Improved Face Recognition in the Myra System

Tibor Kövér, Dénes Vígh, Zoltán Vámossy

Institute of Software Technology, John von Neumann Faculty of Informatics, Budapest Tech
Bécsi út 96/B, H-1034 Budapest, Hungary
Phone: +36 1 666 5550, Fax: +36 1 666 5551
Email: vamossy.zoltan@nik.bmf.hu

Abstract: The MYRA system has been created for face detection and face recognition on video images or live webcam video stream. The face recognition part of the program has been significantly improved since the previously published state of the system. With the implementation of the 2D Gabor wavelets based feature extraction method the user can get more accurate recognition results. The system now is able to detect and track not only the face but the characteristic points of the given face also.

Keywords: face recognition, Gabor wavelets

1 Introduction and Aims

The main goals of the MYRA project to develop a system that is capable to detection and recognition faces. The source of the information is a video camera or a webcam. Face detection means that the program is able to find human faces on a video stream. To make the system more reliable it is necessary to use not only one, but more different techniques that are different enough to supplement one another. The system uses two different techniques, an appearance based method and skin tone detection. Since they work in quite a different way, it can be assumed that they complement each other well. Another important task relating to face detection is face tracking, which means that once a face is found on the video stream the system sticks to it and tries to follow that face even if its orientation or illumination changes.

The second main goal of the project is face recognition. Of course there is a prerequisite for this, the face has to be close enough and be in a good position related to the camera for the system to be able to recognize it. The system is expected to recognize those people once it has seen (and has been able to extract enough features for recognition). So if a person known by the system steps into the view of the camera, then the program has to recognize him or her and supply information stored in the system database about the given person.

The structure of this paper is the following. After a review of the earlier system [1] in the fourth chapter the main features of the 2D Gabor wavelets are summarized. The fifth part of the paper presents our approach to recognize the face that based on the description of Kepenekci in [2]. Finally the results are presented.

2 The Face Detection Part

The Myra system uses two different techniques for finding faces, an appearance based method and a skin tone detection method. These two techniques are meant to make face detection more accurate, so that on the one hand it would be less sensitive for changes of the illumination, on the other hand it should find faces even if the face on the picture does not look directly into the camera, so when the appearance based method can not be applied (which is able to detect frontal or profile faces, but can not cope with those that are between the two pose). The two methods can also be used to check the result of each other.

The appearance based method was implemented according to the method described in source [3], for this purpose the function of the OpenCV library was used that can give accurate results in real time if the direction of the face is frontal on the images. The other basis of face detection is a skin tone detection algorithm [4], which was supplemented with some methods, pre- and post-processing procedures to improve the performance of the detection. Prefiltering mainly consists of a noise filtering algorithm and some color temperature correction, while post filtering is the correction of the skin tone map. After prefiltering the resulting image is used and processed by the two different face detection methods. One is the appearance based technique, which is carried out with the help of the Adaboost method and Haar-like filters [3, 5]. It is implemented using the above mentioned functions and procedures of the OpenCV library.

Following the other branch of the detection a non-linear color space conversion is carried out as the first step of skin tone detection. It is necessary, since the method uses the YCC color space and not the RGB. The great advantage of the YCC color space is that illumination and the real color (luminance and chrominance) of each pixel are separated into different channels. After the color space conversion is done skin tone color pixels are to be found on a definite area of the 2D projection of channels C_b and C_r. It is approximately the shape of an ellipse. Skin tone color pixels are stored in a binary 320x240 matrix (the exact size of one video image frame). The matrix will be referred as the skin map from now on. The purpose of the next step is to eliminate lonely skin color pixels and smaller pixel groups from the map, while filling in the gaps of the larger skin tone areas at the same time.

The results of the two different face detection techniques are compared and combined. It has to be examined whether the face candidates found by the appearance based method contain a large amount of skin tone pixels (as it could be expected) or not. If so, then the candidate can be regarded as a face found on the image, else the face candidate has to be rejected and removed from the list.

There is one more essential task that belongs to the face detection part and that is face-tracking, where the computation speed optimization is important. The main goal here is to spare some time and face-tracking also helps the program to see those faces that are not in an ideal position for face detection. This tracking method also uses the skin map, because even if a face is turned or rolled the cloud of skin tone pixels will remain the same place. So if the appearance based method loses the face found on the previous frames, but the skin tone pixels are still there, then the program has a good reason to think that the face is still there and will show it, improving the continuity of the detection.

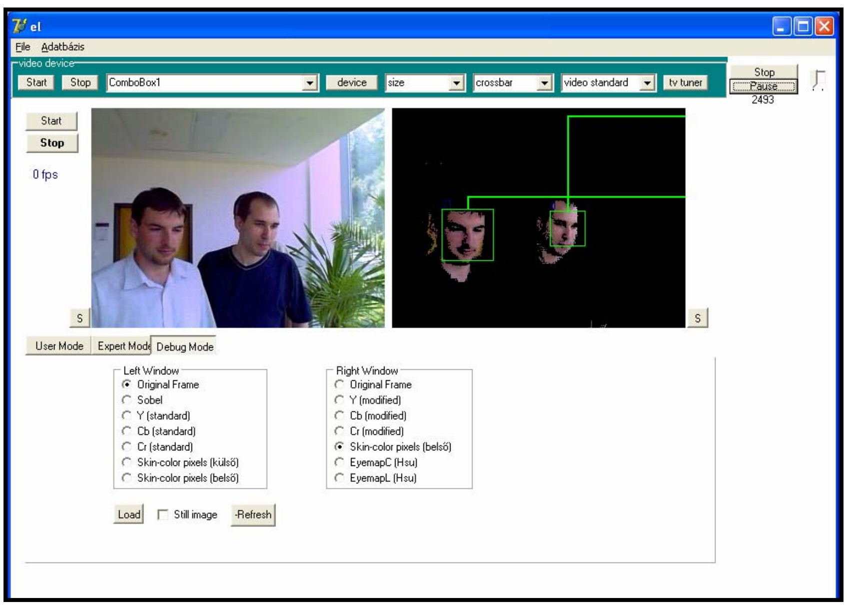

Figure 1
Result of the face detection and the skin tone detection

3 The Earlier Face Recognition Part

To recognize a face on an intensity image a holistic method, only the 'eigenfaces' approach [6, 7] was used in the earlier system. The eigenfaces method is based on efficiently extracting the relevant information from human face images. This information is extracted by calculating the difference between the actual image and an 'average', mean face image. In other words, the main (principal) components of the face images are sought (facespace). The Principal Components Analysis (PCA) is an appropriate method for the evaluation of data sets because it shifts out their differences and equalities.

The facespace can be described by eigenvectors and therefore a smaller number of dimensions is needed than for the description of the whole space. The eigenvectors are the eigenfaces that carry the most characteristic differences from the average image. In order to compute the base of the image space the eigenvectors and eigenvalues of the covariance matrix must be calculated which is composed of the difference vectors of the training set.

The first step of recognition is to transform the input image into the facespace. This means to represent the input image as a linear combination of the eigenfaces' base vectors. The resulting vector will be compared with the other stored image vectors from the database during the process of recognition. The best resemblance is achieved when the Euclidian distance of the weighted vector of the stored image is the smallest to the weighted vector of our image. In order to classify the input picture an appropriate threshold value is used.

4 2D Gabor Wavelets [2]

Unlike the holistic eigenface method that uses the whole face to recognise, the 2D Gabor wavelets assures an feature based, structural classification using local fetures. The Gabor functions proposed by Daugman are local spatial bandpass filters. The Gabor function is a tool for signal detection in noise. Gabor discovered that Gaussian modulated complex exponentials provide the best trade off. For such a case, the original Gabor elementary functions are generated with a fixed Gaussian, while the frequency of the modulating wave varies.

Gabor filters, rediscovered and generalized to 2D, are now being used extensively in various computer vision applications. Daugman generalized the Gabor function to the following 2D form:

$$\Psi_j(x) = \frac{k_j^2}{\sigma^2} \exp(-\frac{k_j^2 x^2}{2\sigma^2}) \left[\exp(ik_j x) - \exp(-\frac{\sigma^2}{2}) \right] \tag{1}$$

Each plane wave characterized by the vector k_j enveloped by a Gaussian function, where σ is the standard deviation of this Gaussian. The center frequency of jth filter is given by the characteristic wave vector,

$$k_j = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos\theta_\mu \\ k_v \sin\theta_\mu \end{pmatrix} \tag{2}$$

having a scale and orientation given by (k_v, Θ_μ) (Figure 1).

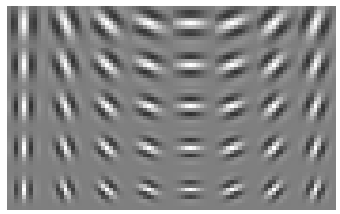

Figure 2
Gabor filters for 5 spatial frequencies and 8 orientation

2D Gabor filters (the convolutions of the image with the 2D Gabor functions) can be used to detect dimensions along a specific orientation, so they can be use to find objects in all orientation by rotating the filter. Since such curves correspond to some low-level salient features in an image, these cells can be assumed to form a low level feature map of the intensity image.

5 Face Representation with Gabor Features [2]

An image can be represented by the Gabor wavelet transform allowing the description of both the spatial frequency structure and spatial relations. Convolving the image with complex Gabor filters with 5 spatial frequency (v = 0,…,4) and 8 orientation (μ = 0,…,7) captures the whole frequency spectrum, both amplitude and phase. In Figure 3 an input face image and in Figure 4 the amplitude of the Gabor filter responses are shown.

Figure 3
Input face image

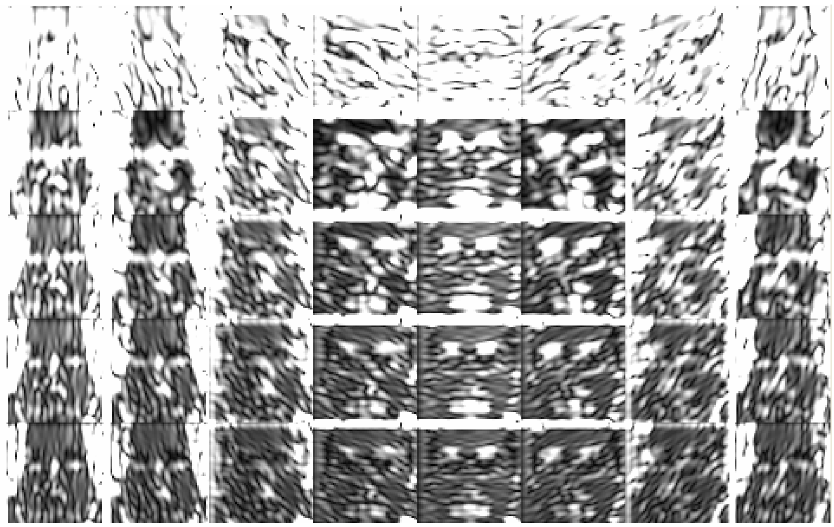

Figure 4
Gabor filter responses of the face image

5.1 Feature Extraction

Feature extraction algorithm for the proposed method has two main steps feature point localization, and feature vector computation. In the first step feature vectors are extracted from points with high information content on the face image.

From the responses of the face image to Gabor filters, peaks are found by searching the locations in a window W_0 of size W x W by the following procedure: A feature point is located at (x0, y0), if

$$R_j(x_0, y_0) = \max_{(x,y)\in W_0}(R_j(x,y)) \quad j = 1..40 \tag{3}$$

$$R_j(x_0, y_0) > \frac{1}{N_1 N_2}\sum_{x=1}^{N_1}\sum_{y=1}^{N_2} R_j(x,y) \tag{4}$$

where R_j is the response of *j*th convolution filter, N_1 and N_2 is the size of the image. A feature map is constructed for the face by applying above process to each of 40 Gabor filters.

Feature vectors are generated at the feature points as a composition of Gabor wavelet transform coefficients, *k*th feature vector of *i*th reference face is defined as,

$$v_{i,k} = \left\{ x_k, y_k, R_{i,j}(x_k, y_k) \quad j = 1..40 \right\} \tag{5}$$

The first two parameters represent the location of the feature points and the other 40 parameters are the responses of the Gabor filters at the given point.

5.2 Similarity Calculation and Face Comparision

To comparise two feature vectors in a given point the following similarity function is used

$$S_i(k, j) = \frac{\sum_l \left| v_{i,k}(l) \right| \left| v_{t,j}(l) \right|}{\sqrt{\sum_l \left| v_{i,k}(l) \right|^2 \sum_l \left| v_{t,k}(l) \right|^2}} \tag{6}$$

$S_i(k,j)$ represents the similarity of jth feature vector of the test face, ($v_{i,j}$), to kth feature vector of ith reference face, ($v_{i,k}$), and l is the number of vector elements. Proposed similarity measure between two vectors satisfies following constrains: $0 < S_i < 1$

After feature vectors are constructed from the test image, they are compared to the feature vectors of each reference image in the database. This comparison stage takes place in two steps. In the first step, we eliminate the feature vectors of the reference images which are not close enough to the feature vectors of the test image in terms of location and similarity. Only the feature vectors that fit the following criterions are examined in the next step.

$$\sqrt{(x_r - x_t) + (y_r - y_t)^2} < th \tag{7}$$

where th is the approximate radius of the area that contains either eye, mouth or nose, (x_r, y_r) and (x_t, y_t) represents the location of a feature point on a reference face and test face respectively. Comparing the distances between the coordinates of the feature points simply avoids the matching of a feature point located verz different position. After such a localization,

$$Sim_{i,j} = \max_{l \in N_{k,j}} (S_i(l, j)) \tag{8}$$

is applied as a second elimination of feature vectors of the reference face in order to guarantee surviving at most only one feature vector of a reference face to be matched with a feature vector of the test face.

The overall similarity of each reference face is computed as the mean of feature vector similarities that passed the step

$$OS_i = \text{mean}\left(Sim_{i,j} \right) \tag{9}$$

OS_i represents the overall similarity of test face to *i*th reference face, and its values change from 0 to 1, but for similair faces its value is tipically greater than 0.95.

6 Results

The new part of the Myra system was tested several different ways. Face recognition was tested on the pictures of a known image database, webcam and stored and live camcorder video streams. During the experimental examinations on the adequate images of the 'Stirling' database, i.e. on faces looking straight into the camera, the face recognition achieved a recognition rate of above 85%. This rates are higher than the results of the earlier implemented eigenface based method. The geometric features based method made face recognition more accurate. Using the 'Gabor wavelets' improved the system a lot, because they are utterly insensitive to light conditions. At the same time it is important to mention that the speed of the 2D Gabor filters based approach works significantly slower in our implementation than the eigenface method. Using the complete MYRA system to process video pictures with face detection and Gabor wavelets face recognition allows only the processing of 1-2 frames per second in contrast with the less accurate but quicker eigenface method (11–12 frames).

The implementation of the feature based 2D Gabor wavelets in the Myra system improved the recognition rate of the face recognition under good lighting condition.

Acknowledgement

The research has been supported by the OM TDK grant under the terms of grant OM FPO 245540/2005.

References

[1] T. Kövér, D. Vígh, Z. Vámossy: MYRA – Face Detection and Face Recognition System, 4th Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence, SAMI 2006, Herlany, Slovakia, pp. 255-265

[2] B. Kepenekci: Face Recognition Using Gabor Wavelet Transform, Thesis, Middle East Technical University, Turkey, 2001, 118 pages

[3] P. Viola, M. J. Jones: Rapid Object Detection Using a Boosted Cascade of Simple Features, in Proceedings of IEEE CVPR, Vol. 1, 2001, pp. 511-518

[4] Rein-Lien Hsu; Mohamed Abdel-Mottaleb; Anil K. Jain: Face Detection in Color Images, in: IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24, No. 5, May 2002, pp. 696-706

[5] Paul Viola; Michael Jones: Robust Real-time Object Detection, Cambridge Research Laboratory, 2001

[6] M. Kirby, L. Sirovich: Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces, in IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990

[7] M. Turk, A. Pentland: Face Recognition Using Eigenfaces, in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1991, Maui, Hawaii, pp. 586-590

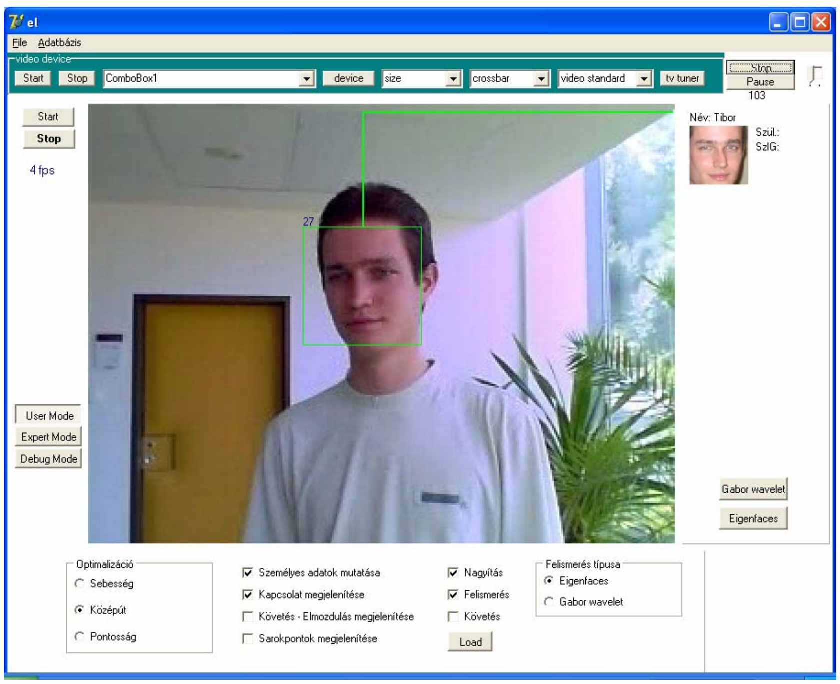

Figure 5
The main (user) view of the system