An Efficient Gait Recognition System Based on PCA and Multi-Layer Perceptron

C.Murukesh [1], Dr.K.Thanushkodi [2]

[1.] Department of EIE, Velammal Engineering College, Anna University, Chennai, Tamilnadu, India
[2.] Director, Akshaya College of Engineering and Technology, Anna University, Tamilnadu, India
pcrmurukesh@gmail.com

Abstract: Gait Recognition System is an emerging technology which aims to identify person at a distance by the way they walk. Human gait is a spatio-temporal phenomenon that typifies the motion characteristics of an individual.The Gaussian Mixture Model (GMM) is used to remove the background of an input video file image and the relevant features are extracted using Principal Component Analysis (PCA). This paper presents an efficient gait recognition system based on Multi-Layer Perceptron (MLP) and compares its performance with Fisher Linear Discriminant Analysis (FLDA) to identify the gait images from the video stream. Experimental results evaluated on public database of video sequences shows that the proposed method increases the recognition rate.
[Murukesh C, Dr.Thanushkodi K. **An Efficient Gait Recognition System Based on PCA and Multi-Layer Perceptron**. *Life Sci J* 2013; 10(7s):1024-1029] (ISSN: 1097-8135). http://www.lifesciencesite.com. 165

Keywords*:* Human gait, Gaussian Mixture Model, Principal Component Analysis, Multi-Layer perceptron, Fisher Linear Discriminant Analysis

1. Introduction

Biometric technology works as a tool to identify a person by their selected features. Each organ in the human body might be used as an identification unit because they have unique characteristics. Many researcher have focused on human identification using various physiological characteristics such as iris, DNA, fingerprint, human face, palmprint, and even behavioral characteristics such as a way of talk, voice and gait walking [1].The traditional methods such as PINs, passwords, tokens, electronic cards used for identification are easily subjected to cracking of password or forgetting the password. Unlike the traditional methods, the biometric authentication system shown in Figure 1 has gained the substantial attention in all the fields and plays the major role in accurately identifying the person based upon the physiological and behavioral characteristics.

Human gait is less unobtrusive biometric to recognize an individual without any interaction. Video-based gait recognition is mainly applicable in surveillance systems such as recognizing an unlawful person from a security camera video [2]. The gait of an individual is known to differ from person to person and to be fairly stable; whereas intentional imitation of another person's gait is complicated. Gait recognition system can be classified into three groups namely, floor sensor based, wearable based and motion vision based. The wearable sensors technique measures different walk style by carrying necessary sensors in any part of the human body and the floor sensors located on the floor is used to detect the required measurement. The motion vision based is divided into appearance based methods and model based methods. The appearance based method can be subdivided into spatio-temporal methods and state space methods [3].The proposed algorithm is based on spatio-temporal method as it deals with time variation [14]. The effect of time plays a great role in gait recognition system by controlling their covariate factors and removing their influences [4].

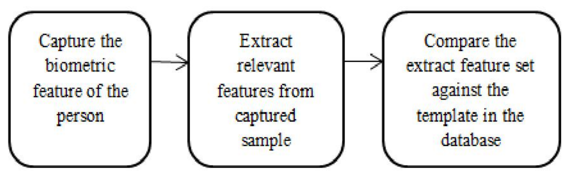


Figure 1. Block diagram of biometric Identification System

Gait recognition is also non-invasive as it requires no direct contact and can be applied from a distance. It has extensive applications in various fields such as social security, biometric authentication and remote surveillance [5]. The behaviour of gait is invariant over long period of time. Unfortunately, it exhibits several variations due to varying surfaces, view angles, clothing, shoe types, elapsed times, moods of the person, etc., often resulting in poor performance of gait recognition algorithms. Gait based gender identification using pattern classifier has attracted many researchers to classify gender from gait. It helps to improve search speed and efficiency of the retrieval system [6-8]. In this paper, MLP classifier is used to classify the gait sequences and also it is compared with the other classifiers like Euclidean distance and Fisher Linear Discriminant Analysis (FLDA) classifier. The classification rate is analysed with the Cumulative Match Score (CMS) [15].

Recognizing the particular person starts with acquiring required data with the help of static sensor fixed at the public places like airport, super malls etc. The sensor used in the gait recognition is the video camera, which acquires with a resolution of 320×256 image pixels. The acquisition speed is 30 frames per second and for the given approach the camera can be fixed far from the person to be recognised. The Figure 2 illustrates the block diagram of the proposed gait recognition system.

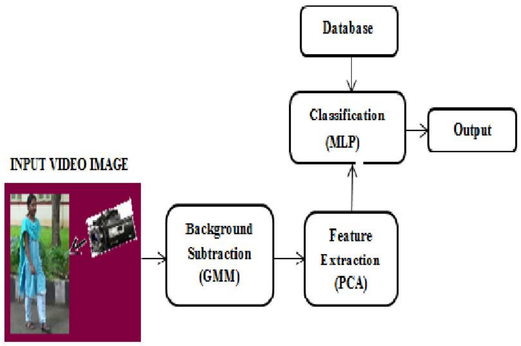


Figure 2. Block diagram of proposed gait recognition system

2. Background Subtraction

The purpose of pre-processing on the input video image is to extract the foreground or moving objects from the background scene. The process of background subtraction remains as an important low-level step in many computer vision applications such as video surveillance and traffic monitoring. In this paper, the Gaussian Mixture Model (GMM) is used to extract the silhouette images using the background subtraction model [10].

The GMM is the probability density function where the density is composed of a sum of Gaussians. It is also described that the background modelling by Gaussian mixtures is a pixel based approach. The pixel distribution is modeled as a mixture of Gaussians. Let k be a random process which represents the value of a given pixel in time. Then the probability density function $p(k)$ of a Gaussian mixture comprising n component densities is described as

$$p(k) = \sum_{n=1}^{N} w_n \eta(k, \mu_n, \sigma_n) \tag{1}$$

Where, w_n denote the weights, and η describes the Gaussian probability density of mean μ_n and covariance matrix $\Sigma_n = \sigma_n I$ (I denotes the identity matrix). The mixture of Gaussian algorithm estimates these parameters over time to obtain a robust representation of the background. Initialization of the parameters is done with $w_n = w_o$, $\mu_k = \mu_o$ and $\sigma_n = \sigma_o$.

The matching of foreground and background objects is described as:

$$\frac{\left\| k - \mu_j \right\|}{\sigma_j}, \quad j \in [1 \ldots N]$$

For some threshold value $\tau(\rangle 0)$, the parameters of the Gaussian mixture are updated as follows:

$$w_n(t) = (1-\alpha) w_n(t-1) + \alpha M_n(t)$$

$$\mu_n(t) = (1-\beta)\mu_n(t-1) + \beta k$$

$$\sigma_n^2(t) = (1-\beta)\sigma_n^2(t-1) + \beta \left\| \left(k - \mu_n(t)\right) \right\|^2$$

For the matching component j, $M_n(t)$ is equal to 1. If the component does not match, then it is equal to 0. The component with lowest weight is re-initialized with $W_n = W_o$, $\mu_n = k$ and $\sigma_n = \sigma_o$. The learning coefficient α is constant and β is defined as

$$\beta = \alpha\, \eta\left(k, \mu_n, \sigma_n\right)$$

The normalization of weights w_n is performed to add up to 1. The weight-to-standard-deviation ratio w_n / σ_n is decreased to sort the Gaussians and represent the background [10]. To find the set {1...B} of Gaussians modelling the background, a threshold λ is applied to the cumulative sum of weights

$$B = \operatorname{argmin}_{MB} \left(\sum_{m=1}^{MB \le m} w_n > \lambda \right)$$

Gaussians with the highest probability of occurrence w_n and lowest variability in the distribution, measured by σ_n indicates a representative mode to model the background. The increase of learning rate β and large variations of the pixel value distribution are considered for the saturation of pixels. Indeed, the update of the variance with the square of the difference between the mean and the pixel value typically, result in large and overestimated values for the variance. Finally, the pixels are classified as either foreground or background, depending on the weight of the Gaussian component. GMM can cope with multimodal background distribution and it is able to filter noise during image differentiation which in turn provides a selective level of detail for the contour of the moving shapes.

3. Feature Extraction

A feature extraction technique deals with the best way of tracking information from of an image for the desired analysis. The features are extracted from foreground objects using PCA. It is also known as eigen analysis, to represent most of the variation of original variables using a few principal components. It is a technique used to reduce the dimensionality of data. It converts a set of observations of possible correlated variables into a set of values of linear uncorrelated variables called principal components by an orthogonal transformation [10]. While retaining the possible variation present in dataset, the PCA will reduce the dimensionality of the data correspondingly. PCA is a statistical analysis of data to compute a linear transformation that maps data from a high to a lower dimensional space. PCA is a powerful tool for analysing data and has been used successfully in recognition techniques. In PCA, most of the information is carried in the variance of the features and the image sequences are aligned using mathematical approaches. It is used to resize the different sizes of images to same size.

Using PCA method, 1-D image vectors are formed by concatenating 2-D gait images. Then the zero mean 1-D training images are obtained. PCA applied on the collection of 1-D zero-mean image set vector will further produce a low-dimensional features vector [11].

PCA is used to identify the patterns in data, and also to express the data to highlight their similarities and differences. The data cannot be represented graphically as patterns in data are in high dimensional space. Simplifying the data structure will account for as much of the total variation in the original data as possible. The Empirical mean is given by

$$u(m) = \frac{1}{N}\sum_{n=1}^{N} X(m, n)$$

Where, the mean vector u is of dimension M x 1. The deviations from the mean can be calculated, by subtracting the empirical mean vector u from each column of the data matrix X which is described as $K = X - \mu h$, where K be the M*N matrix and h is a 1*N row vector of all 1's. The covariance matrix is expressed as:

$$C = \frac{1}{N}\sum K.K^{*}$$

The goal of using PCA is to extract most of the variation of the original variables as feature vectors and reduce the amount of computation needed for accurate reconstruction of the data. To calculate the PCA, a set of feature vectors are created by placing all the original data for a given configuration in a single vector correlation matrix C. The correlation matrix C is a symmetric matrix that helps to reduce the computation when calculating the eigenvectors and eigenvalues.

$$C = \frac{1}{N}\sum_{n=1}^{N}(x_n - x_k)(x_n - x_k)^T$$

The mean value of the vector is defined as

$$X_k = \frac{1}{N}\sum_{n=1}^{N} X_n$$

A set of special vectors is then used to represent the correlation matrix, which satisfies the following equation:

$$C e_x = \lambda_x e_x$$

These vectors are described as eigenvectors and each eigenvector e_x has an associated eigenvalue λ_x. The representation of the largest inherent variation in the original data set is given by the largest eigenvalues of the correlation matrix of the original data. If it is required then the eigenvectors can be rearranged back into the form at the end of the process. The feature vectors are projected into the new eigenvector space and then finally used for classification. The limitation of PCA is that it eliminates the dimension that is best for discriminating positive from negative cases, as it is an unsupervised algorithm.

3. Classification

Once gait information is extracted from gait sequences and projected into a feature space, the next step is to perform the recognition by pattern classification. Two main approaches can be taken, namely, a template-based approach or a stochastic approach. In both cases, an appropriate distance metric between feature vectors must be initially defined. The classical Euclidean distance is the measure that is used in most gait recognition applications.

3.1 Multi-Layer Perceptron

The MLP is a feed forward neural network uses the back propagation algorithm to minimize the mean squared error function. It recognizes the extracted silhouette images from the stored database. It can train the network with differentiable transfer function to perform approximate non-linear functions and pattern classification. The network as shown in Figure 3 is designed with three layers of units typically, one input layer, one hidden layer and one output layer. The artificial neurons are organized such that input layer is connected to the hidden layers and the hidden layers units fully connected to units in the output layer. The error is then back propagated to the hidden unit. The training of the MLP network is done in three stages:

1. Feed forward of input training pattern.
2. Back propagation of the error.

3. Updation of weights.

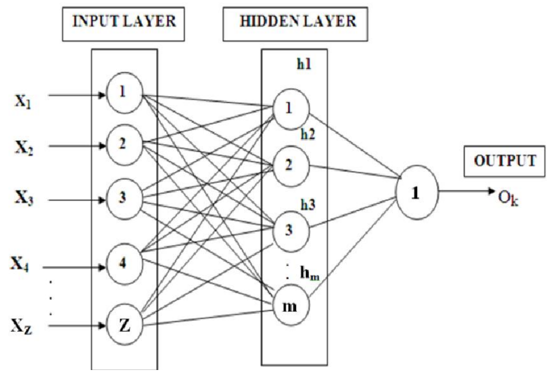


Figure 3. Back Propagation Neural Network

The net Input to the hidden unit is

$$net_m = \sum_z x_z w_{zm}$$

The output of hidden unit after passing through the activation function is

$$h_m = \frac{1}{1+\exp(-net_m)}$$

The net input to the output unit is given by

$$net_k = \sum_m h_m w_{mk}$$

Error function for output of the each neuron is

$$E = \frac{1}{2}\sum_{i=l}^{k}(o_i - t_i)^2$$

Weight change between hidden unit and output unit

$$\Delta w_{mk} = \beta\delta_k h_m$$

Weight change between input unit and hidden unit

$$\Delta w_{zm} = \alpha\delta_m x_z$$

δ_k , δ_m represent the error factors to distribute the error back to the hidden unit and input unit.

Each output and hidden unit updates its weights

$$w_{mk}(\text{new}) = w_{mk}(\text{old}) + \Delta w_{mk}$$

The weights are updated until error reaches a minimum value. The network was trained with the appropriate selection of parameters such as number of input units, number of hidden layers, learning rate and epochs, for their efficient operation. The gradient decent is very small if the learning rate is small and oscillates widely if it is too large. The learning rate parameters of α and β in the range between 0 and 1 is used to scale the adjustments from a previous iteration to avoid major disruption of direction of learning. It influences the speed and quality of learning.

3.2 FLDA classifier

The Discriminant Analysis will classify objects into one of two or more groups on the basis of set of features that describe the objects. It can classify an object to one of a number of predetermined groups based on the analysis made on the object. The feature selection and classification are the main task of FLDA. The dependent variable (Y) and independent variables (X) are the group and object features of discriminant analysis to make it more reliable for classification.

FLDA is used to find a mapping from the high-dimensional space to a low-dimensional space in which the most distinguishing features are preserved [12]. It achieves this by minimizing the variation within the same class and maximizing the variation between classes. The low dimensional vector is represented as x_k, the mean value of class x_i as m_i, and the mean value of all data as m. The between class scatter matrix S_B can be calculated as

$$S_B = \sum_{i=1}^{c} n_i (m_i - m)(m_i - m)^T$$

The number of classes is denoted as c and n_i represents the number of samples in class x_i. The within class scatter matrix is defined as

$$S_W = \sum_{i=1}^{c} \sum_{x_k \in C_i} (x_k - m_i)(x_k - m_i)^T$$

LDA tries to find a projection direction that maximizes the ratio of between-class scatter to within-class scatter.

$$W_{FLD} = \arg\max_W \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|} = [w_1 w_2 .... w_n]$$

Optimal projection Matrix is described as

$$W_{opt}^T = W_{fld}^T W_{pca}^T$$

$$W_{pca} = \arg\max_w \left|W^T S_T W\right|$$

W_{opt} represents the optimal projection matrix. In its columns, it contains the generalized eigenvectors that correspond to the largest eigenvalues.

$$S_B W_i = \lambda_i S_W W_i, i = 1,2,...m$$

If S_W is non-singular, W_{opt} can be calculated by simply computing eigenvectors of S− 1 W_{SB}. Instead of computing eigenvectors of $S-1W_{SB}$, it is preferable to diagonalize S_W and S_B.

4. Results and Discussion

Experimental simulations were conducted to evaluate the performance of the proposed system using CASIA gait database, USF gait database and also on real database. The input video stream contains frame width and height of about 320 and 240 respectively to capture the gait sequence. The frame rate is chosen as 30 frames per second. The extracted binary silhouettes images from the gait sequence using GMM is shown in Figure 7. The isolated foreground objects from input video file which is obtained with the help of Gaussian parameters.

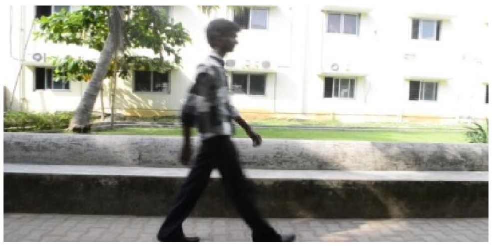


Figure 4. Input video image

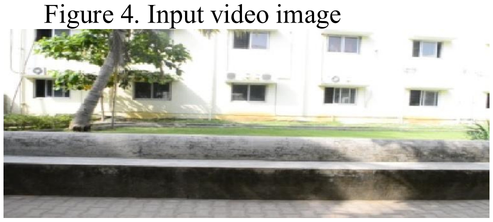
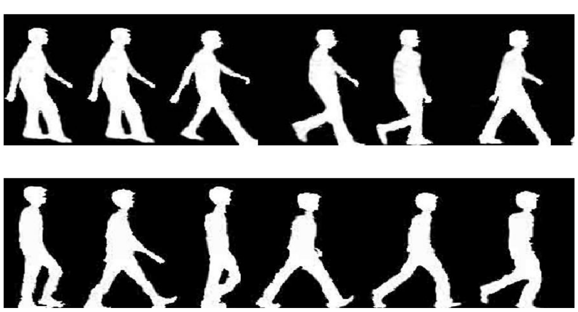


Figure 5. Background image

Figure 6. Difference image



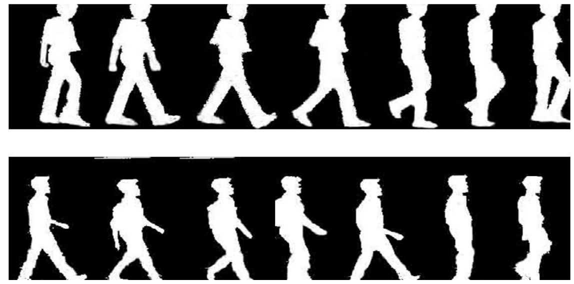


Figure 7. Samples of extracted silhouette images

PCA is then used to extract the silhouette image features and transforms a high dimensional data space into a lower dimensional data space. After updating the number of classes and number of images in each class along with the total number of training images, the mean image shown in Figure 8 is calculated from one dimensional image vectors.

Table 1. Performance comparison of gait recognition with different classifiers

Training Set	Recognition Rate %		
	PCA with Euclidean Distance Classifier	PCA with FLDA Classifier	PCA with MLP Classifier
1	95%	97%	98%

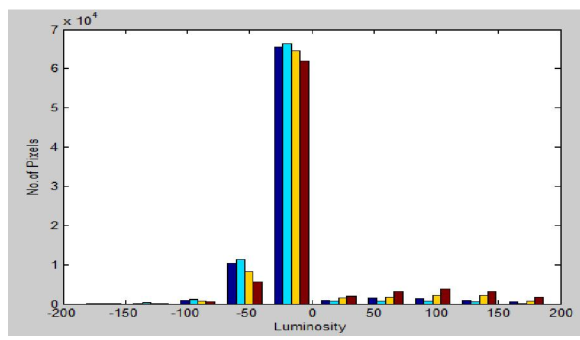


Figure 8. Histogram of mean image

In the proposed recognition algorithm the MLP classifier combined with PCA had better recognition performance compared with the other classifiers as shown in Table.1. The training set I of real database consists of video sequence of 10 subjects captured under normal working conditions. The training set II of USF gait database and training set III of CASIA gait database consists of the sequence of silhouette of 25 subjects captured from different views are recorded with 30 frames/second. The classification rate measured with the Cumulative Match Score (CMS) achieved 98% in the proposed method.

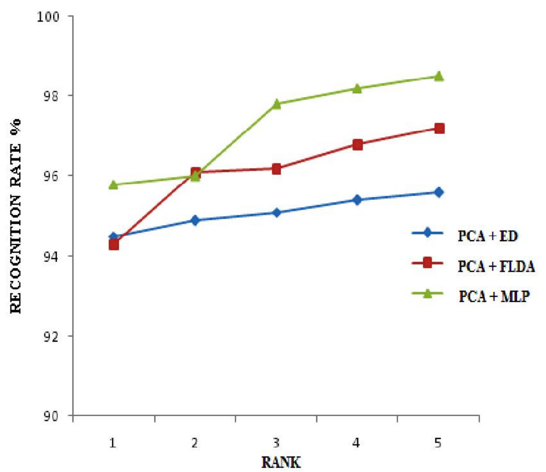


Figure 9. CMS curve of recognition result with different training methods

5. Conclusion

In this paper, GMM model is presented to extract the silhouette images and remove the background image from the video file with the help of Gaussian parameters. Then the features extracted from the image frames are projected into the Eigen space and the dimensionality reduction is done using principal component analysis. MLP classifier combined with PCA provides optimal linear dimensionality reduction and finally results in minimal process time. The proposed method tested using gait database achieved the recognition rate of 98%.The experimental results demonstrated that our proposed method outperforms other gait recognition methods in both accuracy and speed. In the future, we can combine gait with other biometric features to increase the recognition rate and accuracy.

Corresponding Author:
C.Murukesh
Department of EIE
Velammal Engineering College
Anna University, Chennai, India
E-mail: pcrmurukesh@gmail.com

References

1. Changhong Chen, Jimin Liang, Xiuchang Zhu. Gait recognition based on improved dynamic Bayesian networks. Pattern Recognition, 2011; 44: 988-995.
2. U. Chandrasekhar, Tapankumar Das. A survey of techniques for background subtraction and traffic Analysis on surveillance video. Universal Journal of Bioinformatics and Biotechnology, 1, pp: 107-113.
3. G.Venkata Narasimhulu, Dr. S. A. K. Jilani. Gait recognition: A survey. International Journal of Electronics Communication and Computer Engineering. 2012; 3(1):33-38.
4. Darko S. Matovski, Mark S. Nixon, Sasan Mahmoodi, John N. Carter. The effect of time on gait recognition performance. IEEE Trans. Information Forensics and Security.2012; 7: 543-552.
5. Junping Zhang, JianPu, Changyou Chen, Rudolf Fleischer. Low-resolution gait recognition, IEEE Trans. Systems, Man and Cybernetics.2010; 986-996.
6. L.R.Sudha, R.Bhavani. Gait based gender identification using statistical pattern classifiers. International Journal of Computer Applications. 2012; 40(8):30-35.
7. Toby H.W. Lam, K.H. Cheung, James N.K. Liu Gait flow image: A silhouette-based gait representation for human identification, Pattern Recognition. 2011; 44(4):973-987.
8. Chien Wen Cho, Wen Hung Chao, Sheng Huang Lin, You Yin Chena. A vision-based analysis system for gait recognition in patients with parkinson's disease. Expert Systems with Applications.2009; 36:7033-7039
9. Kantipudi M.V.V Prasad, V. Sailaja, A. Jagan. Background subtraction and target classification for gait recognition.Int.J.Comp.Appl.2 (3):515-520.
10. Yang Ran, Qinfen Zheng, Rama Chellappa, Thomas M. Strat. Applications of a simple characterization of human gait in surveillance. IEEE Trans. Systems, Man, and Cybernetics.2009; 1009-1020.
11. M.Pushparani, D. Sasikala. A survey of gait recognition approaches using PCA & ICA.Global Journal of Computer science and Technology.2012; 12(10).
12. Nikolaos V. Boulgouris, Zhiwei X. Chi. Gait recognition using Radon transform and Linear Discriminant Analysis. IEEE Trans. Image Processing.2007; 16(3): 731-740.
13. Maodi Hu, Yunhong Wang, Zhaoxiang Zhang, De Zhang. Gait-based gender classification using mixed conditional random field. IEEE Trans. Systems, Man, and Cybernetic.2011; 41(5): 1429-1439.
14. Sandhitsu Das, MaciejLazarewicz, Leif H. Finkel. Principal Component Analysis of temporal and spatial information for human gait recognition. IEEE Conference.2004; 2: 4568-4571.
15. Jussi Tohka. Introduction to Pattern Recognition, SGN-2506, Tampere University of Technology, 2011.

6/22/2013