International Journal of Advances in Computer Networks and its Security

Person Identification Using Extracted Vectors

Kantipudi MVV Prasad, Dr.V.Sailaja, A. Jagan, Dr.D.S.Rao
Department of Electronics and Communication Engineering
[1,2]GIET, Rajahmundry, [3,4] MMU, Mullana.
Email:prasad.rku@gmail.com

Abstract:-The main objective of my paper is to propose a new spatio temporal gait representation scheme to characterize human walking properties for individual recognition by gait in a single camera-based setup. In the proposed method, width vector of outer contour of binary silhouette and ART (Angular Radial Transform) coefficients are taken as the feature vector. PCA is applied to remove correlation between the features and also to reduce its dimensionality of the features. These extracted feature vectors are used to recognizing the individuals. Hidden Markov Model (HMM) is used to recognize the individual. Some intermediate experimental results shows the effectiveness of proposed system.

I. INTRODUCTION

Today, in metropolitan public transport stations, authentication or verification using conventional technologies is practically infeasible. In such type of applications, biometric authentication methods are more attractive. Biometrics are the unique features of a person. Biometric recognition refers to an automatic recognition of individual based on feature vectors derived from their physiological and/or behavioral characteristic. Biometric characteristics can be classified into two types:

• Physiological: These characteristics are related to the body. Recognition techniques come in to this category are fingerprint, face, iris, DNA and palm geometry.

• Behavioral: These are related to the behavior of the person. Voice and gait recognitions techniques comes in to this category

An important limitation of these biometric recognition is that it require co-operation of the individual that is to be identified.

II. GAIT RECOGNITION:

Biometric systems for human identification at distance have ever been an increasing demand in various significant applications. In such type of applications, generally used biometrics resources in the form of iris, fingerprint, palm print, hand geometry are suffers from two main disadvantages: 1) Failure to match in low resolution images and 2) Necessitates user cooperation for accurate results. For these reasons, innovative biometric recognition methods for human identification at a distance have been an urgent need for surveillance applications. Recognition using gait becomes more attractive in such type of situations. Human gait is a spatio-temporal phenomenon that characterizes the motion of an individual. The definition of gait is "A particular way or manner of walking on foot". Human gait recognition works from the observation that an individual's walking style is unique and can be used for human identification. Using human gait as biometric is a relatively new area for Computer Vision researchers. Extensive work has been carried out in the psychophysics community on the ability of humans to recognize others by their style of walking. In the Computer Vision community, research on gait has concentrated mostly on recognition algorithms. Depending on feature extraction, gait recognition methods are classified as appearance-based and model-based gait recognition. The appearance-based approaches suffer from changes in the appearance owing to the change of the viewing or walking directions. But, model-based approaches extract the motion of the human body by means of fitting their models to the input images. Model-based methods are view and scale invariant.

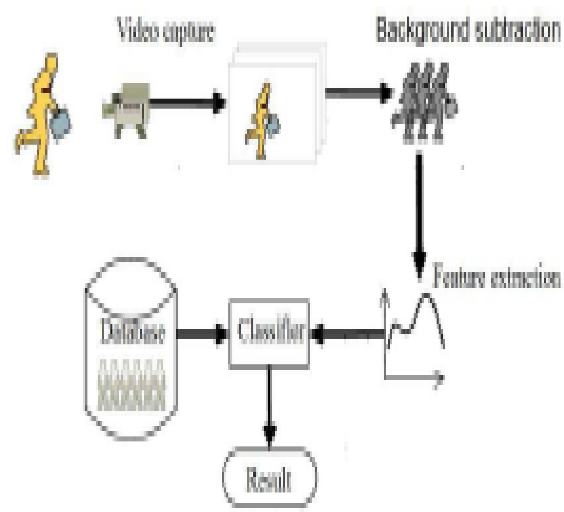

Figure 1: Block diagram of gait recognition system

SCENARIO:-

Gait recognition scenario is shown Figure 2. It can be explained by analysis of video stream obtained from surveillance cameras. If any unauthorized individual walks in front of the camera, system will compare his gait with stored gait sequences and recognize him and alerts the appropriate authorities for necessary action.

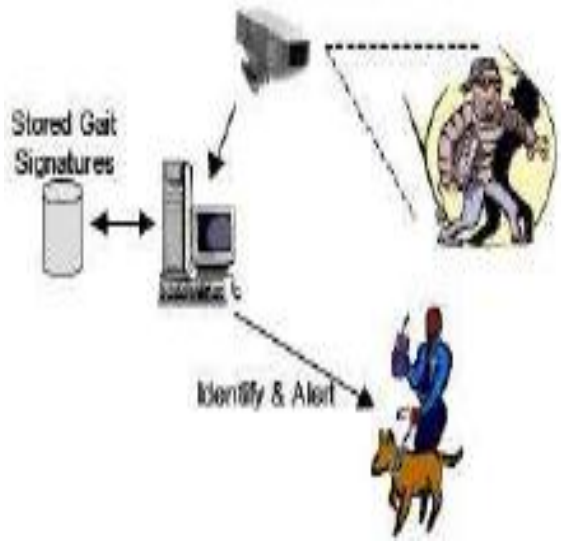

Figure 2: Gait recognition scenario [18]

III. BACKGROUND SUBTRACTION

Identifying moving objects from a video sequence is a fundamental task in Gait recognition. A common approach is background subtraction in which moving objects from background in the scene are identified. Pixels in the current frame that deviate significantly from the stationary background are considered to be moving objects.

Background subtraction techniques are classified in to two types

1. Non-recursive methods

2. Recursive methods

Non-recursive methods: Non-recursive techniques uses a sliding window approach for background estimation. It stores a buffer of the previous L video frames, and estimates the background image based on the temporal variation of each pixel within the buffer. Disadvantage of this method is memory storage requirement.

Recursive methods: In these methods, background model is recursively updated based on each input frame. Recursive techniques require less storage.

IV. FEATURE EXTRACTION

An important step in gait recognition is the extraction of appropriate feature that will effectively capture the gait characteristics. The features must be reasonably robust to operating conditions and should yield good discriminability across individuals. Commonly used feature selection approaches are:

1. Model-based approaches

2. Holistic approaches

Model-based approaches: Model-based approaches employ models whose parameters are determined by processing of gait sequences (binary silhouettes). These methods are scale, view invariant and requires good quality video sequences. In these methods, parameters used as features are the height, the distance between head and pelvis, the maximum distance between pelvis and feet and the distance between feet. In [1], the silhouette of a walking person is divided in to some regions (generally seven regions). Subsequently, ellipses or rectangles are fit to each region and region feature vectors are determined. This includes averages of the centroid and the aspect ratio. Figure 3(a) shows the examples for model-based approaches.

Holistic approaches: Holistic methods operate directly on binary silhouettes without assuming any specific model for the walking human. The contour of the silhouette is the most reasonable feature in this method. For high quality binary silhouettes, width of outer contour of the silhouette was proposed as a suitable feature. For low quality binary silhouettes, the binarized silhouette may be is used as a feature. Example of holistic approaches is shown in Figure 3(b).

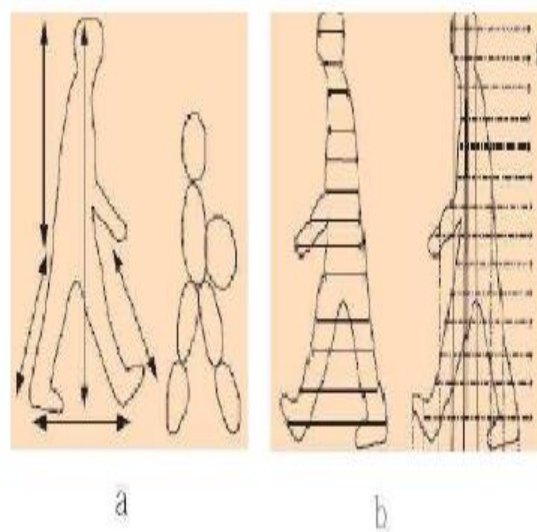

Figure 3: Examples of feature selection methods. (a) model-based approach and (b) Holistic approach [2]

Recognition:- This is the final step of gait-based person identification. Here, input test video sequences are compared with the trained sequence in the database. In general, minimum distance classifier may be used for gait recognition.

Feature extraction and PCA training: Binary silhouette obtained from background subtraction is used as the feature. For computational efficiency, 2D silhouette is converted into associated sequence of 1D signals to approximate temporal pattern of gait. The silhouette representation is shown in Figure 4.

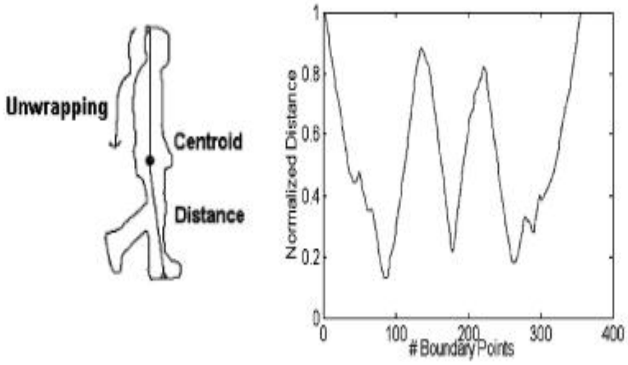

Figure 4: Silhouette representation [10]

Apparently, centroid (x_c, y_c) of the binary silhouette can be calculated from the information of the outer contour of the silhoueete. By choosing the centroid as reference origin, the outer contour counter is unwarped clockwise to get a distance signal S = {d_1, d_2, d_i, d_{nb}} that is composed of all distances di between each boundary pixel (x_i, y_i) and the centroid. The distance is given by:

$$d_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}$$

Again, for training of s classes, each class can be represented by a sequence of distances. Let D_{ij} be the j^{th} distance signals in class i and Ni the number of such distance signals in the i^{th} class. The total number of training samples is $N_t = N_1 + N_2 + ... + N_s$ and the whole training set can be represented by [$D_{1,1}$,$D_{1,2}$, ...,D_1,N_1 ,$D_{2,1}$...,D_s,N_s]. The mean m_d and covariance matrix P are calculated as,

$$m_d = \frac{1}{N_t}\sum_{i=1}^{s}\sum_{j=1}^{N_i} D_{i,j}$$.

$$\sum = \frac{1}{N_t}\sum_{i=1}^{s}\sum_{j=1}^{N_i}(D_{i,j} - m_d)(D_{i,j} - m_d)^T$$

If the rank of the matrix ∑ is N, then N non zero eigenvalues, λ_1, $\lambda 2 \ldots \lambda_n$ and Corresponding eigenvectors e_1, $e2 \ldots e_n$ are obtained based on SVD (Singular Value Decomposition). Generally, first few eigenvectors correspond to large Changes in training patterns. So, those small eigenvalues and corresponding eigenvectors are ignored using a threshold value T_s. Taking only K < N largest eigenvalues and their corresponding eigenvectors, the transform matrix E = [e_1, $e2 \ldots e_K$] can be constructed to project an original distance signal D_{ij} into a point P_{ij} in the k-dimensional eigenspace. For each training sequence, the projection centroid C_i in the eigenspace is given by,

$$C_i = \frac{1}{N_i}\sum_{j=1}^{N_i} P_{ij}$$

Recognition is carried out by measuring similarities between reference patterns and test samples in the parametric eigenspace. In this, nearest neighborhood or spatio-temporal correlation may be used to measure similarities.

Representing silhouette using wavelet descriptors and ICA: Wavelet descriptors have been established and proved as a good method for representing two dimensional shape's boundary. Major advantage is it is robust against rotation, scale and linear transformation. When one particular shape is represented in wavelet domain, its frequency components can be easily obtained. The general features of the shape are located in the lower frequencies; the detail features are located in the higher frequencies. Each point on the contour can be represented by complex number. N is the number of contour points, one dimensional vector can be computed as

$$d_i = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}$$

Where (x_c, y_c) is the centroid of human boundary. Value of N, may be selected as 256. Let us denote observed variable xi as a vector with zero mean random variables X = $(x_1, x_2, x_3, ...x_n)^T$. The component variables s_i as a vector S = $(s_1, s_2, s_3, ...s_n)^T$ with the model is

$$X = AS$$

To reduce computational cost, an algorithm named Fast ICA using a fixed point iteration algorithm finding local extrema of a linear combination of the observed variables may be used [11]. The matrix X contains n individual persons and each person has m frames, a_{ij} represent entry at the ith row, j^{th} column. The value SB_j , which is called the mean of within the class distance in the j^{th} column, is given by

$$SB_j = \frac{1}{mn(m-1)}\sum_{i=1}^{n}\sum_{u=1}^{m}\sum_{v=1}^{m}\left(a_{(i-1)m+u,j} - a_{(i-1)m+v,j}\right)^2$$

$$SI_j = \frac{1}{n(m-1)}\sum_{s=1}^{n}\sum_{t=1}^{m}\rho(\bar{a}_{s,j} - \bar{a}_{t,j})$$

Is the mean between class distance in the j^{th} column.

$$SB_j = \frac{1}{mn(m-1)} \sum_{i=1}^{n} \sum_{u=1}^{m} \sum_{v=1}^{m} \left(a_{(i-1)m+u,j} - a_{(i-1)m+v,j}\right)^2$$

$$SI_j = \frac{1}{n(m-1)} \sum_{s=1}^{n} \sum_{t=1}^{m} \rho(\bar{a}_{s,j} - \bar{a}_{t,j})$$

$$\bar{a}_{i,j} = \sum_{u=1}^{m} a_{(i-1)m+u,j}$$

Recognition: Classification process carried out through two different methods, the Nearest Neighbor (NN) and Support Vector Machine (SVM). In NN classifier, euclidian distance is applied to evaluate discriminatory of the gait Sequences. SVM have high generalization capability in many tasks especially pattern analysis. SVM is based on structural risk minimization which is the expectation of the test error for the trained machine. Several kernel functions are used in SVM. In this, SVM classifier is a 2-class classifier and there are two options for user: one is using N number of SVMs and another one is separating one class from the rest or using N(N − 1)/2 SVMs are for each pair of class.

Proposed Approach:

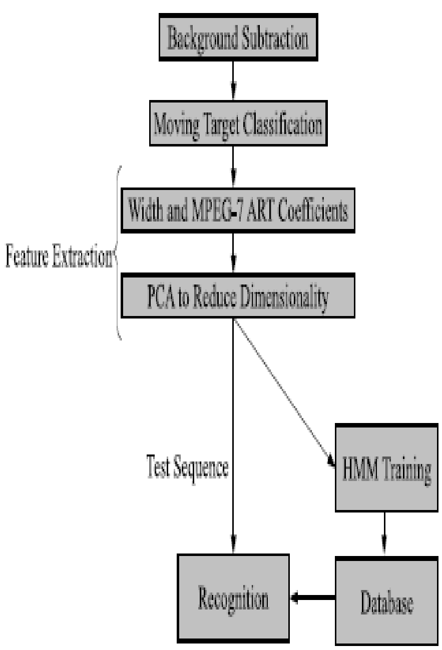

Figure 5: Flow diagram of proposed approach.

V. Moving Target Classification

In this, moving objects obtained from background subtraction are classified in to human, vehicle and background clutter. A classification metric operator ID(x), notion of temporal consistency is used for classification. The metric is based on the knowledge that in general humans are smaller than vehicles. In this dispersedness is taken as classification metric and is given by,

$$Dispersedness = \frac{Perimeter^2}{Area}$$

VI. Feature Selection

Feature selection is a crucial step in gait recognition. The feature must be robust to operating conditions and should yield good discriminability across individuals. Each gait sequence is divided into cycles. Gait cycle is defined as person starts from rest, left foot forward, rest, right foot forward, rest. Figure 6 shows the stances during gait cycle. Gait cycle is determined by calculating sum of the foreground pixels. At rest positions this value is low. By calculating number of frames between two rest positions, gait cycle (period) is estimated. Figure 7 shows the sum some of the foreground pixels of two persons. x axis denotes frame index, y axis denotes sum of foreground pixels. Valley points represent rest position. In the proposed method, two types of features are extracted.

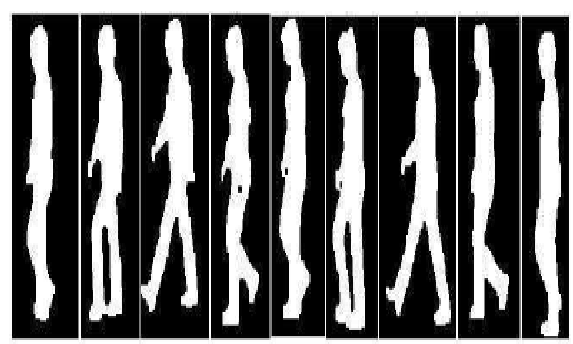

Figure 6: Stances during a gait cycle

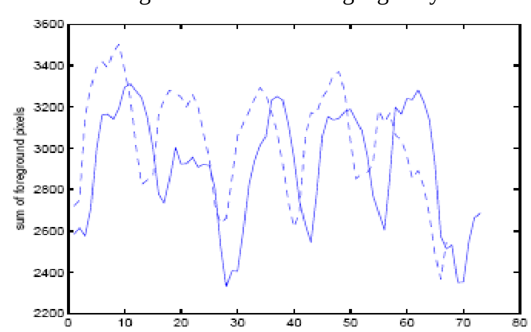

Figure 7: Sum of foreground pixels to estimate gait cycle (gait period)

Width of the outer contour: In this, width of the outer contour of binary silhouette is extracted as feature. Distance between left and right extremities

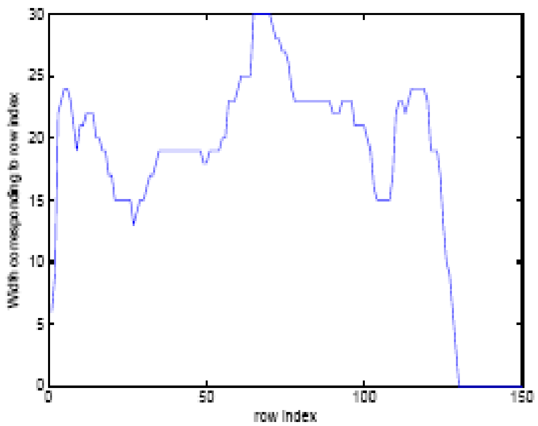

Figure 8: Width vector of person #1

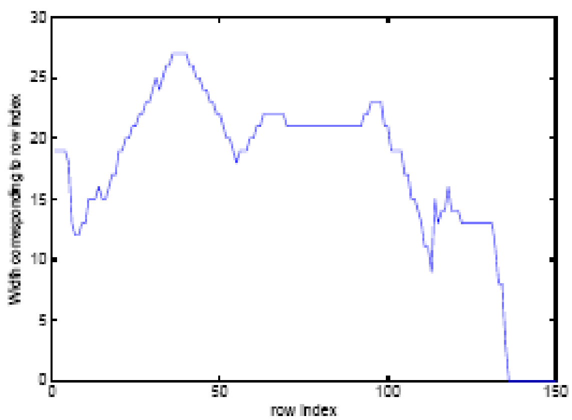

Figure 9: Width vector of person #2

of the silhouette gives the width vector. From the binarized silhouettes, the left and right boundaries are traced. The width along a given row is simply the difference between leftmost and rightmost boundary pixels in that row Figures 8, 9 shows the feature (width) vector for one frame of two persons. X axis denotes the row index and y axis denotes the width associated with that row. Figure 10, 11 shows the feature vectors (width profile) for several gait cycles. X denotes frame index, y denotes index of the width vector (row index). Brightness parts represents maximum width regions such as hands swing, distance between feet. By observing both images, it is cleared that one person has more brightness value at upper region compared to other. Upper regions correspond to swings of the hand, bottom regions correspond to swings of the extremities of the foot. As the distance between the left and right extremities of the silhouettes is used as feature, the two halves of the gait cycle are almost indistinguishable. So, in this project, half gait cycle is considered as one cycle.

MPEG-7 Angular Radial Transform (ART) coefficients: This is the second feature vector used in this project. MPEG-7 based-region-based descriptors used to represent shapes. This descriptor takes into account all pixels constituting the shape that is both the boundary and interior pixels. The descriptor

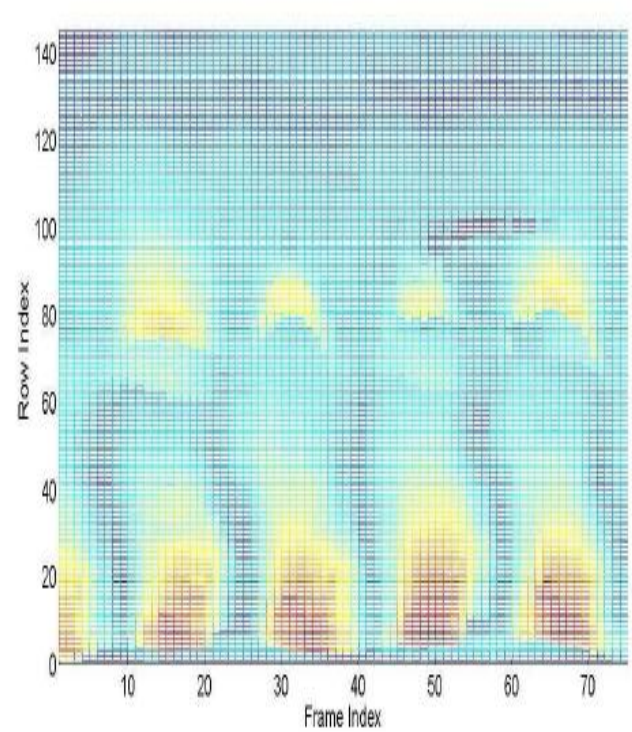

Figure 10: Width vector profile of person #1.

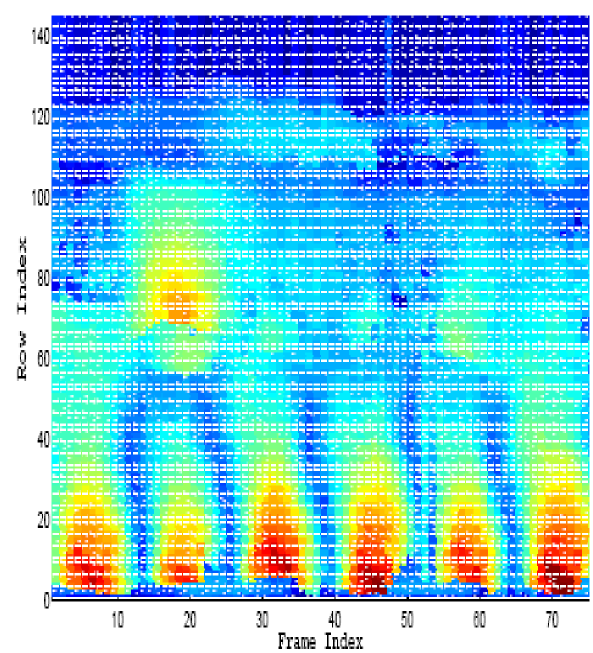

Figure 11: Width vector profile Width vector profile of person #2.

works by decomposing the shape in to 2-D basis functions (complex-valued), defined by ART. The normalized magnitude of coefficients is used to describe the shape. The ART coefficients are defined by:

$$F_{nm} = \langle V_{nm}(\rho,\theta), f(\rho,\theta)\rangle = \int_0^{2\pi}\int_0^1 V^*_{nm}(\rho,\theta) f(\rho,\theta)$$

Where F_{nm} is an ART coefficient of order n and m, f (ρ, θ) is an image function in polar coordinates and V_{nm} (ρ, θ) is the ART basis function that are separable along the angular and radial direction, that is

$$V_{nm}(\rho,\theta) = A_m(\theta)\, R_n(\rho)$$

In order to achieve rotation invariance, an exponential function is used for the angular basis function,

$$A_m(\theta) = \frac{1}{2\pi}\exp(jm\theta)$$

The radial basis function is defined by a cosine function,

$$R_n(\rho) = 1, n = 0$$

$$R_n(\rho) = 2\cos(\pi n\rho), n \neq 0$$

The ART descriptor is defined as a set of normalized magnitudes of complex ART coefficients. Rotational invariance is obtained by using the magnitude of the coefficients. Twelve angular and three radial functions are used (n < 3,m < 12). Total thirty six coefficients obtained to represent particular shape. ART coefficients are normalized by dividing with magnitude of ART coefficient of order n = 0,m = 0. Similarity between two shapes described by the ART coefficients is calculated using L-1 norm. After getting the feature vectors, next step is classification/recognition. In our method, HMM-based recognition is used. For each person one HMM model is developed at training stage. The primary HMM parameters used are number states, initial probability (π), the transition probability (A), output probability (B). From the training sequence feature vectors HMM model is estimated as π = (A,B, π). The parameters of HMM are described below:

1. N, the number states in the model. Number of states used in this are N=5. HMM states are represented as S = {s_1, s_2,.., s_N}.

2. M, the number of distinct symbols (observation features) per each state. Each feature vector is treated as one observation symbol. The number M depends on the number of frames per cycle, the number of states in the model and how to divide one cycle in to clusters. The observation symbols for one HMM state are denoted as V = {v_1, v_2, ..., v_M}.

3. A, the transition probability matrix. A = {a_{ij}}, a_{ij} is defined as

$$a_{ij} = P\left[q_{t+1} = S_j | q_t = S_i\right], 1 \le i, j \le N$$

Where q_t is the state at time t. In this, left to right model is chosen, which allows the transition from j^{th} state to either j^{th} or the $(j + 1)^{th}$ state. The forward HMM is shown in Figure 12.

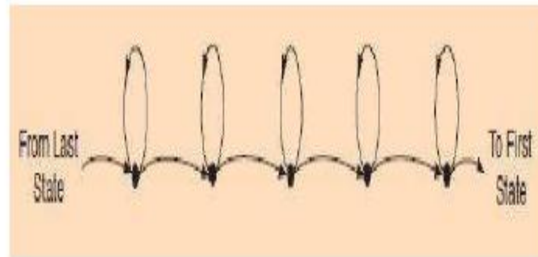

Figure 12: State traverse in forward HMM [2]

4. B, the observation symbol probability matrix. B = {b_j (k)}, where

$$b_j(k) = P\left[v_k \text{ at } t | q_t = S_j\right], 1 \le j \le N, 1 \le k \le M$$

5. π, the initial probability. π= {π_i}, where

$$\pi_i = P\left[q_1 = S_i\right], 1 \le i \le N$$

First frame is assigned to first state only, so the initial probability _1 is set be 1 and remaining π_i are set to be zero. The complete parameter set of the HMM can be denoted as

$$\lambda = (A, B, \pi)$$

In the proposed method, model parameters are estimated by selecting the states of a gait sequence. During gait cycle, every person transits across some phases or stances. These stances are used as states. N exemplars (stances) ∑" = {e1,e2,..., eN} are picked up from a pool training features to get states. Exemplars are show in Figure 13.

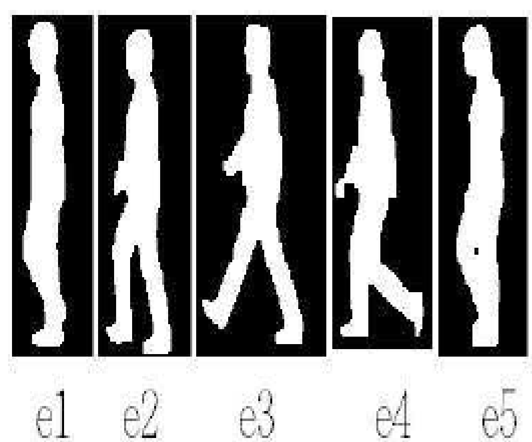

Figure 13: Exemplars (stances) during a gait cycle (gait period)

VII. Experimental Results

The first row shows the the input video frames. Second row in those figures shows the background subtraction results without median filter. Third row shows the result after using median filter. It is cleared that some noises are removed by the median filtering operation.

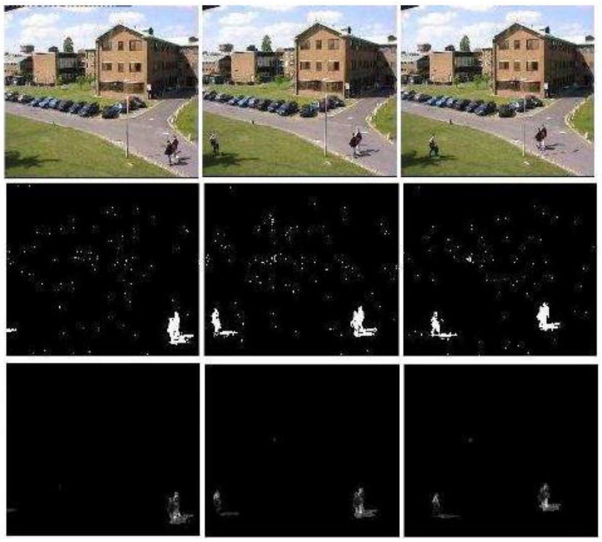

Figure 14: Background subtraction results. The first row shows input videoframes(Pets2001:http://www.cvg.cs.rdg.ac.uk/PETS2001/pets2001-dataset.html), second row shows background subtracted frames and third row shows the background subtraction results after median filtering.

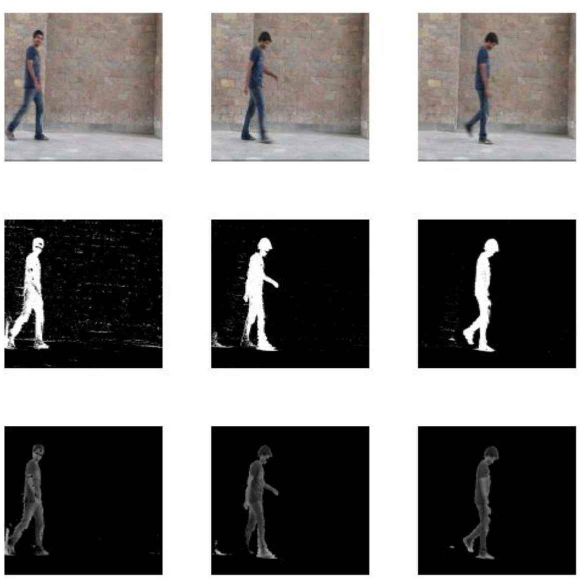

Figure 15: Background subtraction results of data set. The first rowshows input video frames, second row shows background subtracted frames and third row shows the background subtraction results after median filtering.

Connected Component Labeling: Connected component is used to group the pixels which have similar properties and connected in some way. Connected component labeling is applied to background subtraction results. Figure16 shows connected component labeling results.

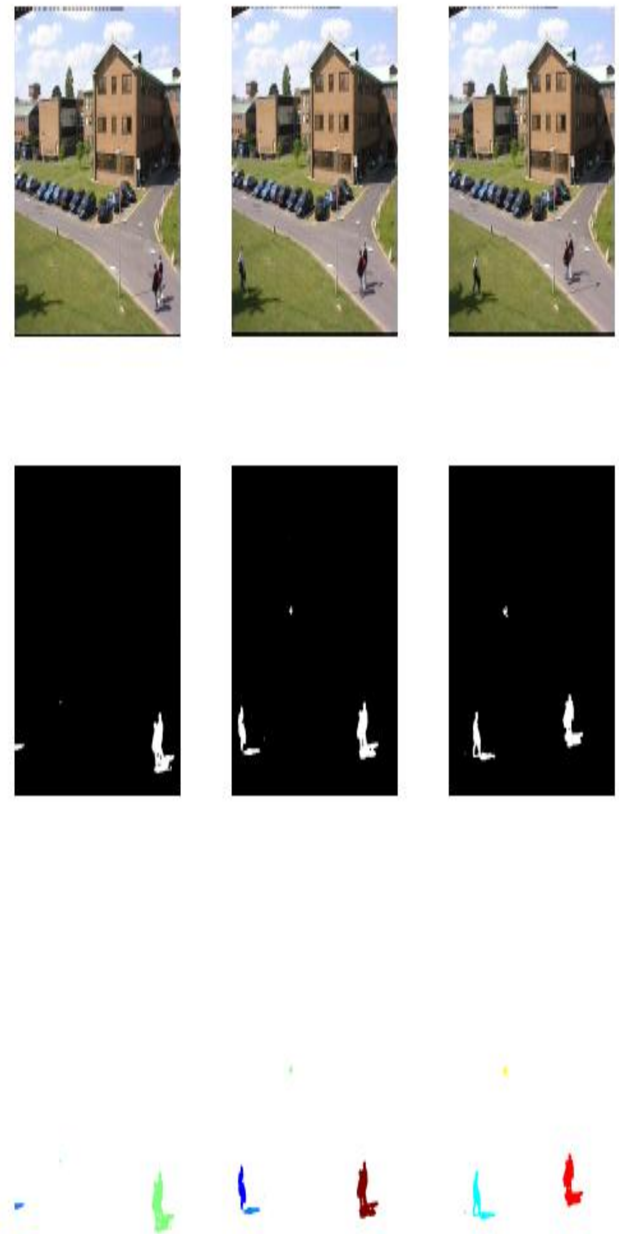

Figure16: Connected component labeling results. The first row shows inputvideoframes(Pets2001:http://www.cvg.cs.rdg.ac.uk/PETS2001/pets2001-dataset.html), second row shows background subtracted frames and third row shows results after the connected component labeling.

Moving Target Classification: Moving target classification is used to classify humans from other moving objects which are obtained from background subtraction. A dispersedness criterion is taken as the classification metric. A figure 17, 18 shows the obtained results of moving target classification**.**

Figure17: Input video frame.

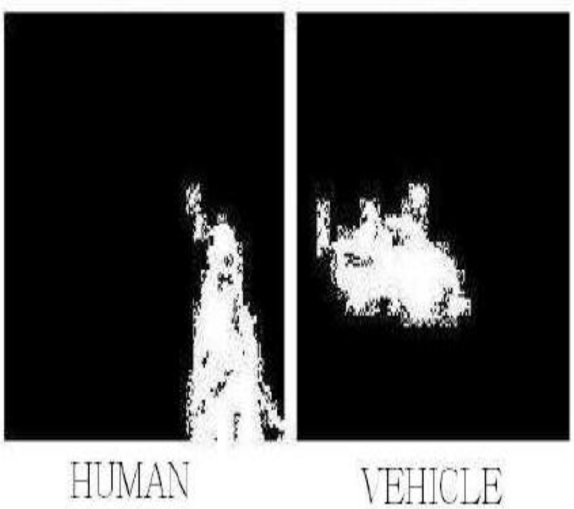

Figure18: Moving target classification results.

test sequence	Recognition						Accuracy	error rate
	DP	SB	MH	KK	SY	CS		
DP	6	0	0	1	0	0	85.7	14.3
SB	0	5	2	0	0	0	71.4	28.6
MH	0	1	5	0	0	1	71.4	28.6
KK	0	0	0	7	0		100	0
SY	0	0	0	0	7	0	100	0
CS	0	1	1	0	0	5	71.4	28.6

Table1: Width vector based classification results for database

test sequence	Recognition						Accuracy	error rate
	DP	SB	MH	KK	SY	CS		
DP	7	0	0	0	0	0	100	0
SB	0	7	0	0	0	0	100	0
MH	0	0	7	0	0	0	100	0
KK	0	0	0	7	0		100	0
SY	0	0	0	0	7	0	100	0
CS	0	0	0	0	0	7	100	0

Table2: Width vector and ART coefficient based classification results

Classification results in terms of average recognition for CASIA database

Feature type	attempts	success	Recognition rate
Width of the contour	72	59	81.9
width+ART coefficients	72	72	100

Classification results in terms of average recognition for database

Feature type	attempts	success	Recognition rate
Width of the contour	42	35	83.3
width+ART coefficients	42	42	100

VIII. Future Work:

This project is limited to acquiring gait characteristics from single camera-based set up. Using multi camera-based set up has more advantages. Performance rate have to be improved for occlusion, clothing style conditions and also for different walking considerations.

IX. Conclusion

Gait based recognition has been described in context of person authentication. Several existing techniques for gait recognition have been discussed. An intermediate result describes the effectiveness of proposed system. A result obtained in all intermediate steps has been discussed. Gaussian mixture model in addition with median filtering has been investigated for background subtraction. Moving target classification algorithm has done to separate human beings from other moving objects. Two types of gait features, width of outer contour of the binary silhouette and ART coefficients to describe gait shape are investigated. Using these shape descriptors (ART coefficients), disconnected objects can also be represented. So, using these shape descriptors we can get features in better way. PCA is used to remove correlation between the features and also for reducing the dimensionality.

Hidden Markov model has been developed for classification/recognition of individual. Even though we are getting promising results with the proposed approach, it has to be improved for large data bases.

REFERENCES

[1] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification,"Proceedings of Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002

.

[2] Nikolaos V. Boulgouris, Dimitrios Hatzinakos, and Konstantinos N. Plataniotis,"Gait recognition: A challenging signal processing technology for biometric identification",IEEE sinal processing magazine, November, 2005.

[3] M. Pushpa Rani and G. Arumugam, "An efficient gait recognition system forhuman identification using modified ICA," International Journal of ComputerScience and Information Technology (IJCSIT) , vol. 2, no. 1, February, 2010.

[4] Massimo Piccardi, "Background subtraction techniques: A review," IEEE Inter-national Conference on Systems, Man and Cybernetics, 2004.[5] P. Wayne Power and Johann A. Schoonees, "Understanding background mixturemodels for foreground segmentation," Proceedings of IEEE International Confer-ence on Image and Vision Computing, 2002.

[6] Chris Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time video tracking," Proceedings of IEEE Computer Society Conference onComputer Vision and Pattern Recognition, vol. 82, pp. 246-252, 1999.

[7] R. Cutler and L. Davis, "View based detections," Proceedings of Fourteenth IEEEInternational Conference on Pattern Recognition, vol. 1, pp. 495-500, (Brisbane,Australia), August, 1998.

[8] R. Cucchiara, M. Piccardi, and A. Prati, "Dtecting moving objects, ghosts and shadows in video streams," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1337-1342, October, 1993.

[9] Alan J. Lipton, Hironobu Fujiyoshi, and Raju S. Patil, "Moving target classificationand tracking from real-time video," Proceedings of Fourth IEEE Workshopon Applications of Computer Vision (WACV'98.

[10] Wang, L., Tan, T., Ning, H., Hu, W., " Silhouette analysis-based gait recognitionfor human identification,'IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1505-1518, 2003.45REFERENCES 46

[11] Lu, J. W., Zhang, E. H, Zhang, Z. G., Xue, Y.X., "Gait recognition using independentcomponent analysis," Proceedings of Second Symposium on Neural Net-works, (Chongqing, china), 2005.

[12] Ping. S. Huan, " Automatic gait recognition via statistical approaches for extendedtemplate features," IEEE Transactions on Systems, Man, and Cybernetics,vol. 31, no. 5, October, 2001.

[13] Qiong Cheng, Bo Fu, and Hui Chen, "Gait recognition based on PCA and LDA",Proceedings of Second Symposium Internationaln Computer Science and Comutational Technology (ISCSCT'09), pp. 124-127, December, 2009.

[14] L. Rabinerr, "A tutorial on hidden markov models and selected applications onspeech recognition," Proceedings of IEEE, vol. 77, no. 2, pp. 257-285, February,1989.

[15] Amit Kale, Aravind Sundaresan, A. N. Rajagopalan, Naresh P. Cuntoor, AmitK. Roy-Chowdhury, Volker Krger, and Rama Chellappa, "Identification of humansusing Gait" IEEE Transactions on Image Processing, vol. 13, no. 9, September,2004.

[16] Changhong Chen, Jimin Liang, Heng Zhao, and Haihong HU, "Gait recognitionusing hidden Markov model" Lecture Notes on Computer Science, pp. 399-407,2006.

[17] Aravind Sundaresan, Amit K. Roy-Chowdhury, and Rama Chellappa, "A hiddenMarkov model based framework for recognition of humans from gait sequences"Proceeding of ICIP, 2003.

[18]www.doc.ic.ac.uk/teaching/projects/Distinguished02/MarkDawson.pdf.