GAIT RECOGNITION SYSTEM

By

Muhammad MuneebUllah Khan

Saleh Muhammad

MuneebIlyasBhatti

Submitted to the Faculty of Electrical Engineering, Military College of Signals National University of Sciences and Technology, Islamabad in partial fulfillment for the requirements of a B.E Degree In Telecommunication Engineering

JUNE 2012

ABSTRACT

Gait is a particular way or manner of moving on foot.Gait recognition is identifying a personby the manner of its walk.This is a marker less and unobtrusive biometric; offering the possibility to identify people from a distance and without interacting with them; this property makes it an attractive method for identification.

The project aimedat developing a system capable of automatic gait recognition. A person'sgait signature was created using a model based approach. Temporal and spatial metricsextracted from the model fitted to the individual; such as changein angles of the limb or the stride length of a person'sgaitwere used to create a "gait signature" of the individual which weretransformed in Eigen Space using Principle Component Analysis and were later used to identifythe subject in videosegments.

The project provided promising results with the technique of Principle Component Analysis and even with the self-similarity plots. Limb angles proved the best way to extract a gaitsignature and angular velocities showed quite degraded results.Principle Component Analysis offered a good way to represent most of the variation in the

data, whilereducing dimensionality. Recognition rates up to 80% were achievedby the project, further strengthening the notion that gait can be used as a biometric.

DEDICATION

We would like to dedicate this project to our project leader whose principles of hard work and determination never let us down and who always stood against all the difficulties.

ACKNOWLEDGEMENT

We would like to thank Dr. Imran Siddiqi, for the initial project proposal andour supervisor Dr. AdilMasoodSiddiqiwho helped us at every stepand never lost hope in us.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Introduction

1.1 Types of Biometrics

Biometrics is used in a variety of applications, which makes a precise definition difficultto be set up. The most general definition of a biometric is, “A physiological or behavioral characteristic, which can be usedto identifyand verify the identity of an individual.”

There are,but many, biometric measures which could be used to identify an individual. They can be classified into two distinct categories; physiological and behavioral.

Physiologicalbiometrics is derived from a direct measurementof a part of a human body. The most successful and eminent of these types ofmeasures to date are fingerprints, face recognition, iris-scans and hand scans [1].

Behavioralbiometrics extracts characteristics based on an action(s) performed by an individual;they indirectly measure the characteristic of the counter human form. The keyfeature metric of a behavioral biometric is time. Establishedmeasures include keystroke-scan and speech patterns.

Biometric identification should be an automated process. Extracting feature extraction manuallyisboth undesirable and time consuming;due to the large amount of data that must be first gathered and later processed in order to extract a biometric signature. Inability to automatically extract therequired characteristics would brandthe process infeasible on realistically large data setsin real-world applications.

A unique biometric signature for an individual does not exist; each time the data is acquired from the individual,it will generate a slightly different signature, there is simply no suchthing as a 100% match in biometrics. This does not mean that these systems are inherently insecure, appreciablyhigh recognition rates have been achieved. Since the recognition is done through a process ofcorrelation and setting up thresholds, systems claiming to offer 100% recognition rates should be greeted with apinch of salt.

1.2 Advantagesand Disadvantages of Gait

Gait,as defined earlier,"A particular way or manner of moving on foot."Using gait as a biometric is a relatively new area of study, within the realms of computervision. Interest from the computer vision community is increasing day by day in gait and anumber of gait metrics have been developed. Early psychological studies into gait by Murray suggested that gait was a unique personal characteristic, with cadence and, was cyclic innature [2]. Johansson studied gait by attaching moving lights to human subjects onall the major body parts and showed these moving patterns to human observers. Theobservers could later recognize the biological patterns of gait from the moving light displays, even when some of the markers were removed, once again indicating gait as apotential candidate as a prominent biometric [3].

We used the term gait recognition to identify the individual from a videosequence of the subject walking. This does not mean that gait is limited to walking, it canalso be applied to running (the only other manner of walking present in human beings). As a biometric gait can have many advantages over other forms of biometric identification techniques for the reasons; unobtrusive, distance recognition, reduced detail, difficult to conceal.

The gait of a person can be extracted without the user knowing they are being analyzed and without any cooperation from the user in the information gathering stage unlike fingerprinting or retina scans. The gait of an individual can be captured at a distance unlike other biometrics such as fingerprint recognition. Gait recognition does not require images that have been capturedto be of a very high quality unlike other biometrics such as face recognition, whichcan be easily affected by low resolution images. The gait of an individual is difficult to disguise, by trying to doso the individual will probably appear more suspicious. With other biometrictechniques such as face recognition, the individuals face can easily be altered orhidden.

Being a biometric, an individual's biometric signature will be affected by certain factors like stimulants, physical and physiological changes and clothing.

Stimulants like drugs and alcohol will affect the way in which a person walks. Physical changes such as a person during pregnancy, after an accidentordisease affectingthe leg, or after severe weight gain (loss) can all affect the movement characteristic ofan individual. Psychological changes affecting a person's mood can also affect an individuals' gait signature [1]. The same person wearing different clothing may cause an anomaly in automaticsignature extraction method.

Although these disadvantages are inherent in a gait biometric signature, other biometricmeasures can also be easily disguised and altered by individuals, in order to eluderecognition.

The process of automatic gait feature extraction is made more difficult from the externalfactors such as lighting conditions, self-occlusion of feature points when the subjects' legscross over and different types of clothing, which can all affect the data acquisition process.

Results from previous work on automatic gait recognition look promising [6, 7, 8, 9, 10, 11], withresult ranging up to 100% successful identification rates on smaller database samples.

Subsequent work has to be carried out on larger databases to ascertain the effectiveness ofthisidentification method with a large data set, but initial reports are promising.

1.3 Project Motivation

The ability to identify an individual, efficiently and accurately is an important task.Controlled environments such as banks, military set ups and even airports needs to quickly detect threats and provide differing levels of access to different user groups.Recent events such as September 11thhave brought biometrics to a lot of attention as identification methods.

As stated earlier,gait as a biometric has many advantageswhich make it an attractiveproposition as an identification method. The main advantage of Gait is its unobtrusive identification from a distance,whichmakes it a veryappealing biometric. The ability to identify a possible threat from adistance, gives the user time to react before the suspect becomes apossible threat.

Another motivation is the readily availability of video footage of suspects, assurveillance cameras are now installed in almost every location, the videos just needs to be checked against that of the suspect.

As well as the inherent advantages of gait, the increase in processor power, along with thefall in price of high speed memory and data storage devices, all having contributed towards theincreased availability and applicability of computer vision and video processing techniques.Real time video processing, required for gait recognition, is a feasible possibility oncurrent home PC technology, making this technology a viable security application.

1.4 Gait Recognition Scenario

Gait recognition can be used in different scenarios. One example would be, analyzing the video stream from surveillance cameras. If an individual walks by the camera whose gait has been previously recorded and he is a known threat, then the system will recognize he and the appropriate authorities can then be automatically alerted and the person can be dealt with before he becomes a threat. Since the threat has been detected from a distance, itcreates a time buffer for the authorities to take appropriate action. Such systems have a large amount of potential application domains, such as airports, banks and general high security areas. This all is being depicted in figure1.1.

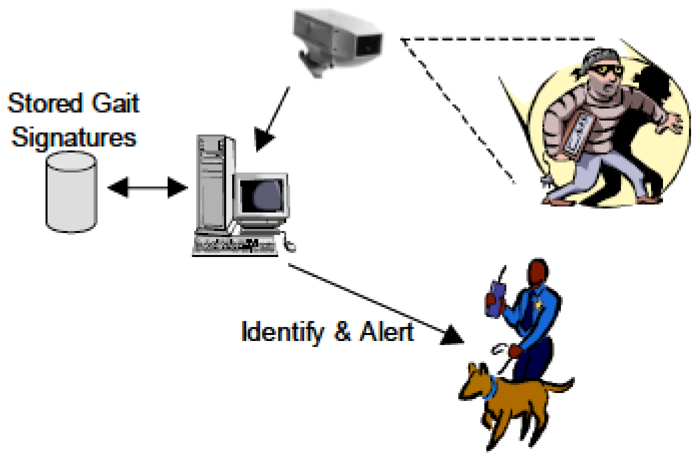


Figure 1.1 Gait Recognition in Action

1.5 Project Scope

The main objective of this project can be divided into two steps; develop a program and perform automatic extraction. In step one a program is developed capable of performing recognition of individuals derived from avideo sequence of a person walking. Automatic extraction of relevant gait feature points should be available from a videosequence in order to automate the classification process.

The project can then be broken down into three main sections, which were completed in theorder shown in figure 1.2.

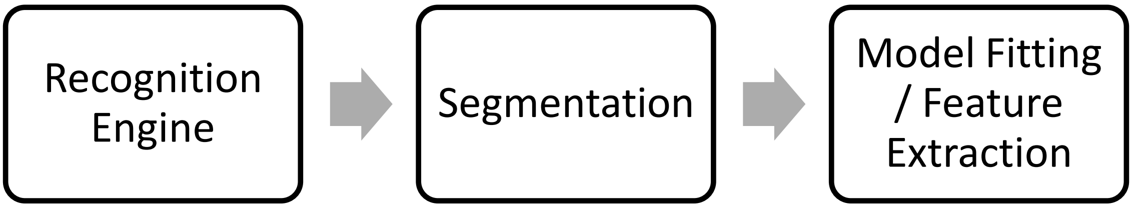


Figure 1.2 Project Process

In section one the recognition engine develops the algorithms and functionality that can classifyindividuals based on extracted gait information. The section two;segmentation,extracts the foreground subjects from the provided video sequence, ready forextracting gait features. Finally, in section three;feature extraction, the segmented image is used to extract the relevant gaitfeatures which wouldfurther be used for classification.

1.6 Report Structure

The rest of this report is structured in nine chapters. Chapter two gives an insight into previous work which has been pursued into gaitrecognition, along with theories, based upon which the methods are used throughout this project. Chapter three explains about the overall conceptual design of the project along with the different technologies, algorithms and development environments used. Chapter fouroutlines the segmentation process used to extract the person's silhouette from thevideo sequence. Chapter fiveexplains the methods used in the automatic model fitting. A model based approach was taken to extract the feature vector from the video sequence. Chapter sixexplains how the gait signatures are generated from the captureddata; these include signatures based on angles, angular velocity and self-similarityplots. Chapter sevendescribeshow the classification process takes place and how thegait signatures are projected into feature space. Chapter eightis the evaluation of all the processes carried out in this project; such assegmentation, model fitting and recognition. Chapter nineprovides an overall conclusion of the outcomes of the project.

Chapter2

Background

In this chapter the current state-of-the-art techniques in gait recognition are reviewed, providing anoverview of the methods which are currently being investigated and providing backgroundinformation on all the techniques which have been used in this project. Gaitrecognition techniques can be divided into two main categories, model based and modelfree. These approaches are described in more detail in upcoming sections.

2.1 Model Free Approach

The advantage of a model free approach is that the methods derived are not linked to one particularobject i.e. it is a holistic approach, and therefore a method detecting human gait could be used alsofor animal gait and vice versa with little modification. A number of model free approaches togait recognition have been investigated. Some of them are detailed in the following paragraphs.

Nixon aimed to use an area based metric for measuring gait signatures.The aim is to combine holistic (concerned with the whole) and model-basedapproaches to recognition, by using statistical data which linked to the gait of asubject. Foster, Nixon and Bennett suggested an implementation where thehuman silhouette is captured from the video sequence and then a gait mask isplaced over the silhouette. Each gait mask isolates a specified portion of theoriginal image; the change in the area of the silhouette inside each of the gaitmasks can be measured and then used to produce series, which can beanalyzed for unique characteristics. Examples of gait masks are shown in figure 2.1.

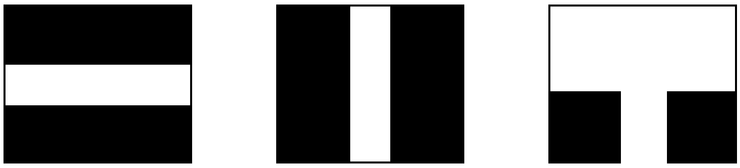


Figure 2.1 Sample Gait Masks

The technique has produced encouraging results with recognition rates of over 80% on a small database, even when noise has been added to the original videosequence and has shown that gait recognition is possible by only using thetemporal components of the human silhouette [8].

Dr. V. Huang performed gait recognition using PCA and Canonical Analysis.He used the silhouette of the subject during motion to derive the gait parameters,this motion was then compressed using PCA. He then applied CanonicalAnalysis to derive the signature from which the subject can be recognized. Arecognition rate on a small database had a success rate of 100% suggesting thatthis technique is reliable and has the potential to be improved and extended [9].

J.E. Boyd and J.J. Little used a technique of identifying individuals by studyingthe variations in the motion description of a subject as they walked. They took ashort video sequence and derived the dense optical flow of the subject in boththe x and y direction. Scalars of these values were then created based onmoments of moving points in order to characterize the shape of the motion notthe shape of individual points. They used least-squares linear predictionspectrum analysis to find the fundamental frequency and phase. These werethen used as the basis for identification comparison. The results from this lookedquite promising with recognition rates of 95% when using the four bestrecognition signature features, based on the nearest neighbor algorithm [10].

BenAbdelkader used a motion based approach, segmenting theperson from the background, and then computing a self-similarity plot of the ofthe captured foreground images over a number of frames. The self-similarity plotis a measure of correlation between two different extracted foreground regions attime $t_i$and t_j. Principle Component Analysis was then performed to reduce thedimensionality of the extracted images followed by clustering analysis. Resultswere promising with recognition rates up to 78% being obtained from a nearfronto-parallel view [18].

2.2 Model Based Approach

Model based approaches to feature extraction, use priori knowledge of the object, which isbeing searched for in the image scene. Models used are typically stick representations eithersurrounded by ribbons or blobs, as shown in figure 4. When modeling the human body, thereare various kinematical and physical constraints we can place on the model which arerealistic i.e. maximum variation in angle of knee joint.

The advantages of a model based approach are that evidence gathering techniques can beused across the whole image sequence before making a choice on the model fitting. Modelscan handle occlusion and noise better and offer the ability to derive gait signatures directlyfrom model parameters i.e. variation in the inclination of the thigh. They also help to reducethe dimensionality needed to represent the data.

The disadvantage of implementing a model based approach to is that the computationalcosts, due to the complex matching and searching that has to be performed are high.

Although this is a limitation, computing power is always increasing so this can be seen as lessof a disadvantage, especially in non-real-time applications, and efficiency improvements canbe found for most algorithmic implementations, which help to reduce the computational costs.

When modeling the human body the method shown in figure 2.2 can also be used. Modeling the human body as a skeleton is a good representation due to the underlying skeleton of the physical body and the restrictions it creates. In this model the skeleton can be thought of as rigid segments between articulating joints. In a 2D environment, rotation not in the x,y direction can cause changes in the length of these bones and has to be taken into account during the modeling stage.

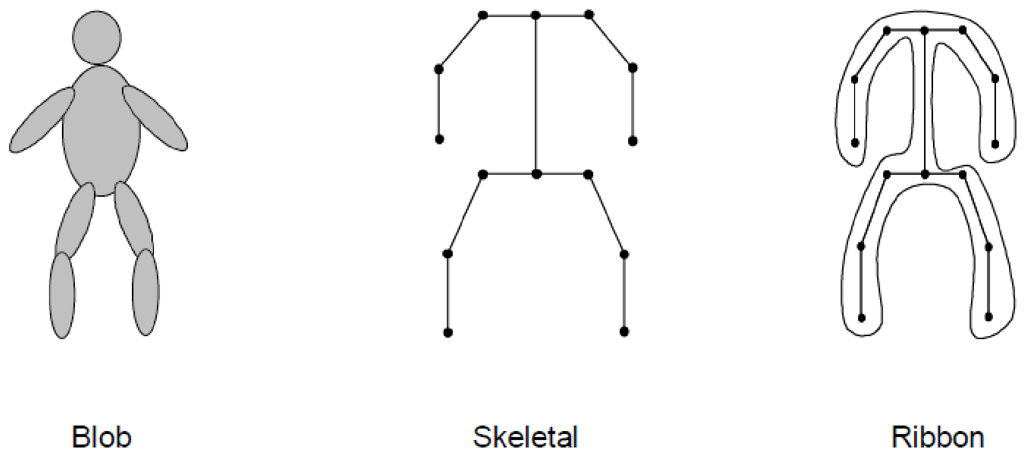


Figure 2.2 Different Modeling Techniques

It is possible to extract a skeletal model from an image using Medial Axis Transformation(MAT). MAT is a process which can be used for reducing foreground regions toa skeletal representation, whilst preserving the connectivity of the original region. Each pointon the skeletal representation contains both time and space information, allowing the originalfigure to be reconstructed. It was found stick figures could be extracted from video sequencesusing Medial axis transformations and these matched quite closely to the skeletal models ofhumans [4].

Wrenproduced a real-time people finder program (PFinder), which models and tracksthe human body using a set of blobs, the blobs correspond to the person’s hand, head, feet,shirt and trousers. The foreground region is extracted from the background and then a modelbuilding process follows where blobs are placed over the foreground region. Trackinginvolves a loop of predicting the appearance of the person in the new image, determining thelikelihood for each pixel to be part of the blob model or part of the background [5].

One of the most important aspects in gait recognition is capturing accurately the positions of the legs;these are the best source for deriving a gait signature and contain most of thevariation in the subjects gait pattern. The legs of a human are usually modeled based uponSingular Harmonic Motion (SHM) in gait applications [6,7,11]. The legs can be considered asa pair of connected pendulums, with a number of constraints imposed that accurately reflectsthe physical constraints of the human legs [6, 8].

The aim of a model based approach is to model the motion of a human, and then fit thismodel to the motion of a human being tracked [7]. Previous projects involvinggait recognition using model based approaches are detailed in upcoming paragraphs.

Dr. D. Cunadomodeled the leg as a pendulum. The method of identification was defined by calculating the difference between SHM and the motion ofthe subjects’ thighs. The gait signature was successfully extracted andcould withstand differing amounts of noise and occlusion. This methodachieved recognition rates of 100% on a database of ten subjects [7].

Nixon considered just the legs to extract the gait signature; the legs weremodeled as the motion of interlinked pendula. The Hough transform wasused to extract the lines which represent the legs in a video sequence [6].

The change in the inclination of these lines was recorded and FourierTransform Analysis is used to reveal the frequency change in theinclination of the legs. Classification rates using phase-weightedmagnitude spectra incorporating the K-Nearest Neighbor Rule had aclassification rate of 90%.

An example of the main modules of a model based approach using image segmentation isshown in figure 2.3.

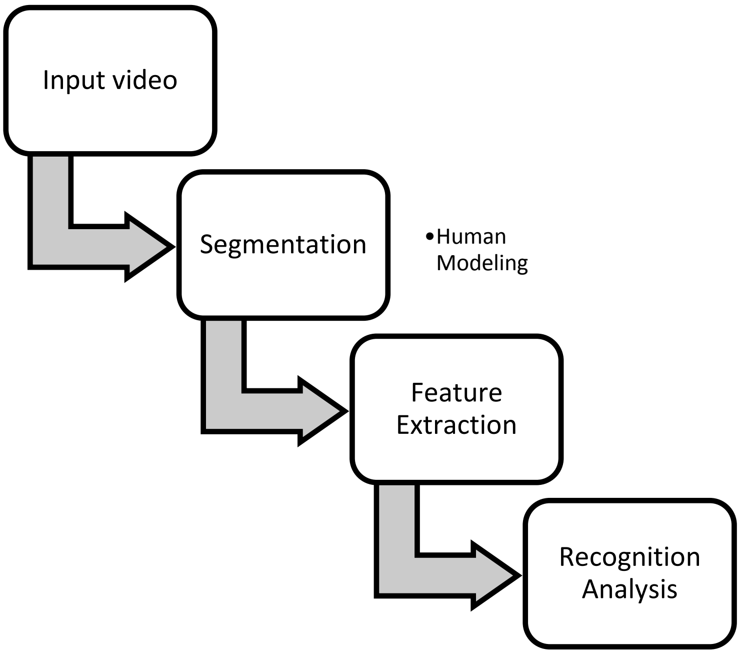


Figure 2.3 Overview of Gait Recognition Process

The basic concept is to first segment the foreground information (i.e. the subject) from the background. The features used to derive the gait signature must be extracted as accuratelyfrom the segmented object as possible. These parameters can then be analyzed and storedfor recognition use at a later stage.

2.3 Principle Component Analysis (PCA)

Principle component analysis (PCA), also known as Eigenanalysis, is a technique used toreduce the dimensionality of data and examine the relationship between a set of correlatedvariables. PCA has been used successfully before in both gait and face recognitiontechniques [9, 16, 18].

Dimensionality reduction is vital to the recognition purposes because the size of recognitionmatrices can be vast and very computationally expensive or infeasible. For example Huangcreate a feature vector by concatenating the columns of each image into onefeature vector, obviously this has a large dimension, possibly greater than 1000, which wouldbe infeasible to use for recognition purposes[13]. PCA extracts the main variation in the featurevector and allows an accurate reconstruction of the data to be produced from only a few ofthe extracted feature values, hence reducing the amount of computation needed.

The aim of using PCA is to be able to represent most of the variation of the original variablesusing only a few “principle components”. This can be seen in the figure 2.4, after performing PCA, the eigenvectors (which are all orthogonal to each other) represent the newaxis, in ascending order of variation in the original data set.

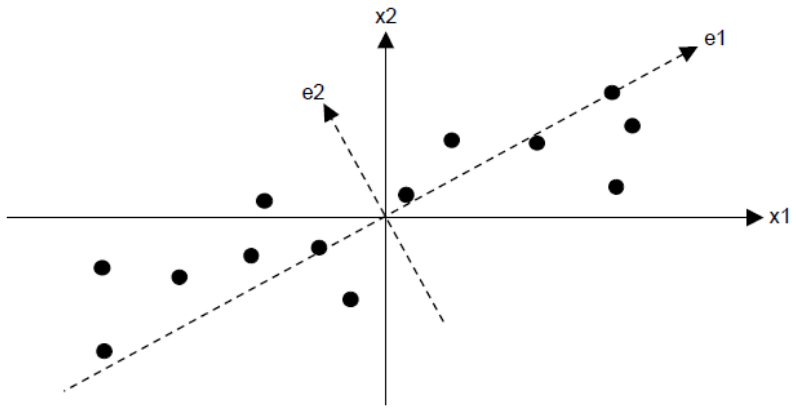


Figure 2.4 Example of PCA in Action

In order to calculate the PC of a given set of feature vectors X = x_1,x_2, …x_n, which are created by placing all of the original data for a given configuration in a single vector, first wefind the correlation matrix C. This is a symmetric matrix that helps to reduce the computation when calculating the eigenvalues and eigenvectors, the formulation is shown in 2.1.

$$C = \frac{1}{N}\sum_{k=1}^{N}(x_k - x_m)(x_k - x_m)^T \quad (2.1)$$

The mean value of the vectors is given by:

$$x_m = \frac{1}{N}\sum_{k=1}^{N} x_k \quad (2.2)$$

The mean is subtracted from all of the vectors in order to produce a set of normalized difference vectors. The correlation matrix can then be represented by a set of specialvectors, which satisfy the following equation:

$$Ce_k = \lambda_k e_k \quad (2.3)$$

These vectors are called eigenvectors, each eigenvector, e_k, has an associated eigenvalueλ_k. The largest eigenvalues of the correlation matrix represent the largest inherent variationin the original data set and tell us most about the original data. The eigenvectors can berearranged back into the form they were derived from at the end process if required (as inface recognition).

The most popular way to derive the eigenvalues and vectors is to use Singular ValueDecomposition (SVD), although may other techniques are also possible. The technique forcomputing these values has been outlined in Numerical Recipes and will not bediscussed any further here. It should be noted that if all of the variables are uncorrelated tobegin with then the PCA will not be of much use and return nearly as many non-zeroeigenvalues as we had variables in the original data.

Given that X is the feature matrix, where the matrix dimensions have the following property, m>> n, calculating the square matrix R=XX^T, which is used for PCA becomes computationallyinfeasible as mentioned earlier, Huang[19] used the following technique based on SVD toreduce the amount of computation needed to computed the eigenvalues and eigenvectors ofR. Given:

$$\lambda_i e_i = R e_i \tag{2.4}$$

Normally the square matrix R which is used for calculating PCA is computed by:

$$R = XX^T \tag{2.5}$$

Instead an alternative square matrixŘcan be computed by:

$$\check{R} = X^T X \tag{2.6}$$

Řis now a square matrix of n x n, which is a much smaller size than m x m. This significantlyreduces the computation needed for calculating the eigenvector. The eigenvectors andvalues of Řare related to R in the following manner:

$$\begin{cases} \lambda_i = \lambda_{\tilde{i}} \\ e_i = \lambda^{\sim\frac{1}{2}} X e_{\tilde{i}} \end{cases} \tag{2.7}$$

Once the principle components have been calculated, the next step is to decide how many of the PCs should be kept, in order to maintain a correct and accurate representation of theoriginal data. One method is to define a threshold value t, such that the total number ofprinciple components kept should be greater than this value usually 80-90%. The totalproportion of the variance of the original variables, accounted by *p* principle components isgiven by:

$$\nu = \frac{\sum_{k=1}^{p} \lambda_k}{\sum_{i=1}^{N} \lambda_i} \tag{2.8}$$

It is normally the case that just the first three or four principle components will be kept, as thefirst PC is usually very large and the drop off of variance representation is quite steep. Fromthe graph shown in figure 2.5, we can see that with relatively few PCs we can represent 80-90%of the total variance.

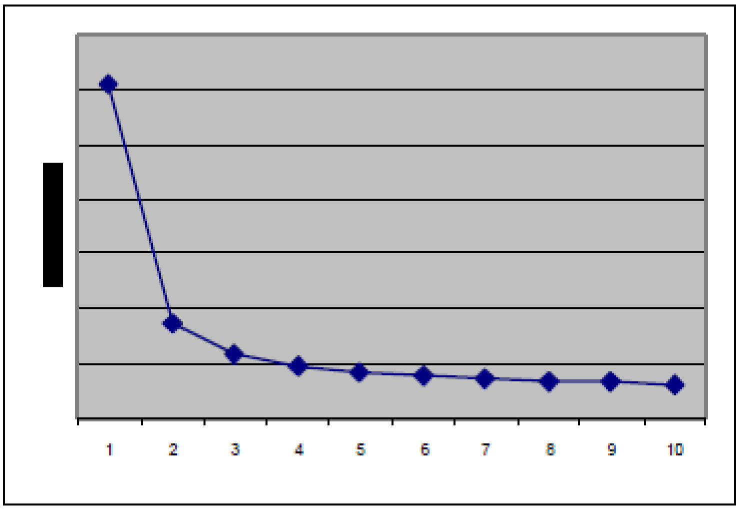


Figure 2.5 Eigenvalues of Principle Components

The final step is to project the feature vectors into the new eigenvector space; these projectedpoints can then be used for classification. Given the feature vectors $x_{i,j}$, they can be projectedinto Eigen space by multiplying the feature vectors by the newly calculated eigenvectors:

$$y_{i,j} = [e_1, e_2, \dots, e_{k-1}, e_k]^T x_{i,j} \tag{2.9}$$

As with all real numbers there will be errors introduced either by rounding (inherent in floatingpoint operations) or as noise in the initial capture of the data set. This will have the effect ofmaking eigenvalues who should have zero values to have very small non-zero values. Thisshould not be a problem if the threshold method is applied.

2.4 Canonical Analysis (CA)

Canonical Analysis offers the ability to express the relationship between two or more sets ofvariables. If items can be classified into one of g groups, then the total variation can be seenas the combination of between-group variation and within-group variation. Ideally we want tobe able to maximize the between-group ratio and minimize the within-group ratio in order toincrease separation of the different classes. Canonical Analysis finds the linear variables(canonical variants), which help maximize this ratio.An example of the outcome of Canonical Analysis can be seen in figure 2.6. The different groups of data are initially quite close and have quite large inner variance.

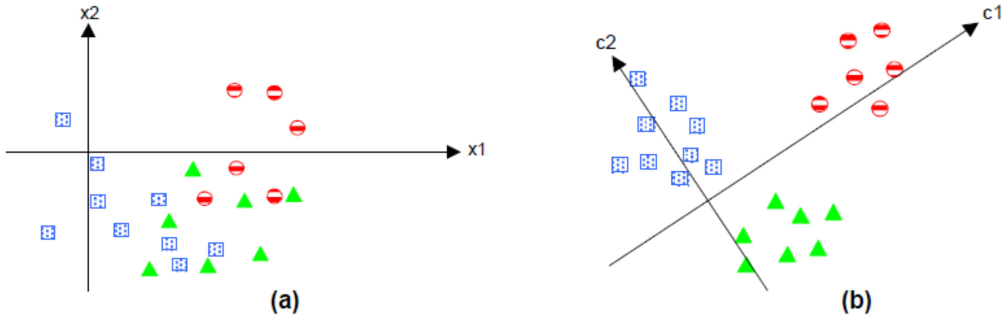


Figure 2.6 Example of Canonical Analysis

Ideally CA aims to separate the classes (as in figure 2.6 b) and minimize the within classvariance, hence improving the classification rates of the different groups.

Canonical Analysis is used in advanced classification techniques, and has been successfullyapplied to gait recognition [19]. Huang used Canonical Space Transformation toseparate different training classes of individuals walking, based on feature points projectedinto Eigen space (created after applying PCA) [19]. Computing the within $S_w$and between $S_b$classvariance of the projected Eigen points, $y_{i,j}$represents feature point j of class I (total of cdifferent classes), is done by:

$$S_w = \frac{1}{N_T}\sum_{i=1}^{c}\sum_{y_{i,j}\,\varepsilon\,\psi_i}(y_{i,j} - m_i)\,(y_{i,j} - m_i)^T \quad (2.10)$$

$$S_b = \frac{1}{N_T}\sum_{i=1}^{c} N_i(m_i - m_y)\,(m_i - m_y)^T \quad (2.11)$$

Where

$$m_i = \frac{1}{N_i}\sum_{y_{i,j}\,\varepsilon\,\psi_i} y_{i,j}$$

$$m_y = \frac{1}{N_T}\sum_{i=1}^{c}\sum_{j=1}^{N_i} y_{i,j}$$

Once $S_b$and $S_w$have been calculated the generalized eigenvector problems (2.11) needs tobe solved in order for us to find the new set of canonical axis. This will give use a set ofeigenvectors which represent the orthogonal axis in canonical space where classification cantake place:

$$S_b w_i^* = \lambda_i S_w w_i^* \tag{2.12}$$

This problem can be solved by initially preparing the $S_b$and $S_w$matrix in order to be able tocalculate the eigenvalues and eigenvectors. Firstly we perform Singular Value Decompositionon S_w:

V is orthogonal (i.e. V = inv(V)), and S is a diagonal matrix, where the diagonal elementsrepresent the singular values. The next step is to calculate a matrix U:

$$U = \left(VS^{\frac{1}{2}}\right)^T S_b \left(VS^{\frac{1}{2}}\right)^T \tag{2.13}$$

Next the matrix U is split using Singular Value Decomposition:

$$U = ABA^T$$

$$delta = VS^{\frac{1}{2}} A$$

Delta is used to help diagonalize the between and within matrices, this is needed to be able tocompute the eigenvectors and eigenvalues. Finally the matrix Eigen problem is computed,and this is the matrix which we derive the eigenvectors and eigenvalues for, which willrepresent the axis of the canonical space. There will be c-1 (c is the number of differentclasses, i.e. number of different people caught on video), non-zero eigenvectors which can beused for projecting the points into canonical space:

$$eigenproblem = delta * B * inv(delta)$$

Canonical Analysis is like PCA, the eigenvectors who have large eigenvalues represent theaxis of most variation in the data set, we only need to use the larger eigenvectors in order tobe able to represent most of the variation in the original data set.

2.5 Image Segmentation

In order to be able to perform analysis on the gait of the individuals caught on video thesubject needs to be extracted from the video sequence. Image segmentation is used toseparate dynamic objects such as people, which are part of the foreground, from thebackground of the image sequence. It is a very important pre-processing step in manycomputer vision applications, accurate retrieval of the foreground objects are vital in order tominimize distortion or inaccuracies. These may propagate into other parts of the visionsystem, which could affect results at later stages in the processing.

2.5.1 Approaches to Image Segmentation

One of the most common and simplest methods for performing segmentation is to first carryout image subtraction. The known background image is subtracted from the current pictureframe, comparing the intensities of relating pixels, then thresholding is performed. A pixel isconsidered part of the foreground when the current pixel value differs from its mean value bymore than a pre-defined threshold valueϕ:

$$f(x,y) = \begin{cases} background, if\ |I_{current}\ (x,y) - \ I_{known}\ (x,y)| \leq \phi \\ foreground, otherwise \end{cases} \qquad (2.14)$$

Segmentation is a complex process and there are several problems inherent with extractingforeground regions, such as occlusion, shadows (cast by the foreground objects as theymove) and noise. The process is complicated by the fact that the background may not bestatic, i.e. a changing television screen in the background does not want to be considered asa foreground object but is continually changing. Also a foreground objects intensity/color ata given point may be very similar to the background and hence the foreground object may beconsidered part of the background. Several techniques have been developed that includeboth range and color in order to minimize the distortion of these effects [10]. The effect ofthese distortion factors can be seen in figure 2.7 where simple subtraction has takenplace:

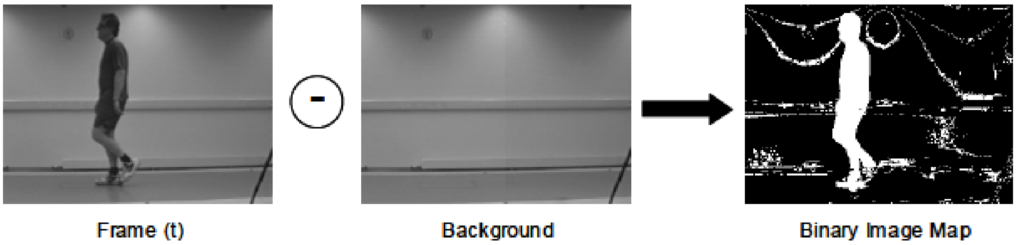


Figure 2.7 Background Subtraction

In these pictures, although the background is static, the flicker of the light, which isundetectable to the human eye causes large changes in intensities, along with the subjects'shadow, resulting in poor segmentation, the binary image shows just the subtraction process.

Segmentation can be improved by building a model of the background pixel intensities. Themodel of the background can be built on a combination of statistical range and color values for each pixel in the scene. If the background is continually changing gradually over a periodof time, the model would have to be updated over time to reflect these changes. Varioussimplification assumptions can be made in controlled environments to enhance theperformance of segmentation.

As well as creating a better model of the background in the image, it is also possible to applyimage filters to the foreground image map, which help to reduce noise (pixels which havebeen misclassified as foreground pixels) and classify pixels into groups.

2.5.2 Morphological Filters

Such filters are called Morphological Filters. They usually take a binary image and processthe objects in the input image based on characteristics of its shape, which are derived fromthe overlayed structuring element, using basic set operations such as intersection and union.

A good example of these filters in use is the W^4(Who? When? Where? What?) Project[20].This project tracks individuals and body feature points in numerous settings. They usedmorphological filters to remove background noise and classify blobs. The main filters used forimage pre-processing are Erosion, Dilation and Connected Component Labeling; these areall subsequently discussed in more detail.

2.5.2.1 Connected Component Labeling

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

The basic pseudo algorithm for 8-neighbour connectivity is given in figure 2.8.

for(each foreground pixel (p_i)){

1. examine neighboring pixels which have already been labeled (i.e. top left, top, top right and left)

2. Based on the information above:

a)If all 4 neighboring pixels are 0 then give p_i a new label.

b)If only 1 neighbor (p_j) is a foreground pixel, assign p_i p_j's label.

c)If more than 1 neighbor is a foreground pixel, assign one of the labels to pi, make a note of the equivalences of the existing labels.

}

Figure 2.8 Basic Pseudo Algorithm for 8-Neighbour Connectivity

After completion of initial scan loop through all the pixels once again replacing equivalentlabels with one single label.

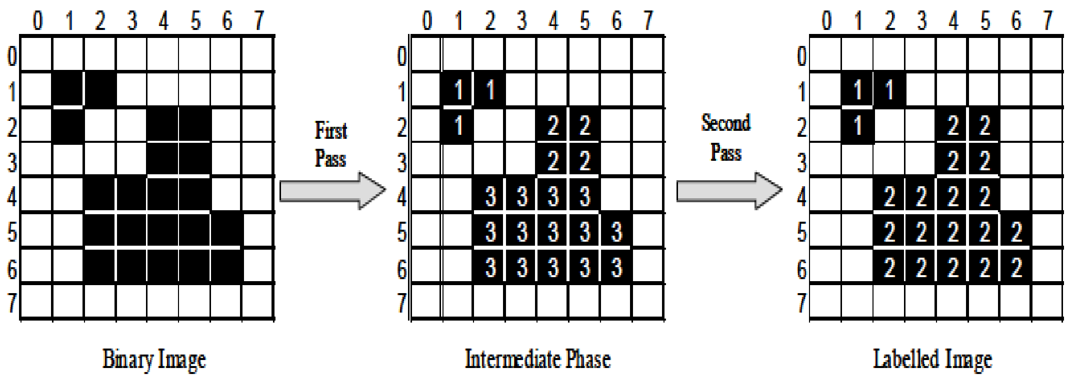


Figure 2.9 Example of Connected Component Labeling

A two pass scan can be seen. Initially a binary image is passed into the function, theimage is then scanned and each foreground pixel is given a label, note for pixels (4,3), (4,4)and (4,5) they have the choice of being assigned label 2 or label 3, therefore this equivalence(i.e. label2 == label3) must be noted. In the final pass all equivalent labels are replaced by oneunique label.

Once all regions have been labeled, it is then possible to remove regions which are smallerthan a certain threshold from the binary image as it is most likely these do not belong to themain foreground blobs.

2.5.2.2 Erosion

The basic effect of the erosion operator is to erode away at boundaries of foreground pixels,thus in effect shrinking foreground regions. This is an excellent way of removing small noisespots in an image.

It works by placing a kernel, a pre-defined search area shape, on top of each foregroundpixel. All foreground pixels which fall inside the kernel are counted and if the number of pixelsis less than a predefined threshold then the center foreground pixel is eroded and now isconsidered to be part of the background.

The 3x3 kernel is the most common structure used in erosion, but other kernels of varyingsizes, orientations and shapes can be used, depending on the nature of the image.

$$e(i,j) = \sum_{(m,n)\epsilon\Omega_{i,j}} f(m,n) \quad (2.15)$$

Where *f(m,n)* returns 0,1

Ω = all pixels belonging to kernel

$$f'(i,j) = \begin{cases} f(i,j), if\ e(i,j) \geq T \\ 0, otherwise \end{cases}$$

An example of erosion on a binary image using a square 3x3 kernel, and applying a thresholdof five foreground pixels can be seen in figure 2.10.

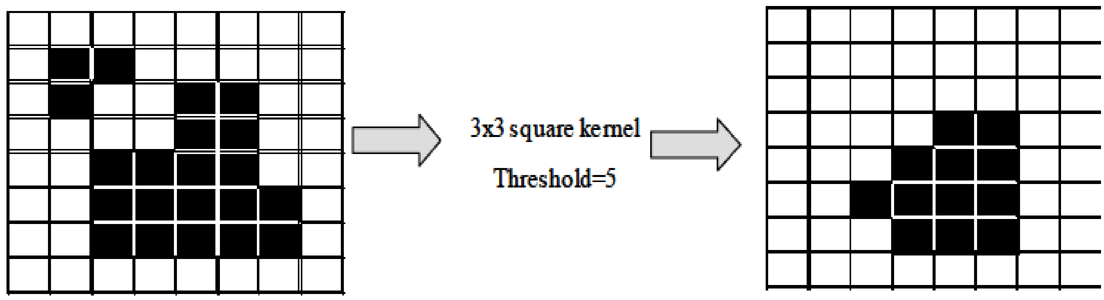


Figure 2.10 Example of Erosion Filter on Binary Image

2.5.2.3 Dilation

Dilation is the complementary operation to erosion. The effect of dilation is to graduallyenlarge the foreground region by a factor,eachiteration of dilation which is performed.

Dilation is useful for enlarging regions which may have been eroded by the erosion process.For each background pixel in the binary image, overlay the kernel (as described in the erosionsection), if any of the pixels inside the kernel are foreground pixels the make the center pixelpart of the foreground. It is also possible to add a threshold to the equation which allowsareas to be dilated only if the number of foreground pixels is greater than a certain threshold.

An example of dilation is shown in figure 2.11, using a 3x3 square kernel and 8-neighbourhoodconnectivity.

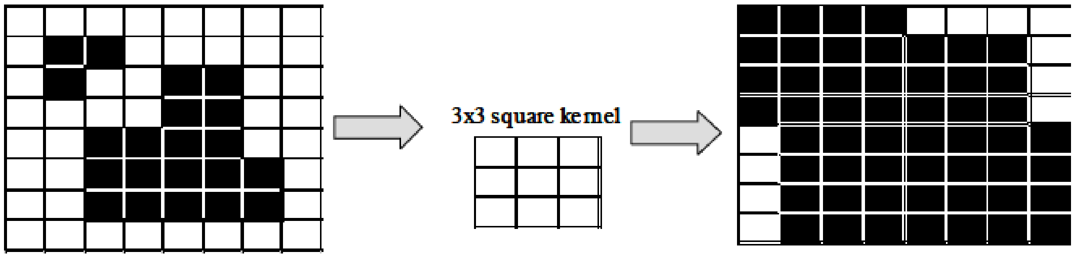


Figure 2.11Example of Dilation Filter on Binary Image

2.6 Self Similarity Plots

A self-similarity plot is a measure of how similar a measured parameter is over a period oftime, it measures the correlation of the variable over a number of recorded readings. Forexample, in relation to gait recognition the similarity plot may measure the similarity of theangle of incline of the thigh, or the similarity of extracted foreground regions over N frames.The latter was investigated by BenAdbelkader [18], they successfully used a self-similarity plot for gait recognition. The plot was constructed by comparing the similarity of thepixel intensities of segmented foreground region blobs. An example of one of the plotsproduced can be seen in figure 2.12.

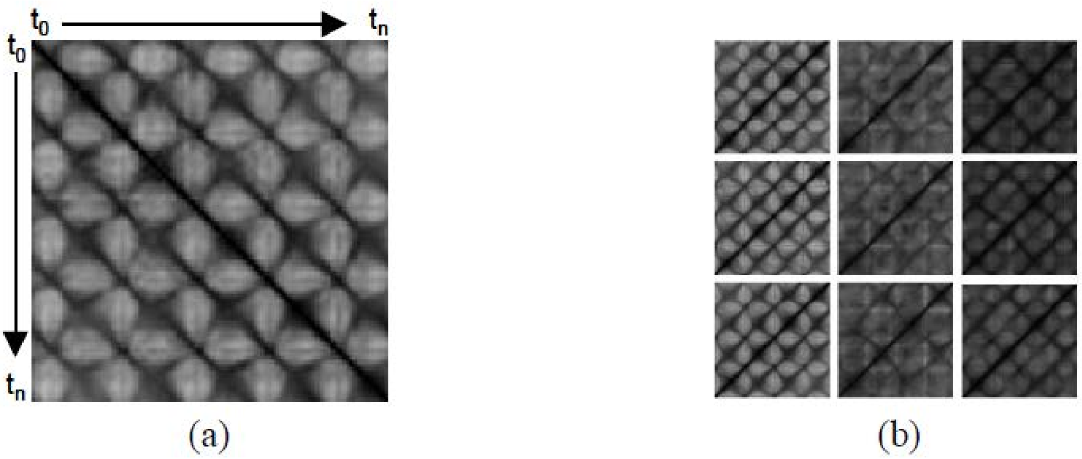


Figure 2.12 Examples of Self Similarity Plots

The dark areas signify high correlation, and the lighter areas signify a drop in the correlationof the measured parameter, each pixel represents one period in time. The plots in figure 14 arefor people walking, it is evident that the frequency and phase of the person walking isdisplayed in the plot.

Figure 14 (b) shows the plots produced from three different subjects. Each different subjectrepresents one column, and each row represents a different video sequence of the individual.It can be seen that the different people maintain a similar plot on different video sequences,and that the plots differ from person to person, giving evidence that they could be used as agait classifier.

Calculating the plot based on image intensities can be done as follows:

$$S(t_1, t_2) = \sum_{(x,y)\epsilon B_{t_1}} \left| O_{t_1}(x,y) - O_{t_2}(x,y) \right| \quad (2.16)$$

Where $O_t$are scaled image templates, representing the extracted object at time t and $B_t$represents the minimum bounding box surrounding the extracted object. In order to accountfor small errors we can introduce error tracking in a small radius around the target area:

$$\hat{S}(x,y) = \min_{|dx,dy|<r} \sum_{(x,y)\in B_{t_1}} \left| O_{t_1}(x+dx, y+dy) - O_{t_2}(x,y) \right| \quad (2.17)$$

Properties of self-similarity plots are; S(t,t) = 1, i.e. the main diagonal has high correlation and, $S(t_1,t_2) = S(t_2,t_1)$, i.e. it is symmetrical along the main axis.

Obviously self-similarity plots are not limited to being produced in just this way; they can becomputed from any parameter which can be measured over a period of time.

2.7 Relational Modeling

Modeling in computer vision is strongly reliant on the use of prior knowledge of the desiredscene in order to be able to build a model for feature extraction. Using models in the featureextraction process helps to improve the accuracy of the final results, as well as helping tonarrow down the search space during model fitting, due to the constraints placed upon themodel.

Relation modeling is used to make assumptions about the structure of objects in a scene,based upon the relations of the different objects which make up the whole scene. Thegranularity of the model can be at any level of detail, depending on the nature of the imagebeing analyzed and the requirements of the system. Relational models help vision systemsrelate to real world scenarios where perfect data extraction is not always possible.

Rules can be incorporated that help to reduce the objects extracted in the image, examples ofconstraints for a human relational model could be; If a circle (i.e. head) is not located above an ellipse (i.e. body) then remove it from theextracted feature list. If an ellipse is connected to a circle, but the length of its major axis is smaller thanthe circle, then remove the ellipse from the extracted feature list

Although relations may have the same name, such as “connected”, they may have differentimplied physical meanings. For example, head connected to body means that the headtouches the body but does not overlap, whereas leg connected to body, would mean the legdoes overlap the body, because the leg is connected to the hip joint.

One way to express a relational model is in the form of a semantic net. They offer a way ofexpressing the relational structures with a high level of generality. An example of a semanticnet at various levels of detail, representing the relation between different parts of the humanbody can be seen in figure 2.13.

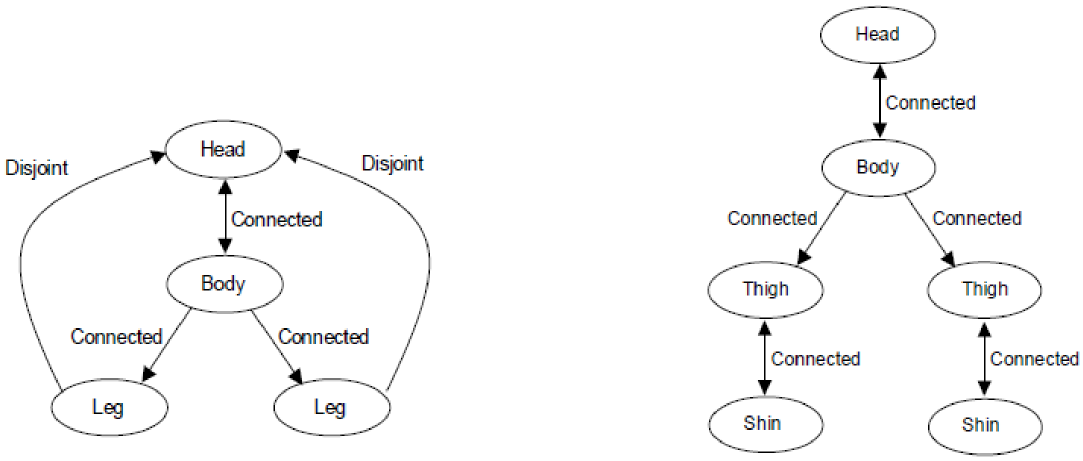


Figure 2.13 Examples of Relational Models

Semantic nets aim to be as general as possible, so that they can be applicable in manydifferent situations. An important point to note is that semantic nets only store the relationalstructure of an object; they don't store any size or location information about the individualcomponents, as this would severely limit their usefulness to extract objects from differentscenes.

One of the main advantages from using semantic nets is the ability to infer about a scene.Even if all the components of the object are not successfully matched in the image, but themajority is with the correct relation, then it is possible to infer that the object is indeed in thescene, but some parts have not been extracted, possibly due to occlusion or noise in theoriginal image. We assume that there are functions available for extracting the geometric oriconic shapes represented in the nodes.

2.8 Pattern Classification

Once the gait feature has been extracted from the person, it would be projected into a featurespace and it would then have to be classified. This means we have to determine which group inthe feature space (i.e. which person) the unknown feature point should belong to.

A classifier defines boundaries in a feature space which are used to separate different sampleclasses from each other in the data. The simplest type of classifier is a linear classifier.

This is a straight line which is defined in the feature space, points above the line are in oneclass, and points below the line are placed into another class. This is a simple method and doesnot provide the best results due to small non-linear fluctuations around the boundary regionwhich result in poor classification results. An improvement on this method is called the K-NearestNeighbor (KNN) classification rule.

A feature vector of unknown class can be classified as belonging to a class by using the K- NearestNeighbor rule. A training set of points (i.e. feature vectors projected in Eigen space)T are used to determine the classification of feature vector X, by the following method.

In step one calculate the K-nearest points to the unclassified feature vector X in the featurevector set T. There are a number of distance measures which can be used (asdescribed in section 2.11) to calculate the separation of two points in n-dimensionalspace.

In step two,determine the class which has the most points in the k selected points, from set TThe class which has the most points is chosen as the class which point X now will belong to.This is called a voting classifier, and clearly it is able to handle non-linear classificationspaces. An example of a linear classifier can be seen in figure 2.14 (a), it shows all the pointsbelow the line being classified as class ‘a’, and the points above the line as being in class ‘b’.

Figure 2.14 (b) shows the K-Nearest Neighbor approach to classification. The five closest training points to unknown point U have been chose. Out of the five points four are classifiedas belonging to class ‘a’, therefore by majority vote the unknown point U will be classified asbelonging to class ‘a’.

This voting clearly has the advantage of being more resilient to noise, with spurious trainingprojection points not dominating the classification process.

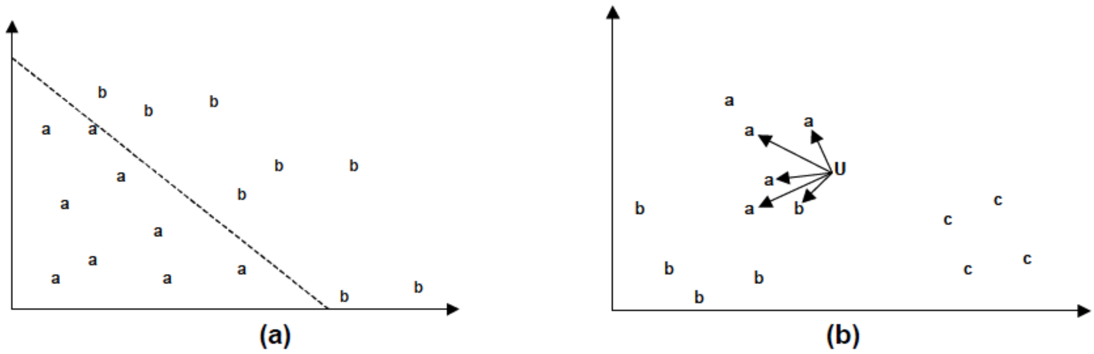


Figure 2.14 Pattern Classification

Chapter3

Conceptual Design andDevelopment Environment

This chapter featuresconceptual design of the gait recognition system,with debating the reasons for choosing a particular development environment and the designmethodologies.

3.1 Overall Program Architecture

3.1.1 Input to the Program

The system could take various inputs; image sequences, video frames or output directly from camera. To enhance the computational efficiency, the program was to accept a video sequence as input. However, it was decided that the video sequence would be preprocessed to further aid the system. All these preprocessing techniques are discussed in detail, in the order; segmentation, model fitting, recognition engine. Segmentationmodule takes the video sequence as aninput, then processes it to determine pixels are part ofthe foreground or background.Thebinary image produced from the segmentationprocess is used as input for the model fitting. This module fits a model of the human form onto the segmented area of the image. Oncethe model has been fitted to the image, features,which can be used to create a gait signature, are derived fromthe model i.e. variation in thigh angles over N frames.Therecognition engine takes data, either from a newly captured subject via the feature extraction module, or a previously stored signature, then performs recognition based on a database. Various different metric, are incorporated in the recognitionprocess to facilitate this procedure.

3.1.2 Outputfrom the Program

After the recognition is done and the necessary gait features are extracted from the input video it is then compared against the database and the closest match is found.A message is displayed on GUI showing the closest match from the available dataset.

3.2 Gait Recognition Design

This section deals with the design characteristics of the recognition engine as well as its methodology and input parameters. The actual algorithm and its implementation with its problems are discussed in later sections.

The method used is to classify different people's gait using PrincipleComponent on the extracted feature matrix of each person. It was chosen because it hasled to good results [18,19] and to explore the new avenues of using it withdifferent gait feature vectors.

The feature vectors used for gait recognition could be of two types; primary and secondary.Primary features are features that are directly extracted from the humanmodel, such as the angles of the limbs at each frame in the video sequence, or theposition of each joint over the video sequence.Secondary features are derived from primary features. In thisproject, self-similarity feature plots (mentioned earlier) are used to derive a gaitsignature based on the self-similarity of the angle of the limbs over the N frames of video.

3.3 Development Environment

This project was developed using MATLAB R2010a and was run on an Intel ® Core$^{(TM)}$2 Duo CPU 2.93 GHz, with 3GB RAM. MATLAB was chosen to implement, becauseof the speed of its simplicity and user friendly interface.

It was possible to develop the program using Java, butthis was not chosen because of its slower execution speed becauseJava is compiled into an intermediate form, Java Bytes rather than machine code directly, which hasto be interpreted at run time.

It was also possible to develop a program using C++, but because it limits the application to the Windowsenvironment it was dropped.

Chapter4

Image Segmentation

The main aim of image segmentation was to extract a good quality binary foreground image, from a video frame. This was required to perform modelfitting on the image, which is later used to extract the gait featuresfrom the video sequence.

4.1 Assumptions for Image Segmentation

The main assumption made for image segmentation, was that the background of the videosequence was static i.e. the only moving object is the person itself. Had it not been the case then changes in pixel intensities would be incorrectly classified as foreground regions, thisassumption was made because a non-static background is a non-trivial problem.

Other than having a static background it wasalso assumed that all external conditions, such aslighting remain constant, otherwise this can affect results.The other assumption was static camera position. This was assumed to be able to perform simple frame subtraction. Had the camerabeen moving, the image subtraction would not have worked, since the temporalcorrespondence between equivalent pixels would have been lost.

4.2 Segmentation Design

The segmentation module performs a number image pre-processingoperations, such as; erosion, background generation, median calculation and connected component labeling.

Erosion reduces the noise introduced during background subtraction. Background generationis the ability to generate a background from a sequence ofvideo images. Median calculationcalculates the median positions of the person walking inthe video sequence. Connected Component Labelingis the ability to group pixels togetherunder a single label.

4.3 Background Generation

The data set used for this project [21], does not have any images in the video sequencewithout the human subject. Additional functionality was added to theSegmentation which automatically generates a static background image from asequence of moving images.

This was done by generating a model of the pixel intensities over 50 frames.The intensity of each pixel was stored i.e.

$$P(x,y)_{T_{1-N}} = [P(x,y)_1, P(x,y)_2, \dots, P(x,y)_N]$$

Once these all were recorded aQuicksort was performed on all the pixel values to bring them to order and then the median value wastaken as the value of the background pixel. An example output from the program is shownin figure 4.1.

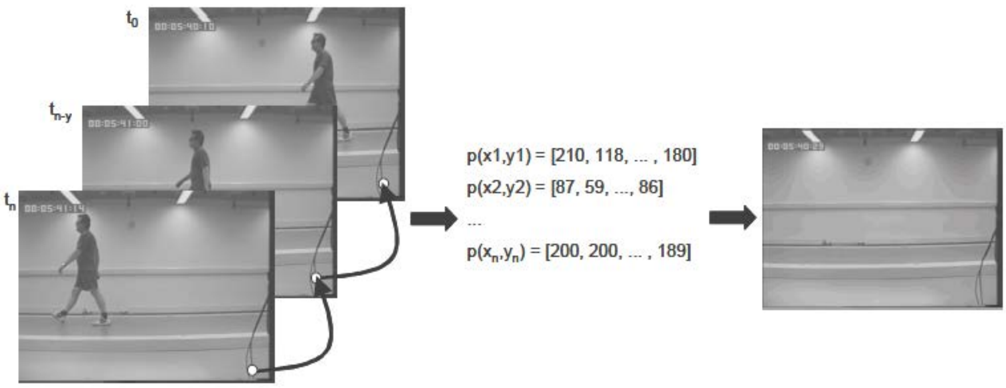


Figure 4.1 Background Frame Generation

A median value was taken rather than the mean value, because a mean value is distorted by the large change in pixel intensities when theperson moves past that pixel. The median value is unaffected, as long as large a number offrames were recorded.

The main assumption made in this module was that the person does not stand at any point over the frameswhich wereanalyzed, had this been the case the background generation would have classified the personas part of the background.

4.4 Foreground Segmentation

It is necessary to remove as much noise and distortion as possible from the segmentedforeground image, in order to reduce the chances of misclassification of body parts during model fitting. Main segmentation algorithm is based on a method used by [20], which had successfully extracted foreground regions from a video.The process is performed in a two pass approach. First pass generates a static background from the images, and performsglobal background subtraction, filtering and small blob removal to locate the smallestbounding region encapsulating the person. Second pass is similar to pass one, but the filtering takes places in a smaller localized regionaround the person. The subtraction; erosion is performed at even lower thresholds toproduce a binary image, which is only made possible by using a smaller searchregion.

The individual steps in the algorithm can be seen in figure 4.2.

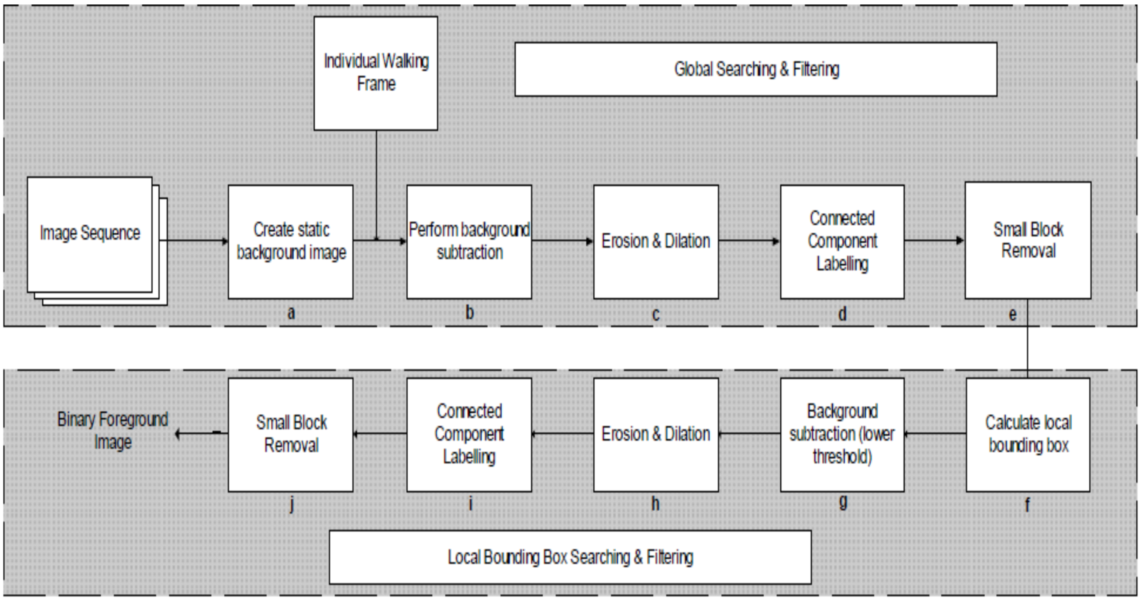


Figure 4.2 Individual Steps in Algorithm

To perform background subtraction for extracting the foreground regions, a static background image is must. Background subtraction between the current image and background imageprovides a foreground image, but this has a large amount of salt-and-pepper noise and was not suitable for model fitting.

Erosion was then performed on the subtracted foreground image to remove salt-and-pepper noisewhich wasintroduced to the binary image during segmentation. It was crucial to perform a pairederosion-dilation otherwise this couldenlarge smaller noiseregions. Connected componentlabelingtags each group on the foreground image with a unique label,a table was then produced of all the labels the correspondingnumber of pixels having these label, the list was arranged in the descending order of group size. Stages e and j in figure 4.3removedlabeledgroupsfrom the foreground region on threemain criteria; relative and absolute size and the distance of the groupfrom the centroidof the main group. Physical size of the groupwas also compared to othergroupsin the same image.

```
if(listPosition(labelx) > listThreshold || blobSize(labelx) < blobSizeThreshold ||
distanceFromCentroid(labelx) > distanceThreshold ){
remove all pixels with labelx
}
where:
listPosition(x) – gives the position of the blob in the list of ordered blobs
blobSize(x) – returns the number of pixels associated with label x
```

Figure 4.3Basic Pseudo Algorithm for Foreground Segmentation

It was unrealistic to define a threshold for the position of a label in the ordered table. This was because sometimeslabels appearing at the bottom of the table represented quite large parts of human foreground region. Hence, the blobSizeThreshold and distance wereused as a criterion for inclusion of a group in the foreground region. ThisblobSizeThreshold value was dependent on the quality of the images and theamount of noise present in the data capture. For the gait dataset consisting of frames of dimension 320x160 pixels, a threshold of 10 pixels was found to be a satisfactory cut off point.

Figure 4.4demonstrates this method for removal of small group and helps to tightenthe bounding box enclosing the subject.

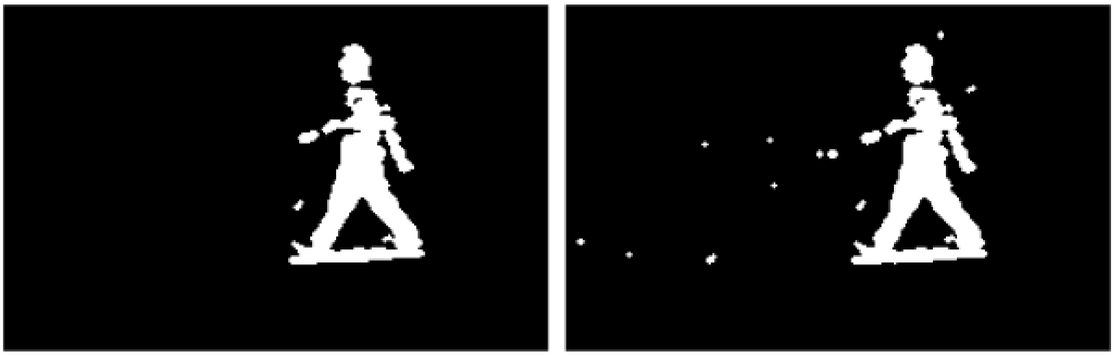


Figure 4.4 Blob Size Thresholding

4.5 Gait Data Set

The gait data set used for this project was obtained fromBoyd and Little [21], it has been used as the data set for many previous gaitexperiments [19,18]. The gait laboratory setup for this data can be seen in the figure 4.5.

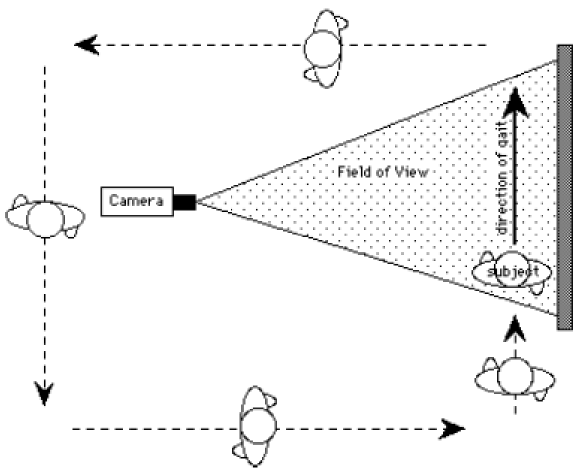


Figure 4.5 Courtesy of Boyd and Little

A static camera was positioned, facing towards a static background. Each gait subject walked in front of the camera several times, during a 15 minute time period. The video sequences were initially digitized in 24-bit color, at a resolution of640x480 pixels. They were then cropped to a resolution of 320x160 and converted to greyscale images.It is important to note that these video were of very high quality and this helped much in segmentation.

Chapter 5

Model Fitting

Model fitting matches the model of the human to the video. When this is done, gait signatures can be extracted from the fitted model for using in the recognition process. This chapter describes the methods used for trackingvarious parts of the body; head, torso and legs.

5.1 Assumptions and Limitations

There were a number of assumptions which were made for the methods to work; only one person was present in the video footage at any time. The camera lens was always perpendicular to the person's gait. This implies that the length of the parts of the body can be assumed to remain constant throughout thevideo; otherwise if the gait was viewed at any angle, then the parts of body would appear to change size depending on the perspective of view.The camera must always be in the same position for thewhole video capture.

5.2 Human Model Design

There are two main parts for the model used for recognition, which helps thetracking, the physical representation of the model and the semantic representation.

5.2.1 Human Model Representation

There could be a number of ways in which a human bodycan be modeled. In this project a blob representation was chosen; to allow correlation of model parts against thesegmented binary image. An example of the model used formodel fitting is depicted in figure 5.1.

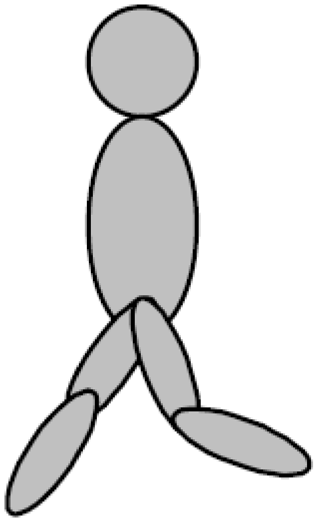


Figure 5.1 Six Segment Human Model

A number of kinematical constraints can be impliedfrom the structure of the human body which enables to further reduce the possible search space; body part dimensions, constant velocity, joint dynamics.

The size of the all the body parts remainsthe same throughout the data capturing process. This was a fair assumption because of thesimplifications made earlier in the data capture process, namely that all thesubjects walk parallel to the camera. It could also be assumed that the subjects in the video walk at almost aconstant velocity; this velocitycan be estimated and used fornarrowing down the search space.Other constraints that could be madeare that the lower legs are limited in their freedom by the knee joint. It is impossible for the lower section of limb of lower limb to bend past 180 degrees with the upper section of the leg. This constraint was very helpful in reducing the search space as well asproviding a more accurate model fitting result

5.2.2 Human Relational Model

As mentioned earlier in the background, relational modeling is quite a useful tool for modeling real worldstructures, and also for narrowing down on the possible search spaces with the help of the constraints placed on the model. A relational model was used in this modeling which uses thesemantic netas depicted in figure 5.2.

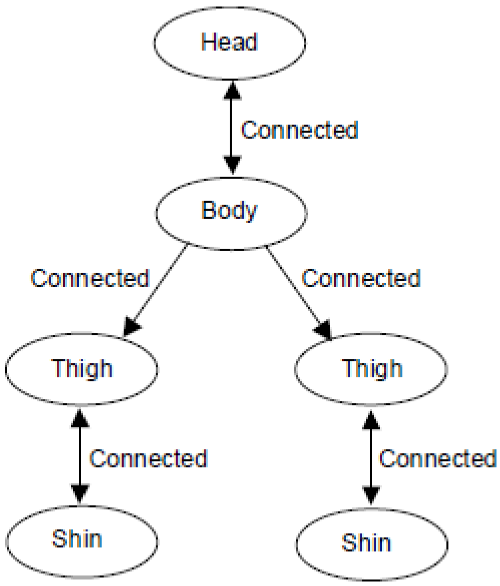


Figure 5.2 Semantic Net of Human Form

The relations between the various body parts have these meanings; Head connected-to Torso;the top of the body is touching the bottom of the circlerepresenting the head, they do not overlap at all. Thigh connected-to Torso**;** the thigh overlaps the torso a little, but the top of thethigh is positioned in the middle of the torso. Shin connected-to Thigh; the shin slightly overlaps with the bottom of the thigh, as both are interconnectedtogether at the knee joint.

5.3 Search Space Estimation

After segmentation, the program returns a bounding boxtermed as search space inside, which the person is located. In order to speed up theexecution time, the search space size is to be further reduced, because the larger thesearch space the more it has to be processed. The search spacewas split into four distinct areas, as shown in figure 5.3.

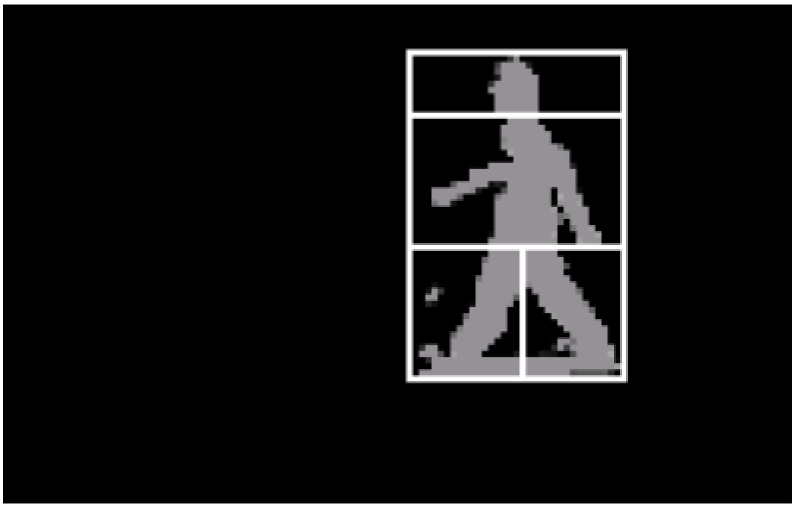


Figure 5.3 Search Space Reduction

The four main areas are head, torso, front leg and back leg.

Head is the top 25% of the bounding box, Torso is the middle section of the bounding box, Front leg is the bottom 50% of box, to the left hand side of the box while Back leg, is the bottom 50% of box, on the right hand side of the box.

5.4 Body Dimension Estimation

Another important point about the search space is that all of the dimensions of the model werederived inproportionto the height of the search space. To ensure that they are notmisjudged due to a search space being large, probably due to noise, the height of thesearch space was recorded for N frames, and then the median height of thesearch space was taken tobe the fixed height of the search space. The median height was used rather than the meanheight, because the median height is not as affected by spurious values as the mean value is.

5.5 Head Tracking

From segmentation, an initial search space which is fairlyconfident in enclosing the moving subject s obtained. As shown in figure 24 the search space for head can be cut down to a sub-region of the bounding box. To locate theposition of the head first calculate the amount of coverage of the head, overlapping withforeground pixels. The region with the best coverage for the circle would be allotted theposition of head. The size of the head can be estimated from the search space height.To correlate between the circle, which represents the head and theforeground pixels, the following formula wasapplied:

$$E(x_c, x_y, r) = \sum_{x,y \in Bhead} e(x_c, x_y, r, x, y) \tag{5.1}$$

Where $e(x_c, x_y, r, x, y) = \begin{cases} 1, if\ (x - x_c)^2 + (y - y_c)^2 \ \leq r \\ 0, otherwise \end{cases}$

Noting that x,y are foreground pixels inside the head search space. Pixelsinside the reduced head search space were designated as being x_c,$y_c$and the best match position was recordedas the position of head.

5.6 Torso Tracking

Correlation of the torso uses the same type offormula as (5.1) apart from the fact that now circle is replaced by an ellipse orientated aroundthe y-axis:

$$E(x_c, x_y, a, b) = \sum_{x,y \in Btorso} e(x_c, x_y, a, b, x, y) \tag{5.2}$$

Where $e(x_c, x_y, a, b, x, y) = \begin{cases} 1, if\ (\frac{x-x_c}{a})^2 + (\frac{y-y_c}{b})^2 \ \leq 1 \\ 0, otherwise \end{cases}$

5.7 Head Body Location

Using the semantic net it wasobservedthat the body was connected to the bottom of the head. Therefore, the head and body search for the bestpositions were not independent rather they work hand in hand and find the bestcombined coverage with the body located directly below the head. The body wasallowed to movealong the x axis for a small range on either side of the center point of the head in order to find thebest position.

The general pseudo code for locating the position of the head and the body is shown in figure 5.4.

```
for each possible location of the head
calculate the coverage of the head
for each possible constraint body position
calculate body coverage
if body coverage + head coverage > best_combined_total
update best coverage
make not of the location of the head and the body
endif
endfor
endfor
draw the head and body at the position where the best combined coverage was
obtained
```

Figure 5.4 General Pseudo Code for Locating the Position of the Head and the Body

5.8 Legs Tracking

The legs were by far the most difficult part of the model to trackdue to non-linear motion and self-occlusion problems. This section aims to explain how the legs were modeled.

There were three main constraints placed on the movement of the legs; The value of β(the angle of incline of the shine) was always less than or equalto α(the angle of incline of the thigh). The angle of $\alpha$was always less than ±50 degrees. This stops the thigh frombeing in a position which was obviously not possible due to noise orincorrect tracking parameters. The value of 50 degrees was determined afterconcluding the maximum amplitude of the leg expansion on the various subjects.

The angle of incline of the front thigh must exactly be opposite to the angle ofincline of the back thigh, i.e. angleThigh1 = - angleThigh2. This forces the legs tomove harmonically and forward propulsion for the subject.

Figure 5.5shows a snapshot of the leg with angles alpha and beta. Zero degrees isparallel with the y-axis.

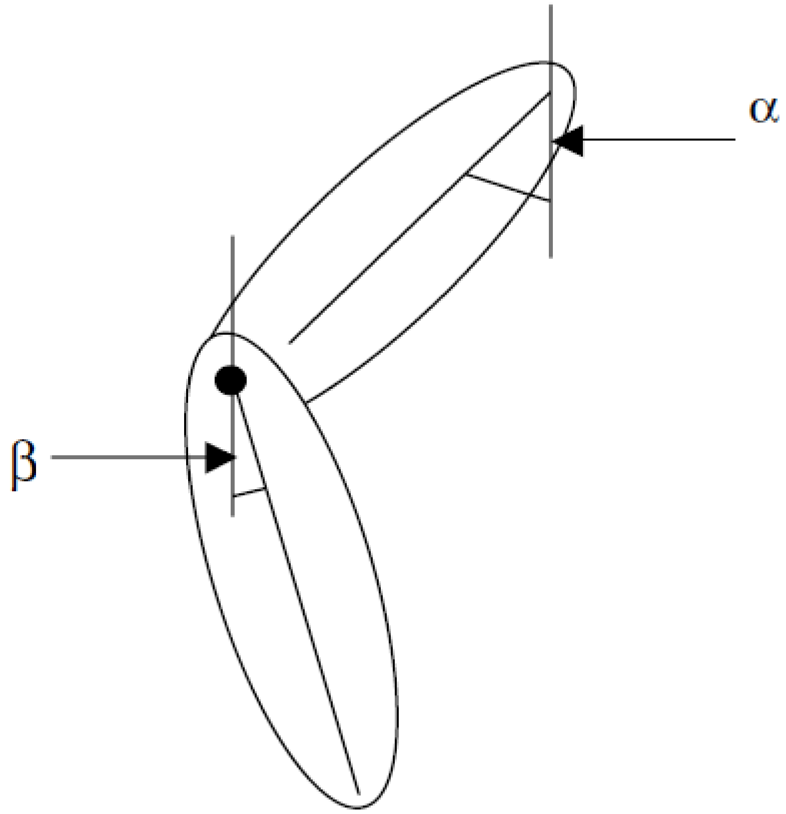


Figure 5.5 Snapshot of Leg with Angles

5.8.1 Initial Pose Location

After the initialization period, where the dimensions of thesearch space were calculated, need to find the best pose for the leg was needed to be found, but it required some starting point. This was done using the moments (5.3) to find the angle of incline of thelongest edge of the bounding box around the leg in the either legs search spaces. The formula for finding the angle is given in (5.4):

$$\mu_{p,q} = \sum_x^N \sum_y^M x^p y^q f(x,y) \qquad (5.3)$$

$$\theta = \frac{1}{2} tan^{-1}(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}) \quad (5.4)$$

This angle found the first and second thigh and then the angle of the shins were initially given the same angle as the thigh, so the legs are initially fully stretched. It can thenbe applied to advanced searching for refine the position of the legs into the best fit point.

5.8.2 Leg Position Tracking

For trackingthe position of each part of the leg, consider how to track one leg andthen generalize this for two legs.

The first step was to calculate the coverage of the foreground pixels by the model of the legs atconsecutive time t and t-1. The legs can be seen as two sections (refer to figure 24).Therefore, twodegrees of freedom were present in the legS; angles alpha and beta. A function can be defined which wouldcalculate the coverage of one leg:

$$E(\alpha, \beta) \quad (5.5)$$

Taking partial derivative of the equation (5.5) would give the rate of change of coverageof the leg due to changes in angles.

$$\frac{dE}{d\alpha} , \frac{dE}{d\beta} \quad (5.6)$$

Gradient of the function can be calculate from (5.6)

$$\nabla (\frac{dE}{d\alpha} , \frac{dE}{d\beta}) \quad (5.7)$$

From these gradients values of the alpha and beta attime t+1 can be found:

$$\alpha_{t+1} = \alpha_t + \frac{dE}{d\alpha}\delta \quad (5.8)$$

$$\beta_{t+1} = \beta_t + \frac{dE}{d\beta}\delta \quad (5.9)$$

Iterating these steps and changing the values of the angles using formula(5.9) tillabsolute maxima of the function E(α,β) is obtained.At absolute maxima, the leg has movedinto the best position and that point onwards it would start to decreasing the coverage.It can then be applied for both legs, and the model legs should stay inside the foreground regions as the legs move.

The total coverage of the legs was taken into account when finding the best positionfor the legs, combining the coverage values for the front and back, shin and thighs. When thishas a maximum value, it was the best place for the legs to match the foregroundregion.

Chapter 6

Gait Signature Extraction

Once the feature points are extracted from the video, a gait signature has to be extracted which coulduniquely identify, or at leastprovide a high recognitionrate for the subject. This chapter describes in detailthe methods used to produce a gait feature vector for the different subjects.

6.1 Signature Design and Implementation

The data used for creating a gait signature was the joint position in thehuman model, over Nvideo frames. Vector P contains all these instantaneous points, such that:

$P = [p_1, p_2, \dots, p_n]$ Where $p_i = [x_i, y_i]$ (6.1)

Feature extraction returns a vector of human poses which could be used to create the gait signature. The more eminent of these features uses acombination of temporal and spatial information; absolute joint positions, limb angles and angular velocities and self-similarity plots.

Normalizedjoint position in eachframe can form basis for a gait signature. Limb angles over N framescan also be used. Also, the limbs angular velocity can be used to extract thesignature. Self-similarity plotsas described earlier in section 2.7, can also beused to correlate between a gait variable over a period of time with itself.Self-similarity plots offer a good way to express the phase and frequency ofthe subjects'gait.

The rest of this section explains each of these processes in greater detail.

6.1.1 Joint Position Signature

A possible gait signature for an individual is the joints position,over the video sequence. Over the time the position pattern may be uniqueenough to act as a gait signature. The humanSequence data type stores all thejoint positions over the video. It function is to returns a feature matrix, whose columnsare equivalent to the joint positions in one frame. Therefore in a feature matrix, thenumber of columns is equalto the number of captured frames.

$$\begin{matrix} p_{1x}(t) & p_{1x}(t+1) & p_{1x}(t+1) \\ p_{1y}(t) & p_{1y}(t+1) & p_{1y}(t+1) \\ p_{2x}(t) & p_{2x}(t+1) & p_{2x}(t+1) \\ p_{2y}(t) & p_{2y}(t+1) & p_{2y}(t+1) \\ \vdots & \vdots & \vdots \\ p_{8y}(t) & p_{8y}(t+1) & p_{8y}(t+1) \end{matrix}$$

Figure 6.1 Layout of Joint Position Feature Matrix

Producing a feature matrix has the advantage that each individual person can be plotted in Eigen space as a cluster of points, allowing more advanced cluster recognition techniques toanalyze them.

6.1.2 Limb Angle / Angular Velocity Signature

The HumanSequencedata type has the functionto extract a matrix from thevideo consisting of either the limb angles or the angular velocities overN captured frames. As with the position of joints method earlier described, the output here is a matrix too, in which columns represents the limb angles in a single frame, and rowelement are the angles for one of the limbs.

To compute these values the limb end points are to be known, as shown in figure6.2.

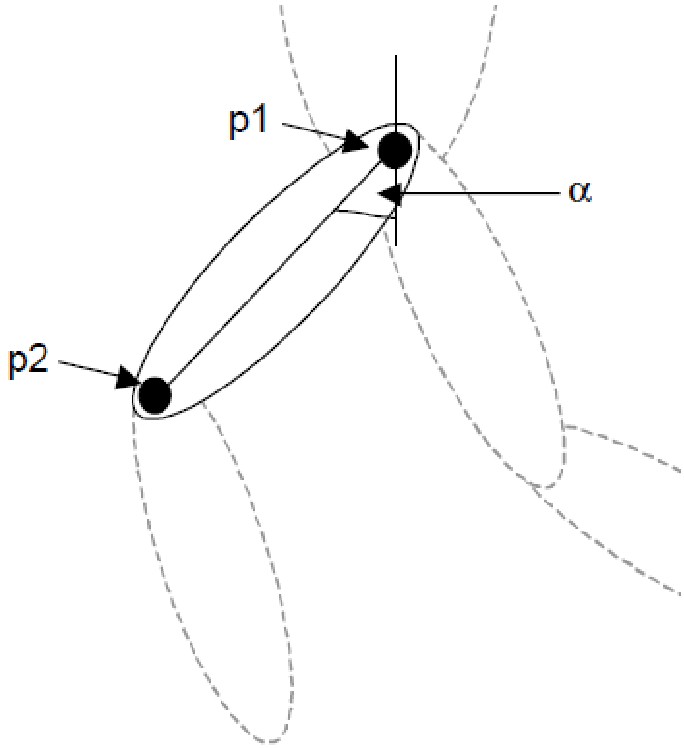


Figure 6.2 Limb Angle Calculation

$$\alpha = \tan^{-1}(\frac{p2_x - p1_x}{p2_y - p1_y})$$

Limb anglesare used to expressthe rate of change in each limb angle.An angular velocitiesmatrix for each joint isproduced in the same fashion; the slight difference is to take into account the timedifference between frames to calculate the rate of change of angles.

6.1.3 Self Similarity Plots

Self-similarity plots are a good way to extract the phase and frequency of the gait of a subject. BenAdbendulker used them successfully for recognizing gait[18]. He created a self-similarity plot based on the pixel image of the foreground. A number of self-similarity plots were produced for this project as well, based on the angles ofvarious body parts over N frames. Five self-similarity plots were created for eachsubject, for; Left Shin, Right Shin, Left Thigh, Right Thigh.

Initially a self-similarity plot for the torso was also produced, but small variations in the torsoresulted in to large amount of error in the plots, making it not very useful for gait identification.

6.1.3.1 Data Normalization

Data has to benormalized into a formwhich could allow non-bias comparison, before generating the self-similarity plots. For the videos, all were edited so they start with each subject's legs at a maximum distance apart. Thisensured that these plots were in phase,and hence the recognition results were saved from being false negative.

6.1.3.2 Creating Self Similarity Plot

The self-similarity plot for limb $J_i$was calculated using the formula (6.2):

$$S(J_i) = \min_{-d\alpha \leq r \leq d\alpha} \left| J_i(t_1) + r - J_i(t_2) \right| \tag{6.2}$$

The formula (6.2)was carried over $t_1$to $t_n$frames, which measures correlation of the desired features, such as, $J_i$could be thelimb angle i, or the limb angular velocity. The value r is an error minimizing term which was used to reduce theamount of error in the plot. The value of d$\alpha$was fixed to 5 degrees based upon experimentation.

6.1.3.3 Creating the Feature Vector Representation

Once the self-similarity matrix was produced can reduce the amount of informationneeded can be reduced because symmetry is a feature of self-similarity plots.Since thematrix is symmetrical so can discard half of the matrix can be discarded as redundant data, principal diagonal entries could also be left because it is a diagonal matrix i.e. S(t,t) = 1, meaning thereby that an item is correlated with itself.

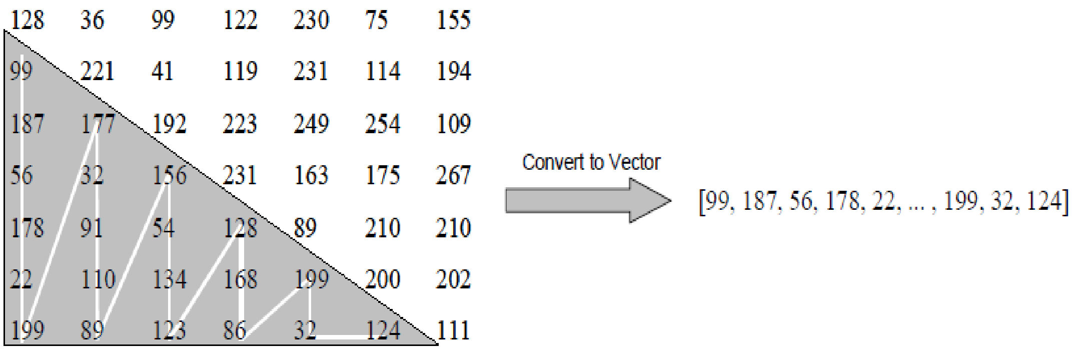


Figure 6.3 Converting Matrix to Vector

So instead of having a vector length n^2items, it now has length:

$$length = \frac{n(n+1)}{2} - n \qquad (6.3)$$

This reduced the computational complexity for calculation of the Eigenvectors, at the same time preserving the gait information of the subject.Finally, these feature vectors which were produced earlier for different body parts were concatenated to form one large vector which formed the gaitsignature for each subject. Five self-similarity plots which were created fora subject aredepicted in figure 6.4.

Figure 6.4 Example of Self Similarity Plot

Chapter 7

Gait Signature Recognition

Once the gait signature has been created from the individuals, it further neededto be compared and recognize between the different subjects. After this, classification would be usedon the gait data sets to perform recognition. This project implemented twomain classification methods and a number of feature transformation operators.

7.1Gait Signature Recognition Design

The whole recognition process wascondensed within a main module.This module provided the functionalities of; performing PCA, Canonical Analysis, projecting feature vectors to a different space, K-Nearest Neighbor classification, recognition based upon Spatio Temporal Correlation and its analysis.

7.2 Gait Signature Feature Spaces

Comparing the extracted feature vectors in their raw extracted form did not lead to appreciable recognition rates; because different subject classes' were overlapping, they were poorly separated andalso the data had a high dimensionality.To achieve higherrecognition rates feature vectors needed to be projectedinto a different featurespace which couldensure maximum separation. As mentioned earlier in thebackground, two main ways exist for reducing the data required for comparisonor separating classes of data, be itEigen feature space or Canonical feature space.

Once feature vectors are created each vector is projected intoEigen space to aid recognition rate. This was done by first computing theeigenvectors from the covariance matrix of the extracted feature points. Allpoints were concatenated to form a vector and then all instances were merged to producea feature matrix, as depicted in figure 27.

To project feature matrix into Eigen spaceit was multiplied by eigenvectors of equation 2.9. Hence, the PCA could preserve most of the data variation, but needed lesser information to be stored.

7.3 Training Data Set

For performing recognition a set of data is required to which the query canbe compared against and classified. Spatio temporal correlation method (see section7.4.1) the training data was created byrandomly choosing one sequence of each of the different subjects andstoring their details in the database. Queries were then compared against this database.With K-Nearest Neighbor classification (see section 7.4.2) the training database set wascreated by placing all of the extracted gait signatures of all the video sequences away froma sequence. The left over sequence was then compared against thisscatter plot. This was done for all the video sequences and is known as the leave-one-out cross validationmethod.

7.4 Recognition through Classification

Once the feature vectorsare extracted from thesubject, and projected into a feature space of choice; which gives the best of recognition, the nextstep was to recognize them by pattern classification. Two different methods ofclassification are used in the project, depending upon the method of gait feature extraction.First is spatio temporal correlation and the second is K-nearest neighbor with leave-one outcross-validation to estimate the rate of error.

7.4.1 Spatio Temporal Correlation

As mentioned earlier in the chapter, two different types of signature were producedfor a subject, a cluster or a point. For spatio temporal correlation cluster comparison was used to ascertain to which training class the query should belong.

The idea being that each corresponding temporal point of the classwas created from the same number frame in the video sequence was compared again and the distancecomputed. So the point projection of x^{th} subject in i^{th}framewas compared against the projectionpoint of y^{th} subject in i^{th}frame. The total distance of all these points was calculated:

$$d_i = \sum_{t=1}^{N_i} |c_i(t) - q(t)| \tag{7.1}$$

As can be seen in figure 31, each element of the class cluster one is being compared with the corresponding temporal instances of other class, and the distance is computed. Thetotal distance between all corresponding elements was summed and a measure of thedistance of the two classes was calculated. In order to recognize, the query wascompared against all of these training classes and the separation distance based upon spatio-temporal (as in figure 7.1) was calculated. The training class which has the smallest distancefrom the query was chosen to be that subject.

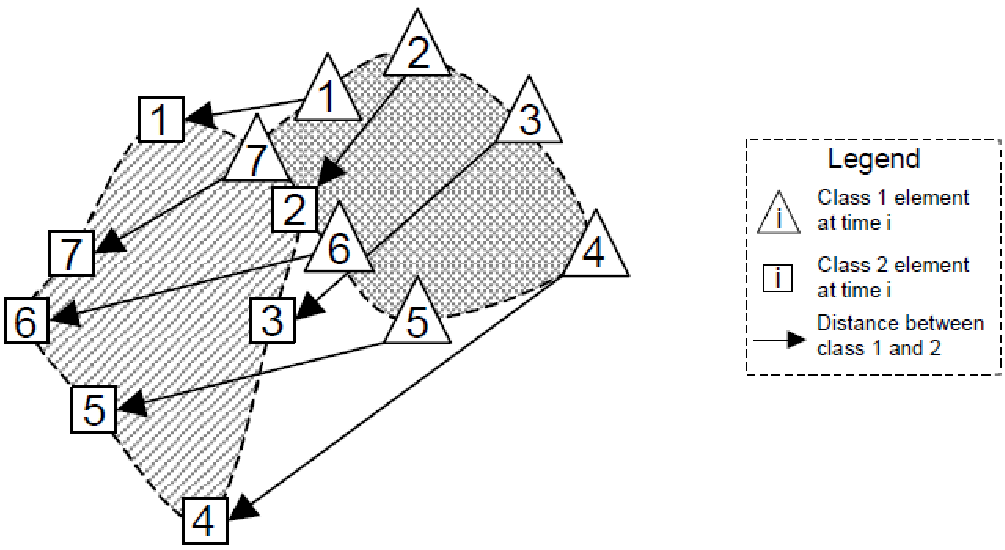


Figure 7.1 Spatio Temporal Correlation

Once all of the training classes have been compared against the query the class having greatest distance wasdeclared asthe match.

This method of classification was used for recognition for Positions of Joints using PCA and Angles of Limbsusing PCA.

7.4.2 K-Nearest Neighbor

Another classifying method for classes is the K-Nearest Neighbormethod. This was used when single queryfeature point was compared against the data set.This classification method was used to recognizeSelf-Similarity Plot using PCA feature extraction and recognition.

Gait recognition by the K-Nearest Neighborhad the following steps (see figure 7.2):

In step a query point was compared against all training points (all points had already been projected into feature space). In step two, an ordered list of class elements and their distance from the querypoint was generated. In step three, the top k elements from the ordered classification list were extracted. In step four, the number of occurrences of each person in the kextracted points was counted and the class with the most occurrences was declared as the classwhich the unknown query belonged. This is known as theMajority Vote Method, and it’s quite resilient to the training set noise.

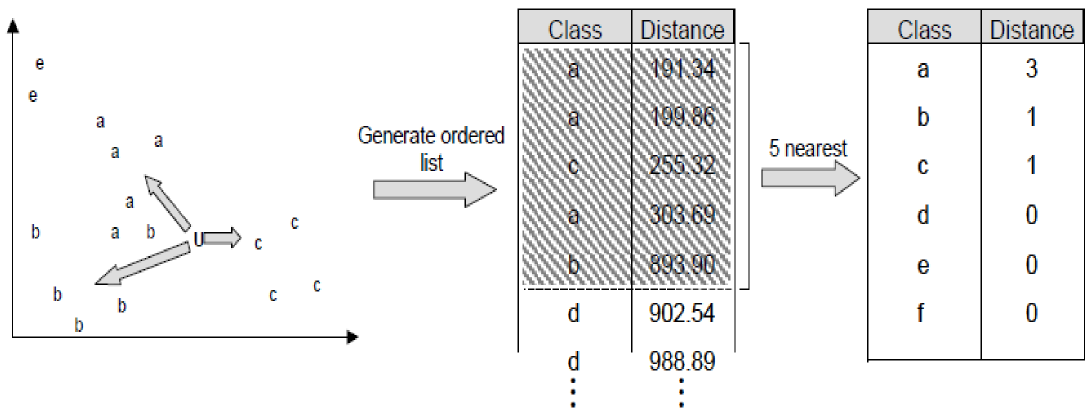


Figure 7.2 K-Nearest Neighbor Classification

Chapter 8

Evaluation

This chapter gives an overview of all of the main sections, segmentation, model fitting and recognition and explains the results of each section, and then ageneral evaluation of the whole process is given.

8.1 Segmentation Results

Results of segmentation could be divided into two parts, the foreground extraction,where the person was extracted and the backgroundgeneration, which was required for image subtraction.

8.1.1 Foreground Extraction

Segmentation process described earlier provided good segmentation results, with a number of people and videos.Figure 8.1 shows the binary image which was produced from the original image, it can be seen there are still small blobs of noise present, but these did not cause anynegative effects as they were relatively small.

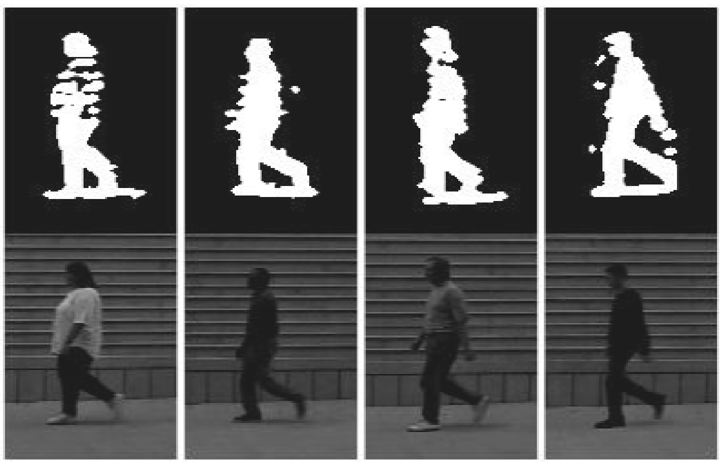


Figure 8.1Binary Image Map which is produced from the Original Image

One of the main problems was the shadow cast by the walking person. Due to the change in pixel intensities it causes, it was incorrectly classified as a foreground region, whichaffected the width of the minimum bounding box around the person. Since the imageswere greyscale, the removal of shadow was much more difficult that if the images were colored,further extension wereadded to this process to perform pre-processing on theimage to remove shadow.

Another problem was clothing; some of the subjects were wearing clothes that were close in intensityto the background. It resulted in part of person being classified as thebackground. An example of this can be seen in the first column of figure 8.1 and infigure 8.2.

To reduce this effect threshold used for subtraction has to be lowered butthis makes salt-an-pepper noise more pronounced. Hence a tradeoff between the two effects was accepted. Figure 8.2depictsthe segmentation process results with different thresholds. All the images are of same person, the first row is using a threshold oftwenty, the second row is a threshold value of ten and the last uses a threshold value five.

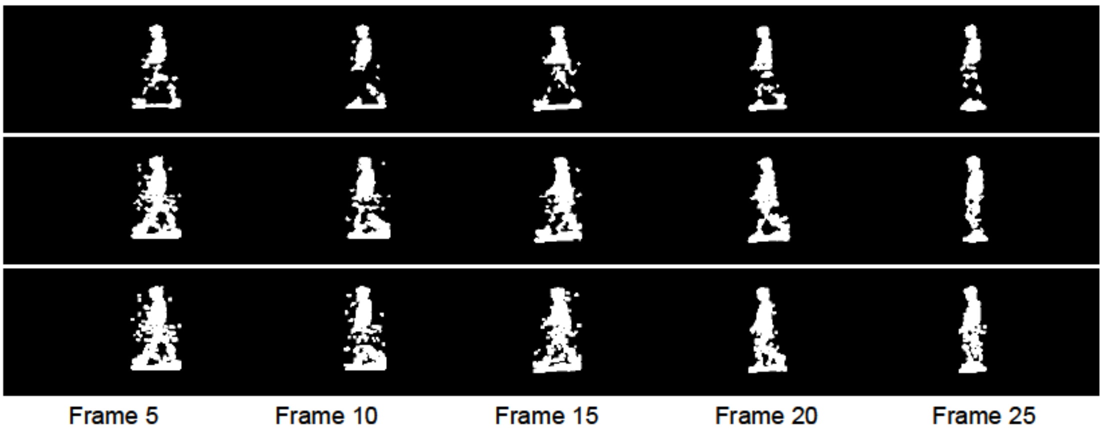


Figure8.2Segmentation at Various Threshold Levels

It can be seen the higher the value of the threshold,the more the subject is classified as background instead of the foreground. Lowering thethresholding value reduces this effect but it introduces more noise to image.

Another example of noise was in the initial dataset we had used. The images hadbeen taken and the lighting was a florescent light, the flickering of this light caused largeamounts of noise to be produced using this subtraction technique and the background model being used. This is depicted in figure 8.3.

Figure8.3Noise in Binary Image

A better background model can be used for this data set with temporal and range distance included in the model to help reduce the effect of the lighting noise.

8.1.2 Background Generation

The background was generated from a number of frames of the chosen video sequence. Intensity of each pixel was recorded over a period of N frames and then the median valuewas taken as the background intensity.

This method had good results, with a good background being producedfrom each video sequence used. The only obstacle was choosing therequired number of frames to make up the model. As it is depicted in figure 36,a video sequences differing in lengths were used to generate the background. Theshorter the length of the video sequence, the more the subject was classified as a part of background.

Also, the speed of walking of the person affected how many frames were to be used, slower thesubject walked, more frames were required to remove the person from the background.In figure 36it is evident that for frames less than 30, some of the subject is still visible in the background (as circled in 20 and 30 frames snapshots). This is due to insufficient readings taken to differentiate the foreground and background using the median calculation. After a number of experiments, 50 frames were found to be a reasonable value for background generation.

Figure 8.4 Background Frame Generation

8.3 Gait Signature Extraction

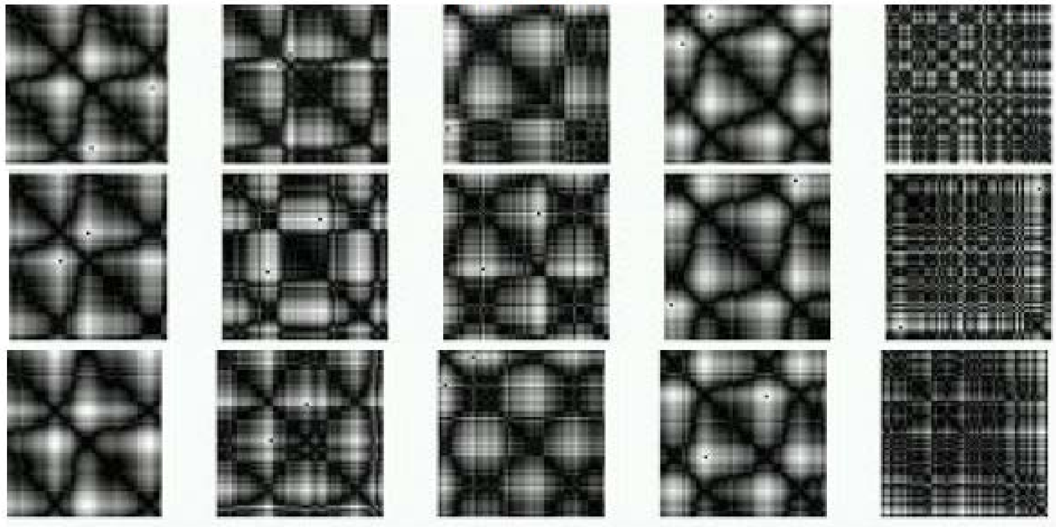
Gait signature extraction extracts a signature which is unique to a person's gait. A number of signatures were proposed, and this section exemplifies onthe outputs generated from them.

8.3.1 Self-Similarity Plots

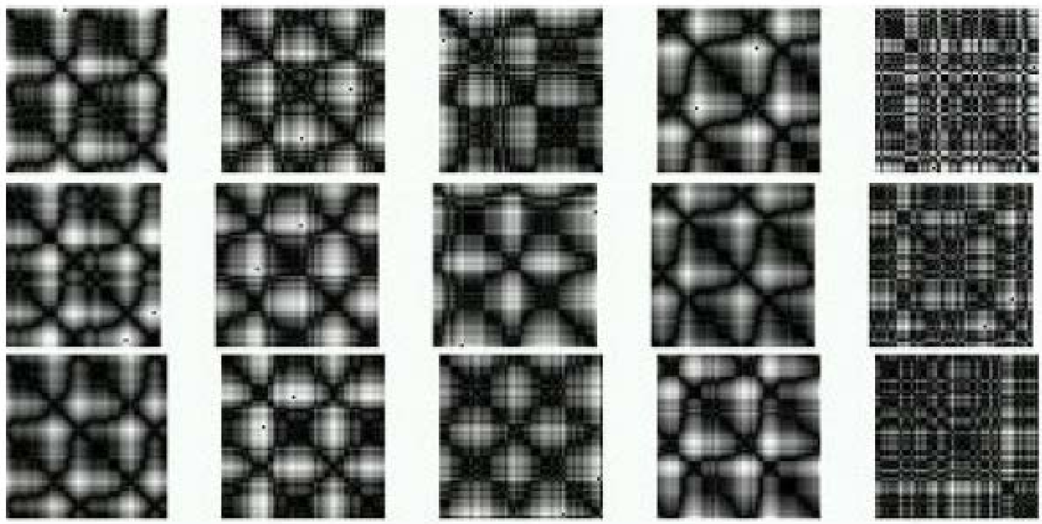
Self-similarity plot gives a measure of correlation of a variableover time with itself. Effectiveness of self-similarity plotsas a gait signature was evaluated by creating them from the angles of various extremities of the individuals' namely; front shin, back shin, front thigh, back thigh and torso.

Having a glance at the self-similarity plots, it was evidentthat the variation in the anglesof the torso was insignificant and random, as it change from sequence tosequence quite dramatically for the very same subject, probably due to error in the data acquisitionstages. Self-similarity plots for six subjects are shown.Column one is the front shin, column two is the front thigh, while column three is the backthigh, column four is the back shin and column five is the torso. Respective limb angles were the variable used to produce these plots.

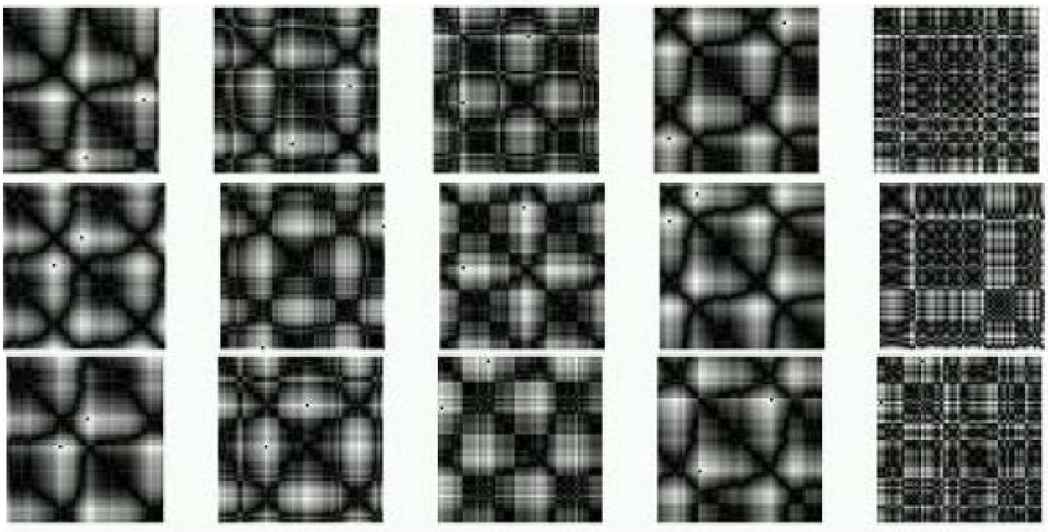
Person 1



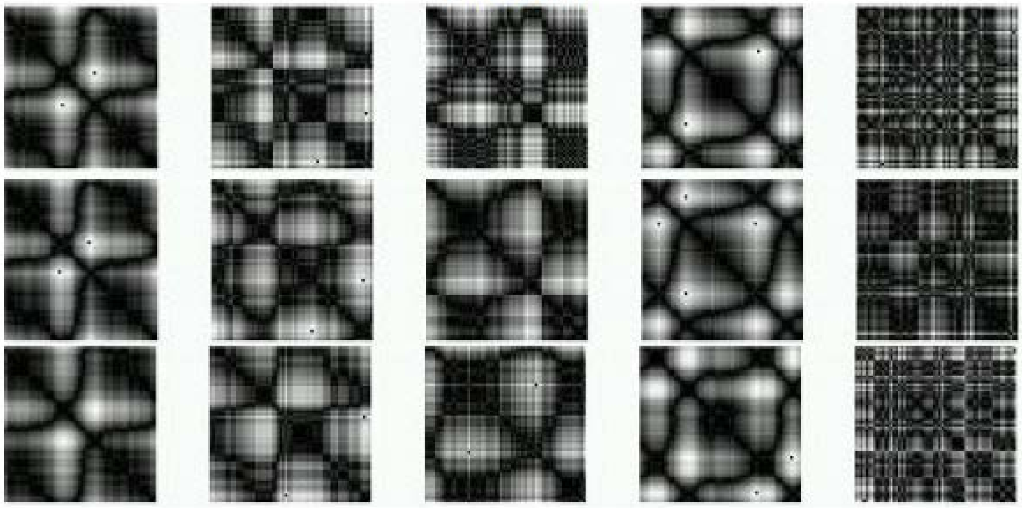
Person 2



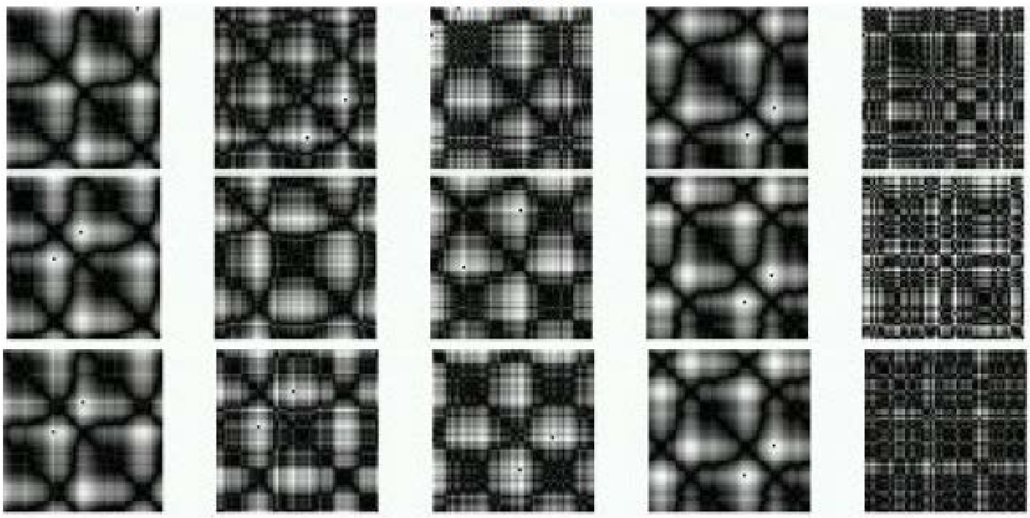
Person 3



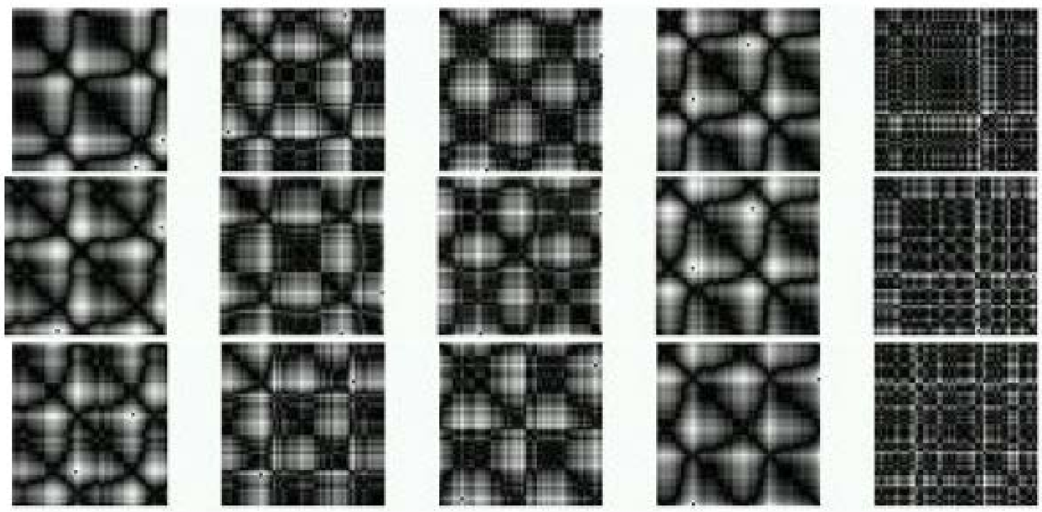
Person 4



Person 5



Person 6



By seeing these plots we can conclude that the -similarity plot for torso self-similarity plot is very different even for the same person and different videos. It is also apparent that the same personhas similar plots for different videosequences. Also, plots for different people vary; especially subject4, who hadchanged plots than other subjects.

From this we conclude that self-similarity plots offer us the prospect to produce a gait signature, as stated earlier by BenAbdelkader [18].

8.3.2 Limb Angles

The other measure used to extract a gait signature was the of the differentbody anglesfrom the video. Figure 8.5 is a graph plottedfor the rate of change of theshin angle ofsubject 1.

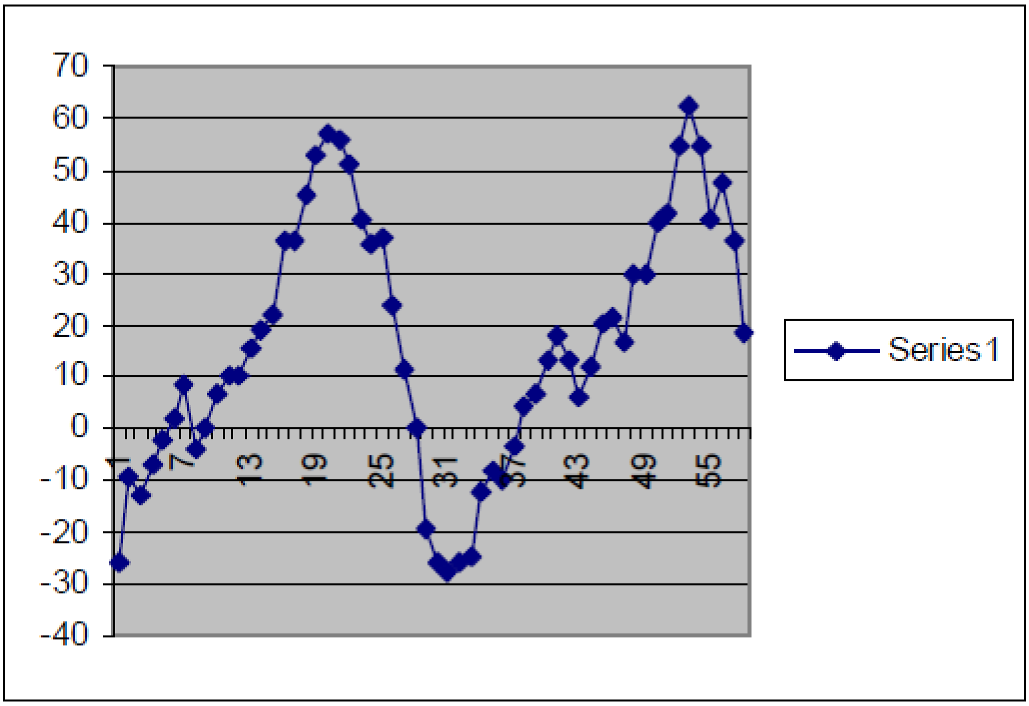


Figure 8.5 Graph Plot for Change in Angles of Shin for one Person

It is possible to make out the cyclic nature of the gait cycle from figure 37, and when compared against other subjects, these graphs showed a marked difference.

8.3.3 Limb Angular Velocities

The limb angular velocities provided a poor measure for a gait signature, probably due to errors introduced inthe earlier processes of data acquisition, which adversely affected the recognition. Figure 8.6shows the plot of the angular velocity for a person over 60 frames. It is jagged and could not offer much help for gait identification.

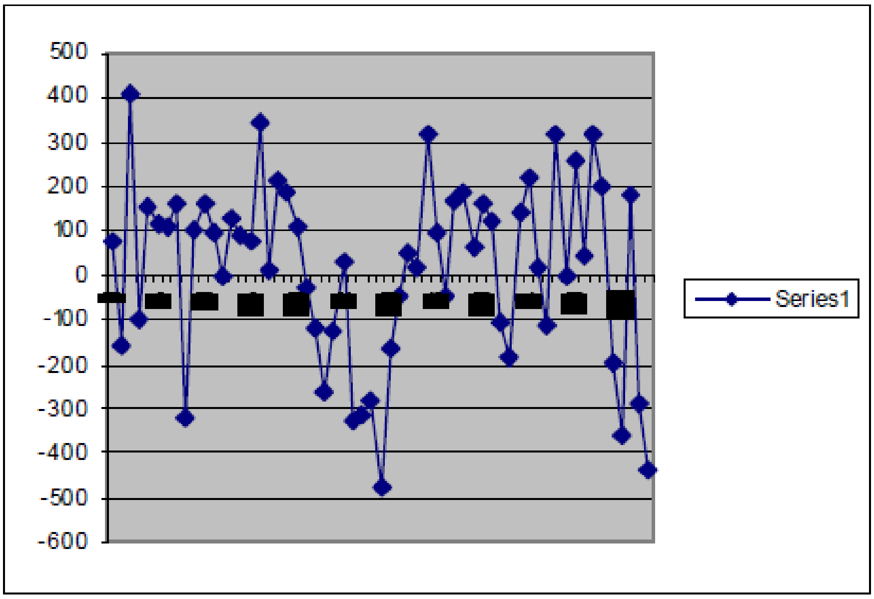


Figure 8.6 Plot of the Angular Velocity for a Person over 60 Frames

8.4 Recognition Rate

The most important measure of this project and also the aim of the project were the recognition rates, for various features. A table showing the successrates for recognition for the different types of gait signatures using PCAis provided.

The attempts column counts the times a person was given as a query for the variousfeatures. The success column shows how many of these queries were classified correctly bythe different methods.It could be seen that using the angles of the limbs proved to be the best way to derive a gait signature, with angular velocities being the worse.

Signature type	**Attempts**	**Success**	**Percentage**
Joint Positions	50	35	70%
Limb Angles	50	44	88%
Limb Angular Velocity	50	30	60%
Self-Similarity Plot	50	36	72%

Chapter 9

Conclusion

This chapter gives an overview of the whole project, its successes, limitations, and future improvements.

9.1 Achievements of the Project

The aim of this project was to develop an automatic system capable of recognizing peoplebased upon their gaits. We tried differing features to be used for the “gaitsignatures”, including; physical position of joints on the body over time, limb angles over a period of time, limb angular velocities and self-Similarity plots created from them.

PCA offered good representation of most of the data, whilst reducing dimensionality. Recognition rates up to 80% of correctly classified people out of 50 attempts were achieved by the project, further strengtheningthe notion that gait can be used for biometric identification. With more advanced classification techniques, higher recognitionrates can be achieved.

9.2 Limitations of the Project

Not all the fields of the project were successful. We extracted the featurepoints by model fitting. Segmentation went well but tracking the legs took a lot moretime than anticipated and eventually, only a rudimentary model fitting procedure wasemployed. The head and the body being tracked well, but thelegs were not being tracked to a well enough standard.

Gait as a recognition meanagain proved to be accurate, but is susceptible to factorssuch as clothes, change in mood, etc.There were also a lot of limiting simplifications assumed, which restrictedcertain features ofthe project in the real time applications. The subjects in the video were always walking perpendicular to the camera. Thiscould not be the case in real life as people would be walking at all angles to the videocamera.The classifiers used were rather simple, and recognition rates could be improved with advanced classifiers. Use of canonical analysis to reduce dimensionality canalso be improved.

9.3 Future Work

There are anumber of improvements, which can be made in thefuture, including improved pattern classifiers, advanced model fitting algorithms, code optimization, and different camera angles and creating a larger gait data set.

Both the classifiers used for this project, K-Nearestneighbor and Spatio Temporal Classification are good but there are moresophisticated classifiers available that could possibly offer better recognition rates. Non-linear classifiershave the ability to segment groups in a nonlinear fashion in the feature space, leading to increased success in pattern matching. Using them could potentially increase the rate of recognition.

Model fitting certainly needs future work to improve the tracking methods usedfor leg movement modeling. This is necessary if independent identification is to be performed which is a must for a real world gait recognitionapplication.

During the course ofthe development, good design andproviding functionality had preference to the optimization. After running the code it was experiencedthat bothsegmentation and recognition were computationally complexand overall performance would be enhanced from code optimization. This could either be done by refining the algorithms employed, or bycreating smaller functional data types which are not memoryhungry.

This program was limited to acquiring gait signatures from astatic camera position; it would be helpful to have the capability ofrecognize people from several angles, rather than being bounded to just one. These experiments werecarried out on a limited number of subjects, and it would be a good extension to determine a larger data set to get a better measure of how effective the gaitmechanisms are for recognition.

Appendix A – References

[1] Chris Kirtley"Psychological Influences on Gait",*Clinical Gait Analysis*, Sep 2011.

[2] M.P. Murray,"Gait as a Total Pattern of Movement."*American journal of Physical Medicine*vol. 46 pp. 290-333, 1967.

[3] G. Johansson "Visual Perception of Biological Motion and a Model for its Analysis," *Perception and Psychophysics*, 1973

[4] A.G. Bharatkumar,"Lower Limb Kinematics of Human Walking with Medial Axis Transformations,"*IEEE Workshop on Motion of Nonrigid and Articulated Objects*,1994

[5] Christopher Wren, Ali Azarbayejani, trvor Darrell and Alex Pentland,"PFinder: Real-time Tracking of the Human Body" *IEEE transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 97-99.

[6] D. Cunado, M.S. Nixon and J.N. Carter,"Using Gait as a Biometric, via Phase-weighted Magnitude Spectra,"*1st International Conferenceon Audio and Video Based Biometric Person Authentification*,Springer-Verlag,pp.95 – 102, 1997.

[7] Dr. David Cunado,"Model Based Gait Recognition – Variation in Hip Inclination"*Proceeding of the 3rd International Conference on Informatics and Technology*, 2009

[8] J.P. Foster, M.S. Nixon and A. Prugel-Bennett, "New Area Based Gait Recognition," *Audio and Video Based Biometric Person Authentication*,Springer – Verlag, pp. 311-317, June 2001

[9] P.S. Huang, C.J. Harris, and M.S. Nixon, “Recognizing Humans by Gait via Parametric Canonical Space,”*Artificial Intelligence in Engineering,*vol. 13 no. 4, pp. 359–366, October 1999.

[10] James J. Little and Jeffrey E. Boyd, “RecognisingPeople by their Gait: the Shape of Motion,” *MIT Press Journal Videre*, 1996“

[11] G. Gordon, T. Darrell, M. Harville, J. Woodfill. “Background Estimation and Removal Based on Range and Color,” *Computer Vision and Pattern Recognition*, vol. 2, pp. 459-464, June 1999.

[12] ImedBouchrika and Mark S. Nixon, “Model-based Feature Extraction for Gait Analysis and Extraction” *Proceedings of the 3rd International Conference on Computer Vision/Computer Graphics Collaboration Techniques, pp.* 150-160, Springer-Verlag Berlin, Heidelberg © 2007.

[13] JuergenLuettin, Neil A. Thacker and Steve W.Beet, University of Sheffield,“Active Shape Models for Visual Speech Feature Extraction” by, Electronic Systems Group Report No. 95/44

[14] T. Cootes, G.J. Edwards and C.J. Taylor, In H.Burkhardtand B. Neumann,“Active Appearance Models,”*5th European Conference on Computer Vision,*vol. 2, pp.484-498. Springer, 1998.

[15] Ross Cutler, *Face Recognition using Infra-redImages and Eigen Faces (Computer Science Technical Report Series),*University of Maryland, 1996.

[16] C. BenAbdelkader, R. Culter, H. Nanda, and L. Davis, “EigenGait: Motion-Based Recognition of People Using Image Self-Similarity,” *Proceedings of International Conference on Audio- and Video-Based Biometric Person Authentication*, pp. 284-294, 2001.

[17] Ping S. Huang, Chris J. Harris and Mark S. Nixon“Canonical Space Representation for Recognizing Humans By Gait and Face,”*IEEE Southwest Symposium on Image Analysis and Interpretation*,1998.

[18] I. Haritaoglu, D.HarwoodandL.Davis, “W4: Who, When, Where, What: A Real Time System for Detecting and Tracking People,”*International Conference on Face and Gesture Recognition*, April 14-16, 1998.

[19] J.BoydandJ.Little “Gait Data Set,”(Database on Department of Computer Science, University of Calgary System)

[20] “Gait Database”, Institute of Automation, Chinese Academy of Sciences,(former NLPR Gait Database) Dec. 10, 2001.