Facial Feature Extraction and Principal Component Analysis for Face Detection in Color Images

Abstract. A hybrid technique based on facial feature extraction and Principal Component Analysis (PCA) is presented for frontal face detection in color images. Facial features such as eyes and mouth are automatically detected based on properties of the associated image regions, which are extracted by RSST color segmentation. While mouth feature points are identified using the redness property of regions, a simple search strategy relative to the position of the mouth is carried out to identify eye feature points from a set of regions. Priority is given to regions which signal high intensity variance, thereby allowing the most probable eye regions to be selected. On detecting a mouth and two eyes, a face verification step based on Eigenface theory is applied to a normalized search space in the image relative to the distance between the eye feature points.

Keywords: face detection, facial feature extraction, PCA, color segmentation, skin detection

1 Introduction

Face detection is an important task in facial analysis systems in order to have a priori localized faces in a given image. Applications such as face tracking, facial expression recognition, gesture recognition, etc., for example, have a pre-requisite that a face is already located in the given image or the image sequence. Numerous face detection techniques have been proposed to address the challenging issues associated with this problem in the literature. These techniques generally fall under four main categories of approach: knowledge-based, feature invariant, template matching, and appearance-based [1]. Some algorithms rely solely on low-level image properties such as color and image contours from which image blobs are detected and compared with predefined shapes (elliptical shape) [1][2]. Combining facial features, which are detected inside the skin color blobs, helps to extend the above type of approach towards more robust face detection algorithms [3][4]. Facial features derived from gray scale images along with some classification models have also been used to address this problem [5]. Menser and Muller presented a method for face detection by applying PCA on skin tone regions [6]. Using the appearance-based properties in more efficient ways to classification, upright frontal face detection in gray scale images through neural networks has proved to be a promising solution to this problem [7]. Chengjun Liu proposed a face detection technique based on discriminating feature analysis, statistical modeling of face and non-face classes, and a Bayes classifier to detect frontal faces in gray scale images [8]. Existing face detection algorithms in the literature, however, all indicate that different levels of success have been achieved with varying algorithm complexities and detection performance.

In this paper, we present a hybrid approach for frontal face detection in color images based on facial feature extraction and the use of appearance based properties of face images. This is followed by the face detection algorithm proposed by Menser and Muller, which attempted to localize the computation of PCA on skin-tone regions. Our approach begins with a facial feature extraction algorithm which illustrates how image regions segmented using chrominance properties can be used to detect facial features, such as eyes and mouth, based on their statistical, structural, and geometrical relationships in frontal face images. Applying statistical analysis based on PCA to a smaller search space (a normalized search space) of the image then performs the face detection task.

The block diagram of the proposed system is shown in Fig. 1. The first task of face detection in this system is skin detection which is carried out using a statistical skin detection model built by acquiring a large training set of skin and non-skin pixels. A skin map is generated through direct reference to the pre-computed probability map in the RGB space and using a simple threshold criterion. A face-bounding box is then obtained from the skin map to which the RSST segmentation algorithm is applied for creating a segmentation partition of homogeneous regions. Possible mouth features are first identified based on the redness property of image pixels and the corresponding RSST regions. Eye features are then identified relative to the position of the mouth, by searching for regions which satisfy some statistical, geometrical, and structural properties of the eyes in frontal face images. On detecting a feature set containing a mouth and two eyes, PCA analysis is performed over a normalized search space relative to the distance between the two eyes. The image location corresponding to the minimum error is then considered the position of the detected face.

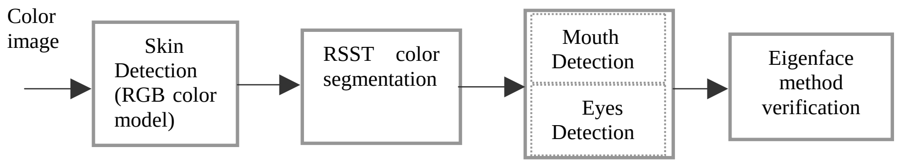


Fig. 1. Face detection system

The paper is organized as follows. In section 2, a region-based facial feature extraction technique is described. Starting with a brief description of the RSST segmentation algorithm in section 2.1, the mouth detection and eye detection tasks are described in sections 2.2 and 2.3 respectively. The combined approach to face detection using facial features and PCA analysis is described in section 3. Experimental results are presented in section 4. Some conclusions and future work are then given in section 5.

2 Facial Feature Extraction

A general view of a frontal face image containing a mouth and two eyes is shown in Fig. 2. E_l and E_r represent left and right eyes respectively, while M represents the mouth feature. The distance between the two eyes is w, and the distance from the mouth to the eyes is h. In frontal face images, structural relationships such as the Euclidean distance between the mouth, and the left and right eye, the angle between the eyes and the mouth, provide useful information about the appearance of a face. These structural relationships of the facial features are generally useful to constrain the facial feature detection process. A search area represented by the square of size (3w x 3w) is also an important consideration in order to search for faces based on the detected eye feature positions in the image.

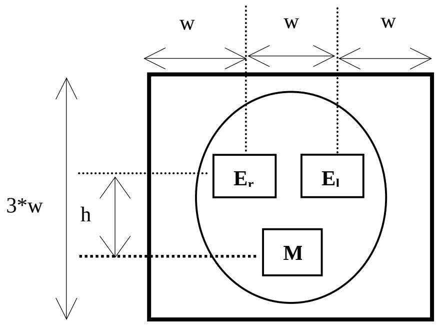


Fig. 2. A frontal face view

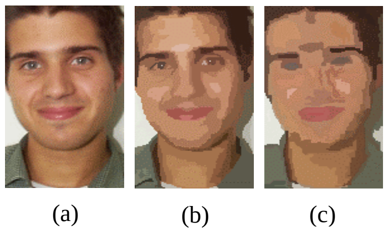


Fig. 3. RSST color segmentation (a) face-bounding box (b) segmentation based on luminance and chrominance merging (c) segmentation based on chrominance merging

2.1 Recursive Shortest Spanning Tree (RSST) Color Segmentation Algorithm

The process of region merging in the conventional RSST algorithm is defined by the merging distance that is related to both the luminance and chrominance properties in the YC_bC_r color space. However, it was found in our experiments that the eye detection task could be performed using only the chrominance components in the merging distance. The conventional RSST merging distance and the modified merging distance are defined by equations (1) and (2) respectively. The distance d(R1,R2) represents the merging distance between the two regions R1 and R2 with their mean luminance, mean chrominance, and spatial size represented by Y(R), C_b(R), C_r(R) and N(R) respectively. Two segmentations shown in Fig. 3b and Fig. 3c are obtained from these two distance measures when RSST is performed on the face-bounding box shown in Fig. 3a. This highlights the fact that distinct eye regions can be obtained more accurately from the chrominance-based merging.

$$d(R1,R2)=\left\{[Y(R1)-Y(R2)]^2+[C_b(R1)-C_b(R2)]^2+[C_r(R1)-C_r(R2)]^2\right\}\times\frac{N(R1)\times N(R2)}{N(R1)+N(R2)} \tag{1}$$

$$d(R1,R2)=\left\{[C_b(R1)-C_b(R2)]^2+[C_r(R1)-C_r(R2)]^2\right\}\times\frac{N(R1)\times N(R2)}{N(R1)+N(R2)} \tag{2}$$

2.2 Mouth Detection

The mouth detection task is performed based on the redness property of lips (mouth). After extracting a face-bounding box from the skin detection process, the red color lips are detected using the criterion defined in equation (3), and represented as a mouth map [4].

$$MouthMap = C_r^2 \cdot (C_r^2 - \eta \cdot C_r / C_b)^2 \tag{3}$$

$$\eta = 0.95 \times \frac{(1/N)\sum Cr^2}{(1/N)\sum (Cr/Cb)}$$

where N represents the spatial size of the face-bounding box.

Regions in the segmentation that correspond to the detected mouth map are first identified. In case of the presence of multiple regions, they are then merged based on their proximity and represented as a single mouth map. The center of gravity of the combined regions is then considered to be the mouth feature position.

2.3 Eye Detection

An important statistical property of eye image regions is that they correspond to high intensity variance as a result of the fact that human eyes generally contain both black (near black) and white (near white) regions. Such regions can be identified by computing their variances using equation (4). This principle is illustrated in the graph shown in Fig. 4, which shows the distribution of different regions' variances for the segmentation given in Fig. 3c against their 2D positions in the original image. This shows the important feature that only a few regions show high variance.

$$Variance = (1/N)\sum (Y - \overline{Y})^2 \tag{4}$$

where Y represents the intensity value of each pixel in the region, $\overline{Y}$ is the mean intensity of the region, and N is the spatial size of the region.

However, in practice, a variance measure alone will not be sufficient to find eye regions in a given image, although it can provide some useful clues. This fact leads us to constrain the eye search process by relating it with geometrical and structural properties of eye regions in frontal face images. Hence, the following heuristic rules are applied in the eye detection process (geometrical parameters are given with reference to Fig. 2).

- Eye region should be at least 10 pixels above the mouth level, i.e. h>=10 pixels.
- Width/height ratio of eye regions should be at least 0.4.
- Distance from the mouth to the left and right eyes should be within a pre-defined range, i.e. $1.4 \times ME_l >= ME_r >= 0.6 \times ME_l$.
- Angle between the mouth and the eyes should be within a pre-defined range, i.e. 35 degrees $<= \angle E_l ME_r <=$ 80 degrees.
- Eye region should correspond to a dark blob in the image.

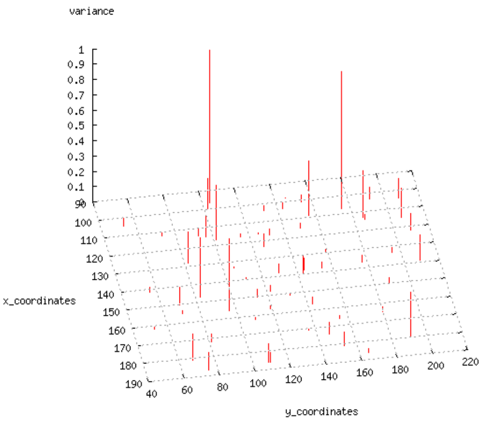


Fig. 4. Variance distribution of regions corresponding to a face segmentation containing 61 image regions

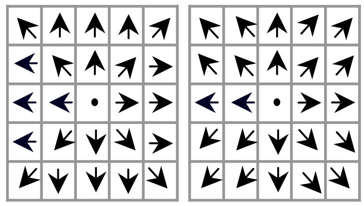
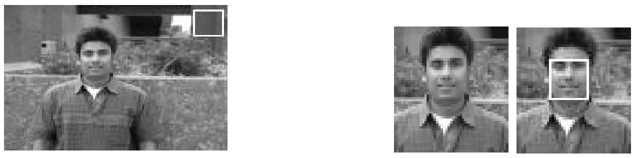


Fig. 5. Gradient orientation of pixels for 5X5 masks

While the first four conditions are simple heuristic rules and require no description, the fifth condition uses the feature of dark blobs (corresponding to pupils) present in human eyes. A dark/bright blob detection technique, specifically proposed for facial feature detection by Lin and Lin [9], is used for identifying dark blobs (gradients pointing away from dark to bright) in this system. It computes the radial symmetry about a center pixel considering the gradient distributions of local neighbors of a pre-defined mask. Fig. 5 shows the possible orientations of pixel gradients with respect to a 5X5 mask. Lin and Lin pointed out that the algorithm produces quite dense radially symmetric feature points at this stage, most of them corresponding to spurious or non-facial feature points. Thus, two inhibitory mechanisms are used to suppress the spurious points: regions of uniform gradient distribution should not be considered, and a line sketch of the input image should be used to eliminate areas which are not covered by sketch lines. However, we can avoid the use of these two additional inhibitory conditions by applying the algorithm on C_r chrominance image rather than on luminance image.

3 Combining facial features with appearance-based face detection

A difficulty of using PCA as a face classification step is due to the inability of properly defining an error criterion on face images. As a result of this, inaccurate face images can signal a smaller error, resulting in an image block with the minimum error converging to a wrong face image. However, when the search space is reduced to a smaller size, the intended results can be achieved in most cases. This phenomenon is illustrated in Fig. 6.



PCA on the original image PCA on a cropped image

Fig. 6. Detected faces using PCA alone

Due to PCA being a sub-optimal solution for face classification, we use an approach to perform face detection based on the use of both facial features and PCA. The objective of using facial features in this system is to localize the image area on which the PCA analysis is performed. In this context, the detected eye facial feature points are used to define a normalized search space. A search area is first defined according to the eye-eye distance, and it's of size 3w x 3w (see Fig. 2). A scale factor is then calculated according to the positions of the detected eye feature points and the eye-eye distance of the predefined face model (8 pixels). Hence, a normalized image search space is obtained on which PCA is performed based on 16X16 image blocks. It should be noted that the use of a normalized search space in this system avoids the requirement of analyzing the image at different resolution levels to locate faces of different scales.

The principal theory of detecting faces in images using Eigenfaces is defined by how far a reconstructed image is from the face space. This distance, known as Distance from Face Space (DFFS), is defined by equation (5) [6][10].

$$DFFS = \| x - \bar{x} \|^2 - \sum_{i=1}^{M} y_i^2 \quad \textbf{(5)} \qquad\qquad DIFS = \sum_{i=1}^{M} \frac{y_i^2}{\lambda_i} \quad \textbf{(6)}$$

where x is the current image vector, $\bar{x}$ is the mean face vector, and y_is are the M principal components corresponding to the M eigenvectors.

Menser and Muller also used a modified error criterion (e) by incorporating another distance measure called “Distance in Face Space (DIFS)”, which is defined by equation (6). This is used to obtain a measure of the difference between the reconstructed test image and the mean face image. The modified error criterion is defined by equation (7) [8].

$$e = DIFS + c \times DFFS \tag{7}$$

where c is a constant which is defined as $c = \frac{1}{k\lambda_M}$. The smallest computed eigenvalue, represented by λ_M, and an empirically chosen constant k, are the two parameters that define the value of constant c.

In our system, a total number of 500 face images were used as the set of training images – taken from the ECU database [11]. The set of 500 training images was obtained by doubling the original number of images with their mirror images. The 250-image training set is a collection of 100 frontal upright images, 50 frontal images with glasses, and 100 slightly rotated face images selectively chosen from the second set of face patterns of the ECU database.

4. Experimental Results and Analysis

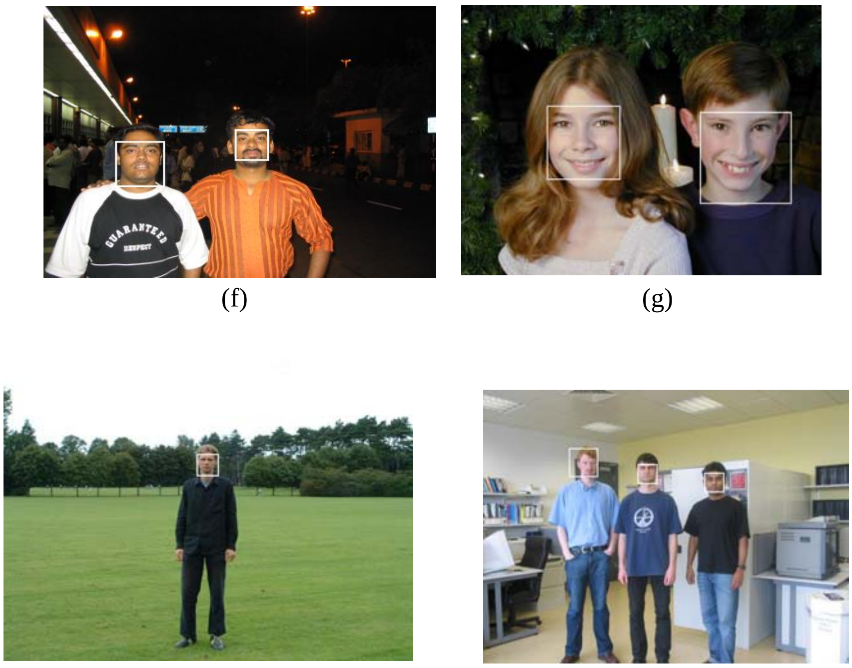
Fig. 7 shows some examples of the detected facial features superimposed on the original images. Mouth feature points are shown by green crosses while eye feature points are shown by blue crosses. Fig. 7a is a bright frontal face image whereas Fig. 7b is a frontal face image with glasses causing bright reflections. A half frontal face image is shown in Fig. 7c. A randomly selected image of two faces is shown in Fig. 7d. Accurate results are reported in the first, third, and the fourth image while slightly inaccurate eye feature points have been detected in the second case. Bright reflections caused by glasses in the second image have led to errors in the eye detection process.

Face detection results are shown in Fig 8. Our experiments were carried out on test images taken from the HHI face database and various other randomly selected sources. We noted that the slight inaccuracies occurred in the facial feature extraction process did not affect the performance of face detection. However, we observed in a few cases that, despite being able to detect facial features accurately, the precise face location was not found (see Fig. 8e for example). The reason for this is because PCA does not always guarantee that the minimum error coincides with the exact location of the face image.

(a) (b) (c) (d)

Fig. 7. Facial feature extraction results

(a) (b) (c) (d) (e)



(f) (g)

(h) (i)

Fig. 8. Face detection results

5 Conclusion and future work

A hybrid solution to frontal face detection using facial features and Eigenfaces theory is presented. Using a facial feature extraction step prior to performing PCA analysis helps to address two requirements for this system. Firstly, the search for faces does not need to be carried out at every pixel location in the image since a small search space can be obtained using the detected facial feature points. Secondly, the face detection process can be carried out in one cycle over a normalized search space, thereby avoiding the requirement of processing the image at multiple scales. However, due to the fact that PCA is a sub-optimal solution to face classification, detection of inaccurate facial features in poor quality images results in performance degradation in the system. For this reason, we believe that the performance of this system can be improved by extending the face classification step towards a two-class classification problem with the use of a carefully chosen set of non-faces as the second class.

Acknowledgment

This work is carried out as part of the MobiSyM project funded by Enterprise Ireland. Their support is gratefully acknowledged.

References

1. Ming-Hsuan Yang, David J. Kriegman and Narendra Ahuja, "Detecting Faces in Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 24, no. 1, pp. 696-706, January 2002.
2. A. Eleftheriadis and A Jacquin, "Automatic Face Location, Detection and Tracking for Model Assisted Coding of Video Teleconferencing Sequences at Low Bit Rates", Signal processing: Image Communication, vol. 7, no. 3, pp. 231-248, July 1995.
3. Karin Sobottka and Ioannis Pitas, "A Novel Method for Automatic Face Segmentation, Facial Feature Extraction and Tracking", Signal Processing: Image Communication, vol. 12, no.3, pp. 263-281, 1998.
4. Rein-Lien Hsu, Mohamed Abdel-Mottaleb and Anil K. Jain, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 696-706, May 2002.
5. Kin Choong Yow, Roberto Cipolla, "Feature-Based Human Face Detection", Image and Vision Computing, vol. 15, pp. 713-735, 1997.
6. B. Menser and F. Muller, "Face Detection in Color Images using Principal Component Analysis", IEE Conference Publication, vol.2, no. 465, pp. 620-624, 1999.
7. Henry A. Rowley, Shumeet Baluja and Takeo Kanade, "Neural Network-Based Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, January 1998.
8. Chengjun Liu, "A Bayesian Discriminating Features Method for Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, January 1998.
9. Cheng-Chung Lin and Wei-Chung Lin, "Extracting Facial Features by an Inhibitory Mechanism based on Gradient Distributions", Pattern Recognition, vol. 29, no. 12, 1996.
10. Matthew Turk and Alex Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, vol.3, no. 1, pp. 71-86, 1991.
11. Son Lam Phung, Douglas Chai and Abdesselam Bouzerdoum, "Adaptive Skin Segmentation in Color Images", International Conference on Acoustic, Speech and Signal processing (ICASSP), vol. 3, pp. 353-356, April 2003.