Profile-based Face Recognition using the Outline Curve of the Profile Silhouette

Keramat Allah Ghaffary
University of Kurdistan
Department of Computer Engineering, Sanandaj, Iran

Fardin Akhlaghian Tab
University of Kurdistan
Department of Computer Engineering, Sanandaj, Iran

Habibollah Danyali
University of Kurdistan
Department of Computer Engineering, Sanandaj, Iran

ABSTRACT

Face recognition technology has received significant attention in the past several years due to its potential for a wide variety of applications. However most of proposed face recognition systems are designed to work with frontal face images, there are several works that aimed to identify human faces from a profile view. In this paper we present a robust method to identify an individual based on the outline curve of the front portion of the silhouette that bounds the profile face images. The proposed method uses the center of ear and the tip of nose as the two reference points needed in the 2D space. The profile curve is then extracted by a segmentation method based on 2D histogram of the individual's face skin in H and S channels of the HSV color space. Having the two reference points and the segmented profile area, an angular sampling method is used to extract final normalized feature vector from the facial profile curve. In the matching phase, the Hausdorff distance metric makes the algorithm robust to small displacements in reference points. The experimental results on the GTAV face database show that the proposed method is promising and can operate reliably under illumination variations.

General Terms

Pattern recognition, Image processing, Biometrics.

Keywords

Profile based face recognition, Biometrics, Face profile segmentation, Ear detection.

1. INTRODUCTION

Face recognition has become one of the most important biometrics authentication technologies in the past few years due to its non-intrusive nature and aptness in various applications. Most of proposed face recognition methods are 2D based and designed to work with frontal face images. However the human face has a 3D nature, thus all of information about the structure of the face is not present in the frontal view. In the other words, facial profile curve contains different information of the face that is not present in the frontal view. Moreover in some applications such as the situation of a driver entering a gated area [1], the individual walking in a corridor or surveillance purpose, a frontal face image may not be available or hard to acquire. In these situations a profile based method is more applicable and maybe the only way to identify the individuals. In addition it is relatively easy to analyze, more foolproof [2] and more robust to illumination variations comparing with the frontal based face recognition systems.

Within the last decade, several algorithms have been proposed for automatic person identification using face profile curve. These methods can be classified into two main categories: appearance-based and silhouette-based methods. Appearance-based methods use all intensity information of an input side-view face image, where silhouette-based methods use the outline curve of the front portion of the silhouette that bounds the face image. The silhouette-based methods are less sensitive to illumination changes In contrast with appearance-based methods. In addition, they have less computational complexity and memory usage. These methods can be classified into two sub categories: curve-based methods and feature-based methods.

Curve-based methods try to use all information of the profile curve instead of limited fiducial points. Bhanu and Zhou in [2] proposed a method using dynamic time warping (DTW) to match face profiles. The similarity score between the probe and each profile in the gallery are computed by the DTW based on curvature. Their method was evaluated on two side-view face databases, reporting a recognition rate of almost 90%. Gao and Leung in [3] proposed a string matching method which the face profile is first transformed into a series of line segments. Then each line segment is represented by its attributes such as the length, orientation, and midpoint. After performing the merge domain string matching method, the distance score of the probe and the gallery profiles were computed. Pan, Zheng and Wu in [4] proposed to use metrics for the comparison of probability density functions on properly rotated and normalized profile curves.

Feature-based methods use some predefined fiducial points extracted from the facial profile curve. In the 1970's, Harmon Kuo, Ramig and Raudkivi in [5] published their seminal work in face profile recognition. According to the 9 extracted fiducial points on the profile, they formed a 10-dimensional feature vector. In their later work [6], the number of fiducial points was increased to 11 to form a 17-dimensional feature vector for each profile and a 96% recognition accuracy rate was reported. Wu and Huang in [7] used a B-spline to find six landmarks and extracted 24 features from the resulting segments. Liposcak and Loncaric in [8] used scale-space filtering to locate 12 landmarks and extracted 21 distances based on those landmarks. The Euclidean distance between vectors of features was used for identification.

Accurate localization of the fiducial points significantly affects the recognition results [1]. To the best of our Knowledge, there is no fiducial point's detection algorithm which is able to work perfectly and robustly on most of face images at different condition. The other short come of the proposed methods is that they do not include profile face detection and segmentation phase of real world data. Most of these methods including [2], [4], [8] and [9] process the pre-extracted profile face database such as university of Bern dataset which contains binary and segmented profile face images.

To overcome these limitations, we propose a new automatic and comprehensive profile based face recognition system which includes both the face segmentation phase and more reliable recognition algorithm. In the profile segmentation phase, at first and as a clue of face, the location of ear is detected based on the cascaded AdaBoost algorithm. Then using the color information in H and S channels of individual's face skin, the profile face area can be segmented. According to the ear location, the nose tip can be detected in the boundary of the segmented area. Considering the reference vector from the center of ear point to the tip point of nose, angular sampling of the face profile curve is started from the $+40^{o}$ to -50^{o} of the reference vector with 2^{o} steps. The distances of each sampled points from center of ear then normalized using norm of the reference vector. This feature vector can be stored and matched using a modified Hausdorff dissimilarity method. The extracted features rely on the all points of the face profile curves; therefore it is not sensitive to the correct location of fiducial points. Comparing feature vectors using a modified Hausdorff dissimilarity method, makes the algorithm more robust against various conditions.

The structure of the rest of this paper is as following: In the next section the proposed system is explained in details in four different subsections. In section 3 experimental results are presented and analyzed and finally in section 4 the conclusion is derived.

2. TECHNICAL APPROACH

2.1 Ear Detection

We adapted the cascaded AdaBoost [10] approach to detect ear from a 2D profile face images. AdaBoost is a method for selecting the best classifiers among a set of too many weak classifiers. The selected classifiers are then combined to produce a strong boosted classifier. In the cascaded AdaBoost method [11], a cascade of selected classifiers is built as the final detector, where a classifier is run only when all previous classifiers in the cascade have accepted a particular input data.

In this work, to adapt the AdaBoost method to the ear detection task, Haar-like rectangular features representing the grey-level differences as the weak binary classifiers are used. To reduce the computational complexity of the algorithm and perform it in the real-time, the integral image representation of the input image is used [11].

The train set composed of more than 2000 manually extracted ears as positive samples and 5000 images as negative set which are randomly chosen from a set of around 25,000 non-ear images. Examples of some ear images used in the training phase are shown in Figure 1. Testing on a set of 150 profile face images, the method achieved a 96.67% detection rate with a false positive rate (FPR) of $5\times10{-}6$.

The introduced ear detection method is automatic, in the sense that it does not require any manual intervention for the detection process.

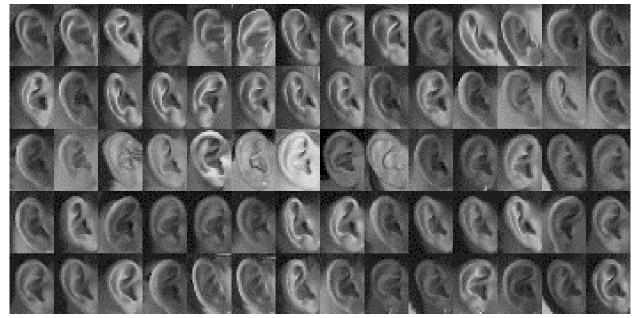

Fig 1: Examples of some manually extracted ear images used in training phase

2.2 Face Profile Extraction

Most of proposed methods for extracting the face area is generally based on skin color segmentation [12], [13], [14]. These methods use color components of a known skin model to classify each pixel of image as skin or non-skin. However, Kakumanu in [12] shows that using a predefined skin color model to extract the face region is a very challenging task as the skin color in an image is sensitive to various factors such as: illumination variations, camera characteristics, ethnicity, individual characteristics and some other factors.

Overcoming these challenges, instead of using a fixed predefined skin color model, we propose a new method which adapts the skin color model in face extraction phase. The introduced method, utilizes the skin sample of each individual as him/her skin color model. This sample can be extracted automatically from a predefined area according to the location of previously detected ear in previous step as shown with a blue rectangular area in Figure 2. This window is obtained according to the detected ear area by scaling and horizontal shift as is shown in Figure 2.

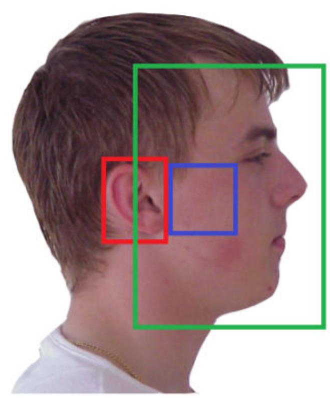

Fig 2: The three rectangular area used in face profile segmentation phase

Instead of examining all the image pixels by the obtained skin model, the searching area is limited to a window which surely contains the profile face. This window is also defines according to the detected ear region size and location, as is shown in Figure 2 by a green rectangular window.

Classification of each pixel in the green rectangular area as skin or non-skin is based on the estimated conditional probability

P(S|C). According to the Bayes' theorem, P(S|C) can be written as:

$$P(S \mid C) = \frac{P(S)}{P(C)} P(C \mid S)$$

Which C is the color of the pixel and P(S) is the probability that a pixel is skin. The method used in this approach to compute these probabilities is based on 2D histogram back projection. The back projection of a histogrammed image is the re-application of the modified histogram to the original image, functioning as a look-up table for pixel brightness values. In other word it is a way of recording how well the pixels fit the distribution of pixels in a histogram model. In this application, if we have a histogram of skin color then we can use back projection to find skin colored areas in the image. We use the HSV color space to compute the 2D histogram. HSV is one of the most interesting color space used in face detection methods [15], [16], [17], [18]. Since this color space separate the luminance and chrominance color components, using only H and S channels makes the proposed method more robust to the effects of illumination variations [12].Then using these two color channels, the ratio of the 2D histogram of the blue rectangular area as the skin model image ($H_b(h,s)$) and the green rectangular window ($H_g(h,s)$) is computed.

$$\forall h \in [0,180], \forall s \in [0,255]: H(h,\ s) = \min\left(\frac{H_b(h,\ s)}{H_g(h,\ s)}, 1\right)$$

According to back projection method, for each image pixel the probability of it belonging to the marked initial model (P(S|C)) can be computed as the count of H(h,s) bin that the pixel indexes into. Replacing each pixel with the corresponding probability, a grey scale image is computed as Figure 3(a). By thresholding this image, any examined pixel is classified as skin or non-skin (Figure 3(b)).

The extracted region may still include some holes in the eye or eyebrow locations. To fill these holes, it needs to apply morphological operations to have perfect segmented binary profile face image. We use a consecutive dilation and erosion with a same kernel window of size 3x3. The kernel size should be selected as small as possible to have minimal effects on the curve of profile boundary. The final segmented face profile is shown in Figure 3(c).

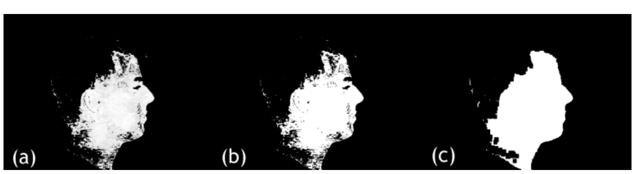

Fig 3: The probability image: (a) Raw grayscale mode, (b) thresholded and (c) after applying morphological operations

2.3 Angular Sampling

Having the extracted face profile and the ear center point, the nose tip can be easily extracted as the second reference point. For this purpose, we search along the facial profile curve for the point with maximum distance from the center of ear. Considering the vector through these two fiducial points as the initial side of angle, profile curve sampling procedure can be stared from $+40^o$ to -50^o with 2^o steps, as shown in Figure 4.

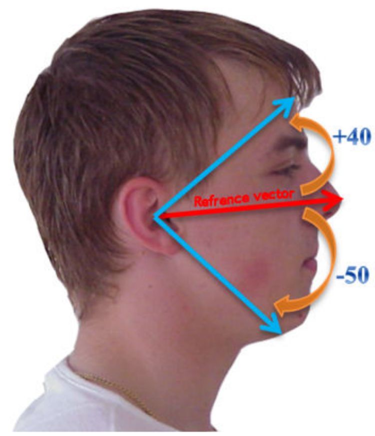

Fig 4: The reference vector and the sampling area

Each sample vector is the meeting point of the second side of the angle and the boundary of the segmented area from the center of ear. This sampling procedure results in 45 vectors meeting face profile curve as shown in Figure 5.

The distance between each sampled points and the center of ear, is normalized by dividing the norm of each sampled vector (d_i) by the norm of the reference vector(d_{ref}) which is shown in Figure 4 and Figure 5 with red color.

$$v_i = \frac{\|d_i\|}{\|d_{ref}\|} \forall 1 \le i \le 45$$

The obtained normalized distances form a 45D vector which represents the extracted feature vector from the individual facial profile curve. Since we have used an angular sampling this procedure is completely scale invariant.

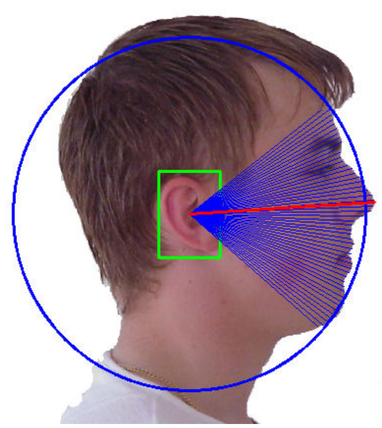

Fig 5: Angular sampling of the profile curve according to the reference vector

2.4 Feature Matching Using Hausdorff Distance

Each element of the extracted feature vector is corresponding to a point on the profile face curve. However the similar point of different individuals, such as tip of noses, start point of upper lips and etc., may appear at different feature vector indices. This can be due to different face shapes of different individuals or even variations in the face of a person at different conditions such as face pose or appearance. Therefore a simple matching measure such as Euclidean distance doesn't result in a robust face profile classification method.

The dissimilarity measure used in feature matching phase of this approach is based on Hausdorff distance. The Hausdorff distance is the maximum distance of every point in either feature set to the nearest point in the other feature set. This method is robust to partial shifts and displacements in the input sets. The classical Hausdorff distance between two given sets A and B which $A=\{a_1, a_2, \ldots, a_n\}$ and $B=\{b_1, b_2, \ldots, b_m\}$ is defined as:

$$D_H(A,B) = \max\{d_H(A,B), d_H(B,A)\}$$

Where:

$$d_H(A,B) = \max_{a\in A}(\min_{b\in B}\|b-a\|)$$

And

$$d_H(B,A) = \max_{b\in B}(\min_{a\in A}\|a-b\|)$$

The $\|a-b\|$ indicates any appropriate distance function such as Euclidean distance between a and b. Two sets are close in the Hausdorff distance if every point of either set is close to some point of the other set. The basic Hausdorff distance is very sensitive to even a single outlying point of A or B. Thus rather than $D_H(A, B)$, we use a modified versions of the Hausdorff distance $D'_H(A,B)$, which is insensitive to small perturbations of the point sets and allows for small positional errors in point sets. A lot of modified versions of Hausdorff distance are presented [19]. In this work a modified Hausdorff distance is used, where the distance between two sets A and B in this method is defined as follow:

$$D'_H(A,B) = \max\{d'_H(A,B), d'_H(B,A)\}$$

Where:

$$d'_H(A,B) = \frac{1}{n}\sum_{a\in A}(\min_{b\in B}\|b-a\|)$$

And

$$d'_H(B,A) = \frac{1}{m}\sum_{b\in B}(\min_{a\in A}\|a-b\|)$$

Dubuisson in [19] shows that, among the class of distance measures based on the Hausdorff distance, this method is more robust to various noses and its value increases monotonically as the amount of difference between the two sets increases. Using this method, extracted feature vectors are compared and according to a threshold, it can be decided that each two input profile face are belonging to a same individual or not. The value of threshold is determined according to the false accept rate (FAR) and the false reject rate (FRR) as shown in Figure 6. These two error rates are more discussed in the next section.

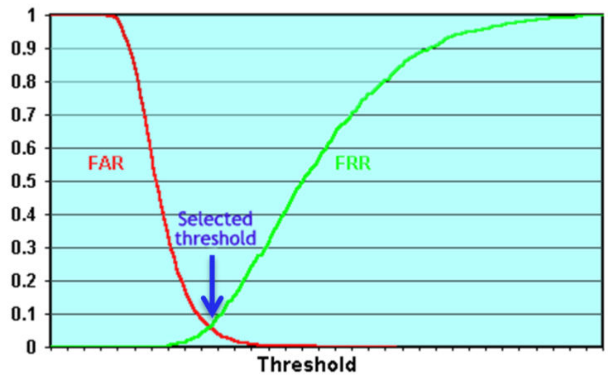

Fig 6: Finding the suitable threshold value according to the FAR and the FRR curve

3. EXPERIMENTAL RESULTS AND ANALYSIS

To evaluate the proposed method, the GTAV Face Database [20] is used which includes a total of 44 persons with 27 pictures per each individual which correspond to different pose views (0º, ±30º, ±45º, ±60º and ±90º) under three different illuminations. The resolution of each image is 240x320 pixels. Some sample image of an individual in this database is shown in Figure 7.

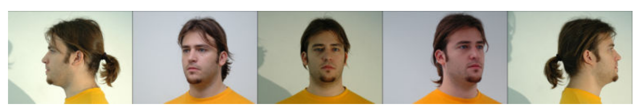

Fig 7: Some image samples of an individual in the GTAV face database

The method is implemented in C++ using the OpenCv2.2 library. To test this method using GTAV database, first a subset of 40 individual is selected. Since this database includes three pair of side-view images per each individual in three different illuminations, the first -90º image is used as gallery and the five other profile images is used as probes. The method is simulated using these datasets on a machine with 1.86 GHz Intel PIV processor and 2MB of RAM. For each recognition task, including feature extraction and matching phases, it takes an average of 0.045 second. Thus it is suitable for real time applications.

To measure respectively the performance of the proposed method for verification and identification tasks, we use equal error rate (EER) and cumulative match characteristic (CMC) [21], which are frequently used in face recognition community.

The receiver operating characteristics (ROC) curve of the data set is shown in Figure 8 and the CMC is shown in Figure 9.

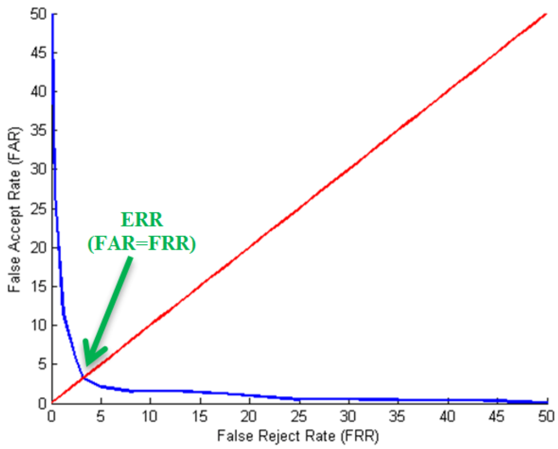

Fig 8: The receiver operating characteristics (ROC) curve of the method on the GTAV dataset

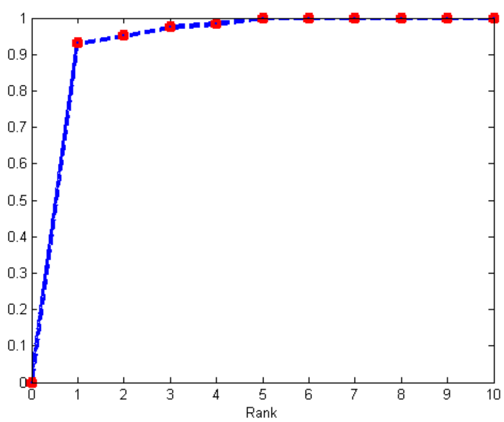

Fig 9: The cumulative match characteristic (CMC) curve of the method on the GTAV dataset

The EER is about 3.17% for the GTAV dataset. As the ROC and CMC curves show, the method has a promising performance on the dataset. However it has not been achieved 100% recognition rate. This is due to the size of Images in database. Each profile area in the GTAV database takes an average of 177x186 pixels in each samples, where is relatively a small size for recognition task. Since smaller images have fewer details about the profile curve, they have fewer discriminating power and lead to lower recognition rate.

To show the robustness of the proposed approach in different conditions, some experimental evaluations are performed. Since the GTAV dataset contains images in different lighting conditions, the promising result on this dataset shows that this approach is robust to illumination variations. The sampled area is shown for a set of images of an individual in three different lighting conditions in Figure 10.

To show the robustness of the proposed method against cluttered backgrounds, the experimental evaluation is done by adding various backgrounds to some randomly selected images. The results illustrated in Figure 11 show that method is also robust to variations in background image.

Since the training set for the cascaded AdaBoost is composed of ears with various rotations between -25^o and +25^o, the algorithm performs robustly for the images rotated between this range as shown in Figure 12. For the input images rotated more than 25^o (in either clockwise or counter-clockwise directions) the ear detection phase will fails and the algorithm cannot proceed.

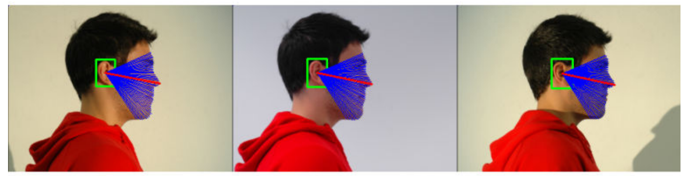

Fig 10: Robustness of the proposed method to illumination variations

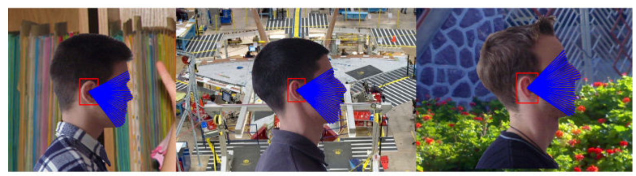

Fig 11: Robustness of the proposed method against cluttered backgrounds

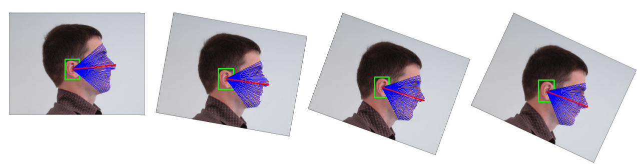

Fig 12: Robustness of the proposed method to 10^o, 20^o and 25^o rotations of the input image

4. CONCOLUSION

In this paper a new profile-based face recognition method is presented which uses the geometric information of the outline curve of profile silhouettes. The procedure starts with detecting the ear location by a cascaded AdaBoost method. Based on the ear location, the tip of nose and a skin sample area is defined. Using the 2D histogram back projection method on H and S color channels, the face area is segmented. The method utilizes the vector through the center of ear and the tip of nose to perform the angular sampling method and extract the final normalized feature vector from the face profile curve. A modified version of Hausdorff distance is used in matching phase to measure the similarity of each pair of individual feature vectors. As the experimental results show, the method is accurate and robust enough to do recognition task in various backgrounds and lighting conditions and also fast enough for real-time applications.

5. REFERENCES

[1] I.A. Kakadiaris, H. Abdelmunim, W. Yang, and T. Theoharis, "Profile based face recognition,"FG2008, 8th IEEE Int. Conf. on Automatic Face and Gesture Recognition, Amsterdam, September 2008.

[2] B. Bhanu, X. Zhou, "Face recognition from face profile using dynamic time warping," Int. Conf. on Pattern Recognition (ICPR) 4 (2004) 499–502.

[3] Y. Gao and M. Leung, Human face profile recognition using attributed string. Pattern Recognition, 35(2):353–360, February 2002.

[4] G. Pan, L. Zheng, and Z. Wu, "Robust metric and alignment for profile-based face recognition: An experimental comparison," 7th IEEE Workshop on Applications of Computer Vision, pages I: 117–122, Breckenridge, CO, USA, 2005.

[5] L. Harmon, S. Kuo, P. Ramig, and U. Raudkivi, Identification of human face profiles by computer. Pattern Recognition, 10(5-6):301–312, 1978.

[6] L. Harmon, M. Kaun, R. Lasch, and P. Ramig, Machine identification of human faces. Pattern Recognition, 13(2):97–110, 1981.

[7] C.J. Wu and J.S. Huang, Human face profile recognition by computer, Pattern Recognition, vol. 23, pp. 255–259, 1990.

[8] Z. Liposcak and S. Loncaric, "A scale-space approach to face recognition from profiles," 8th International Conference on Computer Analysis of Images and Patterns, London, UK, 1999, pp. 243–250.

[9] Z. Liposcak, and S. Loncaric, "Face recognition from profiles using morphological operations," pp. 47-52, IEEE, 1999.

[10] R. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. Mach. Learn., Vol. 37, No. 3:297–336, 1999.

[11] P. Viola and M. Jones, Robust real-time face detection. Int'l Journal of Computer Vision, 57(2):137–154, 2004.

[12] P. Kakumanu, S. Makrogiannis and N. Bourbakis, A Survey of Skin Color Modeling and Detection Methods, Pattern Recognition, Vol. 40, No. 3, pp. 1106-1122, March 2007.

[13] E. Hjelmas, B.K. Low, Face detection: a survey, J. Computer. Vision Image Understanding 83 (2001) 236–274.

[14] M.H. Yang, D.J. Kriegman, N. Ahuja, Detecting faces in images: a survey, IEEE Trans. Pattern Anal. Mach. Intell. 24 (1) (2002) 34–58.

[15] C. Garcia, G. Tziritas, Face detection using quantized skin color regions merging and wavelet packet analysis, IEEE Trans. Multimedia 1 (3) (1999) 264–277.

[16] S. McKenna, S. Gong, Y. Raja, Modeling facial colour and identity with Gaussian mixtures, Pattern Recognition 31 (12) (1998) 1883–1892.

[17] Q.H. Thu, M. Meguro, M. Kaneko, Skin-color extraction in images with complex background and varying illumination, Sixth IEEE Workshop on Applications of Computer Vision, 2002.

[18] Y. Wang, B. Yuan, A novel approach for human face detection from color images under complex background, Pattern Recognition 34 (10) (2001) 1983–1992.

[19] M. P. Dubuisson, A.K. Jain, "A modified Hausdorff distance for object matching," Proceedings of the 12th IAPR International Conference on Pattern Recognition, vol. 1, October 1994, pp. 566–568.

[20] F .Tarrés, A. Rama, "GTAV Face Database" available at: http://gpstsc.upc.es/GTAV/ResearchAreas/UPCFaceDatabase/GTAVFaceDatabase.htm.

[21] P Jonathon Phillips, Patrick Grother, Ross J Micheals, Duane M Blackburn, Elham Tabassi, Mick Bone, "Face recognition vendor test 2002:evaluation report," 2003.