Facial Features and Appearance-Based Classification for Face Detection in Color Images*

Saman Cooray#, Noel O'Connor##

School of Electronic Engineering, Dublin City University, IRELAND
email:coorays@eeng.dcu.ie

##Centre for Digital Video Processing, Dublin City University, IRELAND
email: oconnorn@eeng.dcu.ie

***Abstract* - A technique is presented for frontal face detection in color images based on facial feature extraction and appearance-based classification. Salient facial features are used to define a search space that is then used in a classification step in order to find the best position of the face in the image. Mouth feature points are identified using the redness property of image pixels whilst eye feature points are detected using a search strategy applied to a subset of regions in a fine region-based segmentation of the candidate face. Face class modeling based on a multivariate normal distribution and discriminating feature analysis is used as the face classification method. The utilization of facial features in this system avoids analyzing the image at every pixel location as well as at multiple scales when detecting faces of different sizes.**

Keywords: Facial feature extraction, face detection, multivariate normal distribution, PCA, skin detection.

I. INTRODUCTION

Face recognition, facial expression analysis, gesture recognition, and multimedia indexing and retrieval are just some application areas that have stimulated research in the field of face detection. When dealing with real-life images, robust face detection can be an extremely challenging task. To date, solutions proposed generally fall under four main categories: knowledge-based, feature invariant, template matching, and appearance-based [1]. Low-level image features such as color and shape have been used for face detection with a certain degree of accuracy and robustness. Color is used to detect skin blobs, thereby enabling a process of finding blobs shaped like a human face [1][2]. Other approaches [3][4], avoid the use of shape by considering facial features. In fact, some approaches also avoid the use of color using facial features derived from grey-scale images [5]. Template matching, although relatively straightforward, can have drawbacks due to its inability to properly accommodate real-world faces [6] and whilst deformable templates are a solution to scale invariance, the initial positioning of such templates must be known a priori [6]. Richard and Thomas [7] proposed a probabilistic model for face detection comprising template matching and facial features whereby scale variations are addressed using image pyramids and orientation variations are addressed using a relatively small number of templates and feature models. Menser and Muller [8] presented an appearance-based technique based on PCA analysis and skin detection. A neural network-based approach to upright frontal face detection in grey-scale images has also proven to be a promising solution [9]. Face detection using a two-class classification model was proposed by C. Liu to detect multiple frontal faces in grey-scale images [10].

In this paper, we present a technique to detect frontal faces in color images using facial features and statistical classification. Skin detection is first performed to derive a skin map and associated bounding box

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

from the given color image using a statistical skin detection model which was constructed using a larger number of skin and non-skin pixels. A facial feature extraction technique based on the use of color image regions is then used to detect mouth and eye feature points within the candidate face. A color segmentation algorithm is applied to the face-bounding box to obtain a set of homogenous image regions, and detection of mouth feature points is carried out using the redness of regions. Eye feature points are detected using statistical, geometrical, and structural properties of eyes in frontal face images. Upon detecting a pair of eyes, a normalized search space is defined relative to the distance between the two eye positions. A statistical face classification step applied to this region is the final step in detecting a face.

The paper is organized as follows. In section II, the facial feature extraction algorithm is described. The color segmentation algorithm is described in section II-A. Mouth and eye detection are described in sections II-B and II-C respectively, whilst face classification is described in section III. Experimental results are presented in section IV, followed by conclusions in section V.

II. FACIAL FEATURE EXTRACTION

Following skin detection, the mouth is first extracted using the redness property of lips, searched over the face-bounding box. The inherent structural property of the frontal view of a face is then exploited to facilitate eye detection.

A. *Recursive Shortest Spanning Tree (RSST) Colour Segmentation Algorithm*

RSST is a region-based colour segmentation algorithm that can be used to segment a given image into a desired number of regions. In the conventional algorithm, the region merging sequence is defined using both the luminance and chrominance properties of regions. However, we noted that the task of eye detection can be better facilitated using only chrominance data, resulting in a modified merging distance. The original and modified merging distances are defined by equations 1 and 2. The distance *d(R1,R2)* represents the merging distance between two regions *R1* and *R2* where their mean luminance, chrominance, and spatial size are represented by *Y(R)*, *Cb(R)*, *Cr(R)* and *N(R)* respectively. Fig. 1(b) and Fig. 1(c) illustrate the different types of regions obtained from these two distance measures. It can be noted that when chrominance only merging is used, eye regions appear as grey blobs and are distinctively separated from other image regions.

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

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

(a) (b) (c)

Fig. 1: RSST Segmentation (a) face-bounding box (b) segmentation from luminance and chrominance (c) segmentation from chrominance merging only

B. *Mouth Detection*

Mouth detection is performed using the redness property which generally exists around the lips. A mouthmap is detected by applying equation 3 over the face-bounding box as proposed by Hsu *et al* [4].

$$Mouthmap = C_r^2.(C_r^2-\eta.C_r/C_b)^2 \qquad (3)$$

where

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

, and N is the spatial size of the face-bounding box. Upon calculating the mouth map, the corresponding RSST regions are combined to form the final mouth map based on their proximity. The centre of gravity of the combined regions is considered to be the mouth feature position.

C. *Eye Detection*

The modified RSST algorithm ensures that dark and bright pixels are quickly merged, and that the resulting regions contain high intensity variance [11]. This condition is generally satisfied with human eyes since they contain both black (near black) and white (near white) regions. The luminance variance within these images, however, cannot be used alone for detection, but it can be combined with the following geometrical and structural constraints:

- An eye region should be at least 10 pixels above the mouth level.
- The width/height ratio of eye regions should be at least 0.4.
- The distance from the mouth to the left and right eyes should be within a pre-defined range.
- The angle between the mouth and the eyes should be between 35 degrees and 80 degrees.

- The mouth feature point should be located in between the eye feature points.
- An eye region should correspond to a dark blob in the image.

The dark/bright blob detection technique proposed by Lin and Lin [12] is used in our system for detecting dark blobs. This technique computes the radial symmetry about a centre pixel considering the gradient distributions of local neighbours of a predefined mask. Lin and Lin pointed out that the algorithm produces quite dense radially symmetric feature points, most of them corresponding to spurious or non-facial feature points and thus inhibitory mechanisms are used to suppress the spurious points. However, we avoid the use of these additional inhibitory conditions by applying the algorithm on the C_r chrominance image rather than on the luminance image. Illustrative results of feature detection are presented in Fig 2.

Fig. 2: Detected mouth and eye facial features

III. STATISTICAL FACE CLASSIFICATION

Face discrimination makes use of face class modelling based on multivariate normal distribution and discriminating feature analysis, which was originally proposed by C. Liu [10] for face detection in grey-scale images. It uses a normalized feature vector derived by combining the normalized input image, its normalized 1D Harr representation, and its normalized amplitude projections [10]. The combination of normalized vectors provides a more enriched image feature vector from which the face class parameters can be derived using PCA. The conditional probability density function of face class, ω_f, can be estimated through a multivariate normal distribution as defined in equation 4.

$$p(Y/\omega_f)=\frac{1}{(2\pi)^{N/2}\left|\Sigma_f\right|^{1/2}}\exp\left\{-\frac{1}{2}(Y-M_f)^t\Sigma_f^{-1}(Y-M_f)\right\} \tag{4}$$

, where the mean and covariance matrices of face class ω_f are denoted by M_f and Σ_f respectively. It was shown in [10] that an error term, δ_f, for the face class considering only the first M principal components can be calculated as:

$$\delta_f = \sum_{i=1}^{M}\frac{z_i^2}{\lambda_i} + \frac{\left\|Y - M_f\right\|^2 - \sum_{i=1}^{M} z_i^2}{\rho} + \ln\left(\prod_{i=1}^{M}\lambda_i\right) + (N-M)\ln\rho \tag{5}$$

where z_i are the principal components of Z defined by equation 7 below, ρ, is the average sum of the remaining eigenvalues defined by equation 6 below, and λ_i are the eigenvalues of the face class.

$$\rho = \frac{1}{N-M}\sum_{k=M+1}^{N}\lambda_k \tag{6}$$

$$Z = \phi_f^t(Y - M_f) \tag{7}$$

The final stage of our face detection system uses δ_f to perform face classification over a normalised search area. Eye feature positions are used to find a (*3d X 3d*) search area in the image, where *d* is the distance between the detected eye feature points in the test image. A normalised search area is then obtained with respect to the eye feature distance of the training image model, which is 8 pixels in our system. We then check if δ_f is less than a predefined threshold to decide if the test image is a valid face or not.

The face training model comprises a total number of 1100 face images taken from the ECU face database [13]. We use 16×16 normalised face images for this task. The training set is a collection of upright frontal faces, faces with glasses, beards (both grey and dark), and slightly rotated face images selectively chosen from the face patterns of the ECU database.

IV. Results and Analysis

Experiments were carried out on test images taken from the HHI color face database, our own test data, and numerous other images downloaded from the Internet. Some example face detection results are shown in Fig. 3, with white squares showing the detected faces.

Table 1. Face classification performance for 125 single face images of correctly identified eye features.

Threshold on δ_f	detected	false rejections/detections
T1=2700	119	0/6
T2=2200	112	7/6
T3=1700	68	57/0

In this research, we are interested in analysing the performance of the above method given accurately detected eye feature points. In this context, we use 125 test images of correctly detected eye features for the analysis and results are provided in Table 1. Face classification performance is presented in terms of correct detections, false rejections and false detections against three different threshold values on δ_f. "Correct detection" is deemed to occur if the square marked in white surrounds all three facial features. If the face is not detected at all, this is deemed a "false rejection". It can be noted that a higher detection rate and a lower false rejection rate are reported for larger threshold values. When the threshold value decreases, the number of correct detections also decreases while the number of false rejections increases. An interesting feature of smaller threshold values is that the number of false detections is reduced. This analysis shows that using a high threshold value helps to obtain a better face detection rate. However, it also indicates that some other measure is required to avoid the increase of false detections which would otherwise result in non-faces being identified as faces.

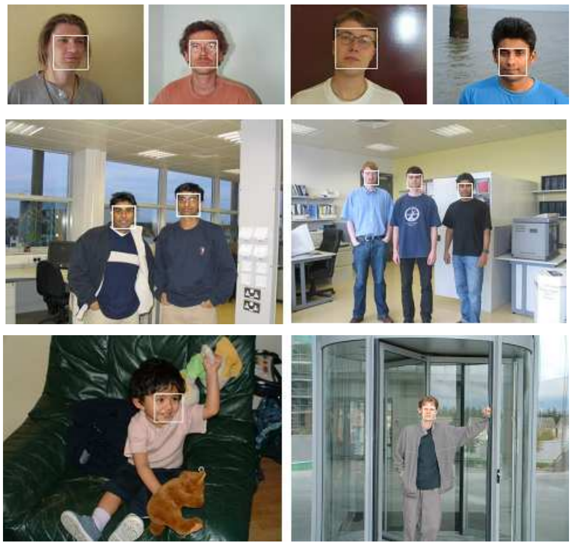

Fig. 3: Face detection results

V. CONCLUSION

We have presented a technique that combines facial feature extraction and statistical face classification to detect frontal faces in color images. The use of eye facial feature points enables us to derive a normalized search space thereby eliminating the requirement of analysing the image at multiple scales for detecting different sized faces.Furthermore, the search for faces needs to be carried out only within the normalized search space, thus the requirement of analysing the image at every pixel location in the image is also eliminated. While the method of face classification employed provides a promising classification solution, the need for face/non-face discrimination knowledge in the classification model can be foreseen and will be the target of future research.

References

[1] M.-H. Yang, D. J. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 1, pp. 34-58, 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] K. Sobottka and I. 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] R.-L. Hsu, M. Abdel-Mottaleb and A. K. Jain, "Face Detection in Color Images", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 696-706, 2002.

[5] K. C. Yow, R. Cipolla, "Feature-Based Human Face Detection", Image and Vision Computing, vol. 15, pp. 713-735, 1997.

[6] V. Govindaraju, "Locating Human Faces in Photographs", International Journal of Computer Vision 19(2), pp. 129-146, 1996.

[7] R. J. Qian and T. S. Huang, "Object Detection using Hierarchical MRF and MAP Estimation," Computer Vision and Pattern Recognition, pp. 186-192, 1997.

[8] 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.

[9] H. A. Rowley, S. Baluja and T. Kanade, "Neural Network-Based Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, 1998.

[10] C. Liu, "A Bayesian Discriminating Features Method for Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, 1998.

[11] S. Cooray and T. Curran, "Region-Based Facial Feature Extraction for Face Detection in Color Images", 1st Indian International Conference on Artificial Intelligence, pp. 600-609, Hyderabad, India, Dec 18-20 2003.

[12] C.-C. Lin and W.-C. Lin, "Extracting Facial Features by an Inhibitory Mechanism based on Gradient Distributions", Pattern Recognition, vol. 29, no. 12, 1996.

[13] S. L. Phung, D. Chai and A. Bouzerdoum, "Adaptive Skin Segmentation in Color Images", International Conference on Acoustic, Speech and Signal processing (ICASSP), vol. 3, pp. 353-356, April 2003.