Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN BIOINGEGNERIA

Ciclo XXIII - ING-INF/06

ANALYSIS OF HUMAN BODY KINEMATICS USING A HYBRID MARKERLESS VIDEO ACQUISITION AND PROCESSING

ELIF SURER

Coordinatore Dottorato	Relatore
Prof. ANGELO CAPPELLO	**Prof. UGO DELLA CROCE**

Esame finale anno 2011

Alma Mater Studiorum – Università di Bologna

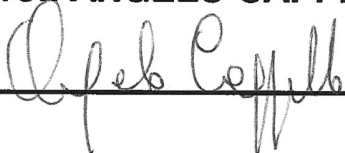## DOTTORATO DI RICERCA IN

BIOINGEGNERIA

Ciclo XXIII

Settore scientifico-disciplinare di afferenza: ING-INF/06

ANALYSIS OF HUMAN BODY KINEMATICS USING A HYBRID MARKERLESS VIDEO ACQUISITION AND PROCESSING

Presentata da: ELIF SURER

Coordinatore Dottorato

Prof. ANGELO CAPPELLO

Relatore

Prof. UGO DELLA CROCE

Esame finale anno 2011

To the beloved memory of Dr. Ali Vahit Şahiner

To my grandparents

SUMMARY

A main objective of the human movement analysis is the quantitative description of joint kinematics and kinetics. This information may have great possibility to address clinical problems both in orthopaedics and motor rehabilitation. Previous studies have shown that the assessment of kinematics and kinetics from stereophotogrammetric data necessitates a setup phase, special equipment and expertise to operate. Besides, this procedure may cause feeling of uneasiness on the subjects and may hinder with their walking. The general aim of this thesis is the implementation and evaluation of new 2D markerless techniques, in order to contribute to the development of an alternative technique to the traditional stereophotogrammetric techniques.

At first, the focus of the study has been the estimation of the ankle-foot complex kinematics during stance phase of the gait. Two particular cases were considered: subjects barefoot and subjects wearing ankle socks. The use of socks was investigated in view of the development of the hybrid method proposed in this work. Different algorithms were analyzed, evaluated and implemented in order to have a 2D markerless solution to estimate the kinematics for both cases. The validation of the proposed technique was done with a traditional stereophotogrammetric system. The implementation of the technique leads towards an easy to configure (and more comfortable for the subject) alternative to the traditional stereophotogrammetric system.

Then, the abovementioned technique has been improved so that the measurement of knee flexion/extension could be done with a 2D markerless technique. The main changes on the implementation were on occlusion handling and background segmentation. With the additional constraints, the proposed technique was applied to the estimation of knee flexion/extension and compared with a traditional stereophotogrammetric system. Results showed that the knee flexion/extension estimation from traditional stereophotogrammetric system and the proposed markerless system were highly comparable, making the latter a potential alternative for clinical use.

A contribution has also been given in the estimation of lower limb kinematics of the children with cerebral palsy (CP). For this purpose, a hybrid technique, which uses high-cut underwear and ankle socks as "segmental markers" in combination with a markerless methodology, was proposed. The proposed hybrid technique is different than the abovementioned markerless technique in terms of the algorithm chosen. Results showed that the proposed hybrid technique can become a simple and low-cost alternative to the traditional stereophotogrammetric systems.

SOMMARIO

Uno dei principali obiettivi dell'analisi del movimento umano è la descrizione quantitativa della cinematica e della dinamica delle articolazioni. Questa informazione può avere grandi potenzialità nell'individuazione di approcci clinici sia in ortopedia che in riabilitazione motoria. Alcuni studi hanno mostrato che la stima della cinematica e dinamica da dati stereofotogrammetrici richiede una fase di preparazione, della strumentazione molto specifica e delle competenze per utilizzarla. Inoltre tali procedure possono creare nei pazienti un senso di impaccio e possono modificarne il cammino naturale. Lo scopo generale di questa tesi è l'implementazione e la valutazione di alcune nuove tecniche markerless in due dimensioni, come contributo allo sviluppo di tecniche alternative alle tradizionali tecniche stereo fotogrammetriche.

Inizialmente, lo studio è stato focalizzato sulla stima della cinematica del complesso caviglia-piede durante la fase di appoggio del cammino. Sono stati considerati due casi particolari: soggetti scalzi e soggetti che indossano calzini sportivi alla caviglia. L'uso dei calzini è stato analizzato in previsione degli studi successivi che richiedono l'uso di marker segmentali. Sono stati analizzati diversi algoritmi, valutati e implementati per avere una soluzione markerless in due dimensioni per la stima della cinematica in entrambi i casi. La validazione della tecnica proposta è stata svolta con un sistema stereofotogrammetrico tradizionale. L'implementazione della tecnica si muove verso un'alternativa ai tradizionali sistemi stereo fotogrammetrici, che sia di facile configurazione e meglio accettata dal paziente.

La tecnica sviluppata è stata poi migliorata in modo che la stima della flesso/estensione del ginocchio potesse essere svolta con la tecnica markerless a due dimensioni. Le modifiche principali nell'implementazione hanno riguardato la gestione delle occlusioni e la segmentazione dello sfondo. Con l'aggiunta di altri vincoli, la tecnica proposta è stata applicata alla stima della flesso/estensione del ginocchio e confrontata con un sistema stereo fotogrammetrico tradizionale. I risultati hanno mostrato che la stima della flesso/estensione ottenuta con un sistema stereo fotogrammetrico.

E' stato anche sviluppato un contributo per la stima della cinematica dell'arto inferiore durante il cammino di bambini con paralisi cerebrale infantile. In questo caso, è stata sviluppata una tecnica markerless ibrida che utilizza la diversa colorazione della maglieria intima e dei calzini per identificare dei marker segmentali. Gli algoritmi utilizzati in quest'ultima applicazione sono diversi dai precedenti. I risultati hanno mostrato che la tecnica ibrida proposta può diventare una alternativa ai tradizionali sistemi stereo fotogrammetrici di semplice uso e a basso costo.

GLOSSARY OF TERMS

The following nomenclature is used throughout the thesis:

AF: anatomical frame

AL: anatomical landmark

CA: calcaneous

CAST: calibrated anatomical system technique

CCD: charged-coupled device

CP: cerebral palsy

CSG: constructive solid geometry

CSP: colored surface points

DLT: direct linear transformation

DOF: degree of freedom

GF: global frame

GLT: Gauss–Laguerre transform

HF: head of fibula

Hyb: hybrid

LE: lateral epicondyle

LED: light emitting diode

LM: lateral malleolus

MAT: medial axis transform

Mb: marker-based

ME: medial epicondyle

Ml: markerless

MoG: mixture of gaussians

NSS: nonlinear spherical shells

OBE: oriented bounding ellipsoid

RI: reference image

RMSD: root mean square deviation

$RMSD_v$: intra-subject variability

ROI: region of interest

SMAC: simultaneous multi-frame analytical calibration

SPM: scaled prismatic models

STS: sit-to-stand

TF: technical frame

TOE: big toe

VH: visual hull

VI: visual intersection

VM: fifth metatarsal head

INTRODUCTION

Clinical gait analysis performed with video systems usually requires the use of markers to be positioned to the patient's body surface. In some occasions, the presence of markers may represent a source of uneasiness and discomfort, may interfere with natural walking. Moreover, operators are required to spend some time to set-up the patient, increasing the cost of the evaluation. To overcome the abovementioned limitations, markerless techniques are proposed. The main goal of this thesis is to develop new 2D markerless approaches for the analysis of gait.

The thesis is organized as follows.

Chapter 1 is a summary of the history of human movement analysis. A brief chronology of the devices and methods used throughout the history are presented.

Chapter 2 is about the state of the art of the methods to analyze human movement. Human movement is analyzed under the headings of marker-based and markerless human movement analysis. Theoretical background, applications and limitations of both methodologies are presented thoroughly.

Chapter 3 defines the aims of the thesis.

Chapter 4 presents a review of the algorithms used in the image processing implementation of this thesis. Besides, a brief comparison of the algorithms in the literature and their limitations are taken into consideration.

Chapter 5 presents a study focusing on the analysis of the 2D kinematics of the ankle-foot complex during the stance phase of gait from markerless images. The proposed technique is explained in detail with the sections of material and methods, results and discussion.

Chapter 6 presents an extension to the study proposed in Chapter 5. The proposed technique is applied to the knee flexion/extension estimation of children with CP.

Chapter 7 describes a hybrid technique, which is a combination of a markerless methodology and "segmental markers". The proposed technique is used in analyzing the lower limb kinematics of the children with CP. The purpose of the study, material and methods and results are explained in detail.

Finally, Chapter 8 sums up the discussions of the methods presented in this thesis.

INDEX

CHAPTER 1

HUMAN MOVEMENT ANALYSIS: HISTORICAL NOTES

Cinematography has been an essential instrument for the study and interpretation of animal motion. Experimental physiology has served as a catalyst for this technique, which has become a key tool for the progress in biological research. Two early scientists, German physiologist Karl Ludwig and French physiologist Etienne-Jules Marey - influenced this development.

Development of cinematography had an important effect on the development of the analysis of animal locomotion. During the 17th century, the new physics became embedded in Alfonso Borelli's work on animal motion. This work (Figure 1-1), which is a complete textbook of Physiology, claims that "every function in the living body, animal or vegetable, manifests itself through movement: macroscopic and apparent, as in locomotion, or microscopic, on an atomic dimension, as the movement in which atoms come in contact to form living matter" (Borelli, 1681). The fundamental aim of Borelli was to integrate physiology and physical science.
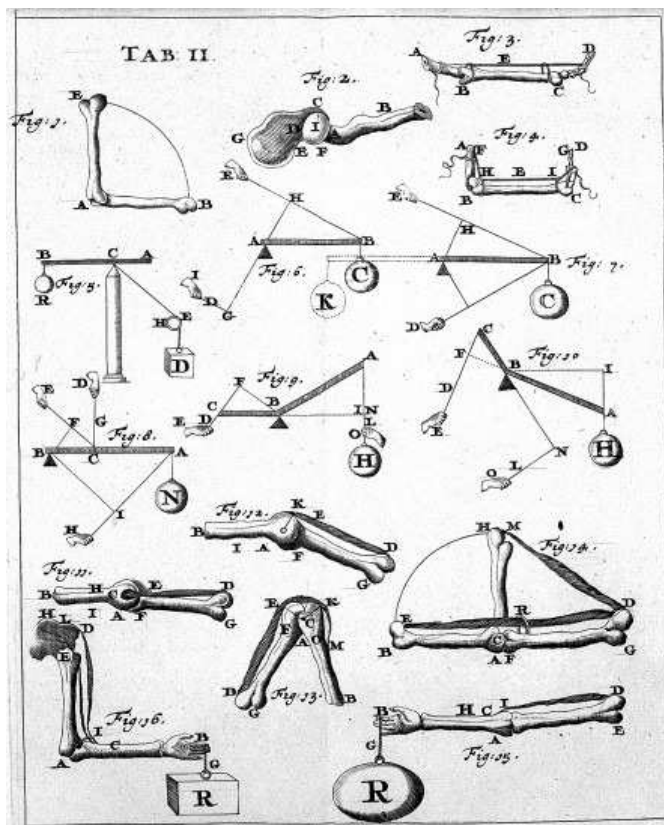
Fig. 1-1. Sample page from the book "De Motu Animalium" that shows the illustrations of biomechanical studies of Giovanni Borelli (from De Motu Animalium, 1681).

Adoption of the graphic method by Marey in 1857 was a breakthrough for the studies of animal locomotion. He used this approach during the following 20 years and applied it to humans, animals such as horse and dove together with the mechanical detectors he had designed to complex movement of locomotor acts. Marey published what he obtained using this approach in his books La Machine Animale", which was published in 1873, and "La Methode Graphique", published five years later.

Precisely at that time photography was begun to be used in the physiology in order to advance further the studies of biolocomotion. Leland Stanford has been claimed to be the first person to propose using photography to prove the real positions of a horse's leg during galloping.

English photographer Edward Muybridge made use of the idea of Stanford's with a series of cameras whose shutters were triggered by running horses, which was the beginning of his studies on biolocomotion. The most important contribution of Muybridge was the 781 plates he created, each having 1, 2 or 3 dozens of serial shots for a total of 20,000 images. These were published with the title "Animal Locomotion" (Figure 1-2).
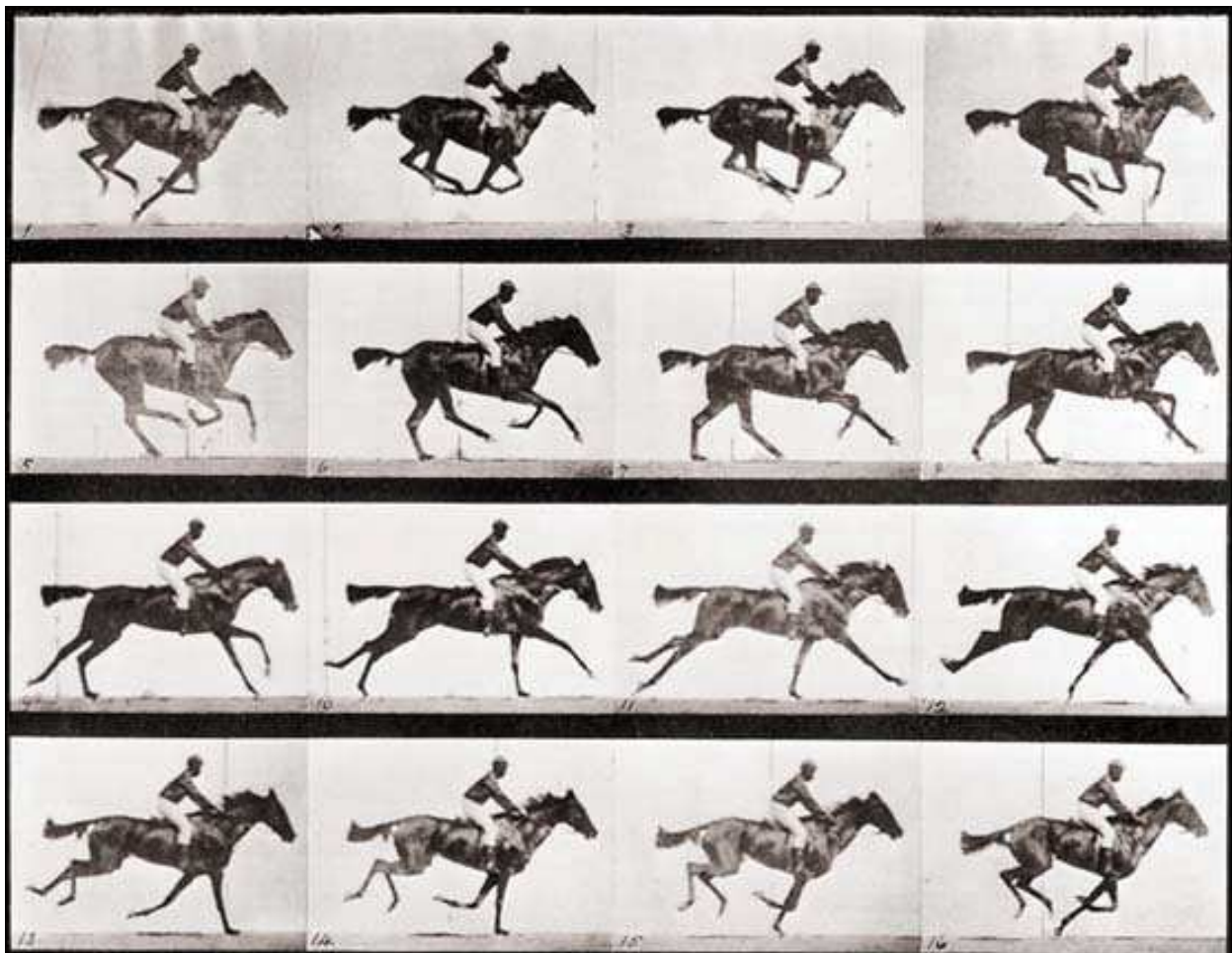
Fig. 1-2. Galloping horse (from Muybridge, 1878).

Using photography for the biolocomotion studies also inspired Marey and he envisioned a new device in order to obtain multiple images of moving objects at equal time intervals on a single photographic plate. This device - "fusil photographique" – was the first device of photographic apparatuses he invented for the study of locomotion. The invention of the process was named as "chronophotographie" by him (Figure 1-3).
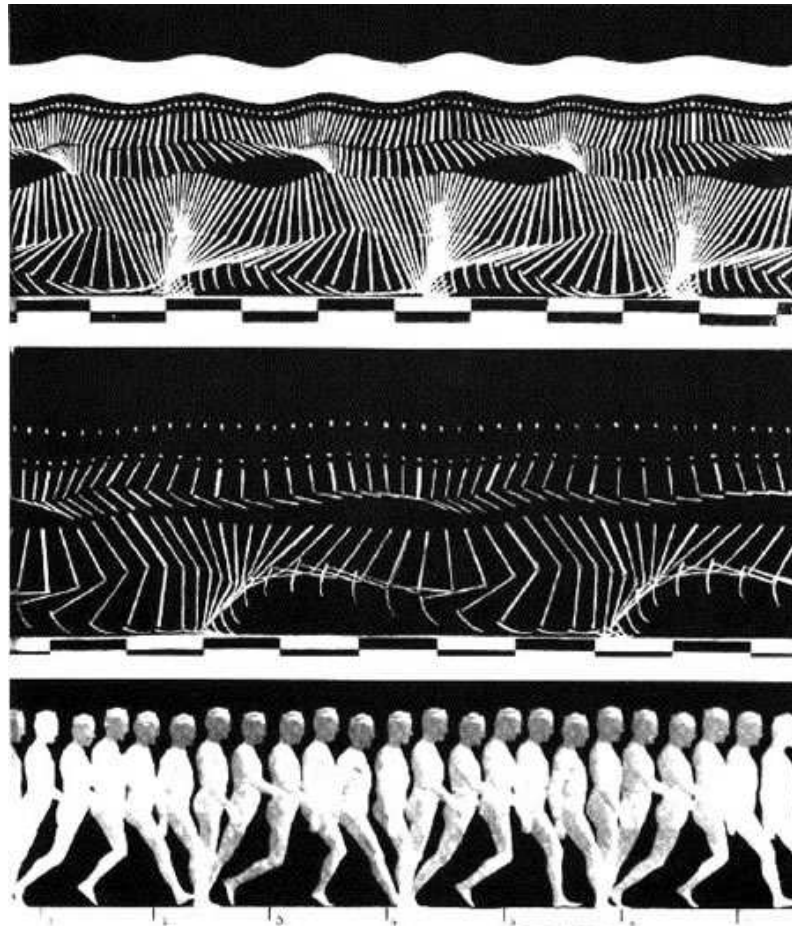
Fig. 1-3. Serial images of a man walking acquired on a still plate by Marey with his "chronophotograph". The model wears a black suit with white stripes and radiant points to indicate the position of one arm, one leg and the joints (from Musée Marey, Beaune).

Another French scientist, Jules Jannsen, had also contributed to his field with the photographic device he invented. This device – "revolver astronomique" - was able to record, on a single, circular-shaped, photosensitive plate having a rotating motion with regular intermittence, up to 48 consecutive images, spaced by constant time intervals, of an object in motion. Jannsen applied this device to the telescope and obtained the permanent recording of the transit of the planet Venus across the sun on December 8, 1874.

On October 29, 1888 Marey presented the device "chambre chronophotographique" which is claimed to contain all the principal components of the modern cine-camera. In 1893, Marey

constructed his projector, which was an important inspiration for the improvement of the cinematographic technique.

When the history of motion picture is analyzed, it can be realized that this invention cannot be attributed to a single person; it is collective work of optical, chemical and mechanical studies. The integration of these interdisciplinary contributions by the science of physiology was a great catalyst for the improvements in motion picture. Applying the graphic method to the study of biolocomotion was the key point in this contribution.

The growth of cinematography also contributed to physiology and introduction of cinematography can be considered as the "Renaissance" of biomechanics. Marey's publications have contributed to the knowledge of motor phenomena, but the results were largely qualitative. In 1895, two physiologists, Wilhelm Braune and Otto Fischer, started publishing their works on human locomotion which were very important for the science of biomechanics. Besides being able to estimate the locomotor act, they were the first to present a three-dimensional analysis of human movement by using stereophotogrammetry. They were also the first measure the forces acting on the human body.

In the years between 1927 and 1936, Nikolai Bernstein improved the work of Braune and Fischer by increasing the shutter frequency from 26 to 70-100 and in some cases 120-156 frames per second. This improvement enabled the details of the human movement to be observed. The analysis of coordinated movements became the study of biomechanics which involve kinematics and dynamics.

Research handled at the University of California at Berkeley between 1945 and 1947 on both normal and pathological human locomotion was also a great contribution to biomechanics. Electromyographic apparatus and the force platform were begun to be used in biomechanics laboratories.

In the mid-1960s many specialized biomechanics laboratories were founded both in Europe and in the United States of America. Bioengineering Unit of Strathclyde University in Glasgow, Scotland and Institute of Human Physiology of the University of Milan, Italy were among the most important of these institutions.

Starting in the 1970s, the optoelectronic technique began to be used which resulted in easier movement recording and faster data reduction. The data can easily elaborated by digital computer which enables the analyses to be performed conveniently and fast.

CHAPTER 2

HUMAN MOVEMENT ANALYSIS: STATE OF THE ART

## 2.1 Marker-Based Human Movement Analysis

2.1.1 Theoretical Background

"Acquisition of quantitative information about the mechanics of the musculo-skeletal system during the execution of a motor task is the main goal of the human motion analysis" (Cappozzo et al., 2005). In order to pursue this goal, motion capture is frequently used in biomechanics. Human motion capture is widely used in order to study musculoskeletal biomechanics and clinical problems. In this context, estimating joint kinematics is of extreme relevance. For this purpose, video-based optoelectronic systems are commonly preferred among the human motion capture systems.

Gait analysis is generally carried out by mounting retro-reflective markers on the body of the subjects and reconstructing their 3D position using video-based optoelectronic systems (Figure 2-1). Retro-reflective markers and infrared illumination - produced by light-emitting diodes (LEDs) around the lens of the cameras - are used for the 3D reconstruction. By adjusting the camera thresholds, reflective markers are sampled and the recognition of the markers in the video frames is performed.

The 3D position of a marker in a reference frame fixed to the laboratory (global frame - GF) can be reconstructed if the marker is visible from at least two cameras at the same time. Visibility from additional cameras is usually beneficial (Chiari et al., 2005). Additional reference frames associated to body segments (technical frame - TF) can be defined from the position in the GF of cluster of markers attached to the same body segments.

The pose of the TFs in the GF can then be determined. However, although considered fixed to the underlying bone, the TFs are not necessarily representative of the anatomy of the body segment they are attached to. For this reason for each body segment being analyzed an additional frame is defined: the anatomical frame (AF). The AF has a time invariant relationship with the respective TF. To define AFs, it is necessary to determine the location of selected anatomical landmarks (ALs) with respect to the relevant TF (Cappozzo, 1995). Standards for the definitions of AFs have been proposed (Wu et al., 2002; Wu et al., 2005).

The pose of an AF is the orientation and position in space of a body segment. Given the pose of the AFs of two adjacent body segments, the kinematics of the joint between the two body segments can be determined.
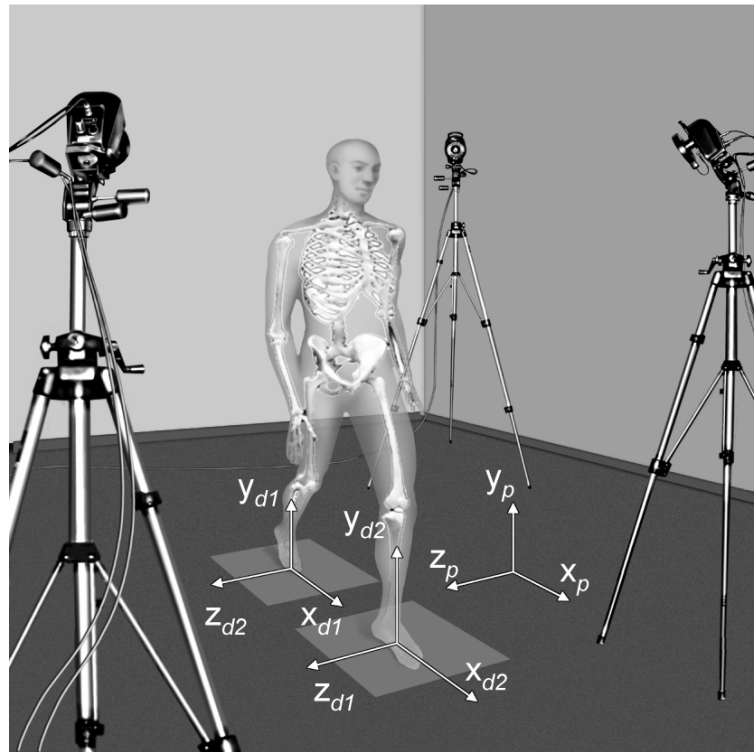
Fig. 2-1. The human movement analysis laboratory with basic measurement instruments, with their systems of axes (p: photogrammetry; d: dynamometry). When level walking is analysed, the motor task frame may overlap with the frame of one of the two force plates (from Cappozzo et al., 2005).

2.1.2 Calibration of Anatomical Landmarks

ALs are either bony prominences or bone points of geometrical relevance. In the first case they are normally identified by palpation, in the second case, they can be identified using imaging, regression equations or functional movements (Cappozzo et al., 2005). In any case, once identified, their location with respect to the relevant TF has to be determined. Once the location of ALs in their relevant TF is determined, it is possible to reconstruct their position in the GF by simple coordinate transformations (Figure 2-2). The Calibrated Anatomical System Technique (CAST) is an experimental methodology that formalizes the concept of AL calibration and allows the implementation of various calibration methods.

The AL calibration can be implemented using a) a marker positioned on the AL during a static acquisition, b) a pointer where a minimum of two markers are mounted with a known distance from its tip, pointing at the AL during a static acquisition, c) determining the centre of rotation of recorded functional movements (for joint centres, such as the hip centre), d) by imaging of the bone and the relevant TF (Cappozzo et al., 1995; Benedetti et al., 1998).

Recently, the CAST methodology was updated, by adding information on the subject-specific bone geometry. By determining the position of unlabelled points (UPs) situated over all prominent

parts of the bone surface, initial estimation is employed. After the estimation step, a digital model of a template-bone is matched to them. The estimated subject-specific bone contains all relevant anatomical landmark locations. The technique, UP-CAST, is evaluated in terms of repeatability and accuracy on average weight subjects (Donati et al., 2007).
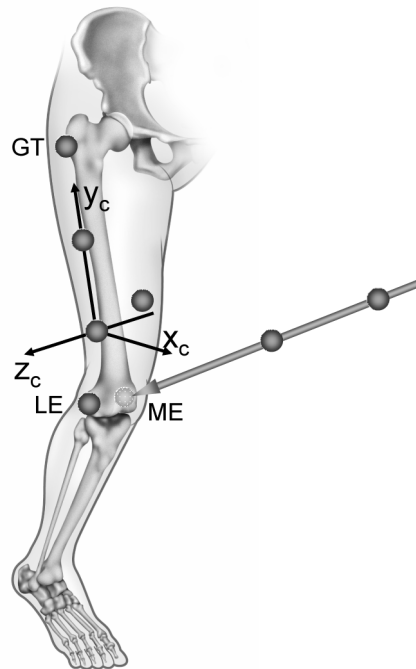
Fig. 2-2. Anatomical calibration using stereophotogrammetry. The following external, palpable, anatomical landmarks are indicated: prominence of the greater trochanter external surface (GT), medial epicondyle (ME), lateral epicondyle (LE). The location of the external anatomical landmarks relative to the marker cluster technical frame (x_c, y_c, z_c) may be reconstructed using markers denoting the anatomical landmarks, or using a wand which carries a cluster of at least two markers. Prior to recording, the end point of the wand, the position of which relative to the latter cluster of markers is accurately known, is made to coincide with the target anatomical landmark (from Cappozzo et al., 2005).

2.1.3   Protocols

Human movement analysis and gait analysis in particular, typically makes use of the theory of multi- rigid body systems. The portion of human body is modelled with a number of rigid segments. Adjacent segments are connected by joints. The number of degrees of freedom of the each modelled joint characterizes the multi-body system model employed. Protocols – data collection and reduction practices – have been proposed in gait analysis offering various ways of modelling the system of rigid bodies of interest. Often, in clinical gait analysis, all model joints are rotational (either cylindrical or spherical) and AFs are defined based also on this assumption. The rationale behind this choice is related to the errors affecting the human movement recordings (see following section).

Proposed protocols also differ in the marker-sets used to identify AFs and joint centre locations. Typically data acquired with different protocols cannot be compared.

From these protocols, "Newington model" is the pioneer and the most commonly used practice for gait data acquisition and reduction which has also been used by the commercial applications like Plug-in Gait (PiG—Vicon Motion Systems, Oxford, UK). "Servizio di Analisi della Funzioni Locomotoria" developed their protocol named "SAFLo" - which differs from the Newington model in terms of segmental anatomical references and anatomical marker configurations. Then, "Calibration Anatomical System Technique" (CAST) was introduced which standardizes and defines references, internal anatomical landmarks and external technical markers. Then, protocols of "Laboratorio per l'Analisi del Movimento nel Bambino" (LAMB) and "Istituiti Ortopedici Rizzoli Gait" were proposed, of which the latter was the basis of the software "Total 3D Gait" (T3Dg-Aurion s.r.l, Milan, Italy) (Ferrari et al., 2008; Baker, 2006).

Ferrari et al. compared these commonly used protocols and find out that same gait cycles revealed good intra-protocol repeatability. Regardless of the known significant differences among the techniques, reasonable correlations are observed for most of the gait variables. It was pointed out that model conventions and definitions seem to be more important than the design of the relevant marker-sets. Sharing the model conventions and definitions can be sufficient for worldwide clinical gait analysis data comparison (Ferrari et al., 2008).

2.1.4 Sources of errors

Human movement analysis performed with stereo-photogrammetry is affected by three major sources of errors.

- Instrumental errors: these errors are the results of both instrumental noise and volume calibration inaccuracies. These errors have been studied intensively in the 80s and 90s (Fioretti and Jetto, 1989; Chen et al., 1994), tests for estimating them have been proposed (Della Croce and Cappozzo, 2000). The instrumental noise can be substantially reduced by low pass filtering. The volume calibration inaccuracies stem from the inadequate number of cameras and the volume calibration algorithm chosen for the application. Direct linear transformation (DLT) algorithm (Abdel-Aziz and Karara, 1971) is broadly used, but when the volume of interest is large, the construction of a suitable calibration object to be used with DLT becomes restrictive. Simultaneous multi-frame analytical calibration (SMAC) (Woltring, 1980) - a technique based on a planar calibration object with a grid of known control points - suffices the recording of the calibration object by at least two convergent cameras. SMAC allows covering larger volumes when compared with DLT, but for very large volumes, analytical self-calibration is more appropriate (Chiari et al.,

2005). Thus, the volume calibration inaccuracies can be remarkably lowered by increasing the number of cameras and improving the volume calibration algorithms.

The contribution of the instrumental errors to the total error is currently considered negligible.

- Soft tissues artefacts: the markers captured by the cameras can be directly attached to the skin or arranged in clusters and positioned with fixtures over a body segment. Due to its origin, this error has the same frequency content as the bone movement. Thus, there is no way of separating the artefact from the actual bone movement by simply using a filter, as opposed to most instrumental errors. However, its effect on the end results may be reduced in the following ways. First of all marker locations (marker points) must be chosen so that the above-mentioned relative displacement is minimal, and secondly through a proper choice of the mathematical operator which estimates position and orientation of the bone from skin marker positions (Lucchetti et al., 1998; Alexander and Andriacchi, 2001). Operators that cope with this problem in an optimization context have been proposed" and their use in movement analysis is being developed.

Knowledge regarding the characteristics of the artefact movement in different body segments is required in order to utilize the mentioned countermeasures against the experimental artefacts.

- Anatomical landmark misplacement: The incorrect location of subcutaneous bony ALs through palpation can be caused by three main factors: (1) the palpable ALs are not points but surfaces, sometimes large and irregular; (2) a soft tissue layer of variable thickness and composition covers the ALs; (3) the identification of the location of the ALs depends on which palpation procedure was used. Studies showed that AL position uncertainty and consequently the erroneous determination of AF axes may result in erroneous clinical interpretations of the estimation (Della Croce et al., 2005).

In addition to the abovementioned sources of errors, marker based movement analysis is affected by the influence of markers attached to the body on the subject's movement and the need of an extended setup time for marker placement (Corazza et al., 2006).

2.2 Markerless Human Movement Analysis

Some of the limitations of marker based systems can be overcome using a completely different approach. Markerless systems of human motion capture have been proposed where cameras can be utilized without the necessity of using special clothing or hardware (Deutscher et al., 2000). Markerless motion capture ensures an important reduction of the amount of time for setup preparation in comparison to marker-based techniques. Besides, the problem of inter-operator

variability is removed because no specialized operator is needed to place markers on the skin (Corazza et al., 2006).

Markerless techniques can be classified into model-based and model-free techniques both for the cases of 2D and 3D applications. Model-based approaches utilize an a priori human body model and are composed of two stages: modelling and estimation.

Modelling is the building step of a likelihood function by taking camera model, image descriptors, body model, and matching function into consideration. Estimation step is fitting the optimum pose in the likelihood domain planned in the modelling step. Model-free approaches do not use an a priori model, but "implicit model variations in pose configuration, body shape, camera viewpoint and appearance" (Poppe, 2007).

In the next section, current markerless techniques, brief algorithm explanations and their limitations are analyzed and exemplified with current applications.

2.2.1   2D Markerless Techniques

Much of the work on motion analysis uses detailed 3D kinematic models and 3D motion estimation. These techniques require multiple camera viewpoints, but motion analysis can also be operated using a single camera input (Cham and Rehg, 1999).

Motion capture with a single camera is a significant task since data acquisition is very simple, besides being an interesting computer vision challenge that focuses on inference as much as movement (Howe et al., 2000).

Hence, 2D markerless techniques, which can be classified as model-based and model-free, have been proposed. In the following section, the algorithms and the applications of these techniques are introduced.

2.2.1.1   Model-Based Techniques

Ju et al. (Ju et al., 1996) uses a cardboard model to define the human body as a set of connected planar patches and to approximate the limbs as planar regions (Figure 2-3). The main assumption behind this model is that, the motions of the limb planes are assumed to be the same at the points of articulation. The motion of each patch is estimated using the energy minimization (annealing) concept and estimated motions are called "absolute motions". After the estimation of the absolute motions, it is necessary to estimate the articulated motions. To estimate the articulated motions, the motions of limbs which are relative to their preceding (parent) patches should be recovered. The relative motion of the patch is calculated using the displacements of the connected patches.

The estimated articulated motion between two frames is used in the tracking step in order to predict the location of each patch in the next frame. In the first frame, each patch is manually defined by its four corners. For every patch, the first two corners are defined as the articulated points, whose corresponding points are the last two corners of its previous patch. This shows that two connected patches share an "edge". After the "chain" structure definition step is over, automatic object tracking starts. The articulated motion between two frames is used to predict the location of each patch in the next frame. Then, the location of each of the four corners of each patch is updated by applying its estimated planar motion to it.

The experiments demonstrate that the image motion models are able to track motion correctly during long sequences. In this study, optical flow is estimated with the parameterized model, 3D model is not necessary and edges are not used.

In the study of Deutscher et al. (Deutscher et al., 2000), the idea of annealing is adapted to perform a particle based stochastic search. The adapted algorithm is called annealed particle filtering and is capable of recovering full articulated body motion. The authors focused on the problem of constraining the search space along which the real posture of the subject is investigated. Other studies (Hogg, 1983; Goncalves et al., 1995; Bregler and Malik, 1998) in the literature do this making the following assumptions: 1) assuming that the subject is walking, 2) assuming a constant angle of view, 3) performing a hierarchical search using color cues. The study of Deutscher et al. does not depend on these assumptions and reduces the dimensionality of the search space through annealed particle filtering (Condensation algorithm).
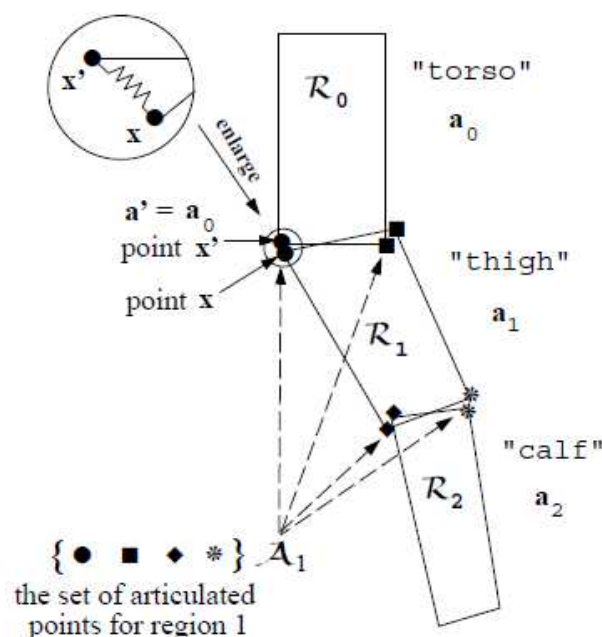

Fig. 2-3. The "chain" structure of a three-segment articulated object (from Ju et al., 1996).

13

This algorithm includes an edge detection procedure followed by a particle filtering, developed to match a 3D human body model of 29 DOF to the 2D image edge (Figure 2-4). However, this matching process has a very high computational cost. Also, a number of particles are required to the posterior density representations, which increase with the size of the model's configuration space. In order to solve these problems, a multi-layer method using simulated annealing approach was implemented. Even if knee and ankle joints are modelled as simple hinge joints, the algorithm performance is in general satisfactory. The tracking performance of the algorithm was compared to the standard Condensation algorithm (i.e. Particle filtering algorithm which is used for tracking objects in clutter (Isard and Blake, 1996)) and resulted to perform better even if it uses fewer particles.

2.2.1.2 Model-Free Techniques

When there is not a priori human body model, there has to be a mapping between the image output and pose. Model-free algorithms, which do not suffer from (re)initialization problems, can be used for initialization of model-based pose estimation approaches (Poppe, 2007).
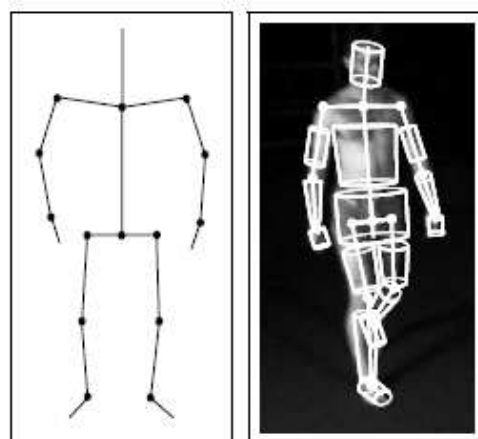
Fig. 2-4: The model is based on a kinematic chain consisting of 17 segments (a). Six degrees of freedom are given to base translation and rotation. The shoulder and hip joints are treated as sockets with 3 degrees of freedom, the clavicle joints are given 2 degrees of freedom and the remaining joints are modelled as hinges requiring only one. This results in a model with 29 degrees of freedom. The model is fleshed out by conical sections (b) (from Deutscher et al., 2000).

Mori and Malik (Mori and Malik, 2006) estimate body pose in 3D by placing the joint points in a single 2D image with a human figure. First, a number of example views of the human body in different viewpoints with respect to the camera, are acquired. Each of the views are manually marked from the body joints and labelled. Then, the input figure is matched to each stored view using the shape context matching with a kinematic chain-based deformation model. By extracting

external and internal contours of an object, shape contexts are employed to encode the edges (Figure 2-5).
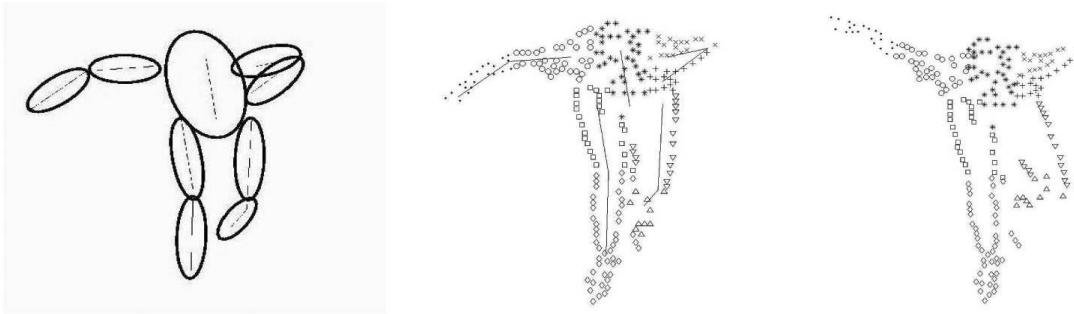
Fig. 2-5. The deformation model. (a) Underlying kinematic chain. (b) Automatic assignment of sample points to kinematic chain segments on an exemplar. Each different symbol denotes a different chain segment. (c) Sample points deformed using the kinematic chain (from Mori and Malik, 2006).

Following the correspondence step, the locations of the body joints are then moved from the example views to the test figure. The 3D body configuration and pose are then estimated using the existing algorithm of Taylor with the 2D joint locations (Taylor, 2000) which uses point correspondence in a single image. In an estimation step, the stored example images are deformed to match the image observation. The 2D joint estimate is found by enforcing 2D image distance consistency between body parts. This technique can be applied to each frame of a video sequence so that tracking recognition becomes repeated for every frame. The experiments of this study are performed with CMU MoBo Database (Gross and Shi, 2001) and the main contribution of this study was demonstrating the use of deformable template matching to example views in order to localize human body joint positions.

In another 2D model-free markerless application, Elgammal and Lee (Elgammal and Lee, 2004) use human silhouettes extracted from a single camera to derive 3D poses. The purpose of this study is to recover the intrinsic body configuration and reconstruct the silhouette excluding the outliers from the visual input. To recover intrinsic body configurations from the silhouette, manifolds are learned from the visual input and afterwards mappings are learned from manifolds to visual input and 3D poses (Figure 2-6).

The experiments demonstrate that the model can be learned from the data of one person and successfully adapt to recovering poses for other people from noisy data. When compared to previous approaches for inferring 3D body pose from visual input, this approach has certain advantages and limitations. This framework makes interpolation of intermediate 3D poses easy even if they are not part of the training data. This approach constrains the mapping to the learned manifold which facilitates robust pose recovery from noisy inputs as well as for reconstruction of

the input. It is based on learning activity manifold and so its application is limited to recovery of poses for the learned activities only. Although in this study the focus is gait, the framework is general and can be applied to other activities by learning their manifolds. In the experiments, validation was done with a sequence obtained from the Georgia Institute of Technology (Atlanta, GA, USA) data compared to relevant motion capture data. CMU Mobo Gait database was used to demonstrate that the proposed approach, which is based on "learning" from the data of a single person, is also applicable to different people. CMU Mobo Gait database contains six views of each walking person. Five views from one person were used for the learning process of the study. In order to evaluate the 3D reconstruction, five sequences (five people with five views, each) were used. Overall correct classification rate from a single frame was 93.05%, while it increased to 99.63% after five frames were used for the classification.
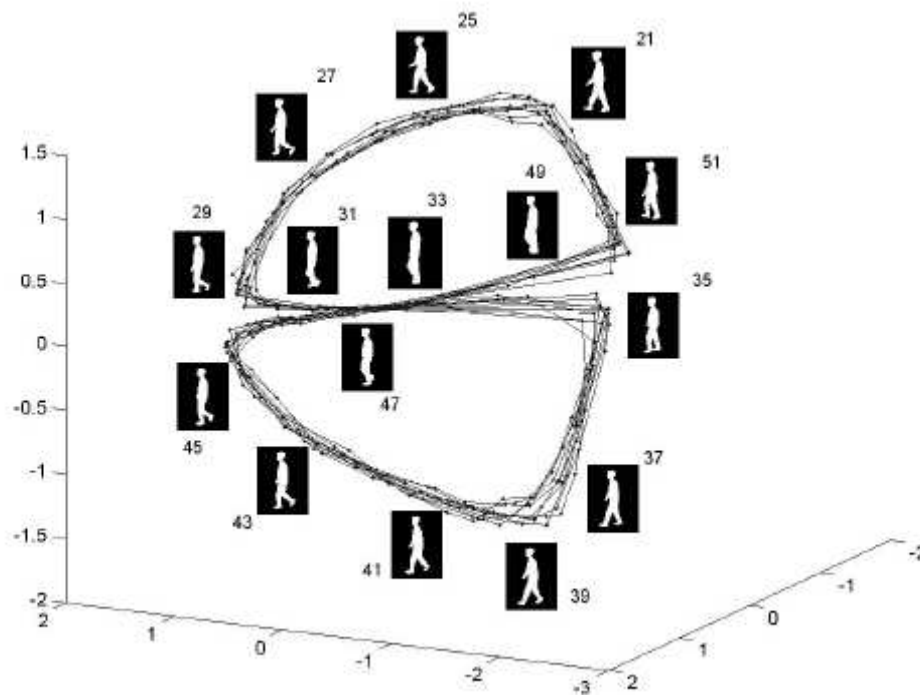
Fig. 2-6. Embedded gait manifold for a side view of the walker. Sample frames from a walking cycle along the manifold with the frame numbers shown to indicate the order. Ten walking cycles are shown (from Elgammal and Lee, 2004).

In another markerless 2D study, Goffredo et al. (Goffredo et al., 2009) use a region of interest (ROI) based tracking approach for the kinematic analysis of sit-to-stand (STS) tasks. Their approach uses Gauss–Laguerre transform (GLT) since image features such as edges, lines and orthogonal crosses are enhanced easily regardless of their orientations. A 4-segment human body

model was applied (only for the purpose of pose reconstruction of the body segments); each segment was assumed to be rigidly connected the next segment with an ideal rotational joint. Four body markers on ankle, knee, hip and shoulder joints were selected in the initial frame of every sequence and the GLT algorithm was applied to these ROIs. By computing the corresponding candidate points via calculating the log-likelihood function between the textures of consecutive frames, points relevant to motion estimation were tracked (Figure 2-7). The authors used GLT-based motion estimation method in a 2D setting, but this method could be used in stereo vision applications where GLT coefficient matching can be applied for motion and disparity field estimations. Estimating both the translations and the rotations of related anatomical segments in the transform domain with a pattern algorithm appeared to be a good solution for the movement reconstruction. For validation purposes, marker-based results from the study of Gross et al. (Gross et al., 1998) were used. The results of the validation are reported in Figure 2-8.
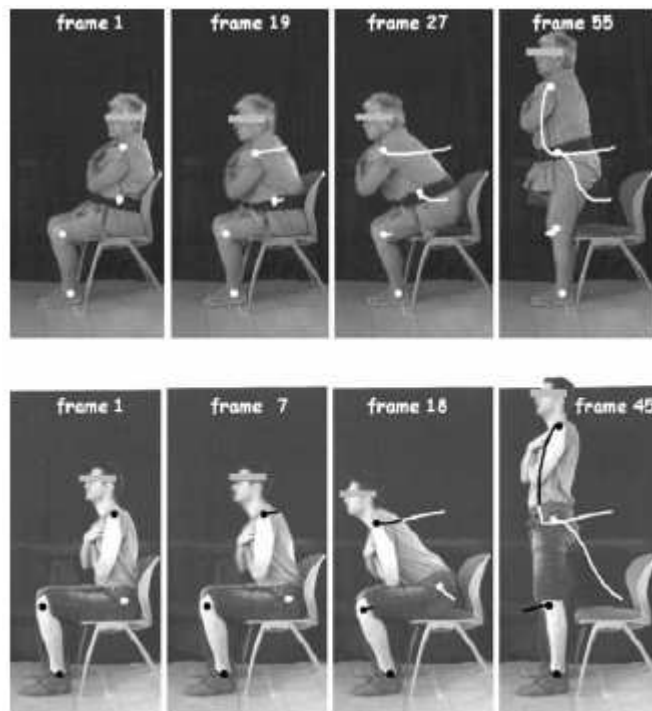
Fig. 2-7. Estimated trajectories obtained with the proposed method (upper panel: elderly subject I; lower panel: young subject D). The circles are the points of interest on which the GLT algorithm has been applied. The lines are the estimated trajectories at the end of the phases characterizing the STS task (from Goffredo et al., 2009).

### 2.2.2   3D Markerless Techniques

When the observation is limited to a single camera, the 3D motion of humans is not determined thoroughly, due to the inherent 3D ambiguity of 2D video (Howe et al., 2000). In order to overcome this ambiguity, 3D markerless techniques are proposed.
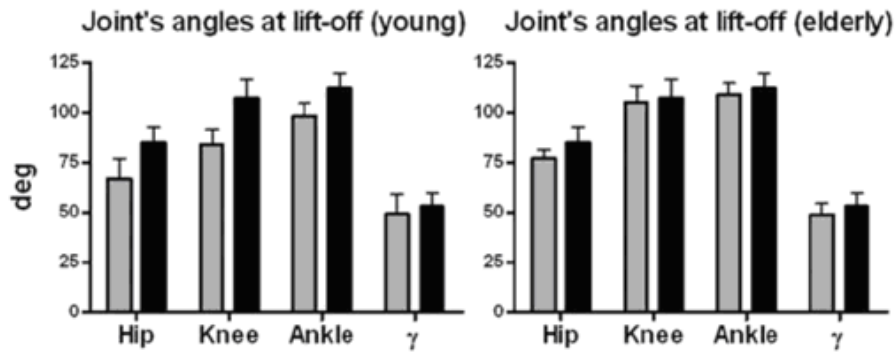
Fig. 2-8. Hip, knee, ankle, and γ (shin orientation) angles at liftoff during chair rise. The results obtained with the proposed markerless system (gray bars) are compared with the one obtained by Gross et al. (black bars) with a marker-based system (from Goffredo et al., 2009).

2.2.2.1 Model-Based Techniques

An important study exemplifying 3D model-based markerless technique is the study of Bottino and Laurentini (Bottino and Laurentini, 2001). Bottino and Laurentini presented a technique to reconstruct unconstrained motion from multiple-view images, which were computed using volume intersection data. First, views of the human body were acquired using different cameras and their 2D silhouette was extracted from each of the images. Then, a volumetric description was formed by intersecting the cones derived by back-projecting from each viewpoint of the corresponding silhouette. This step, called the volume intersection (VI), provided the final voxel representation. Then, a human body model was fitted to the extracted volume (Figure 2-9). Model fitting was done with the minimization of a distance function between the volume and the model via a search through the space of pose parameters. Pose recovery was based on a search through the 32 dimensional space of pose parameters and entailed finding the pose of the model closest to the actual appearance of the moving subject. The approximation accuracy was measured by a similarity function between the current model pose and the volume obtained by VI. This function was obtained by summing the squared distance between each voxel center to the closest segment of the model. Each segment was approximated with an oriented bounding ellipsoid (OBE) at the first step of the reconstruction algorithm in order to reduce the number of computations. The size of the axes of each OBE was the same as the dimensions of the boundary box of the related segment. The posture recovery was a two-step process: first, the OBEs were fitted to the volume reconstructed and then, fitting was employed to the model. In order to recover the motion of the model, the procedure mentioned above was applied to every frame, followed by the implicit filtering to avoid the phantom volumes.
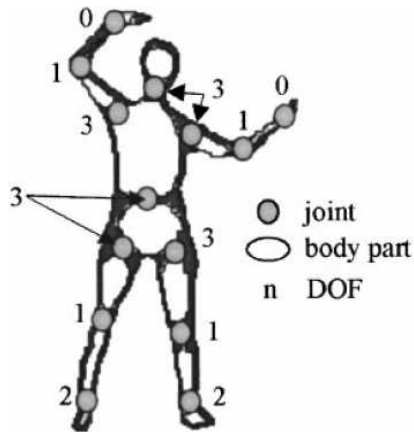
Fig. 2-9. Human body model used in the study of Bottino and Laurentini (from Bottino and Laurentini, 2001).

The experimental setup is composed of two phases. First, the system was tested in a virtual environment in order to investigate the precision of 3D direct reconstruction. Second, the proposed approach was applied to real image sequences (Figure 2-10). Results showed that the proposed approach could reconstruct unconstrained human motion without using markers or external devices attached to the subject's body.
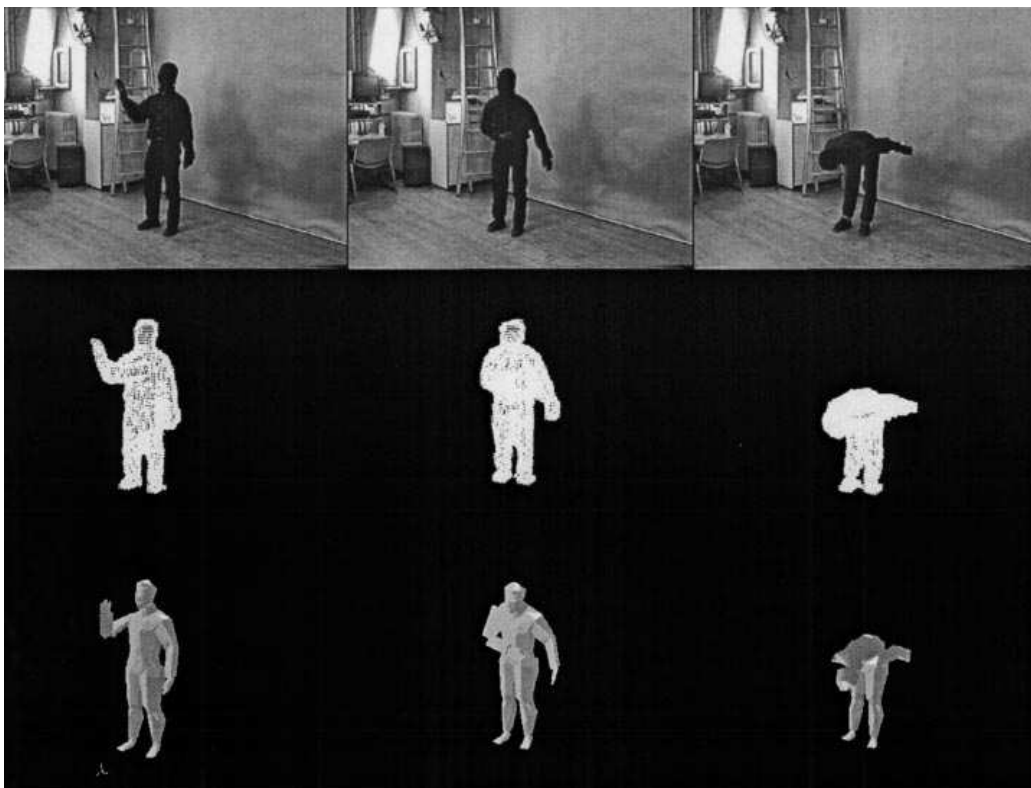
Fig. 2-10. Original images, reconstructed voxel models, and parameterized shape models for a bow sequence (from Bottino and Laurentini, 2001).

Corazza et al. (Corazza et al., 2006) also used the annealing approach in order to implement a markerless technique making use of visual hull reconstruction. The full body model contained morphological information (surface with 1600 points) and kinematics information on the possible movements of the model. The morphological information was provided by a laser scan of a reference pose of the subject. The model was segmented into parts corresponding to twelve anatomical segments (pelvis, thighs, shanks, feet, arms, forearms, and combined torso and head). The full body kinematic model had 33 degrees of freedom (DOF) where joints were modelled as ball-and-socket joints or as hinge joints. The geometrical formulation of the model was flexible such that each joint model could be modified separately without readjusting the others. The completed model was formed by rigidly combining the morphological representations to the kinematic model. (Figure 2-11). Visual hull was reconstructed, and matched to the model with an adapted fast simulated annealing approach. The validation of the technique was done in a virtual environment where an animated virtual character with known kinematics, provided the gold standard. The validation results are reported in Table 2-1.
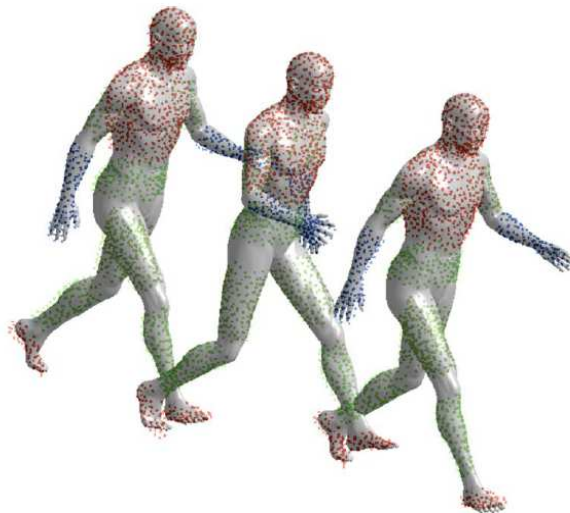
Fig. 2-11. Results of the matching algorithm (colored points) applied to the virtual environment sequence superimposed over the virtual character (gray surface) (from Corazza et al., 2006).

	Mean Error ($^\circ$)	Standard Deviation ($^\circ$)	RMS Error ($^\circ$)
Hip flexion/extension	2.0	3.0	3.6
Hip adduction/abduction	1.1	1.7	2.0
Knee flexion/extension	1.5	3.9	4.2
Knee adduction/abduction	2.0	2.3	3.1
Ankle plantar/dorsiflexion	3.5	8.2	9.0
Ankle inversion/eversion	4.7	2.8	5.9
Shoulder flexion/extension	1.2	4.2	4.4
Shoulder adduction/abduction	3.8	1.2	4.0

Tab. 2-1. Summary of validation results for joint angles at the hip, knee, ankle and shoulder (from Corazza et al., 2006).

Another application to 3D model-based application is proposed by Bregler and Malik (Bregler and Malik, 1998). Bregler and Malik presented a motion estimation technique capable to extract high degree-of-freedom articulated human body configurations from complex video sequences using exponential maps and twist motions. The product of exponential maps and twist motions and their integration into differential motion estimation is a significant parameterization. The pose of each body segment was defined with respect to its "parent" segment to which is attached through a revolute joint. Moreover, the formulation is a very simple linear representation of the motion model. The visual tracking was based on an initial frame, in which the angular configuration was known beforehand. The 2D joint locations in all views were manually marked by a user. The 3D poses and the image projection of the matching configuration was found by minimizing the sum of squared differences between the joint locations of the projected model and of the marked model. The study provided a new technique for articulated visual motion tracking. The tracking results were qualitatively tested on video recordings of moving subjects, and on the Muybridge photographic sequences (Figure 2-12), but no quantitative information on the tracking parameters was given.

In the studies of Cheung et al. (Cheung et al., 2005 (Part I and Part II)) voxel based surface from silhouette algorithms combined with a new colour based approach, were used. Colored Surface Points (CSP) – multi-view stereo points – were extracted from the surface of the object and used in a 3D alignment algorithm, rigid motion between visual hulls determined and recursively refined and the silhouette images were used to refine the object's shape. Figure 2-13 shows the CSPs on the visual hull of the subject.
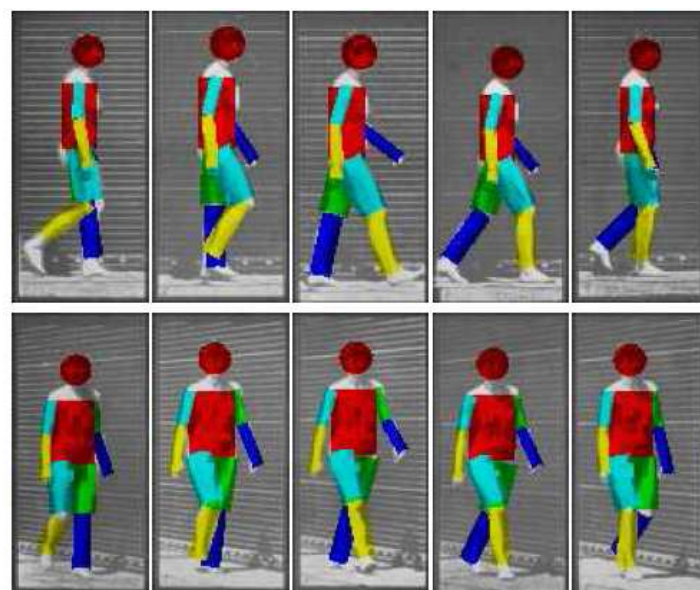
Fig. 2-12. Muybridge's Woman Walking: Motion Capture outputs. This shows the tracked angular configurations and its volumetric model projected to two example views (from Bregler and Malik, 1998).

A sample of shape reconstruction and digital model rendering are reported in Figure 2-14. Tracking algorithm is evaluated with ground-truth data, obtained from synthetic sequences generated with OpenGL. Figure 2-15 shows the results of this validation. The use of the method requires a controlled environment since the algorithm is based on color information.
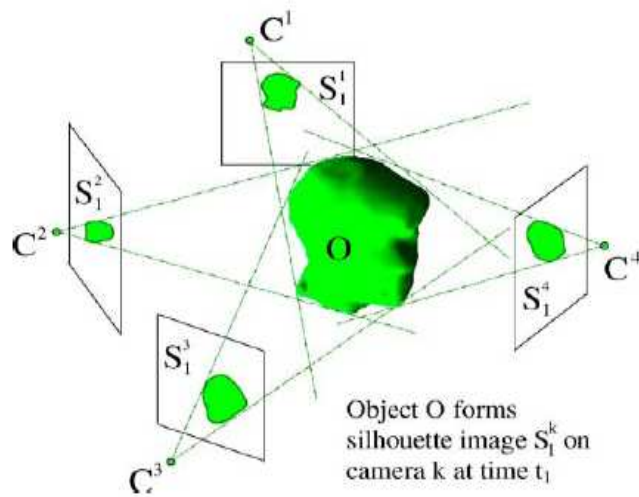
Fig. 2-13. The Shape-From-Silhouette problem scenario: a head shaped object O is surrounded by four cameras at time t_1. The silhouette images and camera centres are represented by S_{kj} and C_k respectively (from Cheung et al., 2005).
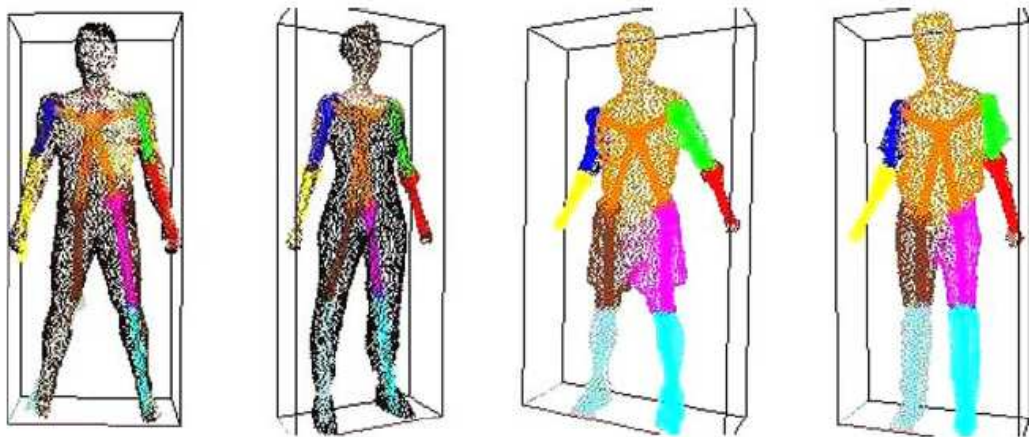
Fig. 2-14. Articulated model of (a) synthetic virtual person, (b) Subject E, (c) Subject G and (d) Subject S. In (a) and (b), the CSPs are shown with their original colours. In (c) and (d), the CSPs of different body parts are shown with different colours. For display clarity, the CSPs drawn are down-sampled in the ratio of one in two in total number of points (from Cheung et al., 2005).
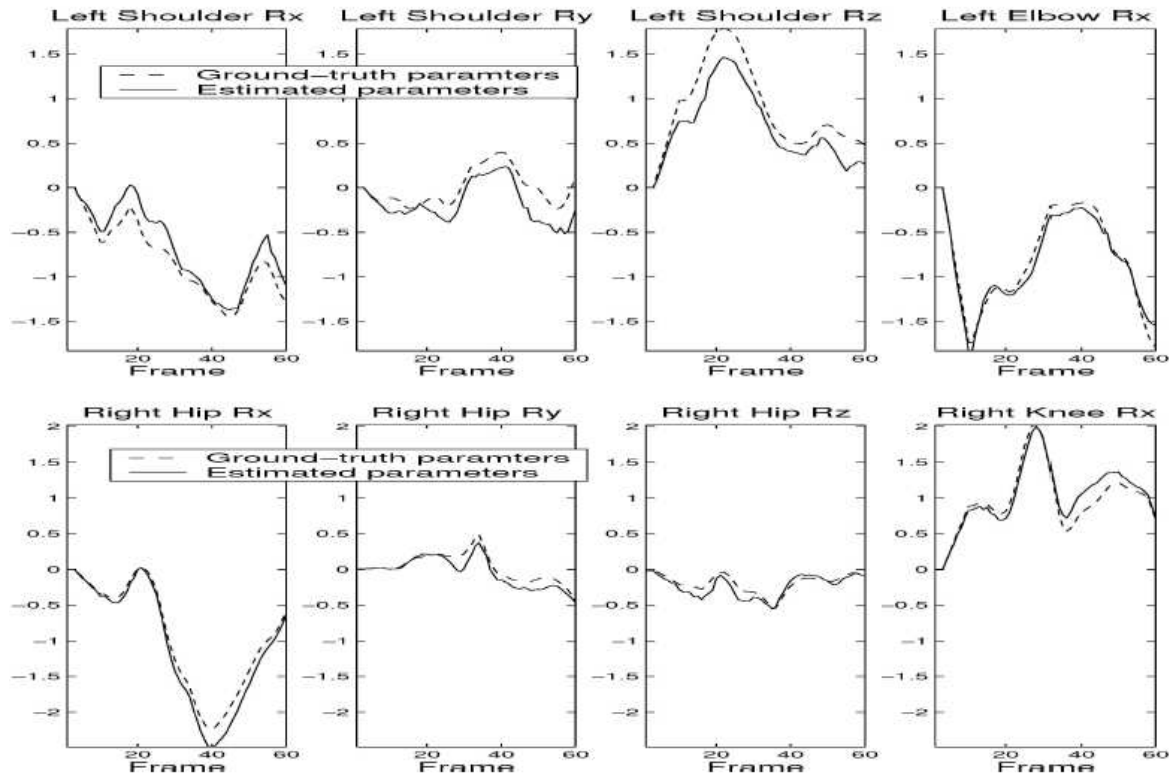
Fig. 2-15. Graphs comparing ground-truth and estimated joint angles of the left arm and right leg of the synthetic sequence. The estimated joint angles closely follow the ground-truth values throughout the whole sequence (from Cheung et al., 2005).

Sundaresan and Chellappa (Sundaresan and Chellappa, 2006) modelled the human body with a set of articulated super-quadrics and proposed algorithms in order to estimate the model parameters from video sequences. The model was a combination of different body segments and some labelled joints. Each body segment was modelled with a tapered super-quadric (i.e. 3D parametric objects which give the flexibility of defining a large variety of shapes in geometric modelling (Barr, 1981)). The trunk segment was the base, and together with neck, head and four limbs formed the kinematic chain of the human body. Figure 2-16 shows the 3D body model used in the study. A 3D scanned model was used to obtain the dimensions of the super-quadrics. The trunk segment had 6 DOF while the other segments had at most 3 rotational DOF with respect to the trunk. The body model involved the shape and the joint locations of the body segments. Given the pose at time t, the pose at time $t+1$ was computed by using the images at time t and $t+1$. The pose estimation required the prediction step and the correction step. Pixel displacement estimation, pose prediction and silhouette-motion combination were necessary to estimate the pose at $t+1$. The authors claimed the method to be accurate and robust using a visual feedback. Accuracy of the method is strongly dependent on the quality of the estimation of joint location during the model acquisition. The flexibility of the model on some joints (e.g. shoulder joint) affects the performance of the method.

23
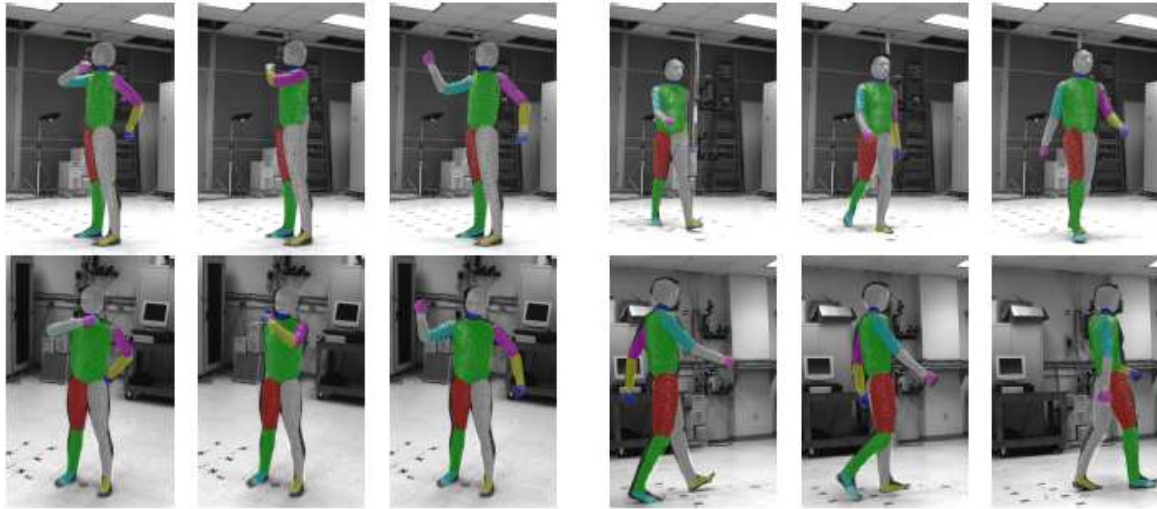
Fig. 2-16. Tracking results using both motion and spatial cues (from Sundaresan and Chellappa, 2006).

Gagalowicz and Quah (Gagalowicz and Quah, 2009), proposed a novel 3D model-based framework and algorithm that can manage clutters and occlusions, is proposed. This method uses a 3D geometrical human model similar to the subject, in order to synthesize the candidate posture producing the image minimizing the matching error with the real image. In this approach, segmentation is performed through the direct projection, texturing and shading via the 3D geometrical human model onto the images (Figure 2-17). The use of analysis-by-synthesis and error feedback allows avoiding the ill-posed problem of standard segmentation. Results of tracking the arms in the presence of occlusions and clutters were presented.
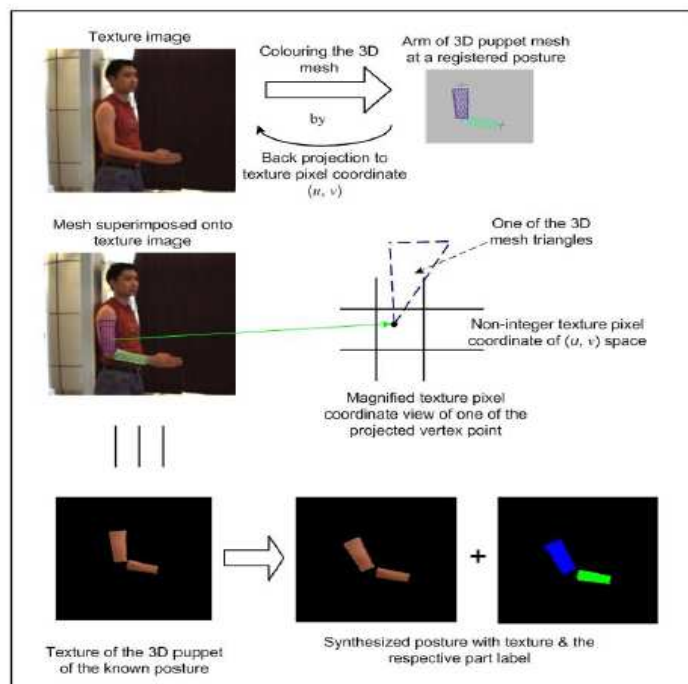
Fig. 2-17. Generating the human pose for matching (from Gagalowicz and Quah, 2009).

2.2.2.2 *Model-Free Techniques*

An example study in the 3D model-free markerless application is the study presented by Chu et al. (Chu et al., 2003). They proposed an approach that generates "underlying nonlinear axes" (or skeleton curve) from a volume of a human subject. Multiple cameras were used for human volume capture and skeleton curves estimated the kinematic model and posture for each volume. Skeleton curves were used to automatically produce kinematic motion. Isomap transformation was used in order to map a set of 3D points describing a human body volume into a "pose-invariant intrinsic space posture". This transformation allowed finding a correspondence between volume points in both Euclidean and intrinsic spaces. By building principal curves in intrinsic space and mapping back to the volume feature produces a skeleton curve. An a priori body model is not used (Figure 2-18). This is a fast technique to be applied to image sequences and manages to define the posture without the help of an a priori model. However, it is not known if the technique gives position and orientations of body segments accurately enough for clinical purposes since no validation is presented. The technique can be used as the initialization step of the marker-based techniques.
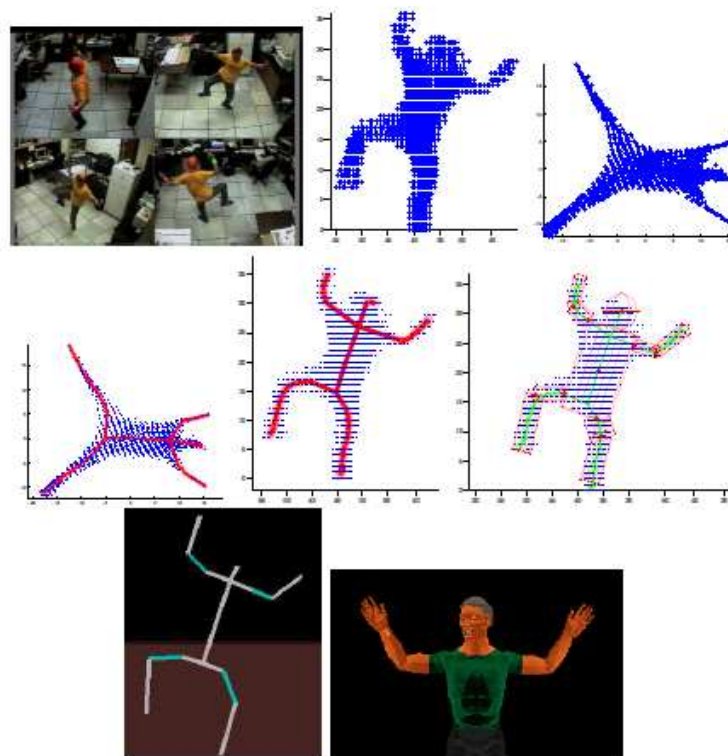
Fig. 2-18. The outline of the approach. (1) A human viewed in multiple cameras is used to build (2) a Euclidean space point volume of the human. This volume is transformed into (3) an intrinsic space pose invariant volume and its (4) principle curves are found. The principal curves are projected back into Euclidean space to provide (5) a skeleton curve. The skeleton curve is used to determine (6) the posture of the human. Using the postures of a volume sequence, (7) the kinematic motion of the human is found and (8) actuated on the Adonis humanoid simulation (from Chu et al., 2003).

In another application, Grauman et al. (Grauman et al., 2003) presented an image-based approach to infer 3D structure parameters. By using a probability density of multi-view silhouette images with known 3D structure parameters, a probabilistic shape and structure model was created (Figure 2-19). This probabilistic model was merged with a model of the observation uncertainty of the silhouettes seen in each camera to compute Bayesian estimate of structure parameters. This was the first study where an image-based statistical shape model was used for the inference of 3D structure. Besides, by using a computer graphics model of articulated bodies, a database of views augmented with the known 3D feature locations were formed in order to learn the image-based models from known 3D shape models. This synthetic training set removed the necessity of labelled real data. The study's novelty was the use of a probabilistic multi-view shape model to narrow the possible object shape and configurations to those that are more "probable" given the class of the object and the current observation.
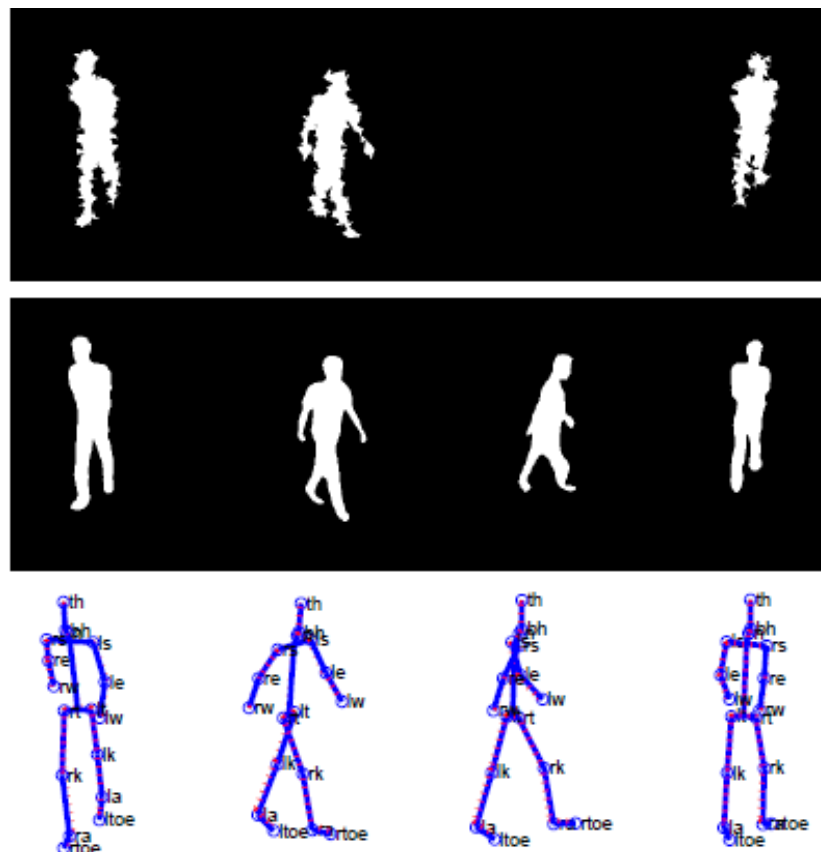
Fig. 2-19. Top row shows noisy input silhouettes, middle row shows contour reconstructions, and bottom row shows inferred 3D joint locations (solid blue) and ground truth pose (dotted red) (from Grauman et al., 2003).

2.2.3 Goal-Oriented Classification of the Markerless Studies

In this review, the mentioned markerless studies were analyzed in terms of their shape representations, existence of models and application space. Yet it is also possible to make a classification based on the goal of these studies.

Markerless studies can be classified in terms of the goal of the applications: 1) studies that merely provide graphical representations of the human body, 2) studies that aim to estimate joint kinematics. The table below (Table 2-2) depicts this classification and provides an overview of the abovementioned markerless studies together with their validation information.

	Graphical Representation	Joint Kinematics Estimation	Notes	Validation
2D	Ju et al., 1996	-	-	-
	Deutscher et al., 2000			Comparison with Standard Condensation Algorithm
	Mori and Malik, 2006			CMU MoBo database
	Elgammal and Lee, 2004			Georgia Tech gait data with ground-truth
		Goffredo et al., 2009	Hip, knee and ankle angles are estimated.	Comparison with the results of the marker-based study by Gross et al. (Gross et al., 1998)
3D	Bottino and Laurentini, 2001			
		Corazza et al., 2006	Shoulder, hip, knee, ankle angles are estimated	Comparison with the ground-truth provided by virtual environment
	Bregler and Malik, 2002			Qualitatively validated by the Muybridge sequence, but quantitative information regarding the parameters are not presented.
		Cheung et al., 2005	Shoulder, hip, and elbow angles are estimated	Synthetic sequences with ground-truth
	Sundaresan and Chellappa, 2006			
	Gagalowicz and Quah, 2009			
	Chu et al., 2003			
	Grauman et al., 2003			

Tab. 2-2. Classification of the previously presented markerless techniques and their validation information.

2.2.4   Limitations

Markerless techniques are quite promising in terms of proposing an alternative to marker-based techniques, easy setup and low-cost solution. However, the use of markerless techniques to capture human movement for biomechanical or clinical applications has been restricted by the complexity of acquiring accurate 3D kinematics. The general problem of estimating the free motion of the human body or more generally of an object without markers, from multiple camera views, is under-constrained when compared with marker-based systems.

Existing computer vision approaches focusing on markerless movement analysis may have great potential to be used in biomechanical applications, but most of them have not been validated for these applications. Evaluation of these approaches in terms of applicability to clinical applications is essential.

For the purpose of enhancing computational performance, simple or generic models of human body with fewer joints or reduced number of degrees of freedom are frequently used. Nevertheless, detailed and accurate representation of 3D joint mechanics is required in biomechanical and clinical applications.

Another challenge for the whole-body movement capture is the non-rigid nature of human body segments and the variability of human motion, the presence of self-occlusion or occluding objects (Mündermann et al., 2006). This diversity causes the some predefined parameters to be created or assumptions to be made, which restrict the analysis (Poppe, 2007; Bray, 2001).

To sum up, the field of markerless movement analysis is a promising and active research which will continue to evolve by considering the abovementioned challenges as a roadmap.

CHAPTER 3

AIMS OF THE THESIS

Marker-based human movement analysis suffers from the instrumental errors, soft tissues artefacts and anatomical landmark misplacement. Besides, markers hinder with the subject's movement and an extended setup time is required. To solve these limitations, markerless human movement analysis has been introduced. Even though the markerless techniques mentioned in the previous chapter are promising, they often lack validation and accurate representations, which are crucial for the clinical applications. Therefore, this thesis aims to provide new 2D markerless techniques to overcome the abovementioned difficulties of both systems.

In particular, the following issues are dealt with listed in order of relevance:

1) Development of two different markerless techniques to determine joint kinematics: Two different markerless techniques (cross-correlation-based and skeletonization-based) are implemented.

2) Validation of the proposed markerless techniques with the marker-based techniques: The proposed markerless techniques are validated with traditional stereophotogrammetric marker-based systems by recording the same trials at the same time.

3) Extraction of additional information from anthropometric measurements and garments used during the acquisition (high-cut underwear and ankle socks used as "segmental markers") to be combined to the information extracted with a markerless methodology, implementing a hybrid technique applied to children with CP.

4) Analysis of the influence of the presence of socks on the performance of the markerless technique.

The present work was conducted at the Biomedical Sciences Department of University of Sassari, Sassari, Italy. The experiments of the study in Chapter 5 were done in University of Rome, "Foro Italico", Rome, Italy while the experiments of the studies in Chapter 6 and 7 took place in the Motion Analysis Lab at Spaulding Rehabilitation Hospital, Boston, MA, USA.

CHAPTER 4

IMAGE PROCESSING: REVIEW OF THE ALGORITHMS
EMPLOYED

4.1   Image Segmentation

Vision is a deduction problem where the aim is to find out the source of the outputs from a given model and measurements. The main difficulties in the deduction problem of vision are the abundance of data and the ambiguity on whether a specific data item is a part of the deduction problem or not. To overcome these difficulties, image data is generally represented by grouping the features that highlight its main properties, as in segmentation. There are several algorithms for image segmentation depending on the application. In an application where there is a static background, removing an estimate of the background from the image would be functional as image segmentation. However, when the backgrounds change over time, this approach would not work fruitfully (Figure 4-1 and Figure 4-2) (Forsyth and Ponce, 2003).
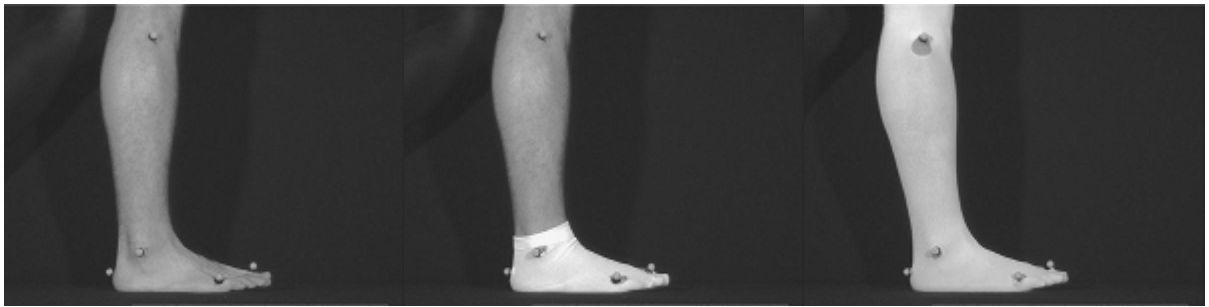
Fig. 4-1. Subject with barefoot, short socks and long socks.

Fig. 4-2. Output of Fig. 4-1 after background subtraction.

Estimating the background using a moving average is a better solution compared to simple background subtraction. Instead of removing the static background, the value of the background is computed by calculating the weighted average of the previous values of the background pixels. Thus, the pixels from the initial frames have a weight of zero and the moving average adapts to the changes in the background. Even though this method can be useful for coarse scale

images, in the dynamic scenes the performance of the adaptation is low (Forsyth and Ponce, 2003).

The motion of the objects in 3D space causes motion on the image plane which can be computed by the displacements of the image points or displacement vectors of the entire image place, which are defined as "optical flow". Even though optical flow is highly used on the images where there is motion, ambiguity stemming from 3D to 2D projection is still a major problem (De Micheli et al. 1993; Fermuller et al., 2001).

Many segmentation methods suffer from the changes in the lighting. Ridder et al. (Ridder et al., 1995) used Kalman Filter to model the pixels, which enabled the lighting changes to be handled. Even though this methods has a "pixel-wise automatic threshold", recovering from the light changes is slow and not effective with bimodal backgrounds. Yet, this method has been successfully used in an automatic traffic monitoring application by Koller et al. (Koller et al., 1994). In another application, Pfinder (Wren et al., 1997), background is modelled by a single Gaussian per pixel and a multi-class statistical model is preferred for the object tracking. In the Pfinder application, the system works well after an initialization phase where the room is empty. However, the performance of the tracker in outdoor scenes is not reported.

Expectation-Maximization (EM) is highly used for image segmentation; Friedman and Russell (Friedman and Russell, 1997) used this method to develop a pixel-wise EM framework for vehicle detection. With this method, pixel values are classified into three separate distributions based on the road color, the shadow color and vehicle color. This system manages the effects of the shadows, but the behaviour of the system when these three distributions are not available, is not known. The performance of the method would be affected if there is a single background or a multiple colored background stemming from motion, shadow or reflectance.

Mixture of Gaussians (MoG) (Hu et al., 2004) is a commonly used method which computes the dynamic features from the image sequence and in this thesis, it has been used as the segmentation method. In this structure, the underlying principle is to describe each single pixel in the image statistically, through a set of Gaussian probability distributions. With this model, the variability of each pixel over time is characterized.

In the study of Stauffer and Grimson (Stauffer and Grimson, 2000), the value of a particular pixel is modelled as a mixture of Gaussians, instead of using a single Gaussian distribution. Gaussians that form the background are determined by calculating the variance of each of the Gaussians of the mixture. Pixels that do not fit to this estimated background are

grouped as foreground pixels. This approach manages "lighting changes, slow-moving objects and introducing or removing objects" successfully (Figure 4-3).
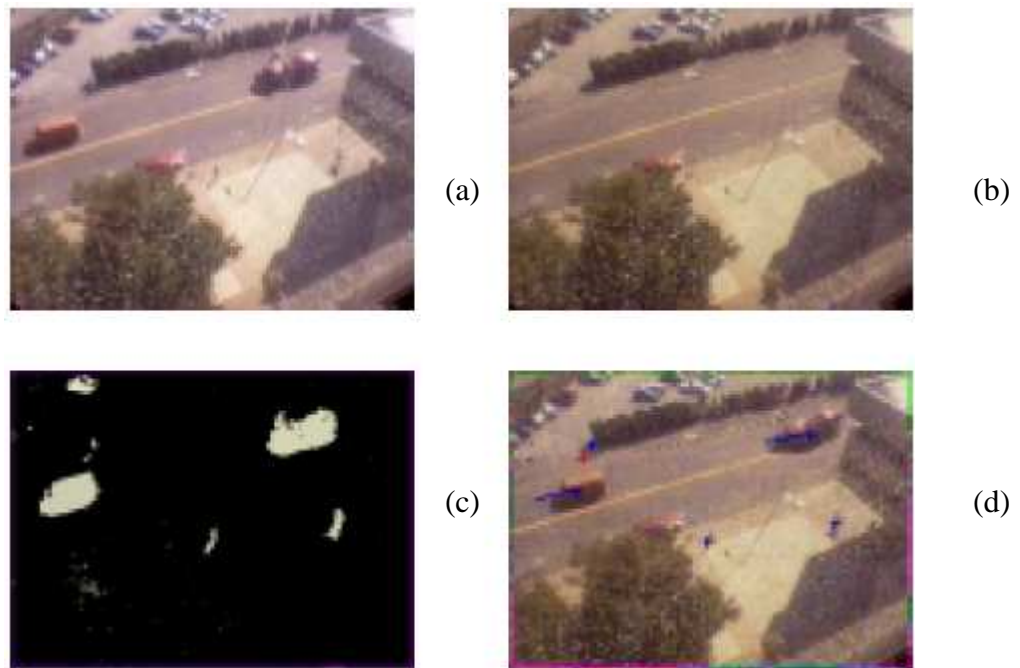
Fig. 4-3. The execution of the program. (a) the current image, (b) an image composed of the means of the most probable Gaussians in the background model, (c) the foreground pixels, (d) the current image with tracking information (from Stauffer and Grimson, 1999).

In this approach, there are two important parameters – α, the learning constant and T, the percentage of the data that has be used for by the background. Every image pixel in the image sequence can be statistically described as a series of values changing over time:

$$\left\{ X_1, \ldots, X_t \right\} = \left\{ I(x_0, y_0, t) : 1 \le t \le T \right\}$$

where $I(x,y,t)$ represents the intensity value of the pixel at position (x,y) and time t, in the image sequence. The latest changes of the intensities can be modelled as the mixture of K Gaussian probability density distributions:

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} \cdot \eta\left(X_t, \mu_{i,t}, \sum\nolimits_{i,t}\right)$$

where K is the number of Gaussian distributions, $\omega_{i,t}$ is the individual weight of each Gaussian at time t (the sum being equal to 1), $\mu_{i,t}$ and $\Sigma_{i,t}$ are the mean the covariance matrix associated with the ith Gaussian at time t.

The general formula of the Gaussian distribution is:

$$\eta\,(X_t, \mu_t, \Sigma_t) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma_t|^{\frac{1}{2}}}e^{-\frac{1}{2}(X_t-\mu_t)^{\mathrm{T}}\Sigma_t^{-1}(X_t-\mu_t)}$$

At every frame, pixel densities are updated due to their intensity values. This update is calculated by changing the Gaussian distribution with a mean value at a distance less than 2.5 times its standard deviation, which is found to be optimal.

The probability density distribution is calculated by modifying the weight associated to each Gaussian according to the following formula:

$$\omega_{k,t} = (1-\alpha)\,\omega_{k,t-1} + \alpha\big(M_{k,t}\big)$$

where $M_{k,t}$ is a binary value which checks if the pixel belongs to one of the Gaussians, α is a parameter which determines the learning capability of the distribution, and the sum of all K weights is equal to 1. The value of α identifies the adaptation speed of the distribution which is an important factor to update the model to the desired temporal frequency. The parameters of the Gaussian are modified according to the following formula:

$$\mu_t = (1-\rho)\mu_{t-1} + \rho X_t$$
$$\sigma_t^{\,2} = (1-\rho)\sigma_{t-1}^{\,2} + \rho(X_t - \mu_t)^{\mathrm{T}}(X_t - \mu_t)$$
$$\rho = \alpha\eta\,(X_t\mid\mu_k,\sigma_k)$$

After the update step is over, the Gaussians are ordered according to the ratio ω/σ. The first Gaussian, which has the lower weight and higher variance, represents the background in the image.

MoG is a highly effective model since it handles the new objects or changes and adapts its background model. The parameter α defines the period to adapt to new objects. Initialization phase is a key step to determine the initial model which can be achieved either by calculating the mean values of each Gaussian or by choosing the most repeated Gaussian as the initial Gaussian. Both of these methods perform well.

Fig. 4-4. Output of Fig. 4-1 after using MoG algorithm.

Learning the correct scene characterization takes few frames, where slowly moving vehicles are merged with the steady objects in the scene, but slowly vanish as the model is updated. The maximal velocity of the objects to be included in the background is calculated by the learning speed of the Gaussian model.

4.2 Cross-Correlation

Cross-correlation, also known as sliding dot product, is a similarity measure of two signals and is commonly used in signal processing, computer vision and image processing. Normalized cross correlation, as in the study of Lewis JP (Lewis, 1995), can be applied in the image processing template matching applications where the image brightness and template can change due to lighting and exposure

First, the distance is calculated as follows:

$$d_{f,t}^2(u,v) = \sum_{x,y} \left[f(x,y) - t(x-u, y-v) \right]^2$$

where f is the image and the feature t positioned at u,v.

$$d_{f,t}^2(u,v) = \sum_{x,y} \left[f^2(x,y) - 2f(x,y)\, t(x-u, y-v) + t^2(x-u, y-v) \right]$$

the term $\sum t^2(x-u, y-v)$ is constant. If the term $\sum f^2(x,y)$ is constant then the cross-correlation term is as below:

$$c(u,v) = \sum_{x,y} f(x,y)t(x-u, y-v)$$

Cross-correlation term represents the similarity factor between the feature and the image. However, this approach has the following disadvantages:

1) When the image energy $\sum f^2(x,y)$ varies with position, matching may fail.
2) The range of $c(u,v)$ is dependent on the size of the feature.
3) The formula is not invariant to lighting changes.

To overcome these disadvantages, the correlation coefficient is calculated:

$$\gamma(u,v) = \frac{\sum_{x,y}\left[f(x,y)-\overline{f}_{u,v}\right]\left[t(x-u,y-v)-t\right]}{\left\{\sum_{x,y}\left[f(x,y)-\overline{f}_{u,v}\right]^2 \sum_{x,y}\left[t(x-u,y-v)-\overline{t}\right]^2\right\}^{0.5}}$$

where $\overline{t}$ is the mean of the feature and $\overline{f}_{u,v}$ is the mean of $f(x,y)$. The formula above is referred as "normalized cross-correlation".

Normalized correlation is a template matching method in which the aim is to find the incidences of a pattern in an image. It is widely used in computer vision application such as stereo vision, motion tracking and image mosaicing. This method is a simple, yet effective method as a similarity measure and easy to be implemented for real-time applications. In this thesis, normalized cross correlation has been used for the studies in Chapter 5 and Chapter 6.

4.3 Skeletonization

Thinning is a morphological operation which removes the selected pixels from the binary images. It is widely used for skeletonization applications where the aim is to sort out the edge detection outputs by reducing the thickness of the lines to single pixels. The main advantages of thinning in image processing are reduction of the data amount of an input binary image and the preservation of the skeleton (Vanajakshi et al., 2010).

Euclidean Distance transforms are commonly used for extracting the medial axis or skeleton of the image. It is the simplest approach for the skeletonization algorithm and is based on extracting the skeleton by finding the pixels furthest from the boundary. Euclidean distance is used for the distance measurements. Even though this approach is faster than the thinning operation, the output may not preserve connectivity (Daya, 2008).

Blum (Blum, 1967) introduced the medial axis transform (MAT) in order to explain biological shape. MAT can be considered as the locus of the center of a maximal disc rolling inside an object. It has been widely used in pattern and image analysis, mold design and path planning. There is a unique MAT for each object and it is possible to reconstruct an object by using its MAT. The MAT can be used in constructive solid geometry (CSG), boundary representation (B-rep) and in applications requiring the abstract representation of the geometry.

The mathematical definition of the MAT is as follows:

"The medial axis (MA), or skeleton of the set D, denoted M(D), is defined as the locus of points inside D which lie at the centers of all closed discs (or balls in 3D) which are maximal in D, together with the limit points of this locus. A closed disc (or ball) is said to be maximal in a subset D of the 2D (or 3D) space if it is contained in D but is not a proper subset of any other disc (or ball) contained in D. The radius function of the MA of D is a continuous, real-valued function defined on M(D) whose value at each point on the MA is equal to the radius of the associated maximal disc or ball. The MAT of D is the MA together with its associated radius function." The boundary and the corresponding MAT of an object are shown in Figure 4-5. (Ramanathan and Gurumoorthy, 2002).

Fig. 4-5. Boundary and its medial axis (from Ramanathan and Gurumoorthy, 2002).

In the present work, MAT has been used for the study in Chapter 6. Below, Figure 4-6 shows the skeletonised outputs of Figure 4-1.

**Fig. 4-6.** Skeletonized outputs of Fig. 4-1 after MAT algorithm was applied.

4.4   Convex Hull

The definition of convex hull is as follows: "A subset S of the plane is called convex if and only if for any pair of points $p, q \in S$ the line segment $\overline{pq}$ is completely contained in S. The convex hull $CH(S)$ of a set S is the smallest convex set that contains S. To be more precise,

38

it is the intersection of all convex sets that contain S " (Figure 4-7 and Figure 4-8) (De Berg et al., 2008).
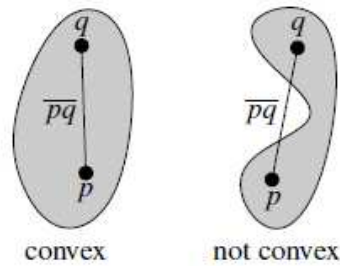
Fig. 4-7. Illustrations depicting the concepts of convex and not convex (De Berg et al., 2008).

The convex hull is an important concept which is used mainly in mathematics and computational geometry. It is used in the analysis of spectrometry data, power diagrams, halfspace intersection, Delaunay triangulation, Voronoi diagrams (Barber et al., 1996).
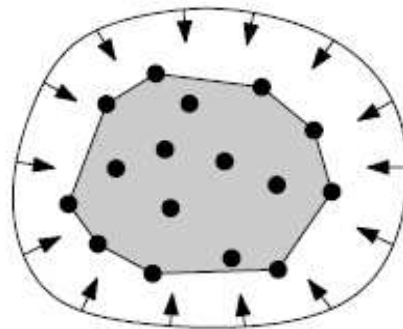
Fig. 4-8. Illustration of the working principle of convex hull method (De Berg et al., 2008).

In the review article by Aurenhammer, other applications of convex hull are listed as: mesh generation, file searching, cluster analysis and image processing (Aurenhammer, 1991) .In the present work, convex hulls are used as described in Chapter 7.

CHAPTER 5

A MARKERLESS ESTIMATION OF THE ANKLE-FOOT COMPLEX 2D KINEMATICS DURING STANCE

1) Surer E, Cereatti A, Grosso E, Della Croce U (2011) A Markerless Estimation of the Ankle-Foot Complex 2D Kinematics during Stance. Gait & Posture, doi: 10.1016/j.gaitpost. 2011.01.003, 2011 (in press).

2) Surer E, Cereatti A, Della Croce U, Grosso E (2009) Evaluation of 2-D Joint Kinematics During Stance Using Markerless Shank and Foot Segmentation, Modelling and Tracking. Gait & Posture;, 30(1).

3) Surer E, Cereatti A, Della Croce U, Grosso E (2009) Wearing Socks Does Not Affect Markerless Estimation of Ankle and Foot Kinematics. Gait & Posture;30(1).

4) Surer E, Grosso E, Della Croce U (2009) Shank-foot complex 2-D kinematics during stance using markerless segmentation and body-segment anatomical axes identification. Gait & Posture;30(1).

5) Surer E, Grosso E, Della Croce U (2008) Lower Limb Kinematics during Stance Using Markerless Shank and Foot Segmentation, Skeletonization and Modelling. In: Proc. of 10th International Symposium on 3D Analysis of Human Movement, Santpoort, Netherlands, 28-31 October 2008.

6) Surer E, Grosso E, Della Croce U (2008) Markerless Shank and Foot Segmentation and Skeletonization for Lower Limbs 2-D Movement Analysis During the Stance Phase in Gait. In: Proc. of Primo Congresso Nazionale di Bioingegneria, Pisa, Italy, 3-5 July 2008.

5.1 Introduction

Quantitative gait analysis is generally carried out by mounting retro-reflective markers on the skin of subjects and reconstructing the three-dimensional (3D) position in the laboratory space by means of stereophotogrammetric systems. The use of stereophotogrammetry requires the placement of markers on selected points of the body segments. Typically, an expert operator spends a considerable amount of time in attaching the markers. In order to do so, subjects are often asked to remove their clothing, including shirts, shoes and socks, sometimes causing feelings of uneasiness. A technique less time consuming, requiring less expertise, discomfort-free to the subject would be favorably accepted in clinical applications.

Markerless techniques (*Ml*) have been recently presented (Mündermann et al., 2006) and may potentially play an important role in this respect. Different approaches have been proposed for estimating the human body kinematics based on an *Ml* approach. Corazza et al. (Corazza et al., 2006) employed a full 3D body model of the subject to be matched with the visual hull by using Simulated Annealing. Bregler and Malik (Bregler and Malik, 1998) used twist and exponential maps to define the motion of their model. Chu et al. (Chu et al., 2003) proposed a model-free approach by describing the human body with a set of points to be mapped to a pose-invariant intrinsic space posture. The use of 3D *Ml* techniques in the clinical and research fields has been so far limited due to the high computational cost (Azad et al., 2006; Deutscher et al., 2000) and equipment requirements (Mündermann et al., 2006), especially in the full body analysis.

In two-dimensional (2D) quantitative analysis of joint kinematics, *Ml* approaches could possibly be effectively implemented in clinical applications. By using a Cardboard kinematic model, Howe et al. (Howe et al., 2000) modeled the limbs as planar patches and enforced 2D constraints on capturing and analyzing the motion To make the model representation independent of the original image, image descriptors such as silhouettes, edges, color and texture are frequently used in 2D *Ml* approaches (Poppe, 2007). In answering to some specific clinical questions, 3D gait analysis showed that the most significant differences between groups were concentrated in the sagittal joint kinematics (Calhoun et al., 2010; Picelli et al.,

2000; Galli et al., 2010), therefore, in such cases, the information provided by a 2D quantitative sagittal joint kinematics analysis may be sufficient, assuming that the main joint axis remains approximately perpendicular to the image plane throughout the recording of the motor task.

Since the shank and foot complex is key for propulsion and support during locomotion, the analysis of its kinematics provides important information for the diagnosis and treatment of pathologies affecting locomotion (Gage et al., 1995; Chin et al., 2009).

Based on the considerations above, in this study we focus to the sagittal kinematics of the shank and foot complex during the stance phase of walking, with two aims: a) to propose a 2D *Ml* technique and b) to verify if the performance of the proposed technique is affected when the subject walks with socks on (as opposed to barefoot). The proposed method was validated (both in barefoot and socked conditions) by acquiring the same walking trials both with a single camera (used in the *Ml* approach) and a simple marker-based (*Mb*) system.

5.2 Materials and methods

5.2.1 Acquisition Setup

Three healthy subjects (one male and two females; 27, 28 and 28 years old, respectively) were asked to walk at self-paced speed (approximately 0.7 m/s) in two different conditions: barefoot and wearing ankle sport socks. Five trials for each condition were recorded for each subject.

The *Mb* data were acquired simultaneously with the *Ml* data using a six-camera stereophotogrammetic system (Vicon MX, 1.3 Mpixel, 120 frames/s). The measurement volume was $1.5m^3$ ($1.5m \times 1m \times 1m$). The markers were positioned on the head of fibula (HF), on the calcaneus, on the lateral malleolus and on the first and fifth metatarsal heads. The marker positions were projected to 2D in the estimation of the joint kinematics. A force platform (AMTI, Watertown, MA) was also used to detect heel strikes and toe offs.

The *Ml* estimate of the sagittal plane kinematics of the shank and foot complex required the execution of the following steps (Figure 5-1).
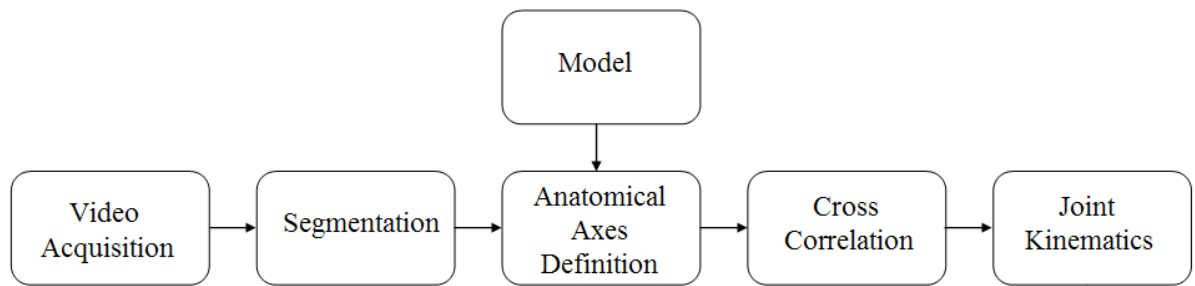
Fig. 5-1. The block diagram of the processing steps used in the proposed markerless technique.

### 5.2.2   Video Acquisitions

Sagittal view images of the shank and foot complex of the subjects were acquired during the stance phase with a single digital camera (Basler A101f, resolution: 800×600 pixels). The camera, acquiring at 15 frames/s, was positioned laterally to the subject to obtain a sagittal view of the shank and foot during the stance phase. The measurement plane was $1.5m^2$ (1.5m × 1m). Sample frames of acquisitions with the subject barefoot and wearing ankle socks are shown in Figure 5-2 (a and b, respectively).
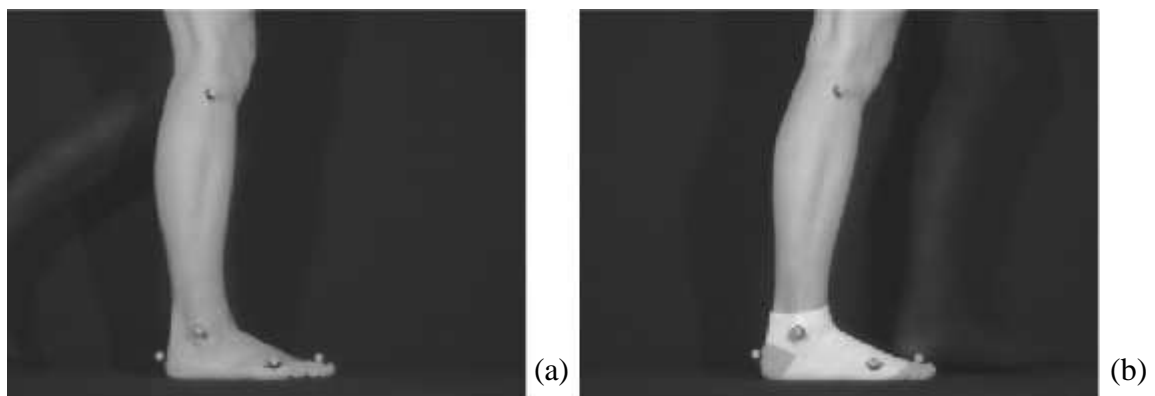
Fig. 5-2. Reference images of a subject barefoot (a) and wearing ankle socks (b).

5.2.3   Segmentation

The objective of the segmentation procedure is to subtract the background from the moving body parts on the acquired image frames. To accomplish this aim, the Mixture of Gaussians method (MoG) (Stauffer and Grimson, 1999) was applied. The MoG is a widely used statistical method, particularly effective when dealing with moving objects and illumination changes (Lagorio et al., 2008).

Using the combination of a finite number of Gaussian distributions rather than a single Gaussian distribution, enables the MoG method to optimally characterize the statistical content of image sub-parts, handling sequences of images with overlaps and occlusions. By calculating the variance of each Gaussian in the mixture, the correct statistical distribution of the corresponding background is determined. Pixel values that do not fit the background distributions form the foreground.

The segmented outputs of the images can be seen in Figure 5-3 (a and b, respectively).

(a)

(b)
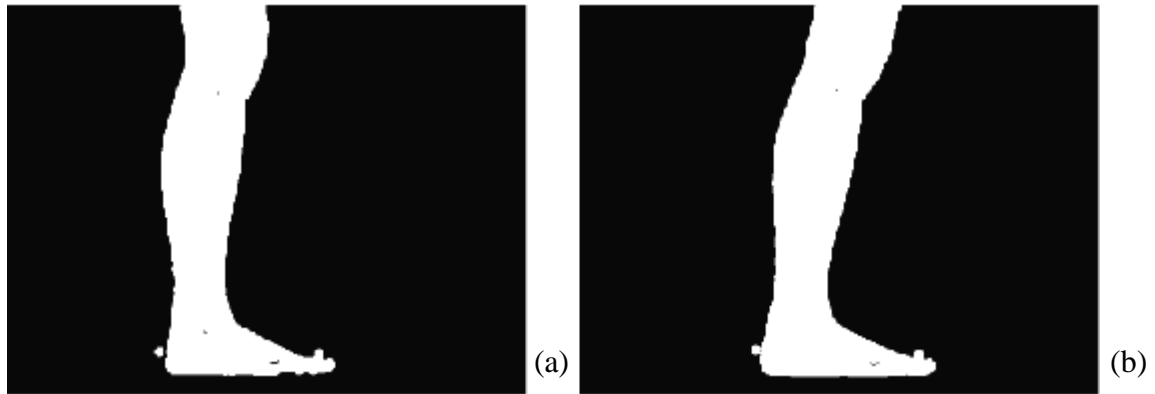
Fig. 5-3. Relevant segmented output images (a and b).

5.2.4   Multi-Segment Model

The multi-rigid body model adopted for the kinematic analysis consisted of three rigid body segments: shank (tibia and fibula), rearfoot (tarsus and metatarsus) and forefoot (phalanges), connected by cylindrical hinges. The model was characterized by two degrees of freedom: the ankle plantar/dorsi-flexion angle (α) and the rearfoot-forefoot flexion/extension angle (β) as described in Figure 5-4a.

5.2.5   Anatomical Axes Definition

For each subject, anatomical axes were defined using a reference image (RI) extracted from the video recordings of the shank and foot complex with markers on at mid stance. In general, this phase does not require the use of markers, although it might be helpful to mark the anatomical landmarks on the subject's skin after palpation and prior to the RI acquisition. In this study, to validate the proposed *Ml* technique, the same markers used for the *Mb* acquisitions were also used to identify anatomical landmarks in the RI.

(a)   (b)
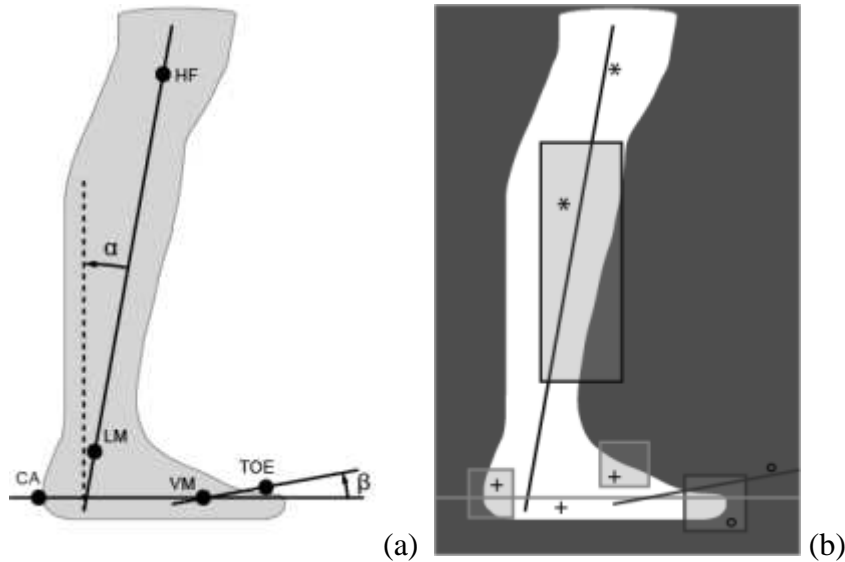
Fig. 5-4. Definition of ankle joint angle (α) and rearfoot-forefoot joint angle (β) (HF=head of fibula, LM=lateral malleolus, CA=calcaneous, VM=fifth metatarsal head, TOE= big toe) (a). Reference patches and anatomical axes (b). Symbols *, + and □ represent the anatomical axes and reference patches of shank, rearfoot and forefoot, respectively.

An axis for each of the three segments was identified from the RI: the shank axis was identified as the axis passing through the head of fibula and the lateral malleolus, the rearfoot axis was made to join the calcaneous to the fifth metatarsal head, the forefoot axis passed through the fifth metatarsal head and the toes (Figure 5-4a). Moreover, in the RI, patches containing portions of the body segments, expected to show minimum changes in shape during movement, were identified (Figure 5-4b). Axes and patches belonging to the same body segment were assumed to be rigidly connected.

5.2.6 Cross-Correlation

Image cross-correlation was applied to the selected image patches to track the movement of the body segments. In image processing, cross-correlation is a well known and effective technique for template matching (Goshtasby et al., 1984). The cross-correlation coefficients, usually normalized in the range [0;1], express similarity between two different images: 1 represents full similarity, 0 no similarity.

The patches identified in the RI were searched in all the images of the sequence, one at the time, over a search space stemming from the possible translations and rotations. The patch in the searched image that showed the highest cross-correlation value was selected (Lewis, 1995). Cross-correlation coefficients were first computed translating the template along the

vertical and horizontal axes of the whole picture. When a first maximum was found, then the patch was rotated by up to ten degrees and a second maximum was computed for each rotated patch over a limited search area. The highest cross-correlation value defined the position and orientation of the searched patch. The selected patch then became the new reference image for the succeeding frame and the whole procedure was repeated until the last frame.

5.2.7   Data Analysis

The duration of the stance phase was defined as the number of the *Ml* frames between heel strike and toe off. Since *Ml* data and the *Mb* data were not synchronized, an *ad-hoc* procedure was implemented. For each trial, heel-strike and toe-off event frames were selected. This was done through visual inspection on the *Ml* frames and using force platform data for the *Mb* acquisitions. Since the *Ml* frame rate was $1/8^{th}$ of the *Mb* frame rate, the *Mb* frame best matching the *Ml* event was defined as the *Mb* frame with α and β values most similar to the values obtained from the marker locations in *Ml* event images.

Once the *Ml* and *Mb* frames were aligned in time, the *Mb* sagittal joint kinematics were down-sampled to the *Ml* frame rate for comparison purposes.

Ml and *Mb* joint kinematics were compared as follows.

- In order to account for offsets between the *Ml* and the *Mb* ankle kinematics, the absolute difference between their mean values ($\overline{\alpha}_{Ml}$ and $\overline{\alpha}_{Mb}$, respectively) over the stance phase was determined.

$$\Delta(\alpha) = \left| \overline{\alpha}_{Ml} - \overline{\alpha}_{Mb} \right|$$

- In order to account for pattern differences, for each time series the deviation from the mean values were determined:

    $\alpha'_{i\_Ml} = \alpha_{i\_Ml} - \overline{\alpha}_{Ml}$, and $\alpha'_{i\_Mb} = \alpha_{i\_Mb} - \overline{\alpha}_{Mb}$

    where the subscript *i* refers to the ith frame;

    and the Root Mean Square Deviation (RMSD) of the α'_{i_Ml} values from the α'_{i_Mb} values, was estimated:

$$RMSD(\alpha'_{Ml}, \alpha'_{Mb}) = \sqrt{\frac{\sum_{i=1}^{n}(\alpha'_{i\_Ml} - \alpha'_{i\_Mb})^2}{n}} \ .$$

The same processing was applied to the rearfoot-forefoot joint kinematics (β).

To verify if measurements obtained with the two techniques were comparable to the intra-subject variability, similar indexes were introduced to estimate the intra-subject variability of

the ankle kinematics obtained with the *Mb* measurements: $\Delta_V(\alpha) = \left| \overline{\alpha}_a - \overline{\alpha}_b \right|$ and

$$RMSD_V(\alpha'_a, \alpha'_b) = \sqrt{\frac{\sum_{i=1}^{n} (\alpha'_{i\_a} - \alpha'_{i\_b})^2}{n}}$$, where *a* and *b* represent any two of the five trials

performed per condition per subject. Similar indexes were introduced for the β angle measurements.

The maximum values of $\Delta_V(\alpha)$ and $RMSD_V(\alpha'_a, \alpha'_b)$ were compared to the maximum values of $\Delta(\alpha)$ and $RMSD(\alpha'_{Ml}, \alpha'_{Mb})$, and similarly was done for the indexes regarding the angle β.

5.3   Results

A representative sagittal joint kinematics during the stance phase, estimated with the *Ml* technique and the *Mb* technique is reported in Figure 5-5 for the three subjects.

The differences $\Delta(\alpha)$ and $\Delta(\beta)$ between the mean angle values for each trial and the maximum values of the intra-subject variability indexes $\Delta_V(\alpha)$ and $\Delta_V(\beta)$ are reported in Table 5-1. Table 5-2 shows the values of the *RMSD* of the joint angles α and β, $RMSD(\alpha'_{Ml}, \alpha'_{Mb})$ and $RMSD(\beta'_{Ml}, \beta'_{Mb})$, the maximum values of the intra-subject variability indexes $RMSD_V(\alpha'_a, \alpha'_b)$ and $RMSD_V(\beta'_a, \beta'_b)$ obtained for the three subjects in barefoot and socked conditions. No noticeable differences were found between barefoot and socked trials.

5.4   Discussion and Conclusion

3D marker based motion capture systems are commonly used for estimating joint kinematics in clinical contexts. This approach requires expensive equipment and a high level of expertise to operate, limiting its use in clinical routine. Unfortunately, valid alternatives are not available yet. However, in some specific clinical issues, the determination of 2D joint kinematics is sufficient.
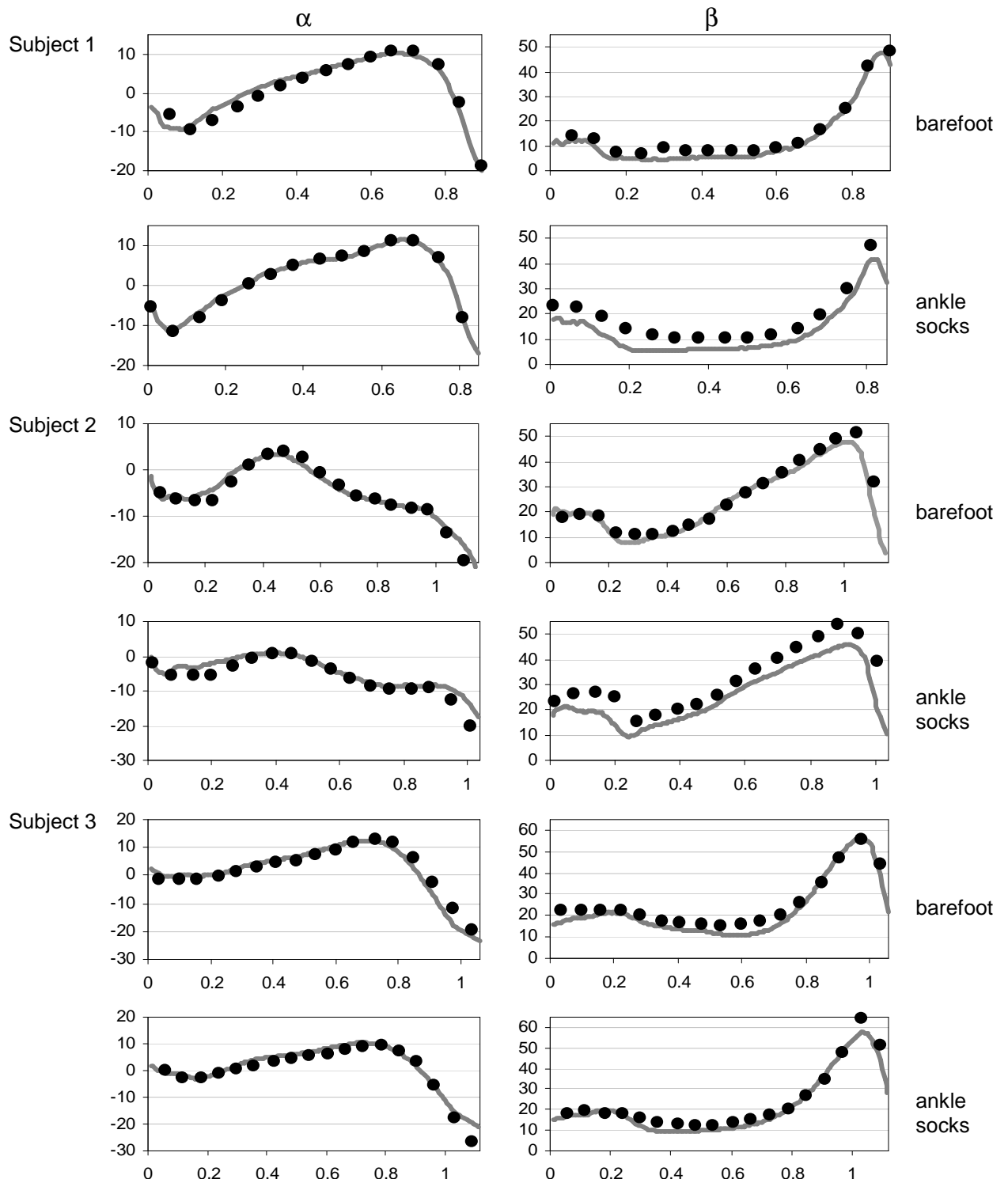
Fig. 5-5. Ankle kinematics (α) and rearfoot-forefoot joint kinematics (β) of three subjects (#1, #2, #3), obtained both in barefoot and socked conditions during the stance phase of walking. The solid lines are the joint kinematics obtained using the marker-based technique while dots are the joint kinematics obtained using the proposed markerless technique. Quantities are expressed in degrees.

[deg] angle		trial	Subject #1 barefoot	ankle socks	Subject #2 Barefoot	ankle socks	Subject #3 barefoot	ankle socks
A	Δ	1	0.2	0.2	1.7	1.1	0.4	1
		2	0.2	0.7	2.5	0.8	0.2	0.6
		3	0.9	0.2	0.3	1.5	0.2	0.6
		4	1.5	0.8	0	2.2	1.1	1.4
		5	0	1.6	0	0.4	0.8	1.7
		average	0.6	0.7	0.9	1.2	0.6	1.0
	Δ_V	max	2.5	3.3	1.6	1.9	2.1	1.8
β	Δ	1	4.2	5.2	0.7	5.6	0.8	5.8
		2	2.4	3	6.9	8.2	2.4	3.1
		3	5.8	3.4	2.7	6	2.5	1.4
		4	5.4	3.2	0.6	6.9	1.5	4.2
		5	2.4	4.5	1.8	4.5	0.4	3.7
		average	4.0	3.9	2.5	6.2	1.5	3.6
	Δ_V	max	2.9	2.5	5.9	6.5	4.4	3.4

Tab. 5-1. The absolute difference (Δ) between the mean values during stance of ankle (α) and rearfoot-forefoot (β) angles for each trial, condition and subject, obtained with marker-based and markerless techniques. Trial averages are reported in bold (maximum values in gray cells). In the grayed area the maximum intra-subject variability values (from marker-based measurements) of the absolute difference (Δ_V) are reported.

[deg] angle		trial	Subject #1 barefoot	ankle socks	Subject #2 barefoot	ankle socks	Subject #3 barefoot	ankle socks
A	RMSD	1	1.6	0.6	1.7	1.6	3.4	1.7
		2	1.8	1.4	1.7	1.5	2.8	3
		3	2.6	0.7	2.1	1.3	2.4	2
		4	1.1	1.5	1.3	0.8	3.1	2.9
		5	0.7	2.7	1.1	2.4	2.7	2.2
		average	1.6	1.4	1.6	1.5	2.9	2.4
	RMSD$_V$	max	4.1	2.7	2.8	3.0	2.6	4.2
β	RMSD	1	2	1.2	2.8	4	5	2.3
		2	1.3	3.2	7.1	3.6	3	3.5
		3	3.7	1.7	3.7	5.1	2.3	2.6
		4	2.8	3.8	3.2	3.4	3.5	3.1
		5	1.6	3	3.8	3.4	3	2.7
		average	2.3	2.6	4.1	3.9	3.4	2.8
	RMSD$_V$	max	3.2	3.9	6.6	6.5	3.1	6.5

Tab. 5-2. Root Mean Square Deviation (RMSD) estimated during stance of the markerless joint kinematics values from the marker-based joint kinematics (α and β) values. Trial averages are reported in bold (maximum values in gray cells). In the grayed area the maximum intra-subject variability values (from marker-based measurements) RMSD$_V$ values are reported.

In this study, we presented a low-cost, low-discomfort markerless technique along with a preliminary validation and reliability assessment in estimating the sagittal kinematics of the shank and foot complex during the stance phase of normal walking.

The proposed technique requires a series of steps: video acquisition, segmentation, multi-rigid body model definition, anatomical axes definition and cross-correlation.

The segmentation outputs indicated that the algorithm chosen for segmenting the images had limited sensitivity to the presence of socks, suggesting that if subjects wore socks during the trials, results would not be affected. This can represent an advantage of the use of the *Ml* technique when analyzing the gait of subjects feeling more comfortable walking with socks than barefoot.

The ankle kinematics (α) estimated with the *Ml* and *Mb* techniques showed very similar results throughout the stance phase (both Δ and *RMSD* values are in general within the intra-subject variability). Differently, for the rearfoot-forefoot kinematics (β), Δ showed values about two degrees higher than the relevant intra-subject variability values ($Δ_V$) in all three subjects (except for the barefoot condition of subject #3). On the converse, the *RMSD* values for the β angle were in most cases within the relevant intra-subject variability index ($RMSD_V$). The different results obtained for the angle β are most probably due to the small size of the forefoot segment and consequently, to the lack of reliability in identifying the anatomical axis using either technique. The larger intra-subject variability determined for the angle β increases the chances of having larger differences in the joint kinematics estimated with the two techniques.

In general, the accuracy and precision of both *Ml* and *Mb* methods suffer from body segments of reduced size. *Mb* techniques perform better in identifying body segments orientation when segment markers are farther from each other (i.e. larger body segments). Similarly, *Ml* techniques may use a larger number of pixels to estimate the orientation of a large body segment. However, in general, while *Mb* techniques use a minimum number of points to describe the segment kinematics (typically three or four points), it is reasonable to expect that future *Ml* techniques may fruitfully use the redundancy of the information carried by the hundreds of pixels used to estimate the segment kinematics and may increase its precision. For instance, since markers are often located over a layer of soft tissues, near a joint or over an active muscle, their movement relatively to the underlying bone introduces errors in the estimation of joint kinematics (Cappozzo et al., 1995; Leardini et al., 2005). The redundant number of points used by *Ml* techniques in determining segment kinematics could potentially reduce such errors.

The *Ml* technique employed in this study suffers of the same limitations of any 2D kinematics analysis of gait performed with a single camera. They are mostly related to the impossibility a) of describing the out-of-plane joint kinematics, b) of obtaining a bilateral analysis, c) of describing segment deformity and d) of keeping image plane and sagittal plane parallel. The last limitation has a limited effect on the resulting joint kinematics for small angles between the two mentioned planes (a 10 degrees angle between planes generates a 1.5% difference in the sagittal joint kinematics estimate).

From an algorithmic standpoint, the chosen *Ml* technique shows limitations to be overcome for increasing its potential in clinical applications. To analyze the sagittal kinematics of pelvis and lower limbs during the entire gait cycle, the improved *Ml* technique should cope with complex backgrounds, shadows and occlusions. In this study, background subtraction was simplified by covering the contra-lateral leg with a long black sock. A more robust cross-correlation and/or image processing technique such as deformable contours (Shahrokni et al., 2005) could help in solving the problem. Moreover, the processing time of the proposed *Ml* technique in the current version implemented in MATLAB® (MathWorks, Natick, MA, USA) needs to be reduced to be fruitfully used in clinical applications (currently about 15 minutes are required to process a trial).

Finally, in order to assess the differences in the sagittal joint kinematics, the two techniques had to be registered at a reference point in time. This required the use of the markers located on anatomical landmarks for the definition of anatomical axes in the *Ml* technique. As a consequence, this study does not provide information regarding the discrepancy (an offset) due to different ways of calibrating anatomical landmarks (from a reference image in an *Ml* technique as opposed to palpation in an *Mb* technique). A reliable automatic anatomical axes identification procedure from the *Ml* images would increase the robustness of the proposed technique.

The performance of the *Ml* technique proposed to estimate 2D joint kinematics is promising for future use in clinical settings. In fact, the acquisition of movement data without the need of attaching markers to the subject's skin, and yet obtaining results comparable to those obtained with a simple marker based technique, represents an important step toward the design of an acquisition system for clinical use. Such a system could also be easy to configure and operate and most probably relatively affordable.

CHAPTER 6

MEASUREMENT OF KNEE FLEXION/EXTENSION USING A 2D MARKERLESS TECHNIQUE

1) Surer E, Kasi P, Cereatti A, Bonato P, Della Croce U (2010) Measurement of Knee Flexion/Extension Using a 2-D Markerless Technique. In: Proc. of Secondo Congresso Nazionale di Bioingegneria, Turin, Italy, 8-10 July 2010.

6.1 Introduction

Several measurement tools have been used to quantify knee joint angles and knee movements. The selection of a measurement tool depends on the purpose of testing (Miller, 1985) and psychometric properties such as reliability and validity (Rothstein et al., 1983). To be valid, tools for quantifying knee position and movement need to produce minimal measurement error. Measurement error can arise from the tool, the tester or from variability in the performance of the individual (Stratford et al., 1984; Russek 2004).

Assessment of knee flexion/extension is typically done with a number of measurement tools: electro-goniometers, inertial sensors, 2D and 3D marker-based motion analysis systems (Piriyaprasarth and Morris, 2007). Most of the traditional clinical analysis methods involve applying sensors to the patients, which may cause discomfort and hinder the natural movement. Besides, they necessitate expertise to operate and expensive hardware.

To overcome the abovementioned limitations, in this study, a 2D markerless technique is proposed to measure knee flexion/extension. This study is an enhancement to our previous study (Surer et al., 2011), which is explaind in detail in the Chapter 5 of this thesis.

The focus of the first study was the analysis of the 2D kinematics of the ankle-foot complex using a single lateral view by defining a three-segment model and tracking its movement using selected patches rigid to the segment. To simplify the problem, a black background was used and the subjects wore long black socks to cover the contra-lateral leg.

In this study, a 2D markerless technique is developed, in which the presence of a more complex background was dealt with by using an adaptive statistical background subtraction model and the occlusions were dealt with by defining an additional patch per segment to be used when the occlusion occurs on the main patch. The proposed technique is validated with a traditional stereophotogrammetric system and the results of the proposed technique are comparable to those obtained with the stereophotogrammetric system.

6.2 Materials and methods

The material and methods used in the proposed technique is very similar to those used in the previous study. Still, for the purposes of integrity, an overall summary of the similarities between the studies will be presented in this chapter as well.

6.2.1 Acquisition Setup

Two CP children (one female and one male, 14 and 7 yrs, respectively) were asked to walk at self-paced speeds. Five trials from each direction were recorded for each subject. During the acquisitions, subjects wore white high-cut underwear and white ankle socks. One of the subjects prepared for the study can be seen in Figure 6-1.

Fig. 6-1. Subject prepared for the study. For validation purposes, retro-reflective markers were mounted on the subject.

Marker-based (*Mb)* data were recorded simultaneously with the *Ml* data using a six-camera stereophotogrammetic system (BTS® SMART-D stereophotogrammetric system, 640x480 pixels, 60 Hz). The infrared filter of one of the cameras was removed and modifications on the shutter and strobe timing were made by the manufacturer. The remaining five cameras

were used for the marker based acquisitions only while the modified camera acquired the lateral view of the subject as grayscale image. During the analysis, only the leg in focus was taken into consideration.

"Simple Davis" protocol of the BTS® SMART-D system were used for the positioning of the markers, so they were positioned on the sacrum (between PSIS), right and left ASIS, greater trochanter (GT), femoral condyle, head of fibula, lateral malleolus and fifth metatarsal joint. Lateral bars were used between the greater trochanter and femoral condyle and between head of fibula (HF) and lateral malleolus. A force platform (AMTI, Watertown, MA) was also used in order to detect heel strikes and toe offs.

The *Ml* sagittal plane kinematics estimate of the lower limbs necessitated the execution of the following steps.

6.2.2   Segmentation

As in the previous study, segmentation of the lower limbs was performed to separate the background information from the regions of interest using MoG method. MoG is an adaptive background modelling method where the parameters of Gaussians are updated to separate the background from the moving foreground (Stauffer and Grimson, 1999).

While a simple black background was used for the previous study, in this study, a more complex background is preferred. Besides, the contralateral leg is not covered with black socks. Despite of the complex background, the results of the segmentation were quite successful. The occlusions stemming from the contralateral leg were handled in the model and axes definition step.

Figure 6-2 shows a segmentation sample. As for the previous study, the proposed technique was validated acquiring the same trials with both the single camera used in the proposed markerless technique and a traditional stereophotogrammetric system.

6.2.3   Model and Axes Definition

The multi-segment model adopted for the kinematics analysis consisted of two different rigid segments – shank and thigh. Anatomical axes were visually identified and patches in a reference frame were selected.

To manage occlusions – portions of segmented images composed by the overlapping of the two legs – two patches rigidly connected are defined (instead of one) per segment, one on the front edge of the segmented image and one in the rear edge.

The patch on the front edge of the segments is normally considered active except when an occlusion occurs in the front edge. Then, the rear patch automatically becomes the active patch until the occlusion ends (Figure 6-3). Both patches are also rigidly connected to the segment anatomical axis and are the portions of the images used in the cross correlation.

Fig. 6-2. Sample image and its segmentation output. As can be seen, complex background did not diminish the performance of the segmentation algorithm.

6.2.4   Cross Correlation

The cross-correlation (Goshtasby et al., 1984), which is a template matching approach, calculates similarity coefficients between two different images. The coefficients express similarity between the images (value 1 represents full similarity, 0 value no similarity) in the search space derived from the possible translations and rotations. It was used in order to track the movements of thigh and shank segments of the leg under analysis.

The patches identified in the reference image (RI) were searched in all the images of the sequence, one at the time, over a search space stemming from the possible translations and rotations. The patch in the searched image that showed the highest cross-correlation value was selected. Cross-correlation coefficients were first computed translating the template along the vertical and horizontal axes of the whole picture.

When a first maximum was found, then the patch was rotated by up to ten degrees and a second maximum was computed for each rotated patch over a limited search area. The highest cross-correlation value defined the position and orientation of the searched patch. The selected patch then became the new reference image for the succeeding frame and the whole procedure was repeated until the last frame.

6.3   Results and Discussion

Table 6-1 shows the results of the two subjects, analyzed for this study. In order to compare the pattern differences between the proposed technique and the marker-based, RMSD results, after removing the mean differences of the curves (offsets). Besides intra-subject variability and correlation outputs of the two techniques are presented for the knee-flexion extension angle (β).
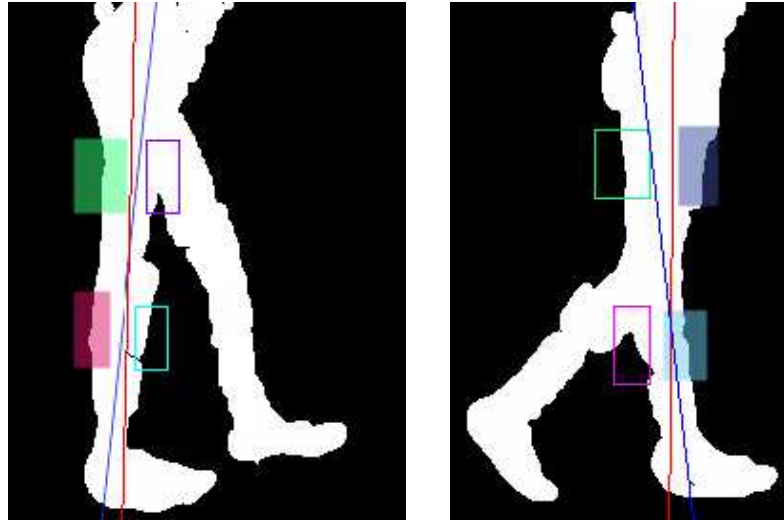

Fig. 6-3. Segmentation outputs with patches and axes are presented. Green and pink are the active patches in the left image. In the right image, purple and blue patches are the active ones; notice that the active patches are swapped.

	Subject #1	Subject #2
	B	
Correlation	0.98	0.98
RMSD	4.4	4.6
RMSDv	26.5	2.3

Tab. 6-1. The correlation between the two techniques, their Root Mean Square Deviation (RMSD), and intrasubject variability ($RMSD_V$).

Figure 6-4 shows the knee-flexion kinematics curves of the two subjects.

Results show that lower limb sagittal kinematics estimates from marker-based and proposed markerless methods are highly comparable, making the latter a potential alternative for clinical use.
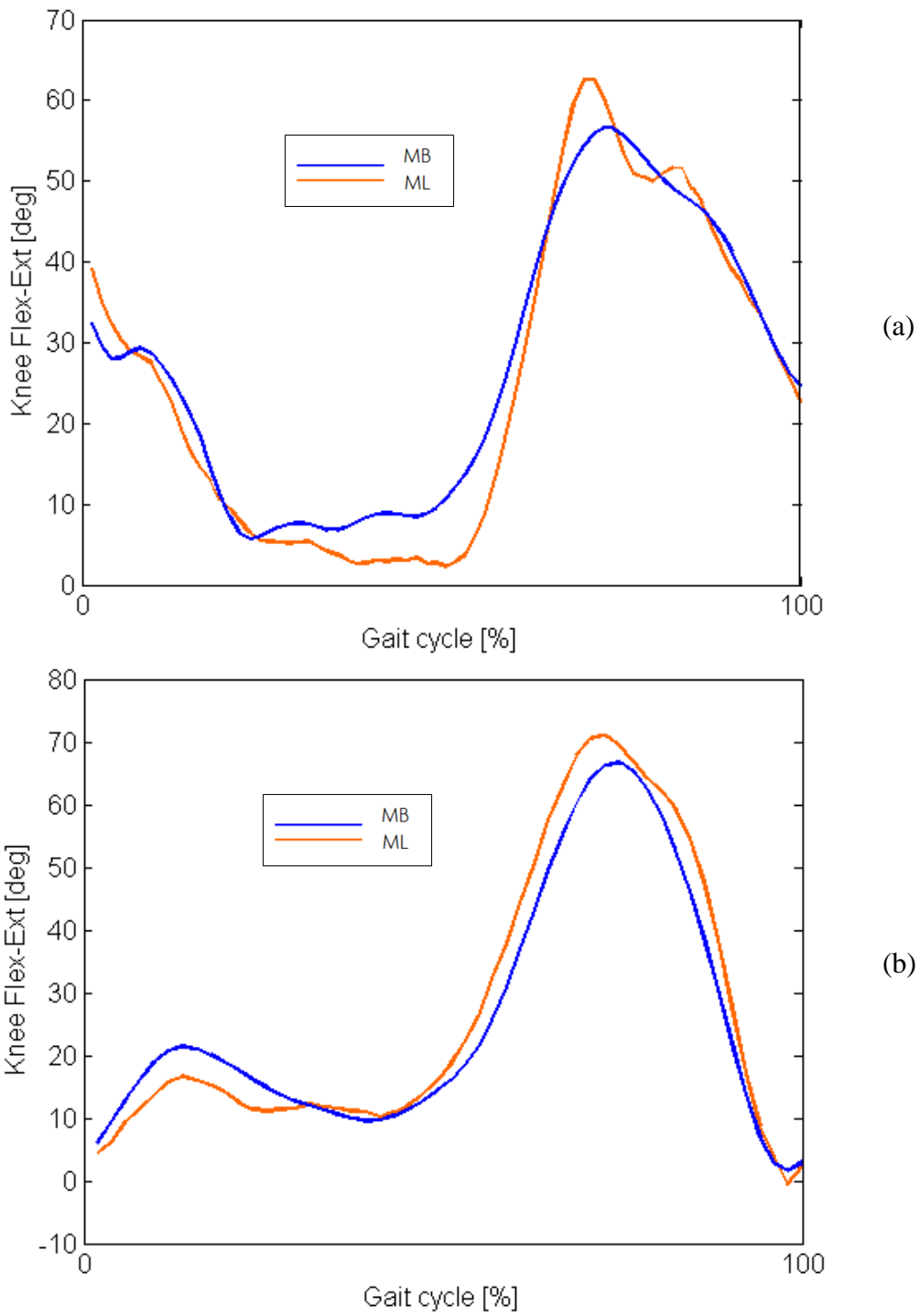
Fig. 6-4 Knee flexion/extension angles of the two subjects are shown in the figure below respectively (Fig. 6-4a – subject #1; Fig. 6-4b – subject #2. Blue lines represent the estimates obtained using the marker-based system (MB) and the orange lines those obtained using the markerless method (ML).

CHAPTER 7

2D GAIT ANALYSIS OF CHILDREN WITH CEREBRAL PALSY USING SEGMENTAL MARKERS AND A MARKERLESS APPROACH

1) Surer E, Kasi P, Cereatti A, Bonato P, Della Croce U (2011) A Hybrid Markerless Approach for 2D Gait analysis: Application to Gait of Children with Cerebral Palsy. (Being submitted to Gait and Posture).

2) Surer E, Kasi P, Cereatti A, Bonato P, Della Croce U (2011) A Hybrid Markerless Approach for 2D Gait analysis: Application to Gait of Children with Cerebral Palsy. (submitted to ISB Conference 2011).

3) Surer E, Kasi P, Cereatti A, Bonato P, Della Croce U (2010) 2-D Gait Analysis of Children with Cerebral Palsy using Markers and a Markerless Approach. In: Proc. of XI Congresso Nazionale della SIAMOC, Ferrara, Italy, 4-7 October 2010.

7.1 Introduction

Cerebral palsy (CP) is defined as a group of movement and posture disorders stemming from the abnormalities that occurred in the developing fetus or infant (Bax et al., 2005). The motor disorders of CP are often accompanied by impairments of sensation, cognition, communication and perception (Marlow, 2004).

The standard assessment approach in CP is to use kinematics, kinetics, electromyography (EMG) and clinical examinations for a comprehensive assessment (Whittle, 1996). In determining the joint kinematics of patients with CP during walking, marker-based movement analysis is frequently used (Baker, 2006).

Marker-based movement analysis has played a significant role in the assessment of CP with its use in documentation, operative planning and postoperative evaluation (Gage et al., 1995). However, attaching markers to the subjects, especially to the children, may cause feelings of uneasiness and may hinder the walking. In addition, they require expensive setup and expertise to operate. For these reasons, a technique less discomfort-free, low-cost and easy to use would be appealing in determining the joint kinematics of patients with CP.

Markerless techniques may provide solutions to overcome the abovementioned difficulties. Several algorithms, such as Simulated Annealing (Corazza et al., 2006) and twist and exponential maps (Bregler and Malik, 1998) in 3D have been proposed for estimating the joint kinematics based on markerless approaches, but they have not been validated for clinical applications. Also, up to now, the use of 3D markerless techniques in the clinical applications has been limited due to the high computational cost (Deutscher et al., 2000) and equipment requirements (Mündermann et al., 2006).

Despite the fact that most studies use sophisticated 3D measurement systems to collect gait data in three planes of motion, using sagittal plane kinematics is sufficient as in the cases of classification of CP (Dobson et al., 2007) and comparison of right and left hemiplegia (Galli et al., 2010).

Markerless approaches could be successfully applied to the clinical applications determining the two-dimensional (2D) sagittal joint kinematics. Cham and Rehg (Cham and Rehg, 1999) used scaled prismatic models (SPM) – i.e. a class of 2D kinematic models – to model the human body in 2D. Howe et al. (Howe et al., 2000) modeled the limbs as planar patches and enforced 2D constraints on capturing and analyzing the motion by using a

Cardboard kinematic model. Even though these studies have great potential for clinical applications, they have nor been validated either.

In this study, we propose a 2D markerless technique defined as hybrid (*Hyb*) since it uses garments (high-cut underwear and ankle socks) to extract additional information as "segmental markers" in defining pelvis and foot segments. The validation of the proposed technique was done by acquiring the same trials also with a marker-based system.

7.2   Materials and methods

7.2.1   Acquisition Setup

Ten CP children (five females, ages = 9 ± 4 yrs) were asked to walk at self-paced speeds. Five trials from each direction were recorded for each subject. During the acquisitions, subjects wore white high-cut underwear and white ankle socks.

Marker-based (*Mb*) data were acquired simultaneously with the *Hyb* data using a six-camera stereophotogrammetic system (BTS® SMART-D stereophotogrammetric system, 640x480 pixels, 60 Hz). The markers were positioned according to the "Simple Davis" protocol of the BTS® SMART-D system, and therefore were positioned on the sacrum (between PSIS), right and left ASIS, greater trochanter (GT), femoral condyle, head of fibula, lateral malleolus and fifth metatarsal joint. Lateral bars were used between the greater trochanter and femoral condyle and between head of fibula (HF) and lateral malleolus. A force platform (AMTI, Watertown, MA) was also used in order to detect heel strikes and toe offs.

The *Hyb* sagittal plane kinematics estimate of the lower limbs necessitated the execution of the following steps.

### 7.2.2   Anthropometric Measurements and Calibration

In order to extract additional information for the use of high-cut underwear and ankle socks as "segmental markers", the following measurements (Figure 7-1a) were made before the acquisitions:

M1: distance between the high-cut of the underwear (hip) and the femoral condyle;

M2: distance between the femoral condyle and the edge of the ankle sock (ankle);

M3: foot length

The camera was positioned laterally and calibrated in order to map *M1* and *M2* to their pixel-lengths on the images. *M3* was used to validate the mapping calculation. A checkerboard of known geometry and size (Figure 7-1b) was used in order to calibrate the camera.

| (a) | (b) |
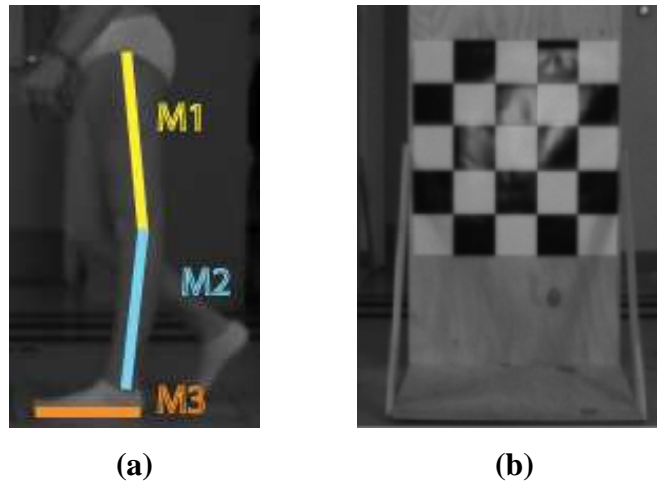
Fig. 7-1. The measurements M1, M2 and M3 (a). Checkerboard (b).

7.2.3   Video Acquisitions

Sagittal view images of the lower limbs of the subjects were acquired with a single BTS® SMART-D camera. In order to acquire the sagittal view of the lower limbs, infrared filter of the camera was removed and modifications on the shutter and strobe timing were made by the manufacturer. The camera, also acquiring at 60 Hz, was positioned laterally to the subject to obtain a sagittal view of the lower limbs. Sample frame of an acquisition with the subject is shown in Figure 7-2a.

7.2.4   Segmentation

As in the studies of Chapter 6 and 7, Mixture of Gaussians (MoG) method (Stauffer and Grimson, 1999) was applied in order to subtract the background from the moving parts on the acquired image frames. The MoG is a widely used statistical method, particularly fruitful when handling moving objects and illumination changes (Lagorio et al., 2008).

Using the combination of a finite number of Gaussian distributions instead of a single Gaussian distribution enables the MoG method to optimally characterize the statistical content of image sub-parts, dealing with sequences of images with complex backgrounds. By

calculating the variance of each Gaussian in the mixture, the correct statistical distribution of the corresponding background is determined.

Example of the output of the segmentation procedure is reported in Figure 7-2b.

7.2.5 Skeletonization

Medial Axis Transform (MAT) is a shape model that characterizes an object by the set of maximal circles that are completely contained in the object. The medial axis consists of the centers of the circles, and can be intuitively thought of as the skeleton of the object (Lam et al., 1997). In order to determine the medial axis, the pixels are eroded from the boundary until a skeleton is formed (Tam and Heidrich, 2003). MAT has numerous applications in visualization, computer graphics and computer vision.

MAT was applied to the segmented image frames in order to "skeletonize" the lower limbs so that joint segments could be extracted.

The skeleton of the segmented frame is shown in Figure 7-2c.

**(a)**   **(b)**   **(c)**
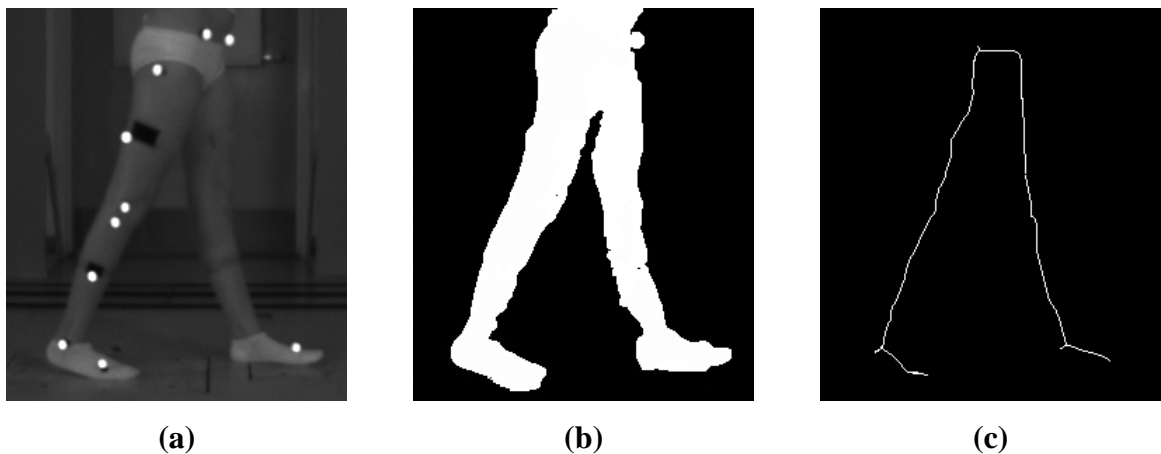
Fig. 7-2. Sample frame (a). Segmented frame (b). Skeletonized frame (c).

7.2.6 Thresholding and Labeling of the Garments

To distinguish the high-cut underwear and the ankle socks, intensity values of the garments were used in thresholding. By checking the visibility of the garments during a sequence of a proper number of frames, the sock on the contra-lateral leg was eliminated. Thresholded garments can be seen in Figure 7-3a.

Labeling is a process which finds the connected components on a binary image and groups them as an object (Haralick and Shapiro, 1992) and is especially practical to operate on related image parts. In order to access garments easily, labeling process was applied.

(a)   **(b)**   **(c)**
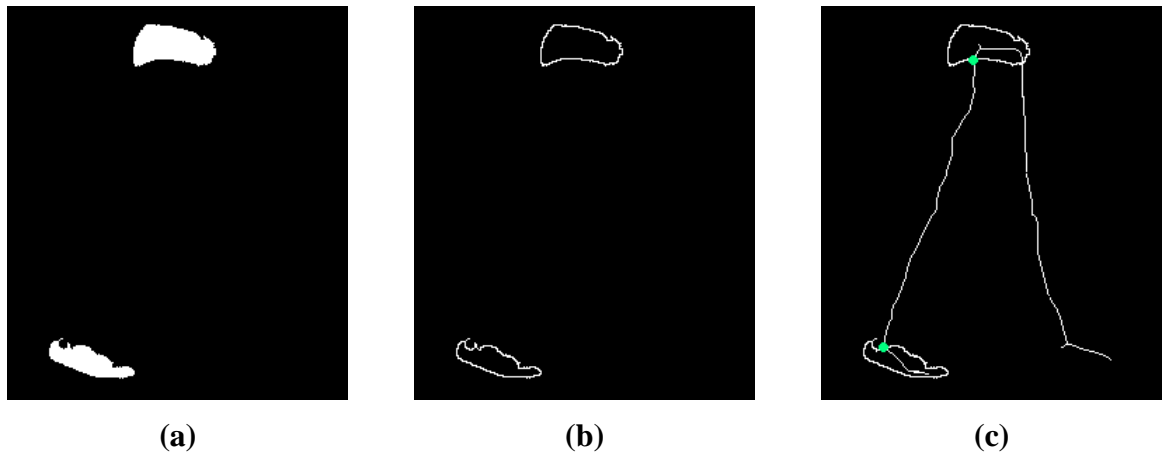
Fig. 7-3. Thresholded garments (a). Edged garments (b). Intersection points (c).

7.2.7   Edge Detection and Extracting Body Segments

Following the labeling of the garments, Robert's edge operator (Roberts, 1965) was applied to the distinctive garments in order to find the edges. The outputs of the edge detection were overlapped with the skeletonization outputs so that the intersection points were identified. Intersection points were significant in body segment extraction step. Edged garments and intersection points can be seen in Figure 7-3b and 7-3c, respectively.

Body segments were extracted after the edge detection. To locate the thigh in the images, a line was drawn from the hip to the intersection of the circle of *M1* radius and the skeletonized line of the leg. Similarly, to locate the shank, a line was drawn from the ankle to the intersection of the circle of *M2* radius with the skeletonized leg (Figure 7-4a). Using the morphological operators, the side edges of the garments were automatically eroded and labeled. After the erosion, only the upper and lower edges remained, i.e. two edges for each garment. The pelvis reference axis was extracted by fitting a line to the upper edge of the underwear. The foot reference axis was extracted by fitting a line to the rear part of the lower edge of the sock (Figure 7-4b). The four axes can be seen in Figure 7-4c.
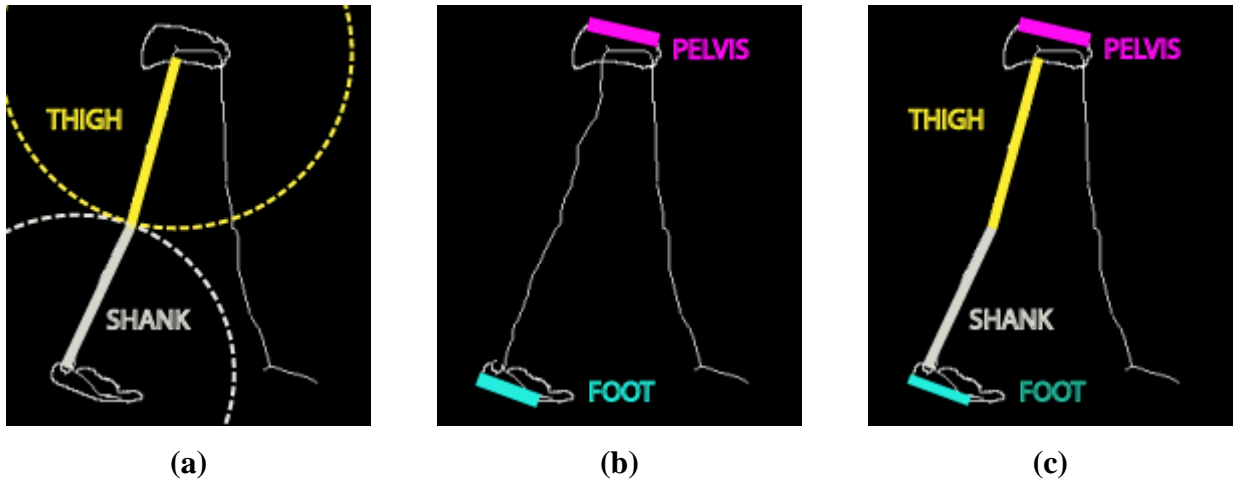
Fig. 7-4. Thigh and shank segments. **(a)** Pelvis and foot segments. **(b)** The four reference axes fitted on the skeleton (c).

7.2.8 Occlusion Handling

During walking, swing of the arms caused occlusions that affected the visibility of the high-cut underwear. Because of the occlusions, in some of the image frames the garment was seen partially or as two blobs and the intersection point on the high-cut underwear became invisible.

In order to solve the two blobs problem, labeling and convex hull operations (Barber et al., 1996) were used. Since the ankle sock in focus was always visible on the image frames, it was labeled first. It was not possible to correctly label the high-cut underwear by simply using the thresholding outputs, since two labels were created for the blobs. In order to overcome this setback, labeled ankle sock was used. Adding the already known shank and thigh lengths to the labeled ankle sock, the approximate height of the high-cut was estimated. Then, all the blobs around that height were considered as parts of the high-cut underwear. Then, the blobs were merged by using convex hull operation, i.e. minimal region containing the two blobs and labeled again.

In order to estimate an intersection point when it was invisible, the edge detection was performed on the convex hull and the output was intersected with the skeleton.

7.2.9 Data Analysis

The proposed hybrid technique was compared with a marker-based technique using a single acquisition system recording the same trials. Root Mean Square Deviation is used as an

indicator of differences between the pattern of curves of proposed hybrid technique and the marker-based.

Hyb and *Mb* joint kinematics were compared as follows.

- In order to evaluate pattern differences of ankle plantar/dorsi-flexion (α), for each time series the deviation from the mean values were determined:

$$\alpha'_{i\_Hyb} = \alpha_{i\_Hyb} - \overline{\alpha}_{Hyb} \text{, and } \alpha'_{i\_Mb} = \alpha_{i\_Mb} - \overline{\alpha}_{Mb}$$

where the subscript *i* refers to the ith frame;

and the Root Mean Square Deviation (RMSD) of the α'_{i_Hyb} values from the α'_{i_Mb} values, was estimated:

$$RMSD(\alpha'_{Hyb}, \alpha'_{Mb}) = \sqrt{\frac{\sum_{i=1}^{n}(\alpha'_{i\_Hyb} - \alpha'_{i\_Mb})^2}{n}} \ .$$

The same processing was applied to the knee flexion/extension (β) and hip flexion/extension (γ), for both directions.

To verify if measurements obtained with the two techniques were comparable to the intra-subject variability, similar indexes were introduced to estimate the intra-subject variability of the ankle kinematics obtained with the *Mb* measurements: $\Delta_V(\alpha) = \left| \overline{\alpha}_a - \overline{\alpha}_b \right|$ and

$$RMSD_V(\alpha'_a, \alpha'_b) = \sqrt{\frac{\sum_{i=1}^{n}(\alpha'_{i\_a} - \alpha'_{i\_b})^2}{n}} \ , \text{ where } a \text{ and } b \text{ represent any two of the five trials}$$

for each direction performed per subject. Similar indexes were introduced for the β and γ angle measurements.

7.3 Results and discussion

Tables 7-1, 7-2, 7-3 and 7-4 show the values of the *RMSD* of the joint angles α, β and γ, $RMSD(\alpha'_{Hyb}, \alpha'_{Mb})$ and $RMSD(\beta'_{Hyb}, \beta'_{Mb})$, the maximum values of the intra-subject variability indexes $RMSD_V(\alpha'_a, \alpha'_b)$, $RMSD_V(\beta'_a, \beta'_b)$ and $RMSD_V(\Gamma'_a, \Gamma'_b)$ obtained for the 10 subjects. Tables 7-1 and 7-2 show the results from trials when subjects walked from the right side, while in Tables 7-3 and 7-4 the results are from trials when the subjects walked from the left side.

[deg] angle		trial	Subject #1	Subject #2	Subject #3	Subject #4	Subject #5
α	RMSD	1	3.7	4.5	3.7	5.2	2.1
		2	2.5	3.3	3.2	6.0	2.6
		3	2.8	4.1	3.0	6.3	2.8
		4	3.4	6.1	4.5	5.4	3.5
		5	3.1	4.7	4.7	6.9	3.2
		average	**3.1**	**4.5**	**3.8**	**6.0**	**2.8**
	$RMSD_V$	max	*3.9*	*6.4*	*3.8*	*15.7*	*2.2*
β	RMSD	1	3.9	4.1	9.6	6.2	5.8
		2	3.7	3.9	8.3	5.5	6.1
		3	3.1	4.5	8.5	5.8	5.2
		4	4.7	5.4	7.3	7.1	4.4
		5	3.9	5.6	8.8	5.4	4.6
		average	**3.9**	**4.7**	**8.5**	**6.0**	**5.2**
	$RMSD_V$	max	*7.9*	*4.5*	*13.7*	*10.2*	*5.0*
γ	RMSD	1	3.4	3.7	7.6	7.1	3.9
		2	4.9	3.2	7.1	7.7	3.2
		3	5.0	5.7	7.8	7.1	4.0
		4	5.1	2.8	7.0	6.5	4.7
		5	5.3	5.2	8.5	7.2	3.7
		average	**4.7**	**4.1**	**7.6**	**7.1**	**3.9**
	$RMSD_V$	max	*5.3*	*3.8*	*12.3*	*6.3*	*3.3*

Tab. 7-1. Root Mean Square Deviation (RMSD) estimated during stance of the markerless joint kinematics values from the marker-based joint kinematics ($α$, $β$ and $γ$) values, from the data of the subjects (#1 - #5) walking from the right side. Trial average values are also reported (in bold). Colored cells represent the maximum values of the trials. In the grayed area the maximum intra-subject variability values (from marker-based measurements) $RMSD_V$ values are reported.

[deg] angle		trial	Subject #6	Subject #7	Subject #8	Subject #9	Subject #10
α	*RMSD*	1	2.5	7.6	6.9	8.1	3.3
		2	2.4	11.2	6.2	10.5	2.9
		3	3.5	9.2	4.1	12.4	2.7
		4	2.8	8.5	4.2	8.3	3.6
		5	3.7	12.5	5.5	9.7	3.1
		average	**3.0**	**9.8**	**5.4**	**9.8**	**3.1**
	RMSD$_V$	*max*	*3.0*	*12.9*	*7.4*	*13.4*	*3.4*
β	*RMSD*	1	4.8	10.2	8.8	8.7	6.5
		2	4.7	11.4	9.1	10.7	5.0
		3	4.9	14.8	8.9	12.5	4.4
		4	3.9	10.3	7.3	11.8	4.6
		5	4.7	8.9	8.3	10.9	5.3
		average	**4.6**	**11.1**	**8.5**	**11.0**	**5.2**
	RMSD$_V$	*max*	*7.1*	*22.0*	*9.1*	*18.8*	*8.9*
γ	*RMSD*	1	5.6	7.9	6.2	8.7	6.1
		2	5.7	9.9	5.9	11.9	5.8
		3	4.5	12.2	4.7	8.1	5.0
		4	4.9	12.0	4.6	12.6	4.8
		5	6.5	8.5	3.9	9.3	4.1
		average	**5.4**	**10.1**	**5.1**	**10.1**	**5.2**
	RMSD$_V$	*Max*	*5.4*	*11.7*	*11.0*	*13.5*	*6.7*

Tab. 7-2. Root Mean Square Deviation (RMSD) estimated during stance of the markerless joint kinematics values from the marker-based joint kinematics (α, β and γ) values, from the data of the subjects (#6 - #10) walking from the right side. Trial average values are also reported (in bold). Colored cells represent the maximum values of the trials. In the grayed area the maximum intra-subject variability values (from marker-based measurements) $RMSD_V$ values are reported.

[deg] angle		trial	Subject #1	Subject #2	Subject #3	Subject #4	Subject #5
α	RMSD	1	7.3	7.4	5.4	10.4	2.9
		2	6.4	5.8	6.2	7.8	3.6
		3	7.9	6.5	7.4	10.3	2.8
		4	8.6	6.2	5.0	8.4	2.5
		5	4.0	7.7	6.2	9.6	3.2
		average	**6.8**	**6.7**	**6.0**	**9.3**	**3.0**
	$RMSD_V$	max	12.2	7.3	6.5	16.1	3.1
β	RMSD	1	8.2	4.1	8.9	6.1	6.8
		2	7.8	5.7	8.3	10.1	5.1
		3	5.6	5.2	6.3	9.6	5.5
		4	9.0	4.5	6.0	8.2	4.8
		5	8.3	3.9	7.1	7.0	5.6
		average	**7.8**	**4.7**	**7.3**	**8.2**	**5.6**
	$RMSD_V$	max	24.48	4.8	12.1	15.3	5.2
γ	RMSD	1	9.2	4.8	9.4	11.0	4.1
		2	8.3	3.9	7.1	10.8	5.0
		3	10.5	4.6	3.7	13.0	3.6
		4	9.8	3.8	6.6	9.7	3.8
		5	6.4	4.1	9.2	8.6	4.0
		average	**8.8**	**4.2**	**7.2**	**10.6**	**4.1**
	$RMSD_V$	max	19.9	4.7	9.0	12.5	3.9

Tab. 7-3. Root Mean Square Deviation (RMSD) estimated during stance of the markerless joint kinematics values from the marker-based joint kinematics (α, β and γ) values, from the data of the subjects (#1 - #5) walking from the left side. Trial average values are also reported (in bold). Colored cells represent the maximum values of the trials. In the grayed area the maximum intra-subject variability values (from marker-based measurements) $RMSD_V$ values are reported.

[deg] angle		trial	Subject #6	Subject #7	Subject #8	Subject #9	Subject #10
α	RMSD	1	3.8	11.8	7.2	9.6	4.5
		2	2.5	13.4	8.9	10.7	4.3
		3	2.5	12.8	7.4	10.2	5.3
		4	2.1	13.7	8.5	11.4	5.8
		5	2.4	14.5	9.1	11.9	5.9
		average	**2.7**	**13.2**	**8.2**	**10.8**	**5.2**
		max	**3.8**	**14.5**	**9.1**	**11.9**	**5.9**
	$RMSD_V$	max	2.9	21.1	10.5	13.2	5.5
β	RMSD	1	4.5	8.7	12.9	14.7	5.8
		2	5.8	11.5	13.3	10.1	4.9
		3	3.1	9.4	11.8	12.6	3.9
		4	3.1	10.2	14.0	12.9	6.1
		5	3.6	8.8	12.5	13.4	5.9
		average	**4.0**	**9.7**	**12.9**	**12.7**	**5.3**
	$RMSD_V$	max	5.7	13.2	16.6	18.4	8.7
γ	RMSD	1	4.1	9.8	11.7	9.4	5.8
		2	3.5	8.7	12.6	10.7	5.5
		3	4.2	10.8	12.8	9.9	6.6
		4	3.7	12.0	13.7	11.2	6.4
		5	3.8	9.3	13.2	13.8	7.5
		average	**3.9**	**10.1**	**12.8**	**11.0**	**6.4**
	$RMSD_V$	max	3.4	10.3	16.8	13.7	5.9

Tab. 7-4. Root Mean Square Deviation (RMSD) estimated during stance of the markerless joint kinematics values from the marker-based joint kinematics (α, β and γ) values, from the data of the subjects (#6 - #10) walking from the left side. Trial average values are also reported (in bold). Colored cells represent the maximum values of the trials. In the grayed area the maximum intra-subject variability values (from marker-based measurements) $RMSD_V$ values are reported.

In the estimation of joint kinematics of children with CP, 3D marker-based motion capture systems are generally used. Marker-based analysis necessitates the use of expensive specialized hardware and expertise. Besides, the existence of markers may cause feelings of uneasiness and interfere with walking.

In this study, we proposed a low-cost, low-discomfort hybrid technique along with a validation in estimating the sagittal kinematics of the lower limbs of children with CP. The technique requires a series of steps: anthropometric measurements and calibration, video acquisition, segmentation, skeletonization, thresholding and edge detection, labeling the garments and extracting body segments and occlusion handling. The lower limb kinematics estimated with the *Hyb* and *Mb* techniques showed very similar results throughout the gait cycle (*RMSD* values are in general within the intra-subject variability).

This is, to authors' knowledge, the first attempt to apply a markerless technique for the gait analysis of CP children. The absence of markers represents a valuable advantage in terms of patient discomfort. The performance of the proposed hybrid markerless approach is promising for its future use. Another important strength of this study is the fact the proposed hybrid technique was validated with a marker-based system.

As a future work, segmentation and occlusion handling mechanisms of the technique will be improved so that CP children with assistive devices can also be analyzed. Besides, with an additional sagittal camera, bilateral analysis can also be performed. Finally, the processing time (about 60 s for each frame) needs to be reduced in order to perform clinical analysis in an efficient way.

CHAPTER 8

CONCLUSIONS

Quantitative gait analysis is generally carried out by mounting retro-reflective markers on the skin of subjects and reconstructing the 3D position by means of stereophotogrammetric systems. The use of stereophotogrammetry requires the placement of markers on selected points of the body segments. Typically, an expert operator spends a considerable amount of time in attaching the markers. A technique less time-consuming and requiring less expertise would be preferred in clinical applications. Markerless techniques (*Ml*) may potentially play an important role in this respect. The present thesis aims at providing a contribution towards the development of new 2D markerless approaches which can be fruitfully used in clinical applications.

The first markerless technique which uses cross-correlation has been used to analyze the ankle-foot kinematics of healthy subjects. The analysis has been done for two conditions: subjects wearing ankle socks and subjects barefoot. The results have shown that, the proposed markerless technique is comparable to the current marker-based techniques and may represent an important step towards the design of an acquisition for clinical purpose. Besides, the segmentation outputs showed that the algorithm chosen for segmenting the images had limited sensitivity to the presence of socks, suggesting that if subjects wore socks during the trials, results would not be affected.

The proposed markerless technique was improved and used in a clinical context, by mainly on the analysis of the knee flexion/extension angle of the children with CP. In order to do so, additional patches were added to the technique and more complex backgrounds were handled. This technique is an important step towards estimating the knee flexion/extension angle with an easy-to-use and low-cost setup.

The final study combined a markerless methodology with "segmental markers" i.e. high-cut underwear and ankle socks. For this reason, the proposed technique was defined as "hybrid". The lower limb joint kinematics of the children with CP is analyzed with a skeletonization-based markerless technique which also combines the information extracted from the garments. Results show that the proposed technique produced comparable results with the marker-based system and can be fruitfully used in the lower limb kinematics estimation of children with CP.

The main contribution of this thesis is that it proposes three new markerless techniques, applied to ankle-foot complex, knee and lower limbs, respectively. The strength of each of the proposed techniques is validated with gold standard data; i.e. the data acquired by the marker-based system working synchronously with the markerless system. The effectiveness of the

proposed techniques for biomechanical/clinical applications is better understood considering the fact that most of the markerless studies are not validated.

As a future work, the processing time of the segmentation algorithm in the current version implemented in MATLAB® (MathWorks, Natick, MA, USA) needs to be reduced to be fruitfully used in clinical applications. Besides, by adding an extra sagittal camera to the setups, acquisitions can be done for both directions. Finally, the proposed hybrid technique should be developed so that the CP children with assistive devices can also be analyzed.

ACKNOWLEDGEMENTS

REFERENCES

Abdel-Aziz YI, Karara HM (1971) Direct linear transformation into object space coordinates in close range photogrammetry. In: Proceedings of the ASP Symposium on Close-Range Photogram, Urbana, IL, 1–18.

Alexander E, Andriacchi TP (2001) Correcting for deformation in skin-based marker systems. J Biomech;34:355–61.

Aurenhammer F (1991) Voronoi Diagrams – A Survey of a Fundamental Geometric Data Structure. ACM Computing Surveys;23:345-405.

Azad P, Ude A, Asfour T, Cheng G, Dillmann R (2006) Image-Based Markerless 3D Human Motion Capture Using Multiple Cues. In: Proceedings of the International Workshop on Vision Based Human-Robot Interaction, Palermo, Italy.

Baker R (2006) Gait analysis methods in rehabilitation. Journal of NeuroEngineering and Rehabilitation;3:4.

Barber CB, Dobkin DP, Huhdanpaa HT (1996) The Quickhull Algorithm for Convex Hulls. ACM Transactions on Mathematical Software; 22:4:469-483.

Barr AH (1981) Superquadrics and Angle-Preserving Transformations. IEEE Computer Graphics and Applications;1:11-22.

Bax M, Goldstein M, Rosenbaum P, Leviton A, Paneth N (2005) Proposed definition and classification of cerebral palsy. Developmental Medicine & Child Neurology;47:571–576.

Benedetti MG, Catani F, Leardini A, Pignotti E, Giannini S (1998) Data management in gait analysis for clinical applications. Clinical Biomechanics;13:3:204-215.

Blum H (1967). A transformation for extracting new descriptors of shape. In: Proceedings of the Symposium on Models for the Perception of Speech and Visual Form. MIT Press, Cambridge (MA); 362-380.

Borelli GA (1681) De motu animalium. Rome.

Bottino A, Laurentini A (2001) A Silhouette Based Technique for the Reconstruction of Human Movement. Computer Vision and Image Understanding;83:79-95.

Bray J (2001) Markerless Based Human Motion Capture: A Survey. Department Systems Engineering Brunel University 2001.

Bregler C, Malik J (1998) Tracking people with twists and exponential maps. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition;8-15.

Calhoun M, Longworth M, Chester VL (2010). Gait patterns in children with autism. Clinical Biomechanics;26(2):200-6

Cappozzo A, Marchetti M, Tosi V, editors (1992) Biolocomotion: a century of research using moving pictures. Rome: Promograph.

Cappozzo A, Catani F, Della Croce U, Leardini A (1995) Position and orientation of bones during movement: anatomical frame definition and determination. Clinical Biomechanics;10:4:171-178.

Cappozzo A, Della Croce U, Leardini A, Chiari L (2005) Human movement analysis using stereophotogrammetry Part 1: theoretical background. Gait and Posture;21:186-196.

Cham TJ, Rehg JM (1999) A Multiple Hypothesis Approach to Figure Tracking. In: Proc. Computer Vision and Pattern Recognition;239-245.

Chen L, Armstrong CW, Raftopoulos DD (1994) An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique. J Biomech;27(4):493–500.

Cheung K, Baker S, and Kanade T (2005) Shape-from-silhouette across time Part I: Theory and algorithms. The International Journal of Computer Vision;2(3): 221–247.

Cheung K, Baker S, Kanade T (2005), Shape-from-silhouette across time Part II: Applications to Human Modeling and Markerless Motion Tracking, International Journal of Computer Vision;63(3):225–245.

Chiari L, Della Croce U, Leardini A, Cappozzo A (2005) Human movement analysis using stereophotogrammetry Part 2: Instrumental errors. Gait and Posture;21:197-211.

Chu CW, Jenkins OC, Matari MJ (2003) Towards model-free markerless motion capture. In: Proceedings of Computer Vision and Pattern Recognition, Madison, WI.

Corazza S, Mündermann L, Chaudhari AM, Demattio T, Cobelli C, Andriacchi TP (2006) A Markerless Motion Capture System to Study Musculoskeletal Biomechanics: Visual Hull and Simulated Annealing Approach. Annals of Biomedical Engineering;34:6:1019-1029.

Daya Bhavya (2008) Parallelization of Two-Dimensional Skeletonization Algorithms. Journal of Undergraduate Research,University of Florida;9:4.

De Berg, Van Kreveld M, Oyermars, Schwarzkopf O (2000) Computational Geometry: Algorithms and Applications. 2nd edition, Springer-Verlag.

De Micheli E, Torre V, Uras S (1993) The accuracy of the computation of optical flow and of the recovery of motion parameters. IEEE Trans. Pattern Anal. Mach. Intell.;15(5):434-447.

Della Croce U, Leardini A, Chiari L, Cappozzo A (2005) Human movement analysis using stereophotogrammetry Part 4: assessment of anatomical landmarks misplacement and its effects on joint kinematics. Gait and Posture;21:226-237.

Deutscher J, Blake A, Reid I (2000) Articulated body motion capture by annealed particle filtering. Computer Vision and Pattern Recognition, South Carolina, June 13-15 2000.

Dobson F, Morris ME, Baker R, Graham HK (2007) Gait classification in children with cerebral palsy: A systematic review. Gait & Posture;25:140-152.

Donati M, Camomilla V, Vannozzi G, Cappozzo A (2007) Enhanced anatomical calibration in human movement analysis. Gait & Posture;26:179-185.

Elgammal A, Lee CS (2004) Inferring 3D body pose from silhouettes using activity manifold learning. Computer Vision and Pattern Recognition, Washington, DC, USA, 2004.

Fermuller C, Shulman D, and Aloimonos Y (2001) The Statistics of Optical Flow. Computer Vision and Image Understanding;82:1–32.

Ferrari A, Benedetti MG, Pavan E, Frigo C, Bettinelli D, Rabufetti M, Crenna P, Leardini A (2008) Quantitative comparison of five current protocols in gait analysis. Gait & Posture;28:207-216.

Fioretti S, Jetto L (1989) Accurate derivative estimation from noisy data: a state space approach. Int J Syst Sci;20:33–53.

Forsyth D and Ponce J (2003) Computer Vision—A Modern Approach, Prentice Hall.

Friedman N and Russell S (1997) Image segmentation in video sequences: A probabilistic approach, In: Proc. of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI), Aug. 1-3, 1997.

Gagalowicz, A.; Chee Kwang Quah (2009) 3D model-based marker-less human motion tracking in cluttered environment. In: Proc. of IEEE 12th International Conference on Computer Vision (ICCV'09) 1042-1049.

Gage JR, DeLuca PA, Renshaw TS (1995) Gait analysis: principles and applications: emphasis on its use in cerebral palsy. Journal of Bone and Joint Surgery;77:1607-1623.

Galli M, Cimolin V, Rigoldi C, Tenore N, Albertini G (2010) Gait patterns in hemiplegic children with Cerebral Palsy: Comparison of right and left hemiplegia. Research in Developmental Disabilities; 31:1340–1345.

Goffredo M, Schmid M, Conforto S, Carli M, Neri A, D'Alessio T (2009) Markerless Human Motion Analysis in Gauss-Laguerre Transform Domain: An Application to Sit-To- Stand in Young and Elderly People. IEEE Transactions on Information Technology in Biomedicine;13:2: 207-216.

Goncalves L, Di Bernardo E, Ursella E, and Perona P (1995) Monocular tracking of the human arm in 3D. In: Proc. 5th Int. Conf. on Computer Vision;764-770.

Goshtasby A, Gage SH, Bartholic JF (1984). A two-stage cross-correlation approach to template matching. IEEE Transactions on Pattern Analysis and Machine Intelligence;6:374-378.

Grauman K, Shakhnarovich G, Darrell T (2003) Inferring 3D structure with a statistical image-based shape model. In: Proceedings of International Conference on Computer Vision;1:641-647.

Gross M, Stevenson P, Charette S, Pyka G, and Marcus R (1998) Effect of muscle strength and movement speed on the biomechanics of rising from a chair in healthy elderly and young women. Gait & Posture;8:3:175-185.

Gross R, Shi J (2001) The CMU Motion of Body (MoBo) Database. Technical Report CMU-RI-TR-01-18, Robotics Inst., Carnegie Mellon University.

Haralick RM and Shapiro LG (1992) Computer and Robot Vision, Volume I, Addison-Wesley,28-48.

Hogg D (1983) Model-based vision: a program to see a walking person. Image and Vision Computing;1(1):5-20.

Howe NR, Leventon ME, Freeman WT (2000) Bayesian reconstruction of 3D human motion from single-camera video. Advances in Neural Information Processing Systems 12; MIT Press, Cambridge, MA, 2000

Hu W, Tan T, Wang L, and Maybank S (2004) A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE Trans. Systems Man and Cybern;34(3):334-352.

Isard M, Blake A (1996) Visual tracking by stochastic propagation of conditional density. In: Proc. 4th European Conf. Computer Vision (Cambridge, England, Apr 1996);343-356.

Ju S, Black M, Yacoob Y (1996) Cardboard people: A parametrized model of articulated image motion. In: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, Killington;38-44.

Koller D, Weber J, Huang T, Malik J, Ogasawara G, Rao B, and Russel S (1994) Towards robust automatic traffic scene analysis in real-time. In: Proc. of the International Conference on Pattern Recognition, Israel, November 1994.

Lagorio A, Grosso E, Tistarelli M (2008) Automatic detection of adverse weather conditions in traffic scenes. In: Proceedings of 2008 IEEE International Conference on Advanced Video and Signal based Surveillance, Santa Fe, NM, USA, 1-3 September 2008; 273-279.

Lam L, Lee SW, Suen CY (1997) Thinning methodologies: a comprehensive survey. IEEE Transactions on Pattern Analysis and Machine Intelligence;27:5:553-568.

Leardini A, Chiari L, Della Croce U, Cappozzo A (2005) Human movement analysis using stereophotogrammetry Part 3. Soft tissue artifact assessment and compensation. Gait and Posture; 21:212-225.

Lewis JP (1995). Fast normalized cross-correlation. Vision Interface;1:120-123.

Lucchetti L, Cappozzo A, Cappello A, Della Croce U. Skin movement artefact assessment and compensation in the estimation of knee-joint kinematics. J Biomech;31:977–84.

Miller PJ (1985) Assessment of joint motion. In: J.M. Rothstein, Editor, Measurement in physical therapy, Churchill Livingstone, London.

Marlow N (2004). Neurocognitive outcome after very preterm birth. Archives of Disease in Childhood;89:224-228.

Mori G, Malik J (2006) Recovering 3D Human Body Configurations Using Shape Contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence;28:7.

Mündermann L, Corazza S, Andriacchi TP (2006) The evolution of methods for the capture of human movement leading to markerless motion capture for biomechanical applications. Journal of NeuroEngineering and Rehabilitation;3:6.

Mündermann L, Corazza S, Chaudhari AM, Andriacchi TP, Sundaresan A, Chellappa R (2006) Measuring human movement for biomechanical applications using markerless motion capture. In: Proceedings of SPIE - Three-dimensional image capture and applications VII, San Jose, CA, USA, 16-17 January 2006; 1-10.

Piriyaprasarth P, Morris ME (2007) Psychometric properties of measurement tools for quantifying knee joint position and movement: A systematic review. The Knee;142–8.

Picelli A, Camin M, Tinazzi M, Vangelista A, Cosentino A, Fiaschi A, Smania N (2010) Three-dimensional motion analysis of the effects of auditory cueing on gait pattern in patients with Parkinson's disease: a preliminary investigation. Neurological Sciences;31:423-430.

Poppe R (2007) Vision-based human motion analysis: An overview. Computer Vision and Image Understanding;108:4-18.

Ramanathan M and Gurumoorthy B (2003) Constructing medial axis transform of planar domains with curved boundaries. Computer-Aided Design;35(7):619-632.

Rothstein JM, Miller PJ, Roettger RF (1983) Goniometric reliability in a clinical setting: elbow and knee measurements. Phys Ther;63:1611–5.

Ridder C, Munkelt O, and Kirchner H (1995) Adaptive Background Estimation and Foreground Detection using Kalman-Filtering. In: Proceedings of International Conference on recent Advances in Mechatronics;193-199.

Roberts L (1965) Machine Perception of 3-D Solids. Optical and Electro-optical Information Processing, MIT Press.

Russek L (2004) Factors affecting interpretation of reliability coefficients. J Orthop Sport Phys Ther;34:341–9.

Shahrokni A, Fleuret F, Fua P (2005) Classifier-based contour tracking for rigid and deformable objects. In: Proceedings of the British Machine Vision Conference;699–708.

Stauffer C, Grimson WEL (1999) Adaptive background mixture models for realtime tracking. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition; 2246-2252.

Stauffer C , Grimson WEL (2000) Learning Patterns of Activity Using Real-Time Tracking. IEEE Trans. Pattern Anal. Mach. Intell.;22(8):747-757.

Stratford P, Agostino V, Brazeau C, Gowitzke BA (1984) Realibility of joint angle measurement: discussion of methodology issues. Physiother Can;36:5–9.

Sundaresan A, Chellappa R (2006) Multicamera Tracking of Articulated Human Motion Using Motion and Shape. In: Proc. Seventh Asian Conf. Computer Vision (ACCV '06);2131-140.

Surer E, Cereatti A, Grosso E, Della Croce U (2011) A markerless estimation of the ankle-foot complex 2D kinematics during stance. Gait & Posture. doi: 10.1016/j.gaitpost. 2011.01.003, 2011 (in press).

Tam R, Heidrich W (2003) Shape simplification based on the medial axis transform. In: Proceedings of the Conference on Visualization;481-488.

Taylor CJ (2000) Reconstruction of Articulated Objects from Point Correspondences in a Single Uncalibrated Image. Computer Vision and Image Understanding;80:349-363

Vanajakshi B, Sujatha B, Krishna KSR (2010). An Analysis of Thinning & Skeletonization for Shape Representation. In: Proceedings of the International Joint Journal Conference on Engineering and Technology (IJJCET 2010).

Whittle MW (1996) Clinical gait analysis: a review. Human Movement Science;15:369-387.

Woltring HJ (1980) Planar control in multi-camera calibration for three dimensional gait studies. J Biomech;13(1):39–48.

Wren CR., Azarbayejani A, Darrell T, and Pentland A (1997). Pfinder: Real-Time Tracking of the Human Body, In IEEE Transactions on Pattern Analysis and Machine Intelligence;19:7:780-785.

Wu G, Siegler S, Allard P, Kirtley C, Leardini A, Rosenbaum D, Whittle M, D'Lima D, Cristofolini, L, Witte H, Schmid O, Stokes I (2002) ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion—part I: ankle, hip, and spine. Journal of Biomechanics;35(4):543–548.

Wu G, Van der Helm FC, Veeger HE, Makhsous M, Van Roy P, Anglin C, Nagels J, Karduna AR, McQuade K, Wang X, Werner FW, Buchholz B (2005) ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion--Part II: shoulder, elbow, wrist and hand. Journal of Biomechanics;38:981-992.