustrates the good eye candidates for the segmented regions in Fig. 1(b). The locations of the possible eye candidates are stored in

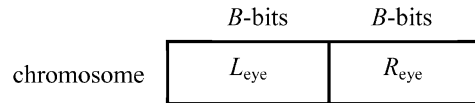


Fig. 2. Structure of a chromosome.

a buffer. In the genetic algorithm, two entries are selected from the buffer to form a possible face candidate. Therefore, the search space is limited to the possible eye candidates, which can then greatly reduce the required runtime.

2.2. Structure of a chromosome

Each solution generated for a problem using the genetic algorithm is called a chromosome or string, which is represented in binary format. In our approach, two components are used to specify a face region in a chromosome. The two components which represent the position of the left eye (L_{eye}) and the right eye (R_{eye}) are the index numbers to the buffer. The structure of the chromosome is illustrated in Fig. 2. The number of bits required to represent the L_{eye} or R_{eye} is $B = \lceil \log_2 N \rceil$, where N is the total number of detected eye candidates. Thus, the total number of bits in each chromosome is $2B$.

Since the size of a human face is proportional to the distance between the two eyes (d_{eye}), a possible face region which contains the eyebrows, eyes, nose, and mouth can be formed based on this relationship. In our method, a square block is used to represent the detected face region. Fig. 3 shows an example of a selected face region based on the location of an eye pair. The line passing through the centers of the eye pair is called the base line. The extracted possible face regions are subsampled and interpolated to a resolution of 28×31 . A low resolution has been proved to be sufficient for face identification. Moreover, the required computation is also reduced due to the fact that fewer pixels need to be manipulated. The relationships between the eye pair, the face size, and the orientation angle θ between the base

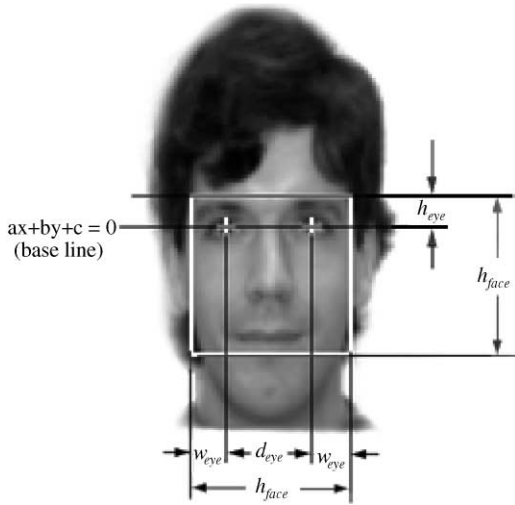


Fig. 3. The defined geometry of our head model.

line and the x -axis are defined as follows:

$$h_{face} = 1.8d_{eye}, \tag{4a}$$

$$h_{eye} = \frac{1}{5}h_{face}, \tag{4b}$$

$$w_{eye} = 0.225h_{face}, \tag{4c}$$

$$a = y_2 - y_1,$$

$$b = x_2 - x_1,$$

$$c = x_2y_1 - x_1y_2,$$

$$\theta = \tan^{-1}\left(-\frac{a}{b}\right), \quad -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}. \tag{4d}$$

Based on the locations of the eye pairs, a population of possible face regions of different locations, sizes, and orientations can be generated. An initial population of the chromosomes is generated by pairing the possible eye candidates, depicted as white dots in Fig. 1(c). If the total number of detected eye candidates is N , the total number of pairing combinations for the initial chromosomes is $N(N - 1)/2$. Therefore, members of the initial population are produced by randomly selecting from the $N(N - 1)/2$ chromosomes.

2.3. Normalization of the possible face regions

The orientation angle of a face candidate can be determined based on the gradient of the eye pair. However, the human face is not a rigid object; it will suffer from a shirring effect if the head is rotated too much, as

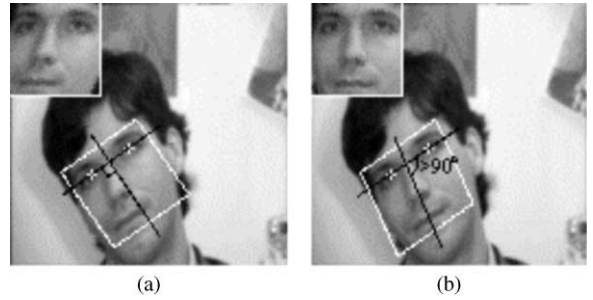


Fig. 4. (a) A normal face candidate, and (b) adjusted possible face candidate.

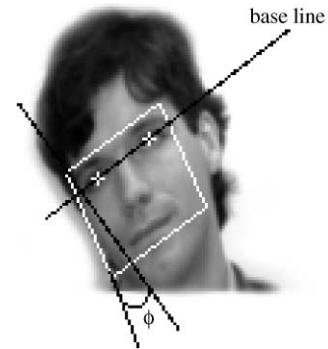


Fig. 5. Shirring angle approximation.

illustrated in Fig. 4. In Fig. 4(a), if the face region is considered to be rectangular, the extracted face will be distorted. However, if the face region is a parallelogram, as shown in Fig. 4(b), the shirring effect is alleviated and a more upright face can be extracted. In our approach, the shirring effect will be compensated when the rotation angle $\theta > 10^\circ$. In this case, the shirring is estimated to be $\theta/3$, which is based on the measurement of over 50 rotated human faces. If the rotation angle is less than 10° , the shirring effect may be neglected. If the rotation angle is larger than 10° , two possible face candidates will be generated for a chromosome in the calculation of the fitness values. One candidate uses a rectangular face region, while the other one is adjusted based on the shirring angle of the face. The shirring angle, ϕ , is defined as shown in Fig. 5. This normalization process for the shirring effect is performed using the following transformation:

$$\begin{bmatrix} xr \\ yr \end{bmatrix} = \begin{bmatrix} \tan \phi \sin \theta + \cos \theta & -\tan \phi \cos \theta + \sin \theta \\ -\sec \phi \sin \theta & \sec \phi \cos \theta \end{bmatrix} \times \begin{bmatrix} x \\ y \end{bmatrix}, \tag{5}$$

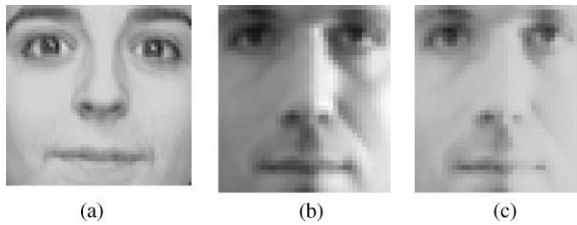


Fig. 6. (a) Reference face image, (b) original face image with half in shadow, and (c) the histogram normalized face image.

where xr and yr denote the coordinates of x and y after compensating for the shirring effect. The derivation of this equation is shown in the appendix.

The detection performance is also affected by the external environment, such as the direction of the lighting source. The uneven lighting conditions make a face become asymmetrical; a true face may not be detected. In order to reduce the lighting effect, the possible face candidates would be normalized by transforming their histograms to the histogram of a reference face image [22] before calculating the fitness value. This can be achieved due to the fact that all human faces have basically the same shape and general illumination properties. The advantage of the histogram normalization is that the size of the reference face image and the input region can be unequal. Thus, it is unnecessary to resample the face candidate to the size of the reference face. Fig. 6 shows an example of the histogram normalization of a face region. After the shirring effect and the histogram normalization processes, the fitness value of the face region will then be computed.

2.4. The fitness function

To determine whether the normalized face candidate is a face or not, the fitness value of the possible face regions is computed by means of the eigenfaces [15]. The eigenfaces are obtained by extracting the principle components from a training-set of pre-processed face images. The training images are also pre-processed by a histogram normalized to reduce the lighting effect. The normalized possible face region is then projected onto the eigenface space in order to calculate the fitness. The fitness function is a measure of the distance between its projection and that of the training-set face images. The distance between the mean adjusted faces Φ of the training images and the projection of the mean adjusted input region $\Phi_f(n)$ for the n th chromosome on the face space is calculated by

$$\varepsilon(n) = \|\Phi - \Phi_f(n)\|. \quad (6)$$

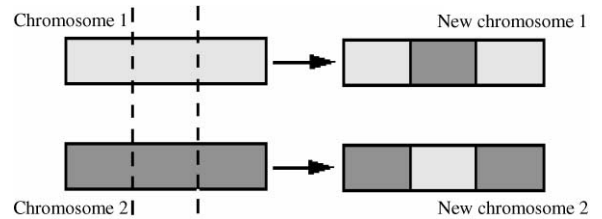


Fig. 7. The crossover process.

Table 1
Parameter settings

Population size	100
Selection probability	0.9
Crossover probability	0.8
Mutation probability	0.08
Chromosome length	2B bits

The value of $\varepsilon(n)$ is a measure of the distance between the input candidate and the training images. Thus, the fitness function of the possible face region for the n th chromosome is defined as

$$f(n) = \frac{1}{\varepsilon(n)}. \quad (7)$$

From Eq. (7), it follows that a chromosome with a smaller distance will have a larger fitness value. A new population is then generated by means of the genetic operations: selection, crossover, and mutation. A chromosome with a higher fitness value will have a better chance of being chosen for the next generation. In the crossover process, two chromosomes are selected from the mating pool. In our method, the two-point crossover method is employed. Two cutting points are selected randomly within the chromosome for exchanging the contents. The crossover process is illustrated in Fig. 7. Since the probabilities of a chromosome being selected for the crossover and mutation processes are proportional to its computed fitness value, these good offspring will probably be passed on to the next generation. In the order to increase the success rate, the best candidate in one generation can pass directly to the next generation. After a number of iterations, those good candidates will be further verified as to whether they are human faces. The parameter settings used in our approach are shown in Table 1. The extracted good face candidates are then input to the next stage for further verification and facial feature extraction.

2.5. Verification of face regions and facial feature extraction

The possible face candidates with a high fitness value are passed on to the second stage. The functions of

the second stage are to verify whether the candidates are human faces or not, and to extract the respective facial features in the face region. The verification process is based on the characteristics of the projected face images.

At this stage, the symmetry of a face candidate is measured. As every face region is normalized for the shading effect and the illumination effect, the difference between the left half and the right half of a face region should be small due to its symmetry. In our method, the size of a face region is normalized to 28×31 , and the symmetrical measure is calculated as follows:

$$T_s = \frac{1}{434} \sum_{y=0}^{30} \sum_{x=0}^{13} |f(x, y) - f(27 - x, y)|, \quad (8)$$

where $f(x, y)$ represents a possible face candidate. If the value of T_s is smaller than the threshold, the face candidate will be selected for further verification. In any overlapping region, the one with the lowest value of T_s is chosen.

After measuring the symmetry of a face candidate, the existence of the different facial features is also verified. The position of the facial features is determined by analyzing the projection of the normalized face candidate region. The facial feature regions will exhibit a low value on the projection. A normalized face region is divided into three parts; each of which contains the respective facial features. In our method, the y -projection is performed in each part to determine the vertical position of the facial features. The y -projection is the average of the gray-level intensities along each row of pixels in a window. In order to reduce the effect of the background in a face region, only the white windows as shown in Fig. 8 are considered in computing the projections. The two top windows contain the eyebrows and the eyes; the middle window contains the nose; and the bottom window contains the mouth. In each of the windows, the

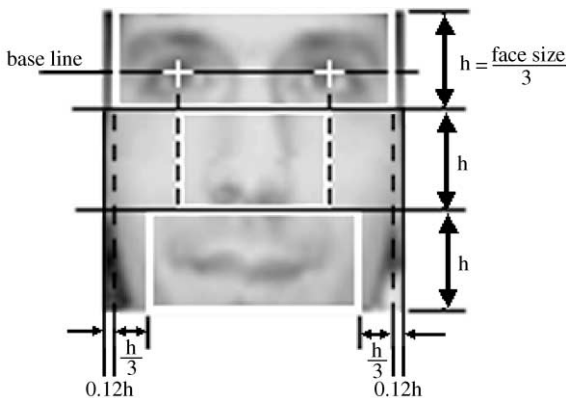


Fig. 8. Windows for facial feature extraction.

position where the projection value is a minimum is identified. For each of the two top windows, two significant minima will be detected due to the eyebrow and eye, respectively. These minima indicate the vertical position of the eyebrow and the eye. Similarly, the minima in the middle and the bottom windows represent the vertical position of the nostril and the mouth, respectively. The results of the y -direction for the windows in Fig. 8 are shown in Fig. 9. A valid minimum is identified by measuring the difference between the minimum and its neighboring maximum. If the vertical position of any of the facial features cannot be found, the face candidate is then declared as a non-facial image, and is rejected from the x -projection process.

Having obtained the vertical position of the respective facial features, the horizontal position of the facial features is then determined by the x -projection. The x -projection is computed by averaging the gray-level intensities on each column in a window. The position of the eyes can be estimated by performing an x -projection around their vertical position and identifying the location of the two minimal points of the projection. For the eyebrows, sudden changes in the x -projection values signify the end points of the eyebrows. To detect the horizontal position of a nostril in the middle window, two significant minima and a maximum between the two minima will be obtained. The first minimum represents the horizontal position of the left nostril, while the second minimum represents the right nostril. Fig. 10(a) shows the x -projection for determining the nostrils. For the bottom window, the mouth corner can be detected based on two assumptions; the mouth corners are close to the horizontal position of the corresponding iris and the gray-level intensity changes significantly at the mouth corner. Fig. 10(b) illustrates the x -projection and the determination of the detected mouth corners. The detection result for the respective facial features is shown in Fig. 11. Similarly, if any horizontal position of the facial features cannot be located, the candidate is assumed to be a non-facial image. Otherwise, a true face region is declared, as are the different facial features being located.

3. Experimental results

In our approach, if the fitness value of the chromosome is greater than the threshold, it is assumed to be a possible face candidate. These possible face candidates will pass into the second stage for further verification. In the second stage, the symmetry of a face candidate will be calculated. If the difference between the left- and right-half regions of the candidate is greater than the threshold, it is declared a non-facial image. Otherwise, the projection processes will be applied to detect the respective facial features. If the projection results of the face

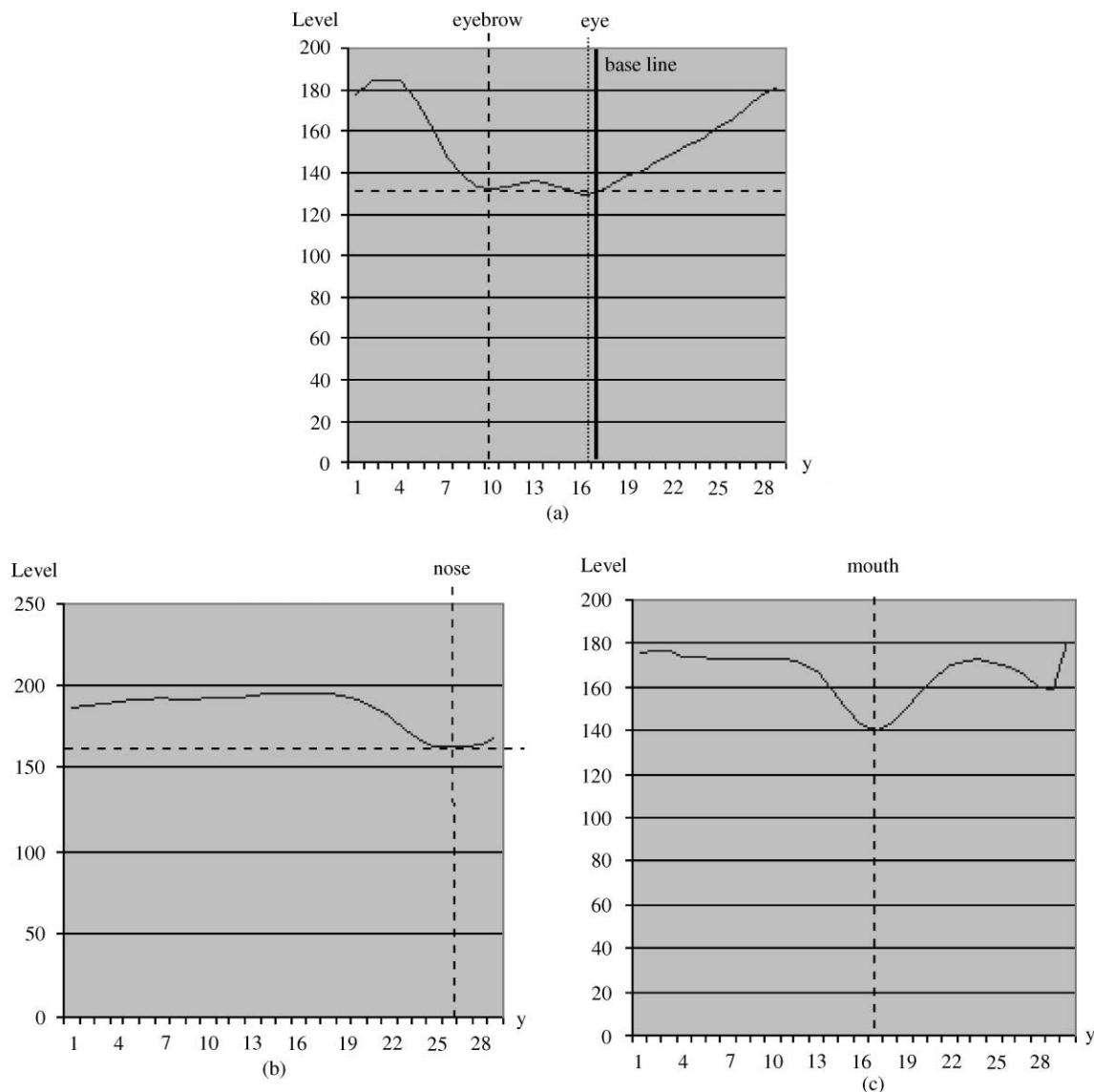


Fig. 9. The y-projection results of (a) eye region, (b) nose region, and (c) the mouth region.

candidate do not fulfill the defined rules for facial features, the face candidate will also be declared a non-facial image.

The detection performance of our method is tested using the face database from MIT and some images with a number of faces. In the experiment, the training set images are different from the test images. Table 2 shows both the hit and miss rates of our method of face detection under different conditions. This approach can achieve an overall hit rate of 100% without head tilt and under head-on lighting. When the heads tilt to the left or right, the hit rate is 95.3%. When the light source to the

faces is 45°, the hit rates for the upright and tilted faces are 87.5 and 82.8%, respectively. When the lighting is 90°, the hit rate for the upright face is 93.75% and the hit rate for the tilted face is 81.25%. The experiment shows that the hit rates for a tilted face after performing shirring normalization have a great improvement over our previous work on face detection.

The hit rates for facial feature detection are tabulated in Table 3. In this part, only those faces detected successfully are considered. The reasons for the failure in detecting the facial features can be summarized as follows: facial images with glasses may affect the determination of

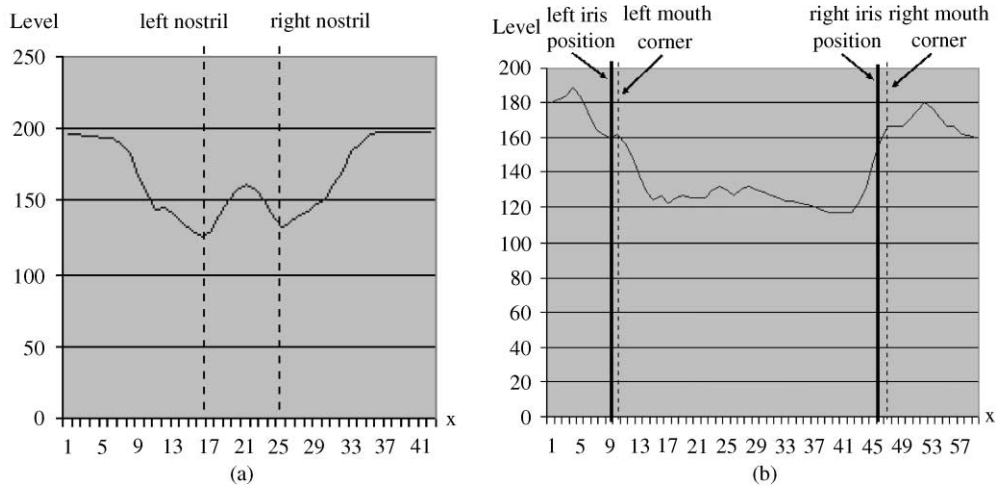


Fig. 10. The x-projection results of (a) nose region, and (b) the mouth region.



Fig. 11. An example of facial feature extraction by analyzing the projection of normalized face region.

the eyebrows; nostril detection is highly affected by the lighting conditions; and a moustache in a facial image covers the mouth corners. Fig. 12 shows the detection results under different lighting conditions and different angles of rotation, while Fig. 13 illustrates some errors in locating the facial features.

Our method is extended to the detection of multiple faces in an image. A user may choose to make either a single-face or a multiple-faces detection. The respective processes for detecting a single face and multiple faces are very similar. The major difference is in the threshold setting in stage one: the threshold value for single-face detection is greater than that for multiple-faces detection. This means that more face candidates may pass into the next stage in multiple-faces detection. Thus, false alarms will happen in this case. We have tested 20 images with multiple-faces (2–3 faces in each of the images). The total number of false alarms is 6, while the hit rate is 92%. The

experiments were performed on a Pentium II 400 MHz computer. The average processing time for locating faces and the facial features in a picture of size 128×120 is about 2.18 s. In conclusion, this method outperforms those used in our previous work.

4. Conclusion

In this paper, we have proposed a more reliable face detection approach based on the genetic algorithm and the eigenface technique. Firstly, possible eye candidates are obtained by detecting the valley points. Based on a pair of eye candidates, possible face regions are generated by means of the genetic algorithm. Each of the possible face candidates is normalized by approximating the shirring angle due to head movement. Furthermore, the lighting effect is reduced by transforming their histograms into the histogram of a reference face image. The fitness value of a face candidate is calculated by projecting it onto the eigenfaces. Selected face candidates are then further verified by measuring their symmetries and determining the existence of the different facial features. The advantages of our approach are that a tilted human face can still be detected robustly even if the face is shirred, under shadow, of a different scale, under bad lighting conditions, and is wearing glasses. In conclusion, this method can achieve a high-performance level in detecting human faces and extracting facial features in complex and simple backgrounds.

Appendix

A face region under shirring effect is illustrated in Fig. 14(a), where θ and ϕ are the angle of rotation and

Table 2
Experimental results for face detection

	Lighting	Head on		45°		90°	
	Head tilt	No	Tilt	No	Tilt	No	Tilt
Full scale	Hit	16	30	15	26	16	26
	Miss	0	2	1	6	0	6
Medium scale	Hit	16	31	13	27	14	26
	Miss	0	1	3	5	2	6

Table 3
Experimental results for the facial feature extraction

	Lighting	Head on		45°		90°	
	Head tilt	No	Tilt	No	Tilt	No	Tilt
Full scale	Hit rate of first part (eyebrow and iris)	$\frac{16}{16}$	$\frac{30}{30}$	$\frac{15}{15}$	$\frac{25}{26}$	$\frac{16}{16}$	$\frac{25}{26}$
	Hit rate of middle part (nostril)	$\frac{16}{16}$	$\frac{29}{30}$	$\frac{15}{15}$	$\frac{24}{26}$	$\frac{15}{16}$	$\frac{23}{26}$
	Hit rate of bottom part (mouth corner)	$\frac{15}{16}$	$\frac{28}{30}$	$\frac{14}{15}$	$\frac{24}{26}$	$\frac{14}{16}$	$\frac{22}{26}$
Medium scale	Hit rate of first region (eyebrow and iris)	$\frac{16}{16}$	$\frac{31}{31}$	$\frac{13}{13}$	$\frac{26}{27}$	$\frac{13}{14}$	$\frac{24}{26}$
	Hit rate of middle part (nostril)	$\frac{16}{16}$	$\frac{31}{31}$	$\frac{12}{13}$	$\frac{26}{27}$	$\frac{13}{14}$	$\frac{23}{26}$
	Hit rate of bottom part (mouth corner)	$\frac{15}{16}$	$\frac{29}{31}$	$\frac{11}{13}$	$\frac{24}{27}$	$\frac{12}{14}$	$\frac{22}{26}$

the shirring angle, respectively. Rotating the region about the point O by an angle θ , the rotation transformation is as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{A1}$$

The rotated region is illustrated in Fig. 14(b). Let pt1 be the point before normalization and pt2 be the point after shirring normalization, then

$$\frac{x' - x''}{y'} = \tan \phi \Rightarrow x'' = x' - y' \tan \phi,$$

$$(y'')^2 = l^2 = y'^2 + (x' - x'')^2,$$

$$(y'')^2 = (y')^2 + (x' - (x' - y' \tan \phi))^2,$$

$$\Rightarrow y'' = y' \sec \phi,$$

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} 1 & -\tan \phi \\ 0 & \sec \phi \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

From Eq. (A1), we have

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} 1 & -\tan \phi \\ 0 & \sec \phi \end{bmatrix} \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} \cos \theta + \tan \phi \sin \theta & \sin \theta - \tan \phi \cos \theta \\ -\sin \theta \sec \phi & \cos \theta \sec \phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$



Fig. 12. (a) Experimental results under head-on lighting, (b) experimental results when the lighting is 45° , (c) experimental results when the lighting is 90° , and (d) some more experimental results.

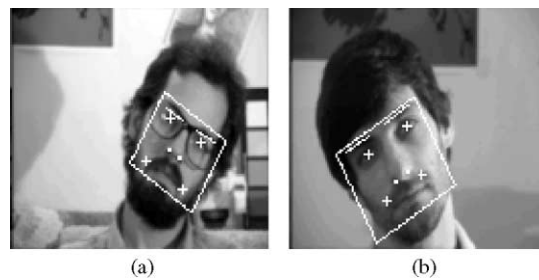


Fig. 13. Error in facial feature extraction.

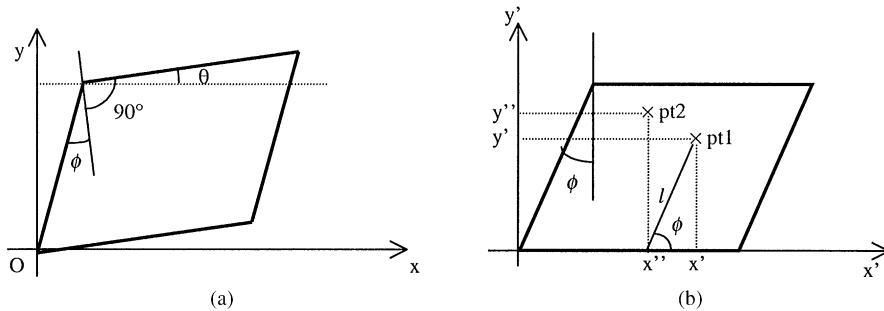


Fig. 14. (a) A region under shearing effect, and (b) the rotated region.

References

- [1] K.-M. Lam, H. Yan, An analytic-to-holistic approach for face recognition based on a single frontal view, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (1998) 673–686.
- [2] L. Zhang, Automatic adaptation of a face model using action units for semantic coding of videophone sequences, *IEEE Trans. Circuits Systems Video Technol.* 8 (6) (1998) 781–795.
- [3] K. Sobottka, I. Pitas, A novel method for automatic face segmentation, facial feature extraction and tracking, *Signal Process. Image Commun.* 12 (1998) 263–281.
- [4] H. Wang, S.F. Chang, A highly efficient system for automatic face region detection in MPEG video, *IEEE Trans. Circuits Systems Video Technol.* 7 (4) (1997) 615–628.
- [5] A.M. Mohamed, A. Elgammal, Face detection in complex environments from color images, *Proceedings of International Conference on Image Processing 3* (1999) 622–626.
- [6] K.K. Sung, T. Poggio, Example-based learning for view-based human face detection, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (1) (1998) 39–51.
- [7] G. Yang, T.S. Huang, Human face detection in a complex background, *Pattern Recognition* 27 (1) (1994) 53–63.
- [8] F.C. Wu, T.J. Yang, M. Ouhyoung, Automatic feature extraction and face synthesis in facial image coding, *Sixth Pacific Conference on Computer Graphics and Applications*, 1998, pp. 218–219.
- [9] A.M. Alattar, S.A. Rajala, Facial features localization in front view head and shoulders images, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 6, 1999, pp. 3557–3560.
- [10] S.H. Jeng, H. Yuan, M. Liao, C.C. Han, M.Y. Chern, Y.T. Liu, Facial feature detection using geometrical face model: an efficient approach, *Pattern Recognition* 31 (3) (1998) 273–282.
- [11] A. Al-Oayed, A.F. Clark, An algorithm for face and facial-feature location based on gray-scale information and facial geometry, *International Conference on Image Processing and Its Applications*, Vol. 2, 1999, pp. 625–629.
- [12] K.M. Lam, A fast approach for detecting human faces in a complex background, *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 4, 1998, pp. 85–88.
- [13] K.W. Wong, K.M. Lam, A reliable approach for human face detection using genetic algorithm, *Proceedings of the IEEE International Symposium on Circuits and Systems*, Vol. 4, 1999, pp. 499–502.
- [14] D.E. Goldberg, *Genetic algorithms in search, optimization, and machine learning*, Addison-Wesley, Reading, MA, 1989.
- [15] M. Turk, A. Pentland, Eigenfaces for recognition, *J. Cognitive Neurosci.* 3 (1) (1991) 71–86.
- [16] D.L. Swets, B. Punch, J. Weng, Genetic algorithms for object recognition in a complex scene, *Proceedings of the International Conference on Image Processing*, Vol. 2, 1995, pp. 595–598.
- [17] Y. Suzuki, H. Saito, D. Ozawa, Extraction of the human face from natural background using GAs, *Proceedings of the IEEE TENCON, Digital Signal Processing Applications*, Vol. 1, 1996, pp. 221–226.
- [18] Y. Yokoo, M. Hagiwara, Human faces detection method using genetic algorithm, *Proceedings of IEEE International Conference on Evolutionary Computation*, May 1996, pp. 113–118.
- [19] C.H. Lin, J.L. Wu, Automatic facial feature extraction by genetic algorithms, *IEEE Trans. Image Process.* 8 (6) (1999) 834–845.
- [20] C.H. Lin, J.L. Wu, Genetic block matching algorithm for video coding, *Proceedings of the Third IEEE International Conference on Multimedia Computing and Systems*, 1996, pp. 544–547.
- [21] P. Maragos, Tutorial on advances in morphological image processing and analysis, *Opt. Engng.* 26 (7) (1987) 623–632.
- [22] P.J. Phillips, Y. Vardi, Efficient illumination normalization of facial images, *Pattern Recognition Lett.* 17 (1996) 921–927.

About the Author—KWOK-WAI WONG received his B.E. degree in Electronic Engineering from the Hong Kong Polytechnic University, Hong Kong, in 1998. He is currently pursuing his M.Phil. degree in the Department of Electronic and Information Engineering, the Hong Kong Polytechnic University. His research interests include very low bit rate video coding, motion tracking, human face detection and recognition.

About the Author—KIN-MAN LAM received his Associateship in electronic engineering from The Hong Kong Polytechnic University in 1986 and the M.Sc. degree in communication engineering from the Department of Electrical Engineering, Imperial College of Science, Technology and Medicine, England, in 1987. Then, he joined TechTrend E. & C. Ltd. as an Application Engineer working on micro-supercomputers and CAD/CAM. From 1990 to 1993, he was with the Department of Electronic Engineering at The Hong Kong Polytechnic University as a lecturer. In 1993, he started working for a Ph.D. degree in the Department of Electrical Engineering at the University of Sydney, Australia. He completed his Ph.D. studies in August 1996. He joined the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University again as an Assistant Professor in October 1996, and has become an Associate Professor since February 1999. His current research interests include image processing, pattern recognition, video technology, and digital TV.

About the Author—WAN-CHI SIU received the Associateship from The Hong Kong Polytechnic University (formerly called the Hong Kong Polytechnic), the M.Phil. degree from The Chinese University of Hong Kong and the Ph.D. degree from the Imperial College of Science, Technology and Medicine, London, in 1975, 1977 and 1984, respectively. He was with the Chinese University of Hong Kong between 1975 and 1980. He then joined The Hong Kong Polytechnic University as a Lecturer in 1980 and is the Chair Professor since 1992. He took up administrative duties as Associate Dean(Eng) of Engineering Faculty between 1992–1994 and Head of Department of Electronic and Information Engineering between 1994–2000. Prof. Siu has been the Dean of Engineering Faculty of the same university since September 2000. Prof. Siu has published over 200 research papers. His research interests include digital signal processing, fast computational algorithms, transforms, video coding, computational aspects of image processing and pattern recognition, and neural networks.

Prof. Siu is a Member of the Editorial Board of the Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, and an overseas member of the Editorial Board of the IEE Review. He was a Guest Editor for a Special Issue of the IEEE Transactions on Circuits and Systems, Pt.II, published in May 1998. He was also an Associate Editor of the IEEE Transactions on Circuits and Systems, Pt.II between 1995–1997. Prof. Siu was the General Chairman of the International Symposium on Neural Networks, Image and Speech Processing (ISSIPNN'94), and a Co-Chair of the Technical Program Committee of the IEEE International Symposium on Circuits and Systems (ISCAS'97) which were held in Hong Kong in April 1994 and June 1997, respectively. Prof. Siu is the General Chair of the 2003 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) which is to be held in Hong Kong. Between 1991 and 1995, Prof. Siu was a member of the Physical Sciences and Engineering Panel of the Research Grants Council (RGC), Hong Kong Government, and in 1994 he chaired the first Engineering and Information Technology Panel to assess the research quality of 19 Cost Centers (departments) from all the universities in Hong Kong.

Prof. Siu is a Honorary Professor of the South China University of Technology, China. He is a Chartered Engineer, a Fellow of the IEE and the HKIE, and a Senior Member of the IEE, and has also been listed in Marquis Who's Who in the World, Marquis Who's Who in Science and Engineering and other citation biographies.