Face Detection using Color and Symmetry Information

A.H.M. Sajedul Hoque, Md. Moynul Haque Bhuiyan and Md. Al-Amin Bhuiyan

Abstract—The localization of human faces in digital images is a fundamental step in the process of face recognition. This paper presents a face detection method based on color, local symmetry and geometry information of human face. The algorithm first detects most likely face regions or ROIs (Region- Of-Interest) from the image using skin color model and face outline model and produces a face color similarity map. Then it performs local symmetry detection within these ROIs to obtain a local symmetry similarity map. These two maps are fused to obtain potential facial feature points. Finally similarity matching is performed to identify faces between the fusion map and face geometry model under affine transformation. The output results are the detected faces with confidence values. Experimental results demonstrate its validity and robustness to identify faces under certain variations.

Index Terms— Face detection, Face color symmetry, Local symmetry information

1 INTRODUCTION

FACE detection is concerned with finding whether or not there are any faces in a given image and, if present, returns the image location and content of each face. This is the first step of any fully automatic face recognition system that analyzes the information contained in faces (e.g., identity, gender, expression, age, race and pose).

Face detection is a key problem in face and facial expression recognition. It has become a popular area of research due to its emerging applications in human-computer interface, surveillance systems, secure access control, video conferencing, financial transaction, forensic applications, image database management systems and so on. Various approaches to face detection and facial feature extraction have been reported in literature over the last few decades, ranging from the geometrical description of salient facial features to the expansion of digitized images of the face on appropriate basis of images [1]. Different techniques have been introduced recently, for example, principal component analysis [2], neural networks [3], color analysis [4] and so on. Face detectors based on Markov random fields and Markov chains [5] make use of the spatial arrangement of pixel gray values. Color based approaches reduce the search space in face detection algorithm. The neural network-based approaches require a large number of face and non-face training examples, and are designed primarily to locate frontal faces in grayscale images.

Different approaches to face detection can be classified into two categories. The first is called feature-based approach where the face is located by first locating some of its important features. Once the features are identified, the overall location of face is determined using face geometric information[6-9]. Its main disadvantage is, under different imaging conditions, it is difficult to detect those facial features robustly and reliably. In second type of approaches, the face is examined as a whole, usually using model based vision techniques. Model based approaches assume that the initial location of the face is known. The models used often involve color, geometric shape, motion information etc. Actually, the model-based approach is simpler than feature-based approach in implementation and the face detected is less reliable than feature-based approach in implementation and the face detection is less reliable than feature-based approach if a simple model is selected, or it will take much longer time if a complex model is employed. We believe that if we combine the feature-based method and the model-based method, and use a coarse-to-fine processing procedure, a better detection result should be achieved.

This paper explores a face detection system which integrates the detection of human faces in complex backgrounds and localization of facial features on it. The approach is based on the combination of the feature-based method and the model-based method. The algorithm first detects ROIs potentially containing faces from the image using face color model and face outline model, producing a face color similarity detection within

- *A.H.M Sajedul Hoque is with the Department of Computer Science & Engineering, Northern University Bangladesh. E-mail: sajidiuk@yahoo.com*
- *Md. Moynul Haque Bhuiyan is with the Department of Physics, Eden Mohila College, Dhaka. E-mail: mhbhuiya@gmail.com.*
- *Md. Al-Amin Bhuiyan is with the Department of Computer Science & Engineering, Jahangirnagar University, Savar, Dhaka – 1342, Bangladesh. E-mail: alamin@juniv.edu.*

those ROIs detected by color segmentation to get a local symmetry similarity map, fuses these two maps to obtain potential facial feature points. Finally, it performs similarity matching and identifies faces based on face geometry model under affine transformation. The output results are the detected faces associated with the confidences and descriptions. Experimental results indicate that the system is capable of detecting and locating the face parts from complex backgrounds with a high degree of variability in expression, pose, and facial details.

2 FACE DETECTION METHODOLOGY

Face detection is concerned with determining which part of an image contains face. This is the first step of face recognition which requires both high and low-level visual and geometric information processing. Fig. 1 illustrates the information processing flow of the whole system.

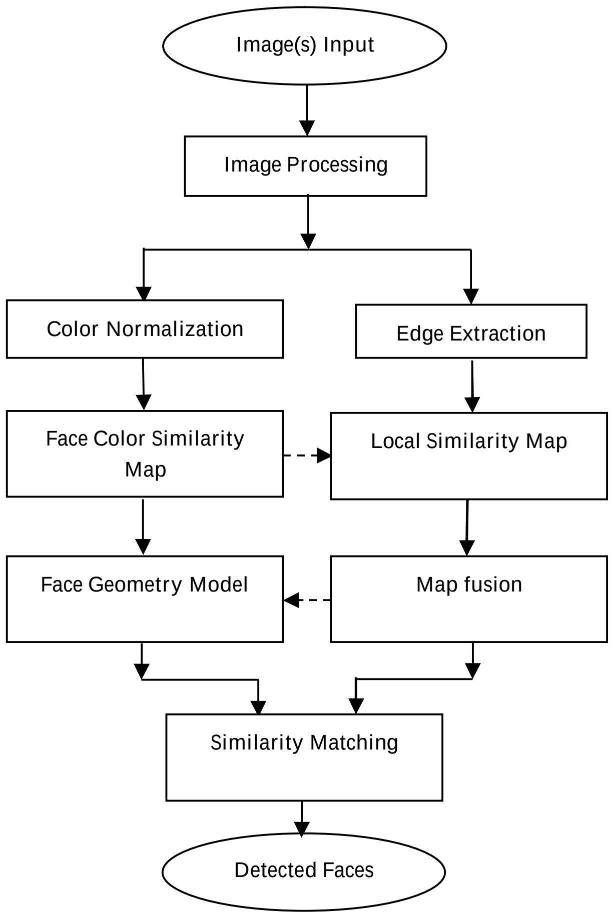


Fig. 1 Organization of the System

First, preprocessing is essential to reduce the influences caused by noise and lighting. It is followed by two parallel processing channels: color similarity map formation and local symmetry similarity map formation. In color similarity map formation, face Gaussian color model and normalized color RGB values are employed to compute a distance map. Where at each pixel the distance value represents the possibility of the pixel belonging to a face region. On this distance map, a threshold operation is performed to identify the ROIs over the potentially containing faces. Then the region clustering in L-a-b color space and region grouping are conducted within those ROIs. Each of resulted regions is assigned a confidence value according to the shape of the region. The more elliptic the region is, the higher confidence value it is assigned. At last, this confidence value is refined on each pixel according to the possibility that the pixel belongs to a facial feature region.

In local symmetry similarity map formation, local symmetry detection on Gaussian edge image within the ROIs is performed. To ensure a consistent local symmetry similarity value on each pixel, the size of the local symmetry filter is chosen according to the size of those ROIs. Normally the potential facial features will have a higher local symmetry value.

The two similarity maps from its respective processing channels are fused by a fuzzy-like operation to form a new fusion map. The pixel with high value in this fusion map shows high possibility being a facial feature. The similarity matching is to match between the fusion map and the face geometry model under affine transformation. After these processing steps, the faces associated with the confidence values are detected and some necessary descriptions about those detected faces are also obtained.

3 IMAGE PREPROCESSING

The imaging conditions of the face detection process may vary from time to time or from place to place. These changes include changes in scale, position and orientations of the faces as well as the lighting condition. Therefore various kinds of image processing operations are needed. The fundamental steps employed for the image processing operations, as shown in Fig. 2, are described as follows.

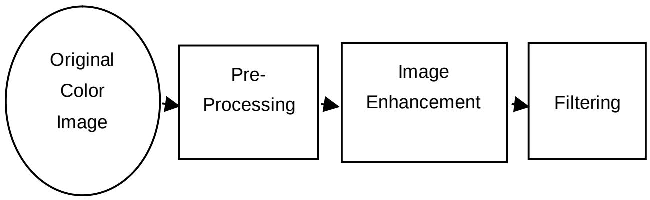


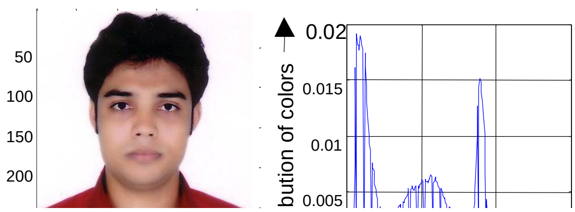
Fig. 2 Fundamental steps employed for image processing for face detection.

3.1 Pre-processing

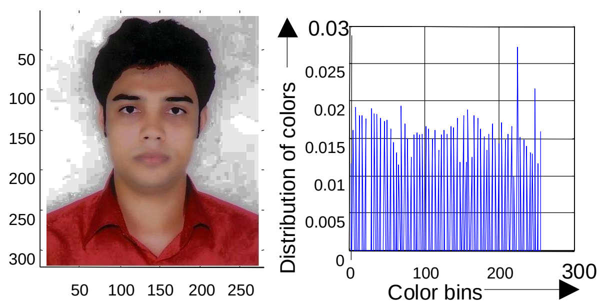
The original image is obviously a color image. It is first converted into gray scale image. While registering images, the eyes, tip of the nose, and the corners of the mouth of each face is labeled. These points are then used to normalize each face to same scale, orientation and position. The normalization is performed by mapping the facial features to some fixed locations in an $M \times N$ image. Each normalized image is then subjected to some image processing operations to account for different lighting conditions and contrast.

3.2 Image Enhancement

The face images may be of poor contrast because of the limitations of the lighting conditions. So histogram equalization is used to compensate for the lighting conditions and improve the contrast of the image [10]. Let the histogram $h(r_i) = \frac{p_i}{n}$ of a digital face image consists of the color bins in the range $[0, C-1]$, where r_i is the *i-th* color bin, p_i is the number of pixels in the image with that color bin and *n* is the total number of pixels in the image. For any *r* in the interval [0,1], the cumulative sum of the bins provides with some scaling constant. Histogram equalization is performed by transforming the function $s = T(r)$, which produces the mapping with the allowed range of pixel values, i.e., a level *s* for every pixel value *r* in the original image and $0 \le T(r) \le 1 \text{ for } 0 \le r \le 1$, as shown in Fig. 3.



(a) A face image (b) Histogram of the face image



(c) Histogram equalized image (d) Histogram of the equalized image

Fig. 3 Histogram equalization of a face image.

3.3 Filtering

Various sources of noise may exist in the input image. The fine details of the image represent high frequencies which mix up with those of noise. So low-pass filters are used to obliterate some details in the image. In this experiment, Prewitt filter is used to suppress the noise.

4 Color Similarity Map

Although skin colors of different people vary over a wide range in color space, the variation of human face color with respect to hue and saturation is much less than that of brightness. Furthermore, when the skin color values are normalized, the distribution of face skin color values follow a 2D Gaussian distribution function even in the case of different races. Another advantage of color normalization is that it can reduce the lighting effect because the normalization process is actually a brightness elimination process. For input images represented in RGB color space, the normalized values are using Eq. 1.

$$r = \frac{R}{R+G+B}, \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \tag{1}$$

where *R*, *G*, *B* are the three primary colors, red, green and blue and *r*, *g*, *b* are normalized red, green and blue color values, respectively. Depending on the normalized color values, a face color model represented by a Gaussian model $N(m, \Sigma^2)$ is obtained in advance, where $m = (\bar{r}, \bar{g})$ with

$$\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i, \quad \bar{g} = \frac{1}{N}\sum_{i=1}^{N} g_i, \quad \Sigma = \begin{bmatrix} \sigma_{rr} & \sigma_{rg} \\ \sigma_{gr} & \sigma_{gg} \end{bmatrix} \tag{2}$$

where *N* is the total number of pixels measured in face regions, and σ is the covariance of the 2D Gaussian distribution model.

Thus a distance map is obtained by measuring the distance between the normalized pixel color value and the Gaussian face color model. Fig. 2 illustrates the brighter pixel area meaning that it is more likely to locate in a face area. Since an arbitrary image may make it impossible to distinguish faces from background reliably only by normalized color values, a further color clustering process is essential. Assuming that in a particular image, the pixels within a face has consistent color values in L-a-b color space, especially in it's a and b dimensions. The color clustering process is done in L-a-b color space and the initial cluster centres are assigned at the peaks of the histograms. Thus we can obtain the results of consistent region partitions and treat each as a potential face region.

The process of forming color similarity map is actually a process of weighting the pixel values on the color distance map. The pixel with its high value in the color similarity map corresponds to potential facial feature regions. The weighting operation is done three times. At the first time, the weighting is determined according to the following rule:

Rule 1: $$D_1 = \begin{cases} 0.5+\alpha_1(d-T_1), & \text{if } d \geq T_1, \\ \alpha_2(T_1-d), & \text{if } d < T_1 \end{cases} \quad (3)$$

Where T_1 is a threshold, and $\alpha_1 \geq 1, \alpha_2 \leq 1$, d is the distance value to face color model. This rule is applied to each ROIs, so even the face regions with different races should get almost the same weights at this step. Here we obtain a segmented binary image by T_1. After region grouping operation (Rule 2), this binary image has two functions. First it can be used to guide the region of local symmetry operation to shorten the processing time. Second, we can adjust the values of color similarity map again based on the shapes of the grouped regions in that binary image. A higher value is assigned if it is more elliptic.

Rule 2: Merge the face and non-face regions if

$$D_{merge} < T_2,\ D_{merge} = \alpha_1 d_{color\ value} + \alpha_2 d_{special\ position} \quad (4)$$

Finally, in each grouped region, we re-assign the color similarity value to each pixel mainly based on its distance to face color model in L-a-b space (Rule 3) and a final color similarity map is obtained. In real implementation, considering the fact there may exist errors in ellipse fitting, and furthermore, face region may be occluded by eyeglasses, beard, or moustache, once the region is fitted by an ellipse, all pixels within this ellipse should be processed as potential face areas. The pixels near the ellipse boundary will also be assigned high similarity values in the way of descending gradually. Now the pixels most likely in facial regions should have high values. This is also an output result of model-based face detection using color and geometry information.

Rule 3: The color similarity value $f_{color} \propto S_1 S_2 S_3$, where S_1 is related to the distance between the normalized color pixel value and Gaussian face color model, S_2 is related to the shape of grouped face-like region, S_3 is related to the distance between L-a-b color value and the current face color model.

5 Local Similarity Map

There are some symmetric attributes on human faces. For instance, the two eyes and two ears are symmetrical about the face middle-line, as well as one nose and one mouth locate at the middle-line. Rather than these global symmetry attributes on human faces, we only focus on the local symmetry attributes on human faces which have been used for facial feature detection. The former approach provided a theoretic measure on local symmetry which both magnitudes and orientations of the image gradients around are utilized, the latter proposed a simplified detection mechanism using the orientations of the gradients only. We propose a new method for local symmetry operator and then scan whole gradient image to get a local symmetry similarity map. To obtain better detection performance our modifications are as follows:

To ensure that all potential facial features can be detected effectively, the size of local symmetry operator is determined by the size of ROI obtained from color segmentation and ellipse fitting, normally 3×3 or 5×5. If the size of the operator needs to be very large, we scale down the image to speed the detection process either. In order to obtain accurate results, real number is used to calculate the orientation of the gradient image although the orientation mask is still and index type (integer). Besides the information of the orientation map, we also utilize the information of the magnitude of image gradient. Similar to the orientation mask operator, a distance-weighted operator is formed at the beginning of local symmetry detection. These two operators are used jointly to detect the attributes of local symmetry within the ROIs.

6 Fusion of Color & Symmetry Information

Once both similarity map and local symmetry similarity map are obtained, the detection process is to fuse these two maps to get a fusion map where each point with high value will be considered as potential facial feature point. The fusion operation actually is a fuzzy-like operator:

After a threshold and grouping operation on the obtained fusion map, we can get potential facial features points. Taking each possible combination, the face geometry model is affinely transformed based on the current hypothesis on facial feature positions. Then matching operation is performed between the transformed face geometry model and the fusion map directly according to the following matching criterion:

$$\varepsilon(map, mdel) = \frac{\sum_{i=1}^{N}\left((map(i)mdel(i)\right)}{\left[\sum_{i=1}^{N}\left(map(i)\right)^2 \sum_{i=1}^{N}\left(mdel(i)\right)^2\right]^{1/2}} \quad (5)$$

where *N* is the total number of matching pixels involved, and map(*i*), model(*i*) are the corresponding pixel values in the fusion map and transformed face geometry model, respectively. This normalized measure is used to adjust the confidence of the detected face. A ROI is identified to be a face if its overall confidence value is higher than some threshold value.

7 Experimental Results & Performance

The effectiveness and robustness of this approach have been justified using different images with various kinds of expressions. Experiments are carried out on a Pentium Dual Core 1.80GH PC with 2 GB RAM. The algorithm has been implemented using Visual C++. When a complex image is subjected in the input, the face detection result highlights the facial part of the image, as shown in Fig. 4. The system can also cope with the problem of partial occlusion of mouth and wearing sunglasses. Images of different persons are taken at their own work places and at different environments both in shiny and gloomy weather. Most of the images are taken using a digital camera, but some are from scanner, and some from video tapes recorded from different television channels. The algorithm is capable of detecting single face in an image. For multiple faces, the system finds the dominant face only. A total of 360 images, including more than 80 different persons, are used to investigate the capacity of the proposed algorithm. Among them only 6 faces are found false. Experimental results demonstrate that the success rate of approximately $98\%\left(\frac{354}{360}\times 100\% = 98.33\%\right)$ is achieved. The main reason behind the failure of those images in finding face regions is the occlusion. Face detection is performed by using both color and grayscale modes. The detection result is summarized in Table 1.

TABLE 1
EXPERIMENTAL RESULTS

Image Condition	No of Images	Correctly Detected	Accuracy (%)
Bright Sunny	120	119	99.2%
Foggy Weather	120	117	97.5%
Dark Environment	120	118	98.3%

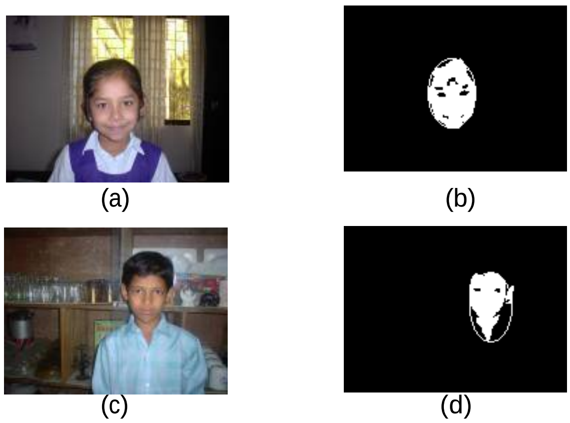


Fig. 4 Face detection for the children at their own environment: (a)-(c): original image, and (b)-(d): detected face image.

7 CONCLUSION

Detection of human faces is a problem that appears time. Although human beings accomplish these tasks countless times a day, they are still very challenging for machine vision. Most of the researchers attack this kind of problem with face localization and feature selection with frontal view faces and without facial expression and normal lighting conditions although the variation between the images of the same face is too large due to facial expression, hair style, pose variation, lighting conditions, make-up, etc. In this paper, face detection has been implemented using color symmetry algorithm to search for the face of a particular individual in an image. The effectiveness of the face detection algorithm has been tested both in simple and complex backgrounds for different types of face and non-face images of 320×240 resolution. The algorithm is capable of detecting the faces in the images with different backgrounds and lighting conditions. Our next approach is to extend the algorithm for multi-face detection and overlapping faces in images and to detect facial poses.

REFERENCES

[1] H.E. Low and B. Kee, "Face detection: a survey", *Computer Vision and Image Understanding,* vol. 83, no. 3, pp. 236-274, 2001.
[2] M. Turk and A. Pentland, "Eigenfaces for recognition", *Journal of Cognitive Neuroscience,* vol. 3, no. 1, pp. 71-86, 1991.
[3] H. Rowley, S. Beluga and T. Kanade, "Neural network-based face detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 20, no. 1, pp. 23-37, 1998.
[4] A.A. Bhuiyan, V. Ampornaramveth, S. Muto, and H. Ueno, "Face detection and facial feature localization from human-machine interface", *NII Journal,* vol. 5, no. 1, pp. 25-38, 2003.
[5] W. Freeman, E. Pasztor and O. Carmichael, "Learning low level vision", *International Journal of Computer Vision,* vol. 40, no. 1, pp. 25-47, 2008.
[6] K.C. Yow and R. Cipolla, "Feature-based human face detection", *Technical Report,* no. 249, University of Cambridge, 1996.
[7] C. Lin and W. Lin, "Extracting facial features by an inhibitory mechanism based on gradient distributions", *Pattern Recognition,* vol. 29, no. 12, pp. 2079-2101, 1996.
[8] A.A. Bhuiyan and H. Hama, "Identification of actors drawn in Ukiyoe pictures", *Pattern Recognition,* vol. 35, no. 1, pp. 93-102, 2002.
[9] P. Viola and M.J. Jones, "Robust Real-Time Face Detection", *International Journal of Computer Vision,* vol. 57, no. 2, pp. 137–154, 2004.
[10] J. Wang, and T. Tan, "A new face detection method based on shape information", *Pattern Recognition Letters,* vol. 21, no. 6, pp. 463-471, 1999.