A

Ph.D. THESIS

on

Machine Learning Techniques for Vision based Human Gait Analysis

submitted for partial fulfillment for the degree of

Doctor of Philosophy

in

Computer Science and Engineering

Academic Session

(2014-2018)

Supervisors:
Dr. Namita Mittal
Prof. Rajesh Kumar

Submitted By:
Chandra Prakash
(2014RCP9512)

MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY, JAIPUR, RAJASTHAN

Machine Learning Techniques for Vision based Human Gait Analysis

Ph.D. Thesis

Chandra Prakash
(2014RCP9512)

Department of Computer Science and Engineering
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR
November 2018

Dedicated to
My beloved and caring Parents,
Loving Wife Vandana,
Inspirational Brothers and
My sweet daughter Ravya

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

CANDIDATE'S DECLARATION

Declaration

I hereby declare that the thesis titled, **"Machine Learning Techniques for Vision based Human Gait Analysis"** is my own work conduted under the supervision of Dr. Namita Mittal and Dr. Rajesh Kumar. I confirm that:

- This work is done towards the partial fulfillment of the degree of "*Doctor of Philosophy*" at MNIT, Jaipur.
- Wherever any part of this thesis has previously been submitted for a degree or any other qualification at MNIT or any other institution, this has been clearly stated.
- Wherever I have consulted the published work of others, this is always clearly mentioned.
- I have acknowledged all main sources of help.

Signed: _____

Date: _____

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
MALAVIYA NATIONAL INSTITUTE OF TECHNOLOGY JAIPUR

CERTIFICATE

This is to certify that the thesis entitled, **"Machine Learning Techniques for Vision based Human Gait Analysis"** submitted by Chandra Prakash (2014RCP9512) to Malaviya National Institute of Technology, Jaipur for the award of the degree of ***Doctor of Philosophy*** in Computer Science and Engineering is a bonafide record of original research work carried out by him under our supervision.

It is further certified that:

1. The results contained in this thesis have not been submitted in part or in full to any other university or institute for the award of any degree.
2. Mr. Chandra Prakash has fulfilled the requirements for the submission of this thesis.

Dr. Namita Mittal
(Supervisor)
Associate Professor
Department of Computer Science
& Engineering
MNIT Jaipur, India

Dr. Rajesh Kumar
(Supervisor)
Professor
Department of Electrical Engineering
MNIT Jaipur
India

Date:

ACKNOWLEDGMENTS

It gives me immense pleasure to express gratitude and regards to all those people who supported me during the course of this doctoral research work at MNIT Jaipur. I acknowledge the involvement and contribution of each one of them.

I deeply owe a debt of gratitude to my supervisor Dr. Rajesh Kumar and Dr. Namita Mittal for all the support and motivation that they showered upon me throughout the doctoral program. They kept the faith in me and always backed my research proposal and ideas. I attribute the successful realization of this research work to their sincere and consistent guidance, versatile and unconventional thinking and their penchant for trying something new and taking risks.

My special thanks to the members of Doctoral Research Evaluation Committee (DREC), Dr. Santosh Kumar Vipparthi, Dr. Arka Prokash Mazumdar, Dr. Rajesh Kumar, Dr. Namita Mittal and Dr. Girdhari Singh for their constructive criticisms and valuable suggestions. I am also grateful to all other faculty members of the department for their inspiration and suggestions at various stages. I am also thankful to the staff members of the department for their kind help in official work.

Let me express special thanks to Head of the Department, Dr. Girdhari Singh , for the keen support and consistent encouragement in our academic activities. I am extremely thankful to Prof. Udaykumar R Yaragatti, Director, MNIT Jaipur for providing me infrastructural facilities to work in, without which this work would not have been possible.

Special regards to colleagues and friends at RAMAN Lab, MNIT namely Sujil, Venu, Vishu, Shalini, Lokesh, Srikanth, Akanksha, Om Ji, Rahul, and Harish for their constant support, warmth and affection. I would like to give my special thanks to Dr Shashank Vyas, for the long discussion and exchange of ideas. I learn a lot from him and always feel motivated after discussion with him. Thanks to all members of the robotics group ZINE for being a source of inspiration. Special thanks to Anshul, Kritika, Manas, Uddeshya and Avinash.

I would like to thank my co-scholars Maroti Deshmukh, Avani, Vijay Sharma, Mithlesh Arya, Ankit, Vimal, Tanvi and Kuwar pal, other research colleagues of the department for their loving cooperation, positive criticism, excellent advice, consistent support and consideration during the preparation of this thesis and to spend quality time together. I express my gratitude to my estimable friends Nitin

Umesh, A. R. Quadri and Vijay Meena for their perpetual support, motivation, love and care.

I would also like to thank all the volunteers who helped me to collect the database for this study. I am very thankful to the referees who reviewed this work as pieces of it were submitted to Journals and conferences. Their detailed reviews, constructive criticism and excellent advice have improved both the presentation and content of this thesis.

I feel a deep sense of gratitude for my beloved parents Smt. Rama Meena and Shri Shankar Lal Meena, and my siblings (Anupam kuamr and Deepak kumar and Rajnee), who formed part of my vision and taught me good things that really matters in life. Their infallible love and support has always been my strength. Their patience and sacrifice will remain my inspiration throughout my life. I am very thankful to my better half Vandana and my sweet little girl "Ravya" for their extraordinary and highly appreciated patience day and night during my Ph.D.

Finally, I gratefully, acknowledge one and all who are directly or indirectly involved to shape this research work.

Contents

Abstract **v**

List of Figures **vi**

List of Tables **xi**

1 Introduction **1**
1.1 Gait Analysis . . . 3
1.2 Motivation . . . 4
1.2.1 Primary Motivation . . . 7
1.2.2 Secondary Motivation . . . 7
1.3 Research Gaps . . . 7
1.4 Research Objectives . . . 8
1.5 Contributions . . . 9
1.6 Thesis Organization . . . 9

2 Literature Review **13**
2.1 Basics of Human Gait . . . 13
2.2 Parameters in gait analysis . . . 16
2.2.1 Normal Gait . . . 19
2.2.2 Pathological Gait . . . 19
2.3 Human Gait Analysis Approaches . . . 20
2.3.1 Vision Based Approach . . . 21
2.3.2 Sensor based Approach . . . 22
2.3.3 Other Technologies . . . 23
2.3.4 Hybrid Approaches . . . 24
2.4 Application domain . . . 24
2.4.1 Analysis based Applications . . . 24
2.4.2 Biometric Trait . . . 26
2.4.3 Artificial Gait . . . 27
2.4.4 Control based Applications . . . 28
2.4.5 Other Applications . . . 28
2.5 Machine Learning Techniques . . . 29
2.5.1 Supervised Learning . . . 29
2.5.2 Unsupervised Learning . . . 32

2.5.3 Reinforcement Learning . . . 32
2.5.4 Rule Based Learning . . . 33
2.5.5 Evolutionary Learning . . . 33
2.5.6 Probabilistic Learning . . . 33
2.5.7 Hybrid Learning . . . 34
2.6 Dataset for Gait Analysis . . . 39
2.7 Chapter Summary . . . 43

3 Development of Gait Parameter Extraction (GPE) Approach 45
3.1 Experimental Setup . . . 46
3.1.1 System Description . . . 47
3.1.2 Subjects/Participants . . . 47
3.1.3 Clinical Gait Model . . . 49
3.2 Passive Marker based Gait Parameter Extraction Approach (PM-GPEA)- Model 1 . . . 51
3.2.1 Proposed Methodology for Passive marker based Gait Analysis System . . . 52
3.2.1.1 Preprocessing . . . 52
3.2.1.2 Marker Identification and Tracking . . . 54
3.2.2 PM-GPEA Model: Preliminary Experimental Results . . . 56
3.2.2.1 Spatial Temporal Parameters Estimation . . . 56
3.2.2.2 Kinematic Parameters Estimation . . . 60
3.3 Marker-less/Holistic Based Gait Parameter Extraction Approach (Ml-GPEA)- Model 2 . . . 61
3.3.1 Methodology for Ml-GPEA . . . 61
3.3.1.1 Silhouette image extraction . . . 64
3.3.1.2 Body segmentation . . . 69
3.3.1.3 Identification of joint co-ordinates . . . 70
3.3.2 Ml-GPEA Model: Preliminary Results and Discussion . . . 74
3.3.2.1 Joint coordinates . . . 74
3.3.2.2 Spatio-temporal parameters estimation . . . 76
3.3.2.3 Kinematic parameters estimation . . . 77
3.4 MNIT Gait Dataset Analysis: A step towards INDIAN Gait Norms 81
3.4.1 MNIT Gait dataset: Adult Gait Norms . . . 81
3.5 Chapter Summary . . . 86

4 Development of Gait Pattern Prediction (GPP) Models 89
4.1 Individual-Specific GPP Models from historical Gait Data (GPP-HD) . . . 90
4.1.1 Data Selection from MNIT Gait Dataset . . . 90
4.1.2 Linear Time Series (LTS) Model . . . 92
4.1.2.1 Univariate Linear Time Series (ULTS) Model . . . 92
4.1.2.2 Multivariate Linear Time Series (MLTS) Model . . . 92
4.1.3 Back Propagation Artificial Neural Network (BP-ANN) . . . 93
4.1.3.1 Parameter Settings . . . 94

4.1.4 Support Vector Regression (SVR) . . . 95
4.1.5 Results and Discussion: GPP-PD Models . . . 96
4.1.5.1 Sample Selection from MNIT Gait Dataset . . . 96
4.1.5.2 Evaluation Parameters . . . 97
4.1.5.3 Discussion . . . 97
4.2 GPP Models from Anthropometric Data (GPP-AD) . . . 101
4.2.1 Methodology for the GPP-AD Models . . . 102
4.2.1.1 Data selection for GPP-AD Models . . . 102
4.2.2 Principal Component Analysis (PCA) and BP-ANN . . . 103
4.2.3 Gaussian Process Regression (GPR) . . . 106
4.2.4 GPP-AD : Results and Discussions . . . 106
4.3 Chapter Summary . . . 107

5 Machine Learning Techniques for Gait Abnormality Detection 109
5.1 Fuzzy Logic based Gait Phase Detection From Clinical Perspective (FL-GPD) . . . 110
5.1.1 Proposed Methodology for Gait Phase Detection . . . 112
5.1.1.1 Input parameters . . . 114
5.1.1.2 Fuzzy Inference System (FIS) . . . 118
5.1.2 Experimental Result :Fuzzy-based Approach for Phase identification . . . 120
5.1.2.1 CASE 1: Healthy Gait Pattern . . . 120
5.1.2.2 CASE 2: Unhealthy Gait Pattern . . . 122
5.1.2.3 Discussion . . . 122
5.2 Optimized Clustering Techniques for Identification of Abnormal Gait Profile . . . 123
5.2.1 Experimental Methodology . . . 126
5.2.1.1 Gait Data-set Description . . . 126
5.2.1.2 Data Clustering Techniques . . . 127
5.2.1.3 Proposed Grey Wolf Optimization (GWO) Clustering Technique . . . 134
5.2.1.4 Optimization Parameter Settings . . . 140
5.2.2 Case Studies Considered . . . 140
5.2.3 Result Analysis: Optimization based Clustering Techniques 142
5.2.3.1 Clustering Evaluation Indices . . . 142
5.2.3.2 Discussion . . . 144
5.3 Chapter Summary . . . 151

6 Conclusion 153
6.1 Contribution . . . 153
6.2 Limitation and Future work . . . 154

Appendix 160

Abstract

Although walking is natural phenomena in human, but its significance is felt when it is distorted. The research associated with human walking is known as gait analysis. Gait analysis has come into picture since the last decade because of its numerous applications in biometric, artificial gait, analysis (normal gait, clinical, geriatric care and sports monitoring & tactics), animation industry, control applications, etc. Now, with the new advanced techniques, clinicians can exploit vision, sensor, and other hybrid instrumental based gait analysis approaches, in their regular clinical practice to assess a patient's status for complex musculoskeletal and neurological disorders. However, the cost associated with gait analysis systems are very high, as the analysis is conducted in highly specialized laboratories. In such constraint environment, the subject is not able to reflect the natural gait pattern. These are some of the factors that limit the popularity of the gait analysis in the medical arena. Thus there is need of alternate Gait Parameter Extraction Approaches (GPEA). Gait movements of 120 healthy and 80 unhealthy subjects are collected and analyzed. This collected dataset is named as MNIT Gait Dataset. In this thesis, two vision-based methodologies to extract gait parameters (kinematic and spatiotemporal) are developed; first is Passive Marker-based Gait Parameter Extraction Approach (PM-GPEA), and second is Marker-less Gait Parameter Extraction Approach (Ml-GPEA). The patterns of both proposed models are consenting when compared with the available standard gait setup. The promising results of the PM-GPEA are used to analyze the MNIT Gait dataset. Proposed PM-GPEA based gait parameters results when compared with the western gait standard, shows the deviation as with the nationality, the gait patterns also change. In this thesis, two Gait Pattern Prediction (GPP) models are proposed to increase the efficiency of gait rehabilitation approaches. The first model is individual-specific GPP Models from historical Gait Data (GPP-HD). The second model developed is GPP models from anthropometric data (GPP-AD) considering body parameters such as height, weight, age, etc. Fuzzy Logic based Gait Phase Detection (FL-GPD) technique is proposed for the detection of unhealthy gait pattern in the collected dataset. Later, a vision based gait data of children having Cerebral Palsy (CP) disease is also analyzed. To analyze the dataset, a novel nature-inspired meta-heuristic algorithm, Grey Wolf Optimized Clustering (GWOC) approach is proposed to find the optimal number of gait profiles. The Experimental results show the effectiveness of all the proposed approaches.

List of Figures

1.1 Gait analysis applications (a) Robotics (b) Resistive Rehabilitation (c) Smart prosthetic leg (d,e) Orthotic (f) Surveillance (g) Elder Person care (h) Sports monitoring & tactics 2
1.2 Available gait parameters extraction techniques 3
1.3 Gait Lab Setup 4
1.4 Thesis Organization 11

2.1 Link segment model a) Three anatomical positions in coordinate system, b) Simplified 2D model of the lower limb 14
2.2 Fundamental gait phases with expected interval of gait cycle 15
2.3 Comprehensive parameter tree in gait analysis 16
2.4 Hip, knee and ankle joint angles for a Healthy subject [1] 18
2.5 Healthy gait parameter for different age group 19
2.6 Available gait analysis approach: A comprehensive tree diagram . . 21
2.7 Application of human gait analysis 25

3.1 Proposed models for the gait parameter extraction. 46
3.2 Gait lab setup at RAMAN Lab, MNIT Jaipur 46
3.3 Few sample of subjects of MNIT Gait dataset 48
3.4 MNIT Gait sample acquisition (a) Male subject walking from left to right (b) Female subject walking from right to left 49
3.5 Canadian Society of Biomechanics (CSB) Clinical Gait Model: marker locations at anatomical points on subject body (in red color) . . . 51
3.6 CSB Gait standard angle calculation method 52
3.7 Passive marker based proposed methodology for vision based Gait analysis 53
3.8 (a) Marker at anatomical points on subject's body identification in real time environment (in red color) (b) identification of coordinates for each marker 55
3.9 Result of Passive marker based Gait event detection (PMGED) System for a random subject and marker trajectory. (a) Trajectory plot of difference of malleolus and lateral metatarsal coordinatesand (b) and (c) shows the Heel Strike (HS) and Toe Off (TO) with the marker trajectory. 58
3.10 Result of Passive marker based gait Analysis approach. (a)Joint coordinates (b) Segments angles for the whole video sequence 62

3.11 PMGPEA result for S22, dotted lines are angles for the different gait cycle and dark line represents the average angle; a)Knee angle (b) Hip angle (c) Ankle angle 63
3.12 Flow diagram for parameter extraction in markerless gait analysis . 64
3.13 (a) Foreground extraction techniques on the subject frames in different color space result on RBG color space (b)result on HSV color space 66
3.14 Silhouette image extraction result (a) Foreground extraction (b) Morphological operations (c) Segmentation 68
3.15 Knee Coordinate detection for bent and staight leg cases 72
3.16 Preliminary result of proposed Ml-GPEA (a) Identification of left and right leg from silhouette images (b) Joint position estimation by the proposed method for female subject (S22) and male subject (S43) 73
3.17 Bar graph for horizontal and Vertical coordinates 75
3.18 Average knee and hip angle for a female subject 77
3.19 Average knee and hip angle for a male subject 78
3.20 Female and Male Left and right Knee and hip Angle respectively in a marker-free environment 80
3.21 Kinematic comparison of Passive marker based and approach with the golden standard for male subject 82
3.22 Knee-hip and Ankle-knee cyclogram at self selected speed 82
3.23 ANOVA analysis on male angles for age group 18-49 years. (a) Result using two σ male knee angle.(b) Result using two σ male hip angle. (c) Result using four σ male knee angle. (d) Result using four σ male hip angle. (e) Result using six σ male knee angle. (f) Result using six σ male hip angle. 84
3.24 ANOVA analysis on female angles for age group 18-49 years. (a) Result using two σ female knee angle.(b) Result using two σ female hip angle. (c) Result using four σ female knee angle. (d) Result using four σ female hip angle. (e) Result using six σ female knee angle. (f) Result using six σ female hip angle. 85

4.1 Plot of male subject (S22) knee angle, seven samples of Gait cycle (GC=7) 91
4.2 Architecture of Back Prorogation Neural Network (BP-ANN) for GPP-HD model 94
4.3 Performance plot for knee angle prediction using various model for SS21 99
4.4 Co-relation coefficient of (a) Knee angle dataset (b) Hip angle for subject S32 100
4.5 Methodology for Gait Pattern Prediction Model from Anthropometric parameters (GPP-AD) 102
4.6 Aggregate knee and hip angle of 120 subject's of MNIT Gait Dataset 104
4.7 Result for kinematics parameters prediction model from anthropometric parameters 107

5.1 Fundamental gait phases and expected interval phases and sub-phases in complete gait cycle 110
5.2 Gait phase detection methodology; (a) Marker setup and data acquisition; (b) Gait analysis system; (c) Gait kinematics extraction; (d) Fuzzy inference system for gait phase identification; (e) Phase detection 112
5.3 Modules of fuzzy logic expert system 114
5.4 Average knee and hip joint angle for one gait cycle of subject S22 . 115
5.5 Membership function for input variable: hip, knee, time and stage and output variable gait phase 116
5.6 Membership function for output variable (gait phase) 118
5.7 Result after fuzzy-inference for normal gait phase plots for two strides 120
5.8 Fuzzy-logic based phase detection in one gait cycle 121
5.9 Kinematics signals and phase detection 121
5.10 Unhealthy subject (US22) fuzzy-logic based phase detection for one gait cycle 122
5.11 Clustering based methodology used for gait profiling of CP Children 126
5.12 Grey Wolf Optimization 135
5.13 Flow chart for Grey Wolf Optimization (GWO) approach 137
5.14 Two dimensional plot of gait data for (a) CP and Normal Children natural data; (b) Polynomial Normalization and scale gait data with CP and Normal Children and (c) Gait Profiling after clustering approach. 145
5.15 Bar plots of clustering performances indices from the different clustering approaches 147
5.16 MSE Convergence plot for the different clustering approach 148
5.17 Composite error and indices plot for case 4 in terms of MSE, SC and DI for K-means, GA, H-GA and PSO clustering in case of $2 \leq k \leq 7$. 149
5.18 Gait profile plots for gait profile considering 5 clusters using affiliation probability 150

List of Tables

1.1 Analysis of previous gait analysis surveys 5

2.1 Definitions of some of terminologies used in gait analysis 17
2.2 Machine learning techniques highlighting potential gait application . 31
2.3 Applications based machine learning techniques 35
2.4 Survey of available database 40

3.1 MNIT Gait dataset based on age-groups and genders 49
3.2 Anthropometric parameters of the MNIT Gait dataset 50
3.3 An overview of the subjects considered for the experiment 57
3.4 Statistical result for the proposed Passive marker based Gait event detection (PMGED) system 60
3.5 Spatio Temporal parameter observed in the study 60
3.6 Mean, minimum, maximum and RMSEs value for shoulder, hip and knee coordinates by marker based, proposed markerless system and ground truth 75
3.7 Extracted Spatio-temporal Gait Parameters using MI-GPEA Model 76
3.8 Kinematic performance for a female subject (S22) male subject (S43) in a marker-free environment 79
3.9 Spatio-Temporal parameter data of the MNIT Gait dataset for age group 21-40 83
3.10 Comparisons of Sagittal based angle range of hip, knee and ankle joints during walking published in previous studies 83

4.1 Sample gait pattern data of a specific subject for p gait cycle . . . 91
4.2 Subject selection from MNIT Gait dataset based on age-groups and genders 96
4.3 MAE, RMSE, MAPE measures for the predicted knee Angle 98
4.4 Mean and standard deviation of 120 experiment participantBP-ANNs body parameters from MNIT Gait dataset. 103
4.5 Result for kinematics parameters prediction model from anthropometric parameters using Artificial Neural network with and without PCA and gaussian regression 106

5.1 Fuzzy membership function for input parameters 116
5.2 Set of fuzzy rules used in the study 119
5.3 Extracted gait phases using proposed system for subject S34 122

5.4 Anthropometric and the spatio-temporal gait parameters of the participating subjects 127
5.5 Cases considered in this study 140
5.6 Testing gait data of four subjects A, B, C and D 141
5.7 Clustering result for case 1, when all four gait parameters (stride length, cadence, leg length and age) without normalization and scaling for k=2 are considered. 144
5.8 Clustering result for case 2, when first two gait parameters (stride length and cadence) in their original form for k=2 144
5.9 Clustering result for case 3, when considering stride length and cadence, after Polynomial normalization with leg length and age respectively for k=2 146
5.10 **Result of t-test considering MSE on case 3** 147

Chapter 1

Introduction

Walking is one of the natural and common traits of the human being. However, from the perspective of analysis, it is among one of the most complex phenomena [1]. It is the combined effort of the brain, nerves, and muscles. Walking indicates liberty and individuality in humans, and thus any deviation from the normal pattern can sternly reduce our quality of life.

Nutt et al. define walking as the coordinated movement of lower limbs with spanned flexion-extension in an involuntary and recurring fashion [2]. The pattern of locomotion (walks, runs, crawls, etc.) combined with their posture is known as gait. As a species, the human is bipedal, i.e., we move on two extremities. Gait is a way of locomotion, rather than the walking process itself. Most of the people tend to use the words gait and walking interchangeably. Whittle suggests that gait comparison makes more sense than walking [1].

Research associated with human walking is known as gait analysis. Gait analysis is a way to reveal the mechanisms of human movement by quantifying factors governing the functionality of lower extremities. It assists in recognizing deviations in the gait pattern and determining their reason and effects [3, 4]. Gait analysis has come into picture since the last decade because of its numerous application in different fields such as surveillance, biometrics, sports, medical, rehabilitation, animation industry, etc. as illustrated in figure 1.1. Gait can be utilized as one of the biometric traits and can be used in surveillance where a large number of people pass through. In the last decade, there is a remarkable growth in both lower and upper extremity based rehabilitation assistive devices. Nguyen et al. [5] proposed a three dimension system for mixed reality, and it can be used in the entertainment industry to create 3D human avatars. Gait analysis techniques are

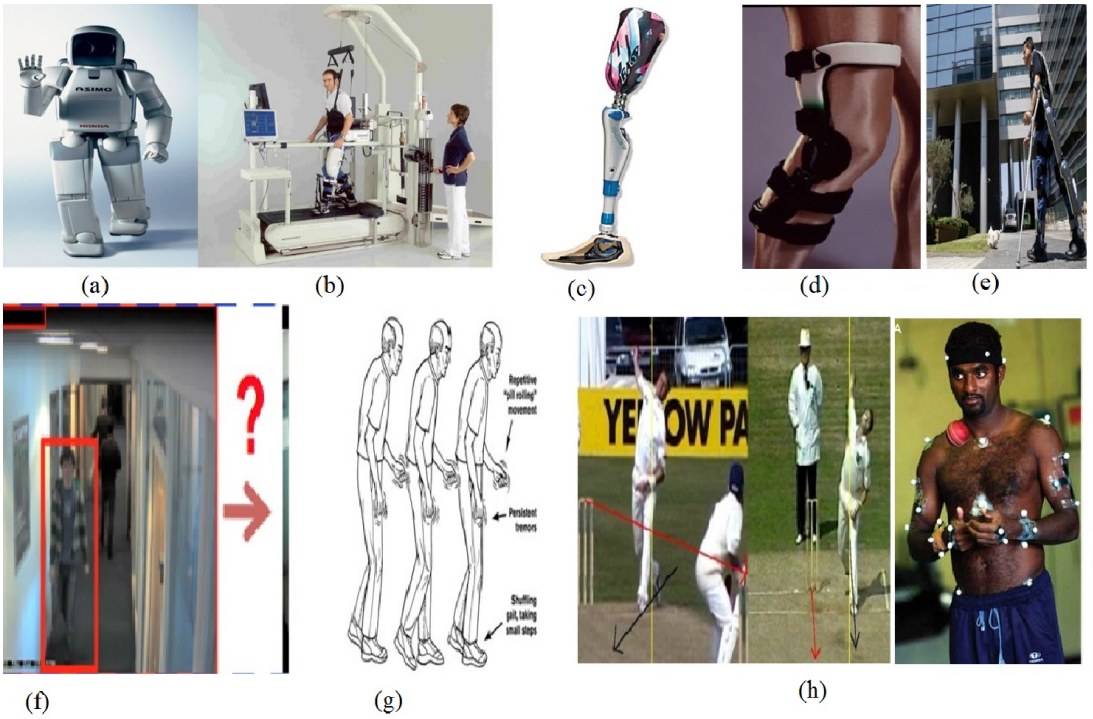


Figure 1.1: Gait analysis applications (a) Robotics (b) Resistive Rehabilitation (c) Smart prosthetic leg (d,e) Orthotic (f) Surveillance (g) Elder Person care (h) Sports monitoring & tactics

used to optimize and increase athletic performance. Analysis of walking pattern may suggest potential injury, and accordingly, preventive action can be taken.

According to United Nations [6], Asia is a home to close to two-thirds of the world's population of physically challenged and this figure is expected to climb over the next decade. However, now with advanced measurement technology and biomechanical modeling, there are more opportunities for better quality of life for the disabled community worldwide. The desire to improve the quality of life and reliability of rehabilitation techniques drives new research and development activities in several countries. Motion analyzer systems have been widely used to monitor a patient's response to the gait rehabilitation systems.

Physiotherapists and neurologists can examine gait movement variables, i.e., stride length, step length, cadence, stance and swing phase, etc. and evaluate a patient's status, treatment, and rehabilitation [1, 7]. The empirical and quantitative analysis of gait variability using kinematics and kinetic characterizations can be helpful to healthcare professionals in both predicting onsets of a condition and to monitor a patient's recovery status through clinical approach. Moreover, these quantitative results may help to strengthen their confidence in the rehabilitation through orthoses, prosthetics and surgical procedures.

1.1 Gait Analysis

Human motion analysis is a challenging problem due to variations in human movement and appearance, camera viewpoint, and environment settings. A larger section of researchers, fascinated by this area proposed new techniques.

Contemporary human gait analysis can be broadly classified into four different approaches; vision-based, sensor-based, other technologies based and the combination of aforementioned approaches [8]. The different gait parameters extraction techniques are shown in figure 1.2.

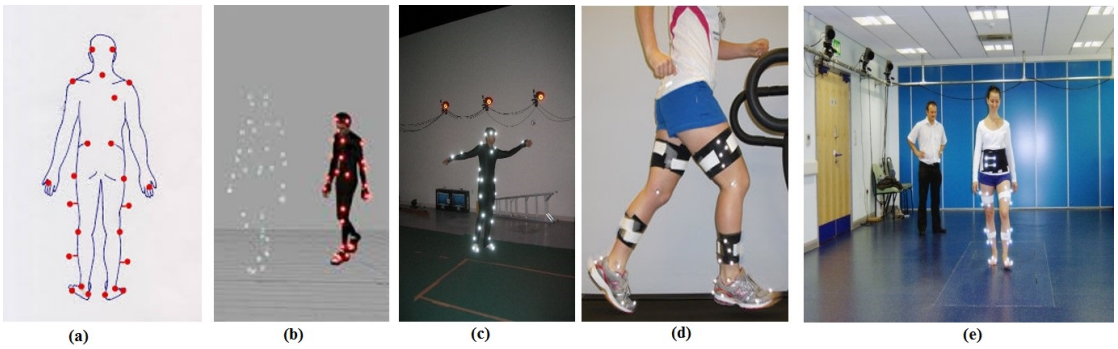


Figure 1.2: Available gait parameters extraction techniques

At present, numerous commercial motion capture systems for both 2D and 3D analyses are available (Visual3D, 3D Gait, Vicon motion capture system, Microsoft Kinect™sensor, OptiTrack, Kistler, BTS Bioengineering, BIOPAC MP-150, etc.). Vision-based kinematic gait analysis can be subdivided into two types; contact and vision based [8, 9, 10, 11] and shown in figure 1.2(a-e). Contact-based techniques include accelerometers and goniometers sensors, while in vision-based techniques active or passive markers are developed to perform real-time kinematics gait analysis. The active marker is triggered to illuminate as they use light-emitting diode (LEDs). This signal is used to specify the position of the marker. Each marker has its own identity through predefined frequencies. Passive markers are spheres covered with reflective Scotchlite™tape that reflects incident light. These LED-based systems are highly accurate however such benefits come at a price and subject has to carry many cables or other components that can affect the subject walking pattern. Ground Reaction Force (GRF) plates are used to calculate force magnitude and direction when the foot contacts the force plate as shown in figure 1.2(e). Pressure sensors are placed inside the sole of shoes to find the kinematic parameter in the gait analysis [12, 13]. Surface or middle based electromyography (EMG) and inertial systems are placed on the body of the subject. Force Platform is also used

Figure 1.3: Gait Lab Setup

to obtain the kinetics of the subject's movement. Electromagnetic, magnetic and medical imaging based systems can be used as a human gait analysis approach as other technologies based analysis. Accelerometers, gyrometer, and magnetic sensors are used to measure the spatiotemporal gait parameters, but these increase the cost of the gait analysis [14]. In India, it may not be within the budget of most hospitals [15]. So there is an urgent need for an alternative cost-effective and reliable system.

Availability of quantitative standard gait parameters is essential for the identification of balance features, detection of gait disorders, and assessment of medical gait interventions. Table 1.1 epitomizes visual overview of gait relevant surveys and articles published in the past two decades, on human walking analysis. Review by Aggarwal et al. [16] is perceived as the first survey paper on gait analysis study. Papers are arranged in order of year of publication and approach either Vision Based (VB) or Sensor based (SB) are also mentioned in table 1.1. It suggests that gait analysis is an emerging research area with an application in clinical pathology and biometrics.

1.2 Motivation

Conventionally, the human gait has been considered subjectively through visual observations, but now with technological advancement, human gait analysis can be done objectively and empirically. Following are the two motivation factors that drives this study:-

Table 1.1: Analysis of previous gait analysis surveys

Year	Author	Comment/Focus	VB	SB
1994	Aggarwal et.al	Priori shape and model based Articulated and elastic non rigid motion	Yes	Yes
1995	Cedras and Shah	Extraction of Motion information from frame and recognize activity	Yes	No
1996	Ju	Motion estimation and recognition	Yes	No
1996	Whittle	Gait analysis considering video based , EMG, kinematics , kinetics parameters for clinical purpose	Yes	Yes
1998	Aggarwal et al.	Articulated and elastic non rigid motion	Yes	Yes
1999	Aggarwal and Cai	Human motion tracking, analysis and recognizing activity using single/multiple cameras	Yes	No
1999	D. M. Gavrila	Recognize humans and their activities , applications	Yes	No
2001	Chau	Explore application of fuzzy, multivariate statistical and fractal technique for gait data	Yes	Yes
2001	Chau	Neural network and wavelet method for gait data	Yes	Yes
2001	Moeslund and Granum	Initialization, tracking, pose estimation, and recognition techniques in last 2 decade	Yes	No
2001	Sutherland	Discuss EMG methods to gait applications	No	Yes
2002	Sutherland	Focus on Kinematics methods for gait applications	No	Yes
2003	H. Buxton	Model for interpretation of a dynamic scene	Yes	No
2003	Wang and Singh	Human tracking and modeling behavior related approaches	Yes	No
2003	Wang et al.	Human detection, tracking, and activity recognition	Yes	No
2004	Aggarwal and Park	Human Recognition	Yes	No
2004	Hu et al.	Gait analysis in Surveillance	Yes	No
2005	Sutherland	Focus on kinetics and energy assessment to gait applications	No	Yes
2006	Baker	Methodologies for gait analysis in a clinical rehabilitation	Yes	Yes
2006	Moeslund et al.	Initialization, tracking, pose estimation and recognition	Yes	No

Table 1.1-*Analysis of previous gait analysis surveys Continued from previous page*

Year	Author	Comment/Focus	VB	SB
2007	Owusu	Computational technologies for sports	Yes	Yes
2007	R. Poppe	Modeling and Pose Estimation	Yes	No
2007	Kruger et.al	Classification of human action representation, recognition, synthesis and understanding of action	Yes	No
2008	Turaga et al.	Human activity analysis - approaches based on complexity	Yes	No
2008	Vasconcelos et al.	Computational techniques in human motion analysis	Yes	No
2009	Liu et al.	Video based gait recognition	Yes	No
2010	Mannini and Sabatini	Classification of human physical activity	Yes	Yes
2010	Wang et al	Gait recognition Techniques	Yes	No
2011	Aggarwal et al.	Compare different types of complex human activity recognition approaches	Yes	No
2011	Chai et al.	Human gait recognition datasets approaches	Yes	No
2012	C.B Ng et al.	Gender recognition	Yes	No
2012	Tao et al	Human Kinematics kinetics parameters from wearable sensors	No	Yes
2014	Gowsikhaa et al.	Human behavior recognition	Yes	No
2014	Shirke et.al	Model free gait recognition approach	Yes	No
2014	Tracey K. M. Lee et al.	Technologies in Gait analysis and recognition	Yes	No
2015	Wright and Jordanov	Computational techniques for legged locomotion	Yes	Yes
2015	Connie et al.	Cross and multi-view gait recognition	Yes	No
2016	Chen et al.	Quantitative gait analysis techniques with wearable sensors	No	Yes
2017	Benjamin et al.	Machine-based human emotion perception	Yes	Yes
2018	Chandra et al.	Gait analysis approaches, machine learning techniques, applications and available datasets for gait	Yes	Yes
2018	Berretti et al.	Representation, Analysis and Recognition of 3D Humans	No	Yes

1.2.1 Primary Motivation

As illustrated in figure 1.3, the gait analysis is conducted in highly specialized laboratories or physicians consulting rooms with expensive instruments. The placement of sensors (active markers) on the body is very crude (in term of the clothing). In such a constraint environment, the subject is not able to reflect the natural gait pattern. The cost associated with the gait analysis system is very high [14]. These are some of the factors that limit the popularity of the gait analysis systems among people[7]. Researcher around the globe, are developing the approaches to get the data with minimal sensor placement on the subjects body or without a marker. Accordingly, the gait analysis system should be of low cost and robust and adaptable by the users.

1.2.2 Secondary Motivation

India has a population of more than 1.5 billion, the second largest population, but no population-based lower extremity gait parameter study has been found for the Indian scenario. Thus due to the lack of Indian gait norms, the standard of the western gait kinematic parameters are frequently used in clinical applications across India.

1.3 Research Gaps

Even the current state of the art in data collection is more accurate, but still, the system is not able to capture the actual gait pattern of subjects. Another issue in gait analysis is that walking pattern is affected by a large number of extrinsic, intrinsic, physical, psychological and pathological factors. Researchers are still not able to find the correlation between these influencing factors on normal walking. The problem with the existing vision-based gait analysis approaches and techniques are as follows:

- As per the current state of the art, Marker-based systems are accurate; however, subjects are not able to exhibit their natural gait pattern under controlled environment setup. Existing literature reveals that fixed viewpoints and hardware limit gait applications. Thus, there is the scope of more research in marker-less approaches in two and three-dimensional space.

- There are two approaches in marker-less based gait analysis, i.e., model based and model free. In the former, human body/structure/motion is modeled, and features are extracted to match them to the model component. While in model-free approach eigen gait/key frame/ kinematic features etc. are analyzed. Still these markerless is not fully explored.

- Gait data is highly heterogeneous, high dimensional, temporal dependent, variable. It is not easy to process this data. Conventional Machine Learning techniques are limited in their ability to process natural gait data in their raw form.

- Current Supervised Machine learning Techniques are based on the training-testing based approach. A large amount of training data is required for higher accuracy. Thus, there is a need to find an alternate approach such as semi-supervised learning.

- No standard Indian gait dataset is found. There is a scope for a new dataset that can be used for defining normal gait pattern among gender and age groups and can be used for recognition of human behavior and activity.

1.4 Research Objectives

The goal of this study is to propose cost-efficient yet reliable, robust and automatic framework for vision-based gait analysis with enhanced machine learning techniques. The research objectives considered in this thesis are as follows:

1. To develop gait parameters extraction methodologies with minimal sensor placement on the subject's body and without markers, from image sequences in an unconstrained environment.

2. To develop machine-learning based gait parameter prediction frameworks/models that can help in rehabilitation based on the past anthropometric /kinematic/spatiotemporal data from MNIT Gait dataset, collected during the study.

3. To perform extensive quantitative and statistical analysis of MNIT gait dataset to discover the scatter profile and ascertain its applicability as a standard dataset from Indian Perspective.

4. To explore machine learning based gait abnormality detection techniques. To create a machine-learning based framework for enabling a predictive strategy for detection of posture deviation based on knowledge of predicting parameters/gait profiles/features/signals/symptoms/signatures.

1.5 Contributions

We are motivated by the fact that joint angles can be extracted either with minimal use of the marker or without using marker techniques. Following are the contribution of this thesis.

1. Passive Marker-based Gait Parameter Extraction Approach (PM-GPEA) is developed to extract the kinematic and spatiotemporal parameters.
2. To extract the gait parameters in an unconstrained environment, Marker-less Gait Parameter Extraction Approach (Ml-GPEA) is proposed.
3. Two Gait Pattern Prediction (GPP) models have been developed based on the past gait patterns(joint angles) and body parameters, data collected during the study.
4. The fuzzy logic based gait phase detection technique is proposed for the detection of unhealthy gait pattern in the collected dataset.
5. A novel nature-inspired meta-heuristic algorithm, Grey Wolf Optimized Clustering (GWOC) approach is proposed to find the optimal number of gait profiles in the considered dataset that aids in the rehabilitation.

1.6 Thesis Organization

Figure 1.4 provides a flow of the thesis in the form of a block diagram highlighting the major portions covered in each chapter. Chapter 1 introduces the definition and applications of the gait analysis, and it also highlights the motivations and research gaps in the vision-based gait analysis. Finally, in this chapter, research objectives and contributions of the thesis are presented. Chapter 2 presents a summary of recent developments in human gait research. It summaries the various parameters, approaches, machine learning techniques, and datasets used in gait

analysis. Chapter 3 is a detailed description of the two proposed gait parameter extraction approach. The first approach is passive marker based, and it minimizes the use of sensors on the subject body. The second proposed approach is marker-free that extracts the gait parameter/features in a constraint-free environment. Statistical analysis of the MNIT Gait dataset is also presented in this chapter. Chapter 4 discuss gait pattern prediction models based on the past data. It also presents the joint angle prediction model based on the anthropometric data such as age, gender, weight, height, etc. using supervised learning techniques. In chapter 5, gait abnormality detection using two proposed machine learning techniques is presented and discussed. Fuzzy-logic based gait phase identification is proposed followed by the hybrid optimization based clustering technique. Chapter 6, concludes the machine learning approaches for vision-based gait analysis and suggests the future research directions in this field.

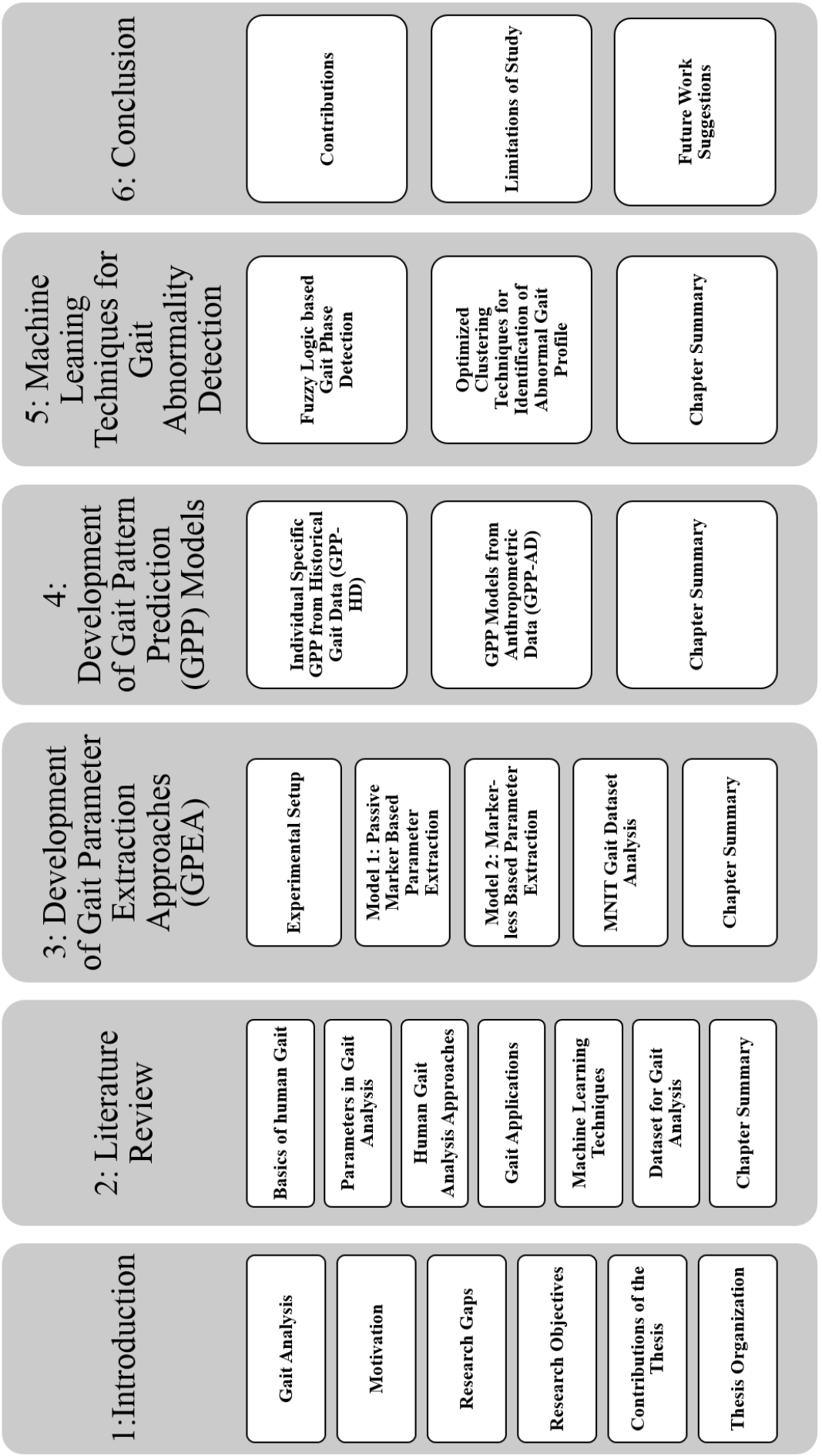


Figure 1.4: Thesis Organization

Chapter 2

Literature Review

This chapter provides an overview of essential taxonomies and representation techniques of human gait available in research articles in section 2.1, and 2.2, followed by details human gait analysis approaches in section 2.3. A detailed discussion on different application domains of gait analysis is presented in section 2.4. Section 2.5 focuses on the available machine learning techniques. Section 2.6 presents the datasets available for gait analysis. A summary of the chapter and its implications is presented in section 2.7.

2.1 Basics of Human Gait

For understanding pathology, normal gait pattern is essential to be able to detect alteration in gait. Figure 2.1a) illustrates the anatomical position with three reference planes in gait analysis [1]. YZ plane is known as the frontal plane (divides the body into front and back segments) while the XZ plane is corresponding to the traverse plane that septate the body into upper and lower segments. XY plane is known as the sagittal plane and divides the body part into right and left portions. Figure 2.1(b) shows the 2D model for the lower limb to extract the joint angles and torque.

Gait Cycle Walking is considered as a series of cyclic events known as gait cycles. In this thesis, the gait cycle and gait patterns are used interchangeably. Weber brothers used the concept of the gait cycle and calculated the timing of gait in 1836 [17]. A gait cycle comprises of the activities that happen from the point of initial contact of one lower limb to the point at which the same limb contacts the

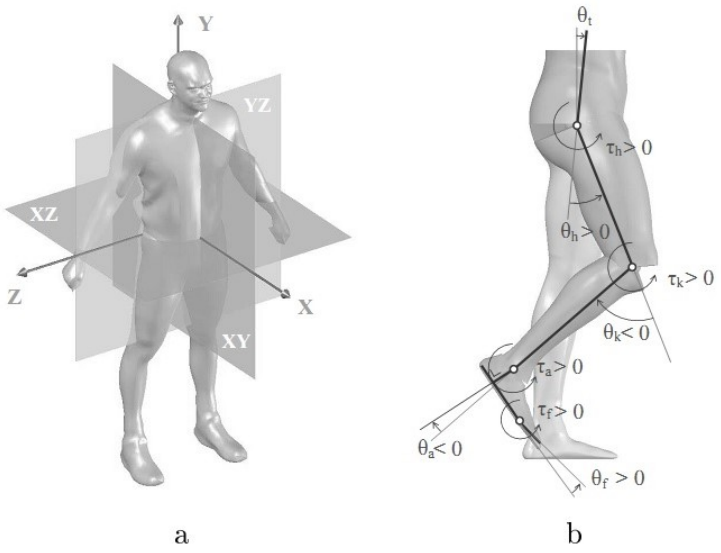


Figure 2.1: Link segment model a) Three anatomical positions in coordinate system, b) Simplified 2D model of the lower limb

ground again. Gait cycle is a combined function of the lower extremity, pelvis, and spinal column. It is a single sequence of functions by one limb. It begins when reference foot contacts the ground and ends with subsequent floor contact of the same foot. A single gait cycle is known as a stride.

Gait cycle begins with heel contact of either foot and ends with the heel contact of the same foot. Therefore, one complete gait cycle consists of two steps one of either right foot and then left or vice versa. By convention, normal gait cycle is the period in which heel of one foot contacts the ground to when the heel contact of the same foot takes place, and forward propulsion of the center of gravity is involved.

A single gait cycle consists primarily of two phases: a swing phase and a stance phase [18]. In general, stance phase begins with the heel contact and ends with the toe off of the same foot. That duration when foot remains in contact with the ground is known as stance phase and accounts for approximately 60% of the normal gait cycle. The duration when the foot is off the ground is known as swing phase and accounts for 40% of the gait cycle. Swing phase begins with the toe off of the delete foot and ends with the heel contact of that same foot.

Figure 2.2 presents the Perry division of gait cycle into two phases Stance and swing and eight sub-phases (five stance and three swing)[18].

Phase 1: Initial Contact (IC): It is considered as the point at which the heel of

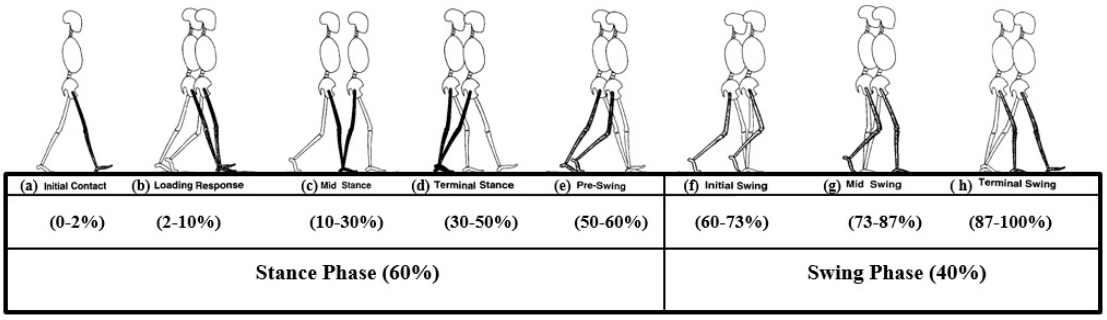


Figure 2.2: Fundamental gait phases with expected interval of gait cycle

the reference foot touches the ground. Thus, it is also known as the heel strike (HS). It is the beginning of the loading response, which constitutes 0-2% of the total gait cycle as shown in figure 2.2(a).

Phase 2: Loading Response (LR): It begins with the initial contact of the reference foot and continues until the other foot is lifted for the swing. During this period the reference foot comes in full contact with the floor and body weight is fully transferred onto the stance limb and accounts for 2-10% of the total gait cycle, as shown in figure 2.2(b).

Phase 3: Mid-Stance (MSt): It begins with the contralateral toe (one opposite the reference foot toe) off and ends when the center of gravity is directly over the reference foot. It begins when tibia of the swing leg is upright to the ground, and its share is 10-30% in the total gait cycle.

Phase 4: Terminal Stance (TSt): Its share is 30-50% in the gait cycle and is initiated by moving into hip extension until just before pre-swing as shown in figure 2.2(d). It is the point at which COG is over the reference foot and ends when the contralateral foot touches the ground.

Phase 5: Pre-Swing (PSw): It accounts for 50-60 % of the total gait cycle as shown in figure 2.2(e). It is considered to begin when the contralateral toe is in Initial Contact and ends with toe-off (TO).

Phase 6: Initial Swing (ISw): This is the first phase of swing and accounts for 60-73% in the gait cycle. In this phase, hip of the reference foot is in, flexion and the corresponding knee stays in flexion mode.

Phase 7: Mid Swing (MSw): This is considered to be the second phase of swing, and in this phase, the flexion of the reference foot knee is at the extreme. It constitutes 73 to 87 % of the total gait cycle.

Phase 8: Terminal Swing (TSw) : This is the phase when the tibia is perpendicular to the ground. Its share is 87-100 percentage in the gait cycle and ends at IC.

Based on the complete gait cycle one can analyze the abnormal gait behavior [19].

2.2 Parameters in gait analysis

Dysfunctional gait can arise from acute or chronic injury or either because of improper biomechanics. Physiotherapists and orthopedists can monitor and analyze gait movement variables. Thus, it is essential to provide a brief discussion of the parameters used in gait analysis. Literature survey suggests that there are six categories of possible parameters used in gait analysis. Muybridge et al.[20] are the first one to study the gait mechanism. They studied gait of a running horse, known as Horse in Motion.

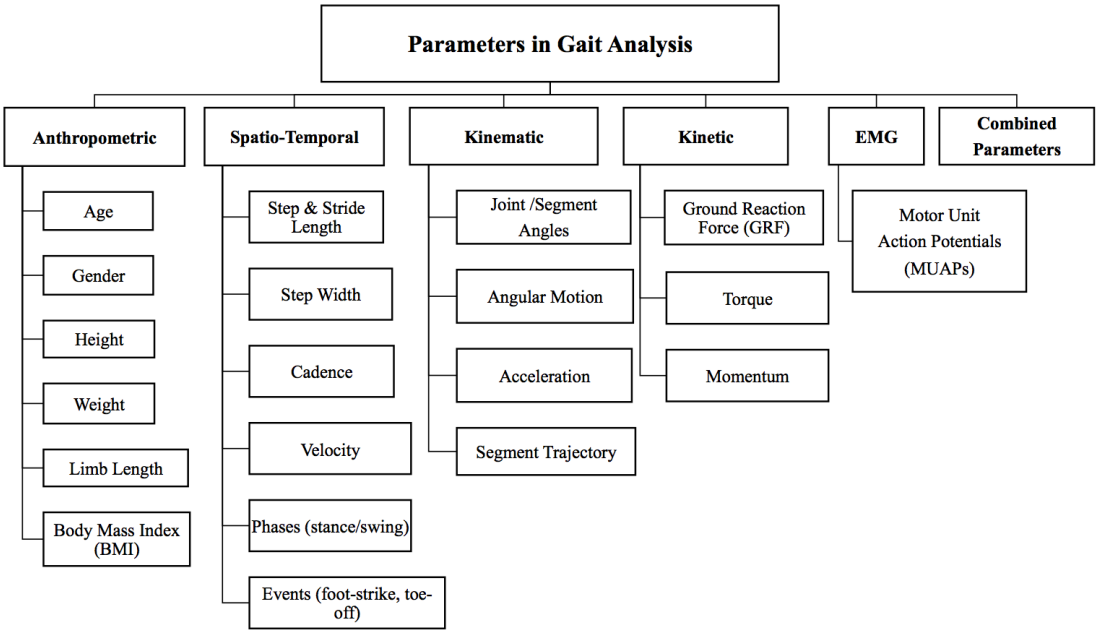


Figure 2.3: Comprehensive parameter tree in gait analysis

Gait analysis includes the measurement of temporal, Kinematics, Kinetics and Dynamic Electromyography(EMG) based parameters. Figure 2.3 shows the comprehensive parameter tree in gait analysis.

Anthropometric Parameters: The anthropometric parameters usually consider corporeal dimensions of the human including age, gender, weight, height, limb length and Body Mass Index (BMI) of a subject. Researchers advocate group based

analysis on the similar anthropometric parameters for gait analysis [1, 21]. It is essential to isolate anthropometric effects on gait analysis.

Spatio-Temporal Parameters: Spatio-Temporal is also known as general gait parameters. They are responsible for providing the simplest form of objective gait evaluation (gait specifications) [1]. Gait analysis system considers time-distance parameters such as step & stride length, step width, cadence, velocity, phases (stance & swing) and foot strike and toe-off events for spatiotemporal. It also includes walking base, toe-out angle, foot contact pattern. Among them, the stride length is the most important and useful gait parameter for both medical and computing field. Table 2.1 lists some of the definitions of terms used in gait analysis.

Table 2.1: Definitions of some of terminologies used in gait analysis

Terms	**Definition**
Body Mass Index (BMI)	It is the ratio of body weight (kg) divided by the square of the body height (m)
Single-Support Duration	Interval when one foot is in contact with the ground.
Double-Support Duration	Interval when both feet are on the ground.
Step Length	Distance between corresponding successive points of heel contact of the opposite feet.
Stride Length	Distance between consecutive points of heel contact of the same foot. In normal gait, it is double of step length.
Cadence	Rate at which a person walks. It is the number of steps per unit time.
Velocity	Distance covered by the body in unit time. Usually measured in m/s.
Phase	Walking can be considered as a set of repetitive components known as phases. Classical gait model by Perry divides gait cycle into 2 phases (stance and swing)
Event	Two Types: Foot strike and toe-off. Foot strike is the point of contact of the foot with the ground while toe-off is the point when the foot is off the ground.

Kinematic Parameter: Kinematic parameters are the measures of movement or geometric description of the motion of body segments. It is the movement of the body in space without any reference to forces. It includes joint angles by considering the motion of body landmark selected for analysis. Along with angle of joints (trunk angle, hip angle, knee angle, ankle angle, etc.), it also comprises angular

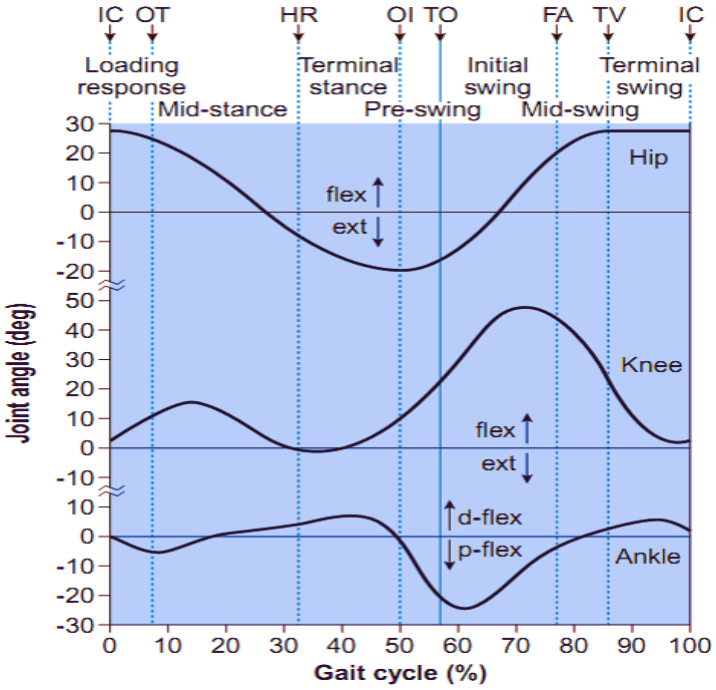


Figure 2.4: Hip, knee and ankle joint angles for a Healthy subject [1]

motion, acceleration, and segment trajectory. Usually, markers and sensors are used to measure these parameters [4, 22]. Electromagnetic systems track coordinates (X, Y, and Z) and orientation are also considered. Figure 2.4 illustrates the normal plot of hip, knee and ankle angle of a healthy subject [1].

Kinetic Parameter: It is the group of forces involved in producing ground reaction forces (GRF), torque, pressure patterns and joint forces. Force platforms are used to measure three-dimensional force components. Researchers consider mass, the center of gravity and momentum [23, 24].

EMG Parameter: It is used to study the muscular activity during walking. Needle or surface EMG electrodes are used to record motor unit action potentials (MUAPs) [25].

Combined Parameters: For better analysis and visualization, researchers attempt to combine the aforementioned parameters such as joint angles and ground reaction forces with the anthropometric measures [26].

Selection of accurate gait parameters is very crucial in gait analysis as the outcome of research depends greatly on choosing the most appropriate gait features. The factor of interest among the discussed parameters depends on the field of research for example in the field of sports; EMG parameter is best suited to analyze force

applied on muscle.

2.2.1 Normal Gait

Physiotherapists, orthopedists, and neurologists use quantitative gait parameters (stride length, step length, cadence, stance and swing phase, etc.) for better realization of patients gait pathology [15, 27]. These gait parameters are compared to age and sex-matched normal gait population distributions to decide whether the patient is exhibiting normal representation or not.

Parameters (Self-Selected Pace)	Young (1-7)	Adult	Elderly (> 65)
Walking Velocity (m/s)	0.64-1.14	1.30-1.46	Declines 15% per decade
Stride length (m)	0.23-0.57	1.68-1.72	1.66-1.70
Step length (m)	0.20-0.32	0.68-0.85	0.44-0.60
Stance phase (s)	0.32-0.54	0.62-0.70	0.68-0.72
Swing phase (s)	0.19-0.27	0.36-0.40	0.42-0.44
Cadence - Fast Walking (steps/min)	176-144	113-118	58-70
Single-Support (% of stride)	64.4-65.6	60.6-62.0	61.7-62.9
Double-Support (% of stride)	22.5-23.9	21.2-23.8	23.4-25.8

Figure 2.5: Healthy gait parameter for different age group

Defining normal gait pattern is a very tedious task. Research suggests for group-wise analysis to find the normal pattern. Figure 2.5 presents healthy gait parameters for different range proposed in the literature. Ranges are indicative of the subject population and do not necessarily hold for the general population. It may vary from region to region [28].

2.2.2 Pathological Gait

Dysfunctional gait prohibits normal weight-bearing competence on the bipedal and influences stresses placed on joint surfaces. Lai et al. suggest clinical and elderly gait disorders as two types of common gait disorders [28]. Clinical gait disorders include limping movements such as antalgic gait patterns arise due to lower segment pathologies such as knee osteoarthritis tendon rapture and patellofemoral pain syndrome. The pain is due to abnormal weight-bearing competence either on ankle or knee joints. Patients exhibit a short stance phase. The second type of disorder is apraxic gait which, arises due to deterioration of neurons(nerve cell) and motor system responsible for human locomotion control. Alzheimer, Parkinson, Cerebral Palsy (CP), Freezing of gait and dementia are the examples of this

apraxic based gait pathology and is characterized by loss of ability to move properly [29, 30]. Ataxic gait is considered as the third general disorder where there is a loss of sense of relative limb positions and characterized by unsteady and uncoordinated walking feet pointed outward. It is due to chemical intoxication such as medial or alcohol or cerebral disease.

Nearly 50% of people over age 65 have gait problem. Senile gait pattern is the best example of gait disorder arising due to normal aging [31, 32]. Subject exhibits broad stance and reduced gait velocity. Frontal lobe gait pattern arises due to injury in frontopontocerebellar tract Cerebrocerebellar system and Arnold's bundle and is characterized by balance disorders.

In sensory ataxic gait pattern, sensory inputs are averted to the brain. This pattern results in the decrease in visual activity, proprioception and depth perception. Also, elderly people exhibit waddle gait patterns due to either back or pelvis muscle breakdown. Such type of pattern is known as myopathic gait. The elderly population is faced with gait disorder progression, which increases the risk of death precipitated by falls and bone fractures.

2.3 Human Gait Analysis Approaches

Gait analysis is not a new research area. A systematic study of gait started with the description of walking principle by Leonardo da Vinci, Galileo, and Newton. In 1682, a student of Galileo, Borelli described how balanced walking could be achieved using the concept of center of gravity of the body. In 1878, Edward Muybridge and Leland Stanford were the first ones to study the gait mechanism [1]. Even though considerable research has been done in gait analysis area, still it has not yet been fully utilized. With the evolution of new models and methods, gait analysis research is an ongoing activity. Since 1960, clinical gait analysis gained momentum with the advent of new computational techniques. The researcher has proposed numerous approaches for the gait analysis [9].

Contemporary human gait analysis approaches can be broadly classified into four types: a vision based or image processing based using a video camera, sensor-based, and other technologies and combined approaches as shown in figure 2.6.

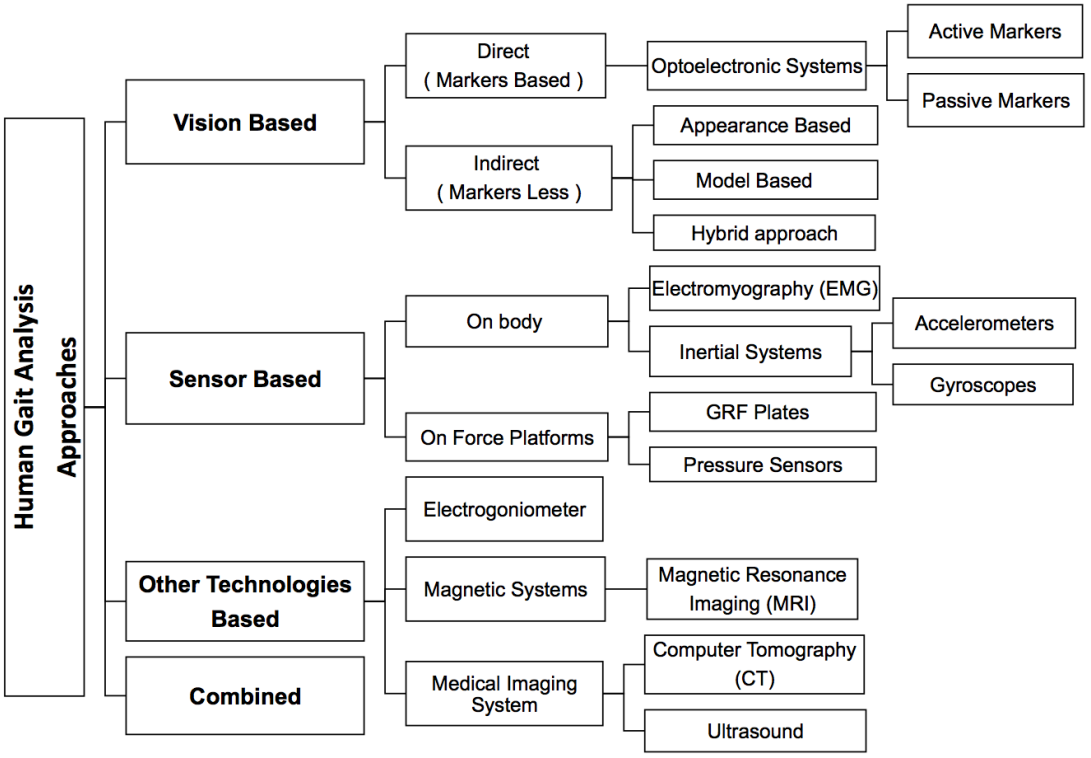


Figure 2.6: Available gait analysis approach: A comprehensive tree diagram

2.3.1 Vision Based Approach

The video camera is used to capture frames in vision-based gait analysis approaches. This analysis can be accomplished by two methods; with (direct) or without markers (Indirect) on the subject. Analog or digital cameras are used to analyze gait. Vision-based approaches are used to recognize gait, segment position and gait phase detection as presented by [1, 4, 11, 33].

The marker-based approach is also known as direct vision-based approach [34]. It is a popular method to find the kinematics of a human. Active or passive markers are developed to perform real-time direct vision based gait analysis[35]. Direct gait analysis technique uses an optoelectronic system. Optoelectronic systems convert light signals into electrical signals and track the light emitted or reflected by markers. Active markers emit a signal and come in the form of small light-emitting diodes (LEDs) that can be attached to the subject. The cameras send out infrared light signals and detect the reflection from the markers attached to the body. This signal is used to locate the position of the marker as individual markers work at predefined frequencies. The accuracy of these systems is commendable, given the fact that they can often locate markers uniquely within a distance as small as 1 mm. However, such systems have their disadvantages such as the subject's nat-

ural motion is hindered due to the presence of cables or other components that can affect their gait pattern. Passive markers reflect light emitted by the cameras back to them, and near-infrared light is used to illuminate the markers. Passive markers are spheres covered with reflective scotchlite™tape. They are specially designed to reflect incident light directly back along its line of incidence [8, 36].

The researchers propose numerous viewpoints for observation of gait abnormality, but all agree on two best views (sagittal or lateral and frontal or back) for observation of gait abnormality. For 2D and 3D analysis, many camera-based motion capture systems are available in the market [8, 37, 38]. One camera is decently enough for two-dimensional analysis. The camera is positioned in sagittal view (perpendicular to the walkway), but this approach has a limitation that it is unable to capture all out of plane motion. Three-dimensional analysis can overcome this limitation, but in that, there is a need for more than one camera. To project the three-dimension model, a point of interest region of the subject should be visible by at least two cameras simultaneously [39]. Across the world, leading gait labs use four to eight cameras.

The indirect approach in vision-based gait analysis doesn't use markers. From the video captured, through video camera either based on the model, appearance or hybrid approach can be used to analyze the features of the subjects [40, 41]. This is the most common approach to recognize human based on the human walking trait in surveillance.

2.3.2 Sensor based Approach

Gait analysis can be performed, by using the sensor, placed either on the subject body or, on the floor [24, 42, 43]. EMG is used to study the muscle electrical activity during walking, and it can be used to detect gait phase. Needle or surface EMG electrodes are used to record Motor Unit Action Potentials (MUAPs) [44]. Even for a single movement, a group of muscles is involved. The amplitude of EMG signals derived during gait may also be interpreted as a measure of relative muscle tension but there is a need for specific knowledge on electrode setup, and they are sensitive to interference.

Inertial systems are based on a principle of resistance towards change in motion. Accelerators and gyroscopes are used to measure inertia and segment orientation respectively. The sampling rates used for gyroscopes are similar to that used in accelerometers. Some researchers propose to use gyroscopes with accelerators to

get the kinematics of the subject's movement [45, 31]. They can be used to find segment position, step detection, and stride length. It is sensitive to interference, and the algorithms are complex.

Floor platform based sensors are used to obtain forces involved in producing ground reaction force, force pattern, foot plantar pressure distribution, step and gait phase detection [22, 12]. Ground Reaction Force (GRF) plates are used to calculate force magnitude and direction when the foot contacts the force plate. Force applied on GRF plates is sensed by transducers (with steel plate) attached to each corner of the plate. This force is converted to an electrical signal, and it is used to calculate the center of pressure, three orthogonal force component of the subject's movement.

Pressure sensors are used to get the load details applied on sensors [13]. They are placed inside insole of shoes. When mechanical strain is applied to these piezoelectric-based sensors, they generate an electrical signal. Healthcare professionals can identify the gait phase with these methods, but there is a limitation of space, in these techniques. For correct measurement, subjects need to keep their foot on the center of a plate of floor platforms; this makes the subjects conscious, and they are not able to exhibit their normal pattern. Another limitation is regarding cost. Even the well-equipped gait labs have only two or three plates at maximum. Thus healthcare professionals are not able to capture the regular pattern of the subjects.

2.3.3 Other Technologies

Other technological based approaches such as Electrogoniometer, magnetic and medical imaging based systems can be used as a human gait analysis approach. By evaluating the change in resistance of potentiometer with two rotating arms in Electrogoniometer, one can perform the analysis of the joint angle change and step detection. There should be precise calibration of potentiometer before using it for analysis. This Electrogoniometer based approach is inappropriate in a time-constrained environment as it consumes time to attach to a subject [1, 8].

Magnetic system based approach doesn't need a line of sight for the markers as is required in vision-based marker approach as it operates using a magnetic field to track ferromagnetic markers. Movement of the segment and anatomical data of subjects segment can be obtained by using Magnetic resonance imaging (MRI), computer tomography (CT) and ultrasound. It is then used to customize

a computational model of the subject to which kinematic and kinetic data can be applied [46, 47]. These systems are also sensitive to interference.

2.3.4 Hybrid Approaches

Besides these approaches, there is also another hybrid approach that uses a combination of two or more aforementioned approaches. Baker in 2006 used both the vision and sensor-based techniques for the clinical rehabilitation [48]. Similarly, Owusu implements hybrid approaches for the sports tactics analysis[7]. Recently in 2017 Benjamin et al. proposed machine based human emotion perception using the hybrid approach (vision and sensor-based) Researchers have used vision and EMG and force platform for better analysis of human gait [46, 49, 50, 51, 52, 31].

2.4 Application domain

Early work by Vasconcelos et al. discusses the applications domains in gait analysis [9]. In this section, we discuss the state of the art within the possible application of gait analysis. Gait analysis applications can have five grouped: Analysis, Biometric, Artificial gait, Control based and other application as shown in figure 2.7.

2.4.1 Analysis based Applications

Identification of normal gait pattern, medical diagnostic, geriatric or older person care and sports monitoring & tactics are part of the gait analysis based application. Health-care professionals can utilize the gait phase recognition concept in their routine practice to identify the abnormality [26]. Extrinsic factors (terrain, footwear, clothing, and cargo), intrinsic (sex, weight, age), physical, psychological (a type of personality and emotion) and pathological factors (such as neurological disease, trauma, musculoskeletal and psychiatric) influence Normal walking [52, 13]. Even a single individual retains a wide-ranging gait behavior, thus determining this normal gait parameters range is a very complex and intricate task. It is difficult in pathological gait diagnosis [53].

The technique that deals with the identification of hidden impairments and affects gait patterns is known as Clinical Gait Analysis (CGA). Identification of pathological gait is the most unswerving implementation of gait analysis. Clinicians

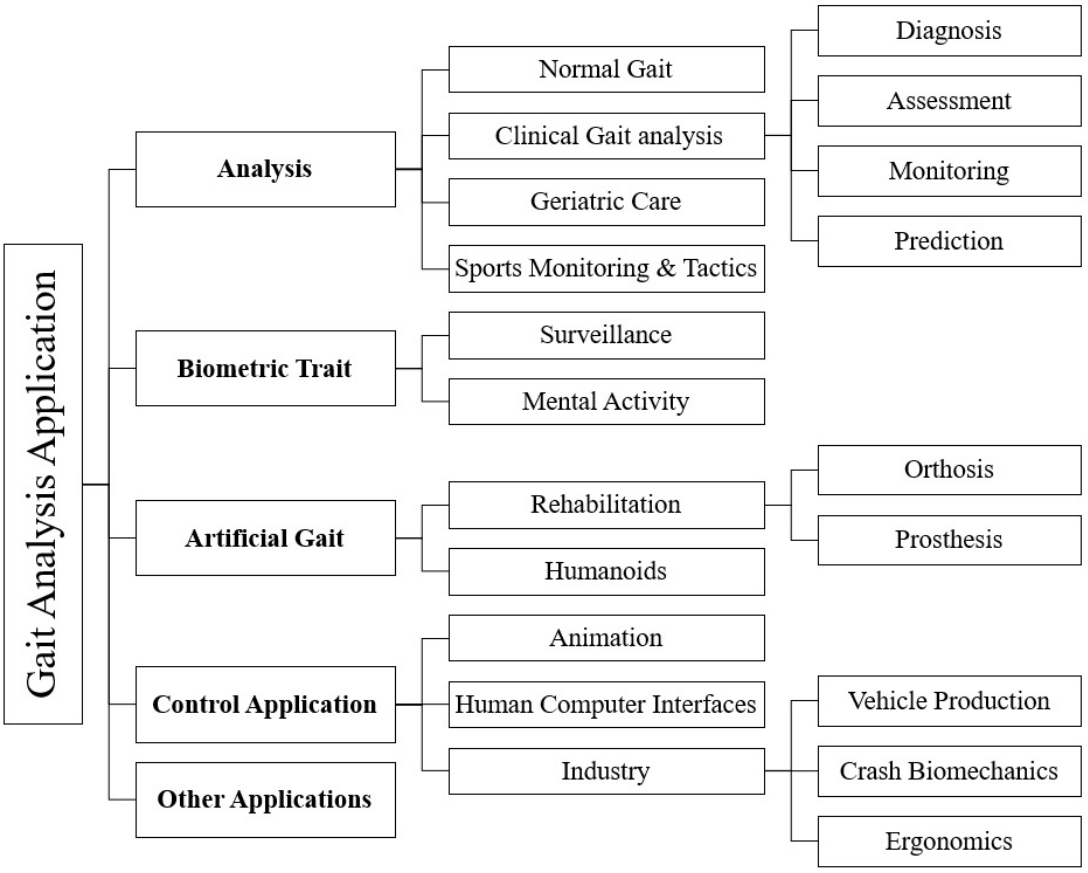


Figure 2.7: Application of human gait analysis

can utilize gait segmentation concept in their routine clinical practice to evaluate a patients status, treatment, and rehabilitation for complex musculoskeletal and neurological disorders using the spatiotemporal and kinematics parameters. Healthcare Professionals can monitor the progress of the patients by comparing with the standard norms gait pattern and their records. By comparing the normal and current gait cycle phase pattern, healthcare professionals can suggest effective treatment. Gait analysis can be implemented to cure pathologies such as Cerebral Palsy (CP), Parkinson's disease (PD), dementia stroke, patellofemoral pain syndrome, knee osteoarthritis and tendon rapture, Freezing of gait [54, 55, 56, 29, 30].

It can be used as an appliance for kinesiology for movement disorder. Problems of postural defects and walking Gait in elderly adults (also known as senile gait) have become more substantial and of greater concern [52].

Each year approximately 33% of Australians aged over 65 years experience a fall and half this number fall more than once. Rapid growth in the proportion of the population over 65 years will accelerate falls-related healthcare costs. This results in loss of independence and increase dependability and thus increase overall cost

for public heath and social service of a country. Therefore, scientifically diagnostics of the poor postural condition is receiving more and more attention from health, medical and sports academy in the past decades since it can detect an incipient fault at an early stage. Though no authenticated statistics for elderly Indian population can be found, it can be safely presumed that situation is no more different in India as increased urbanization and nuclear family concept which has led to the breakdown of traditional healthcare of elderly as discussed by [6, 57].

In injury, gait analysis can aid in selecting the best available treatment options [58]. [59] discusses that movement analysis of sports player prevents possible sports injury. Silva et al. demonstrate that wearable sensors based analysis can effectively prevent many injuries from overuse or incorrect posture and motion in games [60]. [61] discuss the use of Artificial Intelligence in sports such as cricket, soccer, short putters, golf, baseball, running, jumping and weightlifting.

2.4.2 Biometric Trait

Biometric systems are based on identifying and classifying individuals using biological traits. Biometric systems use fingerprints, palm, iris, face and voice recognition but are not limited to only these traits [62]. The aforementioned biometric systems, have certain weaknesses to counter including unreadable fingerprints, iris scans obscured by eyelashes, eyelids, reflections from cornea or face obstruction due to glasses, different hairstyles, and facial expressions.

One issue with these systems is the fact that people are concerned while touching the scanning device or waiting in line for using the biometric person recognition system. People often feel that they can identify a familiar person from afar simply by recognizing the way a person walks. This common experience, combined with the recent interest in biometrics, has to lead to the development of gait recognition as a form of biometric identification.

Gait recognition is an emerging biometric technology, which involves people being identified purely through the analysis of the way they walk. It has recently attracted interest as a method of human identification because it is non-invasive and does not require the subject's cooperation. Gait recognition could also be used from a distance, making it well-suited to identifying perpetrators at a crime scene. Also, since this approach only analyzes a person's walking style, it can also work on masked subjects.

In today work identification of abnormal activity in crowd play a vital role in avoiding terrorist attacks. Gait recognition is an emerging biometric technology, which involves people being identified purely through the analysis of the way they walk. It has recently attracted interest as a method of human identification because it is non-invasive and does not require the subjects cooperation [63, 64]. Presently the methods being used for gait recognition can be broadly divided into two categories model based and model free [33, 65, 38].

Recently various computational techniques have been proposed to analyze the actions, activities and behavior of crowd and individuals [66, 67, 41]. Abnormal activities can be monitored to alert the related authority for potential criminal activity using vision-based gait analysis [34]. Gait analysis has a huge potential for identifying the psychology (mental activity analysis) of the individual without contacting the concerned individual [45, 68]

2.4.3 Artificial Gait

The desire to improve the quality of life and reliability of rehabilitation drives new research and development activities in several countries. Researchers are evaluating the causes and consequences of various faults in postural deformities and walking gait conditions. Gait analysis can supply useful information about kinematics and kinetics of human walking to be used in the design and manufacturing of humanoid robots. Now, with the new techniques, clinicians can exploit instrumental gait analysis as their regular clinical practice to assess a patient's status, healing, and recovery for complex musculoskeletal and neurological disorders. Treadmills based rehabilitation systems are used by many hospitals for walking on a level surface or climbing stairs. Lokomat developed by Jezernik et al. has four degrees of freedom (DoF) [69]. HapticWalker developed by Schmidt et al., Stair Master, GaitMaster are some of the examples of robotic rehabilitation systems [70]. The walking speed and weight support can be adjusted automatically as per the severity of the disease in case of robotic rehabilitation systems. Gait studies have contributed significantly to the design of artificial limbs (prostheses and orthosis) for amputees and inspired artificial locomotor controllers used in Exoskeletons and robotics. It can be used to evaluate orthosis and prosthesis alignment to the lower limbs as discussed by Sigal et al. [71]. Juang and Jih-Gau proposed a robotic gait synthesis model based on kinematics using the fuzzy neural network [72]. Kim et al. explore Hybrid genetic algorithm and neural network for joint angles prediction based on accelerator data [73].

Gait analysis is important in normal gait behavior of humanoids. Numerous artificial intelligence based algorithms is available for the same [48, 74, 75]. Popovic et al. proposed design of control for a neural prosthesis for walking using artificial neural networks [76]. Heinen et al. proposed a humanoid gait optimization approach that automatically creates and controls stable gaits for legged robots into a physically based simulation environment [50]. [77] explores reinforcement learning with experience for humanoid gait optimization.

2.4.4 Control based Applications

Computational techniques and gait parameters can be used to control events in the entertainment industry through human-computer interfaces [78, 79]. Animation industry also uses the gait analysis concept to pose parameters to make it more realistic [80]. Based on mood detection from gait analysis, one can control room environment. Liu et al. proposed an advance mixed reality human interaction approach for better learning and playing experience for children at Singapore Science Center [78]. These approaches can play a major role in live interaction based applications such as virtual reality and video games.

The industry has much potential to utilize gait analysis in vehicle production, crash biomechanics, and pedestrian detection and to design a comfortable place to work in, using ergonomics techniques. Gait analysis is used in the automotive industry for the proper placement of handles on doors in cars for better customer sanctification. The work on designing for ergonomics, proposed by Webber et al. has been followed by many researchers and the industry in recent years [81]. Mavrikios et al. used virtual reality and simulation models to improve customer satisfaction on ergonomics in automotive industry [82]. Untaroiu et al. proposed vehicle-to-pedestrian crash (VPC) model, for analysis of pre-impact pedestrian posture using various optimization techniques [83].

2.4.5 Other Applications

There are numerous other possible applications of gait analysis based on human activity individually and in the group as well. Abnormal activity in a crowd or at home can play a vital role. Recently researchers also explored harvesting energy during human movement [84, 85]. Raziel et al. analyze human gait cycle using EMG, Qualysis Tracking Markers (QTM), WinDaq, insole pressure sensors

and V3D to provide the basic understanding of human locomotion to design a mechanism to harvest energy from heel strike [86]. The field of gait is still not fully explored to its full potential.

2.5 Machine Learning Techniques

The objective of machine learning techniques is to design algorithms that learn either from experience in the form of the labeled data or automatically discover useful patterns from given data-points. Approaches such as Statistical and Machine Learning techniques have been used for gait data representation and classification [28]. Statistical techniques can be used in gait models, the study of the effects of various independent variables on dependent variables. It can be used for gait kinetic as discussed in [87, 88]. Multivariate statistical techniques such as Principle component analysis (PCA) and Linear Decrements Analysis (LDA) can be used to represent gait data for find analysis such as linear relationship [89] and surveillance [90, 91]. These techniques are known as data reduction techniques. But when the nature of the problem is nonlinear or complex, they are not a good choice.

Machine learning techniques can be further subdivided into Supervised (classification based), Unsupervised (clustering based), Reinforcement, Rule-based, Evolutionary (Genetic Algorithm, Particle Swarm Optimization), Probabilistic and Hybrid approaches. Table 2.2 lists the use of Machine Learning techniques highlighting potential applications.

Machine Learning techniques such as supervised (Neural Network, K-Nearest Neighbor (k-NN), Supporting Vector Machine (SVM), Ensembles (Bagging, Boosting, Random Forest) , etc) ; clustering based (Self-Organizing Map, k means, Fuzzy c means, hierarchical), reinforcing learning, fuzzy logic, and Evolutionary approaches are used for gait analysis by Wright et al. [92] as shown in table 2.2. It highlights machine learning techniques with potential gait application.

2.5.1 Supervised Learning

Supervised Learning is considered as a task driven approach where we have input and desired output. A mathematical model is created to map the inputs to the desired outputs. Unseen data should be assigned a class as accurately as possible

based on this model. The main objective is to minimize the risk or error. In the gait analysis, the data has to be labeled by healthcare professionals. Neural Networks, Radial Basis Function (RBF), Ensembles (Bagging, Boosting, Random Forest), Decision Tree, K-Nearest Neighbor (k-NN), Support Vector Machine (SVM), etc. are included in Supervised Learning.

The neural network is an attempt to computationally replicate the working of biological neural networks. The earliest of the research on mathematically modeling the neurons is carried out by Warren McCulloch and Walter Pitts in 1943 to formulate artificial neurons. Since then various more sophisticated models of artificial neurons have been proposed by researchers. The neural network proceeds until it reaches to a certain minimum error or predefined epochs. Neural Network is most widely used for Normal Gait Analysis, Robotic Rehabilitation, Sports Monitoring and Tactics, Geriatric Care Surveillance, Activity Recognition as discussed by [46, 49, 51, 93, 94]. A learning model is developed using the existing data through training and then to validate the same, testing is carried out with either the unknown or known data. For validation the K-fold validation, leave one out etc approaches are used.

Radial basis function (RBF) is know to be the simplest variant of the neural network. Joshi et al. use the spatially localized basis function in feature space for disorder detection [29]. Naive Bayes is a probabilistic approach used in case of availability of prior knowledge and is thus considered as a supervised learning approach. It assumes the Gaussian distribution of values for each class. From a given set of training data, it learns the conditional probability of each feature $x_i \in$ X having class name c $\in$ class label set C. Zhao et al. use naive Bayes for prosthesis intention detection [75]. Ensembles are used for clinical analysis and prediction model.

Support Vector Machine (SVM) is based on the concept of optimally separating hyperplanes and part of supervised learning [95]. It is a non-probabilistic method that projects the input space onto a high-dimensional feature space non-linearly. The SVM is a robust classifier that essentially uses a kernel function to define the distribution of data points inside the feature space. Authors such as [49, 32, 94, 89, 31, 84, 96] uses SVM for study of normal and age-related differences, normal and abnormal gait pathology. SVM is a powerful classifier for small to the medium dataset.

Supervised learning techniques is a dominant methodology when labeled training data points are available.

Table 2.2: Machine learning techniques highlighting potential gait application

Technique	**Learning**	**Potential Application**	**Reference**
RBF	Supervised	Disorder Detection	[29]
Naive Bayes	Supervised	Prosthesis- Intention Detection	[75]
Neural Networks	Supervised	Normal Gait Analysis; Robotic Rehabilitation; Sports Monitoring and Tactics; Geriatric Care Surveillance ;Activity Recognition	[51, 46, 94, 61, 68]
Ensembles (Bagging, Boosting, Random forest)	Supervised	Clinical Gait Analysis	[29]
Decision trees	Supervised	Disorder Detection ; Prediction - Model	[53]
K-Nearest Neighbor	Supervised	Surveillance	[93]
Support Vector Machine (SVM)	Supervised	Normal /abnormal gait pathology ; Gender and Age-Related Differences ; Geriatric Care ; Normal Gait Analysis; Clinical Gait Analysis	[102, 55, 31, 84, 94, 89]
Self-Organizing map (SOM)	Unsupervised	Clinical Gait Analysis	[98]
K-means Clustering	Unsupervised	Clustering of Disorders	[103]
Deep Learning	Semi-Supervised	Human Pose Estimation; Human and group Activity	[104, 105]
Markov Decision Process	Reinforcement Learning	Humanoid Gait Optimization; Gait Training	[106]
Fuzzy Logic	Rule Based	Inexact Gait Classes ; Muscle activity pattern model; Phase identification; Prosthetic control	[54, 107, 13, 108]
Genetic Algo	Evolutionary	Gait Feature Selection ; Heel-toe running model ; Humanoid Gait Optimization; Surveillance	[50, 109]
PSO	Evolutionary	Optimal Classifier Selection; Markerless Motion Capture.	[110]
Nave Bayesian	Probabilistic	Artificial Gait	[75]
Gaussian Process Regression	Probabilistic	Kinematics predication model	[111]
HMM	Probabilistic	human physical activity	[45]
Fuzzy c-mean	Hybrid	Multiple Gait disorder Detection	[112]

Table 2.2-*Continued from previous page*

Technique	Learning	Potential Application	Reference
Neuro Fuzzy	Hybrid	Artificial Gait	[72]
Fuzzy SVM	Hybrid	Gait monitoring platform	[113]
Self Organisation ANN	Hybrid	Clinical Gait Analysis	[114]

2.5.2 Unsupervised Learning

In unsupervised learning, no labeled examples are available to train the model. In this learning, agent attempts to find the similarity among the given data points each having a set of attributes and a similarity measure. Based on this similarity measure, each of the data points is grouped in a cluster. The number of clusters is predefined. The similarity measure can be taken to be Manhattan, Euclidean, Minkowski, cosine distance, etc. The objective is to minimize the intra-cluster distance and maximize the inter-cluster distance. K-mean, Fuzzy k-mean, Hierarchical Clustering, Self-Organizing Map (SoM)are some of the examples of clustering techniques. Clustering techniques are most widely used for categorizing and classifying gait data of disorders based on some common features of the disease [97, 98]. These techniques help in diagnosing gait abnormalities or classification of everyday physical activities. OMalley et al. explored the fuzzy clustering approach on cerebral palsy children based on temporal distance parameters [99]. Xu et al. explore k mean, SOM and Hierarchical Clustering to differentiate Normal and pathological gait pattern based on stride length and cadence [100]. Phinyomark et al. use hierarchical clustering for clustering healthy group from pathological patients [101].

2.5.3 Reinforcement Learning

Reinforcement learning is very similar to human learning without a teacher. It is motivated by behaviorist psychology of human, where an agent learns to take actions based on the experience so that it can maximize the rewards, which guides the learning algorithm. The major application of reinforcement learning is in humanoid gait optimization as presented by Wawrzyski [77]. Reinforcement learning paradigm optimizes humanoid step-size using actor-critic experience learning. Hasson et al. also use reinforcement learning for gait training with directional error feedback rehabilitation [106]. The agent on the treadmill learns gait patterns

using reinforcement learning and this experience can be extended to walking over ground.

2.5.4 Rule Based Learning

Fuzzy logic is best suited for the representation of information extracted from inherently imprecise data. The fuzzy logic approach is considered to be a rule-based approach. Fuzzy logic handles imprecision, vagueness, and insufficient knowledge. Fuzzy logic can work in this scenario with reasoning algorithms to simulate human reasoning and judgment making capability in machines. These procedures let researchers build intelligent systems in the areas where data cannot be represented in binary form. Fuzzy logic lets intelligent systems perform optimally with uncertain or ambiguous data and knowledge. Gait phase identification activities are often vague or based on intuition, as one can't differentiate all phases. Skelly et al. explore fuzzy logic for the prosthetic controller using electrically stimulated locomotion for paraplegic subjects [54, 19]. Kyoungchul implemented the fuzzy-based approach to detect gait phases from foot pressure patterns. The outcome highly depends on the selection of optimal membership value and function, manipulation of linguistic operators and rules.

2.5.5 Evolutionary Learning

Evolutionary algorithms (Genetic Algorithm (GA), Particle Swarm Optimization(PSO), etc.) can play an important role in optimization based problems. The concept of these algorithms is that they are motivated by biological evolution, such as selection, mutations, crossover, etc. A fitness function is evaluated and based on the problem it is either minimized or maximized. Optimization can be applied to feature selection in gait; the relationship among gait parameters, surveillance, heel-to-running model and optimal classifier selection [50, 110]. Slow and uncertain convergence and high computational complexity are some of the limitations of evolutionary approaches.

2.5.6 Probabilistic Learning

The probabilistic model is used to express noise and uncertainty using mathematics of probability. Hidden Markov Model (HMM), Gaussian process regression and

Bayesian are examples of probabilistic models used in gait analysis for human physical activity recognition, kinematics predication model and artificial gait in [45, 75, 111] respectively.

2.5.7 Hybrid Learning

For better feature recognition, two or more machine learning techniques are combined for pattern recognition. This new technique is known as Hybrid learning. The combination of the neural network with the rule-based paradigm (Neuro-fuzzy and fuzzy-neuro model) is proposed by Juang et al. for creating artificial gait [72]. Zhang et al. combine rule-based, and k means algorithm (fuzzy c-mean) to identify multiple gait disorder detection [31]. Fuzzy with support vector machine (FSVMs) and self-organization ANN are proposed for clinical gait analysis as listed in table 2.2. Researchers combined supervised and unsupervised and constructed a semi-supervised paradigm. With the help of limited labeled data, the model is trained for better results. This is useful when the labeled dataset is either expensive or scarce.

Human crafted feature representation limits the machine learning techniques performance. In the past decade, researchers have designed advances form of neural networks known as deep Learning techniques. It constructs feature from raw data and goes with an end to end training. Deep neural nets with a large number of parameters and hyperparameters are very powerful machine learning systems. Convolutional Neural Network (CNN) is a form of deep learning used for classification problems for image, video, Nature Language Processing (NLP), etc. Recently deep learning has been used in monitoring crowd flux, flow and congestion analysis.

Gait analysis has a huge potentiation in identifying the psychology (mental activity analysis) of the individual without contacting the concerned individual [45, 115]. It is not yet fully explored for the gait analysis. Table 2.3 highlights the field of gait analysis, pointing out the purpose and implementations of different AI techniques with input gait parameter and approaches adopted have been added with an additional column of remarks.

Table 2.3: Applications based machine learning techniques

Author	Application	Machine Learning	Input Parameters	Approach	Remarks
[51]	Normal Gait Analysis	NN	Kinematics and EMG	Hybrid	Gait Muscle Modeling
[49]	Normal Gait Analysis	NN Back-propagation with Bayesian Regularization (BR)	Spatio-temporal, GRF and Joint angular	Hybrid	12 young and 12 elderly; Change in gait due to ageing from their respective gait-pattern characteristics
[55]	Normal Gait Analysis	SVM	GRF and kinematic features	Hybrid	14 normal & 13 Patellofemoral pain syndrome patients; Automatic recognition of gait pattern
[107]	Normal Gait Analysis	Fuzzy	Kinematic	Hybrid	Prediction Model - Muscle activity pattern from kinematic parameters
[53]	Normal Gait Analysis	Probabilistic graphical model, EM algorithm, Decision tree - HMM CRF	Anthropometeric	NA	Prediction Model - From segment length predict kinematic parameters; Correlated static(leg length, etc) -dynamic(actual figure) model
[89]	Normal Gait Analysis	PCA + SVM	Kinematic	Vision Based	483 subjects on treadmill ; Gender and Age-Related Differences during walking
[111]	Normal Gait Analysis	Gaussian process regression	Anthropometeric	Vision	Prediction Model - 12 body parameters to gait kinematic; 113 healthy subjects (50 males and 63 females
[108]	Normal Gait Analysis	Fuzzy	kinematic	Vision Based	Fuzzy based gait phase identification
[54]	Clinical Gait Analysis	Fuzzy Rules	GRF	Sensor Based	2 insole FSRs are used for gait event detection. Heel strike and toe off detected in real time
[103]	Clinical Gait Analysis	k- Mean and Hierarchical Clustering	stride length and step frequency/cadence	Vision Based	Diagnosis - Normal or pathological gait pattern

Table 2.3-*Applications based Machine Learning Techniques, Continued from previous page*

Author	Application	ML	Input Parameters	Approach	Remarks
[114]	Clinical Gait Analysis	Self Organization ANN	Kinematic and Kinetics	Hybrid	129 Gait quality assessment
[97]	Clinical Gait Analysis	Hierarchical Clustering	Kinematics	Vision Based	Clustering CP children from sigittal vision
[96]	Clinical Gait Analysis	SVM, KStar (KNN) and Random Forest	5 features - kinematic	Sensor Based	Assessment neuro-degenerative diseases - 15 Amyotrophic lateral sclerosis,20 Parkinson's disease, 13 Huntington's disease and 16 normal
[116]	Clinical Gait Analysis	SVM	Spatiotemporal , GRF	Hybrid	6 healthy , 11 - patient undergo surgery; Monitoring recovery from knee replacement surgery
[29]	Clinical Gait Analysis	Decision tree , Bagging, BF tree, RF,RBF networks, and Multilayer Perception	11 attributes such as genes, age, Alcohol, BMI	Hypertension, diabetes,Family history etc.	487 ADRC dataset - for classification of Alzheimer's and Parkinson's disease
[13]	Clinical Gait Analysis	Fuzzy logic	Limb accelerations and 3D GRFs	Sensor Based	10 healthy &4 unhealthy Relapsing remitting multiple sclerosis
[32]	Geriatric Care	SVM	Kinematic Variable	Vision Based	17 young and 17 elder persson treadmill - Age related changes
[30]	Geriatric Care	supervised and unsupervised features + PCA	3D accelerometers	Sensor Based	DAPHNet datase ; Freezing of Gait - detection and Prediction
[84]	Geriatric Care	SVM - Multi-class SVM	Images	Vision Based	UCF-Sports dataset; Activity Recog
[94]	Geriatric Care	NN - BP ANN , SVM	Accelerometers and Gyroscopes data	Sensor Based	322 subjects, Developed Wireless Gait Analysis Sensor ; used for fall detection

Table 2.3-*Applications based Machine Learning Techniques, Continued from previous page*

Author	Application	ML	Input Parameters	Approach	Remarks
[102]	Geriatric Care	SVM	kinetic	Sensor Based	60 persons - identify the small significant difference between lower limbs (left and right)
[61]	Sports Monitoring and Tactics	Neural Network and Evolutionary Computing	N/a	na	performance of cricket bowlers, soccer players and shot putters
[46]	Sports Monitoring and Tactics	Neural Network	Kinematic and Kinetics	Hybrid	Monitoring; Sports Model using NN
[60]	Sports Monitoring and Tactics	Neural Network	Kinematic and Kinetics	Sensor Based	Monitoring; AI in Weight lifting training
[76]	Robotic Rehabilitation	Neural Network	sensors	Sensor Based	Design of a control for a neural prosthesis for walking
[73]	Robotic Rehabilitation	Hybrid Genetic Algorithm + Neural Network CPG	kinematic from accelerometer	Sensor Based	from sensor data create joint angles
[75]	Robotic Rehabilization	Bayesian classification and clustering, C4.5	Kinematic and Kinetics	Hybrid	Movement prediction and judgment of lower limb
[72]	Robotic	Fuzzy Neural Network	Kinematic	Hybrid	Model for robotic gait synthesis
[50]	Humanoids	GA	Joint angular velocity for each joint	Hybrid	Humanoid Gait Optimization- Automatically create and control stable gaits for legged robots
[77]	Humanoids	Reinforcement Learning	sensor	Sensor Based	Humanoid Gait RL Optimization
[117]	Surveillance	Genetic Fuzzy SVM	Silhouette	Vision Based	70; Gait Recognition

Table 2.3-*Applications based Machine Learning Techniques, Continued from previous page*

Author	Application	ML	Input Parameters	Approach	Remarks
[90]	Surveillance	LDA	Silhouette	Vision Based	Human Recognition
[118]	Surveillance	feature selection mask	Silhouette	Vision Based	CASIA Gait database; Human Recognition
[119]	Surveillance	vPCA +LDA+ Deep Learning	Silhouette	Vision Based	UCMG and CASIA Database; Human Recognition
[109]	Surveillance	Genetic Algorithm - KPCA	Silhouette	Vision Based	CASIA Database; Gait Recognition
[84]	Surveillance	Sparse reconstruction based metric learning (SRML)	Silhouette	Vision Based	20 subject dataset- human and gender recognition
[105]	Surveillance	Deep Neural Network	Images	Vision Based	collective activity and Nursing Home Dataset ; Group activity
[68]	Surveillance	Neural Network - modified Clockwork RNNs	human kinematic	Sensor Based	Project Abacus from google - data from mobile phone ; Human Recognition
[120]	Activity Recognition	NN -SOM	Kinematic and Kinetics	Hybrid	22 subject ; Recognize emotion(normal,happy, sad, angry) from gait patterns ; Effect of music on gait pattern
[45]	Activity Recognition	Hidden Markov Model	Sensors	Sensor Based	20 subjects with 20 activity
[104]	Activity Recognition	Deep Learning - Convolutional NN	Silhouette + Accelerometer	Hybrid	3 database - Skoda, Actitracker and Opportunity

2.6 Dataset for Gait Analysis

Large and publicly available datasets are very vital for performance comparison and consistent evaluation. Table 2.4 summarizes some of major publicly available gait datasets since 1997. The datasets are arranged in order of publication year.

Approaches used for gait dataset collection involves vision, sensor, and combination of both. These dataset are different in term of number of subjects, subject viewpoint (single & Multiview and number of cameras (1-25)), scene type (indoor and Outdoor), walking style (walking, jumping, running, straight, circular etc), variable walking speed, carrying condition (with and without baggage etc), footwear (with different and without shoes etc.), variable surface condition (incline, treadmill) and clothing and time(in term of months, day time) and night vision. Multiple cameras are mostly synchronized in these datasets. Walking environment considers in datasets include indoor, outdoor, slope based environment, on the treadmill. Discussed Dataset have different applications. The survey illustrates that most datasets collected are used for gait recognition, others are for the clinical purpose. Now, there is the scope of the new dataset that can be used for recognition human behavior and activity in robust conditions.

For clinical analysis, PhysioBank collected dataset for Neuro-degenerative disease (e.g., Parkinson) in 1997 [121]. It is collected indoor using 64 patients by using force sensitive resistors. Later in 1999 and 2004 with 15 and 166 subjects, they study co-relation of aging and disease such as Parkinson. UCSD dataset from Visual Computing Group of the University of California San Diego is collected for gait recognition using six subject walking patterns in outdoor condition [38]. UCSD is considered as the first publically available dataset.

Researchers are also collecting dataset to identify and quantify natural gait pattern and characteristics according to gender and age-group from clinical and biomechanical perspective [122]. But it is not as simple because gait characteristics not only depend on anthropometric parameters but it also varies with cultural and social variations. Several gait studies have been carried out in different countries to get this normal gait characteristic. Spatiotemporal, kinematic and kinetic parameters are considered in study in United States ([123]; [124], [1]). In 1998, Benedetti et al. consider kinematic and kinetic data to define the normal characteristics of Italian people. Auvinet et al. consider spatiotemporal parameter in their study for France [125].

Tanenbaum et al. analyzed spatiotemporal, kinematic and kinetic characteristics

Table 2.4: Survey of available database

Year	Dataset	From	Subjects	N Seq	WE	Approach	Parameters	Setup	Syn*	Application
1997	Gait Dynamicse	Physio Bank	64	-	Indoor	Sensor Based	Ground reaction time , stride interveral	Force-sensitive resistors, VGRF	N/A	Disease and dynamics of the stride time
1998	UCSD	University of California	6	42	Outdoor	Vision Based	View	Sony Hi8 camera ; 1view ; 30 pfs, 640x480 pixel	N/A	Gait Recognition
1999	Gait in Aging and Disease	Physio Bank	15	-	Indoor	Sensor Based	Stride interval	Force-sensitive resistors, VGRF	N/A	Normal gait and Parkinson's disease analysis
2001	MIT Database	MIT	24	194	Indoor	Vision Based	View, time	Sony Handycam, 720x480, 15	N/A	Gait Recognition
2001	Georgia Tech Database	Georgia Tech, Atlanta, GA	20	188	Outdoor, indoor	Vision Based	View, time, distance	-		Gait Recognition
2001	CMU Mobo Database	Carnegie Mallon University	25	600	Indoor - treadmill	Vision Based	4 walking pattern, speed, carrying condition, incline surface	6 color camera; 60 degree each	Y	Gait Recognition
2001	HID-UMD Database 1	University of Maryland	25	100	Outdoor	Vision Based	View	4 viewpoints	N	Gait Recognition
2001	HID-UMD Database 2	University of Maryland	55	220	Outdoor	Vision Based	view	2 viewpoints- front, side , top mounted	N	Gait Recognition
2001	SOTON Small Database	University of Southampton	12	-	Indoor	Vision Based	5 Shoe, 3 cloths, view, 3 speed , 5 bag (without bag)	green background		Gait Recognition
2001	SOTON Large Database	University of Southampton	115	2128	outdoor, treadmill	Vision Based	View	6 viewpoints	Y	Gait Recognition
2001	USF Gait Database	University of South Florida	122	1870	Outdoor	Vision Based	Surface, show, carrying condition, time, walked around an ellipse in front of cameras.	2 viewpoints - left & right		Gait Recognition
2001	CASIA Database (A)	IACAS	20	240	Outdoor	Vision Based		3 viewpoints		Gait Recognition
2004	Gait in Parkinson's Disease.	PhysioBank	166	-	Indoor	Sensor Based	Stride interval , demographic information	16 Force sensitive sensor	N/A	Normal gait and Parkinson's disease analysis ; stride-to-stride dynamics
2005	CASIA (B)	IACAS	124	13640	Indoor	Vision Based	clothing, carrying 4 walking condition	11 viewpoints,		Gait and gender Recognition
2005	CASIA (C)	IACAS	153	1530	Outdoor, at night	Vision Based	Speed , carrying condition			Gait Recognition
2006	HumanEva I	Brown University	4	57	In door	Vision Based	Walking, Jogging, Gesturing Combo , Throwing and Catching a ball, Boxing and Combo	7 view points	yes	human motion tracking and pose estimation
2008	BUAA-IRIP	Beihang University, China	86	3010			multi view	7 view points		Gender classification
2008	GaitaBase Web repository	Royal Children's Hospital, Australia	-	-	-	Repository				web-accessible repository system
2009	CASIA Database (D)	(NLPR), China	88		Indoor	Hybrid- Vision and Sensor Based	Camera and foot scan			Gait Recognition
2010	HumanEva II	Brown University	2		In door	Vision Based	Combo of Walking, Jogging,Gesturing Combo, Throwing and Catching a ball, Boxing	4 viewpoints	Yes	human motion tracking and pose estimation
2010	TUM-IITKGP Database	TUM and IIT kharagpur	35	840	indoor	Vision Based	Include dynamic occlusion, walking style , carrying condition	6 view points		Gait Recognition
2011	SOTON Temporal Database	University of Southampton	25	-	Indoor	Vision Based	Time (0,1,3,4,5,8,9 and 12 months), View	12 viewpoints	Y	Effect of time on Gait Recognition
2011	OU-ISIR (Inertial Sensor)	Osaka University	744		Indoor , slop based	Vision Based	Age varation (2-78)	center IMUZ	N/A	gait-based human identification
2012	OU-ISIR (Treadmill)	Osaka University	122	1870	Indoor	Vision Based	Variable walking speed; clothing	25 views		Gait Recognition
2012	OU-ISIR (Large Population)	Osaka University	4007			Vision Based	Age variable (1-94)	2 cameras , 30pfs; 640 x 480pixels		Gender and age-group classification and recognition
2012	INDONESIAN GAIT	Institut eTeknologi Bandung	212		Indoor	Vision Based	View point, carrying condition, surface, shoe, time (months)	1 camera 90 fps,	N/A	Spatiotemporal & Kinematics Parameters
2013	AVA Multi-View (AVAMVG)		20	1200	Indoor	Vision Based	Multiview	6 cameras	Yes	Gait Recognition
2013	KIST Human Gait Pattern	Korea Institute of Science and Technology	113		Indoor , on treadmill	Vision Based	Multiview	8 camers	yes	predicting human gait pattern kinematics
2014	OU-ISIR (Gait Speed Transition)	Osaka University	179		Indoor, on treadmill	Vision Based	Variable Speed	1 camera	N/A	Gait Recognition under Speed Transition

of Korean people [126]. OU-ISIR gait dataset is considered as world's largest gait dataset with 4007 subjects including 2,135 male and 1,872 female with ages ranging from 1 to 94. Indonesian gait dataset suggests the standard spatiotemporal and kinematic parameters for Indonesian. KIST Human gait pattern dataset predicts kinematics parameter for Korean people.

As discussed earlier most of the dataset is collected either for gait or activity recognition considering different environmental conditions. CMU Dataset is constructed by Carnegie Mellon University, with 25 subjects walking on a treadmill in a room [127]. Six cameras, each placed at 60 degrees is used to capture gait pattern under different environment condition (4 walking pattern, speed, carrying condition and inclined surface).

Soton dataset from the University of Southampton has two databases namely small and large database. The small dataset contains 12 subjects walking in the door with a green background. The subject is filmed by walking at a different speed and wearing different clothes and shoes, with or without various bags is constructed in 2001. The Large dataset is the first dataset filmed over more than 100 subjects under three scenarios (indoor and outdoor track, treadmill)[128]. In 2011 they published temporal dataset containing data of 25 subjects over a large period up to 12 months from 12 views. This database accelerates research for inter-subject recognition and exploring gait recognition technique in different conditions

USF database also is known as HumanID Gait challenge data is the widely used gait database. It is collected by University of South Florida in 2001 with 122 subjects walking outdoor in a different walking environment (view, shoes, bag, grass or concrete surface and time up to 6 months) containing 1870 sequences [129]. The characteristic of considering the highest number of factors in this database among all existing database at that time makes it suitable for inter-factor analysis on gait recognition.

CASIA from Institute of Automation Chines Academy of Science has four datasets (A, B, C and D)[130]. Dataset A filmed 20 subjects from 3 view point (0, 45 and 90 degrees). It has 19139 images with 2.2 GB. Dataset B created in 2005 contains 124 subjects recorded from 11 views with and without baggage and with and without coat wearing condition. Human silhouettes are extracted and provided to the research community for analysis effect of view angles in gait recognition. Dataset C is collected using an infrared camera in 2005 in an outdoor environment at night. It consists of 153 subjects including variation in walking speed (slow, normal, fast and normal waking carrying baggage). This opens the research for

gait recognition in a dark environment (at night). Dataset D is a collection of 88 Chines subject's camera image and foot scan. Data is collected by synchronizing camera and Rescan Foot scan in 2009.

The Technical University of Munich and IIT Kharagpur collected TUM-IITKGP Dataset, in the year 2010. There are 35 individuals recorded from six viewpoints. Total 840 sequences are filmed during normal different walking style including static and dynamic occlusion with various carrying condition. This database is suitable for gait recognition in occlusion environment.

In 2011, Osaka University collected [131] Gait inertial sensor Dataset [132]. It includes 744 subject Spatio-temporal data using center IMUZ walking on slop surface indoor. They consider a larger variation in the age group from 2 to 78. In 2012 collected Treadmill dataset with 25 views (largest so far), walking speed (2-10 km/h) and clothing variations. Due to these variations, it accelerates research for the view, clothing, and speed-invariant gait recognition. In 2012 OU considers 4007 subjects including 2,135 male and 1,872 female with ages ranging from 1 to 94 with two viewpoints, it can use to analysis age-group or gender-based gait classification.

Other important dataset for gait recognition includes UCSD database, MIT database, Georgia Tech database, HID-UMD database 1 and 2 [21, 3, 133, 134]. To generalize the gait pattern over a larger set of population researchers [27, 1] collected data.

The available dataset, have a larger diversity in term of the walking environment but still insufficient for reliable for various gait analysis. In spite of all, we can't generalize the gait pattern of a given person. The number of subjects is still insufficient to generalize a standard gait pattern for a given age-group and given gender. Besides, for biometrics, the number of individuals are very limited in comparison to other biometrics such as face and fingerprints. There are not enough samples so that the dataset can be said as bias free on gender and age for example in CASIA Dataset B the ratio of male to female is 3 to 1 while USF dataset is biased toward an age group between twenty to thirty.

Currently, there is the scope of a new dataset that can be used for defining normal gait pattern among gender and age groups and can recognition human behavior and activity.

2.7 Chapter Summary

This chapter summarizes the development of gait analysis. Different aspects of gait analysis such as basic taxonomy (gait cycle, normal and abnormal gait) and parameters used have been discussed. A comprehensive review of the major survey article published in reputed journals and relevant conferences has been presented. This chapter also categorizes the available gait approaches into four types. Vision, sensor, others and a hybrid approach are discussed in details. Gait can be used in the analysis (normal gait, clinical, geriatric care and sports), biometric (surveillance and activity recognition), artificial gait (rehabilitation and humanoids), control applications and other applications. A comparative study of machine learning techniques such as supervised (Neural network, K-Nearest Neighbor (k-NN), Supporting Vector Machine (SVM), Bagging, Boosting, Random forest, etc) ; clustering based (Self Organizing map, k mean, hierarchical) reinforcing learning; rule-based fuzzy logic, evolutionary and Hybrid approaches for gait analysis is presented for a clear understanding. Recently Deep learning based researches are being carried out to explore new applications. Deep learning requires a large dataset to work with. In the end, some relevant dataset available for gait analysis has been discussed. There is a need for alternative parameters extraction and machine learning techniques to analyze the gait and extend the reachability to the common user.

The next chapter describes two proposed gait parameter extraction techniques using both marker and marker-less (also known as holistic) approach. The Proposed methodology is less complex and tedious in terms of marker placement, where there is no requirement of a controlled environment to acquire high-quality data, as compared to existing systems. The dataset collected during the study is analyzed statistically to find the norm of the dataset is discussed in the next chapter.

Chapter 3

Development of Gait Parameter Extraction (GPE) Approach

Gait parameters such as kinematics and spatio-temporal variables (stride length, step length, cadence, stance and swing phase etc.) of patients, play a significant role in the clinical gait analysis. Health-care professionals can examine the anomalous walk characteristics of the injured individual and assess the variables related to pathological gait, restoration and treatment mediations. But the cost associated with these clinical gait investigation framework are high, as the analysis is directed in exceedingly concentrated research centers or doctors counseling rooms. The position of the sensors (active markers) on the body is extremely rough (in terms of the clothing). In such imperative conditions, the subject's natural walking pattern cannot be captured. These elements constrain the popularity of the gait analysis systems [14, 7]. Thus there is need of alternate Gait Parameter Extraction Approaches (GPEA).

In this chapter, two methodologies to extract gait parameters with minimal sensor placement on subject's body and one without marker is proposed as delineated in figure 3.1. First proposed approach is with the passive marker based gait parameter extraction techniques. The GPEA which gives better results will be utilized in chapter 4 and 5 for developing the proposed Gait Parameter Prediction Models (GPPM) and Machine learning techniques for gait abnormality detection models.

The description of the experimental set-up used in this study is described in section 3.1. Section 3.2 discuss the passive marker based methodology with minimal sensor has been proposed. The proposed marker-based approach uses simple, comfortable to wear, error-free instrument, fast to wear, free from anatomical landmark

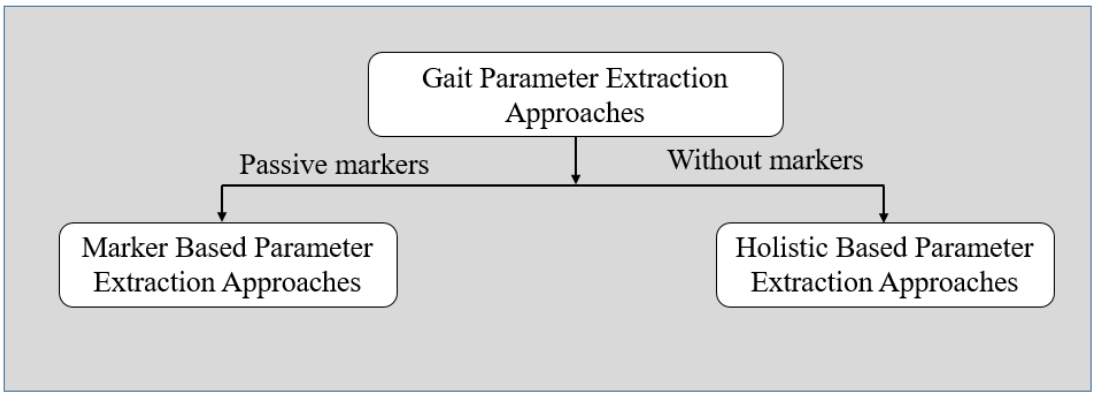


Figure 3.1: Proposed models for the gait parameter extraction.

misplacement, and low-cost markers. Later in section 3.3, a marker-free gait feature extraction methodology has been proposed. The MNIT Gait dataset collected during this study is analyzed statistically in section 3.4.

3.1 Experimental Setup

The gait analysis, presented in this study is carried out at *Robotics And Machine ANalytics Laboratory* (RAMAN Lab), MNIT Jaipur, INDIA. The experimental environment is presented in figure 3.2. The proposed gait analysis system consists of a digital video camera for recording and Personal Computer (PC) for data acquisition and processing. Joint movement tracking is done using an optical motion-capture system as discussed in section 3.2.

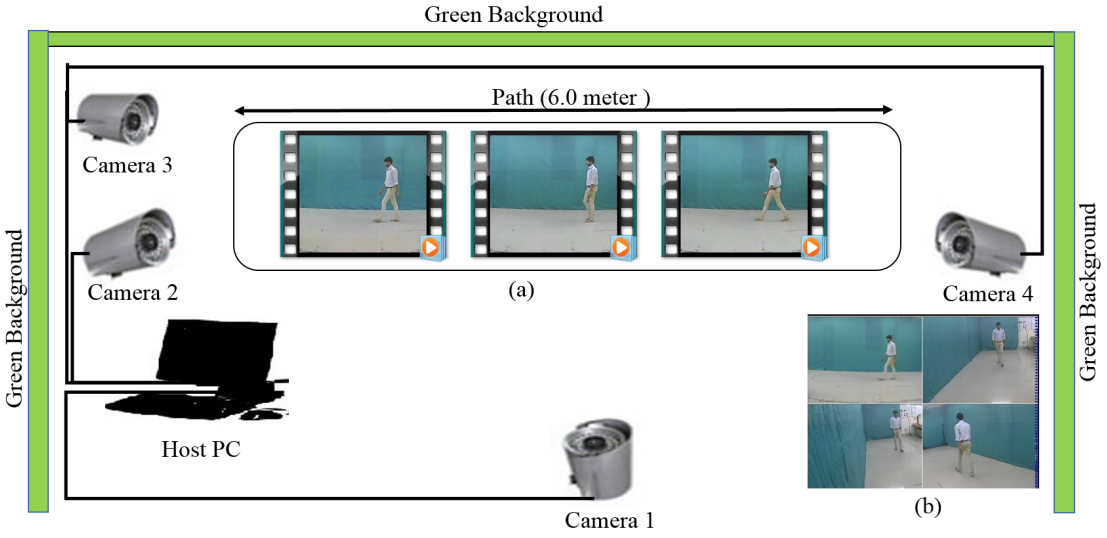


Figure 3.2: Gait lab setup at RAMAN Lab, MNIT Jaipur

3.1.1 System Description

The video capture is through different digital recorder with different frame rates and resolutions. A home digital video recorder, Nikon D5200 DSLR camera, and BIOPAC® 4 cameras were used to acquire the motion of the markers in the sagittal plane at 29, 50 and 30 frames per second (fps) respectively.

The home digital video recorder used for video acquisition is Sony Cybershot(SNY W830PS SIR) DSC W830/S, 20.1 MegaPixel (MP) with max ISO as 3200, focal length as 4.5-36 mm, the sensor size of 1/2.3 Type(S) (7.76 mm) and with 8x Optical Zoom. Initial experiment is carried out with the data collected using this configuration.

Nikon D5200 DSLR camera is of 24.1 MP with lens 18-55 mm, max ISO as 6400, CMOS image sensor, 1.5x lens focal length, the sensor size of 23.5 X 15.6 mm, maximum shutter speed of 1/1400 sec.

BIOPAC® system Camera System 4 (CAMSYS4) is used to record the four channels of video chaptering camera integration with existing MP150 System and Acq*Knowledge* software. A camera is of 1/4 Color CMOS image sensor, 6mm lens. It consists of tripod stands, gooseneck adapters, and A/D converter. The Gait lab setup with BIOPAC® based configuration as illustrated in figure 3.2 (a). Figure 3.2 (b) is the output screen at the host PC with this configuration.

In case of a passive marker such as a beacon, the background should be selected carefully. In RAMAN Lab, a green color background is used for the dataset collection. The walking path, used in this experiment, has a total length of 6 meters, which gives the subject enough space to take at least four steps after breaking inertia. In this study, the subject walks at a self-selected normal pace from right to left on the walking path and then left to right. The image frames in the collected video sequences vary from 90-300 frames.

3.1.2 Subjects/Participants

Optical Motion Analyzer developed during the study at RAMAN LAB is employed to collect Indian gait norms of 120 subjects from different age groups and genders using Nikon D5200 DSLR camera. In this thesis, now onwards it is called by *"MNIT Gait Dataset"*. Figure 3.3 presents the MNIT Gait dataset, collected at RAMAN Lab. The age classification follows that of Whittle [1] and presented in

Figure 3.3: Few sample of subjects of MNIT Gait dataset

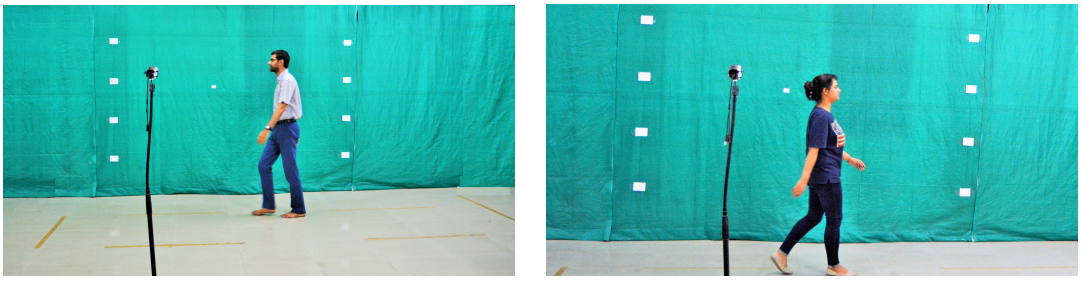
table 3.1. The range of the population of the dataset is from 5-60 years. Majority of the subjects are from the age group 18-49 years. Table 3.2 presents the anthropometric data of the male and female subjects within the 18-49 age-groups; which is the bulk of the participating subjects, along with the Standard Deviation (SD).

All of the subjects had no past record of injury or pain to the lower limbs, any type of lower extremity surgery, or balance problems that would affect their performance. Prior to data collection, the body parameters of each subject including age, height, leg length and weight are measured. This study is designed in accordance with the Declaration of Helsinki on experiments involving human beings [135]. Before participating, all subjects considered for the study signed a consent form.

Ten walking video sequences; five from left to right direction and another five from right to left are recorded for 120 different subjects. Thus total ($120 \times 10 = 1200$)

video sequences are collected in MNIT Gait dataset. Figure 3.4 illustrates one male and female subject respectively in the lab environment.

Later, subjects were instructed to act like if they had abnormal gait (assuming pain in either of the legs). After manual analysis with an expert on the dataset collected, 80 subjects seem to exhibit unhealthy gait pattern. The collection of these 80 subjects are considered to have unhealthy gait pattern. Each video consist of minimum three gait cycle. Thus MNIT Gait dataset consists of 3600 healthy and 240 unhealthy gait cycle.



(a) (b)

Figure 3.4: MNIT Gait sample acquisition (a) Male subject walking from left to right (b) Female subject walking from right to left

Processing of the data acquired from the conducted experiment is done using a Gait Analysis model that is developed using a simulation framework, for kinematic and spatiotemporal analysis of human gait. MATLAB® Toolbox and Python are used to implement and simulate the gait analysis process.

3.1.3 Clinical Gait Model

Davis [136], Helen Hayes Hospital [137], Kadaba [138], Vicon Clinical Manager (VCM) models[139], etc. are examples of the conventional clinic gait model used by almost all commercially available clinical systems [140]. These models are based on the minimum number of markers possible in lower extremity, at a time when

Table 3.1: MNIT Gait dataset based on age-groups and genders

Age-Group (years)	**Male**	**Female**	**Total**
< 18	5	5	10
18-49	78	24	102
>50	7	1	08
Total			120

systems were developed for detecting the limited markers to determine kinematics and kinetics parameters [141].

Canadian Society of Biomechanics (CSB) [142] is considered in this study as clinical gait model and presented in figure 3.5. The multi-rigid body model adopted for the kinematic analysis consisted of the rigid body segments such as trunk (shoulder and hip), thigh (knee and hip) and shank (ankle and knee) are connected by a line segment. The model is illustrated by four degrees of freedom: the trunk angle (trunk), thigh angle (thigh), the shank angle (shank) and foot angle.

The x and y locations of the segment endpoints are considered. These are digitized from video captured by a digital camera. Figure 3.5 shows the Canadian Society of Biomechanics (CSB) Gait Standard angle calculation methods. Absolute or segment angle is the angle between a segment and the right horizontal of the distal end. It should be consistently measured in the same direction from a single reference either horizontal or vertical. In this study, the horizontal direction is considered, thus trunk angle is 90 degree for the normal human gait. In this model, we used segment angle to calculate trunk, thigh, leg, and foot angle as described in illustrated in figure 3.6(a). While the joint or relative angle is used to calculate hip, knee and ankle angle as described in figure 3.6(b) and 3.6(c). Anatomical position or straight fully extended position is generally defined as 0 degrees in this case.

These joint regions (marker coordinates) contain significant information that helps in the identification and extraction of gait kinematics of individuals.

Table 3.2: Anthropometric parameters of the MNIT Gait dataset

Variable	Male		Female	
	Mean	**SD**	**Mean**	**SD**
Age (year)	29.92	8.78	23.42	6.28
Height (m)	1.72	0.06	1.56	0.13
Weight (kg)	70.53	13.37	56.48	13.84
BMI	22.38	6.757	21.86	5.82
Leg length(m)	0.90	0.05	0.83	0.09
Foot Length (m)	0.26	0.014	0.22	0.01
Bi-iliac width (m)	0.85	0.07	0.78	0.10

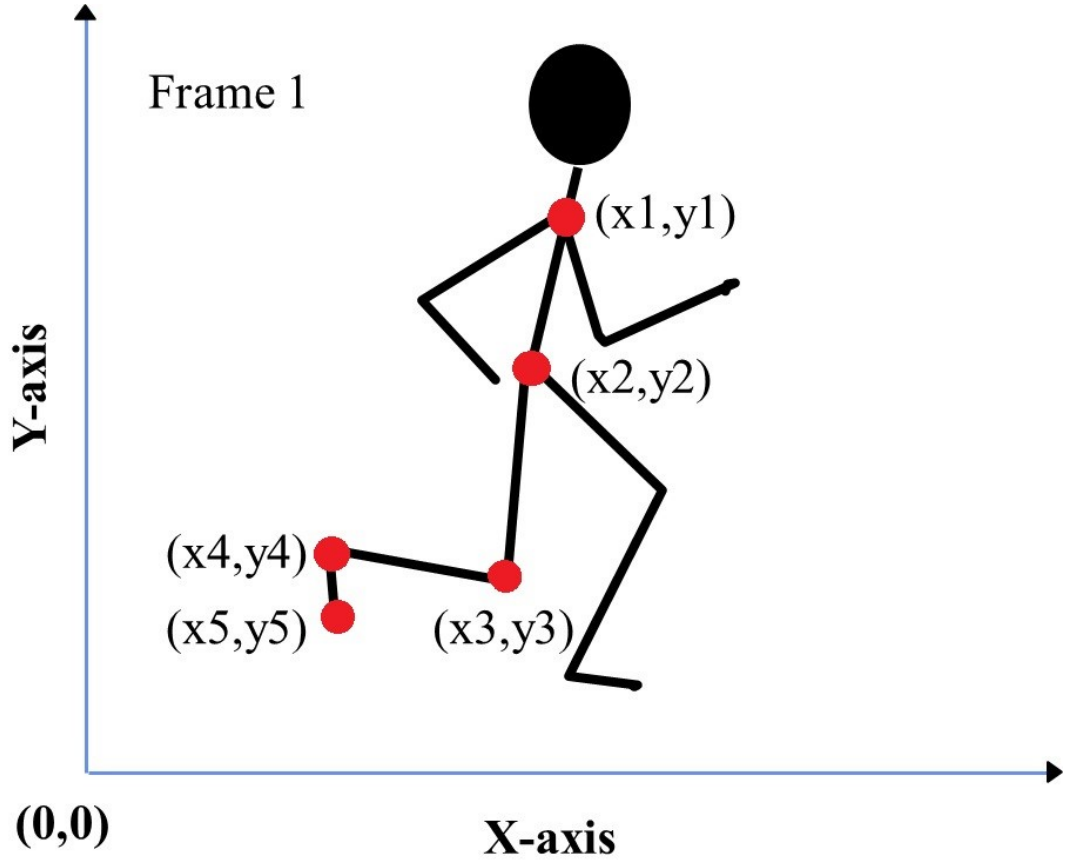


Figure 3.5: Canadian Society of Biomechanics (CSB) Clinical Gait Model: marker locations at anatomical points on subject body (in red color)

3.2 Passive Marker based Gait Parameter Extraction Approach (PM-GPEA)- Model 1

Most gait capture systems use direct measurement techniques to acquire specific motion information, but at the high cost for hardware and the subject's natural motion is hindered due to the presence of cables or other components that can affect subject's gait pattern. One of the main objectives of this work is to propose gait parameter extraction methodology with minimal sensors on the subject body so that the subject can exhibit the natural gait pattern. To overcome these limitations, an alternative approach has been introduced by using five single-color passive markers. The position of the markers is obtained by capturing and processing digital video images of markers that are fastened to the subject's body. The methodology is presented in the next section.

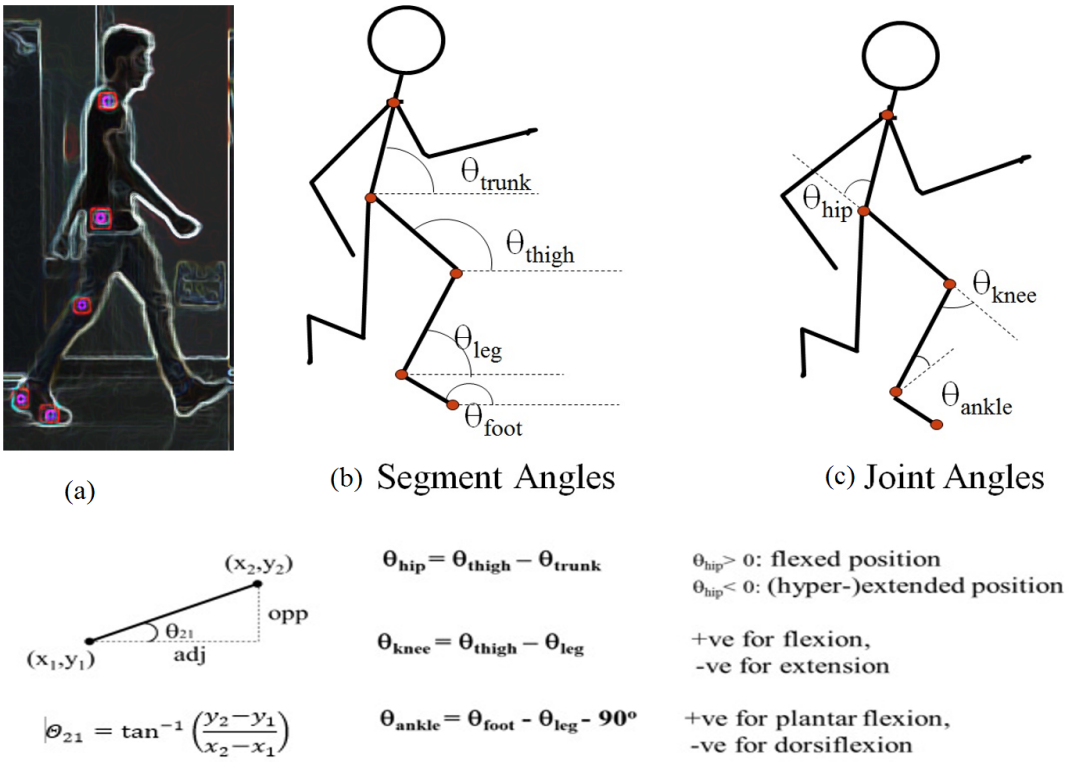


Figure 3.6: CSB Gait standard angle calculation method

3.2.1 Proposed Methodology for Passive marker based Gait Analysis System

In view of the problems prevalent in the state of the art discussed above, the present work provides a visual observation based system and method for analyzing gait. In this section, the approach to develop a 2D Gait analysis system is discussed. Our methodology of passive marker based gait analysis as shown in figure 3.7 involves three phases: Preprocessing, marker detection & tracking, Gait parameters Estimation (spatiotemporal and kinematics parameters). The main benefits for the presented system are that it does not consume excessive time and complexity required for marker placement, the need for a controlled environment to acquire high-quality data, the high cost for the markers, and also the effect of the markers on the subject's movement is reduced.

3.2.1.1 Preprocessing

In the pre-processing phase, a set of five red colored passive markers are attached to the clothes of the target subject at anatomical points of concern i.e. shoulder,

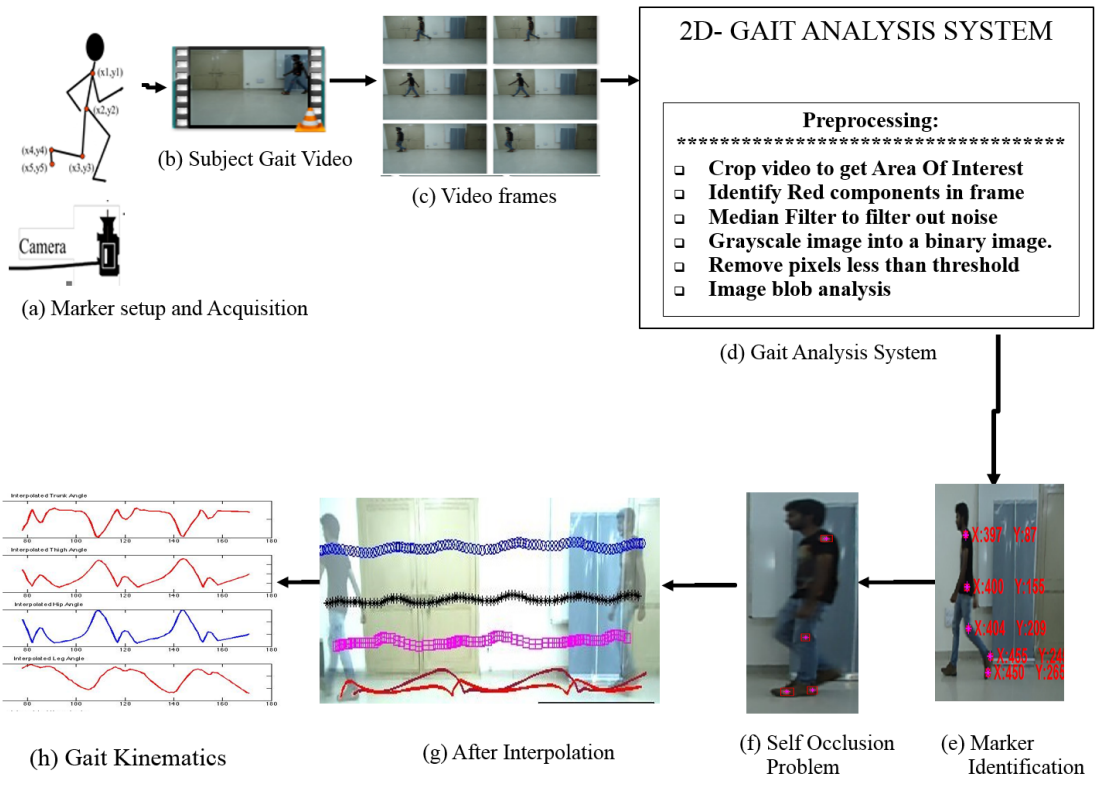


Figure 3.7: Passive marker based proposed methodology for vision based Gait analysis

hip, knee, ankle and toe as shown in figure 3.7(a). All these markers are placed on the left and right side of the subject's body.

The lateral or sagittal plane has been considered for the experiment. The pattern of the left side is similar to the right side of a healthy human, thus in this study, we have considered the sagittal plane for the analysis of the human subjects. Then the walking video is extracted into frames as shown in figure 3.7 (a),(b), and (c). It contains 180-220 frames in each sample video. Area of interest is extracted by cropping the frame. The extracted frame image data is in Red Green Blue (RGB) format as shown in following equation (3.1).

$$P = p_{i,j} = (r_{i,j}, g_{i,j}, b_{i,j}) \tag{3.1}$$

where P is color image, $p_{i,j}$ is the color information at the point (i,j), r, g and b represent red, green and blue respectively. The values are in the range of 0-255.

The approach adopted is to search for red colored regions in frame. RGB image $P = (r_{i,j}, g_{i,j}, b_{i,j})$ is converted into intensity format (grey scale) $Q = (q_{i,j})_{m,n}^{i,j=1}$, then into binary format $B = (b_{i,j})_{m,n}^{i,j=1}$, by introducing a threshold value T. The

algorithm is illustrated in algorithm 1

Algorithm 1 Conversion from color image to binary image

1: First convert Color image $P = (r_{i,j}, g_{i,j}, b_{i,j})$ to gray level image $Q = (q_{i,j})_{m,n}^{i,j=1}$

2: For $\forall$ pixels use one of the following operations:
3: 1. Set $q_{i,j} = max(r_{i,j}, g_{i,j}, b_{i,j})$ for all i and j
4: 2(a). Compute $q_{i,j} = r_{i,j} + g_{i,j} + b_{i,j}$
5: 2(b). Normalize all the pixel values in the interval [0,1] using $q_{i,j} = \frac{q_{i,j}}{max_{k,l} q_{k,l}}$

6: 3(a). Compute $q_{i,j} = \sqrt{r_{i,j} + g_{i,j} + b_{i,j}}$
7: 3(b). Normalize all the pixel values in the interval [0,1] using $q_{i,j} = \frac{q_{i,j}}{max_{k,l} q_{k,l}}$

8: End
9: For $\forall$ pixels, consider the threshold value T and binary image $B = (b_{i,j})_{m,n}^{i,j=1}$
10: if $(b_{i,j} > T)$ $b_{i,j} = 1$;
11: else $b_{i,j} = 0$;
12: End

Every pixel having an intensity below the threshold value is converted to black, and everything above becomes white. After fixing the threshold process, the white markers representing shoulder, hip, knee and toe ankle from the images could then be reliably extracted. The x, y coordinates of each marker are then calculated in the image plane (in pixels) on the basis of the marker's centroid. Binary Large OBject (BLOB) is referred to a group of connected pixels in a binary image and is used to extract the feature of the image. The purpose of the BLOB is to extract the objects in a binary image. It is a group of connected pixels, mainly 4 and 8-connectivity are popular techniques [143]. In this study, 8-connectivity is used using recursive grass-fire algorithm [144] as 8-connectivity is more accurate than the 4-connectivity.

3.2.1.2 Marker Identification and Tracking

At this point, the marker's positions in image plane have been identified. Marker detection and tracking module are used to obtain image coordinates of each marker in each frame as presented in figure 3.7(e). The five passive marker position M with coordinates (x, y) of subject were recorded. The vertical position coordinate is represented as y and x is the horizontal coordinate.

$$M_n(x, y); \quad \forall n = (1, 2 \ldots 5)\,; \tag{3.2}$$

where M is the coordinator of nth passive markers number mounted on the body.

Though the marker recognition process began from the top to the bottom of the image, the order of marker position in each frame is not always steady especially the toe and ankle marker. Thus on the basis of co-ordinates in 2D, toe and ankle are uniquely identified. In the current sagittal record setting, when the left hand blocks the marker attached to the hip joint, self-occlusion problem is identified as shown in figure 3.7(f). The break in the path of joint co-ordinates shows the occlusion problem. Linear interpolation technique is used to calculate or fill the location of the missing marker (hip) co-ordinates and thus overcome the occlusion problem as shown in following equation (3.3)

$$y_2 = \frac{(x_2 - x_1)(y_3 - y_1)}{x_3 - x_1} + y_1 \tag{3.3}$$

where $x_1, x_3, y_1 \& y_3$ are known variables, x_2 is the point to perform the interpolation and y_2 is not known.

A linear interpolation technique is able to find the correct marker position after the overlapping occurs and changes the marker data arrangement [145]. Marker tracking is used to track the position of each marker in correct arrangement after interpolation process. Figure 3.7(g) shows the position of each marker on the subject's body. Then kinematics parameter is extracted using this system as illustrated in figure 3.7(h). Figure 3.8(a) shows the output of marker identification phase in an unconstrained environment. Figure 3.8(b) shows the coordinates position for each marker in the frame when the green background is there.

(a) (b)

Figure 3.8: (a) Marker at anatomical points on subject's body identification in real time environment (in red color) (b) identification of coordinates for each marker

3.2.2 PM-GPEA Model: Preliminary Experimental Results

The preliminary results of gait analysis parameters are discussed in this section. A reframework for the detection of the Heel strike (HS) and toe-off (TO) using passive marker from conventional videography has been proposed followed by the spatiotemporal and kinematic parameters. Automated gait event detection result are discussed in section 3.2.2.1.

3.2.2.1 Spatial Temporal Parameters Estimation

Spatial-temporal parameters include step length, cadence, stride rate, velocity, etc. phases (stance/swing) and events (foot-strike, toe-off). Also includes walking base, toe out the angle, foot contact pattern. Among them, the stride length is most important and useful gait parameter for both medical and computing field. The first step to identify the gait events.

Passive Marker based Gait Event Detection (PMGED)

In this section, the result of the proposed PMGED is presented and analyzed. Gait event is considered to consist of Heel strike (HS) or foot strike and toe-off (TO). Heel strike is point of contact of the foot with the ground while toe-off is the point when the foot is off the ground. These events play a vital role for bio-mechanics and clinicians in diagnosis and rehabilitation. Detection of gait events also aids in estimation in some spatio-temporal and kinematics parameters such as stance and swing phase duration, stride time without heavily relying on gait laboratory. [146, 147].

For analysis, ten random healthy volunteers (seven male and three female) from MNIT Gait dataset were selected. Table 3.3 presents the anthropomorphic parameters of the subjects that participated in the study; (age: 23.5 ± 3.50 years old; height: 1.7019 ±0.11 m; weight: 62.99±7.92 kg and BMI:21.80 ±2.52). The experimental environment is same as discussed in section 3.1

The detail of the subject is presented in Table 3.3. Seven male volunteers (age: 23.85 ± 4.05 years old; height: 1.75 ±0.07 m; weight: 65.63±7.48 kg and BMI: 21.23 ±2.55) and three female volunteers (age: 22.67 ± 2.08 years old; height:

1.57 ±0.03 m; weight: 56.83±5.79 kg and BMI: 23.16 ±2.31) were requested to walk along the walkway at self-selected normal speed (SSNS).

Table 3.3: An overview of the subjects considered for the experiment

Subject	**Gender**	**Age (year)**	**Height (m)**	**Weight (kg)**	**BMI**
S1	M	20	1.70	65.90	22.80
S2	M	18	1.70	54.00	18.69
S3	M	30	1.74	61.00	20.15
S4	M	26	1.72	41.3	13.96
S5	F	21	1.54	53.00	22.35
S6	F	22	1.59	54.00	21.36
S7	M	23	1.69	67.90	23.63
S8	M	24	1.77	78.80	25.04
S9	M	27	1.85	65.90	19.15
S10	F	36	1.6	81.5	31.83

The five passive marker position coordinates (x, y) of the subject were recorded. The proposed PMGED can be identified from the plot of the difference between the vertical position coordinate (Y) for the malleolus and lateral metatarsal at each frame or time.

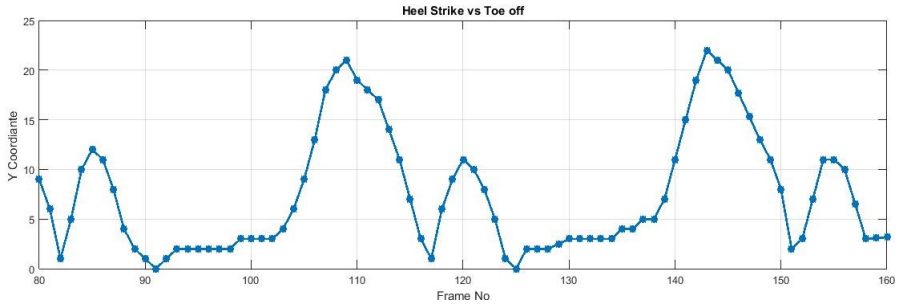
$$Y_n(t); \quad \forall n = (1, 2 \ldots 5)\,; \tag{3.4}$$

where $Y_n(t)$ represent the Y coordinate of nth passive markers number mounted on the body at time t. The difference in the malleolus and lateral metatarsal coordinates (Y_2 and Y_1) respectively can be calculated as

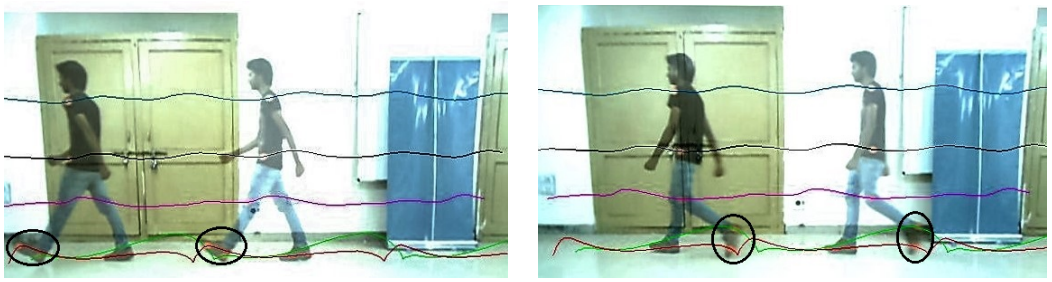
$$Y\_Event(t) = Y_2(t) - Y_1(t) \tag{3.5}$$

The plot of the difference of malleolus and lateral metatarsal coordinates is shown in figure 3.9(a). The minimum value of Y coordinates of the markers as illustrated in figure 3.9(a) and can be defined as Heel Strike (beginning of stance phase) and the maximum value of markers (during upward movement of the foot) can be defined as toe-off. It can be marked as the beginning of swing phase. Mathematically the gait events HS and TO can be represented as equation (3.6) and (3.7) respectively.

$$HS = Y(t) < Y(t+1) \quad \& \quad Y(t) < Y(t+1) \quad \& \quad Y(t) < C_1 Max(Y) \tag{3.6}$$



(a)



(b) (c)

Figure 3.9: Result of Passive marker based Gait event detection (PMGED) System for a random subject and marker trajectory. (a) Trajectory plot of difference of malleolus and lateral metatarsal coordinatesand (b) and (c) shows the Heel Strike (HS) and Toe Off (TO) with the marker trajectory.

$$TO = Y(t) > Y(t+1) \quad \& \quad Y(t) > Y(t+1) \quad \& \quad Y(t) > C_2 Min(Y) \tag{3.7}$$

where C_1 and C_2 are the the thresholds for HS and TO respectively. After regression analysis, C_1 and C_2 value are considered as 0.6.

On applying the PMGED algorithm, the system is not only able to detect the gait events automatically but also calculate the swing and phase in term of percentage and timing. Figures 3.9 (b) and (c) show two HS event and TO events with the marker trajectory coordinates. A standard is required to compare with, is developed using the expert.

Expert based visual event detection System An expert is responsible for manual identification and visual detection of the gait events in the capture video dataset. No software is used for the same. This manual video analysis act as a gold standard, to evaluate the accuracy of the proposed algorithm for the gait events detection system. The video frame number where foot makes the first contact with the ground is stored as Heel strike (HS), and the video frame where the foot is off the ground is recorded as foot off (FO). Each person reports exhibiting two and a half cycle, contributing three HS and two TO frame as Ground Truth (GT). Thus

total 30 GT_{HS} and 20 GT_{TO}, cases were recorded during the experiment.

Statistical Analysis for PMGED From the Statistical analysis, prospective detection rate and F-score value are used. In this study detection rate is the ratio of PMGED algorithm based detected gait events to total gait events identified manually by the expert. A temporal tolerance of ± 5 frame is used to match the gold standard Ground truth frame identified by the expert. To evaluate the overall performance of the system F-score was computed as

$$F = 2\left(\frac{Precision \times Recall}{Precision + Recall}\right) \tag{3.8}$$

where,

$$Precision = \frac{True \quad positives}{True \quad positives + False \quad positives}$$
$$Recall = \frac{True \quad positives}{True \quad positives + False \quad negatives}$$

The detection rate and F-score are calculated individually and presented in table 3.4. The statistical results suggest that PMGED can be used as the event detection technique. It is able to identify correctly both HS and TO in eight cases out of ten. The average detection rate is 96.00 and mean F score is 0.93. The average stance and swing percentage reported as 40.33 and 59.67 respectively. This is in accordance to the standard limit of 40 and 60 respectively [7].

The spatio-temporal parameters focused in this study are Step Length, Stride Length, Cadence, Velocity, Stance Phase, Swing Phase, Single-Support Duration and Double Support Duration (DS). The detail of these is same as presented in table 2.1.

For data validation for the proposed methodology, a 20-year-old male subject (175 cm height, 56 kg weight) with no known neurological or orthopedic gait impairments is selected randomly from the MNIT Gait dataset.

Table 3.5 shows the spatio-temporal parameters observed in the study. They are compared with the Standard data. Field of view of the camera is considered to calculate the stride length. The observed data is somewhat different from standard data [1] this may be due to cultural or social variations.

Table 3.4: Statistical result for the proposed Passive marker based Gait event detection (PMGED) system

Subjects	#of HS	# of TO	Stance %	Swing %	Detection rate	F-Score
S1	3	2	41.38	58.62	100	1
S2	3	2	39.65	60.35	100	1
S3	3	2	39.88	60.12	100	1
S4	3	2	41.06	58.94	100	1
S5	3	2	38.36	61.64	100	1
S6	2	2	41.47	58.53	80	0.66
S7	3	1	42.67	57.33	80	0.66
S8	3	2	38.98	61.02	100	1
S9	3	2	40.67	59.33	100	1
S10	3	2	39.21	60.79	100	1
Overall	**29**	**19**	**40.33**	**59.67**	**96**	**0.93**

Table 3.5: Spatio Temporal parameter observed in the study

Gait Parameters	PMBA result	Standard data [1]
Stride length(m)	1.09	1.68-1.72
Step Length (m)	0.54	0.68-0.85
Stance phase (s)	0.80	0.62-0.70
Swing phase (s)	0.367	0.36-0.40
Cadence	140	113-118
Single-Support	68.57	60-62.0
Double-support	31.143	21.2-23.8
Walking Velocity	1.17	0.64-1.14

3.2.2.2 Kinematic Parameters Estimation

For data validation of kinematic result, the same subject has been considered as taken in section 3.2.2.1. The real coordinates of markers attached on subject's body is in 2D sagittal plane is shown in graph, plotted in figure 3.10, consists of 3 complete walking steps. Therefore, the subject needs 35 frames to walk in 1 step. With 29 frame per second (fps) camcorder, the walking period of the subject can be calculated, i.e. 1.206 seconds. Extracted knee, ankle, hip and Trunk region in the gait sequence for the subject are shown in figure 3.10(a). It is also known as the gait signature. M1 represents the shoulder coordinates, M2 is for hip, knee coordinates are represented as M3, M4 and M5 are ankle and toe respectively.

Considering the Canadian Society of Biomechanics (CSB) Gait clinical model as discussed in 3.1.3, the joint angles are extracted. The preliminary result of the proposed system provides decent quantitative kinematics gait parameters i.e. joint

angles. In this model, joint coordinates are used to calculate trunk, thigh, leg, and foot angle as illustrated in figure 3.10(b). The x-axis corresponds to one gait cycle, and the y-axis is the angle. This figure is for all gait walking sequences captured in the video.

Figure 3.11(a) is the extracted knee angle for subject S21. The bold line represents the average knee angle for this subject. The dashed line in the plot is the knee angle for different gait cycles of the same person.

The limitation of the proposed approach is that it is not marker-free based approach. In the next section, a holistic based approach is proposed to explore the markerless way to extract the kinematics and spatio-temporal parameters.

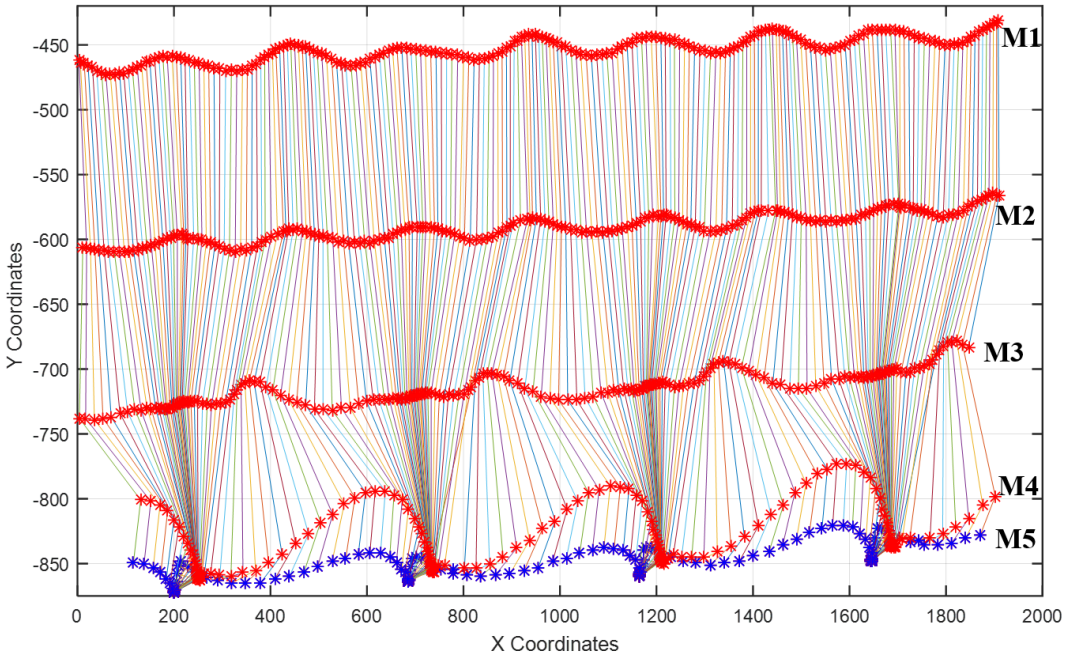
3.3 Marker-less/Holistic Based Gait Parameter Extraction Approach (Ml-GPEA)- Model 2

The other objective of our work is to propose a method that can analyze the gait parameters without using the markers on the body so that the natural gait parameter can be extracted without any interference to the subject. In this section, a heuristic based approach is proposed to extract the key joint coordinates (shoulder, hip, left and right knee, left and right ankle) and angles (knee and hip angles for left and right legs) for MNIT Gait dataset. The lab-setup is same as discussed in the section 3.1.

The marker-less or holistic approach is based on analyzing the shape of human gait and extracting potent feature. The marker-less approach typically uses the subject's silhouette as an input. From the silhouette image, joint coordinates are estimated based on the position of legs in the frame. After estimating the coordinates, the joint angles for hip and knee are then compared with the ground truth (it is a term used in various fields to refer to information provided by direct observation (i.e., empirical evidence) as opposed to information provided by inference). The methodology for the proposed system is discussed in detail in next section.

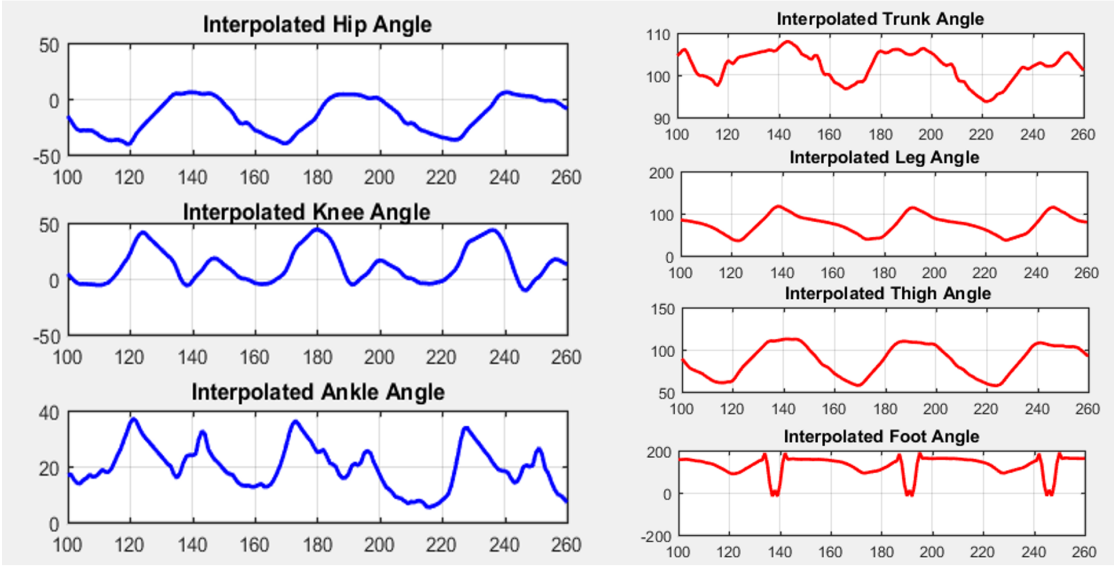
3.3.1 Methodology for Ml-GPEA

The proposed approach is a single-camera based method, adequate for the representation of gait in 2D and provides for a simplified laboratory setting diminishing



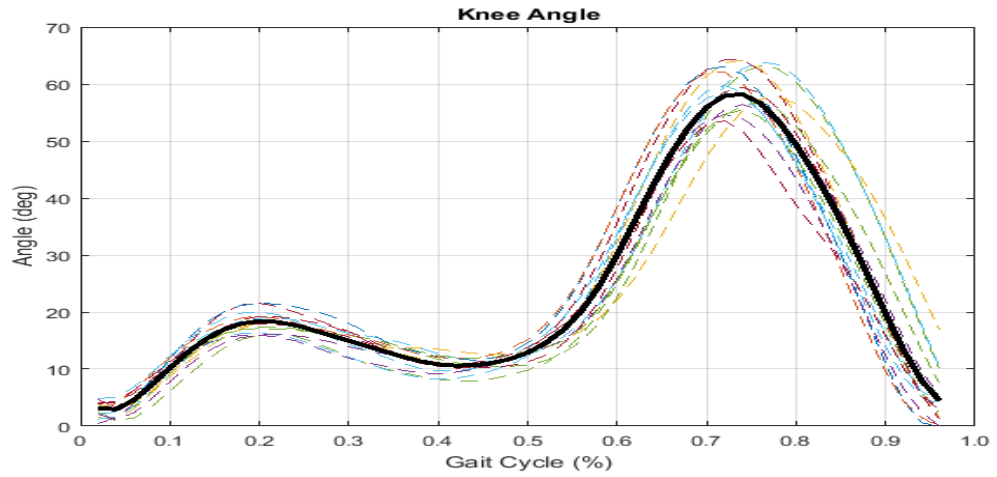
M1- Shoulder Marker, M2- Hip Marker, M3- Knee Marker, M4- Ankle Marker, M5-Toe Marker

(a)

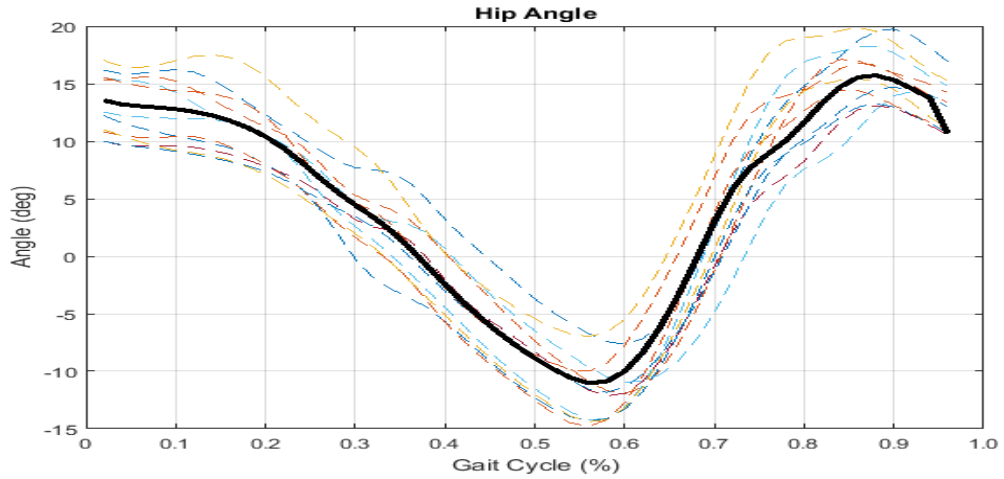


(b)

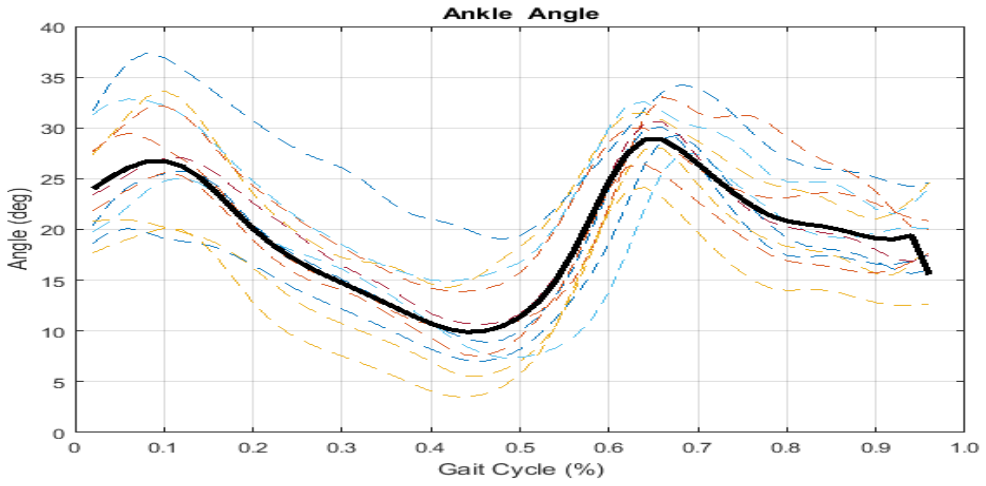
Figure 3.10: Result of Passive marker based gait Analysis approach. (a)Joint coordinates (b) Segments angles for the whole video sequence



(a)



(b)



(c)

Figure 3.11: PMGPEA result for S22, dotted lines are angles for the different gait cycle and dark line represents the average angle; a)Knee angle (b) Hip angle (c) Ankle angle

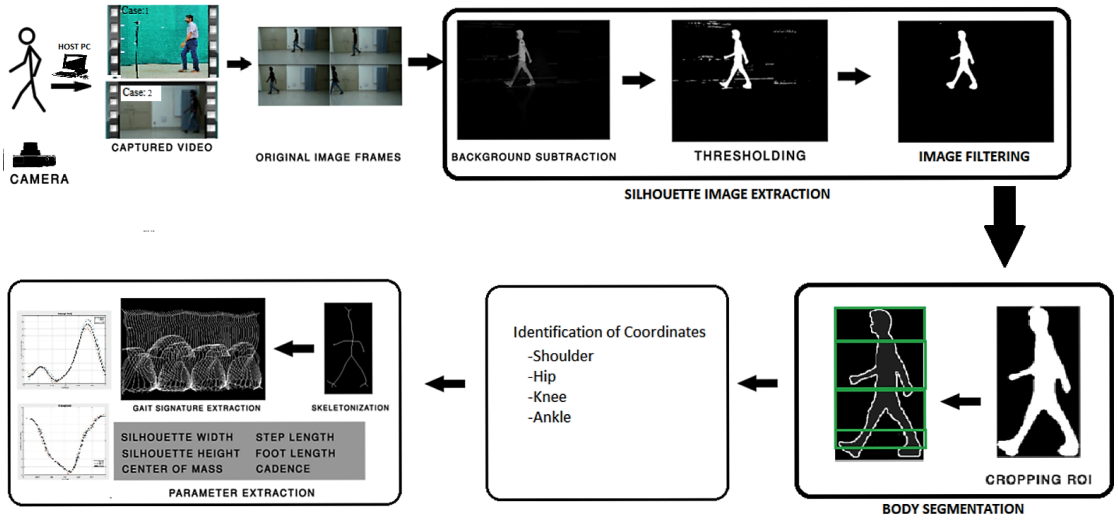


Figure 3.12: Flow diagram for parameter extraction in markerless gait analysis

the space required for the devices, the number of cameras, and thus the associated costs. The proposed marker-less methodology is used to extract the joint coordinates from the videos captured. The block diagram for gait parameter extraction is illustrated in figure 3.12. Two cases are considered for the study; in first the subjects walk in the constrained lab condition without marker; the second one is when subjects walk across random background without marker using Nikon D5200 DSLR camera. For both the cases, gait parameters can be extracted using silhouette image extraction phase, body segmentation, identification of the joint coordinates.

The accuracy of marker-less studies relies on silhouette images. Thus the first step is an extraction of a noise-free silhouette. Image segmentation and filtering are used to extract the joint coordinates and then extract the gait spatiotemporal and kinematic parameters angle from the subject image.

3.3.1.1 Silhouette image extraction

Silhouette extraction is vital to extract the parameters from the image. Selection of color space model, foreground extraction modeling, and object detection are implemented to extract the silhouette. In the preliminary study, it is assumed that camera is static and the subject is the only object moving in front of background (green or random) in the video. The video is a sequence of images. Each image underwent a pre-processing stage where the silhouette images are extracted from the frame with minimum noise and shadow in it.

Color space model

The selection of best color space models is vital for silhouette image extraction approach. Red-Green-Blue (RGB) and Hue-Saturation-Value (HSV) are among the commonly available color space models [148].

HSV stands for Hue, Saturation and Value where Hue is the attribute of the color, Saturation is the measure of how different a color appears from a gray of the same lightness and Value is the other name for lightness. It is the Hue component that makes approach less sensitive toward lighting conditions [148]. From a RGB image, the H, S, and V component are calculated by the following equations:

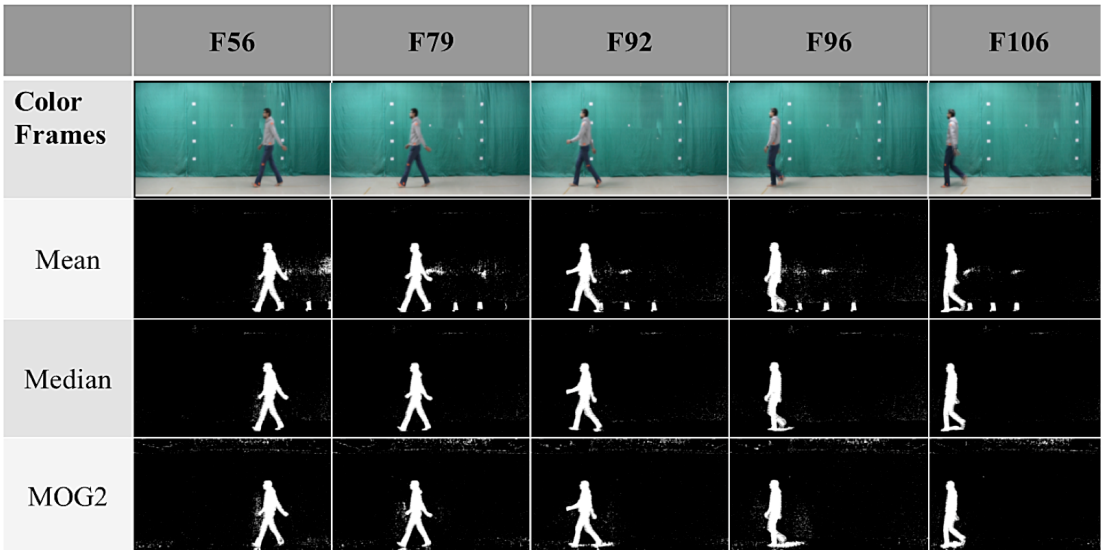
$$\begin{aligned} Value &= max(R_{value}, G_{value}, B_{value}) \\ Saturation &= Diff/Value \\ Hue &= \Delta \times 60^{\circ} \end{aligned} \tag{3.9}$$

Where

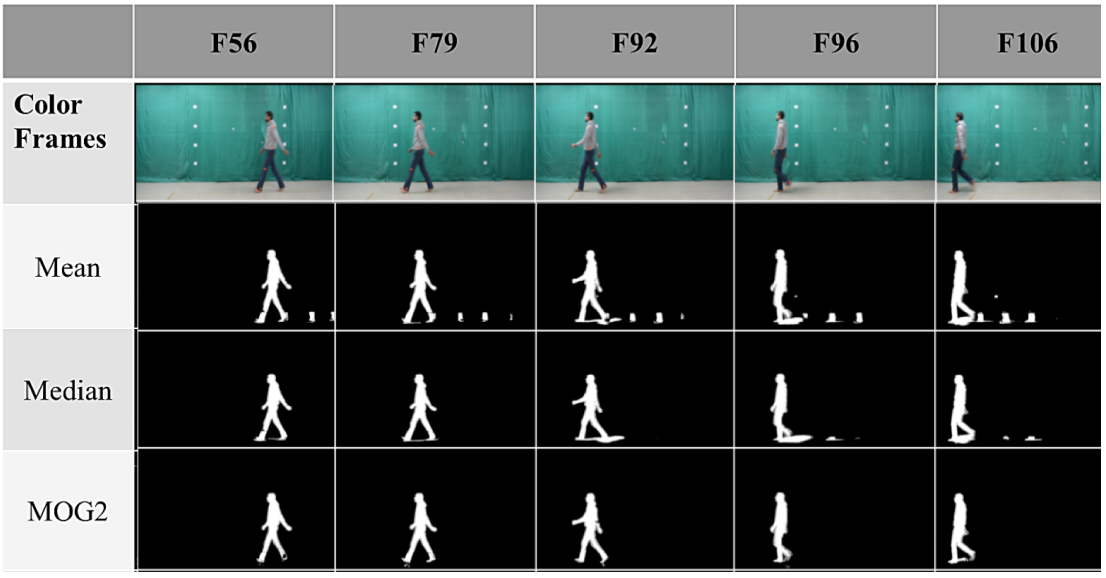
HSV separates image intensity from color information. Prerequisites are done to minimize the shadow content of the image frame. A human perceives color similar to HSV, against, RGB approach. In HSV color space, shadows hold different spectral properties like high saturation due to Rayleigh scattering effect, high hue values because shadow areas are comparatively darker and high value because of light occluded by the moving subject. Problems like false color and shape distortions are minimized by the HSV color space [149].

Foreground Extraction Modeling

Foreground extraction modeling is considered to be a fundamental and critical pre-processing step in many vision-based applications. It extracts the moving foreground from static background [150, 151]. In foreground extraction subtraction technique a static background is subtracted from the frame, and only the dynamic part of the frame (the person in this study) is left. In case of fundamental background extraction technique, it is assumed that video sequence (I) is with static background (B), devoid of any moving object and foreground pixel value differentiable from B. Each pixel value is classified into foreground and background:



(a)



(b)

Figure 3.13: (a) Foreground extraction techniques on the subject frames in different color space result on RBG color space (b)result on HSV color space

$$\begin{aligned} Foreground(I_{x_i,y_i} : d(I_{x_i,y_i}, B_{x,y}) > T) \\ Background(I_{x_i,y_i} : d(I_{x_i,y_i}, B_{x,y}) < T) \end{aligned} \tag{3.10}$$

Where T is a threshold, $I_{x_i,y_i}, B_{x,y}$ is pixel value.

Normally an image with background frame is not available, so the idea here is to extract the background from the given frames. It becomes more difficult when there is the presence of noise in the moving object as in case 2, where subject

walks in a random background. After intensive literature survey, it is found out that foreground extraction or background subtraction techniques can be divided into four main classes: Basic Techniques, statistical techniques, machine learning techniques, and other techniques [151].

In this study, a comparative analysis of Mean, Median, Mixture of Gaussian (MOG), and MOGG (MOG2) [151] are used as the foreground extraction techniques on RGB and HSV color space demonstrated in figure 3.13(a) and 3.13(b) respectively.

1. Mixture of Gaussian (MOG) - A robust approach is proposed as Mixture of Gaussian(MOG) by Stauffer et al. [152] for moving background such as moving trees or running water. Each pixel is represented as a combination of $K(X_1, X_2 \cdots X_k)$ Gaussian distributions.

$$\begin{aligned} P(X_t) &= \sum_{i=1}^{K} \omega_{i,t} N\left(X_t \mid \mu_{i,t}, \Sigma_{i,t}\right) \\ N\left(X_t \mid \mu_{it}, \Sigma_{i,t}\right) &= \frac{1}{(2\pi)^{D/2}} \frac{1}{|\Sigma_{i,t}|^{1/2}} J \\ J &= \exp\left(-\frac{1}{2}(X_t - \mu_{i,t})^T \Sigma_{i,t}^{-1}\left(X_t - \mu_{i,t}\right)\right) \end{aligned} \tag{3.11}$$

where $\omega_{i,t}$ is the weight of i^{th} Gaussian distribution at time t. $\mu_{i,t}$ and $\Sigma_{i,t}$ are mean and covariance(diagonal) matrix for corresponding distributions.

It uses a function to model each background pixel by a mixture of K Gaussian distributions (K = 3 to 5). The weights of the mixture represent the time proportions that colors should stay in the scene.

2. Mixture of Gaussian2 (MOG2) - It is similar to MOG i.e. based on Gaussian Mixture-based background subtraction [153]. The key difference between MOG and MOG2 is that MOG2 selects the suitable number of Gaussian distribution for each pixel while in MOG, K Gaussian distributions are considered throughout the algorithm [153]. It helps to achieve better adaptability to the scenes due to illumination changes in the video. Figure 3.13(b) presents the result of MOG on HSV image.

Above algorithms are applied to the randomly selected test video, and the key insight is that MOG2 approach performed better than mean, median and MOG as delineate in figure 3.13(a) and 3.13(b). In the case of MOG, most of the body parts

of the moving person are absent due to fixed value Gaussian distribution whereas this issue is resolved using MOG2. Thus in this study MOG2 is considered for the foreground modeling and object detection. To further enhance the quality of silhouette, morphological operations are applied.

Image Filtering

Images may be corrupted with the different type of noise (Gaussian, uniform, or salt and pepper). So based on the noise, smoothing is applied. Smoothing, also called blurring is an image processing operation which acts as a filter used to reduce noise. In our case salt and pepper noise exists in the image. In this study, a nonlinear digital filter known as the median filter has been used. Median filtering is a nonlinear process useful in reducing impulsive or salt and pepper noise. The function smooths an image using the median filter with the $k \times k$ size aperture. To reduce the noise, an aperture of 5×5 median filter is used and the quality of an image obtained is enhanced as shown in the figure 3.14. This size is selected after rigorous trials.

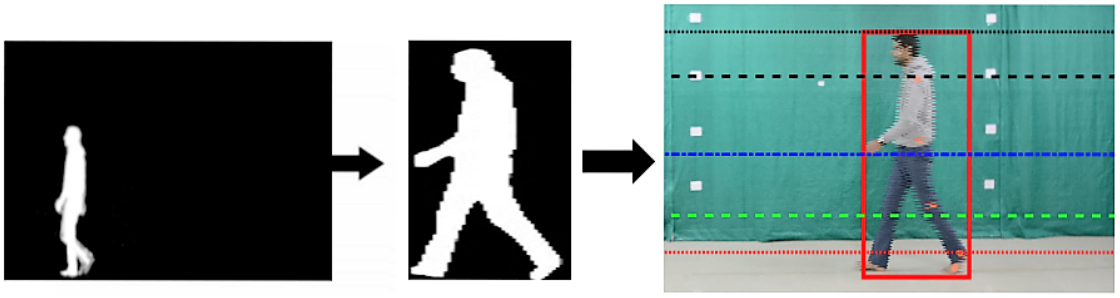


Figure 3.14: Silhouette image extraction result (a) Foreground extraction (b) Morphological operations (c) Segmentation

Morphology removes the unwanted white noise by shrinking its area. Mathematical morphology is a theory for the analysis and processing of geometrical structures. Its operators include erosion, dilation, opening, and closing. The opening of A by B is obtained by the erosion of A by B, following by dilation of the resulting image by B.

$$\begin{aligned} open(image, element) &= dilate(erode(image, element)) \\ Final\ Image &= open(image, element) \end{aligned} \tag{3.12}$$

Finally, a dilation is applied which is the reverse of erosion, used to dilate the object to reshape the object without noise interference. A repeated trial of multiple

combinations of erosion and dilation can also be applied depending upon the image under processing. The result of the opening is delineated in figure 3.14(b).

3.3.1.2 Body segmentation

The Region Of Interest (ROI) is a selected set of samples within a larger set of data. The video captured by Nikon DSLR is used for analysis. In the study, the image frame is of the size 1920×1080 pixels and the region of the moving subject is of the size 200×400 pixels. So, the region of interest is generated using the basic algorithm of determining the farthest edge where pixel's value in the binary image is changing in vertical as well as in horizontal direction. This creates a rectangular region on which further operations are performed. The result of this operation is shown in figure 3.14(c).

The ROI detection helps us in tracking the moving object in the silhouette image extracted out of the video. Based on the extracted ROI, the height of the image is calculated as the vertical length given by:

$$h = y^{max} - y^{min} \tag{3.13}$$

The dimension of various body segments are extensively studied [154] is shown in 3.14(c).

After extracting the silhouette image, segmentation of the image is done according to the values given in the study [155, 156]. These value will give a rough idea about the positions of different joints. A bounding box is created around the silhouette which gives the height of the person in pixels. The silhouette is divided horizontally into 5 regions.

Based on their anatomic studies [157, 155, 156], the row positions of neck, shoulder, hip, knee, and ankle are estimated as follows:

$$\begin{aligned} Y_{neck} = 0.13 \times h \quad Y_{shoulder} = 0.182 \times h, \\ Y_{hip} = 0.530 \times h, \\ Y_{Knee} = 0.285 \times h, \quad Y_{ankle} = 0.90 \times h, \end{aligned} \tag{3.14}$$

Where h is the height of the subject in silhouette image. These segments can be defined into different regions. Region 1 is from $Y_{top} - Y_{neck}$, Region 2 is from

$Y_{neck} - Y_{shoulder}$, Region 3 is from $Y_{shoulder} - Y_{hip}$, Region 4 is $Y_{hip} - Y_{ankle}$. This segmentation will provide a rough idea of the positions of the joints as shown in figure 3.14 (c).

The mean body height can be used to determine the topological position of each body parts in a human figure. Based on this, the extracted ROI silhouette image can be divided horizontally into several regions to obtain the better realization of different body segments and the coordinates corresponding to them in the actual image frame.

3.3.1.3 Identification of joint co-ordinates

For extracting the gait parameters, there is a need for identification of key joint coordinates. First shoulder and hip coordinates are estimated, followed by knee and ankle. For identification of knee and ankle coordinates, the position of legs in the frame is identified and then based on cases, methods are proposed accordingly.

Shoulder and hip detection

The y coordinate for the shoulder is considered as $0.182 \times h$, where h is body height. For the x coordinate mean of all the columns where the silhouette image exists for the particular y coordinate is taken.

The y coordinate for the hip is estimated as $0.530 \times h$ using (3.14). Arm swing acts as noise for identifying the horizontal positions of the hip. For this region, hand outlines are removed by subtracting standard deviation (σ) of vertical coordinate from the mean(μ) from both sides. This σ act as the threshold. Now, for the x coordinate, mean of all the columns is taken for the particular y coordinate.

Knee and ankle detection

A two-fold approach is proposed for knee and ankle detection for left and right leg. In the first fold front and rear leg's knee and ankle coordinates are identified, and then these coordinates are assigned to the left and right leg respectively. To identify knee and ankle coordinates, we first need to identify the position of legs in the frame as shown in figure 3.15(a).

During heel strike and toe-off phase both the legs (both knees and ankles) are not overlapping. During the initial swing, only knees are overlapping. In mid swing phase, both knees and ankles are overlapping. When only knees are overlapping in a frame, then the knee coordinates are determined using previous frame's coordinates, and when both knees and ankles are overlapping, then interpolation is done to find out the coordinates. To identify the pose of legs in a frame, the number of horizontal lines between the knees and ankles are estimated. In this study, a threshold value of ten pixels is set to identify the leg position. This threshold is selected after hit and trial method in regression iterations. The knee and ankle coordinates are stored separately for front leg and back leg.

Knee and Ankle positions during heel strike and toe off stages
This stage is identified when both the legs are not overlapping. Two sets S1 and S2 are considered, where S1 and S2 contain set of points for front leg and back leg respectively.

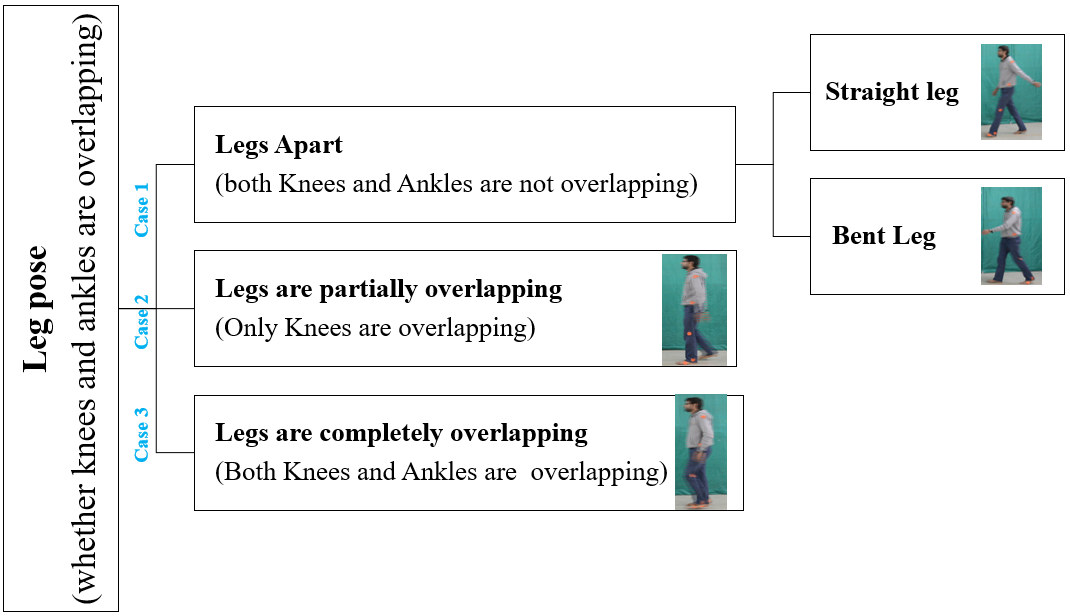
During walking, leg position can be subdivided into two categories, i.e., leg straight and leg bent. To determine whether the leg is straight or bent, linear regression on the sets S1 and S2 is performed as shown in figure 3.15(b). After applying linear regression, the mean square error is identified, and a threshold (D_Knee) of 10.5, in this case, is set. If the error is more than the threshold, then that leg is bent otherwise it is straight.

For straight-leg case: The y coordinate is defined as 0.285*height of the body using (3.14). For the x coordinate mean value of all the columns for the defined y coordinate is taken. Figure 3.15(b) shows case when the leg is straight.

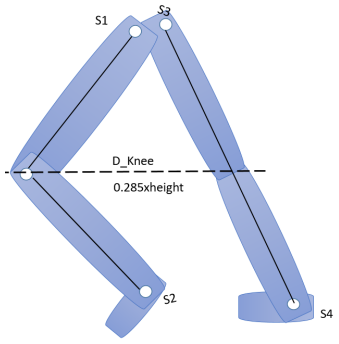
For bent leg case: The equation of a straight line is found, using the first and last points of the leg. Now distance of the pixels on the leg from the straight line is calculated. The coordinates with maximum distance (Dmax) will give knee position for the leg as shown in figure 3.15(c). For the ankle coordinates in both the cases the end points of the linear regression line are taken.

Knee and ankle coordinates during initial swing & mid stance
During mid swing stage, knees are overlapping, and ankles are not. One leg is in mid stance phase, and another leg is in mid-swing.

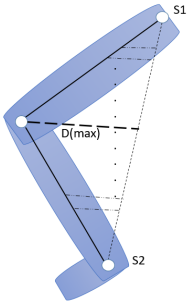
Stance limb: The knee and ankle movements are minimal for the stance limb. Therefore, knee coordinate is



(a) Knee Coordinate detection work flow



(b) Leg state bent or straight



(c) Bent case knee identification

Figure 3.15: Knee Coordinate detection for bent and staight leg cases

$$(x, y, t) = (x - 1, y, t - 1) \tag{3.15}$$

where (x, y) is the previous knee position.

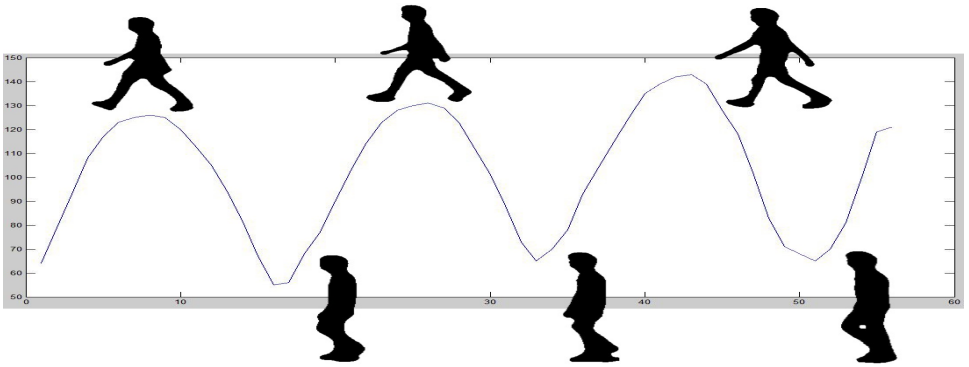
The y coordinate for ankle during the mid-stance phase is Y_{ankle} and x coordinate is the midpoint of all the columns for row Y_{ankle} where silhouette exists.

Swinging Limb: The swinging limb moves faster than the stance limb, but it still depends on the previous frame knee coordinates. The knee coordinates can be determined as:

$$(x, y, t + 1) = (x - k, y, t) \tag{3.16}$$

Where (x, y) corresponds to the previous knee position. The k value depends on

the speed of the subject and after regress analysis considered as ten. For the ankle coordinates, a line segment is obtained by finding the rightmost points of the calf of the rear leg. This line segment is then shifted to the median. The last point of this line segment gives the location of the swinging limb ankle.



(a)

(b)

Figure 3.16: Preliminary result of proposed Ml-GPEA (a) Identification of left and right leg from silhouette images (b) Joint position estimation by the proposed method for female subject (S22) and male subject (S43)

During mid swing phase both the legs are completely overlapping, knee and ankle coordinates are obtained by interpolation [4]. Interpolation is a legitimate choice

as this overlapping condition remains only for two-three frames in the entire gait cycle.

Left and right knee detection

After determining the knee coordinates for both the legs in each frame, knee coordinates to left leg and right leg are assigned. After regression analysis, it is found that switching, between legs, takes place when the width of the bounding box is minimum. For example, the minimum width of the silhouette is shown in figure 3.16(a) for the randomly chosen subject. So at these frames, front and back knee coordinates will keep switching between left and right knee's coordinates.

3.3.2 Ml-GPEA Model: Preliminary Results and Discussion

This section discusses the results for joint coordinates and angles, after identification of coordinates proposed in the previous section. A random female (S24) and a random male (S43) subject are chosen from the MNIT Gait dataset for the analysis purpose. The selection fo these subjects are purely random.

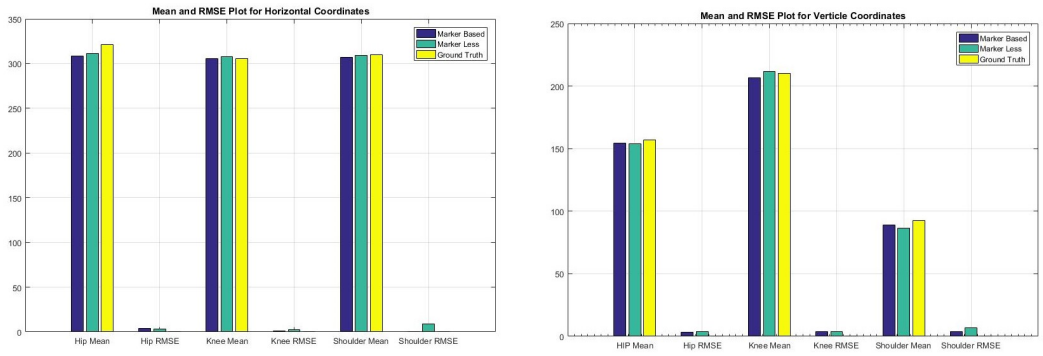
3.3.2.1 Joint coordinates

Figure 3.16(b), demonstrates the coordinates position of the shoulder, hip, knee and ankle identified by the proposed Ml-GPEA technique, for the subjects. The circles in green color represent the joints at the left side, and the red color circles represent the right side of the subject joints.

After finding the coordinates by both passive marker and markerless approach, a comparative analysis is presented in this thesis on ten randomly selected frames from the video.

Root Mean Square Error (RMSE) [158] is an estimator to quantify the difference between values estimated by the technique followed with the ground truth value and presented as follows

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \widehat{y})^2}, \tag{3.17}$$



(a) Horizontal coordinates

(b) verticle coordinates

Figure 3.17: Bar graph for horizontal and Vertical coordinates

Where y_i and $\widehat{y_i}$ are the actual and predicted gait movements at frame i and and N is the no of frames.

Ground truth is calculated manually for these randomly selected 10 frames. Mean, minimum, maximum and RMSE calculated for both horizontal and vertical coordinates of the shoulder, hip, and left knee are presented in table 3.6 for the selected video frames. It is found that the RMSE value for the shoulder for the vertical coordinates is 6.618 while that of passive marker based approach is 3.807 as shown in figure 3.17. Statistically, the passive marker result is better than ML-GPEA approach, but the difference of 2.6 pixels is not very significant in this case.

Table 3.6: Mean, minimum, maximum and RMSEs value for shoulder, hip and knee coordinates by marker based, proposed markerless system and ground truth

Segments		PM-GPEA		Proposed Ml-GPEA		Ground truth	
		Horizontal	Vertical	Horizontal	Vertical	Horizontal	Vertical
Shoulder	Mean	307.2	88.8	309.05	86.3	310.10	92.5
	Min	162	85	172	82	162	89
	Max	448	93	444.5	91	449	97
	RMSE	0.54	3.807	8.67	6.618	0	0
Hip	Mean	308.6	154.2	311.5	153.86	320.9	157.1
	Min	167	149	177	150.36	174	154
	Max	451	158	446.5	156.66	451	161
	RMSE	4.06	3.42	3.5	3.77	0	0
Left Knee	Mean	305.6	206.5	307.52	211.5	305.3	210.2
	Min	161	199	163.75	209.5	162	209
	Max	457	210	456	213	456	214
	RMSE	0.836	4.037	2.58	3.97	0	0

The Ml-GPEA approach is proven to be significant in identifying the hip and left knee coordinates in term of RMSE value. Similarly result of hip and knee are

shown in 3.6. The RMSE for hip coordinates using PM-GPEA and proposed Ml-GPEA are 3.42 and 3.77 respectively. 4.03 and 3.9 are the values of RMSE for knee coordinates by respective methods.

The result suggests that the marker less based model can be used to identify the coordinates of joints. Thus it can remove the barrier for gait analysis under well-equipped laboratory and markers based systems.

3.3.2.2 Spatio-temporal parameters estimation

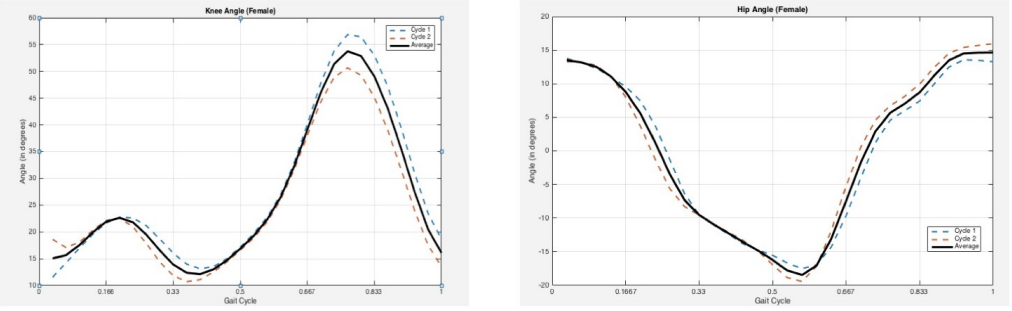
One gait cycle consists of two steps Left to Right and Right to Left. Bounding is formed around the silhouette image in each frame and ratio of width/height of the bounding box is calculated.

After processing all the frames, the ratio of the Gait cycle is presented as figure 3.16(a). The plot illustrates a smooth periodic wave with a series of local maxima and minima. Two consecutive peaks represents one step. The difference between first peak and third peak gives the number of frames in one gait cycle. Other gait parameters can be obtained by the method proposed in 3.2.2.1.

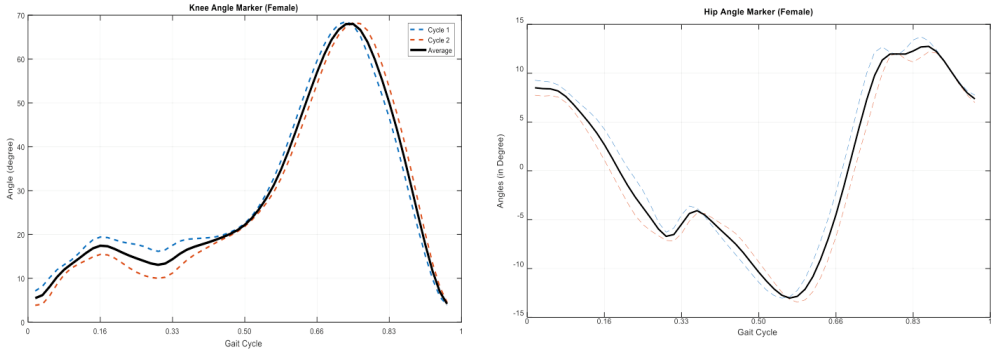
The subject considers for the result analysis is same as in 3.2.2.1. The gait parameters of subject i.e. height, width, foot length, cadence, step length, stride length, cycle time, velocity and stance and swing phase are presented in table 3.7. They are compared with actual and standards as specified by Whittle [1]and Mahyuddin [27]. Satisfactory results confirm the practicality of system. It can be seen that the step size calculated is 0.727 m, and it lies within the range given by Whittle. The cadence, and Gait Cycle Time values are also within the ranges specified by

Table 3.7: Extracted Spatio-temporal Gait Parameters using MI-GPEA Model

Gait Parameter	Ml-GPEA	**Ground Truth**	**Whittle [1]**	**Mahyuddin [27]**
Height (cm)	171.2	170	-	1.58-1.72
Width (cm)	95.75	-	-	-
Foot Length (cm)	22.98	24.1	-	-
Step Size (m)	0.727	0.705	0.68-0.85	-
Cadence (steps/min)	102	105-110	91-135	101-117
Stride Length (m)	1.45	1.41	1.25-1.85	1.12-1.28
Gait Cycle Time (s)	1.2	1.206	0.89-1.32	0.97-1.23
Velocity (m/min)	9.99	1	1.10-1.82	0.98-1.20
Stance Phase (%)	64	62	-	-
Swing phase (%)	36	38	-	-



(a) Average Knee and Hip angle for a Female Subject (S24) in constraint free evnironment using proposed marker free gait analysis



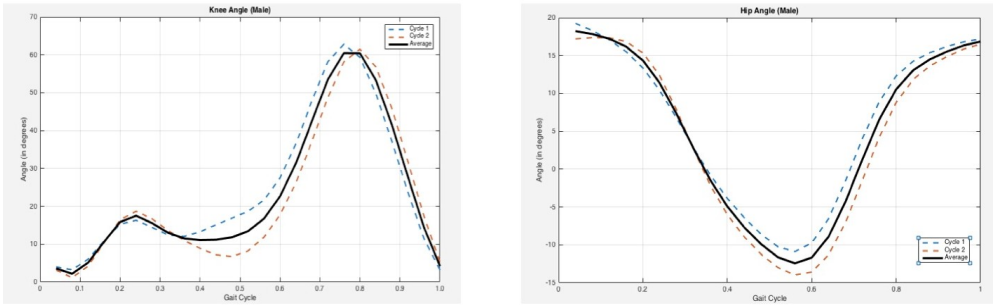
(b) Average Knee and Hip angle for a Female Subject (S24) using passive marker gait analysis

Figure 3.18: Average knee and hip angle for a female subject

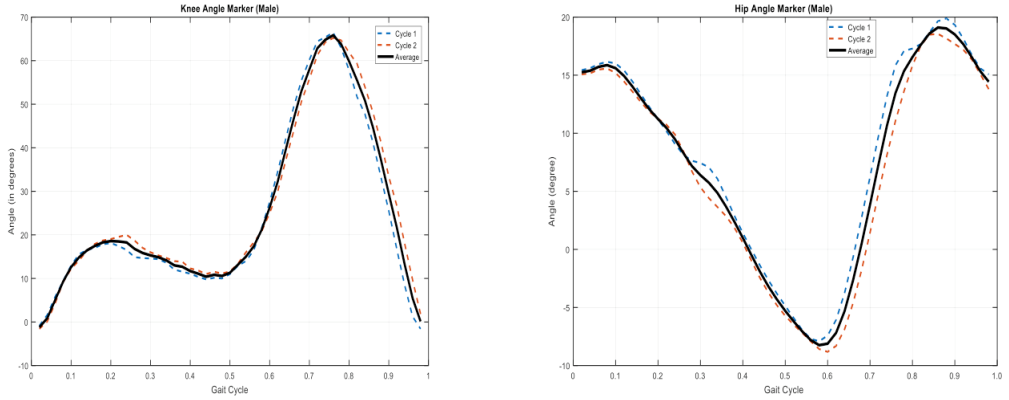
Whittle and Mahyuddin. The stance phase and swing phase are seen to be close to the ideal 60-40 ratio. However, the stride length obtained by the experiment is seen to be larger than the ideal range for an Indonesian individual, as suggested by Mahyuddin [27]. To formulate a generalized gait characteristic of an Indian individual, a broad database is required.

3.3.2.3 Kinematic parameters estimation

The kinematic parameters include different angles. In this thesis, angles focused for the gait analysis are hip and knee angles for both the legs. To find the joint angles, the Canadian Society of Biomechanics (CSB) Gait Standards is considered, as discussed in section 3.1.3. The preliminary results obtained using Ml-GPEA model for Knee and Hip angles are shown in figure 3.18. Figure 3.18(a) illustrates the knee and hip angle for the randomly selected female subject. The line in black is the average after two gait cycle. Figure 3.18(b) illustrates the angle extracted for the female subject using passive marker based approach (PM-GPEA) as discussed in section 3.2.



(a) Average Knee and Hip angle for a Male Subject (S43) in constraint free evnironment using proposed marker free gait analysis



(b) Average Knee and Hip angle for a Female Subject (S43) using PM-GPEA

Figure 3.19: Average knee and hip angle for a male subject

Similarly figure 3.19(a) represents the knee and hip angle for randomly selected male subject. The results shown in figure 3.19 are obtained after averaging two gait cycles and applying the Gaussian filter to get a smooth curve. Figure 3.19(b) illustrates the angle extracted for the male subject PM-GPEA as discussed in section 3.2 .

Cross-validation with ground truth

The plots of trajectory and angles obtained by marker less approach are cross verified. The ground truth is generated manually with the help of an expert for cross validation purpose.

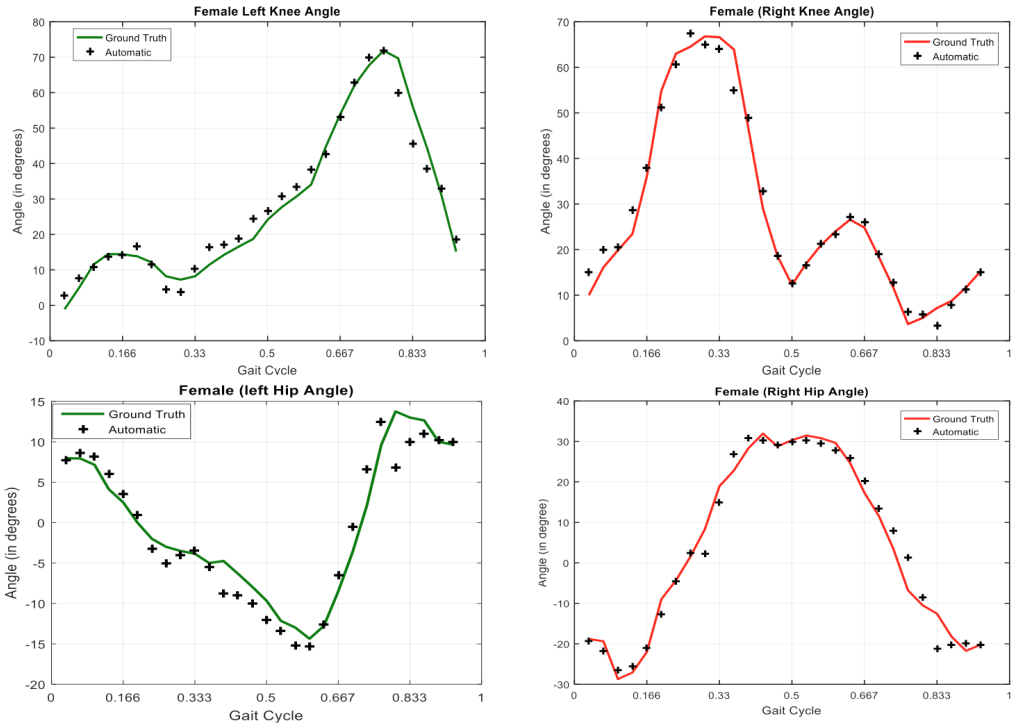
For angle evaluation, ground truth is obtained from two videos randomly chosen; frames 64 to 91 for video2 (female) and frames 68 to 92 for video1 (male). These frames include one complete gait cycle. Comparisons of the joint angle results for the proposed method using marker-less approach (Ml-GPEA) versus the ground truth are shown in figure 3.20. The x-axis corresponds to one gait cycle, and

Table 3.8: Kinematic performance for a female subject (S22) male subject (S43) in a marker-free environment

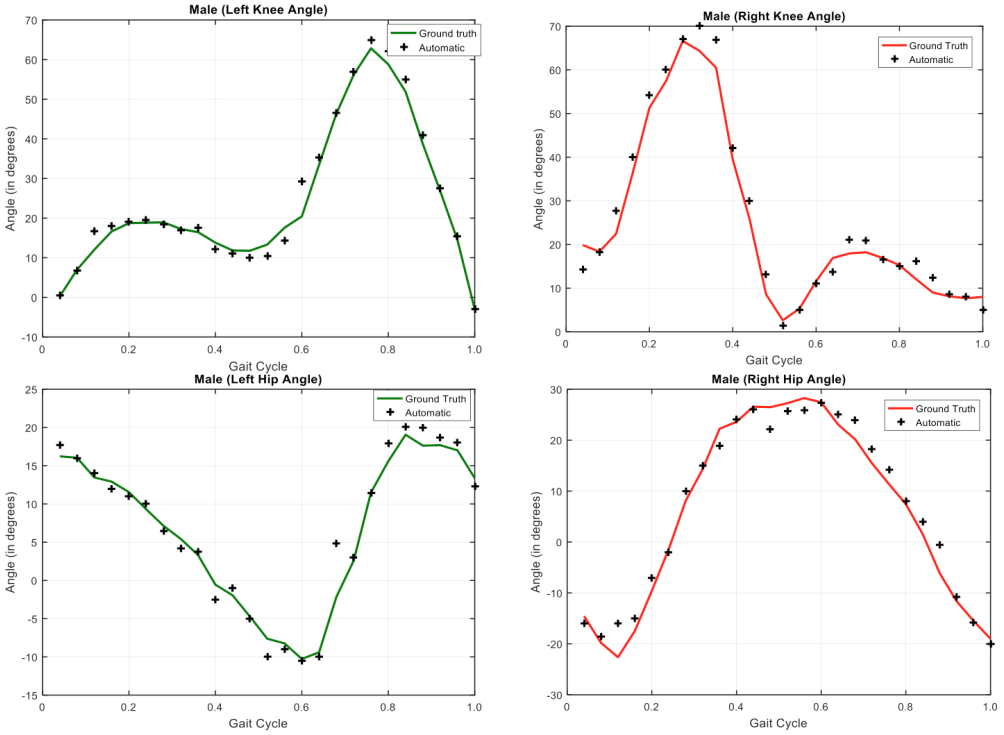
	Evaluation Parameters	**Left Hip**	**Right Hip**	**Left Knee**	**Right Knee**
Female (S22)	Total Frames	28	28	28	28
	Frames difference < 5°	26	22	24	26
	Frames difference >5 °	2	6	4	2
	Min Error	0.287	0.062	0.63	0.05
	Max Error	6.9	8.68	10.56	5.5
	Mean Error	1.859	2.926	3.805	2.141
	RMSE	2.361	3.750	3.960	2.946
	SDE	0.291	0.451	0.477	0.389
	Accuracy	92.8	78.5	85.7	92.8
Male (S43)	Total Frames	25	25	25	25
	Frames difference <5°	22	236	23	21
	Frames difference >5 °	3	2	2	4
	Min Error	0.037	0.151	0.106	0.111
	Max Error	5.53	6.62	5.94	6.59
	Mean Error	1.510	2.1069	1.861	2.652
	RMSE	2.319	2.669	2.781	3.274
	SDE	0.359	0.336	0.420	0.400
	Accuracy	88	92	92	84

the y-axis is the angle in degree. The green color represents the left side angles and red color represents the right side gait parameters. The dark line in the plot demonstrates the ground truth and the plus (+) in black is the result of the automated marker-less approach.

Table 3.8 shows the performance of the proposed method for female and male. Accuracy is calculated after considering all the frames with an error less than 5°angle. For calculating the overall accuracy, the average of left-right hip and left-right knee angles is estimated. For the female subject coordinates accuracy achieved is 87.45%. Although, this is not as good as the male subject is, but quite close to the male subject's accuracy. From table 3.8, it can be seen that worst case accuracy is for right hip where only 22 out of 28 frames have an error less than 5°angle. The mean for the average difference is 2.69°angle. The results for the male subjects' video achieved are the best with an accuracy of 89%. This shows that 89% of the total frames in the video is correctly identified with an error less than 5°angle. The mean for the average difference for all the angles is 2.032°angle.



(a) Female subject (S22) Left and right Knee and hip Angle respectively



(b) Male subject (S43) Left and right Knee and hip Angle respectively

Figure 3.20: Female and Male Left and right Knee and hip Angle respectively in a marker-free environment

3.4 MNIT Gait Dataset Analysis: A step towards INDIAN Gait Norms

Literature is full of the investigations about changes in gait which occur according to the age, gender, and nationality [7]. The result discussed so far is preliminary in nature as all subjects are not considered to generalize the gait pattern for the collected dataset. In this section, the data of the MNIT Gait dataset collected during the study has been analyzed and presented.

India population is 17 percent of the total world [159], but no record of Indian gait standard has been found out so far. Thus due to the lack of Indian gait norms, the standard of the western gait kinematic parameters are frequently used in clinical applications across India. As per the knowledge of the authors of this thesis, this can be considered as India's first gait dataset.

As discussed in section 3.1.2, MNIT gait database formed by 2400 samples (120 healthy subjects, 20 gait cycle each) is analyzed. Mean and standard deviations has been computed for the entire group and presented in table 3.2.

The INDIAN Gait dataset discussed here is using the model 1, PM-GPEA. For the validation of the results from passive markers on MNIT Gait dataset, it is compared with the result from the standard Gait lab at Bhagwan Mahaveer Viklang Sahayata Samiti (BMVSS), Jaipur as shown in figure 3.21 for the male gait sample. The lab has a 7.5m walkway and the BTS Bioengineering GAITLAB setup for the gait analysis. The results from the proposed PM-GPEA are comparable the result of standard gait lab. The result concludes the hypothesis that PM-GPEA has a wide scope for the clinical purpose. In our setup, during movement capturing, subject agreed to be in conformable and amicable state.

Figure 3.22 illustrates gait angles using cyclograms (also known as angle-angle diagrams or cyclokinograms) [160]. This diagram along with the artificial intelligence has a wide scope in prosthesis control systems.

3.4.1 MNIT Gait dataset: Adult Gait Norms

People in the age-group 18-49 are considered to have established gait, while beyond 49 years of age people tend to have degenerative aspects that affect gait parameters [7]. The result of gait spatio-temporal parameter has been presented in table 3.9.

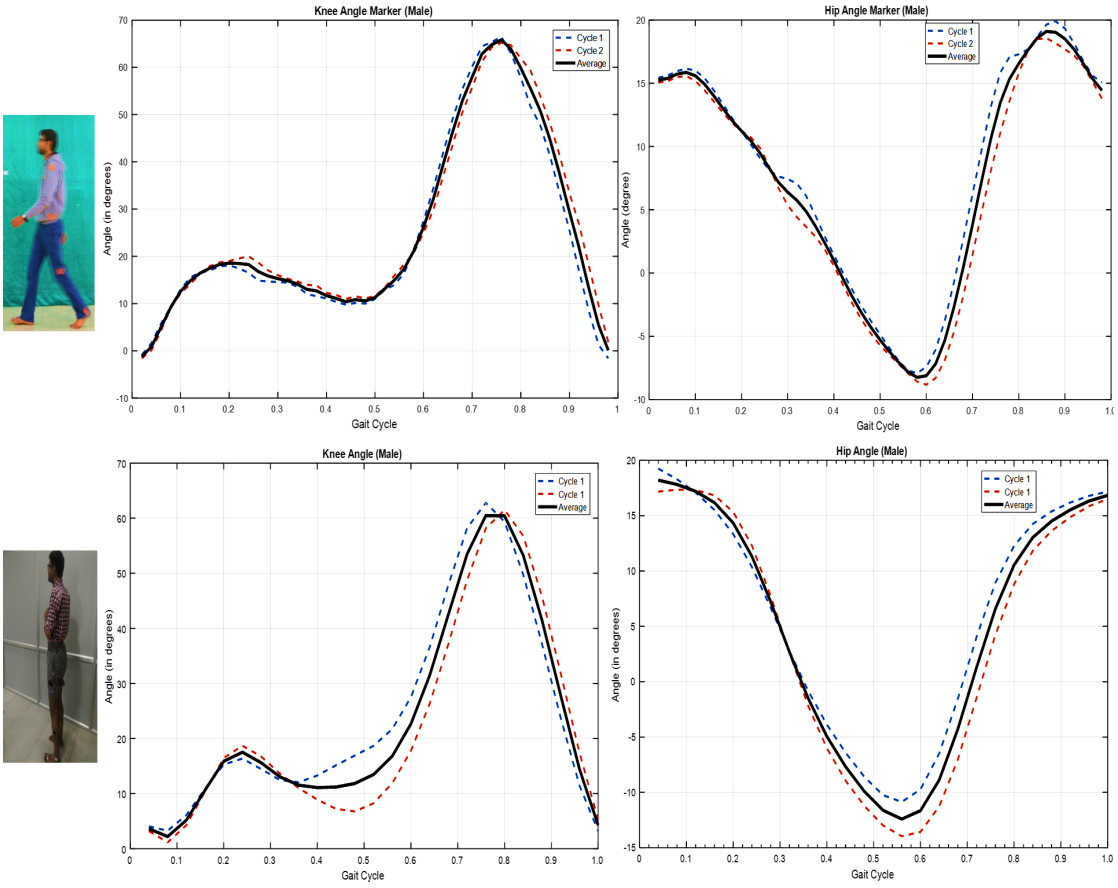


Figure 3.21: Kinematic comparison of Passive marker based and approach with the golden standard for male subject

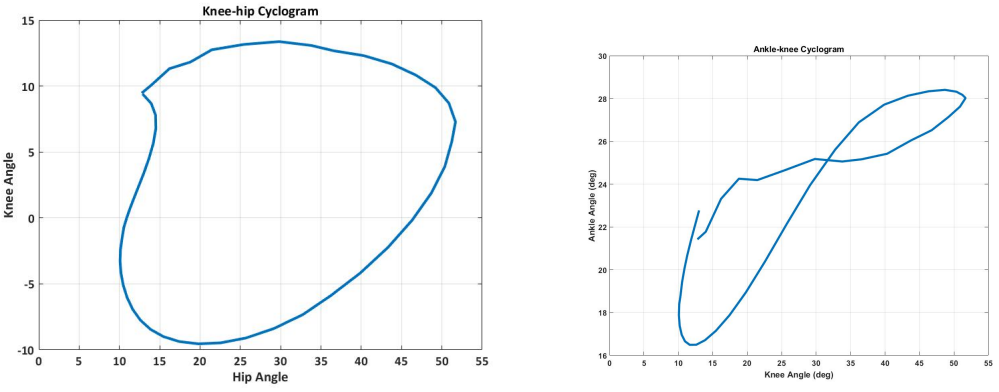


Figure 3.22: Knee-hip and Ankle-knee cyclogram at self selected speed

The stance phase in 67.32 with a standard deviation of 3.7 for the Indian scenario, while for the east it is 57.5-60.6 in the male category. The similar deviation is found for the female sample where stance phase for MNIT gait dataset is 67.73 with a standard deviation of 4.58. The result also projects that Indian people (male and female) walk faster as compared to the Mahyuddin [27] and Whittle [1] projections. The male subjects involved in this study walked with a mean speed of 1.20m/s. Table 3.9 summarizes the average values of their spatiotemporal gait parameters.

Table 3.9: Spatio-Temporal parameter data of the MNIT Gait dataset for age group 21-40

S-T Parameter	**MNIT Gait dataset**				**Mahyuddin, [27]**		**Whittle [1]**	
	Male		**Female**		**Male**	**Female**	**Male**	**Female**
	Mean	Range	Mean	Range				
Stance (%)	65.32	**63-67**	65.73	**63-67**	-	-	57.5-60.5	57.5-60.5
Swing(%)	34.59	**32-37**	34.25	**32-38**	-	-	39.0-42.0	39.0-42.0
Velocity(m/s)	1.20	**0.82-1.62**	1.11	**1.02-1.22**	0.98 - 1.2	0.9 - 1.14	1.10-1.82	0.94-1.66
Cadence (steps/minute)	136.78	**100-105**	139.40	**93-103**	101.45 - 117.13	100.58 - 120.14	91-135	98-138
Stride Length (m)	1.24	**1.12-1.26**	1.19	**1.14-1.22**	1.12 - 1.28	1.01 - 1.21	1.25-1.85	1.06-1.58
Step Length (m)	0.62	**0.56-0.62**	0.59	**0.51-0.65**	0.56-0.64	0.501-0.605	0.62-0.92	0.53-0.76
Cycle time (s)	0.95	**0.79-1.11**	0.97	**0.1-1.14**	0.97 - 1.23	0.96-1.24	0.89-1.32	0.87-1.22

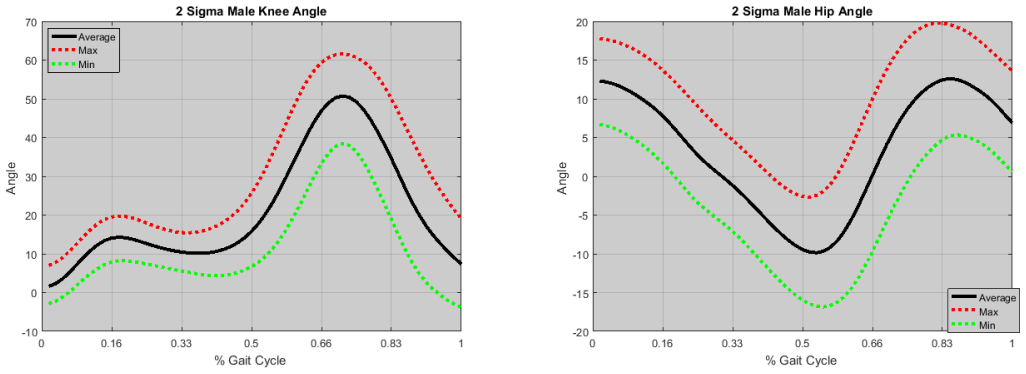
Table 3.10 presents the quantitative comparison of gait kinematics (hip,knee and ankle angle) for sagittal view, obtained in this study with the previous studies [161]. S Han et al. extract the kinematics of Chinese during activity of daily living and compared with the studies from UK, Australia, Italy, Korea [161]. The average joint angles of this study agreed well with those obtained from other same build population. The possible reason may be of the similar lifestyle.

Table 3.10: Comparisons of Sagittal based angle range of hip, knee and ankle joints during walking published in previous studies

Subjects	**UK**	**Australia**	**Italy**	**Korea**	**Korea**	**China**	**MNIT Gait dataset**
Hip	-6.3∼40.1	-13.7∼32	-10∼29.8	-4.2∼39.0	-18.3∼27.4	-5.5∼35.0	-9.5∼30.95
Knee	-8.3∼61.3	5.8∼73.7	0.4∼65.7	2.3∼59.0	1.4∼59.4	3.9∼65.6	1.8∼52.6
Ankle	-10∼16.6	-22.7∼12.4	-22.6∼10.9	-14∼11.5	-17.5∼10.8	-16.5∼31.6	-14.03∼30.75

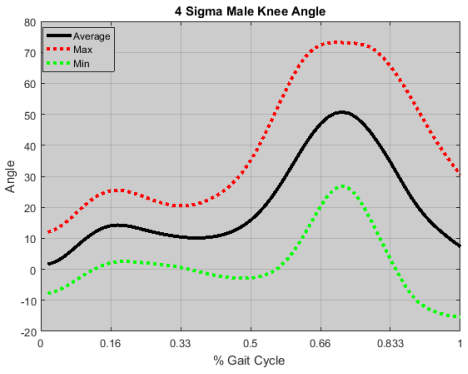
Analysis of variance (ANOVA) is a gathering of factual models and their related strategies, (for example, "variation" among and between gatherings) used to break down the distinctions among group mean [162, 163]. For the statistical analysis of experiments, a random male and female has been selected. As the information about the population is completely known by mean of its parameters, thus parametric test can be implemented.

A series of two-way repeated measures ANOVA tests were conducted at two, four and six standard deviations for the hip and knee joint angles. The degree certainty

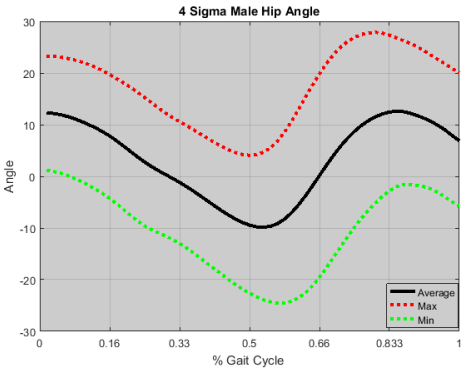


(a)

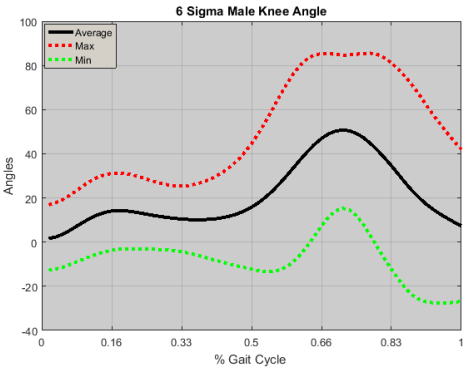
(b)



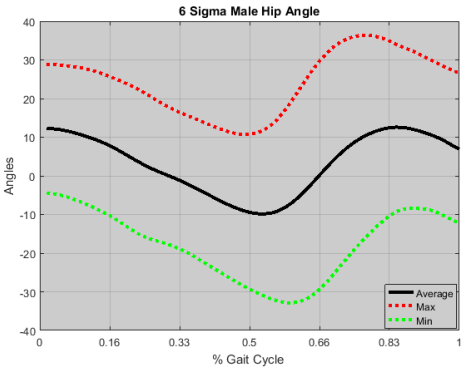
(c)



(d)

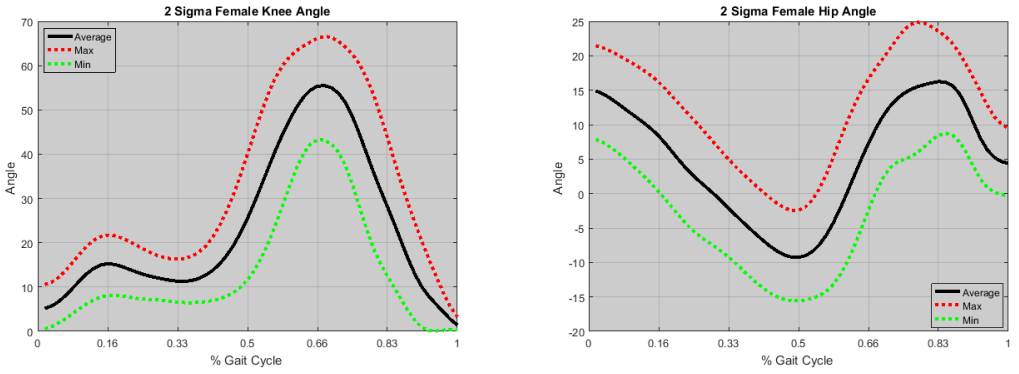


(e)



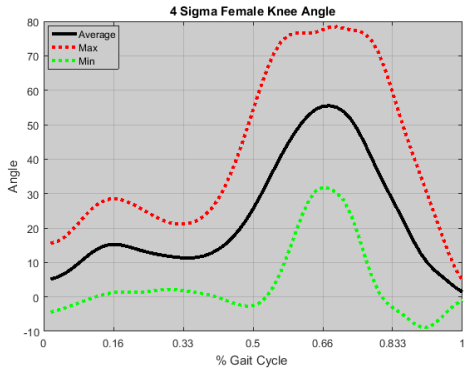
(f)

Figure 3.23: ANOVA analysis on male angles for age group 18-49 years. (a) Result using two σ male knee angle.(b) Result using two σ male hip angle. (c) Result using four σ male knee angle. (d) Result using four σ male hip angle. (e) Result using six σ male knee angle. (f) Result using six σ male hip angle.

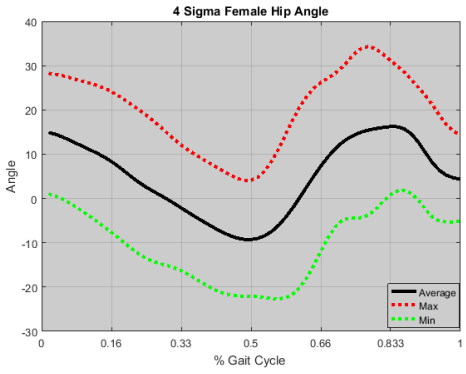


(a)

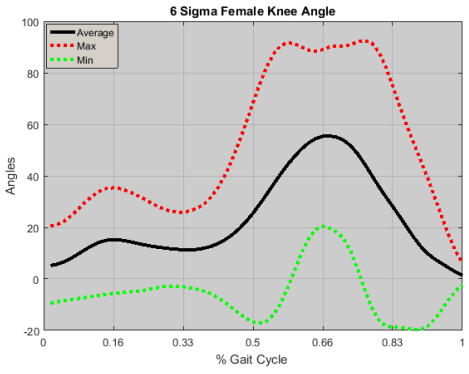
(b)



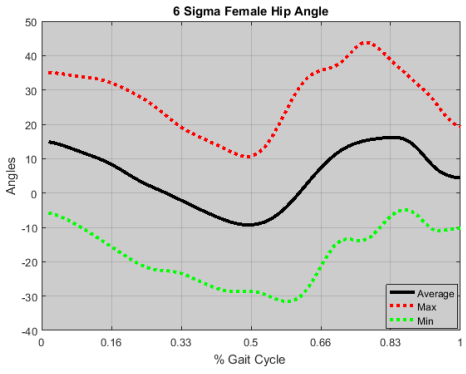
(c)



(d)



(e)



(f)

Figure 3.24: ANOVA analysis on female angles for age group 18-49 years. (a) Result using two σ female knee angle.(b) Result using two σ female hip angle. (c) Result using four σ female knee angle. (d) Result using four σ female hip angle. (e) Result using six σ female knee angle. (f) Result using six σ female hip angle.

evaluated utilizing the standard error and difference of means is correct when samples are chosen from a normal population.

Two, four and six standard deviations (sigma (σ)) has been carried out to find the minimum and maximum range of the knee and hip angle for male and female in the age group of 18-49 years. The result has been presented in figure 3.23 and 3.24 respectively. Larger value of sigma (σ) reflect the more confidence in the sample. That means larger the probability that a random sample lie in the given distribution. One standard deviation away from the mean (σ) in either direction on the horizontal axis accounts for around 68 percent of the data. Two standard deviations away from the mean accounts for roughly 95 percent of the data with three standard deviations representing about 99 percent of the data.

These separated parameters can assume a fundamental part in reconnaissance framework as a behavioral biometric. Analyst can utilize these highlights for model, fragment and group human movement. It has a noteworthy application in observing senior citizens with constrained independence at elderly reverence.

3.5 Chapter Summary

Gait analysis can be used as a valuable clinical tool to quantify the normal and pathological locomotion. But to determine gait of a subject expensive lab setup is required which limits the use of gait analysis in general environment.

Thus in this chapter, two gait parameter extraction approach (GPEA) is proposed. One with the passive marker based gait analysis with the five markers on the subject body. Another proposed approach is a markerless based joint angle identification model using camera have been presented and discussed. Markerless gait analysis system is user-friendly, and the subject is comfortable while in observation when compared with the marker based gait analysis. This technique is fully autonomous and does not require corporation of the subject. The result of the coordinates for the shoulder, hip knee and ankle and hip and knee angles are compared with the ground truth. The result confirms the hypothesis that marker-free vision based gait analysis has a broad scope of the clinical and biometric purpose. The major challenge in the markerless based gait analysis system is the extraction of silhouette images and segmentation. The results are compared with the golden standard gait setup available. The result of the passive marker based analysis is promising. PM-GPEA is used to analyze the 120 people dataset also known by

MNIT Gait dataset. The result of thus collected dataset can serve as the gait standard for Indian norms. Proposed passive marker base results when compared with the western gait standard, shows the deviation as with the nationality, the gait patterns also changes. The standard thus extracted from MNIT Gait dataset can be used in clinical applications in Indian scienerio.

Prediction models directly influence the efficiency of gait rehabilitation methods and tools. Various factors affect gait co-ordinates trajectory and joint angle prediction, are age and gender; history of injury of gait, etc. factor. Thus there is need of accurate and reliable techniques for joint co-ordinates trajectory and angle prediction.

In Chapter 4, prediction models for gait parameters has been discussed. Based on the past gait data cycle, next gait cycle pattern is estimated using the Linear Time Series (LTS), BP-NN and SVM models. Then supervised learning approaches has been explored for the gait kinematics (joint angles) that takes the body parameters as input.

Chapter 4

Development of Gait Pattern Prediction (GPP) Models

As discussed earlier, one of the objectives of this thesis work is to develop the Machine learning based Gait Pattern Prediction (GPP) framework or model for lower extremity based on the past anthropometric, kinematic and spatiotemporal data collected during the study (MNIT Gait dataset) as discussed in section 2.2. This chapter proposes two GPP models for individual specific gait parameters to be used in rehabilitation medicine.

Prediction is finding the future value for new event based on previously known or observed event data. Prediction model directly influences the efficiency of gait rehabilitation based intelligent systems [164]. In the last decade, there is a remarkable growth in both lower and upper extremity based rehabilitation assistive devices. Gait rehabilitation based robots can be classified into end effector based and exoskeleton based devices [165]. Haptic walker and Lokmat®are the examples of each class respectively. Other devices includes powered prosthetic (C-leg®), Human performance augmenting exoskeletons (HAL), assistive exoskeletons (Re-Walk ®), and etc [8, 165, 146]. These prediction model can play a vital role in the control systems in robotics rehabilitation systems. There is an acute need for more accurate prediction model for online control in case of powered based rehabilitation assistive devices [164, 166].

Section 4.1 presents a new prediction model for lower extremity joint movements (joint co-ordinates trajectory and angles) based on less amount of past gait data pattern. Later in section 4.2, GPP model has been discussed to generate gait parameters (joint angles (knee, hip), stride length, cadence, etc.) based on an-

thropometric data of targeted subjects. The results indicate the effectiveness and accuracy of the proposed prediction model for joint movements.

4.1 Individual-Specific GPP Models from historical Gait Data (GPP-HD)

Gait can be considered as stereotyped activity in healthy humans for locomotion from one position to another. Gait analysis discloses the pattern of human movement by quantifying factors responsible for the functionality of the legs. Gait analysis, not only applicable to medical, Geriatrics care, biometrics, sports, while it has been widely used in rehabilitation sector [7].

Prediction modules, using hard and soft-computing techniques have been suggested for hip, knee and ankle trajectories and joints angles in sagittal plane [167, 168, 164]. Model-based prediction, Hidden Markov Models (HMM), Gaussian Mixture Models (GMM), Gaussian Processes Latent (GPL), Probabilistic Principle Component Analysis (PPCA) are some of the developed prediction models [169, 170, 171]. Kutilek and Viteckova use Artificial Neural Network (ANN) for human gait using angle-angle diagram [172, 160]. Yun et al. suggest Gaussian process regression (GPR) for the prediction of gait kinematics [111]. Nogueira et al. explored Kalman filter and genetic algorithms for the lower limbs parameters prediction [173]. Salavka et al. use the fuzzy system for gait modeling for lower limb joints angle [174]. But these models require a significant amount of training gait data.

To overcome this limitation, two new simple linear time series models based-on the least square method, using minimal previous historical gait pattern data are proposed. Persistent supervised models (BP-ANN and SVM) are also explored.

4.1.1 Data Selection from MNIT Gait Dataset

A total of 120 healthy subjects data is collected in the RAMAN Lab as discussed in section 3.1. The output of the MNIT Gait dataset, i.e., spatial-temporal and kinematics are considered in the study. Two gait cycle outputs for each of hip and knee angle time series are used in this analysis. Later, the coordinates of the markers (joint trajectories) are considered in this thesis.

Table 4.1: Sample gait pattern data of a specific subject for p gait cycle .

Gait Cycle⟶ **Sample ↓**	GC_1	GC_2	...	GC_{p-1}	GC_p
1	$X_{GC_1}(1)$	$X_{GC_2}(1)$	...	$X_{GC_{p-1}}(1)$	$X_{GC_p}(1)$
2	$X_{GC_1}(2)$	$X_{GC_2}(2)$	...	$X_{GC_{p-1}}(2)$	$X_{GC_p}(2)$
3	$X_{GC_1}(3)$	$X_{GC_2}(3)$	...	$X_{GC_{p-1}}(3)$	$X_{GC_p}(3)$
	...	...	...	...	...
n-1	$X_{GC_1}(n-1)$	$X_{GC_2}(n-1)$	...	$X_{GC_{p-1}}(n-1)$	$X_{GC_p}(n-1)$
n	$X_{GC_1}(n)$	$X_{GC_2}(n)$	...	$X_{GC_{p-1}}(n)$	$X_{GC_p}(n)$

In the MNIT gait dataset, we have a total of 20 Gait Cycle (GC) for each subject; thus total GC or walking pattern is 2400 walking trials. The purpose of this work is to build gait pattern for the specific subject. A gait pattern of target subject S22 having seven GC is shown in table 4.1. n represent the total no of time-frame/samples in one GC. For the representation, we have considered p sample as seven for the selected subject S22 as shown in figure 4.1. In the collected dataset, one Gait Cycle varies from 55-64 samples/ time-frames. The X in the table 4.1 presents the predicted gait patterns (either knee angle, hip angle, joint coordinates or the target gait parameter) for the specific subjects.

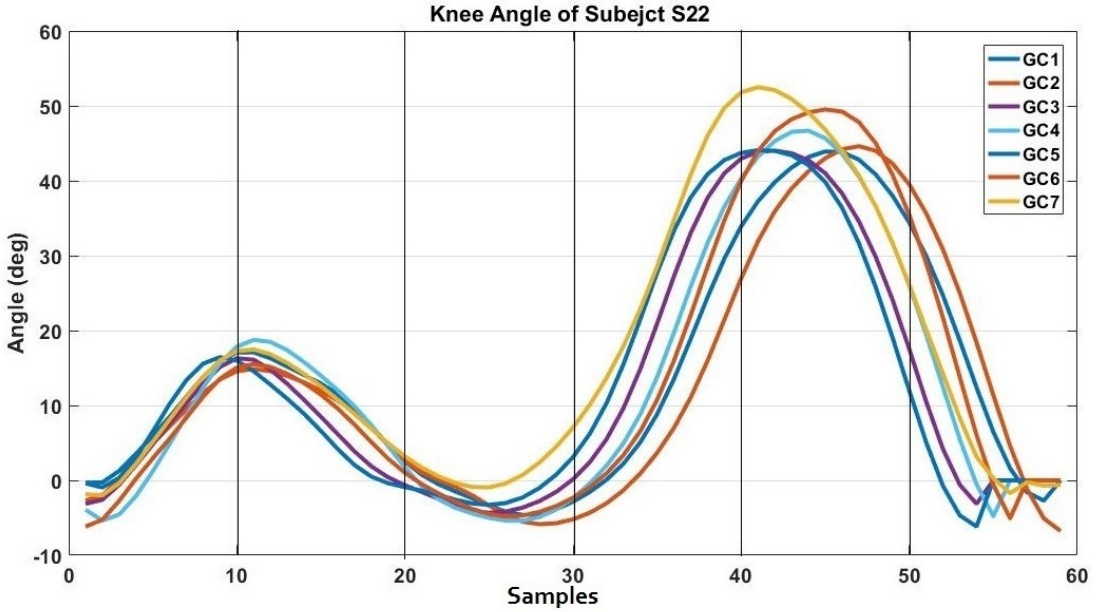


Figure 4.1: Plot of male subject (S22) knee angle, seven samples of Gait cycle (GC=7)

4.1.2 Linear Time Series (LTS) Model

Two models (univariate and multivariate based LTS) are developed and analyzed in this thesis.

4.1.2.1 Univariate Linear Time Series (ULTS) Model

A univariate time series model is based on using the p historical gait cycle data (knee, hip angles, and coordinates) from the previous GC pattern (GC_p) and predicts the coordinates for the next GC pattern (GC_{p+1}) for the same time interval. Thus by previous n frames of (GC_p), next (GC_{p+1}) is predicted. In this study, one gait cycle is composed of 59 data samples. The univariate model can be expressed as (4.1)

$$Y_{GC_{p+1}}(i) = a + bX_{GC_p}(i) \quad where \quad i = 1, 2, ..., n \tag{4.1}$$

where i is the no of samples in the GC, $Y_{GC_{p+1}}(i)$ is the pth gait cycle pattern sample, predicted gait pattern (GC_{p+1}) co-ordinates and joint angles), $X_{GC_p}(i)$ represents the corresponding actual data-points for the X_{GC_p} joint movement (co-ordinates and angles) data-points, i represents the data samples. Model variables a and b are evaluated from least-square techniques as shown in eq(4.2)

$$\begin{gathered} C = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \beta^T & \beta \end{bmatrix}^{-1} .\beta^T .X_{GC_p} \\ where \quad \beta = \begin{bmatrix} 1 & X_{GC_p}(1) \\ 1 & X_{GC_p}(2) \\ : & : \\ 1 & X_{GC_p}(n) \end{bmatrix} \qquad X_{GC_p} = \begin{bmatrix} X_{GC_p}(1) \\ X_{GC_p}(2) \\ : \\ X_{GC_p}(n) \end{bmatrix} \end{gathered} \tag{4.2}$$

In ULTS, GC_{p+1} can be predicted using the previous GC_p Gait cycle of the required prediction parameter.

4.1.2.2 Multivariate Linear Time Series (MLTS) Model

The second model is the generalized form of the univariate time series based prediction model. This model is uses p previous series to predict the next $p + 1$ data

samples for the GC using (4.3)

$$\begin{gathered} Y_{GC_{p+1}}(i) = a + b_1 X_{GC_1}(i) + b_2 X_{GC_{p-1}}(i) + b_k X_{GC_p}(i) \\ where \quad i = 1, 2, ..., n \end{gathered} \tag{4.3}$$

where $Y_{GC2}(i)$ is the predicted points for the GC_2 co-ordinates and angles data values, $X_{GC_1}(i)$ represents the corresponding actual data-points for the GC_1 Joint movement (coordinates and joint) data-points, i is the number of frames, k is taken as $p+1$. Model variables a and b are evaluated from least-square techniques as shown in expressed in(4.4)

$$\begin{gathered} C = \begin{bmatrix} a \\ b_1 \\ b_2 \\ : \\ b_{p+1} \end{bmatrix} = \begin{bmatrix} \beta^T & \beta \end{bmatrix}^{-1} .\beta^T .X_{GC_p} \\ where \quad \beta = \begin{bmatrix} 1 & X_{GC_p}(1) & X_{GC_{p-1}}(1) & \dots & X_{GC_1}(1) \\ 1 & X_{GC_p}(2) & X_{GC_{p-1}}(2) & \dots & X_{GC_1}(2) \\ : & : & : & : & : \\ 1 & X_{GC_p}(n) & X_{GC_{p-1}}(n) & \dots & X_{GC_1}(n) \end{bmatrix} \end{gathered} \tag{4.4}$$

GC-6,and GC-7 in table 4.1 are defined as the input test vectors of LTS variant and the other five GC are defined as the input vectors LTS in training set with $p = 5$.

4.1.3 Back Propagation Artificial Neural Network (BP-ANN)

The concept of the neural network was developed with the idea of imitating biological neuron for taking the intelligent decision. Back-propagation is a variant of the supervised learning based artificial neural network, invented in 1970. Late in 1986, Rumelhart, Hinton, and Williams describe several neural networks where backpropagation works far faster than earlier approaches to learning [175]. Figure

4.2 presents the architecture of BP-ANN with n input, two hidden layer, and m output.

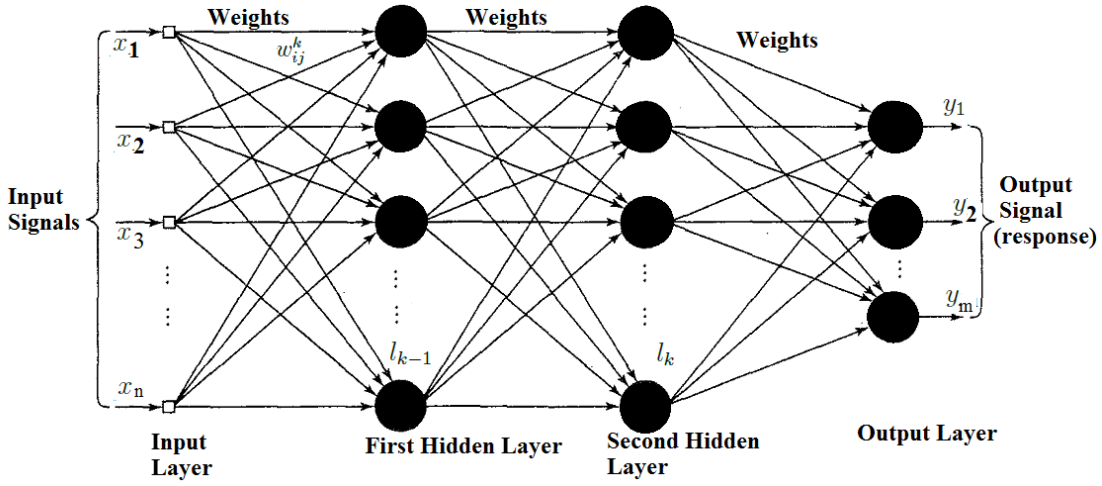


Figure 4.2: Architecture of Back Prorogation Neural Network (BP-ANN) for GPP-HD model

BP-ANN follows below steps :

STEP 1: DATASET: Set of input-output pairs of size N is denoted as $[X, Y] = \{(x_1, y_1), \ldots, (x_N, y_N)\}$ input-output pairs (x_i, y_i), where y_i is the desired output of the network on input x_i.

STEP 2: A feedforward neural network: w_{ij}^k , the weight between node j in layer l_k and node i in layer l_{k-1}, the bias (b_i^k) for node i in layer l_k

STEP 3: Error Function: $E(X, \theta)$, define error between the desired output$y$ and the actual output y_i' on input x_i.

Training a neural network with gradient descent requires the calculation of the gradient of error function $E(X, \theta)$ with respect to weights w_{ij}^k and biases b_i^k. Then, according to the learning rate α , each iteration of gradient descent updates the weights and biases collectively denoted θ according to following equation

$$\theta^{t+1} = \theta^t - \alpha \frac{\partial E(X, \theta^t)}{\partial \theta} \tag{4.5}$$

4.1.3.1 Parameter Settings

Trail and error method choose the number of hidden layer neurons. The training set consists of p past patterns. All other network and training parameters are carefully chosen to reduce the sum squared error. Hidden layer selected is one with

two neurons, and the presented evaluation measure is the average of 30 iterations as error converges at this point.

4.1.4 Support Vector Regression (SVR)

Supervised Learning method performs binary classification and regression analysis. SVM differentiates between two classes by constructing a hyperplane. SVM tries to maximize the distance between the hyperplane constructed and the closest training samples on either side of the hyperplane.

In 1996 Vladimir et al. proposed the regression version of SVM, known as Support Vector Regression (SVR) [176]. The concept behind SVR is to map nonlinearly the original data x into a higher dimensional feature space [177]. Consider a set of data of gait pattern D.

$$D = (x_i, d_i) \tag{4.6}$$

where, x_i and d_i and are input vector and desired values respectively. The SVR function is shown in Eq.(4.7)

$$Y = g(x) = W\psi(x) + b \tag{4.7}$$

where, $\psi(x)$ is the feature obtained after mapping. Regularized risk function $R(C)$ is presented in Eq.(4.8) is used to find the coefficients W and b

$$R(C) = (C/N)\sum_{i=1}^{N} P_\epsilon(d_i, y_i) + \frac{|W^2|}{2} \tag{4.8}$$

where,

$$P_\epsilon(d, y) = \begin{bmatrix} 0 & if|d-y| <= \epsilon \\ |d-y| - \epsilon & otherwise \end{bmatrix} \tag{4.9}$$

C is trade-off between empirical risk and model flatness and error ϵ are prescribed parameters and N is the number data pattern.

The Eq. (4.8) can be written as

$$minR(W, \mu, \mu^*) = \frac{|W^2|}{2} + C(\sum_{i=1}^{N}(\mu_i + {\mu^*}_i)) \tag{4.10}$$

With following constraints,

$$W\psi(x_i) + b - d_i \leq \epsilon + \mu^*{}_i \tag{4.11}$$

$$d_i - W\psi(x_i) - b \leq \epsilon + \mu_i \tag{4.12}$$

$$\mu_i, \mu^*{}_i \geq 0 \tag{4.13}$$

where i= 1, 2, 3.....N

Different kernels like linear, polynomial, radial basis, Gaussian, Laplacian kernel, exponential kernel are used. Gaussian kernel is giving the best result in forecasting of the gait pattern [177]. Thus Gaussian kernel function is used in this study.

4.1.5 Results and Discussion: GPP-PD Models

In this section the result generated using the proposed models has been presented.

4.1.5.1 Sample Selection from MNIT Gait Dataset

For the statistical analysis of the proposed model, eight subject's gait patterns from MNIT Gait dataset are considered as below groups: two from the first group, six from 18-49 age group and two from third age group are selected. The age classification follows that of Whittle [1] and shown in table 4.2. One male subject (S59) and one female subject (S65) is selected for the analysis in the below 18 age group. S21, S32 and S100 ; three male subjects and three female subjects S36, S71 and S88 are selcted in the 18-49 age group. One male (S123) and one female subject (S132) are considered for analysis in above 50 age group.

Table 4.2: Subject selection from MNIT Gait dataset based on age-groups and genders

Age-Group (years)	**Male Selected**	**Female Selected**	**Total Selected**
< 18	S59	S65	02
18-49	S21, S32, S100	S36, S71, S88	06
>50	S123, S132	-	02
	Total		10

4.1.5.2 Evaluation Parameters

To demonstrate the efficiency and accuracy of the model, a comparison is made to examine the difference between the joint angle predicted by this model and the actual angle using following evaluation parameters:

- Mean Absolute Deviation (MAD) :

$$MAD_j = \sum_{n=0}^{N-1} \frac{|y_i - \widehat{y}|}{N} \tag{4.14}$$

  where N is the no of samples in one gait cycle , n is the joint index, $\widehat{y}$ is the predicted angle and y is the actual angle of the joint j.

- Mean Absolute Error (MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |y_i - \widehat{y}| \tag{4.15}$$

- Root Mean Square Error (RMSE):

$$RMSE = \sqrt[2]{\frac{1}{n} \sum_{i=1}^{M} (y_i - \widehat{y})^2} \tag{4.16}$$

- Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \frac{|y_i - \widehat{y}|}{y_i} \tag{4.17}$$

Where y_t and $\widehat{y_t}$ are the actual and predicted gait movements at frame, i and n is the no of frames.

4.1.5.3 Discussion

Multivariate time series with the $p = 6$ model has the lowest MAD, MAE, RMSE, and MAPE for all the selected subjects when compared to other models (BP-ANN, SVR) model. Figure 4.3 illustrates the performance plot for hip angle using the proposed and BP-ANN based model for subject S21.

.

Table 4.3: MAE, RMSE, MAPE measures for the predicted knee Angle

Subjects	Subjects	Prediction Models	MAD	MAE	RMSE	MAPE
Group 1	S59	BP-ANN	2.8408	3.2521	2.9034	17.7204
		SVR	3.3488	4.1653	5.222	44.9816
		Univariate Time Series	2.1154	2.1085	2.4316	13.1979
		Multivariate Time Series (p=3)	1.7385	1.4753	1.6264	12.7006
		Multivariate Time Series (p=6)	1.6395	1.3853	1.5003	11.6185
	S65	BP-ANN	3.071	3.4823	3.1336	17.9506
		SVR	3.579	4.3955	5.4522	45.2118
		Univariate Time Series	2.3456	2.3387	2.6618	13.4281
		Multivariate Time Series (p=3)	1.9687	1.7055	1.8566	12.9308
		Multivariate Time Series (p=6)	1.8697	1.6155	1.7305	11.8487
Group 2	S21	BP-ANN	2.0765	2.4878	2.1391	16.9561
		SVR	2.5845	3.401	4.4577	44.2173
		Univariate Time Series	1.3511	1.3442	1.6673	12.4336
		Multivariate Time Series (p=3)	0.9742	0.711	0.8621	11.9363
		Multivariate Time Series (p=6)	0.8752	0.621	0.736	10.8542
	S32	BP-ANN	5.0764	5.4877	5.139	19.956
		SVR	5.5844	6.4009	7.4576	47.2172
		Univariate Time Series	4.351	4.3441	4.6672	15.4335
		Multivariate Time Series (p=3)	3.9741	3.7109	3.862	14.9362
		Multivariate Time Series (p=6)	3.8751	3.6209	3.7359	13.8541
	S36	BP-ANN	2.5452	2.9565	2.6078	17.4248
		SVR	3.0532	3.8697	4.9264	44.686
		Univariate Time Series	1.8198	1.8129	2.136	12.9023
		Multivariate Time Series (p=3)	1.4429	1.1797	1.3308	12.405
		Multivariate Time Series (p=6)	1.3439	1.0897	1.2047	11.3229
	S88	BP-ANN	2.9331	3.3444	2.9957	17.8127
		SVR	3.4411	4.2576	5.3143	45.0739
		Univariate Time Series	2.2077	2.2008	2.5239	13.2902
		Multivariate Time Series (p=3)	1.8308	1.5676	1.7187	12.7929
		Multivariate Time Series (p=6)	1.7318	1.4776	1.5926	11.7108
	S100	BP-ANN	4.0641	4.4754	4.1267	18.9437
		SVR	4.5721	5.3886	6.4453	46.2049
		Univariate Time Series	3.3387	3.3318	3.6549	14.4212
		Multivariate Time Series (p=3)	2.9618	2.6986	2.8497	13.9239
		Multivariate Time Series (p=6)	2.8628	2.6086	2.7236	12.8418
Group 3	S132	BP-ANN	3.3461	3.5698	3.9756	17.1979
		SVR	3.0876	3.9405	4.7945	46.2376
		Univariate Time Series	1.7115	1.3442	1.6673	12.4336
		Multivariate Time Series (p=3)	0.8342	0.7110	0.8621	11.9363
		Multivariate Time Series (p=6)	0.7215	0.6614	0.7830	8.5652

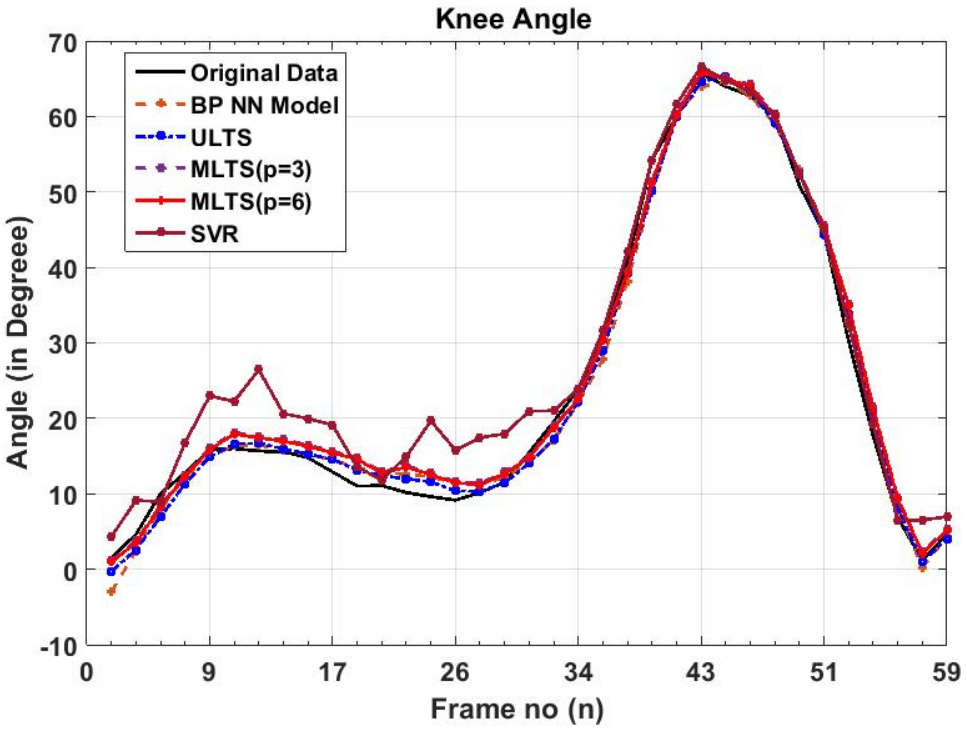
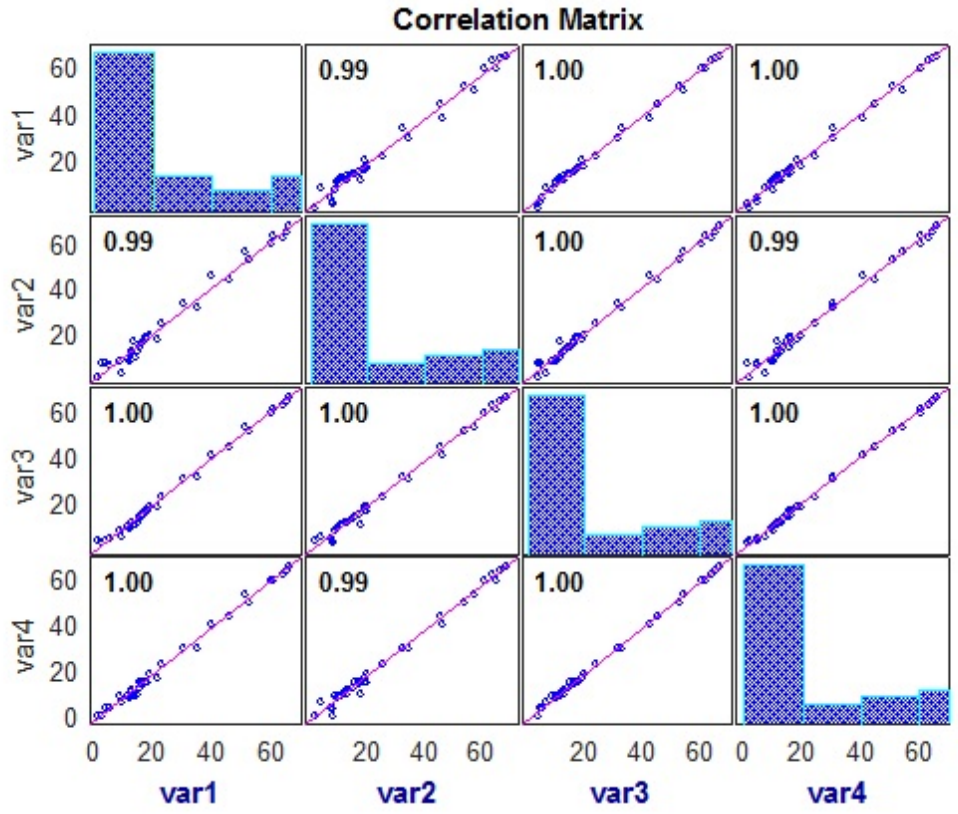


Figure 4.3: Performance plot for knee angle prediction using various model for SS21

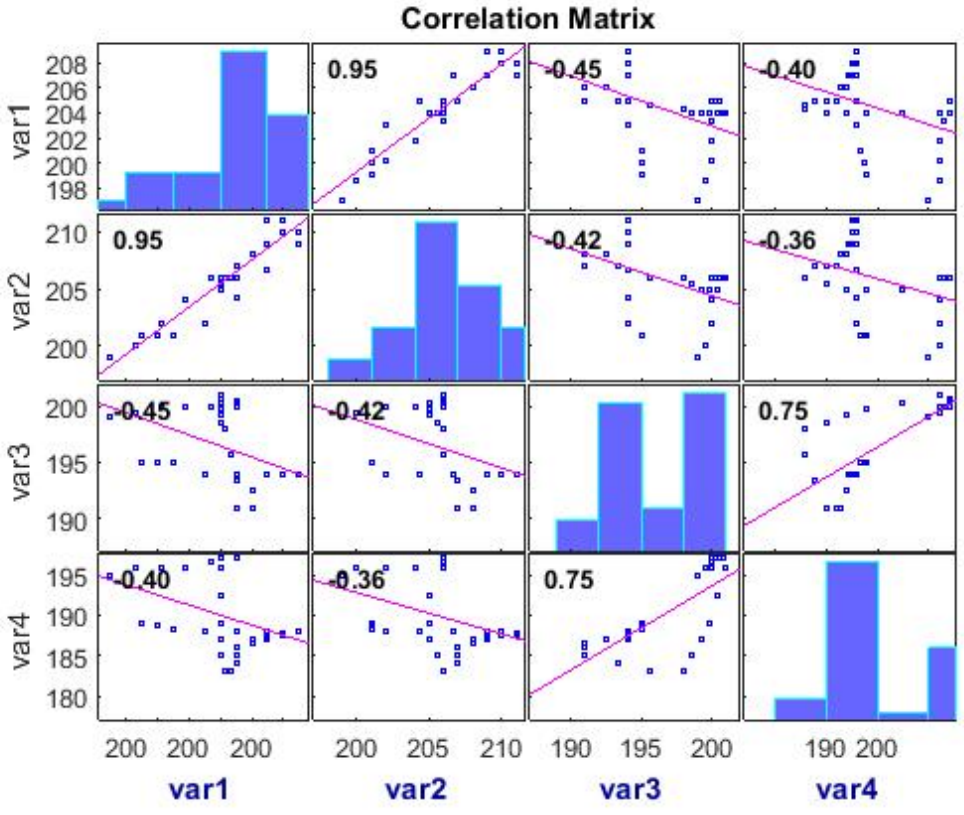
The Multivariate time series model predicts the knee angle more precisely when compared with other models. BP-ANN, MAPE value is 16.9561 while that of the proposed MTS with P=3 is 11.9363 and with p=6 is 10.8542 for subjct S21. There is an improvement of around 5% in MAPE. The possible reason is higher correlation coefficient among the dataset as presented in Fig 4.4. Similarly for the knee cordinates, univariate time series prediction model outperform than other models. MLTS with p=6 shown higher MAPE in all the cases. The reason is that the historical data is not linearly correlated as shown in Fig 4.4.

The results reveal the higher accuracy of the proposed model for gait movement prediction for lower extremities. After rigorous analysis of the result and justifying the better accuracy obtained, it is found out the previous data used in the study are highly co-related for all the subjects. The correlation coefficient is very high in the selected subjects gait pattern. Thus it can be concluded that the proposed model works better with linearly co-related data.

The proposed model is individual specific and based on past historical data of the subject under consideration. But the historical data is not always available to us. Another limitation of person-specific model that utilizes the predefined gait trajectory and kinematics. In the proposed GPP model, the predefined gait pat-



(a)



(b)

Figure 4.4: Co-relation coefficient of (a) Knee angle dataset (b) Hip angle for subject S32

tern considers height, age, weight, etc. body parameters. These body parameters influence the gait patterns [178]. Thus there is a requirement of an alternative approach that can generate the gait pattern from the anthropometric data of the subject as these parameters are easy to calculate. In the next section gait pattern prediction model is explored that generate the gait profile of the subject by considering the body parameters.

4.2 GPP Models from Anthropometric Data (GPP-AD)

The anthropometric parameters usually consider corporeal dimensions of the human includes age, gender, weight, height, limb length and Body Mass Index (BMI) of a subject as mentioned previously in section 2.2 i.e. height, weight, limb length, etc. The gait pattern possesses inherent variability, which depends on differences in body parameters such as age, gender, weight, height and so on [179, 180]. Researchers and engineering rehabilitation community are trying to find the relation between body parameters and human gait kinematics for several years because various variables influence gait kinematics. Previous studies have employed statistical methods to investigate the effects of body features on the gait motion (e.g., gender and age) [178]. Samson et al. measured walking speed, stride length, and cadence of 118 women and 121 men over a wide range of age (19-90 years). Linear regression was used to analysis the effect of age, gender, height and weight on these spatiotemporal parameters [181]. These research provide new insights into walking dynamics, and health care professionals can utilize these facts in taking decisions during gait-related rehabilitation. Diagnosis of gait abnormalities highly depends on the gait profile of a given demographically match cohort subjects [7].

Model-based optimization techniques utilize the human limb dynamics modeling to find the gait pattern [182, 111]. But the issue with gait dynamics models is that everybody walks differently. Mechanical and biomechanical models needed to predict the human gait pattern, and thus it is not possible to capture the randomness in the walking pattern. Researchers advocate group based analysis on the similar anthropometric parameter for gait analysis. It is essential to isolate anthropometric effects on gait analysis.

Statistics-based methods are also investigated for the gait pattern prediction model [111]. However, these studies, however, analyze the effect of a single factor, for

example, gender, or age or weight. Song et al. proposed prediction model from segment length and predict kinematic parameters using correlated static and dynamic using probabilistic graphical model, EM algorithm, Decision tree [53]. Recently in 2017, Liu et al. propose a Deep Spatial-Temporal Model (DSTM) for generating knee joint trajectory of lower-limb exoskeleton [183].

4.2.1 Methodology for the GPP-AD Models

Figure 4.5 shows the methodology used in this study for the gait profiling based on the anthropometric parameters. The anthropometric data of the MNIT gait dataset is used in the prediction models for the gait pattern. In this study we have explored BP-ANN (with and with Principal Component Analysis (PCA)) and Gaussian Process Regression models for the gait pattern prediction based on anthropometric data parameters (GPP-AD), using the MNIT Gait Dataset.

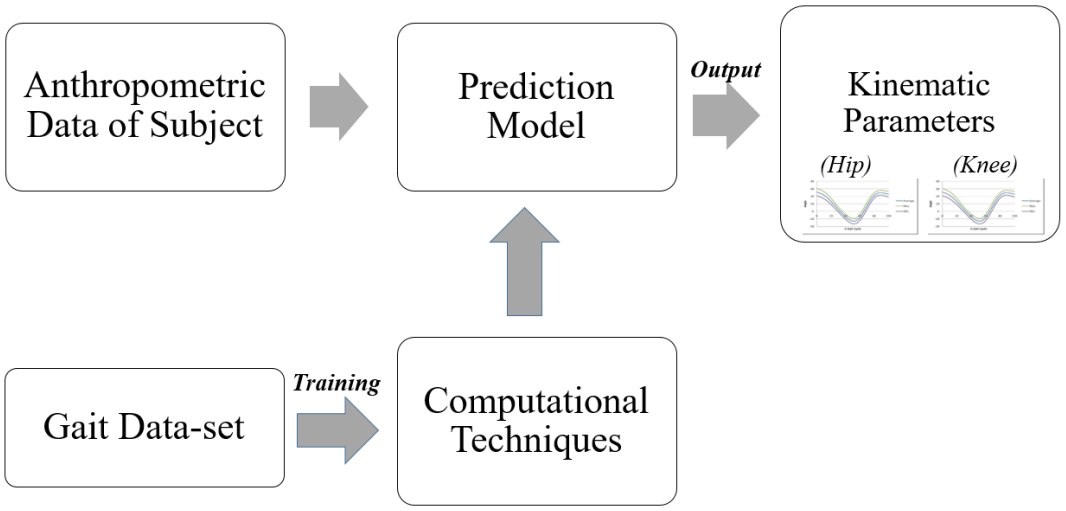


Figure 4.5: Methodology for Gait Pattern Prediction Model from Anthropometric parameters (GPP-AD)

4.2.1.1 Data selection for GPP-AD Models

A large number of elements are reported to influence gait pattern for a subject, including age, sex, height, weight, muscle to fat ratio, muscle quality, and even mental condition. It is difficult and moreover unwanted to incorporate all one of the variables in the investigation and forecast of gait patterns with a measurable technique.

Anthropometric information collected during the collection of MNIT Gait Dataset is used to the explored GPP-AD models. Ten delegates human body parameters which significantly influence the gait profile are considered in this thesis. Table 4.4 lists the mean values and standard deviations for all the body parameters for the 120 MNIT gait subjects. Besides gender, age, height, weight, Body Mass Index (BMI), thigh and calf length, malleolus height, foot length, waist width and brilliance width is considered in the study.

Table 4.4: Mean and standard deviation of 120 experiment participantBP-ANNs body parameters from MNIT Gait dataset.

Body parameters	Mean	Std. deviation
Age (Year)	28.10	8.83
Height (m)	1.68	0.11
Weight (Kg)	67.05	14.59
BMI	23.55	3.85
Thigh Length (m)	0.48	0.04
Calf Length	0.39	0.08
Mallelous height (cm)	0.10	0.14
Foot Length (cm)	0.25	0.02
Waist width (inches)	32.88	3.07
Bi-illiac width (m)	0.83	0.08

The outputs of the GPP-AD model considered in this study are the joint angle (hip, knee, ankle) and spatiotemporal parameters(stride length, velocity, and cadence). These parameters are required to generate the gait profile of the target subject. Figure 4.6 shows the aggregate knee angles of the 120 subject considered in the study.

4.2.2 Principal Component Analysis (PCA) and BP-ANN

We have considered 10 feature as input for learning the model, but these input or features are passed for learning the model. The larger number of the inputs slows down the learning of the model. Thus, there is need for a suitable approach that keeps the information or pattern in the data and reduces the number of redundancies or data duplicacy for unbiased and robust model learning. Thus, before building a predictive model, it is advocated to reduce the dimensionality of the data.

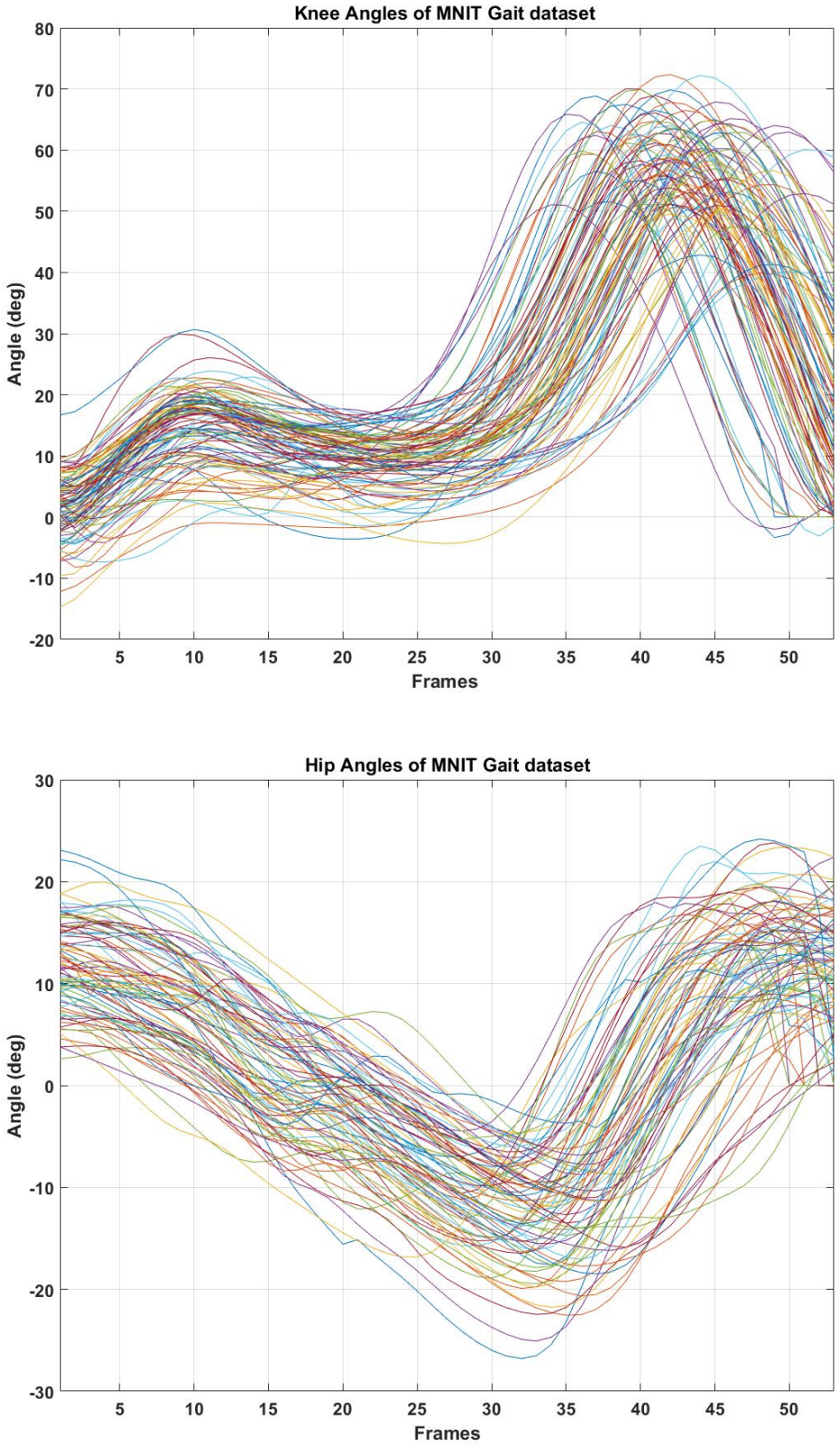


Figure 4.6: Aggregate knee and hip angle of 120 subject's of MNIT Gait Dataset

Principal Components Analysis (PCA) is an important multivariate statistical technique that is often used for feature reduction by considering the higher principal components [184]. It is a linear-algebra based projection technique that maps highly correlated features into a low dimensional space. The resultant feature vectors in this sub-space are non-correlated and orthogonal to each other. Mathematically, it represents an orthogonal basis transformation wherein the input data vectors are projected onto new basis vectors of m dimensional feature space, where m represents the number of features in the original data.

Let $X \epsilon R^m$ be a real-valued matrix having n samples or where each sample vector is a column. Hence X is arranged as a $m \times n$ data matrix which has to be represented by Y, of the same dimensions. Let P be a transformation matrix that converts X into a representation matrix Y such that the features of Y follow the two conditions mentioned above. Accordingly, P must be a $m \times m$ matrix that transforms X as

$$PX = Y \tag{4.18}$$

The covariance matrix of Y is given by equation 4.2:

$$Cov(Y) = \frac{1}{n-1} YY^T \tag{4.19}$$

the diagonal element of $Cov(Y)$ is the variance of each dimension while the off-diagonal elements are the covariance between different dimensions. It can be represented by the algorithm shown in 2.

Algorithm 2 Principal Component Analysis (PCA)

1: Arrange data sample vectors into a data matrix X = samples $\times$ observation
2: Mean centering of X $\longrightarrow$ mean center X by subtracting column means from each observation along the columns to have zero means along the columns.This results in matrix A
3: Compute the covariance matrix of the mean-centered data matrix A by $\frac{A^T \times A}{n-1}$. It is a square symmetric matrix of $p \times p$ dimension
4: Find the eigen vectors of the covariance matrix of A
5: Sort the eigen vectors in decreasing order of eigen values
6: Project the sample vectors of A onto these eigen vectors to find their feature vectors for building prediction models

After PCA, highly correlated features are mapped into low dimensional space, the BP-ANN model, as shown in section 4.1.3.

4.2.3 Gaussian Process Regression (GPR)

Linear regressions describes relationships between variables but sensitive to outliers. A Gaussian Process (GP) is defined as a probability distribution over functions $y(x)$, such that the set of values of $y(x)$ evaluated at an arbitrary set of points $x_1, x_2 \ldots x_n$ jointly have a Gaussian distribution. GP governed by prior covariances, as opposed to a piecewise-polynomial spline chosen to optimize smoothness of the fitted values.

Gaussian Process Regression (GPR) is a variant of nonlinear regression that provides probabilistic prediction and an estimate of uncertainty in the prediction [185]. Thus there are fewer chances of over fitting the data. GPR is nonparametric kernel-based probabilistic models. The reason to choose the GPR is that general GPR is equivalent to Bayesian linear regression with an infinite number of basis functions.

4.2.4 GPP-AD : Results and Discussions

After predicting the gait profile for the test subject using their anthropometric data, the actual and predicted gait patterns are quantitatively compared using the evaluation parameters (MAE, MSE, RMSE and MAPE) as discussed earlier in section 4.1.5.3.

Table 4.5: Result for kinematics parameters prediction model from anthropometric parameters using Artificial Neural network with and without PCA and gaussian regression

	ANN	**PCA+ANN**	**GPR**
Hidden Layer 1	11	2	-
Hidden Layer 2	4	0	-
Performance	96.17	95.56	-
MAPE	6.06	18.17	12.98776612
MSE	4.6	14.02 8	7.302842802
RMSE	2.16	3.70	2.70237725
MAE	1.50	3.07	2.249703887

Result indicates that Neural network with two hidden layers with eleven and four neurons respectively yields better result than Neural network with PCA and Gaussian progression regression (GPR) in terms of Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean

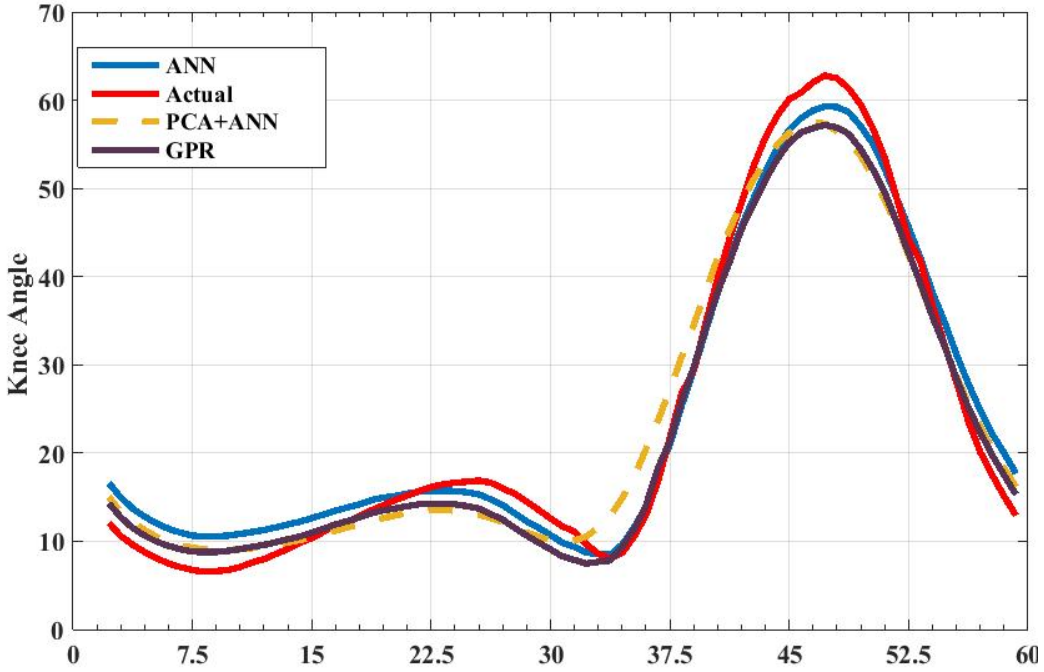


Figure 4.7: Result for kinematics parameters prediction model from anthropometric parameters

Absolute Error (MAE). as indicated in Table 4.5. In the study eight components has been considered in PCA.

The accuracy of the ANN model with and without PCA is 95.56 and 96.17 respectively when compared with the true angle. The MAPE value for ANN is best with 6.06 while PCA with ANN is 18.17 and for GPR it is 12.98. Similarly, MSE is best with ANN model. There is approx. The difference of 10 and 3 when compared with PCA with ANN and GPR. Similar trends are also found in when compared using RMSE and MAE. Figure 4.7 shows the plot of ANN, PCA with ANN and GPR when compared with the original value of knee angle.

The prediction of the gait pattern based on the body parameters plays a vital role in case of rehabilitation systems.

4.3 Chapter Summary

This chapter presented the prediction model for the gait kinematics parameter estimation model. Based on the previous gait cycle, using the univariate and multivariate time series model, new gait cycle has been proposed. This play a vital

role in the robotic rehabilitation system. Then a prediction model for the gait kinematics based on the body parameters has been presented and discussed. Each has a specific gait profile as per their body built. Based on these gait profile, health care professionals decide gait-related rehabilitation. In this study, a framework has been presented for the prediction of the gait angle profile from 14 body parameters. To validate the model, BP-ANN with and without PCA, Gaussian process regression model has been explored over the MNIT Gait dataset. It has been observed that the proposed model can perform better in term of Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Standard deviation of errors (SDE).

As it is needed to identify the gait abnormality from the gait pattern. So in Chapter 5, different machine learning techniques have been proposed for the gait abnormality detection.

Chapter 5

Machine Learning Techniques for Gait Abnormality Detection

Although walking is natural phenomena, but its importance is felt, when it is distorted. Availability of quantitative gait parameters is essential for the detection of gait disorders, identification of balance features, and assessment of medical gait interventions and rehabilitation developments. It can be used to analyze healthy subject's gait pattern and can be used as a preventative health screening context, for example, it can be used to imply potential injury or detect the risk of fall in the elder people [29, 30]. In case of injury, gait analysis may assist the doctors to select the correct treatment options, preventing unnecessary surgical operations.

For the clinical gait analysis, identification of accurate gait phase is of paramount importance. Section 5.1 puts forward the proposed approach using the passive marker that automatically recognizes gait sub-phases using fuzzy logic approach. In addition to stance phase and swing phase, the approach is capable of detecting all the sub-phases such as initial swing, mid-swing and terminal swing, loading response, mid stance, terminal stance and pre-swing. The prototype of the system provides effective and accurate gait phase that could be used for understanding patients' gait pathology (abnormal gait pattern) and in control strategies for active lower extremity prosthetics and orthotics.

A vision based gait data of children with neuro-development disease (Cerebral Palsy (CP)) has been considered and analyzed in section 5.2. It is quite an intricate task to categorize gait pattern into normal and CP based pathology. Five cases are considered to explore the feature selection criteria before applying clustering technique. Finding the optimal number of clusters is a challenging task

in the unsupervised learning area. Thus, an optimal number of gait profiles in the datasets are identified based on voting from mean square error, silhouette coefficient, and Dunn index. The result demonstrate that optimized based gait profile clusters could assist quantitatively in clinical rehabilitation evaluation for the children affected by CP.

5.1 Fuzzy Logic based Gait Phase Detection From Clinical Perspective (FL-GPD)

Walking is considered as a series of cyclic events known as Gait Cycle (GC). It starts with the heel contact and ends with the heel contact of the same foot. It consists of two phases, swing and stance [18]. Stance phase begins with the heel contact and ends with the toe off of the same foot. The share of stance phase is 60% in a normal gait cycle. Swing phase begins with the toe off of the delete foot and ends with the heel contact of that same foot. It is 40% of the complete gait cycle.

Classical gait model by Perry [18], divides gait cycle into eight sub-phases (5 stance and 3 swing) as shown in figure 5.1 .

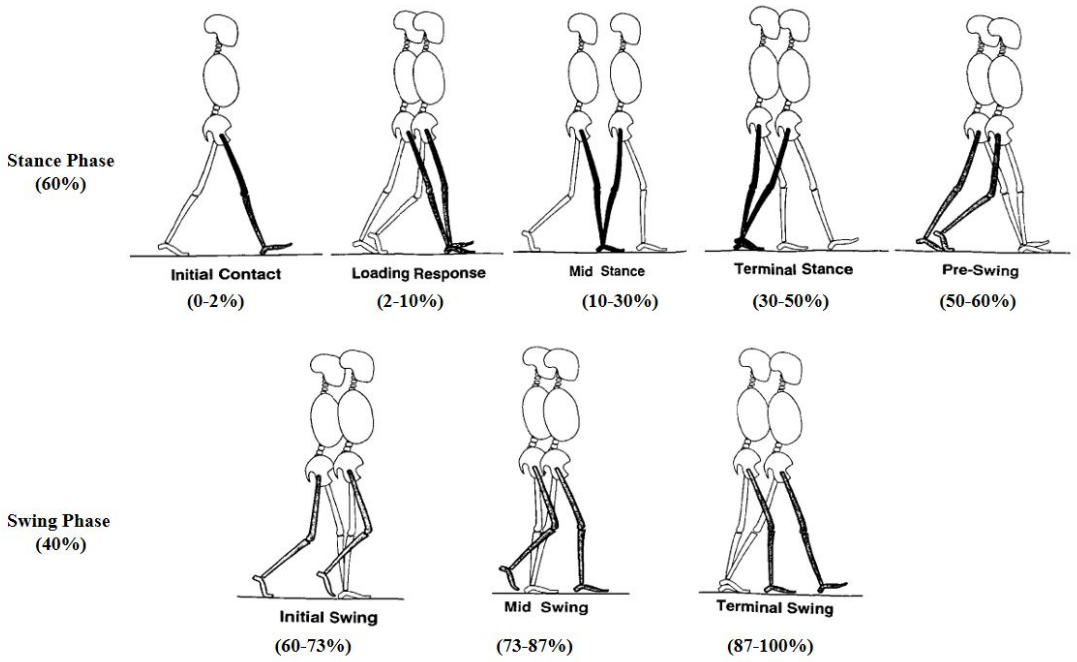


Figure 5.1: Fundamental gait phases and expected interval phases and sub-phases in complete gait cycle

The stance phase is divided among five sub-phases: Initial Contact, Loading Response, Mid Stance, Terminal Stance and Pre-Swing whereas swing phase has three sub-phases: Initial Swing, Mid Swing, and Terminal Swing. In this figure, the right foot is considered as the reference foot for gait phases cycle and is shown in the black shade. For the analysis, the initial contact and loading response is considered as the same phase as former is an instance of loading response only [1].

Each gait phase has a functional objective and a critical pattern of selective synergistic motion to accomplish its goal. It is seen that different pathologies affect different segments of either swing or stance phase [18, 186, 187]. Any abnormality suggests that there is a pathology, which should be identified by the examiner. Recognition of gait cycle phases is extensively useful to spot the time instance at which feedback should be applied for safety and effective response by the patient undergoing rehabilitation or physical therapy.

For gait phases detection, wearable sensors like gyroscopes, accelerometers, EMG sensors, Force Sensitive Resistors (FSRs), inertial sensors, force contact sensors, foot switches, load cells, etc. are used to collect the gait parameters [188, 189]. Alternative approaches such as force plate and vision-based methods are used to compute quantitative parameters of interest [190, 4]. Wang et al. identified only initial contact, stance phase, and swing phase by using 3-axis accelerometers fixed on the ankle [19]. Pappas et al. used gyroscope attached to rear end of the shoe along with force sensitive resistors to detect heel Strike (Initial Contact), Stance, Heel-off and Swing phase. Only 4 phases are identified by this approach [188].

Computational-based techniques have also been proposed for real and precise gait phase recognition. Researchers have explored Fuzzy Inference System (FIS) to segment gait phase. Liu et al. used gyroscope and accelerometers to detect four gait phases using fuzzy logic due to its robustness to noise [15]. Kyoungchul et al. implemented a fuzzy based approach to detect gait phases from foot pressure patterns [16]. DeRossi et al. have used Hidden Markov Model for identification of six gait phases [17].

But in all these techniques, one or more sensors need to be attached to one, or both legs, which is not an appropriate approach as the presence of sensors, cables or other components hinders subject's natural motion. In contrary, vision-based analysis systems can be used to obtain gait kinematics smoothly and continuously without affecting the natural motion. The technology associated with this measurement approach has continued to change over the past decade.

The proposed passive marker based gait kinematic parameters extraction technique from chapter 3 is used to obtain the gait phase identification using fuzzy logic. The automated phase segmentation algorithm employed can identify the deviation or missing gait phases and thus it could be used as a preventative health screening context. It can also be incorporated to obtain the timing of feedback in control strategies for active prosthetics.

5.1.1 Proposed Methodology for Gait Phase Detection

The proposed methodology consists of a digital video camera for recording and a computer for data acquisition and processing. Figure 5.2(a) illustrates the marker position in red color. Figure 5.2(b) shows an optical motion-capture system developed to detect and track the markers, fastened to the subject's body at anatomical points of concern.

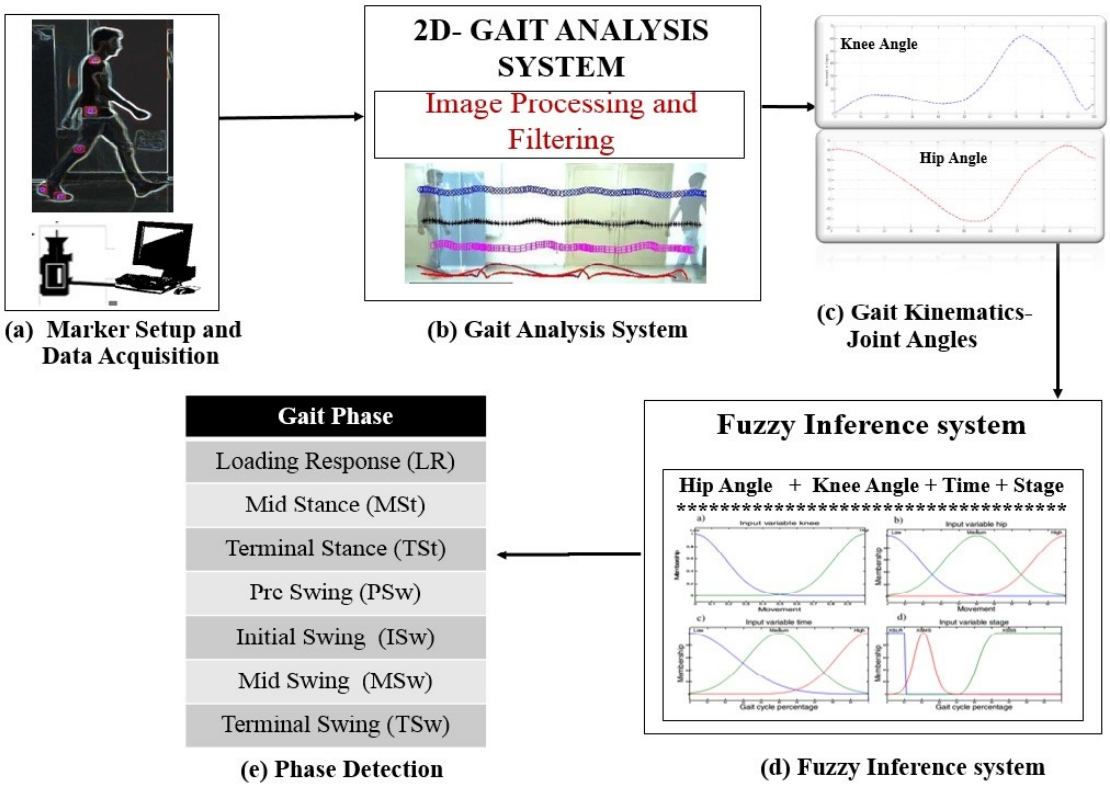


Figure 5.2: Gait phase detection methodology; (a) Marker setup and data acquisition; (b) Gait analysis system; (c) Gait kinematics extraction; (d) Fuzzy inference system for gait phase identification; (e) Phase detection

The proposed passive marker-based approach as discussed in chapter 3, is used to extract knee and hip angles as shown in figure 5.2(c). The Fuzzy Inference System

(FIS) shown in figure 5.2(d) maps inputs (hip, knee angles, time and stage) to outputs (gait phases) using a predefined set of fuzzy rules. Finally, the output of this system is a segmented (with sub-phases) gait cycle.

At a given stance of time, kinematic parameters such as hip, ankle and knee angle can be put to use to detect the gait phase [191]. One possible approach for gait phase identification is by setting the threshold for discrete event analysis. This is only expedient when the change is visible. Change in phases is not visible in the knee, and hip angle parameter as gait is not a set of isolated actions, although phases keep changing continuously and smoothly. Thus, there is need of a new approach that can work with imprecise data to efficiently detect the gait phase.

Lotfi A. Zadeh introduced the concept of fuzzy logic in 1965 [192]. Fuzzy logic is best suited for representation of information extracted from inherently imprecise data. Fuzzy logic handles imprecision, vagueness and insufficient knowledge. Conventional boolean logic is for the crisp sets that have either ON (1) or OFF (0) value as represented mathematically with the indicator function.

$$\chi_A(x) = \begin{cases} 1, x \in A \\ 0, x \notin A \end{cases} \tag{5.1}$$

where the symbol $\chi_A(x)$ indicates an unambiguous membership of element x in a set A.

The fuzzy system can have any real number as membership value between zero and one as shown in the equation.

$$\mu_A(x) \in [0, 1] \tag{5.2}$$

where the symbol $\mu_A(x)$ is the degree of the membership of element x in fuzzy set A. Gait phase identification activities are often vague or based on intuition, as one can not differentiate between all phases. Fuzzy logic can work in this scenario with reasoning.

Fuzzification, inference, knowledge base and defuzzification are four modules of the fuzzy expert system as shown in figure 5.3. Fuzzification is the process of crisp input to fuzzy set by using Membership Function (MF). To draw the inference, fuzzy logic necessitates knowledge which is stored in the fuzzy system and provided by either an expert (have experience and knows the process of that specific domain) or extracted from numerical data. IF-THEN rules inferred from the ex-

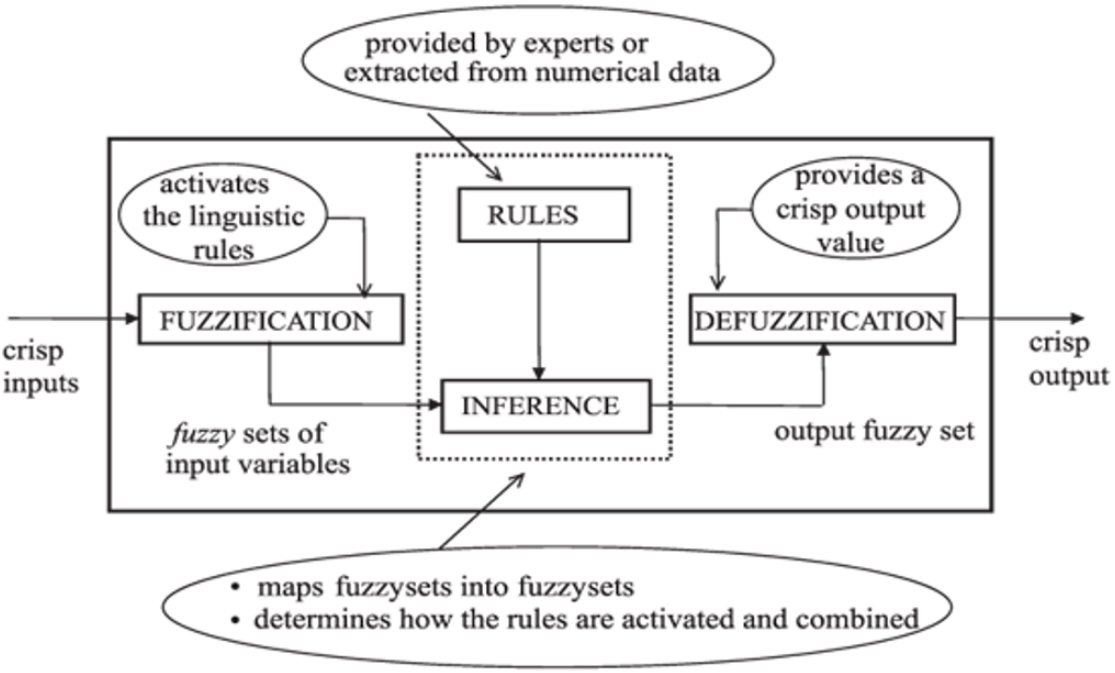


Figure 5.3: Modules of fuzzy logic expert system

pert are stored in the knowledge base. Using these rules inference engine simulates reasoning process similar to human. The output is in fuzzy form. Thus there is need of conversion of the fuzzy set to the crisp value, this process is known as defuzzification.

5.1.1.1 Input parameters

The inputs to this proposed fuzzy system are the hip angle, knee angle, Time (gait cycle percentage) and Stage (a function of time used to distinguish the phases loading response and initial swing and between mid-stance and terminal swing). The information of kinematic parameters averages hip and knee used in this system are acquired by the proposed passive marker based gait technique described in chapter 3. The fuzzy-logic based gait phase segmentation system can be formalized from given kinematic parameters of the hip and knee, $x_{hip}(n)$, $x_{knee}(n)$, and two time variables $x_{Time}(n)$ and $x_{Stage}(n)$ to design the mapping as shown in equation 5.3.

$$F : \{x_{hip}(n), x_{knee}(n), x_{Time}(n), x_{Stage}(n)\} \Rightarrow Phase \tag{5.3}$$

where $x_{hip}(n)$, $x_{knee}(n)$, $x_{Time}(n)$ and $x_{Stage}(n)$ and $Phase$ will be fuzzified.

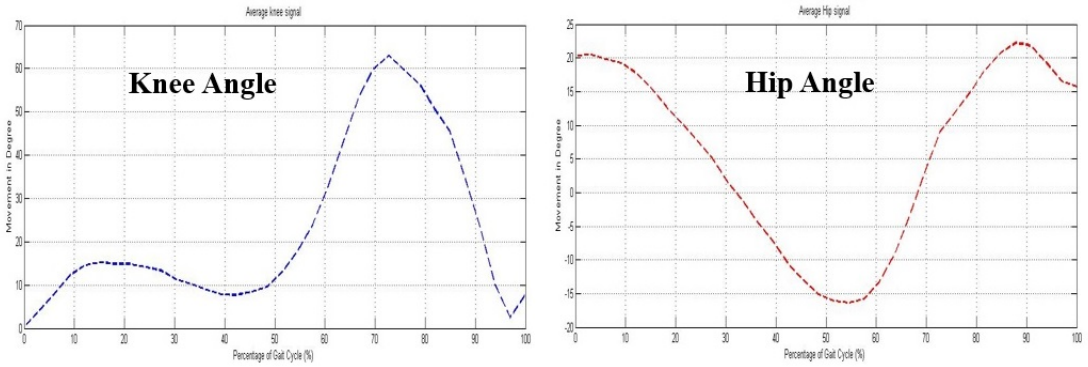


Figure 5.4: Average knee and hip joint angle for one gait cycle of subject S22

Kinematic parameter considered for the study i.e., average hip and knee angle is shown in figure 5.4. Their membership functions depict the different intervals in the y axis. Hip angle membership can be divided into three intervals Low, Medium and High represented as x_{hLow}, $x_{hMedium}$ and x_{hHigh} respectively. Knee angle movement is also divided into three intervals Extension $x_{kExtension}$, Low Flexion x_{kLowF} and High Flexion x_{kHighF}).

Similarly input variable Time is divided into Low x_{TLow}, Medium $x_{TMedium}$ and High x_{THigh} intervals. Stage input variable is employed to differentiate between the phases mid-stance and terminal swing and between loading response and initial swing. Stage variable membership functions are defined by phase occurrences such as Loading response (SLR), Mid-stance (SMS) and Swing stance (SSS). Its Membership functions are x_{SLR}, x_{SMS4} and x_{SSS}.

The membership value of input (hip,knee, Time and Stage) and output(Phase) is represented as following by equation 5.4

$$\begin{aligned} x_{hip}(n) &= \{x_{hLow}, x_{hMedium}, x_{hHigh}\} \\ x_{knee}(n) &= \{x_{kExtension}, x_{kLowF}, x_{kHighF}\} \\ x_{Time}(n) &= \{x_{TLow}, x_{TMedium}, x_{THigh}\} \\ x_{Stage}(n) &= \{x_{SLR}, x_{SMS}, x_{SSS}\} \\ x_{Phase}(n) &= \{P_{LR}, P_{MSt}, P_{TSt}, P_{PSt}, P_{ISw}, P_{MSw}, P_{TSw}\} \end{aligned} \tag{5.4}$$

The membership function values for hip, knee, Time and Stage input are defined based on normative data presented by J.Perry [18, 193] and is shown in table 5.1.

A fuzzy set is completely characterized by its membership function. The membership functions are chosen by the data distribution. The membership function

Table 5.1: Fuzzy membership function for input parameters

Input	Quantity
Hip	Extension [0 to -30], Low Flexion [0 to 15], High Flexion [15 to 30]
Knee	Low [0-20], Medium [20-40], High [40-70]
Time interval (% of Gait)	Low [0-25], Medium [25-75], High [75-100]
Stage	Loading response (SLR) [0-10%], Mid-Stance (SMS) [10-40%], and Swing Stance (SSS)[40-100%]

considered for the three input variables (hip, knee and Time) is a Gaussian function. For Stage variable, SLR uses Polynomial based Z-function, the Gaussian function is used for SMS membership, and S-function is used for SSS. The MFs of the input variable is shown in figure 5.5

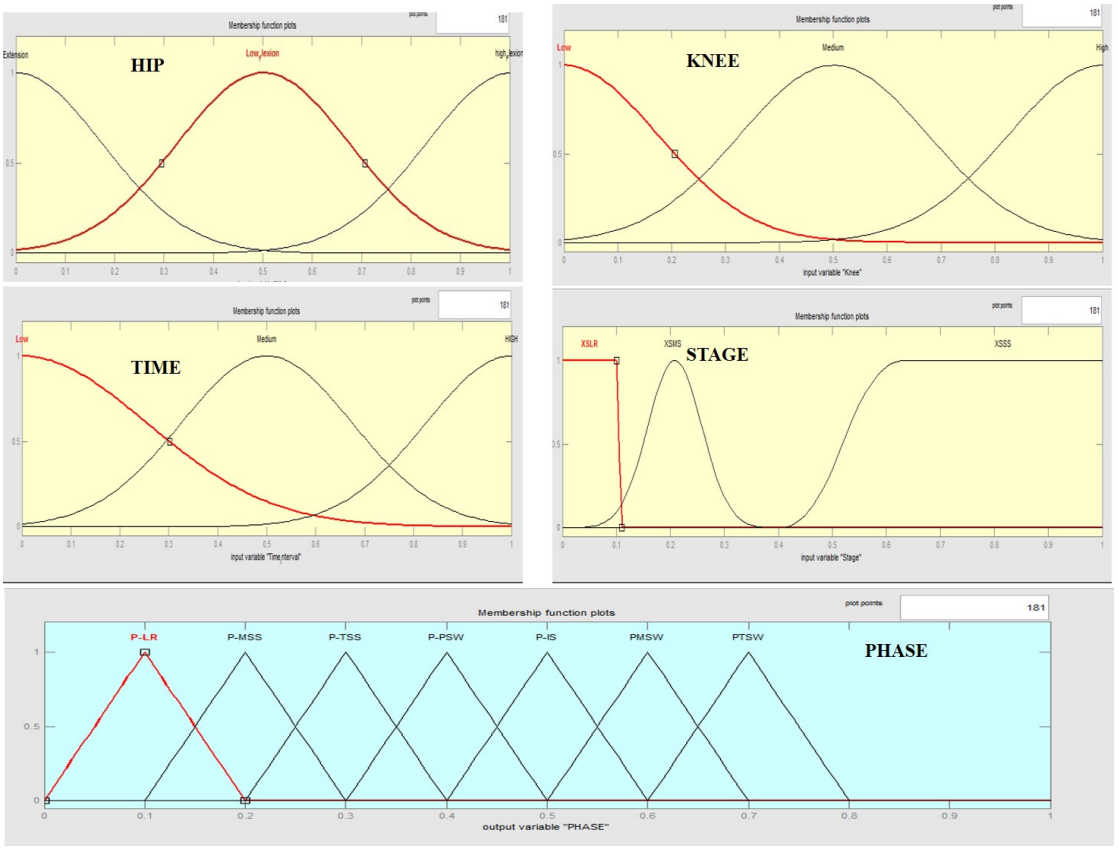


Figure 5.5: Membership function for input variable: hip, knee, time and stage and output variable gait phase

A Gaussian MF is determined complete by mean μ, (the MFs centre) and variance σ, (the MFs width). Mathematically it can be represented as :

$$Gaussian(x, \mu, \sigma) = e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{5.5}$$

Hip, knee and Time variables can be represented using the possibility distributions; thus Gaussian function is used as the membership function. For the hip and knee

input the Gaussian function parameters are as following: (0,0.175), (0.5, 0.175) and (1, 0.175) for each member respectively. For the Time input it is (0, 0.256), (0.5, 0.175) and (1, 0.175) for Low, Medium and High fuzzy membership respectively.

Spline-based function is also known as Z shaped function and it can be represented as

$$Z - shaped \quad (x, a, b) = \begin{cases} 1, x \leq a \\ 1 - 2(\frac{x-a}{b-a})^2, a \leq x \leq \frac{(a+b)}{2} \\ 2(\frac{x-b}{b-a})^2, \frac{(a+b)}{2} \leq x \leq b \\ 0, x \geq b \end{cases} \tag{5.6}$$

Where a and b parameters, locate the extremes of the sloped portion of the curve. In case of Stage input, x_{SLR} Z function is used as membership value with (0.1, 0.11) parameter. This Z shaped function is used to segregate loading response and mid-swing. Stage variable x_{SMS} (used to segregate among mid-stance and terminal stance) is represented by the Gaussian function with (0.208, 0.05) as parameter.

Similar to x_{SLR}, function x_{SSS} (used to differentiate between swing and stance phases) can be represented with another Spline based function (S function) with paramter a and b as 0.405, 0.632 respectively as illustrated in figure 5.5(d). S function can be represented similar to Z function as following :

$$S - shaped \quad (x, a, b) = \begin{cases} 0, x \leq a \\ 2(\frac{x-a}{b-a})^2, a \leq x \leq \frac{(a+b)}{2} \\ 1 - 2(\frac{x-b}{b-a})^2, \frac{(a+b)}{2} \leq x \leq b \\ 1, x \geq b \end{cases} \tag{5.7}$$

If the data pattern increased to a peak and starts to fall immediately then triangular Membership Function may be a choice. A triangular MF is specified using three parameters a, b, c such that $(a < b < c)$ and mathematically represented as follows:

$$\Delta(x; a, b, c) = \begin{cases} 0, x \leq a \\ \frac{x-a}{b-a}, a \leq x \leq b \\ \frac{c-x}{c-b}, b \leq x \leq c \\ 0, c \leq x \end{cases} \tag{5.8}$$

Thus, the membership function for output variable (Phase) is considered as triangular as shown in figure 5.6. The function parameters, a, b and c for Phase

vairable are as following:

$$P_{LR}(0.0, 0.1, 0.2), P_{MSt}(0.1, 0.2, 0.3), P_{TSt}(0.2, 0.3, 0.4),$$
$$P_{PSt}(0.3, 0.4, 0.5), P_{ISw}(0.4, 0.5, 0.6), P_{MSw}(0.5, 0.6, 0.7), P_{TSw}(0.6, 0.7, 0.8)$$

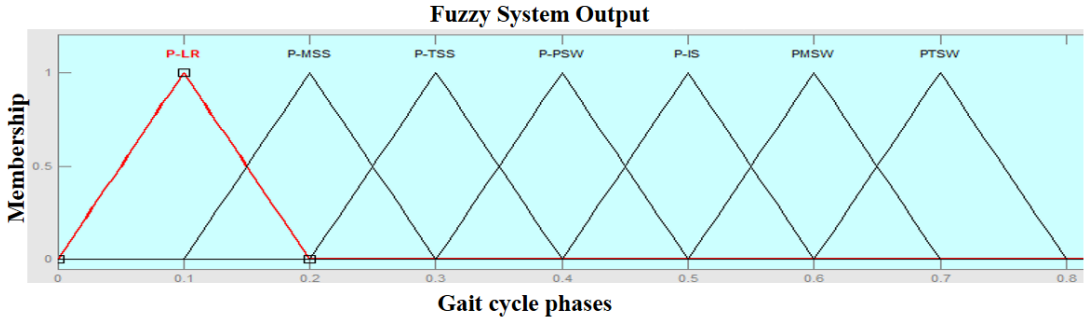


Figure 5.6: Membership function for output variable (gait phase)

5.1.1.2 Fuzzy Inference System (FIS)

Fuzzy inference is the method of mapping of input to the output using fuzzy logic. It is analogous to the reasoning.

Mamdani and Sugeno are the two type of fuzzy inference system. In 1975 Ebrahim Mamdani proposed Mamdani-Type Fuzzy Inference as an attempt for the control system [194]. The output of Sugeno type inference system is either linear or constant mathematical expression, while Mamdani type fuzzy inference gives a fuzzy set as output.

Mamdani-type Fuzzy Inference System (FIS) uses the technique of defuzzification of a fuzzy output. Max-Min composition and Max-Product composition are two types of Fuzzy relation composition for Mamdani-type FIS. In the proposed system for phase identification Max-Min composition is used. Let R be a fuzzy relation in $X \times Y$, and S be a fuzzy relation in $Y \times Z$. The Max-Min composition of R and S, $R \circ S$, is a fuzzy relation in X × Z and mathematically it is represented as following.

$$R \circ S \leftrightarrow \mu_{R \circ S}(x, y) = \vee \{\mu_R(x, y) \Lambda \mu_S(y, z)\} = Max\{Min\{\mu_R(x, y), \mu_S(y, z)\}\} \tag{5.9}$$

Due to the interpretable and intuitive nature of the rule base, Mamdani-type FIS is widely used in particular for decision support application. In the proposed system

Table 5.2: Set of fuzzy rules used in the study

	Hip	Knee	Time	Stage	Gait Phase
1.	High Flexion	Low	Low	SLR	Loading Response (LR)
2.	Not High Flexion	Low	Low	SMS	Mid Stance (MSt)
3.	Low Flexion	Low	Low	SMS	Mid Stance (MSt)
4.	Extension	Medium	Medium	—	Terminal Stance (TSt)
5.	Extension	Not Low	Medium	SSS	Pre-Swing (PSw)
6.	Low Flexion	Not Low	Medium	SSS	Initial Swing (ISw)
7.	Flexion	High	High	SSS	Mid-Swing (MSw)
8.	High	Flexion Low	High	SSS	Terminal Swing (TSw)

the output is fuzzy output, thus Mamdani type FIS is considered. Mamdani type fuzzy system is used as the mapping F as

$$R(\{x_{knee}(n), x_{hip}(n), x_{Time}(n), x_{Stage}(n)\}, Phase) \tag{5.10}$$

The rules for this system are determined by an intuitive reasoning by considering the literature [1, 18, 47] and shown in table 5.2. For example, a condition for the LR phase in table 5.2 suggests that as "*If hip angle is of 20 degrees while Knee is 18 and Time period is 18% and Stage is also 18% of total time, then the motion is in the LR phase*".

Fuzzy logic expresses the statement as "*Hip angle is of High Flexion nature AND Knee angle has Low MF AND Time interval is also Low AND Stage is SLR then PHASE is Loading Response*". Based on these rules, gait phase can be determined at any given instance of time.

The defuzzification scheme used in the proposed system is the Centroid of Area (COA), also known as the center of gravity. It is widely adopted in the calculation of expected value in probability distributions. The total area of the membership function distribution using Max-Min composition is divided into many sub-areas. In COA, the summation of the centroid of each sub-area is calculated, and then the defuzzified value is calculated. For discrete membership function, the defuzzified value for COA is defined as:

$$COA = \frac{\sum_{i=1}^{n} x_i \cdot \mu(x_i)}{\sum_{i=1}^{n} \mu(x_i)} \tag{5.11}$$

Here x_i indicates the sample element, $\mu(x_i)$ is the membership function, and n represents the number of elements in the sample.

When the membership function is contionous, COA is represented as following :

$$COA = \frac{\int x\mu_A(x)dx}{\int \mu_A(x)dx} \tag{5.12}$$

5.1.2 Experimental Result :Fuzzy-based Approach for Phase identification

Experimental conditions are similar as discussed earlier in chapter 3. The Kinematics parameters of the subjects are visualized using MATLAB. For the analysis of the result, two different cases have been considered.
CASE 1) : Healthy subject gait phase analysis, where all the phases are present.

CASE 2): Unhealthy subject gait phase analysis, where the phases are either misplaced or are not present.

5.1.2.1 CASE 1: Healthy Gait Pattern

In the case 1, 120 healthy subjects are analyzed. The result for fuzzy based gait identification approach of one of the healthy subject S34 (age 20, weight 54) for two strides is shown in figure 5.7. It can be seen that the sequence of gait phases obtained is in the natural order. Detected phases in each gait cycle had a maximum membership value of one, stipulating that within a gait cycle, all phases are fully identified. From the figure 5.8 it is reflected that some portion of the gait cycle is spent in transition from one phase to another phase in one gait cycle. It is found out that a maximum mean deviation of 1.04 % is there about normative data.

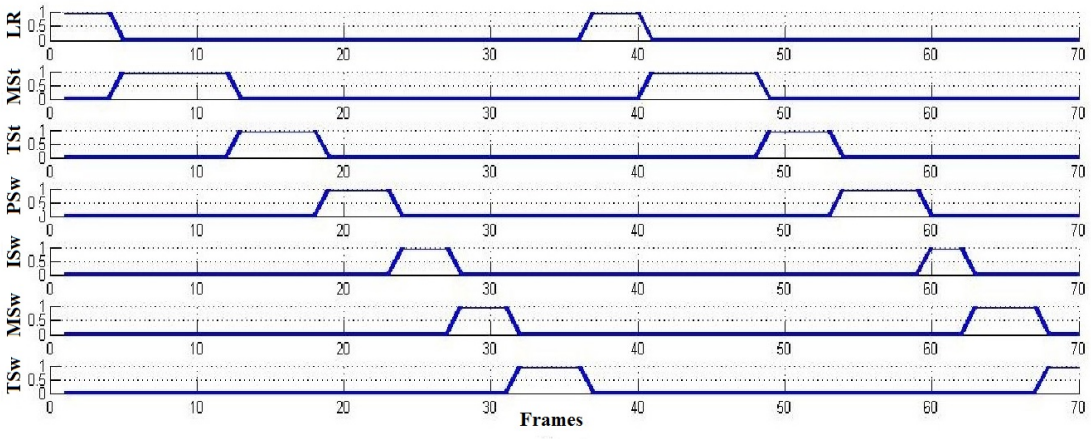


Figure 5.7: Result after fuzzy-inference for normal gait phase plots for two strides

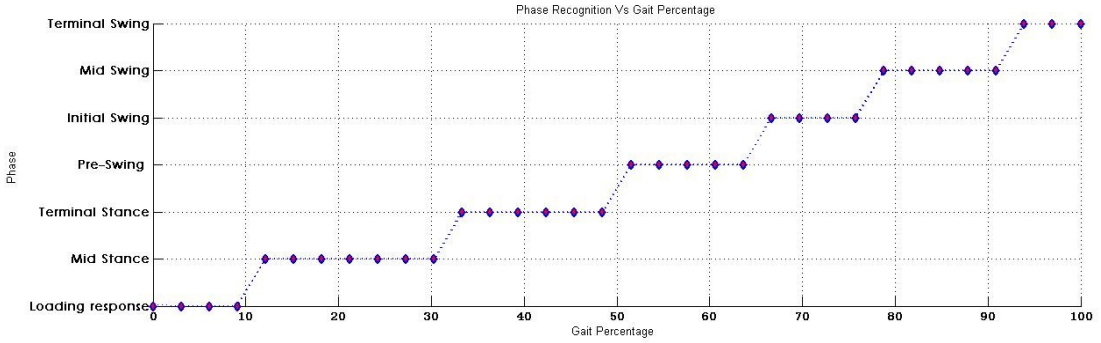


Figure 5.8: Fuzzy-logic based phase detection in one gait cycle

The result of fuzzy logic based phase detection is illustrated in figure 5.9. Based on fuzzy rules, the result follows the natural order of phase sequence, and at a time only one gait phase has been detected. The first phase, loading response is found in interval 0% to 10% of the complete gait cycle, which indicates a normal behavior occurrence of this phase. Mid-stance is detected in 10% to 32%; terminal stance phase is detected in the percentage 33% to 50%, and pre-swing is found in the range of 50% to 64% of a complete gait cycle. The initial swing is detected in the percentage of 64 to 75, mid-swing is located in 75% to 87%. The terminal swing phase is correctly detected in the percentage 87% to 100% of the gait cycle.

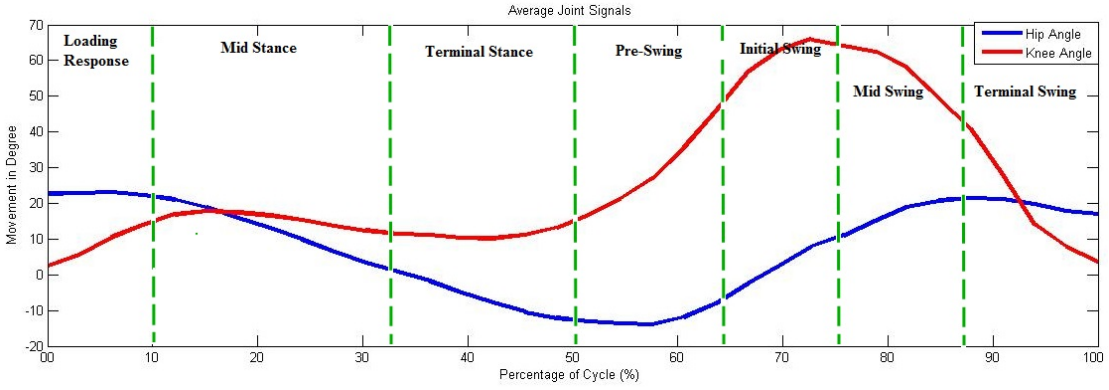


Figure 5.9: Kinematics signals and phase detection

Cross-validation of these results with figure 5.1 is shown in table 5.3 indicates that the proposed approach reports very similar quantitative outcome. Detection of gait phase is in natural order, i.e., LR, MSt, TSt, PSw, ISw, MSw, and TSw, thus it can be concluded that subject S34's gait pattern is free from gait abnormalities and prove the hypothesis that the subject is healthy.

Table 5.3: Extracted gait phases using proposed system for subject S34

Gait Phase	Proposed Fuzzy based Approach	Standard Values [1]
Loading Response (LR)	0-10 %	0-10
Mid Stance (MSt)	10-32 %	10-30%
Terminal Stance (TSt)	33-50%	30-50%
Pre-Swing (PSw)	50-64%	50-60%
Initial Swing (ISw)	64-75%	60-73%
Mid-Swing (MSw)	75-87%	73-87%
Terminal Swing (TSw)	87-100%	87-100%

5.1.2.2 CASE 2: Unhealthy Gait Pattern

80 unhealthy subjects as discussed in 3.1.2 are considered for the cross-validation of the proposed fuzzy-based gait system.

Figure 5.10 illustrates the result of fuzzy logic approach for the unhealthy subject (US22) considered for the analysis. The result disproves the hypothesis that the subject is having normal gait pattern as the gait phases identified are not in the normal order and also not in the range of the standard values suggested by Whittle [1].

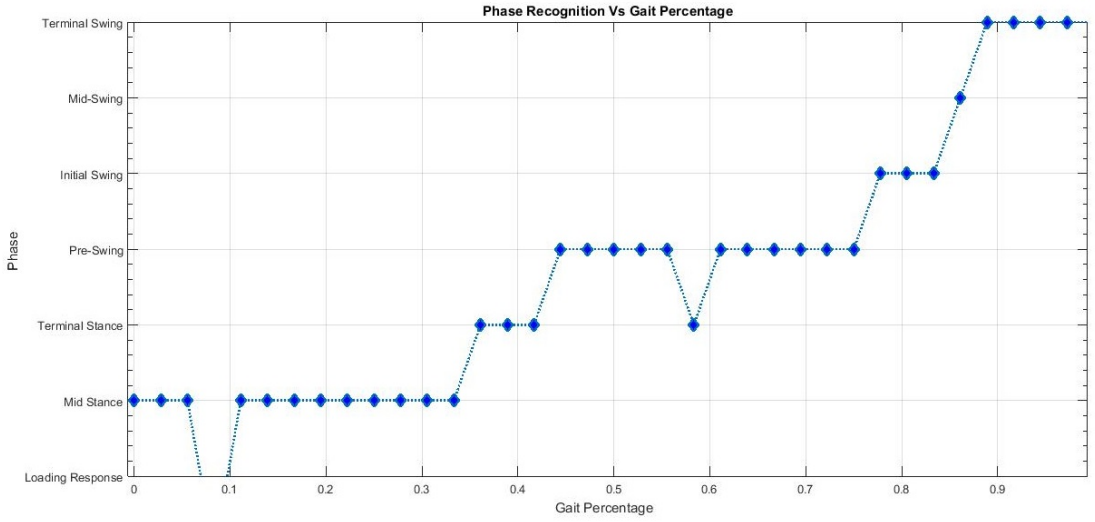


Figure 5.10: Unhealthy subject (US22) fuzzy-logic based phase detection for one gait cycle

5.1.2.3 Discussion

Findings of the experiments indicate that this technique can correctly segment the gait phases at very low cost using fuzzy-based approach because of the use of pas-

sive based parameter extraction approach. For all the healthy subjects considered in this research, identification rate in the natural sequence is 100%. The proposed algorithm was able to detect the abnormal gait patterns. This approach of gait phase detection has the potential to be used in rehabilitation activities and gait analysis.

If we have large labeled training data, i.e., conditions and classes, the fuzzy logic approach is a dominant technique. The limitations of these rule-based and supervised approaches are that their accuracy depends on the training sample, larger the training sample, higher is the accuracy of the system. Another limitation of fuzzy logic is that if there is a change in the rules, then we have to change the MF and vice-versa. The limitation of the rule-based system is that it gives the same importance to all factors.

Supervised learning is a popular choice for the research community to classify the gait pattern and CP pathology for automated analysis. Neural Network is most widely used for normal gait analysis, robotic rehabilitation, sports monitoring, and tactics, geriatric care surveillance, activity recognition [4, 60, 108, 195]. A learning model is developed using the existing data through training and then to validate the same, testing is carried out with either the unknown or known data.

The primary limitation of the rule-based and supervised techniques is that they cannot be applied to the data where no prior information of the data is given. In diagnosis related cases, there is no clear boundary between healthy and unhealthy subject cases. Each cluster represents a gait profile in the chosen dataset. Let us say, if we have only two clusters then the former represents the healthy one and latter represents unhealthy. But in real time the subject can lie in different groups at a time. So there is a need for the optimal number of gait profiles in treatment intervention for individuals with a disability. In the next section unsupervised learning has been explored for finding the gait abnormality. A vision based gait data of children with neuro-development disease (Cerable Palsy (CP)) has been considered and analyzed in the next section.

5.2 Optimized Clustering Techniques for Identification of Abnormal Gait Profile

All pathological pattern cannot be apprehended through just visual observation. There is a need to take the aid of statistical (frequency domain characteristics,

spectral components, harmonic content, the coefficient of variation, etc.) and machine learning techniques [196]. In this study, nature-inspired meta-heuristic algorithms are explored on a publicly available gait dataset of 156 subjects for automatic gait profiling of children with cerebral palsy.

Cerebral Palsy (CP) is a neurodevelopment disease, common in children. It is associated with floppy or rigid limbs, exaggerated reflexes and involuntary motion, poor speech and learning ability which, is considered as a non-progressive disease [197, 198]. The major reason for this motor disability is abnormal brain development, which often occurs before the birth of a child. On behalf of motor impairment of the limb, CP is classified into three major categories: Spastic, Ataxic, and Athetoid [199]. Different studies, around the globe, highlight the severity of this disease and projects that worldwide CP cases vary from 1.5 to 4 per 1,000 children. The condition is more severe in the developing countries [197, 200]. Thu, considerable amount of money is involved in the rehabilitation and intervention policies related to the diagnosis of gait pathology related to cerebral palsy [201, 202].

The diversity in the gait pattern from children suffering from CP makes it a challenging task for researchers in the physical therapy and surgical community. There are several attempts made by researchers to characterize the gait pattern and categorize them into healthy and gait pathology using computational techniques [203, 204, 108, 4, 205].

Supervised learning is a popular choice for the research community to classify the gait pattern of CP pathology for automated analysis. The Neural Network is most widely used for normal gait analysis, robotic rehabilitation, sports monitoring, and tactics, geriatric care, surveillance, activity recognition [7, 206, 60, 207, 195]. Supervised learning technique is a dominant methodology when labeled training data points are available. Zheng et al. [96] implemented the trained neural network classifier to identify the abnormality in older person gait. Gait Pathology can be identified using ground reaction force with Learning Vector Quantization (LVQ). Multilayer perceptron (MLP), Linear Discriminant Analysis (LDA) classifiers, Kernel Fisher Discriminant (KFD) and Bayesian classifier have been used to classify gait pattern in [112, 208]. Support vector machine (SVM) is used for the study of normal and age-related differences, normal and abnormal gait pathology. The authors in [94, 89] showed that it exhibits good results in CP pathology. SVM is a powerful classifier suitable for small to the medium dataset.

The limitation of these classification techniques is that they are not applicable

to the unlabeled data. To overcome these limitations, researchers have explored unsupervised, or clustering approaches [209, 99, 100]. Cluster models are not only popular in signal and image processing, but they also have wide applications in web mining and pattern recognition, sensor network, robotics, seismology, medical science in disease clustering, etc. [210, 211].

In gait analysis, clustering techniques are most widely used for categorizing gait data into groups of disorders based on common hidden features in the subject dataset. K-means, fuzzy c-means, hierarchical clustering, Self-Organizing Map (SOM) are some of the examples of clustering techniques that have been explored in diagnosing CP related gait abnormalities [7].

In 1997, O'Malley et al. examined fuzzy clustering approach on cerebral palsy children based on temporal distance parameters (stride length and cadence) [99]. They explore this clustering technique on 156 subjects (88 children with spastic diplegia CP disease and 68 normal). Xu et al. explore k-means, self-organizing feature map (SOM) and hierarchical clustering to differentiate normal and pathological gait pattern based on stride length and cadence [100]. Angkoon et al. use hierarchical clustering for clustering healthy group from pathological patients [101]. Carriero et al. demonstrates the possibility of Principle Component Analysis (PCA) in quantitative classification of CP gait Pattern [212]. Fuzzy C-means analysis is implemented to cluster the data of 40 subjects (20 healthy and 20 soastic diplegic CP patients) with 27 parameters each. In 2007 Taro et al. apply hierarchical clustering on the sagittal kinematic gait data of 67 subjects (56 CP, 11 normal). Thirteen gait cluster are formulated using sagittal plane hip, knee, and ankle kinematics [213].

In K-means clustering approach, the result depends on the choice of the randomly selected cluster centroids. If they are close, it takes time to converge, while if the distance is on two extreme sides, then better cluster formation takes place. Thus the selection of the first cluster head points can be optimally selected using meta-heuristic approaches [210].

Selection of the number of clusters is considered as obscure as the interpretations not only depends on the shape and distribution of the data-points but also on the desired clustering resolution of the user. In this thesis, it is proposed that nature-inspired clustering approaches can help in finding the optimal number of gait profiles in the datasets. A novel nature-inspired meta-heuristic algorithm, Grey Wolf Optimized Clustering (GWOC) approach is proposed to cluster CP gait patterns using a publicly available CP gait dataset of 68 normal healthy and

88 with spastic diplegic, a form of CP. The performance of the proposed clustering methods is evaluated using internal and external cluster performance index. The proposed clustering technique outperforms the traditional clustering approaches.

5.2.1 Experimental Methodology

In this section, the methodology used and nature-inspired optimized based clustering approaches are discussed followed by the five different cases, clustering evaluation indices and parameter settings. Clustering techniques are employed on the CP dataset and based on the best possible number of cluster identified in this study, the condition after the post-surgery is examined to validate the clustering result. Figure 5.11, illustrates the work flow of the proposed methodology.

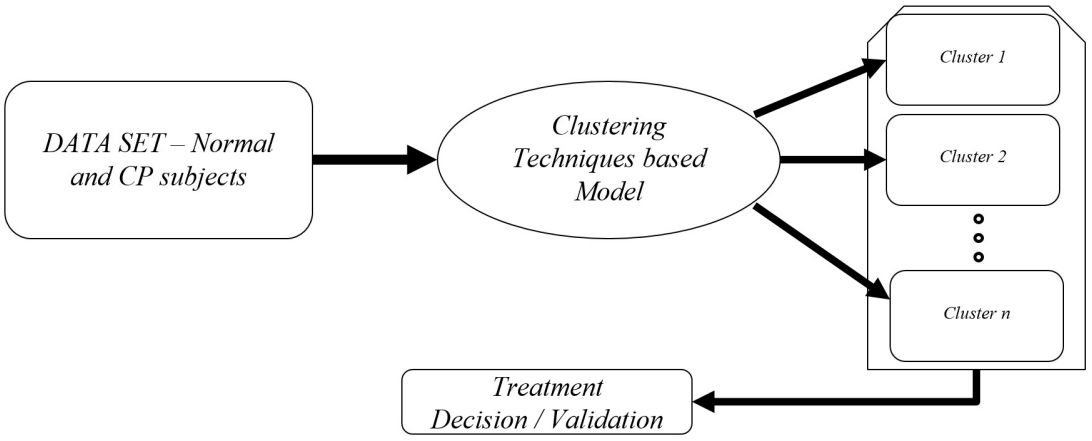


Figure 5.11: Clustering based methodology used for gait profiling of CP Children

5.2.1.1 Gait Data-set Description

The gait dataset considered, is taken from the publicly accessible data from O'Malley et al. [99]. This gait data consists of two groups, first consisting of 88 children with the spastic diplegia, CP and another is the collection of 68 neurologically intact children and have no history of any motor disease. Data is collected using six cameras of Vicon System and processed with Vicon Clinical Manager at Motion Analysis Laboratory, University of Virginia [99].

The spatiotemporal gait parameters considered in the study are stride length and cadence.Age and limb length are considered as anthropometric parameters and are used to normalize the stride length and cadence. O'Malley et al. used fuzzy

Table 5.4: Anthropometric and the spatio-temporal gait parameters of the participating subjects

	88 Children with Spastic-diplegia, CP				**68 Neurologically Intact Children**			
	Min	**Max**	**Mean**	**Std**	**Min**	**Max**	**Mean**	**Std**
Age (years)	2	20	9.89	4.34	2	13	7.09	2.89
Leg length (m)	0.41	0.94	0.67	0.14	0.34	0.86	0.57	0.12
Stride Length (m)	0.31	1.25	0.74	0.21	0.67	1.44	1.03	0.19
Cadence (step/min)	10.46	210.24	120.00	33.56	104.88	174.24	136.84	15.81

based clustering techniques for the classification of control and CP children based on stance and cadence parameters. For study and comparison, in this thesis, the authors have used the same gait parameters as in [99].

Authors considered these parameters after averaging of data measured through at least three trials of both legs. The mean age of children with CP is 9.89 years while it is 7.09 years for the control group. Table 5.4 presents the overview of the CP gait dataset considered in this study.

5.2.1.2 Data Clustering Techniques

The objective of clustering problem is to divide a given dataset $X = (X_1, X_2, ...X_N)$ of N objects into K groups such that $K \leq N$. Let each object having F features, then dataset X can be represented as a matrix of size $N \times F$, where N is number of rows and P is number of column and represented as follows

$$X_{N\times F} = \begin{bmatrix} X_{1,1} & X_{1,2} & \cdots & X_{1,F} \\ X_{2,1} & X_{2,2} & \cdots & X_{2,F} \\ \vdots & \vdots & \vdots & \vdots \\ X_{N,1} & X_{N,2} & \cdots & X_{N,F} \end{bmatrix} \tag{5.13}$$

Following conditions should be satisfied while forming a cluster.

$$\begin{gathered} C = (C_1, C_2 \cdots C_K); \\ C_i \neq \phi \quad \forall i = (1, 2, \cdots K); \\ C_i \cap C_j \neq \phi \quad \forall i; j = (1, 2, \cdots K) \quad \& \quad i \neq j \\ \bigcup_{i=1}^{K} C_i = X \end{gathered} \tag{5.14}$$

where C_i is instances in cluster i and C is the set of K Clusters.

The purpose of clustering is to find K groups each having a centroid point μ_k in such a way that the similarity among the data objects within the cluster to the centroid of same cluster is maximum. In case of objects in the different cluster, this similarity should be minimum. Mathematically it can be modeled as a single objective optimization problem to find a C_k that has minimum similarity. The fitness function $f()$ can be represented as :

$$\begin{gathered} f(X_{N\times F}, C_k) = \sum_{i=1}^{N} Min\left\{\|X_{i\times F} - \mu_{k\times F}\|^2\right\} \\ \forall k = 1, 2\cdots K \\ Minimize \quad C_k^{Optmize} \end{gathered} \tag{5.15}$$

where $\|X_{i\times F} - \mu_{k\times F}\|^2$ represents the closeness between ith objects and centroid of kth cluster.

This similarity can be measured by distance within cluster variance or mean square error (MSE). In [210], Nanda et al. presents the potential similarity measures used in cluster analysis. The general form of similarity matrix is Minikowski metric as presented below.

$$d(X_i, \mu_k) = \left(\sum_{m=1}^{F}(X_{im} - \mu_{km})^r\right)^{\frac{1}{r}} \tag{5.16}$$

where c_j is the center of jth cluster μ_j, m is subset in F. Euclidean distance is special case with r as two is represented as equation (5.17).

$$d(X_i, \mu_k) = \sqrt{\sum_{m=1}^{F}(X_{im} - \mu_{km})^2} \tag{5.17}$$

Partitional, overlapping and hierarchical are the three main types of clustering approaches. Automated clustering techniques play an important role in image classification, intrusion detection, document clustering, medical imaging. Researchers have also explored the nature-inspired optimization algorithm for the clustering purpose [210].

K-means

The K-means algorithm is the simplest and most popular partition clustering approach. It is also known as hard clustering approach. In K-means, the overall

distance between the cluster members and the cluster centroids (intra-cluster distance) need to be optimized [214]. Mean square error (MSE) is considered as the objective function illustrated in eq (5.18).

$$\frac{1}{N}\sum_{k=1}^{K}\sum_{x_i \epsilon c_k} || \, x_i - C_k \, ||^2 \tag{5.18}$$

Where K is the total number of pre-defined clusters, k represents the clustered index, C_k is the centroid of the cluster k. This expression is needed to be minimized. The algorithm of K-means is presented in algorithm 3.

Algorithm 3 K-means Algorithm

1: Initialize number of cluster as (K) and *max_iter*
2: Select initial centroids randomly
3: For $iter = 0 : max\_iter$ or other termination criteria
4: Allocate each data-point to a cluster based on the similarity measure
5: Update the centroids by taking means of all points in the particular cluster
6: End

Data points assigned to a cluster are based on their degree of closeness, measured by the Euclidean distance from the points to the cluster's center. K-means approach is easy to implement but the time complexity increases with larger datasets and is sensitive to the initially provided centroids.

Fuzzy c-means (FCM)

Fuzzy logic was introduced by Zadeh during 1960s for handling the uncertain and imprecise knowledge in real-world applications [215]. The fuzzy c-means algorithm uses the reciprocal of distances between data in instances to decide the cluster centers. It is considered as the extension of the k-means and also known as soft clustering technique. The centroid of a cluster in a fuzzy c-means method is calculated as the mean of all points value, weighted by their degree of belongingness to the cluster. When the nature of cluster is overlapping, then fuzzy clustering is preferred [215].

In FCM, as illustrated in algorithm 4, to determine the best value of partition matrix U, the following objective function $J_k(U, Cj)$ is minimized as

Algorithm 4 Fuzzy c-means algorithm

1: Initialize number of cluster as k
2: Initialize μ_{ij} for all i,j from (5.21)
3: For Until termination criteria is met
4: Calculate cluster centers V_j
5: Update μ_{ij} for all i,j
6: End

$$J_k(U, Cj) = \sum_{i=1}^{k} \sum_{j=1}^{c_j} \mu_{ij}^m (||X_i - C_j||)^2 \tag{5.19}$$

$$1 < m < \infty$$

where m is any real number greater than one, μ_{ij} is the degree of membership of X_i in the cluster k^{th}. Degree of membership μ_{ij} and cluster center C_j is given by equations (5.20) and (5.21) respectively.

$$\mu_{ij} = \frac{1}{\sum_{p=1}^{k} \left(\frac{||X_i - C_j||}{||X_i - C_p||} \right)^{\frac{2}{m-1}}} \tag{5.20}$$

$$C_j = \frac{\sum_{i=1}^{N} \mu_{ij}^m X_i}{\sum_{i=1}^{N} \mu_{ij}^m} \tag{5.21}$$

In K-means clustering approach, the result depends on the choice of the randomly selected cluster centroids. If they are close, it takes time, while if the distance is on two extreme sides, then better cluster formation takes place. Thus the selection of the first cluster head points can be optimally selected using meta-heuristic approaches. In this study, nature-inspired optimization algorithms are used for the clustering purpose [210]. Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) are the most common in Evolutionary Algorithms. In this study GA, PSO and a hybrid version of both are proposed as clustering techniques with the following objective function [216].

Objective function : The objective function used in this study is the combination of intra and inter-cluster distance. Intra-cluster distance ($D_{inter}c_n$) is the overall distance between the cluster members and its centroid. While inter-cluster distance $D_{intra}c_n$ is the overall distance between the centroids.

Mathematically, it can be modeled as a multi-objective optimization problem to find a cluster centroid C_k that has maximum similarity. The fitness function $f()$

can be represented as (5.22).

$$
\begin{gathered}
f(X_{N\times F}, C_k) = \sum_{n=1}^{K} \{w_1 * D_{inter}(c_n) - w_2 * D_{intra}(c_n)\} \\
where \\
D_{inter}c_n = max \,||\, C_i - C_j \,||\,, \\
Here\ i, j\ \epsilon[1, 2, 3...k], i \neq j, \\
\forall k = 1, 2 \cdots K \\
D_{intra}c_n = \sum_{i=1}^{N} Min \left\{ \|X_{i\times F} - C_{k\times F}\|^2 \right\} \\
Max \quad C_k^{Optmize}
\end{gathered}
\tag{5.22}
$$

Where i is the data point of the kth cluster, K is the total number of pre-defined clusters, k represents the cluster index, C_k is the centroid of the cluster k. w_1 and w_2 are weighted parameters. In this thesis, we have considered both w_1 and w_2 as 0.5 for the study purpose [216]. This objective function is to be maximized.

Genetic Algorithm (GA) and Variant (H-GA)

Evolutionary techniques are a good approach to the optimization problem. Based on Darwin's principle of natural selection, Genetic Algorithm (GA)was proposed by Holland et al. in 1975 [217].

In 1994, Bezdek et al. explored basic GA to be used as a clustering approach [218]. Similar to basic GA, individual features are characterized in the form of strings called chromosomes. Based on the number of clusters (K) and the population size (N), the chromosomes are considered as the cluster centroids candidate. The algorithm initiates by creating the random set of chromosomes in the search space followed by an iterative process of selection (maximum number of iterations $maxiter$), crossover probability (C_{prob}) and mutation probability (M_{prob}) to find the optimal K solution (cluster center (centroids)) in the search space. It is depicted in Algorithm 5.

In 1999, Krishna and Murty proposed a novel hybrid Genetic K-means algorithm. K-means help in the search operation during crossover [219].

In this study, GA is hybridized with K-means algorithm to get the optimal cluster centers and effective fitness values as shown in algorithm 6 [220].

Algorithm 5 Pseudo-Code for the GA based Clustering

1: Initialize population size (N), maximum number of iterations $maxiter$, crossover probability (C_{prob}) and mutation probability (M_{prob})
2: Generate initial population randomly where each chromosome acts as cluster centers
3: For $iter = 0 : maxiter$
4: Apply selection in the population
5: Apply crossover in the population
6: Apply mutation in the population
7: End

Algorithm 6 Pseudo-Code for the hybrid-GA based Clustering

1: Initialize K, N, $maxiter$, C_{prob} and M_{prob}
2: Generate initial population randomly where each chromosome act as set of cluster centers
3: For $iter = 0 : maxiter$
4: Apply selection in the population
5: Apply crossover in the population
6: Apply mutation in the population
7: End
8: Output is the K optimal cluster center with minimum MSE
9: Apply K-means by initializing max_iter_k
10: Select initial centroids as output in step 8
11: For $iter = 0 : max\_iter\_k$ or other termination criteria
12: Allocate each data-point to a cluster based on the similarity measure
13: Update the centroids using eq. (5.22)
14: End

Hybrid GA optimizes the location of starting centroid (cluster centers) from the population with the minimum fitness function. Use these cluster centers as the initial centroids in the k-mean and iterate the steps until no significant changes in consecutive cluster centers.

Particle Swarm Optimization (PSO) and Variant (H-PSO)

Particle Swarm Optimization technique mimics swarm (birds, fish, etc.) social behavior for food searching and was developed by Eberhart and Kennedy in 1995 [221]. As it has a fast convergence rate and is easy to implement, so it is a popular choice among the researchers. Omran et al. in 2002, proposed cluster analysis using PSO for image processing [222]. One year later, Merwe explored PSO for data clustering on different datasets [223]. The PSO based clustering algorithm is shown in algorithm 7.

Algorithm 7 PSO based Clustering Algorithm

1: Initialize no of cluster as k
2: Initialize population size (N), maximum number of iterations $maxiter$ and inertia weight w
3: Generate initial population randomly with centroids as its parameters
4: For $iter = 0 : maxiter$
5: Find fitness function value for all particles
6: Update velocity of particles by Eq (5.26)
7: Update position of particles by Eq (5.27)
8: Update G_b and P_{b_i}
9: End

In PSO based clustering algorithm, swarm is defined in search space according to the constraints. The pseudo-code of PSO is presented. Population, initial weights, iteration is defined initially. Each particle position is defined as

$$x_i = [\mu_{i1}, \mu_{i1}, \cdots, \mu_{iK}] \tag{5.23}$$

where K cluster centroid vectors. The cluster C_{ik} has μ_{ik} as the cluster center point (centroid).

Fitness function is considered as the eq (5.22). Based on this, previous position is considered as particle best position P_b as

$$P_b = [P_{b1}, P_{b2}, \cdots, P_{bK}] \tag{5.24}$$

In the beginning P_b is consider as X_i as $X = (X_1, X_2, ...X_N)$. Global solution G_b is the best position of swarm in the next iteration (t) and represented as follows.

$$G_b = [G_{b1}, G_{b2}, \cdots, G_{bt}] \tag{5.25}$$

Cluster centroids' positions are updated with updation of velocity (v_{ik}) and positions (x_i) of the particles as shown below.

$$v_{ik}(t+1) = w \times v_{ik}(t) + c_1 \times r_1 \times P_{b_k}(t) - x_{ik}(t)) + c_2 \times r_2 \times (G_b(t) - x_{ik}(t)) \tag{5.26}$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \tag{5.27}$$

Algorithm 8 Hybrid-PSO Algorithm

```
Initialize number of clusters as k
Initialize population size (N), maximum number of iterations maxiter and
inertia weight w
Generate initial population randomly with centroids as its parameters
    For iter = 0 : maxiter
        Find fitness function value for all particles
        Update velocity of particles by eq. (5.26)
        Update position of particles by eq. (5.27)
        Update g_b and p_{b_i}
    End
Apply K-means Algorithm to return the centroids
```

where t is the number of iterations, r_1 and r_2 are random numbers between [0,1], w is inertia weight, c_1 and c_2 are acceleration constants. In this thesis, inertia weight is taken as a function of time and it varies from w is 0.9 to 0.5. This iteration continues till

i) it reaches the predefined number of iteration

ii) there is no further change in the centroid of the clusters in the next iteration.

Hybrid PSO - Variant for Clustering Approach

Van der Merwe and Engelbrechtdeveloped a hybrid algorithm based on K-means and PSO in 2003 [223]. In this work, a combination of PSO and k-means is used as expressed in algorithm 8.

5.2.1.3 Proposed Grey Wolf Optimization (GWO) Clustering Technique

This study presents a nature-inspired novel approach, Grey Wolf Optimization for clustering. To validate our proposed methods, we have compared with the results with those obtained from traditional clustering technique (k-means) and two most common nature-inspired optimization based clustering approach (GA and PSO and Hybrid GA and PSO).

Grey Wolf Optimization (GWO)

GWO is comparatively new nature-inspired, population-based meta-heuristic algorithm proposed by Mirjalili et al. in 2013 [224]. It is another example of swarm optimization. Grey wolves inhabit the mountains, forests, plains of North America, Asia, and Europe. The social hierarchy (alpha, beta, delta, and omega are the four types of grey wolves) and hunting approach (searching, encircling and attacking) observed in Grey wolf (Canis lupus) are the sources of inspiration for the optimization algorithm as shown in figure 5.12 [224].

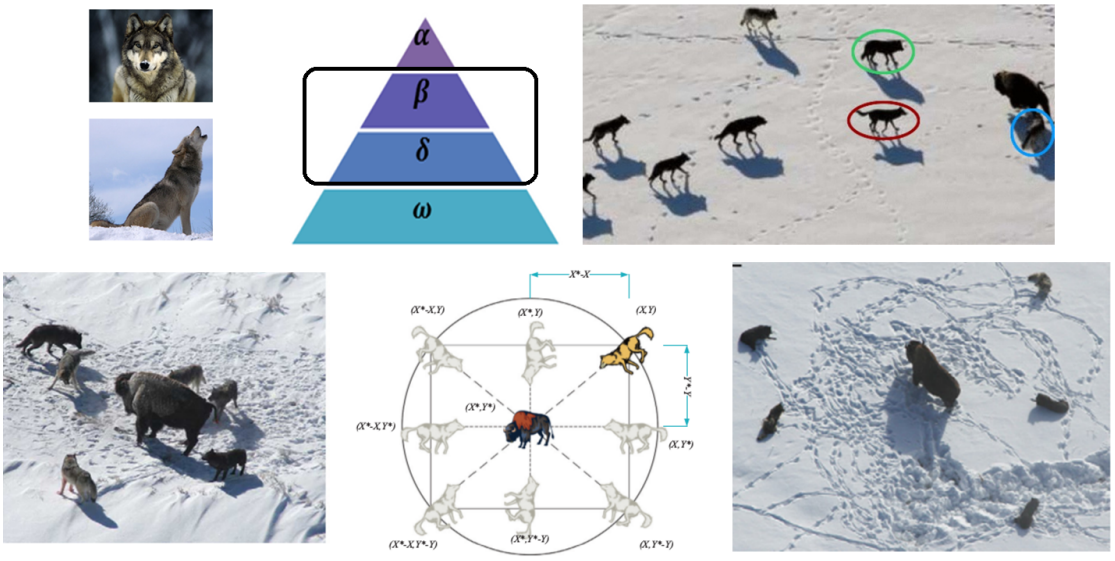


Figure 5.12: Grey Wolf Optimization

Grey wolves prefer to live in an average group size of 5-12 and have a very strict social dominant hierarchy consisting of alpha, beta, delta, and omega. Alpha wolf (male/female) is considered as the most dominant and decision marker in the group and serve as a leader. Alpha is responsible for managing the team and making an important decision such as hunting, time to wake, etc. Beta wolf (male/female) acts as an advisor and helps the leader (alpha wolf) in decision making and other group activities. It may be the next one to serve as the leader in case of either death of alpha or alpha become very old. Beta is considered as the discipliner for the group and gives feedback to the leader.

The lowest level in the social hierarchy of wolves consists of omega. They follow the commands of alpha and other dominant wolves and sometimes act as the babysitter in the group. They have to eat last and act as the scapegoat. The remaining wolves are the delta one and also known as subordinate. They may be elders, caretakers, scouts, sentinels, and hunters. Elders are the experienced

wolves who used to be alpha or beta. Delta wolves dominate the omega but have to submit to alpha and beta. The hunting approach of the wolf consists of searching, tracking and chasing prey, followed by pursuing, encircling, and harassing the prey until it stops moving. At the end, when the prey stops, wolves make an approximate regular polygon around it and lay down.

Mathematically the Grey wolf social hierarchy can be formulated in GWO. The fittest candidate is considered as alpha (α) followed by second and third best solution as beta (β) and delta (δ) respectively. The remaining solutions act as omega (ω). The optimization problem can be considered as the hunting problem where ω follows the α, β and δ.

Similarly the hunting steps such as encircling prey, hunting, attacking (exploitation), search for prey (exploration) can be modeled mathematically [224]. Grey wolf can update its location around the prey in any random location by using (5.28) and (5.29)

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A}.\vec{P} \tag{5.28}$$

where $\vec{X}$ is updated in the search agent position (grey wolf in our case), t is the current iteration number and $\vec{X}_p$ is prey position, P is position of the search agent and is defined in (5.29)

$$\vec{P} = |C.\vec{X}_p(t) - \vec{X}(t)| \tag{5.29}$$

where A and C are coefficient vectors and updated iteratively using equation (5.30) and (5.31)

$$\vec{A} = 2ar_1 - a \tag{5.30}$$

$$\vec{C} = 2r_2 \tag{5.31}$$

where r_1 and r_2 are random vectors between zero and one used by agents (wolves) to reach new position. The value of a is linearly decreasing from two to zero [224] as presented in equation (5.35).

In hunting, alpha is responsible for guiding other wolves. Seldom beta and delta also contribute as leaders. It is considered that alpha, beta, and delta have information of the potential position of the prey. α, β and δ have the best solution to the problem and α is the best candidate solution among them.

Hunting behavior of grey wolves can be simulated as follows. New positions of

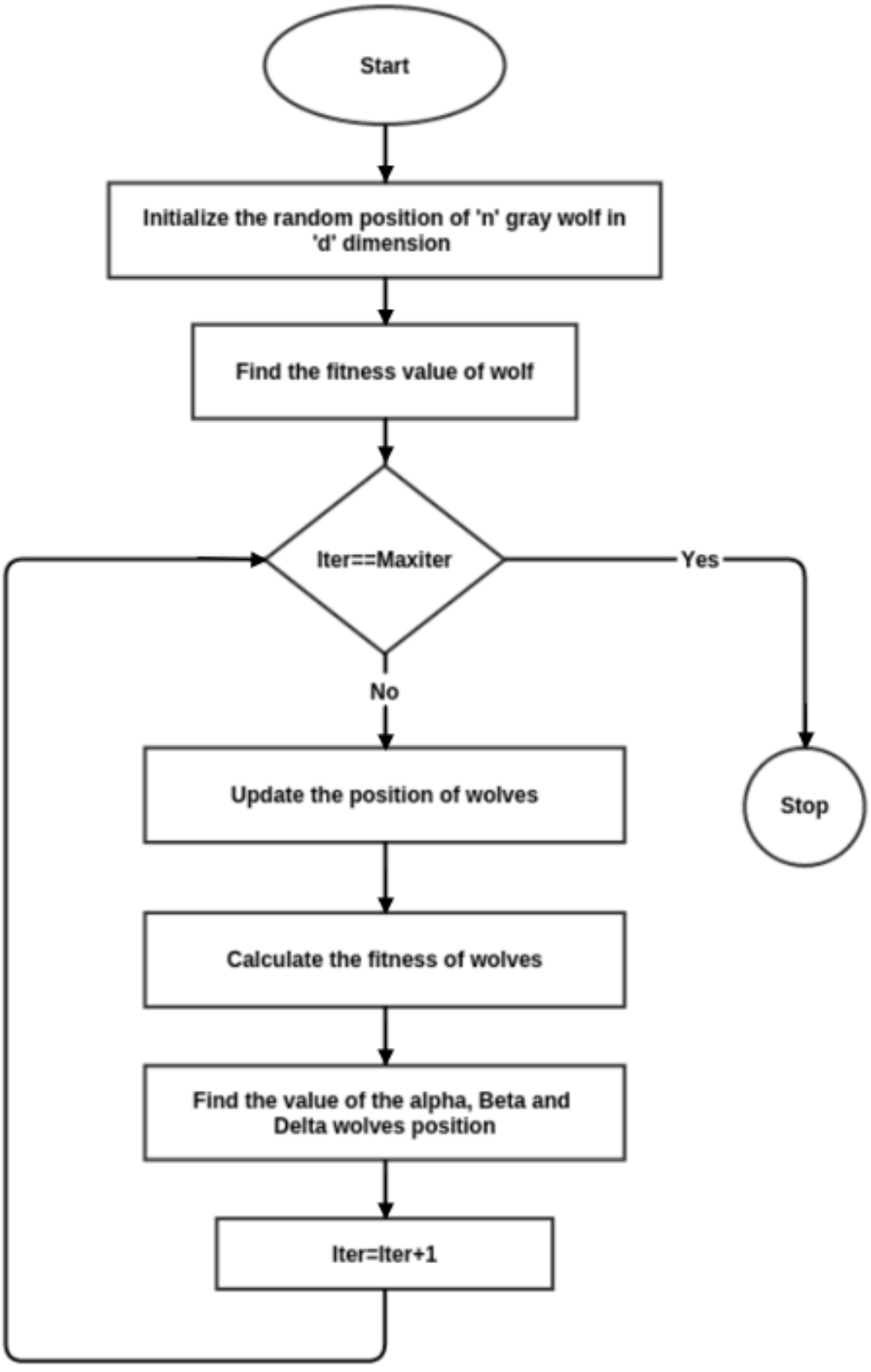


Figure 5.13: Flow chart for Grey Wolf Optimization (GWO) approach

wolf during hunting can be models as (5.32)

$$X(\vec{t}+1) = \frac{(\vec{X}_1 + \vec{X}_2 + \vec{X}_3)}{3}. \tag{5.32}$$

where $X(t+1)$ is the position of the search agent at $(t+1)th$ iteration; X_1,X_2 and X_3 represent the best three solutions (positions) and are represented as (5.33)

$$\begin{aligned} \vec{X}_1 &= \vec{X}_\alpha - \vec{A}_1.\vec{P}_\alpha, \\ \vec{X}_2 &= \vec{X}_\beta - \vec{A}_2.\vec{P}_\beta, \\ \vec{X}_3 &= \vec{X}_\delta - \vec{A}_3.\vec{P}_\delta \end{aligned} \tag{5.33}$$

where X_α, X_β and X_δ is the position of alpha, beta and delta wolves in the population respectively. The search agents are updated using (5.34) as follows:

$$\begin{aligned} \vec{P}_\alpha &= \left|\vec{C}_1.\vec{X}_\alpha - \vec{X}\right| \\ \vec{P}_\beta &= \left|\vec{C}_2.\vec{X}_\beta - \vec{X}\right| \\ \vec{P}_\delta &= \left|\vec{C}_3.\vec{X}_\delta - \vec{X}\right| \end{aligned} \tag{5.34}$$

where P_α, P_β and P_δ are the positions of the search agent with respect to the alpha, beta and delta wolves respectively.

Based on this newly updated position in equation (5.32), other search agents are forced to upgrade their positions according to the location of the best search agents around the prey.

Algorithm 9 Pseudo-code of Grey Wolf Optimization Approach

1: Initialize population size N, control vectors $\vec{a}$, $\vec{A}$ and $\vec{C}$ and maximum number of iterations $Maxiter$
2: Generate initial population randomly
3: Assign the values of the first, second and third best solution X_α, X_β and X_δ respectively
4: For $iter = 0 : maxiter$
5: Update each search agent using (5.32)
6: Update $\vec{a}$, $\vec{A}$ and $\vec{C}$
7: Evaluate fitness of each search agent
8: Update the vectors X_α, X_β and X_δ
9: End

Attacking and searching for prey can be considered as the exploitation and exploration of the optimization problem respectively and hunting stops when the prey does not move. $|A| < 1$ forces the wolves to attack towards the prey. $|A| > 1$

forces the grey wolves to diverge from the prey to hopefully to find a fitter prey. The trade-off between exploitation and exploration is controlled by a in equation (5.35)

$$a = 2 - t\frac{2}{MaxIter} \tag{5.35}$$

Where t is the iteration number and $MaxIter$ is the maximum number of attempts for optimization. The flow chart of the Grey Wolf optimization approach is shown in figure 5.13.

Grey Wolf Optimization for Clustering (GWOC)

The methodology for GWO approach for Clustering has been discussed. To use GWO as a clustering approach, following two modification are required in the GWO for the numerical optimization problem. The problem should be formulated considering the centroids and the fitness function, as the MSE as given in Eq. (5.22). A brief description of these is discussed below.

Problem Formulation:

There is a difference in the GWO optimization based numerical solution and GWO based clustering solution. Each prey location is the possible solution for the first case, while for the latter case it is one of the possible set of cluster center as presented in Eq. (5.36)

$$X_i = x_1, x_2, \cdots, x_p, x_{p+1}, \cdots, x_{K\times F} \tag{5.36}$$

Pray can be considered to the cluster centers as shown in Eq. (5.37)

$$c_m = x_{(m-1)\times(F+1)}, x_{(m-1)\times(F+2)}, \cdots, x_{(m\times F)} \tag{5.37}$$

Where X_i presents a prey location in Grey Wolf optimization approach. K is the number of clusters, and F is the number of features in the dataset. The size of the data matrix of GWO is $K \times F$.

Objective function: MSE from equation (5.18) is used to evaluate the quality of the clusters partition. The pseudo code of GWO algorithm for solving clustering problems is shown in algorithm 10. The other parth of the approach is similar to above section.

Algorithm 10 Grey Wolf Optimized Algorithm

Initialize population size N, control vectors $\vec{a}$, $\vec{A}$ and $\vec{C}$ and maximum number of iterations $Maxiter$
Generate initial population randomly
Assign the values of the first, second and third best solution X_α, X_β and X_δ respectively
    For $iter = 0 : maxiter$
        Update each search agent using (5.32)
        Update $\vec{a}$, $\vec{A}$ and $\vec{C}$
        Evaluate fitness with **Input** : Pray Location X, Data D
            For $eachsearchagent(wolf)locationX_i$
                Decode Xi to the k cluster centers using Eq. (17)
                Calculate the distance between all objects in D and each cluster center using Eq. 5.31 with r=2
                Assign objects to the nearest clusters' centers
                Compute the total MSE using Eq. (11)
                $Fit_i = MSE_i$
                Output : Fitness fit
            End
Update the vectors X_α, X_β and X_δ
    End

Table 5.5: Cases considered in this study

CASE	Gait Parameter/s Considered	Data Size	Normalization	Scaling
Case 1	Stride Length, Cadence, Leg length and Age	156×4	×	×
Case 2	Stride length and Cadence	156×2	×	×
Case 3	Stride length and Cadence	156×2	✓ Polynomial	✓
Case 4	Stride length and Cadence and Different cluster size $2 \leq K \leq 7$	156×2	✓ Polynomial	✓
Case 5	Validate Test case using Stride length and Cadence and cluster size from case 4.	156×2	✓ Polynomial	✓

5.2.1.4 Optimization Parameter Settings

The parameters affecting the performance of optimized clustering are taken as follows. In this study, population size is taken as 100 and 40 for GA and PSO respectively. Crossover rate and mutation are chosen as 0.3 and 0.2. In this study, a time-varying inertia weight w varying from 0.9 to 0.5 is considered. w_1 and w_2 are 0.5 in objective function. All these parameters are tuned by using sensitivity analysis [225].

5.2.2 Case Studies Considered

For the analysis, five different cases are considered as shown in table 5.5, to explore the best possible features selection.

Case 1: Considering all the four gait parameters (stride length, cadence, leg length, and age) without normalization and scaling.

Case 2: The first two gait parameters (stride length and cadence) in their original form are considered. In [99, 203, 212] authors suggest that stride length and cadences are clinically more significant for CP analysis than another kinematics parameter when focused on single joints, in both classification and clustering study. Stride length and cadences could be affected by age and leg length. Thus before applying any machine learning techniques, normalization model is required to remove trends that/if exist, in the CP and normal dataset concerning age and leg length.

Case 3: Considering stride length and cadence, after Polynomial normalization with leg length and age respectively. O'Malley et al. suggest that detendring normalization technique is better than offset and decorrelation based normalization methods [99]. They suggested first and second order polynomial model to normalize stride length for leg length and cadence concerning age, respectively for each subject. In our study, we consider stride length and cadence after normalization and scaling procedure suggested by [99]. They are statistically independent and significant in CP analysis. Table 5.6 presents an overview of the case considered in the study.

Table 5.6: Testing gait data of four subjects A, B, C and D

Subject	State/Pathology	Stride length (m)	Cadence (m)	Leg length (m)	Age (year)	Treatment Condition
A	Neurologically Intact	1.29	112.8	0.78	13	Normal
B	Neurologically Intact	1.29	122.6	0.79	19	Normal
C1	Spastic-diplegia, CP	0.59	134.0	0.66	8	Prior to Surgery
C2	Spastic-diplegia, CP	0.89	110.0	0.67	9	1 year post Surgery
C3	Spastic-diplegia, CP	1.04	119.0	0.71	11	3 year Post Surgery
D1	Spastic-diplegia, CP	0.20	49.5	0.45	3	Prior to Surgery
D2	Spastic-diplegia, CP	0.51	74.0	0.47	4	1 year Post Surgery
D3	Spastic-diplegia, CP	0.76	131.0	0.52	5	2 year Post Surgery

Case 4: Finding the optimal number of clusters in a dataset is an open research area. In the case of rehabilitation, the number of gait profiling of the patients can play a vital role. In case 4, the optimal number of gait profiles in the dataset is identified.

Case 5 Evaluation of the test subject is necessary to demonstrate the significance of the surgery. To validate the clustered gait profile, four subjects are taken as case 5 as presented in table 5.6 from [99]. A and B are normal subjects while C and D are examined before and after surgery.

5.2.3 Result Analysis: Optimization based Clustering Techniques

5.2.3.1 Clustering Evaluation Indices

The clustering performance is evaluated using six clustering performance indices, including both internal (based on intrinsic characteristics, thus considered as unsupervised) measures and external (based on previous knowledge about data, thus can be considered as supervised) cluster validity indices. One external validity index (Cluster Purity Index (CPI)) and five internal validity measures (Distance Measures (intra-cluster distance and inter-cluster distance), mean square error (MSE), silhouette coefficient (SC) and Dunn index (DI)) [210] are considered in this study. The external cluster validity index is considered to be close to user's semantic [211].

The clustering evaluation indices considered in this study are discussed below.

(a) Cluster Purity Index (CPI): It is considered as the external validity measure and applicable only when true class labels of each sample are pre-known. The purity of the clustering concerning the known categories is given by (5.38)

$$CPI = \frac{1}{N} \sum_{k=1}^{K} \max_{1 \leq j \leq l} N_k^j \tag{5.38}$$

where N is the total number of samples; K is number of the clusters to be formed; N_k^j is the number of samples in cluster k that belongs to original class j. The purity index is a real number in [0,1]. A high value of CPI indicates a better clustering performance.

(b) Distance Measures: Intra-cluster and inter-cluster distance are the internal indices to validate the clustering technique. Intra-cluster distance is the distance between data objects and its centroid within a cluster as presented in Eq (5.33). The inter-cluster distance is the distance between the centroids of the clusters and expressed as

$$Inter\ Cluster\ Distance = min\,||\,\mu_i - \mu_j\,|| \tag{5.39}$$

where $i, j\ \epsilon[1, 2, 3...k], i \neq j$

(c) Mean Square Error (MSE): It is the error about the origin given by (5.18)

$$MSE = \frac{\sum_{k=1}^{k} \sum_{x_i \epsilon c_k} \frac{||x_i - \mu_k||^2}{N_k}}{K} \tag{5.40}$$

(d) Silhouette Coefficient (SC): It is an internal measure for evaluating clustering performance. It is mathematically formulated after combining the measures of cohesion and separation but for individual points. Cluster cohesion represents the sum of the weight of all links within a group while separation is the sum of the weights between nodes in the cluster and nodes outside the cluster.

$$SC = \frac{1}{N} \sum_{i=1}^{N} \frac{b(i) - a(i)}{max[a(i), b(i)]} \tag{5.41}$$

where N is the number of objects, $a(i)$ is the distance between object i and its centroid, and $b(i)$ is the smallest distance between object i and the other centroids. The value of SC varies in the interval $[-1, +1]$ and is normalized to $[0, 1]$, where 1 indicates the best possible clustering

(e) Dunn index (DI) : It is another internal cluster validity index introduced by J. C. Dunn in 1974 [226]. It is expressed as following

$$\begin{gathered} DI(K) = \min_{i=1 \subset K} \left\{ \min_{j=1 \subset K, j \neq i} \left\{ \frac{d\left(c_i, c_j\right)}{\max_{k \subset K(Diam(c_k))}} \right\} \right\} \\ where \quad d\left(c_i, c_j\right) = \min_{x \subset c_i, y \subset c_j} \{d(x, y)\} \; and \\ diam(c_i) = \max_{x, y \subset c_i} \{d(x, y)\} \end{gathered} \tag{5.42}$$

(f) Affiliated Probability Index (API): It is defined as the probability of belongingness of a given test sample Ts_i over the given set of cluster C and represented as

$$AP(Ts_i, Cj) = \frac{\left(\frac{1}{D(Ts_i, Cj)}\right)}{\sum_{m=1}^{K} \left(\frac{1}{D(Ts_i, Cj)}\right)} \tag{5.43}$$

where $D(Ts_i, Cj)$ is the affiliation distance between a testing subject and a given cluster.

Table 5.7: Clustering result for case 1, when all four gait parameters (stride length, cadence, leg length and age) without normalization and scaling for k=2 are considered.

Algorithm	Cluster purity index	Intra-cluster	Inter-cluster	MSE	Silhouette Coefficient	Dunn Index
K-means	0.647435897	727.414639437	2249.398883024	1454.829278874	0.534981035	0.021440854
FCM	**0.660256410**	747.354298589	2218.037446776	1464.736588728	0.516406269	0.004728326
GA	0.636538462	**720.143215298**	2397.168829263	**1440.286430596**	0.543759910	**0.046664080**
H-GA	0.639743590	722.851700278	**2453.155978148**	144 5.703400555	**0.553784374**	0.045306570
PSO	0.626923077	723.916326896	2254.787022798	1447.832653792	0.523427378	0.036291725
H-PSO	0.639743590	722.851700278	**2453.155978148**	1445.703400555	**0.553784374**	0.045306570
GWO	0.625000000	725.947356230	2141.823025944	1440.486430593	0.513647761	0.033415870

Table 5.8: Clustering result for case 2, when first two gait parameters (stride length and cadence) in their original form for k=2

Algorithm	Cluster purity index	Intra-cluster	Inter-cluster	MSE	Silhouette Coefficient	Dunn Index
K-means	0.637179487	702.160532015	2190.654305937	1404.321064030	0.546380125	0.011359487
FCM	**0.666666667**	722.581050726	2221.344900625	1451.894712461	0.532371425	0.005553906
GA	0.647435897	689.652285216	2636.076130618	**1379.304570433**	**0.589329956**	**0.026969577**
H-GA	0.647435897	**689.520056309**	2640.796659385	1379.040112617	**0.589329956**	**0.026969577**
PSO	0.647435897	689.766697516	**2642.220564311**	1379.533395032	**0.589329956**	**0.026969577**
H-PSO	0.647435897	**689.520056309**	2640.796659385	1379.040112617	**0.589329956**	**0.026969577**
GWO	**0.647435897**	689.989739760	**2658.670056246**	**344.760028154**	**0.589329956**	**0.026969578**

5.2.3.2 Discussion

In this section, the analysis of the cases considered is discussed. From analysis of table 5.7, it is found that the result of H-GA and H-PSO is the same as the first cluster is optimally selected and is the same for both cases.

Case 1: Table 5.7 presents clustering result for case 1.

FCM yields the best external evaluation index; cluster purity index when compared with the mean. Considering other internal evaluation measures, GA reports the best result in this case. The best one among each is highlighted in table 5.7. The value stated here is the mean of 25 runs.

Case 2: H-GA based optimized clustering algorithms performed best on four internal clustering performance indices. The best cluster purity index is given by FCM as reported in table 5.8 case 2. The overlapping nature of CP gait data with the normal children may be reasoned for this.

Considering other evaluation indices, H-GA based clustering outperforms other traditional partitioning clustering techniques. Case 2 reported Minimum MSE than case 1.

Case 3: Figure 5.14, illustrates the significance of the polynomial normalization for leg and age on stride length and cadence respectively. For this study, stride

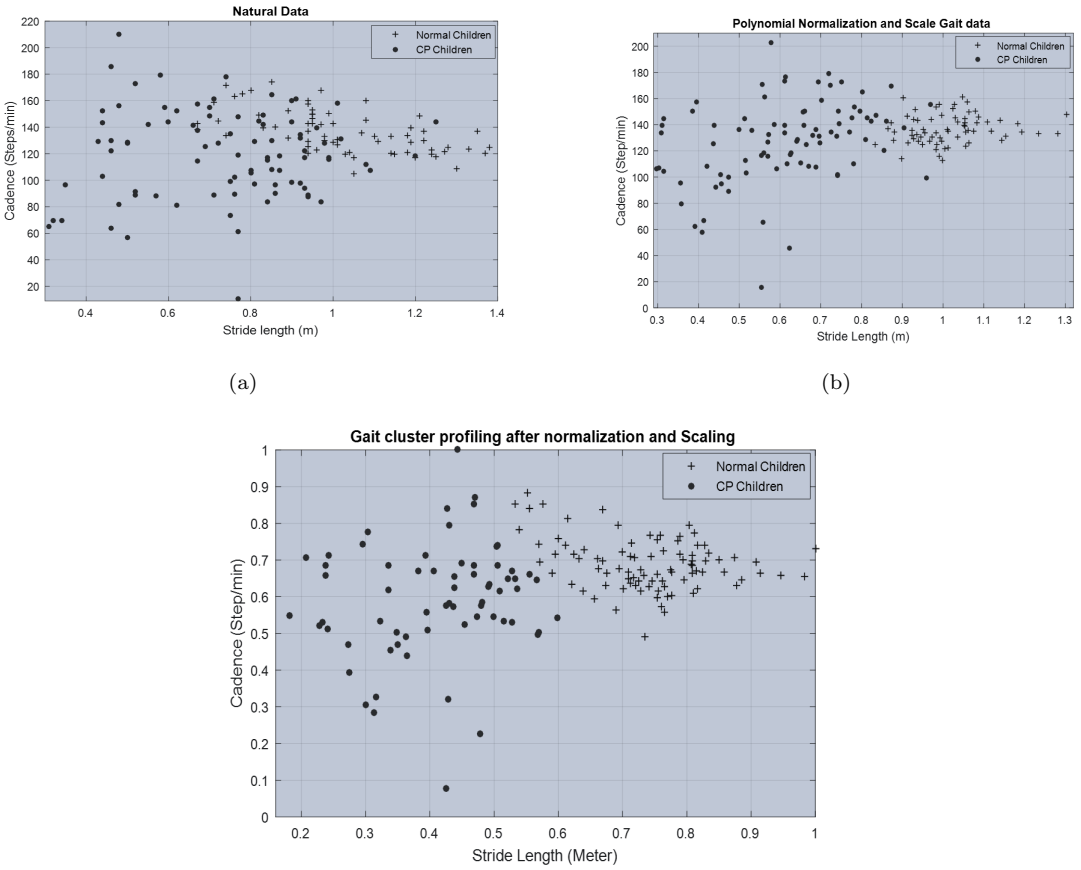


Figure 5.14: Two dimensional plot of gait data for (a) CP and Normal Children natural data; (b) Polynomial Normalization and scale gait data with CP and Normal Children and (c) Gait Profiling after clustering approach.

length can be considered a significant factor in discriminating the children with CP from the control group (neurologically intact children) when visualized by plotting a figure. The red dot represents the children with CP, and blue is neurologically intact children. A clear separation in the data set is observed after normalization, considered case 3 for $k = 2$.

Table 5.9 presents the result of K-means, GA, PSO, H-GA and PSO optimized clustering on gait data for k=2. As shown in table, GWO based clustering obtained the best cluster purity index. The best one among the indices is highlighted in the mentioned table. GWO outperforms other clustering approaches in term of external validation index, cluster purity index.

Bar plots of clustering performance indices from the different clustering approaches are presented in figure 5.15. The mean square error convergence profiles of the Kmeans, GA, PSO H-GA/PSO and GWO algorithms are shown in figure 5.16.

Table 5.9: Clustering result for case 3, when considering stride length and cadence, after Polynomial normalization with leg length and age respectively for k=2

Algorithm		Cluster Purity Index	Intra-Cluster	Inter-Cluster	MSE	Silhouette Coefficient	Dunn Index
K-means	Mean	0.8551	**0.0484**	**0.1176**	0.0242	**0.4996**	0.0332
	Stan Dev.	0.0033	0.0005	0.0014	0.0003	0.0038	0.0017
	Min	0.8526	0.0474	0.1156	0.0237	0.4957	0.0327
	Max	0.8590	0.0491	0.1202	0.0245	0.5056	0.0381
GA	Mean	0.8776	0.0489	0.1167	0.0245	0.4902	**0.0334**
	Stan Dev.	0.0332	0.0002	0.0058	0.0001	0.0038	0.0044
	Min	0.8269	0.0487	0.1067	0.0243	0.4849	0.0280
	Max	0.9167	0.0494	0.1257	0.0247	0.4939	0.0379
PSO	Mean	0.8955	0.0489	0.1130	**0.0244**	0.4887	0.0290
	Stan Dev.	0.0157	0.0002	0.0049	0.0001	0.0019	0.0031
	Min	0.8526	0.0487	0.1072	0.0243	0.4880	0.0280
	Max	0.9038	0.0495	0.1215	0.0247	0.4939	0.0379
H-GA	Mean	0.8590	0.0491	0.1175	0.0246	0.4951	0.0327
	Stan Dev.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Min	0.8590	0.0491	0.1175	0.0246	0.4951	0.0327
	Max	0.8590	0.0491	0.1175	0.0246	0.4951	0.0327
H-PSO	Mean	0.8590	0.0491	0.1175	0.0246	0.4951	0.0327
	Stan Dev.	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Min	0.8590	0.0491	0.1175	0.0246	0.4951	0.0327
	Max	0.8590	0.0491	0.1175	0.0246	0.4951	0.0327
GWO	Mean	**0.8994**	0.0488	0.1140	**0.0244**	0.4881	0.0280
	Stan Dev.	0.0043	0.0001	0.0046	0.0000	0.0002	0.0000
	Min	0.8910	0.0487	0.1061	0.0243	0.4880	0.0280
	Max	0.9038	0.0489	0.1211	0.0244	0.4888	0.0281

GWO converge to 0 in 25th iteration. Even it was high in the starting, but after 23rd iteration, it Mean Square Error is converged to 0.024395. Figure 5.15, presents the bar plots of clustering performances indices from the different clusters. The results are from the average of 25 tests.

Although classical performance metrics such as mean square error, silhouette coefficient, and Dunn index are suitable methods of comparing the algorithms, they are not sufficient to find a difference in performance of the algorithms. To aggregate the performance comparison and statistical significance of computational intelligence algorithms, the popularity of parametric and nonparametric tests has increased in last few years. T-test (parametric) carried out for comparing different algorithms. The t-test assesses whether the mean of two groups of results is statistically different from each other or not. For testing, the two-tailed t-test is adopted with 5% significance level. The negative t- value with PSO as base algorithm along with low p-value and h-value of 1 w.r.t. all the other algorithms proves PSO to be significantly better than other algorithms including GA. The further comparison is made in table 5.10.

Case 4: The result of different cluster sizing on three internal cluster validity indices (MSE, SC and DI) in case of $2 \leq k \leq 7$ is presented in figure 5.17. Figure

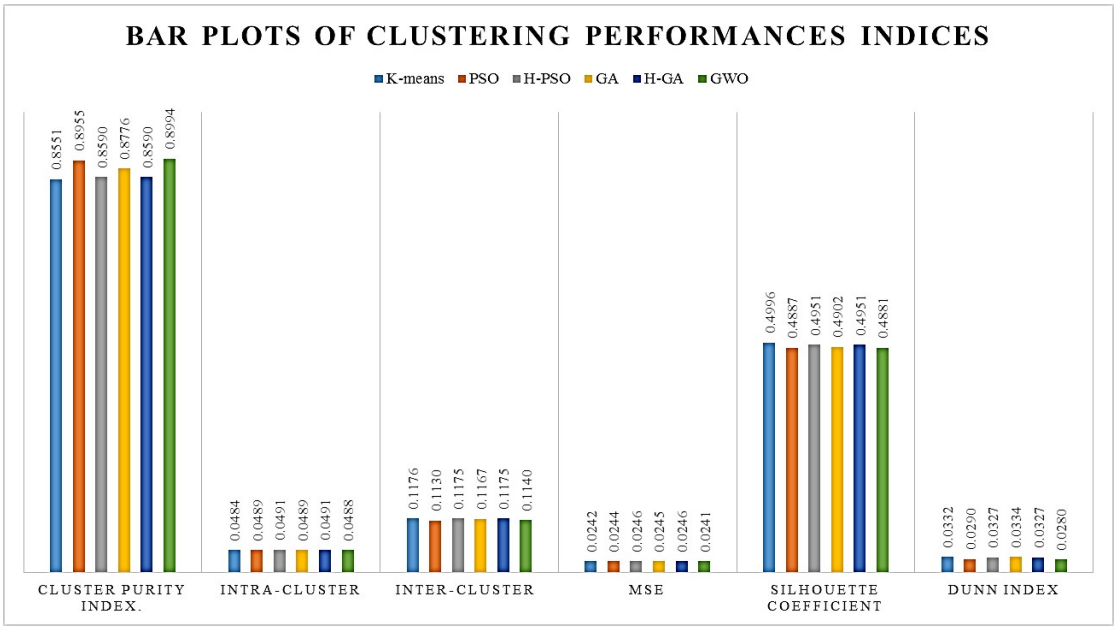


Figure 5.15: Bar plots of clustering performances indices from the different clustering approaches

Table 5.10: **Result of t-test considering MSE on case 3**

Algorithms		**K-means**	**FCM**	**GA**	**H-GA**	**PSO**	**H-PSO**
K-means	t	0.0000	-2.2820	28.8308	-6.3298	31.9412	-6.3298
	h	0.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	p	1.0000	0.0349	0.0000	0.0000	0.0000	0.0000
FCM	t	2.2820	0.0000	24.6364	-1.7860	26.2677	-1.7860
	h	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	p	0.0349	1.0000	0.0000	0.0910	0.0000	0.0910
GA	t	-28.8308	-24.6364	0.0000	-80.2971	1.2556	-80.2971
	h	1.0000	1.0000	0.0000	1.0000	0.0000	1.0000
	p	0.0000	0.0000	1.0000	0.0000	0.2253	0.0000
H-GA	t	6.3298	1.7860	80.2971	0.0000	192.4919	0.0000
	h	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	p	0.0000	0.0910	0.0000	1.0000	0.0000	1.0000
PSO	t	-31.9412	-26.2677	-1.2556	-192.4919	0.0000	-192.4919
	h	1.0000	1.0000	0.0000	1.0000	0.0000	1.0000
	p	0.0000	0.0000	0.2253	0.0000	1.0000	0.0000
H-PSO	t	6.3298	1.7860	80.2971	0.0000	192.4919	0.0000
	h	1.0000	0.0000	1.0000	0.0000	1.0000	0.0000
	p	0.0000	0.0910	0.0000	1.0000	0.0000	1.0000

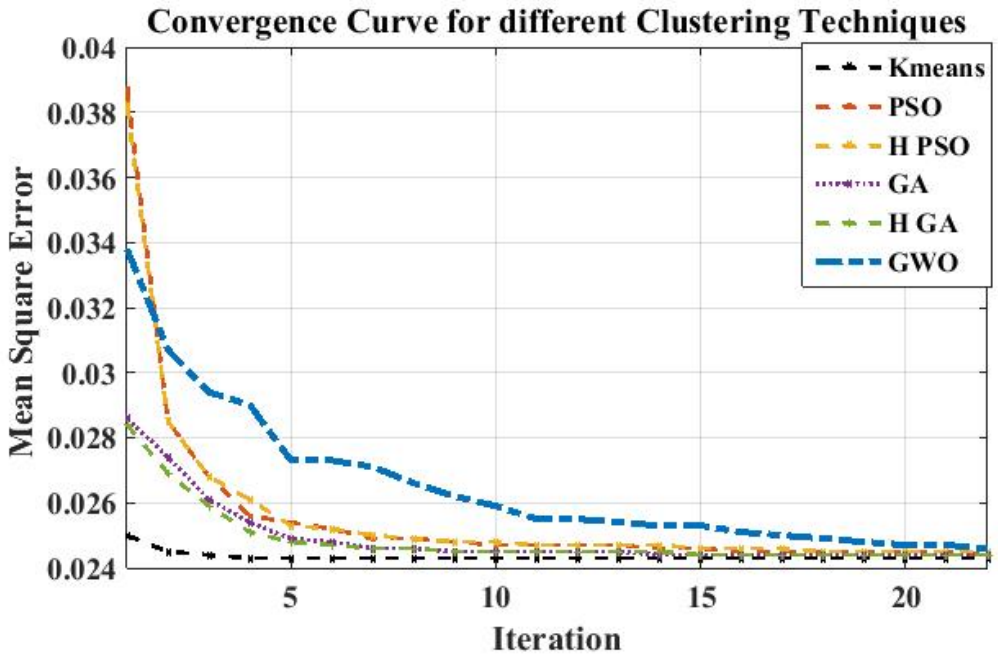
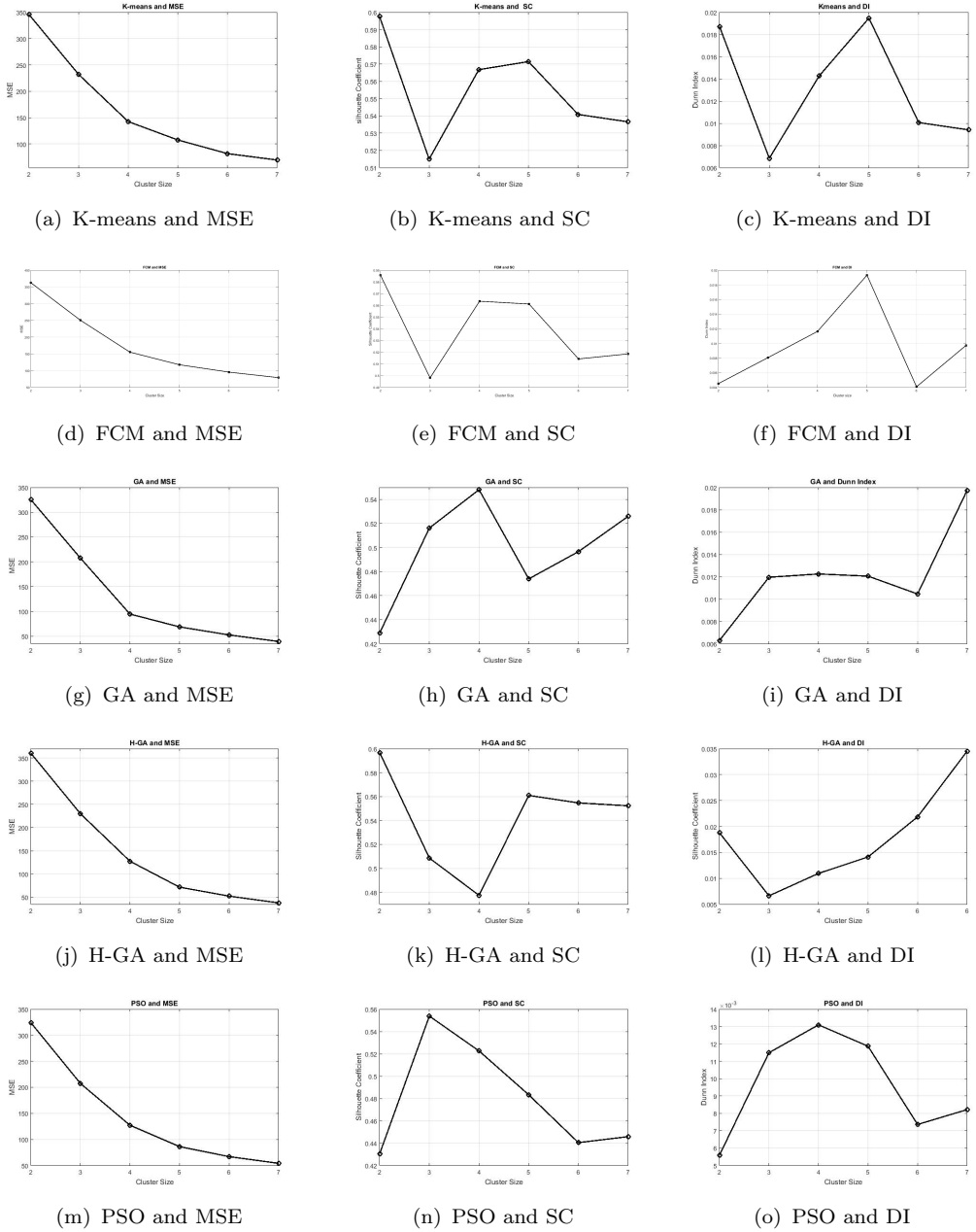


Figure 5.16: MSE Convergence plot for the different clustering approach

5.17(a), 5.17(b) and 5.17(c) are the plots of K-means for MSE, SC, and DI respectively. Figure 5.17(d),5.17(e) and 5.17(f) are the plots of FCM for MSE, SC, and DI respectively. Figure 5.17(g), 5.17(h) and 5.17(i) plots are for Genetic Algorithm (GA)based clustering with MSE, SC, and DI respectively. Figure 5.17(j), 5.17(k) and 5.17(l) are the plots of Hybrid GA clustering for MSE, SC, and DI respectively. Figure 5.17(m), 5.17(n) and 5.17(o) are the plots of Particle Swam Optimization based clustering for MSE, SC and DI respectively. For each approach, cluster size is decided by majority voting. Lower the value of the MSE, better is the quality of clustering, and larger is the SC and DI value, the better is the cluster quality. Optimal cluster number (gait profiles) is chosen by majority voting among these validity indices [210].

K-means, FCM and H-GA based clustering votes for 5 clusters as optimal choice, while GA and PSO based clustering indicate four as the optimal clustering size. Figure 5.17 indicates that cluster size five exhibits a better quality than other given k setting. Thus conclusively, to validate test subject profile, assessment is carried out by considering k as five in the following analysis.

Case 5: Evaluation of the test sample is necessary to demonstrate the significance of the surgery as considered in case 5. This section discusses the analysis of generated CP gait profile, considering the optimal cluster sizing as $k = 5$ from figure 5.17, on four test subjects (brief description is in table 5.6(b)) using affiliated



(a) K-means and MSE (b) K-means and SC (c) K-means and DI

(d) FCM and MSE (e) FCM and SC (f) FCM and DI

(g) GA and MSE (h) GA and SC (i) GA and DI

(j) H-GA and MSE (k) H-GA and SC (l) H-GA and DI

(m) PSO and MSE (n) PSO and SC (o) PSO and DI

Figure 5.17: Composite error and indices plot for case 4 in terms of MSE, SC and DI for K-means, GA, H-GA and PSO clustering in case of $2 \leq k \leq 7$.

probability index (API).

Figure 5.18 presents the affiliated probability distribution for two neurological intact subjects A, B and two patients C and D from pre and post surgery in case of five clusters. Extensive analysis shows that cluster 5 in the figure 5.18 is for the subjects with the no gait pathology, the control group, followed by Cluster 3, Cluster 1, Cluster 4 and Cluster 2. Cluster 4 and 2 can be classified as gait pattern with CP cases.

The result confirms the gait pattern of subject A as normal. The study shows that subject B exhibits deviated gait pattern. After intensive analysis, it is found that the mean age of control group is 7.09 (2-13 age range), while the test subject considered is of 19-year-old. Thus this gait pattern is misclassified as cluster 4 instead of cluster 5; the control cluster.

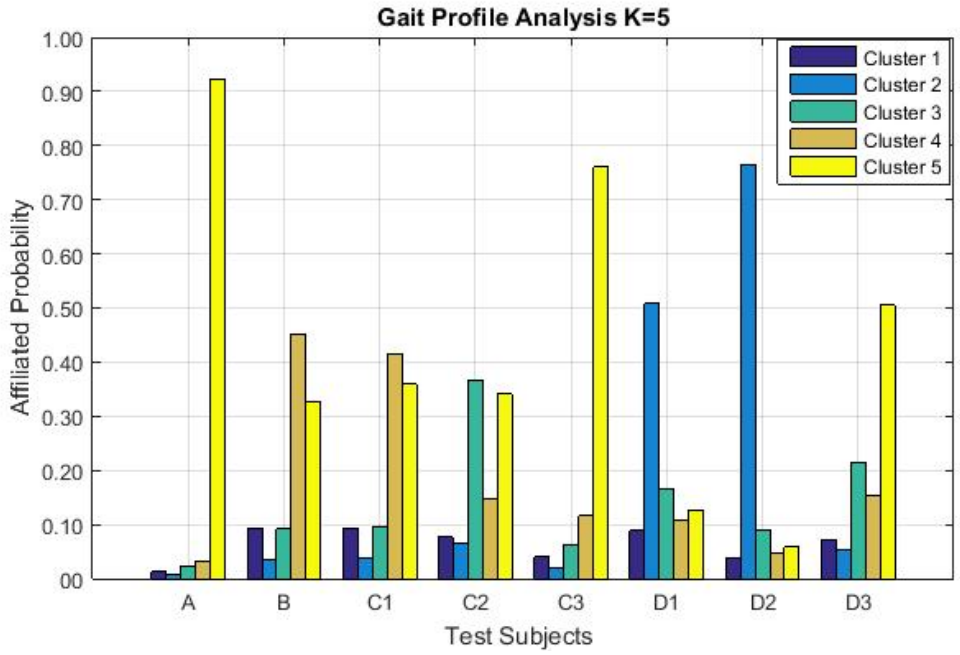


Figure 5.18: Gait profile plots for gait profile considering 5 clusters using affiliation probability

Three observations are taken for both patients C and D. First observation of the third subject(C1) is performed before surgery, affiliated probability based gait profile also confirms the CP based pattern. In the second (C2) and third (C3) observations of the subject, C shifted from cluster 4 to cluster 3 and then to cluster 5 (the second and third observation are post-surgery case). Gait profile based on affiliated probability index confirms that surgery helps the subject C to shift toward normal gait profile. Similar kind of observation is observed in patient D, where the first observation is before surgery, and D2 and D3 are post-surgery observations. The D1 and D2 in the plot exhibit that the problem is increased even after surgery. But after two years of surgery, observation D3 presents that the gait of the subject could be classified as normal gait profile.

Thus health-care professionals can take help of optimized clustering approaches to evaluate the recovery progress of patients after surgery. Doctors can monitor the recovery progress and if required change the strategy for treatment. But here

it is important to note that the gait profile from automated optimization based clustering is an indicator of the gait state, thus aid doctors in making decisions regarding rehabilitation.

5.3 Chapter Summary

The work discussed in this study is of implementing fuzzy based techniques for gait phase detection (FL-GPD) for the 2-D vision based system using the passive marker. The benefit is that these markers don't need high cost and excessive time in place. The subject's opinion on these marker enactments further simplified that the setup did not affect the gait performance. Findings of the experiments indicate that this technique was able to correctly segment the gait phases and can differentiate among the normal and healthy gait pattern. For all the healthy subjects considered in this research identification rate in the natural sequence is 100%. Unhealthy cases, are also identified using FL-GPD. This approach of gait phase detection has the potential to be used in rehabilitation activities and gait analysis.

A comparative study of clustering methods (kmeans, FCM, PSO, GA, H-GA and H-PSO) are explored for the gait analysis using CP data and show how different clustering approaches may be employed for CP gait identification. In this chapter, a novel nature-inspired meta-heuristic algorithm, Grey Wolf Optimized Clustering (GWOC) approach is proposed to cluster CP gait patterns using a publicly available CP gait dataset of 68 normal healthy and 88 with spastic diplegic form of CP. The performance of the proposed clustering methods is evaluated using quantization error, inter and intra distance, purity index, mean square error, and silhouette coefficient. The Proposed clustering technique outperforms the traditional clustering approaches.

The conclusion drawn from the thesis work is discussed in Chapter 6. It also presents the limitation and scope of the future work in the end.

Chapter 6

Conclusion

A lot of research is present in literature for vision-based gait analysis. Still, there is immense scope for improvement of the existing gait analysis models. In this chapter, the conclusion drawn from the current research work is presented along with the possible directions of future work.

6.1 Contribution

Two feature extraction approaches have been proposed for the gait analysis to find the hidden pattern (the pattern that cannot be captured through visual observation only) in human gait such as joint angles, stride length, step length, cadence, stance and swing phase, etc. We are motivated by the fact that gait pattern can be extracted either with minimal use of the marker or without using marker techniques. Following are the major contributions of this thesis.

1. Passive Marker-based Gait Parameter Extraction Approach (PM-GPA) is developed with the minimal use of markers, to extract the gait pattern/parameters (kinematic and spatiotemporal).

2. PM-GPA is used to retrieve the gait pattern for the 120 healthy subjects of MNIT Gait dataset, collected during the study. The result explicates that there is a significant difference between the Indian and western people gait pattern. Thus, the generated pattern can be used as Indian gait norms from the clinical perspective.

3. Another approach for the gait pattern extraction without placement of any marker on the subject's body is proposed as Marker-less Gait Parameter Extraction Approach (Ml-GPEA). The patterns and results of the proposed approach are promising, and it added one more step toward the clinical application without any marker.

4. Based on the gait parameters extracted using PM-GPA, two Gait Pattern Prediction (GPP) models namely, Individual Specific-GPP model from Historical Data (GPP-HD) and GPP Models from Anthropometric data (GPP-AD) are developed. GPP-HD model gives better results when the similarity between the past pattern in co-related. GPP-AD model uses the body parameters as input to predict the gait pattern.

5. To identify an abnormality in the gait pattern, Fuzzy-Logic based Gait Phase Detection (FL-GPD) is proposed. FL-GPD model identifies the healthy and unhealthy gait pattern in the dataset.

6. The patients gait profile cannot shift from abnormal to the normal one in one go. Thus, to identify the optimal number of gait profile in the dataset, hybrid optimization based clustering technique is proposed. It is validated with the unseen test subjects, who underwent surgery during the rehabilitation process. From the patients history, healthcare specialists can clinically correlate the significance of proposed gait profiling to treat the patients condition.

6.2 Limitation and Future work

Due to a large number of real-world applications, this field of gait analysis research has been attractive for the researchers, industrialists, and health-care professionals. Following are some of the limitations and future work:

1. It is observed that walking pattern is affected by a large number of extrinsic, intrinsic, physical, psychological and pathological factors. In this thesis work, it is assumed that the subjects are in a neutral emotional state. Thus, researchers can find the correlation between these influencing factors on normal walking.

2. A careful engineering and considerable domain expertise are required to design a feature extractor that transformed the raw data into a suitable internal representation or feature vector. Thus, there is a scope to explore feature

representation approach at multiple levels (e.g. using Deep learning) that automatically discover the best representation needed.

3. One can analysis nature inspired based optimized fuzzy clustering approach for treatment intervention for individuals with a disability. Selection of the optimal objective function can be further investigated for better cluster purity index.

4. There is an obvious opportunity to continue this research and to use these methods to study and design a new hydraulic or pneumatic knee prosthesis

PUBLICATIONS

Patent

P1. Rajesh Kumar, Chandra Prakash and Namita Mittal, CHAAL: The smart gait analyzer, Journal of the Indian Patent Office, 201711014157A, May 2017.

International Journals

J1. Chandra Prakash, Rajesh Kumar, Namita Mittal, Recent Developments in Human Gait Research: Parameters, Approaches, Applications, Machine Learning Techniques, Datasets and Challenges, *Artificial Intelligence REview (AIRE)*, Volume 49, Pages 1-40, 2018. [SCI Indexed; Impact Factor 2.627]

J2. Chandra Prakash, Sneha Choudhary, Rajesh Kumar, Namita Mittal. Identification of Joint Angles for Lower Extremity in Constraint-Free Environment, *IET Image Processing*, 2018 (Accepted, SCI Indexed, Impact Factor 1.401])

J3. Chandra Prakash, Rajesh Kumar, Namita Mittal. Optimized Clustering Techniques for Gait Profiling in Children with Cerebral Palsy for Rehabilitation, *The Computer Journal*, Oxford University Press,2018. [SCIE Indexed; Impact Factor 0.711]

J4. Chandra Prakash, Rajesh Kumar, Namita Mittal. Gray Wolf Optimized Clustering for Gait Pattern Analysis in Children with Cerebral Palsy, *Expert Systems with Applications* , Springer, 2018. (Under Review)

J5. Chandra Prakash, Rajesh Kumar, Namita Mittal. Vision based Identification of Joint Coordinates for Marker-less Gait Analysis. Procedia Computer Science, Volume 132, Pages.68-75, 2018. (Scopus index)

J6. Chandra Prakash, Kankia Gupta, Anshul Mittal, Rajesh Kumar, Vijay Laxmi, Passive Marker Based Optical System for Gait Kinematics for Lower Extremity, *Procedia Computer Science*, Volume 45, Pages 176-185, 2015. (Scopus index)

International Conferences

C1. Chandra Prakash, Uddeshya Mishra, Manas Jain, Rajesh Kumar and Namita Mittal. Automated Kinematic Analysis using Holistic based Human Gait Motion for Biomedical Applications, *8th International Conference on Cloud Computing, Data Science & Engineering. (Confluence-2018)*, pp. 700-706. IEEE, 2018.

C2. Chandra Prakash, A. Sujil, Rajesh Kumar, Namita Mittal. Linear prediction model for joint movement for lower extremity, *5th International Conference on Advanced Computing, Networking, and Informatics, (ICACNI 2017)*, NIT Goa, AISC-Springer.

C3. Chandra Prakash, Kankia Gupta, Rajesh Kumar, and Namita Mittal. Fuzzy Logic-Based Gait Phase Detection Using Passive Markers, Chapter in Proceedings of Fifth International Conference on Soft Computing for Problem Solving: SocProS 2015, Volume 1 , pp. 561-572. Springer Singapore, 2016 **(Best Paper)**

C4. Chandra Prakash, Gourav. Takhar, Namita Mittal, Rajesh Kumar. A Review on Background Subtraction Techniques and Applications, *Recent Advances and Innovations in Engineering (ICRAIE-2016)*, 23-25, December, 2016, IEEE.

C5. Chandra Prakash, Rajesh Kumar, Namita Mittal. Automated Detection of human Gait events from Conventional Videography, *International Conference on Emerging Trends in Communication Technologies (ICETCT 2016)*, pp. 1-4, Nov18-19, 2016, IEEE.

C6. Chandra Prakash, Anshul Mittal, Rajesh Kumar, and Namita Mittal. "Identification of spatio-temporal and kinematics parameters for 2-D optical gait analysis system using passive markers." *International Conference on Advances in Computer Engineering and Applications (ICACEA)*, pp. 143-149. IEEE, 2015. **(Best Paper)**

C7. Chandra Prakash, Anshul Mittal, Rajesh Kumar, and Namita Mittal. "Identification of gait parameters from silhouette images." *Eighth International Conference on Contemporary Computing (IC3)*, pp. 190-195. IEEE, 2015.

Bibliography

[1] M. W. Whittle, *Gait analysis: an introduction*. London: Butterworth-Heinemann, 2014.

[2] J. Nutt, C. Marsden, and P. Thompson, "Human walking and higher-level gait disorders, particularly in the elderly," *Neurology*, vol. 43, no. 2, pp. 268–268, 1993.

[3] A. Y. Johnson and A. F. Bobick, "A multi-view method for gait recognition using static body parameters," in *Audio-and Video-Based Biometric Person Authentication*, pp. 301–311, Springer, 2001.

[4] C. Prakash, K. Gupta, A. Mittal, R. Kumar, and V. Laxmi, "Passive marker based optical system for gait kinematics for lower extremity," *Procedia Computer Science*, vol. 45, pp. 176–185, 2015.

[5] T. H. D. Nguyen, T. C. T. Qui, K. Xu, A. D. Cheok, S. L. Teo, Z. Zhou, A. Mallawaarachchi, S. P. Lee, W. Liu, H. S. Teo, *et al.*, "Real-time 3d human capture system for mixed-reality art and entertainment," *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 6, pp. 706–721, 2005.

[6] M. Fuhrer and ESCAP, "Disability inclusive disaster risk reduction," *Planet @ Risk*, vol. 2, no. 3, 2014.

[7] C. Prakash, R. Kumar, and N. Mittal, "Recent developments in human gait research: Parameters, approaches, applications, machine learning techniques, datasets and challenges," *Artificial Intelligence Review*, vol. 2016, 2016. (in press).

[8] C. Prakash, R. Kumar, and N. Mittal, "Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges," *Artificial Intelligence Review*, pp. 1–40, 2016.

[9] M. J. M. Vasconcelos and J. M. R. Tavares, "Human motion analysis: methodologies and applications," *CMBBE 2008*, 2008.

[10] 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.

[11] J. Aggarwal and Q. Cai, "Human motion analysis: A review," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428 – 440, 1999.

[12] S. Frenkel-Toledo, N. Giladi, C. Peretz, T. Herman, L. Gruendlinger, and J. M. Hausdorff, "Effect of gait speed on gait rhythmicity in parkinson's disease: variability of stride time and swing time respond differently," *Journal of neuroengineering and rehabilitation*, vol. 2, no. 1, p. 1, 2005.

[13] M. Alaqtash, H. Yu, R. Brower, A. Abdelgawad, and T. Sarkodie-Gyan, "Application of wearable sensors for human gait analysis using fuzzy computational algorithm," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 6, pp. 1018–1025, 2011.

[14] S. Chen, J. Lach, B. Lo, and G.-Z. Yang, "Toward pervasive gait analysis with wearable sensors: A systematic review," *IEEE journal of biomedical and health informatics*, vol. 20, no. 6, pp. 1521–1537, 2016.

[15] C. Prakash, A. Mittal, R. Kumar, and N. Mittal, "Identification of spatio-temporal and kinematics parameters for 2-d optical gait analysis system using passive markers," in *2015 International Conference on Advances in Computer Engineering and Applications (ICACEA)*, pp. 143–149, IEEE, 2015.

[16] J. K. Aggarwal, Q. Cai, W. Liao, and B. Sabata, "Articulated and elastic non-rigid motion: A review," in *Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects, 1994.*, pp. 2–14, IEEE, 1994.

[17] J. Paul, "The history of musculoskeletal modelling in human gait," *Theoretical Issues in Ergonomics Science*, vol. 6, no. 3-4, pp. 217–224, 2005.

[18] J. Perry, J. R. Davids, *et al.*, "Gait analysis: normal and pathological function.," *Journal of Pediatric Orthopaedics*, vol. 12, no. 6, p. 815, 1992.

[19] J.-S. Wang, C.-W. Lin, Y.-T. C. Yang, and Y.-J. Ho, "Walking pattern classification and walking distance estimation algorithms using gait phase information," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 10, pp. 2884–2892, 2012.

[20] E. Muybridge, Mozley, Anita, and Ventura, *Muybridge's Complete Human and Animal Locomotion: All 781 Plates from the 1887 Animal Locomotion.* 1887.

[21] R. T. Collins, R. Gross, and J. Shi, "Silhouette-based human identification from body shape and gait," in *Proceedings. Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002.*, pp. 366–371, IEEE, 2002.

[22] D. H. Sutherland, "The evolution of clinical gait analysis: Part ii kinematics," *Gait & posture*, vol. 16, no. 2, pp. 159–179, 2002.

[23] D. H. Sutherland, "The evolution of clinical gait analysis part iii–kinetics and energy assessment," *Gait & Posture*, vol. 21, no. 4, pp. 447–461, 2005.

[24] A. Muro-de-la Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, "Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications," *Sensors*, vol. 14, no. 2, pp. 3362–3394, 2014.

[25] N. Kunju, N. Kumar, D. Pankaj, A. Dhawan, and A. Kumar, "Emg signal analysis for identifying walking patterns of normal healthy individuals," *Indian Journal of Biomechanics*, vol. 1, pp. 118–122, 2009.

[26] D. T. H. Lai, R. K. Begg, and M. Palaniswami, "Computational intelligence in gait research: A perspective on current applications and future challenges," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, pp. 687–702, Sept 2009.

[27] A. I. Mahyuddin, S. Mihradi, T. Dirgantara, M. Moeliono, and T. Prabowo, "Development of indonesian gait database using 2d optical motion analyzer system," *ASEAN Engineering Journal*, pp. 62–72, 2012.

[28] D. T. Lai, R. K. Begg, and M. Palaniswami, "Computational intelligence in gait research: a perspective on current applications and future challenges," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 687–702, 2009.

[29] S. Joshi, D. Shenoy, G. Vibhudendra Simha, P. Rrashmi, K. Venugopal, and L. Patnaik, "Classification of alzheimer's disease and parkinson's disease by using machine learning and neural network methods," in *Machine Learning and Computing (ICMLC), 2010 Second International Conference on*, pp. 218–222, IEEE, 2010.

[30] S. Mazilu, A. Calatroni, E. Gazit, D. Roggen, J. M. Hausdorff, and G. Tröster, *Feature Learning for Detection and Prediction of Freezing of Gait in Parkinson's Disease*, pp. 144–158. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.

[31] J. Zhang, T. E. Lockhart, and R. Soangra, "Classifying lower extremity muscle fatigue during walking using machine learning and inertial sensors," *Annals of biomedical engineering*, vol. 42, no. 3, pp. 600–612, 2014.

[32] R. K. Fukuchi, B. M. Eskofier, M. Duarte, and R. Ferber, "Support vector machines for detecting age-related changes in running kinematics," *Journal of Biomechanics*, vol. 44, no. 3, pp. 540–542, 2011.

[33] C. Prakash, A. Mittal, R. Kumar, and N. Mittal, "Identification of gait parameters from silhouette images," pp. 190–195, Aug 2015.

[34] T. K. Lee, M. Belkhatir, and S. Sanei, "A comprehensive review of past and present vision-based techniques for gait recognition," *Multimedia Tools and Applications*, vol. 72, no. 3, pp. 2833–2869, 2014.

[35] 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*, pp. 320–327, IEEE, 2010.

[36] R. Poppe, "Vision-based human motion analysis: An overview," *Computer vision and image understanding*, vol. 108, no. 1, pp. 4–18, 2007.

[37] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer vision and image understanding*, vol. 104, no. 2, pp. 90–126, 2006.

[38] 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.

[39] R. Morris and S. Lawson, "A review and evaluation of available gait analysis technologies, and their potential for the measurement of impact transmission," *Newcastle University*, 2010.

[40] J. Lu, G. Wang, and P. Moulin, "Human identity and gender recognition from gait sequences with arbitrary walking directions," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 1, pp. 51–61, 2014.

[41] W. Hu, T. Tan, L. Wang, and S. Maybank, "A survey on visual surveillance of object motion and behaviors," *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 34, no. 3, pp. 334–352, 2004.

[42] W. Tao, T. Liu, R. Zheng, and H. Feng, "Gait analysis using wearable sensors," *Sensors*, vol. 12, no. 2, pp. 2255–2283, 2012.

[43] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," *Pattern Recognition*, vol. 47, no. 1, pp. 228–237, 2014.

[44] D. H. Sutherland, "The evolution of clinical gait analysis part l: kinesiological emg," *Gait & posture*, vol. 14, no. 1, pp. 61–70, 2001.

[45] A. Mannini and A. M. Sabatini, "Machine learning methods for classifying human physical activity from on-body accelerometers," *Sensors*, vol. 10, no. 2, pp. 1154–1175, 2010.

[46] W. Schöllhorn, J. Jäger, and D. Janssen, *Artificial neural network models of sports motions*, pp. 50–64. London: Routledge, 2008.

[47] M. Nordin and V. H. Frankel, *Basic biomechanics of the musculoskeletal system*. Lippincott Williams & Wilkins, 2001.

[48] R. Baker, "Gait analysis methods in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 3, no. 1, p. 1, 2006.

[49] R. Begg and J. Kamruzzaman, "A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data," *Journal of biomechanics*, vol. 38, no. 3, pp. 401–408, 2005.

[50] M. R. Heinen and F. S. Osório, "Gait control generation for physically based simulated robots using genetic algorithms," in *Advances in Artificial Intelligence-IBERAMIA-SBIA 2006*, pp. 562–571, Springer, 2006.

[51] S. Prentice, A. Patla, and D. Stacey, "Artificial neural network model for the generation of muscle activation patterns for human locomotion," *Journal of electromyography and kinesiology*, vol. 11, no. 1, pp. 19–30, 2001.

[52] R. Begg and J. Kamruzzaman, "Neural networks for detection and classification of walking pattern changes due to ageing," *Australasian Physics & Engineering Sciences in Medicine*, vol. 29, no. 2, pp. 188–195, 2006.

[53] Y. Song, J. Zhang, L. Cao, and M. Sangeux, *On Discovering the Correlated Relationship between Static and Dynamic Data in Clinical Gait Analysis*, pp. 563–578. Berlin, Heidelberg: Springer Berlin Heidelberg.

[54] M. M. Skelly and H. J. Chizeck, "Real-time gait event detection for paraplegic fes walking," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on*, vol. 9, no. 1, pp. 59–68, 2001.

[55] D. T. Lai, P. Levinger, R. K. Begg, W. L. Gilleard, and M. Palaniswami, "Automatic recognition of gait patterns exhibiting patellofemoral pain syndrome using a support vector machine approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 810–817, 2009.

[56] O. Tirosh, R. Baker, and J. McGinley, "Gaitabase: Web-based repository system for gait analysis," *Computers in biology and medicine*, vol. 40, no. 2, pp. 201–207, 2010.

[57] O. Beauchet, G. Allali, G. Berrut, C. Hommet, V. Dubost, and F. Assal, "Gait analysis in demented subjects: Interests and perspectives," *Neuropsychiatric disease and treatment*, vol. 4, no. 1, p. 155, 2008.

[58] G. Owusu, "Ai and computer-based methods in performance evaluation of sporting feats: an overview," *Artificial Intelligence Review*, vol. 27, no. 1, pp. 57–70, 2007.

[59] B. De Silva, A. Natarajan, M. Motani, and K.-C. Chua, "A real-time exercise feedback utility with body sensor networks," in *Medical Devices and Biosensors, 2008. ISSS-MDBS 2008. 5th International Summer School and Symposium on*, pp. 49–52, IEEE, 2008.

[60] H. Novatchkov and A. Baca, "Artificial intelligence in sports on the example of weight training," *Journal of sports science & medicine*, vol. 12, no. 1, p. 27, 2013.

[61] R. Bartlett, "Artificial intelligence in sports biomechanics: New dawn or false hope?," *Journal of sports science & medicine*, vol. 5, no. 4, p. 474, 2006.

[62] C. Prakash, A. Mittal, S. Tripathi, R. Kumar, and N. Mittal, "A framework for human recognition using a multimodel gait analysis approach," in *International Conference on Computing, Communication and Automation (ICCCA2016)*, IEEE, April 2016.

[63] J. K. Aggarwal and S. Park, "Human motion: Modeling and recognition of actions and interactions," in *Proceedings. 2nd International Symposium on 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004.*, pp. 640–647, IEEE, 2004.

[64] S. Shirke, S. Pawar, and K. Shah, "Literature review: Model free human gait recognition," in *Communication Systems and Network Technologies (CSNT), 2014 Fourth International Conference on*, pp. 891–895, IEEE, 2014.

[65] Z. Zhang, M. Hu, and Y. Wang, "A survey of advances in biometric gait recognition," *Biometric Recognition*, pp. 150–158, 2011.

[66] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea, "Machine recognition of human activities: A survey," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 18, no. 11, pp. 1473–1488, 2008.

[67] V. Krüger, D. Kragic, A. Ude, and C. Geib, "The meaning of action: a review on action recognition and mapping," *Advanced Robotics*, vol. 21, no. 13, pp. 1473–1501, 2007.

[68] N. Neverova, C. Wolf, G. Lacey, L. Fridman, D. Chandra, B. Barbello, and G. Taylor, "Learning human identity from motion patterns," *arXiv preprint arXiv:1511.03908*, 2015.

[69] S. Jezernik, G. Colombo, and M. Morari, "Automatic gait-pattern adaptation for treadmill training with robotic orthosis Lokomat.," in *Congress of the Int. Society of Biomechanics*, (Zürich, Switzerland), p. 204, July 2001.

[70] H. Yano, S. Tamefusa, N. Tanaka, H. Saitou, and H. Iwata, "Gait rehabilitation for stair climbing with a locomotion interface," in *Rehabilitation Robotics, 2009. ICORR 2009. IEEE International Conference on*, pp. 218–223, IEEE, 2009.

[71] L. Sigal and M. J. Black, "Humaneva: Synchronized video and motion capture dataset for evaluation of articulated human motion," 2006.

[72] J.-G. Juang, "Fuzzy neural network approaches for robotic gait synthesis," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 30, no. 4, pp. 594–601, 2000.

[73] J.-J. Kim and J.-J. Lee, "Gait adaptation method of biped robot for various terrains using central pattern generator (cpg) and learning mechanism," in

Control, Automation and Systems, 2007. ICCAS'07. International Conference on, pp. 10–14, IEEE, 2007.

[74] D. Zhang, Y. Wang, Z. Zhang, and M. Hu, "Estimation of view angles for gait using a robust regression method," *Multimedia tools and applications*, vol. 65, no. 3, pp. 419–439, 2013.

[75] C.-y. Zhao, X.-g. Zhang, and Q. Guo, "The application of machine-learning on lower limb motion analysis in human exoskeleton system," in *International Conference on Social Robotics*, pp. 600–611, Springer, 2012.

[76] D. B. Popovic and M. B. Popovic, "Design of a control for a neural prosthesis for walking: Use of artificial neural networks," in *Neural Network Applications in Electrical Engineering, 2006. NEUREL 2006. 8th Seminar on*, pp. 121–128, IEEE, 2006.

[77] P. Wawrzyński, "Autonomous reinforcement learning with experience replay for humanoid gait optimization," *Procedia Computer Science*, vol. 13, pp. 205–211, 2012.

[78] W. Liu, O. N. N. Fernando, A. D. Cheok, J. P. Wijesena, and R. T. Tan, "Science museum mixed reality digital media exhibitions for children," in *Digital Media and its Application in Museum & Heritages, Second Workshop on*, pp. 389–394, IEEE, 2007.

[79] S.-R. Ke, H. L. U. Thuc, Y.-J. Lee, J.-N. Hwang, J.-H. Yoo, and K.-H. Choi, "A review on video-based human activity recognition," *Computers*, vol. 2, no. 2, pp. 88–131, 2013.

[80] N. Savva, A. Scarinzi, and N. Bianchi-Berthouze, "Continuous recognition of player's affective body expression as dynamic quality of aesthetic experience," *Computational Intelligence and AI in Games, IEEE Transactions on*, vol. 4, no. 3, pp. 199–212, 2012.

[81] B. L. Webber, C. B. Phillips, and N. I. Badler, "Simulating humans: Computer graphics, animation, and control," *Center for Human Modeling and Simulation*, p. 68, 2015.

[82] D. Mavrikios, V. Karabatsou, K. Alexopoulos, M. Pappas, P. Gogos, and G. Chryssolouris, "An approach to human motion analysis and modelling," *International Journal of Industrial Ergonomics*, vol. 36, no. 11, pp. 979–989, 2006.

[83] C. D. Untaroiu, M. U. Meissner, J. R. Crandall, Y. Takahashi, M. Okamoto, and O. Ito, "Crash reconstruction of pedestrian accidents using optimization techniques," *International Journal of Impact Engineering*, vol. 36, no. 2, pp. 210–219, 2009.

[84] T. Lan, L. Chen, Z. Deng, G.-T. Zhou, and G. Mori, "Learning action primitives for multi-level video event understanding.," pp. 95–110, Springer International Publishing, 2015.

[85] F. Multon, R. Kulpa, L. Hoyet, and T. Komura, "Interactive animation of virtual humans based on motion capture data," *Computer Animation and Virtual Worlds*, vol. 20, no. 5-6, pp. 491–500, 2009.

[86] R. Riemer and A. Shapiro, "Biomechanical energy harvesting from human motion: theory, state of the art, design guidelines, and future directions," *Journal of neuroengineering and rehabilitation*, vol. 8, no. 1, p. 22, 2011.

[87] T. Takahashi, K. Ishida, D. Hirose, Y. Nagano, K. Okumiya, M. Nishinaga, Y. Doi, and H. Yamamoto, "Vertical ground reaction force shape is associated with gait parameters, timed up and go, and functional reach in elderly females," *Journal of rehabilitation medicine*, vol. 36, no. 1, pp. 42–45, 2004.

[88] B. A. Jones and I. D. Walker, "Kinematics for multisection continuum robots," *IEEE Transactions on Robotics*, vol. 22, no. 1, pp. 43–55, 2006.

[89] A. Phinyomark, B. A. Hettinga, S. T. Osis, and R. Ferber, "Gender and age-related differences in bilateral lower extremity mechanics during treadmill running," *PloS one*, vol. 9, no. 8, p. e105246, 2014.

[90] N. V. Boulgouris and Z. X. Chi, "Gait recognition using radon transform and linear discriminant analysis," *IEEE Transactions on Image Processing,*, vol. 16, no. 3, pp. 731–740, 2007.

[91] I. Gaba and S. P. Ahuja, "Gait analysis for identification by using bpnn with lda and mda techniques," in *MOOC, Innovation and Technology in Education (MITE), 2014 IEEE International Conference on*, pp. 122–127, IEEE, 2014.

[92] J. Wright and I. Jordanov, "Intelligent approaches in locomotion-a review," *Journal of Intelligent & Robotic Systems*, vol. 80, no. 2, pp. 255–277, 2015.

[93] S. Choi, I.-H. Youn, R. LeMay, S. Burns, and J.-H. Youn, "Biometric gait recognition based on wireless acceleration sensor using k-nearest neighbor

classification," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*, pp. 1091–1095, IEEE, 2014.

[94] B. T. Nukala, N. Shibuya, A. Rodriguez, J. Tsay, J. Lopez, T. Nguyen, S. Zupancic, and D. Y.-C. Lie, "An efficient and robust fall detection system using wireless gait analysis sensor with artificial neural network (ann) and support vector machine (svm) algorithms," *Open Journal of Applied Biosensor*, vol. 3, no. 04, p. 29, 2015.

[95] J. A. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural processing letters*, vol. 9, no. 3, pp. 293–300, 1999.

[96] H. Zheng, M. Yang, H. Wang, and S. McClean, "Machine learning and statistical approaches to support the discrimination of neuro-degenerative diseases based on gait analysis," in *Intelligent Patient Management*, pp. 57–70, Springer, 2009.

[97] B. Toro, C. J. Nester, and P. C. Farren, "Cluster analysis for the extraction of sagittal gait patterns in children with cerebral palsy," *Gait & posture*, vol. 25, no. 2, pp. 157–165, 2007.

[98] V. Cimolin and M. Galli, "Summary measures for clinical gait analysis: a literature review," *Gait & posture*, vol. 39, no. 4, pp. 1005–1010, 2014.

[99] M. J. O'Malley, M. F. Abel, D. L. Damiano, and C. L. Vaughan, "Fuzzy clustering of children with cerebral palsy based on temporal-distance gait parameters," *IEEE transactions on rehabilitation engineering*, vol. 5, no. 4, pp. 300–309, 1997.

[100] G. Xu, Y. Zhang, and R. Begg, "Mining gait pattern for clinical locomotion diagnosis based on clustering techniques," in *Advanced Data Mining and Applications*, pp. 296–307, Springer, 2006.

[101] A. Phinyomark, S. Osis, B. A. Hettinga, and R. Ferber, "Kinematic gait patterns in healthy runners: A hierarchical cluster analysis," *Journal of biomechanics*, vol. 48, no. 14, pp. 3897–3904, 2015.

[102] J. Wu and B. Wu, "The novel quantitative technique for assessment of gait symmetry using advanced statistical learning algorithm," *BioMed research international*, vol. 2015, 2015.

[103] S. M. Van Rooden, W. J. Heiser, J. N. Kok, D. Verbaan, J. J. van Hilten, and J. Marinus, "The identification of parkinson's disease subtypes using

cluster analysis: a systematic review," *Movement Disorders*, vol. 25, no. 8, pp. 969–978, 2010.

[104] M. Zeng, L. T. Nguyen, B. Yu, O. J. Mengshoel, J. Zhu, P. Wu, and J. Zhang, "Convolutional neural networks for human activity recognition using mobile sensors," in *Mobile Computing, Applications and Services (MobiCASE), 2014 6th International Conference on*, pp. 197–205, IEEE, 2014.

[105] Z. Deng, M. Zhai, L. Chen, Y. Liu, S. Muralidharan, M. J. Roshtkhari, and G. Mori, "Deep structured models for group activity recognition," *arXiv preprint arXiv:1506.04191*, 2015.

[106] C. J. Hasson, J. Manczurowsky, and Yen, "A reinforcement learning approach to gait training improves retention," *Frontiers in human neuroscience*, vol. 9, 2015.

[107] H. Yu, M. Alaqtash, E. Spier, and T. Sarkodie-Gyan, "Analysis of muscle activity during gait cycle using fuzzy rule-based reasoning," *Measurement*, vol. 43, no. 9, pp. 1106–1114, 2010.

[108] C. Prakash, K. Gupta, R. Kumar, and N. Mittal, "Fuzzy logic-based gait phase detection using passive markers," in *Proceedings of Fifth International Conference on Soft Computing for Problem Solving*, pp. 561–572, Springer, 2016.

[109] F. Tafazzoli, G. Bebis, S. Louis, and M. Hussain, "Improving human gait recognition using feature selection," in *International Symposium on Visual Computing*, pp. 830–840, Springer, 2014.

[110] E. Yeguas-Bolivar, R. Muoz-Salinas, R. Medina-Carnicer, and A. Carmona-Poyato, "Comparing evolutionary algorithms and particle filters for markerless human motion capture," *Applied Soft Computing*, vol. 17, pp. 153–166, 2014.

[111] Y. Yun, H.-C. Kim, S. Y. Shin, J. Lee, A. D. Deshpande, and C. Kim, "Statistical method for prediction of gait kinematics with gaussian process regression," *Journal of Biomechanics*, vol. 47, no. 1, pp. 186–192, 2014.

[112] Z. Zhang, H. S. Seah, and C. K. Quah, "Particle swarm optimization for markerless full body motion capture," in *Handbook of Swarm Intelligence*, pp. 201–220, Springer, 2011.

[113] R. Prasad, S. Babu, N. Siddaiah, and K. Rao, "A review on techniques for diagnosing and monitoring patients with parkinsons disease," *J Biosens Bioelectron*, vol. 7, no. 203, p. 2, 2016.

[114] G. Barton, P. Lisboa, A. Lees, and S. Attfield, "Gait quality assessment using self-organising artificial neural networks," *Gait & posture*, vol. 25, no. 3, pp. 374–379, 2007.

[115] D. Gowsikhaa, S. Abirami, and R. Baskaran, "Automated human behavior analysis from surveillance videos: a survey," *Artificial Intelligence Review*, vol. 42, no. 4, pp. 747–765, 2014.

[116] P. Levinger, D. T. Lai, R. K. Begg, K. E. Webster, and J. A. Feller, "The application of support vector machines for detecting recovery from knee replacement surgery using spatio-temporal gait parameters," *Gait & posture*, vol. 29, no. 1, pp. 91–96, 2009.

[117] X. Chen, Y. Zhao, Y.-Q. Zhang, and R. Harrison, *Combining SVM Classifiers Using Genetic Fuzzy Systems Based on AUC for Gene Expression Data Analysis*, pp. 496–505. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.

[118] I. Rida, A. Bouridane, G. L. Marcialis, and P. Tuveri, "Improved human gait recognition," in *International Conference on Image Analysis and Processing*, pp. 119–129, Springer, 2015.

[119] M. Hofmann, S. Sural, and G. Rigoll, "Multimodal biometric gait database: A comparison study," pp. 71–82, 2014.

[120] D. Janssen, W. I. Schöllhorn, J. Lubienetzki, K. Fölling, H. Kokenge, and K. Davids, "Recognition of emotions in gait patterns by means of artificial neural nets," *Journal of Nonverbal Behavior*, vol. 32, no. 2, pp. 79–92, 2008.

[121] J. M. Hausdorff, S. L. Mitchell, R. Firtion, C.-K. Peng, M. E. Cudkowicz, J. Y. Wei, and A. L. Goldberger, "Altered fractal dynamics of gait: reduced stride-interval correlations with aging and huntingtons disease," *Journal of applied physiology*, vol. 82, no. 1, pp. 262–269, 1997.

[122] T. P. Andriacchi and E. J. Alexander, "Studies of human locomotion past, present and future," *Journal of biomechanics*, vol. 33, no. 10, pp. 1217–1224, 2000.

[123] M. P. Kadaba, H. Ramakrishnan, M. Wootten, *et al.*, "Measurement of lower extremity kinematics during level walking," *Journal of orthopaedic research*, vol. 8, no. 3, pp. 383–392, 1990.

[124] K. C. Moisio, D. R. Sumner, S. Shott, and D. E. Hurwitz, "Normalization of joint moments during gait: a comparison of two techniques," *Journal of biomechanics*, vol. 36, no. 4, pp. 599–603, 2003.

[125] B. Auvinet, G. Berrut, C. Touzard, L. Moutel, N. Collet, D. Chaleil, and E. Barrey, "Reference data for normal subjects obtained with an accelerometric device," *Gait & posture*, vol. 16, no. 2, pp. 124–134, 2002.

[126] T. Ryu, H. S. Choi, H. Choi, and M. K. Chung, "A comparison of gait characteristics between korean and western people for establishing korean gait reference data," *International journal of industrial ergonomics*, vol. 36, no. 12, pp. 1023–1030, 2006.

[127] R. Gross and J. Shi, "The cmu motion of body (mobo) database," tech. rep., Robotics Institute, Pittsburgh, PA, June 2001.

[128] N. J. S. M, C. J, and G. M, "Experimental plan for automatic gait recognition," tech. rep.

[129] 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.

[130] C. for biometrics and security research, "Casia gait dataset," 2016.

[131] Y. Makihara, H. Mannami, A. Tsuji, M. A. Hossain, K. Sugiura, A. Mori, and Y. Yagi, "The ou-isir gait database comprising the treadmill dataset," *IPSJ Transactions on Computer Vision and Applications*, vol. 4, no. 0, pp. 53–62, 2012.

[132] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi, "The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1511–1521, 2012.

[133] A. Kale, N. Cuntoor, and R. Chellappa, "A framework for activity-specific human identification," in *Acoustics, Speech, and Signal Processing*

(ICASSP), 2002 IEEE International Conference on, vol. 4, pp. IV–3660, IEEE, 2002.

[134] N. Cuntoor, A. Kale, and R. Chellappa, "Combining multiple evidences for gait recognition," in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03)*, vol. 3, pp. III–33, IEEE, 2003.

[135] W. M. Association *et al.*, "Declaration of helsinki (human experimentation: Code of ethics)," *BMJ*, vol. 2, p. 177, 1964.

[136] R. B. Davis III, S. Ounpuu, D. Tyburski, and J. R. Gage, "A gait analysis data collection and reduction technique," *Human movement science*, vol. 10, no. 5, pp. 575–587, 1991.

[137] D. Tabakin, *A comparison of 3D gait models based on the Helen Hayes Hospital marker set.* PhD thesis, University of Cape Town, 2000.

[138] M. P. Kadaba, H. Ramakrishnan, and M. Wootten, "Measurement of lower extremity kinematics during level walking," *Journal of orthopaedic research*, vol. 8, no. 3, pp. 383–392, 1990.

[139] B. Smith, K. M. Ashton, D. Bohl, R. C. Clark, J. B. Metheny, and S. Klassen, "Influence of carrying a backpack on pelvic tilt, rotation, and obliquity in female college students," *Gait & posture*, vol. 23, no. 3, pp. 263–267, 2006.

[140] R. Baker, "Gait analysis methods in rehabilitation," *Journal of neuroengineering and rehabilitation*, vol. 3, no. 1, p. 4, 2006.

[141] S. Ounpuu, J. Gage, and R. Davis, "Three-dimensional lower extremity joint kinetics in normal pediatric gait.," *Journal of pediatric orthopedics*, vol. 11, no. 3, pp. 341–349, 1991.

[142] "Lab manual; lab 3 introduction to angular kinematics," 2013.

[143] T. Lindeberg, *Discrete scale-space theory and the scale-space primal sketch.* PhD thesis, KTH Royal Institute of Technology, 1991.

[144] F. Leymarie and M. D. Levine, "Simulating the grassfire transform using an active contour model," *IEEE transactions on pattern analysis and machine intelligence*, vol. 14, no. 1, pp. 56–75, 1992.

[145] A. M. A.S. Jaya, T. Dirgantara and S. Mihradi, "Robust algorithms of marker image processing in automatic human gait analysis," in *Regional Conference on Mechanical and Aerospace Technology, Bali, Indonesia*, pp. 33–39, 2010.

[146] D. Gouwanda and A. A. Gopalai, "A robust real-time gait event detection using wireless gyroscope and its application on normal and altered gaits," *Medical engineering & physics*, vol. 37, no. 2, pp. 219–225, 2015.

[147] K. Tong and M. H. Granat, "A practical gait analysis system using gyroscopes," *Medical engineering & physics*, vol. 21, no. 2, pp. 87–94, 1999.

[148] S. L. Phung, A. Bouzerdoum, and D. Chai, "Skin segmentation using color pixel classification: analysis and comparison," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 1, pp. 148–154, 2005.

[149] G. Takhar, C. Prakash, N. Mittal, and R. Kumar, "Vision based gender recognition using hybrid background subtraction technique," in *International Conference on Next Generation Computing Technologies*, pp. 1–6.

[150] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: An overview," *Computer Science Review*, vol. 11, pp. 31–66, 2014.

[151] G. Takhar, C. Prakash, N. Mittal, and R. Kumar, "Comparative analysis of background subtraction techniques and applications," in *Recent Advances and Innovations in Engineering (ICRAIE), 2016 International Conference on*, pp. 1–8, IEEE, 2016.

[152] P. KaewTraKulPong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," *Video-based surveillance systems*, vol. 1, pp. 135–144, 2002.

[153] Z. Zivkovic and F. Van Der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," *Pattern recognition letters*, vol. 27, no. 7, pp. 773–780, 2006.

[154] W. T. Dempster and G. R. Gaughran, "Properties of body segments based on size and weight," *Developmental Dynamics*, vol. 120, no. 1, pp. 33–54, 1967.

[155] R. Drillis, R. Contini, and M. Bluestein, *Body segment parameters*. New York University, School of Engineering and Science Research Division, NY, 1966.

[156] J.-H. Yoo and M. S. Nixon, "Automated markerless analysis of human gait motion for recognition and classification," *Etri Journal*, vol. 33, no. 2, pp. 259–266, 2011.

[157] S. L. Dockstader, K. A. Bergkessel, and A. M. Tekalp, "Feature extraction for the analysis of gait and human motion," in *Proceedings. 16th International Conference on Pattern Recognition, 2002.*, vol. 1, pp. 5–8, IEEE, 2002.

[158] A. Castelli, G. Paolini, A. Cereatti, and U. Della Croce, "A 2d markerless gait analysis methodology: validation on healthy subjects," *Computational and mathematical methods in medicine*, vol. 2015, 2015.

[159] P. D. UNDESA, "World population prospects: The 2017 revision."

[160] P. Kutilek and J. Hozman, "Prediction of lower extremities movement using characteristics of angle-angle diagrams and artificial intelligence," in *E-Health and Bioengineering Conference (EHB), 2011*, pp. 1–4, IEEE, 2011.

[161] S. Han, G. Cheng, and P. Xu, "Three-dimensional lower extremity kinematics of chinese during activities of daily living," *Journal of back and musculoskeletal rehabilitation*, vol. 28, no. 2, pp. 327–334, 2015.

[162] L. S. Kao and C. E. Green, "Analysis of variance: is there a difference in means and what does it mean?," *Journal of Surgical Research*, vol. 144, no. 1, pp. 158–170, 2008.

[163] S. E. Lloyd, "Kinematic analysis of hip and knee joints between barefoot and shod treadmill running," 2013.

[164] J. Sun, *Dynamic modeling of human gait using a model predictive control approach*. PhD thesis, Marquette University, 2015.

[165] P. Beyl, M. Van Damme, R. Van Ham, B. Vanderborght, and D. Lefeber, "Design and control of a lower limb exoskeleton for robot-assisted gait training," *Applied Bionics and Biomechanics*, vol. 6, no. 2, pp. 229–243, 2009.

[166] S. Jiang, B. Zhang, and D. Wei, "The elderly fall risk assessment and prediction based on gait analysis," in *Computer and Information Technology (CIT), 2011 IEEE 11th International Conference on*, pp. 176–180, IEEE, 2011.

[167] D. Aoyagi, W. E. Ichinose, S. J. Harkema, D. J. Reinkensmeyer, and J. E. Bobrow, "A robot and control algorithm that can synchronously assist in naturalistic motion during body-weight-supported gait training following neurologic injury," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 15, no. 3, pp. 387–400, 2007.

[168] E. Aertbeliën and J. De Schutter, "Learning a predictive model of human gait for the control of a lower-limb exoskeleton," in *Biomedical Robotics and Biomechatronics (2014 5th IEEE RAS & EMBS International Conference on*, pp. 520–525, IEEE, 2014.

[169] L. Ren, R. K. Jones, and D. Howard, "Predictive modelling of human walking over a complete gait cycle," *Journal of biomechanics*, vol. 40, no. 7, pp. 1567–1574, 2007.

[170] I. DE INGENIEURSWETENSCHAPPEN, "Design and control of a knee exoskeleton powered by pleated pneumatic artificial muscles for robot-assisted gait rehabilitation,"

[171] M. Ackermann and A. J. van den Bogert, "Optimality principles for model-based prediction of human gait," *Journal of biomechanics*, vol. 43, no. 6, pp. 1055–1060, 2010.

[172] P. Kutilek and S. Viteckova, "Prediction of lower extremity movement by cyclograms," *Acta Polytechnica*, vol. 52, no. 1, p. 51, 2012.

[173] S. L. Nogueira, R. S. Inoue, M. H. Terra, and A. A. Siqueira, "Estimation of lower limbs angular positions using kalman filter and genetic algorithm," in *Biosignals and Biorobotics Conference (BRC), 2013 ISSNIP*, pp. 1–6, IEEE, 2013.

[174] S. Viteckova, P. Kutilek, Z. Svoboda, and M. Jirina, "Fuzzy inference system for lower limbs angles prediction," in *Telecommunications and Signal Processing (TSP), 2012 35th International Conference on*, pp. 517–520, IEEE, 2012.

[175] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *nature*, vol. 323, no. 6088, p. 533, 1986.

[176] H. Drucker, C. J. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support vector regression machines," in *Advances in neural information processing systems*, pp. 155–161, 1997.

[177] R. K. Begg, M. Palaniswami, and B. Owen, "Support vector machines for automated gait classification," *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 5, pp. 828–838, 2005.

[178] D. C. Kerrigan, M. K. Todd, and U. C. Della, "Gender differences in joint biomechanics during walking: normative study in young adults.," *American journal of physical medicine & rehabilitation*, vol. 77, no. 1, pp. 2–7, 1998.

[179] A. Escalante, M. J. Lichtenstein, and H. P. Hazuda, "Walking velocity in aged persons: its association with lower extremity joint range of motion," *Arthritis Care & Research*, vol. 45, no. 3, pp. 287–294, 2001.

[180] S. Hussain, P. K. Jamwal, M. H. Ghayesh, and S. Q. Xie, "Assist-as-needed control of an intrinsically compliant robotic gait training orthosis," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 2, pp. 1675–1685, 2017.

[181] M. Samson, A. Crowe, P. De Vreede, J. Dessens, S. Duursma, and H. Verhaar, "Differences in gait parameters at a preferred walking speed in healthy subjects due to age, height and body weight," *Aging Clinical and Experimental Research*, vol. 13, no. 1, pp. 16–21, 2001.

[182] C. R. Guerrero, V. Grosu, S. Grosu, A. Leu, D. Ristic-Durrant, B. Vanderborght, and D. Lefeber, "Torque control of a push-pull cable driven powered orthosis for the corbys platform," in *Rehabilitation Robotics (ICORR), 2015 IEEE International Conference on*, pp. 25–30, IEEE, 2015.

[183] D.-X. Liu, X. Wu, W. Du, C. Wang, C. Chen, and T. Xu, "Deep spatial-temporal model for rehabilitation gait: optimal trajectory generation for knee joint of lower-limb exoskeleton," *Assembly Automation*, vol. 37, no. 3, pp. 369–378, 2017.

[184] S. Feng, J. Lin, Z. Huang, G. Chen, W. Chen, Y. Wang, R. Chen, and H. Zeng, "Esophageal cancer detection based on tissue surface-enhanced raman spectroscopy and multivariate analysis," *Applied Physics Letters*, vol. 102, no. 4, p. 043702, 2013.

[185] T. Pfingsten, M. Kuss, and C. E. Rasmussen, "Nonstationary gaussian process regression using a latent extension of the input space," 2006.

[186] M. W. Whittle, "Clinical gait analysis: A review," *Human Movement Science*, vol. 15, no. 3, pp. 369–387, 1996.

[187] R. C. Wagenaar and R. E. van Emmerik, "Dynamics of pathological gait," *Human Movement Science*, vol. 13, no. 3-4, pp. 441–471, 1994.

[188] I. P. Pappas, M. R. Popovic, T. Keller, V. Dietz, and M. Morari, "A reliable gait phase detection system," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 9, no. 2, pp. 113–125, 2001.

[189] J. F. Knight, H. W. Bristow, S. Anastopoulou, C. Baber, A. Schwirtz, and T. N. Arvanitis, "Uses of accelerometer data collected from a wearable system," *Personal and Ubiquitous Computing*, vol. 11, no. 2, pp. 117–132, 2007.

[190] D. Gouwanda and S. Senanayake, "Emerging trends of body-mounted sensors in sports and human gait analysis," in *4th Kuala Lumpur International Conference on Biomedical Engineering 2008*, pp. 715–718, Springer, 2008.

[191] C. M. Senanayake and S. A. Senanayake, "Evaluation of gait parameters for gait phase detection during walking," in *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2010 IEEE Conference on*, pp. 127–132, IEEE, 2010.

[192] L. A. Zadeh, "Fuzzy sets," in *Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh*, pp. 394–432, World Scientific, 1996.

[193] M. I. Chacon-Murguia, O. Arias-Enriquez, and R. Sandoval-Rodriguez, "A fuzzy scheme for gait cycle phase detection oriented to medical diagnosis," in *Mexican Conference on Pattern Recognition*, pp. 20–29, Springer, 2013.

[194] E. H. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International journal of man-machine studies*, vol. 7, no. 1, pp. 1–13, 1975.

[195] C. Prakash, R. Kumar, and N. Mittal, "Vision based gait analysis techniques in elderly life -towards a better life," tech. rep., CSI Communications, Dec 2015.

[196] S. E. Crouter, E. Kuffel, J. D. Haas, E. A. Frongillo, and D. R. Bassett Jr, "A refined 2-regression model for the actigraph accelerometer," *Medicine and science in sports and exercise*, vol. 42, no. 5, p. 1029, 2010.

[197] M. Oskoui, F. Coutinho, J. Dykeman, N. Jetté, and T. Pringsheim, "An update on the prevalence of cerebral palsy: a systematic review and meta-analysis," *Developmental Medicine & Child Neurology*, vol. 55, no. 6, pp. 509–519, 2013.

[198] P. Rosenbaum *et al.*, "A report: the definition and classification of cerebral palsy april 2006," 2007.

[199] J. R. Gage, M. H. Schwartz, S. E. Koop, and T. F. Novacheck, *The identification and treatment of gait problems in cerebral palsy*, vol. 4. John Wiley & Sons, 2009.

[200] C. L. Arneson, M. S. Durkin, R. E. Benedict, R. S. Kirby, M. Yeargin-Allsopp, K. V. N. Braun, and N. S. Doernberg, "Prevalence of cerebral palsy: autism and developmental disabilities monitoring network, three sites, united states, 2004," *Disability and health journal*, vol. 2, no. 1, pp. 45–48, 2009.

[201] A. G. Vyas, V. K. Kori, S. Rajagopala, and K. S. Patel, "Etiopathological study on cerebral palsy and its management by shashtika shali pinda sweda and samvardhana ghrita," *Ayu*, vol. 34, no. 1, p. 56, 2013.

[202] R. Palisano, P. Rosenbaum, S. Walter, D. Russell, E. Wood, and B. Galuppi, "Development and reliability of a system to classify gross motor function in children with cerebral palsy," *Developmental Medicine & Child Neurology*, vol. 39, no. 4, pp. 214–223, 1997.

[203] F. Dobson, M. E. Morris, R. Baker, and H. K. Graham, "Gait classification in children with cerebral palsy: a systematic review," *Gait & posture*, vol. 25, no. 1, pp. 140–152, 2007.

[204] G. Cola, M. Avvenuti, and A. Vecchio, "Real-time identification using gait pattern analysis on a standalone wearable accelerometer," *The Computer Journal*, vol. 60, 2017.

[205] R. E. Cook, I. Schneider, M. E. Hazlewood, S. J. Hillman, and J. E. Robb, "Gait analysis alters decision-making in cerebral palsy," *Journal of pediatric orthopaedics*, vol. 23, no. 3, pp. 292–295, 2003.

[206] B. Barshan and M. C. Yüksek, "Recognizing daily and sports activities in two open source machine learning environments using body-worn sensor units," *The Computer Journal*, vol. 57, no. 11, pp. 1649–1667, 2014.

[207] Y. Chai, J. Ren, H. Zhao, Y. Li, J. Ren, and P. Murray, "Hierarchical and multi-featured fusion for effective gait recognition under variable scenarios," *Pattern Analysis and Applications*, vol. 2015, pp. 1–13, 2015.

[208] B.-l. Zhang and Y. Zhang, "Classification of cerebral palsy gait by kernel fisher discriminant analysis," *International Journal of Hybrid Intelligent Systems*, vol. 5, no. 4, pp. 209–218, 2008.

[209] M. A. Wong, S. Simon, and R. A. Olshen, "Statistical analysis of gait patterns of persons with cerebral palsy," *Statistics in medicine*, vol. 2, no. 3, pp. 345–354, 1983.

[210] S. J. Nanda and G. Panda, "A survey on nature inspired metaheuristic algorithms for partitional clustering," *Swarm and Evolutionary computation*, vol. 16, pp. 1–18, 2014.

[211] H. P. Lai, M. Visani, A. Boucher, and J.-M. Ogier, "An experimental comparison of clustering methods for content-based indexing of large image databases," *Pattern Analysis and Applications*, vol. 15, no. 4, pp. 345–366, 2012.

[212] A. Carriero, A. Zavatsky, J. Stebbins, T. Theologis, and S. J. Shefelbine, "Determination of gait patterns in children with spastic diplegic cerebral palsy using principal components," *Gait & posture*, vol. 29, no. 1, pp. 71–75, 2009.

[213] B. Toro, C. J. Nester, and P. C. Farren, "Cluster analysis for the extraction of sagittal gait patterns in children with cerebral palsy," *Gait & posture*, vol. 25, no. 2, pp. 157–165, 2007.

[214] O. Maimon and L. Rokach, "Decomposition methodology for knowledge discovery and data mining," *Data mining and knowledge discovery handbook*, pp. 981–1003, 2005.

[215] L. A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy sets and systems*, vol. 90, no. 2, pp. 111–127, 1997.

[216] M. G. Omran, *Particle swarm optimization methods for pattern recognition and image processing*. PhD thesis, University of Pretoria, 2005.

[217] D. E. Goldberg and J. H. Holland, "Genetic algorithms and machine learning," *Machine learning*, vol. 3, no. 2, pp. 95–99, 1988.

[218] J. C. Bezdek, S. Boggavarapu, L. O. Hall, and A. Bensaid, "Genetic algorithm guided clustering," in *Proceedings of the First IEEE Conference on*

Evolutionary Computation, 1994. IEEE World Congress on Computational Intelligence.,, pp. 34–39, IEEE, 1994.

[219] K. Krishna and M. N. Murty, "Genetic k-means algorithm," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 29, no. 3, pp. 433–439, 1999.

[220] C. Ozturk, E. Hancer, and D. Karaboga, "Improved clustering criterion for image clustering with artificial bee colony algorithm," *Pattern Analysis and Applications*, vol. 18, no. 3, pp. 587–599, 2015.

[221] R. C. Eberhart, J. Kennedy, *et al.*, "A new optimizer using particle swarm theory," in *Proceedings of the sixth international symposium on micro machine and human science*, vol. 1, pp. 39–43, New York, NY, 1995.

[222] M. Omran, A. Salman, and A. P. Engelbrecht, "Image classification using particle swarm optimization," in *Proceedings of the 4th Asia-Pacific conference on simulated evolution and learning*, vol. 1, pp. 18–22, Singapore, 2002.

[223] D. Van der Merwe and A. P. Engelbrecht, "Data clustering using particle swarm optimization," in *Evolutionary Computation, 2003. CEC'03. The 2003 Congress on*, vol. 1, pp. 215–220, IEEE, 2003.

[224] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[225] A. E. Eiben and S. K. Smit, "Parameter tuning for configuring and analyzing evolutionary algorithms," *Swarm and Evolutionary Computation*, vol. 1, no. 1, pp. 19–31, 2011.

[226] J. C. Dunn, "Well-separated clusters and optimal fuzzy partitions," *Journal of cybernetics*, vol. 4, no. 1, pp. 95–104, 1974.