IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 16, Issue 1, Ver. 5 (Jan. 2014), PP 88-93
www.iosrjournals.org

Use of Gait Energy Image in Implementation of Real Time Video Surveillance System

Sonali Vaidya[1], Dr. Kamal Shah[2]
[1]*(ME-IT student, Department of Information Technology, TCET- Mumbai, India)*
[2]*(Professor, Department of Information Technology, TCET- Mumbai, India)*

Abstract: *The significance of Real Time video surveillance System has been raised up due to increased importance in safety and security. The Proposed Real Time video surveillance system is capable of detecting, classifying and tracking objects of interest (human). This can be done by using Gait Analysis for tracking scenarios and generating notification to authoritative person. Gait Analysis helps us to identify people by their walking style. Because of this exquisite feature we used gait recognition in proposed real time video surveillance system. In this proposed system for depicting human walking properties for individual recognition i.e. for performing feature extraction we have used a new spatio-temporal gait representation called as Gait Energy Image (GEI). GEI represents a human motion sequence in a single image while preserving temporal information. For human recognition we have used Statistical GEI feature matching, wherein to reduce dimensionality problem of GEI's, we used two approaches they are Principal Component Analysis (PCA) and its variants Multiple Discriminant Analysis (MDA).*

Keywords: *Feature extraction, Gait analysis, GEI (Gait Energy Image), GEI Template Matching, MDA (Multiple Discriminant Analysis), PCA (Principle Component Analysis)*

I. INTRODUCTION

A video surveillance system is the system proposed to identify an individual efficiently and accurately. Gait as a biometric has many advantages which makes it an attractive proposition as a method of identification. A main advantage of Gait is its unobtrusive and distance recognition. So even from far distance without any cooperation from the user we can identify them. Video footage of suspects are readily available with user, as surveillance cameras are comparatively low cost and installed in prime locations like school, colleges, airport etc, which require security.

The term gait recognition signifies the identification of an individual from a video sequence of the subject walking. This does not mean that gait is restricted to walking; it can also be applied to running or various types of movement on foot.

The objective of the proposed system is to develop a system which is capable of performing human identification from a video sequence from his/her walking pattern also system should be able to store and retrieve gait signature as per requirement[21] [22].

II. REAL TIME VIDEO SURVEILLANCE SYSTEM

A. System Scenario

In the below scenario (Fig. 1), we need to analyze the video stream from surveillance cameras. Surveillance camera records a video which then send to video surveillance system. If some individual walks by the camera whose gait signature has been previously stored and they are a known threat, then the system will identify them and the appropriate authorities will be inform and the person can be dealt with before they are allowed to become a threat. If the threat has been detected successfully from a distance, then it will give a time span for authorities to take an action.

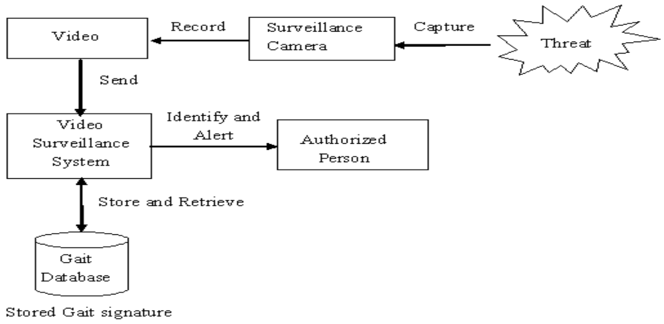

Fig. 1 . System Scenario

B. System Phase

System consists of 3 phases

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

III. PHASE 1: HUMAN DETECTION AND TRACKING

The advanced video surveillance system needs to analyze moving object i.e. human and track him in video frames. So main aim of Human Detection and Tracking in this system is, to extract a good quality Human silhouette image also track this silhouette(s) from a video frame.

A. Background Subtraction

The background subtraction method is nothing but foreground segmentation which is nothing but difference between the current image and reference image. In this system, for background subtraction we have used Gaussian Model, as this method can handle most of tough situations like sudden light change, heavy shadow etc.

Using extended expectation maximization (EM) algorithm, Friedman *et al.* [24] 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 [21] [22].

B. Connected Components Labeling

The idea of connected component labeling is grouping of pixels which possess similar properties and then connect them 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 components 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 equivalence and then assign temporary labels. In case of second pass, it replaces each temporary label by the label of its equivalence class [11]. Here, the background classification is specific to the data, used to distinguish most important elements from the foreground. The two-pass algorithm will treat the background as another region when background variable is absent [10].

C. Object Tracking

In order to track an individual we need to create the human model for each individual. For that purpose we are using Appearance based tracking method. In this method we use 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 [7].

D. Object Classification

In video surveillance system, main target of interest is human. Moving regions detected in video may correspond to different objects in real-world such as pedestrians, vehicles, etc. In order to track it reliability, it is very important to recognize the type of a detected object. Currently, there are two approaches for finding moving object classification which are motion-based and shape-based methods. Shape-based methods make use of the object's 2D spatial information whereas motion-based methods use temporally tracked features of objects for the classification solution. In this system, for human recognition we used Shape-Based Approach to implement object classification using Jianpeng Zhou and Jack Hoang Algorithm's based on codebook theory which classifies the human from other objects. 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 [18]. Based on this theorem, we used distortion sensitive competitive learning (DSCL) algorithm to design the codebook, which is explained in [7].

IV. PHASE 2: FEATURE EXTRACTION

Feature extraction can be performed in a 2 ways:-

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

For the Model-Based Feature Extraction, we use a new method to extract human joints with better accuracy then blobs via incorporating prior knowledge to refine accuracy.
This technique has capability to extract moving joints of human body with high accuracy for indoor data as well as outdoor data filmed in an unconstrained environment. Model based approaches are difficult to follow in low resolution images. Also implementing a model based approach includes high computational costs, due to the complex matching and searching that has to be performed are high which makes it highly complex, also processing takes more time. So, there more appropriate to be used in non real time application. Therefore instead of using Model-Based Feature Extraction, for feature selection we used Gait Energy Image (GEI).
Gait Energy Image (GEI) is selected for gait representation, which is a spatio-temporal gait representation, constructed using silhouettes. GEI represents gait using a single image which contains information about both body shape and human walking dynamics. GEI is thus a compact representation which makes it an ideal starting point for feature selection since it is computational expensive if the number of features to select is high. In spite of its compactness, it has been demonstrated that GEI is less sensitive to noise and able to achieve highly competitive results compared to alternative representations [15].

V. GAIT ENERGY IMAGE (GEI) FEATURE EXTRACTION

Following are the ways by using which we can perform feature extraction:-

A. 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 [22]. Gait Periodicity can be estimated by counting the number of foreground pixels in the silhouette in each frame over time. The number will reach the maximum 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) [17]. 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 [21].

B. Size Normalization and Horizontal Alignment

Before extracting features, we should normalize all silhouette images to be the fixed size, then centroids of an image is calculated [21].

C. Gait Representation

Given a human walking sequence, a human silhouette is extracted from each frame using the method in [14] [17]. After applying size normalization and horizontal alignment to each extracted silhouette image, gait cycles are segmented by estimating gait frequency using a maximum entropy estimation technique presented in [14][17].
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 [15] [20].

VI. PHASE 3: TRAINING AND RECOGNITION

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 [21].
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 have the minimal distance. Gait recognition can be performed by matching a probe GEI to the gallery GEI that has the minimal distance between them [21].

A. Human Recognition using GEI Template

Human walking sequences for training are limited in real surveillance applications. Because each sequence is represented as one GEI template, the training/gallery GEIs for each individual might limited to several or even one template [22].

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 [14]. 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.

B. 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. [19] 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 [20]. 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 [14].

VII. EXPERIMENTAL SETUP AND DISCUSSION

There are 2 types of database Standard and Regular (Non- Standard) Database. Here for experiment we used Regular (Non- Standard) Database. For both the types of database, we need to have 2 sets training as well as testing.

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) Walk with different clothing, d) Walk with bag, e) Slow Walk, f) low light.

Here we took 2 scenarios with different training and testing conditions.

A. Scenario One

Out of above mentioned conditions for training set we took first 3 conditions (Walk with bag, slow walk and low light) Whereas as for a testing set we took next 3 conditions (Normal walk, Fast walk and walking with different clothing)

For all 6 condition view angle is 90^0.

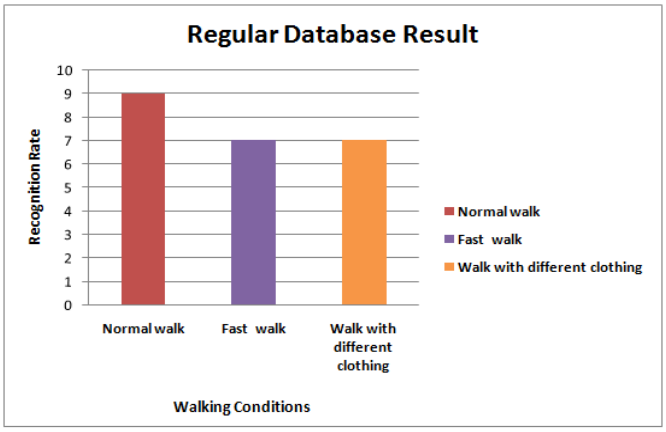

Fig. 2 . Result of Scenario one for various walking conditions on Regular Database

Above graph (Fig.2) depicts that, for normal walk condition, out of 10 persons system recognized 9 persons, on the other hand for fast walk condition out of 10 persons system recognized 7 persons and for walk with different clothing condition out of 10 persons system recognized 7 persons. This in turn achieves 76.66% of efficiency.

B. Scenario Two

Out of above mentioned conditions for training set we took first 3 conditions (Normal walk, Fast walk and walking with different clothing).Whereas as for a testing set we took next 3 conditions (Walk with bag, slow walk and low light).
For all 6 condition view angle is 90^0.

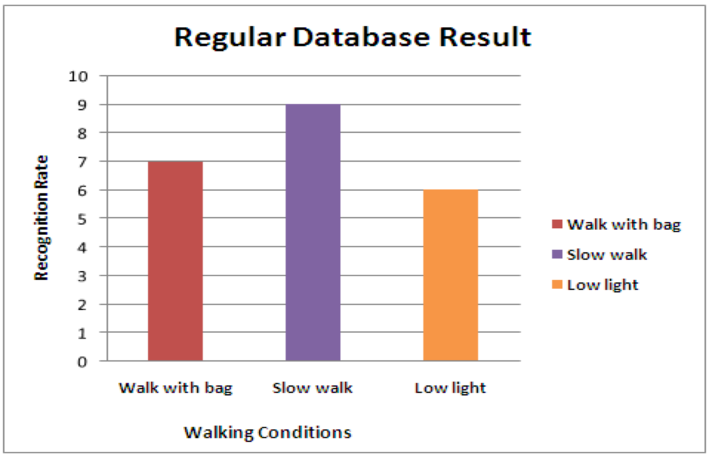

Fig. 3 . Result of Scenario two for various walking conditions on Regular Database

Above graph (Fig.3) depicts that, for walk with bag condition, out of 10 persons system recognized 7 persons, on the other hand for slow walk condition out of 10 persons system recognized 9 persons and for low light condition out of 10 persons system recognized 6 persons. This in turn achieves 73.33% of efficiency.
However the efficiency achieved in above two scenarios 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

VIII. CONCLUSION

In this paper, for depicting human walking properties for individual recognition i.e. for performing feature extraction we are using a new spatio-temporal gait representation called as Gait Energy Image (GEI). In case of GEI human motion sequence has been represented in a single image even though it is also preserving temporal information. For human recognition we are GEI template matching technique. There are two approaches - Direct GEI Matching and Statistical GEI feature matching. Out of which we have used 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 the proposed gait recognition approach achieves good 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] 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.

[8] S.J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H.Wechsler, Tracking Groups of People, *Computer Vision and Image Understanding, vol. 80*, pp.42-56, 2000.
[9] P.L.Rosin and T.Ellis, Image Difference Threshold Strategies and Shadow Detection, *Proc. of 6thBritish Machine Vision Conference*, pp.347-356, 1995.
[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] Sunny Kumar, Pratibha Sharma, Offline Handwritten & Typewritten Character Recognition using Template Matching, *International Journal of Computer Science & Engineering Technology (IJCSET),vol. 4, no. 6*, pp.818-825, June 2013.
[12] Imed Bouchrika and Mark S. Nixon, Model-Based Feature Extraction for Gait Analysis and Recognition, *Proc. of 3rd international conference on Computer vision/computer graphics collaboration technique*s, pp.150-160, 2007.
[13] Yigithan Dedeoglu, *Moving Object Detection, Tracking And Classification For Smart Video Surveillance*, Bilkent University, 2004.
[14] 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.
[15] 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.
[16] A.Aguado,M.S. Nixon, M .E. Montiel, Parameterising arbitrary shapes via Fourier descriptors for evidence-gathering extraction, *computer vision and image understanding, vol. 69, no. 2,* pp.202–221,1998.
[17] 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.
[18] A. Gersho, Asymptotically optimal block quantization, *IEEE Trans. on Information Theory*, vol.25, pp. 373–380, July1979.
[19] 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.
[20] J. Han and B. Bhanu, Individual recognition using gait energy image, *IEEE Trans. on PAMI*, vol. 28, no. 2, pp.316–322, Feb 2006.
[21] Sonali Vaidya, Dr. Kamal Shah, ,Real Time Surveillance System, *International Journal of Computer Application (IJCA), vol. 86, no. 14*, pp. 22-27, January 2014.
[22] Sonali Vaidya, Dr. Kamal Shah, Implementation of Real Time Surveillance System using Gait Analysis, *International Journal Of Scientific & Engineering Research(IJSER),* Volume 5, Issue 2, February-2014 .[Submitted for publication]
[23] Supreet Kaur, Er. Shikha Chawla**,** 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.
[24] 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.