Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Science and Engineering

Elixir Comp. Sci. & Engg. 55A (2013) 13255-13258

ISSN: 2229-712X

Face detection and recognition system in bank robbery using CFV modules based on video recordings

C. JayaMohan, M. Saravana Deepak, M.L. Alphin Ezhil Manuel and D.C Joy Winnie Wise
Department of CSE, Alpha College of Engineering, Chennai, T.N, India.

ARTICLE INFO

Article history:
Received: 12 September 2012;
Received in revised form:
1 February 2013;
Accepted: 19 February 2013;

Keywords
Face Verification Module,
Eye template verification module,
Video Image Processing,
Feature Extraction,
Pattern Recognition.

ABSTRACT
In this paper, It examines the Face detection and recognition system using three fast CFV modules which are face skin verification module, face symmetry verification module, and eye template verification module. The three verification modules can eliminate the tilted faces, the backs of the head, and any other non-face moving objects in the video. Only the frontal face images are sent to face recognition engine. The frontal face detection reliability can be adjusted by simply setting the verification thresholds in the verification modules. Then three hybrid feature sets are applied to face recognition. Experiments demonstrated that the frontal face detection rate can be achieved as high as 95% in the low quality video images. The overall face recognition rate and reliability are increased at the same time using the proposed ensemble classifier in the system. An ensemble classifier scheme is proposed to congregate three individual Artificial Neural Network classifiers trained by the three hybrid feature sets.

© 2013 Elixir All rights reserved.

Introduction

The research on the video-based face detection and recognition can be considered as the continuation and extension of the still-image face recognition, which has been extensively researched for years and some good results have been reported in the literature. Face detection and recognition in a video sequence has become an interesting research topic due to its enormously commercial and law enforcement applications. For example, the well-known methods such as Principal Component Analysis (PCA) [1], Linear Discriminant Analysis (LDA) [1], eigenfaces and Fisherfaces methods [2], Elastic Graph Matching (EGM) [3], robust Handsdorff distance measure for face recognition [4], eigenspace-based face recognition [5], a novel hybrid neural and dual eigenspaces methods for face recognition [6], etc.

Another important task for face recognition in a video clip is face detection. In order to capture the frontal face images accurately and timely, many face detection methods have been proposed, such as the discriminating feature analysis and Support Vector Machine (SVM) classifier for face detection [7], neural network-based face detection [8], face detection in colour images based on the fuzzy theory [9], etc. Face colour information is an important feature in the face detection. In reference [10], authors used the quantized skin colour regions for face detection; statistical colour modules with application to skin detection was reported in reference [11]; A latest survey of skin-color modelling and detection methods can be found in reference [12]. Eye is another important feature for face detection and recognition. For example, a robust method for eye feature extraction on the colour image was reported in reference [13]. Using optimal Wavelet packets and radial basis functions for eye detection was introduced in reference [14]. Some face detection applications in videos can be found in references [15, 16]. The development of cheap, high quality video cameras has generated new interests in extending still image-based recognition methodologies to video sequences. In recent years, research on this area has attracted great interests from scientists and researchers worldwide. In this paper, we propose a simple and efficient cascade face detection scheme:

The tester's head is automatically detected by the body moving information between adjacent frames. Then the possible face area is serially verified by face skin verification module, face symmetry verification module, and eye template verification module. Only the frontal face image, which has passed the three verification modules, is sent to face recognition engine for recognition. By doing so, there are two advantages: 1) computer can save a plenty of time for video process; 2) face recognition performance is increased as only frontal face image is sent for recognition.

The schematic flowchart of face detection and recognition is presented. In Next Section, three novel and fast face verification modules are proposed. Then, three face recognition algorithms are summarized in order to pursue a higher recognition rate, a novel ensemble classifier scheme is proposed. The face detection and recognition experiments conducted on five video clips are reported in this Section.

Face detection and recognition scheme

The flowchart of the proposed video-based cascade face detection and face recognition system is shown in Fig.1

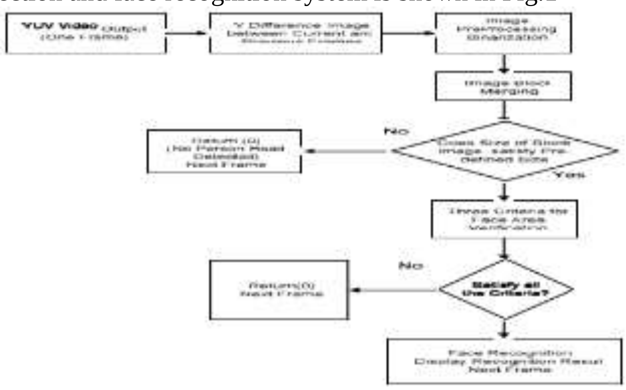

Fig. 1: Face Detection and Recognition Schematic Flowchart

Tele:
E-mail addresses: alphacse138@yahoo.com

© 2013 Elixir All rights reserved

Face Verification Modules

In this section, three face verification modules are proposed. Firstly, face skin and non-face spectra are analysed. The skin spectra are given. Secondly, a fast face symmetry verification algorithm is developed. Finally, three eye templates are chosen to verify the frontal face.

Face Skin Verification

It has been reported that human skin only occupies certain spectra in the colour space regardless of the different ethnic groups. In order to obtain an accurate face colour spectrum,1024 face images were used for colour spectrum analysis. The following colour spaces have been analysed: (r, g, b), (H, S, V), (S, T, L), (Y, Cb, Cr,), (L, U, X), and (I, Rg, By). We used the k-n cluster algorithm to classify skin colour space into four categories and construct four multivariate normal distribution modules for face colour analysis.

Based on (R, G, B,) colour space, formula used in the analysis are listed as follows:

r = R /(R+ G+ B), g = G /(R+ G+ B),b = 1.0− r − g

H = a cos(0.5*(2* R − G − B) / sqrt ((R − G) 2+ (R − B) 2))*180/ PI S

= (Max (R, G , B) − Min(R, G , B)) / Max (R, G , B)

V = Max (R, G , B) / 255.0

Here, PI=3.1415. After analysis, we decide to use only two colour spaces (r, g, b) and (H, S, V) in our face verification module due to the facts that the two colour spaces are comparatively stable for different face skins and are less sensitive to illumination, intensity and the partial occlusion of the lights. Fig. 2 shows the distributions of four categories of skin spectra in the (r,g,b) and (H,S,V) colour spaces.

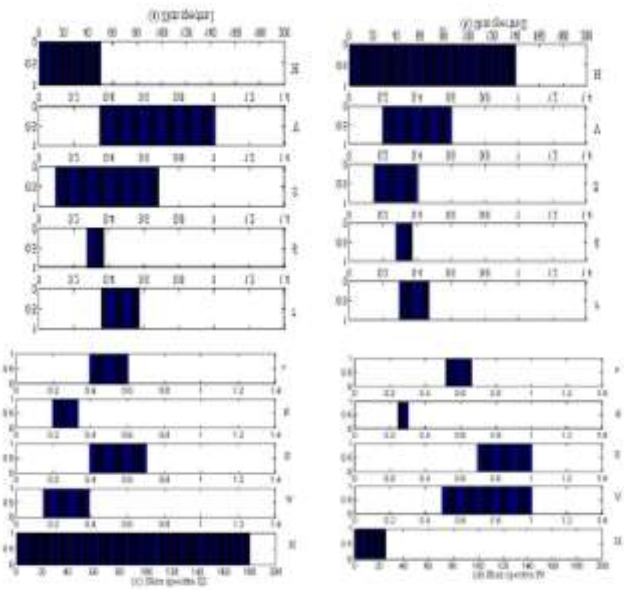

Fig. 2: Skin Spectra Distributions in the (r,g,b) and (H,S,V) Color Spaces

Although skin spectra can be used to detect human skin, sometimes hands, other parts of the body, and even some background objects which have similar colour spectra to the human skin's are also misclassified. In the next two sub sections, other two criteria will be introduced.

Face Symmetry Verification

Eyes may be the most important feature in the verification module [13, 14]. Before the eye template verification module is applied, in order to highlight eye area in the face area image, the Y grayscale image in the (Y, U, V) colour space is preprocessed as follows:

1) Face image is scaled down the size of 64x64;

2) Morphological dilation operation on Y:

$$I_m = Y \oplus Str_1 \tag{2}$$

Here Str_1 is a 3x3 matrix with all elements being set to 1.

3) Morphological erosion on I_m :

$$In = I\ m{:}Str2 \tag{3}$$

Here Str2 is a 2x2 matrix and all elements are set to 1.

4) Calculate Io I m / In in the pixel wise operation, and scale the image Io into a grayscale image Ig with scale of [0,255].

Fig. 3 shows four original face images and the processed images using our proposed algorithm. The experiments demonstrated that the eye areas can be extracted regardless of ethnic groups, wearing of glasses, and illuminations.

It is a common sense that the human's face is symmetric. In this section, the Y grayscale image in the (Y, U, V) color space is used for face symmetry verification. The following steps are proposed for face symmetry verification:

1) Segment the face area from the shoulder and other parts of human body using face color information and the width information of the human body;

2) Divide the face area into two parts vertically and evenly: left sub-block and right sub-block;

3) Calculate the histogram of the left sub-block L(i) and the histogram of the right sub-block R(i), i=0, 1,2, …, N-1; N is the length of the face area;

4) Compute the Symmetric Similarity Coefficient (SSC) using the flowing formulae:

SSC = 2*L(i)*R(i)/(L(i) 2 + R(i)2) (1) The threshold for SSC can be adjusted according to applications. In our experiments, if it is set to 0.75~0.85, then most face images in the video can be verified. The higher the threshold is set, the less frontal faces will be captured and the more symmetric frontal faces are extracted.

Eye Template Verification

After a face area image passed the face skin verification module and the face symmetry verification module, the face area image will be scaled to 64x64 in order to use the next verification module: the eye template verification module.

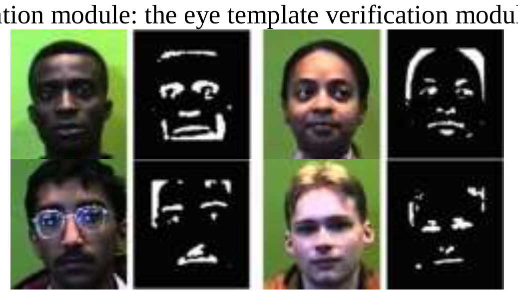

Fig. 3: Original Face Images and Processed Face Images

Three eye templates are chosen as shown in Fig. 4. One template is horizontal, other two are with 15 degree angles.

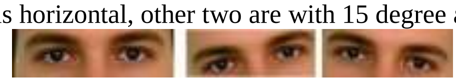

Fig. 4: Three Eye Templates

The above eye pre-process is applied to the three eye template images. Then, the processed eye image data are saved as the eye templates. A face image in the video, which has successfully passed first two verification modules, will be processed by the eye preprocess algorithm in order to enhance the eye areas. Then, three eye templates correlate with the processed face area image, respectively. Three maximum correlation coefficients for three templates are denoted as Coe1, Coe2, and Coe3.

An overall maximum correlation coefficient is calculated as follow:

MaxCoe = max {Coe1,Coe2,Coe3} (4) If the value of MaxCoe is higher than the predefined threshold, the testing frontal image passes this module. In our experiments, if the predefined threshold is set to 0.90. The eye templates can detect frontal face with maximum 15 degree angles.

Fature extraction for face recognition

Y component in (Y, U, V) colour space in a frontal face image is used to extracted three sets of features for face recognition as follows:

Haar Wavelet Feature

The Haar wavelet is the simplest and the fastest wavelet transformation, which operates on data by calculating the sums and the differences of the adjacent elements. The Haar wavelet operates first on adjacent horizontal elements and then on adjacent vertical elements. Fig. 5 shows the diagram of Haar feature extraction.

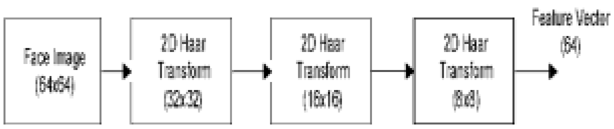

Fig. 5: Haar Wavelet Feature Extraction Scheme

A feature vector of 64 dimensions is extracted from each 64x64 face image.

Gradient Based Directional Feature

We assume an input image as Iz, the templates of the Sobel operators Sx and Sy [17] are listed in Tables 1 and 2.

Table 1: Template of Sobel Operator Sx

Table 2: Template of Sobel Operator Sy

1	2	1
0	0	0
-1	-2	-1

The X-gradients of the frontal face image can be calculated by:

$Ix=Iz*Sx$ (5)

and the Y-gradients of the frontal face image is calculated by:

$Iy=Iz*Sy$ (6)

The gradient magnitude and phase are then obtained by:

$$r(i,j) = \sqrt{Ix^2(i,j) + Iy^2(i,j)}$$

$$\theta(i,j) = \tan^{-1} Iy^2(i,j) / Ix^2(i,j) \quad (7)$$

Then, we can count the gradient direction of each pixel of the convoluted image with nonzero gradient magnitude values as a directional feature.

In order to generate a fixed number of features, each gradient direction is quantized into one of eight directions at $\pi / 4$ intervals. Each normalized gradient image is then divided into 16 sub-images. The number in each direction of each sub-image is counted as a feature. In total, the number of features is 4x4x8=128. The algorithm used in the recognition of handwritten numerals with good discriminant ability was reported in [18].

Gradient-Based Wavelet Feature

The directional-based feature extraction is implemented as follows: firstly, the Kirsch nonlinear edge enhancement algorithm is applied to a 64x64 face image to extract horizontal, vertical, right-diagonal and left-diagonal directional feature images and the global feature image; 2- D wavelet transform is used to filter out the high frequency components of each directional feature image and the global feature image, and to convert the feature matrix into The details of the algorithm can be found in [19]. Three extracted feature sets combine with five colour components (r,g,H,S,V), respectively, in the face areas to form three hybrid feature sets. The extracted three sets of hybrid features are used to train Artificial Neural Networks (ANNs) with Back Propagation (BP) algorithms as three classifiers.

Ensemble Classifier

A novel classifier combination scheme is proposed in order to achieve the lowest error rate while pursuing the highest recognition rate for the video based face recognition.

the confidence values c0,0~c0,R-1 of ANN1, and so on). A gating network is used to congregate the weighted confidence values.

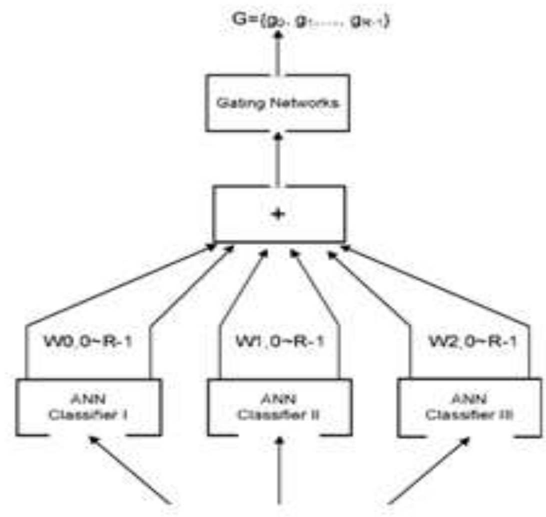

Fig. 6: An Ensemble Classifier Consisting of Three ANNs andOne Gating Network

A genetic algorithm is used to evolve the optimal weights for the gating network from the confidence values of three ANNs.Suppose the outputs of three ANNs are represented as:

{c0,0, c0,1,…,c0,R-1}, {c1,0, c1,1,…,c1,R-1}, {c2,0, c2,1,…, c2,R-1}

Our goal is to pursue a lowest misrecognition rate and at the same time to seek the highest recognition performance. We can create a vector Otarget with R elements, which represent the number of persons in the test video. In the vector, the value of the corresponding label is set equal to 1.0, while others are set equal to 0.0. A fitness function f is chosen to minimize the difference between the output G and the corresponding training sample vector Ot arg et , as follows:

f =| G − Ot arget |2 (11) By minimizing (11) through a genetic evolution, the eights tend to be optimal. Then, the recognition criteria are set as follows:

A recognition result is accepted if:

1) three ANN classifiers vote for the same person at the same time, where the sum of the confidence values is equal to or larger than 1.8; or
2) the gating network votes for a person, where the confidence value of the gating network is larger than
0.65; or
3) the sum of the confidence values of any two ANNs is larger than 1.30 and they both vote for the same person and the gating network votes for the same person; otherwise, the person is rejected.

Face detection and recognition experimental results

In our experiments, five surveillance video clips taken indoors are used to test the face detection and recognition performance. Some of detected face images from the videos are shown in Fig. 7.

Fig. 7: Detected Frontal Face Images from Video

Table 3: Overall Recognition Rates on Video Images with Different Feature Sets

Feature Set		V1	V2	V3	V4	V5
Hybrid Feature	A	90.00	86.70	88.80	84.00	80.00
Hybrid Feature	B	90.00	93.30	88.80	84.00	82.20
Hybrid Feature	C	90.00	86.70	83.30	84.00	82.20
Ensemble classifier		100.00	93.30	94.40	88.00	88.90

Note: V1~5 denotes videos from 1 to 5.

Experiments demonstrated that the frontal face detection rate can be achieved as high as 95% in the low quality video images. A comparative face recognition experiments have been conducted on three ANN classifiers trained by three hybrid feature sets and the proposed ensemble classifier using the same five video clips as used in the face detection experiment. The overall recognition results are shown in Table 3.

In the security applications, reliability is a crucial issue. The reliability is defined as:

RE = Total number of testers − Number of misrecognized testers

Conclusions

In this paper, Face detection and recognition system is presented. Three fast and efficient face detection verification modules are proposed to detect and verify the frontal faces in the video clips. Only the frontal faces, which have serially passed three verification modules, are sent to the recognition engine for face recognition. Furthermore, the frontal face detection reliability can be adjusted through the setting of the verification thresholds. Therefore, this adaptable mechanism can be used to the different applications. For example, we can set the lower verification thresholds for the face tracking purpose only. Experiments conducted on five video clips have demonstrated that the frontal face detection rate can be achieved as high as 95% in the low quality videos.

Three hybrid feature sets are successfully applied to the face recognition system. A novel ensemble classification scheme is proposed to congregate the outputs of three ANN classifiers, which are trained by three sets of hybrid feature sets. The experiment shows that the overall face recognition rate and reliability are increased if the ensemble classifier is used.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, “Face recognition: a literature survey, ”ACM Computing Surveys, vol. 35, no. 4, 2003, pp.399-458.
[2] P. N. Bellhumeur, J. P. Hespanha, and D. J.Kriegman, “Eigenfaces vs. fisherfaces: recognition using class specific linear projections,” IEEE Trans. Pattern Anal. and Machine Intell., vol. 19, no. 7, 1997, pp. 711-720.