FACE RECOGNITION USING MODIFIED KALMAN FILTER AND EIGEN FACE METHOD

J Narendra Babu 1 *
1. Assoc. Prof, Dept of ECE, SIETK, PUTTOR, INDIA.

ABSTRACT:

The human face is central to our identity. It plays an essential role in everyday interaction, communication and other routine activities. Thus, face detection and tracking algorithms are of great importance for human-machine interaction. It is an extensively studied field and a large body of research has been done. This paper gives a broad overview of basic principles and some representative methods for face detection, tracking and facial feature extraction.

KEYWORDS: face tracking, facial feature localization, face detection.

I. INTRODUCTION

Face Recognition is vividly researched area of image analysis, pattern recognition and more precisely biometrics with numerous commercial and law enforcement applications being developed. This technique is an effective means of authenticating a person. There are various approaches and methods for face recognition proposed by researchers till date. One of the most popular methods that yield promising results on frontal face recognition is the principal component analysis (PCA), which is a statistical approach where face images are expressed as a subset of the eigenvectors, and hence called eigen faces. Matthew Turk and Alex Pentland [1] used this PCA method for face recognition. Their work projected face images onto a feature space that spans the significant variations among known face images. The significant features were known as “eigenfaces” because they were the eigenvectors (principle components) of the set of faces. But the results show that false reject rate is quite high for the faces if face image is even a little different from the one in the database. However the Eigen-face method is not able to differentiate between an unknown face image and the one, who is of the same person as in the database, with a little variation in its gesture. An improved face recognition technique had been given by Rajkiran Gottumukkal, Vijayan K. Asari [2]. They proposed a modular PCA method, which was an extension of the PCA method for face recognition. They claimed that modular PCA method performed better than the PCA method. But again the deficiency of PCA for frontal face recognition cannot be overcome using this modular PCA approach. Guan-Chun Luh and Ching-
Chou Hsieh [3] proposed a method using PCA and immune network for machine learning. Within the last few decades, numerous novel face recognition algorithms had been proposed [4]. A central issue to these approaches was the feature extraction. The most well-known technique for linear feature extraction was the linear discriminant analysis (LDA). Its basic idea was to seek an optimal set of discriminant vectors by maximizing the Fisher criterion. Then, Jin et al. [5] proposed the uncorrelated LDA technique (ULDA), which tried to find the optimal discriminant vectors by maximizing the Fisher criterion under the conjugated orthogonal constrains. Zhongkai Han et al. [6] proposed a method which used discriminated correlation classifier and addressed mainly the classification problem in face recognition. However, the ULDA and discrimination techniques suffer from the so-called small sample size problem (SSSP) which is often encountered in face recognition. The face detection technique proposed by Lamiaa Mostafa and Sherif Abdelazeem [7] was based on Skin Color. Their method used the normalized RGB color space. To build the model, samples were collected of human skin from different races like Africans, Americans, Arabs etc. They proposed that the first step in skin detection was pixel-based skin detection, where the skin detector tested every pixel of the input image and computed its normalized red value \( r \) and normalized green value \( g \). If \( r \) and \( g \) values of the pixels satisfied some predefined inequalities, then this pixel was considered skin. The drawback of this approach is that it produces good results for different races people but for the same race it does not produce satisfactory results. Various other methods based on finding skin and skin color [8, 9, 10] were also proposed for face recognition and face detection. Further Sanjay Kr. Singh et. al. [11] proposed a Robust Skin Color Based Face Detection Algorithm in which they used the three popular color models named RGB, YCbCr and HSI color models to detect skin color features. But again this method suffers with the problem of robustness as no other parameter was taken into the account except the skin color for the faces of the persons belonging to the same country. Further Yuehui Chen and Yaou Zhao [12] proposed a face recognition method which used Discrete Cosine Transformation (DCT) and Hierarchical Radial Basis Function (HRBF) model. The DCT was employed to extract the input features to
build a face recognition system, and the HRBF was used to identify the faces. Fabrizia M. de S. Matos et. al. [13] proposed a method which used three coefficient selection criterion. The drawback of this approach is that DCT is not a very efficient transform to extract features of the image. Wavelet transform is considered better than DCT as wavelets can be used for Multi Resolution Analysis (MRA) while DCT can not be used for the same and further there is no way to use the DCT for lossless compression, since outputs of the transformation are not integers while wavelet transform has efficient implementation, both in lossy and lossless case as stated by Øyvind Ryan [14]. Futher Cunjian.chen and Jiashu.zhang [15] proposed a new feature extractor for face recognition which used wavelet energy entropy. They stated that wavelet energy entropy can be used as new facial features to recognize faces. Mrinal Kanti Bhowmik et. al. [16] performed a comparative study on fusion of visual and thermal images using different wavelet transformations like Haar and Daubechies (db2). They found wavelet a useful transformation for face recognition. Boqing Gong et. al. [17] proposed a method based on extracting facial features for 3-D faces. They emphasized on two components one for corresponding basic face structure and another corresponding expression change in the face based on the change in shape of 3D face. N. Sudha et al. [18] proposed an efficient self-configurable systolic architecture for very large scale integration implementation of a face recognition system. The proposed system applied principal component neural network (PCNN) with generalized Hebbian learning for extracting eigenfaces from the face database. Further Wonjun Hwang et al. [19] proposed a method which has two stages. First, in the preprocessing stage, a face image was transformed into an illumination-(insensitive image, called an “integral normalized gradient image,” by normalizing and integrating the smoothed gradients of a facial image. The hybrid Fourier features were extracted from different Fourier domains in different frequency bandwidths and then each feature was individually classified by linear discriminant analysis. Further Arindam Biswas et. al. [20] proposed a novel algorithm for extracting the region of interest of a face like eye pair, nostrils, and mouth area. Bart Karoon et. al. [21] addressed the influence of eye localization accuracy on face matching performance in
the case of low resolution image and video content. Jun-ying zeng et. al. [22] proposed a model for partially occluded face recognition based on Biometric Pattern Recognition. An experiment based on biomimetic pattern recognition adopting a PCA and LDA feature extraction method was performed and results were quite satisfactory. Brian Cheung [23] trained a convolutional neural network to distinguish between images of human faces from computer generated avatars as part of the ICMLA 2012 Face Recognition Challenge. It can be inferred that the regions of our interest in a face can be located correctly even if the face image is of low quality. In an attempt to further improve the performance of the face recognition method we have presented an improved algorithm by designing a feature vector comprising of wavelet horizontal, vertical and diagonal coefficients, determined from the wavelet energy function capturing global features of the face and, the facial parameter capturing local features of the face. The result of this algorithm when compared with that of the PCA method is found more satisfactory.

II. RELATED WORK
The fast evolution of computer technologies including hardware and software has advanced the state of computing machinery in the past two decades to the point where human life has been significantly improved by machine intelligence. This trend has resulted in an active development in information technology and artificial intelligence, where more friendly and efficient approaches for human computer interaction are developed based on new devices. Computer vision, which is one aspect of machine intelligence, focuses on duplication/emulation of human vision. Traditionally, computer vision systems have been utilized in specific applications such as assembly line inspections and quality control in automated manufacturing. The ever decreasing cost of computing systems and video image acquisition equipment has resulted in computer vision systems advancing towards more generalized vision applications such as face detection and face tracking techniques. Face detection, which is the first step in any face processing system, attempts to determine whether there are any faces in a single image. If any faces exist, the processing system provides the image location and extent of each face (Yang et al., 2002). Face detection is important in any human face related system, such as any fully
automatic face recognition system, warning and surveillance system, or face tracking and human tracking system. Face detection algorithms can be typically extended to generic object detection and recognition (Zhao et al., 2003), which leads to automatic target recognition (ATR). So far, face detection in computer vision is still a challenging task, even though it is easy for humans to perform effortlessly (Hjelmås and Low, 2001). The various face detection related problems include face localization, facial expression recognition, face recognition, face authentication, and face tracking. Traditionally, the solutions to the problems are based on image segmentation, facial feature extraction, and face verification in the presence of complicated background. The challenges associated with face detection are contributed by changes in scale, location, orientation, pose, facial expression, occlusion, and illumination. Face tracking aims to keep account of face in a video sequence i.e., determine if there are any faces in a single frame, and continuously estimate the locations and possibly the orientations of the faces in the video sequence in real time (Darrell et al., 2000; Crowley and Berard, 1997; Edwards et al., 1998). Face tracking belongs to the larger area of visual object tracking pursued by the computer vision community, where the object of interest is the face. Object tracking has been studied extensively by researchers in the context of computer vision because of many vision applications such as autonomous robots (Davison and Murray, 1998), video surveillance (Borg et al., 2005), human eye tracking (Hansen and Hammoud, 2005) and human face tracking (Nummiaro et al., 2003). Generally speaking, an image sequence, which is collected in real time, does not change rapidly from one frame to the next frame. This results in a large redundancy of object information over consecutive frames spanning a certain time interval. This redundancy can be utilized to disambiguate the appearances of the visual objects and track the individual objects. Since the human visual system may not distinguish a camouflaged object from a complicated background, the exploitation of the redundancy in a sequence of images is still regarded as a challenging problem in the computer vision community (Isard, 1998).

III. PROPOSED METHODOLOGY
The face recognition process is accomplished in the following steps stated
in chronological order as shown in block diagram in figure 1.

**Figure 1. Block diagram of the Methodology of the Proposed Work**

### 3.1 Forming Database of Face Images
Database will consist of a number of images depending on the number of persons and the number of images captured per person with different gestures. For each image calculate the wavelet’s horizontal, vertical, diagonal coefficients and hue component as explained in subsection 3.2 as following:

### 3.2 Extraction of Feature Vector
It comprises of the wavelet’s horizontal, vertical and diagonal coefficients named h, v and d are used as three coefficients of the feature vector estimated from equation (4), (5) and (6) as described in subsection 3.2.1. The fourth coefficient of the feature vector is a ratio corresponding to facial parameters from equation (7) as described in subsection 3.2.2. Thus the feature vector comprises of these four coefficients.

#### 3.2.1 Estimating Wavelet’s Coefficients of Face
(i) Decompose the face image using ‘Haar’ wavelet.

(ii) Extract detail horizontal, vertical and diagonal coefficients named H, V and D.

(iii) As H, V and D are two dimensional matrices. In order to generate a single valued horizontal, vertical and diagonal coefficients use wavelet energy function for each horizontal, vertical and diagonal detail coefficient named WHE, WVE and WDE as in equations (1), (2) and (3).

In above equations m is the number of rows and n is the number of columns of the two dimensional matrices H, V and D.

(iv) Take the square root of the equations (1), (2) and (3) which finally gives the horizontal, vertical and diagonal coefficients named h, v and d respectively:

\[ h = \sqrt{WHE} \]  
\[ v = \sqrt{WVE} \]  
\[ d = \sqrt{WDE} \]

#### 3.2.2 Estimating Facial Parameters of Face
(i) Calculate the euclidean distance \( d_1 \) between the center of both the pupil (the center of iris of the eye) of face.
(ii) Calculate the length of the nose $d_2$. $d_1$ and $d_2$ together form a ‘T’ shape.

(iii) Calculate $d_2/d_1$ ratio named $fp$.

$$fp = \frac{d_2}{d_1} \quad (7)$$

### 3.3 Learning of Recurrent Neural Network using Extended Kalman Filter

It consists of the training of the recurrent neural network (RNN) with the coefficients of the feature vector using Extended Kalman filter. Recurrent neural network is the network with one or more feedback loops [24]. In present case the error $(\cdot)$, the difference between the desired output and estimated output, is used as a feedback to the neural network which makes it a recurrent neural network.

#### 3.4 Testing of Face Images

After machine learning the testing is performed with the test cases. The three test cases for the presented work have been designed as following:

- **Case A:** Testing with the same image as stored in database.
- **Case B:** Testing with the different image of the same person (whose images is stored in the database) with variation in gesture.
- **Case C:** Testing with an unknown image (image of a person whose image is not at all present in the database).

### IV. RESULTS

A 13 stage cascaded classifier was trained to detect frontal upright faces. Each stage was trained to eliminated 50% of the non-face patterns while falsely eliminating only 0.2% of the frontal face patterns. In the optimal case, we can expect a false alarm rate about $0.00213 = 8 \cdot 10^{-36}$ and a hit rate about $0.99813 = 0.97$ (see Fig. 2). To train the detector, a set of face and non-face training images were used. The face training set consisted of over 4,000 hand labelled faces scaled and aligned to a base resolution of 24 x 24 pixels. The non-face sub-windows used to train the detector come from over 6,000 images which were manually inspected and found to not contain any faces. Each classifier in the cascade was trained with the 4,000 training faces and 6,000 non-face windows using Adaboost.

**Speed of the Final Recognition System**

The introduction of the Kalman filter to reduce the search region has demonstrated its value. Actually on a 1GHz PIII laptop, the detection and recognition runs at a 8.6 fps whereas with the Kalman improvement its processing rate depends on the area occupied by the face. Naturally the larger improvements are observed when the user’s face occupies the least detectable area on the
image. In this case processing speeds of 24 fps are obtained.

V. CONCLUSION

During this work first the well known Eigen-face method has been implemented and it has been found that its performance is not satisfactory for the different image which has variation in gesture, but is of the same person, whose image is already stored in the database. Further in this proposed work, a wavelet analysis is used to decompose the original image. A feature extraction approach, called “wavelet energy function” is proposed to extract wavelet horizontal, vertical and diagonal coefficients of the face image. The derived coefficients are later fused to form a new feature vector along with the facial parameter, which will be used as the inputs to the recurrent neural network and machine learning is performed using Extended Kalman Filter. Results of the experiments on the face database as shown are the evidence that the new algorithm is able to perform much better than traditional methods like Eigen-face method and its variants.

VI. REFERENCES


