CHAPTER 2
REVIEW OF LITERATURE

2.1 INTRODUCTION

This chapter presents literature review on gait analysis using different methods. The performance of different techniques used for gait analysis by the previous researchers are presented.

2.2 LITERATURE REVIEW

Gait recognition can be related to human movement analysis methods i.e. vision methods that detect, track and / or recognize classes of human motion and more specific methods that deal with whole-body movement (without regard to its individual parts). A comprehensive survey of gait recognition techniques can be found in [Nixon and Carter, 2006].

2.2.1 Human Movement Analysis

It has been reported by different researchers that one of the main benefit of gait recognition over other biometrics is its non-intrusive nature [Wang et al., 2002]. Hence forth, the analysis of the covariate factors becomes essential to the understanding of the uniqueness of gait recognition. Several areas of computer vision are interested in the analysis and modeling of human motion in video, including pedestrian detection / counting, gesture recognition, action / activity recognition, lip reading and person identification. Existing methods can be grouped into structural methods, which recover the structure of the human body prior to recognizing its movement [Hogg, 1983; Akita, 1984; Rohr, 1994; Gavrila and Davis, 1995; Bregler, 1997; Feng, 2000] and structure-free methods which directly model
the 2D motion patterns generated in the image sequence by any particular type of movement [Debrunner, 2000; Huang et al., 1998; Cuntoor et al., 2002].

2.2.2 Structural Methods

A 2D or 3D structural model of the human body is assumed and body pose is recovered by extracting image features and mapping them to the structural components of the model (i.e. body part labelling). Hence, a human is detected in the image, if there exists a labelling that fits the model well enough (based on some measure of goodness of fit). Once a person has been detected and tracked in several images, motion recognition is done based on the temporal trajectories of the body parts, typically by mapping them to some low-dimensional feature vector and then applying standard pattern classification techniques [Keiter, 1994]. Tracking body parts in 3D over long sequences remains an unsolved computer vision problem, the effectiveness and applicability of these methods remain limited. 

Yam et al., 2004, analyzed the performance of the bottom-up model-based approach which is satisfactory for extracting the most important dynamic gait features for people recognition. Wang et al., 2003, derived binary silhouettes of walking subjects which are converted into a one-dimension normalized distance signal by contour unwrapping with respect to the centroid position. Wang et al., 2004, extracted dynamic and static features of gait motion using a model-based method based on the condensation framework.

Most vision systems for human motion analysis rely primarily on the extraction of markers attached to the joints of moving people. However, most applications such as visual surveillance require the deployment of an automated marker-less vision system to recover the
joints’ trajectories. It is claimed that automated extraction of the joints’ positions from videos of walking subjects is an unsolved problem. This is because, non-rigid human motion encompasses a wide range of possible motion transformations due to the highly flexible structure of the human body and to self-occlusion.

Veres et al., 2005, showed that fusing those parts of the gait signature which are invariant with time are beneficial in gait identification. There is currently not much work that investigates the effects of speed on the performance of gait recognition methods and the relationship between the gait features and the varying walking speed [Tanawongsuwan and Bobick, 2003].

Tanawongsuwan and Bobick, 2004, observed that appearance-based features derived from silhouette of walking people are speed-dependent and therefore, a preprocessing stage for feature adjustment is suggested to improve the recognition performance.

Sundaresan et al., 2003, proposed an adaptive filter that is used to filter the foreground. This is done by summing the signal prior to the measurement of the gait cycles using the minima of the signal. However, most of the gait cycle detection algorithms suffer from accuracy problems due to the varying walking speed as well as badly segmented silhouettes. This indicates that fusion of dynamic features yields more discriminative features which would boost the recognition performance. The same conclusion is confirmed by the work of [Tanawongsuwan and Bobick, 2001] using maker-based solutions for feature extraction.

For the recovery of the angular measurements and deriving view-invariant gait signature from the different viewing planes into the sagittal plane, the approach described by [Spencer and Carter, 2005] was used.
A project by name HumanID included the following research institutions: University of Southampton, University of Maryland, Georgia Institute of Technology and Massachusetts Institute of Technology. In this research, they used the SOTON gait database developed by the Information: Signals, Images, Systems (ISIS) Research Group at the University of Southampton [Shutler et al., 2002] for the analysis and evaluation of automated extraction of gait features as well as gait recognition. Several gait databases were developed primarily for the HumanID at a Distance program sponsored by the Defence Advanced Research Projects Agency (DARPA). The program was aimed to improve technologies for facial and gait recognition as well as new technologies for people identification.

Rahmatalla et al., 2005, claimed that restrictive clothing can impose constraints on the relative joint angle limits of the walking subject and therefore affect their gait pattern.

Despite the fact that static features were proved by experiments to achieve promising gait recognition rates, their use for the development of a biometric system was found to be impractical. This is mainly because, static features are dependent on clothing [Wagg and Nixon, 2004], bags, and other factors [Phillips et al., 2002; Veres et al., 2005] which would certainly affect the recognition performance.

Various different ways of extracting discriminative features from gait sequences have been proposed [Nixon and carter, 2004, 2006] for the purpose of people identification. An important issue in gait recognition is the derivation of appropriate features that can capture the discriminative individuality from the subject’s gait. Such features should respond to crucial criteria such as robustness and invariance to weather conditions,
Lynnerup and Vedel, 2005, affirmed the usefulness of gait analysis in forensics. They were able to identify two bank robbers by matching surveillance images with images of the suspects and this evidence was later used to convict two suspects. Based on body features, gait and anthropometric measures, Lynnerup argued that there was a strong resemblance between the suspect and one of the perpetrators.

The Baseline algorithm has been used as a benchmark for performance comparison in the works in gait recognition. The performance of our method is much higher than the Baseline algorithm. The Baseline algorithm performs very poorly on the GTech database. The performance of the proposed feature set is 41.4% higher at rank 1 and 32.47% higher at rank 5 than that of Baseline algorithm. However, the mass vector approach reported in [Hong et al., 2007] performs slightly better than our method at both rank 1 and rank 5. The dimension of the gait feature vector is lower in our technique which is 115 compared to $128 \times N$ for the mass vector method. Here $N$ represents the number of frames in the silhouette sequence.

Boulgouris et al., 2006, developed a methodology for gait recognition based on the concept of discrete time wavelet (DTW). They exploited the periodicity of walking to partition the gait sequence into cycles. They located the frame indices where the sum of the pixels is minimized corresponding to the half gait cycle. Autocorrelation function of the sum of the pixels was used to determine the cycle length because of the noisy nature of the original function. DTW was performed between all cycles of probe and gallery sequences after the partitioning of sequences into gait cycles. The recognition rate showed increased performance over the Baseline algorithm on the GC
database. They also developed an angular transform which gives the average distance from the centroid of the silhouette to a small group of pixels on the contour of the silhouette. The silhouettes were pre-processed before the application of the angular transform to remove isolated errors and artifacts. All the silhouettes were aligned so that their centers were at the center of respective frames. The method was evaluated using the gait challenge database and performed considerably better than the Baseline algorithm. Linear Time Normalization (LTN) with angular transform features was used. They found that linear time normalization worked better than DTW and succeeded in achieving 8-20% increase in recognition performance when compared techniques on GC database. They studied the recognition performance of 4 sets of gait features using 5 different recognition methods, [Boulgouris et al., 2005]. An average CMS of 36% was obtained by using DTW, LTN, HMM and structural matching at rank 1. The performance at rank 5 was about 60% for all these 4 pattern recognition techniques. The only exception was noted using the frequency domain distance which performed significantly lower achieving only 20% and 41% average CMS value at rank 1 and rank 5 respectively.

DTW was used for pattern matching between gallery and probe sequences. The performance of the features for gait recognition was not very impressive and it further degraded when automatic silhouettes were used. The average rank 1 performance of 25% was achieved for manually extracted silhouettes which dropped to 18% for the automatically calculated silhouettes. A 3D human body model consisting of 11 body segments was developed by [Gu et al., 2010]. The head was represented by a sphere and other segments were cylindrical. The model contains 10 joints with 24 DOF. The kinematic
structure of the model was estimated by employing anthropometric constraints between ratios of limb lengths. After the body segmentation, adaptive particle filter was used to track the body segments. Gait features were extracted from pose parameters and joint position sequences. Two gait models were obtained from normalized joint sequence of the whole body and the normalized joint sequence of two legs using an exemplar-based Hidden Markov Model (HMM). MAP estimation was used for pattern classification. The test database consisted of multiple video streams of 12 subjects that were simultaneously captured from multiple static calibrated cameras. Volumetric representation sequences were created using visual hull method after foreground extraction. An average recognition rate of 94.4% was reported on the test database.

A subspace approach based on the matrix representation of gait data was developed by [Xu et al., 2006]. The image matrix is concatenated into a single dimensional vector to apply PCA and Linear Discriminant Analysis (LDA). The well-known curse of dimensionality due to large dimension compared to much smaller number of samples give rise to errors. The proposed matrix based coupled subspace analysis and discriminant analysis with tensor representation attempted at resolving the dimensionality issue.

A CMS of 89% was achieved for probe A at rank 1 compared to 73% for the Baseline algorithm.

Ioannidis et al., 2007, designed three new feature extraction methods for gait recognition. Two methods described as radial integration transform and circular integration transform are based on radon transform. Their third approach for feature extraction was based on weighted Krawtchouk moments. The recognition results were the highest for the Krawtchouk moments followed by radial integration
transform and circular integration transform on GC database. They also used a feature fusion scheme based on genetic algorithm to improve the recognition performance. An improvement of 1-8% was obtained using all three types of features.

Boulgouris and Chi, 2007, used Radon transform of the binary silhouettes to generate templates. LDA and subspace projection was applied to obtain a low-dimensional feature vector consisting of selected Radon template coefficients. These selected Radon features were used for gait based recognition. Each gait sequence was represented by a sufficiently compact signature of 40 coefficients. A significant improvement in recognition performance was achieved by them over the Baseline benchmark results.

Tao et al., 2007, developed general tensor discriminant analysis to apply directly to tensor data without any vectorization. This resulted in overcoming the under-sampling problem and also provides more robust features. Each gait sequence was partitioned into gait cycles after determining the gait period. The silhouettes in each gait cycle were averaged to one image that represented the whole cycle. These images were then used as features directly for classification using general tensor discriminant analysis. Gabor based features were also used in combination with general tensor discriminant analysis. The highest average CMS of 60.58% was achieved for a combination of Gabor, general tensor discriminant analysis and LDA.

Lu et al., 2008, proposed a layered deformable 2D body model for gait recognition. Their model is a full body model consisting of 10 body segments specified by 22 parameters. These 22 parameters define the size, position and orientation of the body segments. The limb orientation and position was estimated using mean shift algorithm for manually labelled silhouettes. The joint angles were then calculated
from limb orientations and positions using simple geometry. A coarse to fine estimation based on the ideal human body proportions (eight-head height) was proposed for automatically extracted silhouettes.

Multilinear PCA (MPCA) was developed by [Lu et al., 2008]. MPCA was introduced to apply on the 3D gait data directly by representing it as tensors. The application of subspace projection directly to 3D gait tensor data mitigates the famous curse of dimensionality problem. It also preserves structural information which is lost when data is vectorized for processing with traditional PCA and LDA. The tensor data was first normalized to make all tensors equal dimension.

MPCA is then applied to obtain eigen tensors. Classification was performed using different distance functions. The GC database was used for performance evaluation. The average recognition performance of 54% and 76% was obtained at rank 1 and rank 5 respectively.

This is a significant improvement over Baseline algorithm performance of 42% and 79% at rank 1 and rank 5 respectively. In a latter work [Lu et al., 2009], they introduced uncorrelated multi-linear discriminant analysis which was shown to perform even better than MPCA.

Arai and Andrie, 2012, reported gait recognition results on Chinese Academy of Sciences (CASIA) data set consisting of 31 male and 31 female subjects. They extracted silhouettes by simple background subtraction. The skeleton was then extracted using thinning and other morphological operations. Eight important feature points were then determined on the extracted skeleton structure. The skeleton was reconstructed by connecting 8 points with straight lines. Motion was also estimated using simple frame subtraction method. Discrete wavelet transform was used on skeleton data and motion
signals to extract features for recognition. They achieved an average correct recognition rate of 95.97% on the test database.

The idea of using more than one view to extract gait features has been attempted in several works such as the one reported in [Huang and Boulgouris, 2008]. The use of frontal view video instead of the usual side view for gait recognition was also reported in some works. Goffredo et al., 2008, used frontal view camera image sequence for gait recognition. After calculating the gait period, 3D gait volume was constructed using silhouettes from one gait cycle. The feature vector composed of 3D central moments of the gait volume and some scale dependent gait features including the number of frames for one gait cycle and the silhouette’s height and width maximum increment. They performed experiments on 3 data sets including CASIA-A and CASIA-B data. A correct classification rate of 91% and 97.92% was obtained for CASIA-A and CASIA-B respectively.

Yang et al., 2008, decomposed gait energy image using Gabor wavelet kernels with 5 different scales and 8 orientations. The feature vector was constructed using the Gabor phase and LDA was applied to reduce the dimension of the feature space. Comparative performance on GC database showed that Gabor phase possesses more discriminatory power than Gabor magnitude. An average CMS of 62.25% was achieved using Gabor phase compared to 51.88% for magnitude.

Chen et al., 2010, proposed a layered time series model which is a two level model combining HMM and dynamic texture model. The gait cycle was first partitioned into temporally adjacent clusters of equal number of frames. Frieze feature and wavelet feature were then extracted from these clusters. Individual linear dynamic texture models were trained for each cluster that represent the states of the
HMM. The evaluation was done using CASIA gait dataset B consisting of 124 subjects recorded from 11 views [center for biometrics, 2003].

Wavelet features outperformed the frieze features in their experimental analysis. An average recognition rate of 95.7% was obtained using layered time series model technique which was higher than that of dynamic texture model and HMM results of 58.6% and 93.9% respectively. Wang et al., 2012, modified the gait energy image and constructed chrono gait image to include temporal information. After gait period detection, they used local information entropy to obtain the gait contour from the silhouette images. Synthetic chrono gait images were also constructed to avoid over fitting due to smaller number of real chrono gait images. LDA and PCA were applied for dimensionality reduction. A comprehensive experimental evaluation was reported using 3 major gait databases. An average CMS value of 48.64% and 66.81 was achieved at rank 1 and rank 5 respectively using all 12 probe sets of GC database. These results did not show marked improvements over related gait energy image method and were only marginally higher.

A channel coding method based on distributed source coding principles was adopted for human gait recognition by [Argyropoulos et al., 2009]. The framework is different from the traditional pattern recognition approach that is used for feature matching in gait recognition works. They experimented with features extracted using radial integration transform, circular integration transform and Krawtchouk moments. The gait features are then coded using SlepianWolf encoder implemented by using a systematic low density.
2.2.3 Structure-free Methods

These methods characterize the motion of the body, without regard to its underlying structure. They can be further divided into two main classes. The first class of methods consider the movement to be comprised of a sequence of poses of the moving person and hence recognize it by recognizing a sequence of static configurations of the body corresponding to each pose [Murase and Sakai, 1996]. The second class of methods characterizes the spatiotemporal distribution generated by the motion in its continuum, i.e. without decoupling the spatial and temporal dimensions of the motion.

2.2.4 Whole-body Movement Analysis

The study of whole-body human movement spans several areas of computer vision mainly, including:

1. Human detection that detect and locate any moving person in video.
2. Activity recognition, that recognize different movement or activity types such as walking, running, limping, dancing and throwing.
3. Biometric identification and verification that determine or validate the identity of a moving person from their gait in some database.

2.2.5 Holistic Approach

Body movement is characterized by the statistics of the spatiotemporal patterns generated in the image by the moving silhouette. They hence obviate the need to track image features by instead computing these correspondence-free features. These features are inherently appearance-based.

Niyogi and Adelson, 1994, extracted four silhouette signatures of a moving person, two of which correspond to the outer boundaries of
the person and the other two to the inner edges of each leg. Each signature is normalized, via spatial and temporal alignment and scaling (i.e. so that it is stationary in the image and has a fixed height, a fixed period and a fixed phase). These normalized signatures define a spatiotemporal sheet over the entire image sequence. Gait recognition is done by matching these sheets for the model gait and input gait.

There are two main approaches for gait recognition, namely model-based and appearance based. The model-based approach explicitly models the human body based on body parts such as foot, torso, hand and leg. Model matching is usually performed in each frame to measure the shape or dynamics parameters.

2.2.6 Model-based Approach

Model-based approaches, on the other hand, model either the shape or the motion of the body, in order to recover explicit features of gait mechanics such as stride dimensions and the kinematics of joint angles [Nixon et al., 1999, 2001, 2002].

Cutting, 1978 studied human perception of gait using moving light displays (MLD) and showed human person identification results.

The method of Johnson and Bobick, 2001, is the most similar parametric gait recognition technique. Their method extracts four static parameters namely the body height, torso length, leg length and step length and used them for person identification. These features are estimated as the distances between certain body parts when the feet are maximally apart i.e. at the double-support phase of walking. They used stride parameters (step length only) and height-related parameters (stature, leg length and torso length) for identification. They considered stride length to be a static gait parameter, while in fact it varies considerably for any one individual over the range of their
free-walking speeds. The typical range of variation for adults is about 30cm [Ralston et al., 1981]. Hence, both cadence and stride length are used. Their approach for estimating the step length does not exploit the periodicity of walking and is hence not robust to systematic tracking and calibration errors.

Barton and Lees, 1997, applied Artificial Neural Network (ANN) to differentiate simulated gait (e.g., leg length discrepancy), using features from lower-limb joint-angle measures.

Holzreiter and Kohle, 1993, applied ANNs for classification of normal and pathological gait, using force platform recordings of foot-ground reaction forces. Gait classification between young and older adults was carried out using ANN. In their experiment, they used a three-layer ANN with weights adjusted, using the Scaled Conjugate Gradient Algorithm [Moller, 1993] to train relationship between gait features and the respective gait class. The ANN model had an input layer consisting of 24 neurons corresponding to the input gait features, one hidden layer and an output layer with two neurons representing gait types. After training a NN model, its generalization ability was evaluated using the three cross-validation test samples. Previous research on automated gait classification has used NNs and fuzzy clustering techniques for applications in diagnosis of pathological gait [Malley et al., 1997].

Polana and Nelson, 1993, detected periodicity in optical flow and used these to recognize activities such as frogs jumping and human walking.

Given the ability of human to identify persons and classify gender by the joint angles of a walking subject, Goddard, 1992, developed a connectionist algorithm for gait recognition using joint locations obtained from MLD. Computing joint angles from video
sequence is still a difficult problem, though several attempts have been made on it [Sidenbladh et al., 2000]. The particular difficulties of joint angle computation from monocular video sequence are occlusion and joint angle singularities. Self-occlusion of a limb from the camera view causes difficulties in tracking the hidden limb.

Rehg and Morris, 1997, pointed out the singularity in motion along the optical axis of a camera. Yam et al., 2002, recognized people based on walking and running sequences and explored the relationship between the movements that were expressed as a mapping based on phase modulation. The model-based approaches are easy to understand. However, these methods require high computational cost due to the complex matching and searching processes involved. On the contrary, the appearance-based approach is more straightforward. These methods generally apply some statistical theories to characterize the entire motion pattern using a compact representation without considering the underlying motion structure.

There are a number of appearance based algorithms for gait and activity recognition. Cutler and Davis, 2000, used self-correlation of moving foreground objects to distinguish walking human from other moving objects such as cars.

Shutler et al., 2000, used higher order moments summed over successive images of a walking sequence as features in the task of identifying persons by their gait. Instead of using moment descriptions and periodicity of the entire silhouette and optical flow of a walker, they divide the silhouettes into regions and compute statistics on these regions. They studied the capacity of features in tasks beyond person identification such as gender classification.

Bobick and Davis, 2001, used a time delayed motion template to classify activities. Huang, 2001, implemented a template matching
approach by combining transformation based on canonical analysis, with Eigen space transformation for feature selection.

Liang Wang et al., 2002, represented static pose changes of silhouettes over time as a sequence of associated complex configurations in a common coordinate, and then analyzed using the Procrustes shape analysis method to obtain gait signature. The k-nearest neighbor classifier and the nearest exemplar classifier based on the full Procrustes distance measure were adopted for recognition.

Liang Wang et al., 2002, implemented an improved background subtraction procedure to accurately extract spatial silhouettes of a walker from the background. They used eigenspace transformation to time-varying silhouette shapes to realize feature extraction. The nearest neighbor classifier using spatio-temporal correlation or the normalized Euclidean distance measure is finally utilized in the lower-dimensional eigenspace for recognition and some additional personalized physical properties are selected for the validation of final decision.

Liang Wang et al., 2003, applied eigenspace transformation based on Principal Component Analysis (PCA) to time-varying distance signals derived from a sequence of silhouette images to reduce the dimensionality of the input feature space. Supervised pattern classification techniques were finally performed in the lower-dimensional eigenspace for recognition. This method implicitly captures the structural and transitional characteristics of gait.

Liang Wang et al., 2003, implemented a visual recognition algorithm based upon fusion of static and dynamic body biometrics. For each sequence involving a walking figure, pose changes of the segmented moving silhouettes are represented as an associated sequence of complex vector configurations and were then analyzed
using the Procrustes shape analysis method to obtain a compact appearance representation called static information of body. A model-based approach is presented under a condensation framework to track the walker and to recover joint-angle trajectories of lower limbs called dynamic information of gait. Both static and dynamic cues are respectively used for recognition using the nearest exemplar classifier. They were also effectively fused on decision level using different combination rules to improve the performance of both identification and verification.

Liang Wang et al., 2003, presented a comprehensive survey of research on computer vision based human motion analysis. The emphasis is on three major issues involved in a general human motion analysis system namely human detection, tracking and activity understanding. Various methods for each issue are discussed in order to examine the state of the art. Finally, some research challenges and future directions are discussed.

A gait sequence between each frame and the specimen FED (Frame to Exemplar Distance) vector, complete gait recognition using HMM. The FED vector captures both structural and dynamic traits of each individual. For compact and effective gait representation and recognition, the gait information in the FED vector sequences is captured in a Hidden Markov Model (HMM). In the first method referred to as the indirect approach, the high-dimensional image feature is transformed to a lower dimensional space by generating what they call the FED Distance. In the second method, referred to as the direct approach, we work with the feature vector directly (as opposed to computing the FED) and train an HMM. The HMM parameters is estimated based on the distance between the exemplars and the image features. HMM is suitable for gait recognition because of
its statistical feature and it can reflect the temporal state-transition nature of gait. HMM was applied to human identification in [Iwamoto et al., 2003; Sundaresan, 2003]. They implemented HMM-based approach for gait representation and recognition.

Utsumi, 2004, transformed few-view features to arbitrary-view features. The method successfully adapted to view changes and illustrated the effectiveness by experiments, using various-view straightwalk gait sequences.

Kale, 2004, considered the width of the outer contour of the binarized silhouette of the walking person as the gait signature. They modeled the temporal state transition nature of gait by using Hidden Markov Model (HMM).

Yu, 2006, discussed the effects of view angle variation on gait identification and reported a performance drop when view difference is large.

Makihara et al., 2006, integrated a frequency-domain gait feature [Sagawa, 2006] and a bilinear model as View Transformation Model (VTM).

Han and Bhanu, 2006, extracted the GEI of the walking person. Lu and Zhang, 2007, exploited Independent Component Analysis (ICA) to extract gait feature from human silhouettes.

2.2.7 Human Tracking

To track same object between two consecutive frames some correspondence should be there. If correspondence is found then mark that object with same color rectangle and classify that object as human being. To find correspondence, few features [Suzaimah et al., 2011] of previous frame are stored and compared with the next frame, if features matches, then same object found in both frames. And these
features can be color, orientation, speed and intensity etc. For tracking any object these features play an important role. Tracking can be categories in two types, one is region based on tracking and second one is contour based tracking. In region based tracking features of the blob detected from two consecutive images and if both matches then these two frames object gets related. In contour based tracking the energy of the boundary/contour of blob detected from both frames and if that energy of the boundary matches, then same object present in both images is declared.

Lee et al., 1995, calculated color, orientation and intensity using center-surround method. After the feature map has been created, to find out which feature more uniquely identifies the object, these features are weighted. Using their weight the saliency map created for the detected object. Finally from the previous image feature weight vector is calculated for most salient region and this feature weight vector is matched with the subsequent next frame’s feature weight vector. If match is above the threshold value, then there is a match and can predict that the same object is present in both images. If match is not found, the search area is doubled and the same salient feature matching procedure is repeated. If the detected object match is not found after doubling the search area, then object is expected to be occluded by some other object or any stationary background object.

Dong Xu et al., 2012, proposed patch distribution feature for gait recognition referred as Gabor-PDF. In order to extract the PDF they have represented a gait Energy Image GEI as a set of Gabor feature and learned the image in two-stage approach and exploited the image distribution. They also developed a classification method LGSR locality-constrained group sparse representation. They carried experiments on
USF database and demonstrated the effectiveness of Gabor-PDF and LGSR method.

Michelle Karg et al., 2010, recognized gait at distance. They have addressed inter-individual and person dependent recognition. They have compared PCA, kernel PCA, linear discriminant analysis and general discriminant analysis to extract relevant features for classification and to reduce temporal information in gait. They have attained 95% of accuracy in person-dependent recognition by observing the stride length.

Jin Wang and Abbas Kouzani, 2010, have provided a survey in recent development approaches on gait recognition. They have also discussed three major issues of gait recognition system namely gait image representation, feature dimensionality reduction and gait classification. They have also discussed the gait datasets, research challenges and future direction of the gait.

Changhong et al., 2009, has refined a framework of factorial hidden Markov model which fused the multiple gait features without concatenating it into single feature. They also refined the parallel HMM which combined recognition results at decision level. They also compared the feature-level fusion and decision-level fusion statistically. They have used CASIA database. They have concluded feature level fusion FHMM performs better only when few gait cycles are available for recognition.

Dadashi et al., 2009, have proposed gait recognition by describing body outer contour using wavelet packets. Feature has been selected by the matching coefficients of the wavelet packet to that of the signal. Transductive Support vector Machine (TSVM) used to classify gait. This is a semi supervised identification. The proposed methodology has given results for orthogonal view of the individual.
Nikolaos et al., 2013, proposed a combination of holistic and model based features for recognition. Features of gait silhouette are extracted from both holistic to capture gait dynamics and model based to capture sub-dynamics. Two hidden Markov model is used to combine the gait features.

Liang Wang et al., 2003, proposed spatial-temporal silhouette analysis. They have proposed for each sequence of image a background subtraction algorithm, procedure for segmentation, Eigen space transformation based PCA is used to reduce dimensionality of the image sequence and supervised algorithm is used for classification of gait. They have also concluded more work has to be done on gait with viewing angle, unconstrained environment, and clothing. Their Future enhancement is to develop more sophisticated gait classifier.

Shiqi Yu et al., 2009, proposed classifier for gender based on gait analysis. Human able to recognize people based on gait. The prior knowledge from experiments used for automated classification to improve accuracy. They have used numerical analysis on human components and identified discriminative human components. They have done cross-race experiments to prove gait based classification of gender is achievable in controlled environments. They have used CASIA data set and given correct classification rate with other algorithms. They have noted that gait recognition suffers from many difficulties like view variation, clothing shoes change and carrying objects.

Junping Zhang et al., 2010, deals with the two issues of gait recognition. One issue is low resolution which reduces gait recognition. The issue is gait sequence data complexity which also reduces recognition. They have proposed super resolution with manifold sampling to reduce redundancy of the image and back projection.
which deals with high-resolution of the image. They have also reduced dimensionality without adding any new parameter. They have used theoretical analysis of the algorithm for recognition. Statistic is more on analysis of data but not with the statistical inference.

2.3 HUMAN BEING DETECTION

In data acquisition phase different kinds of sensing and capturing input devices are used for data gathering from the working environment. Video capturing and monitoring devices are used to record video stream. For image framing, video must be divided into sequence of frames. Video must be divided mostly into 20-30 frames, which are sent to the next phases for further processing. Background subtraction method applies on the static background images [Connell, 2004]. System can detect ‘motion area’ (only useful data) by comparing reference frame with current frame and get the extraction of a (new) moving object from the background.

2.3.1 Tracking-by-detection

While many tracking methods relay on background subtraction from one or several static cameras [Berclaz et al., 2006; Khan and Shah, 2006], many progress in object detection has stimulated the interest in combining tracking and detection procedures. In contrast to data association based tracking approaches which link detection responses to trajectories by global optimization based on position, size and appearance similarity [Andriluka et al., 2008], the combination of object detectors and particle filtering results in algorithms, that are more suitable for time-critical, online applications, Ess, 2009.

Similarly, tracking can be performed by exploiting a classifier trained to distinguish between object and background. Similar
approaches exist that apply classifiers with different confidence thresholds or accumulate detection probabilities temporally, Zhao et al., 2004. However, the extension of these methods to robust multi-target tracking is not trivial, relaying on the detector confidence in every situation can cause tracking errors, particularly during occlusions between interacting targets and in complex, cluttered scenes.

2.3.2 People tracking system

In an image sequence, people detection can be implemented frame by frame. The output of the detection system such as position, shape and color of the object can be used to construct a tracking model. Because there exist a large correlation between consecutive video frames, the location and appearance of people do not undergo a dramatic change within a certain period. Therefore the state of tracking method is focused on how to use these kinds of properties to solve feature matching problem in video sequences.

Based on the acquired video data tracking systems can be divided into monocular and multi-view approaches. Monocular approaches use the video image from a single view to perform tracking. Motion model is used for each person to estimate its position in subsequent frames, additionally a dynamic appearance model is constructed by combining gray-scale textural appearance and shape information of people. It is capable of tracking multiple persons simultaneously and people can still be tracked through interaction such as occlusion. The assumption of this method is that the motion of people is predictable such as walking and running. It is primarily designed for outdoor surveillance. As for the random motion of people in an indoor scene their motion model does not fit well. The distribution of color on human body is utilized to build 2D blob model
of people. The body parts can be reliably labeled and located by 2D contour shape analysis. The color tracker can still track the hands in front of the body, although in this situation there is no contour evidence of hands. Tracking performance is significantly increased by taking color into account. The limitation of this method is that it describes the blob model as a single human which will cause problem for multiple user gesture recognition.

Multiple-view approaches utilize multiple cameras to record the video data which is effective to calculate the precise 3D position of people in complex environment. In the indoor video sequences are shot with four synchronized cameras and outdoor with three cameras. With these data, they build up a stochastic model to compute the optimal individual trajectories and then use optimization schemes to process the trajectories one after another. It is shown that this algorithm can follow up to six individuals with significant occlusion, but in the outdoor surveillance they do not mention the influence of other moving objects, for instance, a car that is parked in the scene, this sudden change in the background will degrade the performance dramatically.

Cai and Aggarwal, 1998, used a Bayesian tracking scheme to locate the most likely match of the object in next frame. The automatic camera switching selects the camera with the best view during tracking period, although they have developed techniques of multiple view transition tracking to deal with occlusion at a certain level by using distributed monocular cameras, it still remains a major obstacle in their tracking system.

Many current methods for detecting and tracking people relay on color contrast or movement to segment the image. Using color, requires the target and the back ground to be significantly different
and motion segmentation requires the target to be in constant motion relative to the background often requiring stationary cameras.

Sidenbladh and Kraigi, 1999, used binary skin classification to locate faces within the image. A closed-loop controller is then used to maintain tracking of the person by cameras. A more statistical approach comes from Tarokh and Ferarri, 2003, who use clothing color to segment the image. Statistical analysis is then used to determine which color blobs belong to a person and which do not. A more sophisticated approach was discussed by Kwon et al., 2005, who used color histograms to locate a person in a pair of stereo images followed by triangulation between the images to find the distance between the camera and the tracked person. Schlegel et al., 1998, used color histograms. They combined the histogram data with an edge-based detector to improve robustness of the algorithm at the expense of computation time.

Piaggio et al., 1998, approached the problem by threshold the velocity data and assuming that the person moves differently from the background.

Chivil`o, 2004, viewed velocity changes as a disturbance to be minimized by regulation (i.e. moving the camera or the robot to minimize the relative motion of the person, thereby providing tracking).

Munkelt et al., 1998, attempted to detect people in a video by calculating correspondence between features in the image and a 3D model of the internal structure of the target. They presented a complex articulated model of a person and showed how it could be matched to a person in an image using color and 3D information. 3D scene information is provided by either a stereo camera setup or
through matching basic shapes in the image with basic shapes in the model and determining depth and relative position information from this.

Inamura et al., 1998, presented a method of person detection and tracking that combines motion detection, edge detection and color detection with a voice recognition system. Krinidis et al., 2005, implemented a database of audio and visual data as a standard for testing of tracking systems. Their database covers a range of situations from a single person in a simple scene to multiple peoples moving in different directions. Each of these are repeated for various lighting conditions.

2.3.3 Feature selection for people detection

Perez et al., 2004, presented the data fusion method to include color, stereo sound and motion cues to a human model. It is beneficial in a visual tracking system for the richness and complexity. When a predefined class of objects needs to be tracked other information such as shape cues can be involved to extend the model for various scenarios.

Harasse and Bonnaud, 2006, constructed a human model by using a shape characteristic which described explicitly spatial relationship inside the human body. In their application the heads were well separated from each other. This model multiple persons who can be detected under this specific situation. However, for the occlusion of the whole human body including the head, this algorithm cannot identify different people because of its model limitation.
2.3.4 Neural network in gait recognition

Morris et al., 2004, used a Gait Shoe system consisting of several sensors (bend sensors, ultrasound, accelerometers etc.) attached to a shoe. The main goal was to extract clinical relevant information from the sensor. But in addition to distinguish between subjects with Parkinson's disease and without, he also analyzed if subjects can be recognized based on their gait. 25 features were extracted from the different sensor outputs and a correct classification rate of 97.4% was obtained using Neural Networks and data of ten subjects.


2.4 SUMMARY

The literature has shown the usage of gait analysis in clinical analysis. The human movement analysis with fixed labeling of body parts and free body movement labeling are used. The angular movements of the different parts of the body considering leg and hand are very much used for gait analysis. Mostly close laboratory experiments have been considered in many PhD thesis. Standard
database like CASIA have been used. Many researchers have used hidden Markov model, Gabor, discrete wavelets, LDA, PCA for extracting representative features.

All these analysis have been carried out only for controlled movement of the human walk in laboratory. Each feature extraction methods claim superior over other methods in different literature. Most of the work reported based on the algorithms like Gabor, wavelets, LDA, PCA which have been used in many another applications other than gait analysis. With regard to artificial neural network (ANN) training, radial basis function algorithm, back-propagation algorithm have been used by many researchers. As such counter propagation algorithm has been used by one researcher. However, the extraction of features have not been mentioned. Hence, in this thesis, CPN is considered for implementation. This chapter has presented the different methods used for gait recognition. In each method, the quality of gait recognition is achieved to a limited extent. This has given scope for implementing Artificial Neural Networks for gait recognition. Chapter 3 presents gait data collection.