Multiple Face Detection and Recognition: Automated Attendance System – A Survey

1Manoj D, 2Nethra N, 3Manjunatha D, 4Girish Kumar L, 5Jyothi S Nayak
1,2,3,4 Students, 5Associate Professor
Department of Computer Science,
BMSCE, Bengaluru, India

Abstract: The existing attendance systems are time consuming, inefficient and error-prone. Face detection and recognition is the focus of cutting edge research in today's era due to its wide applicability in areas such as Surveillance, Automated Access Control, People Tracking, etc. One relevant application of multiple face detection and recognition is Automated Attendance System. In this paper, we present a survey on the existing methods and technologies used for detecting and recognizing faces in images and the applicability of these in automated attendance system.

Keywords: Face Detection; Face Recognition; Principal Component Analysis; Linear Discriminant Analysis; Support Vector Machine; Ensemble Classifiers.

I. Introduction

The attendance systems currently in use require human intervention, be it calling out names, biometric methods like iris scanning, fingerprints, RFID etc. In iris scan method, images of the eyes are given as input to the iris scanner which recognizes the iris of the particular student and marks the attendance. Fingerprint method scans the fingerprint of all the students, recognizes them and attendance is marked. Students are provided with identification (ID) card which contains RFID tags that contains electronic information. The ID cards of the students are scanned and attendance is marked.

The proposed system involves three main steps: Detection of multiple faces in the images acquired by the instructor, Recognition of faces detected and marking attendance of the same. Face detection involves detecting candidate regions in an image that are possible faces. It is a challenging task because of differences in shape, illumination, lighting conditions, size, orientations, complex background occlusions and skin tone. Next step, Face recognition involves identifying the person by comparing the detected face with a reference dataset and then classify or recognize his/her identity. Even though there are no standard algorithms for recognition of faces there are some statistical, classification techniques for face recognition like Principal Component Analysis, Linear Discriminant Analysis, Support Vector Machine, etc. In the final step, once a person is recognized, his/her attendance is marked in the database consisting of enrolled students.

This paper presents a survey on existing attendance systems, various methods used for face detection, different recognition and classification techniques. The methods used for preprocessing of the image acquired is also discussed. Thus an overview of the existing techniques and how they can be used for the attendance system is presented.

II. Existing Attendance Systems

Srinidhi et al. describes a system based on RFID (Radio Frequency Identification) [1]. This system is web enabled and can be used for real time location tracking. Here, the registered RFID card holders need to flash their RFID tag embedded ID cards in front of the RFID reader. This method is insecure i.e., can be spoofed and also time consuming as users need to stand in queue to get their cards scanned by the reader and also expensive.

Attendance can be marked based on Bluetooth and EyeOS [2]. This system provides scalability using cloud computing. It is a client/server interaction style of database which is fast and has efficient storage. However, user needs to have Bluetooth enabled device with him/her, spoofable and fault-intolerant.

Benyo proposed a system based on NFC (Near Field Communication) which provides fast and secure identification, data exchange features [3]. It includes student card which must be NFC capable, a terminal which is an interface and then a backoffice which generates identification policies, collects and stores data. This system however, is time consuming and performs poorly.

Attendance system based on iris recognition was proposed by kadry et al. and is implemented using Daugman's algorithm [4]. It consists of three steps: Image acquisition and pre-processing, Iris feature extraction and Sign encoding, Signature matching and verification. This system is stable, unique however a rejection rate of 9.2% and noisy transmission are some of its disadvantages.
Manjot Singh et al. implemented an attendance system using Raspberry pi [5]. Two devices are used here namely, AIMEGA16 micro-controller as a hand-held device and Raspberry Pi web server as a local server. Hand-held devices work on 5V DC is basically used for taking and sending the attendance. A local server (Raspberry Pi) with Zigbee (application) for monitoring all devices in the network and updates. This system is affordable and secure.

The authors in [6] describes a system based on face recognition algorithms. Image acquisition, Database development, Face detection, Preprocessing, Feature extraction and classification are the stages in this system. Frontal images of the students are detected by face geometry based methods. Viola-Jones algorithm is used for face detection. Histogram equalization technique (Resizing of extracted images) is used as a preprocessing method. The performance of a system depends on the Feature extraction and classification. PCA (Principal Component Analysis) Algorithm is used to represent the face economically, while LDA (Linear Discriminant Analysis) can be used in well illuminated condition. This system is fast and robust. Spoofing is one disadvantage of the system.

III. Face Detection Techniques

Face detection has been extensively researched in past decades. It is a specific case of object detection which determines the location and size of candidate faces in an image. It is a process of designing a system by giving input consisting of images that contains faces and non-faces and then training a classifier to identify a face in an image. The main focus of here is to decrease false positive rate thereby increasing accuracy.

Yang et al. [7] describes that there are four main approaches used in face detection: Feature based, Appearance based, Knowledge based and Template Matching.

Feature based methods uses features like skin color, eyes, nose and ears for Face Detection. Rotation independency, Scale independency and quick execution time are the main advantages of this method. Feature restoration is the key challenge in Feature-based methods.

Appearance based methods are learning based method and uses classification and features in search window for Face Detection. By the values of pixels features are calculated. Models are learned from set of training images and these learned models are then used for face detection. Classifiers are created using statistical learning between the enormous instances. The main advantage of this method is that the non-face window will be rejected at the early stage and hence the execution time will be decreased.

Knowledge based methods require some simple rules for Detecting the Faces from images. The rules can also be extended to detect faces from complicated background. Facial features are extracted from the input image and then the face candidates are recognized based on the rules defined. The challenge here is to convert the human knowledge into rules so that the can be applied on various inputs with different conditions.

Template Matching based methods are used to obtain regions with high probability of being human face. A template is constructed from several training images, the input image is correlated with face template for detecting whether the region is face or not. In this method some assumptions are to be considered. Simple implementation is its main advantage.

The Viola – Jones algorithm [8] which can be considered as a milestone in the field of object detection. It makes use of haar features[9] for feature extraction. Haar-like features considers adjacent regions of rectangular shapes at a specified location in window of the image, sums pixel intensities and calculates the difference between these sums. The values obtained are used to identify whether the window contains a face or not. Integral images to reduce computational time, Adaboost for picking relevant number of features that best distinguishes face from non-face, Adaboost constructs a strong classifier as a linear combination of weak classifiers and then a cascade of classifiers are used in this algorithm. However, this traditional method requires large training data set. The disadvantages of classifier can be overcome by using neural networks once the haar features are extracted.

The method in [10] divides each image into small sized sub images and then each sub image is tested separately using a fast neural networks. This fails in the case of detecting occluded faces or tilted faces in image as it only detects upright frontal faces.

Skin color segmentation method can be used to detect faces [11]. This method removes face pixels from image thereby eliminating large portion of an image that are non-face. HSV color space is used here as it perceives different shades of color. A face template is constructed from several training images. Thus Eigen image with template matching is used here. This method can be further extended to even classify genders by correlating with two average faces template made of male and female faces respectively. This method works better with increase in variety of datasets and also detects occluded faces in image.

Mingxing [12] proposes a face detection system that makes use of Support Vector machine (SVM) and Adaboost which significantly improves detection rate and also reduces computation time in training. This method shows improvement in the extraction of several Haar-like features, higher accuracy in different lighting environment.

The work described in [13] is an interesting approach for video based face detection where video is converted into pairwise image set. Both query and reference gallery will contain set of images for each individual. This is a M versus N strategy where M denotes face sets detected and N denotes number of reference faces. Here, to train recognizer annotation of generated tracklets to face image pool for each person is done manually. Each image is encoded with histogram of local binary pattern. K-means is used.
to cluster face images in each pool. However, since this is video based the memory requirement will be high and time is required to generate tracklets even though it achieves better recognition rate.

The authors in [14] describes a face detection system by combining multiple classifiers using locally assembled histograms of oriented gradients. This method combines eight neighboring HOG features to capture the co-occurrence. The selected LAH features focus on eyes, mouth, etc. This method is proved to better than methods that make use of edge oriented histograms, Haar features. The detection rate also changes according to face size.

Venugopal et al. [15] has given some of the preprocessing steps to be done for the images. Segmentation to separate human face regions from non-face regions by using different color spaces. Background subtraction to eliminate background objects from the images thus eliminating non face pixels from the image. Thresholding is used to create binary image which is suitable for morphological operations. These operations simplify image information and removes irrelevance. De-noising where low pass filters are used to remove noise.

Histogram normalization method is used in [16] for face detection. The captured image is converted to gray scale as it may have brightness or darkness in it. This technique is also good for contrast enhancement. In Binarization, whole features are used to segment possible skin color regions and non-skin color regions. Pixels containing facial features like eyes, eyebrows, mouths are assigned 1 and background pixels are assigned 0.

Preprocessing is an important step for face detection because it saves lot of time required for computation, reduces dimensions of the original data and makes the system robust to images with variations in illumination and other lighting conditions.

Jian-qing Zhu et al. proposed a Face Detector based on gentle AdaBoost and nesting cascade structure [17]. To avoid the reduction in the Face Detection speed caused by the weak classifiers in the cascade classifier, nesting cascade classifier is used. Node classifiers are trained using gentle AdaBoost Algorithm on a Haar like feature set to improve the generalization ability of the node classifiers and also the performance of the Face Detector. Weak classifiers are combined to form a strong classifier called node classifier and the node classifier are combined to construct a cascade Face Detector. This paper focuses mainly on accelerating the Face Detection speed without pre-processing. The experimental results conclude that the proposed detector reaches very high detection speed and accuracy.

A Real Time Multi-Face detection system based on hardware design [18]. Other existing system focuses mainly on the software algorithms to improve the detection rate and decrease the false alarm. Complex algorithms require more computational time which hinders the real-time applications. The proposed model uses Feature based methods and is implemented on Altera DE2-70 development board which is used to test the feasibility of the hardware design. This model requires 15,223 logic elements and operates in real-time at a frame rate of 30fps and detects up to five faces simultaneously. It can be concluded from the experimental results that the proposed Face Detection Architecture attains a real-time reliable system with low cost and high detection rate.

IV. Face Recognition and Classification Techniques

Face recognition is the process of identifying a particular individual present in the image by analysis and comparing patterns. Face Recognition has a wide range of applications and is mainly used in security purposes.

1. Principal Component Analysis (PCA)

PCA is simplest of true Eigen vector based and multivariate analysis. It reveals internal structure of data in a way that best explains the data i.e., major features and directions in the data. Set of images can be represented in high dimensional data space. PCA can supply the user with lower dimensional picture when viewed from most informative viewpoint. This technique uses orthogonal transformation to convert set of values possibly correlated M variables into a set of values of K uncorrelated variables K< M called principal components or Eigen faces. This transformation is defined in such a way that first Eigen face shows most dominant features of dataset and each succeeding Eigen face shows next most possible dominant features of dataset. In this only first few principal components are selected and rest may be discarded as they may represent noise and other less interesting features.

Let $Y = \{y_1, y_2, \ldots, y_k\}$ be the random vectors.

- Compute mean, $\mu$

$$\mu = \frac{1}{k} \sum_{i=1}^{k} (y_i)$$
• Computation of Covariance Matrix, $S$

$$S = \frac{1}{k} \sum_{i=1}^{k} (y_i - \mu)(y_i - \mu)^T$$

• Compute Eigen Values $\lambda_i$ and Eigen Vectors $\mathbf{v}_i$ of $S$,

$$S \mathbf{v}_i = \lambda_i \mathbf{v}_i$$
where $i=1,2,3,\ldots,k$

2. Linear discriminant analysis (LDA)

LDA [19] is one of the statistical technique that can be applied for face recognition and requires a large dataset. Let’s say the given dataset contains $N$ images of $M$ persons equally i.e., equal number of images are present for each subject. This images of each subject can be termed as class. Mean is calculated for each class and overall mean is also calculated for projection matrix training. Scatter matrix is computed for each class and also among classes. This can be called as within class scatter matrix and between class scatter matrix respectively. The goal here is to reduce the scatter within class and increase the distance or scatter between the classes. LDA uses face images as feature vectors for face recognition.

Let $C$ represent the number of classes and $Y$ random vector,

$$Y = \{Y_1, Y_2, \ldots, Y_C\}$$

$$Y_i = \{y_{i1}, y_{i2}, \ldots, y_{ik}\}$$

• Compute the Scatter matrix $S_B$ and $S_W$

$$S_B = \sum_{i=1}^{c} N_i (\mu_i - \mu)(\mu_i - \mu)^T$$

$$S_W = \sum_{i=1}^{c} N_i \sum_{y \in \mathcal{Y}_i} (y - \mu_i)(y - \mu_i)^T$$

Where $\mu$ is total mean,

and $\mu_i$ is the mean of the class $i \in \{1,2,\ldots,c\}$

3. Support Vector Machine (SVM)

SVM [27] is a non-linear classifier that provides great classification by mapping data into high dimensional space by a kernel function. It constructs a hyperplane between two classes in the transformed space. Say constructing a hyperplane between face class and non-face class. It requires training images containing faces and also non-faces.

Let the training dataset contain $k$ points,

$$(\mathbf{m}_1, n_1), (\mathbf{m}_2, n_2), \ldots, (\mathbf{m}_k, n_k)$$

Where $\mathbf{m}_i$ is a vector and $n_i$ is either 1 or -1.

$$\mathbf{w} \cdot \mathbf{m}_i - b \geq 1 \quad \text{if} \quad n_i = 1$$

Or

$$\mathbf{w} \cdot \mathbf{m}_i - b \leq -1 \quad \text{if} \quad n_i = -1$$

Where $\mathbf{w}$ is normal vector to the hyperplane.
An efficient Face Recognition system was proposed by Chunjing Li et al. [20]. Principal Component Analysis (PCA) and Bidirectional Principal Component Analysis (BDPCA) methods are used to recognize a grayscale face image. K-L Transform is used to extract main parts of face to form Eigenface pace. Singular Value Decomposition (SVD) is proposed to solve the eigenvalue and eigenvector. Eigenvector are extracted from the testing image and compared with the eigenvectors of the training face images. PCA is used to determine whether it is a human or inhuman face and BDPCA determines whether the testing face image is present in the training set or not. BDPCA method improves the speed of feature extraction and the face recognition rate. But the disadvantage of BDPCA is that the compression rate of high-dimensional data is smaller with single dimensionality reduction and hence a lot of storage space is required.

Gender classification can increase the accuracy of face recognition [21]. The gender of the testing face image is determined and is compared with only the images of same sex. PCA face recognition system based on single image has been used. The proposed method for gender estimation is based on training processes using several samples for each person such as Support Vector Machine, Neural Networks and Adaboost system. The main advantage of gender classification is the enhancement in accuracy of face recognition using PCA which is up to 7%. By comparing the input image only with the same sex images in the database ensures reduction in the time required for recognition.

Parvathy et al. proposed an efficient Face Recognition system [22]. This system recognises faces in different views and poses based on a novel algorithm which utilizes the combination of texture and depth information based on modular PCA to overcome the problem of pose variation and illumination changes. It combined 2D and 3D systems in the feature level to achieve higher performance. Recognition rate of this system is 86% which is well above the recognition rate of PCA (71.45%).

Kamencay et al. put forward a method for Face Recognition using SIFT-PCA [23]. This paper shows the impact of graph based segmentation algorithm on recognition rate. The important information from the face data are extracted and represented as a set of orthogonal variables called principal components. Pre-processing of face images is carried out by using segmentation algorithm and Scale Invariant Feature Transform (SIFT) descriptor. Image segmentation is the process of splitting the entire image into set of segments. SIFT features are the features that are extracted from images which help in reliable matching of the same objects in different views. The accuracy rate is increased from 87% to 92% by using SIFT-PCA rather than PCA alone.

A Face Recognition algorithm using 2D-LDA (Linear Discriminant Analysis) and polynomial-based radial basis function neural networks (P-RBF NN) [24]. It consists of two modules: pre-processing and recognition. P-RBF NN is used for recognition and 2D-LDA for pre-processing. 2D-LDA overcomes the singularity problem implicitly. Differential Evolution (DE) optimizes the essential design parameters such as learning rate, feature selection etc.

The method proposed in [25] deals with the variance Gaussian noise. Gaussian noise can cause declination in the Face Recognition rate. To overcome this problem, Pulse Coupled Neural Network (PCNN) combined with LDA is used. The Region clustering feature of PCNN is good. Reconstruction of variance information of a destroyed image can be done using clustering feature. LDA has a very low variance Gaussian noisy facial recognition rate.

Verma et al. put forward a Face Recognition system using Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) [26]. The dimension of the training database set is reduced by PCA, feature extraction is by LDA and nearest mean classifier for testing. Orthogonal linear space representation of a class is provided by PCA. Class discriminatory information in a linear separable space which is not necessarily orthogonal is generated by LDA. The recognition rate by employing this technique is 96.35%.

Combining an image in different ways with its original version so that it contains more relevant information was proposed by authors in [28]. Fusing an original image with its corresponding diagonal images can be done in two ways. In first method from original face image we generate diagonal face image and then the original image and the diagonal image are appended horizontally side by side. Second, we append original face image with the corresponding diagonal face image vertically to get two sets of fused image matrices. G-2DFLD method is applied on both of these large fused images for extraction of discriminant features extracted feature matrices are applied on Radial Basis Function-Neural Networks (RBF-NN) for classification and recognition. Performance of this technique is better than PCA, 2DPCA, PCA and FLD combined.

The method for face recognition using SVM as a classification technique was put forward in [29]. Wavelet transform is applied for image preprocessing in order to reduce the impact of expression change on face recognition. PCA method is followed here to map original face image to Eigen faces axis to achieve dimensionality reduction of Eigen. Then, SVM classification model is used to identify the projection vector of human face image in Eigen face axis. SVM improves generalization ability of model and also in this method the recognition rate is competitively better than PCA, Wavelet and PCA combined.
Global feature and local feature are very essential for face recognition and detection. Yu et al. proposed a method which uses both local and global features for face recognition [30]. Low-frequency coefficients of Fourier transform gives the global features which are extracted from the whole face images. This gives a holistic facial information. Gabor Wavelets are used for extraction of local features. Next we apply Fisher’s Linear Discriminant (FLD) is separately applied to both global and local features. Hierarchical Ensemble Classifier is obtained by combining all these classifiers. FERET and FRGC version 2.0 databases are used evaluating the proposed method. Results proves that the proposed method is an efficient face recognition method.

In the Proposed method, we detect faces from the image, recognize them and mark attendance. Face Detection is done by extracting Haar-like Features and using Viola Jones Algorithm. After the candidate face regions are extracted, Statistical LDA is used for face recognition. To increase the efficiency of the system we are combining local features with holistic approach using ensemble of classifiers.

V. Conclusion

The traditional systems as well as biometric methods are time consuming, error-prone, costly and lack accuracy. These systems require the intervention of the students to mark the attendance causing inconvenience to students. Automating the system will decrease the time required to mark attendance and also eliminates the need for student intervention. We propose an automated attendance system which marks attendance by detecting and recognizes the faces of the students captured in the image and thus eliminating the student intervention.

VI. ACKNOWLEDGMENT

The work reported in this paper is supported by our college [BMS College of Engineering, Bengaluru] through the TECHNICAL EDUCATION QUALITY IMPROVEMENT PROGRAMME [TEQIP-II] of the MHRD, Government of India.

VII. REFERENCES


