

AN APPROACH TO ANALYZING AND RECOGNIZING HUMAN GAIT

by

Nalini Vishnoi
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
In Partial fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Computer Science

Committee:

_____	Dr. Zoran Duric, Dissertation Director
_____	Dr. Naomi Lynn Gerber, Dissertation Co-Director
_____	Dr. Jana Kosecka, Committee Member
_____	Dr. Jyh Ming Lien, Committee Member
_____	Dr. Sanjeev Setia, Department Chair
_____	Dr. Kenneth Ball, Dean, Volgenau School of Engineering

Date: _____ Summer Semester 2014
George Mason University
Fairfax, VA

An Approach to Analyzing and Recognizing Human Gait

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy at George Mason University

By

Nalini Vishnoi
Master of Science
University of Allahabad, 2007
Bachelor of Science
University of Allahabad, 2005

Director: Dr. Zoran Duric, Associate Professor
Department of Computer Science

Co-Director: Dr. Naomi Lynn Gerber, Professor
College of Health and Human Services

Summer Semester 2014
George Mason University
Fairfax, VA

Copyright © 2014 by Nalini Vishnoi
All Rights Reserved

Dedication

To my loving parents, Dr. Amarendra Nath Vishnoi and Mrs. Usha Vishnoi.

Acknowledgments

The support and guidance from a number of people have made this dissertation possible. I would first like to express my deepest gratitude to my dissertation directors Prof. Zoran Duric and Prof. Naomi Lynn Gerber for their constant motivation and advice throughout the course of this dissertation. Dr. Duric introduced me to the wonders of computer vision and helped me in understanding and solving the problems whenever I got stuck. He helped me understand what it takes to pursue research goals and always kept my best interests in mind. Dr. Gerber has been a wonderful inspiration throughout who patiently taught me the aspects of biomechanics. Her expert experience in the field helped me in coming up with new ideas and kept me on track. Dr. Gerber's emphasis on the applications of my dissertation work always kept me motivated.

I am also grateful to my other committee members Prof. Jana Kosecka and Prof. Jyh-Ming Lien for their suggestions on my work. Dr. Kosecka helped me with the understanding of intelligent robotics system and computer vision. She helped me explore the implications, applications and the technical depth of my work. Dr. Jyh-Ming taught me computational geometry, which got me interested in the robotic motion planning and helped me in developing critical thinking.

I would also like to thank all the members who have been associated with our Laboratory for the Study and Simulation of Human Movement at different times. Cody Narber, Sam Gelman, Ed Lawson, Dan Fleck provided valuable discussions, suggestions and feedback about my work. Keith Sullivan, Christopher Vo, Kevin Andrea and Katherine Russell in the Robotics lab were always there to discuss the technical matters, share the life experiences and world news. I extend my special thanks to Yanyan Lu, with whom I shared my office space. She has been a very close and dear friend, with whom I could discuss almost everything. We had a lot of fun together inside and outside the laboratory during these years.

I can never thank my husband, Anish, enough who encouraged me with the use of Kalman filters in my work and with whom I shared innumerable discussions and arguments about everything ranging from my dissertation to my daily activities. He provided constructive comments whenever I asked for his feedback on any issue. I would like to thank all my family members and friends for always being a pillar of support, especially my brother, Ambreesh, who took care of me and attentively listened to my long stories. He was always there whenever I needed him.

Lastly but most importantly, I am deeply indebted to my parents who have been the constant source of love, inspiration, strength and support throughout the course of my studies. They always stood beside me in the moments of disappointment and frustration and helped me regain my confidence and the ability to believe in myself.

Table of Contents

	Page
List of Tables	viii
List of Figures	ix
Abstract	xii
1 Introduction	1
1.1 Phases of Gait Cycle	2
1.2 Modeling and Tracking Human Segmental Motion	4
1.3 Gait Recognition	5
1.4 Contributions	7
1.5 Dissertation Outline	8
2 Related Work	9
2.1 Biomechanics and Gait Kinematics	9
2.2 Analyzing Images of Human Motion and Gait	13
2.3 Gait Recognition	15
3 Technical Background	17
3.1 Image Flow	17
3.1.1 The Imaging Models and the Image Motion Field	17
3.1.2 Normal Flow	19
3.1.3 Parametric Flow Estimation	20
3.2 Nonlinear dynamic systems	21
3.2.1 Van der Pol equation	22
3.2.2 Discrete time dynamic system	23
3.3 Kalman filter	25
3.3.1 Unscented Kalman filter	25
3.4 Classification	29
3.4.1 Support Vector Machines (SVMs)	29
3.4.2 Nearest Neighbors	31
4 Gait Databases	33
4.1 George Mason University (GMU) Gait Database	33
4.1.1 Experimental Design	34

4.2	Carnegie Mellon University 3D Motion Capture (CMU MoCap) Database	35
4.2.1	Experimental Design	35
4.3	University of Southampton (USH) Gait Database	36
4.3.1	Experimental Design	37
5	Identifying Phases of Gait Cycle Using Image Motion	38
5.1	Double support and mid swing	39
5.1.1	Motion Models	39
5.1.2	Synchronization of Human Gait	43
5.2	Toe off and heel strike	45
5.2.1	Assumptions	45
5.2.2	Finding the lower leg motion from images	45
5.3	Finding the hand position from images	52
5.4	Experiments	53
5.4.1	Double support and mid swing	54
5.4.2	Toe off and heel strike	56
5.5	Validation	56
5.6	Discussion	59
6	Modeling and Estimating Motion of Unobserved Segments	60
6.1	Transforming 3D data to sagittal plane	60
6.2	Modeling Nonlinear Dynamic Systems	61
6.2.1	Upper leg and lower leg	63
6.2.2	Foot and lower leg	64
6.2.3	Upper arm and forearm	64
6.3	Estimation	65
6.3.1	Upper leg	66
6.3.2	Foot	66
6.3.3	Upper arm and Forearm	68
6.4	Angular position estimates	69
6.5	Initialization of body segments	69
6.6	Experiments	70
6.7	Validation	71
6.8	Discussion	71
7	Gait Recognition	73
7.1	Foreground Extraction	74
7.2	Normalizing motion with respect to torso motion and body height	74

7.3	Histograms of Normal Flow	75
7.4	Comparing two sequences	76
7.5	Gait recognition using similarity matrix	77
7.5.1	Gait recognition using SVM	78
7.5.2	Gait recognition using Dynamic Programming and NN	78
7.6	Experiments	80
7.6.1	Recognition using SVM	81
7.6.2	Recognition using Dynamic Programming and NN	82
7.6.3	Effect and Contribution of Slices on Recognition	82
7.7	Discussion	83
8	Conclusions and Future Directions	84
8.1	Conclusions	84
8.2	Applications	85
8.3	Future Directions	86
	Bibliography	87

List of Tables

Table	Page
5.1 Table comparing the time of double-support and mid-swing occurrences (Optotrak, S_o , vs Video camera, S_c)	58
5.2 Table comparing the time of toe-off and heel-strike occurrences (Optotrak, S_o , vs Video camera, S_c)	59
6.1 Table listing the average error (in degrees) in the computation of angular positions of different lower limbs	71
7.1 Specificity and sensitivity measures for 114 subjects in the USH gait database. . . .	82
7.2 ANOVA: The respective F-measures in the increasing order of importance.	83

List of Figures

Figure	Page
1.1 Planes of Human movement: Transverse, Frontal and Sagittal, from [1].	2
1.2 Subdivision of the gait cycle illustrating the phases of stance and swing and periods of single and double limb support (taken from [2]).	3
1.3 Center of Mass (CoM) displacement during gait. The vertical and the side-to-side displacements of the CoM are illustrated in A and B, respectively. The CoM is at its lowest and most central position, in the side-to-side direction, in the middle of double-limb support (5% and 55% of the gait cycle)- a position of relative stability with both feet on the ground. Conversely, the CoM is at its highest and most lateral position at mid-stance (30% and 80% of the gait cycle- a position of relative instability (taken from [2]).	4
2.1 Anthropomorphic measures of human body expressed as a fraction of body height, H , from [3].	10
2.2 Sagittal plane hip motion (thigh relative to pelvis). Normal range during free walking is 40° [4].	11
2.3 (a) Sagittal plane knee motion. Normal range during a gait cycle for free walking. Black line denotes the mean and dotted lines represent one standard deviation. (b) Ankle motion: Normal range during a gait cycle.	12
2.4 Arcs of elbow and shoulder motion during arm swing while walking [4].	12
3.1 The plane perspective projection image of P is $F = f(X/Z, Y/Z, 1)$; the weak perspective projection image of P is obtained through the plane perspective projection of the intermediate point $P_1 = (X, Y, Z_c)$ and is given by $G = f(X/Z_c, Y/Z_c, 1)$. . .	18
3.2 Multiple patterns of the Van der Pol oscillator based on value of μ	23
3.3 Left: Separation of a set of 2D points belonging to two different classes using lines. Right: Optimal hyperplane separating the points.	29
4.1 Selected frames from the GMU gait database A in the frontal and sagittal planes. .	34
4.2 Selected frames from the GMU gait database B in the sagittal plane.	35
4.3 Pictures showing the accurate placement of different markers in CMU MoCap Database.	36

4.4	Pictures showing the frames from USH gait database.	37
5.1	Foreground detection in the image sequence using background subtraction.	39
5.2	Camera position with respect to the walking person with translational velocity T_t	40
5.3	Example of the image flow from a frontal plane image sequence followed by the figure showing the focus of expansion (FOE) computed using voting.	42
5.4	Example of the image flow from a sagittal plane image sequence followed by the figure showing the results of voting which determines the amount of translation in x and y directions.	43
5.5	The first figure shows the the change in the angle of FOE with respect to the center of subject's body in the frontal plane. The second figure shows the variation of the translational value in the vertical direction of motion in the sagittal plane. The zero crossings indicate the key frames of the image sequences.	44
5.6	Examples of normal flow and their respective histograms on the right. The red flow vectors represent the resultant flow vector in each superpixel computed using the least squares solution given by Eq.5.2	49
5.7	(a) Spatial Separation found for lower leg segments. The dividing point is shown in red(b) The segmented lower legs are shown in red and green based on the dividing point.	50
5.8	Segmentation of lower legs during mid swing using superpixels and histograms of normal flow	51
5.9	Instantaneous rotational velocities computed for the two lower leg segments using our proposed method. The zero crossings have been marked using black and magenta diamonds.	52
5.10	(x, y) positions of the tracked hand in one of the gait sequences.	53
5.11	Figure showing the frames from the original sequence corresponding to the zero crossings in Fig. 5.5. The rows correspond to the alternating maximum and minimum head excursions.	54
5.12	Figure showing the key frames from the original sequence of frontal and sagittal plane sequences. The rows correspond to the alternating maximum and minimum head excursions.	55
5.13	Synchronized frames for the same person in multiple view planes.	55
5.14	Synchronized frames for different individuals in the same view plane.	56

5.15	Figure showing the frames from the original sequences corresponding to the zero crossings in Fig. 5.9. The rows correspond to the alternating maximum and minimum lower leg angles.	57
5.16	Figure showing the settings for the simultaneous recording of video and Optotrak data.	58
6.1	Principal Component Analysis (PCA) gives components 1 and 2 and 3 in the order of decreasing variance. Component 1 and 2 together define the sagittal plane of movement. The red trajectories show the progression of various markers on a person's body while walking. The black axes represent the original coordinate system and blue axes represent the new coordinate system.	61
6.2	Observed (3D data) and computed model for upper leg and lower leg.	63
6.3	Observed (3D data) and computed model for foot and lower leg.	64
6.4	Observed (3D data) and computed mode for upper arm and forearm.	65
6.5	Estimation of left (blue) and right (red) upper leg, lower leg and foot angular positions.	67
6.6	Estimation of upper arm and forearm angular positions.	68
6.7	Frames overlaid with our results of motion estimation of lower limb segments and arms.	70
7.1	Normal Flow.	74
7.2	Numbered regions obtained by taking horizontal slices from the subject's foreground.	76
7.3	Histograms taken from different layers/slices of the body.	77
7.4	Similarity Matrices (SM) in two different cases computed between the frame sequences of the: (a) same person (b) different persons.	78
7.5	Vertical instantaneous velocity of a walking person.	79
7.6	Similarity Matrices (SM) in two different cases: (a) perfectly aligned gait cycles (b) gait cycles shifted by half. The white lines indicate the path taken to obtain the minimum cost.	80
7.7	Examples of frames from USH gait databse.	81

Abstract

AN APPROACH TO ANALYZING AND RECOGNIZING HUMAN GAIT

Nalini Vishnoi, PhD

George Mason University, 2014

Dissertation Co-Directors: Dr. Zoran Duric and Dr. Naomi Lynn Gerber

Gait analysis has been an active area of research in computer vision for a long time. It is also important for rehabilitation science where clinicians explore innovative ways helping to analyze gait of different people. The traditional ways to study gait rely on 3D optical motion capture systems which involve the use of cumbersome active/passive markers to be placed on a subject's body. The attachment of markers to the segments hinder natural patterns of movement and may lead to altered gait information. Automated gait analysis has been proposed as a solution to this problem. The aim of automated gait analysis is to provide information about the gait parameters and gait determinants from video without using markers. Gait is a repetitive, highly constrained and periodic activity. Different gait determinants are active in different phases of the gait cycle to minimize the excursion of the body's center of gravity and help produce forward progression with the least expenditure of energy. The motion of limb segments encode information about different phases of gait cycle. However, estimating the motion of limbs from the videos is challenging since limbs are self occluding and only apparent motion can be observed using the images. To add to the issue, the quality of the recorded video (color contrast, cluttered background) and clothing worn by the subject can play a significant role in the computation of that apparent motion.

In this thesis, we present novel methods using image flow to identify different phases (double support, mid swing, toe off and heel strike) of a gait cycle. We use the torso excursion information

and lower legs rotational velocities to identify these phases. The top 30% of the subject's body is used to estimate the torso instantaneous velocity. The zero values of the vertical component of the velocity identifies the double support and mid swing phases. Utilizing these phases, we present approaches to approximate the lower leg motions using translation and rotational motion models. The zero values of the instantaneous rotational velocity of the lower legs determine the toe off and heel strike events. We also apply a modified version of color tracking algorithm to track the hand position during the gait cycle. Some of the limb segments such as upper legs and upper arms get occluded by other segments during majority of the gait cycle. We have presented a method to model the motion synergies between pairs of segments (upper leg–lower leg, foot–lower leg and upper arm–lower arm) using data from a 3D motion capture database. The estimated parameters of the non-linear dynamic models depend on the phases of the gait cycle being considered. We implement an Unscented Kalman Filter that estimates the angular position of the unobserved limb segments (upper leg, foot, upper arm and forearm) based on these models and the motion data obtained for the observed limb segments (lower leg, hand). We compare our results with those obtained from a 3D motion capture system and by manually labeling the images. The small error in the results demonstrates the sensitivity and specificity of our techniques.

Finally, we use histograms of normal flow to represent the motion patterns of different regions of the body. We measure the motion similarity between two image frames using the cosine similarity measure for comparing two histograms. Computing this measure between all the pairs of image frames in the two gait sequences gives a similarity matrix as a feature. These features are used in Support Vector Machines and Dynamic Programming together with the information about the phases of gait to compare two gait sequence. We demonstrate our approach on a publicly available gait dataset and present the analysis.

In summary, we establish that we can capture segmental data using a markerless gait analysis system. These data are sensitive, reliable and provide recognizable clinically relevant information about motion through all phases of gait.

Chapter 1: Introduction

Humans are capable of movement because of articulations that occur between body segments. Quantifying whole body and segmental movement has required a laboratory based approach. This often requires a labor intensive approach towards data capture and analysis and expensive equipment. Additionally, in order to capture kinematic and kinetic data of limbs markers or other sources of encumbrances (sleeves, body jackets etc) are required. These requirements often restrict the setting in which data can be captured, making it a “study” environment rather than a more natural setting, and limiting the nature of the kinds of activities that can be studied. There is a need for the application of existing technology and the development of methods for capturing human movement that is marker-less, inexpensive, quantitative, sensitive and specific to the problems of human motion. We propose novel computer vision techniques to analyze human gait (lower extremity movement). Specifically, we explore whether human gait can be analyzed with sufficient ease, reliability, sensitivity and specificity using computer vision, in order to obtain meaningful quantitative and descriptive data for clinical applications and for soft biometrics.

Lower extremity movements are: highly constrained by human anatomy, highly characteristic of an individual and reproducible. Nonetheless movement occurs in three different planes of movement, presenting the need for analyses that accurately describe the kinematics. The three planes of movement in which the motion can be described are (see Fig. 1.1): the sagittal plane which passes from front to rear dividing the body into two symmetrical halves, the coronal plane (also referred to as the frontal plane) that divides the body into the front and back parts, and finally, the transverse plane which divides the body into the top and bottom parts [1].

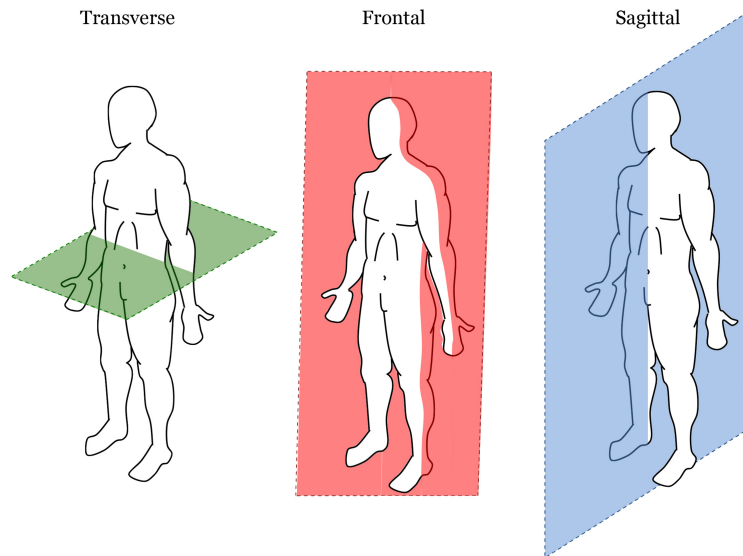


Figure 1.1: Planes of Human movement: Transverse, Frontal and Sagittal, from [1].

1.1 Phases of Gait Cycle

The entire gait cycle is divided into well understood phases of gait which are standard events whose sequence describes the temporal events contributing to locomotion [5]. The complete gait cycle is measured from the heel strike of one foot to the next heel strike of the same foot [6]. It is composed of approximately 40% swing and 60% stance (see Fig. 2.1) [7].

The proportion of time a limb is on the ground (stance phase) compared to it being in the air (swing phase) determines speed of walking. The pattern with which the foot hits the ground and the amount of heel rise, influences the smoothness of gait and weight transfer. The periodicity of the sequences determine whether the movement is smooth, and appears “normal”. “Bilateral Symmetry” occurs when arms are swinging freely [8]. The left arm moves with the right leg and vice-versa. This is most visible in the sagittal plane. An abnormality in phases of gait or stride parameters is defined by a variance in the spatial or temporal sequences. An individual’s gait and its specific phases, can be compared to normal data or can be measured using a variety of walking conditions and speeds. Limb positions in each phase of gait have been carefully analyzed, and

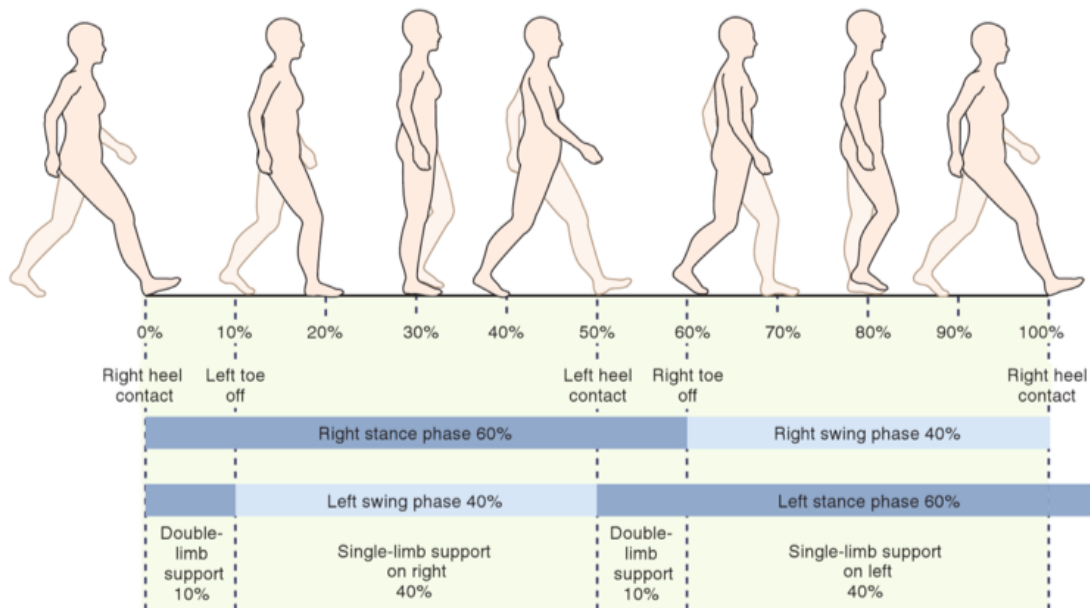


Figure 1.2: Subdivision of the gait cycle illustrating the phases of stance and swing and periods of single and double limb support (taken from [2]).

normal ranges have been established. In addition, one can identify the differences in limb segment kinematics, such as pelvis, hip, knee, and ankle motion.

We have used computer vision techniques to identify the phases of gait from different camera viewpoints, its spatial and temporal parameters in the sagittal plane, much as one would using a marker based video capture system. During the gait cycle, the head moves up and down along a shallow sinusoid (see Fig. 1.3). This may be altered by stride length, degree of joint range of motion and other limb segment characteristics. We have used image flow to detect up/down movement of the head and correlate these movements to specific phases of gait. We have used the instantaneous head velocities to detect the transition points. These transition points mark the key frames for the following phases of gait: first double support, mid stance, second double support and mid swing. We compute the lower leg motion using the information of double support and mid swing phases and use the instantaneous lower leg velocity to further detect the heel strike and toe off events in the gait cycle.

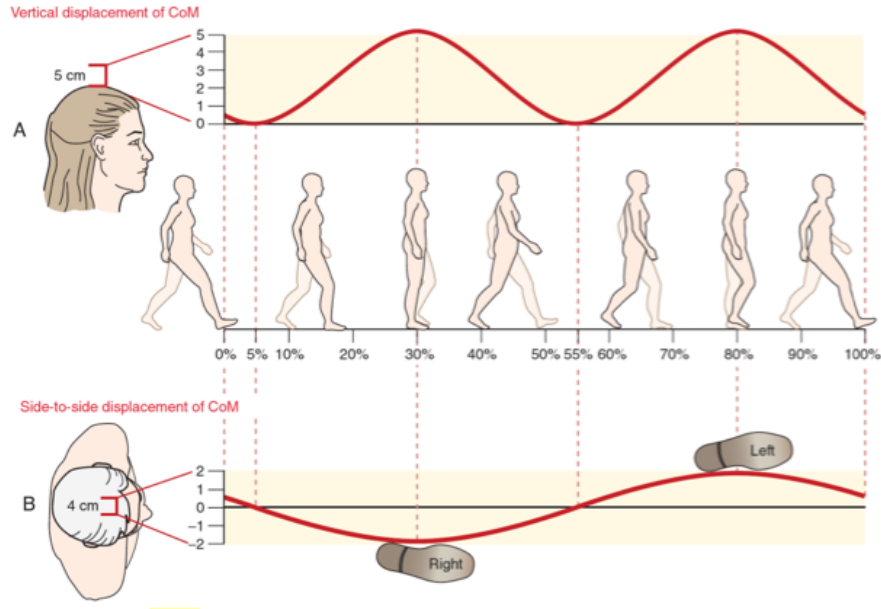


Figure 1.3: Center of Mass (CoM) displacement during gait. The vertical and the side-to-side displacements of the CoM are illustrated in A and B, respectively. The CoM is at its lowest and most central position, in the side-to-side direction, in the middle of double-limb support (5% and 55% of the gait cycle)- a position of relative stability with both feet on the ground. Conversely, the CoM is at its highest and most lateral position at mid-stance (30% and 80% of the gait cycle- a position of relative instability (taken from [2]).

1.2 Modeling and Tracking Human Segmental Motion

Full body pose estimation and tracking has generated a lot of interest over the course of several years because of its varied applications. Tracking the body motion includes the segmentation of background and foreground in the first step. Once the foreground is extracted, a search for correspondences is initiated. These correspondences can be computed using shape, appearance and motion of the body. Few studies have focussed on methods based on the estimation of Center of Mass (CoM) from the silhouettes of the body [9, 10]. The computation of CoM is, therefore, very sensitive to the accuracy of the silhouettes of the moving body obtained from background subtraction. However, there are several problems associated with their use. First, accurate silhouette extraction is a challenging problem. Poor lighting conditions and/or low contrast lead to inaccurate silhouettes, and silhouette filtering operations affect the shape. Second, changes in silhouette are only loosely connected to gait dynamics. Silhouettes can change between trials for reasons that

have nothing to do with gait, such as when a subject changes clothing. Recent work [11] depends on the accurate segmentation of moving body parts to be able to track the outline of the body. To alleviate this problem, people have also proposed model based tracking of human body [12–14]. The human body is assumed to have a kinematic structure comprising fixed number of joints with specified degrees of freedom. While walking, since a number of body segments get self-occluded and are difficult to segment from one another, some a priori knowledge is needed about the way the segments move or about their spatial and motion relationships. [15, 16] have used constant velocity assumptions and enforced it using Kalman filters. Gaussian Process Dynamical Models [17] and Bayesian inference/networks [18, 19] have been used to learn the motion from already gathered data. These learned behaviors are used to predict the human movement in new videos.

As discussed in [19], many of the human activities, for example gait, are highly constrained and we can build strong models for those specific activities. We take advantage of highly repetitive and periodic normal gait and model the strong relationship between upper leg, lower leg and foot, and upper arm and forearm motions using non linear models and use Kalman filters to estimate the upper leg, foot, upper arm and forearm velocities given the lower leg and hand motion data. In a number of articles [15, 16] Kalman filter has been used for constant velocity constraint. We model the synergies between the motion of different body segments using non linear dynamic system and then predict the unobserved motion using the observed motion and Unscented Kalman filter.

1.3 Gait Recognition

Vision based human gait recognition has been of great interest to researchers. Its applications include visual surveillance and biometric identification in forensics; it can be deployed at the airports, shopping malls, and other structures in order to improve security. There is considerable evidence in biomechanics that individuals can be recognized by the way they walk reflecting that gait can be used as a soft biometrics [3, 20]. A unique advantage of gait as a biometric is that it can be easily recognized from a distance. The collection of data is non-intrusive and does not require high resolution images as is required with iris and face recognition. In [21], point light displays were

used to establish that humans can distinguish locomotion from other actions. The studies in [6, 22] have established the ranges for the kinematics of normal gait for the men, women and elderly which permits analysis of a subject's gait in order to identify deviations from that range. During walking, different body parts move in synergy in order to maintain balance and minimize energy consumption. The body divides itself into two parts: passenger and locomotor units. The passenger unit comprises head, arms and trunk (HAT) and two lower limbs and pelvis form the locomotive unit [23]. The motion of different body parts in the passenger and locomotive units can vary greatly from person to person in a gait cycle. This suggests that it may be useful to use different limb motions to improve the sensitivity and specificity of using gait for individual recognition. Different limb motions can form a set of measurements for a subject for which the variation for that subject (intra-subject variation) is less than the variation between subjects (inter-subject variation). An example of this can be seen in the distance traveled by the arms during arm swings among different subjects. The recognition power is not entirely derived from upper body shape, rather the dynamics of the legs also contribute equally to recognition. In [24–26], the authors have discussed the importance of different body parts/components in gender classification. We aim to exploit the motion information contained in different regions of the subject's body for gait recognition.

We propose a novel technique to perform gait recognition utilizing histograms of normal flow. We extract the moving person using background subtraction and thresholding. The normal flow is computed using the method described in [27]. We identify the subject's gait cycle using the vertical component of torso velocity and use this information to automatically align gait sequences. After normalizing the subject's height and computing the relative motion with respect to the torso, the entire body is divided into equally sized overlapping horizontal slices. For each slice, we compute a 16-bin normalized histogram of normal flow. To compare two gait sequences, we determine the similarity matrix by computing the cosine distance between the respective histograms in all frames. The similarity measure is used as a feature to classify the person in the database. We use two different methods to perform gait recognition. In the first one, we have used linear Support Vector Machine (SVM) without incorporating the gait cycle information. In the second one, the information about the gait cycles together with dynamic programming and nearest neighbors (NN)

is used to find the minimum cost path in the similarity matrix. The similarity matrix contains the information about the distance between the motion patterns of two gait sequences. We demonstrate our method using the publicly available University of Southampton (USH) [28] gait database and present the results.

1.4 Contributions

The main contributions of the dissertation are listed below:

- We introduce the concept of histograms of normal flow for superpixels comprising normal flow vectors in different image regions. These histograms are used to compute the image motion using coarse to fine model estimation and as a representation to encode motion of different body segments of a subject's body.
- We estimate the change in the vertical displacement of a walking subject using normal flow in the frontal and sagittal planes. We use the zero crossings of this vertical instantaneous velocity of the torso to identify the double support and mid swing phases in the gait cycle. These phases are used further to align the sequences of different subjects in multiple viewpoints. We estimate the instantaneous lower leg rotational velocities in the sagittal plane videos. The zero crossings of this rotational velocity determines the toe off and heel strike events in the gait cycle.
- We model the synergies between different body segments (upper leg-lower leg, lower leg-foot and upper arm-forearm) using non-linear dynamic models. We predict the unobserved motion from the observed image motion and models derived from the 3D gait data using Unscented Kalman filter.
- Histogram of normal flow are used to compare the motion patterns of two gait sequences and recognize subjects from the database using Support Vector Machines (SVM) and dynamic programming on USH public gait dataset [28].

1.5 Dissertation Outline

In Chapter 2, the related work on biomechanics, human motion tracking and gait recognition is discussed. In Chapter 3, technical background on image flow, nonlinear dynamic systems is presented. It also discusses the Unscented Kalman filter and its advantages when applied to non linear dynamic systems. In Chapter 4, we introduce the GMU gait databses. These include high speed videos of walking subjects captured in the frontal and sagittal planes. We also describe the CMU 3D motion capture database which was used to create the non linear dynamic motion models, and University of Southampton gait database on which gait recognition experiments were conducted.

In Chapter 5, we explain the motion models in the frontal and sagittal planes with respect to the forward translation, up and down movement and focus of expansion of the moving person's image flow. We discuss the use of vertical instantaneous verlocity of the head to identify double support and mid swing events in the gait cycle for both view planes. It also discusses the computation of motion of distal body segments using image normal flow and color cues. In Chapter 6, we explain the formulation of non linear dynamic models between two body segments and implementation of the Kalman filter. The results are presented on GMU gait database. In Chapter 7, we explain the motion representation using histograms of normal flow to compare two gait sequences and present our results of gait recognition in the USH gait database using SVM and dynamic programming.

Chapter 8 presents the conclusions, applications and future work.

Chapter 2: Related Work

This chapter describes previous work on gait analysis, understanding of human movement using markerless motion analysis and people identification using gait as a soft biometric. An overview of the biomechanics of gait and its kinematics is presented in Sec. 2.1. In Sec. 2.2, we provide a review on human motion understanding including human body tracking. Finally, we present the approaches to gait recognition and discuss the features used in the literature.

2.1 Biomechanics and Gait Kinematics

The human body is a very complex system and consists of several segments and joints (206 bones). The length of segments vary with body build, sex and racial origin. An average set of segment lengths expressed as a percentage of body height was prepared by [29] and is shown in Fig. 2.1. These segment proportions serve as a good approximation in the absence of better data or when it is not possible to get the measures directly from the individual. We have used these approximations in our experiments.

Studies on human locomotion are primarily concerned with the interaction among major body segments such as hip, thigh, shin, foot, trunk, shoulder, upper arm, forearm and hand. Marker-based approaches were used to measure the ranges of motion of different joints. The hip joint is located between the pelvis and the upper end of the femur (thighbone). It is an extremely stable ball and socket joint. It also has a great deal of mobility which allows normal locomotion. The motion at the hip joint occurs in all three planes [30]: sagittal (flexion-extension), frontal (abduction-adduction) and transverse (internal and external rotation). The greatest amount of motion is in the sagittal plane, where the range of flexion is from zero to approximately 140 degrees and the range of extension is from zero to 15 degrees. While walking, the hip joint angle is maximally flexed during the late swing phase of gait, as the limb moves forward for heel strike [4]. Maximum extension is reached

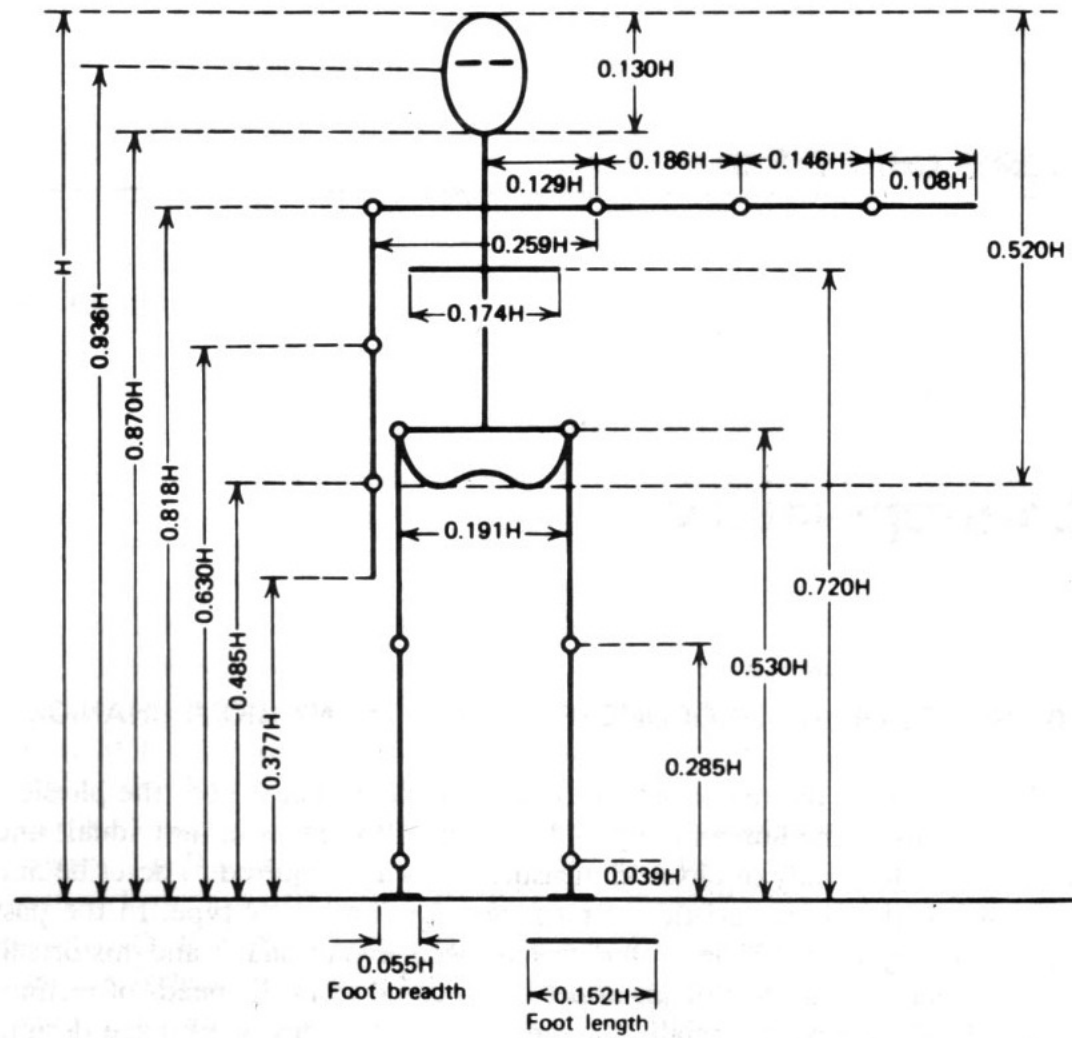


Figure 2.1: Anthropomorphic measures of human body expressed as a fraction of body height, H , from [3].

at heel-off. Fig. 2.2 shows the pattern of hip joint motion in the sagittal plane during a gait cycle. The normal range of hip joint during walking is 40° . Hip flexion of at least 120 degrees, abduction and external rotation of at least 20 degrees are necessary for daily activities.

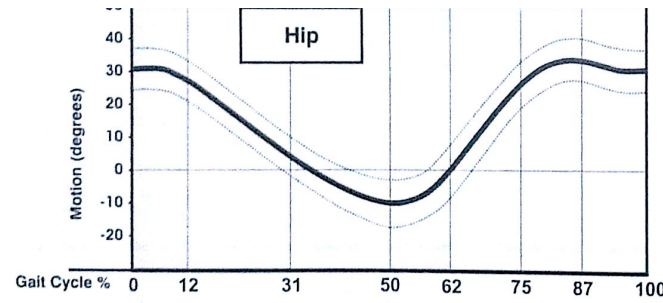


Figure 2.2: Sagittal plane hip motion (thigh relative to pelvis). Normal range during free walking is 40° [4].

The knee is a complex joint, which is made up of the distal end of the femur (the femoral condyles), and the proximal end of the tibia (the tibial plateau). The motion at the knee joint is largely limited to the sagittal plane [1]. The range of motion in this plane for carrying out the activities of daily living is approximately 117 degrees, however normal gait only requires 67 degrees change [1]. During the entire gait cycle the knee never fully extends, nearly full extension (5 degrees of flexion) occurs at the beginning and end of the stance phase. Maximum flexion (75°) was observed during the middle of the swing phase. Kettelcamp noted the longer the leg is, the greater the range of motion. As the pace accelerates from walking slowly to running, progressively more knee flexion is needed during the stance phase [31]. Fig. 2.3a shows 2 waves of flexion experienced by the knee during a gait cycle. The motion at the ankle joint is uniplanar, approximately 45 degrees in the sagittal plane [32]. Ten to 20 degrees of this motion is defined as dorsiflexion and the remaining 25-35 degrees as plantar flexion. The amount of plantar flexion from heel strike to foot flat depends on the height of the shoe heels. The higher the heel is, the more plantar flexion [33]. However, the total amount of ankle joint motion decreases as heel height increases [22]. Unlike the hip and knee, the ankle joint motion tends to decrease as the pace accelerates [33]. Fig. 2.3b shows the normal range of motion of the ankle joint.

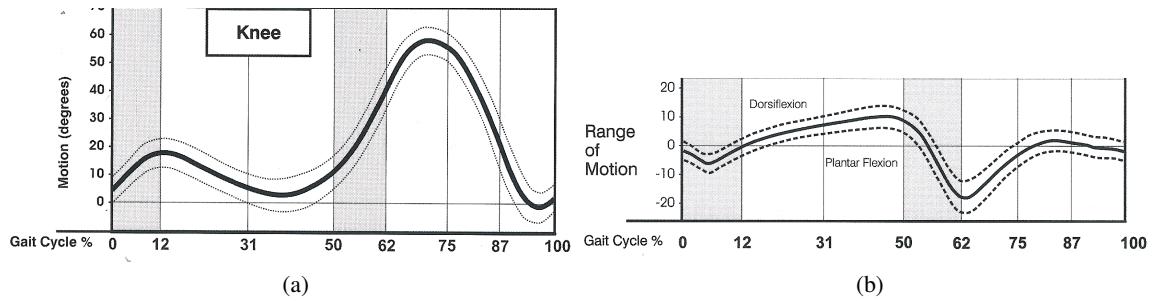


Figure 2.3: (a) Sagittal plane knee motion. Normal range during a gait cycle for free walking. Black line denotes the mean and dotted lines represent one standard deviation. (b) Ankle motion: Normal range during a gait cycle.

During each gait cycle, the arms reciprocally flex and extend. Timing between the 2 arms is offset by 50% in the cycle. The peak extension of each arm occurs during the ipsilateral heel contact whereas the peak flexion happens with contralateral initial contact. The average change in the shoulder joint is 32° during moderate speed walking (see Fig. 2.4). This range increases with faster speed of walking. The pattern of elbow motion during walking is similar to that for the shoulder. Only a small range (18-42 degrees flexion) of the total elbow motion (0 to 147 degrees flexion) is used during walking. During moderate and fast walking, the magnitudes of the arcs are comparable to those occurring in the shoulder, however the elbow remains in flexion throughout the gait cycle.

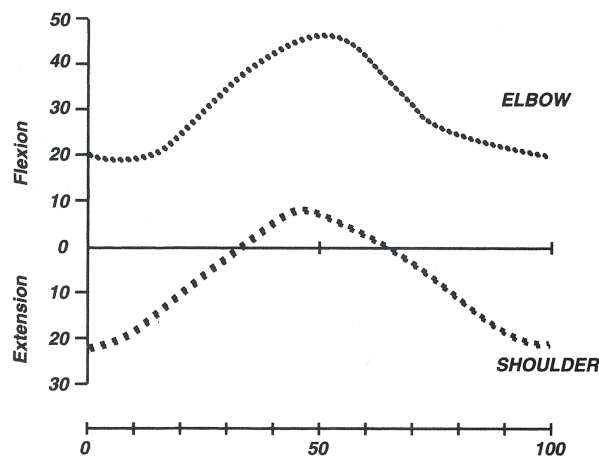


Figure 2.4: Arcs of elbow and shoulder motion during arm swing while walking [4].

2.2 Analyzing Images of Human Motion and Gait

Although marker based analysis has been recognized as clinically useful, the routine clinical use of gait analysis has seen very limited growth. The issue of its clinical value is related to many factors, including the applicability of existing technology to addressing clinical problems and the length of time and costs required for data collection, processing and interpretation. A next critical advancement in human motion capture is the development of a noninvasive and markerless system. Eliminating the need for markers would considerably reduce patient preparatory time and enable simple, time-efficient, and potentially more meaningful assessments of human movement in research and clinical practice. The development of markerless motion capture systems originated from the fields of computer vision and machine learning, where the analysis of human actions by a computer is gaining increasing interest. Extensive work on analyzing images of humans began in the 1990s. A very good review of early work on motion understanding approaches and applications was prepared by Cedras and Shah [34]. Recent reviews on machine analysis of human motions by Gavrilu [35], Aggarwal and Cai [36], Moeslund and Granum [37] and Forsyth et. al. [38] provide excellent coverage of research on detection, tracking, and representation of human motion.

In the remainder of this section we review recent work on motion understanding. Three main criteria can be used to classify research on human motion analysis. First, the research can be classified in terms of the tasks that it focuses on: detection, tracking, or recognition. Second, it can be classified in terms of the models used to represent objects and humans. Third, it can be classified in terms of the control mechanisms used.

Detection of humans in static or video images has usually been addressed through background subtraction and matching. A background subtraction method that uses colors and edges was described in Jabri et. al. [39]. Some authors have used background subtraction as a part of a system combining detection, body labeling, and tracking of humans [40–43]. In some cases cues such as skin color have been used to detect humans in images [44]. Other authors have used motion from single or multiple cameras to detect, label, and track humans or their body parts in video images [12, 16, 43, 45–48]. Other authors have approached this problem as one of matching. Humans

or their parts have been detected and tracked as configurations of points such as light displays, markers, and image features [49]; as configurations of edges [50–56]; and as collections of particularly shaped strips [57], cylinders or superquadrics [19,41,47,48,58]. Recent work [59,60] extend the part-detection paradigm by using kinematic information. Ramanan et.al [59] represents the parts as a Bayesian network and uses kinematic information to search for correct answers. Mikolajczyk et.al [60] uses part assembly information to reject improbable matches.

For tracking some authors have focused on using motions of image points and edges. Human models have been initialized by hand in the first frame of each sequence [61,62]. Some authors have considered the problem of action/activity/gesture recognition for humans using shape and/or motion information [21,63–75]. Recognizing tool actions has been studied by [76–80]. Finally, some authors have implemented systems that combine detection, tracking, and recognition [42,46,65,81–86].

A second set of criteria that can be used for classifying research on human and tool motions is based on the types of models used to model tools and humans. Tools are usually modeled as rigid [76–79] or articulated objects [80]. Humans have been modeled as elongated, blob-like shapes either implicitly [42,46,65,83–85] or explicitly [50,54,55]. Deformable models have been utilized for body part (hands) and facial feature tracking [56,64,74] Some authors have modeled humans as articulated stick figures [49,51–53,68,72,81]; this approach has been particularly effective for moving light display analysis. Finally, humans have been modeled as articulated objects, where parts correspond to blobs [43], strips [12,57,61], tapered superquadrics [41,45], or cylinders [16,19,47,48,58].

A third set of criteria that can be used for classifying research on human motions is based on the mechanisms used to control search in detection and tracking. Kalman filtering has been used frequently; examples include [16,45,47,48,62]. More recently, Bayesian inference has been used [19,55,56,58]; these methods are also known as *Condensation*. More recently, Gaussian Process Dynamic models have been used for human pose estimation and tracking [17,87,88]. These approaches assume an a priori model with kinematic information. Other strategies that have been used include search algorithms such as best-first [41] and/or “winner take all” [62,75,85]. Recent

work by Kolsch et.al [89] builds on search based strategies by dynamically updating the features to be tracked. The approach is to use only those features that are not too close to each other, yet not too far away from the object median.

2.3 Gait Recognition

A unique advantage of gait as a biometric is that it offers potential for recognition at a distance or at low resolution or when other biometrics might not be perceivable. Recognition by gait can be based on the (static) human shape as well as on movement, suggesting a richer recognition cue.

The automated gait recognition system generally consists of the following steps: subject detection, silhouette extraction, feature extraction, feature selection and classification. The feature extraction step can be divided into two broad categories: (1) model free approaches, (2) model based approaches. The model free approaches treat the gait in a holistic manner, the gait is represented as a sequence of binary silhouettes or sequence of the whole body motion. In [90], gait recognition is performed by the temporal correlation of silhouettes. The idea of Motion Energy Image (MEI) and Motion History Image (MHI) for human motion recognition in [65], has been extended to Gait Energy Image (GEI) by [91] and Gait History Image (GHI) by [92]. Procrustes shape analysis has been employed in [93] to extract a mean shape as a gait signature. In [94, 95], features are extracted from the image self-similarity plots for gait and action recognition. These methods rely on the accurate segmentation of the foreground from the background. Optical flow has also been used to derive a gait signature. In [96], periodic structure of the optical flow distribution of centroids of moving points and their moments were utilized for gait representation. [97] builds on the work of [98] and uses intensity and directional optical flow information. The authors compute 5 normalized histograms covering the entire gait cycle which are then used as motion descriptors for recognition. More recently, Gait Flow Image (GFI) has been introduced in [99] in which a binary image is computed where all the pixels with flow magnitude greater than a threshold are represented by 1. The GFI is then directly or indirectly (using dimensionality reduction) matched to the test sequences for gait recognition.

The model based approaches model the human body configuration using structural and motion information to obtain different parameters of body components such as limbs and arms. These parameters are used to derive gait signatures. A few examples of such signatures are stride and cadence [100], body segment widths and gait frequency [101] (structure based), body part shapes and dynamics including positions and orientations [102, 103] (kinematics and dynamical motion-based). For detailed description on the techniques, we refer the readers to surveys on gait recognition [104, 105]. The feature extraction step generates large number of features to which dimensionality reduction algorithms are applied. These include Principal Component Analysis (PCA) [94], Linear Discriminant Analysis (LDA) and Multiple Discriminant Analysis (MDA) [91]. Support vector machines (SVM), Dynamic time warping (DTW), Fourier descriptor analysis, neural networks and Hidden markov models (HMMs) have been used as classifiers for gait recognition [104].

Chapter 3: Technical Background

In this section, we will describe the technical background on image flow estimation, non-linear dynamic systems, Kalman filters and a brief description of the classification methods used in this dissertation (Support Vector Machines and Nearest Neighbors). Sec. 3.1 provides background material on the computation of image motion, i.e. the normal flow. We use normal flow to estimate the apparent image motion of different body segments for gait analysis. Sec. 3.2 and Sec. 3.3 discuss the formulation of a non linear dynamic model and Unscented Kalman filter to solve it. Non linear dynamic models are used to capture the relationship of two body segments while walking. Unscented Kalman filter is used to predict the unobserved motion of a segment using the observed motion of another segment and the model explaining the relationship between the two segments. Sec. 3.4 provides a brief introduction of the classification methods. These methods are used for gait recognition in the University of Southampton gait dataset [28].

3.1 Image Flow

Image flow is the instantaneous velocity vector field for an image of a moving environment. It corresponds to apparent image motion . In Sec. 3.1.1 we describe how a point in the real world projects to a pixel in the image and how we can compute its motion using projection equations. Sec. 3.1.2 describes our method for finding normal flow. In the last Sec 3.1.3 we explain how to fit parametric models to flow values.

3.1.1 The Imaging Models and the Image Motion Field

Let (X, Y, Z) denote the Cartesian coordinates of a scene point with respect to the fixed camera frame (see Fig. 3.1), and let (x, y) denote the corresponding coordinates in the image plane. The

equation of the image plane is $Z = f$, where f is the focal length of the camera. The perspective projection onto this plane is given by

$$x = \frac{fX}{Z}, \quad y = \frac{fY}{Z}. \quad (3.1)$$

For weak perspective projection we need a reference point (X_c, Y_c, Z_c) . A scene point (X, Y, Z) is first projected onto the point (X, Y, Z_c) ; then, through plane perspective projection, the point (X, Y, Z_c) is projected onto the image point (x, y) . The projection equations are then given by

$$x = \frac{X}{Z_c} f, \quad y = \frac{Y}{Z_c} f. \quad (3.2)$$

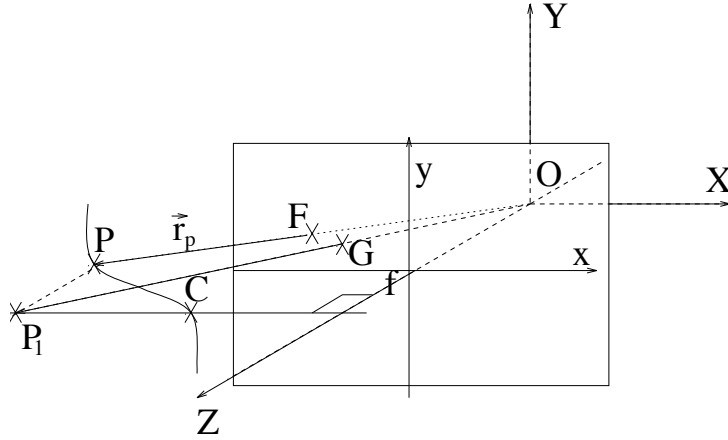


Figure 3.1: The plane perspective projection image of P is $F = f(X/Z, Y/Z, 1)$; the weak perspective projection image of P is obtained through the plane perspective projection of the intermediate point $P_1 = (X, Y, Z_c)$ and is given by $G = f(X/Z_c, Y/Z_c, 1)$.

The instantaneous velocity of the image point (x, y) under perspective projection is given by [106]:

$$\dot{x} = \frac{Uf - xW}{Z} - \omega_x \frac{xy}{f} + \omega_y \left(\frac{x^2}{f} + f \right) - \omega_z y, \quad (3.3)$$

$$\dot{y} = \frac{Vf - yW}{Z} - \omega_x \left(\frac{y^2}{f} + f \right) + \omega_y \frac{xy}{f} + \omega_z x. \quad (3.4)$$

3.1.2 Normal Flow

In this section we briefly describe our method of computing normal flow in color images. Let $\mathbf{i}$ and $\mathbf{j}$ be the unit vectors in the x and y directions, respectively; $\delta\vec{r} = \mathbf{i}\delta x + \mathbf{j}\delta y$ is the projected displacement field at the point $\vec{r} = x\mathbf{i} + y\mathbf{j}$. If we choose a unit direction vector $\mathbf{n}_r = n_x\mathbf{i} + n_y\mathbf{j}$ at the image point $\mathbf{r}$ and call it the normal direction, then the normal displacement field at $\mathbf{r}$ is $\delta\mathbf{r}_n = (\delta\mathbf{r} \cdot \mathbf{n}_r)\mathbf{n}_r = (n_x\delta x + n_y\delta y)\mathbf{n}_r$. The normal direction $\mathbf{n}_r$ can be chosen in various ways; the usual choice is the direction of the image intensity gradient $\mathbf{n}_r = \nabla I / \|\nabla I\|$. Note that the normal displacement field along an edge is orthogonal to the edge direction. Thus, if at time t we observe an edge element at position $\mathbf{r}$, the apparent position of that edge element at time $t + \Delta t$ will be $\mathbf{r} + \Delta t\delta\mathbf{r}_n$. This is a consequence of the well-known *aperture problem*. We base our method of estimating the normal displacement field on this observation.

In color images (RGB) we apply an edge detector to each color band to obtain partial derivatives $r_x, r_y, g_x, g_y, b_x, b_y$ for the (r)ed, (g)reen, and (b)lue bands. Edges in color images can be computed using a standard technique used for processing multi-channel imagery [107]. We first form a matrix S ,

$$S = \begin{pmatrix} r_x^2 + g_x^2 + b_x^2 & r_x r_y + g_x g_y + b_x b_y \\ r_x r_y + g_x g_y + b_x b_y & r_y^2 + g_y^2 + b_y^2 \end{pmatrix}.$$

The trace of S corresponds to the edge strength. If there is an edge at point (x, y) , the larger eigenvalue of S , λ_1 , corresponds to the edge strength. The corresponding eigenvector (n_x, n_y) represents the edge direction. Therefore we can treat color edges in the same manner as we have treated gray level edges. The only difference is that the edge strength and the edge direction correspond to the larger eigenvalue of S and its corresponding eigenvector. For each edge element, say at $\mathbf{r}$, we re-sample the three image color bands locally to obtain three small windows with their rows parallel to the image gradient direction $\mathbf{n}_r = (n_x, n_y)$. For the next image frame (collected at time $t_0 + \Delta t$) we create a larger window, typically twice as large as the maximum expected value of the magnitude of the normal displacement field. We then slide the first (smaller) window along the second (larger)

window and compute the difference between the image intensities in all three color bands. The result is a vector function $(\delta_r, \delta_g, \delta_b)$ of the color differences. The magnitude of this vector has a zero crossing at distance u_n from the origin of the second window; the difference vector changes sign around the zero crossing. We estimate the zero crossing by comparing the magnitudes of the two difference vectors pointing in opposite directions. Our estimate of the normal displacement field is then $-u_n$, and we call it the *normal flow*.

3.1.3 Parametric Flow Estimation

Once the normal flow is computed, several parametric models can be fitted to the flow values as explained below:

Let (x, y) be the image coordinates of a pixel in an image $I(x, y)$ and let the image be centered at $(0, 0)$. We have the following expression for the affine displacement $(u(x, y), v(x, y))$ of the point (x, y) ,

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} w_1 & w_2 \\ w_3 & w_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} w_5 \\ w_6 \end{pmatrix} \quad (3.5)$$

The normal displacement field at (x, y) is given by

$$u_n(x, y) = \delta \vec{r}_n \cdot \mathbf{n}_r = n_x u(x, y) + n_y v(x, y) = \mathbf{w} \cdot \mathbf{p},$$

where $\mathbf{n}_r = n_x \vec{i} + n_y \vec{j}$ is the gradient direction, $\mathbf{p} = (x n_x \ y n_x \ n_x \ x n_y \ y n_y \ n_y)^T$, and $\mathbf{w} = (w_1 \ w_2 \ w_5 \ w_3 \ w_4 \ w_6)^T$ is the vector of parameters. For each edge point $\vec{r}_i$ we have one normal flow value $u_{n,i}$, that we use as an estimate of the normal displacement at the point, a vector $\mathbf{p}_i$ computed from (x_i, y_i) and $\mathbf{n}_{r,i} = n_{x,i} \vec{i} + n_{y,i} \vec{j}$, and an approximate equation $\mathbf{w} \cdot \mathbf{p}_i \approx u_{n,i}$. Let the number of edge points be $N \geq 6$. We need to find a solution of $P\mathbf{w} - \mathbf{b} = \mathbf{e}$, where $\mathbf{b}$

is an N -element vector with elements $u_{n,i}$, P is an $N \times 6$ parameter matrix with rows $\mathbf{p}_i$, and $\mathbf{e}$ is an N -element error vector. We seek the model $\mathbf{w}$ that minimizes $\|\mathbf{e}\| = \|\mathbf{b} - P\mathbf{w}\|$; the solution satisfies the system $P^T P\mathbf{w} = P^T \mathbf{b}$ and corresponds to the linear least squares (LS) solution.

We choose from the following four models that we obtain by setting various elements of $\mathbf{w}$ to zero:

M_1 : *pure translation*, $w_1 = w_2 = w_3 = w_4 = 0$, 2 DoF, $h = 3$;

M_2 : *translation, shear, and rotation*, $w_1 = w_4 = 0$, 4 DoF, $h = 5$;

M_3 : *translation and scaling*, $w_2 = w_3 = 0$, 4 DoF, $h = 5$;

M_4 : *6-parameter affine*, 6 DoF, $h = 7$;

3.2 Nonlinear dynamic systems

The use of mathematical modeling to describe, analyze and predict physical phenomenon is well documented. The advancement of technology has ensured that we can observe newer and more relevant events and use the observations to generate models. A mathematical model is a simplified version of the real world that employs the tools of mathematics - algebraic equations, probability, statistics, graph theory, etc. One of the different classes of mathematical models available is Dynamical Systems. Developed in the first half of the twentieth century it is used to describe the time dependence of a point in multi-dimensional space. Applications of such models include fluid dynamics, chemical reactions, analysis of cardiac rhythms and neuronal spiking as well as description of joint angles in human motion.

In [108] differential equations are used to formally describe dynamic system theory. Typically the independent variable of the equations is time. The basic first order dynamic system contains a linear equation that is used to define the rate of change of the state of the point. Mathematically this can be written as

$$\frac{d\mathbf{x}(t)}{dt} = A\mathbf{x}(t) \tag{3.6}$$

The variable $\mathbf{x}(t)$ represents a vector in n -dimensional space at the time instance t . It is also referred to as the state of the system, or system state. The $n \times n$ matrix A defines the temporal ‘dynamics’ of the vector $\mathbf{x}$. The linear equation is unable to reproduce complex phenomenon such as autonomous oscillations. To achieve this, the rate of change is defined using a nonlinear equation.

$$\frac{d\mathbf{x}(t)}{dt} = f(\mathbf{x}(t)) \quad (3.7)$$

The function $f(\cdot)$ contains nonlinear terms of $\mathbf{x}(t)$. The above equation forms the basis of first order nonlinear dynamic models (in continuous time). The exact form of the equation will govern the dynamics of the point. Often the equation is parameterized and different patterns of the same phenomenon (e.g. oscillatory patterns) can be reproduced by changing the parameters of the equation. We illustrate this using a well known example of an oscillatory system.

3.2.1 Van der Pol equation

The Van der Pol Oscillator [109] is a non-conservative oscillator with nonlinear damping. It’s dynamics are described by a second order nonlinear differential equation.

$$\frac{d^2 y(t)}{dt^2} - \mu[1 - y(t)^2] \frac{dy(t)}{dt} + y(t) = 0 \quad (3.8)$$

This can be redefined as a two dimensional first order nonlinear dynamic system by using the substitutions $x_1(t) = y(t)$ and $x_2(t) = \frac{dy(t)}{dt}$.

$$\begin{aligned} \frac{dx_1(t)}{dt} &= x_2(t) \\ \frac{dx_2(t)}{dt} &= \mu[1 - x_1(t)^2]x_2(t) - x_1(t) \end{aligned} \quad (3.9)$$

As the name suggests, the Van der Pol Oscillator reproduces an oscillating response whose

dynamics and frequency depends on the initial conditions and the parameter μ . Fig. 3.2 shows the responses for several parameter values. This particular aspect of nonlinear dynamic systems makes it applicable to phenomenon where the dynamics are non-deterministic but limited to a pre-defined range. The differential equations used in most practical applications do not have a closed-form solution and need to be solved numerically.

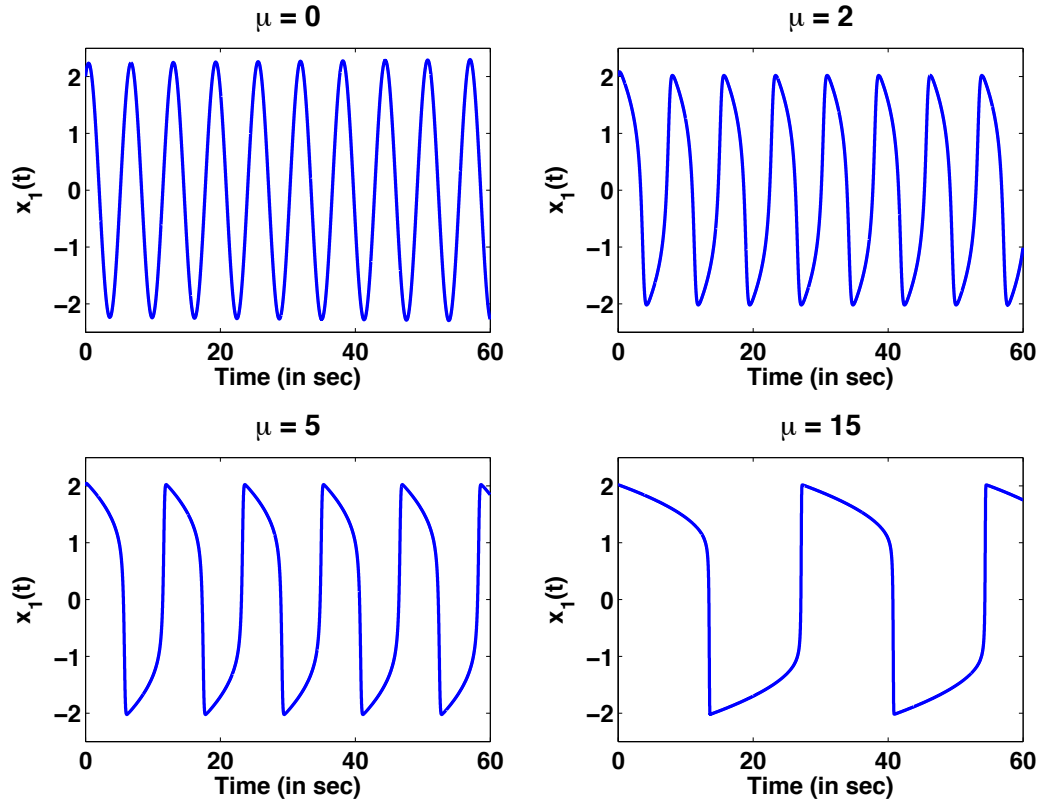


Figure 3.2: Multiple patterns of the Van der Pol oscillator based on value of μ

3.2.2 Discrete time dynamic system

While the dynamical system theory was developed for continuous time mathematical models, they are implemented in discrete time using a defined time-step. Differential equations are replaced by difference equations to describe these models mathematically. For a first order system the value of

the system state at instance $k + 1$ is calculated based on the value at instance k .

$$\mathbf{x}(k + 1) = g(\mathbf{x}(k)) \quad (3.10)$$

One of the techniques used to convert continuous time equations to discrete difference equations is the explicit forward Euler method [110]. The derivative of the state with respect to time is approximated as

$$x'(t_{k+1}) = \frac{x(t_{k+1}) - x(t_k)}{t_{k+1} - t_k} \quad (3.11)$$

The time increment is kept constant, and equal to T . Using the above equation, the Van der Pol system can be written in discrete time as

$$\begin{aligned} x_1(k + 1) &= x_1(k) + T x_2(k) \\ x_2(k + 1) &= x_2(k) + T \{ \mu [1 - x_1(k)^2] x_2(k) - x_1(k) \} \end{aligned} \quad (3.12)$$

The quantization error between the discrete and continuous time model depends on the value of T . A smaller value of T will reduce the error and make the discrete model more accurate. However, it will increase the number of iterations required to simulate the model. T should be chosen carefully after analyzing the trade off between the accuracy and the computational cost for the concerned application.

We have developed a discrete time nonlinear dynamic system to define the relationship between the hip-knee, knee-ankle, ankle-toe, shoulder-elbow and elbow-wrist joint angles. Existing oscillatory models [111] were used to derive the terms of the non linear equation. The coefficients of these terms, also referred to as the parameters of the model, were computed based on the observed data points. The model provided us with a mathematical framework necessary to accurately estimate the joint angles.

3.3 Kalman filter

The Kalman filter [112], developed in 1959, is an estimator that utilizes the knowledge of the system dynamics to compute an unobserved state based on observations of the system. It has been used in various applications of dynamic systems and has proved to be efficient and accurate in being able to track and predict based on past observations. The algorithm is a combination of two steps - a prediction followed by a correction. It adds process noise and observation noise (typically white Gaussian) to model the uncertainty in system dynamics and measurements respectively. Originally the Kalman filter was developed to be used with linear dynamic systems. The extended Kalman filter (EKF) [113] used linearization techniques on nonlinear dynamic systems that enabled the application of the Kalman filter. The accuracy of EKF results depends on the linearization error of the system. For systems with a ‘high’ degree of nonlinearity this is often outside of acceptable limits.

3.3.1 Unscented Kalman filter

One of the latest developments in the Kalman filtering literature has been the Unscented Kalman filter (UKF) [114]. It is a nonlinear filter that utilizes the unscented transform [115] to compute the error covariances and the Kalman gain. It generates an ensemble of sample vectors at every time step of the system and uses the non linear equations to propagate all these vectors to the next time step. Assuming a Gaussian distribution, the updated ensemble is used to compute the covariance matrix and the mean. The ensemble for the next time step is generated based on these statistics. While this technique is computationally more expensive than the EKF, it avoids using the linearized equation of the model to update the state covariances and means.

We define the nonlinear dynamic system as:

$$\mathbf{X}(k+1) = f(\mathbf{X}(k)) + \boldsymbol{\nu}(k) \quad (3.13)$$

The observations are a function of the system states,

$$\mathbf{Y}(k) = h(\mathbf{X}(k)) + \mathbf{w}(k) \quad (3.14)$$

$\boldsymbol{\nu}(k)$ and $\mathbf{w}(k)$ represent the process noise and observation noise.

The augmented state vector is computed by appending to the state vector of the nonlinear dynamic system (Eq. 3.13) the process noise vector, $\boldsymbol{\nu}(k)$.

$$\mathbf{X}^a(k) = \begin{bmatrix} \mathbf{X}(k) \\ \boldsymbol{\nu}(k) \end{bmatrix} \quad (3.15)$$

The mean and covariance of the state estimate at time step k are assumed to be of the following form

$$\begin{aligned} \hat{\mathbf{X}}^a(k|k) &= \begin{bmatrix} \mathbf{X}(k) \\ \mathbf{0} \end{bmatrix} \\ \mathbf{P}^a(k|k) &= \begin{bmatrix} \mathbf{P}(k) & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_\nu(k) \end{bmatrix} \end{aligned} \quad (3.16)$$

The ensemble or sigma points are created based on the present covariance and mean.

$$\begin{aligned} \boldsymbol{\chi}_0^a &= \hat{\mathbf{X}}^a(k|k) \\ \boldsymbol{\chi}_i^a &= \hat{\mathbf{X}}^a(k|k) + (\sqrt{n^a \mathbf{P}^a(k|k)})_i, \quad i = 1, 2, \dots, n^a \\ \boldsymbol{\chi}_i^a &= \hat{\mathbf{X}}^a(k|k) - (\sqrt{n^a \mathbf{P}^a(k|k)})_i, \quad i = n^a, n^a + 1, \dots, 2n^a \end{aligned} \quad (3.17)$$

n^a is the dimension of the augmented state vector $\mathbf{X}^a$ and $(\sqrt{n^a \mathbf{P}^a(k|k)})_i$ is the i^{th} row/column of the matrix square root of $n^a \mathbf{P}^a(k|k)$. Details of different computational methods used to generate

the ensemble, χ , of $2n^a + 1$ sigma points can be found in [116].

The sigma points are each propagated using the nonlinear dynamics of the system, Eq. 3.13 to generate the ensemble of *a priori* state estimates.

$$\chi_i^a(k+1|k) = f[\chi_i^a(k|k)] \quad (3.18)$$

The *a priori* mean and covariance can be computed as a weighted sum of the sigma points. [117] discusses the selection of the weights.

$$\begin{aligned} \hat{\mathbf{X}}^a(k+1|k) &= \sum_{i=0}^{2n^a} W_i \chi_i^a(k+1|k) \\ \mathbf{P}^a(k+1|k) &= \sum_{i=0}^{2n^a} W_i [\chi_i^a(k+1|k) - \hat{\mathbf{X}}^a(k+1|k)][\chi_i^a(k+1|k) - \hat{\mathbf{X}}^a(k+1|k)]^T \end{aligned} \quad (3.19)$$

where

$$W_0 = \frac{\kappa}{\kappa + n^a}, \quad W_i = \frac{1}{2(\kappa + n^a)} \quad i = 1, \dots, 2n^a$$

W_0 is the weight assigned to the mean and W_i are the weights assigned to other vectors in the ensemble. The ratio of W_0 to W_i depends on κ , a constant. It has been observed that estimation results are more accurate when the mean is weighed higher than other vectors. The output function, Eq. 3.14 is used to compute the predicted output from the ensemble of state estimates.

$$\gamma_i(k+1|k) = h[\chi_i^a(k+1|k)], \quad i = 0, \dots, 2n^a \quad (3.20)$$

The mean and the innovation covariance are calculated as

$$\hat{y}(k+1|k) = \sum_{i=0}^{2n^a} W_i \gamma_i(k+1|k) \quad (3.21)$$

$$P_{yy}(k+1|k) = R + \sum_{i=0}^{2n^a} W_i [\gamma_i(k+1|k) - \hat{y}(k+1|k)][\gamma_i(k+1|k) - \hat{y}(k+1|k)]^T$$

where R is the variance of the observation noise.

The cross covariance is calculated according to

$$\mathbf{P}_{xy}(k+1|k) = \sum_{i=0}^{2n^a} W_i [\chi_i^a(k+1|k) - \hat{\mathbf{X}}^a(k+1|k)][\gamma_i(k+1|k) - \hat{y}(k+1|k)]^T \quad (3.22)$$

The *aposteriori* states and covariances are then computed using the Kalman gain, $\mathbf{K}$.

$$\mathbf{K}(k) = \mathbf{P}_{xy}(k+1|k) P_{yy}(k+1|k)^{-1}$$

$$\hat{\mathbf{X}}^a(k+1|k+1) = \hat{\mathbf{X}}^a(k+1|k) + \mathbf{K}(k)[y_{obs}(k) - \hat{y}(k+1|k)] \quad (3.23)$$

$$\mathbf{P}^a(k+1|k+1) = \mathbf{P}^a(k+1|k) - \mathbf{K}(k) \mathbf{P}_{xy}(k+1|k)$$

The *aposteriori* states are used as information that was unavailable from the observations of the system. The UKF can be used in applications where the observations need to be smoothed. In such cases there are no unobservable states, but the measurement noise is very high. It is also used to estimate hidden, or unobservable states based on the measurements.

We use the UKF to estimate the joint angles that are not computed from image analysis. The nonlinear dynamic system defining the relation between the different joint angles is used to describe the states. The observation function is determined based on the relationship of the joint angles being estimated and the observed data.

3.4 Classification

We have used two classification methods for gait recognition: (1) Support Vector Machines and, (2) Nearest neighbors. We provide here a brief review of both methods. A detailed description can be found in [118].

3.4.1 Support Vector Machines (SVMs)

SVMs were developed by Cortes and Vapnik (1995) for binary classification. An SVM is a discriminative classifier formally defined by a separating hyperplane. In other words, given labeled training data (supervised learning), the algorithm outputs an optimal hyperplane which categorizes new examples.

Consider the problem of finding a separating straight line for a linearly separable set of 2D-points which belong to one of two classes (see Fig. 3.3). This is a simplified problem where we deal with lines and points in the Cartesian plane instead of hyperplanes and vectors in a high dimensional space.

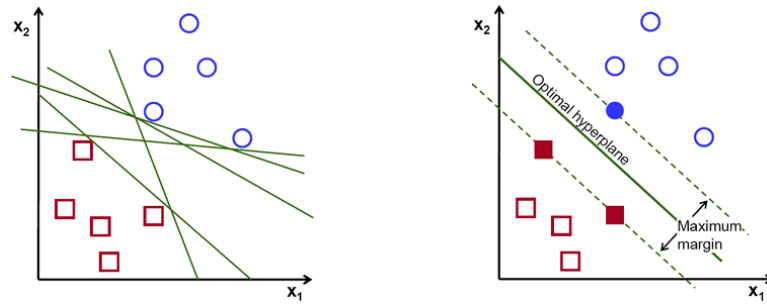


Figure 3.3: Left: Separation of a set of 2D points belonging to two different classes using lines. Right: Optimal hyperplane separating the points.

We note that there exist multiple lines that offer a solution to the problem. We need to define a criterion to estimate the merit of the lines, i.e., which lines can be considered better in partitioning the data into two classes:

A line is bad if it passes too close to the points because it will be noise sensitive and it will not

generalize correctly. Therefore, our goal should be to find the line passing as far as possible from all points.

Then, the operation of the SVM algorithm is based on finding the hyperplane that gives the largest minimum distance to the training examples. Within the SVM theory, *margin* is defined as twice this distance. Therefore, the optimal separating hyperplane maximizes the margin of the training data.

Let's introduce the notation used to define formally a hyperplane:

$$f(x) = \beta_0 + \beta^T x, \quad (3.24)$$

where β is known as the weight vector and β_0 as the bias. The optimal hyperplane can be represented in an infinite number of different ways by scaling of β and β_0 . As a matter of convention, among all the possible representations of the hyperplane, the one chosen is

$$|\beta_0 + \beta^T x| = 1, \quad (3.25)$$

where x symbolizes the training examples closest to the hyperplane. In general, the training examples that are closest to the hyperplane are called support vectors. This representation is known as the canonical hyperplane. We use the result of geometry that gives the distance between a point x and a hyperplane (β, β_0) :

$$\text{distance} = \frac{|\beta_0 + \beta^T x|}{\|\beta\|}. \quad (3.26)$$

In particular, for the canonical hyperplane, the numerator is equal to one and the distance to the support vectors is

$$\text{distance}_{\text{support vectors}} = \frac{|\beta_0 + \beta^T x|}{\|\beta\|} = \frac{1}{\|\beta\|}. \quad (3.27)$$

The margin, M , is defined as twice the distance to the closest examples and is given as:

$$M = \frac{2}{\|\beta\|} \quad (3.28)$$

Finally, the problem of maximizing M is equivalent to the problem of minimizing a function $L(\beta)$ subject to some constraints. The constraints model the requirement for the hyperplane to classify correctly all the training examples x_i . Formally,

$$\min_{\beta, \beta_0} L(\beta) = \frac{1}{2} \|\beta\|^2 \text{ subject to } y_i(\beta^T x_i + \beta_0) \geq 1 \forall i, \quad (3.29)$$

where y_i represents each of the labels of the training examples.

This is a problem of Lagrangian optimization that can be solved using Lagrange multipliers to obtain the weight vector β and the bias β_0 of the optimal hyperplane.

3.4.2 Nearest Neighbors

Nearest neighbor (NN) algorithm is a non-parametric method used for classification. It makes no assumptions about the form of the function, $y = f(x_1, x_2, \dots, x_p)$ that relates the dependent variable, y , to the independent/predictor variables $x_1, x_2, \dots, x_p$. Hence, it does not involve estimation of parameters in an assumed function form such as the linear form. The NN algorithm is among the simplest of all machine learning algorithms. In NN classification, the output is a class membership. The training examples are vectors in a multidimensional feature space, each with a class label. The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. An object (an unlabeled query or test point) is classified by its single nearest neighbor, with the object being assigned to the class of the training sample nearest to that query point. The idea of using a single nearest neighbor to classify observations can be very powerful when we have a large number of observations in our training set.

The generalized form of the NN algorithm is called the k -nearest neighbor (k -NN) algorithm, where the object is classified by the majority vote of its neighbors, with the object being assigned to the class most common among the k training samples nearest to that query point (k is a positive user defined constant, typically small). A commonly used distance metric for k -NN is Euclidean distance. We have used cosine similarity to measure the distance between two normalized histograms

of normal flow.

Chapter 4: Gait Databases

We primarily use three gait databases: George Mason University (GMU) Gait Database which was collected inhouse with the approval of GMU human subjects review board. The consent of all subjects was obtained before doing the video recording. Carnegie Mellon University 3D Motion Capture Database was used to learn the motion models between different body segments, this database is available freely for research use. Finally, we have obtained the permission from Dr. Mark Nixon [28] to use the University of Southampton Gait Database to perform gait recognition experiments. Here we describe each database in detail.

4.1 George Mason University (GMU) Gait Database

The GMU gait database consists of two parts:

(A) First one consisting of 11 healthy adults (males and females) wearing at least 3 different shoe types while walking at both slow and fast walking speeds. We recorded a total of 132 sequences (66 in the sagittal plane and 66 in the frontal plane). These sequences were short and were used to test the gait phases in the walking cycles.

(B) The second one consists of 5 healthy adults walking at normal and fast speed and taken from the sagittal plane only. We recorded a total of 20 sequences (10 for each background). Special care was taken to ensure that the recordings contain more than one gait cycle for each subject. These sequences were used to predict unobserved motion of body segments based on the learned models and the observed motion of other body segments.

4.1.1 Experimental Design

A. In the first part of the GMU gait database recording sessions, we did not vary the filming environments. However, we did ask the subjects to walk in different shoe types (varying heel height) and walking speed (slow, normal and fast). To capture the sequences, we have used a Dragonfly[®] 2 color camera. The data has been recorded in the frontal and sagittal planes at 60 frames/sec with a resolution of 640×480 pixels. The illumination was coming from the overhead fluorescent lights. The selected location also provided a non-reflective surface (i.e. carpet) and did not have any windows (i.e. indoor lighting only). Indoor lighting and non-reflective surface was specifically chosen to simplify the process of background subtraction. The subjects were asked to walk for at least 18 feet before their motion was recorded to acquire their normal walking pattern. Unfortunately the area did not provide enough room to properly capture an entire gait cycle in the sagittal plane. Fig. 4.1 shows some example frames from the collected sequences.



Figure 4.1: Selected frames from the GMU gait database A in the frontal and sagittal planes.

B. The second part of the GMU gait database was captured in the sagittal plane only. The recorded sequences were recorded with a Dragonfly[®] 2 color camera. Images were acquired

at 60 frames/sec at 640×480 pixels using best quality compression (H.264). We used two different backgrounds and asked the subjects to walk in normal and fast walking speeds. The sequences contained more than one gait cycle for each subject and they walked for at least 18 feet to settle into their normal walking pattern. Some example images are shown in Fig. 4.2.



Figure 4.2: Selected frames from the GMU gait database B in the sagittal plane.

4.2 Carnegie Mellon University 3D Motion Capture (CMU MoCap)

Database

The CMU MoCap database consists of a very large 3D motion dataset consisting of 2605 trials in 6 categories and 23 subcategories. We have used the walking motion capture data under the Locomotion category in our experiments. We have used 15 subjects' gait data to learn the motion models. In this database, the same person may appear under more than one subject number.

4.2.1 Experimental Design

The CMU MoCap data was recorded using Vicon motion capture system consisting of 12 infrared MX-40 cameras, each of which is capable of recording at 120 Hz with images of 4 megapixel resolution. Motions are captured in a working volume of approximately 3m x 8m. To capture something, small grey markers are placed on it. Humans wear a black jumpsuit and have 41 markers taped on. The Vicon cameras see the markers in infra-red. The images that the various cameras pick

up are triangulated to get 3D data. This 3D data can be used in two ways:

- Marker Positions.
- Skeleton Movement.

We have used the marker positions on various body segments to extract the data for our models.

Fig. 4.3 shows the placement of various markers on human body.

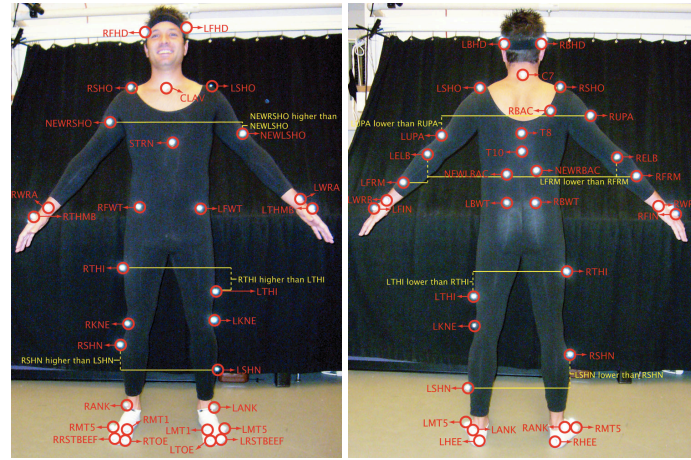


Figure 4.3: Pictures showing the accurate placement of different markers in CMU MoCap Database.

4.3 University of Southampton (USH) Gait Database

The SOTON database is one of the larger available gait databases. The database is composed of 114 subjects. One of the purposes of the SOTON database was to collect a large number of sequences on a treadmill, indoors, and outdoors. The primary purpose of this database is biometric gait recognition, although we are permitted to use it in a soft biometrics setting as well, which is the primary way that we use the database in this thesis. The treadmill sequences are intriguing for several reasons. The treadmill provides a long stable sequence of a subject walking, while not changing position considerably. For this reason, we can collect arbitrarily long sequences for enrollment into biometrics databases. However, there is some question about whether people alter

their gait when they are on the treadmill. That is, whether there is sufficiently similarity between treadmill and non-treadmill walking conditions such that recognition algorithms can be applied.

4.3.1 Experimental Design

The indoor sequences use a green background to permit foreground extraction using chroma-key. Chroma-key provides extremely accurate foreground information for the indoor sequences (treadmill and non-treadmill). The outdoor sequences are realistic surveillance images, with other cars, bikes, trees, etc. moving in the image. The treadmill sequences were recorded with considerable care. The handrail on the treadmill was removed, and a radio was played so that the subjects would not try to fill the empty silence by interacting with researchers. However, there are still many cases of subjects looking down at the treadmill console to check their walking speed. Subjects were asked to walk at a speed that is normal for them. All images were recorded at 24 frames per second, and provided via a hard drive in a .DV format. Fig. 4.4 shows some frames from the treadmill, indoor and outdoor gait sequences.



Figure 4.4: Pictures showing the frames from USH gait database.

Chapter 5: Identifying Phases of Gait Cycle Using Image Motion

Normal human walking is characterized by the smooth up and down movement of the trunk. The six determinants of gait [5]: pelvic rotation, pelvic obliquity, stance-phase knee flexion, foot-ankle mechanisms, ankle and knee interactions, and lateral displacement of the pelvis, minimize the displacement of the body's center of mass (COM) and smooth its trajectory, thereby saving energy. This vertical excursion of COM provides important information about the phases of gait. The vertical excursion is minimum at double support and maximum at mid swing phases, occurring twice in a gait cycle. We have used this information to find the periodicity of subjects' gait cycles. Researchers have used silhouette (shape) based approaches to find the length of the gait cycle. However, these methods suffer as the quality of the background subtraction decreases. This could be due to the effect of poor lighting, contrast or even loose clothing. Some model based approaches use the parameters to estimate the pose directly. However, all these methods use image data at 30 frames/sec, which is insufficient to compute the other major events of the gait cycle such as toe off and heel strike. At lower frame rate, the events overlap with each other and hence are not correctly identified. We use the motion information at 60 frame/sec to compute all the major events in the gait cycle reliably. In Sec. 5.1, we describe image flow based techniques to compute the vertical excursion and identify the double support and mid swing events in a gait cycle. We then show that these events can be used to align the gait sequences of different subjects in different view planes. In Sec. 5.2, we use the above extracted events in the sagittal plane to compute the motion of lower legs. The zero crossings of the estimated lower leg velocities are then used to identify the toe off and heel strike events. We also show the validity of our algorithm by comparing our results with the results from a 3D motion capture system.

5.1 Double support and mid swing

We use the gait data from several subjects in different walking conditions (different types of shoes, two different view planes: frontal and sagittal). The recorded sequences were processed using background subtraction to detect a moving person. The background is learned over time and it is then subtracted from the image frames to obtain the foreground/silhouette of the subject. We compute convex hull of the foreground region containing the person and take the top 30% of that area for further processing. Fig. 5.1 shows the results of foreground detection using background subtraction and the computed convex hull.



Figure 5.1: Foreground detection in the image sequence using background subtraction.

5.1.1 Motion Models

Though people can move the neck to expand their field of vision, during normal gait the head and trunk do not change positions except vertically [5]. The vertical up and down movement is with the body's center of gravity while in locomotion. We are interested in this translational motion of person's torso and head which are at a fixed distance with respect to the camera. We have used normal flow computed from pair of images in a given image sequence to estimate the motion in images. We explain here two types of motion models: body motion model and camera motion model. Fixing the coordinate system to the body, we fix the origin at the sternum, x axis towards the right side, y axis against the gravity (vertically up) and z axis in the direction of motion respectively. The body motion model can then be written as:

$$T_x = 0, \quad T_y = V_t, \quad T_z = T_t$$

where T_t is the forward translation of the person and V_t denotes the up and down motion of person's head.

The equations of motion for translational movement with respect to camera are given by:

$$\dot{x} = (U - xW)/Z, \quad \dot{y} = (V - yW)/Z$$

where $\dot{x}$ and $\dot{y}$ represents the translational components of projected instantaneous velocity, (x, y) are the coordinates in the camera frame, Z is the depth, and (U, V, W) are the translational parameters. The focus of expansion (FOE) [106] is given by $(U/W, V/W)$. For any given frame, at a fixed camera position, depth Z is approximately constant, say Z_0 .

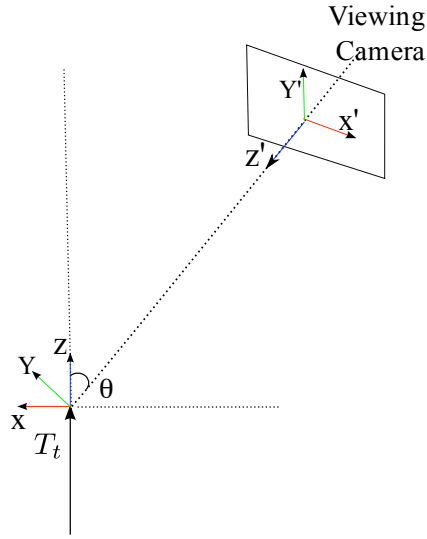


Figure 5.2: Camera position with respect to the walking person with translational velocity T_t .

We assume that the person is moving over a flat surface and the camera viewing direction is at an angle θ relative to the person's direction of movement (see Fig. 5.2). The motion parameters in the camera frame are given as,

- i. Forward translation, $U = T_t \sin \theta$.
- ii. Up and down movement, $V = V_t$.

iii. Expansion, $W = T_t \cos \theta$.

Without loss of generality, we can assume the camera is observing two extreme cases: frontal and sagittal view planes.

5.1.1.1 Frontal Plane View

Here distance of the person from camera is much greater than the variation in the depth, so depth can be assumed to be a constant, Z_0 . For the frontal plane, $\theta = 0$, which implies:

$$U = T_t \sin \theta = 0, \quad V = V_t, \quad W = T_t \cos \theta = T_t$$

Hence, focus of expansion is given by, $(0, V_t/T_t)$. Now, since the expansion, T_t , is approximately constant, the variation in FOE is representative of the head excursion. As an example, for the first sequence given in Fig. 4.1, the expansion is about 6 pixels in 65 frame interval, i.e. about 0.09 pixels/frame. The normal range of maximum head excursion is 5-6 cm for a full gait cycle, i.e. the vertical movement is approximately 0.33 cm/frame. We find the up and down movement as 0.38 pixels/frame. Therefore, in the frontal plane, the up and down movement of the torso dominates the amount of expansion. As the head goes up the FOE shifts down in the image, and as the head excursion decreases, the FOE with respect to the subject's body moves up. The instances where the FOE crosses *zero* value signifies the reversal in the direction of instantaneous velocity of subject's head. These zero crossings in turn determine the key frames in the image sequence.

The FOE is computed by voting over a larger area in the neighborhood of the foreground being processed. Voting calculates for each pixel p_i , the number of pixels that contribute towards p_i being the FOE. The pixel with highest number of votes is determined as the FOE. To speed up the calculations, computations have been done in small overlapping grids (see Fig. 5.3). Finding the angle of FOE with respect to the center of person's body computes the image frames that correspond to the minimum and maximum head excursions. The minimum head excursion relates to the double leg stance of the gait cycle, whereas the maximum excursion corresponds to the single leg stance [5].

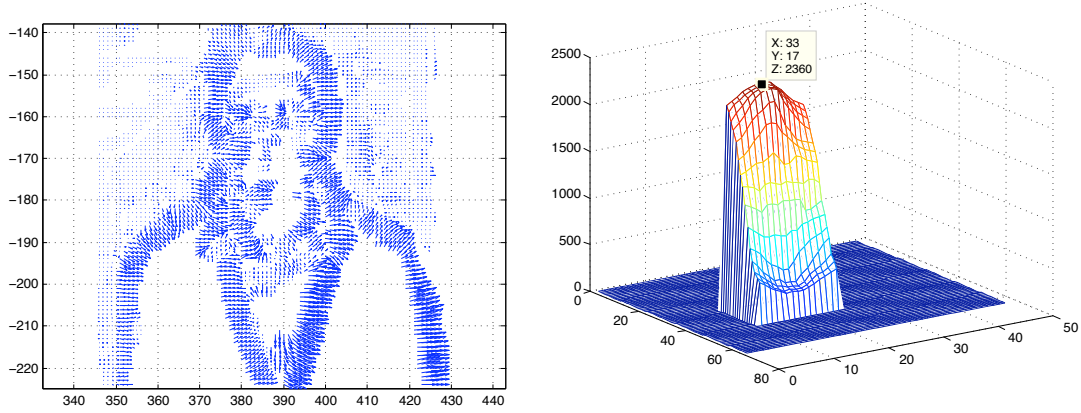


Figure 5.3: Example of the image flow from a frontal plane image sequence followed by the figure showing the focus of expansion (FOE) computed using voting.

5.1.1.2 Sagittal Plane View

The person's motion is parallel with respect to the camera. For the sagittal plane, $\theta = \pi/2$, which implies:

$$U = T_t \sin \theta = T_t, \quad V = V_t, \quad W = T_t \cos \theta = 0$$

Hence, expansion is 0 and focus of expansion is outside the image. In the sagittal plane, the head excursion is comparable to the translational velocity of the person moving forward. Hence, the up and down movement can be computed reliably and accurately. For example, for the fourth sequence in Fig. 4.1, the translational velocity is around 2-3 pixels/frame and the head excursion is 1.2 pixels/frame. Computing the up and down movement from the image sequences enables us to determine the key frames where reversal in the direction of head-velocity occurs. These key frames represent different transition points of the gait cycle.

In the sagittal plane, when the person is moving parallel to the camera, a two parameter translational model can be fitted to the flow. This model is based on voting using the flow values of the pixels. A grid is formed by using the angles $[0, \pi]$ and motion magnitudes in the interval of $[-20, 20]$. Given the translational motion, $\vec{T} = (p_x, p_y)m$, where (p_x, p_y) are the components in the (x, y) direction and m is the magnitude. We project this translational velocity, $\vec{T}$, on normal flow direction (n_x, n_y) as, $\forall(x, y), p_f = (n_x p_x + n_y p_y)m$. The normal flow residual at (x, y) is

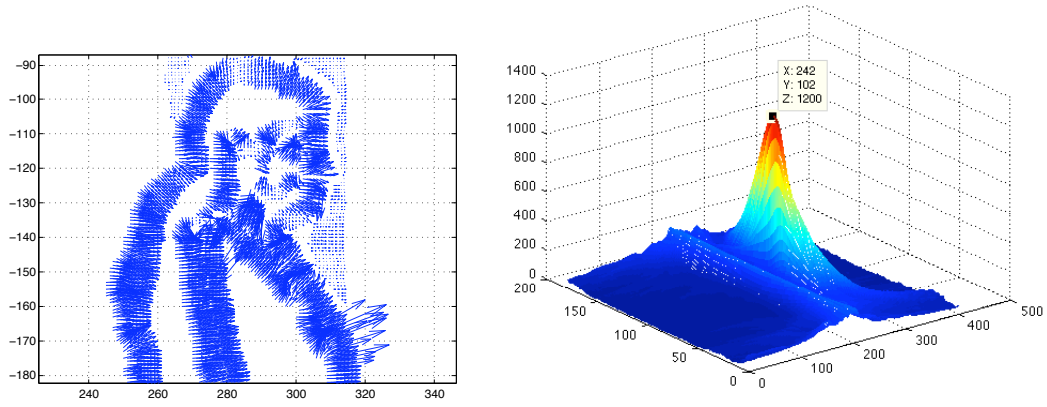


Figure 5.4: Example of the image flow from a sagittal plane image sequence followed by the figure showing the results of voting which determines the amount of translation in x and y directions.

$(p_f - n_f)$. The vote for $\vec{T}$ is given by, $e^{-|p_f - n_f|^2 / 2\sigma^2}$. The total vote for $\vec{T}$ is the sum of votes at all points.

The value of $\vec{T}$ for which the maximum number of votes are obtained is the estimated translational value, (u, v) , (in pixels) (see Fig. 5.4). The values of translation in y (vertically up) direction, v , correspond to the up and down movement of the head. We find the zero crossings of v values over the whole sequence which gives desired single leg and double leg stances (see Fig. 5.5).

5.1.2 Synchronization of Human Gait

Once the phases of gait have been determined using image flow, we can synchronize any two gait videos, even when they correspond to different individuals or have been collected from different viewpoints. The phases of gait (double support and mid-swing) correspond to key frames of the video sequence. We proceed as follows:

1. As a first step to matching, we match the key frames of the videos, double-support to double-support and mid-swing to mid-swing.
- 2a. In the second step, if the lengths of the sub-sequences between two key frames are the same, then we match the frames sequentially.
- 2b. If the subsequences are of different lengths, then a few frames in the longer subsequence are

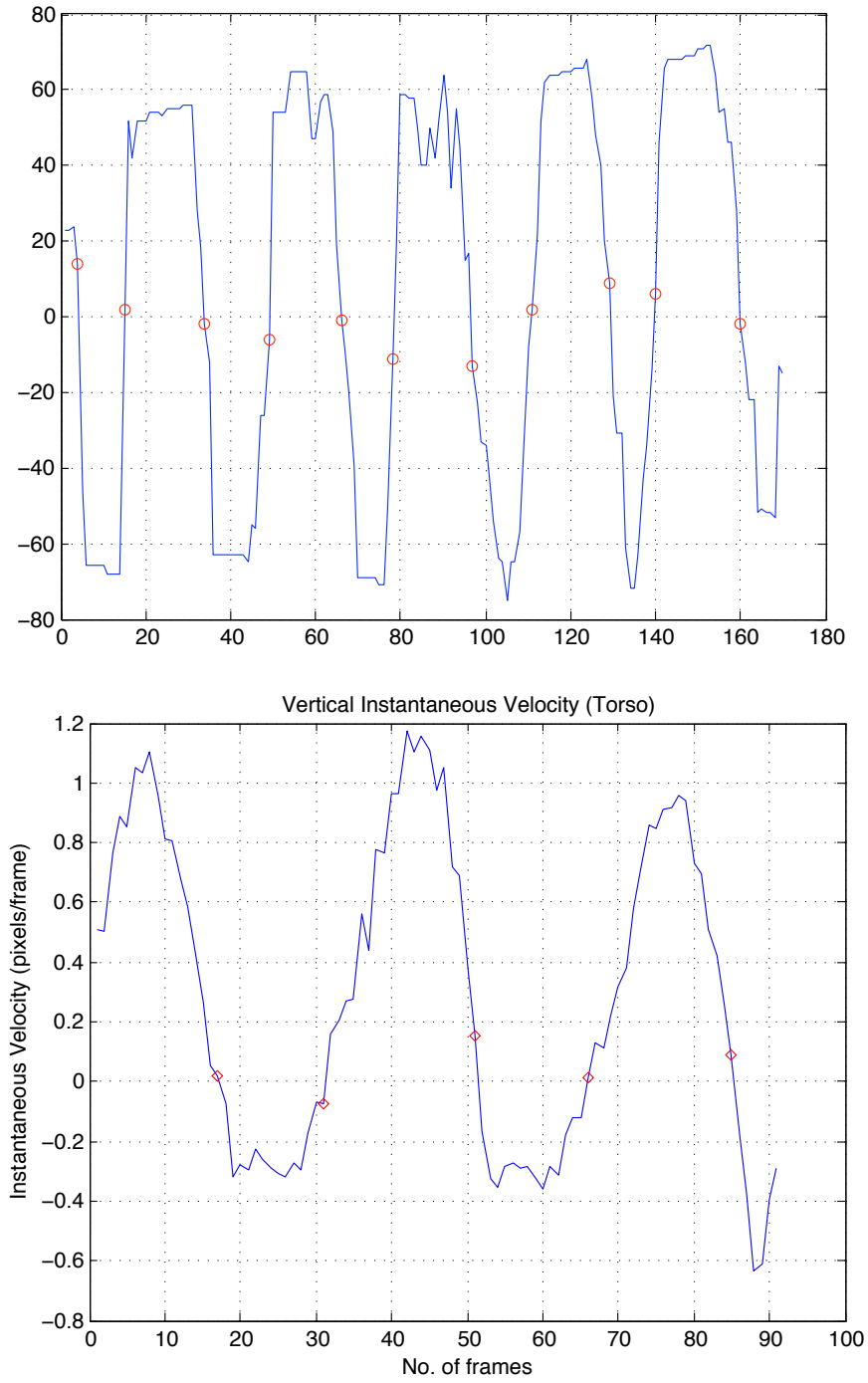


Figure 5.5: The first figure shows the the change in the angle of FOE with respect to the center of subject's body in the frontal plane. The second figure shows the variation of the translational value in the vertical direction of motion in the sagittal plane. The zero crossings indicate the key frames of the image sequences.

dropped and then we match the remaining frames sequentially.

These steps allows us to compare the gait sequences of different individuals and to detect anomalies in their walking patterns.

5.2 Toe off and heel strike

Additional gait events such as, toe-off and heel-strike can be identified by observing the instantaneous rotational velocity of the lower leg segments (shanks). These events occur at zero values of the shank's rotational velocity. For finding these events, we consider the lower 28.5% of the body height H as suggested [3]. We first state the assumptions of our method and then describe the estimation of lower leg motion.

5.2.1 Assumptions

- The videos are captured in the sagittal plane at a high frame rate (at least 60 Hz). Lower frame rate is insufficient to capture all the events of the gait cycle.
- The gait sequences contain more than one gait cycle (preferably two or more).
- All the subjects are within the normal range of gait parameters.

5.2.2 Finding the lower leg motion from images

The motion of the lower leg segments can be described by translation in the sagittal plane, (t_x, t_y) , and rotation, ω , around an axis orthogonal to the sagittal plane. It can be formalized using a reduced version of an affine motion model as described in [27]. We use displacement as an estimate of velocity. The model defines $(\dot{x}, \dot{y})$, i.e. the instantaneous velocity at point (x, y) ,

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \quad (5.1)$$

Using multiple edge points and the corresponding normal flow vectors, the parameters (ω, t_x, t_y) are computed as the linear least squares (LS) solution [27].

Before fitting the models, we need a way to segment the lower legs from each other. We have approached this problem in two ways. In the first method, we use tracking based segmentation with RANSAC [119]. In the second method, we have used spatial and motion information about the lower legs while walking.

5.2.2.1 Tracking based segmentation

We use a slightly modified version of RANSAC algorithm to find motion models for the lower legs in each frame. We first divide the foreground into small, equally spaced overlapping patches (superpixels) to reduce the overhead of dealing with each flow vector thereby saving time and reducing complexity. We compute a motion model, M_{tr} , for each superpixel using Eq. (5.1) and depending on its error being less than a threshold value, the superpixel is marked as good or bad.

We automatically initialize the left and right lower leg boundaries in the double support by using their spatial separation and compute motion models for both of the lower legs using the superpixels inside respective boundaries. We then predict the new position of each pixel in the next frame

Algorithm 1 RANSAC for estimating motion models

INPUT: Boundary B of current region, set of superpixels S , and their respective motion models M of size n , where n is the number of total superpixels.

OUTPUT: Model m_{best} , best describing the motion of all superpixels within B .

repeat

 Randomly sample two superpixels s_i and s_j (within B) for which the models m_i and m_j have error less than a threshold value, t .

 Compute a model $m_{i,j}$ for the combined superpixel $s_{i,j}$.

if $m_{i,j}$ has error value $\leq t$ **then**

 Apply $m_{i,j}$ to all superpixels inside B and compute the error for each superpixel.

 Count the total number of superpixels that have error $\leq t$.

end if

until number of iterations $\leq N$

 Find the model, m_{best} that has the largest number of votes/patches.

return m_{best}

(based on the motion models) and estimate new boundaries. To tackle the noise introduced by noisy measurements of flow and motion estimation, the boundaries are adjusted by applying Principal

Component Analysis (PCA) [120] on the pixels marked as inliers. PCA estimates the orientation of the lower leg segments using those pixels. These boundaries are then used to find new models using the approach given in Algorithm 1 . Once the models are found, the whole process is repeated for the next frame. Our approach allows us to segment and classify the lower leg pixels belonging to either the left or the right leg segment. This computation gives us the estimation of instantaneous rotational velocities of lower leg segments for each frame of the image sequence.

5.2.2.2 Spatial and Motion segmentation

The major limitation of the first approach is that no feedback is used to correct the computation of the noisy boundaries belonging to the lower legs. We propose to use the double support and mid swing information obtained in Sec. 5.1, together with the knowledge of biomechanics to simplify this problem. Once we have obtained the keyframes corresponding to double-support and mid-swing, we divide the whole gait cycle into different sets based on their distance from the keyframes. We use the following observations:

1. The lower legs in the frames belonging to the double-support set, are spatially separated from each other. In other words, there is no occlusion between the left and right lower leg.
2. The lower legs in the frames belonging to the mid-swing set, can be segmented using their motion relative to the motion of the torso. The swinging limb has a velocity greater than the torso velocity whereas the stance (standing) limb's velocity is lower compared to the velocity of torso.

We first introduce the concept of Histograms of normal flow that are used in the segmentation and motion estimation of lower legs.

5.2.2.2.1 Histograms of normal flow

As mentioned in Sec. 5.2.2.1, we divide the image region into overlapping superpixels: first to compute a coarse motion model and then use it to refine the motion model at pixel level thus eliminating

the noise and interference with the correct motion estimation. We use the superpixels of size 20×20 . For each of these superpixels, we compute histograms of normal flow. In each superpixel, we assign the pixels into 16 different bins based on their orientation. The count of pixels, the median value and the standard deviation for each bin is computed. Examples of histogram obtained from the normal flow are shown in Fig. 5.6.

Let $A_{16 \times 2}$ be a collection of unit vectors representing 16 bins. We have used weighted least squares [121] for solving the histogram values. The weight matrix, $W_{16 \times 16}$ is a diagonal matrix in which the terms correspond to the number of flow vectors in each bin. $B_{16 \times 1}$ is a vector containing the median magnitudes of the vectors stored in the particular bin of interest. The resultant flow vector for the superpixel, $\mathbf{v}$, is given by,

$$(A^T W A) \mathbf{v} = A^T W B \quad (5.2)$$

Each superpixel is represented by its center, C and the resultant flow vector, v . Now, we describe the use of these superpixels for segmenting and estimating motion models for the lower legs.

5.2.2.2.2 Spatial segmentation

For frames that are closer to double support, the lower legs are not occluded from one another. This makes the segmentation process easier. We first project all the superpixels belonging to the lower 28.5% of body height on the horizontal axis and find the separation between two limbs based on the projected number of superpixels. We simply find the maximas between the first half and the second half. The minimum value between them gives us our dividing point. We then label the superpixels on different sides of the dividing point as belonging to two different lower leg segments. We compute a motion model, $M_{tr_i}, i = 1, 2$ for both limb segments based on the superpixels being labeled as left or right limb. Fig. 5.7 shows the segmentation of two lower leg segments.

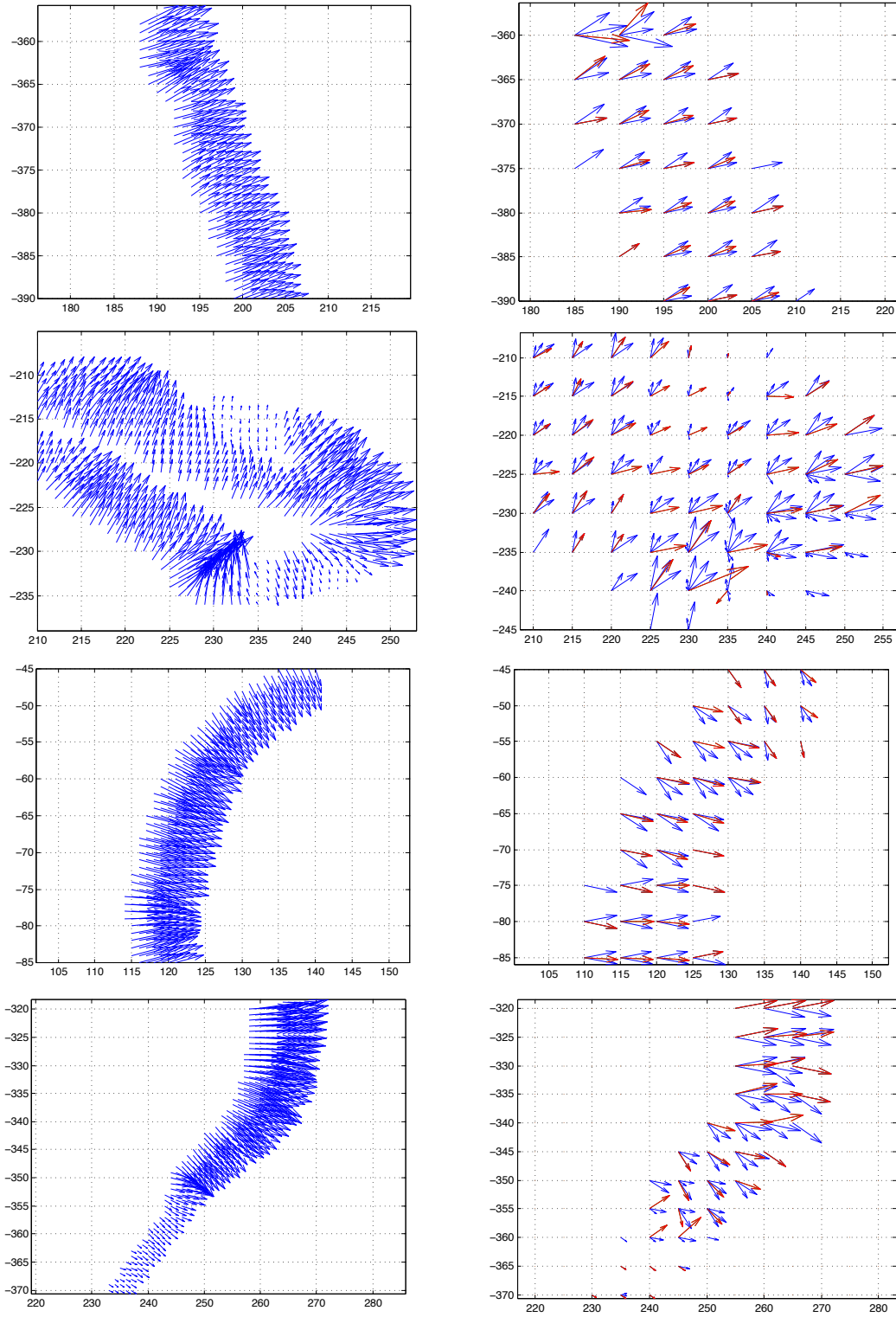


Figure 5.6: Examples of normal flow and their respective histograms on the right. The red flow vectors represent the resultant flow vector in each superpixel computed using the least squares solution given by Eq. 5.2

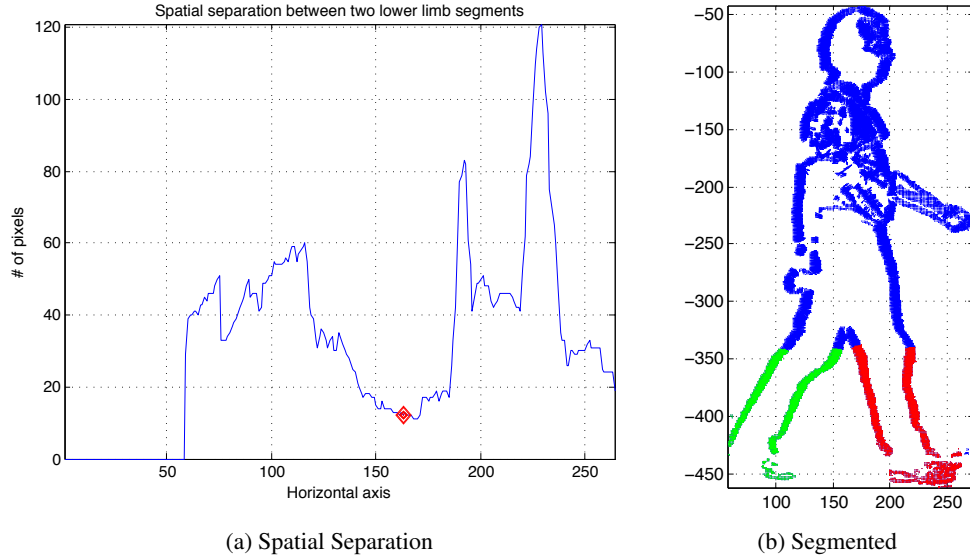


Figure 5.7: (a) Spatial Separation found for lower leg segments. The dividing point is shown in red(b) The segmented lower legs are shown in red and green based on the dividing point.

5.2.2.2.3 Motion Segmentation using Histograms of Normal Flow

In frames closer to mid swing, the lower legs are occluded by different amounts, therefore we compare the motion of the lower leg segments with the torso. Once we compute the resultant flow values of superpixels, we compare its motion magnitude against that of the torso. The superpixels whose motion is lower than that of the torso are classified as belonging to the standing (stance) limb. The superpixels whose magnitude is greater than the torso motion are labeled as belonging to the swinging leg. We then compute two different motion models according to the labeling of superpixels. The models are fitted to the entire lower leg segment region (at the pixel level) and residuals are computed. Based on the value of residuals, two finer motion models are fitted to the lower leg segments. Fig. 5.8 shows the normal flow values for the superpixels computed using our method.

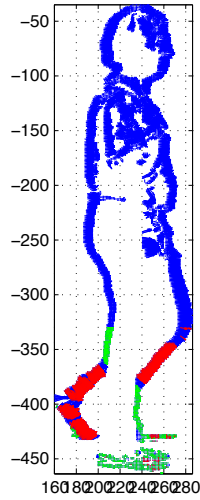


Figure 5.8: Segmentation of lower legs during mid swing using superpixels and histograms of normal flow

5.2.2.2.4 Lower Leg Rotational Velocities and Angular Positions

The parameter ω estimated in Sec. 5.2.2.2.2 and Sec. 5.2.2.2.3 represents the instantaneous rotational velocity of lower legs at each frame. The residual error in model estimation in certain frames can contribute to an inaccurate velocity profile. This error is accumulated over frames. In order to account for this inaccuracy, we find the orientation of lower legs in frames corresponding to double support and mid swing. To achieve this, we fit Radon transforms [120] to the boundaries of the segmented lower legs, which gives the desired orientation. The residual error is then distributed over the intervals between instances of double support and mid swing.

Fig. 5.9 shows the instantaneous rotational velocity profiles of both left and right lower leg segments of a subject for a single sequence. The zero crossings are computed automatically and correspond to the toe-off and heel-strike events of the gait cycle.

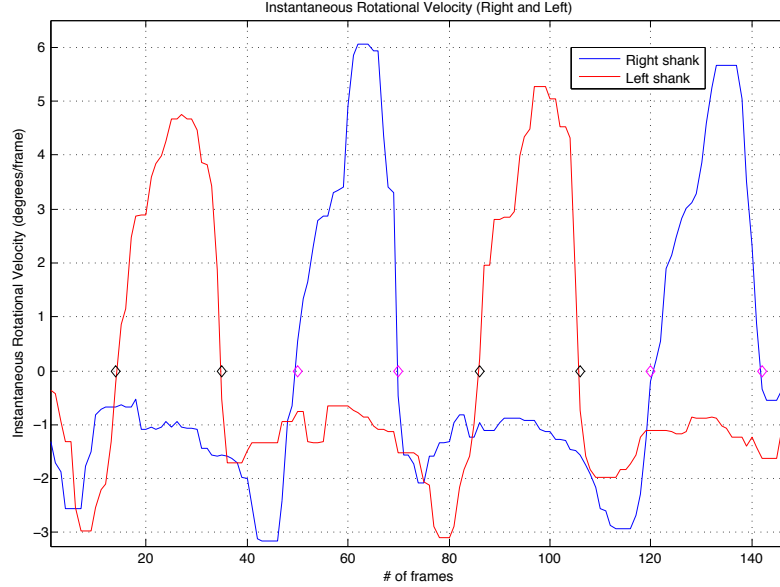


Figure 5.9: Instantaneous rotational velocities computed for the two lower leg segments using our proposed method. The zero crossings have been marked using black and magenta diamonds.

5.3 Finding the hand position from images

We use Mean Shift algorithm for tracking hand in the whole gait sequence. [122] proposed the mean-shift approach to non-rigid object tracking in the joint spatial+color space, (r, g, b, x, y) , where (r, g, b) represents the color and (x, y) represents the spatial location. Given an image, a color model, and a tracking region, the algorithm weights the offsets from the center of the tracking region according to distance from the center and color correspondence. Effectively if the region down and left of the initial center estimate contains colors closer to the color model, the center is moved somewhat in that direction. Once a direction vector is obtained, an offset from the current center estimate is calculated, and the resulting color model compared to the reference color model. If the color model match is worse at the offset than the initial estimate, the length of the offset is halved repeatedly until this is not the case or the offset length falls beneath a threshold. The estimate is then updated and the process repeated until convergence.

[123] proposed a method for updating color models over time to aid tracking, and for resizing the tracked area. We have integrated these improvements with our mean-shift implementation. The

color model update works by comparing the latest color model C_{curr} to the reference model as of L frames before C_{curr} , R_L , and updating the current reference model by $R_{curr} = R_{curr} + 1/L * (C_{curr} - R_L)$. In our case we chose $L = 10$ frames. Resizing the tracking area was supposed to work by resizing the tracked area according to the ratio of the new standard deviation of the offsets weights in each direction to the previous ones, but in practice this ratio was always very close to 1 in our trials. The offsets from the center of each pixel are weighted by a Gaussian. The algorithm takes as inputs: a starting center position of hand, a tracking region size, and the resolution of the color histogram to use (typically $32 \times 32 \times 32$). Fig. 5.10 shows the (x, y) positions of a hand tracked in one of the gait sequences.

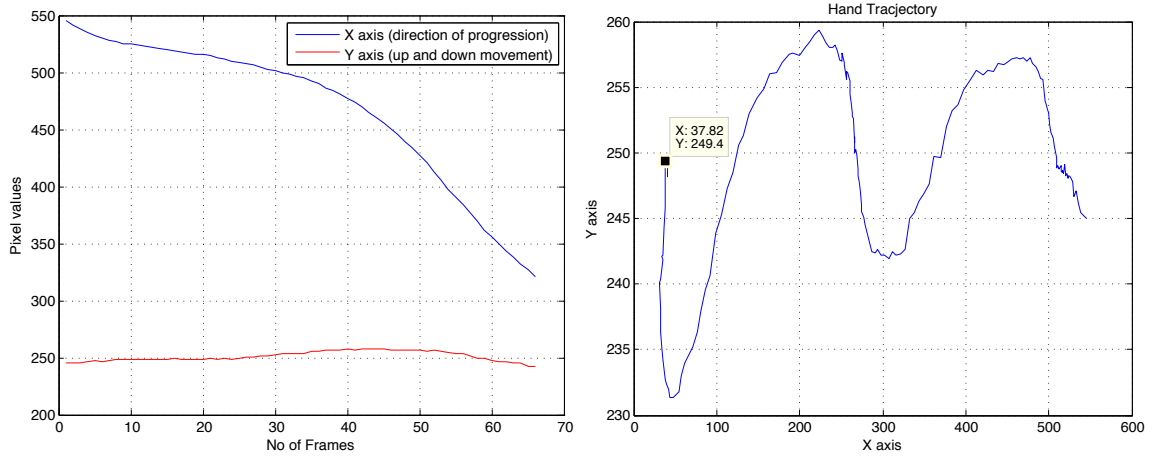


Figure 5.10: (x, y) positions of the tracked hand in one of the gait sequences.

5.4 Experiments

We have tested our methods on GMU gait databases A and B. The foreground is first extracted from the image frames by subtracting the background. We extract double support, mid swing, toe off and heel strike events from the zero crossings occurring in the instantaneous vertical and rotational velocities of torso and lower legs respectively.

5.4.1 Double support and mid swing

We have analyzed GMU A video sequences as explained in Sec. 5.1.1. For ground truth we have extracted the identified image frames showing various phases of gait. Fig. 5.11 and Fig. 5.12 show the results from two of the frontal and sagittal plane sequences.



Figure 5.11: Figure showing the frames from the original sequence corresponding to the zero crossings in Fig. 5.5. The rows correspond to the alternating maximum and minimum head excursions.

5.4.1.1 Synchronization of gait sequences

Fig. 5.13 shows an example of synchronizing two short subsequences from videos of the same person obtained from sagittal and frontal plane views. Fig. 5.14 shows an example of synchronizing two short subsequences from videos of the two different persons obtained from the same view plane. It can be observed from the synchronized sequence that first subject has greater lateral movement of knee as compared to the first subject.



Figure 5.12: Figure showing the key frames from the original sequence of frontal and sagittal plane sequences. The rows correspond to the alternating maximum and minimum head excursions.



Figure 5.13: Synchronized frames for the same person in multiple view planes.



Figure 5.14: Synchronized frames for different individuals in the same view plane.

5.4.2 Toe off and heel strike

We analyzed the video sequences captured in the sagittal plane as explained in Sec. 5.2. We identify the toe off and heel strike events from the zero crossings of the instantaneous zero velocity of the lower leg segments. The extracted frames corresponding to these crossings for several gait sequences are shown in Fig. 5.15.

5.5 Validation

To establish the validity of our method, we have compared our results with the Optotrak motion capture system. We captured a subject's gait data sequence using video camera (sequence S_c) and Optotrak markers (sequence S_o) simultaneously. The markers were fixed on the subject's head and lower leg segments. The sampling rate for the Optotrak system was set as $100Hz$ in the experiment. Fig. 5.16 shows the arrangement of the simultaneous data capture from video and Optotrak. Different gait events were found using our technique and Optotrak data. Table 5.1 and Table 5.2



Figure 5.15: Figure showing the frames from the original sequences corresponding to the zero crossings in Fig. 5.9. The rows correspond to the alternating maximum and minimum lower leg angles.

compare the time of occurrences of all the identified gait events. The numbers reported in the tables are in seconds where DS stands for double support, MS for mid swing, TO for toe off and HS for heel strike. We have used linear interpolation to find exact time occurrences of zero instantaneous velocities for various events.

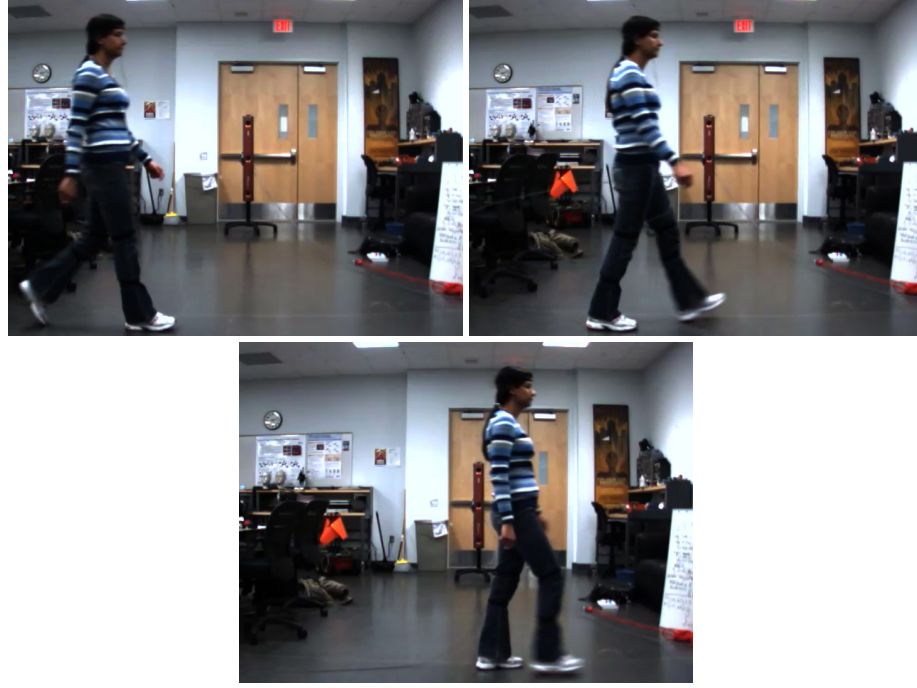


Figure 5.16: Figure showing the settings for the simultaneous recording of video and Optotrak data.

Table 5.1: Table comparing the time of double-support and mid-swing occurrences (Optotrak, S_o , vs Video camera, S_c)

	DS1	MS1	DS2	MS2
Optotrak	4.254s	4.540s	4.835s	5.125s
Camera	4.261s	4.526s	4.883s	5.122s
Abs. Diff	0.007s	0.014s	0.048s	0.003s

The average error computed over all the gait events is 0.0164 ± 0.0196 seconds. Taking into consideration the frame rate of the video camera used, the error values are very small showing that our results are in agreement with the Optotrak data. This reinforces our approach of using image-flow-based technique as a markerless system for obtaining the key events of gait cycle accurately

Table 5.2: Table comparing the time of toe-off and heel-strike occurrences (Optotrak, S_o , vs Video camera, S_c)

	TO1	HS1	TO2	HS2
Optotrak	4.989s	5.317s	4.751s	5.559s
Camera	4.942s	5.311s	4.753s	5.563s
Abs. Diff	0.047s	0.006s	0.002s	0.004s

and reliably.

5.6 Discussion

In this chapter, we have used normal flow to identify the double support and mid-swing events in the gait cycle. We have used the head excursion (instantaneous torso velocity) to find the key frames corresponding to these phases in both the sagittal and the frontal planes. These key frames are then used to synchronize the gait cycles of various subjects in different view planes. Our method allows the detection of double-support and mid-swing phases even with imperfect background subtraction. This makes our method more robust and reliable compared to the methods that use silhouette extraction, which depends on the accurate segmentation of foreground as a pre-processing step. We have also used these events with image motion to compute the lower leg rotational velocities in the sagittal plane. These velocities are estimated by fitting motion models to the lower regions of the body, segmented using the spatial and motion information. The zero crossings in the velocities determine the key frames corresponding to toe off and heel strike events. We have also used mean shift tracking with adaptive color models to track the hand in the gait sequences.

Chapter 6: Modeling and Estimating Motion of Unobserved Segments

In order to estimate the motion of the overlapping and/or occluding segments (upper leg, foot and upper arm), we model their motion relationship with those of their distal counterparts (lower leg and forearm). We then predict the unobserved motion using Unscented Kalman filter. Our approach presents a new way to use non linear dynamic models with Kalman filters which have been previously used only for constant velocity assumptions [15, 16]. We have used CMU 3D motion capture database to capture the motion relationship between different limb segments. In Sec. 6.1, we describe the modification and projection of 3D data from CMU to the sagittal plane using Principal Component Analysis on 3D marker dataset. In Sec. 6.2, we explain the modeling between the 2D motion data of different limb segments using non linear dynamic models. Implementation of Unscented Kalman filter to predict the motion of unobserved limb segments is described in Sec. 6.3. We use color cues from images to correct our predicted motion values. We also discuss automatic initialization of body segments/joint positions in Sec. 6.5. Finally, we present our results and analyses in Sec. 6.6 and Sec. 6.7.

6.1 Transforming 3D data to sagittal plane

We require the 3D gait data to be projected on the sagittal plane for 2D modeling. In order to achieve it, we do the Principal Component Analysis (PCA) on all the 3D markers for the whole walking sequence. Fig. 6.1 shows the graphical illustration of PCA on one of the 3D gait sequences. Components 1, 2 and 3 are numbered in the order of decreasing component variance. Components 1 and 2 together define the sagittal plane of the walking person. Once the equation of the plane is estimated considering the first two components, all the marker points are then projected to the plane by using the following equation: $u' = u - \langle u, n \rangle n$, where u' is the projected point on the sagittal plane P , u is the original 3D point and n is the unit vector normal to P .

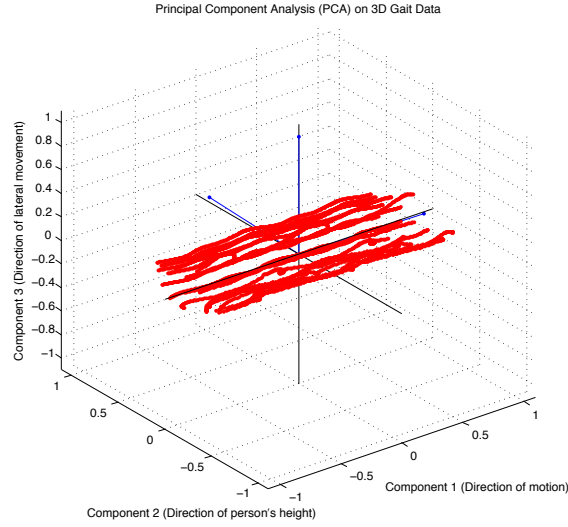


Figure 6.1: Principal Component Analysis (PCA) gives components 1 and 2 and 3 in the order of decreasing variance. Component 1 and 2 together define the sagittal plane of movement. The red trajectories show the progression of various markers on a person's body while walking. The black axes represent the original coordinate system and blue axes represent the new coordinate system.

6.2 Modeling Nonlinear Dynamic Systems

In this section we derive a nonlinear state space system [124] to mathematically model the relationship between a pair of joint angles through differential equations. We use 3D motion capture data with angular positions of the pair of limb segments under consideration to optimize the parameters of the model.

The joint angular positions are each modeled as a two-state nonlinear dynamic model. It has been shown that in applications where the time-varying variables are coupled, the individual velocities can be defined as a nonlinear function of the positions [125]. The function chosen to model the angular velocities in gait is a sum of N^{th} order polynomials of the angular positions as well as

a coupling function that is a product of both angles.

$$\begin{aligned}
\theta_X(k+1) &= f(\theta_X(k), \mathbf{a}_X) + f(\theta_Y(k), \mathbf{b}_X) \\
&\quad + c_X \theta_X(k) \theta_Y(k) \\
\theta_Y(k+1) &= f(\theta_X(k), \mathbf{a}_Y) + f(\theta_Y(k), \mathbf{b}_Y) \\
&\quad + c_Y \theta_X(k) \theta_Y(k)
\end{aligned} \tag{6.1}$$

where $\{X, Y\}$ represents the pair of limbs being modeled. Their angular positions, θ_X and θ_Y , are the states of the system. $f(x, \mathbf{a})$ is an N^{th} order polynomial function of the variable x . $\mathbf{a}$ is an N -dimension vector representing the coefficients of the polynomial. c is the coefficient of the coupling function.

$$f(x, \mathbf{a}) = a_0 + a_1x + a_2x^2 + \dots + a_{N-1}x^{N-1} \tag{6.2}$$

The two equations in Eq. 6.1 represent a coupled nonlinear dynamic system with the state variables representing the motion of the limb segments. The parameter set for limb segment X is represented by $\Phi_X = \{\mathbf{a}_X, \mathbf{b}_X, c_X\}$. We choose a 5^{th} order polynomial for our model ($N = 5$).

Data obtained from the CMU Graphics Lab Motion Capture Database is used to estimate the optimal parameter set by minimizing the least squares error [121] between the angular positions calculated from the model and observed from the 3D data.

$$\Phi_X = \arg \min_{\Phi_X} [\theta_{X,model} - \theta_{X,obs}]^2$$

To further improve the accuracy of the computed angular positions, one gait cycle is divided into two phases based on the gait cycle events related to the limb segments being modeled. Each phase is represented by a different set of parameters ($\Phi_{X,1}, \Phi_{X,2}$) in the model. The onset of each of the phases can be detected using image flow techniques described in Chapter 5 (Fig. 5.5 and Fig. 5.9). The model parameters switch appropriately at the time of transition from one phase to another. The

dynamics of the model is independent of the initial angular positions. It only depends on the correct switching between the different parameter sets.

We now describe the use of the above equations to model the {upperleg, lowerleg}, {foot, lowerleg} and {upperarm, forearm} pairs of limb segments.

6.2.1 Upper leg and lower leg

The upper leg (UL) represents the segment joining the hip to the knee. The lower leg (LL) is the segment connecting the knee to the ankle. The joint angles are measured from the vertical axis. The phases for the non linear dynamic model are decided based on the toe off (TO) and heel strike (HS) events for the same lower limb segment- (i) TO to HS and (ii) HS to TO.

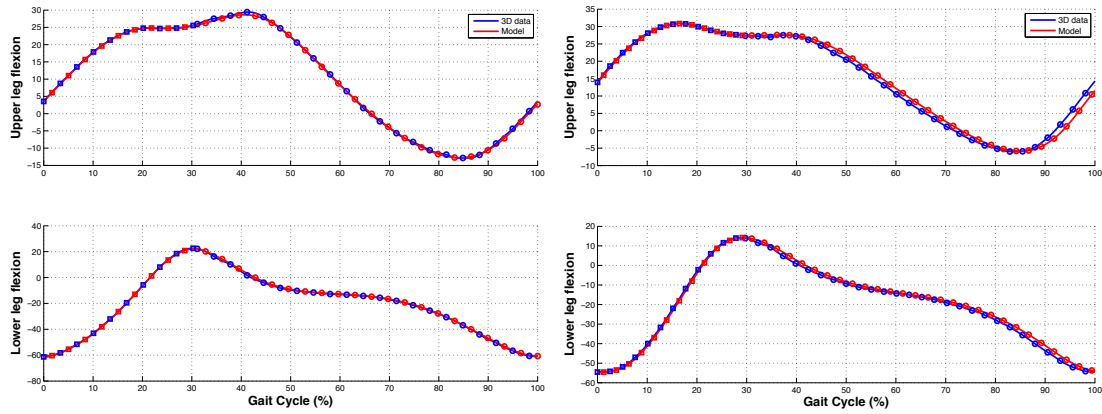


Figure 6.2: Observed (3D data) and computed model for upper leg and lower leg.

Fig. 6.2 shows the comparison of the angular positions recorded from the 3D motion capture system with the angular positions generated from the model for 2 different subjects. The upper leg flexion and lower leg flexion are shown. Phase 1 data points are represented by squares and phase 2 by triangles. The range of motion for each individual varies slightly within an established interval [22, 126]. We develop different models based on the range of motion of the shank, and choose the appropriate parameter set after comparing the results of Sec. 5.2.2.2.4, with the existing database.

6.2.2 Foot and lower leg

The foot (FO) represents the segment joining the ankle to the toe. The phases are again decided on the toe off (TO) and heel strike (HS) events for the same lower limb segment- (i) TO to HS and (ii) HS to TO. Fig. 6.3 shows the comparison of the recorded angular positions with the angular positions generated from the model for 2 different subjects.

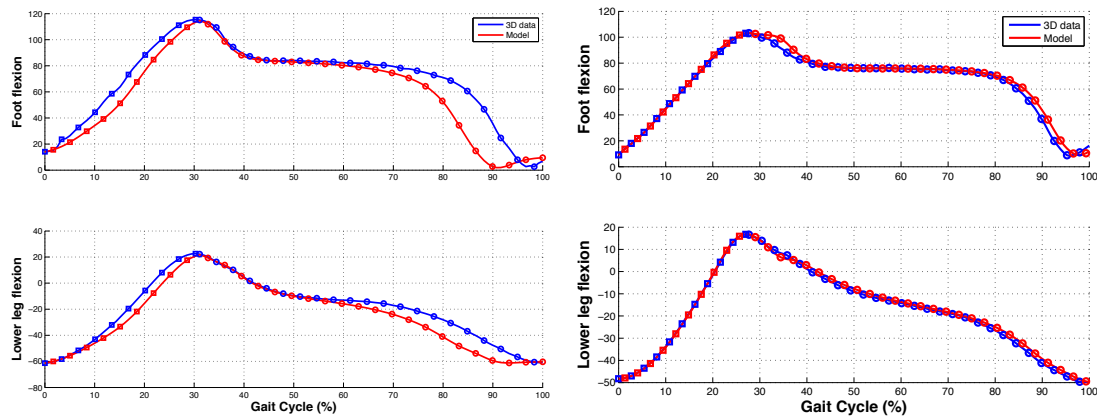


Figure 6.3: Observed (3D data) and computed model for foot and lower leg.

6.2.3 Upper arm and forearm

The upper arm (UA) represents the segment between the shoulder and the elbow. The forearm (FA) connects the elbow and the wrist/hand. The phases for the model between the upper arm and the forearm is based on the consecutive double support (DS) events, we will represent them as DS1, DS2, and... DS_n - (i) DS1 to DS2 and (ii) DS2 to DS3. Fig. 6.4 compares the recorded angular position values with the one estimated by the model.

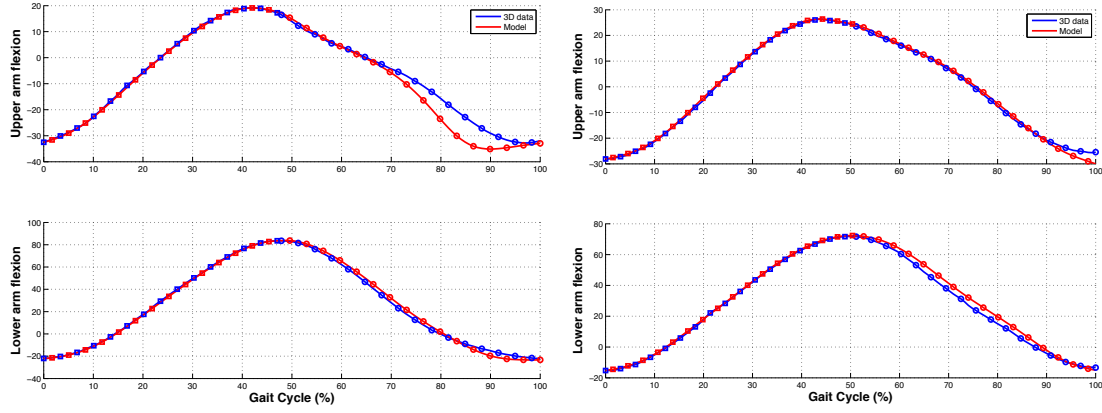


Figure 6.4: Observed (3D data) and computed mode for upper arm and forearm.

6.3 Estimation

The models derived in the above section mathematically define a relationship between a pair of joint angles. We implement an Ensemble Kalman Filter with the derived dynamics to estimate the unobserved joint angles based on observations obtained from image flow analysis. The model can be represented in terms of a state and output equation

$$\begin{aligned} x(k+1) &= g(x(k), \Phi) + w(k) \\ y(k) &= h(x(k)) + v(k) \end{aligned} \tag{6.3}$$

The dynamics of the state vector, $x = [\theta_X, \theta_Y]^T$, is described by Eq. 6.1. The output, y , is the observed variable and the function $h(\cdot)$ depends on how the observation relates to the states. w and v are the process and measurement noises respectively and are modeled by a Gaussian distribution. The unscented transform [117] is applied to generate an ensemble of data points from the *a priori* state estimate, $x^-(k)$. The covariance matrix of this ensemble is represented by $P_{xx}^-(k)$. The ensemble is propagated forward to the next time step according to Eq. 6.3 and the predicted output, y_k^- is calculated. The Kalman gain, *a posteriori* covariance matrix, $P_{xx}^+(k)$ and state estimate, $x^+(k)$

are computed according to the following equations.

$$\begin{aligned}
K(k) &= P_{xy}(k)P_{yy}(k)^{-1} \\
P_{xx}^+(k) &= P_{xx}^-(k) - P_{xy}(k)P_{yy}(k)^{-1}P_{yx}(k) \\
x^+(k) &= x^-(k) + K(k)[y(k) - y^-(k)]
\end{aligned} \tag{6.4}$$

In the following subsections, we discuss the estimation of the specific joint angles and present the results. The advantage of using the ensemble with the unscented transform is that it eliminates the need to linearize the nonlinear dynamic model.

6.3.1 Upper leg

The model used in estimation of upper leg is described in Sec. 6.2.1. The observed variable is the lower leg angular positions, θ_K , computed in Sec. 5.2.2.2.4. The output is, therefore, a linear function of the states and is given as,

$$y(k) = h(x(k)) = \theta_K(k) \tag{6.5}$$

The estimation begins at the toe off event. An identity matrix of the appropriate dimension is used to initialize the covariance of the ensemble. θ_H is initialized at zero degrees and requires 3-4 observations to converge to the true value.

6.3.2 Foot

The model used for estimating the foot angular positions is described in Sec. 6.2.2. Similar to the estimation of upper leg, the observed variable is the lower leg angular positions θ_K , and the output, a linear function of the states, is given by Eq. 6.5. Toe off event marks the start of the estimation process and θ_A is initialized to zero degrees.

Fig. 6.5 shows the estimated angular positions of the lower limbs for a complete gait cycle. It also overlays the results of the manual labeling of the joint angular positions of lower limbs.

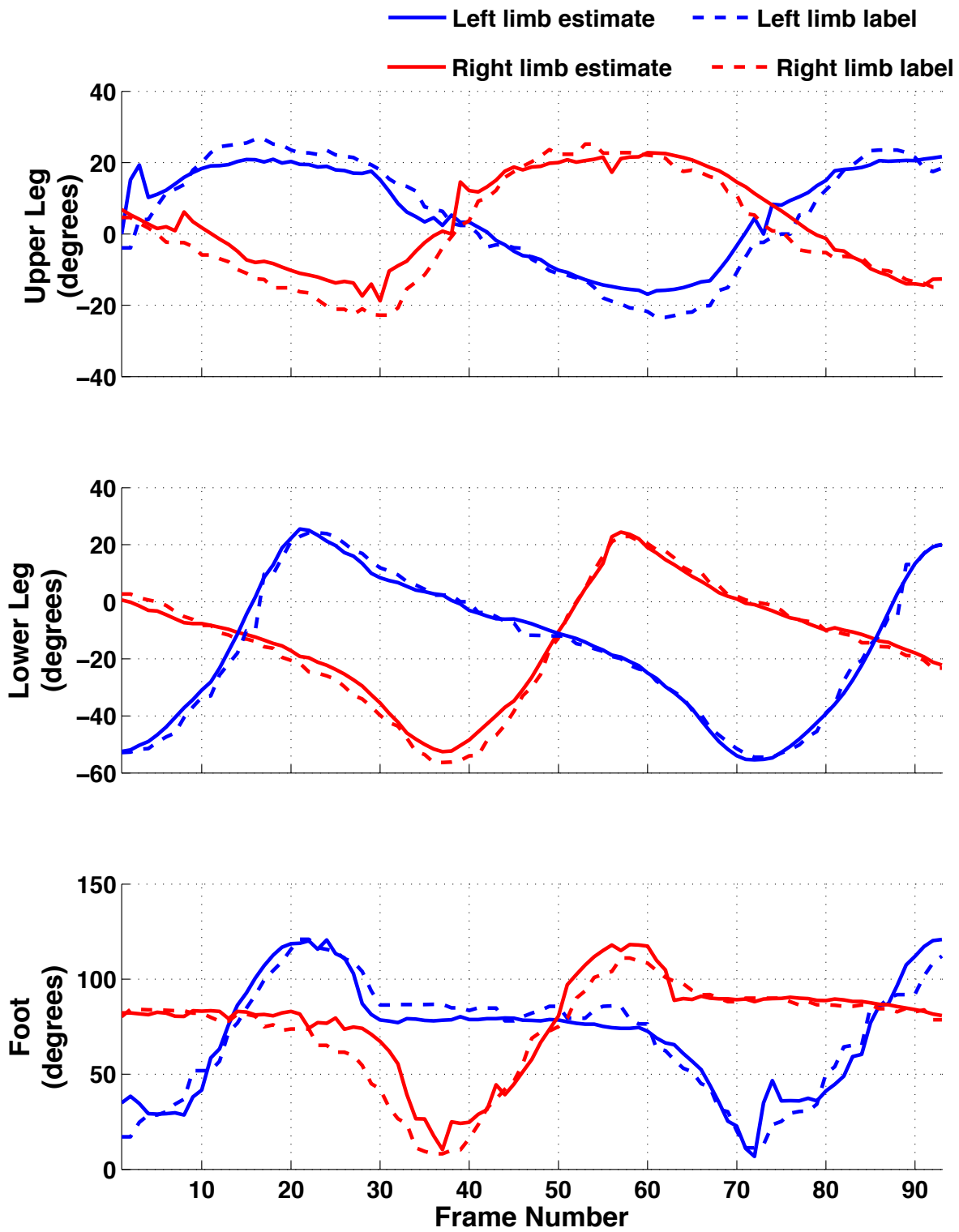


Figure 6.5: Estimation of left (blue) and right (red) upper leg, lower leg and foot angular positions.

6.3.3 Upper arm and Forearm

Sec. 6.2.3 describes the model to be used in the the estimation of upper arm and forearm angular positions. The observed variable is the position (x, y) of the hand/wrist computed in Sec. 5.3. The output function is a non linear function of the states and is given by,

$$x = l_u \sin(\theta_U) + l_f \sin(\theta_F)$$

$$y = l_u \cos(\theta_U) + l_f \cos(\theta_F)$$

where l_u and l_f are the length of upper arm and forearm respectively. θ_U and θ_F represents the upper arm and forearm angular positions with respect to the vertical axis.

Fig. 6.6 shows the estimated angular positions for the upper arm and forearm in a video sequence.

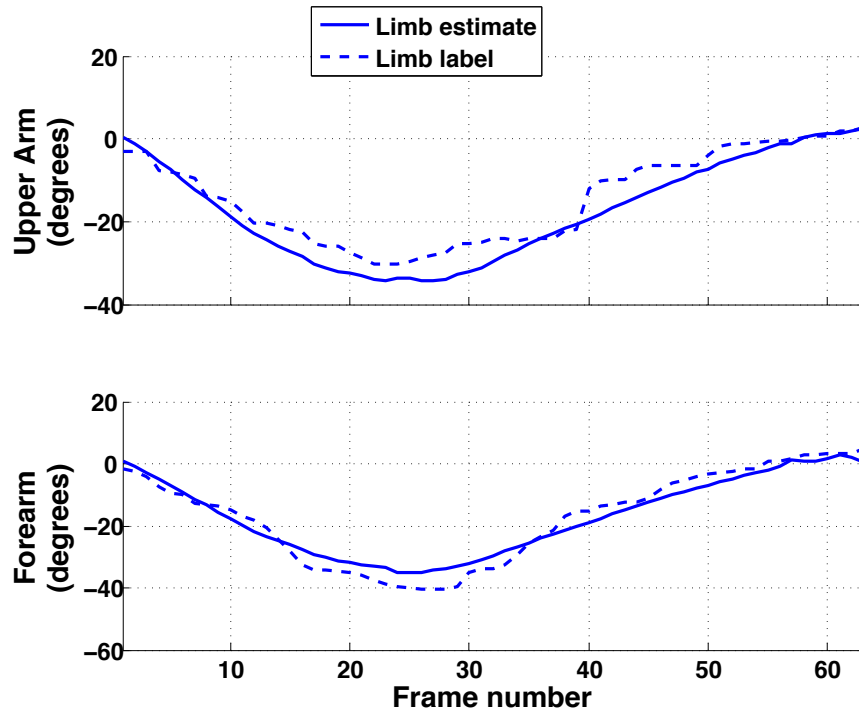


Figure 6.6: Estimation of upper arm and forearm angular positions.

6.4 Angular position estimates

The estimated values of the upper leg flexion and foot flexion from the Kalman filter are corrected using the color distribution models of the upper leg and foot. Color distributions are used as they can provide robustness against non-rigidity, rotation and partial occlusion. We automatically compute the initial color models using $8 \times 8 \times 8$ bin RGB histograms from a double support frame. In each subsequent frame, the color models are computed in the neighborhood of the estimated flexion values and matched against the color models in the previous frame. The flexion value is adjusted corresponding to the values of the best matched model. We have used Bhattacharyya distance to compare two color histograms. The color models are updated in each frame using the following equation:

$$q_t = (1 - \alpha)q_{t-1} + \alpha p_{bm} \quad (6.6)$$

where q_t represents the color model at time t and α weighs the contribution of the color model p_{bm} which gives the best match with q_{t-1} . α is also known as the forgetting factor. We refer the reader to [127] for more details on this method.

6.5 Initialization of body segments

We consider the anthropomorphic measures described in [3] for each body segment characteristics (length and width). Once the bounding box containing the foreground is estimated in each image, we start with a frames of mid swing and double support to aid the initialization process. We initialize the positions of shoulder, elbow, hand, hip, knees and toes. During mid swing, the arms are by the side of the subject and are at almost zero degrees from the vertical axis. We first initialize the shoulder using the body height and width. The shoulder is located approximately $0.818H$ (measured from the foot), where H is the body height. The lengths of upper arm and forearm are approximately $0.188H$ and $0.145H$, which are used to initialize the positions of elbow and hand. The hip point is initialized to $0.530H$. In double support, a subject has both of his/her legs on the ground and are fully extended.

There is minimal overlap and occlusion of leg segments in this pose. We initialize the orientation and position of upper and lower legs using the segmentation process explained in Sec. 5.2.2.2. We fit Radon transforms to the boundaries of the segmented regions to obtain the orientations of lower legs. The orientation of lower and upper legs is same in the double support frame and the positions are initialized using the length of each segment starting from the hip position. The toe orientation is initialized to 90 degrees from the vertical and its position is calculated considering the position of ankle and length of foot segment.

6.6 Experiments

The method was tested on image sequences of 5 different subjects, walking in different environmental conditions (background and speed of walking). Validation of the results is done by superimposing stick figures on original image frames. Fig. 6.7 shows results from a couple of gait sequences.



Figure 6.7: Frames overlaid with our results of motion estimation of lower limb segments and arms.

6.7 Validation

To establish the validity of our method, we have compared our results with manually labeled angular positions of the lower limbs in all the video sequences. Table 6.1 gives the average error (in degrees) in the computation of lower limb angular positions using our method.

Table 6.1: Table listing the average error (in degrees) in the computation of angular positions of different lower limbs

(in degrees °)	Upper Leg	Lower leg	Foot	Average
Left	3.66	2.11	5.76	3.84
Right	4.71	2.45	6.52	4.56
Average	4.18	2.28	6.14	4.20

The average error in the estimation of left upper arm and forearm angular positions was computed as 3.39 and 2.92 degrees respectively.

6.8 Discussion

We have utilized the double support, mid swing, toe off and heel strike phases of a gait cycle and fitted non-linear dynamic models to 3D motion capture gait data defining the relationship between motions of different pair of limb segments. We used Unscented Kalman filter to estimate the upper leg and foot rotation based on the observations of the lower leg rotational velocity. We note that the error in the foot angular position is greater than the lower leg and upper leg angular positions. More distal a segment is, its range of motion is less restricted. The error in the estimation of foot angular positions can also be attributed to the modeling residual error. Upper arm and lower motion is estimated using the observed position of hand. This novel approach, enables accurate computations of the velocities of limb segments. We presented examples showing the overlay of our results on the original frames of the video sequences. We believe this method can be reliably used as an assessment tool to determine individual gait patterns, identify abnormalities in need of treatment

and assess treatment response. In the future, we would like to test our approach with outdoor video sequences and for multiple views. Applications to pathological gait, and to activities in which the sampling is greater than 60 frames/sec, will require additional investigation.

Chapter 7: Gait Recognition

In this chapter we explain our approach to classify people based on the motion information contained in their gait cycle. The feature extraction step in gait recognition consists of model-based and model-free methods. We propose a model free method for gait recognition that utilizes the histograms of normal flow in a gait sequence. We preprocess the images to extract the moving person and divide the region into different horizontal slices after normalizing the person's height and the motion of body parts with respect to his/her torso velocity. Different histograms are computed and normalized for different slices storing the motion information of that region. The cosine similarity measure is used to compare the histograms slice-wise between two different gait sequences. We have used SVMs and dynamic programming for automatic gait recognition on USH gait dataset.

As mentioned in Sec. 2.3, the silhouette based methods suffer from the quality of the foreground and background extraction. We, therefore, focus on the motion information extracted from the gait sequences using normal flow. Our proposed method differs from other optical flow based methods in several ways. 1) Instead of using dense optical flow, we use normal flow. 2) We compute histograms of normal flow for different slices of the body in each frame instead of using the flow distribution of centroids of moving points as in [96] or computing 5 histograms for the entire gait cycle [97]. We normalize the histograms in order to eliminate the effect of walking speed and focus on separating the flow fields of body components instead of using them as a single gait flow image [99]. Our method in part is similar to [128], in which the authors use histograms of oriented flow to classify actions on Weizmann Human Action dataset. Instead of generating a single histogram per frame and performing inter class (action) recognition, we construct multiple histograms in each frame of the gait sequence which enables us to perform intra class recognition. In [128], the authors have used the generalized Binet-Cauchy kernels for nonlinear dynamical systems for the classification of actions, whereas we propose to use dynamic programming to find the minimum cost path in the similarity matrix and use it for gait recognition. The advantage of our method over other methods is

its ability to compare the motion information on frame-by-frame basis rather than representing the entire gait sequence with a single flow image and then performing the comparison.

7.1 Foreground Extraction

We extract the foreground from the image sequence using background subtraction and remove small components by doing connected component analysis. Normal flow is computed for the foreground by the technique described in [27]. Few examples of computed normal flow for the foreground are shown in Fig. 7.1.

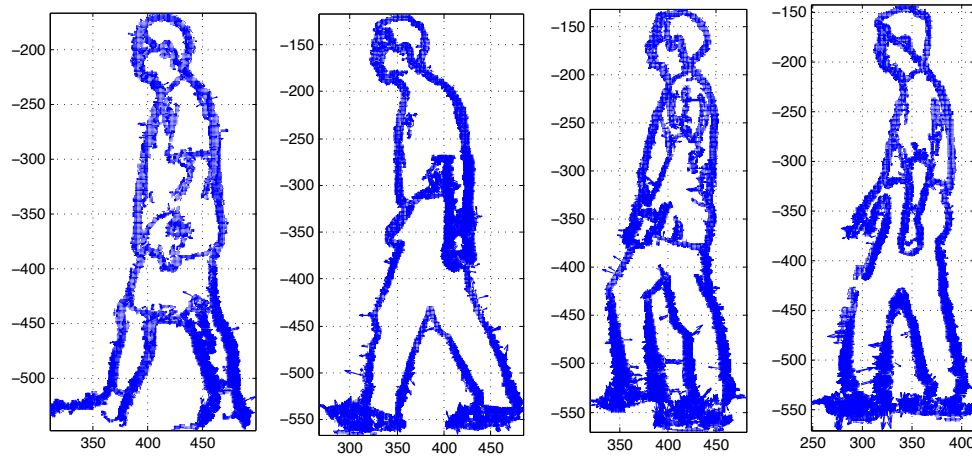


Figure 7.1: Normal Flow.

7.2 Normalizing motion with respect to torso motion and body height

We aim to stabilize/normalize the motion of a person's body with respect to its torso motion. It is known that the height of a human torso approximately corresponds to $0.52H$, where H is the body height [3]. The instantaneous head and torso motion comprises forward translation of the body, t_x , and up/down movement/excursion, t_y . We aim to find this motion, $M_t = (t_x, t_y, 0)$, by "voting" over a range of possible motion values using image flow vectors. We use the top 30% of the whole

body for voting to minimize the effect of arm motion. The flow vectors are binned according to the orientation and magnitude of the flow. The flow direction ranges from 0 to 180 degrees in steps of 1 degree, giving 180 possible orientations. For each orientation, the magnitude ranges from -20 to 20 pixels in steps of 0.1 pixel. This gives a total of 180×400 bins which are used for voting. We count the number of flow vectors in agreement with the estimated motion corresponding to each bin. The bin having the maximum number of flow vectors gives the instantaneous translational velocity of the torso. All the flow vectors are then normalized with respect to the computed torso motion. The resulting image is normalized to compensate for height and is scaled to 250 pixels. The motion of the upper part of the body (torso) is nearly zero in all cases after normalization.

7.3 Histograms of Normal Flow

The entire body is divided into overlapping horizontal slices (size 50 pixels). Fig. 7.2 shows the numbered slice regions. For each slice, we compute n -bin histogram of normal flow, i.e. each flow vector is binned according to its angle with the $+x$ (horizontal) direction. We compute the median value of all normal flow vectors in a bin and use it to represent the value associated with that particular bin. Each bin is weighted by the count of flow vectors in that bin. Thus, all normal flow vectors with direction, $\theta = \tan^{-1}(\frac{dy}{dx})$ in the range,

$$2\pi \frac{(b-1)}{n} \leq \theta < 2\pi \frac{b}{n} \quad (7.1)$$

will contribute by n_f , where n_f is the median normal flow value for bin b , $1 \leq b \leq n$, out of total of n bins. Finally, the histogram is normalized to sum up to 1. Few examples of horizontal slices from the middle and lower part of the body and their corresponding histograms are shown in Fig. 7.3.

Normalization makes the histogram representation scale-invariant. That means, whether a person is walking at a distance away from the camera (small number of pixels) or closer to the camera (large number of pixels), we expect to observe the same histogram. Selecting the median value of normal flow assures that the noisy flow values get rejected as outliers in the process. The parameter,

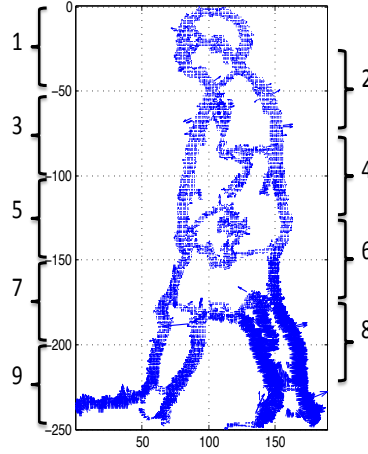


Figure 7.2: Numbered regions obtained by taking horizontal slices from the subject's foreground.

n (number of bins) can be selected depending on the datasets used. In our tests, we were able to get good recognition results with $n = 16$.

7.4 Comparing two sequences

We determine the similarity, D , between two image frames, A_i and B_j by computing the cosine similarity between their respective normal flow histograms, $H(A_i)$ and $H(B_j)$. For k^{th} slice, the similarity between two frames is given by,

$$D_k(A_i, B_j) = 1 - \frac{H_k(A_i) \cdot H_k(B_j)}{\|H_k(A_i)\| \|H_k(B_j)\|} \quad (7.2)$$

We sum the similarity values over all the slices to find the overall similarity $D(A_i, B_j) = \sum_{k=1}^9 D_k(A_i, B_j)$. To compare two sequences, we apply the similarity measure D to all possible pairs of frames, resulting in a similarity matrix $SM(i, j) = D(A_i, B_j)$ where $A_1, A_2, \dots, A_m$ and $B_1, B_2, \dots, B_n$ are the frames of sequences A and B , respectively. Few examples of SMs are shown in Fig. 7.4. In the figure, dark blue regions represent high similarity, and yellow and red regions represent low similarity. In Fig. 7.6a, there are several prominent diagonals, indicating the sequence

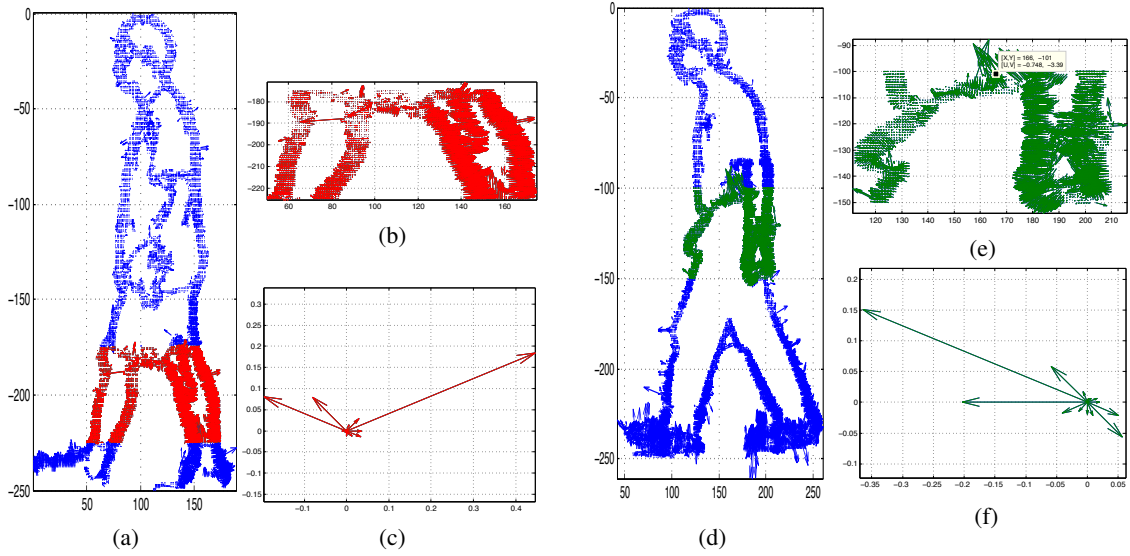


Figure 7.3: Histograms taken from different layers/slices of the body.

of frames where the similarity is high. This sequence of frames is taken from the same person and SM is found for 100 frames. The diagonals represent the gait cycles where the similarity between the corresponding frames is high. In Fig. 7.6b, the SM is computed between the walking sequences of two different persons. The patterns of similar and/or matching gait cycles are clearly missing in this matrix.

7.5 Gait recognition using similarity matrix

We tested two different approaches for identifying people using the similarity matrix based on the histograms of normal flow. In the first approach, we train a linear Support Vector Machine (SVM) by giving the similarity matrices as features. Given a similarity matrix $SM(i, j)$ between a person P_i and P_j , the SVM predicts whether or not person P_i and P_j have similar patterns of motions in their walking sequences. In the second approach we align the gait cycles of two persons and compute the similarity matrix between their gait sequences. Note that this matrix may not be square depending on the walking speed/cadence of the subjects under consideration. We use dynamic programming to find the minimum cost path from the start to the end of the cycle and use it to predict the similarity

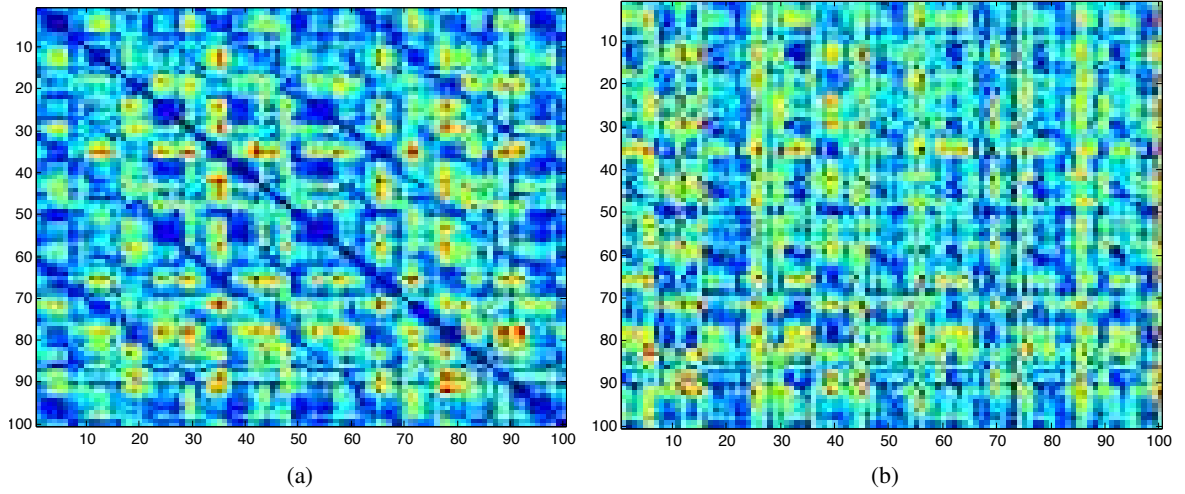


Figure 7.4: Similarity Matrices (SM) in two different cases computed between the frame sequences of the: (a) same person (b) different persons.

between sequences.

7.5.1 Gait recognition using SVM

We use linear SVM to predict whether the pair (P_i, P_j) are similar based on their walking patterns. We use the first 40 frames of a sequence and a moving window of size 5 on the next 80 frames to compute the similarity matrix and input these instances as positive examples to the SVM. The negative examples are generated by comparing the first 40 frames from any two different sequences. We then test the new frames from each subject's gait sequence against all the subjects in the database and this is done by taking each person as a test instance. We create the test cases by randomly choosing the gait frames that have not been used for training instances.

7.5.2 Gait recognition using Dynamic Programming and NN

In this approach, we use the alignment of gait cycles of different people in the database. We use the vertical component of torso motion values computed in Sec. 7.2 to identify the gait cycles. The double-support and mid-swing events occur at the zero value of the instantaneous velocity, i.e.

the minimum excursion occurs when the velocity changes sign from negative to positive, and the maximum excursion occurs as the velocity changes sign from positive to negative. An example of the vertical (up/down) instantaneous velocity is shown in Fig. 7.5.

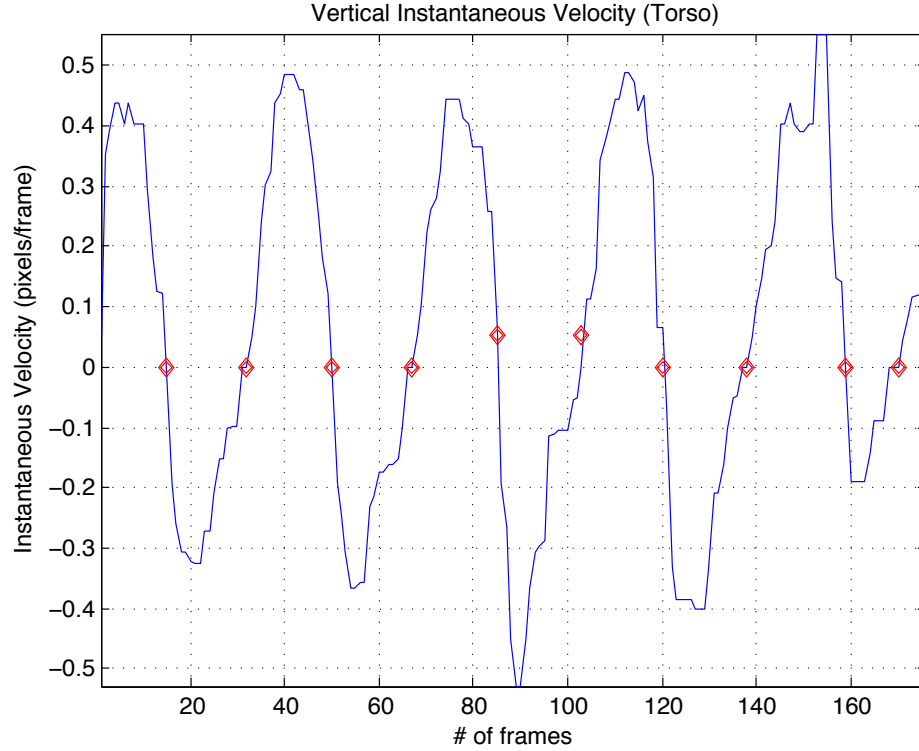


Figure 7.5: Vertical instantaneous velocity of a walking person.

We extract the gait cycles for each person and use them to align the gait sequences and compute similarity matrix between all pairs of frames. We propose that the total cost of the path in the similarity matrix from the start to the end of the gait cycle will be smaller for matching gait cycles (i.e. different gait cycles from the same person) compared to dissimilar gait cycles (i.e. gait cycles from different persons). We use dynamic programming to find the minimum cost path in the similarity matrix. There could be two possible scenarios: (1) The two cycles coincide perfectly, i.e., the left double-support matches with the other cycle's left double-support. (2) The left double-support of one cycle coincides with the right double-support of the second cycle. In that case, the cycles would be shifted by approximately half the length of the gait cycle. We have shown both cases in Fig. 7.6.

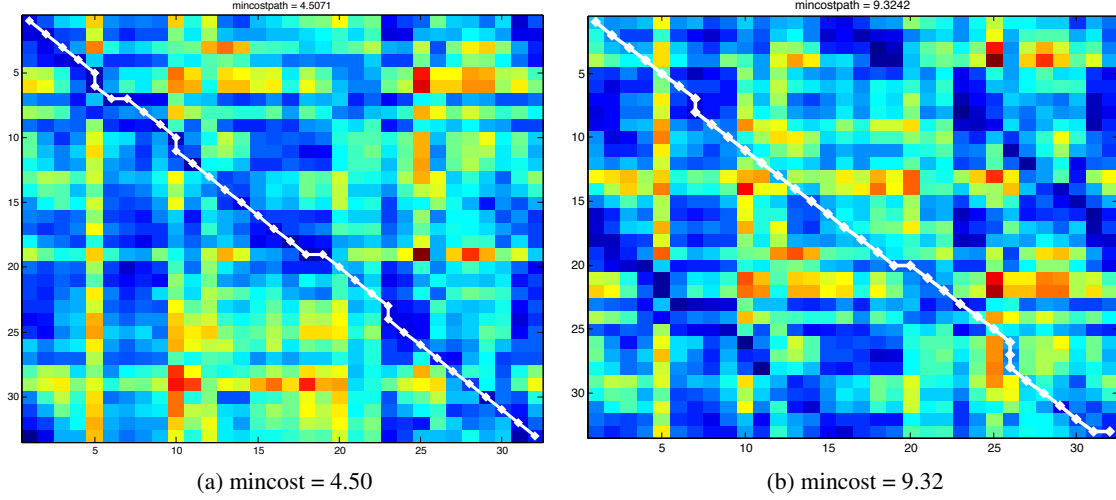


Figure 7.6: Similarity Matrices (SM) in two different cases: (a) perfectly aligned gait cycles (b) gait cycles shifted by half. The white lines indicate the path taken to obtain the minimum cost.

Here the similarity matrices are computed from the same person's different gait cycles. To handle this situation, we choose the smaller value of minimum cost path (mcp) in the two possible cases. For the example shown in Fig. 7.6, we choose the scenario which gives the minimum cost path as 4.50. For matching a person P_i , we match the gait cycle of P_i to P_j , where $j = 1, 2, \dots, n$, n is the size of the database and compute all similarity matrices. The matching subject's index, m , is computed using

$$m = \arg \min_j \{mcp(SM(i, j))\} \quad (7.3)$$

7.6 Experiments

We demonstrate our proposed method of gait identification using the University of Southampton gait database [28]. The sequences were collected indoors. Subjects were walking on a treadmill; some examples are shown in Fig. 7.7. The advantage of treadmill-based data is collection of arbitrarily long sequences in a controlled environment. We process 114 sequences. From each sequence, the first 400 frames are discarded, permitting individuals to settle into their normal gait motion. The next 200 frames are used in this experiment. All the frames are normalized with respect to person's



Figure 7.7: Examples of frames from USH gait database.

height and the motion is computed with respect to the torso motion.

7.6.1 Recognition using SVM

We compute the specificity and sensitivity measures from the results of SVM where they are defined as:

$$specificity = \frac{TN}{FP + TN}; \quad sensitivity = \frac{TP}{TP + FN} \quad (7.4)$$

where, TN , FP , TP , FN are true negatives, false positives, true positives and false negatives respectively. The computed values are given in Table 7.1.

While these values demonstrate promising results, we seek a robust way to improve the recognition rate. SVM does not take into account the gait cycle information of subjects. Studies have shown

that using the gait phases have improved the accuracy of recognition [129]. In the next section, we present the results of recognition obtained after aligning the gait sequences.

Table 7.1: Specificity and sensitivity measures for 114 subjects in the USH gait database.

Specificity	Sensitivity
0.93	0.97

7.6.2 Recognition using Dynamic Programming and NN

The first 3 gait cycles are used for training and the 4th gait cycle is used as a testing instance. We use leave one out cross validation technique. For each test gait cycle, it is pairwise compared against all the gait cycles of the sequences in the database and the minimum cost path for each pair is computed. The top k potential candidates for match P_i are chosen based on the value of minimum cost path from P_i . Since in our case $k = 1$, we choose the person with lowest minimum cost path. We get 100% recognition results using our method.

7.6.3 Effect and Contribution of Slices on Recognition

In order to understand which slices are more distinguishing than the others, we perform analysis of variance (ANOVA) on each slice extracted for each sequence. ANOVA F-statistic gives an estimate of the discriminatory capability of the slices and a higher value indicates high inter-class variance but low intra-class variance between subjects. We have assumed for simplicity that the slices are independent of each other. The analysis indicates that the motion in the slices containing the thighs/upper legs and arms (in accordance with [26]) have higher discriminating power compared to the motion in the region of the foot and head. Note that in our case, since the motion of the whole body is normalized with respect to the torso motion, the contribution of head region is not expected to have large discriminative power. The analysis also showed that a minimum of 3 slices were required to achieve 100% recognition rate.

Table 7.2: ANOVA: The respective F-measures in the increasing order of importance.

Slice number	F-statistic
6	257.54
4	249.41
5	223.38
3	185.38
7	177.91
8	153.58
2	141.39
1	99.1
9	82.96

7.7 Discussion

A novel model free approach to individualized gait analysis based on the normal flow fields computed over a gait sequence has been proposed. We construct the normalized motion histograms of overlapping horizontal slices of the subject’s body containing different body parts. These features are scale-invariant as well as independent of the speed of walking. The similarity matrix is computed using the cosine measure between the respective histograms of two gait sequences. We also explore the importance of gait cycle information while performing gait recognition. We have applied our method to the USH gait database and have achieved good results using SVM and dynamic programming. In the future, we would like to test our method on different gait databases with different gait covariates such as different carrying conditions, shoes and clothing effects.

Chapter 8: Conclusions and Future Directions

In this chapter, we present the conclusions of our presented work. We also discuss the potential application areas where our method could be useful. In Sec. 8.3, we discuss the future directions that we would like to explore.

8.1 Conclusions

We identify various events in a gait cycle reliably using image motion and estimate segmental movement by modeling the synergies between various body parts. We demonstrate the use of vertical component of head/torso excursion in finding the double support and mid swing events in a gait cycle. The vertical excursion of torso is the outcome of the gait determinants which are active at different phases to smooth the walking progression. Hence, it provides crucial and reliable information about the phases of gait. We also estimate the lower leg rotational velocities in the sagittal plane. We use these rotational velocities to determine the toe off and heel strike events of the gait cycle. The toe off represents the start of the swing phase and heel strike represents the start of the stance phase. These events are important to study the time a person is spending in different phases of the gait cycle. Our motion estimates are robust to imperfect background subtraction. We validated our approach by comparing it against the data captured simultaneously using a 3D motion capture system.

We modeled the relationship between different segments with non linear dynamic models using the information about different events of gait cycle. This approach is particularly helpful when limbs are occluding one another and the motion is difficult to estimate. We demonstrated the use of these models to predict the unobserved motion of the limb segments using the observed motion of the related limb segments. This approach works well when the model created to capture the relationship between two segments is mathematically accurate. As can be noted from our presented results, the

estimation of the foot segment angular position is not as accurate as the estimation of the angular positions of other segments. The possible reason could be that the ankle has six degrees of freedom in three dimensions. This gives rise to complex kinematics which makes it more difficult to model them using a dynamic system. As we move vertically up starting from the foot, the body segments are less restricted and have more degrees of freedom. This is illustrated as we go from ankle, knee, hip to shoulders. The shoulders can be unrestricted during general movements but during gait they follow *Bilateral Symmetry* and hence are accurately modeled.

We used the histograms of normal flow to represent the motion patterns of different slices of the human body and used them as features for comparison. Dynamic programming was used to find the minimum cost path in the similarity matrix between two gait sequences in order to determine the overall similarity between two subjects. We demonstrated our method on a public dataset. We presented the analysis that the arms and upper leg motions provide higher discriminating power than other body parts. Our method performed better when gait cycle information was used while comparing the gait sequences. This provides an insight into incorporating the gait cycle information whenever available to improve the matching process.

8.2 Applications

Our approach of markerless gait analysis can be helpful in assisting the clinicians to analyze gait, and monitor elderly and/or disabled people in their homes for any changes in their gait patterns. If the reference/normal gait pattern is available for a person, all the subsequent gait patterns can be compared against it to determine if the person needs medical attention or is at a risk of falling. Our method can also be used to establish the normal gait pattern of an individual. If the ranges of the segmental movement are outside the normal bounds, it may be an indication of abnormal gait pattern. Recent literature shows that there has been a lot of effort in using gait as a soft biometric. For example, it can be used for providing security at public places by detecting any suspicious activity. Our analysis on the gait dataset also demonstrates that the gait patterns of an individual are significantly unique compared to other people. However, our method was tested with a total number

of 114 subjects and this application area would require a larger sample of subjects to establish our hypothesis. Our representation of gait motion patterns can also be used as an assessment tool to determine the gait symmetry by comparing the first half of the cycle with the second half.

8.3 Future Directions

We have dealt with the self occlusion of the limb segments during gait. However, our method can not compensate if there is partial occlusion due to the presence of any object. It would be interesting to explore whether different types of occlusions can be modeled. Our analyses are mostly in the sagittal plane, we would like to extend it to different view points. The identification of double support and mid swing events would be straightforward (as the 'voting' technique would be valid for any viewpoint as long as the camera and the person are on the same ground plane). However, the different viewplanes would have different amounts of occlusions for lower legs and hands thereby imposing additional limitations on their visibility.

As we have mentioned the model to explain the foot kinematics was not very accurate. More rigorous analysis would be required to design a complex non linear dynamic model explaining the relationship between the lower leg and the foot. Another possible extension would be to use additional 3D data for modeling in order to ensure that most of the normal range for the segmental movement is covered by the individuals in the database. Higher frame rate and better resolution would provide the benefit of more accurate image motion analysis. We have demonstrated that we can recognize people from their gait patterns. However, we would like to explore the validity and robustness of our method with different gait covariates for example, different clothes and carrying conditions (backpack and/or bag). It would also be interesting to explore whether the duration of the stance and swing phase (expressed as a percentage of the gait cycle) can provide a discriminative set of features to cluster groups of people or to classify gender and/or ethnicity.

Bibliography

Bibliography

- [1] M. Nordin and V. Frankel, “Biomechanics of the knee,” *Basic Biomechanics of the Musculoskeletal system*, pp. 115–134, 1989.
- [2] D. A. Neumann, *Kinesiology of the musculoskeletal system: foundations for rehabilitation*. Elsevier Health Sciences, 2013.
- [3] D. A. Winter, *Biomechanics and motor control of human movement*. John Wiley & Sons, 2009.
- [4] M. P. Murray, “Gait as a total pattern of movement: Including a bibliography on gait,” *American Journal of Physical Medicine & Rehabilitation*, vol. 46, no. 1, pp. 290–333, 1967.
- [5] J. Perry and J. M. Burnfield, *Gait analysis: normal and pathological function*. Slack, 1993.
- [6] M. P. Murray, A. B. Drought, and R. C. Kory, “Walking patterns of normal men,” *The Journal of Bone & Joint Surgery*, vol. 46, no. 2, pp. 335–360, 1964.
- [7] G. Sammarco, “Biomechanics of the foot,” *Basic biomechanics of the skeletal system*, pp. 193–220, 1980.
- [8] M. S. Nixon, T. Tan, and R. Chellappa, *Human identification based on gait*. Springer, 2010, vol. 4.
- [9] E. E. Stone and M. Skubic, “Evaluation of an inexpensive depth camera for passive in-home fall risk assessment,” in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference On*. IEEE, 2011, pp. 71–77.
- [10] J.-H. Yoo and M. S. Nixon, “Markerless human gait analysis via image sequences,” 2003.
- [11] J. Courtney and A. M. De Paor, “A monocular marker-free gait measurement system,” *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 18, no. 4, pp. 453–460, 2010.
- [12] S. X. Ju, M. J. Black, and Y. Yacoob, “Cardboard people: A parameterized model of articulated image motion,” in *Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on*. IEEE, 1996, pp. 38–44.
- [13] I. A. Kakadiaris and D. Metaxas, “Model-based estimation of 3d human motion with occlusion based on active multi-viewpoint selection,” in *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR’96, 1996 IEEE Computer Society Conference on*. IEEE, 1996, pp. 81–87.

- [14] Y. Yacob and L. Davis, "Learned temporal models of image motion," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 446–453.
- [15] D. Meyer, J. Denzler, and H. Niemann, "Model based extraction of articulated objects in image sequences for gait analysis," in *Image Processing, 1997. Proceedings., International Conference on*, vol. 3. IEEE, 1997, pp. 78–81.
- [16] K. Rohr, "Towards model-based recognition of human movements in image sequences," *CVGIP: Image understanding*, vol. 59, no. 1, pp. 94–115, 1994.
- [17] R. Urtasun, D. J. Fleet, and P. Fua, "3d people tracking with gaussian process dynamical models," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 1. IEEE, 2006, pp. 238–245.
- [18] M. Isard and A. Blake, "Condensation conditional density propagation for visual tracking," *International journal of computer vision*, vol. 29, no. 1, pp. 5–28, 1998.
- [19] H. Sidenbladh, M. J. Black, and D. J. Fleet, "Stochastic tracking of 3d human figures using 2d image motion," in *Computer Vision ECCV 2000*. Springer, 2000, pp. 702–718.
- [20] V. T. Inman, H. J. Ralston, and F. Todd, *Human walking*. Williams & Wilkins, 1981.
- [21] G. Johansson, "Visual perception of biological motion and a model for its analysis," *Perception & psychophysics*, vol. 14, no. 2, pp. 201–211, 1973.
- [22] M. P. Murray, R. C. Kory, and S. B. Sepic, "Walking patterns of normal women." *Archives of physical medicine and Rehabilitation*, vol. 51, no. 11, p. 637, 1970.
- [23] J. Perry, J. R. Davids *et al.*, "Gait analysis: normal and pathological function," *Journal of Pediatric Orthopaedics*, vol. 12, no. 6, p. 815, 1992.
- [24] L. Lee and W. E. L. Grimson, "Gait analysis for recognition and classification," in *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*. IEEE, 2002, pp. 148–155.
- [25] S. Yu, T. Tan, K. Huang, K. Jia, and X. Wu, "A study on gait-based gender classification," *Image Processing, IEEE Transactions on*, vol. 18, no. 8, pp. 1905–1910, 2009.
- [26] X. Li, S. J. Maybank, S. Yan, D. Tao, and D. Xu, "Gait components and their application to gender recognition," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 38, no. 2, pp. 145–155, 2008.
- [27] H. Wechsler, Z. Duric, F. Li, and V. Cherkassky, "Motion estimation using statistical learning theory," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 4, pp. 466–478, 2004.
- [28] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter, "On a large sequence-based human gait database," in *Applications and Science in Soft Computing*. Springer, 2004, pp. 339–346.
- [29] R. Drillis, R. Contini, and M. Maurice Bluestein, "Body segment parameters1," 1966.

- [30] M. Nordin and V. Frankel, "Biomechanics of the hip," *Basic biomechanics of the musculoskeletal system*. Philadelphia: Lea and Febiger Co, pp. 135–51, 1989.
- [31] J. Perry, L. Norwood, and K. House, "Knee posture and biceps and semimembranosus muscle action in running and cutting (an emg study)," *Trans. Orthop. Res. Soc.*, vol. 2:258, 1977.
- [32] V. H. Frankel and M. Nordin, "Biomechanics of the ankle," *Basic biomechanics of the musculoskeletal system*, 7th edn. Lea & Febiger, Philadelphia, pp. 153–161, 1989.
- [33] R. N. STAUFFER, E. Y. CHAO, and R. C. BREWSTER, "Force and motion analysis of the normal, diseased, and prosthetic ankle joint," *Clinical orthopaedics and related research*, vol. 127, pp. 189–196, 1977.
- [34] C. Cedras and M. Shah, "Motion-based recognition a survey," *Image and Vision Computing*, vol. 13, no. 2, pp. 129–155, 1995.
- [35] D. M. Gavrilu, "The visual analysis of human movement: A survey," *Computer vision and image understanding*, vol. 73, no. 1, pp. 82–98, 1999.
- [36] J. K. Aggarwal and Q. Cai, "Human motion analysis: A review," in *Nonrigid and Articulated Motion Workshop, 1997. Proceedings., IEEE*. IEEE, 1997, pp. 90–102.
- [37] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231–268, 2001.
- [38] D. Forsyth, O. Arikan, L. Ikemoto, J. O'Brien, and D. Ramanan, "Computational studies of human motion i: Tracking and animation," *Foundations and Trends in Computer Graphics and Vision*, vol. 1, no. 2/3, pp. 1–255, 2006.
- [39] S. Jabri, Z. Duric, H. Wechsler, and A. Rosenfeld, "Detection and location of people in video images using adaptive fusion of color and edge information," in *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, vol. 4. IEEE, 2000, pp. 627–630.
- [40] I. Haritaoglu, D. Harwood, and L. S. Davis, "Ghost: A human body part labeling system using silhouettes," in *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, vol. 1. IEEE, 1998, pp. 77–82.
- [41] D. M. Gavrilu and L. S. Davis, "3-d model-based tracking of humans in action: A multi-view approach," in *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on*. IEEE, 1996, pp. 73–80.
- [42] I. Haritaoglu, D. Harwood, and L. S. Davis, "W 4 s: A real-time system for detecting and tracking people in 2 1/2d," in *Computer VisionECCV'98*. Springer, 1998, pp. 877–892.
- [43] C. R. Wren, A. Azarbayejani, T. Darrell, and A. P. Pentland, "Pfunder: Real-time tracking of the human body," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 780–785, 1997.
- [44] S. Ioffe and D. Forsyth, "Finding people by sampling," in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2. IEEE, 1999, pp. 1092–1097.

- [45] I. A. Kakadiaris and D. Metaxas, “3d human body model acquisition from multiple views,” in *Computer Vision, 1995. Proceedings., Fifth International Conference on.* IEEE, 1995, pp. 618–623.
- [46] R. Polana and R. C. Nelson, “Detection and recognition of periodic, nonrigid motion,” *International Journal of Computer Vision*, vol. 23, no. 3, pp. 261–282, 1997.
- [47] J. M. Rehg and T. Kanade, “Model-based tracking of self-occluding articulated objects,” in *Computer Vision, 1995. Proceedings., Fifth International Conference on.* IEEE, 1995, pp. 612–617.
- [48] S. Wachter and H.-H. Nagel, “Tracking of persons in monocular image sequences,” in *Non-rigid and Articulated Motion Workshop, 1997. Proceedings., IEEE.* IEEE, 1997, pp. 2–9.
- [49] Y. Song, L. Goncalves, E. Di Bernardo, and P. Perona, “Monocular perception of biological motion-detection and labeling,” in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 2. IEEE, 1999, pp. 805–812.
- [50] D. M. Gavrila and V. Philomin, “Real-time object detection for smart vehicles,” in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 1. IEEE, 1999, pp. 87–93.
- [51] H.-J. Lee and Z. Chen, “Determination of 3d human body postures from a single view,” *Computer Vision, Graphics, and Image Processing*, vol. 30, no. 2, pp. 148–168, 1985.
- [52] M. K. Leung and Y.-H. Yang, “Human body motion segmentation in a complex scene,” *Pattern recognition*, vol. 20, no. 1, pp. 55–64, 1987.
- [53] M. K. Leung and Y.-H. Yang, “First sight: A human body outline labeling system,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, no. 4, pp. 359–377, 1995.
- [54] M. Oren, C. Papageorgiou, P. Sinha, E. Osuna, and T. Poggio, “Pedestrian detection using wavelet templates,” in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on.* IEEE, 1997, pp. 193–199.
- [55] V. Philomin, R. Duraiswami, and L. Davis, “Quasi-random sampling for condensation,” *Computer Vision/ECCV 2000*, pp. 134–149, 2000.
- [56] J. Rittscher and A. Blake, “Classification of human body motion,” in *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, vol. 1. IEEE, 1999, pp. 634–639.
- [57] P. F. Felzenszwalb and D. P. Huttenlocher, “Efficient matching of pictorial structures,” in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 2. IEEE, 2000, pp. 66–73.
- [58] J. Deutscher, A. Blake, and I. Reid, “Articulated body motion capture by annealed particle filtering,” in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 2. IEEE, 2000, pp. 126–133.

- [59] D. Ramanan and D. A. Forsyth, "Finding and tracking people from the bottom up," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, vol. 2. IEEE, 2003, pp. II–467.
- [60] K. Mikolajczyk, C. Schmid, and A. Zisserman, "Human detection based on a probabilistic assembly of robust part detectors," in *Computer Vision-ECCV 2004*. Springer, 2004, pp. 69–82.
- [61] C. Bregler and J. Malik, "Tracking people with twists and exponential maps," in *Computer Vision and Pattern Recognition, 1998. Proceedings. 1998 IEEE Computer Society Conference on*. IEEE, 1998, pp. 8–15.
- [62] T.-J. Cham and J. M. Rehg, "A multiple hypothesis approach to figure tracking," in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on*., vol. 2. IEEE, 1999.
- [63] N. I. Badler and S. W. Smoliar, "Digital representations of human movement," *ACM Computing Surveys (CSUR)*, vol. 11, no. 1, pp. 19–38, 1979.
- [64] C. Bregler, "Learning and recognizing human dynamics in video sequences," in *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*. IEEE, 1997, pp. 568–574.
- [65] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 3, pp. 257–267, 2001.
- [66] K. Gould and M. Shah, "The trajectory primal sketch: A multi-scale scheme for representing motion characteristics," in *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR'89., IEEE Computer Society Conference on*. IEEE, 1989, pp. 79–85.
- [67] S. Gutta, I. F. Imam, and H. Wechsler, "Hand gesture recognition using ensembles of radial basis function (rbf) networks and decision trees," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, no. 06, pp. 845–872, 1997.
- [68] M. Herman, "Computer interpretation of human stick figures." in *AAAI*, 1980, pp. 174–177.
- [69] I. A. Essa and A. P. Pentland, "Facial expression recognition using a dynamic model and motion energy," in *Computer Vision, 1995. Proceedings., Fifth International Conference on*. IEEE, 1995, pp. 360–367.
- [70] I. A. Essa and A. P. Pentland, "Coding, analysis, interpretation, and recognition of facial expressions," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 7, pp. 757–763, 1997.
- [71] G. Johansson, "Spatio-temporal differentiation and integration in visual motion perception," *Psychological research*, vol. 38, no. 4, pp. 379–393, 1976.
- [72] R. F. Rashid, "Towards a system for the interpretation of moving light displays," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, no. 6, pp. 574–581, 1980.

- [73] T. Starner and A. Pentland, "Visual recognition of american sign language using hidden markov models."
- [74] Y. Yacoob and L. S. Davis, "Recognizing human facial expressions from long image sequences using optical flow," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 18, no. 6, pp. 636–642, 1996.
- [75] M.-H. Yang and N. Ahuja, "Recognizing hand gestures using motion trajectories," in *Face Detection and Gesture Recognition for Human-Computer Interaction*. Springer, 2001, pp. 53–81.
- [76] L. Bogoni, "Functional features for chopping extracted from observations and interactions," *Image and vision computing*, vol. 16, no. 11, pp. 765–783, 1998.
- [77] Z. Duric, A. Rosenfeld, and L. S. Davis, "Egomotion analysis based on the frenet-serret motion model," in *International Journal of Computer Vision*. Citeseer, 1995.
- [78] Z. Duric, E. Rivlin, and A. Rosenfeld, "Understanding the motions of tools and vehicles," in *Computer Vision, 1998. Sixth International Conference on*. IEEE, 1998, pp. 925–932.
- [79] Z. Duric, E. Rivlin, and A. Rosenfeld, "Understanding object motion," *Image and vision computing*, vol. 16, no. 11, pp. 785–797, 1998.
- [80] K. Green, D. Eggert, L. Stark, and K. Bowyer, "Generic recognition of articulated objects through reasoning about potential function," *Computer vision and Image understanding*, vol. 62, no. 2, pp. 177–193, 1995.
- [81] K. Akita, "Image sequence analysis of real world human motion," *Pattern Recognition*, vol. 17, no. 1, pp. 73–83, 1984.
- [82] H. Buxton and R. Howarth, "Watching behaviour: The role of context and learning," in *Image Processing, 1996. Proceedings., International Conference on*, vol. 1. IEEE, 1996, pp. 797–800.
- [83] D. Hogg, "Model-based vision: a program to see a walking person," *Image and vision computing*, vol. 1, no. 1, pp. 5–20, 1983.
- [84] S. J. McKenna, S. Jabri, Z. Duric, and H. Wechsler, "Tracking interacting people," in *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*. IEEE, 2000, pp. 348–353.
- [85] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler, "Tracking groups of people," *Computer Vision and Image Understanding*, vol. 80, no. 1, pp. 42–56, 2000.
- [86] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 5, pp. 564–577, 2003.
- [87] J. M. Wang, D. J. Fleet, and A. Hertzmann, "Gaussian process dynamical models for human motion," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 2, pp. 283–298, 2008.

- [88] S. Hou, A. Galata, F. Caillette, N. Thacker, and P. Bromiley, "Real-time body tracking using a gaussian process latent variable model," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1–8.
- [89] M. Kolsch and M. Turk, "Fast 2d hand tracking with flocks of features and multi-cue integration," in *Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on*. IEEE, 2004, pp. 158–158.
- [90] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanid gait challenge problem: Data sets, performance, and analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 27, no. 2, pp. 162–177, 2005.
- [91] J. Han and B. Bhanu, "Individual recognition using gait energy image," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 2, pp. 316–322, 2006.
- [92] J. Liu and N. Zheng, "Gait history image: a novel temporal template for gait recognition," in *Multimedia and Expo, 2007 IEEE International Conference on*. IEEE, 2007, pp. 663–666.
- [93] L. Wang, T. Tan, W. Hu, and H. Ning, "Automatic gait recognition based on statistical shape analysis," *Image Processing, IEEE Transactions on*, vol. 12, no. 9, pp. 1120–1131, 2003.
- [94] C. BenAbdelkader, R. Cutler, and L. Davis, "Motion-based recognition of people in eigengait space," in *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*. IEEE, 2002, pp. 267–272.
- [95] I. N. Junejo, E. Dexter, I. Laptev, and P. Perez, "View-independent action recognition from temporal self-similarities," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 1, pp. 172–185, 2011.
- [96] J. Little and J. Boyd, "Recognizing people by their gait: the shape of motion," *Videre: Journal of Computer Vision Research*, vol. 1, no. 2, pp. 1–32, 1998.
- [97] K. Bashir, T. Xiang, S. Gong, and Q. Mary, "Gait representation using flow fields." in *BMVC*, 2009, pp. 1–11.
- [98] A. A. Efros, A. C. Berg, G. Mori, and J. Malik, "Recognizing action at a distance," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*. IEEE, 2003, pp. 726–733.
- [99] T. H. Lam, K. H. Cheung, and J. N. Liu, "Gait flow image: A silhouette-based gait representation for human identification," *Pattern recognition*, vol. 44, no. 4, pp. 973–987, 2011.
- [100] C. BenAbdelkader, R. Cutler, and L. Davis, "Stride and cadence as a biometric in automatic person identification and verification," in *Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on*. IEEE, 2002, pp. 372–377.
- [101] D. K. Wagg and M. S. Nixon, "On automated model-based extraction and analysis of gait," in *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*. IEEE, 2004, pp. 11–16.

- [102] L. Haiping, P. Konstantinos N, V. Anastasios N *et al.*, “A full-body layered deformable model for automatic model-based gait recognition,” *EURASIP Journal on Advances in Signal Processing*, vol. 2008, 2007.
- [103] J. Gu, X. Ding, S. Wang, and Y. Wu, “Action and gait recognition from recovered 3-d human joints,” *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, vol. 40, no. 4, pp. 1021–1033, 2010.
- [104] J. Wang, M. She, S. Nahavandi, and A. Kouzani, “A review of vision-based gait recognition methods for human identification,” in *Digital Image Computing: Techniques and Applications (DICTA), 2010 International Conference on*. IEEE, 2010, pp. 320–327.
- [105] Z. Zhang, M. Hu, and Y. Wang, “A survey of advances in biometric gait recognition,” in *Biometric Recognition*. Springer, 2011, pp. 150–158.
- [106] E. Trucco and A. Verri, *Introductory techniques for 3-D computer vision*. Prentice Hall Englewood Cliffs, 1998, vol. 201.
- [107] B. Jähne, *Image processing for scientific applications*. CRC press Boca Raton, 1997.
- [108] M. W. Hirsch, S. Smale, and R. L. Devaney, *Differential equations, dynamical systems, and an introduction to chaos*. Academic press, 2004, vol. 60.
- [109] B. Van der Pol, “The nonlinear theory of electric oscillations,” *Radio Engineers, Proceedings of the Institute of*, vol. 22, no. 9, pp. 1051–1086, 1934.
- [110] A. Iserles, *A first course in the numerical analysis of differential equations*. Cambridge University Press, 2009, no. 44.
- [111] J. Guckenheimer and P. Holmes, *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*. New York Springer Verlag, 1983, vol. 42.
- [112] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [113] Y. Song and J. W. Grizzle, “The extended kalman filter as a local asymptotic observer for nonlinear discrete-time systems,” in *American Control Conference, 1992*. IEEE, 1992, pp. 3365–3369.
- [114] E. A. Wan and R. Van Der Merwe, “The unscented kalman filter,” *Kalman filtering and neural networks*, pp. 221–280, 2001.
- [115] S. J. Julier, “The scaled unscented transformation,” in *American Control Conference, 2002. Proceedings of the 2002*, vol. 6. IEEE, 2002, pp. 4555–4559.
- [116] S. J. Julier and J. K. Uhlmann, “Unscented filtering and nonlinear estimation,” *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.
- [117] E. A. Wan and R. Van Der Merwe, “The unscented kalman filter for nonlinear estimation,” in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*. IEEE, 2000, pp. 153–158.

- [118] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.
- [119] M. A. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [120] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*. Thomson-Engineering, 2007.
- [121] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [122] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603–619, 2002.
- [123] S. J. McKenna, Y. Raja, and S. Gong, “Tracking colour objects using adaptive mixture models,” *Image and vision computing*, vol. 17, no. 3, pp. 225–231, 1999.
- [124] H. K. Khalil, *Nonlinear systems*. Prentice hall Upper Saddle River, 2002, vol. 3.
- [125] R. Van Buskirk and C. Jeffries, “Observation of chaotic dynamics of coupled nonlinear oscillators,” *Physical Review A*, vol. 31, no. 5, p. 3332, 1985.
- [126] A. B. D. M P Murray and R. C. Kory, “Walking patterns of normal men,” *Journal of Bone and Joint Surgery*, vol. 46, no. 2, pp. 335–360, 1964.
- [127] K. Nummiaro, E. Koller-Meier, and L. Van Gool, “An adaptive color-based particle filter,” *Image and vision computing*, vol. 21, no. 1, pp. 99–110, 2003.
- [128] R. Chaudhry, A. Ravichandran, G. Hager, and R. Vidal, “Histograms of oriented optical flow and binet-cauchy kernels on nonlinear dynamical systems for the recognition of human actions,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 1932–1939.
- [129] Z. Liu and S. Sarkar, “Simplest representation yet for gait recognition: Averaged silhouette,” in *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 4. IEEE, 2004, pp. 211–214.

Curriculum Vitae

Nalini Vishnoi received her Bachelor of Science in Computer Science and Physics from the University of Allahabad, India in 2005. She completed her Master of Science in Computer Science from the J. K. Insititute of Applied Physics and Technology, University of Allahabad in 2007. In the same year, she joined the Department of Computer Science at George Mason University (GMU) for her PhD study. At GMU, she has been associated with the Laboratory for the Study and Simulation of Human Movement. In addition to being involved in a number of research projects and writing of grant proposals, she has mentored several high school and undergraduate students working in the laboratory during the summer and fall semesters. She has received *Volgenau School Doctoral Fellowship* and *Vision Award* from GMU several times for her outstanding academic performance. She has reviewed for Pattern Recognition Journal, Gait & Clinical Movement Analysis Society and IEEE Engineering in Medicine and Biology Society. Her research interests include markerless human motion analysis, biomechanics, biometrics, robotics & motion planning.

Publication List

Conferences:

1. Nalini Vishnoi, Zoran Duric and Lynn Gerber, **“Using Histograms of Normal Flow for Gait Recognition”**, submitted to IEEE International Joint Conference on Biometrics (IJCB 2014).
2. Nalini Vishnoi, Anish Mitra, Zoran Duric and Lynn Gerber, **“Motion based Markerless Gait Analysis using Standard Events of Gait and Ensemble Kalman Filtering”**, accepted to IEEE Engineering in Medical Biology Society Conference (EMBC 2014).
3. Nalini Vishnoi, Zoran Duric and Lynn Gerber, **“Markerless Identification of Key Events in Gait Cycle using Image Flow”**, Proceedings of IEEE Engineering in Medical Biology Society Conference (EMBC 2012).
4. Nalini Vishnoi, Cody Narber, Zoran Duric and Lynn Gerber, **“Methodology for Translating Upper Extremity Motion to Haptic Interfaces”**, Proceedings of IEEE Human System Interactions (HSI 2012).
5. Nalini Vishnoi, Cody Narber, Lynn Gerber and Zoran Duric, **“Guiding Hand: A Teaching Tool for Handwriting”**, Proceedings of ACM International Conference on Multimodal Interfaces and Machine Learning for Multimodal Interaction (ICML-MLMI 2009).

Abstracts:

1. Qi Wei, Lynn Gerber, Zoran Duric, Paul Otto, Nalini Vishnoi, Siddhartha Sikdar, **“Data-driven Biomechanical Analysis of the Upper Trapezius Muscles and Neck Movement: A Pilot Study”**, in American Society of Biomechanics (ASB 2013).
2. Nalini Vishnoi, Zoran Duric and Lynn Gerber, **“Identifying Events of Gait Cycle Using Image Flow”**, in Gait & Clinical Movement Analysis Society Conference (GCMAS 2012).
3. Lynn Gerber, Cody Narber, Nalini Vishnoi, Sidney Johnson, Zoran Duric, Leighton Chan, **“Feasibility of Haptic Use for Patients with Chronic Traumatic Brain Injury (TBI)”**, in American Academy of Physical Medicine and Rehabilitation (AAPM&R 2012).
4. Nalini Vishnoi, Zoran Duric and Lynn Gerber, **“Analyzing Upper Extremity Movements Using Video Images”**, in Gait & Clinical Movement Analysis Society Conference (GCMAS 2011).
5. Cody Narber, Ivan Avramovic, Nalini Vishnoi, Zoran Duric, Lynn Gerber, **“Application of Haptic Technology to Fine Motor Learning in College Students”**, in American Academy of Physical Medicine and Rehabilitation (AAPM&R 2010).
6. Wallace Lawson, Nalini Vishnoi, Lynn Gerber, Zoran Duric, **“Markerless Identification of Dissimilarities in Gait Sequences”**, in Joint meeting of the Gait and Clinical Movement Analysis Society and the European Society of Movement Analysis in Adults and Children (JEGM 2010).