1-1-2009

Video based human gait bilateral symmetry analysis and its validity for functional assessment after an orthopaedic surgery

Ying Bo Xu
Ryerson University

VIDEO BASED HUMAN GAIT BILATERAL SYMMETRY ANALYSIS AND ITS VALIDITY FOR FUNCTIONAL ASSESSMENT AFTER AN ORTHOPAEDIC SURGERY

By Ying Bo Xu

B. Eng., Northeastern University

Shen Yang, China, 1998

A thesis presented to Ryerson University

in partial fulfillment of the requirement for the degree of Master of Applied Science in the program of Electrical and Computer Engineering

Toronto, Ontario, Canada

August 2009

©YingBo Xu 2009

Author's Declaration

I hereby declare that I am the sole author of this thesis.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

(Ying Bo Xu)

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

(Ying Bo Xu)

Borrower's page

Ryerson University requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Name	Signature	Address	date

Abstract

Video based Human Gait Bilateral Symmetry Analysis and its Validity for Functional Assessment after an Orthopaedic Surgery

© Ying Bo Xu 2009

Master of Applied Science

Department of Electrical and Computer Engineering

Ryerson University

A color marker based computer vision system was proposed to provide spatio-temporal and kinematic information of human gait. By detecting the markers from video sequences, this system provides quantitative gait pattern information for clinicians to evaluate the rehabilitation progress of the patients who had undertaken total knee replacement (TKR) and/or total hip replacement (THR) surgeries. The leg bilateral symmetry as an efficient feature for this evaluation purpose was proposed. To calculate this parameter, leg angle curve as a gait signature was introduced to describe the gait pattern. The symmetry is denoted by dynamic time warping (DTW) distance of this walking signature. Normal and abnormal gait can be distinguished based on the leg bilateral symmetry.

Acknowledgements

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Human Gait Analysis

Walking is the natural way of human locomotion. Meanwhile, it's a complex behavior which is the result of the cooperating movement of almost whole body muscles and the nerve system. The study of human walking styles is called gait analysis. In the Longman Dictionary of Contemporary English, gait is defined as the way someone walks. As early as B. C. 350, Aristotle has discussed about the movement of animal bodies [1]. Eadweard J. Muybridge (1830-1904), an English photographer, was the first one to employ photography to study human and animal locomotion. Video camera systems which profited from the development of computer technology were widely used for gait analysis after 1970s. These video-based systems can efficiently provide detailed kinematics information of human gait.

Gait is an intrinsic human behavioral characteristic and is considered as human signature. With computer processing speed and memory rapid increase, it is possible to process image sequences with reasonable cost. Gait as a biometric was studied from digital video sequences after 1990s. It was reported by psychologists that people can recognize friends by their walk with or without other cues [2, 3]. Gait can be used for human recognition [4], [5], [6], [7], motion tracking [8], gender identification [9], [10], and robot design [11].

Gait analysis is also widely used for clinical purposes. Clinical gait analysis is also called quantitative gait analysis. Kinematics and kinetics information of gait are measured, abnormalities are identified, causes are postulated, and treatments are proposed [12]. Normally, kinematics data will be collected by using video based motion systems and kinetics information will be produced by using force platforms. Gait patterns of pathologically normal people were produced by M. P. Murray *et al* in 1964 [13] and were used to compare with those of pathological abnormal patients in 1967 [14]. Clinical gait analysis are used for rehabilitation program design [15], to evaluate the surgery outcomes [16], to study neurological disorders, such as cerebral palsy [17], to name a few.

1.2 Human Gait Cycle

Gait as human signature is a complex spatio-temporal biometric. It is a repetitive cyclic activity of limb movement, especially lower extremity. The gait cycle is the time interval of motions occurring between two consecutive initial contacts with the floor of the same foot. One gait cycle has seven major components which can be divided into two stages: stance and swing [18], [19], [20]. The stance stage is the period when the one foot is in contact with floor and the swing stage is the period when the same foot is off the floor moving forward. The stance stage comprises initial contact, opposite toe-off, midstance and heel rise, and opposite initial contact. The swing stage has three components which are toe-off, feet adjacent and tibia vertical. Walking is a continuous activity and any time point of the sequence can be set as the start of one gait cycle. In practical application, heel strike is selected as the first component of the gait cycle because it can be recognized easily. Figure 1.1 illustrates an entire gait cycle. The gait cycle starts when the right heel contacts on the floor. The knee is in a stable position and

hip flexion at the maximum position. When the whole body weight is transferred onto right leg, the opposite toe-off phase happens. It is the start point of midstance and the first period of single support. Left heel begins to lift from the floor. Opposite initial contact is the left heel contacts on the floor and it is the start point of preswing. Right hip and knee begin to flex while the ankle is plantar flexing. Right Toe-off is the end of the stance stage and the beginning of the swing stage. The muscles move right leg forward during feet adjacent. The tibia vertical is the last component in which period the tibia of the right leg becoming vertical. When the right heel strike on the floor, next gait cycle begins.
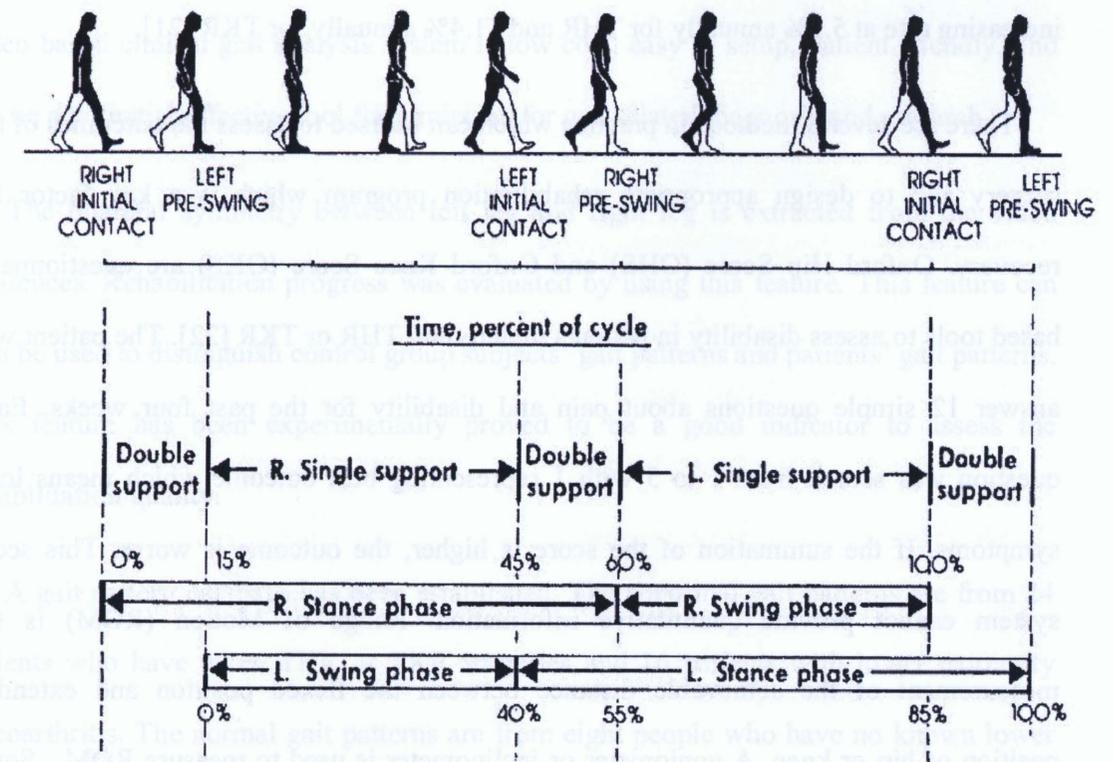

Fig. 1.1 One gait cycle (from [18])

1.3 Total Hip Joint and Total Knee Joint Replacement Surgery

Total Hip Replacement (THR) and Total Knee Replacement (TKR) surgeries are arthroplasties in which the pathological hip joint or knee joint are replaced by artificial joint. The first THR was introduced in 1938 and the first TKR was performed in 1960s [21]. These surgeries can effectively reduce pain and restore the mobility of patients. Although several reasons cause the surgery, such as traffic accidents and athletic injuries, the highly age-related musculoskeletal diseases are the main reasons which cause the surgeries. Based on the data provided in [21], the major population who take this surgery is the people aged from 65 to 84. In France, there are 100,000 THR performed each year. In 2003/04, there are 22724 THR and TKR procedures in Ontario and it has a potential increasing rate at 5.1% annually for THR and 11.4% annually for TKR [21].

There are several methods in practice which can be used to assess the outcomes of the surgery and to design appropriate rehabilitation program which is a key factor for recovery. Oxford Hip Score (OHS) and Oxford Knee Score (OKS) are questionnaire based tools to assess disability in patients undergoing THR or TKR [22]. The patient will answer 12 simple questions about pain and disability for the past four weeks. Each question was scored from 1 to 5 with 1 representing best outcome which means least symptoms. If the summation of the score is higher, the outcome is worse. This score system cannot provide quantitative information. Range of Motion (ROM) is the measurement of the achievable distance between the flexed position and extended position of hip or knee. A goniometer or inclinometer is used to measure ROM. Some hospital data, such as Length of Stay (LOS), blood loss, and analgesia are used to assess the surgery outcome. These tools can provide non-quantitative information or limited

quantitative data to assess the surgery outcome. Fortunately, quantitative gait analysis provides kinematic and kinetic information to assist clinicians in making treatment decision and evaluating the postoperative rehabilitation progress. In some scenario, electromyography (EMG) signals are also collected to produce diagnose information in terms of muscle activity.

1.4 Contributions

In this research project, a color marker based computer vision system was proposed to provide spatial-temporal and kinematic information of human gait. This system has been successfully used to collect the patients gait kinematic information at real clinical environment and provide useful information for rehabilitation progress evaluation. This video based clinical gait analysis system is low cost, easy to setup, patient friendly, and can be a potential effective tool for clinicians for gait related diagnosis and research.

The bilateral symmetry between left leg and right leg is extracted from the video sequences. Rehabilitation progress was evaluated by using this feature. This feature can also be used to distinguish control group subjects' gait patterns and patients' gait patterns. This feature has been experimentally proved to be a good indicator to assess the rehabilitation quality.

A gait pattern database has been established. The abnormal gait patterns are from 24 patients who have taken THR or TKR surgeries and 16 patients with lower extremity osteoarthritis. The normal gait patterns are from eight people who have no known lower extremity diseases. Each person has been filmed twice at each time. One video sequence was filmed with markers on and another video sequence was filmed without markers.

1.5 Organization of This Thesis

The remainder of this thesis is organized as follows.

Chapter 2 examines some existing gait analysis systems. Then, the proposed system setup is shown.

Chapter 3 gives the detailed methodology to obtain the gait data from video sequences. Otsu's thresholding algorithm, K-means algorithm and Mixture of Gaussian model are implemented in order to choose the suitable method for this research. Hip angle and knee angle are extracted from the video sequence. Leg angle as a feature to evaluate the rehabilitation progress is constructed.

Chapter 4 shows the experiment results. By using the bilateral symmetry as a feature, different gait patterns are shown. Rehabilitation progress is evaluated by using this feature. Gait patterns of control group subjects and gait patterns of patients are distinguished based on leg bilateral symmetry.

Chapter 5 draws the conclusions and possible future research directions.

Chapter 2

Vision based Gait Analysis System

2.1 Exiting Systems

Gait analysis can be divided into two classes: biometrics application and clinical diagnosis. Researchers focus on deploying different algorithms to get gait biometrics from video sequences. In this application field, there are two main methods: model based systems and motion based systems. Model based systems concentrate more on the dynamics and motion based systems extract gait information from silhouette shape.

Human walking is a periodic behavior and each gait cycle consists of several adjacent stances. This property makes Hidden Markov Model (HMM) a good choice for gait based human recognition. In [23] [24], human silhouettes are derived from the video sequences. Based on the quality of the silhouettes, the silhouette entirety or out contour was selected as the feature to train the HMM parameters. Each unique stance is an exemplar. Every object s will have a set of exemplars E_s and corresponding HMM parameters θ_s. For a testing video sequence t, the probability of the given sequence t was produced by the sth object in the training database was computed. The testing object was recognized as the sth object when it made the probability have the maximum value.

In [25], Discrete Cosine Transform analyses (DCT) and Support Vector Machine (SVM) method were used to perform automatic gait recognition. The body silhouette was produced by using background subtraction and it was divided into three contiguous

horizontal segments, $R1, R2, and R3$, according to the anatomical fact. The width of each segment was selected as the gait feature because they contain both structural and dynamic information of gait. The thigh angle, front leg knee angle and rear leg knee angle were extracted from the sub segment box, $R1, R2, and R3$. By using a seven-dimension feature space, a multi-class support vector machine (SVM) with a Gaussian kernel was employed to identify the human.

Hayfron-Acquah *et al* [26] derived symmetry maps from optical flow images which were extracted from two successive silhouettes. The Fourier transform was applied to the so-called gait signature which is the average of all symmetry maps from the whole video sequence. Other methods include eigengait described in [27] [31], and mean silhouette in [28].

For model based gait recognition, there are three common human models widely used, Blob, stick and cylinder. Figure 2.1 gives the examples of these models. The stick model is the most often used model to represent the human body.

In [29], a rigid stick model was used to represent the human body, especially the lower extremity. The thigh and the shin are treated as a pendulum and the motion of the thigh and the shin is modeled by the forced coupled oscillator model. A set of sinusoid functions is used to model the hip vertical displacement, thigh rotation, knee rotation and so on. This model provides the estimation of the leg motion which can be used to reduce the data from image sequence. Gait signature was represented by the Fourier description of the thigh and shin motion measured from one gait cycle.
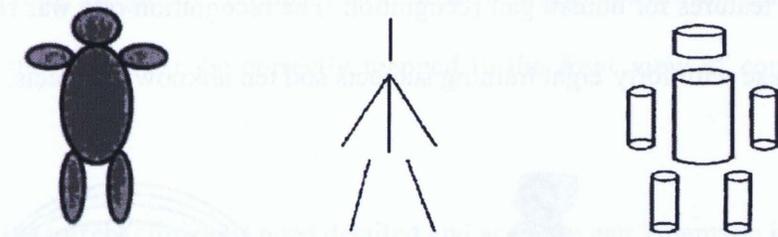
Figure 2.1 Human body models (a) blob, (b) stick, (c) cylinder (from [18])

One disadvantage of the above models is that they cannot accurately fit the contour of the human body. In [30], Green and Guan proposed a more delicate articulated volumetric 3D human model which is called 3D clone-body-model. This model consists of 9 body parts (15 segments), head, clavicle, trunk, upper arms, forearms, hands, thighs, calves, and feet, which connected by joints. There are total thirty two Degrees of Freedom (DOFs). Each body part is considered as a rigid spine with pixels radiating out. Each pixel is denoted by a nine dimensional vector which includes this pixel's cylindrical coordinates (d, θ, r), HIS color space values (h, s, i), accuracy of radius $\propto_r$, accuracy of color $\propto_{HSI}$, and elasticity of radius e_r. Figure 2.2 shows this model and the pixel vector space. Instead of using a generalized cylinder to mimic each body part, this model derives the exact size of each body part. Therefore, this model is robust to somatotype, gender, and age. This model is used in [31] for gait based human recognition. By using the proposed Continuous Human Movement Recognition (CHMR) system, human motion is divided into 35 dynemes which is the smallest contrastive unit of movement. Gait dynemes are used to segment motion vectors in order to get the gait signatures. Knee-hip

angle-angle relationship and left-right asymmetry of gait were proved to be the most significant features for human gait recognition. The recognition rate was reported as 88% for a database with forty-eight training subjects and ten unknown subjects.
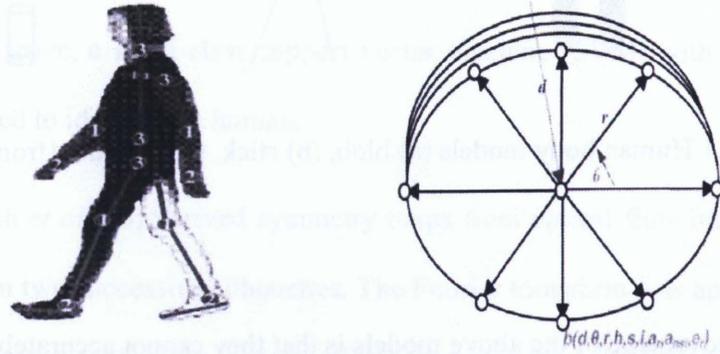

Figure 2.2 3D clone body model (from [30])

Similar 3D articulated human model is also used in [32] for human motion tracking purpose. There are 14 body segments and total 21 DOFs in the model. In general, it is difficult to distinguish different parts of body, especially for self-occlusive lower extremity, when the subject wears loose fit clothing with uniform color and pattern. By using a novel Differential Evolution-Markov Chain (DE-MC) particle filter, authors can track human motions, such as walking, jumping, hopping, and running. The proposed model was reasonably fitted to the side view of corresponding body parts.

Liu and Chellappa [33] tracked human motion by using a similar 3D model. A group of 12 articulated ellipsoids are used to represent the human body. Each joint has one to three degrees of freedom. A 28 degree-of-freedom feature was structured to illustrate the human body. The 3D motion information in the video sequence was projected into 2D

optical flow by using the scaled orthographic projection model. The body motion was extracted by finding the global optimization for a set of linear equations. The results of tracking show the model can be correctly mapped to the front view of corresponding body parts.

For medical research, clinicians need detailed and accurate gait kinematic information to support decision making process. Although the above two papers provide potential methods for markerless video-based clinical gait analysis, marker-based computer vision systems are still the mainstream for clinical gait analysis. There are two types of markers: passive makers and active markers. Passive markers are also known as reflective markers. Cameras pick up the reflections, infrared or visible light, from the markers and the positions of the markers are calculated. Active markers either emit infrared or ultrasound themselves. Most current commercial clinical gait analysis (CGA) systems use passive markers.

Vicon Motion System (VMS) is widely used for clinical gait analysis. This system provides 3D gait kinematic information by using a set of reflective markers and several infrared cameras. It provides kinetic data by force plates and EMG signals are also collected. The cameras emit infrared light as well as record the reflection from the delicately attached markers. The isolated marker dots were identified and tracked based on the anatomy fact and mathematics. 3D data were automatically produced by fusing 2D data from different cameras. A typical VMS needs 3 to 7 infrared cameras. Figure 2.3 gives an example of this system. Other commercial CGA systems include APEC, PEAK.

(a)   (b)

(c)

Figure 2.3 Vicon Motion Capture system (a) infrared camera (b) marker set schedule for lower extremity (c) cameras setup schedule ((a) and (b) , from: http:// www. Vicon .com, (c), from from: http://www.marrc.co.uk/facilities/facilities.html)

In [34], authors proposed a system which uses ultrasound ranging to get 3D gait data. A transponder was worn by the subject and at least three ultrasound receivers are used.

The receivers are mounted evenly apart on a circle with one meter diameter and the center of this circle is the base unit. The distance between the receivers and the position of the base unit need elaborative measurement. The base unit emits infrared pulses to trigger the transponder to emit ultrasound pulses. The distance from the transponder to each receiver is calculated. The motion of body center-of-mass (BCOM) was derived from these distances to obtain gait information.

In [35], authors presented a system using a tracksuit to obtain the gait kinematic information. The tracksuit was designed to have different colors for different parts of the body. Color image segmentation techniques and neural network are used to get the gait information which will be used for the diagnosis of neurological disorders.

In [31], the same tracksuit is used to study the gait of people who has Parkinson's disease (PD) which is characterized by a flexed posture, diminishing armswing, and rigid, small stepped, shuffling gait. The left-right-leg asymmetry was proved to be the most useful feature to distinguish Parkinsonians from healthy people. The proposed system can recognize 95% Parkinsonians from a database which consists of 20 Parkinsonians and 15 healthy people.

One common disadvantage of above systems is that they are not suitable for daily arthroplastic clinical examination. Although VMS can provide accurate gait information, the high cost of the device makes it unaffordable for most clinicians. This system is sensitive to the position of the markers. The markers must be attached obeying a complicated marker set schedule. This is usually accomplished by a professional staff and it costs more than 20 minutes to place the markers on the skin of patients. This is

unacceptable in daily clinical examination environment. Moreover, this system prefers patients wear short pants, such as bike short. This is inconvenient for patients who suffer from lower extremity diseases, such as those come to hospital for follow up exam after THR and TKR surgeries. It is the same scenario for tracksuit. The ultrasound system is difficult to setup and it is not suitable to attach the transponder on the patients, especially aged people with lower extremity diseases.

In this paper, we proposed a new system using color elastic as markers and using regular commercial digital video cameras to get gait kinematic information. By using this information, the clinicians can evaluate the rehabilitation progress after THR and TKR surgeries.

2.2 System Setup

This system is composed of two regular digital video cameras, color elastics as the markers, analysis software and one computer. Fig. 2.4 shows the system setup.

Fig. 2.4 (a) illustrates the distribution of the two cameras, SONY DCR-DVD105. One camera is used to record the side view of the patients and its optical axis is vertical to the walking direction. Another camera is used to record the front view of the patients and its optical axis is parallel to the walking direction. The distance between the side camera and the subject is not specified. A greater distance means more steps can be recorded but it also decreases the accuracy of the markers detection. The cameras record the patients at 29.97 frames per second with 480×720 resolutions. The video sequences were stored on the re-writeable DVD disk. The front view data is not used for this 2D analysis and can be used to obtain 3D model of the gait pattern in the future.

Markers schedule is an important issue for data acquisition. Some common marker sets include Helen Hayes set, Keith Vaughan set and Kadaba set [36]. The choice of the marker set is tricky and the markers need to be carefully calibrated. It needs professional trained staff to calibrate the positions of the markers. These marker sets are necessary for accurate 3D data acquisition. Meanwhile, the natural complexity of them limits the daily clinical usage. Fig. 2.4 (b) shows the markers and the schedule used in this video system. Depending on the patients' pants color, white, black or blue elastic are chosen as markers. The markers can be quickly and reasonably tight placed outside the pants without affecting normal walking manner. For each leg, two markers attached on the thigh and two attached on the shank. The patient walks back and forth in his or her normal walking pace and manner. There is no noticeable displacement of the markers due to the movement of the pants because the elastic markers can change the size with the deformation of the muscles. It is reasonable to ignore the deformation of the markers due to the muscles movement.

By connecting the middle of two relevant markers, a stick articulated model of the human leg is introduced. This model is shown as Fig. 2.4 (c). Lower extremity locomotive information is reserved and all other body movement information is ignored. The hip angle is defined as the angle between thigh and vertical axis. The knee angle is defined as the angle between shank and vertical axis. This definition is slightly different from the definition in [37] which is a common definition. In [37], hip angle was defined as the angle between thigh and horizontal axis. Knee angle was defined as the auxiliary angle of the angle between thigh and shank. The advantage of the definition in Fig. 2.4 (c) is that it distinguishes the movement of thigh and shank. The proposed leg movement is

based on this definition. This system needs to be calibrated to make sure the absolute vertical axis.

(a)

(b)

(c)
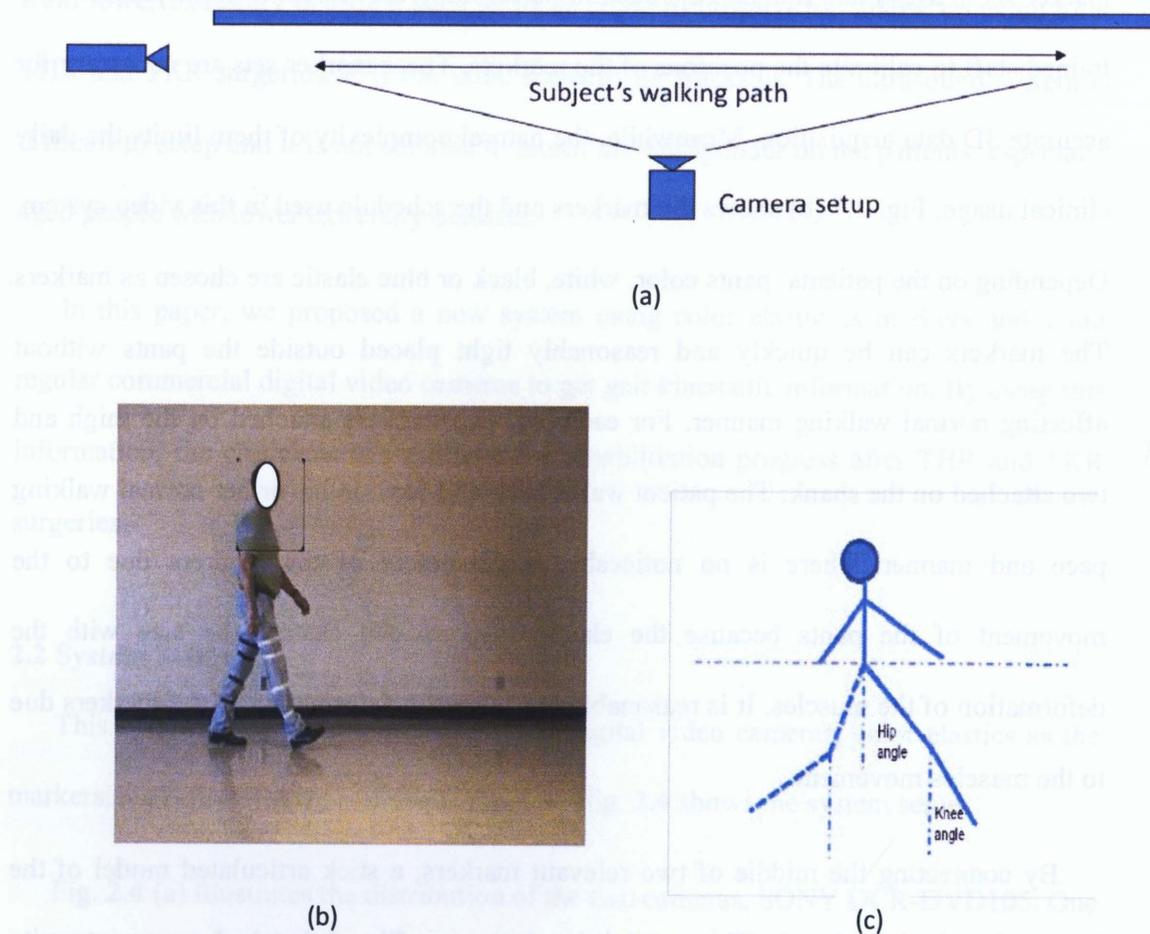
Figure 2.4 Proposed system setup (a) cameras setup schedule (b) markers set schedule (c) human stick articulated model

Chapter 3

Methodology

3.1 Deinterlacing

The video sequences were filmed at two clinical locations, Mt. Sinai Hospital and Oakville Orthopedic Clinic. The footages were recorded with NTSC system which is used by Sony digital camcorders. The video sequence is recorded with interlacing technique which provides good picture quality on the CRT TV and saves the bandwidth at the same time. There are two fields, even field and odd field, in one frame which means the camcorder captures an image by scanning the even horizontal lines and the odd horizontal lines respectively. These two fields are recorded separately in time. This technique takes advantage of the human eye response time which is 0.1 second and benefits the display of traditional CRT by reducing flicker and save bandwidth. When this footage is displayed by a LCD monitor working with progressive scan model, the situation is changed. Saw teeth phenomenon will appear. If there is one high speed moving object in the scene, saw teeth can be seen even on CRT screen. Interlacing also makes it is difficult for image based tracking. Figure 3.1 (a) shows the saw teeth phenomenon around the white markers. The correct positions of white markers on the shank cannot be determined in this situation. Deinterlacing technique is needed to digital quantize the interlaced video footage. The techniques of deinterlacing can be divided into two categories: motion-compensated deinterlacing and non-motion-compensated deinterlacing [38]. For this research project, the purpose of deinterlacing is to eliminate

the saw teeth in order to get the sharp boundaries of the markers. For this purpose, the simplest technique is discard one field. The result of this method is only half height of the original image left. Figure 3.1 (b) shows this effect. In some scenario, the markers may be totally disappeared due to the width of the markers and the object-camera distance. So, this method is not suitable for this research. Figure 3.1 (c) shows the result of another method which is to duplicate one field to reconstruct the whole frame. Due to the time interval between even field and odd field, this method will cause intrinsic errors of the marker position. When an object is moving quickly, this error can't be ignored. We employed the method which is provided by Smart Deinterlacing Filter (Version 2.6) and it is available at http://docs.huihoo.com/transcode/0.6.14/smart.html. Figure 3.1 (d) shows the result by using this cubic spline interpolation. This filter is used as a plug in filter for VirtualDub-MPEG2 which is used to split the footage shot into frames. VitualDub-MPEG2 can be free downloaded from the following website: http://home.comcast.net/~fcchandler/stable/index.html.
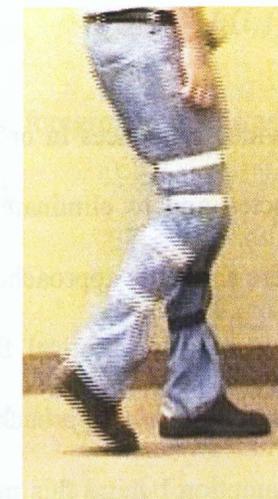
(a) Original   (b) Discard field 1
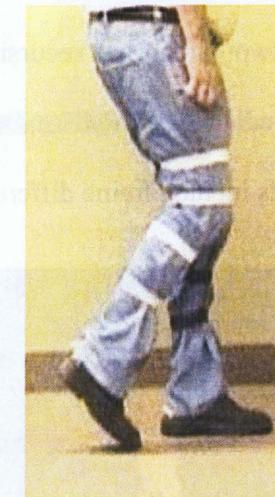
(c) Duplicate field 1   (d) Cubic interpolate

Figure 3.1 Deinterlacing methods (a) original image (b) discard field 1 (c) duplicate field 1 (d) cubic interpolate

3.2 Silhouette Extraction

The walking patient was extracted from the video sequences in order to narrow the searching areas where the markers will be detected and to eliminate the background interference on the markers detection process. There are many approaches to find moving objects from a video sequence, such as blocking matching, optical flow, background subtraction and infer-frame differencing. Among these approaches, background subtraction is a widely adopted method. The assumption behind this method is that the moving objects in current image will significantly change the illumination intensity at their pixels. By subtracting the current image from a background model, the deviation pixels are considered as moving objects. So, background modeling is the kernel of the background subtraction technique. In [39], authors classified the background modeling algorithms into two categories: recursive and non-recursive. Recursive methods include approximated median filter (MF), Kalman filter, and Gaussian mixture model. Non-recursive methods include frame differencing, median filter, and linear predictive filter.

Median filter is a widely used background modeling algorithm [39-41]. This method calculates the median value at each pixel for all the frames which are stored in a buffer. It assumes that the background pixels are visible for more than half the total stored frames. Let $X = (X_1, X_2, ..., X_t)$ denotes the illumination intensity values of a particular pixel(i_0, j_0) over time T, the background pixel value at that point will be recognized as:

$$B(i_0, j_0) = Median(X), \qquad (3-1)$$

The moving object is considered as those pixels whose intensity values satisfy the following criteria:

$$|I_t(i,j) - B(i,j)| > Threshold, \qquad (3-2)$$

By setting the foreground pixels as 1 and background pixels as 0, binary silhouette image is achieved.

$$I(i,j) = \begin{cases} 1 & \forall\, I(i,j) \in foreground \\ 0 & others \end{cases}, \qquad (3-3)$$

Figure 3.2 (a) shows one frame from a video sequence. Fig. 3.2 (b) shows the background image extracted from this sequence by using median filter. The walking patient in (a) is extracted as foreground in (c).

(a)

(b)

(c)
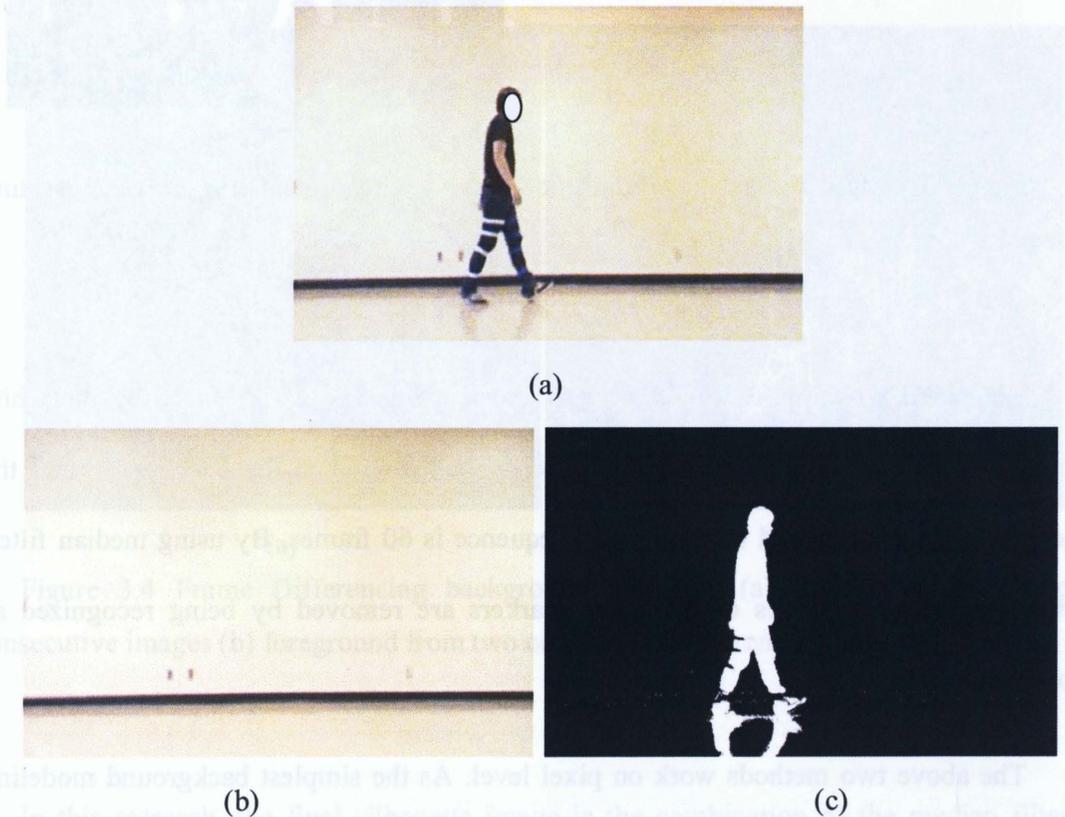
Figure 3.2 Median Filter background modeling (a) original image (b) background image (c) foreground image

20

21

Mean filter uses the mean value at each pixel rather than median values.

$$B(i_0, j_0) = \frac{1}{N}\Sigma_{t=1}^{T} X_t ,\qquad\qquad (3-4)$$

The corresponding background and foreground is shown in Figure 3.3 (a) and (b).

(a)   (b)
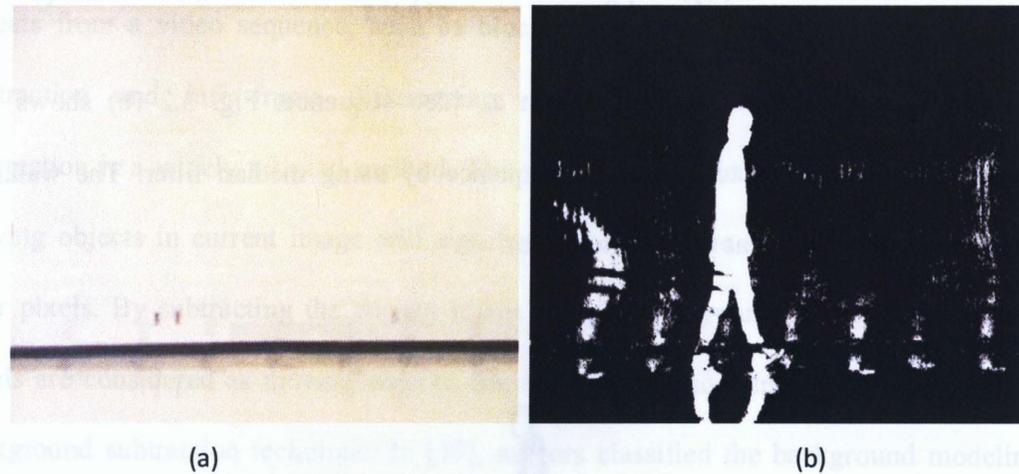
Figure 3.3 Mean Filter background modeling (a) background image (b) foreground image

Compared with Fig. 3.2(a), the white ghost in Fig 3.3 (a) is caused by the white markers. The intensity values of the markers are higher than the background and the length of the background training video sequence is 60 frames. By using median filter, the high intensity values of the white markers are removed by being recognized as outlying values.

The above two methods work on pixel level. As the simplest background modeling method, frame differencing works on frame level. This method uses the frame at time $t-1$ as the background model for the frame at time t. This method is only useful to catch slow moving objects in a nearly still background scenario. The fatal weakness of

this method is the aperture phenomenon which is caused by a slow moving object which has a large uniform color interior area [39]. In [42], the authors extended this method by using the intersection of two consecutive differencing images as the foreground. Let I_t, I_{t+1}, and I_{t-1} represent current frame, previous frame, and future frame respectively. The foreground image, G, is identified by

$$G = (I_{t+1} - I_t) \cap (I_t - I_{t-1}),\qquad\qquad (3-5)$$

Figure 3.4(a) shows the ghost and aperture and figure 3.4 (b) shows the walking people's silhouette by using $(3-5)$.

(a)   (b)
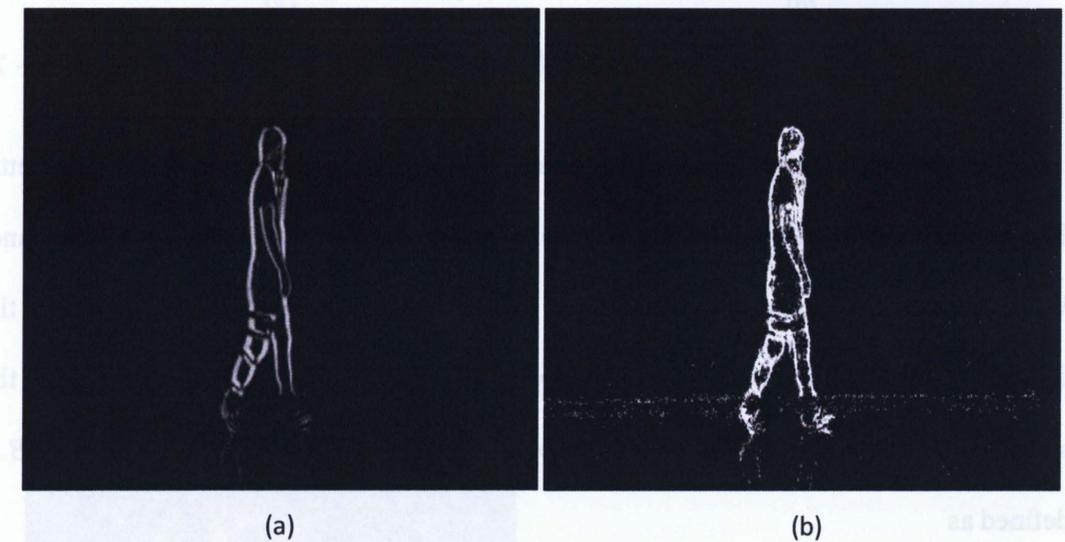
Figure 3.4 Frame Differencing background modeling (a) foreground from two consecutive images (b) foreground from two consecutive differencing images

In this research, the final silhouette image is the combination of the median filter result and frame differencing. This method can minimize the gap on the silhouette image

which is caused by close intensity values between the markers and background. The result is shown in figure 3.5 (a).

$$f(x, y) = \begin{cases} 1 & if \ (I + G) \geq 1 \\ 0 & others \end{cases}, \qquad (3-6)$$

After the silhouette extraction, there are maybe some gaps and holes in the silhouette. In order to recall the whole pixels of walking people, the gaps of the markers and the small holes in the silhouette were filled by using morphological methods: dilation and erosion which are described in [43]. The dilation of binary image f by the structuring element B is defined as

$$f \oplus B = \{ Z | [(\hat{B}_z) \cap f] \subseteq f \}, \qquad (3-7)$$

Dilation operation can bridge the gaps and it also "thickening" the original silhouette. The background pixels around the edge of the silhouette are converted to foreground. When the markers intensity is very close to the background, this effect will cause the failure of further processing. So, erosion operation follows the dilation operation with the same structuring element. The erosion of binary image f by the structuring element B is defined as

$$f \ominus B = \{ Z | (B_z) \subseteq f \}, \qquad (3-8)$$

Figure 3.5 (b) shows the result of (a) after dilation and erosion operation. Now, the silhouette of the waling people is extracted from the video sequence. By setting the background pixels in Fig. 3.2 (a) to zero, the walking people image is extracted and shown in Fig. 3.5 (c). This is the areas where the markers will be detected. Fig. 3.5(d) shows two common structuring elements.
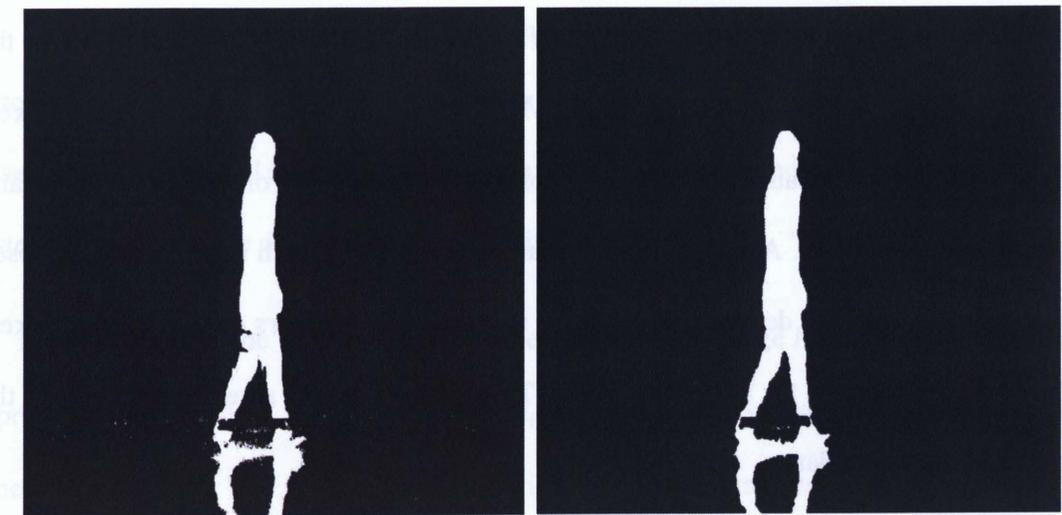
(a)   (b)
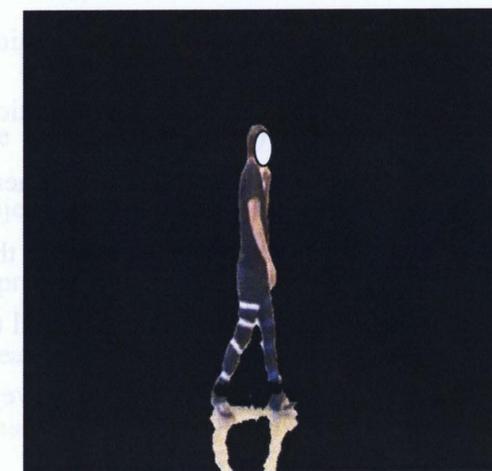
(c)   (d)

Figure 3.5 Walking people extraction (a) original silhouette image (b) final silhouette image after morphological operations (c) walking people (d) 3×3 square structure element and 5×5 diamond structure element

3.3 Otsu's Thresholding Method

To segment the white or black marketers from the image, one assumption is that the white markers are the brightest part in the whole image and black markers are the darkest part. This is true because the marker's color is chosen based on the pants color and illumination condition. At any time, four markers on the leg which is completely exposed to the side camera are detected. This means that the white markers only or black markers only will be detected for any given frame. This strategy makes sure the accuracy of the hip and knee angle data.

The marker pixels should be classified and clustered from the image. This is the field of image segmentation. Image segmentation is very important for computer vision. It provides image information for further level image processing, such as object recognition, image compression, image editing, and image retrieval. The goal of image segmentation is to cluster pixels into different regions based on the visual characteristics. These characteristics can be color, texture, intensity, and motion. Thresholding is probably the most widely used technique to segment the image into a binary image. It's non-trivial to automatically select a threshold to convert an image into binary one. As a good survey, paper [44] described 40 thresholding methods and classified them into six categories:

- histogram shape-based methods
- clustering-based methods
- entropy-based methods
- object attribute based methods
- the spatial methods

- local adaptive threshold

In 1979, Nobuyuki Otsu proposed a nonparametric and unsupervised method to automatically select the optimal threshold in the paper [45]. This method is clustering based and the threshold is set by making each cluster as close as possible. This method is tested to find the white or black markers in this paper.

Let N represents the total number of pixels in a given image I. The gray levels are represented by $L = [0,1,2,\cdots,L-1]$ and n_i is the number of pixels at gray level i where $N = \sum_{i=0}^{L-1} n_i$. The probability distribution function of gray level i is calculated by:

$$p_i = \frac{n_i}{N}, \qquad \text{where } p_i \geq 0, \text{and} \sum_{i=0}^{L-1} p_i = 1, \qquad (3-9)$$

Now suppose that the pixels in the image belong to two groups, G_B and G_M. G_B denotes the background (pants) pixels with gray intensity levels $[0,1,2,\cdots,T]$. G_M denotes the objects (white markers) pixels with gray intensity levels $[T+1,\cdots,L-1]$. Let μ_B and σ_B^2 represent the mean and variance of class G_B, respectively. Let μ and σ^2 represent the mean and variance of the whole image. Let μ_M and σ_M^2 represent the mean and variance of class G_M, respectively. The *within-class* variance is defined as:

$$\sigma_W^2(T) = \omega_B(T)\sigma_B^2(T) + \omega_M(T)\sigma_M^2(T), \qquad (3-10)$$

where:

$$\omega_B(T) = \sum_{i=0}^{T} p_i, \qquad (3-11)$$

$$\omega_M(T) = \sum_{i=T+1}^{L-1} p_i, \tag{3-12}$$

$$\mu_B(T) = \sum_{i=0}^{T} \frac{i p_i}{\omega_B(T)}, \tag{3-13}$$

$$\mu_M(T) = \sum_{i=T+1}^{L-1} \frac{i p_i}{\omega_M(T)}, \tag{3-14}$$

$$\sigma_B^2(T) = \sum_{i=0}^{T} \frac{\left(i - \mu_B(T)\right)^2 p_i}{\omega_B(T)}, \tag{3-15}$$

$$\sigma_M^2(T) = \sum_{i=T+1}^{L-1} \frac{\left(i - \mu_M(T)\right)^2 p_i}{\omega_M(T)}, \tag{3-16}$$

It is always true that:

$$\omega_B(T)\mu_B(T) + \omega_M(T)\mu_M(T) = \mu, \tag{3-17}$$

$$\omega_B(T) + \omega_M(T) = 1, \tag{3-18}$$

The optimal threshold T^* is the value which minimizes the value of *within-class* variance σ_W^2. However, this involves second-order statistics calculation. To simplify the calculation, between-class variance σ_B^2 is defined as:

$$\sigma_C^2(T) = \omega_B(T)(\mu_B(T) - \mu)^2 + \omega_M(T)(\mu_M(T) - \mu)^2, \tag{3-19}$$

$$= \omega_B(\mu_B - \omega_B\mu_B - \omega_M\mu_M)^2 + \omega_M(\mu_M - \omega_B\mu_B - \omega_M\mu_M)^2,$$

$$= \omega_B\omega_M(\mu_B - \mu_M)^2, \tag{3-20}$$

$$= \omega_B\omega_M(\mu_B - \mu_M)^2, \tag{3-21}$$

$$= \omega_B(1 - \omega_B)(\mu_B - \frac{\mu - \omega_B\mu_B}{1 - \omega_B})^2,$$

$$= \frac{\omega_B(\mu - \mu_B)^2}{1 - \omega_B}, \tag{3-22}$$

$\sigma_C^2(T)$ is based on the first-order statistic parameters and the computation is not as heavy as the calculation of within-class variance. Equation $(3-22)$ has different expression form from the original one in [42]. Equation (3-22) is concise in expression and one parameter less than the one used in [42].

The optimal threshold T^* is the one which maximizes the value of *between-class* variance $\sigma_C^2(T)$.

$$\sigma_C^2(T^*) = \max_{0 \leq T \leq L-1} \sigma_C^2(T). \tag{3-23}$$

Figure 3.6 shows the examples of white markers detection and black markers detection by using this method. Fig. 3.6 (a) is the original image. (b) is the grayscale intensity image with all background pixels have been set to zero. (c) is the binary image where all bright pixels denotes those in the white markers class. In this case, the threshold is $T^* = 105$. The pixels are classified as white markers if their intensity values are greater than T^*. It is very clear that all four white markers are extracted from the original frame. However, the fault detection is also clear. Some pixels which belong to the edge of pants are treated as white markers. However, the gray intensity levels between these pixels and white markers pixels are significant. This phenomenon is conspicuous under some illumination conditions and the whole leg is recognized as the marker. Fig. 3.6 (d)-

(f) show the result of black markers detection. The threshold is $T^* = 136$. The pixels belonging to shank skin are classified as black markers due to two reasons. One is because of the close intensity values between these two objects and another reason is that the threshold derived from Otsu's method is loose. To solve this problem, Otsu proposed to use multi-thresholds to segment the image.

| (a) | (b) | (c) |

| (d) | (e) | (f) |

Figure 3.6 Otsu's method for markers detection (a) original image (b) foreground image (c) white markers (d) original image (e) foreground image (f) black markers

The more general expression of equation (3-19) gives the formula:

$$\sigma_C^2(T_1, T_2, \cdots T_M) = \sum_{i=1}^{M+1} \omega_i(T_i)((\mu_i(T_i) - \mu))^2, \qquad (3-24)$$

And the optimal thresholds are set by:

$$\sigma_C^2(T_1^*, T_2^*, \cdots T_M^*) = \max_{0 \le T \le L-1} \sigma_C^2(T_1, T_2, \cdots T_M). \qquad (3-25)$$

where M is the number of thresholds.

This method provides a possible solution for the above false segmentation. When M increases, this method is highly intensive in terms of computation. In this particular research project, there are roughly four classes in a given image: two sets of color markers (white, black, or blue), clothes (pants), and background (can be set as zero or one). So, K-means clustering method is tested for markers detection.

3.4 K-means Clustering Method

K-means clustering method was first proposed by MacQueen in 1967 in [46]. As one simple unsupervised clustering method, it has been widely adopted for data compression and data classification. Suppose the dataset X has n vectors, $X = \{X_1, X_2, \cdots, X_n\}$, where each vector is d-dimensional. The goal of K-means method is to divide the n vectors into K groups, $G = \{G_1, G_2, \cdots G_k\}$, and to minimize the following objective function:

$$J = \sum_{i=1}^{k} \sum_{X_j \in G_i} ||X_j^{(i)} - C_i||^2. \qquad (3-26)$$

where C_i is the centroid of group G_i. $X_j^{(i)}$ is the jth vector which belongs to G_i. This formula shows that the idea behind K-means clustering method is to minimize the sum of within-cluster square error of group G_i.

The traditional K-means clustering method is composed by the following steps:

1. K points from the dataset are randomly selected as the initial centers for each group.

2. Calculate the distance between the vector in X and each centers, the vector is assigned to the group whose center has the smallest distance to this vector.

3. Compute the new centroid of each group, replace the corresponding centroid from last iteration.

4. Repeat step 2 and step 3 until the change of objective function is less than a preset threshold or a maximum number of iteration is reached.

It should be noticed that K-means clustering method does not guarantee a global minimum of the objective function, especially for high dimensional data. Different initial centers may give different clustering results. A tricky way is to select the initial points as much as possible far away from each other or use some *a priori* knowledge about the groups' distribution. Another non trivial thing is to determine the number of k. To segment the markers out in this research, four clusters are assumed. Based on the presumption that black markers have the lowest grayscale and the white markers have the highest grayscale among the testing areas, the initial centers are set as $C = \{20, 90, 160, 220\}$. The final centers are considered as the threshold for different groups. The smallest value is set as the threshold for black markers and the maximum value is considered as the threshold for white markers. Figure 3.7 shows some examples of k-means clustering method. Fig. 3.7 (a) is the k-means result of Fig. 3.6 (a). In this case, the final center is: $C = \{54, 90, 118, 171\}$. All pixels in Fig. 3.6 (b) whose grayscale intensity is greater than 171 are set as 255. Other pixels are set as zero. Compared with Fig. 3.6 (c) where the threshold is 103, Fig. 3.7 (a) has much better segmentation result with less noise for the white markers. However, the lowest marker is not clustered correctly. The right part of this marker is missing because the pixels at these points have lower grayscale intensity. The same phenomena happened at the right part of second marker from top. The gap between the two part pixels will introduce errors for angle calculation. Fig. 3.7 (b) shows the k-means result of Fig. 3.6 (d). In this case, the final center is: $C = \{60, 104, 153, 192\}$. All pixels in Fig. 3.6 (e) whose grayscale intensities are less than 60 are set as 255. Other pixels are set as zero. The third marker is broken by this method.


(a)   (b)
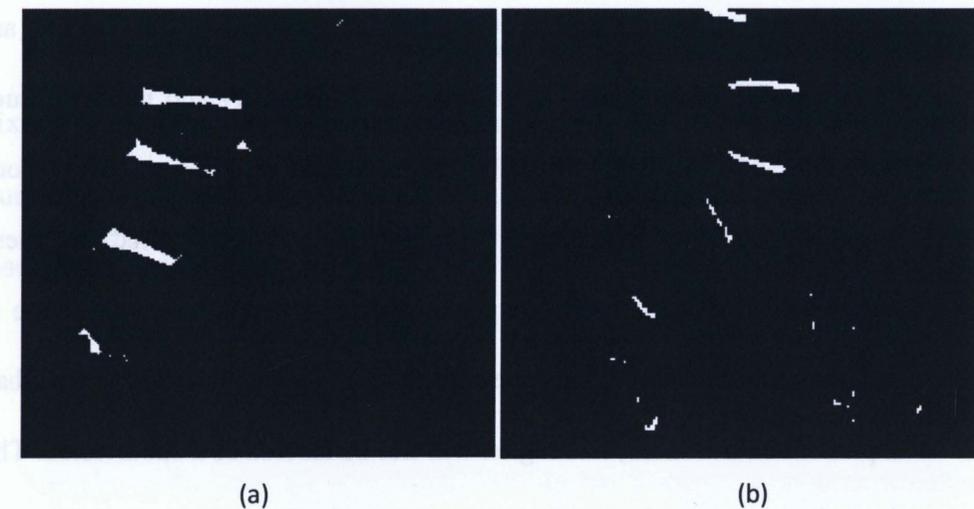
Figure 3.7 K-means method for markers detection (a) white markers (b) black markers

The reason is that traditional K-means method only takes into account the grayscale intensity values from the whole image. The spatial relationship information of the pixels is ignored. Thus, the algorithm of K-means with connectivity constraint (KMC) was proposed by Kompatsiaris *et al* in [47]. Here is the brief expression of this method:

1. The dataset was divided into K regions by performing original k-means method with a small number of iterations. Each region G_k, $k = 1, 2, \cdots, K$, has the center intensity C_k, spatial center $S_k = (S_{k,x}, S_{k,y})$ and area A_k. The mean area of all regions is A.

2. For every pixel p the intensity differences are evaluated between center and pixel as well as the distances between this pixel and each region's spatial center. The distance between p and S_k is generalized as follows:

$$D(p,k) = \frac{\lambda_1}{\sigma_C^2} \|C(p) - \bar{C}_k\| + \frac{\lambda_2}{\sigma_S^2} \bar{A} \frac{\|p - \bar{S}_k\|}{A_k} \qquad (3-27)$$

where $\|\bullet\|$ is the Euclidean distance, σ_C^2 and σ_S^2 are the standard deviations of grayscale intensity and spatial distance, respectively. λ_1 and λ_2 are regularization parameters. Pixels with similar intensity and motion values with those of large object would be assigned to neighboring smaller regions.

3. Eight connectivity component labeling method is applied to these subdivision. All connected components are found and a unique value is assigned to all pixels in the same component. If a region's area is less than a predefined threshold, this region will not be labeled as a subdivision. The intensity center and spatial center are recalculated for each connected region which is labeled by this component labeling algorithm.

4. The iteration will stop when the difference between the new and the old centers is below a threshold or maximum iteration number is reached. The number of connected regions is considered as k. This method provides a way to select the number of cluster which is the shortcoming of the traditional k-means algorithm.

3.5 Gaussian Mixture Model

The K-means algorithm is a nonparametric and unsupervised data clustering method. While, Gussians Mixture Model (GMM) is a popular recursive, model based technique for data clustering. It was first proposed for background modeling in [48]. The idea of this method is that the distribution of a single pixel's values over time can be modeled by a mixture of K Gaussian distributions. Based on the weight and variance of each Gaussian component, it is not difficult to determine which Gaussian may correspond to the background model. GMM has been successfully used for background modeling [49-51].

In this research, GMM is implemented to find threshold for marker detection. A mixture of K Gaussian components is used to model the histogram distribution of illumination intensity values, $X = \{X_1, X_2, \dots, X_N\}$, for a given image. The probability of observing a particular pixel value is

$$P(X|\theta) = \sum_{i=1}^{K} \omega_i * p_i(X|\theta_i), \qquad (3-28)$$

where K is the number of Gaussian distributions, ω_i is the weight of the ith Gaussian in the mixture model and $\sum_{i=1}^{K} \omega_i = 1$. $\theta = (\omega_1, \dots, \omega_K, \theta_1, \dots, \theta_K)$ and p_i is a Gaussian

density function parameterized by $\theta_i = (\mu_i, \Sigma_i)$. In this project, $\Sigma_i = \sigma_i^2$ because the grayscale intensity is used to find the markers.

$$P(X, \mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X_i - \mu_i)^2}{2\sigma^2}\right), \qquad (3-29)$$

In order to estimate the parameters for each ith Gaussian distribution in the mixture, model, EM algorithm is a standard choice. In [49], the author defined the likelihood function $L(\Theta|X)$ as:

$$P(X|L(\Theta|X)) = \prod_{i=1}^{K} P(x_i|\Theta) = L(\Theta|X), \qquad (3-30)$$

To find the Θ^* where

$$\Theta^* = \underset{\Theta}{\operatorname{argmax}} L(\Theta|X), \qquad (3-31)$$

The incomplete-data log-likelihood expression provided in [46] is:

$$\log\big(L(\Theta|X)\big) = \sum_{i=1}^{N} \log\left(\sum_{j=1}^{K} \omega_j * p_j(x_j|\theta_i)\right), \qquad (3-32)$$

By introduce a new unknown, random variable Y which governed by an underlying distribution $f\big(y|X, \Theta^{(i-1)}\big)$, the complete-data likelihood function is defined as:

$$\log\big(L(\Theta|X,Y)\big) = \log(P((X,Y|\Theta)) = \sum_{i=1}^{N} \log\big(\omega_{yi} * p_{yi}(x_j|\theta_{yi})\big), \qquad (3-33)$$

The E-step is to find the expected value of the complete-data log-likelihood $\log(P((X, Y|\theta))$ with respect to the unknown data Y given X and the current parameters estimates. The author provides this formula:

$$Q\big(\Theta, \Theta^{(i-1)}\big) = E\big(\log P(X,Y|\theta) \,|X, \Theta^{(i-1)}\big), \qquad (3-34)$$

$\Theta^{(i-1)}$ is the current parameter and Θ is the new optimized parameters to increase Q.

The M-step is to maximize the expectation Θ in the E-step.

$$\Theta^{(i)} = \underset{\Theta}{\operatorname{argmax}} Q(\Theta, \Theta^{(i-1)}), \qquad (3-35)$$

For GMM parameters estimation, the author provides the final computational formula as follows:

$$\omega_{i,t} = \frac{1}{N} \sum_{j=1}^{N} p(i|x_j, \Theta^{(t-1)}), \qquad (3-36)$$

$$\mu_{i,t} = \frac{\sum_{j=1}^{N} x_j\, p(i|x_j, \Theta^{(t-1)})}{\sum_{j=1}^{N} p(i|x_j, \Theta^{(t-1)})}, \qquad (3-37)$$

$$\sigma_{i,t} = \frac{\sum_{j=1}^{N} p(i|x_j, \Theta^{(t-1)})(x_j - \mu_{i,t})(x_j - \mu_{i,t})^T}{\sum_{j=1}^{N} p(i|x_j, \Theta^{(t-1)})}. \qquad (3-38)$$

where $i = 1, 2, \dots, K$.

The above formulae $(3-36) \sim (3-38)$ do the E-step and M-step simultaneously. However, EM algorithm is only guaranteed to converge to a local maximum of the

likelihood function and we cannot prove whether this local maximum is also the global maximum or not. When $|\theta - \theta^{(t-1)}| < \varepsilon$, the iteration will stop.

The steps for GMM parameters estimation are summarized as follows [52]:

1. Randomly chose original parameter θ^0, where

$$\theta^0 = \left(\omega_{1,0}, \dots, \omega_{K,0}, \mu_{1,0}, \dots, \mu_{K,0}, \sigma_{1,0}^2, \dots, \sigma_{K,0}^2\right);$$

2. Using θ to compute $P_i(x_j)$, $i = 1,2,\dots,K$, $j = 1,2,\dots,N$;

3. Using $(3-36)\sim(3-38)$ to compute new $\omega_{i,t}, \mu_{i,t}, \sigma_{i,t}^2$;

4. If $|\theta - \theta^{(t-1)}| < \varepsilon$, or maximum iteration number is reached, stop, else go to next iteration.

So far, the parameters of each component's mean, variance and weight in the model have been achieved. With the assumption that white markers possess the brightest pixels and black markers have the darkest pixels, the white markers are those pixels whose intensity values satisfy this formula:

$$f(x,y) = \begin{cases} 1, & if \ \ f(x,y) > \max(\mu_k) + \lambda * \sigma_k \\ 0, & others \end{cases} . \quad (3-39)$$

The black markers are those pixels whose intensity values satisfy this formula:

$$f(x,y) = \begin{cases} 1, & if \ \ f(x,y) < \min(\mu_k) + \lambda * \sigma_k \\ 0, & others \end{cases} . \quad (3-40)$$

λ is a constant coefficient whose value can be any number in $[-3,3]$. Figure 3.8 shows the GMM results for markers detection. Fig. 3.8 (a) shows the GMM result for Fig.

3.6 (a). The GMM parameters are: $\mu = \{70, 93, 102, 143\}$, $\sigma = \{22, 9, 22, 35\}$, and $\omega = \{0.23, 0.40, 0.22, 0.15\}$. The threshold used here is 143. Fig. 3.8 (b) shows the GMM result for Fig. 3.6 (d). The GMM parameters are: $\mu = \{85, 128, 166, 184\}$, $\sigma = \{27, 43, 27, 16\}$, and $\omega = \{0.12, 0.18, 0.33, 0.37\}$. The threshold used here is 85. The noise pixels can be removed by using morphological operations. Compared with Otsu's thresholding and K-means method, GMM method provides the tradeoff between recall and precision for markers detections. Therefore, GMM algorithm is employed for markers detection in this research.

(a)   (b)
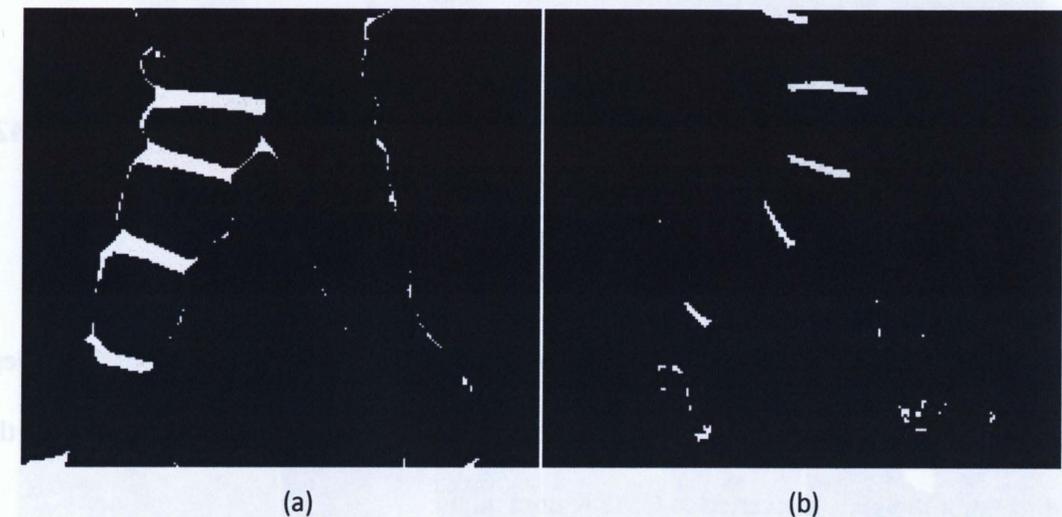
Figure 3.8 GMM method for markers detection (a) white markers (b) black markers

3.6 CMYK Color Space

The selection of color markers is based on the cloth color. Sometimes, black or white markers are not suitable for certain pants color. For example, it will be very difficult or even impossible to find white markers on the light color pants under some illumination conditions. Meanwhile, the grayscale intensity values of black markers are very close to

some jeans color. To solve this problem, blue markers are used and detected in CMYK domain. CMYK color system is used for color printing, where C, M, Y, K represents cyan, magenta, yellow and black, respectively [43]. The simplest operator of conversion from RGB to CMYK is described in [43] as:

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \qquad (3-41)$$

Another conversion algorithm is defined by PostScript language as:

$$\begin{aligned} Ctemp &= 1 - R \\ Mtemp &= 1 - G \\ Ytemp &= 1 - B \\ Ktemp &= \min(Ctemp, Mtemp, Ytemp) \\ C &= \min(1, \max(0, Ctemp - Ktemp)) \\ M &= \min(1, \max(0, Mtemp - Ktemp)) \\ Y &= \min(1, \max(0, Ytemp - Ktemp)) \\ K &= \min(1, \max(0, Ktemp)) \end{aligned} \qquad (3-42)$$

By experimental observations, the use of blue as a color cue cannot benefit from these two algorithms. By using the *makecform* function which is provided by MATLAB, the RGB color image is converted to CMYK domain by:

```
cform = makecform('srgb2cmyk');
CMYK = applycform(RGB,cform);
```

In each CMYK channel, the pixels of blue markers possess significant difference from other pixels. By taking advantage of this property, a new synthetic image is proposed as:

$$I_{syn}(x, y) = \frac{C + M + (1 - Y) + (1 - K)}{4}, \qquad (3-43)$$

The pixels of four blue markers in this synthetic image are clustered by using GMM algorithm which is similar to white markers clustering. Figure 3.9 shows this process. Fig. 3.9 (a) is an original image where blue markers are used. Fig. 3.9 (b)-(e) is CMYK channel, respectively. Fig. 3.9 (f) is the synthetic image after (3-43). Fig. 3.5 (g) shows the result of clustering.

(a)   (b)   (c)

(d)   (e)   (f)   (g)
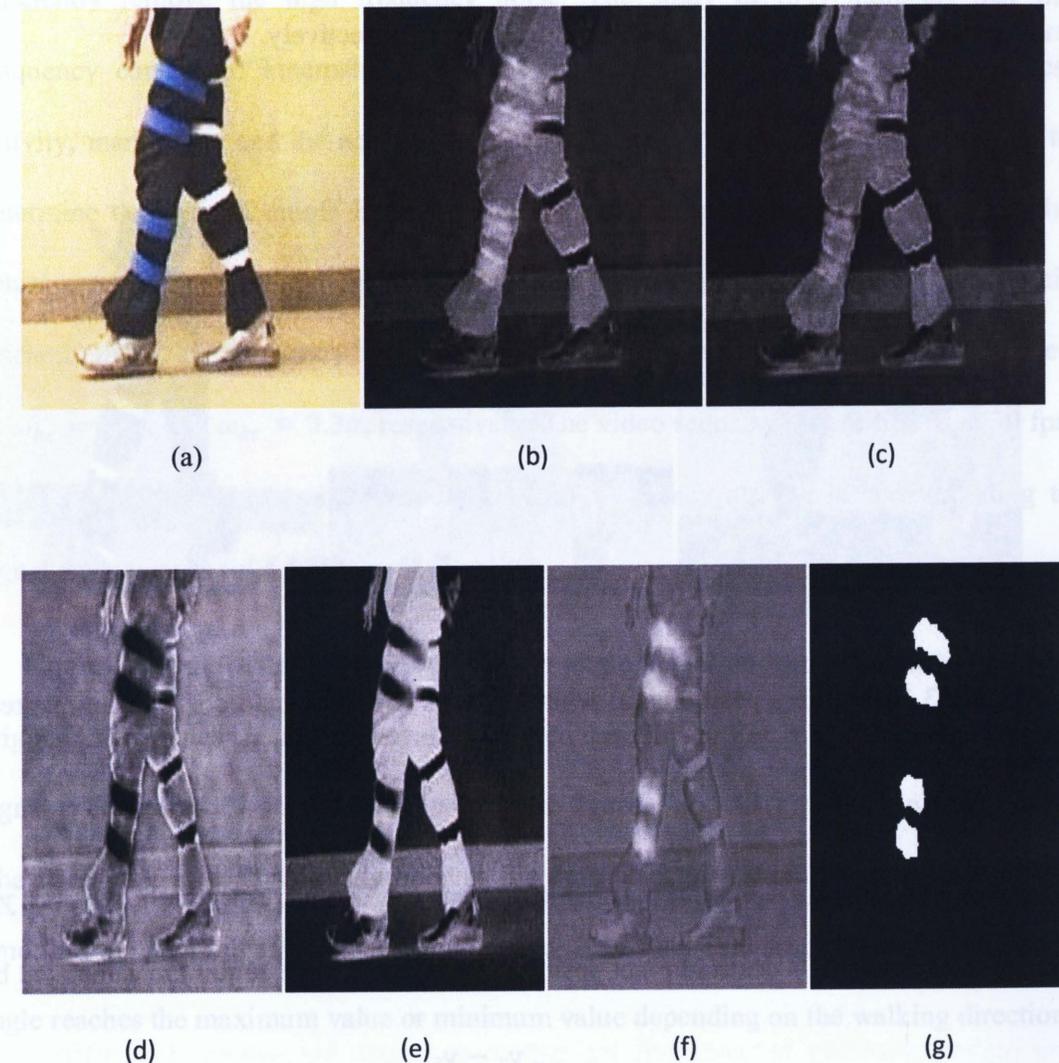
Figure 3.9 Blue markers detection (a) original image (b) C channel image (c) M channel image (d) Y channel image (e) K channel image (f) synthetic image (g) blue markers positions

3.7 Hip and Knee Angle Acquisition

The binary image derived from GMM method may contains some noise pixels from the edge of pants or other parts. Morphological operations are needed to remove these noise pixels. Then the four biggest areas in the binary image are considered as the markers. The centroid point is denoted as the maker's center. Figure 3.10 (a)-(c) show the markers positions in Fig. 3.6 (a), (d), and Fig. 3.9 (a), respectively.


(a)   (b)   (c)
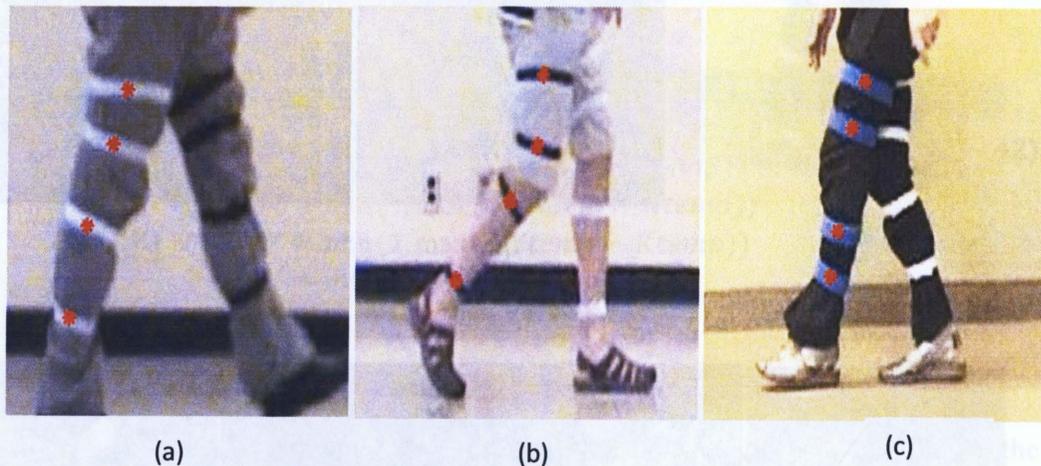
Figure 3.10 Markers positions (a) white markers (b) black markers (c) blue markers

Let M_1, M_2, M_3, M_4 represents the markers from top down, respectively. Let (x_i, y_i) denote the ith marker's centroid pixel position. The hip and knee angles are calculated by:

$$\alpha_{hip} = \tan^{-1}\frac{y_2 - y_1}{x_2 - x_1}, \qquad (3-44)$$

$$\alpha_{knee} = \tan^{-1}\frac{y_4 - y_3}{x_4 - x_3}, \qquad (3-45)$$

The obtained angle signals as a function of time contains lots of noise. The noise may be introduced by illumination changes, measurement errors, quantization errors, etc. A low-pass filter is suitable to remove high frequency noise because low frequency components have the most important signal information. In [53], the authors proved that 99.7% of the signal power lies below the 8th harmonic. A cutoff frequency at 4.8HZ can efficiently remove the high frequency noise. The study of [54] indicates that the frequency content of kinematic data is determined by the body segment dynamics, activity, marker set, and the acquisition system. There are no general rules of thumb to determine the optimal cutoff frequency. The cutoff frequency is selected based on the actual signal data and it is different case by case. To obtain kinematic parameters in this research, the cutoff frequency for hip signal and knee signal was experimentally selected as $\omega_{hc} = 0.3\pi$ and $\omega_{kc} = 0.3\pi$, respectively. The video sequences were filmed at 30 fps. Each frame is one input signal. So, the cutoff frequency at 0.3π is corresponding to signal frequency about 4.5 HZ.

Figure 3.11 shows an example of the hip angle signal and knee angle signal. The original knee angles signal is denoted with green dash line in Fig. 3.11 (a) and the filtered signal is denoted with red solid line in the same figure. Fig. 3.11(c) is the hip angles case. The start point of each step is defined as the time point of 'initial contact' which is the time that the heel initially contacts the ground. It is a suitable assumption that the knee angle reaches the maximum value or minimum value depending on the walking direction. Based on the peak points or valley points of the knee signal, single step is extracted from the continuous signal sequence. The segment points for hip angle are the same time points of the knee angle. These segment points are denoted by red points in Fig. 3.11 (a)

and Fig. 3.11 (c). The knee angle of one step extracted from the sequence in (a) is shown in Fig. 3.11 (b). Fig. 3.11 (d) is the hip angle of this single step.

3.8 Leg Angle Acquisition

Hip angle or knee angle as a gait pattern is a function of time. From these patterns, gait kinematic parameters, such as duration of swing, duration of stance, step length, and stride length, are calculated [16]. Meanwhile, walking is the result of continuously cooperative movement of different muscles and it involves almost whole body muscles. In order to reflect the complexity of this characteristic, the leg integrity can provide more information in terms of walking gait pattern. The leg-angle is defined as the combination of the hip angle and knee angle which is similar to the one in [31]. By defining hip angle as X-axis and knee angle as Y-axis, the curve of leg angle is obtained. This new curve serving as a feature of gait pattern can give more information than single hip or knee pattern. Figure 3.12 shows four people's single step leg angle patterns. Red solid line denotes left leg angle pattern and green dash line denotes right angle pattern. Control group #1 and control group #2 are healthy young males and they do not have any lower extremity disease. The leg angle pattern of patient #1 was obtained before THR surgery. The leg angle pattern of patient #2 was obtained before TKR surgery. Figure 3.12 shows that different people will have different leg angle patterns and the differences between control group's gait patterns and patients' gait patterns are noticeable. The control group has higher symmetry between left leg pattern and right leg pattern. The differences between control group and patients suggest that leg bilateral symmetry has the potential value for rehabilitation progress assessment.


(a) (b)

(c) (d)
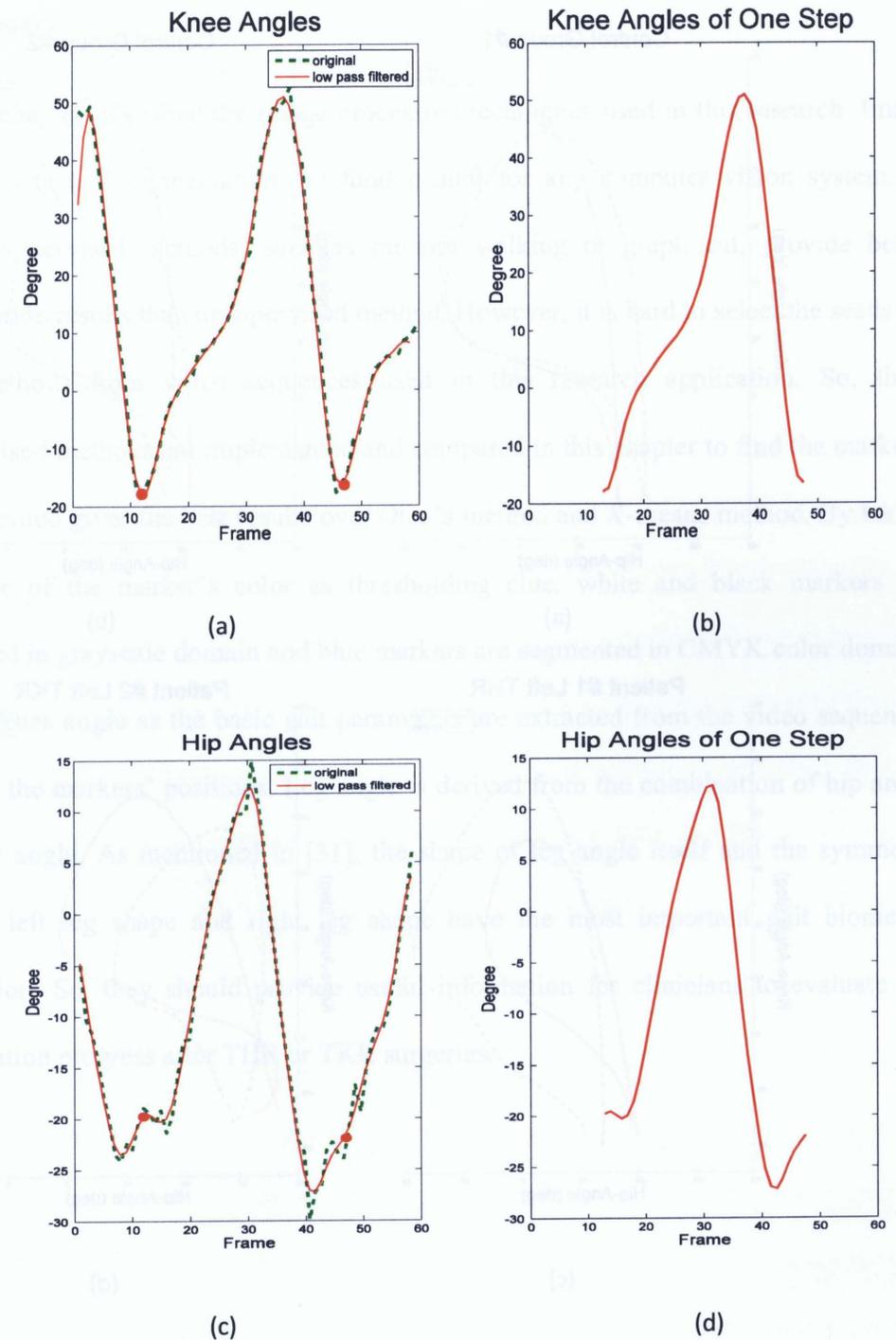
Figure 3.11 Hip angles and Knee angles (a) knee angles (b) knee angles of one single step (c) hip angles (d) hip angles of one single step
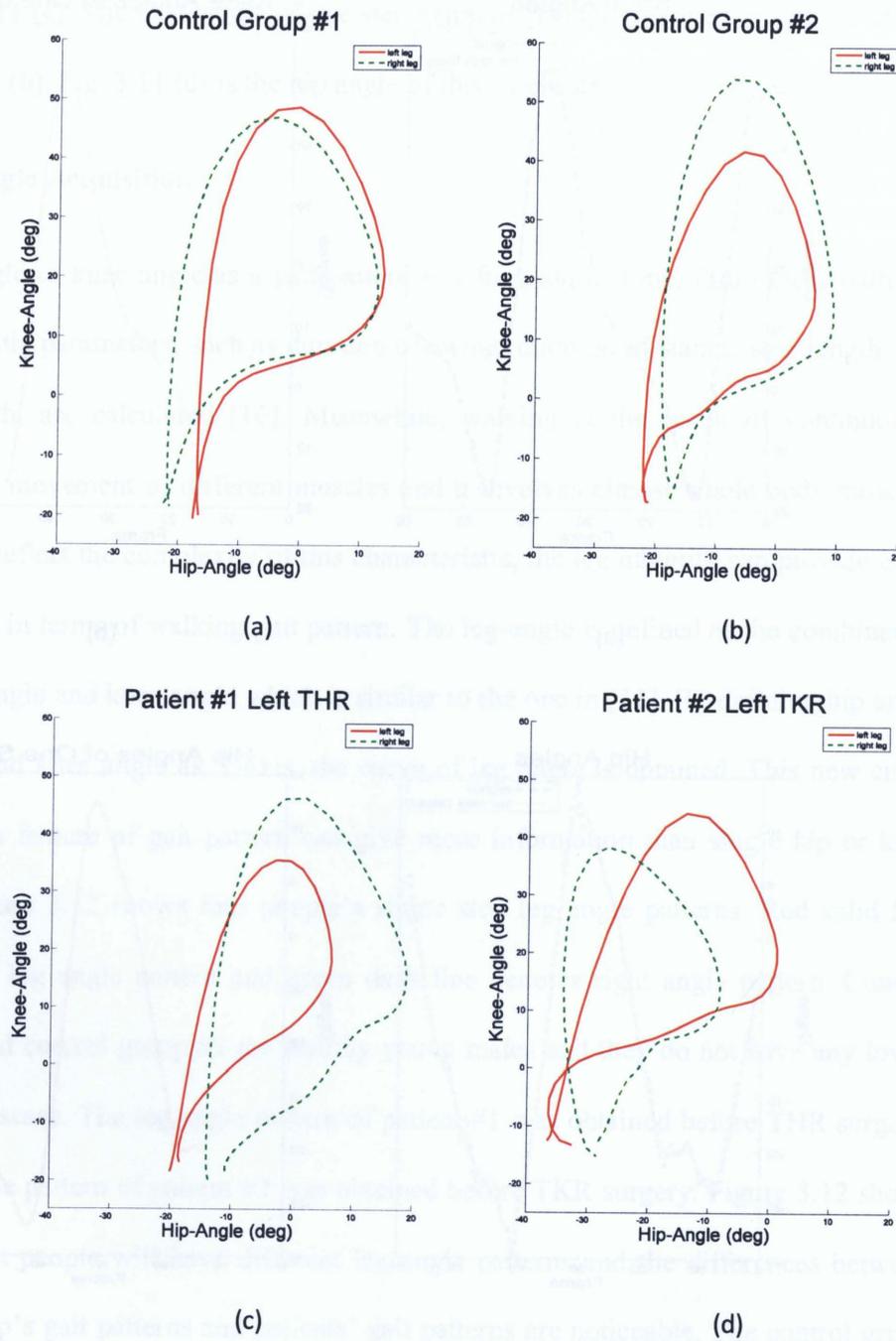
Figure 3.12 Leg angle patterns (a) and (b) examples from control group (c) and (d) examples from patients

3.9 Summary

This chapter described the image processing techniques used in this research. Image preprocessing and segmentation are fundamental for any computer vision system. In general, supervised methods, such as random walking or graph cut, provide better segmentation results than unsupervised method. However, it is hard to select the seeds for these methods from video sequences used in this research application. So, three unsupervised methods are implemented and compared in this chapter to find the markers. GMM method gives the best results over Otsu's method and K-means method. By taking advantage of the marker's color as thresholding clue, white and black markers are segmented in grayscale domain and blue markers are segmented in CMYK color domain. Hip and knee angle as the basic gait parameters are extracted from the video sequences based on the markers' positions. Leg angle is derived from the combination of hip angle and knee angle. As mentioned in [31], the shape of leg angle itself and the symmetry between left leg shape and right leg shape have the most important gait biometric information. So, they should provide useful information for clinicians to evaluate the rehabilitation progress after THR or TKR surgeries.

Chapter 4

Clinical Gait Analysis

Everything in the world shows some levels of symmetry. Symmetry acts a fundamental non-accidental property in computer vision field. It is used for shape representation, shape simplification, characterization, or approximation. In terms of gait based human recognition, symmetry is crucial for model free analysis, especially for silhouette-based approaches. In [26] [55], the authors used the gait symmetry maps extracted from silhouette images to recognize human beings or animals. Instead of using the gait symmetry maps from the video sequences, the bilateral symmetry between left leg and right leg is used in this thesis as an indicator to evaluate the rehabilitation progress after THR or TKR surgeries. The bilateral symmetry itself is also proved to be useful for distinguishing control group subjects' gait patterns and patients' gait patterns.

4.1 Gait Pattern Database

The database is composed with videos from forty patients and eight people in control group. The patients are asked to walk in a straight line for a few steps with their ordinary walking styles. The front view camera and side view camera filmed the subjects with and without markers attached. To evaluate the rehabilitation progress, a patient will be filmed at five scheduled times for this study. One is at the time of pre-admission and four at the time of designated follow up exams which are one month, three months, six months and one year after the surgery. All subjects volunteered to take part in this research. Mt. Sinai Hospital has issued the ethics clearance for this research. All subjects in this research have been fully informed about this research and they have signed the consent forms before the first time recording. All subjects will be required to fill out a brief one page questionnaire regarding any past history that could have had an effect on current gait pattern. The eligible subject should have no neural disease which has effect on gait pattern, such as Parkinson's disease.

The database has 24 patients who had taken THR or TKR surgeries, 16 patients with lower extremity osteoarthritis and eight people as control group who have no known lower extremity diseases. The gait patterns of those patients are treated as abnormal gait patterns. The control group's gait patterns are considered as normal gait patterns. The total abnormal steps are 231. The age of the patients is from 22 to 85, with average age 58.35. The total normal steps are 99. The age span of the control group is 14, from 20 to 34, with average age 27.75.

4.2 Dynamic Time Warping

Dynamic Time Warping (DTW) is an efficient recursive technique to calculate the similarity between two time series. Linear time alignment calculates the similarity between two time series by using the summation of the difference between two corresponding sample points with the same time instant. However, DTW aligns two variable length time sequences with a non-linear warping function. It has been successfully used for gesture recognition [56], signature verification [57] [58] and gait recognition [59].

Given two time series P and Q, where $P = [p_1, p_2, \cdots, p_n]$, and $Q = [q_1, q_2, \cdots, q_m]$. To align P and Q using DTW, a $n \times m$ matrix D is defined as:

$$D(i,j) = \|p_i - q_j\|. \qquad (4-1)$$

where $\|\bullet\|$ is the distance between p_i and q_j. This distance can be any *p-norm*. In this study, Euclidean distance is used which is $p = 2$. A warping path W, $W = [w_1, w_2, \cdots w_K]$, is a contiguous set of matrix elements that define a mapping between P and Q. The kth element of W is $w_k = (i,j)_k$. K is the length of warping path and it satisfies $\max(n,m) \leq K \leq (n+m-1)$. There are some constraints for W.

1. Boundary condition: The start point and endpoint are fixed at $w_1 = (1,1)$ and $w_K = (n,m)$. This makes sure the warping path starts and finishes in diagonally opposite corner cells of the matrix.

2. Monotonicity condition: The warping path should increase monotonically. Given $w_k = (i_1, j_1)$ and $w_{k-1} = (i_2, j_2)$, then $(i_2 - i_1) \geq 0$ and $(j_2 - j_1) \geq 0$.

3. Continuity condition: The allowable steps in the warping path are restricted to adjacent cells. Given $w_k = (i_1, j_1)$ and $w_{k-1} = (i_2, j_2)$, then $(i_2 - i_1) \leq 1$ and $(j_2 - j_1) \leq 1$.

The optimal warping W^* path is the one which minimizes the warping cost:

$$DTW(P,Q) = \min \left\{ \sqrt{\frac{\sum_{k=1}^{K} w_k}{K}} \right\}. \qquad (4-2)$$

To efficiently find this optimal warping path, a cumulative cost matrix C is defined as:

$$C(i,j) = D(i,j) + \min[C(i-1,j-1), C(i-1,j), C(i,j-1)]. \qquad (4-3)$$

where $C(1,1) = D(1,1)$. By using recursive dynamic programming, the minimum optically aligned path between two sequence is obtained as well as the quantized dissimilarity between them.

This is the classical DTW algorithm. There are still some extended versions of it. One is to change the continuity constraints. The continuity constraints are all called local constraints. The continuity constraint described above is illustrated in Figure 4.1(a). Another possible version is illustrated in (b). It also is known as diagonal continuity. With this local constraint, definition in (4-3) will have this expression:

$$C(i,j) = D(i,j) + \min[C(i-1,j-1), C(i-2,j-1), C(i-1,j-2)]. \qquad (4-4)$$
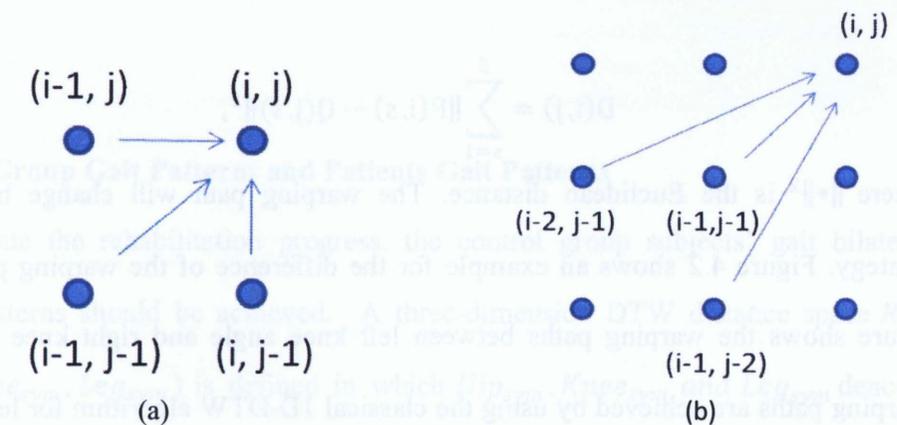
Figure 4.1 Local constraints of DTW (a) adjacent cells continuity (b) diagonal cells continuity

The authors in [57] proposed an extended version which is called Continuous Dynamic Time Warping (CDTW). In this method, a sample point in one sequence is allowed to match a point in between two samples in the other sequence. The value of this in-between point is generated by a linear interpolation model. However, this method does

not show many advantages over the normal DTW method. The authors drew the conclusions that this method has similar results compared with the traditional DTW method for signature verification, but it is computationally intensive and it is three orders of magnitude slower than normal DTW.

Another extended version, multi-dimensional dynamic time warping (MD_DTW) was proposed in [60] [61]. In this method, the distance matrix D was calculated by:

$$D(i,j) = \sum_{s=1}^{S} d(P,Q), \qquad (4-5)$$

where S is the dimension of the data. When calculating the leg bilateral symmetry in this paper, $S = 2$ is used. The equation of $(4-5)$ will be:

$$D(i,j) = \sum_{s=1}^{S} \|P(i,s) - Q(j,s)\|^2, \qquad (4-6)$$

where $\|\bullet\|^2$ is the Euclidean distance. The warping path will change by using this strategy. Figure 4.2 shows an example for the difference of the warping path. The top figure shows the warping paths between left knee angle and right knee angle. These warping paths are achieved by using the classical 1D-DTW algorithm for left knee angle and right knee angle. The bottom figure shows the warping paths between the same knee sequences but the warping paths are achieved from 2D-DTW algorithm for the left leg angle and right leg angle.
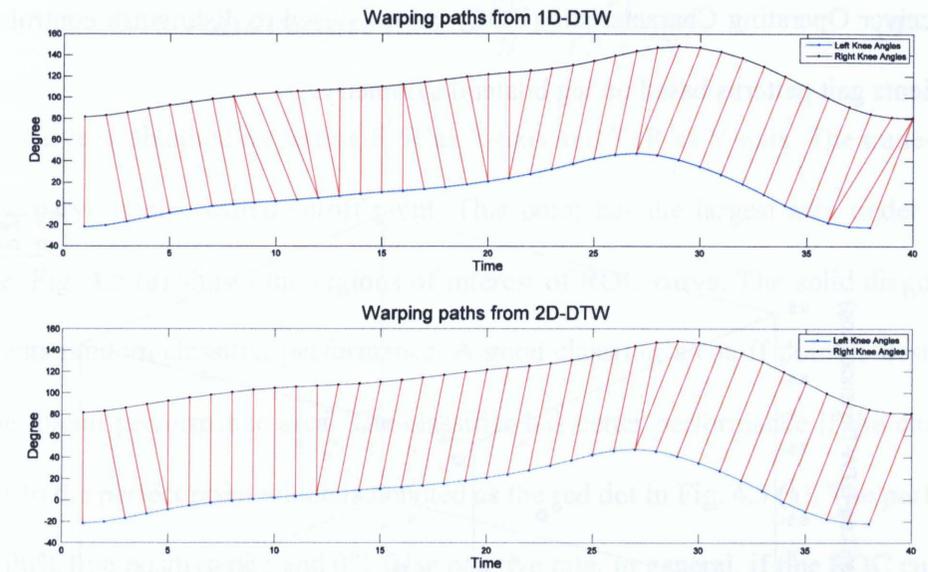
Figure 4.2 Different warping paths from 1D-DTW and 2D-DTW methods

4.3 Control Group Gait Patterns and Patients Gait Patterns

To evaluate the rehabilitation progress, the control group subjects' gait bilateral symmetry patterns should be achieved. A three-dimension DTW distance space $R = (Hip_{sym}, Knee_{sym}, Leg_{sym})$ is defined in which $Hip_{sym}, Knee_{sym}$ and Leg_{sym} denotes the DTW distance between left leg and right leg for hip angle, knee angle, and leg angle, respectively. The median values for each patient on R were calculated and shown in Figure 4.3 as red stars. The bilateral symmetry data from eight control group subjects are used as normal pattern models and they are shown in Figure 4.3 as green circles. The reason of using median values rather than mean values is that it minimize the outlier data which may be noise caused by measurement errors. The control group's gait patterns concentrating on the left lower corner of this figure indicate higher level symmetry.

Receiver Operating Characteristic (ROC) curve is used to distinguish control group and patients gait patterns based on leg bilateral symmetry.
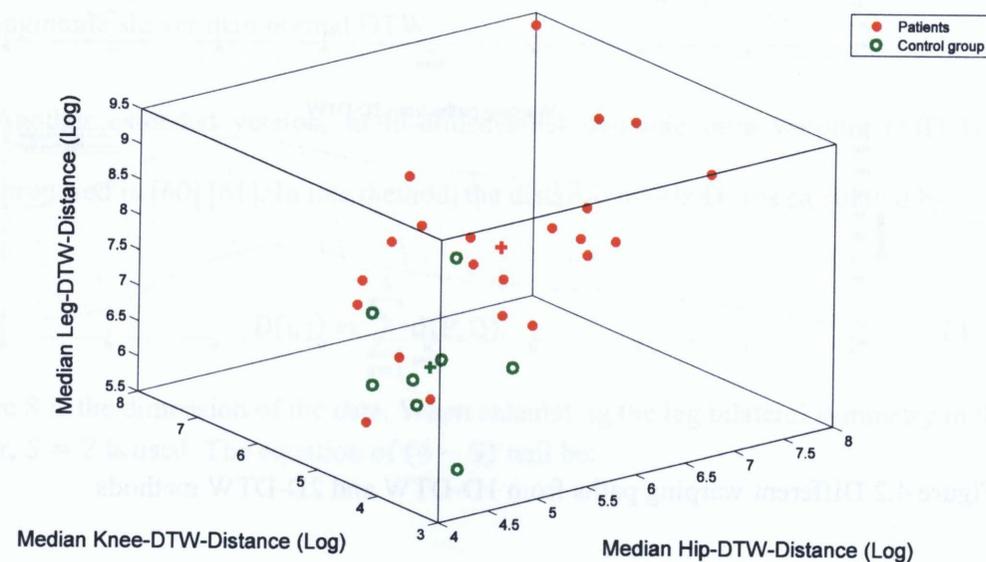

Figure 4.3 Each subject's median symmetry in hip-knee-leg symmetry space

ROC curve is widely used in medical research where it can divide data into two classes to provide information for decision making [62]. It can not only evaluate the performance of a classifier but also select a threshold which can balance the true positive rate and false positive rate. There are four possible outcomes for a classifier: true positive, false negative, true negative and false positive. A confusion matrix in Figure 4.4 is constructed to depict this situation. True Positive Rate (TPR), also called recall, is defined as:

$$True\ Positive\ Rate = \frac{TP}{P} = \frac{TP}{TP + FN}, \qquad (4-7)$$

False Positive Rate (FPR) is defined as:

$$False\ Positive\ Rate = \frac{FP}{N} = \frac{FP}{FP + TN}, \qquad (4-8)$$

A ROC curve is obtained by setting FPR as X-axis and TPR as Y-axis. The trade-off point of the curve is also called cutoff point. This point has the largest area under the ROC curve. Fig. 4.5 (a) shows the regions of interest of ROC curve. The solid diagonal line represents random classifier performance. A good classifier's cutoff point is desired to lie in the liberal performance area. The classifier has better performance if the cutoff point closer to the perfect point which is denoted as the red dot in Fig. 4.5 (a). The perfect point has 100% true positive rate and 0% false positive rate. In general, if one ROC curve C_1 is above and to the left of another ROC curve C_2, the classifier of C_1 has better performance than the one of C_2. Other regions of interest are conservative performance area and worse than random performance area. This is a brief description of ROC method. More details are provided in [63] [64].
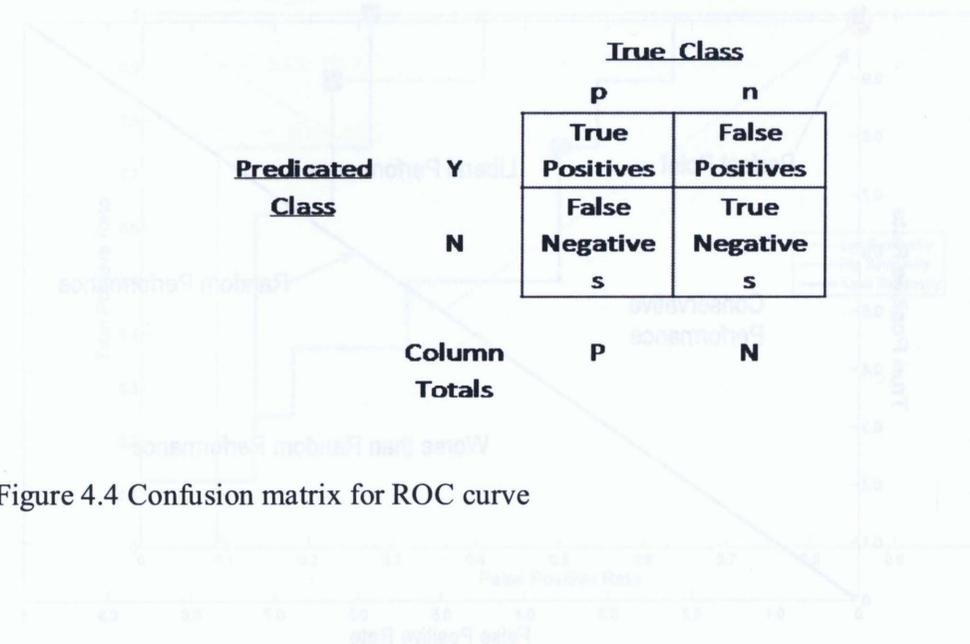

Figure 4.4 Confusion matrix for ROC curve

Figure 4.5(b) shows the result by using ROC curve to distinguish control group's gait patterns and patients' gait patterns. In this figure, each data point is the DTW distance of a single step's symmetry from the database. The box on each curve indicates the trade-off point. Figure 4.6 shows the same ROC curve where each data point represents the median value of each subject's symmetry. The ROC curve shows that leg bilateral symmetry as a classifier has better performance than hip or knee only. Compared to leg bilateral symmetry, knee bilateral symmetry itself is a poor indicator to evaluate the rehabilitation progress and this also means THR or TKR surgery has less effect on the knee symmetry than it does on the hip symmetry. The cutoff point of the leg symmetry is 7.46 in this database. By using this value as the threshold to distinguish control group gait patterns and patients gait patterns, 87.5% control group subjects and 77.2% patients are correctly reported.

(a)

56
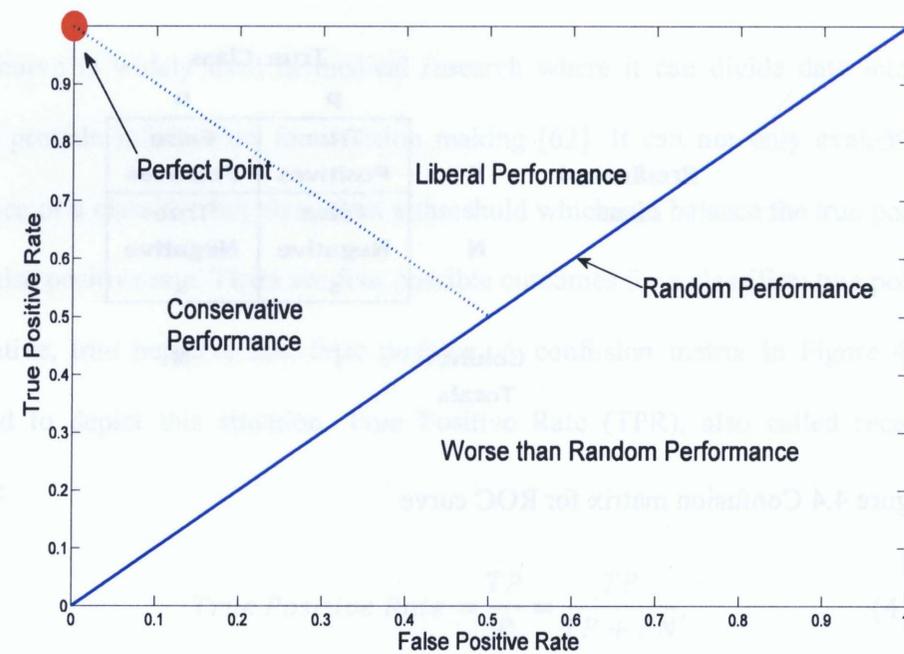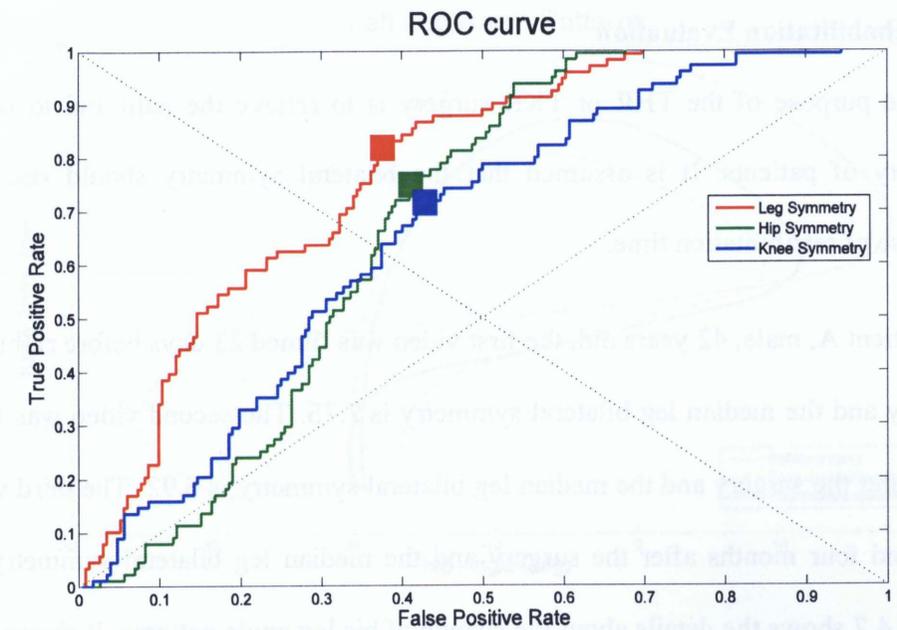
(b)

Figure 4.5 ROC curve (a) regions of interest (b) ROC curve by using symmetry values of all steps
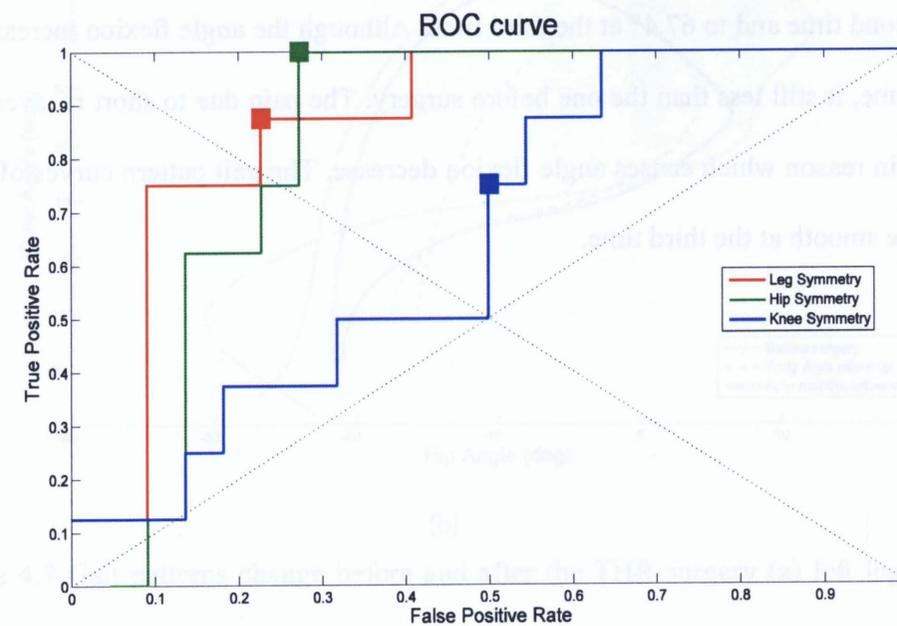
Figure 4.6 ROC curve by using median symmetry value of each subject

57

4.4 Rehabilitation Evaluation

The purpose of the THR or TKR surgery is to relieve the pain and to restore the mobility of patients. It is assumed that the bilateral symmetry should rise with the increase of rehabilitation time.

Patient A, male, 42 years old, the first video was filmed 23 days before right leg THR surgery and the median leg bilateral symmetry is 7.75. The second video was filmed 40 days after the surgery and the median leg bilateral symmetry is 6.92. The third video was recorded four months after the surgery and the median leg bilateral symmetry is 6.36. Figure 4.7 shows the details about the change of his leg angle patterns. It shows that both the angle flexion and gait patterns changed dramatically. For right leg, the range of hip angles changes from 43.1° at the first time to 21.8° at the second time and to 28.3° at the third time. The range of right knee angles changes from 82.5° at the first time to 52.0° at the second time and to 67.4° at the third time. Although the angle flexion increased at the third time, it still less than the one before surgery. The pain due to short recovery time is the main reason which causes angle flexion decrease. The gait pattern curves of right leg become smooth at the third time.
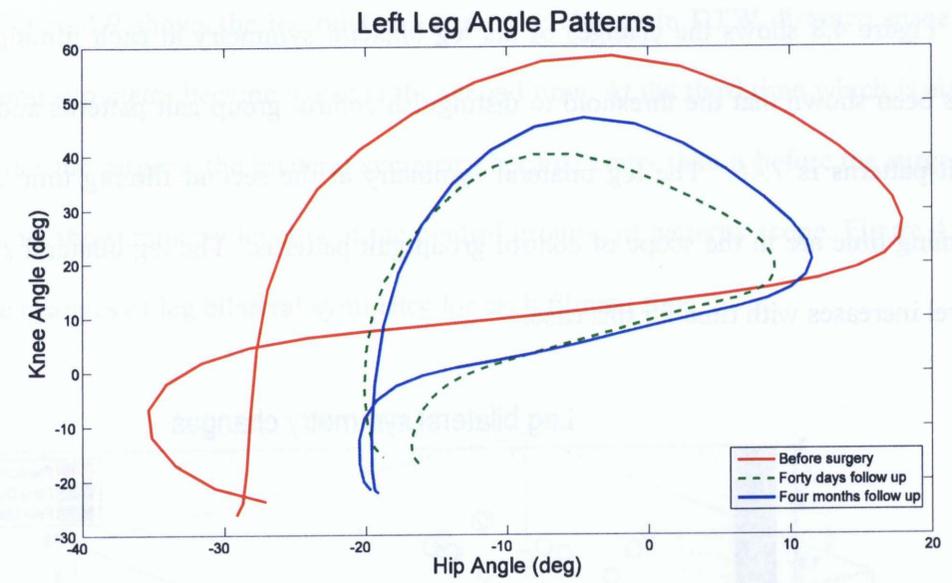
(a)
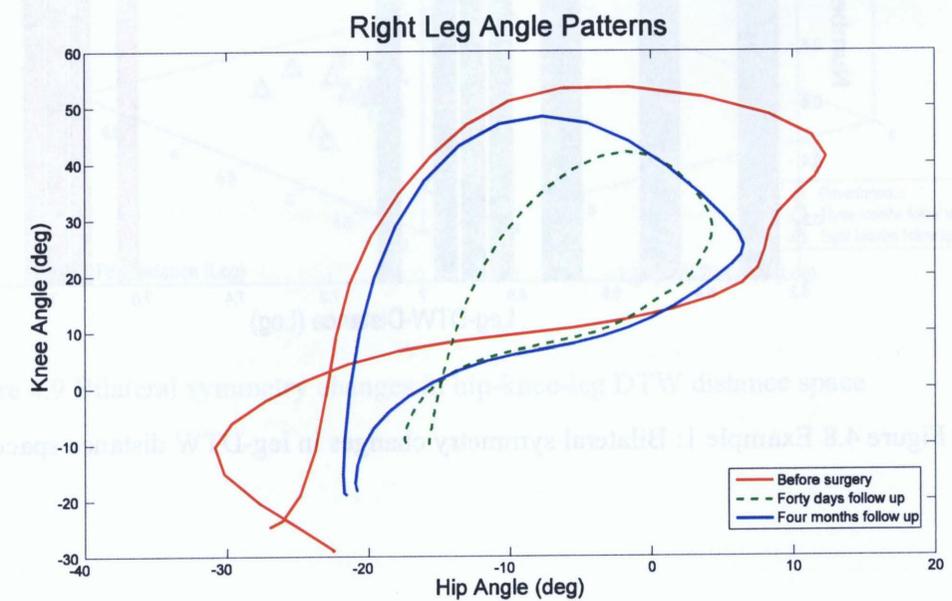
(b)

Figure 4.7 Gait patterns change before and after the THR surgery (a) left leg angle patterns (b) right leg angle patterns

Figure 4.8 shows the changes of his leg bilateral symmetry at each filming time. It has been shown that the threshold to distinguish control group gait patterns and patients gait patterns is 7.46. The leg bilateral symmetry at the second filming time and third filming time are in the scope of control group gait patterns. The leg bilateral symmetry level increases with time for this case.
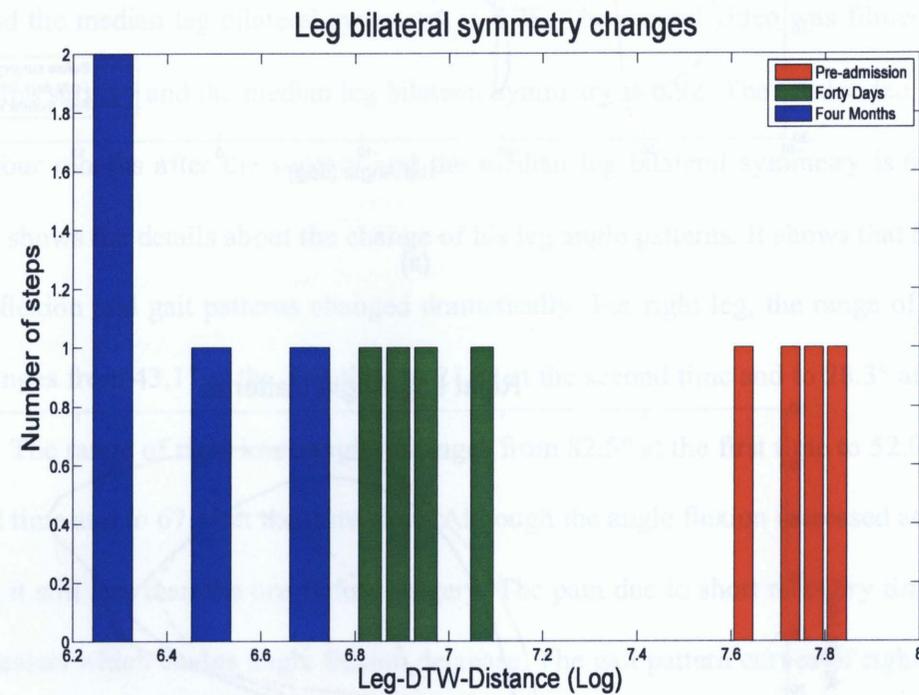
Figure 4.8 Example 1: Bilateral symmetry changes in leg-DTW distance space

Patient B, female, 44 years old, the first video was filmed 12 days before left leg THR surgery and the median leg bilateral symmetry is 6.96. The second video was filmed three months after the surgery and the median leg bilateral symmetry is 8.79. The third video was recorded eight months after the surgery and the median leg bilateral symmetry

is 6.16. Figure 4.9 shows the leg bilateral symmetry change in DTW distance space R. The bilateral symmetry became worse at the second time. At the third time which is eight months after the surgery, the bilateral symmetry became better than it before the surgery. At this time, the symmetry level is in the control group gait patterns scope. Figure 4.10 shows the changes of leg bilateral symmetry for each filming time.
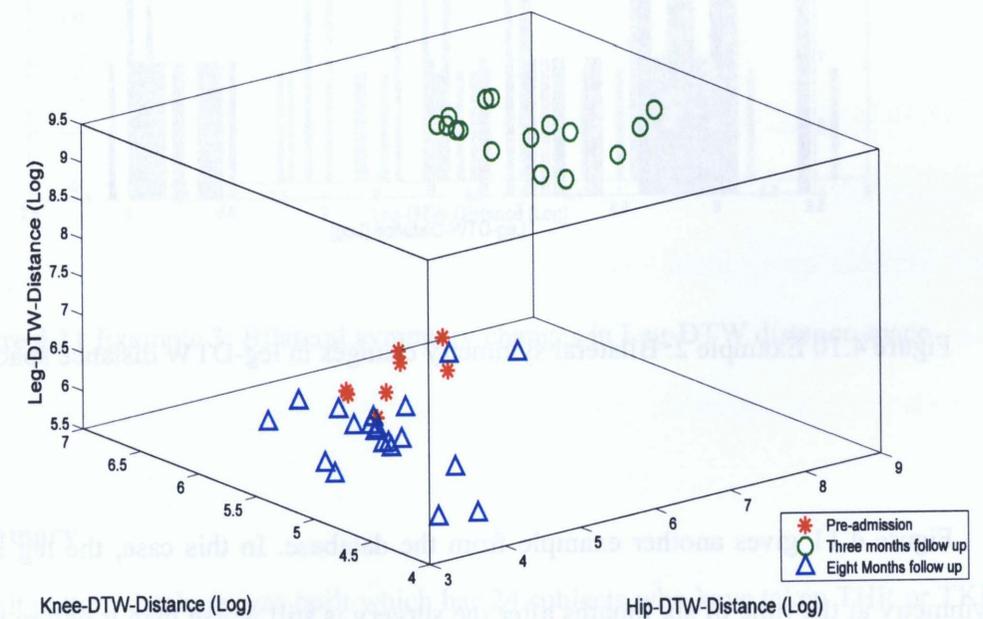
Figure 4.9 Bilateral symmetry changes in hip-knee-leg DTW distance space
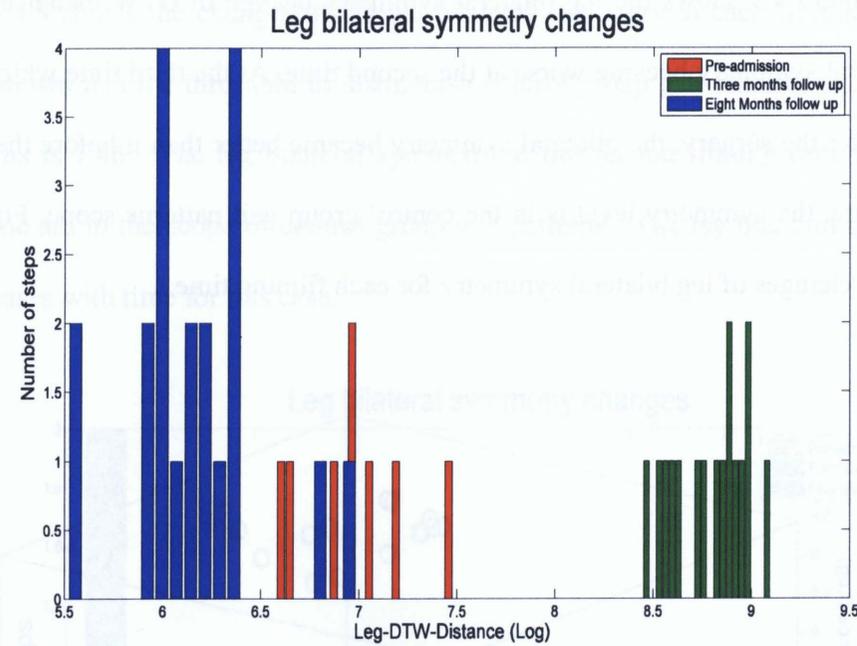
Figure 4.10 Example 2: Bilateral symmetry changes in leg-DTW distance space

Figure 4.11 gives another example from the database. In this case, the leg bilateral symmetry at the time of six months after the surgery is still worse than it before surgery. The median value of leg bilateral symmetry at the second time filming is 8.26. The data for the first follow up exam was missing because the patient was not available for filming at that time due to walking aid used. The using of walking aids, such as cane or walker, is the main reason that most patients cannot be filmed at their first follow up exam which is one month after the surgeries.

From above examples, the leg bilateral symmetry changes dramatically before and after surgeries. These examples show the effectiveness of leg bilateral symmetry as an indicator for rehabilitation progress evaluation.
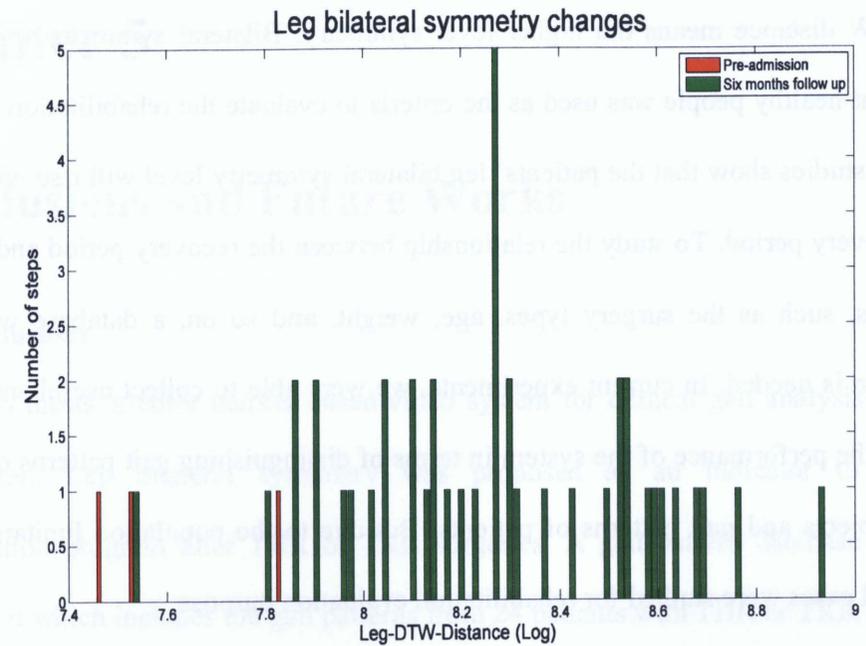
Figure 4.11 Example 3: Bilateral symmetry changes in Leg-DTW distance space

4.5 Summary

A gait pattern database was built which has 24 subjects who have taken THR or TKR surgeries, 16 subjects with lower extremity arthritis, and eight control group subjects who have no known gait related diseases. In terms of the scheduled follow up examination time, the setup of database is a long term project and the uncertain appointment time makes it difficult to record the patients at each follow up examination time. This is the main reason that there are only six datasets which can be used for rehabilitation evaluation in current database.

For rehabilitation progress evaluation, the leg bilateral symmetry is represented by the MD_DTW distance between left leg angle shape and right leg angle shape. A smaller

MD_DTW distance means the higher level symmetry. Bilateral symmetry information from eight healthy people was used as the criteria to evaluate the rehabilitation progress. The case studies show that the patients' leg bilateral symmetry level will rise with a long term recovery period. To study the relationship between the recovery period and medical conditions, such as the surgery types, age, weight, and so on, a database with large population is needed. In current experiments, we were able to collect useful statistics to evaluate the performance of the system in terms of distinguishing gait patterns of control group subjects and gait patterns of patients. But due to the population limitation, only individual cases were studied for rehabilitation evaluation purpose.

Chapter 5

Conclusions and Future Works

5.1 Conclusions

In this thesis, a color marker based video system for clinical gait analysis has been investigated. Leg bilateral symmetry was proposed as an indicator to evaluate rehabilitation progress after THR or TKR surgeries. A gait pattern database has been established which includes the gait patterns from 24 patients with THR or TKR surgeries, 16 patients with lower extremity osteoarthritis, and eight control group subjects. The experimental results show that the proposed color markers based video system can provide quantitative information of gait patterns for clinicians to evaluate the rehabilitation progress.

Background subtraction methods and thresholding techniques for image segmentation were reviewed. In this study, the video recording was performed in the indoor clinical environment and it is relatively static. Median filter algorithm provides good results for background extraction from the video sequences. Three methods, Otsu's method, K-means clustering method, and mixture of Gaussian model, have been compared and discussed in detail. In this particular application environment, GMM model gives the best results for color marker detection.

The symmetry between left leg and right leg in terms of hip, knee and leg were studied. The locomotion of hip and knee was extracted from the video sequences. The

locomotion of leg was derived from the movement of hip and knee. To quantitatively describe the symmetry level, DTW algorithm was reviewed and implemented to calculate the bilateral symmetry. The changes of leg bilateral symmetry have been proved to be a good indicator for rehabilitation progress evaluation. Normal gait bilateral symmetry patterns were established from the control group subjects. The rehabilitation case studies show significant changes of the leg bilateral symmetry before and after THR or TKR surgeries. The bilateral symmetry level rises with time. Moreover, the gait patterns of control group and gait patterns of patients can be distinguished by using the proposed bilateral symmetry feature. By using cutoff point of leg bilateral symmetry ROC curve as a threshold, 87.5% control group subjects and 77.2% patients can be correctly distinguished.

5.2 Future Works

The most likely following research by using this proposed system is to evaluate functional improvement in patients following different types of joint replacement surgeries. There are two types of joint replacement surgeries: conventional joint arthroplasty and minimally invasive surgery (MIS). By using this video clinical gait analysis system, gait kinematic information and rehabilitation progress can be used to quantitatively assess this two types of surgeries.

So far, there is no markerless video system available for clinical gait analysis. One reason is that clinical gait analysis needs accurate gait information which cannot be provide by most silhouette based video system. Current silhouette based video systems are more likely to be used for human recognition or tracking. However, clinical gait analysis is taking place in a well arranged environment where all necessary elements for human model construction are known. By taking this advantage, model based video gait analysis system has the potential to be used for clinical gait analysis, especially for rehabilitation evaluation.

As mentioned at the beginning of this thesis, gait as a biometrics can be used for human recognition. This thesis also shows that gait symmetry can be used to distinguish walking patterns of control group subjects and walking patterns of patients. This character has the potential to be used for health care purpose. In depth study is need to find the relation between gait patterns and certain diseases. Another possible further study about the relationship between gait patterns and emotions can provide valuable information for surveillance system, especially with 3D gait patterns.

Bibliography

1. Aristotle, On the Motion of Animals, B. C. 350, http://classics.mit.edu/Aristotle/motion_animals.html

2. J. E. Cutting and L. T. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues", Bulletin of the Psychonomic Society, Vol. 9 (5), pp. 353-356, 1977

3. W. Wolff, "Expression of personality: experimental depth psychology", http://library.du.ac.in/dspace/handle/1/1332

4. D. Xu, et al, "Human Gait Recognition with Matrix Representation", IEEE Transaction on Circuits and Systems for Video Technology, Vol. 16, pp. 896-902, 2006

5. C. H. Chen, et al, "Factorial HMM and Parallel HMM for Gait Recognition", IEEE Transaction on Systems, Man, and Cybernetics-Part C: Applications and Reviews, Vol. 39, pp. 114-123, 2009

6. M. S. Nixon and J. N. Carter, "Automatic Recognition by Gait", Proceedings of the IEEE, Vol. 94, pp. 2013-2024, 2006

7. H. Su, F. G. Huang, "Human Gait Recognition based on Motion Analysis", Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, pp. 4464-4468, 2005

8. H. Y. Zhou, P. R. Green, and A. M. Wallace, "Efficient Motion Tracking Using Gait Analysis", Proceedings of Acoustics, Speech, and Signal Processing, Vol. 3, pp. 601-604, 2004

9. X. L. Li, et al, "Gait Components and Their Application to Gender Recognition", IEEE Transaction on Systems, Man, and Cybernetics—Part C: Applications and Reviews, Vol. 38, pp. 145-155, 2008

10. C. F. Shan, Sh. G. Gong and P. W. McOwan, "Learning Gender from Human Gait and Faces", IEEE conference on Advanced Video and Signal Based Surveillance, pp. 505-510, 2007

11. S. S. Ha, et al, "Natural Gait Generation of Biped Robot based on Analysis of Human's Gait", International Conference on Smart Manufacturing Application, pp. 30-34, 2008

12. R. B. Davis, III, et al, "Clinical Gait Analysis and Its Role in Treatment Decision-Making", Medscape General Medicine 1(1), 1999, http://www.medscape.com/viewarticle/440148

13. M. P. Muarray, A. B. Drought, and R. C. Kory, "Walking Patterns of Normal Men", Journal of Bone and Joint Surgery, Vol. 46-A, No. 2, pp. 335-360, 1964

14. M. P. Murrary, "Gait as a total pattern of movement", American Journal of Physical Medicine, Vol.46, No. 1, pp. 290-332, 1967

15. B. J. Fregly, et al, "Design of Patient-Specific Gait Modifications for Knee Osteoarthritis Rehabilitation", IEEE Transactions on Biomedical Engineering, Vol. 54, pp. 1687-1695, 2007

16. S. M. Isaac, et al, "Does arthroplasty type influence knee joint proprioception? A longitudinal prospective study comparing total and unicompartmental arthroplasty", The Knee, Vol. 14, Issue 3, pp. 212-217, 2007

17. H. Kuno, *et al*, "Geometrical analysis of hip and knee joint mobility in cerebral palsied children", Gait & Posture, Vol. 8, Issue 2. Pp. 110-116, 1998

18. M. S. Nixon, Tieniu N. Tan, and Rama Chellappa, "Human Identification based on Gait", Springer, 2005

19. J. Perry, "Gait Analysis: Normal and pathological function", SLACK Incorporated, 1992

20. W. Ellis, S. Kishner, and Jams Monroe Laborde, "Gait Analysis After Amputation", eMedicine, http://emedicine.medscape.com/article/1237638-overview

21. A. Gross, "Report of the Total Hip and Knee Joint Replacement Expert Panel", 2005, http://www.health.gov.on.ca

22. The Oxford Orthopaedic Scores, http://phi.uhce.ox.ac.uk/ox_scores.php

23. Z. H. Zhou, A. P.-B., and R. I. Damper, "A Bayesian Framework for Extracting Human Gait Using Strong Prior Knowledge", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, pp. 1738-1752, 2006

24. A. Kale *et al*, "Identification of Humans Using Gait", IEEE Transactions on Image Processing, Vol. 13, pp. 1163-1173, 2004

25. H. Su, F. G. Huang, "Human Gait Recognition based on Motion Analysis", Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, pp. 4464-4468, 2005

26. J. B. Hayfron-Acquah, M. S. Nixon, and J. N. Carter, "Automatic Gait Recognition by Symmetry Analysis", Audio- and Video-Based Biometric Person Authentication, Springer Berlin/Heidelberg, pp. 272-277, 2001

27. C. BenAbdelkader *et al*, "Eigen Gait: Motin-based Recognition of People using Image Self-Similarity", Audio- and Video-Based Biometric Person Authentication, Springer Berlin/Heidelberg, pp. 284-294, 2001

28. L. Wang *et al*, "Automatic Gait Recognition Based on Statistical Shape Analysis", IEEE Transactions on Image Processing, Vol. 12, pp. 1120-1131, 2003

29. Ch. Y. Yam, M. S. Nixon, J. N. Carter, "Automated person recognition by walking and running via model-based approaches", Pattern Recognition, Vol. 37, Issue 5, pp. 1057-1072, 2004

30. R. D. Greeen, and L. Guan, "Quantifying and Recognizing Human Movement Patterns From Monocular Video Images-Part I: A New Framework for Modeling Human Motion", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 2, pp. 179-190, 2004

31. R. D. Greeen, and L. Guan, "Quantifying and Recognizing Human Movement Patterns From Monocular Video Images-Part II: Applications to Biometrics", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 14, No. 2, pp. 191-198, 2004

32. M. Du, and L. Guan, "Monocular Human Motion Tracking with the DE-MC Particle Filter", IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 2, pp. II-205-II-208, Toulouse, May, 2006

33. H. Y. Liu, and R. Chellappa, "Markerless Monocular Tracking of Articulated Human Motion", IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 1, pp. I-693-I-696, April, 2007

34. R. F. Weir, and D. S. Childress, "Portable Devices for the Clinical Measurement of Gait Performance and Outcomes", Proceedings of the 22nd Annual International Conference of IEEE Engineering in Medicine and Biology Society, Vol. 3, pp. 1873-1875, Chicago, July 2000

35. H. Lee, L. Guan, and I. Lee, "Video analysis of human gait and posture to determine neurological disorders", EURASIP Journal on Image and Video Processing, Vol. 2008, Article ID 380867, 12 pages, 2008

36. http://www.univie.ac.at/CGA/faq/

37. M. S. Nixon, J.N. Carter, D. Cunado, P. S. Huang and S. V. Stevenage, "Automatic Gait Recognition", IEE colloquium on Motion Analysis and Tracking (Ref. No. 1999/103), pp. 3/1-3/6, May 1999

38. E. B. Bellers, and G. de Haan, " De-interlacing: A Key Technology for Scan Rate Conversion", North Holland, 2000

39. S.-Ch. S. Cheung and Ch. Kamath, "Robust techniques for background subtraction in urban traffic video", Visual Communications and Image Processing Vol. 5308, No. 1, pp. 881-892, January 2004

40. R. Cucchiara, M. Piccardi, and A. Prati, "Detecting Moving Objects, Ghosts, and Shadows in Video Steams", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 10, page 1337-1342, October 2003

41. R. Cutler and L. Davis, "View-based Detection and Analysis of Periodic Motion", Proceedings Fourteenth International Conference on Pattern Recognition, Vol. 1, pp.495-500, Brisbane, August 1998

42. Q. Zang, and R. Klette, "Robust Background Subtraction and Maintenance", Proceedings of the 17th international conference on Pattern Recognition, Vol. 2, pp. 90-93, August 2004

43. R. C. Gonzalez, and R. E. Woods, "Digital Image Processing", 2nd version,

44. M. Sezgin, and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", Journal of Electronic Imaging, Vol. 13, pp. 146-165, January 2004

45. N. Otsu, "A Threshold Selection Method from Gray-Level Histograms", IEEE Transactions on Systems, Man and Cybernetics, Vol. 9, Issue 1, pp. 62-66, January 1979

46. J. MacQueen, "Some methods for classification and analysis of multivariate observations", Proceedings of Fifth Berkeley Symposium On Mathematical Statistics and Probability, Vol.1, pp.281-297, 1967

47. I. Kompatsiaris, E. Triantafillou, and M. G. Strntzis, "Spatiotemporal segmentation and tracking of objects for visualization of videoconference image sequences", IEEE Transactions on Circuits and Systems for Video Technology, Vol.10, Issue 8, pp. 1388-1402, December 2000

48. N. Friedman and S. Russell, "Image segmentation in video sequences: A probabilistic approach", Proceedings of the Thirteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-97), pp. 175-181, 1997.

49. J. A. Bilmes, "A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models", U.C. Berkely, TR-97-021, April 1998

50. Ch. Stauffer, and W. Eric L. Grimson, "Learning Patterns of Activity Using Real-Time Tracking", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp. 747-757, August 2000

51. D.-Sh. Lee, "Effective Gaussian Mixture Learning for Video Background Subtraction", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 5, pp. 827-832, May 2005

52. R. Zhang, "Data Clustering and Pattern Recognition", http://neural.cs.nthu.edu.tw/jang/books/dcpr/

53. C. Angeloni, P. O. Riley, and David E. Krebs, "Frequency Content of Whole Body Gait Kinematic Data", IEEE Transactions on Rehabilitation Engineering, Vol. 2, Issue. 1, pp. 40-46, March 1994

54. D. A. Winter, H. G. Sidwall, and D. A. Hobson, "Measurement and Reduction of Noise in Kinematics of Locomotion", Journal of Biomechanics, Vol. 7, Issue 2, pp. 157-159, March 1974

55. J. B. Hayfron-Acquah, M. S. Nixon and J. N. Carter, "Recognising Human and Animal Movement by Symmetry", International Conference on Image Processing, Vol. 3, pp. 290-293, 2001

56. A. Kuzmanic, and V. Zanchi, " Hand shape classification using DTW and LCSS as similarity measures for vision-based gesture recognition system", EUROCON, The International Conference on "Computer as a Tool", pp. 264-269, September 2007

57. M. E. Munich, and P. Perona, "Continuous Dynamic Time Warping for translation-invariant curve alignment with applications to signature verification", Proceedings of 7th International Conference on Computer Vision, Vol. 1, pp. 108-115, Korfu, September 1999

58. M. T. Ibrahim, M. Kyan, and L. Guan, "On-line signature verification using most discriminating features and Fisher linear discriminant analysis (FED)," Proceedings of IEEE International Symposium on Multimedia, pp. 172-177, San Diego, December 2008

59. J. Q. Wang, Y. Makihara, and Y. Yagi, " Human Tracking and Segmentation Supported by Silhouette-based Gait Recognition", IEEE International Conference on Robotics and Automation, pp. 1698-1703, Pasadena, May 2008

60. M. Vlachos, M. Hadjieleftheriou, D. Gunopulos, and E. Keogh, "Indexing Multi-Dimensional Time-Series with Support for Multiple Distance Measures", Proceeding of the 9th ACM SIGKDD International conference on Knowledge discovery and data mining, pp. 216-225, Washington, 2003

61. G.A.ten Holt, M. J. T. Reinders, and E. A. Hendriks, "Multi-Dimensional Dynamic Time Warping for Gesture Recognition", http://ict.ewi.tudelft.nl/pub/gineke/DTW-vASCI.pdf

62. M. H. Zweig, and G. Campbell, "Receiver-Operating Characteristic (ROC) Plots: A Fundamental Evaluation Tool in Clinical Medicine", Clinical Chemistry, 39/4, pp. 561-577, August 1993

63. T. Fawcett, "ROC Graphs: Notes and Practical Considerations for Researchers", http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.10.9777

64. L. Hamel, "Model Assessment with ROC Curve", The Encyclopedia of Data Warehousing and Mining, 2nd Edition, Idea Group Publishers, 2008

Publication

- Ying Bo Xu, Chun Hao Wang, Paul Zalzal, Oleg Safir, and Ling Guan, "Analysis of Human Gait Bilateral Symmetry for Functional Assessment after an Orthopaedic Surgery", Image Analysis and Recognition, LNCS, pp. 627-636, Halifax, July 2009