ISSN 2229-5518

Implementation of Real Time Video Surveillance System using Gait Analysis

Sonali Vaidya, Dr. Kamal Shah.

Abstract— Due to increased importance of safety and security, the use of Real Time video surveillance System has been increased. Real Time video surveillance system should be capable of detecting objects of interest, classify and track them. This can be done by using Gait Analysis for tracking scenarios and generating notification to an authoritative person. Gait Analysis helps us to identify people by the way they walk.For that purpose we used a new spatio-temporal gait representation, called Gait Energy Image (GEI), which is proposed for individual recognition by gait. GEI helps to represent a human motion sequence in a single image while preserving temporal information. Principal Component Analysis (PCA) and Multiple Discriminant Analysis (MDA) are used for learning features from the expanded GEI training templates. Recognition is then carried out based on the learned features.

Index Terms— Feature Extraction, Gait Analysis, Gait Energy Image (GEI), Principal Component Analysis (PCA) ,Multiple Discriminant Analysis (MDA), Template matching, Video surveillance system

1 Introduction

The advanced video surveillance systems are more into use due to increasing necessity to provide security. These systems help to analyse the behavior of people to prevent the occurance of potential danger.

In order to provide security, the main aim of the proposed system helps to identify an individual efficiently and accurately. Hence we can then use this system in controlled environmentlike scool, colleges, airport where tsystems need to quickly identify threats.

Gait as a behavioral biometric has many advantages over other biometrics. One of the important advantage of gait is unobtrusive identification, where even from a distance we can identify threat. This facility gives the user enough time to identify the suspect before he could become a possible threat. Video footage of suspects are readily available with user, as surveillance cameras are comparatively low cost and installed in prime locations requiring security ,in these cases the video needs to be checked against that particular suspect [18].

1.1 Gait for visual surveillance

The problem of "personal identification under the area of visual surveillance" is of increasing importance. Such type of personal identification can be treated as a special behavior-understanding problem. Human gait are now regarded as the main biometric features that can be used for personal identification in visual surveillance systems.In this case Gait Analysis is going to be used; Gait refers to the style of walking of an individual. It includes both the appearance and the dynamics of human walking motion. Human gait is an identifying feature of a person that is determinedby his/her weight, limb length, and habitual posture. And Gait recognition is the term typically used for the automatic extraction of visual clue that characterize the motion of a walking person in video and is used for identification purposes in surveillance systems, which can then generate an alarm regarding the object of interest.

Gait is a behavioral biometric source that can be acquired at a distance. Gait recognition is typically used in the computer community to refer to the automatic extraction of visual cues that characterize the motion of a walking person in a video and is used for identification purposes in surveillance systems. Automated surveillance system consists of three phases: detection, tracking and perception. In the perception phase, a high-level description is produced based on the features extracted during the previous phases from the temporal video stream. The main aim of automated surveillance system is to detect and track people in the scene as well as to perceive their behavior and report any suspicious activities tothe authorized person.

Identification systems will undoubtedly play a key role in aiding law enforcement officers in their forensic investigations. More importantly, due to early recognition of suspicious individuals who may pose as threats, the system would be able to reduce future crimes. Human motion perception has been of interest to researchers from different disciplines due to the wide range of applications ranging from activity recognition to people identification. In fact, early studies by Johansson [1] on human motion perception using Moving Light Displays (MLD) have revealed that an observer can recognize different types of human motion based on joint motions. Moreover, the observer can make a judgment of the gender of the person [2], and even further identify the person if they are already familiar with their gait [3]. This leads to the conclusion that gait might be a potential biometric for surveillance systems.A biometric is a descriptive measure based on the human behavioral or physiological characteristicswhich distinguishes a person uniquely among other people; thisunique description should be universal and permanent. Currently, as most biometric systems are still in their infancy, the use of biometrics is limited to identity verificationand authentication [4]. Gait is an emergent

- *Sonali Vaidya is currently pursuing masters degree program in Information Technology from TCET-Mumbai, Mumbai University, India. E-mail: sonalivaidya.we@gmail.com*
- *Dr. Kamal Shah is currently professor in Information Techology in TCET-Mumbai, Mumbai University, India. E-mail: kamal.shah@thakureducation.org*

ISSN 2229-5518

biometric which is increasingly attracting the interests of researchers as well as the industry. Gait is nothing but the way of walking. Early studies by Murray revealed that gait might be a useful biometric for people identification, a total of 20 feature componentsincluding spatial displacement, ankle rotation and vertical tipping of the trunk have been identified to render uniquely the gait signature for every individual, while some of these features are difficult to extract using current computer vision systems, others arenot consistent over time for the same person [5]. In one of the early experiments on gaitrecognition conducted by Cutting et al in 1978, it was demonstrated that people can recognize others just by gait cues [2].

Although gait recognition is still a new biometric and is not sufficiently established to be deployed in real world applications such as surveillance system,but it has the potential to overcome most of the limitations that other biometrics suffer from such as face, fingerprints and iris recognition. Face recognition in many cases has been proven to be unreliable for visual surveillance systems; this is due to the fact that people can disguise or hide their faces as well as that video data being captured can be inadequate at low resolution.Furthermore, another major drawback of face identification in security applications is its low recognition rates in poor illumination. Because most of the facial features cannotbe recovered at large distances even using night vision capability [6]. Although fingerprint and iris recognition have proved to be robust for applications where authentication is required, such biometrics are inapplicable for situations where the subject's consent and cooperation are impossible to obtain.

1.2 Gait for human identification

Gait recognition helps to identify an individual from a video sequence of the subject walking. Gait as a biometric is advantageous over other forms of biometric identification techniques for the following reasons [18]:

1. **Unobtrusive** – Gait of a walking person can be extracted without the user knowing they are being analyzed and without any cooperation with the user.
2. **Distance recognition** – Gait of an individual can be captured at a distance
3. **Reduced detail** – Gait recognition does not require images captured in very high resolution, unlike other biometric techniques such as iris recognition, which can be easily affected by low resolution images.
4. **Difficult to conceal** – Gait of an individual is difficult to cover up, if they try to do so the individual will probably appear more suspicious.

An individual’s gait signature will be affected by certain factors such as:

1. **Stimulants** – Consumption of drug or alcohol will affect the person’s walking style.
2. **Physical Changes** – During pregnancy, after an accident/disease affecting the leg, or after severe weight gain /loss can also affect the movement characteristic of an individual.
3. **Psychological Changes**– A person's mood can also affect an individual's way of walking.
4. **Clothing**– Same person with different clothing may cause change in gait signature.[7]

So by taking into consideration above all points, gait is still a good behavioral biometric for human identification.

2 System Flow

The objective of this proposed system is to develop a system which is capable of performing human identification from a video sequence from his/her walking pattern and system should be able to store the derived gait signature and retrieve it as per requirement.

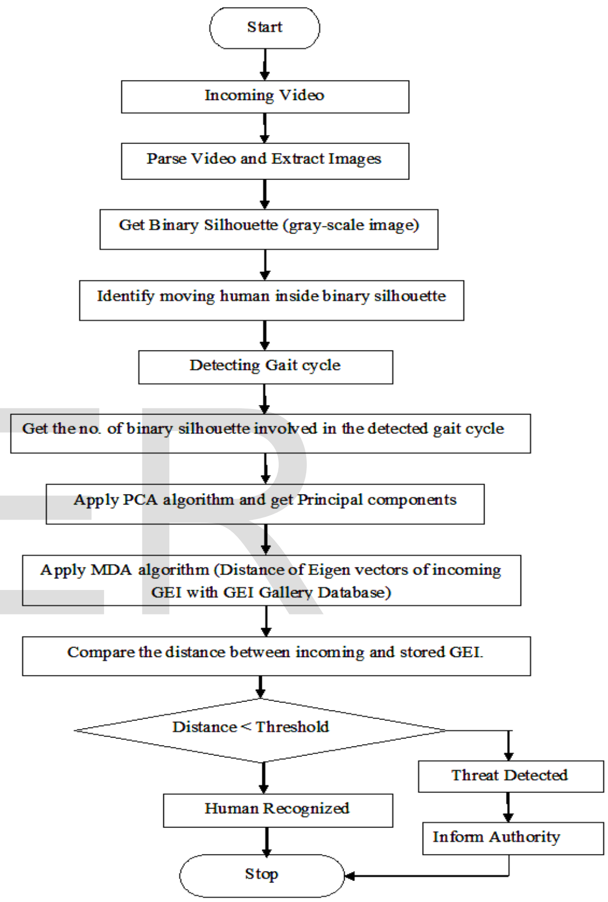


Fig.1: System Flowchart

3 System Phases

For implementation of this system we have divided the system into 3 phases:-

- Human Detection and Tracking
- Feature Extraction
- Training and Recognition

4 Phase 1:- Human Detection and Tracking

In order to prevent the occurrence of the potential threat, the system needs to analyze the behaviors of people.

Aim of Human Detection and Tracking in this system is -

- To extract a good quality Human(s) silhouette image and

- To track this silhouette(s), from a video frame.

This is needed in order to perform the extraction of the gait feature from the walking sequence [18].

4.2 Background Subtraction

The background subtraction method is nothing but extracting foreground image through the threshold of difference between the current image and reference image. In this system, for background subtraction we have used GaussianModel, as this method can handle most of tough sistuations like sudden light change, heavy shadow etc.

Using extended expectation maximization (EM) algorithm, Friedman *et al.* [8] implement a mixed Gaussian classification model for each pixel. This model classifies the pixel values into three separate predetermined distributions corresponding to background, foreground and shadow. It also updates the mixed component automatically for each class according to the likelihood of membership. Hence, slowly moving objects are handled perfectly, while shadows are eliminated much more effectively [18].

4.3 Connected Components Labeling

The purpose of connected component labeling is to group together pixels which have similar properties and are connected in some way. The image is scanned from top to bottom and left to right; pixels which should be grouped together are given the same label.

In this system for finding connected componenets we have used Two Pass algorithm. As name suggests this algorithm consists of 2 passes over a given binary image.In first pass it records equivalenceand then assign temporary labels. In case of second pass, it replaces each temporary label by the label of its equivalence class [9].

Here, the background classification is specific to the data, used to distinguish salient elements from the foreground. If the background variable is omitted, then the two-pass algorithm will treat the background as another region [10].

4.4 Object tracking

In order to identify multiple human at the same time (move at same time in the video) we need to track the individuals.The good human model should be invariant to rotation, translation and changes in scale, and should be able to handle partial occlusion, deformation and light change.In this system we have used Appearance based tracking method.

This method uses the color histogram, velocity, the number of pixels and size as the human model to describe the humans. For tracking, we assume the human always moves in similar direction and similar velocity. During the process of tracking, we will check whether the people stop or change the direction. If the person doesn’t move for period of time, we will check whether this person is false. Once the false person is found, system will learn this false alarm and adjust the background accordingly [11].

4.5 Object classification

After tracking objects and analysing their behavior, it is essential to correctly classify moving objects. Object classification can be considered as a standard pattern recognition issue. In order to track it reliability, it is very important to recognize the type of a detected object. Currently, there are two main categories of approaches for classifying moving objects, they are motion-based and shape-based classification [18].

In this system, For human recognition we have used Shape-Based Approach to implement object classification using Jianpeng Zhou and Jack Hoang Algorithm's based on codebook theory which classify the human from other objects.Here we used classification algorithm based on codebook theory, which works as follows:-

- First step, we normalize the size of object, and then extract the shape of object as the features.
- Second step, we match the feature vector with the code vectors of codebook.
- Thematchprocess is to find a code vector in codebook with the minimum distortion to the feature vector of object. If the minimum distortion is less than a threshold, this object is human else it is not human.

The design of the codebook is critical for the classification. The partial distortion theorem for design codebook is that each partition region makes an equal contribution to the distortion for an optimal quantizer with sufficiently large N [12]. Based on this theorem, we used Distortion Sensitive Competitive Learning (DSCL) algorithm to design the codebook, which is explained in [11].

5 Phase 2:- Feature Extraction

After detecting the object the next step is to extract some useful features.Following are the 3 types by using which feature extraction can be performed.

- Model-Based Feature Extraction
- Model-Free Feature Extraction
- Gait Energy Image (GEI)

Out of which to characterize human walking properties for individual recognition by gait, we used a new spatio-temporal gait representation, called Gait Energy Image (GEI).

GEI is constructed using silhouettes. GEI represents a single image which contains information about both body shape and human walking dynamics due to this compactness they are useful to use and maintain. GEI is less sensitive to noise and able to achieve highly competitive results compared to alternative representations [13].

5.1 Gait Energy Image (GEI)

5.1.1 Gait Cycle Detection

A gait cycle is defined as the time interval between successive instances of initial foot to-floor contact for the same foot, and the way a human walks is marked by the movement of each leg. Gait Periodicity can be estimated by a simple strategy. We need to count the number of foreground pixels in the silhouette in each frame over time. If this count of foreground pixel will reach the max when the two legs are farthest apart (i.e. full stride stance), and drop to a minimum when the legs overlap (i.e. heels together stance) [14]. But it is difficult to get the minimum or maximum number as the frames intensity change frequently. So we calculate the average intensity of k consecutive frames [18].

ISSN 2229-5518

5.1.2 Size Normalization and Horizontal Alignment

Before extracting features, we should normalize all silhouette images to be the fixed size, and then centroid of an image is calculated [18].

5.1.3 Representation Construction

For this step we can use a silhouette extraction procedure and begin with the extracted binary silhouette sequences. After perfroming preprocessing procedure which includes size normalization – fitting the silhouette height to the fixed image height, and sequential horizontal alignment centering the upper half silhouette part with respect to the horizontal centroid.
After this gait cycles are segmented by estimating gait frequency using a maximum entropy estimation technique presented in [14], [15].
A size-normalized and horizontal-aligned human walking binary silhouette sequence $I(x, y, t)$, the grey-level GEI $G(x, y)$ is then computed as follows,

$$G(x, y) = \frac{1}{N} \sum_{t=1}^{N} I(x, y, t) \quad (1)$$

Where N is the number of frames in a complete gait cycle, x and y are values in the 2D image coordinate, and t is the frame number in the gait cycle [16].

6 Phase 3:- Training and Recognition

6.1 Training and Classification

Training - The process of storing the extracted features (i.e. probe GEI) and the information needed about the trained humans (i.e. label, name, address etc.) in the gallery database to be used later for the recognition of walking humans. Training should be performed in a special environment with special conditions to get the best motion patterns [18].
Classification- The process in which individual items are placed into groups based on quantitative information on one or more characteristics inherent in the items (referred to as traits, variables, characters, etc) and based on a training set of previously labeled items. In this phase all GEIs stored in the Gallery will be retrieved and grouped into classes. Then the new features (i.e. probe GEI) will be assigned to one of the classes that has the minimal distance. Gait recognition can be performed by matching a probe GEI to the gallery GEI that has the minimal distance between them [18].

6.2 Human Recognition using GEI Templates

Human walking sequences for training are limited in real surveillance applications. Because each sequence is represented as one GEI template, the training/gallery GEIs foreach individual might limited to several or even one template(s).
There are two approaches to recognize individuals from the limited templates: - Direct GEI Matching and Statistical GEI matching. In case of direct GEI matching approach the features extracted from silhouettes are usually high-dimensional. Working with huge vectors and comparing them and because of which they are sensitive to noise and small silhouette distortions [15]. Even working with huge vectors and comparing them and storing them is a computationally expensive, time consuming and needs a lot of storage space. Due to which dimensionality reduction method which is also called as statistical GEI feature matching is used to find most dominant features and remove redundant or less important once.

6.3 Statistical GEI feature matching

A statistical GEI feature matching approach is used for individual recognition from limited GEI templates. To reduce their dimensionality, there are two classical linear approaches for finding transformations for dimensionality reduction—Principal Component Analysis (PCA) and its variants Multiple Discriminant Analysis (MDA).
First, we generate new templates from the limited training templates according to a distortion analysis. Next, statistical features are learned from the expanded training templates by principal component analysis (PCA) to reduce the dimension of the template and multiple discriminant analysis (MDA) to achieve better class seperatability.As Huang et al. [17] combine PCA and MDA which seeks to project the original features to a subspace of lower dimensionality so that the best data representation and class seperatability can be achieved simultaneously.PCA seeks a projection that best represents the data in a least-square sense, while MDA seeks a projection that best separates the data in a least-square sense.The individual is recognized by the learned features [16].
Finally for individual recognition, we need to calculate the distance between the feature vectors of each gallery GEI and the probe GEI.
If the distance is less than Threshold value then the human is recognized and his information is retrieved and displayed, else the human is not recognized, considered as a stranger and the Authority should be alerted to take an action [15].

7 Result and Discussion

We have performed this experiment on 2 types of database

- Standard Database
- Regular (Non-standard) Database

In both the types of databases we have training and testing sets.

7.1 Standard Database

For this we have used CASIA (The Institute of Automation, Chinese Academy of Science) who provides Gait Database. In this database there are 3 datasets, Dataset A, Dataset B (multiview Dataset), Dataset C (Infrared Dataset).
Out of which we have used Dataset B and Dataset C.

7.1.1 CASIA Gait Database (Dataset B)

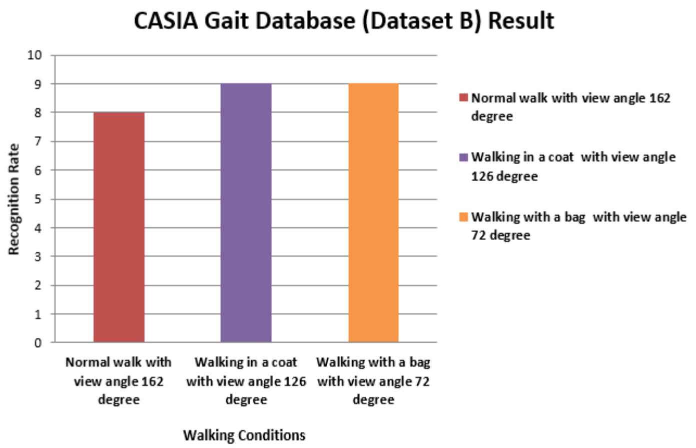
In this experiment, we use the CASIA Gait Database (Dataset B). The database consists of 124 persons and the gait data was captured from 11 views. Every person has 3 walking conditions- a) normal walk, b) walking in coat, and c) walking with bag. In this dataset view angles are ranging from $0^\circ$to 180°. The original image size of the database is 320 x 240.We took 10 persons and analyse them on 6 conditions (3 conditions for training and 3 conditions for testing).
For training purpose we took all 3 conditions with view angle 36^0,90^0, 144^0.

ISSN 2229-5518

And For testing a) Normal walk – View angle 162^0
b) Walking in coat- View angle 126^0
c) Walking with bag- View angle 72^0

Fig.2: Recognition Result of various walking conditions on Dataset B

Above graph (Fig.2) depicts that, for normal walk condition,



out of 10 persons system recognized 8 persons, on the other hand for walking in coat and walking with bag out of 10 persons system recognized 9 persons. From the recognition result we can say that on CASIA dataset B we have achieved 86.66% of efficiency.

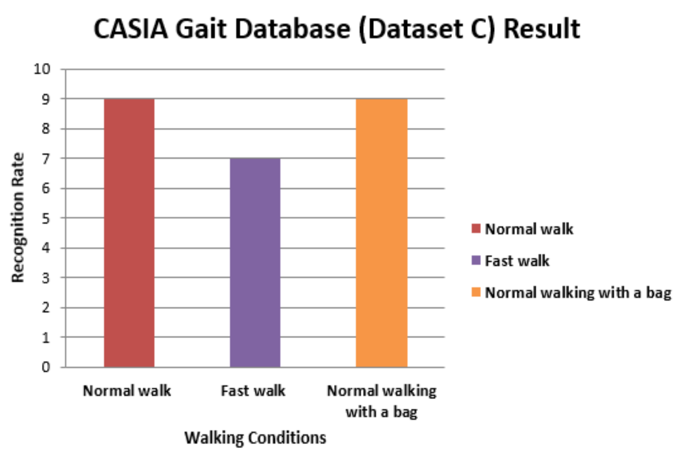
7.1.2 CASIA Gait Database (Dataset C)

In this case, we use the CASIA Gait Database (Dataset C), where images are collected by an infrared (thermal) camera. This database consists of 153 personsand takes into account 3 walking conditions: a) normal walking, b) fast walking, and c) normal walking with a bag. These all videos are captured at night. We took 10 persons and analyse them on 6 conditions (3 conditions for training and 3 conditions for testing).

For training we took sequence 2 and for testing we took sequence 1.

Fig.3: Recognition Result of various walking conditions on Dataset C

Above graph (Fig.3) depicts that, for normal walk condition,

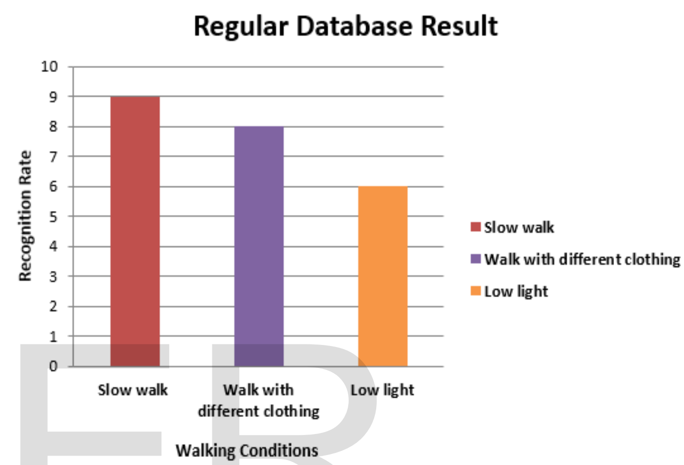


out of 10 persons system recognized 9 persons, for fast walk condition, out of 10 persons system recognized 7 persons and for normal walking with bag out of 10 persons system recognized 9 persons.From the recognition result we can say that on CASIA dataset C we have achieved 83.33% of efficiency

7.2 Regular (Non- standard) Database

In case of regular database, for analyzing system performance we took 10 persons with 6 different walking conditions they are - a) Normal walk, b) Fast walk, c) Walking with bag, d) Slow walk, e) With different clothing, f) low light, Out of these conditions first 3 (a, b, c) conditions are for training purpose and remaining 3 (d, e, f) are for testing purpose. For all condition view angle is 90^0.

Fig.4: Recognition Result of various walking conditions on Regular Database



Above graph (Fig.4) depicts that, for slow walk condition, out of 10 persons system recognized 9 persons, for walk with different clothing condition, out of 10 persons system recognized 8 persons and for low light condition out of 10 persons system recognized 6 persons. From the recognition result we can say that on Regular database we have achieved 76.66% of efficiency.

7.3 Overall Result

Table1. Summary of Overall Result

Type of Database	Regular Database	Standard Database	
		Dataset B	Dataset C
Efficiency	76.66 %	86.66%	83.33%

However the efficiency achieved in both (Regular and standard databases) cannot be generalized as it is performed on less number of test cases and conditions under which they are tested may be changed on other time.

8 Conclusion

In this paper, for individual recognition in real time video surveillance system, a new spatio-temporal gait representation, called Gait Energy Image (GEI), is used. GEI represents a human motion sequence in a single image while preserving temporal information. There are two approaches to recognize individuals from the limited templates: - Direct GEI Matching and Statistical GEI feature matching. Out of which we have ued Statistical GEI feature matching, wherein to reduce dimensionality problem of GEI's, for finding transformations for dimensionality reduction we used two conventional approaches they are Principal Component Analysis (PCA) and its variants Multiple Discriminant Analysis (MDA). For Individual Recognition we have calculated the distance between the feature vectors of each gallery GEI and the probe GEI. If the distance is less than Threshold value then the human is recognized, else the human is not recognized and inform in a form of alarm is given to authoritative person. Experimental results show that (a) GEI is an effective and efficient gait representation (b) the proposed recognition approach achieves high performance as compared to existing gait recognition approaches.

REFERENCES

[1] G. Johansson, "Visual Perception of Biological Motion and a Model for its Analysis", *Perception and Psychophysics*, vol.14, no. 2, pp. 201-211, 1973.

[2] L. T. Kozlowski and J. E. Cutting, "Recognizing the Gender of Walkers from Point-Lights Mounted on Ankles: Some Second Thoughts", *Perception & Psychophysics*, vol. 23, no. 5, pp.459, 1978.

[3] N. H. Goddard, "The Perception of Articulated Motion: Recognizing Moving Light Displays", PhD dissertation, University of Rochester, 1992.

[4] A. K. Jain, R. Bolle, and S. Pankanti, editors, "Biometrics: Personal Identification in Networked Society", Kluwer academic publishers, 1999.

[5] M. P. Murray, "Gait as a Total Pattern of Movement", American Journal of Physical Medicine, vol. 46, no.1, pp.290-333, 1967.

[6] A. Kale, A. N. Rajagopalan, N. Cuntoor, and V. Kruger,"Gait-Based Recognition of Humans using Continuous HMMs", *Proc. of Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 321-326,2002.

[7] Supreet Kaur, Er. ShikhaChawla, "An Approach for Enhancing the Human Identification Using Gait Recognition techniques", International Journal of Application or Innovation in Engineering & Management (IJAIEM), vol. 2, no.9, pp.146-157, September 2013.

[8] N. Friedman and S. Russell, "Image segmentation in video sequences: a probabilistic approach," *Proc. Of 13th Conference on Uncertainty in Artificial Intelligence (UAI)*, pp. 175–181, 1997.

[9] Sunny Kumar,Pratibha Sharma, "Offline Handwritten & Typewritten Character Recognition using TemplateMatching", International Journal of Computer Science & Engineering Technology (IJCSET),vol. 4, no. 6, pp.818-825, June 2013.

[10] Lifeng He; Yuyan Chao, K. Suzuki, "A Run-Based Two-Scan Labeling Algorithm", *IEEE Trans. on Image Processing*, Vol.17, no. 5,pp.749–756, May 2008.

[11] Jianpeng Zhou and Jack Hoang, "Real Time Robust Human Detection and Tracking System", *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops – 3*, pp. 149-156, June 2005.

[12] A. Gersho, "Asymptotically optimal block quantization", *IEEE Trans. on Information Theory*, vol.25, pp. 373–380, July1979.

[13] Khalid Bashir, Tao Xiang, Shaogang Gong,"Feature Selection On Gait Energy Image For Human Identification", *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),* pp.985-988,2008.

[14] S. Sarkar, P. Phillips, Z. Liu, I. Vega, P. Grother, and K. Bowyer., "The humanID gait challenge problem: Data sets, performance, and analysis", *IEEE Trans. on PAMI*, vol. 27, no. 2, pp. 162–177, 2005.

[15] Khalid Bashir, Tao Xiang and Shaogang Gong,"Feature Selection for Gait Recognition without Subject Cooperation", *Pattern Recognition Letters*, vol. 31, no. 13, pp.2052-2060, 2010.

[16] J. Han and B. Bhanu," Individual recognition using gait energy image", *IEEE Trans. on PAMI*, vol. 28, no. 2, pp.316–322, Feb 2006.

[17] P.S. Huang, C.J. Harris, and M.S. Nixon, "Recognizing Humans by Gait via Parametric Canonical Space," *Artificial Intelligence in Engineering*, vol. 13, pp. 359-366, 1999.

[18] Sonali Vaidya, Kamal Shah,"Real Time Surveillance Systeem", International Journal of Computer Application (IJCA), vol. 86, no. 14, pp. 22-27, January 2014.