**UNIVERSITY OF SOUTHAMPTON**

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS

School of Electronics and Computer Science

# RECOGNIZING HUMAN GAIT BY MODEL-DRIVEN STATISTICAL ANALYSIS

by

**Jang-Hee Yoo**

Thesis for the degree of Doctor of Philosophy

May 2004

UNIVERSITY OF SOUTHAMPTON

## ABSTRACT

FACULTY OF ENGINEERING, SCIENCE & MATHEMATICS
School of Electronics and Computer Science

Doctor of Philosophy

# RECOGNIZING HUMAN GAIT BY MODEL-DRIVEN STATISTICAL ANALYSIS

by Jang-Hee Yoo

The ability to recognize humans by computer vision is a very important task, with many potential applications. In this thesis, we present a new method for an automated marker-less system to describe, analyze and recognize the human gait motion. The automated system consists of four stages: *i*) detection and extraction of the moving human body and its contour from image sequences; *ii*) extraction of human gait signatures based on topological analysis guided by known anatomical knowledge; *iii*) description of gait parameters by statistical analysis of the gait signatures; and *iv*) feature extraction and recognition of human gait. The gait signature is represented by a sequential set of 2D stick figures during one gait cycle. A grammatical structure with constraints of the gait sequences has been developed to improve the robustness of the gait signature, together with a new method of step symmetry. In the gait signature, the motion parameters based on biomechanical studies are calculated for characterizing the human gait. The inherent periodicity in gait motion is detected by graphical methods and analyzed by statistical approaches. Also, the periodic gait motion is modelled by interpolation of trigonometric-polynomials. In addition, the features based on motion parameters are extracted from the sequence of gait signatures. Then, a *k*-nearest neighbour classifier and an enhanced back-propagation algorithm is employed to recognize the gait. In experiments, the proposed methods have been successfully demonstrated on the largest available database. The gait signature is a very effective and well-defined representation method for analyzing the gait motion. It can be applied to other areas such as biomechanical and clinical applications, and we have estimated biomechanical parameters on a considerably larger population of subjects, showing that the estimate of variance by marker based techniques appeared generous. Moreover, the features extracted from the gait signatures are useful patterns for identifying human gait. As such, the marker-less approach confirms uniqueness of the gait as in biomechanical studies. In future, we will concentrate on improving the gait signature and on developing more efficient features to deal with the human gait identification with non-studio data.

# Contents

# List of Figures

# List of Tables

# Acknowledgements

My foremost and special thanks goes to my supervisor, Prof. Mark S. Nixon, who consistently provided me with valuable advice, encouragement, inspiration, and friendship in making this study an enjoyable journey of discovery. I also would like to thank my other supervisor, Prof. Chris J. Harris, who allowed me to start my study at Southampton alongside many other excellent researchers. Additionally, I am greatly indebted to all members of the Image, Speech and Intelligent Systems Research Group for providing a friendly and supportive working environment during my years here at Southampton. I am grateful to my colleagues and friends in Southampton and Korea who have made it possible for me to complete this thesis. Finally, my most heartfelt thanks go to my parents, who have supported me in all possible ways. Without their endless love and trust, I would not have dreamt of enjoying my achievement today. I would also like to thank other family members, especially my mother and father-in-law, and my brother and sister. My son Kyung-Hoon who has brought me deep joy and my wife Mi-Jin who has supported me in every way, this achievement is as much theirs as it is mine.

# Chapter 1

# Introduction

## 1.1 Motivation

Recent interest in computer vision has emerged which deals with the analysis of image sequences involving humans. This interest is motivated by the various application domains [39]: visual surveillance, clinical analysis, athletic performance analysis, computer animation, robotics, and biometrics. Moreover, human motion analysis has many challenging issues, because the highly flexible structure and self-occlusion of the human body mandates complicated processes for the measurement and analysis of the motion [71]. The motion of the human body may be defined by the movement of various body parts such as hand or limb segments, and it is known as a form of non-rigid and articulated motion [1][17]. Also, one of the most universal and complex of all human activities is gait motion. Gait is a pattern of human locomotion in which the body moves step by step in the desired direction [57][119]. It has been studied in medical science, psychology, and biomechanics for decades.

In addition, each person appears to have his or her own characteristic gait pattern. There is much evidence from psychophysical experiments [63][109] and medical analysis [57][79] that gait patterns have characteristics of uniqueness for each individual. In computer vision, recognition of humans by their gait has recently become a challenging area [86][95]. As a biometric [61], human gait is defined as a means of identifying individuals by the way they walk. Using gait has many advantages over other biometrics such as fingerprints, iris, and face recognition. Notably, it is available at a distance or at

low resolution, when other biometrics might not be perceivable [83]. It is also a non-invasive biometric technique, which can verify identity without contact and without a subject's cooperation. Although, several biometric techniques have now become practicable alternatives to traditional identification systems [61], human gait identification is still a difficult task.

On the other hand, human gait measurement is crucial in clinical applications, biomechanical analysis, and human identification. At the present, most available measurement systems are generally based on external markers which are attached to key anatomical positions of the human body [41] [110]. Accordingly, trajectories of the gait motion are observed by each marker's 3D position [66], and the trajectories translate into kinematic variables such as body movements and joint angles [28][66]. The use of markers however needs intrusive and expensive specialized hardware and requires contact with a subject and concerns overt rather than covert use. In computer vision, marker-less human motion analysis and recognition methods have recently been investigated. However, many motion analysis systems have been studied for tracking and extracting objects, though not for recognition purposes [2][39][77]. To enable greater application capability, a marker-less system is an essential requirement.

To summarize, the ability to recognize humans and their activities through visual gait measurements is not only a very important task with many potential applications but also a new study area with just a few studied methods. Here we are primarily considering using gait as a biometric. However, current approaches for automated gait recognition have several limitations. Even though higher classification rates have been achieved, a very small number of subjects were used in most experiments. Also, the essence of most approaches is based on visual template matching that does not depend on fundamental properties of the gait model or human body. Thus, our objective in this study is to develop an automated marker-less system for describing, analyzing, and recognizing the human gait motion from image sequences. To achieve this, we will need to extract the gait motion in an image sequence and to model the human body in order to observe how the body parts move in relation to each other. In addition, we may have to show the classification capability on the larger database in order to use gait as a biometric.

## 1.2 Related Work

There is much evidence to support the notion of using gait to recognize people. Psychological study [109] suggested that humans have a remarkable ability to recognize and distinguish between different types of motion. Also, medical studies [57][79][80] suggested that gait appeared unique to subjects, but involved components that can only be derived from an overhead view. In recent years, with advances in computing hardware, gait has become a potential practical biometric, and techniques have been developed [86]. Roughly, these techniques can be divided into model-free (or holistic) and model-based approaches [2][14][85]. The model-based approaches [9][23][81][121] use models of human motion such as a stick figure to represent the human body, whereas the model-free approaches [48][55][72][78][106] only use the shape and/or motion features by the statistics of the spatial-temporal patterns based on the image sequences. Here we will review human gait in medical and psychological studies and automatic approaches to recognize humans by their gait in computer vision.

### *1.2.1 Human Gait Studies*

The potential of gait as a biometric has further been encouraged by the considerable amount of evidence available, especially in medical and psychological studies. The studies of biological motion perception in psychology have progressed from establishing how humans can recognize subjects' motion, to recognizing people [86][109]. Studies using moving light displays (MLDs) [63] have indicated that subjects can be recognized solely by the manner in which they walk, rather than recognition by silhouette shape. Medical studies also support the notion of gait as a biometric. Indeed, gait can be used to diagnose a wide variety of medical conditions [92] from muscular disorders to congenital joint defects. Here we will introduce first the psychological evidence for gait recognition, followed by a review of medical studies for gait recognition in details.

The earliest psychological study of human perception of gait was performed by Johansson [63] using MLDs. Such displays were obtained by filming moving subjects with reflectors attached to their body joints, and filmed them walking in almost dark conditions. Thus, a moving light display contains only information about specific points

of an object undergoing motion. However, Johansson's initial experiments showed that humans are remarkably good at perceiving the human motion from MLDs. Given Johansson's early success, Cutting and Kozlowski [25] showed human perception of gait and their ability to recognize individuals using a dynamic light display of the walking pattern without familiarity cues. They used six subjects filmed walking normal to the plane view of camera. When asked how they recognized subjects, humans tended to mention as clues certain critical features of the displays, such as the speed, bounciness, rhythm of the walker, amount of arm swing or the length of steps.

In another study by Kozlowski and Cutting [68], they examined recognizing the gender of walker from MLDs involving 3 male subjects and 3 female subjects all about the same height. Their results showed that humans were able to correctly identify gender using full body joint markers at 63% correct on average, which is just better than chance (50%). In a later study, Mather and Murdoch [75] showed that frontal or oblique views are much more effective than a side view for gender discrimination, and emphasized that male subjects trend to swing their shoulders more while female subjects tend to swing their hips, the results improved to an accuracy of 79%. In a more recent study by Stevenage [109], she confirmed that gait could be used as a reliable means of recognizing individuals, and can learn their gait for recognition purposes from their video imagery. She also observed that even under adverse conditions, humans can still perceive gait as a cue to identity.

On the other hand, medical studies have been aimed to classify the components of gait for the treatment of pathologically abnormal gait. Murray [79] analyzed the walking patterns of sixty pathologically normal men aged between 20 to 65 years old. The data collection system used markers to be attached to subject. This form of data collection is typical within the medical field, and although practical in that domain, it is not suitable for gait recognition. Each subject was filmed walking for a repeated number of trials. Also, twenty simultaneous gait components were measured, and a mirror was used to note aspects of the overhead view and this was recorded onto film. Murray observed that each movement pattern was strikingly similar for repeated trials of the same subject. They also suggest that if all gait movements are considered then gait is unique. This makes gait an ideal candidate for use as a biometric.

### *1.2.2 Automatic Gait Recognition*

There are two major approaches to gait recognition in computer vision. The first is model-based where the subject's movement is described by a body model. In this approach, a body model is fitted to the human in every frame of the walking sequence, and kinematic parameters are generally measured on the body model as the model deforms over the walking sequence [23][64]. This approach was used by Niyogi [87] where a walking subject was detected by looking at an XT-slice (where X is a slice along the *x*-axis through a stacked image sequence and T is time). The information obtained from multiple XT-slices was then used to create a stick model of the subject for recognition. Bhanu [9] also adopted a body model based on the geometric representation of each part of the human body. The approach estimates 3D human walking parameters by performing a least squares fit of the 3D model to the 2D silhouette. The main advantage of this approach is that it does not rely on subjects walking normal to the plane view of camera.

In addition, Nash [81] used a simple pendulum model as a basis for searching a scene to locate a moving person using the Velocity Hough Transform (VHT). Cunado [23] built on this by using the VHT with double pendulum model to characterize the hip and thigh motion within a gait cycle. The gait features were derived from the Fourier weighted magnitude spectrum. The approach achieved a recognition rate of 90% on a database of 10 subjects. The idea of Cunado was later extended by Yam [121] to include the motion of the lower leg. The gait was modelled as a dynamic coupled oscillator that can be applied to database of running and walking gait sequences. Leg motion during walking and running was extracted using temporal template matching with a model defined by forced coupled oscillators. Fourier analysis of the variations in the motion of thigh and lower leg was used to generate gait features. The approach achieved recognition rates of over 90% on a database of 25 subjects with five image sequences for each subject.

An alternative method is to apply a model-free (or holistic) description to the set of images. Model-free approaches used features based on the motion or shape of subjects. This approach was used by Little and Boyd [72] where the motion of a moving human was described in order to recognize people by variation in the characteristics of the motion description. Murase [78] used eigenvectors for gait recognition where the

silhouettes of subjects were projected into the eigenspace for eignevalue decomposition. The eigenspace approach was extended by Huang [54][55] to use a combination of canonical space transformation, based on the canonical analysis, with the eigenspace transform. Another study also used eigen analysis to characterize gait. BenAbdelkader [7] proposed an approach based on self-similarity plots, and principal component analysis (PCA) is used to reduce the dimensionality of the feature space. A recognition rate of 72% was achieved on a database of 25 subjects with two image sequences for each subject.

Recently, Shutler [106] used the Zernike velocity moments to describe gait motion for recognition. The features were selected using an ANOVA technique, and a recognition rate of 100% was achieved on a database of 6 subjects with seven sequences for each subject. The approach to gait recognition by Foster [35] involves the area masks to measure dynamics of area change within specific regions of the image. As it is a measure of area, not only is it fast in implementation, but it also allows for specificity to gait by choice of the masks used. Also, Hayfron-Acquah [48] has used symmetry of human motion using generalized symmetry operators. This operator locates features by their symmetrical properties rather than relying on the borders of a shape or general appearance. The approach achieved high recognition rates of over 90% on various databases including 100 subjects. In addition, the potential of baseline approach matching silhouette [97], data derivation of stride pattern [64], key frame analysis for sequence matching [22], and ellipsoidal fits [70] have been used to generate gait features for recognition.

The model-based approaches and model-free approaches have their own advantages. Some of the model-free approaches have improved capability over application problems such as noise, because they use more subject information by using the complete silhouette of subject and can be viewpoint invariant. However, one of the main advantages of the model-based approaches is their handling of occlusion, which is of importance in gait as the human body is self-occluding when walking. Also, they can be used to other applications such as clinical analysis. Nevertheless, the use of a silhouette is clothing invariant, but this is more a recognition issue.

## 1.3 Contributions

Several new methods for describing, analyzing and recognizing the human gait motion are developed in this study. The major contributions of this thesis will be made in the following aspects:

- We propose a new method for extracting the gait signature based on topological analysis guided by anatomical knowledge. The gait signature is a sequential set of the 2D stick figures during one gait cycle. It appears to be a very effective and well-defined method for analyzing human gait motion. The gait signature can be used to estimate the gait parameters in biomechanical or medical applications.
- We propose a new method for interpreting the structure of the gait sequence by using grammatical rules with physical constrains. As a constraint, step symmetry is newly defined by the relationship of the joint angles. The structure of human gait motion is also described by regular grammar. These methods are used to improve the robustness of the gait signatures, especially at crossover of the legs.
- We propose a new method for modelling the periodic gait motion via interpolation by trigonometric polynomials. The pattern of rotation angles around the joints is the most important kinematic parameter and defined as a gait time series. The gait time series can be characterized as having a periodic component. This method is efficient for describing the periodic motion and to handle poor quality data.
- We propose a new method for automated gait recognition by a neural network using the features based on motion parameters. The motion parameters can be extracted by analyzing the sequence of gait signatures. A $k$-NN classifier will be used to reveal the discriminatory capability of the feature vector confirming its validity. As a biometric, an enhanced back-propagation algorithm will be employed to recognize the gait.

In addition, early experiments from other approaches were often carried out on small number of subjects, usually up to 10 persons. In this study, the usefulness of the proposed methods will be demonstrated using a much larger database which is currently the largest gait database of its kind consisting of over 100 different subjects.

## 1.4 Thesis Overview

The thesis is divided into six chapters. The remaining chapters will be organized as follows (also see Figure 1.1):

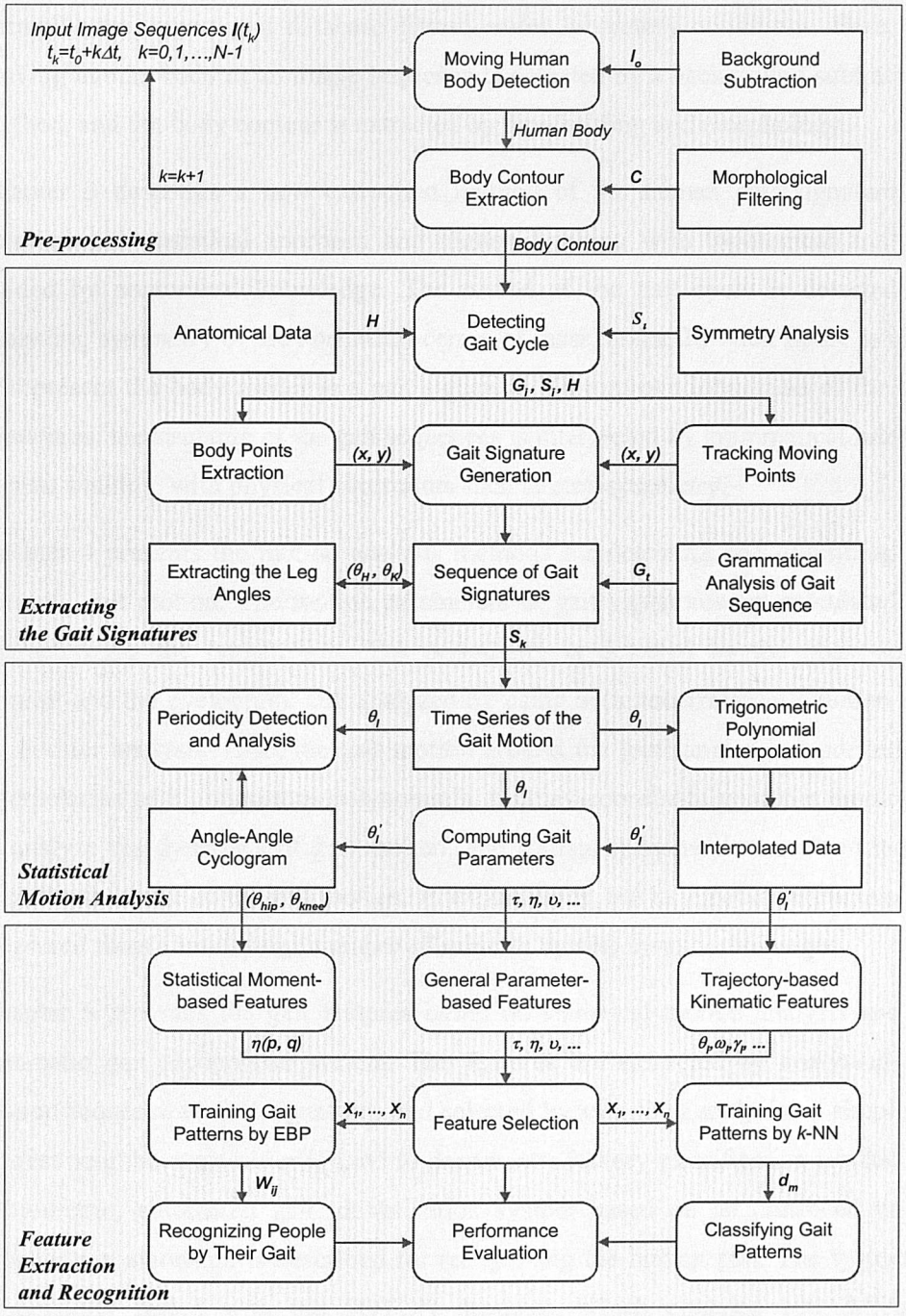



Figure 1.1: System Architecture Used Within this Thesis

- Chapter 2 presents a pre-processing method of image sequence for extracting the moving human body and its body contour. The various motion capture methods and human body models are reviewed, and our gait database (which is called the SOTON database) based on image sequences is described. The SOTON database is currently the largest gait database filmed under laboratory conditions. Here, the moving human body in an image sequence is detected by a background subtraction method, and the body contour is extracted by thresholding and morphology.

- Chapter 3 describes a new extraction method of the human gait signature by combining a statistical approach and motion tracking with topological analysis guided by anatomical knowledge. The period of the gait cycle is detected by analyzing symmetry of the horizontal centre of mass, and a 2D stick figure is used to represent the body model in a gait signature. To improve robustness of the gait signatures, the structure of the gait sequences is interpreted by grammatical rules of human walking, with physical constraints such as step symmetry.

- Chapter 4 presents the motion analysis methods for detecting and describing the periodic gait motion. The motion parameters of gait signatures are calculated for characterizing the human gait. The periodicity is detected by the phase-space portrait and the cyclogram, and analyzed by using an autocorrelation function and by Fourier analysis. Also, the gait motion around the joint angles is modelled by interpolation of trigonometric-polynomials. A delay-coordinate system is employed to analyze the dynamics of gait motion, and statistical moments are also used to describe the scale of the gait motion. In experiments, the kinematic parameters are measured from a much large number of subjects by non-invasive technique.

- Chapter 5 provides the gait features based on statistical motion analysis and an automatic gait recognition system. The features are extracted by analyzing the motion sequence of gait signatures and selected by statistical analysis. A simple *k*-nearest neighbour classifier is used to derive introductory classification results. As a biometric, automated gait identification system based on an enhanced back-propagation algorithm is described for recognizing the human gait. The system is successfully tested with the SOTON database, which contains 100 different subjects with seven image sequences of each subject.

Finally conclusions and future work of this study are discussed in Chapter 6.

## 1.5 List of Publications

Six papers relating to this study have been published. They are:

- Extracting Gait Signatures based on Anatomical Knowledge, *BMVA Symposium on Advancing Biometric Technologies, http://www.bmva.ac.uk/meetings/meetings/02/6March02/soton2.pdf*, London, UK, March 2002.
- Extraction and Description of Moving Human Body by Periodic Motion Analysis, in *Proceedings of the ISCA 17th International Conference on Computers and Their Applications*, pp.110-113, San Francisco, USA, April 2002.
- Extracting Human Gait Signatures by Body Segment Properties, in *Proceedings of the 5th IEEE Southwest Symposium on Image Analysis and Interpretation*, pp.35-39, Santa Fe, USA, April 2002.
- Model-Driven Statistical Analysis of Human Gait Motion, in *Proceeding of the IEEE 2002 International Conference on Image Processing*, pp.285-288, Rochester, USA, September 2002.
- On Laboratory Gait Analysis via Computer Vision, in *Proceedings of AISB'03 Symposium on Biological-Inspired Machine Vision, Theory and Application*, pp.109-113, University of Wales, Aberystwyth, UK, April 2003.
- Markerless Human Gait Analysis via Image Sequences, in *Proceedings of the International Society of Biomechanics XIXth Congress*, Dunedin, New Zealand, July 2003.

In addition, two awards are associated with the last paper:

i) NAC/Miyashita Award for Best Paper on Film/VTR Analysis

ii) NDI Student Award

both at the International Society of Biomechanics XIX$^{th}$ Congress, July 2003.

# Chapter 2

# Video-Based Gait Silhouette Data

## 2.1 Representing the Human Motion

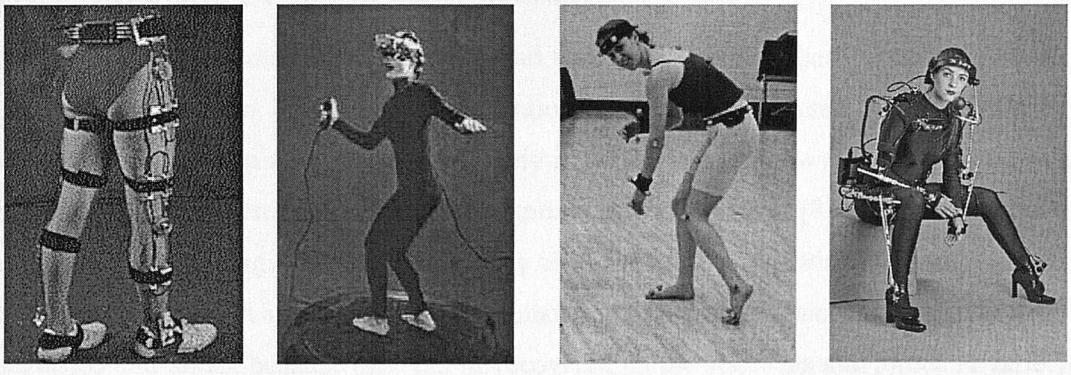
The human body's motion may be defined by the movement of various body segments such as the hand or limb [2][17]. A wide range of methodologies and systems are available for human motion capture and analysis in laboratory environments, and commercially available systems are usually based on markers which are attached to the human body [41][66][110]. Using markers can acquire precise motion information, but requires specialized hardware and subject contact. Therefore, with advances in computing power, marker-less methods have recently been investigated in computer vision, and many human body models have also been developed to describe the human movement from a non-invasive video sequence [77]. In computer vision, the ultimate goal of human motion analysis is to recognize the human body and its activities. However, most of the current body models are mainly developed for tracking humans, and hence the human body and its motion model should be considered with a view to recognition purposes.

### *2.1.1 Human Motion Capture*

Human motion capture is the process of recording human movement and translating it into usable mathematical terms by tracking position relative to a fixed point in the physical space over time [91][110]. In general, a motion capture system consists of subsystems for sensing and processing, respectively. The operational complexity of these subsystems is typically related to the use of active (marker-based) or passive (marker-

less) sensing [77]. In the sensing stage, a group of sensors simultaneously send position and orientation data to a processing subsystem (or computer), and the processing subsystem can represent the observed movement in the form of data-files which describe the three-dimensional (3D) trajectories of the sensors [66]. Accordingly, the human's movements can be recorded by sensors, and the recorded data can be played back by mapping the motion onto a computer model. Several methods have been used successfully to capture human motion. Some methods use cameras that digitize different views of the movement, which are then used to determine the position and orientation of key points or normally reflective markers. Others use magnetic fields or ultrasound to track a group of sensors. Also, mechanical systems based on linked structures use potentiometers to measure the angle of a joint [91][110].

Marker-based technology is the most commonly used motion capture method in commercially available systems, due to accuracy in representation. Figure 2.1 shows examples of motion capture methods based on markers. An electrogoniometer is a device for converting continuous measurements of joint motion into an electrical signal, and the measured output is usually plotted as a graph of joint angle against time [110][117]. Magnetic motion capture uses 6 to 12 or more sensors to measure the magnetic field generated by a transmitter source. The sensors and source are connected to an electronic control unit that is networked with a host computer, and can determine their position and orientation within the space [11]. Also, optical motion capture uses reflective markers (or pulsed-LEDs) attached to the body and a number of special cameras to track the 3D



(a) Electrogoniometer (b) Magnetic-Sensor (c) Optical-Marker (d) Electromechanical

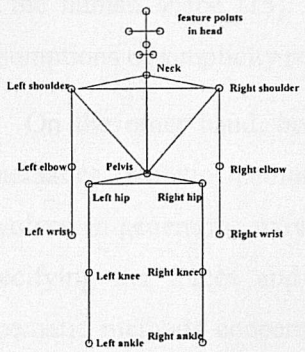
Figure 2.1: Marker-based Motion Capture Methods

location of the markers [41]. The movement of the body segments from the marker data can be represented as a body skeleton with 3D XYZ position [49]. More recently, a mechanical motion capture method uses an exoskeleton incorporating electromechanical sensors for measuring joint angles. The joint angles are then tracked as a human figure motion whilst wearing the exoskeleton [41][91].

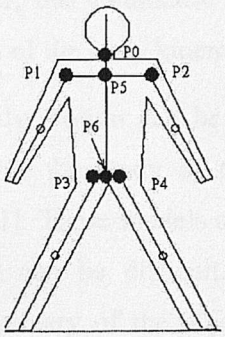
Alternatively, marker-less human motion capture can be achieved by a vision-based method in complex video sequences. This method normally does not require any markers or specialized hardware attached to the human body. The only input needed is an ordinary video recording of the subject [12]. Therefore, the vision-based method is more accessible, more cost-effective, less encumbering of the humans, and works in a wider variety of environments than current marker-based capture methods [41]. The vision-based method however involves segmenting the body parts, tracking the movement of joints, and recovering the body structure in an image sequence [2]. Also, this low-level processing requires complicated vision computing on a high-performance computer system. Recently, a number of methods concerning the motion capture and analysis have been proposed in the computer graphics and vision studies [11][12][77]. To analyze the human motion, a variety of human body models using 3D structure of rigid segments, joints, and constraints have developed [2][39][77][115]. The human body model determined by a vision based approach is a very important component for recovering and interpreting the human movement from a non-invasive image sequence.

### *2.1.2 Modelling Human Body Motion*

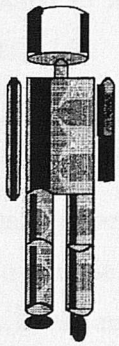
Many studies have considered extracting and tracking the human body's motion, though rarely for recognition. The human body motion is usually represented by different body models. The selection of an appropriate body model is important to efficiently recognize human shapes from an image and analyze human motion properly [83]. The human body consists of several rigid parts connected by the joints, but the motion of a full human-body is non-rigid and articulated [1][5]. Human body models are basically based on body segments and joints, because they can be recovered by the segments and joints. A variety of body models, such as stick figures, 2D contours, and volumetric models have been used to represent the human body and its motion. Figure 2.2 shows examples of these.




(a) Stick Figure Model




(b) 2D Contour Model



(c) Volumetric Model

Figure 2.2: Human Body Models [1][71]

Stick figure models consist of line segments that are connected at joints to represent the human body. The stick figure is obtained by various methods such as the statistical means of the output of the median axis or distance transforms [59]. Lee and Chen's model [21][69] uses 14 joints and 17 segments for the head, torso, hip, arms, and legs. The length of each segment and the coordinates of the joints are the model parameters which are used for tracking. This model is based on the observation that human motion is essentially the movement of the human skeleton brought about by the attached muscles. 2D contour models are another method to describe the human body. This representation is directly relevant to the human body projection in the image plane [115]. Leung and Yang's model [71] consists of five ribbons and a body torso, various joint and mid points, and a number of structural constraints, such as support. The 2D contour model was used to guide the labelling of the image data.

In addition to the basic 2D model, view-based knowledge is defined for a number of generic human postures, to aid the interpretation process. Volumetric models [51][100] are intended to better represent the complexity of the human body, but require more parameters for computation. Rohr's model [100] uses 14 elliptical cylinders to represent the human body as 3D volumes, and the origin of the coordinate system is fixed at the centre of torso. Generalized cylinders, i.e. cylinders with an elliptical cross-section of constant size and shape, are simplified examples of generalized cones. This model can be as refined as necessary, by using a collection of component cylinders representing the different body segments, giving more detailed information about the spatial organization

of the human shape [15]. However, the volumetric model is restricted to impractical assumptions of simplicity regardless of the body kinematics constraints [115].

On the other hand, human body motion can be described in terms of kinetics or kinematics. Kinetic methods involve the study of the forces and torques which are involved in generating movement [1]. These models are computationally expensive, and specifying the forces and torques can be difficult. In contrast to kinetic methods, kinematic methods concern the geometry of the object, such as its linear and angular positions, orientation, and deformation. If movement is explicitly given by time-dependent functions then it is very easy to simulate motion. Most of the model-based tracking approaches in computer vision are concerned with studies of the kinematic patterns [1][100]. Also, human body motion in kinematic methods is usually characterized by joint angles, as extensively studied in medical studies [80].

The movement of human walking can be modelled by using ideas from human motion studies [83]. Hogg [51] and Rohr [100] use flexion and extension curves for the hip, knee, shoulder and elbow joints in their walking person models. Joint angles are also used by Bharatkumar *et al* [10] to represent the walking cycle of the lower limbs in human walking and compare it with the kinematic model. Another approach to modelling the body's motion is to use a sequence of stick figures, called a key frame sequence [3], to model rough movements of the body. This key frame sequence of stick figures is used to indicate the approximate order of the motion and spatial relationships between the body parts. Hence, each figure represents a different phase of the body movement, and the key frame sequence is determined in advance and referred to in the prediction process.

The human body motion is well represented by its joints or skeletal structure (the "stick figure") since it reflects anatomical features of the human [77]. Also, the motion of joints gives the key to motion estimation and analysis. Namely, the stick figure model is closely related to the observation that human motion is essentially the movement of the human skeleton, thus the stick figure can be described as a collection of body segments and joint angles with various degrees of freedom [1][39]. Here, we consider a simplified stick figure model for representing the body structure and describing the body motion. The stick figure model can effectively represent structures and kinematics of the human body motion.

## 2.2 Human Gait Databases

A human gait database is one of the most important components for performance evaluation in the gait analysis and biometric systems. The construction of the gait database can require much time and resources. In clinical or biomechanical applications, the gait data have been collected by using marker-based motion capture systems in the form of data-files, but the gait databases in vision applications are usually constructed by using video camera systems to derive image sequences. The image sequences require a large amount of storage and computational time for processing and analysis. Therefore, early gait databases were collected under very limited conditions and consisted of a relatively small number of subjects. Here we describe the early gait data sets and the recently developed large databases. The main sources of the earliest databases were University of Southampton (SOTON), hereafter referred to as the early SOTON data and University of California at San Diego (UCSD), and the new SOTON database is mainly used in this study.

### *2.2.1 Early Gait Data Sets*

Early approaches to automatic recognition by gait have been evaluated on small data sets, with many researches reporting experimentations using their own data, because there are

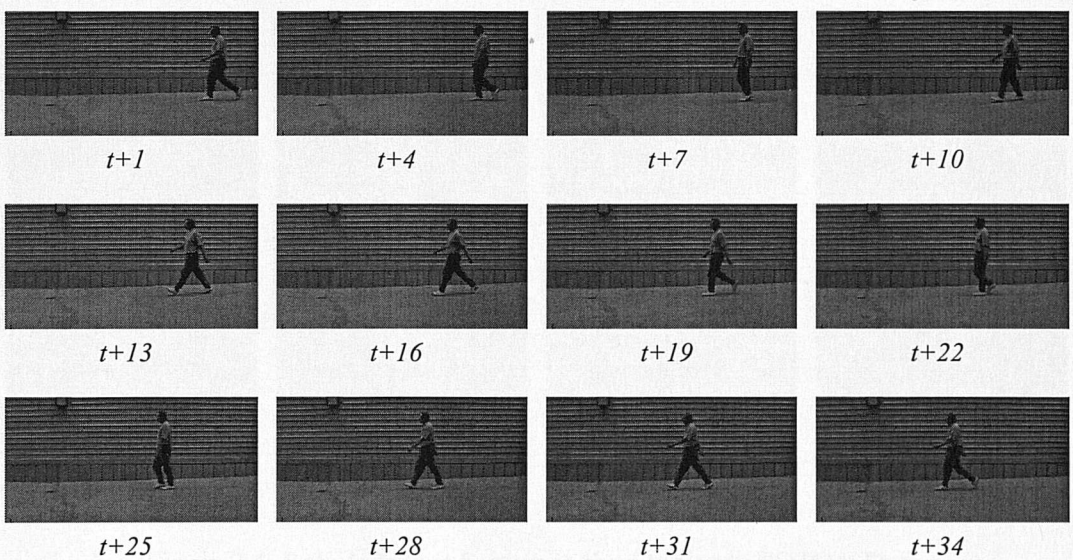



Figure 2.3: Sample Image Sequence from the UCSD Data

as yet no available large databases for vision based systems. There are two available and well-known small data sets: the UCSD data and the early SOTON data. The UCSD data was collected by the Visual Computing Group, University of California at San Diego [72], and the early SOTON data was collected by the ISIS (Image, Speech, and Intelligent Systems) Research Group, University of Southampton. The UCSD data was taken outdoors without lighting control and with complex background, conversely the early SOTON data was taken indoors with lighting control and a plain background. The UCSD data was acquired at 30 fps (frames per second) with 320×160 greyscale pixels, and the data set consists of 6 subjects and 7 sequences for each. The early SOTON data was acquired at 25 fps with 384×288 greyscale pixels, and data set contains 4 subjects and 4 sequences for each. Sample sequences of the original images in the UCSD data are shown Figure 2.3, and the early SOTON data are shown in Figure 2.4.

Those two data sets have been used in most early studies for automatic gait recognition, and several different publications have shown close to 100 percent

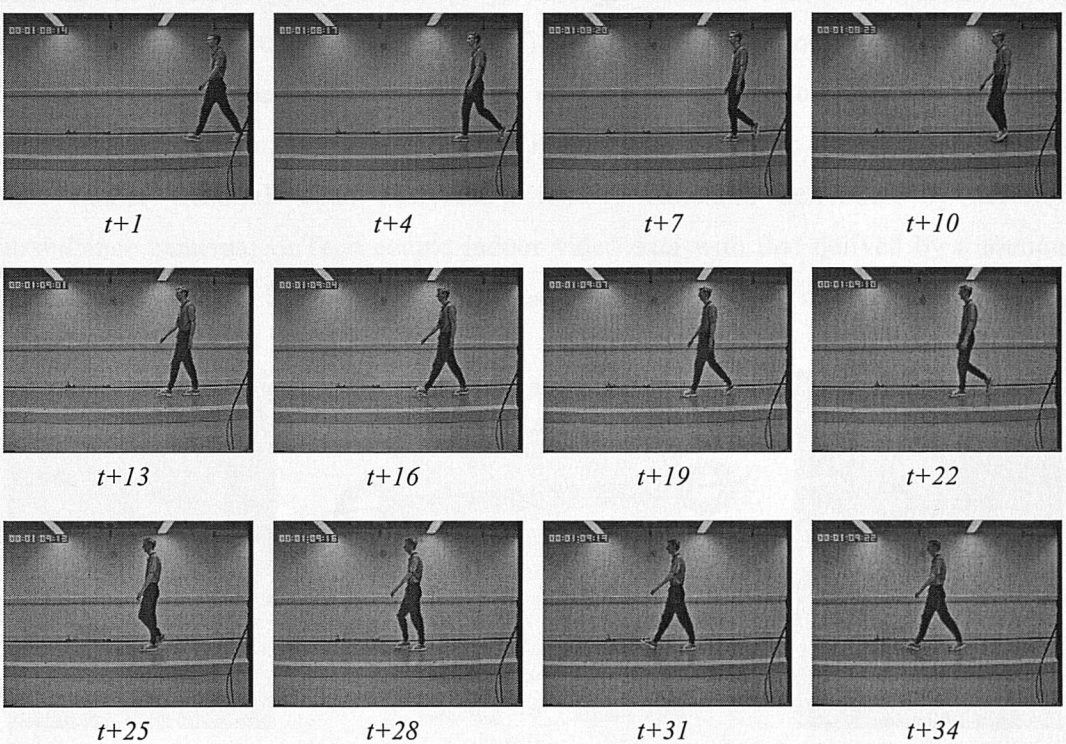
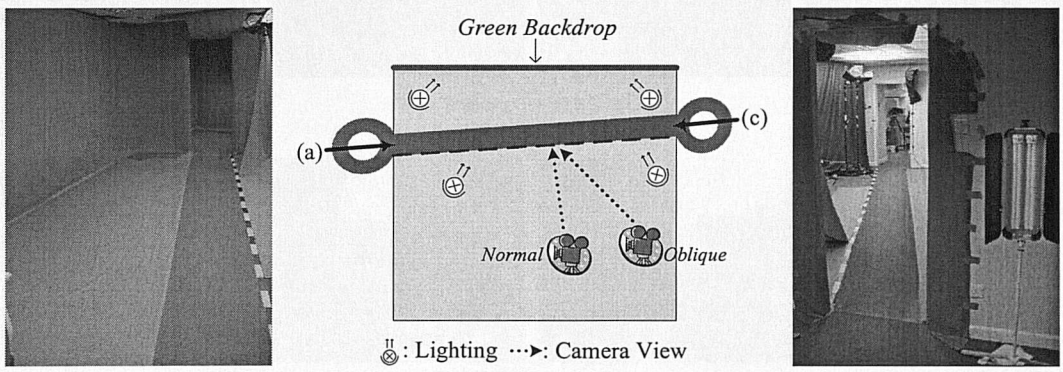



Figure 2.4: Sample Image Sequence from the early SOTON Data

classification capability [23][54][72]. It has been very encouraging to note that similar levels of classification can be achieved on the much larger data sets. As gait is a behavioural biometric, there is much potential for within subject variation such as footwear, clothing and apparel. None of these factors were considered in the early data sets [85]. Application factors concern deployment via computer vision though none of the early data sets allowed facility for such consideration, save for striped trousers in an early SOTON data set (aiming to allow for assessment of validity of a model-based approach). The new SOTON database sought to include more subjects in order to allow for an estimate of inter-subject variation, together with a limited estimate of intra-subject variation thus allowing for better assessment of the potential for gait as a biometric [86].

### *2.2.2 Large Databases: SOTON Database*

Several larger gait databases have recently been developed within the DARPA Human ID at a Distance program [95]. This program includes: University of Maryland [65], Georgia Institute of Technology (GaTech) [64], Carnegie Mellon University (CMU) [45], Massachusetts Institute of Technology (MIT) [70], and University of Southampton [105]. Each site has developed a database and has evaluated techniques on their own database and on other databases. The main difference in database design are: CMU consider multi-view indoor treadmill data; Maryland's footage simulates derivation by security surveillance cameras; GaTech couple indoor video data with that derived by a motion capture system; and Southampton concerns multi-view indoor and outdoor data on

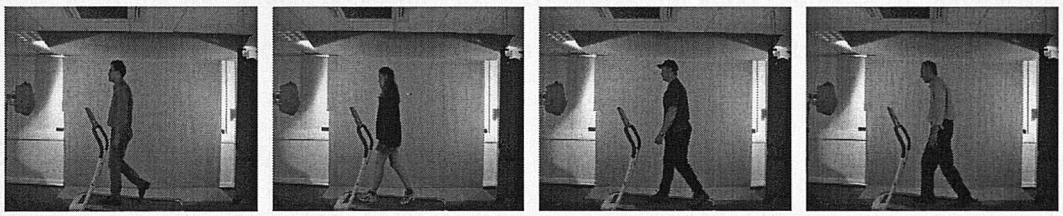



(a) Front View (b) Layout and Views (c) Far View

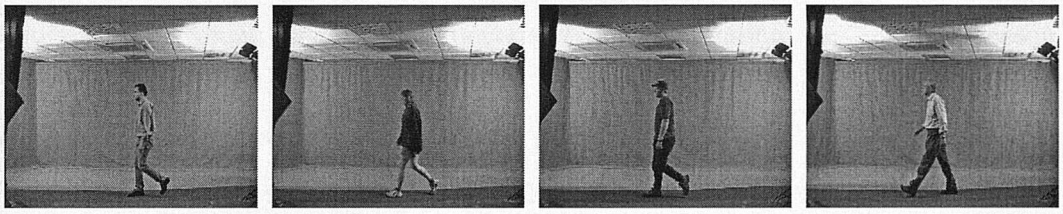
Figure 2.5: Indoor Walking Track used in the SOTON Database

treadmill, and on tracks [85].

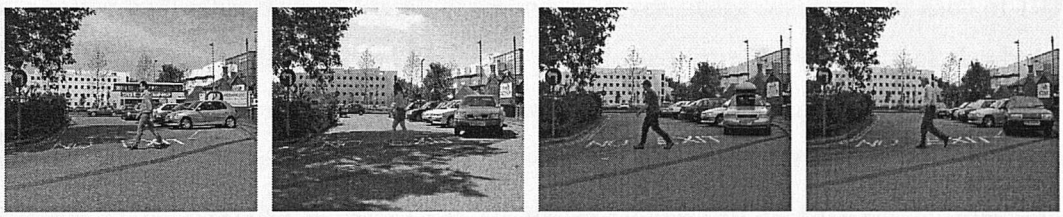
The SOTON database [105] by the ISIS Research Group captures the subjects using good quality progressive scan and interlaced DV (digital video) camcorders. The progressive scan technology provides high-resolution imagery whilst security video often uses interlaced data. In order to provide an approximation to ground truth and to acquire imagery for application analysis, the subjects were filmed indoors (under controlled lighting with a special background) and outdoors (without lighting or background control), respectively. The first form of indoor data is a subject constantly walking on a treadmill, and the second form is subject walking along a specially designed track shown in Figure 2.5. As can be seen in the figure, the track was prepared with chroma-key cloth (bright green, as this is an unusual clothes' colour), and the background was illuminated by photoflood lamps. The same camera view and chroma-key arrangements were used for the treadmill, but subjects were highlighted with diffuse spotlights. The outdoor data used



(a) Indoor Treadmill Data



(b) Indoor Track Data



(c) Outdoor Track Data

Figure 2.6: Sample Images from the SOTON Database

a similar track layout with a greater distance between subject and camera. The background of the outdoor data contained a selection of objects such as foliage, pedestrian and vehicular traffic, buildings as well as occlusion by bicycles, cars and other subjects.

In addition, each subject was filmed walking in both directions, and the database contains more than 100 subjects. All subjects in the database are filmed in fronto-parallel and obliquely viewed imagery (allowing orientation independent analysis), and there is ancillary data which includes subject specific information, camera setups and extraction parameters [105]. Each subject has at least four image sequences and each image sequence contains at least one gait cycle, together with background and other supporting data. Also, an image sequence contains only a single subject walking at normal speed and was acquired at 25 fps with 720×576 colour pixels. The imagery for the SOTON database was completed with a high-resolution still image of each subject in frontal and profile view, allowing for comparison with face recognition and good estimates of body shape and size. Figure 2.6 shows the sample images from the SOTON gait database.

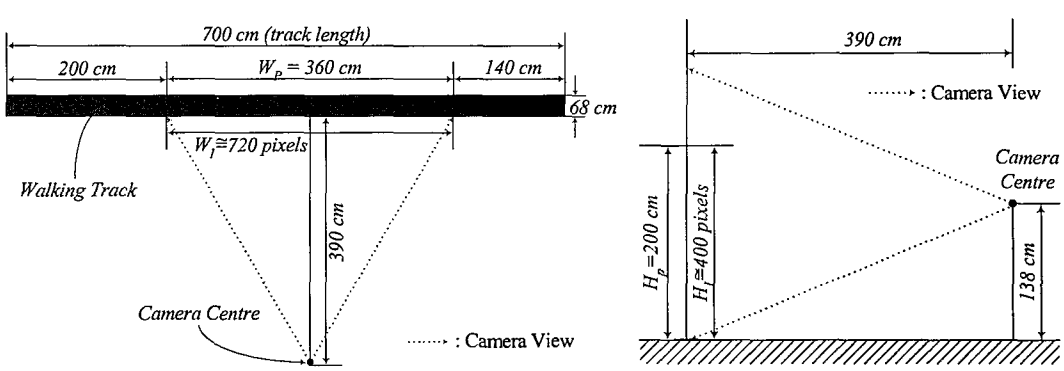
### *2.2.3 Analyzing the Experimental Database*

The most recent version of the SOTON database contains 114 different subjects, and each subject was filmed indoors and outdoors. In addition, the SOTON data were mostly acquired from young and healthy university students during the summer season. However, the SOTON indoor track database is mainly used as an experimental database in this study due to time constraints, and the early SOTON and UCSD data sets are used in the experiment for detecting a human body in greyscale images (*see* Section 2.3.1). In addition, the early SOTON data is also used in the classification experiment for comparing with other studies. For all experiments, seven image sequences of each of 100 different subjects (16 females and 84 males) are selected from the indoor track database. This selection is based on number of sequences of each subject, because some subjects do not have enough image sequences. From now on, we will refer to this large experimental database as the SOTON database. Table 2.1 shows summary of the subjects' information, which is obtained from the SOTON database. In the table, the height and weight data were as stated by each subject without measurement.

Table 2.1: Summary of the Subject Information from the SOTON Database

| Gender | Age Level | #. Of Subject | Ave. Height (Cm) | Ave. Weight (Kg) |
|---|---|---|---|---|
| Male | Adults | 82 | 176.8 ± 6.7 | 70.2 ± 8.0 |
| | Children | 2 | 125.0 ± 7.1 | 43.0 ± 5.7 |
| | All | 84 | 175.5 ± 10.4 | 69.9 ± 8.4 |
| Female | Adults | 16 | 161.8 ± 6.9 | 56.4 ± 9.6 |
| | Children | 0 | - | - |
| | All | 16 | 161.8 ± 6.9 | 56.4 ± 9.6 |
| Total | | 100 | 173.3 ± 11.1 | 67.8 ± 9.9 |

On the other hand, the physical dimension in an image plane can be estimated by analyzing the geometric aspect of image formation in camera models. The most common geometric model is the perspective or pinhole model [34][113]. Also, the 2D intensity image is the result of a perspective projection of the 3D scene. In practice, the real-world and camera coordinate systems are related by a set of physical parameters, such as the focal length of the lens, the size of the pixels, the position of the principal point, and the position and orientation of the camera [113]. However, the set of physical (geometry) parameters is not included with the SOTON database. Therefore, neglecting any geometric distortions, the physical dimension can be simply approximated by linking the position of scene points with that of their corresponding image points.




(a) Plane View (b) Vertical View

Figure 2.7: Physical Mapping between Image Plane and Walking Track

Figure 2.7 shows the physical mapping between the image plane of Figure 2.6(b) and the walking track shown in Figure 2.5(b). As can be seen in the figure, the DV camera is located at the centre position normal to the subject walking track, and the physical dimensions $W_P \times H_P$ are roughly covered by the image plane $W_I \times H_I$ of camera view. Accordingly, one pixel of the image plane can be approximated by the physical dimension $0.5cm \times 0.5cm$. By using this relation, we can estimate the stride and height parameters of gait motion. In future, we aim to calibrate the camera so that recovered image positions can be translated to physical laboratory position. The relative dynamics remain unchanged.

## 2.3 Extracting the Body Silhouette

The detection of moving objects from image sequences is a fundamental and important problem in many vision systems. There are two basic methods for detecting moving objects: temporal differencing and background subtraction. Temporal differencing [67] can adapt to dynamic environments, but cannot robustly extract all relevant object pixels. Thus, the method has been mainly used for tracking moving objects. Conversely, background subtraction [47][52] extracts the most complete representation of an object, but this method is very sensitive to dynamic scene changes due to illumination [37]. However, background subtraction has been successfully applied to many vision systems as a pre-processing phase for object detection and extraction in an image sequence [47][52][82][120]. Here, the moving human body is detected by background subtraction methods. The body region is then determined by analyzing histogram projection profiles, and its location is verified by prior knowledge such as size and shape. Also, thresholding and morphology is employed to extract the body contour of a detected human body.

### *2.3.1 Detecting a Body in Greyscale Images*

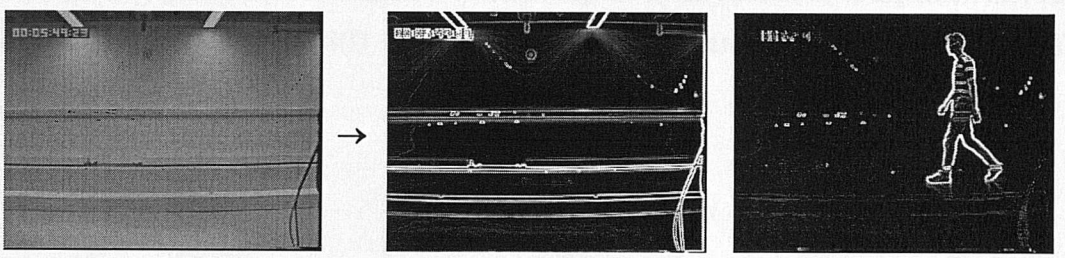
To detect a moving object in an image sequence, the region of interest is typically obtained by background subtraction. The basic idea of background subtraction is to subtract the object image from a reference image, which is acquired from a static background during a period of time. In mathematical terms, a background image $\mathbf{I_B}$ is

denoted by $\mathbf{I_B} = \{i_{b1}, ..., i_{bn}\}$ and the current object image $\mathbf{I_O}$ is denoted by $\mathbf{I_O} = \{i_{o1}, ..., i_{on}\}$ where $n$ is the number of pixels in an image. Then, the difference image $\mathbf{I_D} = \{i_{d1}, ..., i_{dn}\}$ is defined as $\mathbf{I_D} = |\mathbf{I_B} - \mathbf{I_O}|$. In the ideal case, the difference image $\mathbf{I_D}$ is described by following characteristic:

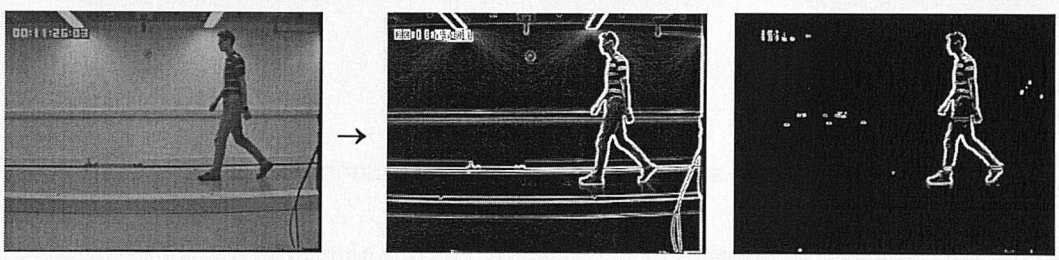
$$i_{dk} = \begin{cases} | i_{bk} - i_{ok} | & \text{if } (p_k \subset N) \\ 0 & \text{otherwise} \end{cases}, \qquad k = 1, ..., n \tag{2.1}$$

where $p_k$ is position of a pixel, and $N$ denotes an object region. Though, many background subtraction methods have been proposed, most are very sensitive to both global and local illumination changes such as shadows and highlights [47][52][120].

To detect the human body in an image of a real scene, background subtraction using an edge difference image is employed to handle changes in illumination. The background image is modelled by taking the median of an image sequence that belongs to the background [84]. The Sobel operator is then applied to each image to obtain the edge image. After that, the edge difference image between the background and the object edge image is obtained by the background subtraction. That is, let $I_{be}(x, y)$ and $I_{oe}(x, y)$ be the background and object edge images with coordinates $(x, y)$ respectively. Then, the edge



(a) Greyscale Image ($\mathbf{I_B}$) (b) Sobel Edge Image ($I_{be}$) (c) Edge Difference Image ($I_{ed}$)



(d) Greyscale Image ($\mathbf{I_O}$) (e) Sobel Edge Image ($I_{oe}$) (f) Noise Filtered $I_{ed}$

Figure 2.8: Edge Difference Image

difference image $I_{ed}(x, y)$ is given by

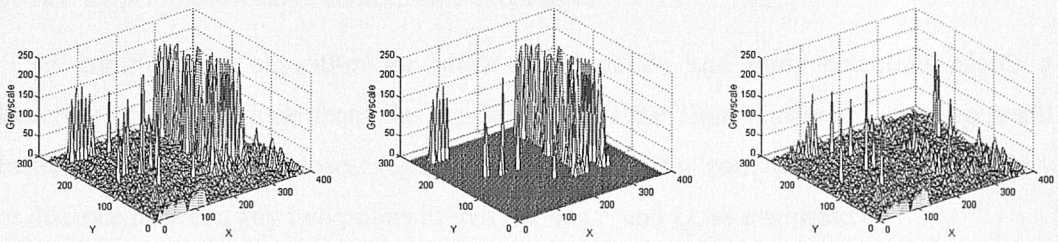
$$I_{ed}(x,y) = \left| I_{be}(x,y) - I_{oe}(x,y) \right|. \tag{2.2}$$

Figure 2.8(c) shows the edge difference image obtained by the Equation 2.2. As can be seen in Figure 2.8, the edge difference image is extracted from real scene images with 384×288 greyscale pixels acquired from an indoor scene of the early SOTON data. In Figure 2.8(c), the edge difference image still has many small isolated areas that are caused by change in illumination. These isolated areas require a more sophisticated algorithm for object detection and segmentation.

By considering these isolated areas as noise, a 3×3 mask operation can remove the noise areas by suppressing small areas of fixed size as

$$g(x,y) = \begin{cases} g(x,y) & \text{if } \left( \sum_{k=-1}^{1} \sum_{j=-1}^{1} f(x-j, y-k) \Big|_{(j,k) \neq (x-j, y-k)} > 2 \right) \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

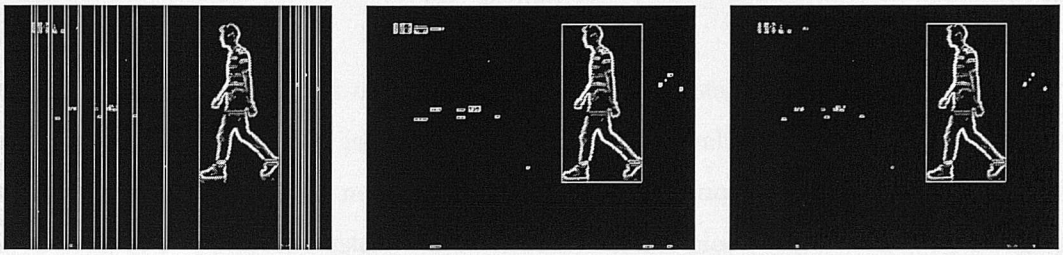
where $f(x, y)$ is set to *1* if it is greater than a threshold value $T_n$, otherwise it is set to *0*. Figure 2.9(a) shows the distribution for greyscale of the edge difference image of Figure 2.8(c), and Figure 2.9(b) is a result of performing the mask operation of Equation 2.3. As can be seen in Figures 2.8(f) and 2.9(b), the mask operation can remove most small isolated areas.




(a) Edge Difference Image (b) Noise Filtered $I_{ed}$ (c) Noise Removed from (a)

Figure 2.9: Distribution for Greyscale of the Edge Difference Image

Also, histogram projection profiles are analyzed to estimate the position of a human body in the edge difference image. The horizontal projection $H(x)$ and vertical projection $V(y)$ in the $n \times m$ image are given by the Equations 2.4 and 2.5



(a) Preliminary Line Segments (b) Candidate Body Regions (c) Body Detection

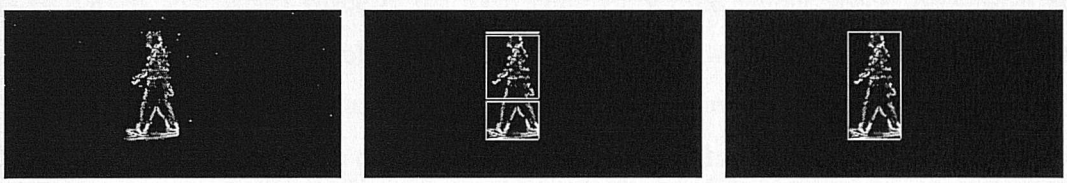
Figure 2.10: Block Segmentation for Human Body Detection

$$H(x) = \sum_{x=0}^{n} O(x, y) \tag{2.4}$$

$$V(y) = \sum_{y=0}^{m} O(x, y) \tag{2.5}$$

where $O(x, y)$ is a pixel value of image with coordinate $(x, y)$. For block segmentation, preliminary lines are defined by a vertical projection profile. After that, to decide if multiple object lines exist in a preliminary line, a horizontal projection profile for each preliminary line is generated. Figure 2.10(a) shows the preliminary segmentation result, and Figure 2.10(b) shows the several candidate body regions that are segmented by connected components included in the horizontal projection on the preliminary lines which means $H(x) \cup V(y)$. Figure 2.10(c) shows the detected human body region that is verified by prior knowledge such as size and shape.

In addition, the algorithm for block segmentation and body detection adopts a clustering procedure done from the single object to the final cluster by merging small clusters. That is, the two closest regions among all possible pairs of regions are found by the distance between any two points in two regions $P$ and $Q$, as evaluated by



(a) Edge Difference Image (b) Fragmented Body Region (c) Merging the Clusters

Figure 2.11: Block Segmentation by Merging the Clusters

$$D_{PQ} = \min_{p \in P, q \in Q} (d_{pq}) \tag{2.6}$$

where $p$ and $q$ are elements of regions, and $d_{pq}$ is the distance between them. The two regions are merged, if $D_{PQ}$ is less than pre-determined value $T_c$. Region size does not affect the distances in this method, and block segmentation is completed by repeated clustering of two closest clusters. Figure 2.11 shows the block segmentation by merging small fragments in the UCSD data. This method can lead to improvement in object segmentation in a low-quality image [114].

### *2.3.2 Detecting a Body in Colour Images*

On the SOTON database, the chroma-key laboratory was used to allow controlled lighting conditions for the indoor data. Due to the nature of both the capture and colour data of the gait database, the use of a colour specific extraction was possible. With the SOTON database, bright green was used as the backdrop colour, and video cameras are generally more sensitive in the green channel. Thus, human body extraction from the image sequences can be easily achieved through background subtraction [105]. The background subtraction based on colour or intensity is a commonly used technique to promptly identify foreground elements. Typical problems in background subtraction include foreground objects with similar colours to the background, and shadows or other variable lighting conditions [42]. Especially, background luminance variations are mainly due to noise and illumination change in indoor sequences. Here, we describe a human body segmentation method based on background estimation using colour information in modified HLS space.

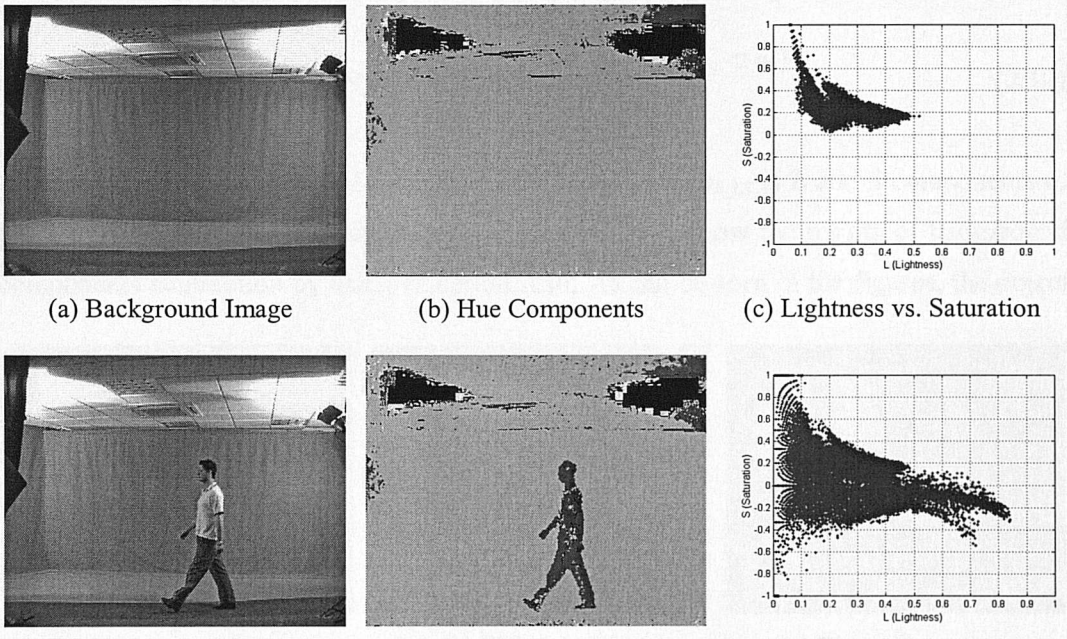
In computer vision and graphics, many different colour models exist, and each model uses its own 3D coordinate system to identify uniquely individual colours [98]. The RGB (Red, Green, and Blue) space is the most commonly used colour space, because it is directly supported by most colour displays and scanners. The HSV (Hue, Saturation, and Value) and HLS (Hue, Lightness, and Saturation) colour models are transformations of RGB space that can describe colours in terms more natural to an artist [33]. To estimate the background of a gait image, the $H$ component is defined as hue value in HSV space, and $L$ and $S$ components are defined as Lightness and Saturation values in modified HLS

space. The conversion method from RGB colour space to modified HLS space is given by

$$
\begin{aligned}
m &= \min(g,b) \\
H &= \begin{cases} -1\ (undefined) & if\ r = m \\ 60 \times (b-r)/(m-r) & if\ g = m \\ 60 \times (r-g)/(m-r) & if\ b = m \\ H + 360 & if\ H < 0 \end{cases} \\
L &= (m+r)/2 \\
S &= \begin{cases} 0 & if\ m = r \\ (m-r)/(m+r) & if\ L \le 0.5 \\ (m-r)/(2-m-r) & otherwise \end{cases}
\end{aligned}
\tag{2.7}
$$

where $r$, $g$, and $b$ are defined as normalized RGB ranges from *0* to *1*. The modified HLS space is very similar to original HLS space but slightly more robust to noise.

In the chroma-key laboratory, subject is captured against a uniform background. To estimate the background, RGB colour space is converted to HSV space, and hue




(a) Background Image (b) Hue Components (c) Lightness vs. Saturation

(d) Object Image (e) Hue Components (f) Lightness vs. Saturation

Figure 2.12: Colour Components for Background Estimation

component of gait images is shown in Figures 2.12(b) and (e). As can be seen in the figures, hue component in the background is very uniform, thus the mode of hue component is simply calculated by

$$\xi = \underset{x \in m, y \in n}{\text{mode}}(h_{x,y}) \tag{2.8}$$

where $h_{x,y}$ is the hue value at coordinate $(x, y)$ in $m \times n$ image region. In this data, HSV space offered a better estimation of hue component $i$. The RGB colour space is also converted to the modified HLS space. The range of lightness and saturation components in the HLS space is defined as

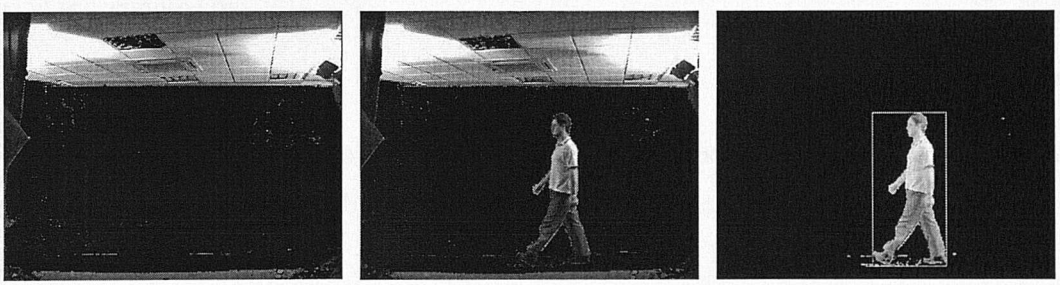
$$\psi = [\sigma L(i),\ \sigma S(i)] \tag{2.9}$$

where $\sigma$ is a variance of each colour component. The background feature ($\psi$) can be calculated by colour clustering method [20] using lightness ($L$) and saturation ($S$) components. Figures 2.12(c) and (f) show the features of background and object images in rectangular coordinate system.

Now, the background features can be removed by using Equations 2.8 and 2.9, namely, the pixel values in HLS space can be re-defined by background estimation as

$$H, L, S = \begin{cases} 0 & if\ (p_{k\_H} \subset \xi) \wedge (p_{k\_LS} \subset \psi) \\ p_k & otherwise \end{cases} \tag{2.10}$$

where $p_{k\_H}$ is $H$ component of pixel $k$ in HSV space, and $p_{k\_LS}$ is $L$ and $S$ components of pixel $k$ in modified HLS space. Figures 2.13(a) and (b) show the results of background components subtraction by using Equation 2.10. As can be seen in the figures, the object



(a) Background Image (b) Object Image (c) Object Detection

Figure 2.13: Background Subtraction and Object Detection

image still has some noise (background components), and the object has also lost some foreground components. Here, to remove this noise and to recover the lost components, noise filtering and histogram logarithm methods are applied to the background subtracted image. The noise filter is defined by Equation 2.3, and histogram logarithm is defined as

$$p'_k = c \times \log(1 + p_k) \tag{2.11}$$

where $p_k$ is pixel value at index $k$, and $c$ is constant. The histogram logarithm increases the dynamic range of greyscale via contrast stretching and is useful to enhance detail in the darker region of the image. However, the object can be detected by the histogram projection method described in the previous section for greyscale images. Figure 2.13(c) shows the detected object by using projection profile in the noise filtered and contrast enhanced object image.

### *2.3.3 Extracting the Body Contour*

To extract the contour of a detected human body, a thresholding and morphological method is used here. Figure 2.14 shows the procedure of thresholding and extracting a human body contour. Thresholding is one of the most important approaches in the field of image segmentation, and choosing a correct threshold is difficult under irregular illumination. However, using the background information in an image can lead to improvement. Accordingly, a thresholding method based on similarity (or dissimilarity) measures between the background and the object image is used. Let $I_b(x, y)$ and $I_o(x, y)$ be the feature (or brightness in greyscale image) of a pixel with coordinate $(x, y)$ in the background image ($I_b$) and the object image ($I_o$). Then, the similarity $\Theta(x, y)$ at coordinate $(x, y)$ is computed by

$$\Theta(x,y) = \left| I_b(x,y) - I_o(x,y) \right|. \tag{2.12}$$

Similarity values close to zero imply a high probability of being background; conversely, large values of similarity imply high probability as an object. Therefore, the binary image $I_{bi}(x, y)$ is thresholded as

$$I_{bi}(x,y) = \begin{cases} 1 & \text{if } (\Theta(x,y) > \tau) \wedge (I_o(x,y) > \lambda) \\ 0 & \text{otherwise} \end{cases} \tag{2.13}$$

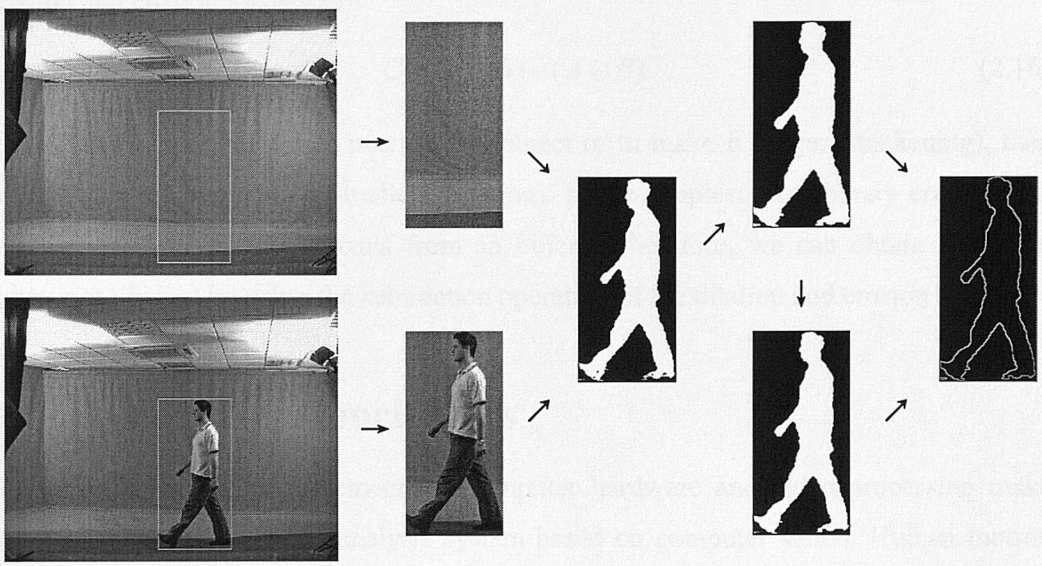


Figure 2.14: Extraction of Human Body Contour

where $\tau$ is the tolerance value, and $\lambda$ denotes the global threshold value. Theoretically, thresholding is a simple image segmentation method, which is very effective and useful for small and low-resolution images, but suffers from difficulty with change in illumination. To improve this method, a more effective algorithm using the probability density of the similarity for determining appropriate values of $\tau$ and $\lambda$ is required.

On the other hand, the binary image can have some noise inside the object which is actually a human body part and some noise outside the object. So, morphological filtering is used to remove the noise and to extract the human body contour, by the dilation and erosion. In mathematical terms, the dilation of a set $A$ by a structuring element $B$ is denoted by $A \oplus B$ and is defined as

$$A \oplus B = \{x \mid B_x \cap A \neq \phi\} \tag{2.14}$$

where $A$ represents the image being operated on, and $B$ is a second set of pixels, a shape that operates on the pixels of $A$ to produce the result. The erosion of image $A$ by a structuring element $B$ can be defined as

$$A \ominus B = \{x(B)_x \subseteq A\}. \tag{2.15}$$

Finally, the human body contour is determined by arithmetic subtraction between the

dilation and erosion images as

$$C = (A \oplus B) - (A \ominus B). \tag{2.16}$$

If dilation can be said to add pixels to an object or to make it bigger (thickening), then erosion will make an image smaller (thinning). In the simplest case, binary erosion will remove the outer layer of pixels from an object. Therefore, we can obtain an object contour easily just by using the subtraction operation of the dilation and erosion image.

## 2.4 Results and Conclusions

The recent technical improvements of computer hardware and video processing make possible marker-less motion analysis system based on computer vision. Human motion can be detected and measured from video cameras at a distance, and the marker-less analysis is essential to enable greater application capability. The study of human motion analysis is related to several research areas of computer vision such as the motion capture, detection, tracking and segmentation of people, and more generally, the understanding of human activity, from image sequences involving humans. Here, the large amount of human gait data, which was collected from DV cameras, is pre-processed, and the human body and its contour is extracted from the image sequences of the SOTON gait database. The success and potential of a new application relies largely on the database used for evaluating the application systems.

### *2.4.1 Experimental Results*

As described in Section 2.2.3, seven indoor image sequences of each of the 100 different subjects are used in the experiments. The detection and extraction of the human body is accomplished by background subtraction and by histogram projection analysis, and thresholding and morphology is then used to extract the contour of a detected human body. The size of the human body in the image sequences is approximately 160×360 pixels in 720×576 image. Figures 2.15 and 2.16 show extracted human body contours during one full stride (or two steps) from an image sequence of the SOTON database. The human body contour is extracted with origin of the segmented body region, and the extracted body contours are visually inspected and graded. The average quality levels of

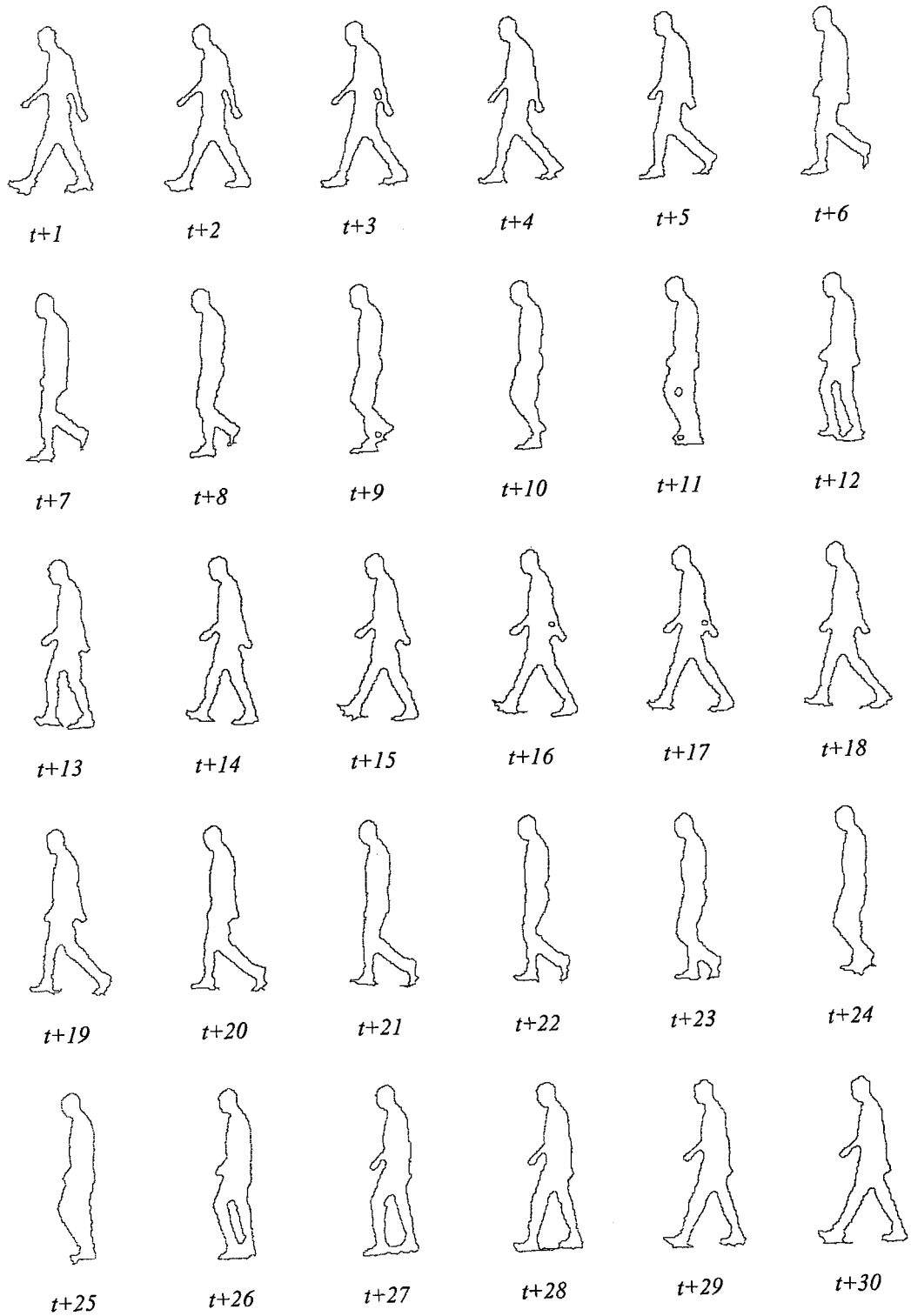



Figure 2.15: Extracted Human Body Contours (Male, Grade A)

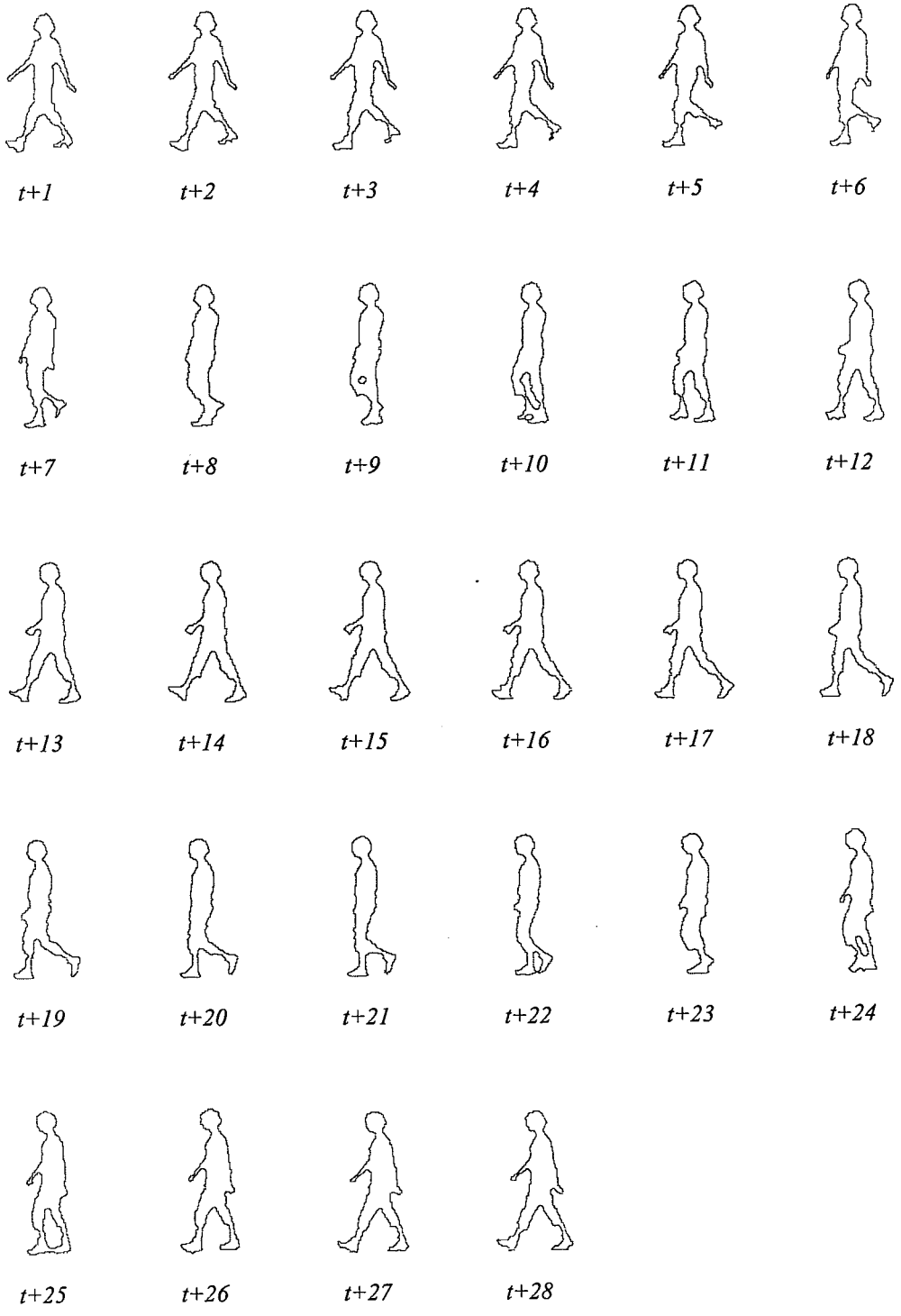



Figure 2.16: Extracted Human Body Contours (Female, Grade C)

the contour data of each subject can be evaluated by comparison with the subject in original images. Here, the quality level is graded as *A* (good), *B* (fair), and *C* (bad) to be 30%, 40%, and 30% of subjects, respectively, and the noise in Figure 2.16 is caused by clothing and apparel. The noise, which is related to quality of the contour data, is mostly caused by the shadow and colour of shoes at ground. The shoes can appear to change in colour due to their reflectance of the walking surface. However, our approach appears to extract the human body contours from image sequences of the SOTON database successfully. Here, there is no ground truth assessment available in this scenario. Meanwhile, the video image analysis software, which is developed for gait motion analysis, has been implemented in Java on a Pentium III 800MHz system. The video image analysis software includes many image processing and motion analysis functions with a GUI environment.

### *2.4.2 Conclusions*

The SOTON database is currently the largest gait database filmed under laboratory conditions. The subjects are filmed in front of a green background, thus human body extraction from the image sequences is easily achieved through background subtraction. To estimate the position of a human body, the histogram projection profiles are analyzed, and the body region is verified by prior knowledge such as size and shape. Also, thresholding method based on similarity measures between the background and the object image is used. The body contour is extracted by subtraction followed by dilation and erosion. However, only the indoor database is used in this work, although the subjects of the SOTON databases were filmed indoors and outdoors. Therefore, the usefulness of the pre-processing methods, which are developed in this study, is not demonstrated in outdoor applications. The quality of pre-processed data filmed under the special laboratory should be better than that of the outdoors, moreover it can be well influenced to results of motion analysis and recognition. Notwithstanding this, the capability of marker-less gait analysis and recognition must be still very attractive technique.

# Chapter 3

# Extracting Human Gait Motion

## 3.1 Describing Human Gait Motion

Human gait (walking and running) is the most common and complex form of all human activities [117][119]. It has been studied by the scientists and artists for a very long time. However, gait has only been quantified very recently, and has been described and analyzed more than any other total movement. The gait motion can be defined as a form of locomotion in which the body's centre of gravity moves alternately on the right side and left side [57][88][118]. It requires the simultaneous involvement of all lower limb joints in a complex pattern of movement. Also, human gait has common patterns of movements and describes a rhythmic and periodic motion by which the body moves step by step in the required direction. A period of the gait cycle exists between the successive heel-strikes, and the gait motion in space and time satisfies spatial and temporal symmetry. Here, gait is described as periodic motion, and the gait cycle is detected by symmetry analysis of human gait.

### *3.1.1 Human Gait as a Periodic Motion*

Human gait is a complex integrated activity with various factors interacting at the same time. It is described by a process of locomotion in which the moving body is supported by first one leg and then other. As the moving body passes over the supporting leg, the other leg is swung forward in preparation for its next stance phase [57][88]. The time interval between two successive occurrences of one of the repetitive events of gait motion is

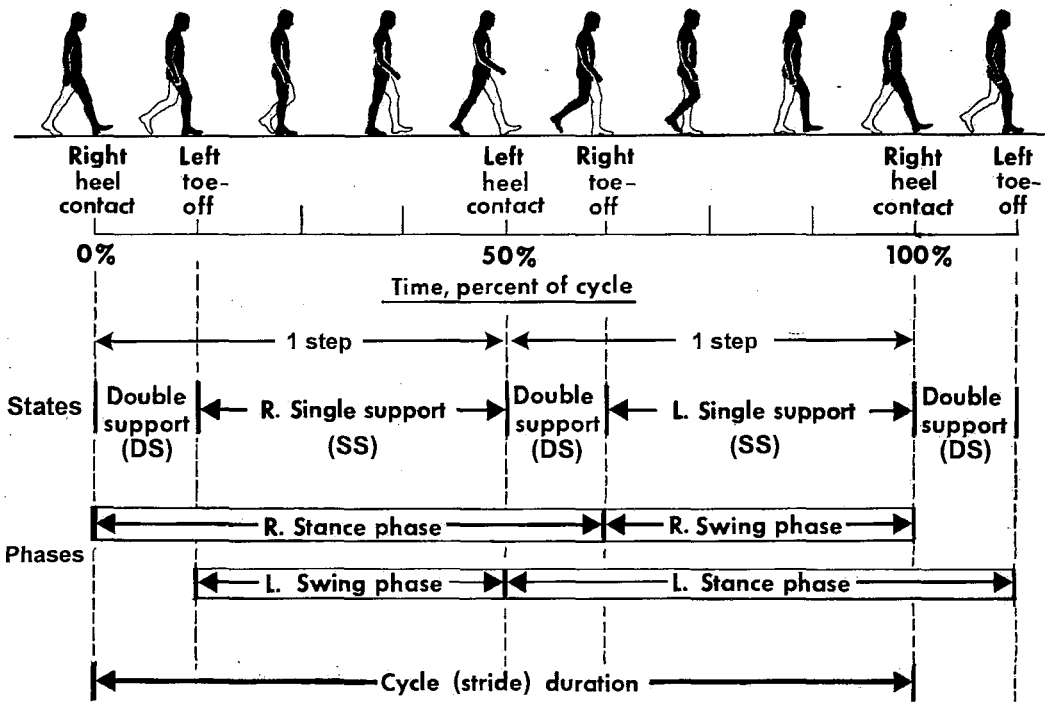



Figure 3.1: Division of the Human Gait Cycle [57]

defined as the gait cycle [117]. Each gait cycle is separated into two distinct periods: the stance and the swing phase. The stance phase is the entire period during which the foot is on the ground, and the swing phase begins as the foot is lifted from floor (toe-off). The usual distribution of the floor contact period is about 60% for stance and about 40% for swing [92]. This phase varies with the speed of gait motion, the swing phase becoming proportionately longer, and the stance phase and double support phases shorter, as the speed increases [80]. The phases of a human gait cycle are described in Figure 3.1.

Early in the studies of gait analysis, researchers recognized that each pattern of motion related to a different functional demand and designated them as the phases of gait [92]. The phases of gait provide a means for correlating the simultaneous action of the individual joints into patterns of total limb function. In general, the stance and swing phases have been divided into eight sub-phases: initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing, and terminal swing [92][117]. Each of the eight gait phases has a functional objective and a different length time interval in gait cycle. Also, the changes in joint motion that occur during each phase

are described by the joint ranges of motion at each segment of gait. Time and distance are basic parameters of motion, and measurements of these variables provide a basic description of gait. These variables provide essential quantitative information about a person's gait [88].

On the other hand, the human body functionally divides into the passenger and locomotor units during walking. The upper body is a relatively passive passenger unit that rides on a locomotor unit. The two lower limbs and pelvis are the anatomical segments that form the locomotor unit. The upper body part above the lower limbs is usually represented by the Head, Arms and Trunk (HAT), and it has a large and heavy mass that represents about 70% of body weight. The aspects of the HAT trajectory have two cycles of upward and downward displacement in each stride [88][92]. The lowest position is reached during the middle of each double support phase, and peak upward deviation occurs in mid-stance of each limb. The amount of displacement varies with the subject's walking speed [119].

### *3.1.2 Detecting Human Gait Cycle*

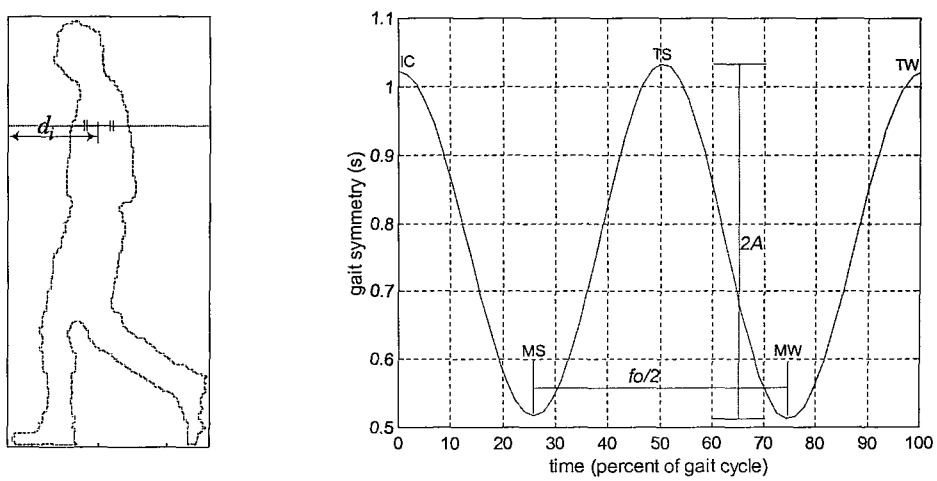
Detecting the period for motions that repeat regularly can produce important information about underlying object or scene properties. In general, human gait decomposes into a repetitive motion and a net translatable component [103]. Several vision based methods have been developed from this to compute the period of human gait from image features [24][73][96][103]. To detect the period of gait cycle, the width of the block segmentation of the corresponding body region, $w(t)$, can be used as a simple method. That is, in each gait cycle there are two periods of double support and two periods of single support. When the person walks parallel to the camera, gait appears bilateral-symmetric and two peaks in $w(t)$ for each gait period are observed. This method is computationally efficient and has proven to work well with background subtraction [7][114]. However, as the camera view point departs from fronto-parallel, one of these two peaks decreases in amplitude with respect to the other, and eventually becomes indistinguishable from noise.

According to biomechanical analysis, large segments such as the HAT must not be only balanced on a joint that is moving in space but also transferred from one leg to other [88]. Namely, the passenger unit is moving in both the plane of progression and the

frontal plane as an inverted pendulum which rotates about the hip joint [119]. Therefore, the upper body's speed varies a little, being fastest during the double support phases and slowest in the middle of the stance and swing phases [117]. Also, the centre of mass of the passenger unit will keep the maximum distance from front foot at initial contact (IC), end of terminal stance (TS) or terminal swing (TW), and it has minimum distance from the front foot at end of mid-stance (MS) or mid-swing (MW). Figure 3.2(a) shows a horizontal centre of mass at the chest region which is guided by anatomical data [29]. To detect the gait cycle, the horizontal centre of mass in the passenger unit is considered as a gait symmetry point. This gait symmetry point at image frame $i$ can be defined as

$$s_i = d_i / (\bar{d} + \sigma) \tag{3.1}$$

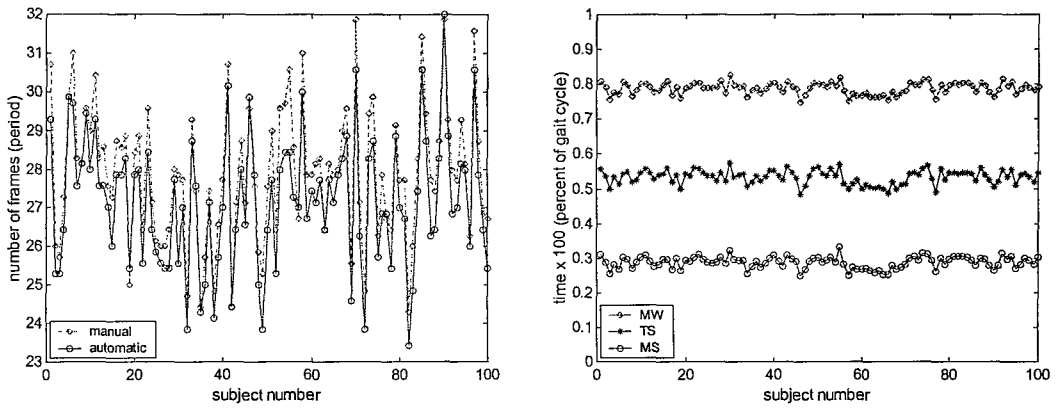
where $\bar{d}$ is the mean value and $\sigma$ is the standard deviation of $d_i$ during a gait cycle. Figure 3.2(b) shows the curves of gait symmetry interpolated by trigonometric-polynomials during a gait cycle. As can be seen in the figure, important gait phases (IC, MS, TS, MW, TW) can be detected from the peaks and troughs.




(a) Centre of Mass (b) Gait Symmetry Analysis

Figure 3.2: Detecting Human Gait Cycle

Also, Figure 3.3(a) shows the average of the detected gait periods derived either by manual labelling or by the automatic method by the symmetry point analysis, from 100 different subjects of the SOTON database. In the figure, the average gait period by manual labelling exceeds the automatically detected periods by 0.74 frames (2.7%). In




(a) Detected Gait Periods (b) Detected Gait Phases

Figure 3.3: Detected Gait Periods and Phases of the 100 Subjects

fact, the gait periods by manual labelling include just over one gait cycle in order to contain certainly one gait cycle. The detected gait phases from the 100 different subjects during one gait cycle are shown in Figure 3.3(b). The average finishing points of mid-stance, terminal stance and mid-swing are observed at the 28.7, 53.3 and 78.7 percentages of gait cycle, respectively. The result has a slight difference (3.1%) compare with medical data [88][92]. Here, some gait features such as gait frequency ($f_0$), amplitude ($A$), and walking velocity ($v$) can be also calculated. Gait cycle analysis determines the frequency and phase of each observed gait sequence, allowing us to perform dynamic time warping to align sequences before matching. Moreover, it provides data reduction by summarizing the sequence with a small number of ideal key-frames [22].

## 3.2 Extracting Human Gait Signatures

Human gait motion is usually described by kinematic motion analysis, and most of the model-based approaches in computer vision are concerned with studies of the kinematic patterns [1][119]. Therefore, a very important aspect in the fundamental study of human gait is the investigation of the kinematics of human body segments [122]. As described in Section 2.1.2, the stick figure model is the most effective and well-defined representation method for kinematic gait analysis, and can be extracted by motion information from human gait. Also, the stick figure is closely related to a joint representation, and the

motion of joints provides a key to motion estimation and recognition of the whole figure [77]. The gait signature can be defined as a sequence of the stick figures obtained from gait silhouette data. Here, the gait signature is extracted from the body contours by determining the body points using the linear regression of body skeleton and motion tracking with topological analysis guide by anatomical knowledge.

### *3.2.1 Body Segment Properties*

The analysis of human motion often requires knowledge of the properties of body segments. The initial interests in the proportions of human beings were presumably motivated by sculptors and prosthesis designers. More recently, interest has arisen from the needs of technological developments such as man-machine interface, computer animation, and motion analysis [94][118]. The dimensions of various body segment-links callipered from cadavers have been extensively studied [27], and an average set of segment lengths represented as a percentage of body height [29]. These proportions of

Table 3.1: Average Link Lengths as Percentage of Stature

| Linkages of Body | Dempster [27] | Drillis [29] | Reynolds [94] | |
|---|---|---|---|---|
| | | | Men | Women |
| 1. Upper arm | 17.35 ± 0.99 | 18.6 | 17.4 | 17.2 |
| 2. Forearm | 15.72 ± 0.52 | 14.6 | 15.6 | 14.9 |
| 3. Hand | 10.54 ± 0.63 | 10.8 | 10.9 | 10.8 |
| 4. Hip-shoulder joint | 27.98 ± 1.95 | 28.8 | 28.8 | 30.4 |
| 5. Thigh | 23.99 ± 1.43 | 24.6 | 24.3 | 24.2 |
| 6. Shin | 25.05 ± 1.22 | 24.5 | 23.6 | 23.0 |
| 7. Ankle above sole | 3.69 ± 0.72 | 3.9 | 4.2 | 4.1 |
| 8. Hip-neck joint | 32.85 | 34.0 | 33.4 | 34.0 |
| 9. Hip-vertex | 47.27 | 47.0 | 47.9 | 48.7 |
| 10. Head | 15.46 ± 1.01 | 13.0 | 14.5 | 14.7 |
| 11. Hip above the SRP | _ | 5.0 | 4.3 | 4.3 |
| 12. Transverse shoulder | 18.34 ± 1.05 | _ | 21.9 | 21.2 |
| 13. Transverse hips | 9.82 ± 1.17 | _ | 9.9 | 10.9 |

body segments are suitable for purposes of a good approximation in the absence of better data [94][118]. Table 3.1 shows the average length of the important linkages of the body expressed as a percentage of stature. In the table, Dempster's work was based on only 8 subjects, and this number was increased to 65 by Reynolds [94].

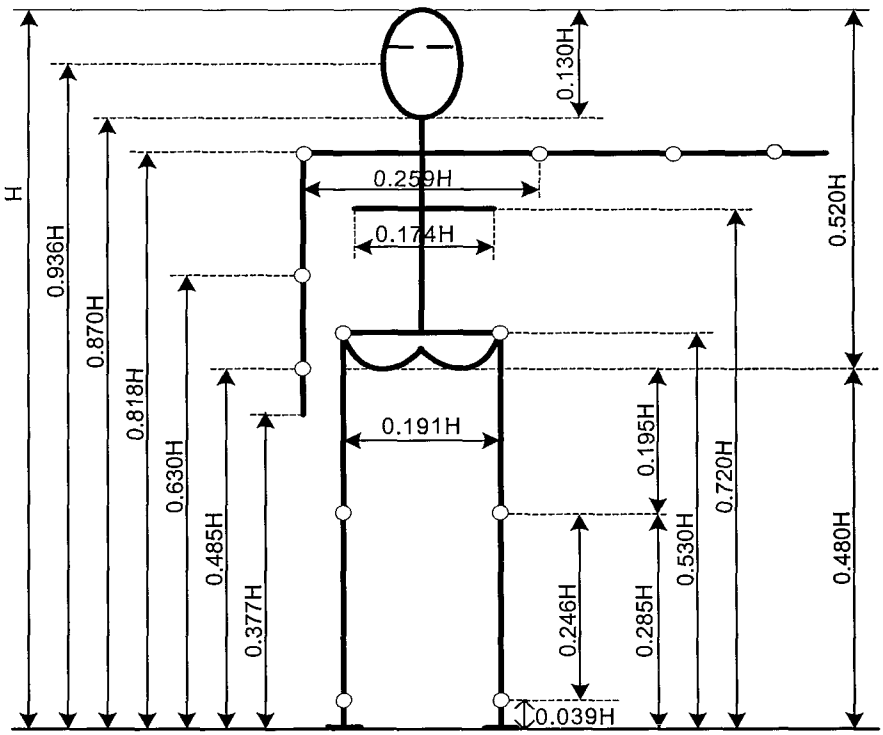



Figure 3.4: Body Segment Properties [118]

In addition, Figure 3.4 shows body segment properties represented as fractions of body height, *H* by Drillis and Contini [29][118]. Their anatomical studies developed a stick figure in which the lengths are shown as a percentage of body height. Most of the results were obtained from only 12 (or less) subjects. The average age of the 12 subjects was approximately 27 years and the average height was approximately 176 *cm* [29]. The mean dimensions of body height can be used to determine topological position of each body part in human figures as such to guide search for the human head and toe, given that humans have a head, neck, shoulder, chest, waist, pelvis, knees, and ankles in known topology. Here, we describe the method for extracting the gait signature by using these body segment properties. The results of their studies have also been used in dynamic models for animating human motion [50].

### *3.2.2 Extracting the Body Angles*

In the previous chapter, the human body and its contour was extracted from the image sequences. To extract body points in a contour image, the skeleton data with body segment properties is used. For a body height $H$, an initial estimate of the vertical position of the neck, shoulder, waist, pelvis, knee and ankle was set by study of anatomical data to be $0.870H$, $0.818H$, $0.530H$, $0.480H$, $0.285H$, and $0.039H$, respectively. The gait skeleton can be simply obtained by two border points of each body segment $p$ with a range constraint as

$$x_{s,p} = (x_{b,p} + x_{e,p})/2 \tag{3.2}$$

where $x_b$ and $x_e$ represent the horizontal position of the begin and the end pixels on the horizontal line respectively. For example, the range constraint for trunk $t$ can be given by mean value of neck's border points (or not labelled, $\varnothing$) as

$$x_{s,t} = \begin{cases} x_{s,t} & if\ (\overline{x}_{b,neck} < x_{s,t} < \overline{x}_{e,neck}) \\ \varnothing & otherwise \end{cases}. \tag{3.3}$$

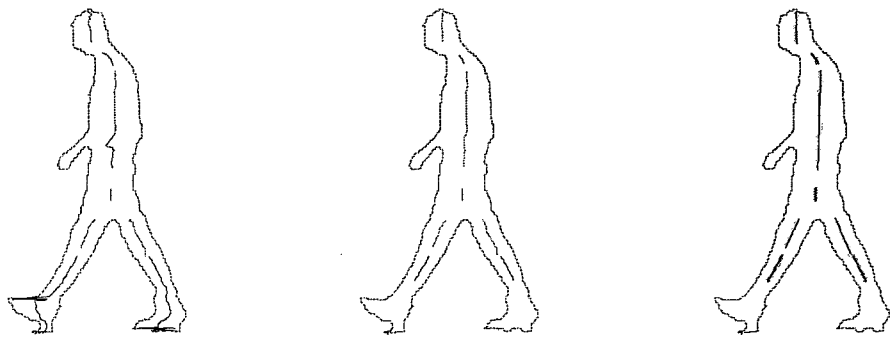
In images, estimation of a primary gait skeleton is highly susceptible to difficulty by the movement of the arm and foot. Therefore, a noise (outlier) removal method, using mean value and standard deviation in each body segment, is employed to select a skeleton point as

$$x'_{s,p} = \begin{cases} x_{s,p} & if\ \left(x_{s,p} \subset \overline{x} \pm \alpha \cdot \sigma\right) \\ \varnothing & otherwise \end{cases} \tag{3.4}$$

where $\sigma$ is standard deviation of $x_{s,p}$ and $\alpha$ is a parameter ($1 \leq \alpha \leq 3$) which depends on the body segments. Figures 3.5(a) and (b) show the gait skeleton and the noise (outliers) removed skeleton data.

Now we can calculate the body angles from the skeleton data by using linear regression. That is, the angle $\theta_p$ of body segment $p$ is approximated by using the slope of the lines in linear regression equation as

$$\theta_p = \tan^{-1}\left(\sum_{i=1}^{n}(x_i - \overline{x})^2 \Big/ \sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x})\right) \tag{3.5}$$



(a) Original Gait Skeleton (b) Outliers Removed (c) Line Fitting

Figure 3.5: Extracting Body Angles at a Key-frame

where $n$ is the number of the skeleton points. Figure 3.5(c) shows the extracted skeleton data with the lines fitted by Equation 3.5. As a certainty factor (CF), the reliability $\gamma_p$ of the angle $\theta_p$ can be also determined by the correlation coefficient as

$$\gamma_p = \left| \sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x}) \Big/ \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (x_i - \bar{x})^2} \right|. \tag{3.6}$$

In practice, linear regression method involves squared error terms, thus it is very sensitive to outliers both in $x$ and $y$ coordinates [34]. The primary outliers can be removed by Equation 3.4, but the substantial errors are still dominated by outliers or misplaced data points. Therefore, the joint angles $\theta_p(k)$ at image frame $k$ are affected by noise in the contour data with outliers. Accordingly, a weighted moving average can be used to reduce the influence of noise as

$$\theta_p(k) = \sum_{i=k-q}^{k+q} (\gamma_p(i) \cdot \theta_p(i)) \Big/ \sum_{i=k-q}^{k+q} \gamma_p(i) \tag{3.7}$$

where $q$ is a size of moving window, here set to *2*. The weighted moving average is an effective method to smooth time series data.

On the other hand, the key-frames (single and double support phases) are determined by gait symmetry analysis during a gait cycle as described in Section 3.1.2. The body angles around double support phases (initial contact, terminal stance, and terminal swing) are clearly extracted, but the angles around single support phases are not as accurate as

those estimated for the double supports. Therefore, a motion tracking method between double supports is used to extract body angles at the lower limbs as well. To track knees and ankles, the left-most skeleton points around the knee region and the right-most skeleton points around the ankle region are considered. Namely, the skeleton points by a size of each body segment are sorted as

$$\boldsymbol{x}_{knee} = \{x \mid x_i < x_{i+1}\} \;\; and \;\; \boldsymbol{x}_{ankle} = \{x \mid x_i > x_{i+1}\}. \tag{3.8}$$

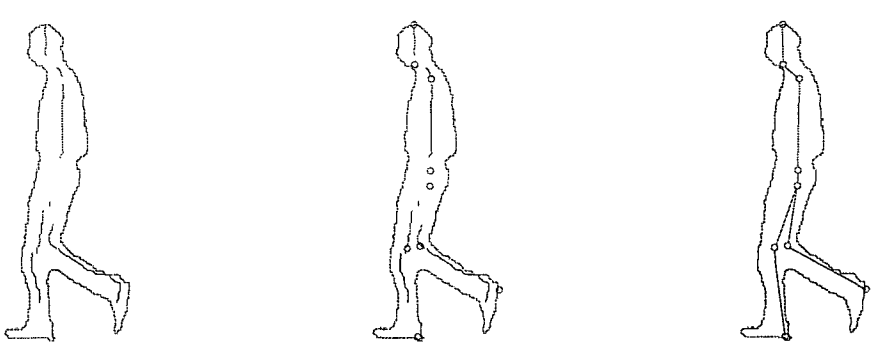
The proper size of body segments is also guided by anatomical knowledge. The knee and ankle have the largest movement of all parts of body. Accordingly, the knee and ankle points for tracking are given by mean value as

$$\boldsymbol{x}_s = \{x_i \mid i = m,...,n\} \;\; and \;\; \bar{x}_s = mean(\boldsymbol{x}_s) \tag{3.9}$$

where $m$ and $n$ are index of the data, here set to *3* and *5* respectively. In a left walking direction, the knee position around single support can be determined by the minimum distance from $x$-axis, and the ankle position can be determined by the maximum distance from $x$-axis.

In normal walking, the knee and ankle positions are moving to the walking direction, hence the forward displacement should be measured as positive value. Finally, the body angles can be calculated by

$$\theta_p = \tan^{-1}\left((x_i - \bar{x}_s)/(y_i - \bar{y}_s)\right) \tag{3.10}$$



(a) Gait Skeleton (b) Tracking Body Points (c) Slope of the Lines

Figure 3.6: Extracting Body Angles at a Inter-frame

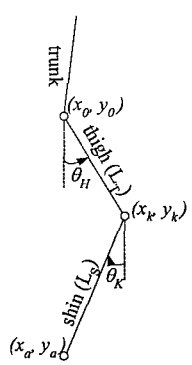
where $x_i$ and $y_i$ are the coordinates of a previously established position such as pelvis and knee. Figures 3.6(a) and (b) show the noise removed skeleton data by Equation 3.4 and the tracking points determined by Equations 3.8 and 3.9, and Figure 3.6(c) shows slope of the lines for computing body angles. As can be seen in the figures, outliers still remain around the rear ankle, but the tracking point is extracted successfully.

### *3.2.3 Gait Signature by Body Points*

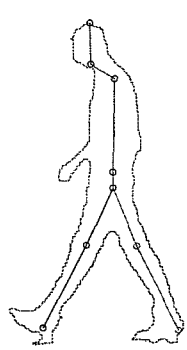
The human gait is typically represented by the movements of the torso and legs, so a stick figure can be used to estimate a human body as a combination of line segments linked by joints. Therefore, the stick figure is obtained from the body points, which are defined by joint angles and body segments. As described in Section 3.2.2, the joint angles are extracted by linear regression analysis and tracking moving points with topological analysis. Also, the length of body segments is determined by the fraction of the body height based on anatomical data. Accordingly, each body point (joint position) can be calculated by using the joint angle and the size of body segments as

$$x_p, y_p = \left[ x_i + L_p \cos(\phi + \theta_p) \quad y_i + L_p \sin(\phi - \theta_p) \right] \tag{3.11}$$

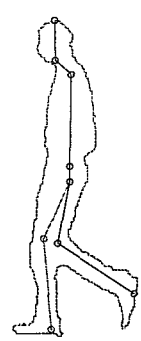
where $\phi$ is the phase shift, $x_i$ and $y_i$ are the coordinates of a previously established position, and $L_p$ is the length of body segments guided by anatomical knowledge. Figure 3.7(a) shows the leg points (knee and ankle) by the joint angles with the segments.




(a) Leg Points by Joints



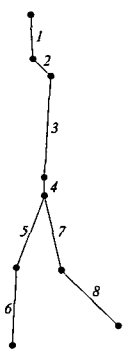
(b) Stick Figure for Figure 3.5



(c) Stick Figure for Figure 3.6

Figure 3.7: Body Points and Stick Figure

There are nine joint coordinates (body points) determined by Equation 3.11 in human body: $(x_{head}, y_{head})$, $(x_{neck}, y_{neck})$, $(x_{shoulder}, y_{shoulder})$, $(x_{waist}, y_{waist})$, $(x_{pelvis}, y_{pelvis})$, $(x_{knee1}, y_{knee1})$, $(x_{knee2}, y_{knee2})$, $(x_{ankle1}, y_{ankle1})$, and $(x_{ankle2}, y_{ankle2})$. The 2D stick figures with the nine body points are shown in Figures 3.7(b) and (c). As can be seen in Figure 3.7, the body points (coordinates) are derived from the joint angles and body segments, and the stick figure is simply obtained by connecting the nine body points. In the figure, the estimate of the joint angle and segment of neck is not very accurate, because the neck position highly depends on head and shoulder position. On the other hand, the legs' motion is extracted very successfully. However, the gait signature is represented by a sequence of the simplified stick figure with 8 sticks and 6 joint angles, and gait motion can be describing the motion in a compact form as sequence of the joint parameters. Namely, each gait signature can be characterized by the body segments (the sticks) and joint angles.




(a) Stick Figure

| *Stick #* | *Body Segments* | *Stick #* | *Body Segments* |
|---|---|---|---|
| *Stick 1* | *head* | *Stick 5* | *left thigh* |
| *Stick 2* | *neck* | *Stick 6* | *left shin* |
| *Stick 3* | *trunk* | *Stick 7* | *right thigh* |
| *Stick 4* | *pelvis* | *Stick 8* | *right shin* |

(b) Sticks vs. Body Segments

Figure 3.8: Simplified Stick Figure Model

The mathematical description of the stick figure requires only a gait direction and nine $(x, y)$ coordinates for body points. That is, the stick figure model at image frame $k$ can be defined by a vector $\boldsymbol{S}_k$ as

$$\boldsymbol{S}_k = \left\{ d_g, x_p, y_p \,|\, p \in head, neck, shld, waist, pelvis, kneeL, kneeR, ankL, ankR \right\} \quad (3.12)$$

where $d_g$ is the gait direction, and is set to *0* or *1* for left or right direction, respectively. Figure 3.8 shows the stick figure model for representing the human body structure by gait signature. A sequence of the stick figures is analyzed as a gait sequence, and this result can be used to enhance the gait signatures.

## 3.3 Structure of the Gait Sequences

Human activities and behaviours may be considered as a stochastically predictable sequence of states [58], and they can be interpreted as a grammatical structure [4]. Thus, the structure of the motion sequences is generally described by a context-free grammar. The grammar and parsing mechanisms provide longer range temporal constraints and allow the inclusion of prior knowledge about structure of temporal events in human motion [58]. However, human gait is an example of highly constrained motion, and physical constraints induce a rhythmic and repetitive pattern of motion [79]. Accordingly, human gait motion can be described as patterns with constrained structural relations. Here, the structure of human gait motion is analyzed grammatically, with physical constraints. The practical reason for analyzing the constraints and grammatical structure of the gait sequences is to improve the robustness of the gait signatures.

### *3.3.1 Concept of Gait Constraints*

Human gait motion has many constraints, which are based on the laws of physics and the functional rules. Some of the constraints are low level, such as the limited range of motion, and the high level are contextual constraints such as that one foot is always on the ground during walking or the body should be in dynamic balance [57][119]. Basically, the constraints represent assumptions about the motion of model, thus the gait motion needs to satisfy the constraints on a movement's trajectory, speed, and energy expenditure. As a low level constraint, human gait has been approximated as the normal range of motion in the joints, and it can be directly detected in the gait signatures if a joint is moved out of the normal range. The Ranges Of Motion (ROM) of the joints are presented in Table 3.2, and the degrees of motion are the values achieved at the end of each phase [88][92]. The approximate total ROM needed for normal gait at the value of the joint angle at each joint throughout the gait cycle.

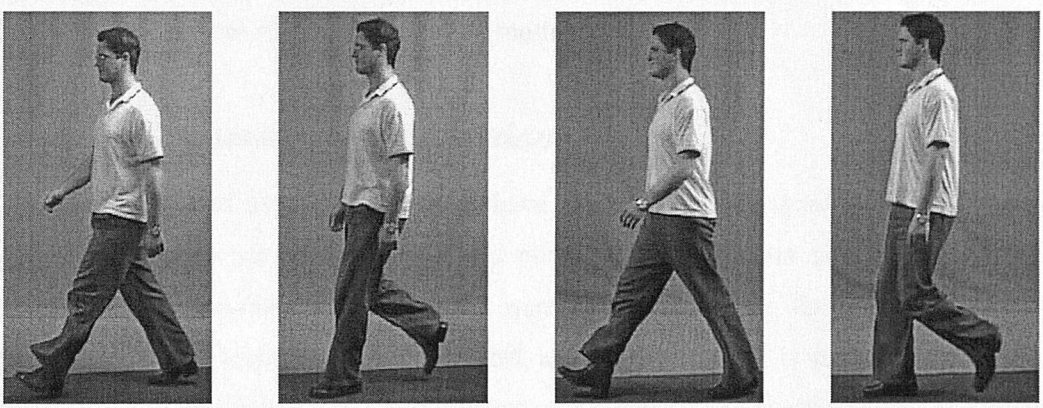
In addition, one of the unique characteristics of the gait motion is bilateral symmetry, which means that the left and right legs perform the same movements, just shifted in time, or if there exists some reflection that is invariant. That is, when a person walks the left arm and right leg interchange direction of swing with the right arm and left leg, and vice

Table 3.2: Range of Gait Motion (RLA*)

| Gait Cycle (GC) | % of Gait Cycle | Thigh | Knee |
|---|---|---|---|
| Initial Contact (IC) | 0% | 30° of flexion | 0° |
| Loading Response (LR) | 12 % | 25° of flexion | 15° of flexion |
| Mid Stance (MS) | 31 % | 0° | 5° of flexion |
| Terminal Stance (TS) | 50 % | (10°~)20° of hyperextension | 0° |
| Pre Swing (PW) | 62 % | 0° | 35°(~40°) of flexion |
| Initial Swing (IW) | 75 % | 20° of flexion | 60° of flexion |
| Mid Swing (MW) | 82 % | 30° of flexion | 30° of flexion |
| Terminal Swing (TW) | 100 % | 30° of flexion | 0° |

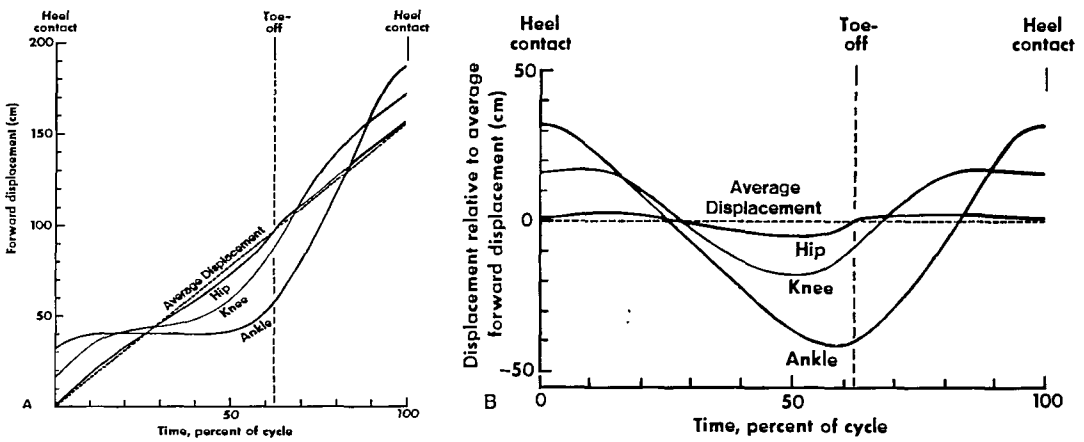
*. Terminology and Data from Gait Laboratory at Rancho Los Amigos (RLA) Medical Centre, California, USA

versa, with half a period phase shift. Figure 3.9 shows an example of the bilateral symmetry in gait motion. In the figure, the second half of the gait cycle shows a reflection of the first half of the gait cycle. This is only a generalization for normal gait, and could be important to a clinical analysis [101]. However, the bilateral symmetry in normal human gait motion satisfies geometrical symmetry but also dynamic symmetry, namely the frequency content of both lower limbs' motion is also similar.



*t+2* *t+6* *t+15* *t+20*

Figure 3.9: Bilateral Symmetry of Gait Motion




(a) Actual Displacements (b) Displacements Relative to Average

Figure 3.10: Forward Displacement at Hip, Knee, and Ankle [57]

On the other hand, a normal gait motion requires continuing ground reaction forces at each step and propelling the body forward in the direction of progression. In normal gait, the horizontal displacement tends to be the same for all parts of the body. Figure 3.10 shows a typical example of the horizontal displacement at the hip, knee and ankle. If there were no variations in velocity, a plot of the horizontal displacement against time would be a straight line as shown by the average displacement line in Figure 3.10 [57]. When the displacement curve of a point is steeper than the average line, that point is moving ahead faster than the average velocity. Conversely, as the velocity of the point becomes less than the average velocity, the slope of the displacement curve becomes less than that of the constant velocity line [57]. These characteristics of the forward displacement can be used as a constraint to reduce noise in gait motion tracking.

### *3.3.2 Step Symmetry of Gait Motion*

Both symmetry and asymmetry of gait have been used to analyze human behaviour [48][101]. Asymmetric gait is frequently considered to indicate gait pathology, and symmetric gait is usually used to analyze normal gait. However, the legs in a compass gait [43] remain straight at a double support, and the step length is similar for both heel strikes. That is, geometrical symmetry exists in a compass gait for different step lengths as shown in Figure 3.11. The angles of the legs measured from the vertical are identical, namely $\theta_1 = \theta_2$, $\theta_1 = \varphi_1$, and $\theta_2 = \varphi_2$ at times $t_1$, $t_2$ and $t_3$. Further, leg angle $\theta_1$, $\theta_2$, $\varphi_1$, and $\varphi_2$

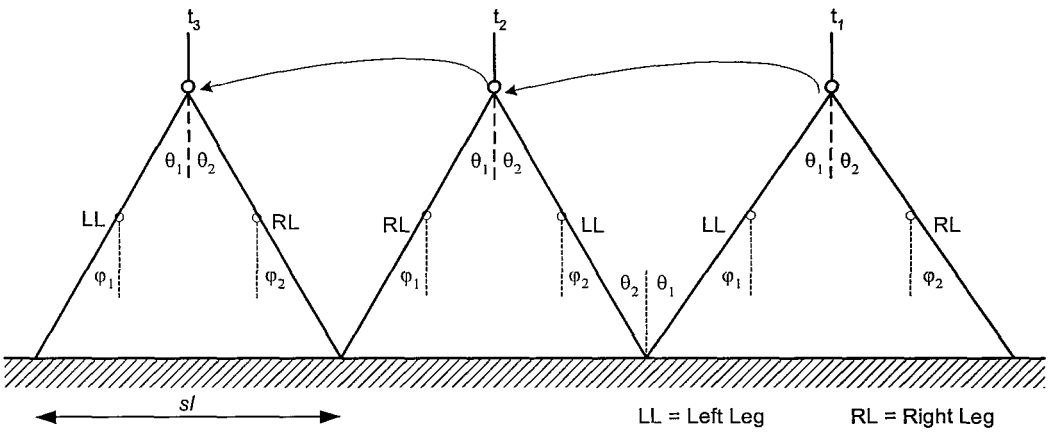



Figure 3.11: Symmetry of Compass Gait for Different Step Lengths [13]

depend only on the step length *sl*. Most importantly, this remains true when the body is accelerating or decelerating, depicted in Figure 3.11 by decelerated step length at time $t_2$ (i.e. the body decelerated from $t_1$ to $t_2$) [13].

However, the step symmetry at double supports $t_1$, $t_2$, and $t_3$ can be described in terms of leg angle $\theta_1$, $\theta_2$, $\varphi_1$, and $\varphi_2$ as

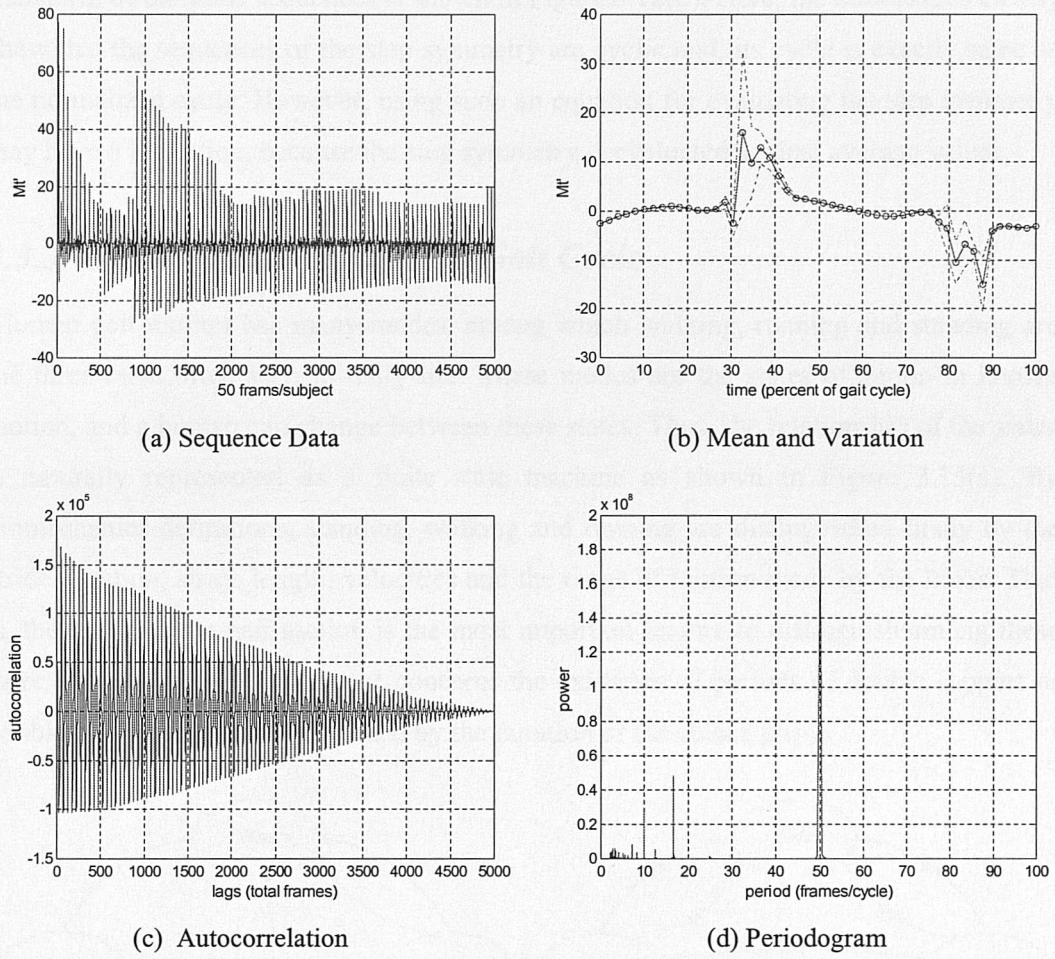
$$M = \theta_1/\varphi_1 - \theta_2/\varphi_2 . \tag{3.13}$$

As described in the previous section, an invariant reflection exists in human gait cycle. Thus, the summation or average of each leg's joint angles during a complete gait cycle should be had equal value to both legs. Accordingly, Equation 3.13 can be extended as Equation 3.14 and can be implemented by Equation 3.15.

$$M_t = \left|\frac{\theta_{1,t} - \bar{\theta}_1}{\varphi_{1,t} - \bar{\varphi}_1}\right| - \left|\frac{\theta_{2,t} - \bar{\theta}_2}{\varphi_{2,t} - \bar{\varphi}_2}\right| \tag{3.14}$$

$$M_t' = \left|\theta_{1,t} - \bar{\theta}_1\right| \times \left|\varphi_{2,t} - \bar{\varphi}_2\right| - \left|\theta_{2,t} - \bar{\theta}_2\right| \times \left|\varphi_{1,t} - \bar{\varphi}_1\right| \tag{3.15}$$

The values of $M_t'$ are equal to *0* around double supports. Now the step asymmetry can be measured by this equation, and symmetry values close to zero imply a high degree of symmetry. Also, the step symmetry shows the degree of weight balance between left and right body parts. Moreover, it is affected by ground slope, because the ground slope modifies the inference of gravity on human body. This equation is also an important




(a) Sequence Data

(b) Mean and Variation

(c) Autocorrelation

(d) Periodogram

Figure 3.12: Analysis of Step Symmetry

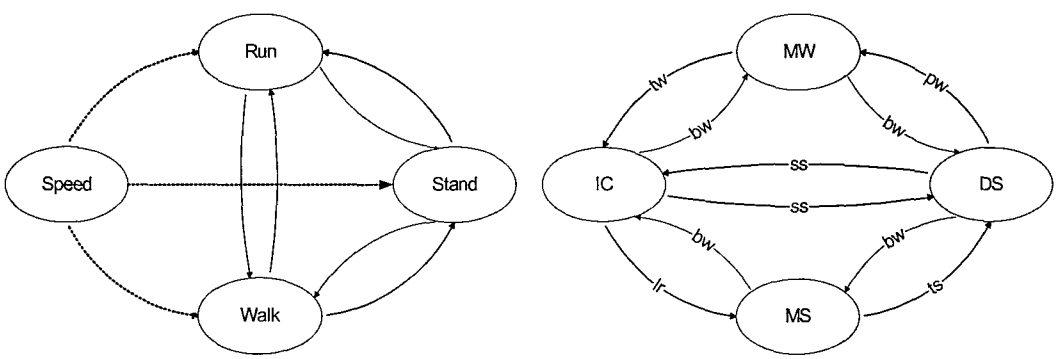
condition for a locomotion step, thus we can use it as the step constraints.

To evaluate the step symmetry, the gait sequences are normalized by 50 frames per sequence, and the step symmetry is calculated by Equation 3.15. After that an average sequence of the step symmetry of each subject is obtained. Figures 3.12(a) and (b) show the sequence data of the step symmetry, and the mean and variation of the data, respectively. These results concern 100 different subjects with seven sequences of each subject, a total of 700 sequences. As can be seen in the figures, the step symmetry values have significantly shown values close to zero around double supports, but the values have high variation around single supports. In addition, the autocorrelation of the sequence data of step symmetry is shown in Figure 3.12(c), and the periodogram of the Fourier

transform of the same sequences is shown in Figure 3.12(d). Here, the both results clearly show that the sequences of the step symmetry are cyclic and, its cycle is exactly same as the normalized cycle. However, using such an equation for evaluating the step symmetry may have a limitation, because the step symmetry is evaluated against average values.

### *3.3.3 Grammatical Analysis of Gait Cycle*

Human gait motion has many modes, among which walking, running and standing are the three most often seen in daily life. These modes are the states of action in human motion, and a human can change between these states. Thus, the relationship of the states is naturally represented as a finite state machine as shown in Figure 3.13(a). By biomechanics definitions, standing, walking and running are distinguished firstly by the stride duration, stride length, velocities and the range of motion made by the limbs. That is, the speed of the gait motion is the most important feature to distinguish among these three states. A second difference concerns the existence of periods of double support or double float, and this is determined by the duration of the stance phase.




(a) Walking, Running and Standing (b) Normal Walking

Figure 3.13: Transition Diagram for Human Gait Model

In general, normal walking is translated to the double support (*DS/IC*) states through the single support (*ss*). A single support in the stance and swing phases contains backward (*bw*), loading response (*lr*), mid-stance (*MS*), terminal stance (*ts*), pre and initial swing (*pw*), mid-swing (*MW*), and terminal swing (*tw*). Figure 3.13(b) shows the finite state diagram within walking mode expanded into their lower level state. This finite state diagram can be represented as context-free grammar parsing mechanism. Figure

G=(Z, V, S, P)

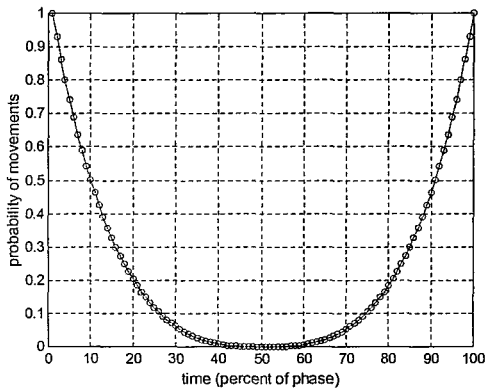
S={IC}

Z={*lr*, *ts*, *pw*, *tw*, *ss*, *bw*}

V={IC, DS, MS, MW}

P={IC→ *ss*DS | *lr*MS | *bw*MW,
DS→*ss*IC | *pw*MW | *bw*MS,
MS→*ts*DS | *bw*IC,
MW→*tw*IC | *bw*DS}

(a) Regular Grammar for Gait Motion

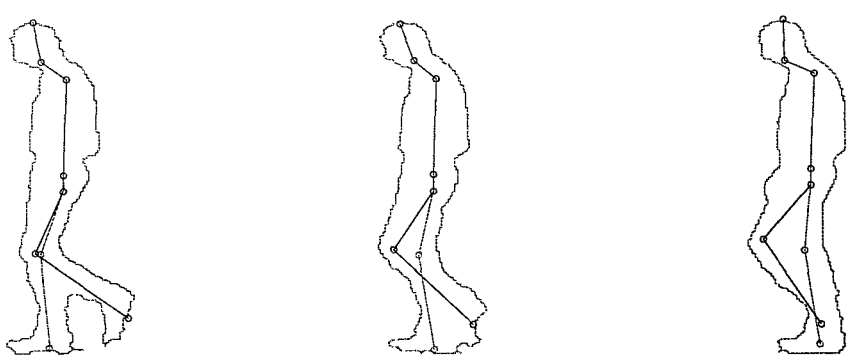



(b) Probabilistic Position of Stationary Foot

Figure 3.14: Grammar and Stochastic for Human Gait Motion

3.14(a) shows the regular grammar for the state diagram in normal walking. The regular grammar is a simple method to describe the human activity by using symbols. However, human movement has been also described by stochastic grammar for understanding human action [58]. In a stochastic grammar, the notion of a production is generalized by the addition of a corresponding probability.

On the other hand, at any point in a gait, the other foot is in contact with the floor (and does not move forwards) during the gait cycle. Thus the shape-of-motion does not look at the full period of motion for a limb, only the forward portion. The cycle of motion analyzed is left foot forward, right foot forward, left foot forward, and so on. This results in a doubling of frequency akin to a full-wave rectifier. Accordingly, we can approximate



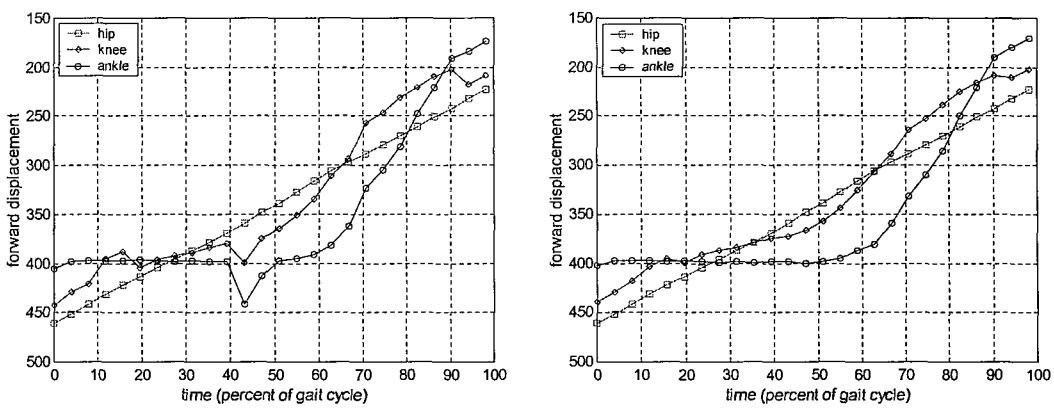
(a) *t-1* (b) *t* (MT/MW) (c) *t+1*

Figure 3.15: Body Points at Crossover of the Legs

the motion of the other foot, which is stationary, as

$$y(t) = |x \cdot \sinh(x^2)|, \quad t = 1,...,n, \quad -1 \le x \le 1 \tag{3.16}$$

where the $x$ is increasing $1/n$ value. The characteristics of Equation 3.16 are shown in Figure 3.14(b). During the gait cycle, the positions of the stationary foot can be estimated by this equation. In addition, the crossover of the two legs is performed on two single support points during one gait cycle. Figure 3.15 shows an example of the body point




(a) Without Grammatical Parsing Method (b) With Grammatical Parsing Method

Figure 3.16: Quality of Gait Signatures in the Forward Displacements

extraction by using the gait constraints with the grammar parsing method. In the figure, the crossover of knee is detected during the three frames of a single support, and ankle crossover is started after a single support. As a result, the forward displacements of the gait signature at hip, knee, and ankle are shown in Figure 3.16. It is an important component for showing quality of the extracted gait signatures. As can be seen in Figure 3.16(a), the quality of the gait signature without the grammatical rules and constraints shows poor result, but this result can be improved by using the grammatical rules as shown in Figure 3.16(b). Using grammatical rules with gait constraints can improve the robustness of the gait signature, and it is especially effective in poor quality data.

## 3.4 Results and Conclusions

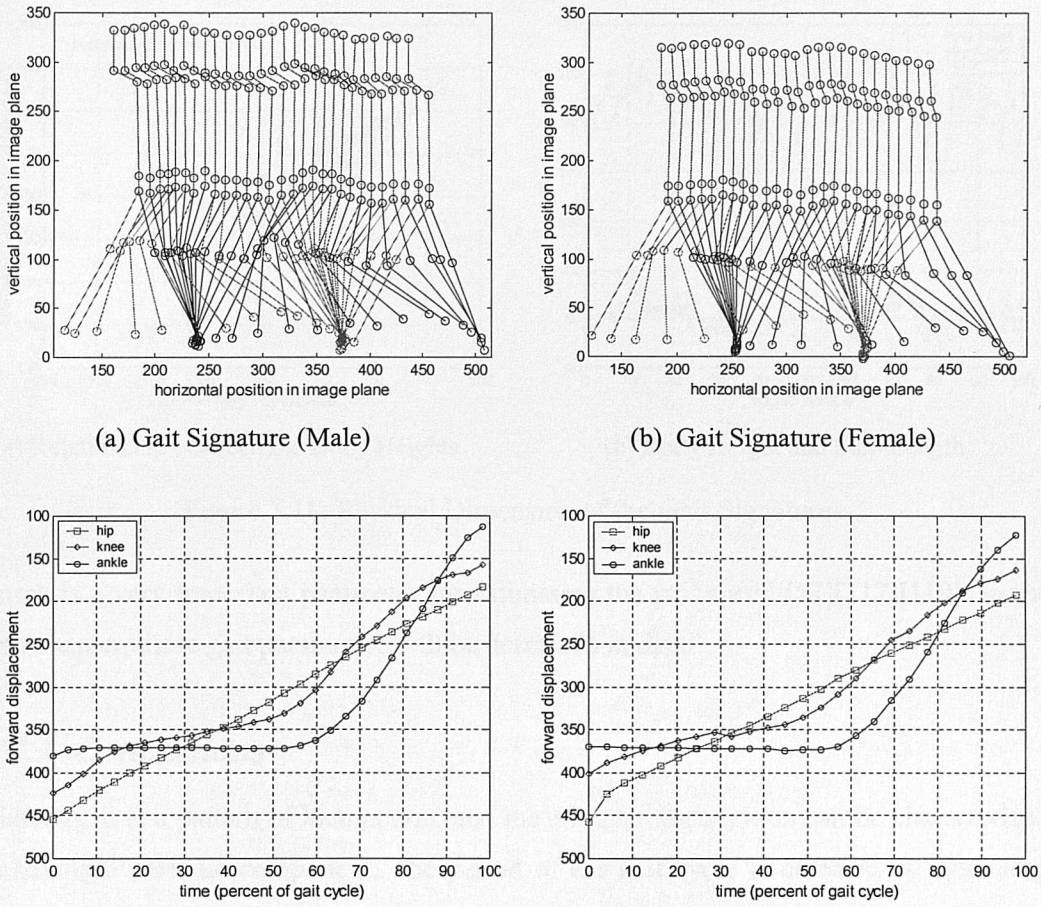
In the previous chapter, a large amount of the image sequences (the SOTON gait database) was pre-processed, and the body contours were extracted from the pre-

processed image sequences. In computer vision, the human gait is usually described by kinematic characteristics such as linear and angular positions with time derivatives. The gait signature based on 2D stick figure is the most effective and well defined representation method for kinematic gait analysis. Here, the gait motion is described as rhythmic and periodic motion, and the gait signature is extracted from body contours by motion information with topological analysis guided by anatomical knowledge. Also, the constraints in gait motion and the grammatical structure of the motion sequences are analyzed to improve the robustness of the gait signature.

### *3.4.1 Experimental Results*

In the experiments, the body contours extracted from seven indoor image sequences of each of the 100 different subjects are used. As described in the section 3.1, human gait is a form of periodic motion. To detect the period of a gait cycle, the symmetry property of the horizontal centre of mass in the HAT is analyzed, and the key-frames (double supports) are also determined by the obvious peaks in the symmetry property. The extracted phases of double supports show very similar result to medical analysis. The average of the detected gait period is 27.16±1.72 frames and 0.74 frames (2.7%) shorter than the average of the manually extracted gait period as shown in Figure 3.3(a). The period of gait cycle is an essential parameter to calculate the various kinematic parameters such as cadence, cycle time, velocity, and frequency.

On the other hand, a simplified 2D stick figure model with six joint angles is used to represent the human body structure in the gait signature. Figures 3.17(a) and (b) show the extracted gait signatures from image sequences during one gait cycle. These figures are enhanced result by the grammatical rules with gait constraints, which is described in Section 3.3. As a gait constraint, the step symmetry is used to verify the double supports and to detect an unstable status of the double supports. The variation (standard deviation), in seven sequences of the same subject, at joint angles is decreased by on average 17% for 100 subjects by using the gait grammar with physical constraints. By these figures, human gait motion can be described in a compact form, as a sequence of the kinematic parameters. In addition, the forward displacements of the extracted gait signatures at hip, knee, and ankle are shown in Figures 3.17(c) and (d). The forward displacements of joints




(a) Gait Signature (Male)

(b) Gait Signature (Female)

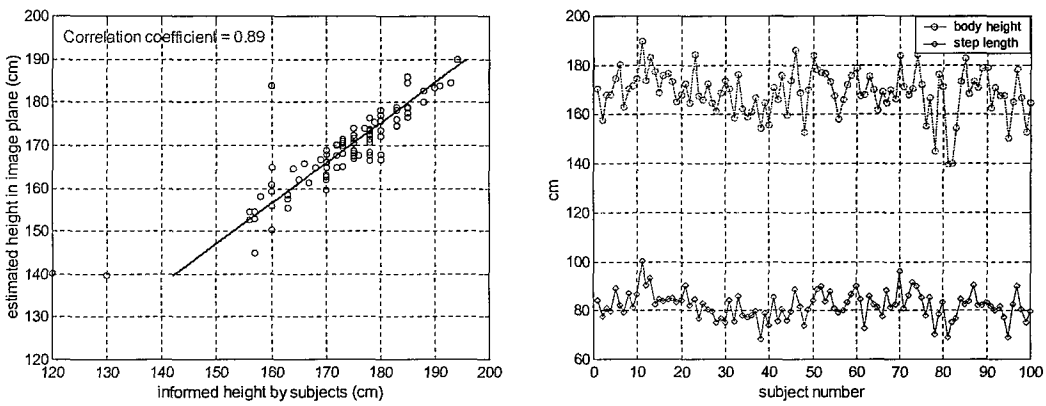
(c) Forward Displacements (Male)

(d) Forward Displacements (Female)

Figure 3.17: Gait Signatures during One Gait Cycle

are consistent with medical data by Inman's [57] analysis as shown in Figure 3.10.

In section 2.2.3, one pixel in the image plane of the SOTON database was approximated by the physical dimension $0.5cm \times 0.5cm$, thus the dimension of human body and its motion parameters can be simply estimated by the body points and forward displacements in the gait signature. The relationship between stated body heights by subjects and estimated body heights in the image plane is shown in Figure 3.18(a). As can be seen in the figure, the correlation coefficient for the relationship between two data sets has a high value (=0.89), and only 2.60 percentage shorter than the stated heights in average. Thus, the physical dimension of the human body by this estimation seems to be very reliable. Figure 3.18(b) shows the estimated body heights and step lengths. The step




(a) Relationship between the Body Heights (b) Body Height and Step Length

Figure 3.18: Physical Dimension of the Gait Signatures

length is a very important parameter for estimating the gait speed [57][117][119]. In the next chapter, these gait parameters will be described in detail.

## *3.4.2 Conclusions*

Human gait is a pattern of locomotion, and most human locomotion can be characterized as having a periodic component. The period of the gait cycle is detected by symmetry property of horizontal centre of mass in the HAT. A 2D stick figure is used to represent the human body structure, and it is extracted from body contour by determining the body points. To extract the body points, joint angles of each segment are extracted from gait skeleton data by linear regression analysis, and gait motion between key-frames is described by tracking the moving points of locomotion. The body segments and moving points are basically guided by topological analysis with anatomical knowledge. A gait signature consists of a sequence of the stick figures and improved by the gait constraints and grammatical analysis. Also, forward displacement of the gait signature is analyzed to show the characteristics of the gait motion. However, the gait signature includes much information for describing the gait motion and is a very effective representation method for analyzing human gait motion. In future, human gait can be analyzed and classified by kinematic features of the gait signatures.

# Chapter 4

# Human Gait by Statistical Analysis

## 4.1 Kinematics of the Gait Motion

Human walking is the most common means of gait motion, and all normal people follow the same basic bipedal pattern, namely all humans walk in the same basic way [92][119]. Also, the normal people will consistently return to the same pattern when tested repeatedly, unless changes in footwear or the walking surface [57]. In addition, gait motion can be decomposed into temporal and spatial components, and these components should be included in gait description. A very important aspect in the fundamental study of gait motion is the analysis of the kinematics of human body segments [122]. Kinematics is the study of movements, or more specifically the geometric description of motion, in terms of displacements, velocities and accelerations [1][117]. Here, gait motion is analyzed by measurement of temporal parameters such as stride, cadence and walking velocity, and kinematics dealing with the analysis of joint movements.

### *4.1.1 Analyzing the Gait Parameters*

Human gait contains numerous parameters, and these parameters can be categorized into general and kinematic parameters. The general gait parameters (also known as the temporal and spatial parameters) are the stride length, cycle time (or cadence) and speed and provide a basic description of the gait motion [117]. These parameters present essential quantitative information about a human gait and give a guide to the walking ability of subject [88]. In addition, the general parameters tend to change together in most

locomotor disabilities, so that a subject with a long cycle time will usually have a short stride length and a low speed. Each parameter may be affected by such factors as age, sex, height, muscle strength, etc. Therefore, they should always be interpreted in terms of the expected values for the subject's age and sex, such as shown in Table 4.1.

Table 4.1: Normal Ranges for General Gait Parameters [117]

| Age (years) | Gender | Cadence (steps/min) | Cycle Time (sec) | Stride Length (m) | Speed (m/sec) |
|---|---|---|---|---|---|
| 13-14 | Male | 100-149 | 0.81-1.20 | 1.06-1.64 | 0.95-1.67 |
| | Female | 103-150 | 0.80-1.17 | 0.99-1.55 | 0.90-1.62 |
| 15-17 | Male | 96-142 | 0.85-1.25 | 1.15-1.75 | 1.03-1.75 |
| | Female | 100-144 | 0.83-1.20 | 1.03-1.57 | 0.92-1.64 |
| 18-49 | Male | 91-135 | 0.89-1.32 | 1.25-1.85 | 1.10-1.82 |
| | Female | 98-138 | 0.87-1.22 | 1.06-1.58 | 0.94-1.66 |
| 50-64 | Male | 82-126 | 0.95-1.46 | 1.22-1.82 | 0.96-1.68 |
| | Female | 97-137 | 0.88-1.24 | 1.04-1.56 | 0.91-1.63 |
| 65-80 | Male | 81-125 | 0.96-1.48 | 1.11-1.71 | 0.81-1.61 |
| | Female | 96-136 | 0.88-1.25 | 0.94-1.46 | 0.80-1.52 |

In the previous chapter, the gait signatures during one gait cycle were extracted from the SOTON database. The trajectories of gait signatures contain the general gait parameters on human movement. In Section 3.1.2, period of the gait is determined by number of frames during one gait cycle in image sequence, and frame rate of the SOTON database was 1/25 seconds. Accordingly, the cycle time can be calculated by

$$cycle\_time(\text{sec}) = gait\_period(\text{frames}) / frame\_rate(\text{frames/sec}). \tag{4.1}$$

The cadence is the number of steps taken in a given time, the usual units being steps per minute. One gait cycle consists of two steps, so the cadence is given by

$$cadence(\text{steps/min}) = 120 / cycle\_time(\text{sec}). \tag{4.2}$$

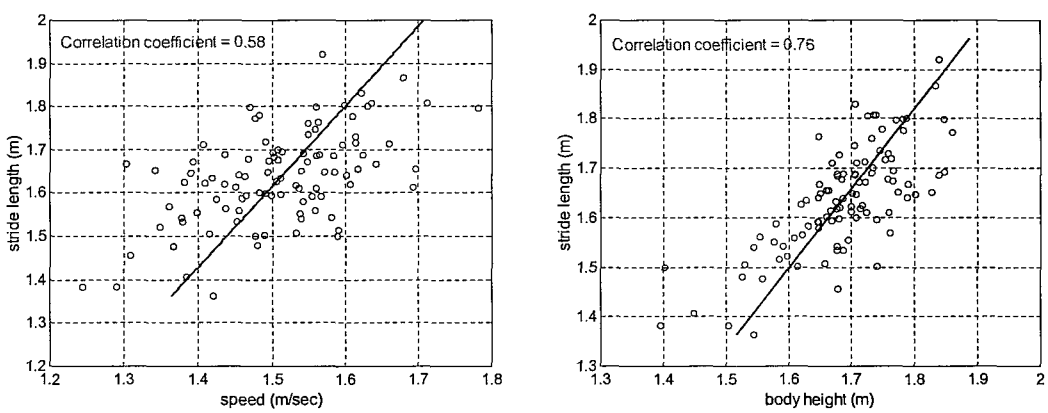
The gait speed is the rate of linear forward motion of the body, which can be measured in metres per second. The speed serves as one of the most important factors in determining gait characteristics, and it can be calculated from the cycle time and stride length as

$$speed(\text{m/sec}) = stride\_length(\text{m})/cycle\_time(\text{sec}). \tag{4.3}$$

The stride length is the distance between two successive placements of the same foot, and the length of one stride is moved during one gait cycle and includes all of the events of one gait cycle [88][117]. If both the cycle time and the speed have been measured separately, the stride length can be calculated by

$$stride\_length(\text{m}) = speed(\text{m/sec}) \times cycle\_time(\text{sec}). \tag{4.4}$$

Also, the stride duration or frequency defines the amount of time in which these motions occur. A stride consists of two steps, a right and a left. The stride length is not always twice the length of a single step because right and left steps may be unequal [88].




(a) Stride Length vs. Speed (b) Stride Length vs. Body Height

Figure 4.1: Relationship between the Gait Parameters

On the other hand, the stride length can be determined by direct measurement. In the previous chapter, the body height and step length in a gait signature were estimated from the physical dimensions of the image plane. The step length is determined by the average distance between two feet at double supports (initial contact and terminal swing), hence the stride length is given by two step lengths. The stride length can be also approximated by the forward displacements in the gait signature during one gait cycle. By medical and biomechanical studies, stride length is not independent of body height and speed [57][119]. Figure 4.1 shows the relationship between stride length and other gait parameters, for 100 different subjects. As can be seen in the figure, the stride length versus gait speed appears linearly related, and stride length versus body height has a

Table 4.2: General Gait Parameters from the SOTON Database

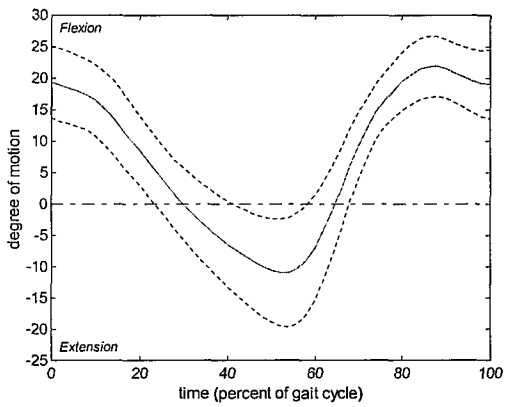
| Age Level | Gender | Cadence (steps/min) | Cycle Time (sec) | Stride Length (m) | Speed (m/sec) |
|---|---|---|---|---|---|
| Children | Male | 109-130 | 0.92-1.10 | 1.36-1.52 | 1.23-1.65 |
| | Female | - | - | - | - |
| Adults | Male | 103-116 | 1.03-1.17 | 1.57-1.76 | 1.42-1.62 |
| | Female | 110-122 | 0.98-1.10 | 1.43-1.62 | 1.38-1.56 |

linear relationship and is more highly correlated. The relationships of stride length have been used for recognizing human [7]. Table 4.2 shows variation of the general gait parameters, which are obtained from the SOTON database in the experiments. In the table, the speed is based on estimated stride length, and all parameters belong to the range of the expected value shown in Table 4.1. In practice, the speed is near the upper limit, because the SOTON database might have relatively much youth in the subject population. Also, the marker-less measurement is a natural method, thus a person may achieve a greater stride length than with a marker-based approach.

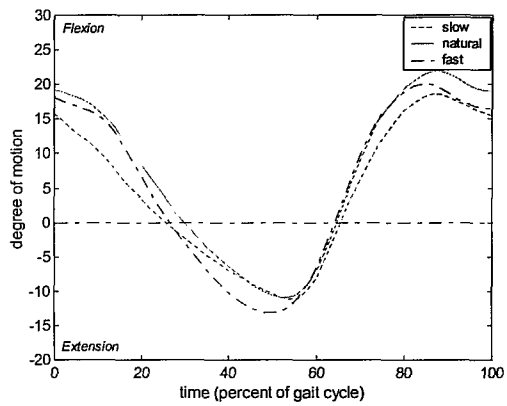
### *4.1.2 Gait Analysis in Angular Kinematics*

Kinematics is the geometric description of motion and does not consider the forces that cause the actions. Kinematic analysis of human gait usually characterizes the joint angles between body segments and their relationships to the events of the gait cycle [80][117][119]. The joint angles are more commonly expressed as flexion and extension [15][89][92]. Flexion occurs when two body segments change their relative position and decrease the angle between them. Similarly, extension is the return from flexion. The joint angles of the hip, knee and ankle have been considered as the most important kinematics of the lower limbs. The hip angle is defined by the relationship between the thigh and trunk or pelvis, and the pattern of the rotation is nearly the same as that of thigh rotation. The angle between the thigh and shin is the knee angle, and the ankle angle is the angle between the shin and foot.

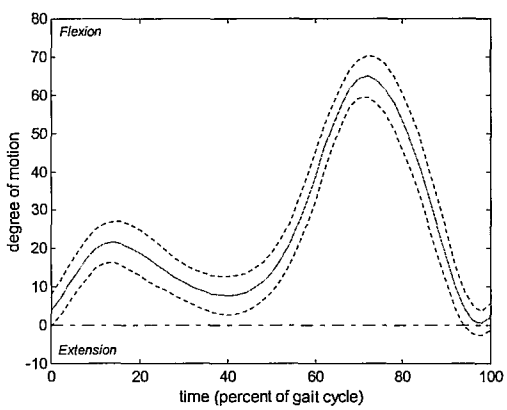
Figures 4.2(a), (c) and (e) show the mean and standard deviations, by medical data of a small number of the subjects (less than 15) from anatomical markers in limb segment,



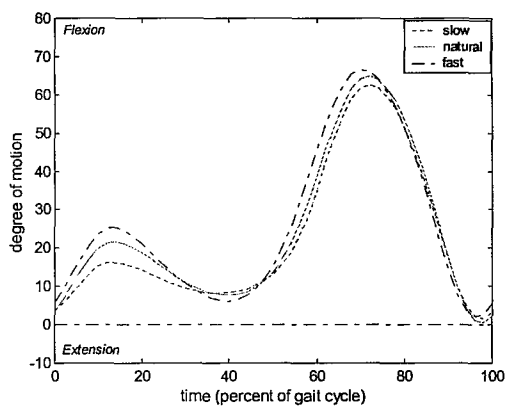

(a) Hip Angles vs. Time (Natural Cadence)



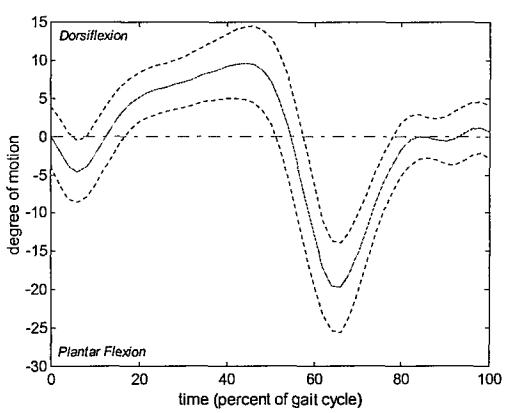

(b) Hip Angles vs. Time (Comparison)



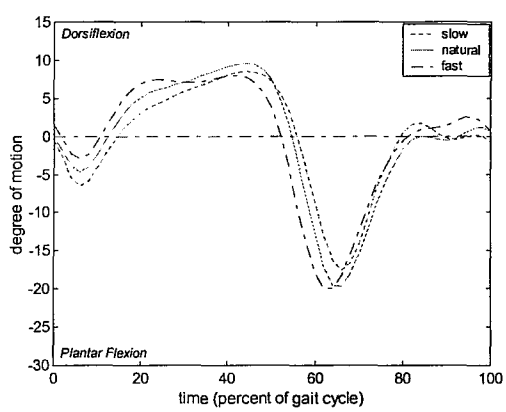

(c) Knee Angles vs. Time (Natural Cadence)




(d) Knee Angles vs. Time (Comparison)




(e) Ankle Angles vs. Time (Natural Cadence)




(f) Ankle Angles vs. Time (Comparison)

Figure 4.2: Joint Ranges of the Motion during Free Walking

for joint angles of the hip, knee and ankle motion during one gait cycle [119]. Figures 4.2(b), (d) and (f) show comparison results of the mean hip, knee and ankle curves across the cadences. The natural cadence was 105±6 steps/min, slow cadence was approximately 20 steps less than natural, and fast cadence was about 20 steps greater than natural [119]. As can be seen in the figures, the plots of these joint angles differ little between the cadence groups. The only minor difference shows up in knee flexion during early stance. At 15% of the gait cycle, the knee reaches maximum flexion and this increases from 15° for slow walkers to 25° for the fast cadence group. The ankle curve also shows very small differences. The ranges of joint angles in normal walking have been shown previously in Table 3.2.

However, the joint angles of a stick figure in the gait signature can be determined by the coordinates of the body segments. Also, kinematics explicitly defines the state vector of a stick figure at a specific time. The state vector is defined by

$$\Theta_t = (\theta_1, ..., \theta_n). \tag{4.5}$$

The state vector is the set of joint angles including independent parameters defining the positions and orientations of all joints belonging to the figure. In the gait signature, a sequence of the state vectors contains all the geometrical and time-related properties of the gait motion. Therefore, the time series of the joint angles provide much kinematic information such as linear and angular velocities, and accelerations, which are derived from the displacements with time interval. Moreover, they have basically the same characteristic with respect to the gait cycle.

## 4.2 Periodicity Detection and Analysis

In the previous chapters, the human body and its contour have been extracted, and the gait signature was also extracted from the body contour by determining the body points. The gait signature represents the movements of human gait with the body structure by a 2D stick figure model. A set of the gait data extracted from the gait signatures provides potentially valuable time-dependent patterns as a gait time series. It has a quasi-periodic temporal dependence, because the data collected during walking or running is at the

subjects' own speed [18]. In addition, human gait is a form of periodic motion, especially when walking laterally; hence we can predict human movement in a gait cycle. Here, the gait signatures are analyzed to derive the trajectory-based kinematic characteristics. The inherent periodicity in gait motion is also analyzed by the graphical representations and statistical methods such as autocorrelation and Fourier analysis.

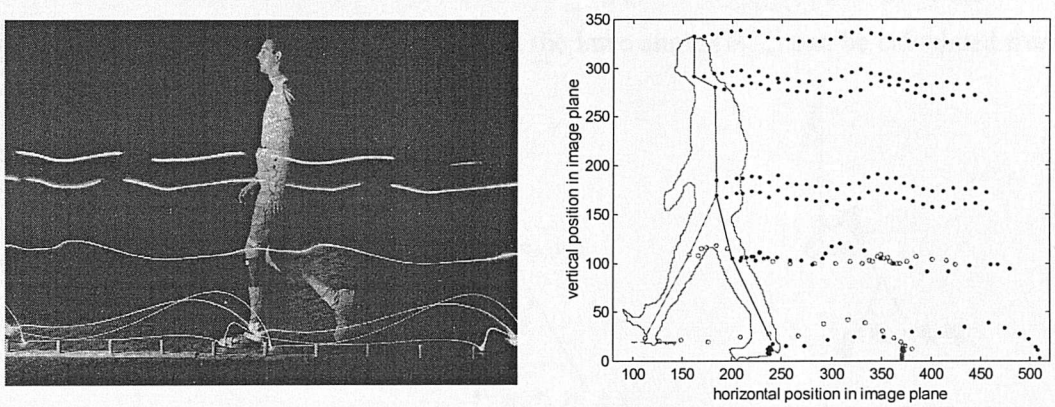
### *4.2.1 Sequence of the Gait Signatures*

Gait motion is the repetitive and well-organized movements generated by a person during actions such as walking. Accordingly, the usual input to analysis is a temporal image sequence. An image sequence is a much richer source of information than a single image. Motion in image sequences refers to the 2D displacement or velocity of the projection of scene. The trajectory of points in each image frame follows a curve in the $(x, y, t)$ space of an image sequence, and the simplest trajectory is linear. Let $\boldsymbol{x} = (x, y)^{\mathrm{T}}$ be the spatial position of a pixel in continuous coordinates, i.e., $\boldsymbol{x} \in \mathrm{R}^2$ within image limits. Then, a linear trajectory can be described as

$$\boldsymbol{x}(\tau) = \boldsymbol{x}(t) + \boldsymbol{v}_t(\boldsymbol{x})(\tau - t) \tag{4.6}$$

where the velocity $\boldsymbol{v}_t(\boldsymbol{x})$ is constant between $t = t_{k-1}$ and $\tau$ ($\tau > t$). In an image sequence, each coordinate can be represented by a linear trajectory.

In the previous chapter, the gait signatures were extracted from the image sequences,




(a) Trajectories by Light Displays [57] (b) Trajectories of Gait Signature

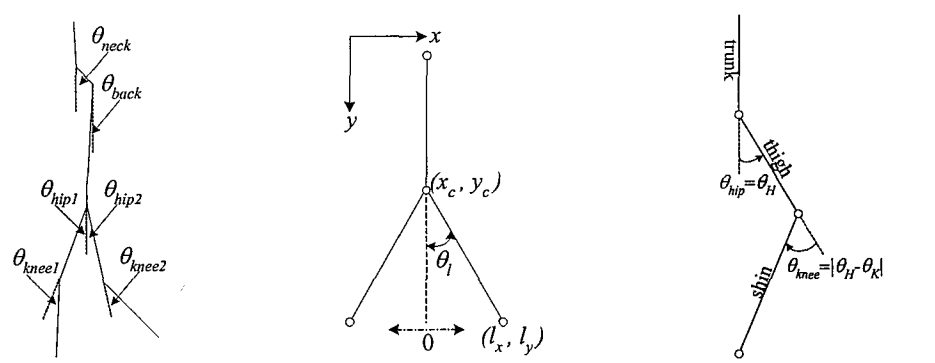
Figure 4.3: Trajectories of Gait Motion Corresponding to the Joints

which are acquired at discrete time instants. Thus, the motion vectors associated with the positions of body joints can be also represented by a linear trajectory. The trajectory plots of one or more coordinates are useful in describing detailed changes of particular motion pointer. A sequence of this plot at equal intervals of time gives a pictorial and anatomical description of the dynamics of the movement [118]. Figure 4.3 shows the trajectories of gait motion during one gait cycle. Figure 4.3(a) shows the result by moving light displays from a medical study [57], and Figure 4.3(b) is extracted from the gait signature shown in Figure 3.17(a). As can be seen in the figures, both results are very similar. Also, a total description in the plane of the movement is defined by the gait signature, and trajectories, velocities, and accelerations can be visualized.

However, we can calculate the six angles which are associated with joint positions (here, the pelvis angle is not considered): $\theta_{neck}$, $\theta_{back}$, $\theta_{hip1}$, $\theta_{hip2}$, $\theta_{knee1}$, and $\theta_{knee2}$. Figure 4.4(a) shows the graphical demonstration of the joint angles, and Figure 4.4(b) shows the definition of the angle $\theta_l$ at position $(l_x, l_y)$. In general, the angle $\theta_{l,k}$ of location $(l_x, l_y)$ at frame $k$ can be calculated by

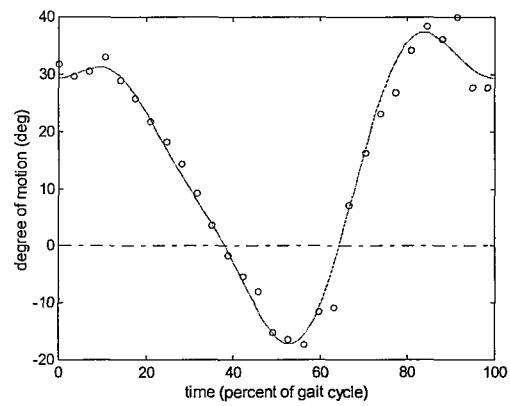
$$\theta_{l,k} = \tan^{-1}\left((l_x - x_c)/(l_y - y_c)\right). \tag{4.7}$$

The leg angles are relative [119], so the relative joint angles in each leg are computed from the extracted angle values by Equation 4.7. Figure 4.4(c) shows the definition of the relative joint angles. In normal walking, the trunk of human body in the SOTON database can be considered to be almost vertical. Accordingly, the relative hip angle ($\theta_{hip}$) is the same as that of the extracted value ($\theta_H$), and the knee angle ($\theta_{knee}$) can be calculated from

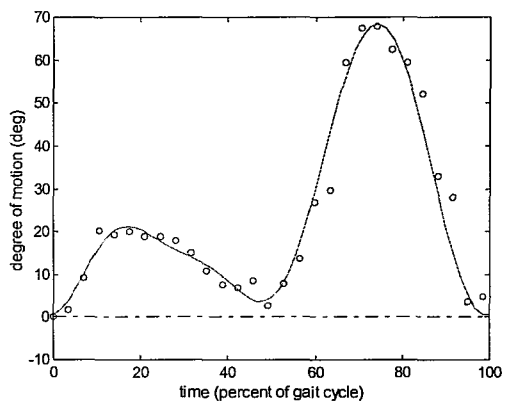



(a) Joint Angles (b) $\theta_l$ at $(l_x, l_y)$ (c) Relative Joint Angles

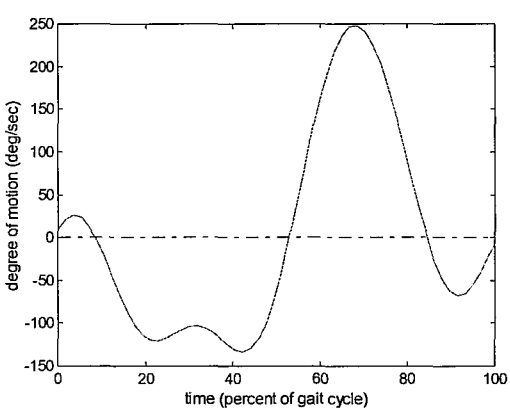
Figure 4.4: Joint Angles in the Gait Signature



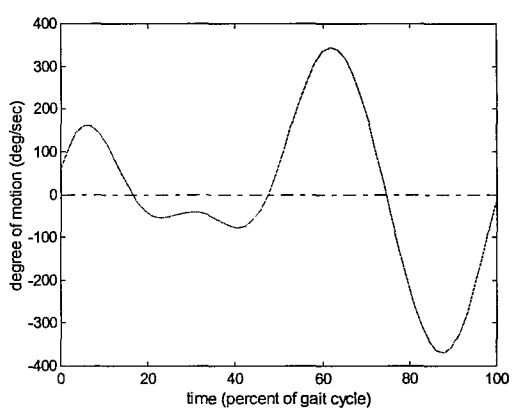

(a) Angles vs. Time (Hip)




(b) Angles vs. Time (Knee)




(c) Angular Velocities vs. Time (Hip)




(d) Angular Velocities vs. Time (Knee)

Figure 4.5: Angular Kinematics of Gait Signature

the extracted hip angle ($\theta_H$) and knee angle ($\theta_K$).

Now we can extract the trajectory-based kinematic characteristics of the gait signatures such as linear and angular position, their displacement and the velocities. Figures 4.5(a) and (b) show the time series of the relative joint angles obtained from the gait signature in Figure 3.17(a), by our new approach. The points in Figures 4.5(a) and (b) are the extracted angle values from the gait signature that were calculated by using Equations 3.5 and 3.7, and these are very similar to the medical data of Figures 4.2(a) and (c). Figures 4.5(c) and (d) show the angular velocities given by angular displacement per inter-frame time 1/25 seconds. In the figures, the lines are interpolated by trigonometric polynomials.

### *4.2.2 Periodicity in the Gait Motion*

Several methods of detecting periodicity in image sequences have been studied [24][73][96][103][104]. Periodicity is a very strong cue in human motion perception, and the periodicity of motion can be used to recognize individuals [72]. In a motion analysis system, the usual data source is a temporal image sequence. An image sequence is defined as a series of $N$ images or frames, acquired at discrete time instants

$$t_k = t_0 + k\Delta t, \quad k = 0,1,...,N-1 \tag{4.8}$$

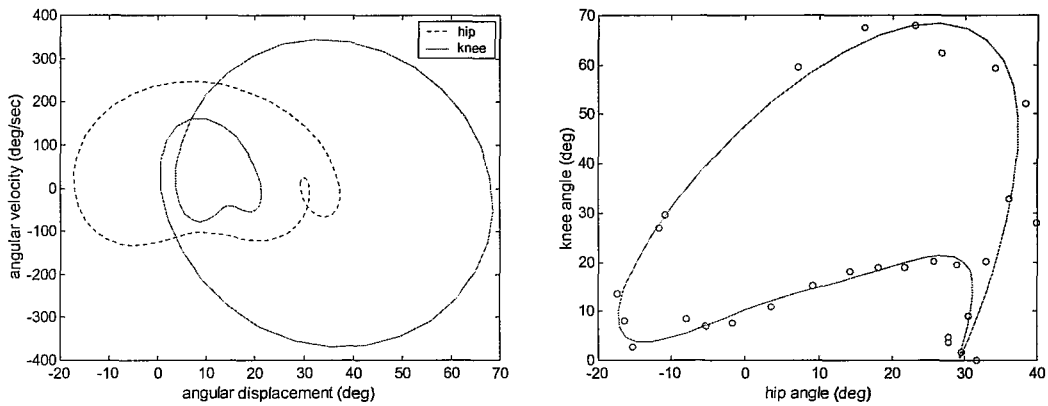
where $\Delta t$ is a fixed time interval and typically defined by frame rate in an image sequence. In general, a motion $f(t)$ can be defined as a function at time $t$. If it repeats itself with a period $T$, then a periodic motion can be described as

$$f(t+T) = f(t), \quad T > 0. \tag{4.9}$$

In the image sequences of gait motion, the period and frequency are respectively the time taken by a step and the number of steps taken per second. Accordingly, gait describes periodic motion if all the spatial-temporal parameters are repeated after $T$ steps, where $T$ is the gait period.

In Section 3.1, gait was described as the periodic motion between the successive heel-strikes, and the period of the gait cycle was detected. Also, Murray [79] considered human gait as "a total walking cycle" – the action of walking can be thought of as a periodic signal. Therefore, the pattern of rotation angles around the joints has been modelled as a pendulum, the motion of which is characterized by simple harmonic motion. This assumes that the pattern of the motion is approximately sinusoidal in nature [14][23][111]. Accordingly, gait motion can be characterized as having a periodic component. Moreover, Bertenthal and Pinto [8] consider three specific dynamic properties in the perception of human gait. First, a frequency entrainment property is observed when two or more components of the gait cycle share the same frequency. Second, phase-lock is observed when phase relationships among the components of the gait cycle are stably related in their position or phases. Third, a periodic attractor property is defined as a stable solution to an equation of motion.

On the other hand, the phase-space portrait is a useful method to represent periodic



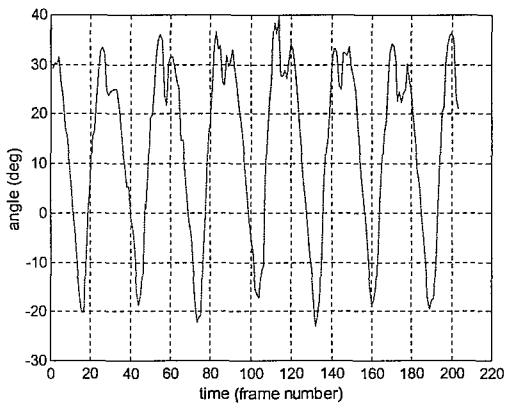

(a) Phase-Space Portrait (b) Hip-Knee Cyclogram

Figure 4.6: Periodicity Detection of Gait Signature

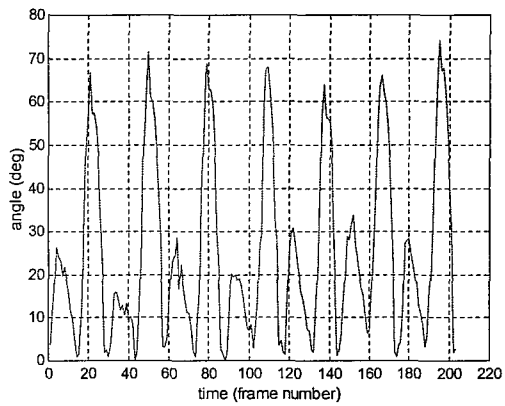
motion. The phase space is described as the space of all possible states of dynamic system, in which the first derivative, or the velocity of movement ($v$), is plotted against its position ($s$). As time varies, the point $P(s, v)$ describing the motion of the system moves along a certain trajectory on the phase space [8]. The curves plotted by these points represent periodic motion because the velocity and position return repeatedly to the same points. Figure 4.6(a) is a phase-space portrait of angular displacement ($\theta$) versus angular velocity ($\omega$) for the hip and knee motion of Figure 4.5. As expected, a periodic phase trajectory is observed in the figure. In addition, a cyclogram [44] is demonstrated by ignoring the time axis of each curve and directly plotting knee angle versus hip angle as shown in Figure 4.6(b). The cyclogram provides information about the posture of the leg and the coordination of two joints but no information about the velocities involved. However, using angle versus angle plots rather than angle versus time curves can improve understanding of gait. The main reason for this is the fact that the shape of the loops is easily recognizable, especially by eye.

### *4.2.3 Analyzing the Gait Periodicity*

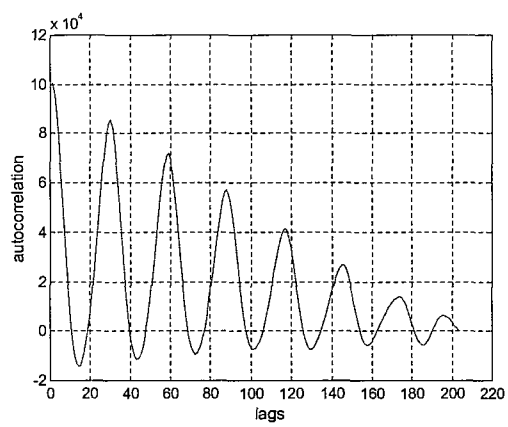
In the previous section, periodicity of the joint angles in the gait signature was detected by graphical analysis. The angles obtained from a gait signature are a form of time series data. A time series consists of values sampled at constant intervals over time, and any quantity that can be measured over time can be analyzed as a time series. Figures 4.7(a)



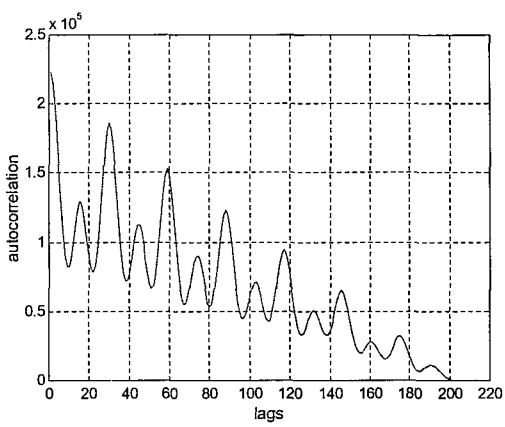

(a) Time Series of Hip Angles



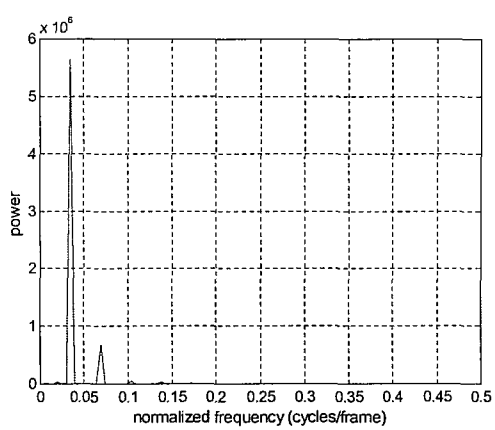

(b) Time Series of Knee Angles



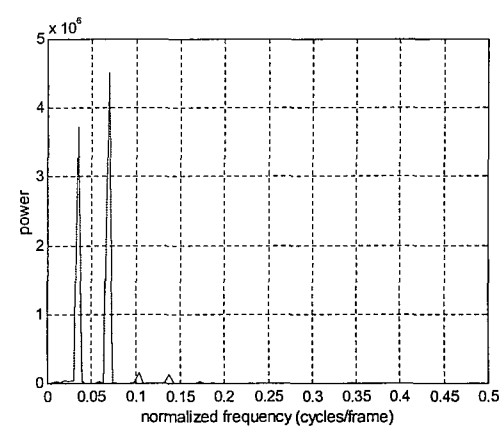

(c) Autocorrelation of (a)




(d) Autocorrelation of (b)




(e) Periodogram of (a)




(f) Periodogram of (b)

Figure 4.7: Periodicity Analysis of Gait Motion

and (b) show the time series of the joint angles, which are extracted from the sequences of the gait signatures. As a sequence of trajectory points, the angle data are ordered by an implicit time dimension. In time series modelling, this kind of data can be understood as a trigonometric model. The trigonometric model is described by

$$X(t) = \alpha \cos(2\pi ft + \phi) + \varepsilon_t \tag{4.10}$$

where the amplitude $\alpha$, frequency $f$, and phase $\phi$ are parameters and $\epsilon_t$ is a white noise series. As the deterministic part of trigonometric, this model is periodic and will be appropriate only for the cyclic data.

On the other hand, periodic motion description requires long sequences of image frames. As can be seen in Figures 4.7(a) and (b), the gait signatures during seven gait cycles are used to describe the sequence data of the joint angles. If the joint motion of human gait is periodic, then there will be some self-similarity within the curvature function which becomes more evident in the autocorrelation function. The autocorrelation function of a periodic waveform is itself periodic, and the periodic waveform $f(t)$ of period $T$ satisfies Equation 4.9. If the waveform is completely random, then the autocorrelation function will have its peak value at zero lag and will reduce to a random fluctuation of small magnitude about zero for lags greater than about unity [56]. Figures 4.7(c) and (d) show the autocorrelations of the time series of the joint angles, which are described in Figures 4.7(a) and (b). Here, the results show that the joint motion of human gait is completely periodic. The autocorrelation describes the general dependence of the values of the samples at one time on the values of the samples at another time.

Fourier analysis has also been used to determine periodicity in an image sequence. In the Fourier analysis, any periodic waveform can be decomposed into a fundamental and harmonics. That is, the energy of a periodic waveform can be concentrated at frequencies which are integral multiples of some fundamental frequency [96]. This implies that peaks at the fundamental frequency and its harmonics can be observed by the Fourier transform of a sampled periodic waveform. Hence, the periodicity of the sequence data of gait motion can be detected by obtaining its Fourier transform and checking whether all the energy in the spectrum is contained in a fundamental frequency and its integral multiples. The time series of gait motion in Figures 4.7(a) and (b) show a periodic oscillation. The

corresponding power spectrums, which are shown in Figures 4.7(e) and (f), exhibit a strong peak at the drive frequency together with some higher frequency harmonics. Consequently, the result shows that the sequence data of gait motion is a periodic signal and has same fundamental frequency.

## 4.3 Time Series Analysis and Prediction Model

The gait signature is a sequential set of the stick figures obtained over time, thus the joint angles extracted from the gait signature are time series data. In the previous section, the time series of gait motion was analyzed and characterized as a form of periodic motion. The time series analysis has three goals: modelling, prediction, and characterization [40]. The goal of modelling is to find a description that accurately captures features of the long-term behaviour. The aim of prediction is to accurately predict the short-term progression. The third goal, characterization, attempts to determine fundamental properties. Here, the time series for periodic gait motion is modelled by interpolation functions of trigonometric polynomials. Prediction method of the gait time series based on delay coordinate embedding is analyzed. In addition, a scale of the gait motion is described by statistical moments.

### *4.3.1 Gait Motion by Interpolated Model*

In many statistical analyses, the main goal is to establish functional relationships which make it possible to predict one or more variables in terms of others. Regression methods are among some of the most widely used methods in statistical approaches. Consider the basic regression model with bivariate observations $(x_1, y_1),\ldots,(x_n, y_n)$ satisfying [31]

$$y_i = \mu(x_i) + \varepsilon_i, \quad i = 1,...,n \tag{4.11}$$

where $\varepsilon_i$ are zero mean random errors and $\mu$ is an unknown regression function, and $x_i$ are assumed to fall in finite interval $[a, b]$. There are many effective methods of estimating $\mu$ including kernel, nearest neighbour, polynomial, Bayesian, and spline estimators.

However, statistical analyses of the gait relationships typically use continuous curves of the time series data measured over the gait cycle. Medical research [79] has shown that

the pattern of gait motion is approximately sinusoidal in nature. The basic trigonometric functions sine and cosine describe sinusoidal functions. Moreover, the trigonometric basis functions $\phi_i$ are all periodic over $2\pi$. The basis functions for trigonometric polynomials on $[-\pi, \pi]$ are described as

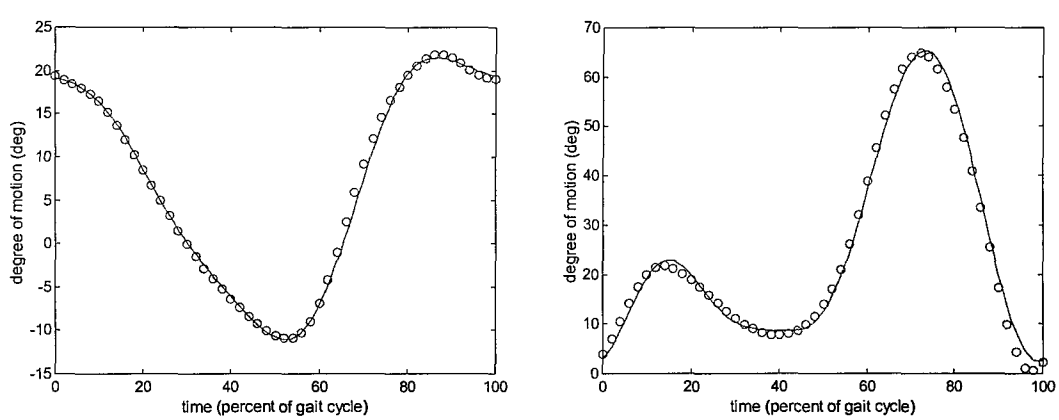
$$\phi_0(x) = \frac{1}{\sqrt{2\pi}}, \quad \phi_{2l-1}(x) = \frac{1}{\sqrt{\pi}}\sin(lx), \quad \phi_{2l}(x) = \frac{1}{\sqrt{\pi}}\cos(lx), \quad \text{for } l = 1,...,\frac{n}{2} \tag{4.12}$$

where $n$ is assumed even, and the basis is orthogonal. The trigonometric functions with period $2\pi$ are naturally suited to estimating a gait curve by time series data. An assumed functional relationship between periodic gait motion and time can be modelled by interpolation of trigonometric polynomials. By the basic regression model of Equation 4.11, an $n^{th}$-order trigonometric polynomial interpolation by least squares approximation can be described as

$$y_n(x) = a_0 + a_n \cos(2\pi nx) + \sum_{k=1}^{n-1}[a_k \cos(2\pi kx) + b_k \sin(2\pi kx)] \tag{4.13}$$

where the $a_0$, $a_n$, $a_k$ and $b_k$ are unknown curve-specific coefficients. As $n \to \infty$, $y_n(x)$ tends to the Fourier series.

The interpolation of much (equally-spaced) data by trigonometric polynomials can make for very accurate results. Figure 4.8 shows the curves for joint angles of the hip and knee motion during one gait cycle. Here, the points are the medical data from anatomical




(a) Hip Angles vs. Time (b) Knee Angles vs. Time

Figure 4.8: Interpolation by Trigonometric Polynomials

markers in a limb segment [118], and the lines are the curves that result from using $4^{th}$-order trigonometric polynomial interpolation. In the figure, the interpolation function is seen to work very well for modelling the gait motion in that the fit over the whole cycle appears close. In addition, future gait motion can be predicted by the periodicity of this interpolation model. To represent or correct measured gait angle, Fourier series and a perspective transformation techniques have also been used in gait studies [14][23].

### *4.3.2 Time-Delay Coordinate Embedding*

The time evaluation of a dynamical system is represented by a function of the time variation or (when sampled at regular intervals) time series of its dynamical variables. The state of many dynamic systems can be accurately reconstructed by a finite window of the time series. Let $A$ denote a compact $n$-dimensional set of states of the system, and $h:A\rightarrow \boldsymbol{R}$ be an observation function which is a measurement of sum quantity of system, i.e., $x=h(\boldsymbol{a})$, and let $\tau$ be a real number greater than zero. For each state $\boldsymbol{a}\in\boldsymbol{R}^n$, one can define the $m$-dimensional vector $\mathbf{X}$ [102]

$$\boldsymbol{X}=[h(\boldsymbol{a}),h(F_{-\tau}(\boldsymbol{a})),\ldots,h(F_{-(m-1)\tau}(\boldsymbol{a}))] \tag{4.14}$$

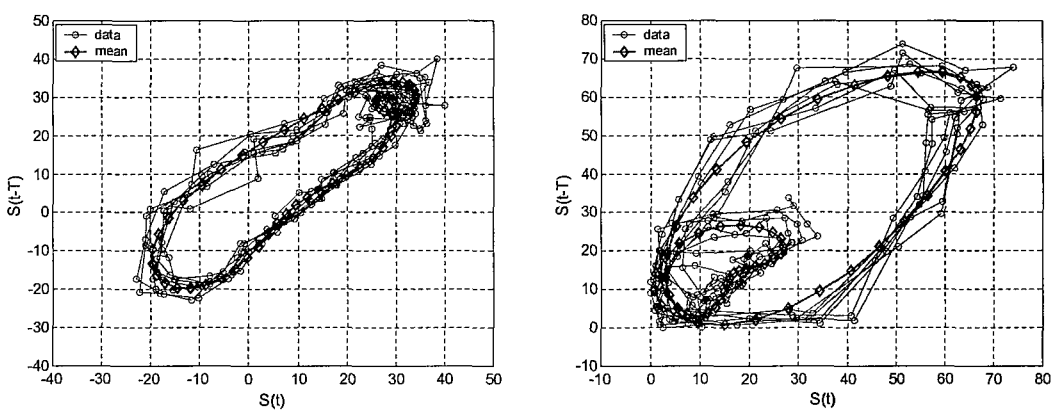
where the function $F_t(\boldsymbol{a})$ is defined through $\boldsymbol{a}(t_T+t) = F_t(\boldsymbol{a}(t_T))$. This vector is called a delay coordinate vector because its components consist of time delayed version of the observable of the system. The vector $\mathbf{X}$ is a segment of a time series with equal spaced data produced by measurement function $h$. This is called a time delay embedding, and it can be obtained as

$$\boldsymbol{X}=[x_t,x_{t-\tau},\ldots,x_{t-(m-1)\tau}] \tag{4.15}$$

where $x_t=h(\boldsymbol{a})$ is the value of the time series at time $t$ and $\boldsymbol{a}$ is the state. Time delay embeddings are widely used as the input vector to dynamic models, both linear and nonlinear [76].

In the previous section, Figures 4.7(a) and (b) show the time series of angle data of human gait. A useful method to plot the time series is to use delay coordinates. Figure 4.9 shows each value of the time series of angles ($\theta_t$) versus a time delayed version, by plotting ($\theta_t$, $\theta_{t-T}$) for fixed delay $T$=0.08. Using delay-coordinates for periodic dynamics

can reproduce the periodic orbit of the true system state space. Namely, the state space defined by $k$ coupled autonomous differential equations is $\Re^k$, however, the dimension $k$ would be very large. In contrast to the complicated differential equations of motion, the behaviour in Figure 4.9 is fairly simple. This graph also represents periodic motion because the same orbit is described, and the graph exhibits symmetry. These trajectories exhibit a tendency to maintain a fixed orbital shape (or limit-cycle attractor). For time series prediction, we need try to identify what state the system is in, look to the past for similar states, and see what ensued at those times. In practice, we can average the predictions to improve the statistical quality of the final prediction and accurately predict where on the curve the system will be in one addition second [8][102].




(a) Delay Plot from Figure 4.7(a) (b) Delay Plot from Figure 4.7(b)

Figure 4.9: Delay-Coordinate Reconstruction for Gait Analysis

Many available methods for time series analysis assume linear relationships among variables. But in the real world, temporal variations in data do not exhibit simple regularities and are difficult to analyze and predict accurately [16]. The motion of human gait is also described by using data of several contemporaneous variables changing with time. For example, the vertical change in head position has two cycles of downward and upward displacement in each stride. These reflect the mechanics of the right and left steps. Namely, there is a strong correlation between the behaviours of entire body motion. To improve prediction accuracy, more sophisticated analyzing methods for non-linear relationships, such as multivariate time series analysis, chaotic time series analysis, or neural networks, will have to be used.

### *4.3.3 Gait Description by Statistical Moments*

In Section 4.2.2, the phase-space portraits and hip-knee cyclograms for detecting the periodicity of human gait motion have been described by analyzing the joint angles, and the periodicity is observed by a two-dimensional closed figure. Both phase-space portraits and hip-knee cyclograms are important signatures of human gait motion, and a typical angle-angle cyclogram has more important features with the characteristics of the gait motion [44]. In gait analysis, hip-knee cyclograms describe the changes in the knee joint as a function of the hip joint. These graphical representations combine the temporal changes of two joint angles, which allow interpretation of the relationships between the two angles. In addition, the hip-knee cyclograms represent the movements of nearly the entire body, thus they can be representative of the subject's gait pattern [6].

The hip-knee cyclograms provide a basis for separating different gait patterns, and they are represented as kinds of plane closed curves. There are several methods for quantifying planar shapes [108]. Here, we consider statistical moments [44][53] as a descriptor for characterizing the gait motion. In general, the statistical moments of order $(p+q)$ in the $x$-$y$ plane are defined by

$$M(p,q) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)dxdy \tag{4.16}$$

where $f(x, y)$ is the membership function of the curve. This function takes the value 1 if the point $(x, y)$ is on the shape boundary, and otherwise its value is 0. For discrete data, the moments can be given by

$$M(p,q) = \sum_{x=1}^{n}\sum_{y=1}^{m} x^p y^q f(x,y) \tag{4.17}$$

where $m$ and $n$ are the finite region of the $x$-$y$ plane. By Equation 4.17, $M(0, 0)$ represents the region of a closed curve. The centre coordinates of the region is given by

$$\mu_x = \frac{M(1,0)}{M(0,0)}, \quad \mu_y = \frac{M(0,1)}{M(0,0)}. \tag{4.18}$$

Translation invariance can be achieved if we use centralized moments. The centralized moments are computed with respect to the centre of gravity by

$$\mu(p,q)=\sum_{x=1}^{n}\sum_{y=1}^{m}(x-\mu_x)^p(y-\mu_y)^q f(x,y). \tag{4.19}$$

To enable invariance to scale, normalized central moments are given by

$$\eta(p,q)=\frac{\mu(p,q)}{\mu(0,0)^{[(p+q)/2]+1}}. \tag{4.20}$$

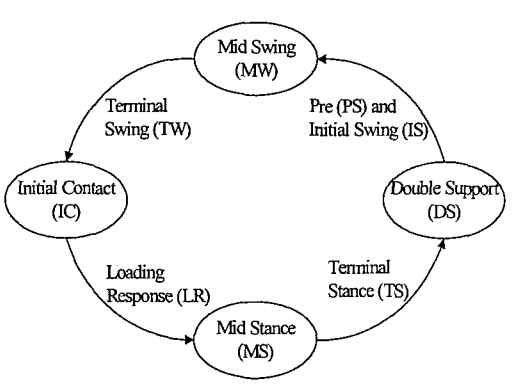
The first seven scale-normalized central moments are described by $\eta(1,1)$, $\eta(2,0)$, $\eta(0,2)$, $\eta(2,1)$, $\eta(1,2)$, $\eta(3,0)$, $\eta(0,3)$.

However, the lower order moments can be used in shape description even if the region is represented by its boundary [108]. Moments are an alternative approach for describing a shape, and they can provide much more information than a single measure of area. Many aspects of moment properties, normalization, descriptive power, sensitivity to noise, and computational cost have been studied. Hu [53] has derived moment expressions that are invariant to translation, rotation and scaling of shapes. They consist of groups of normalized central moment expressions. Some examples are shown below

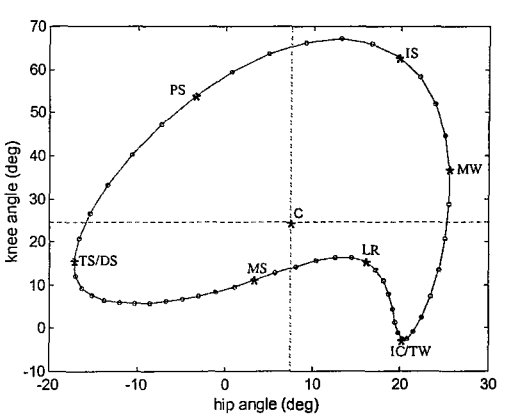
$$\varphi_1=\eta_{20}+\eta_{02} \tag{4.21}$$

$$\varphi_2=(\eta_{20}-\eta_{02})^2+4\eta_{11}^2 \tag{4.22}$$

$$\varphi_3=(\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\eta_{03})^2. \tag{4.23}$$



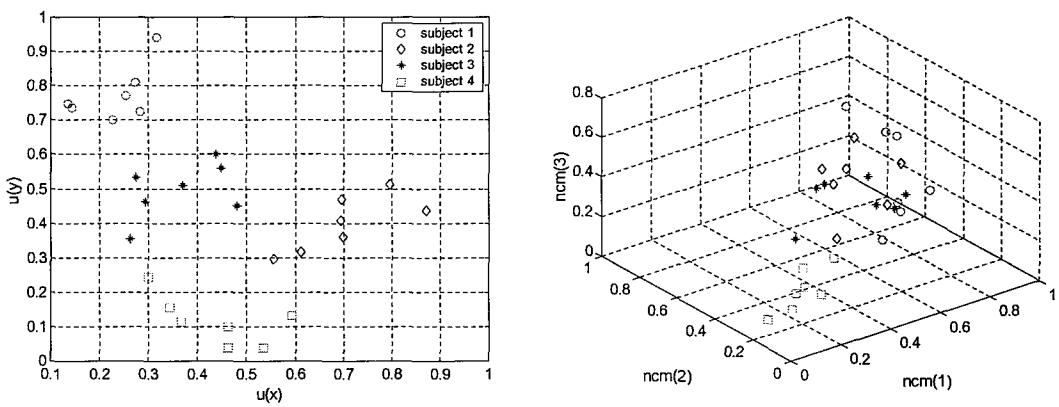

(a) State Diagram of Gait Cycle




(b) Gait State in the Cyclogram

Figure 4.10: Representation of Gait State by Periodic Gait Pattern

These moments are of finite order, thus unlike the centralized moments they do not comprise a complete set of shape descriptors. However, the normalized central moments in a cyclogram can be calculated by assigning $x=\theta_{hip}$ and $y=\theta_{knee}$. Figures 4.10(a) and (b) show the state diagram of gait cycle and the cyclogram with gait state by the data from Figure 4.8. As can be seen in the figures, the gait cycle is divided into seven temporal segments [88][92], and the finish points of each phase are marked by '*' on the cyclogram. As descriptors, the moments are simply calculated by using these finish points or $x$-$y$ coordinates in the angle-angle cyclogram. The descriptors are obviously invariant with respect to translation and scale.




(a) Centre Coordinates of the Cyclograms (b) Normalized Central Moments (2nd Order)

Figure 4.11: Gait Descriptors based on Moments of Cyclograms

The normalized gait descriptors based on Hu moments of the cyclogram are shown in Figure 4.11. For visualization purposes, the central coordinates and only 3 of the moments of the 4 subjects with seven gait signatures of each subject are separately shown. As can be seen in the figure, the centre coordinates appear to have better inter-class variability, but 3 of moments have higher variance. Recently, temporal and Zernike velocity moments have been used to describe motion for recognition [106].

## 4.4 Results and Conclusions

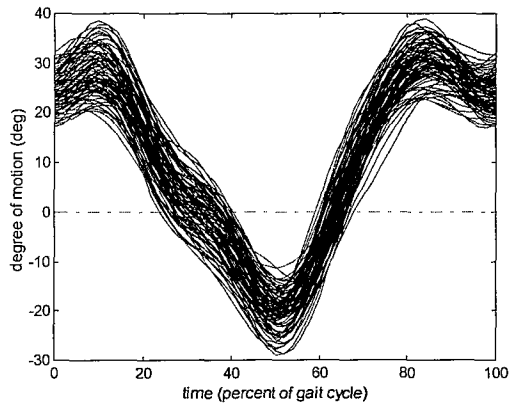
The human gait is a pattern of locomotion and can be described by the general and kinematic parameters. The measured parameters in the gait signatures are stride length,

cycle time (or cadence), speed, and joint angles. The joint angles are the most important kinematic parameters and defined as a gait time series. The time series of gait angles can be characterized as having a periodic component, and hidden periodicity of the gait time series can be detected by the autocorrelation function or by Fourier analysis. In addition, the time series for periodic gait motion is often modelled as harmonic motion such as a pendulum. Here, the gait signatures are analyzed to extract the gait parameters and periodic components. The periodic motion of human gait is modelled by interpolation of trigonometric polynomials. Also, delay coordinates embedding and statistical moments are employed to describe the human gait motion.

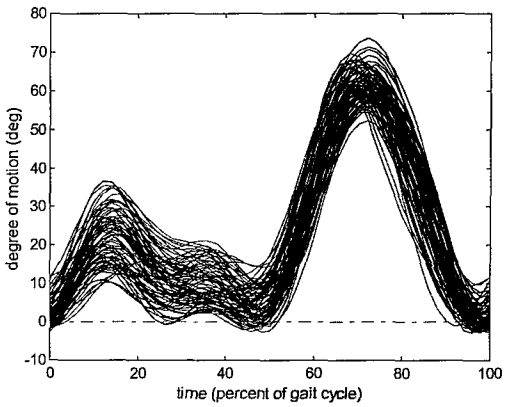
### *4.4.1 Experimental Results*

In the experiments, the gait signatures extracted from seven (body contour) image sequences of each of the 100 subjects are used. The gait parameters are extracted from the gait signatures and the relationship between the parameters is analyzed in Section 4.1.1. Figures 4.12(a) and (b) show the results of measuring relative joint angles obtained from 100 different subjects during one gait cycle. In the figures, the lines are the curves that result from using $4^{th}$-order trigonometric-polynomial interpolants. Also, Figures 4.12(c) and (d) show the mean and standard deviation of the relative joint angles obtained from Figures 4.12(a) and (b). As in medical studies [79][92], the hip and knee at initial contact are flexed by about 25° and 5° from the vertical, respectively. During the loading response, the hip position is relatively stable, possibly losing 2°~3° of flexion, and the hip progressively extends at the same rate after mid stance. Also, peak extension of the knee is attained slightly before the end of the swing phase.

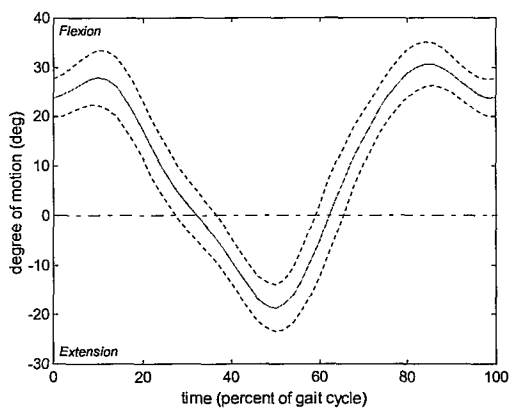
In Figures 4.12(e) and (f), we can observe a periodicity of the gait motion and also predict a gait movement by using the phase-space portrait and delay coordinate embedding. In the figures, the small curves on the left region show the hip motion, whilst the large curves on the right show the knee motion. Clearly, the knee motion is characterized by a large range of motion. However, the results of Figure 4.12 show that the new approach works successfully, comparing well with biomechanical data (*see* Figure 4.2) acquired by a marker-based system. The variance in Figures 4.12(c) and (d) would appear to be smaller than for Winter's [119] analysis. These results concern 100



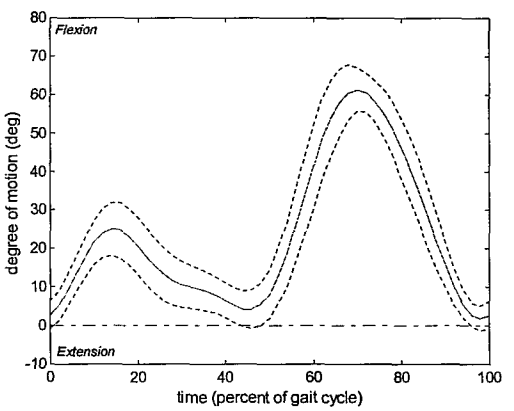

(a) Hip Angles vs. Time (100 Subjects)



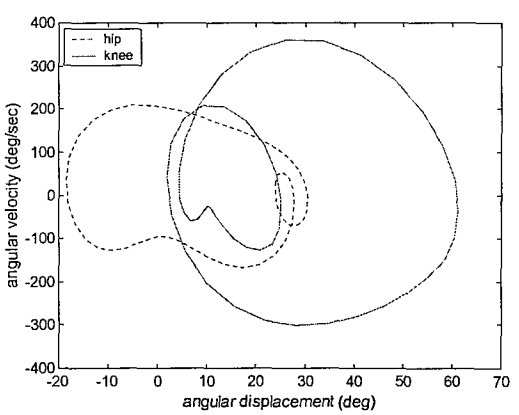

(b) Knee Angles vs. Time (100 Subjects)



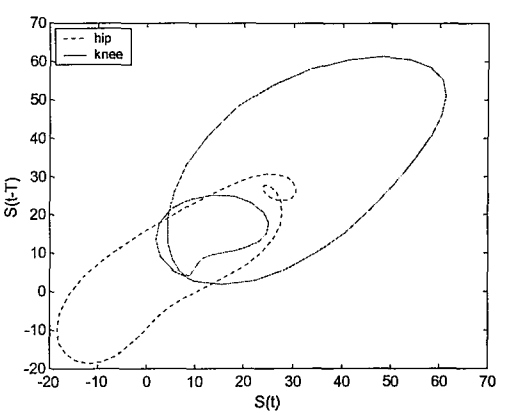

(c) Hip Angles vs. Time (Mean & Std.)




(d) Knee Angles vs. Time (Mean & Std.)




(e) Phase-Space Portrait (Mean)




(f) Delay Coordinate Plot (Mean, T=0.08)

Figure 4.12: Extracted Gait Motion from the SOTON Database

subjects with seven sequences of each subject, a total of 700 sequences, i.e. a much larger volume of data than in that analysis. Further, in this analysis, subjects were not supervised, and carried no markers, allowing for relaxed walking patterns. This is also reflected in the small number of traces that lie outside of the general trend. However, it can clearly be seen that the general trend is followed by most of the traces, suggesting that earlier analyses could be revised by this new approach.

### *4.4.2 Conclusions*

In the previous chapter, the gait signatures are extracted from the SOTON database. The sequence of the gait signatures contains the general and kinematic parameters on human movement. The kinematic parameters are described by rotation angles around the joints over time. To detect the periodicity of the gait time series, phase-space portrait and cyclogram are used, and the periodic components of gait motion are analyzed by the autocorrelation function and Fourier analysis. Also, the gait motion is modelled by trigonometric-polynomial interpolant functions. The prediction of the gait motion is performed by delay coordinates embedding, and statistical moments are used to describe a scale of the gait motion. However, the gait signature has much information for describing the gait motion. In future, the parameters and descriptors extracted from the gait signatures can be used as the features for classifying and recognizing humans by gait.

# Chapter 5

# Recognizing Humans by Their Gait

## 5.1 Feature Extraction of Human Gait

The functionality of a pattern recognition system can be divided into two fundamental tasks: description and classification. The description task extracts features of an object, and the classification task uses a classifier to map the features to a group. Thus, features which truly discriminate among groups will assist in identification, while the lack of such features can impede the classification task from arriving at an accurate identification [90]. Feature extraction may be defined as a process to determine how to explicitly describe the object attributes, generally by constructing a set of features representing objects [62]. There is no general solution for extracting features from object data, so various methods can be used to extract the features for a particular domain and application. Here, the features are extracted by analyzing the sequence of gait signatures in the image sequence during a time period, and the features based on motion parameters are mainly considered.

### *5.1.1 Extraction of the Motion Information*

Extracting motion information from an image sequence is an important step in the identification or recognition of humans. In general, there are two methods for extracting planar motions: motion correspondence and optical flow [15]. The motion correspondence method deals with extracting particular points or characteristic features from an image sequence, and that is concerned with motion trajectory. The motion trajectories can be parameterized using several methods such as position and direction,

velocities, acceleration, and spatial-temporal curvature. The generation of motion trajectories from an image sequence is usually related to the feature points in each frame and correspondence of such feature points from one frame to another. The feature points need to be distinctive enough for robust detection, and stable through time so they can be tracked. Many model-based approaches for the analysis of human motion have been developed by using this method [2][39][77].

Optical flow [32] is very common for evaluating motion from an image sequence. The optical flow can be computed from a sequence by considering the displacement of each pixel between two consecutive images. Optical flow was used to derive a gait signature by analyzing the shape of gait motion [72], and features of the motion were derived as the variation of the first and second-order moments of dense optical flow. Several methods [54][72][83] have been used for recognizing human gait motion. In the optical flow method, accurate and dense measurements are difficult to achieve. Figure 5.1 shows an overview of the extraction of motion information from image sequences. Extracting motion information over a region or a whole image is called a region-based feature. However, most features used in motion and object representation are derived from trajectory-based features and optical flow.

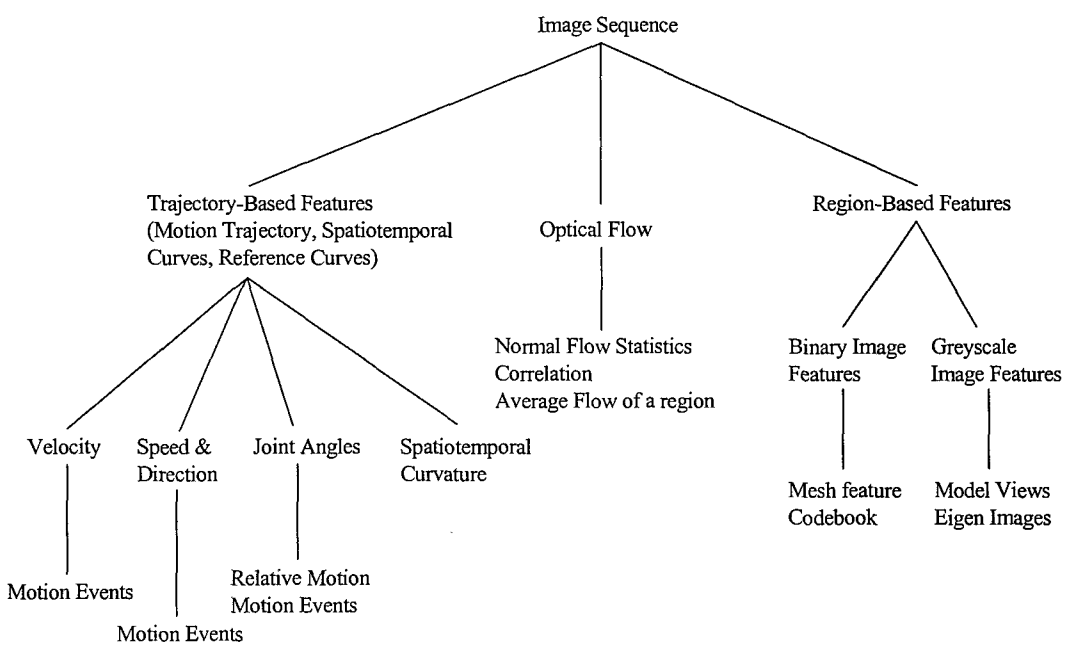


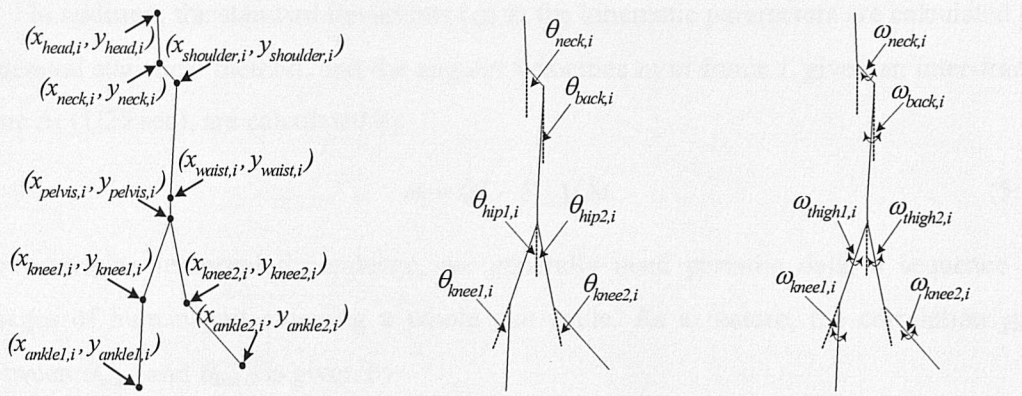

Figure 5.1: Extracting Motion Information from Image Sequences [15]

Human gait data have generally persistent characteristics: high-dimensionality, temporal dependence, curve correlations and non-linear relationships. A set of human gait data may consist of kinematic, kinetic, electromyography (EMG), metabolic and anthropometric variables [18]. The gait variables interact in a complex non-linear form, an observation attributable to the intrinsic non-linear dynamics of human movement [18][19]. However, the vision based approaches have some limitations in the data collection, and hence human gait is usually described by kinematic motion analysis only. In motion analysis of human gait, kinematic characteristics include linear and angular positions, their displacements and the time derivatives, notably the linear and angular velocities and accelerations [89]. These kinematic characteristics can be well represented by trajectory-based features.

### *5.1.2 Extraction of the Gait Features*

In the previous chapters, the sequence of gait signatures has been extracted from an image sequence, and its motion parameters were calculated. Also, the periodic gait motion was modelled by interpolation of trigonometric polynomials. The sequence of gait signatures is assumed to contain the periodic time-dependent information. Therefore, we can extract trajectory-based features from the sequence of gait signatures. The trajectories are basically vector valued functions at each frame of a sequence of gait signatures, and they can be parameterized by topological and parametric feature vectors. The topological feature vector consists of geometry data of human gait motion such as height, width, and position. The parametric feature vector such as cycle time, stride length, linear and angular velocities has variable length since not every component is always visible [17]. In addition, the trajectory-based features have less dimensionality than optical flow.

The nine coordinates that construct the gait signature are the most important topological features, because other parametric features such as the joint angles are calculated based on these coordinates. Therefore, the nine coordinates of a gait signature can be considered as an original feature space. Figure 5.2 shows a graphical demonstration of the spatial-temporal features, such as coordinates of the body points, the joint angles, and the angular velocities. The gait features are basically based on sequences of these spatial-temporal data, and they are extracted by determining an appropriate




(a) Body Points $(x_{l,i}, y_{l,i})$ (b) Joint Angles $\theta_{l,i}$ (c) Angular Velocities $\omega_{l,i}$

Figure 5.2: Graphical Demonstration of Spatial-Temporal Features at Frame $i$

subspace of dimensionality $m$ in the original feature space of dimensionality $d$ ($m \leq d$). Table 5.1 shows the gait features based on motion parameters of the sequence of gait signatures. The general parameters are a basic description of gait motion and can be calculated from each of the gait signatures. Also, the kinematic parameters are obtained from the joint angles, which are interpolated by $4^{th}$-order trigonometric polynomials, during one gait cycle. In general, the kinematic parameters are time series data during the gait cycle, thus mean and standard deviation values of the time series can be used as gait features. Moreover, moments are also used to generate the features, which are invariant to translation and scaling of the hip and knee angles.

Table 5.1: Gait Features based on the Motion Parameters

| Type | Features | Described |
|---|---|---|
| General Parameters | *amplitude* | in Section 3.1.2 |
| | *body height* | in Section 3.4.1 |
| | *cycle time (or cadence), stride length, speed* | in Section 4.1.1 |
| Kinematic Parameters | *mean*($\lvert\theta_{neck}\rvert$), *mean*($\lvert\theta_{back}\rvert$), *mean*($\lvert\theta_{hip}\rvert$), *mean*($\lvert\theta_{knee}\rvert$) | in Section 4.2.1 |
| | *mean*($\lvert\omega_{hip}\rvert$), *mean*($\lvert\omega_{knee}\rvert$), $\sigma(\lvert\omega_{hip}\rvert)$, $\sigma(\lvert\omega_{\_knee}\rvert)$ | in Section 5.1.2 |
| | $\sigma(\lvert\theta_{neck}\rvert)$, $\sigma(\lvert\theta_{back}\rvert)$, $\sigma(\lvert\theta_{hip}\rvert)$, $\sigma(\lvert\theta_{knee}\rvert)$, $\gamma_{knee}$ | in Section 5.1.2 |
| Moments | $\mu_x$, $\mu_y$ | in Section 4.3.3 |
| | $\eta(2,0)$, $\eta(1,1)$, $\eta(0,2)$, $\eta(3,0)$, $\eta(2,1)$, $\eta(1,2)$, $\eta(0,3)$ | in Section 4.3.3 |

In addition, the standard deviations ($\sigma$) in the kinematic parameters are calculated by a general statistical method, and the angular velocities $\omega_i$ at frame $i$, given an inter-frame time $\Delta t$ (1/25 sec), are calculated by

$$\omega_i = (\theta_i - \theta_{i-1})/\Delta t. \tag{5.1}$$

To determine temporal dependence, we generally need periodic data, a sequence of images of human gait covering a whole gait cycle. As a feature, the correlation $\gamma_{knee}$ between $\theta_{knee1}$ and $\theta_{knee2}$ is given by

$$\gamma_{knee} = \frac{\sum_i (\theta_{knee1,i} - \bar{\theta}_{knee1})(\theta_{knee2,i} - \bar{\theta}_{knee2})}{\sqrt{\sum_i (\theta_{knee1,i} - \bar{\theta}_{knee1})^2}\sqrt{\sum_i (\theta_{knee2,i} - \bar{\theta}_{knee2})^2}}. \tag{5.2}$$

Symmetry is one of the characteristics of a walking gait [101], so the correlation between the leading and the following leg is an important factor in gait. Changes in height reflect the magnitude of the right and left steps. Namely, large angles suggest a smaller height than normal, and the characteristics of this relationship need to be verified.

On the other hand, the features contain body height, cycle time, stride length, speed, average angles of the hip and knee, variation of the angles, and moments. Their features vary in dynamic range and units. Thus, features with large values may have a larger influence in the classifier than features with small values [112]. This problem can be overcome by normalizing the features so that their values lie within similar ranges. A straightforward method is linear normalization via the respective estimates of the minimum and maximum of the feature values and can be given by

$$\upsilon_n = T_{\min} + \left( \frac{\upsilon_o - V_{\min}}{V_{\max} - V_{\min}} (T_{\max} - T_{\min}) \right) \tag{5.3}$$

where $T_{max}(1.0)$ and $T_{min}(0.0)$ are the maximum and minimum scaled target values, and $V_{max}$ and $V_{min}$ are the maximum and minimum of the original feature values ($\upsilon_o$). In general, a number of classifiers such as $k$-nearest neighbours and back-propagation neural network require normalized input vectors of some standard range, such as 0 to 1 or -1 to 1. Also, the performance of classification or identification systems depends largely on extracting efficacious features to represent object characteristics.

## 5.2 Feature Selection and Classification

In the previous section, several kinds of gait features are extracted from the motion parameters of the sequence of gait signatures. The gait features may contain information that is redundant or superfluous, in which case it is usually required to select a subset to reduce extraneous noise. This process of removing irrelevant and redundant features is known as feature selection. Using feature selection often improves the accuracy and reduces the computational effort involved in classification. Thus, feature selection is an essential step to enhance correct classification in the presence of many irrelevant features and a small number of samples. In addition, many classification methods are available [61], ranging from basic approaches such as *k*-NN (*k*-Nearest Neighbour) algorithm, to a complicated approach such as hidden Markov models or support vector machines. Here, the gait features are selected by statistical analysis, and the selected features are classified by using a *k*-NN classifier.

### *5.2.1 k-NN Classification Algorithm*

The *k*-Nearest Neighbour (*k*-NN) rule is a well-established and nonparametric pattern classification technique [30], and it provides a basic classification method that is fast to compute and produces easily comprehensible results. Suppose that a sample of $N$ feature vectors $S = \{\boldsymbol{x}_j: j = 1, ..., N\}$ has been grouped into $P$ classes $C_1, ..., C_P$. For any new feature vector $\boldsymbol{x}_i$, the task is to assign $\boldsymbol{x}_i$ to a class, and thus recognize it as belonging to that class [74]. The *k*-NN method is to find the distance from $\boldsymbol{x}_i$ to each of its $k$ nearest neighbours, say $\boldsymbol{x}_{j(1)}, ..., \boldsymbol{x}_{j(k)}$, for some fixed $k > 0$. The Euclidean distance between $\boldsymbol{x}_i$ and $\boldsymbol{x}_{j(m)}$, for each $m = 1, ..., k$, is defined as

$$d_m = d(\boldsymbol{x}_i, \boldsymbol{x}_{j(m)}) = \sqrt{\sum_{q=1}^{Q} \left( x_{i,q} - x_{j(m),q} \right)^2} \tag{5.4}$$

where $Q$ is the dimensionality of the feature vector, and $x_{i,q}$ and $x_{j,q}$ are the values of the $q^{\text{th}}$ feature of the samples $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ respectively. This method allows each $d_m$ to vote for the class $p$ to which $\boldsymbol{x}_{j(m)}$ belongs. By this means, a sample can be associated to the nearest or statistically correct class. Sometimes, the *k*-NN rule is referred to as a lazy classifier, because there is no training phase in the classification process.

Essentially, a *k*-NN algorithm measures the Euclidean distance between a test sample and the training samples in feature space and finds the *k* closest neighbours based on some distance metric [30]. That is, training samples can be used to determine the class of a previously ungrouped sample ***S***. A Euclidean distance metric is used to find the *k* training samples closest to the unknown sample. Of these *k* closest samples, whichever occurs most frequently is used to label the ungrouped sample. In practice, a *k*-NN algorithm can be difficult to beat, even with sophisticated techniques. Unlike other common classifiers, a *k*-NN algorithm does not build a classifier in advance. When a new sample arrives, a *k*-NN classifier finds the *k* closest neighbours to the new sample in the training space based on the Euclidean distance [30][38][74]. However, one of the primary advantages of a *k*-NN algorithm is that it is very fast to compute. By using a *k*-NN algorithm and a simple Euclidian distance metric, we can obtain insight into the baseline performance of the new gait classification system.

### *5.2.2 Feature Selection and Evaluation*

Feature selection has many definitions [26], however it usually attempts to reduce the number of features provided to the classification task. There is more than one reason for the necessity to reduce the features to a sufficient minimum. Computational complexity is the obvious one. A related reason is that although two features may carry good classification results when treated separately, there is little gain if they are combined together in a feature vector, because of a high mutual correlation [112]. Therefore, the goals for feature selection are to retain as much of the original information as possible, remove as much as possible of the redundant and irrelevant information that could degrade the classification performance, and reduce the measurement data to variables that are more suitable for discrimination [74]. Consequently, the task of feature selection is to improve the classification rate, so a natural selection rule seems to be to choose the features with lowest individual classification error values.

However, feature selection is often relegated to classification so that usefulness of each feature can be evaluated using an evaluation function in order to select the optimal subset. An evaluation function typically tries to measure the discriminatory ability of a feature or a subset to distinguish the different class labels [26]. Here, a statistical distance

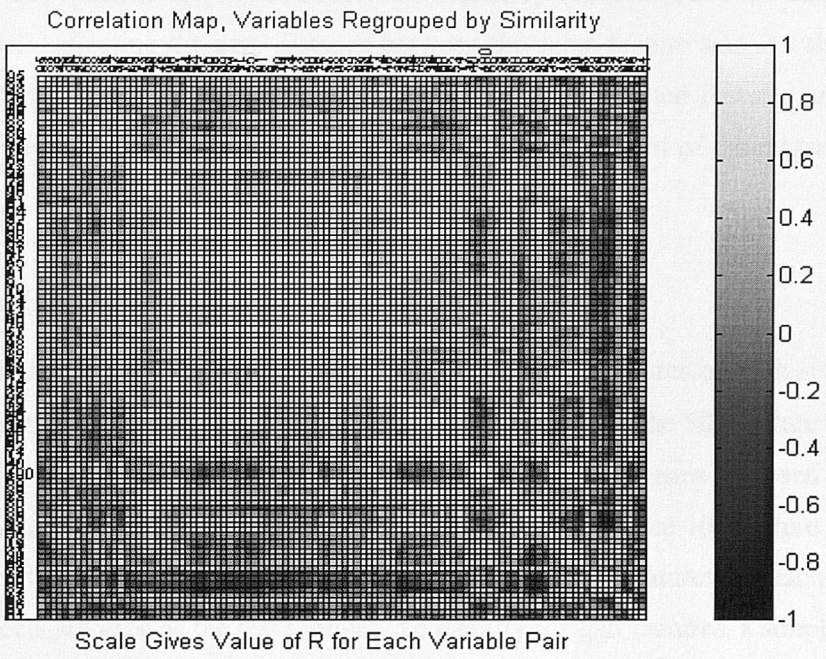
measure that describes the distribution of subjects or classes in the feature space is employed. That is, inter-class separation due to mean-difference with respect to the class covariances is measured by a variation of the Bhattacharyya distance [38][46]. The separation between the two classes $i$ and $j$, for a given feature, is given by

$$\boldsymbol{S}_{i,j} = \left[\boldsymbol{m}_i - \boldsymbol{m}_j\right]^T \left[\frac{\Sigma_i + \Sigma_j}{2}\right]^{-1} \left[\boldsymbol{m}_i - \boldsymbol{m}_j\right] \tag{5.5}$$

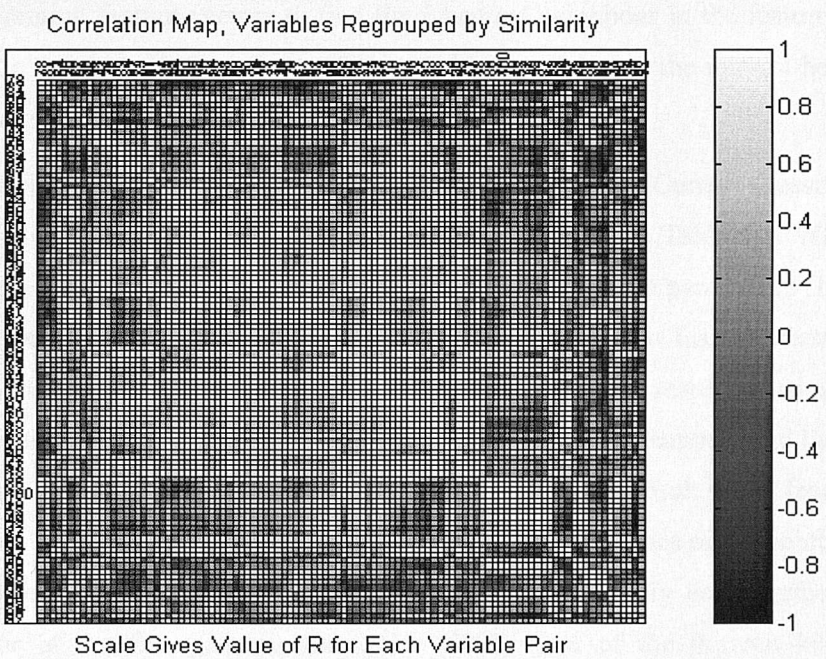
where $\boldsymbol{m}_i$ is the class mean and $\Sigma_i$ is the covariance matrix of class $i$, with equivalent terms for class $j$. To measure a class separability of the given feature, a mean and standard deviation value of $\boldsymbol{S}_{i,j}$ for each feature is calculated, and the larger values of the ratio of mean to standard deviation ($\hat{S}^2/\sigma$) imply a good class separability.

However, treating features individually as scalars has the advantage of computational simplicity but may not be effective for complex problems and for features with high mutual correlation [112]. Therefore, the discrimination effectiveness of feature vectors is evaluated by a classification error rate for each combination of features, although the best features are selected by the class separability measures. In practice, the relationship between feature vector and classification error rate in a classifier can be analyzed by removing relevant input features. In addition, the classification probability for the feature vector combinations may increase the complexity requirements depending on the classifiers. In order to reduce complexity, a number of efficient searching methods have been suggested [26][62], and the sequential forward selection method [112] is used here with $k$-NN classifier. As a result, 17 features are removed from 27 gait features shown in Table 5.1, and 10 features such as *body height*, *cycle time*, *stride length*, *speed*, *mean*($|\theta_{neck}|$), *mean*($|\theta_{back}|$), $\sigma(|\theta_{hip}|)$, $\gamma_{knee}$, $\mu_x$, $\mu_y$ are selected as gait features.

Correlation and covariance matrices provide the basis for all classical multivariate techniques, because they provide sufficient statistics under multivariate normal linear models [36]. Thus, the relationships among the variables may be more readily understood. Figure 5.3 shows the correlation maps for feature vectors of 100 subjects. The correlation map of the correlation matrix $\boldsymbol{R}$ is drawn using Corrgrams [36] and re-ordered by magnitudes of correlation in the correlation map. In the figure, each cell is shaded with the intensity of colour scaled 0 to 100, in proportion to the magnitude of the




(a) Correlation Map using All 27 Features




(b) Correlation Map using 10 Selected Features

Figure 5.3: Correlation Maps for Feature Vectors of 100 Subjects

correlation. Thus, a darker cell shows better discriminatory capability, and the diagonal is the brighter cell reflecting the zero distance between the same feature sets. As shown in Figure 5.3(b), the inter-class correlation decreased when the selected features are used, hence selected feature vectors can lead to performance improvement of the classification system.

### *5.2.3 Gait Classification and Analysis*

In gait classification, 100 different subjects with seven gait signatures of each subject, a total of 700 gait signatures (≅19,534 images), are used. A total of the 500 feature vectors extracted from the front four of the seven signatures and their means for each of 100 different subjects are used as the training samples. Also, a total of the 100 feature vectors extracted from the means of the remaining three of the seven signatures for each of 100 different subjects are used as the test samples. To classify the gait features, a simple $k$-NN algorithm described in the section 5.2.1 is employed as a classifier. This method uses the Euclidean distance between the position of a test feature vector and the position of the surrounding training feature vectors to find the $k$ nearest neighbour in the feature space. Undoubtedly, a more sophisticated classifier would be prudent, but the interest here is to examine the genuine discriminatory ability of the features.

In the $k$-NN with values of $k$ = 1, 3, and 5, the results of CCR (Correct Classification Rate) for using different kinds of the feature vectors are shown in Table 5.2. When not considering the number of features, the features based on kinematic parameters show the best performance in discriminatory capability. By combining two feature vectors, the performance improves to nearly 78% CCR which is similar to the result by using all 27 features. Also, the best performance comes from only 10 selected features, and the result shows an 84% CCR which is 7 percentage points higher then the result by 27 features as shown in Table 5.2. As can be seen in the table, a CCR offer declines only slightly when the value of $k$ increases, so the selected feature vector shows a very good performance. However, four of the five general parameters (80%), four of the thirteen kinematic parameters (31%), and two of the nine moments (22%) are selected as important gait features. This result confirms that individuality by gait concerns a variety of components. Intuitively, classification by gait concerns the human body shape and its motion. That is

Table 5.2: Classification Rate by Selected Feature Vectors

| Feature Vector (Number of Features) | Correct Classification Rate (CCR) | | |
|---|---|---|---|
| | *k=1* | *k=3* | *k=5* |
| General Parameters (5) | 19.0% | 23.0% | 24.0% |
| Kinematic Parameters (13) | 53.0% | 49.0% | 52.0% |
| Moments (9) | 36.0% | 28.0% | 29.0% |
| | | | |
| General + Kinematic Parameters (18) | 78.0% | 71.0% | 75.0% |
| Kinematic Parameters + Moments (22) | 60.0% | 50.0% | 51.0% |
| Moments + General Parameters (14) | 58.0% | 54.0% | 53.0% |
| | | | |
| All Parameters (27) | 77.0% | 75.0% | 73.0% |
| Selected Features A (18) | 83.0% | 79.0% | 76.0% |
| Selected Features B (10) | 84.0% | 80.0% | 82.0% |

confirmed here by selections of the parameters *body height*, $\mu_x$, $\mu_y$ that describe shape and *cycle time*, *stride length*, *speed*, *mean*($|\theta_{neck}|$), *mean*($|\theta_{back}|$), $\sigma(|\theta_{hip}|)$, $\gamma_{knee}$, which describe motion factors. Other studies have clearly confirmed classification capability on this database, using more features than given here to achieve a higher classification rate than our approach gives. However, their approaches are driven purely by biometric capability whereas here we sought to derive confidence in our biomechanical measure indirectly by classification assessment. Since we derive classification capability by these biomechanical measures this gives further confidence in the validity of these marker-less derived measures.

In Section 2.4.1, the quality levels of the pre-processed image sequences, which are body contour data of the SOTON database, were graded as *A* (good), *B* (fair), and *C* (bad) to be 30%, 40%, and 30% of subjects, respectively. To analyze the relationship between the quality level and a CCR, 30 subjects are selected from each of the quality levels and classified separately by using the *k*-NN algorithm. Table 5.3 shows the classification results for each of the quality levels. In the experiments, seven image sequences for each subject with the 10 best features are also used. As can be seen in the table, a CCR is considerably affected by the quality level of the pre-processed image sequences, namely

Table 5.3: Classification Results by Pre-processed Image Quality

| Body Contour Quality | # of Subjects | Correct Classification Rate (CCR) | | |
|---|---|---|---|---|
| | | *k=1* | *k=3* | *k=5* |
| Class A (good) | 30 | 96.7% | 93.3% | 96.7% |
| Class B (fair) | 30 | 93.3% | 90.0% | 86.7% |
| Class C (bad) | 30 | 90.0% | 80.0% | 90.0% |
| | | | | |
| Class A + B | 60 | 91.7% | 86.7% | 85.0% |
| Class A + C | 60 | 90.0% | 83.3% | 90.0% |
| Class B + C | 60 | 83.3% | 80.0% | 80.0% |
| | | | | |
| All Subjects | 90 | 85.6% | 81.1% | 84.4% |

the better performance of a CCR comes from the better quality level of pre-processed images. Note also, that the poorly extracted silhouettes appear to reduce recognition capability (rather than the good extraction can improve that by poor extraction) since the recognition rate of the well and the poorly extracted silhouettes follows that of the poorly extracted silhouettes more closely. Consequently, our new methods are achieved the classification rates of up to 96.7% for 30 subjects and 84.0% for 100 subjects which are very good performance compared with other studies [22][48][64][65][70][72][86][106].

## 5.3 Automatic Human Recognition by Gait

Automated person identification is an important task in many security systems such as video surveillance and access control. It is well-known that biometrics are a powerful tool for reliable automated person identification [39][65]. Automatic gait recognition is one of the newest of the emergent biometrics and has many advantages over other biometrics. The most notable advantage is that it does not require contact with the subjects nor does it require the subject to be near a camera. Various approaches [22][64][65][70][72][86] for the classification and recognition of human gait have been studied, but human gait identification is still a difficult task. Here, the gait feature vectors extracted from the gait signatures of the SOTON database are used for the recognition experiments. Also, an enhanced back-propagation algorithm for training multi-layered neural network, based on

selective retraining and a dynamic adaptation of learning rate and momentum, is employed to recognize the gait.

### *5.3.1 Human Gait Identification System*

In general, a biometric system operates either as a verification system or as an identification system. A verification system authenticates a person's identity by comparing the input biometric characteristic with a person's own biometric data pre-stored in the system, so the system either rejects or accepts the submitted claim for authentication. An identification system recognizes an individual by searching the template database for a match [61]. However, a typical human gait identification system can be divided into training and recognition modules. The training module is responsible for making a trained database to identify a person. During the training phase, the gait

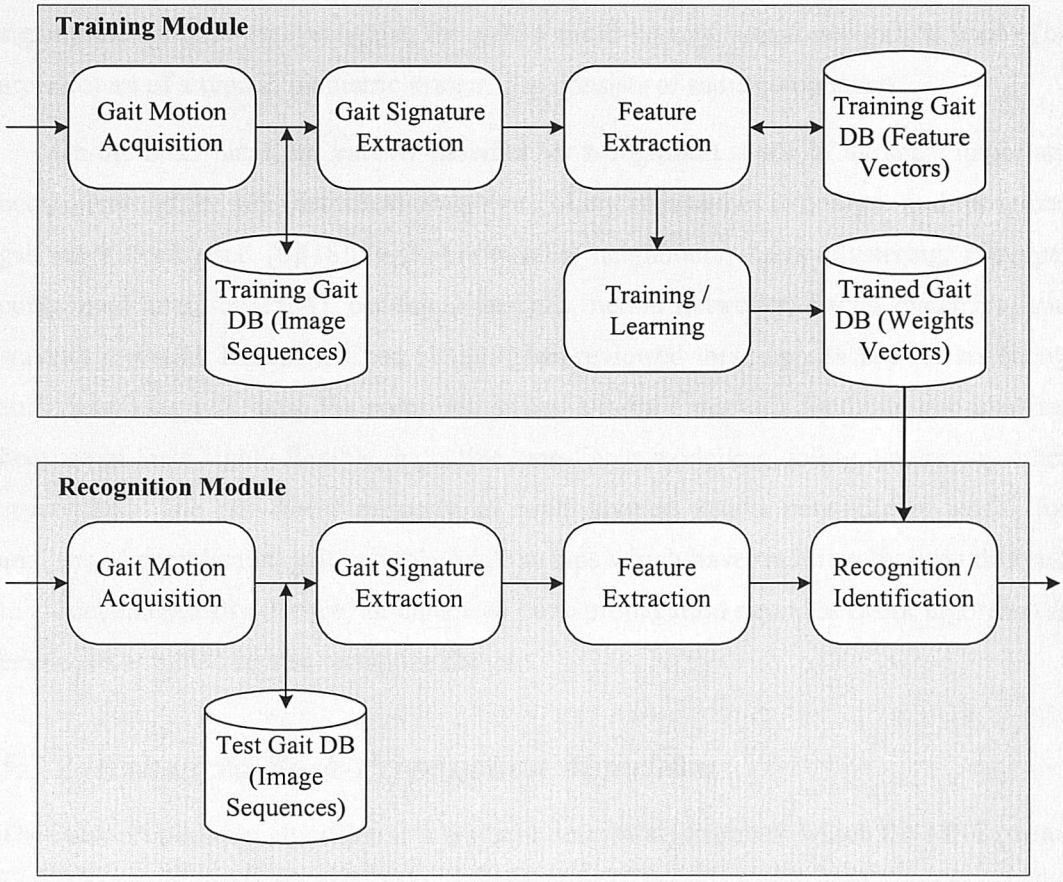



Figure 5.4: Block Diagram of Gait Identification System

motion is captured by a video camera for acquiring a digital representation of the characteristic such as the gait signatures. A feature extractor processes this representation to generate a more compact and expressive representation such as gait feature vector. The feature vectors for each person are then trained by a pattern recognition algorithm, and the trained results will be stored in a gait identification system's database.

In addition, the recognition module is responsible for identifying the person. During the recognition phase, the video camera captures the gait motion of the person to be identified, and it converts into the same sort of feature vector as in training. After that, the feature vector will be submitted to the recognizer, which automatically computes it against the trained database to determine the identity of the individual. Figure 5.4 shows the block diagram of our gait identification system. As described in the previous sections, the gait signatures are extracted from the image sequences of the SOTON database, and the feature vectors are extracted from the motion parameters of the sequence of gait signatures. To train and recognize the gait, a multi-layered neural network is used. The architecture of a typical biometric system also consists of same components.

On the other hand, the pattern classifier (or recognizer) is one of the most important components of the gait identification system. Many approaches to analyze and recognize gait have been used [6][18][19][86]: *k*-nearest neighbours, fuzzy clustering, principal component analysis (PCA), canonical analysis, neural networks, fractal dynamics, and wavelet methods. Recently, Chau [18][19] has reviewed these approaches for analyzing and classifying gait data. He notes that neural network methods facilitate gait analysis because of their highly flexible, inductive, non-linear modelling ability, unlike any other approaches. The non-linear property of multi-layered neural networks is useful for analysis of complicated gait variable relationships which have traditionally been difficult to model analytically. Hence, an enhanced back-propagation neural network algorithm is employed as a classifier to recognize the gait.

### *5.3.2 Enhanced Back-Propagation Algorithm*

The back-propagation algorithm is a gradient descent algorithm in which the MSE (mean square error) is employed [107][116] for minimizing the error in weight-error space. Let $d_{pk}$ be the value of *k-th* output node for *p-th* input pattern, and $o_{pk}$ be the actual output.

Then, an RMS (root-mean square) error measure is derived by normalizing the MSE as

$$E_{rms} = \frac{1}{PK}\sqrt{\sum_{p=1}^{P}\sum_{k=1}^{K}(d_{pk} - o_{pk})^2} \tag{5.6}$$

where $P$ is the number of training patterns, and $K$ denotes the number of nodes in the output layer. In general, $E_{rms}$ is more descriptive than the MSE for comparing the training results of the algorithms and thus is more effective in measuring the accuracy of mapping and association [60]. Consequently, $E_{rms}$ can be used as an error measure in the back-propagation algorithm which continues the training processes until it becomes less than the predetermined tolerance.

However, an algorithm that uses $E_{rms}$ and a predetermined tolerance has two serious problems. First, even though most input patterns are not responsible for the RMS error, we should continue the training processes because of the error caused by some patterns. Second, as $E_{rms}$ is used as an error measurement, the degree of learning obtained for each pattern is not accurately reflected. One of the solutions to the problem might be to calculate the average RMS for all training patterns and the individual RMS for each pattern, and then to train specific patterns which have a greater RMS than the average. In many cases, the weights incorrectly fit the actual output of specific patterns. Incorrect fitting can be detected by identifying the output node $k$ which has the maximum error for pattern $p$, where can be defined as

$$E_{pk_{\max}} = \max_{k=1}^{K}\left(\left|d_{pk} - o_{pk}\right|\right). \tag{5.7}$$

In a conclusion, retraining which reflects the characteristic of each pattern can be achieved by detecting incorrect fittings and by using error measurements of $E_{pkmax}$ and $E_{rms}$. This method may not only reduce training time, but also increase recognition rate by selective retraining.

On the other hand, weights in the back-propagation algorithm are recursively adjusted with a set of pairs (input values and corresponding output values) until the value of the difference between the desired output and the actual output is less than the predetermined tolerance. Weight adjustment is determined based on the generalized formula

$$\Delta W_{ji}(t) = \eta(\delta_j o_i) + \alpha \Delta W_{ji}(t-1) \tag{5.8}$$

where $\eta$ is the learning rate, and $\alpha$ denotes the momentum term [93][107]. As the learning rate becomes larger, the change in the weight becomes larger, and training with a larger learning rate might finish earlier. However, in that case convergence is not guaranteed, because oscillation can arise. Therefore, the learning rate should be maximized for speedy convergence within a range to prevent oscillations [99][116]. The momentum term provides speedy training while preventing oscillation and indicates the size of weight adjustment based on previous changes of weight. Oscillation can be detected by analyzing error curves, namely, by irregular change in the error measurement term $E_{rms}$. In addition, oscillation should be detected within a predetermined interval (number of epochs) to be applicable to dynamic adaptation of learning rate and momentum. This algorithm may be effective in training relatively complex patterns by detecting oscillations and quickly adapting to them.

The usefulness of the proposed method is demonstrated in experiments with the XOR and Encode problems [74][99]. Table 5.4 shows comparison results of the standard algorithm and the proposed algorithm. In the test, after training 100 patterns, investigated was average RMS values when the training iterations reach 500, 1000, and 2000 epochs, average recognition rate measured after 2000 epochs, and recognition rate for 1000 new test patterns. Performance results such as the number of iterations and convergence speed are sensitive to the initial weights. Therefore, the same set of initial weights was used for comparing the two algorithms, and the performance tests were repeated multiple times for preventing statistical biases. As can be seen in the table, the proposed back-propagation

Table 5.4: Performance Results for Enhanced BP Algorithm

| Neural Nets Algorithm | Task | Network Topology | Average RMS | | | Average Correct | Gen. Test |
|---|---|---|---|---|---|---|---|
| | | | 500 | 1000 | 2000 | | |
| Standard BP | XOR | 2×3×1 | 0.0174 | 0.0145 | 0.0133 | 99% | 92.0% |
| | Encode | 8×3×8 | 0.0133 | 0.0100 | 0.0093 | 98% | 72.4% |
| Enhanced BP | XOR | 2×3×1 | 0.0083 | 0.0040 | 0.0024 | 100% | 95.3% |
| | Encode | 8×3×8 | 0.0150 | 0.0138 | 0.0120 | 97% | 74.3% |

algorithm demonstrated better performance than the standard algorithm in solving the XOR problem which has only a single output node. In solving the Encode problem, convergence speed was a little slow, but the degree of generalization was increased. Modification of weights through selective retraining reduced computation complexity and eventually decreased training time. The back-propagation algorithm has been known to be useful in training multi-layered neural networks, and thus has been effectively applied to various fields. However, disadvantages of the algorithm are that it requires a large computational time for training and possibly converges into a local minimum in the training process [99][107][116].

### *5.3.3 Recognition of Human Gait*

An automated pattern recognition system minimally contains an input subsystem that accepts sample pattern vectors and a decision-maker subsystem that decides the classes to which an input pattern vector belongs [74]. If it also classifies, then it has a training phase in which it learns a set of classes of the population from a sample of pattern vectors, namely, it partitions the population into the subpopulations that are classes. As described

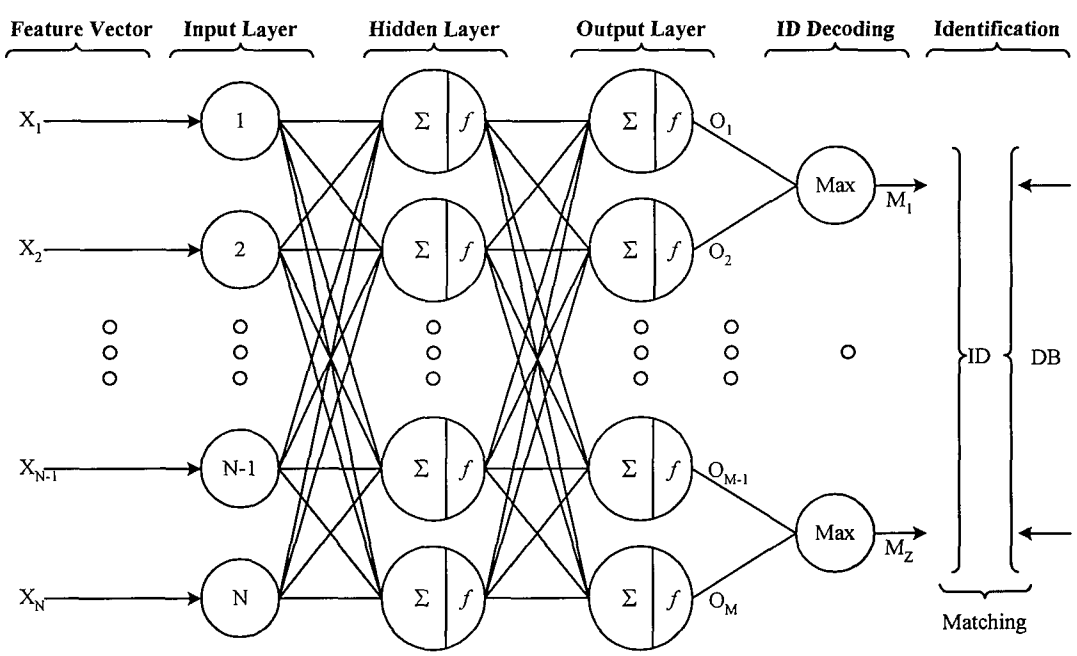



Figure 5.5: A Two-Layer Neural Network for Gait Recognition

in Section 5.3.1, a multi-layered feed-forward neural network is employed here to train and recognize the human gait. Figure 5.5 shows the network architecture used within this study. In the figure, **X** is the input feature vector with $N$ elements, and **O** is the output vector with $M$ elements. The neural network has one hidden layer of sigmoid nodes followed by an output layer of linear nodes. To train the network, the enhanced back-propagation algorithm described in the previous section is used. Also, the nodes of output layer are divided into two groups, and information about a maximum output node of each group is used to decode the output to an identification code of gait.

In the experiments, 90 different subjects with seven gait signatures of each subject, a total of 630 gait signatures, are used. As described in Section 2.4.1, the subjects are categorized into three different groups according to the quality levels of the pre-processed body contour data, and each group contains 30 different subjects. To apply the neural network, four gait signatures of each of the 30 subjects for each group are used to generate the 150 training feature vectors, and the arithmetic means of the other three gait signatures of each of the 30 subjects for each group are used to generate the 30 test feature vectors as described in Section 5.2.3. Also, the 10 gait features for each subject are used as input data, and the numbers of hidden nodes and output nodes are set to 28 and 13 respectively. To decode the output to identification code, the first digit of the code is determined by a maximum output node among the first three output nodes, and the second digit of the code is determined by the other output nodes.

Table 5.5: Recognition Results by Pre-processing Quality

| Body Contour Quality (# of Subjects) | # of Feature Vectors | | Recognition Rate | |
|---|---|---|---|---|
| | Training | Test | Training Phase | Recog. Phase |
| Class A - Good (30) | 150 | 30 | 100.0% | 90.0% |
| Class B - Fair (30) | 150 | 30 | 100.0% | 83.3% |
| Class C - Bad (30) | 150 | 30 | 100.0% | 83.3% |

To recognize humans by their gait, the neural networks are trained until recognition on the training data reached 100 percent, thus the classification rates for each group of the training sets were 100%. Table 5.5 shows the recognition results for each group of the different quality levels. As can be seen in the table, the better the quality of results of the

pre-processing stage, the better the performance of the recognition rate, and the neural network approach achieved a recognition rate of up to 90% for 30 subjects. Although the recognition rate does not reach 100%, this preliminary study does suggest the possibility of this gait biometric. Naturally we seek to improve the recognition rate in future. Notwithstanding this, the marker-less gait recognition can clearly handle a small number (30) of subjects successfully as shown in other studies, often for an even smaller population. By this, these approaches show that people are unique by their walking pattern, according with earlier biomechanical suggestions, and buttressing other similar results.

## 5.4 Results and Conclusions

Although there are extensive studies on the biomechanics of human gait motion, they have been mainly interested in analyzing movements for clinical application and athletic performance, and not for recognition purposes. However, one of the most distinctive characteristics of human gait is the fact that it is individualistic. As a biometric, human gait may be defined as a means of identifying individuals by the way they walk. For a gait biometric to be efficacious, the feature selection is a critical task, because improperly selected features frequently lead to low classification rate and require complex classification algorithms. In addition, there have been several attempts to automatically recognize a person's gait or diagnose a walking condition with neural networks. Here, a simple $k$-NN classifier is used in attempt to reveal the genuine discriminatory capability of the selected feature sets. The enhanced back-propagation algorithm is also employed for recognizing the gait.

### *5.4.1 Experiment Results*

As described in Section 5.2.3 and 5.3.3, the SOTON database, which equals in size the largest contemporaneous gait database, is used in the experiments. The database contains 100 different subjects with seven image sequences of each subject, and the gait signatures are extracted from the image sequences of the database as described in Chapter 3. By analyzing the sequence of gait signatures, 27 gait features based on motion parameters are

considered first, and only 10 important features are selected by the feature selection method. The selected feature set includes most of general parameters, the neck and back angles, dynamic of the hip angles, the correlation coefficient between the left and right knee angles, and the centre coordinates of the hip-knee cyclogram. To apply the *k*-NN classifier, the selected feature vector shows 84% CCR for 100 subjects, which is better performance than original 27 features in discriminatory capability as shown in Table 5.2.

Table 5.6: Classification and Recognition Results

| # of Subjects | # of Image Sequences | | # of Feature Vectors | | Classification Rate (%) | | |
|---|---|---|---|---|---|---|---|
| | Training | Test | Training | Test | *k*=1 | *k*=3 | *k*=5 |
| 30 | 120 | 90 | 150 | 30 | 96.7 | 93.3 | 96.7 |
| 60 | 240 | 180 | 300 | 60 | 91.7 | 86.7 | 85.7 |
| 100 | 400 | 300 | 500 | 100 | 84.0 | 80.0 | 82.0 |
| | | | | | | | |
| 30 | 120 | 90 | 150 | 30 | 100.0♦ | | 90.0♠ |

♦, ♠: classification rate (♦) and recognition rate(♠), both by neural network approach

On the other hand, the classification and recognition results for each of the subject using the *k*-NN and enhanced back-propagation algorithms described in the previous sections are summarized in Table 5.6. As can be seen in the table, our new approaches achieved the classification rates of up to 96.7% for 30 subjects by the *k*-NN classifier. To apply the enhanced back-propagation algorithm, the classification rate of 100% in training phase and the recognition rate of 90% were achieved for 30 subjects. Also, the early SOTON data, which contains 4 subjects with four image sequences, is used here. In the experiments, three image sequences of each of the four subjects are used as the training set, and the fourth from each as the test data. To classify the gait by both the *k*-NN and the enhanced back-propagation algorithm, the results show a 100% CCR and recognition rate. As such, a marker-less gait recognition system confirms uniqueness – as earlier suggested in biomechanical studies, thus confirming validity of the new methods. In future, we require more efficient features of gait motion for classifying and recognizing the gait activities, and a more sophisticated recognition algorithm.

### *5.4.2 Conclusions*

Gait is the most common human motion, and each person appears to have his or her own characteristic gait pattern. To recognize the human gait, three groups of motion parameters, a total of 27 parameters, are considered as gait features. These are including general (temporal and spatial) parameters, kinematic parameters, and moments. By measuring a class separability of the given feature, only 10 important features for classifying the gait are selected from these feature sets. To analyze the discriminatory capability of the selected features, a *k*-NN algorithm is employed as classifier. Here, higher gait classification performances, which are 97% CCR for 30 subjects and 84% CCR for 100 subjects, have been achieved on the larger SOTON database. In addition, the enhanced back-propagation neural network algorithm is applied to the SOTON database, and recognition rate of 90% for 30 subjects is accomplished. The results achieved give promising performance and higher recognition rates than those of an earlier gait recognition approach.

# Chapter 6

# Conclusions and Future Work

## 6.1 Overall Conclusions

In this thesis, we have presented a new approach to automated marker-less system for describing, analyzing and recognizing the periodic gait motion. Essentially, the marker-less system was achieved by the gait signatures extracted from the image sequences involving human walking. To evaluate our approach, the SOTON database, which was constructed by using normal DV (Digital Video) camera to acquire image sequences, was used in this work. The SOTON database is currently the largest database of its kind, consisting of 114 subjects with at least eight image sequences of each subject. In addition, the subjects were not supervised, and carried no markers, allowing for relaxed walking patterns. From the indoor SOTON database, seven image sequences of each of 100 different subjects (16 females and 84 males), a total of 700 image sequences, were selected, and they have been mainly used for the experiments.

The gait signature has been extracted by combining a statistical approach and topological analysis guided by anatomical knowledge. As a pre-processing stage, background subtraction was used to detect moving human body in an image sequence, and thresholding and morphological filters were used to extract the body contour. To extract the gait signature, a 2D stick figure model with 8 sticks and 6 joint angles was used to represent the human body structure. A stick figure was extracted from the body contour by determining the body points using known anatomical data. The gait signature is a sequence of the stick figures during one period of the gait cycle which is detected by

the symmetry property of the horizontal centre of mass. To improve the robustness of these gait signatures, the grammatical structure of the gait sequence was also analyzed with physical constraints.

The gait signature had a small weakness when one leg occludes the other, and the ankle part in the body contour was affected by noise such as caused by shadows and colour of shoes at the ground. However, this weakness was somewhat overcome by the grammatical analysis of gait sequence, and the kinematic parameters were calculated in the gait signature. The trajectories of the joint angles followed with the earlier results of medical studies. Also, the periodicity of the gait motion was observed by using a phase-space portrait, a cyclogram and delay-coordinates, and analyzed by the autocorrelation function and Fourier analysis. In addition, the periodic gait motion was efficiently modelled by trigonometric-polynomial interpolant functions. Consequently, the gait signature appears to be a very effective and well-defined representation method for describing the human gait, and includes much information for analyzing the gait motion.

On the other hand, the gait features based on several kinds of motion parameters were extracted by analyzing the sequence of gait signature. By using feature selection methods, only 10 important features were selected from the 27 gait features. Then, the *k*-NN classifier was used to analyze the discriminatory ability of the selected features. The results have produced very good classification rates which were 97% CCR for 30 subjects and 84% CCR for 100 subjects. In addition, the enhanced back-propagation algorithm was employed for recognizing the gait, and recognition rates of 90% for 30 subjects have achieved. Moreover, both classifiers have accomplished classification and recognition rates of 100% on the early SOTON data of 4 subjects. Consequently, our new approaches gave very promising and comparable analysis and recognition results, on both the smaller and the larger databases, to those achieved in other studies.

Our main contribution in this thesis has been to present a new marker-less method for analyzing and recognizing the gait motion. The gait signature based on the human body model can describe gait motion in a compact form and can be used in biomechanical and medical applications. The usefulness of the proposed methods has been demonstrated successfully in experiments. Also, the results for the gait classification and recognition show that human gait can indeed be used to recognize people via image sequences.

## 6.2 Future Work

Automatic gait recognition as a biometric is still a difficult task mainly because of the variability of image quality and illumination, complexity of the gait extraction task, occlusion and change of the human body, correlations and non-linear relationships of the gait motion, and high-dimensionality of gait data. The results presented in this work show that the gait signatures appear to have considerable potential in automatic gait recognition and marker-less gait analysis for biomechanical and medical applications. However, the performance evaluation for extracting the gait signatures has not been sufficiently demonstrated in real-world data such as the outdoor SOTON database. Therefore, a more complex vision algorithm has not used in pre-processing stage of image sequences. Also, a theoretical description of the grammatical structure of gait sequence is not enough to improve the gait signature. Furthermore, the kinematic characteristics, which are a very important aspect of gait motion, could be used more efficiently to describe the gait features.

In future, we could improve the grammatical structure with the step symmetry of the gait sequence by using prior knowledge with uncertainty related to animated gait model and the behaviour of the human movements. The animated gait model might be developed by the normal range of human movement guided by the gait signatures. Also, we require a more efficient feature vector for classifying and recognizing the gait activities and humans. The probabilistic characteristics of the gait kinematics such as gait dynamics and their entropy could be considered in future as features from the gait signature. To identify humans by their gait, a more sophisticated recognition algorithm such as hybrid classifier by fusion of multiple classification algorithms should be considered. For practical applications, we will need to consider subjects walking at different angles to the cameras plane of view. Furthermore, various classification and recognition experiments using a real-world database are essential for further performance evaluation in these studies.

In addition, the marker-less approach should be extended to the technique in terms of medical and biomechanical application capability. In this thesis, the gait motion was mostly analyzed by the statistical methods rather than biomechanical approaches. Also,

the anatomical knowledge, which was used to extract the gait signature, is statistical data, thus it does not match to all persons. Alternatively, we can directly measure the body segments for each person, and then the measurement data can be used to extract a more accurate gait signature. The alternative approach can however be achieved by a non-invasive marker-less system. Also, this approach might be applied to security system such as gait authentication system. The databases used in this work were mainly captured under controlled environment, so that the simple vision algorithms were employed here. However, sophisticated computer vision algorithms are very important components for further performance of applications.

In summary, further work should concentrate both on improvement of the gait signatures and on development of the efficient feature vector with evaluation and experiments using real-world data to expose gait as a biometric. Also, we seek to extend the marker-less technique in practical applications of clinical and biomechanical areas.

# References


[1] J. K. Aggarwal, Q. Cai, W. Liao, and B. Sabata, "Nonrigid Motion Analysis: Articulated and Elastic Motion", *Computer Vision and Image Understanding*, **70**(2), pp.142-156, May 1998.

[2] J. K. Aggarwal and Q. Cai, "Human Motion Analysis: A Review", *Computer Vision and Image Understanding*, **73**(3), pp.428-440, March 1999.

[3] K. Akita, "Image Sequence Analysis of Real World Human Motion", *Pattern Recognition*, **17**(1), pp.73-83, 1984.

[4] Y. Aloimonos and C. Fermuller, "Action Representation", in *Proceedings of AISB'03 Symposium on Biological-Inspired Machine Vision, Theory and Application*, pp.3-23, University of Wales, Aberystwyth, UK, April 2003.

[5] N. I. Badler and S. W. Smoliar, "Digital Representations of Human Movement", *Computing Survey*, **11**(1), pp.19-38, 1979.

[6] J. G. Barton and A. Lees, "An Application of Neural Networks for Distinguishing Gait Patterns on the Basis of Hip-Knee Joint Angle Diagram", *Gait and Posture*, **5**(1), pp.28-33, 1997.

[7] C. BenAbdelkader, Ross Cutler, and Larry Davis, "Stride and Cadence as a Biometric in Automatic Person Identification and Verification", in *Proceedings of the 5th IEEE International Conference on Face and Gesture Recognition*, pp.357-362, Washington D.C., USA, May 2002.

[8] B. I. Bertenthal and J. Pinto, "Complementary Process in the Perception and Production of Human Movements", *A Dynamic System Approach to Development: Applications*, L. B. Smith and E. Thelen, Eds., pp.209-239, MIT Press, 1993.

[9] B. Bhanu and J. Han, "Individual Recognition by Kinematic-based Gait Analysis", in *Proceedings of International Conference on Pattern Recognition*, Vol.3, pp.343-347, Quebec, Canada, August 2002.

[10] G. Bharatkumar, K. E. Daigle, M. G. Pandy, Q. Cai, and J. K. Aggarwal, "Lower Limb Kinematics of Human Walking with the Medial Axis Transformation", in *Proceedings of IEEE Nonrigid and Articulated Motion Workshop*, pp.70-76, TX, USA, October 1994.

[11] B. Bodenheimer, C. Rose, S. Rosenthal, and J. Pella, "The Process of Motion Capture: Dealing with the Data", in *Proceedings of Computer Animation and Simulation-Eurographics Animation Workshop*, pp.3-18, September 1997.

[12] C. Bregler and J. Malik, "Video Motion Capture", *Technical Report: CSD-97-973*, University of California at Berkeley, 1997.

[13] A. Bruderlin and T. W. Calvert, "Goal-Directed, Dynamic Animation of Human Walking", *ACM Computer Graphics*, **23**(3), July 1989.

[14] J. N. Carter and M. S. Nixon, "On Measuring Trajectory-Invariant Gait Signatures", in *Proceedings of International Archives of Photogrammetry and Remote Sensing*, **33**(B5), pp.114-121, Amsterdam, 2000.

[15] C. Cédras and M. Shah, "Motion-Based Recognition: A Survey", *IEE Proceedings – Image and Vision Computing*, **13**(2), pp.129-155, March 1995.

[16] K. Chakraborty, K. Mehrotra, C. K. Mohan, and S. Ranka, "Forecasting the Behavior of Multivariate Time Series using Neural Networks", *Neural Networks*, Vol.5, pp.961-970, Pergamon Press, 1992.

[17] C. Chang and C. L. Huang, "The Model-Based Human Body Motion Analysis System", *Image and Vision Computing*, **18**(14), pp.1067-1083, November 2000.

[18] T. Chau, "A Review of Analytical Techniques for Gait Data. Part 1: Fuzzy, Statistical and Fractal Methods", *Gait and Posture*, **13**(1), pp.49-66, 2001.

[19] T. Chau, "A Review of Analytical Techniques for Gait Data. Part 2: Neural Network and Wavelet Methods", *Gait and Posture*, **13**(2), pp.102-120, 2001.

[20] T. Q. Chen and Y. Lu, "Color Image Segmentation-An Innovative Approach", *Pattern Recognition*, **35**(2), pp.395-405, February 2002.

[21] Z. Chen and H. J. Lee, "Knowledge-Guided Visual Perception of 3-D Human Gait from Single Image Sequence", *IEEE Transactions on Systems, Man, and Cybernetics*, **22**(2), pp.336-342, 1992.

[22] R. T. Collins, R. G. Gross, and J. Shi, "Silhouette-based Human Identification from Body Shape and Gait", in *Proceedings of the 5th IEEE International Conference on Face and Gesture Recognition*, pp.351-356, Washington D.C., USA, May 2002.

[23] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic Extraction and Description of Human Gait Models for Recognition Purposes", *Computer Vision and Image Understanding*, **90**(1), pp.1-41, April 2003.

[24] R. Cutler and L. Davis, "Robust Real-Time Periodic Motion Detection, Analysis, and Applications", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **22**(8), pp.781-796, August 2000.

[25] J. T. Cutting, D. R. Proffitt, and L. T. Kozlowski, "A Biomechanical Invariant for Gait Perception", *Journal of Experimental Psychology: Human Perception and Performance*, **4**(3), pp.357-372, 1978.

[26] M. Dash and H. Liu, "Feature Selection for Classification", *Intelligent Data Analysis*, **1**(3), pp.131-156, March 1997.

[27] W. T. Dempster and G. R. L. Gaughran, "Properties of Body Segments Based on Size and Weight", *American Journal of Anatomy*, **120**, pp.33-54, 1967.

[28] J. Deutscher, A. Blake, and I. Reid, "Articulated Body Motion Capture by Annealed Particle Filtering", in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Vol.2, pp.126-133, South Carolina, USA, June 2000.

[29] R. Drillis and R. Contini, *Body Segment Parameters*. TR. 1163-03, New York University, Under Contract with Office of Vocational Rehabilitation, Department Health, Education and Welfare, New York, 1966.

[30] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, 1973.

[31] R. L. Eubank and P. Speckman, "Curve Fitting by Polynomial-Trigonometric Regression", *Biometrika*, **77**(1), pp.1-9, March 1990.

[32] C. Fermüller, D. Shulman, and Y Aloimonos, "The Statistics of Optical Flow", *Computer Vision and Image Understanding*, **82**(1), pp.1-32, April, 2001.

[33] J. D. Foley, A. Van Dam, S. K. Feiner, and J. F. Hughes, *Computer Graphics: Principles and Practice*, Addition-Wesley, 1990.

[34] D. A. Forsyth and J. Ponce, *Computer Vision: A Modern Approach*, Prentice Hall, 2003.

[35] J. P. Foster, M. S. Nixon, and A. Prügel-Bennett, "Automatic Gait Recognition using Area-based Metrics", *Pattern Recognition Letters*, **24**(14), pp.2489-2497, October 2003.

[36] M. Friendly, "Corrgrams: Exploratory Displays for Correlation Matrices", *The American Statistician*, **56**(4), pp. 316-325, 2002.

[37] H. Fujiyoshi and A. J. Lipton, "Real-Time Human Motion Analysis by Skeletonization", *IEEE Workshop on Application of Computer Vision*, pp.15-21, Princeton, NJ, USA, October 1988.

[38] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd eds, Academic Press, 1990.

[39] D. M. Gavrila, "The Visual Analysis of Human Movement: A Survey", *Computer Vision and Image Understanding*, **73**(1), pp.82-98, January 1999.

[40] N. A. Gershenfeld and A. S. Weigend, "The Future of Time Series: Learning and Understanding", in *Proceedings of the NATO Advanced Research Workshop on Comparative Time Series Analysis,* pp.1-70, Santa Fe, NM, USA, May 1992.

[41] M. Gleicher, "Animation from Observation: Motion Capture and Motion Editing", *Computer Graphics*, **33**(4), pp.51-55, November 1999.

[42] G. Gordon, T. Darrell, M. Harville, and J. Woodfill, "Background Estimation and Removal based on Range and Color", in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Vol.2, pp.2459-2464, Fort Collins, CO., USA, June 1999.

[43] A. Goswami, B. Thuilot, and B. Espiau, *Compass-like Biped Robot Part I: Stability and Bifurcation of Passive Gaits*, INRIA Research Report No. 2996, October 1996.

[44] A. Goswami, "A New Gait Parameterization Technique by Means of Cyclogram Moments: Application to Human Slop Walking", *Gait and Posture*, **8**(1), pp.15-36, 1998.

[45] R. Gross and J. Shi, "The CMU Motion of Body (MoBo) Database", *Technical Report: CMU-RI-TR-01-18*, 2001.

[46] X. Guorong, C. Peiqi, and W. Minhui, "Bhattacharyya Distance Feature Selection", in *Proceedings of the 13th International Conference on Pattern Recognition*, pp.195-199, Vienna, Austria, August 1996.

[47] Haritaoglu, D. Harwood, and L. S. Davis, "$W^4$: Who? When? Where? What? – A Real-Time System for Detecting and Tracking People", in *Proceedings of the International Conference on Face and Gesture Recognition*, pp.222-227, Nara, Japan, April 1998.

[48] J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter, "Automatic Gait Recognition by Symmetry Analysis", *Pattern Recognition Letters*, **24**(13), pp.2175-2183, September 2003.

[49] L. Herda, P. Fua, R. Plankers, R. Boulic, and D. Thalmann, "Using Skeleton-Based Tracking to Increase the Reliability of Optical Motion Capture", *Journal of Human Movement Science*, **20**(3), pp.313-341, March 2001.

[50] J. K. Hodgins and W. L. Wooten, "Animating Human Athletes", in *Robotics Research: The Eighth International Symposium,* Y. Shirai and S. Hirose Eds., Springer-Verlag, pp.356-367, 1998.

[51] D. Hogg, "Model-Based Vision: A Program to See a Walking Person", *Image and Vision Computing*, **1**(1), pp.5-20, 1983.

[52] T. Horprasert, D. Harwood, and L. S. Davis, "A Statistical Approach for Real-Time Robust Background Subtraction and Shadow Detection", in *Proceedings of IEEE International Conference on Computer Vision: Frame-Rate Workshop*, Kerkyra, Greece, September 1999.

[53] M. K. Hu, "Visual Pattern Recognition by Moment Invariants", *IRE Transactions on Information Theory*, **IT-8**(8):179-187, February 1962.

[54] P. S. Huang, C. J. Harris, and M. S. Nixon, "Human Gait Recognition in Canonical Space Using Temporal Templates", *IEE Proceedings – Vision, Image and Signal Processing*, **146**(2), pp.93-100, April 1999.

[55] P. S. Huang, C. J. Harris, and M. S. Nixon, "Recognising Human by Gait via Parametric Canonical Space", *Artificial Intelligence in Engineering*, **13**(4), pp.359-366, 1999.

[56] E. C. Ifeachor and B. W. Jervis, *Digital Signal Processing: A Practical Approach*, 2nd Eds., Prentice Hall, 2002.

[57] V. T. Inman, H. J. Ralston, and F. Todd, *Human Walking*, Williams & Wilkins, 1981.

[58] Y. A. Ivanov and A. F. Bobick, "Recognition of Visual Activities and Interactions by Stochastic Parsing", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, **22**(8), pp.1-21, August 2000.

[59] S. Iwasawa, K. Ebihara, J. Ohya, and S. Morishima, "Real-time Estimation of Human Body Posture from Monocular Thermal Images", in *Proceedings of the IEEE CS Conference on Computer Vision and Pattern Recognition*, Puerto Rico, pp.15-20, 1997.

[60] R. A. Jacobs, "Increased Rates of Convergence through Learning Rate Adaptation", *Journal of Neural Networks*, **1**(2), pp.295-307, 1988.

[61] A. K. Jain, R. M. Bolle, and S. Pankanti, "Introduction to Biometrics", in *BIOMETRICS – Personal Identification in Networked Society*, A. K. Jain, R. M. Bolle, and S. Pankanti, Eds., pp.1-41, Kluwer Academic Publishers, January 1999.

[62] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(1), pp.4-37, January 2000.

[63] G. Johansson, "Visual Perception of Biological Motion and a Model for Its Analysis", *Perception and Psychophysics*, **14**(2), pp.201-211, 1973.

[64] Y. Johnson and A. F. Bobick, "A Multi-View Method for Gait Recognition using Static Body Parameters", in *Proceedings of the 3rd International Conference on Audio-and Video-based Biometric Person Authentication*, pp.301-311, Halmstad, Sweden, June 2001.

[65] A. Kale, A. N. Rajagopalan, N. Cuntoor, and V. Kruger, "Gait-based Recognition of Humans using Continuous HMMs", in *Proceedings of the 5th IEEE International Conference on Face and Gesture Recognition*, pp.336-341, Washington D.C., USA, May 2002.

[66] R. F. M. Kleissen, J. H. Buurke, J. Harlaar, and G. Zilvold, "Electromyography in the Biomechanical Analysis of Human Movement and Its Clinical Application", *Gait and Posture*, **8**(2), pp.143-158, 1998.

[67] I. Koprinska and S. Carrato, "Temporal Video Segmentation: A Survey", *Signal Processing: Image Communication*, **16**(5), pp.477-500, January 2001.

[68] L. T. Kozlowski and J. T. Cutting, "Recognizing the Sex of a Walker from a Dynamic Point-Light Display", *Perception and Psychology*, **21**(6), pp.575-580, 1977.

[69] H. J. Lee and Z. Chen, "Determination of 3D Human Body Postures from a Single View", *Computer Vision, Graphics, and Image Processing*, **30**(2), pp.148-168, 1985.

[70] L. Lee and W. E. L. Grimson, "Gait Analysis for Recognition and Classification", in *Proceedings of the 5th IEEE International Conference on Face and Gesture Recognition*, pp.155-162, Washington D.C., USA, May 2002.

[71] M. K. Leung and Y. H. Yang, "First Sight: A Human Body Outline Labeling System", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **17**(4), pp.359-377, April 1995.

[72] J. J. Little and J. E. Boyd, "Recognizing People by Their Gait: The Shape of Motion", *Videre: Journal of Computer Vision Research*, **1**(2), pp.2-32, Winter 1998.

[73] F. Liu and R. W. Picard, "Finding Periodicity in Space and Time", in *Proceedings of the International Conference on Computer Vision*, pp.376-383, Bombay, India, January 1998.

[74] C. G. Looney, *Pattern Recognition using Neural Networks: Theory and Algorithms for Engineers and Scientists*, Oxford University Press, 1997.

[75] G. Mather and L. Murdoch, "Gender Discrimination in Biological Motion Displays based on Dynamic Cues", in *Proceedings of the Royal Society of London*, Vol.B, pp.273-279, 1994.

[76] J. McNames, "A Nearest Trajectory Strategy for Time Series Prediction", in *Proceedings of the International Workshop on Advanced Black-Box Techniques for Nonlinear Modeling*, pp.112-128, Belgium, July 1998.

[77] T. B. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture", *Computer Vision and Image Understanding*, **81**(3), pp.231-268, March 2001.

[78] H. Murase and R. Sakai, "Moving Object Recognition in Eigenspace Representation: Gait Analysis and Lip Reading", *Pattern Recognition Letters*, **17**(2), pp.155-162, February 1996.

[79] M. P. Murray, A. B. Drought, and R. C. Kory, "Walking Patterns of Normal Men", *Journal of Bone and Joint Surgery*, **46A**(2), pp.335-360, 1964.

[80] M. P. Murray, "Gait as a Total Pattern of Movement", *American Journal of Physical Medicine*, **46**(1), pp.290-333, 1967.

[81] J. M. Nash, M. S. Nixon, and J. N. Carter, "Extraction of Moving Human Articulated-Object by Evidence Gathering", in *Proceedings of British Machine Vision Conference*, Vol.2, pp.609-618, Southampton, 1998.

[82] A. Neri, S. Colonnese, G. Russo, and P. Talone, "Automatic Moving Object and Background Separation", *Signal Processing*, **66**(2), pp.219-232, April 1998.

[83] M. S. Nixon, J. N. Carter, D. Cunado, P. S. Huang, and S. V. Stevenage, "Automatic Gait Recognition", in *BIOMETRICS – Personal Identification in Networked Society*, A. K. Jain, R. M. Bolle, and S. Pankanti, Eds., pp.231-249, Kluwer Academic Publishers, January 1999.

[84] M. S. Nixon and A. S. Aguado, *Feature Extraction and Image Processing*, Butterworth-Heinemann, 2002.

[85] M. S. Nixon, J. N. Carter, J. D. Shutler, and M. G. Grant, "New Advances in Automatic Gait Recognition", *Information Security Technical Report*, **7**(4), pp.23-35, December 2002.

[86] M. S. Nixon, J. N. Cater, M. G. Grant, L. Gordon, and J. B. Hayfron-Acquah, "Automatic Recognition by Gait: Progress and Prospects", *Sensor Review*, **23**(4), pp.323-331, 2003.

[87] S. A. Niyogi and E. H. Adelson, "Analyzing and Recognizing Walking Figures in XYT", in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.469-471, Seattle, WA, USA, June 1994.

[88] C. C. Norkin and P. K. Levangie, *Joint Structure and Function: A Comprehensive Analysis*, 2nd Eds., F. A. Davis Company, 1992.

[89] S. J. Olney and C. L. Richards, "Hemiparetic Gait Following Stroke. Part I: Characteristics", *Gait and Posture*, **4**(2), pp.136-148, 1996.

[90] R. T. Olszewski, *Generalized Feature Extraction for Structural Pattern Recognition in Time-Series Data*, CMU-CS-01-108, Carnegie Mellon University, February 2001.

[91] F. J. Perales, "Human Motion Analysis and Synthesis using Computer Vision and Graphics Techniques: State of Art and Application", *Interdisciplinary Workshops*, Centre of Computer Graphics and Visualization, University of West Bohemia, Czech Republic, April 2002.

[92] J. Perry, *Gait Analysis: Normal and Pathological Function*, Slack, 1992.

[93] M. Pfister and R. Rojas, "Speeding-Up Backpropagation – A Comparison of Orthogonal Techniques" in *Proceedings of the International Conference on Neural Networks*, Nagoya, Japan, Vol.I, pp.517-523, October 1993.

[94] S. T. Pheasant, *Bodyspace: Anthropometry, Ergonomics and Design*, Taylor & Francis, 1988.

[95] P. J. Phillips, "Human Identification Technical Challenges", in *Proceedings of the International Conference on Image Processing*, pp.49-52, Rochester, USA, September 2002.

[96] R. Polana and R. Nelson, "Detection and Recognition of Periodic, Nonrigid Motion", *International Journal of Computer Vision*, **23**(3), pp.261-282, June/July 1997.

[97] P. J. Phillips, S. Sarkar, I. Robledo, P. Grother, and K. Bowyer, "Baseline Results for the Challenge Problem of Human ID using Gait Analysis", ", in *Proceedings of the 5th IEEE International Conference on Face and Gesture Recognition*, pp.137-142, Washington D.C., USA, May 2002.

[98] C. A. Poynton, "Frequently Asked Questions about Color", *http://www.poynton.com/notes/colour_and_gamma/ColorFAQ.txt*, 1998.

[99] J. C. Principe, N. R. Euliano, and W. C. Lefebvre, *Neural and Adaptive Systems: Fundamentals through Simulations*, John Wiley & Sons, 2000.

[100] K. Rohr, "Towards Model-Based Recognition of Human Movements in Image Sequences", *Computer Vision, Graphics, and Image Processing*, **59**(1), pp.94-115, 1994.

[101] H. Sadeghi, P. Allard, F. Prince, and H. Labelle, "Symmetry and Limb Dominance in Able-Bodied Gait: A Review", *Gait and Posture*, **12**(1), pp.34-45, 2000.

[102] T. Sauer, "Time Series Prediction by Using Delay Coordinate Embedding", in *Proceedings of the NATO Advanced Research Workshop on Comparative Time Series Analysis*, pp.175-193, Santa Fe, NM, USA, May 1992.

[103] S. M. Seitz and C. R. Dyer, "Affine Invariant Detection of Periodic Motion", in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp.970-975, Seattle, WA, USA, 1994.

[104] S. M. Seitz and C. R. Dyer, "View-Invariant Analysis of Cyclic Motion", *International Journal of Computer Vision*, **25**(1), pp.1-23, 1997.

[105] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter, "On a Large Sequence-based Human Gait Database", in *Proceedings of Recent Advances in Soft Computing*, pp.66-71, Nottingham, UK, December 2002.

[106] J. D. Shutler and M. S. Nixon "Zernike Velocity Moments for Description and Recognition of Moving Shapes", in *Proceedings of British Machine Vision Conference 2001*, pp.705-714, Manchester, UK, September 2001.

[107] M. Smith, *Neural Networks for Statistical Modeling*, Van Nostrand Reinhold, 1993.

[108] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision*, 2nd Eds., ITP, 1999.

[109] S. V. Stevenage, M. S. Nixon, and K. Vince, "Visual Analysis of Gait as a Cue to Identity", *Applied Cognitive Psychology*, Vol.13, pp.513-526, 1999.

[110] D. J. Sturman, "A Brief History of Motion Capture for Computer Character Animation", *ACM SIGGRAPH 94: Course 9*, 1994.

[111] G. Taga, "A Model of the Neuro-Musculo-Skeletal System for Human Locomotion", *Biological Cybernetics*, **73**(2), pp.97-111, 1995.

[112] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*, Academic Press, 1999.

[113] E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision*, Prentice Hall, 1998.

[114] J. H. Yoo, M. S. Nixon, and C. J. Harris, "Extraction and Description of Moving Human Body by Periodic Motion Analysis", in *Proceeding of ISCA 17th International Conference on Computer and Their Applications*, pp.110-113, San Francisco, USA, April 2002.

[115] L. Wang, W. Hu, and T. Tan, "Recent Developments in Human Motion Analysis", *Pattern Recognition*, **36**(3), pp.585-601, March 2003.

[116] P. J. Werbos, "Backpropagation Through Time: What It Does and How to Do It", *Proceedings of the IEEE*, **78**(40), pp.1550-1560, October 1990.

[117] M. W. Whittle, *Gait Analysis: An Introduction*, 3rd eds., Butterworth Heinemann, 2002.

[118] D. A. Winter, *The Biomechanics and Motor Control of Human Movement*, 2nd Eds., John Wiley & Sons, 1990.

[119] D. A. Winter, *The Biomechanics and Motor Control of Human Gait: Normal, Elderly and Pathological*, 2nd Eds., Waterloo Biomechanics, 1991.

[120] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-Time Tracking of the Human Body", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**(7), pp.780-785, July 1997.

[121] C. Y. Yam, M. S. Nixon, and J. N. Carter, "Automated Person Recognition by Walking and Running via Model-based Approaches", *Pattern Recognition*, **37**(5), pp.1057-1072, May 2004.

[122] V. Zanchi, V. Papić, and M. Cecić, "Quantitative Human Gait Analysis", *Simulation Practice and Theory*, **8**(1~2), pp.127-139, April 2000.