Human Motion Analysis Using Very Few Inertial Measurement Units

Dissertation

Zur Erlangung des Doktorgrades (Dr. rer. nat)

der

Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Qaiser Riaz

aus Rawalpindi, Pakistan

Bonn, December 2015

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Erstgutachter: Prof. Dr. Andreas Weber

Zweitgutachter: Prof. Dr. Bernhard Eberhardt

Tag der Promotion: 17.12.2015
Erscheinungsjahr: 2016

Abstract

Realistic character animation and human motion analysis have become major topics of research. In this doctoral research work, three different aspects of human motion analysis and synthesis have been explored.

Firstly, on the level of better management of tens of gigabytes of publicly available human motion capture data sets, a relational database approach has been proposed. We show that organizing motion capture data in a relational database provides several benefits such as centralized access to major freely available mocap data sets, fast search and retrieval of data, annotations based retrieval of contents, entertaining data from non-mocap sensor modalities etc. Moreover, the same idea is also proposed for managing quadruped motion capture data.

Secondly, a new method of full body human motion reconstruction using very sparse configuration of sensors is proposed. In this setup, two sensor are attached to the upper extremities and one sensor is attached to the lower trunk. The lower trunk sensor is used to estimate ground contacts, which are later used in the reconstruction process along with the low dimensional inputs from the sensors attached to the upper extremities. The reconstruction results of the proposed method have been compared with the reconstruction results of the existing approaches and it has been observed that the proposed method generates lower average reconstruction errors.

Thirdly, in the field of human motion analysis, a novel method of estimation of human soft biometrics such as gender, height, and age from the inertial data of a simple human walk is proposed. The proposed method extracts several features from the time and frequency domains for each individual step. A random forest classifier is fed with the extracted features in order to estimate the soft biomet-

rics of a human. The results of classification have shown that it is possible with a higher accuracy to estimate the gender, height, and age of a human from the inertial data of a single step of his/her walk.

Acknowledgments

First of all, I would like to express my sincere appreciation and gratitude for my supervisor Prof. Dr. Andreas Weber for his continuous support and help in completing this research. Prof. Andreas Weber, your invaluable guidance, suggestions, and feedback offered to me during the course of this work were of immense help and refined my research immensely. You made available your support in a number of ways and always guided me to the right direction. I cherish all these moments that I shared with you, and indeed, it was an honor to work under your kind supervision. I shall always remain obliged to you for your generous interest and help in my work.

Besides my supervisor, I would like to thank Dr. Björn Krüger for his insightful suggestions that helped me in broadening my research by looking at it from various perspectives. I have learned a lot from him and I highly appreciate his continuous support throughout the course of my study. I am also very grateful to Prof. Dr. Bernhard Eberhardt who agreed to evaluate my work as an external reviewer. I am also very thankful to Prof. Dr. Rainer Manthey and Priv.-Doz. Dr. Rainer Surges for accepting my request to join the PhD commission.

I am grateful to the Higher Education Commission of Pakistan (HEC) for selecting me and funding me throughout my stay here in Germany. Had I not been given this opportunity, I would have not been able to do it all by myself. I also extend my gratitude to the friendly assistance provided by the people at Deutscher Akademischer Austauschdienst (DAAD). This work would not have been possible without their support. Additionally, I would like to thank all my colleagues, both former and current, including Guanhong Tao, Jan Baumann, Hashim Yasin, and Anna Vögele for their invaluable inputs.

Finally, I dedicate my work to my father, Muhammad Riaz, whom I lost during my master's studies. It was his love, wisdom, and trust in my abilities that always encouraged me to work hard and achieve my goals. I would like to thank my family: my mother, my brother and sister, and my wife for being there with me and for me in the time of thick and thin. It was your unconditional love and undaunted support that helped me achieve my dream of completing this doctorate.

To anyone I forgot to mention, thank you as well!!!

Contents

List of Figures

List of Tables

1

Introduction

Motion capture has gained significant popularity in the last decade. Motion capture or mocap is a process of recording high quality 3D movements of objects using multiple sensors in a special environment. The recorded information is used to create 2D or 3D digital animations of the characters. Mocap data sets has been successfully used in a variety of fields including human motion analysis, sports, movies and entertainment, video games, medical applications, military, and robotics. A special recording studio, equipped with several sensors, is typically used to record mocap data. An actor usually wears a special suit with markers attached near to each joint. The actor performs different tasks or movements according to a pre-written script. The sensors record all movements of the actor with high resolution and high frame rate (up to several hundred frames per second). Higher resolution is required for precise tracking of the markers and higher frame rate is necessary to minimize motion blurring which occurs due to low frame rate.

A mocap system can either be optical or non-optical. An optical mocap sys-

tem uses a set of calibrated high resolution image sensors or cameras to capture motions. The markers on the actor's body are tracked by an array of cameras and triangulation is used to compute the 3D position of the markers. There are two different types of markers used in an optical based mocap system: passive markers and active markers. A passive marker does not have it's own illumination source but relies on the illumination provided by the environment. It is a special marker which is made of or coated with retro-reflective material, which reflects light when illuminated. Brighter lights are usually used to unambiguously detect each marker. The retro-reflective property of a passive marker makes it cheap in price, light in weight, easy to attach, and does not require any batteries or wires. Active markers, on the other hand, have their own source of illumination, usually a LED, which is actively emitted from the sensors in the form of flashing light. These markers can be tracked more robustly than the passive markers, however, they are comparatively expensive and bulky to wear. Markerless mocap systems are alternative to marker-based mocap systems. These systems do not require any special markers for detection and tracking.

Mocap data can also be recorded with non-optical motion capture systems such as mechanical systems, magnetic systems, and inertial systems. Mechanical motion capture systems are used to directly track body joint angles. In this case, an actor wears a skeleton like structure and the joint angles are tracked through the mechanical parts. In a magnetic motion capture systems, sensors are attached to the actors body which measure the low-frequency magnetic field generated by the transmitter source. Special algorithms are used to calculate 3D positions and rotations of the joints. However, latency and jitter are two major weaknesses of these systems [Len90]. Latency occurs due to asynchronous measurements by the system whereas jitter occurs due to the presence of ferrous or electronic devices in the environment.

We are living in an era where science and technology is shaping our lives. The development of low-cost sensors has made it possible to invent and derive new gadgets with cheaper cost. Availability of low-cost sensors has also motivated the research community in the field of computer graphics and animation to develop new techniques and methods using these sensors. For the last several years, such sensors have been successfully used in the fields of motion capture, motion

reconstruction, and motion analysis. A standard mocap system is comparatively expensive to purchase and requires special environment to operate. Other than high cost, standard mocap systems are not useable in outdoor environments. For example, passive maker based mocap systems require special suits to wear in order to unambiguously detect and track markers, which is not possible in dynamic outdoor environments.

Accelerometers are a good example of low-cost sensors which are used in a number of digital devices such as smart phones and smart watches, tablets, laptops, cameras etc. Better and high-cost accelerometers are used in automobiles, air planes, rockets etc. In general, most of the modern gadgets are equipped with low-cost accelerometers. They are preferred for being cheap in price, small in size, easily wearable, and can be connected to a receiving device wirelessly. Gyroscopes are another example of low-cost sensors which are used in many of the modern day digital devices. Same as accelerometers, gyroscopes are also cheap, small and easy to wear, and can be connected to receiving devices wirelessly.

(a) ***(b)***

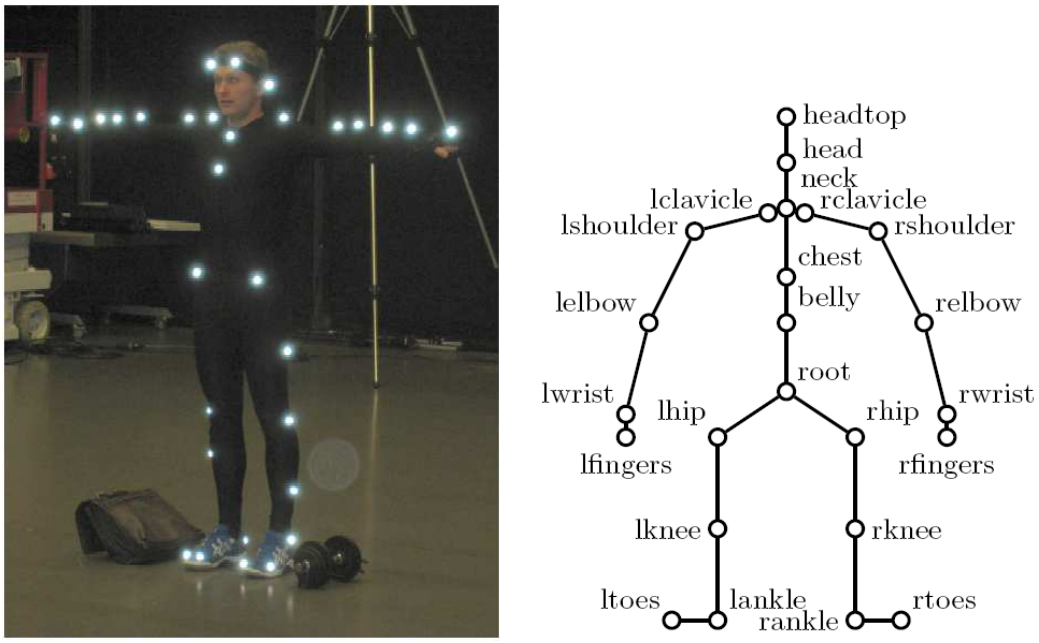
Figure 1.1: *Examples of two inertial measurement units (IMUs) used motion reconstruction and classification. a) Xsens MTx, b) APDM Opal.*

Inertial mocap systems, which use inertial measurement units (IMUs) to capture motion data, are a cheap and low-cost alternative to optical mocap system. An IMU is an electronic device which houses accelerometers, gyroscopes, and sometimes magnetometers. Accelerometers record accelerations of the moving object while gyroscopes record angular velocities of the moving object. Magnetometers

measure earth's magnetic field and are typically used to correct orientation drift. IMUs are essential components of inertial navigation systems which are used in spacecrafts, unmanned aerial vehicles, and satellites etc. Unlike optical based mocap systems, inertial mocap systems do not require any cameras, illuminations, or markers. However, positional drift over time is a major limitation of these systems [WF02]. In figure 1.1, two different types of IMUs are shown. Figure 1.1a shows an Xsens MTx sensor [Xse14] while figure 1.1b shows APDM Opal sensors [Opa15], which are 3-DOF orientation trackers which provide 3D orientation and 3D kinematic data. Both of these sensors have been actively used during the course of our research.

Many human motion data sets, recorded with passive marker based optical mocap systems, are publicly available for research purposes. An example of such data sets is HDM05 motion capture database [MRC+07]. This database consist of 210 minutes of high resolution recordings of roughly 100 different human motions performed by five non-professional actors. A Vicon motion capture system [Vic14] is used to record mocap data. A set of 40–50 passive markers are used, which are tracked at a frame rate of 240 Hz by an array of six to twelve calibrated high resolution cameras. The recorded motion are grouped into five categories including *locomotion*, *grabbing and depositing*, *sports*, *sitting and lying down*, and other *miscellaneous* motions. Figure 1.2 shows an example of placement of passive markers used in the recordings of HDM05. Figure 1.2a shows an actor wearing a special suit with passive markers attached on different joints. Figure 1.2b shows the joint based standard skeleton used in the recordings. Both images are taken from the documentation of HDM05 [MRC+07].

CMU graphics lab motion capture database [CMU03] is another example of freely available mocap data sets captured with a passive marker based optical mocap system. It is a large database consisting of 2605 trials performed by 109 subjects in 6 categories and 23 subcategories. Main categories include *human interaction*, *interaction with environment*, *locomotion*, *physical activities and sports*, *situations and scenarios*, and several *test motions*. Motion data are recorded with a vicon motion capture system. A set of 12 infrared MX-40 cameras are used to track 41 passive markers attached to a special black suit worn by the actor. Each of the cameras has a resolution of 4 megapixel and a frame rate of 120 Hz.



(a) ***(b)***

Figure 1.2: *An example of passive markers set used in the recordings of HDM05. a) An actor wearing a special suit with passive markers attached on different joints. b) The joint based standard skeleton used in the recordings. Both images are taken from the documentation of HDM05 [MRC+07]*

Further examples of human motion data sets include HumanEva [SBB10] for human pose estimation and tracking; Biological motion library [MPP06] for identification of motion properties such as gender, identity, and emotions; Korea University Gesture (KUG) database [HKL06] for gesture recognition and human tracking; ICS action database [MST03] for human action recognition and segmentation; Human Identification at Distance (HID) database [Geo15] for gait recognition; Berkeley Multimodal Human Action (MHAD) database [OCK+13] for recognizing human movements and movement styles in different human activities; KIT whole-body human motion database [MTD+15] for human motion analysis, imitation learning, action recognition and motion generation; and Cologne motion capture database [GF15] for human motion analysis.

The data recorded with any passive marker based optical mocap system are usually available as collections of flat files of different formats. Some of the fa-

mous file formats are C3D, BVH, and ASF/AMC. A C3D file format is a binary format which is used to record synchronized 3D and analog data. It is an old format which is in use since mid 1980s [Mot15]. A biovision hierarchical (BVH) data file format is used to save skeleton hierarchy information along with the motion data [Bio15]. The ASF/AMC or the Acclaim file format is made up of two files. The ASF file (Acclaim Skeleton File) is used to save skeleton data while the motion data are saved into the AMC file (Acclaim Motion Capture data) [Acc15]. For an overview of other motion capture file formats, we refer the reader to the article of Meredith and Maddock [MM01].

The meta-data or annotations about the contents of a mocap data file are given as rough textual description. Such meta-data is important to find the description of the recorded motion saved in a particular file. In case of CMU data sets, the textual descriptions for each motion are given on their website[1]. The name of each file contains the subject number of the actor who acted in the trial, and an event number. Further details of the motions are traced from the information provided on the website. For HDM05, the meta-data is provided in the form of a technical report available on their website[2]. The name of each file contains information about the subject and a reference to one of the chapters of the technical report.

In summary, mocap data sets are currently available as collections of data of various file formats e.g. C3D, BVH, and ASF/AMC etc. The meta-data is provided as rough textual descriptions either online or in technical reports. This method of organization of mocap data has several limitations. Firstly, several file formats make its usage difficult and require considerable amount of time in search and retrieval of contents. One has to learn and understand different file formats and how data are organized in these file formats. Secondly, it is not trivial to extract contents of interest because of a lack of annotation policy. Moreover, since the data are organized in flat files, therefore it is not possible to search and retrieve contents directly. The whole file must be loaded into the system's memory in order to search its contents. In addition to this, there is no centralized and single point access to data. All these limitations make it difficult to search and retrieve contents of interest in minimum time with minimum efforts.

[1] `http://mocap.cs.cmu.edu/subjects.php`
[2] `http://resources.mpi-inf.mpg.de/HDM05/index.html`

A solution to minimize these limitations is using relational database systems to organize mocap data. The idea is to design a flexible database schema and convert all freely available mocap data sets into a single relational database. This approach has several advantages. There is no need to learn different file formats as whole data are organized in a standard relational database. It is possible to search and retrieve contents of interest using standard SQL statements. The flexible schema also provides an easier way of defining hierarchical annotations. Thus, searching and retrieval of motions containing a specific keyword is possible. The hierarchical annotations also allow retrieval of similar type e.g. retrieving all motions containing *left hand waving*. With the flexible design of the database, it is possible to add different freely available mocap data sets into a single database allowing a centralized access to them. The database schema can be extended to include not only additional sensor modalities e.g. inertial measurements, but also synthesized motions. The standard SQL queries can be used to query actual and/or synthesized motions. The proposed relational database for human motion data is discussed in detail in chapter 2 of this report.

The idea of a relation database for human motion data can also be applied to non-human motion data sets such as quadruped motion data. However, the locomotion of the quadruped is much different than the locomotion of humans. The difference is not only in the number of limbs but also in different gait types [Ale84]. A typical human walk step is mainly divided into two phases: a) a swing phase; where the feet leaves ground surface and the leg swings forward and b) a support or stance phase; where the foot stays in contact with the ground supporting body weight [HKH07]. However, quadruped can have different motion phases depending upon the gait type. Horses, for example, have four different gait types: a) walk, b) trot, c) canter, and d) gallop [Ale84]. The movement of limbs differ significantly in each of these gait types.

The difference in gait types and gait mechanism of humans and quadruped indicates that the database schema of the human motion database cannot be used to store quadruped motion data. In this context, we propose a different relational database schema to store motion data of quadruped. Although the database schema is different but the general idea and the benefits of using a relational database management system remains the same. Similar to the relation database

for human motion data, the relational database for quadruped motion data will provide centralized and single point access to the free available mocap data sets, can be queried through standard SQL statements, and will allow storing data from non-mocap sensor modalities as well as synthesized motions. The proposed database schema for quadruped motion data is also discussed in detail in chapter 2 of this dissertation.

Mocap data sets serve as a knowledge base in many data driven methods for full body motion reconstruction [LS99, LS01, AF02, PB02, RSH+05, SH08]. Generally, the knowledge base is searched during the reconstruction process to find best matches of a low dimensional input. Several methods of human motion reconstruction using accelerometers have been proposed. Researchers have shown that whole body human motion reconstruction is possible by using a very sparse configuration of sensors i.e. attaching sensors only to the end-effectors; both hands and both feet [CH05, KCHO10, TZK+11, KL13]. Kim and Lee [KL13], on the other hand, have shown that it is possible to reconstruct many motions from the data of wrist trajectories only. However, their method is limited to only those motions which can be unambiguously distinguished by the position and velocity data.

Taking a step forward from the existing approaches of human motion reconstruction using accelerometers on end-effectors, we have proposed a full body motion reconstruction method using very sparse configuration of sensors: two accelerometers attached on both wrists and one accelerometer attached on the lower trunk. We show that for a number of motions, it is possible to estimate ground contacts from the lower trunk sensor. We also show that visually appealing motion can be reconstructed using the ground contacts information along with the low dimensional input from the wrists accelerometers only. Our proposed algorithm along with results of human motion reconstruction are discussed in detail in chapter 3 of this dissertation.

Sparse representation of human motions is also an important and active area of current research in human motion analysis. It provides a solid base for different inter-linked fields of computer vision and computer graphics such as human activity recognition, user authentication, motion reconstruction etc. Several previous works such as point light walker [Joh73], have shown that several aspects

of human motion are encoded in the sparse kinematic information of a subject. Troje [Tro02] has also used the concept of point light walker to investigate the question of how socially relevant information is encoded in the human biological motion patterns and how such information can be retrieved. The analysis indicates that the gender information is encoded more in the dynamic part of the motion than the motion-mediated structural cues. Researchers have also shown that the video and motion capture data can be used to estimate poses and skeletons which can be used for recognition and analysis of human movement (Lv et al. [LN07], Junejo et al. [JDLP11], Barnachon et al. [BBBG14], Oshin et al. [OGB14]). On the other hand, Venture et al. [VAN08], Kirk et al. [KOF05] have shown that the analysis of representations of motions can provide information on the biometric properties of objects.

The idea of the point light walker can be further extended to investigate different aspects of human motion. An interesting focus of investigation is whether it is possible to estimate soft biometrics i.e. gender, height, and age of a human from the inertial sensor data of a simple human walk. In this context, we have proposed a method of features extraction from the inertial data of a single step. The method extracts a set of features from both time and frequency domains. Additionally, a random forest based classification method is used to estimate soft biometrics of a human. Our results show that the soft biometrics of a human can be estimated with a higher accuracy from the inertial data of an individual step. Further details of the proposed method and related results are presented in chapter 4 of this dissertation.

1.1 Main Contributions

During the course of this doctoral research work, following main contributions are made to the scientific community.

- Our research work on a relational database for human motion data has been presented in the 15^{th} international conference on computational science and its applications (ICCSA 2015). The title of the article is *A relational database for human motion data* [RKW15] and it is published as confer-

ence proceedings in the *Springer's Lecture Notes in Computer Science*.

- Our research work on a relational database for quadruped motion data is partially completed and presented in this thesis report.
- Our research work on full-body human motion reconstruction using very few accelerometers and ground contacts is published in the *Graphical Models*, which is a peer-reviewed journal. The title of the article is *Motion Reconstruction Using Very Few Accelerometers and Ground Contacts* [RTKW15].
- Our research work on estimation of human soft biometrics from the inertial measurements of a single step is unpublished and will be submitted to some peer-reviewed journal for technical reviews.

1.2 Organization of the Thesis

The rest of the report is organized as follows: We start with a discussion on using relational databases for motion capture data in chapter 2. Here we discuss our proposed relational database schemas for human motion and quadruped motion data. In chapter 3, we formulate our work on full body human motion reconstruction using very few accelerometers and ground contacts. In chapter 4, we explain our work on estimation of soft biometrics (gender, height, and age) using inertial measurement of a single step. Chapter 5 is devoted to conclusion and future work followed by a list of literature consulted throughout the course of this study.

2

Relational Databases for Motion Data

Motion capturing is a well-known and standard technique of recording high quality data in various formats. These data sets, which serve as a back bone of data-driven character animation, are usually available in flat files. Different file formats, organization of data in flat files, and a lack of a centralized data access makes it difficult to efficiently use them with minimum efforts. Similarly, it is not easy to manage meta-data and annotations, which describe the nature of the data. These limitations, which are mainly due to organization of data in flat files, can be improved by using relational database management systems (RDBMS). In this chapter, we will show how relational databases can overcome all these limitations. We also discuss the additional benefits of the using relational databases such as design flexibility, scalability, and faster data retrieval. We purpose a relational database approach to two types of mocap data sets: 1) human motion data – discussed in the first section of this chapter and 2) quadruped motion data – discussed in the second section of this chapter.

2.1 A Relational Database for Human Motion Data

Motion capture data have been widely used in applications ranging from video games and animations to simulations and virtual environments. Moreover, all data-driven approaches for analysis and synthesis of motions are depending on motion capture data. Although multiple large motion capture data sets are freely available for research, there is no system which can provide a centralized access to all of them in an organized manner. In this section we show that using a relational database management system to store data does not only provide such a centralized access to the data, but also allows to include other sensor modalities (e.g. accelerometer data) and various semantic annotations. We present two applications for our system: A motion capture player where motions sequences can be retrieved from large data sets using SQL queries and the automatic construction of statistical models which can further be used for complex motion analysis and motions synthesis tasks.

2.1.1 Background

A number of data-driven applications such as for action recognition, motion synthesis, motion analysis etc. have been developed which heavily depend upon high quality motion capture data. For research purposes multiple data sets [CMU03, MRC+07, SBB10, GFB11] containing tens of gigabytes of mocap data are published publicly. These so called *databases* are usually available as collections of files in different formats, like C3D, BVH and ASF/AMC. Currently, the two largest, well established, and freely available collections of mocap data are the CMU [CMU03] and HDM05 [MRC+07] databases where the data are organized in flat files. The file names do not include any information about the nature of the event stored in the file. An indication of the content of each file is given as rough textual description only.

For the CMU data collection, textual descriptions are given for each motion file on the web page. In addition, links to motion files are sorted by subjects and categories, such as *Human Interaction*, *Locomotion*, *Physical Activities & Sports*. Each file name contains the subject number and an event number only.

For the HDM05 data collection, a technical report is available, where a script with the description of the tasks the subjects had performed is published. Here, the file-name refers to the subject, a chapter in this script and thus, gives an indirect indication of the actual content. Additionally a frame rate is stored in the HDM05 file names.

On the one hand there exists no centralized policy for frame accurate annotations of motion capture data. Annotations on the frame level are only available for few example motion sequences, as shown for comparison in the context of motion segmentation [ZTH13, VKK14]. This makes it difficult to search the available data for contents of interest. It is not directly possible to extract a sub part of a relevant information e.g. extracting those frames where subject is in T-pose. To this end, whole files have to be parsed and relevant contents have to be extracted by hand. Another disadvantage of flat files is the inability to query, search, and retrieve information in a structured manner. On the other hand a huge amount of data-driven techniques were developed (see [KTMW08, MLC10, MC12, BWKW14] for some examples) that make use of carefully selected motion data. Our goal is to provide tools that simplify and speed-up the work flow of data selection and model building for such data-driven approaches.

2.1.2 Related Work

To handle the growing amount of mocap data, many alternative methods for fast search and retrieval have been proposed by the research community. Based on frame by frame comparison, *Match Webs* [KG04] were introduced to come up with an efficient index structure. Due to the quadratic complexity it is not possible to build index structures on all currently available mocap data with this method. To avoid quadratic complexity of frame by frame comparisons, segmentation based methods were developed [Lin06]. Cluster based methods classify motion database into small clusters or groups. The classification is either based on similar frames in different motions [BSC05] or on clustering similar poses [LZWM05]. Binary geometric features [MRC05] can be used to tackle this kind of segmentation problems. While methods based on boolean features can not de-

scribe close numerical similarity, this is the case when low dimensional representations of the motion data are used for comparison. To compute low-dimensional representations of mocap data principal component analysis (PCA) is a well established method [FF05, AFO03].

Another way to come up with low dimensional representations is to compute feature sets in a normalized pose space [KTWZ10]. Dimensionality reduction is achieved by using a subset of joint positions only. Typically, a feature set consisting of hand, feet and the head positions is chosen. They also show that it is possible to search these feature sets efficiently by employing a kd-tree as index structure. This retrieval technique is adopted for the database architecture proposed by [WLZ+11]. In this work the authors focus on both, data compression and retrieval of motion capture data. A compression ratio of 25% of the original database size is achieved with this method. To this end a curve simplification algorithm is applied in order to reduce the number of frames to 20% of the original ones. An adaptive k-means clustering algorithm is used to group similar motion clips together. In addition, a three-step motion retrieval method is used which accelerates the retrieval process by filtering irrelevant motions.

A database architecture consisting of data representation, data retrieval and selection modules has been proposed [ACG09]. An in-memory internal representation of the motion data consisting of a collection of poses is created via data representation module. The data retrieval and selection module queries this in-memory representation using PhaseQuery and results are expressed as sequences of segments. The search and retrieval time is almost real-time for small data sets but it increases dramatically for larger data sets. The in-memory condition requires enough physical memory to load large data sets.

For annotation purposes, semi-supervised techniques were employed. Based on a support vector machine (SVM), a database of football motions was annotated [AFO03]. Such kind of annotations are transferred from motion capture data to video sequences in [RF03]. To visualize and explore huge data sets, hierarchical clustering on normalized pose features [BWK+13] was used to obtain an overview on huge data collections. In opposite, Lin [Lin15] presents a system where Kinect queries are used to search and explore mocap data sets that are modeled as sets of 3-attribute strings. Liang and Miao [LM15] have proposed a mocap

data retrieval method based on Labanotation. Key-frames are hierarchically clustered to build a directed graph. A motion metric is used to measure similarity between two motions based on Laban Movement Analysis (LMA). The proposed method works well for simple locomotion activities, however, a considerable decrease in average precision is observed for complex activities such as boxing and dancing.

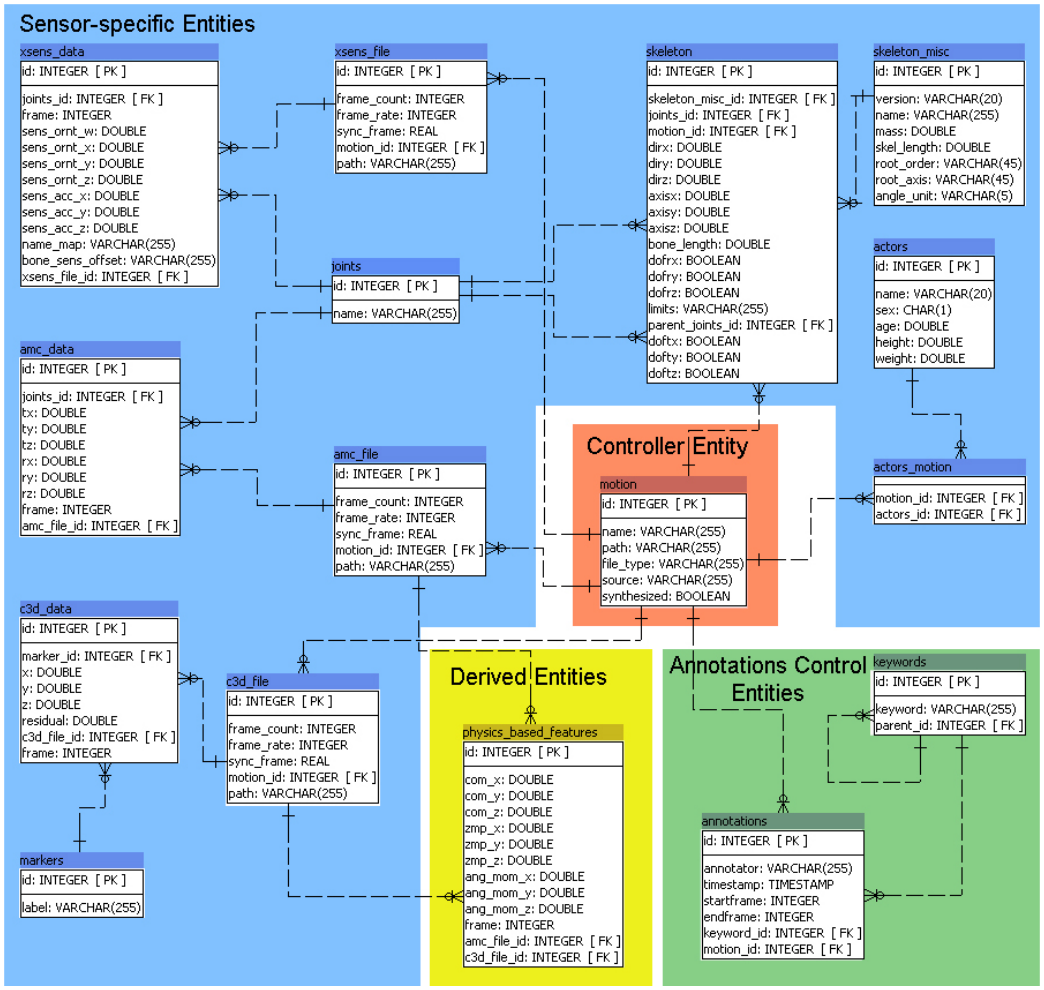


Figure 2.1: *Entity Relationship model of the proposed database architecture for human motion data. The core schema is divided into four different categories, each of which handles an aspect of the proposed schema.*

2.1.3 Database Architecture

The Entity-Relationship (ER) Model of the proposed database architecture is shown in Figure 2.1. The core schema is divided into four different categories:

1. **Controller Entity:** The heart of the proposed schema, which controls the flow of information.
2. **Sensor-specific Entities:** To handle sensor-specific data for each sensor.
3. **Annotations Control Entities:** To control annotation policy mechanism.
4. **Derived Entities:** To entertain non-mocap data which is computed from the mocap data.

We will briefly explain each of these categories in the following subsections.

2.1.3.1 Controller Entity

The purpose of the controller entity is to control the flow of information in a logical manner. It functions as the heart of the schema providing a logical bounding among different entities. All vital information passes through the controller entity. In our proposed database architecture, the *motion* entity acts as the controller entity (Figure 2.1). The controller entity has a simple schema with certain general attributes of a file such as 1) *name:* stores actual name of the file, 2) *file_type:* stores type of the file e.g. amc, c3d, 3) *source:* stores the source of the file e.g. HDM05, CMU, and 4) *synthesized:* a boolean flag to indicate whether the motion data is an actual recording or an artificial motion obtained by some motion synthesis procedure.

2.1.3.2 Sensor-specific Entities

Most of the entities in our proposed database schema are sensor-specific. Sensor-specific entities, as the name indicates, are used to store sensor specific information in the database. In order to achieve a flexible design of the database schema, general properties of each recording are stored in separate entities (name of the entity in a format: *sensor name* + *an underscore* + *'file'* e.g. *c3d_file*) and actual

Table 2.1: *Sensor-specific entities, their attributes and description of each attribute.*

Entities	Attributes	Description
amc_file, c3d_file, xsens_file	frame_count, frame_rate	Total frames and frame rate
	sync_frame	Synchronization frame
	path	Physical path on HDD
amc_data	tx, ty, tz	Translation (x, y, z)
	rx, ry, rz	Rotation (x, y, z)
	frame	Frame number
c3d_data	x, y, z	3D coordinates
	residual	Residual
	frame	Frame number
xsens_data	sens_ornt_w, sens_ornt_x, sens_ornt_y, sens_ornt_z	Sensor orientation (w, x, y, z)
	sens_acc_x, sens_acc_y, sens_acc_z	Sensor acceleration (x, y, z)
	name_map	Name map
	bone_sens_offset	Bone sensor offset
	frame	Frame number
skeleton	dirx, diry, dirz	Direction (x, y, z)
	axisx, axisy, axisz	Axis (x, y, z)
	dofrx, dofry, dofrz	Degree of freedom - rotation (x, y, z)
	doftx, dofty, doftz	Degree of freedom - translation (x, y, z)
	bone_length, limits	Bone length, limits
skeleton_misc	version, name, mass	Sensor's version, name, and mass
	skel_length	Skeleton length
	root_order, root_axis	Order (rotation, translation) and Axis (x,y,z) of the root
	angle_unit	Unit of the angle

data is stored in separate entities (name of the entity in a format: *sensor name + an underscore + 'data'* e.g. *c3d_data*). Each sensor can have any number of additional supporting entities. For example to store AMC data, the general properties

are stored into *amc_file* table and the actual data is stored into *amc_data* table. The supporting entity in this case is *joints* table. We have processed and stored data from different sensors in our proposed database, which includes:

ASF Data: The ASF skeleton data is distributed into two entities namely *skeleton_misc* and *skeleton*. The *skeleton_misc* entity stores the general attributes of skeleton while the *skeleton* entity stores specific skeleton data of each joint in each frame. The attributes of both entities are described in Table 2.1.

AMC Data: The AMC motion data is stored into two mandatory entities *amc_file* and *amc_data* and a supporting entity *joints*. The *amc_file* entity stores general information about the data such as *frame count*, *frame rate*, *synchronization frame* etc. A synchronization frame is used to overcome synchronization problem amongst different sensor systems, which occurs when a single motion is simultaneously recorded by multiple motion capture devices. The *amc_data* stores rotation and translation data for each joint in each frame. The *joints* entity has a *one-to-many* relationship with the *amc_data* and provides an easy mechanism of joint-based data search using standard SQL statements. The attributes of AMC data entities are described in Table 2.1.

C3D Data: The C3D data is stored into two mandatory entities *c3d_file* and *c3d_data* and a supporting entity *markers*. Like the *amc_file*, the *c3d_file* entity also stores general information about the data such as *frame count*, *frame rate*, *synchronization frame* etc. The *c3d_data* entity stores 3D information of each marker in each frame. The *markers* entity has a *one-to-many* relationship with the *c3d_data*. The database can be queried based on markers to fetch data of a particular marker using standard SQL statements. The attributes of C3D data entities are explained in Table 2.1.

Accelerometer Data (Xsens): The accelerometer data is not available in CMU or HDM05 mocap data sets. However, some recordings of accelerometer data were captured later on and we have included these data sets in our database schema. This shows the flexibility and extensibility of our database architecture

that any new sensor can be easily integrated within the existing schema. The data has been recorded using Xsens's MTx accelerometer [Xse14]. In order to store data; two mandatory entities *xsens_file* and *xsens_data* and a supporting entity *joints* are used. The *xsens_file* has same attributes as *amc_file* and *c3d_file*. The *xsens_data* entity stores orientation and acceleration data of each joint in each frame. The *joints* entity has a *one-to-many* relationship with the *xsens_data*. The attributes of accelerometer data entities are explained in Table 2.1.

Table 2.2: *Annotations control entities, their attributes and description of each attribute.*

Entities	Attributes	Description
annotations	annotator	Name of the annotator
	timestamp	Record creation timestamp
	startframe	Starting frame number of the annotation
	endframe	Ending frame number of the annotation
keywords	keyword	Keyword
	parent_id	A self relation, *null* if no parents

2.1.3.3 Annotations Control Entities

Annotations control entities are one of the important entities in the proposed database architecture. These entities define an annotation policy mechanism and provide an easy way to query the database based on an event keyword. In the proposed database architecture, two entities have been introduced to handle annotations. The *annotations* entity stores the general attributes of an annotation such as *start frame*, *end frame*, *timestamp* etc. the *keywords* entity serves as a dictionary of annotations and has a *one-to-many* relationship with the *annotations* entity. It also has a *self-relation* to maintain a *hierarchical relationship* between different keywords. The *hierarchical relationships* are parent- child relationships and define annotations from high-level general terms to low-level more specific terms. For example, in HDM05, a *'jumping jack'* motion is a child of the *'work-out'* event and a grand child of the *'sports'* event. So the *parent_id* of the *'jumping*

jack' will be the *id* of the *'workout'* and the *parent_id* of the *'workout'* will be the *id* of the *'sports'*. The attributes of annotations control entities are expressed in Table 2.2.

2.1.3.4 Derived Entities

In the proposed database architecture, there are certain entities which are derived from the existing mocap data sets. These entities do not exist in any freely available mocap data set. However, they are required in many research activities and researchers have to manually compute them whenever required. A good example of derived entities is physics based features such as center of mass, zero moment point, and angular momentum, which can be computed through kinematics [RW03]. To entertain these features, a separate entity namely *physics_based_features* has been created. The attributes of the *physics_based_features* are explained in Table 2.3. This table has no real data at the moment and computing and dumping physics based features into the database will be carried out in near future.

Table 2.3: *Derived entities - physics based features. Physics based features are derived from mocap data sets and do not exit on their own.*

Entity	Attributes	Description
physics based features	com_x, com_y, com_z	Center of Mass (x, y, z)
	zmp_x, zmp_y, zmp_z	Zero Moment Point (x, y, z)
	ang_mom_x, ang_mom_y, ang_mom_z	Angular Momentum (x, y, z)
	frame	Frame number

2.1.4 Basic Database Operations

2.1.4.1 Processing and Storing Mocap Data into Database

We have used PostgreSQL, version 9.0 - 64 bit, to store extracted data from mocap data sets. PostgreSQL is a widely used object-relational database management

system which is freely available under PostgreSQL license. A database has been created using our proposed database schema. In order to extract data from different formats and store into database, special scripts are written in Visual C++ and Matlab. These scripts read each file of a particular format, extract data of interest, and generate text files with structured queries. These queries are then executed under PostgreSQL environment to store data into different tables. In order to optimize insertion process by minimizing data insertion time, concepts of bulk data insertion are used.

2.1.4.2 Retrieving Collections

Collections can be retrieved using standard SQL statements. In the upcoming subsections, we will give some examples of retrieving collections using standard SQL queries.

Retrieving All Actors: This is a simple example of retrieving data of all *actors*. Each trial in mocap data is performed by one or more actors and motions can be retrieved based on actor's information.

```
select * from actors;
```

Retrieving All Annotation Keywords: This is another simple example of retrieving all *keywords*. Each trial in mocap data is annotated through a keyword, which explains the nature of the event.

```
select * from keywords;
```

Retrieving Annotations of a Specific Event Group: Sometimes one is interested to find all annotations of a specific group of events e.g. finding all '*sports*' annotations. In this example we show how one can retrieve annotations of a specific event group. The event group of interest, in this case, is '*sports*'.

```
select keyword from keywords
 where parent_id = (
  select id from keywords
   where keyword=''sports'')
```

Retrieving Motion Information of an Event: This example shows how to retrieve motion IDs of all motion records for *'dancing'* event. These IDs can be used in later steps to retrieve actual motion data.

```
select m.id from motion m,
 annotations a, amc_file af,
 keywords k
 where m.id=a.motion_id
  and a.keyword_id=k.id
  and af.motion_id=m.id
  and k.keyword=''dancing''
```

Retrieving Synchronized C3D Data: In this example, we show how to retrieve synchronized data based on *syn_frame* value. The synchronization frame, *syn_frame*, is used to overcome synchronization problem amongst different sensor systems, which occurs when a single motion is simultaneously recorded by multiple motion capture devices. The data-type of the *syn_frame* attribute is *real* and stores synchronization time in seconds. In the presence of this time, retrieving synchronized data is very easy and straight forward as shown in the following query.

```
select * from c3d_data
 where c3d_file_id=1 and frame >
   (select sync_frame*frame_rate
    from c3d_file where id=1)
```

Retrieving AMC Data: In this query we extract all AMC data for *'throwing'* event where actor is *'mm'* and source of mocap data is *'HDM05'*. This is a complex query as it involves multiple joins among various entities such as *actors*, *motion*, *amc_file* etc.

```
select * from amc_data
 where amc_file_id in
  (select af.id from amc_file af,
   motion mo where
   mo.source =''HDM05''
   and mo.id=af.motion_id
   and mo.id
```

```
in (select m.id from motion m,
 annotations a,
 actors_motion am
 keywords k, actors ac
  where m.id=a.motion_id
   and a.keyword_id=k.id
   and ac.id=am.actors_id
   and am.motion_id=m.id
   and ac.name=''mm''
   and k.keyword=''throwing''))
```

2.1.5 Database Performance Evaluation

2.1.5.1 Performance Optimization

Before we evaluate the performance of the presented database scheme, we give some insights into the steps taken for optimization of the database structure.

The size of the database on hard disk is approximately 61 GB after parsing and inserting data from all ASF/AMC and C3D files for both *HDM05* and *CMU*. Entities *amc_data* and *c3d_data* are the largest entities having approximately *90 million* and *130 million* records respectively. Hence, an optimization policy is required in order to minimize database search and retrieval time and maximize the system's response time. Indexing is one of the widely used optimization techniques in relational database management systems. PostgreSQL provides several built-in indexing algorithms such as B-tree, Hash, GiST and GIN [Pos13b]. PostgreSQL uses B-tree as default indexing algorithm [Pos13b]. We have created indexes using B-trees on primary keys of both tables. We have also created indexes using B-trees on foreign keys to minimize search and retrieval time.

The trade-off of using indexes is slow data insertion as indexes are updated upon each insertion. However, mocap data sets are not frequently updated so one can compromise on slow data insertion over fast find and fetch. Alternatively, indexes can be dropped during insertion to speed up the insertion process and can be regenerated afterward. We have executed all queries listed in the section *Retrieving Collections* 2.1.4.2 with and without indexing and the results are presented in Table 2.4. The comparison clearly indicates substantial decrease in data search and retrieval time after introducing B-tree based indexes.

The performance of a database can be analyzed based on how much time it takes to

Table 2.4: *A comparison of performance optimization with and without indexing. The data search and retrieve time has substantially decreased by introducing binary tree based indexes.*

Retrieving Collections	Execution Time (ms)	
	Without Indexing	With Indexing
Retrieving all actors	51	11
Retrieving all annotation keywords	13	10
Retrieving annotations of a specific event group	12	11
Retrieving motion information of an event	111	30
Retrieving synchronized C3D data	244,770	18,029
Retrieving AMC data	217,795	8,132

Table 2.5: *Database performance in terms of execution time.*

Entity (Size)	Total Records	Retrieved Records	Trials Count	Exec Time (ms)	Retrieval Criteria (Event, Actor, Source)
amc_data (28 GB)	$164x10^6$	2,581	1	202	T-pose, bd, HDM05
		237,771	2	8,134	Throwing, mm, HDM05
		751,042	20	24,576	Walking, bd, HDM05
		1,505,390	13	51,182	Dancing, All Actors, HDM05
		126,701	12	4,293	Walk, 07, CMU
		522,058	19	17,656	Modern Dance, 05, CMU
		744,662	7	26,081	Swimming, 125, CMU
c3d_data (32 GB)	$230x10^6$	360,756	2	15,978	Throwing, mm, HDM05
		1,139,512	20	48,121	Walking, bd, HDM05
		2,196,786	13	98,768	Dancing, All actors, HDM05
		179,129	12	8,395	Walk, 07, CMU
		738,082	19	34,928	Modern Dance, 05, CMU
		1,052,798	7	47,209	Swimming, 125 , CMU

search and retrieve records against simple and complex queries. As said earlier, we have a particularly large database with a disk size of approximately 61 GB. The two largest entities in our database are *'amc_data'* and *'c3d_data'* having a disk size of *28 GB* and *32 GB* respectively. In section 2.1.5.1, we have outlined our strategy to optimize performance

of the two entities by means of indexing. To test the performance of the database, we have executed several queries on these two large entities. In general, we have found minimum search and retrieval time when the retrieved collections are small in count and maximum search and retrieval time when the retrieved collections are large in count.

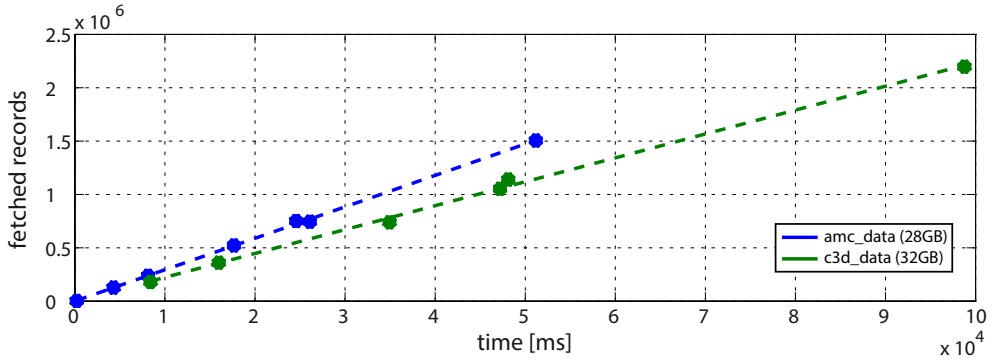


Figure 2.2: *Scatter plot of the timings, when querying the AMC and C3D data sets. We observed a linear relation between the number of retrieved records and the execution time.*

In order to test the performance of the database, several retrieval criteria are used to fetch data and the results are presented in Table 2.5. All tests are performed locally on the database server machine. The database took only *202* ms to fetch *2581* records of *'T-pose'* data of the actor *'bd'* for *HDM05*. The database took *4293* ms to fetch *126701* records of the *'Walk'* event of the actor *'07'* (12 trials in total) from *CMU*. One the other hand, it took *51,182* ms to fetch *1,505,390* records of the *'dancing'* event for all actors (13 trials in total) from *HDM05*. During experimentation, we have observed that the execution time increases as the size of the retrieved records increases and a linear tendency is seen as shown in Figure 2.2. From this, we conclude that the performance of the database is optimal for small record sets. Most applications work in cycles of retrieving small chunks of data from the database and processing these records instead of retrieving the whole data at once. With small execution time (such as *202 ms*), it is possible to achieve interactive processing by fetching and processing data frame by frame.

Complexity Analysis of SQL Queries: The complexity of an SQL statement depends upon a number of factors such as number of tables involved, number of joins, unions, intersections, where/having clauses, sub-queries and so on. The complexity of an SQL statement directly effects its execution cost. The execution plan of any SQL statement can be analyzed in PostgreSQL using the '*explain*' command. "The execution plan

shows how the table(s) referenced by the statement will be scanned by plain sequential scan, index scan, etc., and if multiple tables are referenced, what join algorithms will be used to bring together the required rows from each input table" [Pos13a].

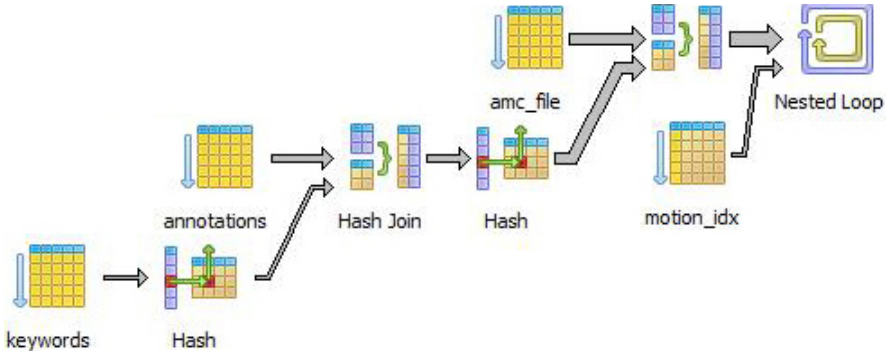


Figure 2.3: *Complexity analysis of the query 'retrieving motion information of an event' analyzed by PgAdmin Query Tool. This query involves inner-joins between four entities: 'keywords', 'annotations', 'amc_file', and 'motion'. It took 12 ms to fetch 13 records.*

Retrieving Motion Information of an Event: The SQL query of *retrieving motion information of an event* is given in number 4 of the Section 2.1.4.2 (*Retrieving Collections*). A visual complexity analysis of this query is presented in Figure 2.3. This query retrieves motion IDs of all motion records for the '*dancing*' event. It consists of four inner-joins between entities: '*motion*', '*annotations*', '*amc_file*', and '*keywords*'. In order to relate entities, PostgreSQL uses indexes for those entities which are indexed and hash joins are used for non-indexed entities. In this example only '*motion*' entity is indexed so its index (*motion_idx*) is used to retrieve records. This query took *12 ms* to fetch *13* records.

Retrieving AMC Data: The SQL query of *retrieving AMC data* is outlined in number 6 of the Section 2.1.4.2 (*Retrieving Collections*). This is one of the most complex queries in our schema. A visual complexity analysis of this query is presented in Figure 2.4. This query retrieves AMC data records of all '*throwing*' events performed by the actor '*mm*'. This query consists of two sub-queries and seven inner-joins between entities: '*keywords*', '*annotations*', '*actors_motion*', '*actors*', '*motion*', and '*amc_file*'. Entities '*motion*' and '*amc_data*' are indexed and their indexes (motion_idx, *motion_data_amc_file_idx*) are used to retrieve records. The query took *8,066 ms* to fetch

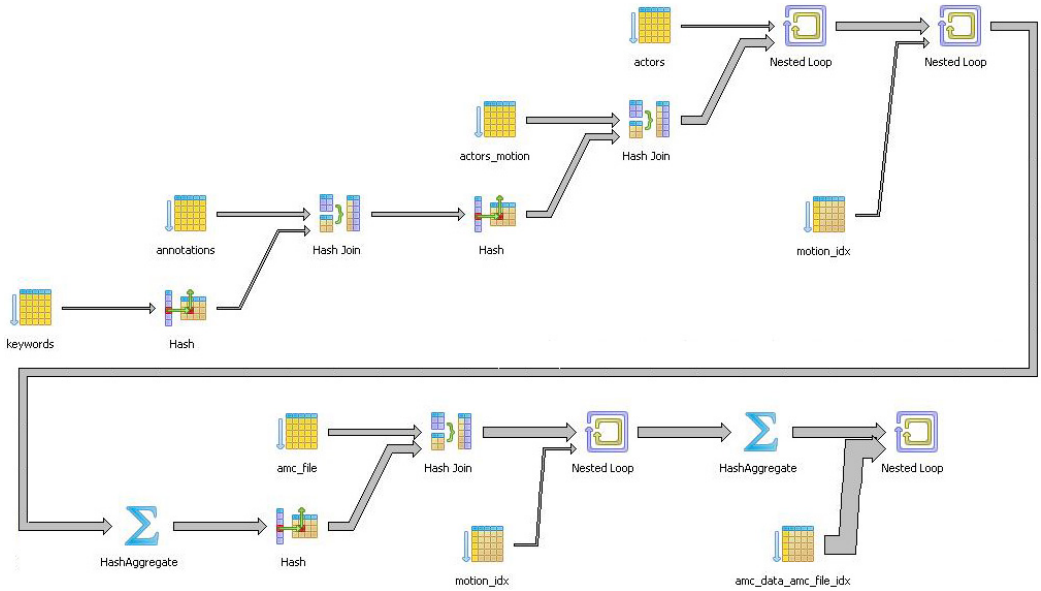


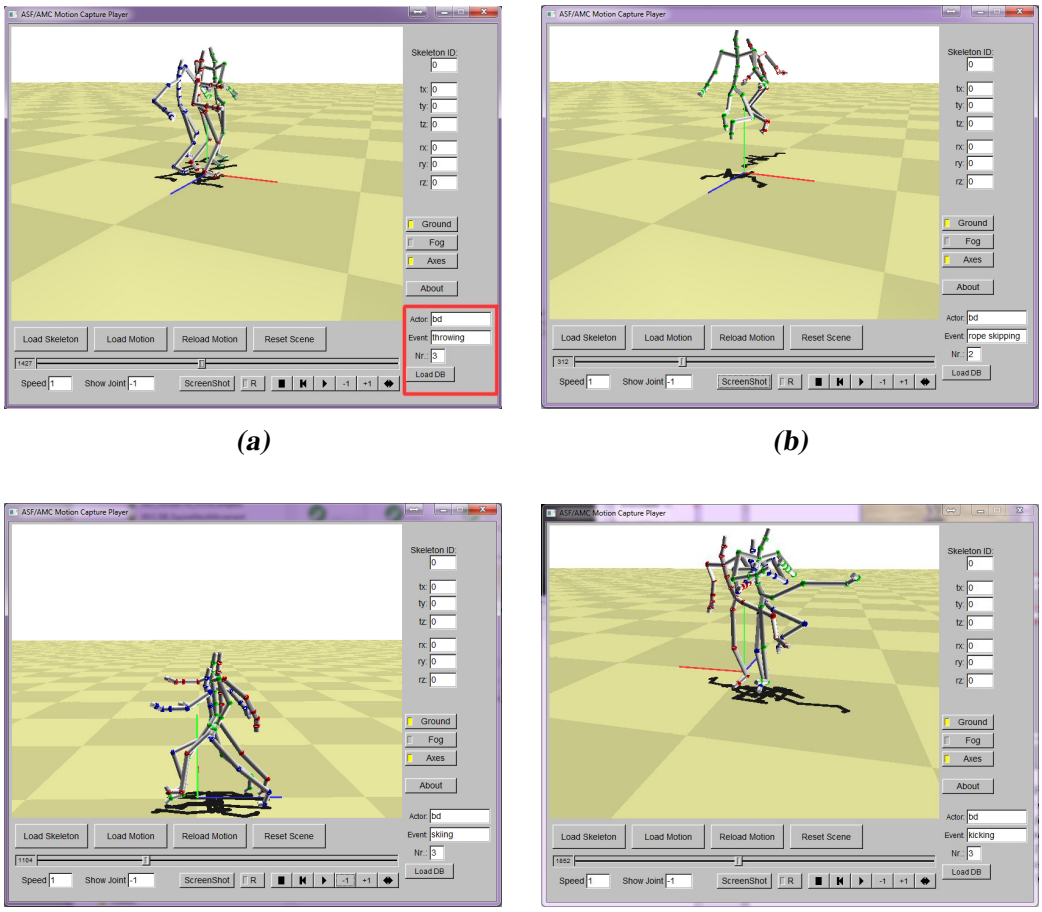
Figure 2.4: *Complexity analysis of the query 'retrieving AMC data' analyzed by PgAdmin Query Tool. This query consists of two sub-queries and seven inner-joins between entities: 'keywords', 'annotations', 'actors_motion', 'actors', 'motion', and 'amc_file'. It took 8,066 ms to fetch 237,771 records.*

237,771 records. This execution time is fairly acceptable considering the size of the entity '*amc_data*' (approx. *90 million* records) and the complexity of this query which involves two sub-queries and multiple inner-joins.

2.1.6 Applications

2.1.6.1 Extended Motion Capture Player

An ASF/AMC motion capture player is one of many ways to visualize motion frames as animation sequence. For this purpose skeleton and motion data for each frame are required. A basic ASF/AMC motion capture player is available with the CMU motion capture dataset [BZ13], which reads skeleton and motion data from flat files and plays them as an animation. We present an extended version of this motion capture player which is capable of playing motions imported from our database. A new GUI element was added, where the variable parts of the *'retrieve AMC data'* SQL query (see Sec. 2.1.4.2) can be filled. As input an *'actor'*, an *'event'*, and an optional *'motion number'* can be given. A database search is carried out for the specified input parameters and if the data is



(a) ***(b)***

(c) ***(d)***

Figure 2.5: *Extended version of ASF/AMC motion capture player originally developed by [BZ13]. This extended version can be used to fetch data from the database and play it. In (a), the rectangle (right side bottom) highlights the extended part. The user provides as input an 'actor name', an 'event', and an optional 'motion number'. The example shows three 'throwing' actions (a), two 'rope skipping' actions (b), three 'skiing' actions (c), and three 'kicking' actions (d). All motions are performed by the actor 'bd'.*

found, it is loaded into the player. A user can then use various control buttons provided in the player to play the animation. Multiple motions can be loaded and played at the same time. With this type of interface it is simple to search for individual motion sequences without having knowledge, or even touching the actual motion capture files. Figure 2.5 shows four different types of motions loaded in the extended motion capture player. In Figure 2.5a, the rectangle (right side bottom) highlights the extended part. The example

shows three *'throwing'* actions (Figure 2.5a), two *'rope skipping'* actions (Figure 2.5b), three *'skiing'* actions (Figure 2.5c), and three *'kicking'* actions (Figure 2.5d). All motions are performed by the actor *'bd'* in these examples. This simple example already shows how simple selections can be made on the basis of SQL queries. More sophisticated data visualization techniques [BWK+13, WVZ+15] could use such selections to allow rapid drill down to important parts of motion data sets for further exploration and analysis.

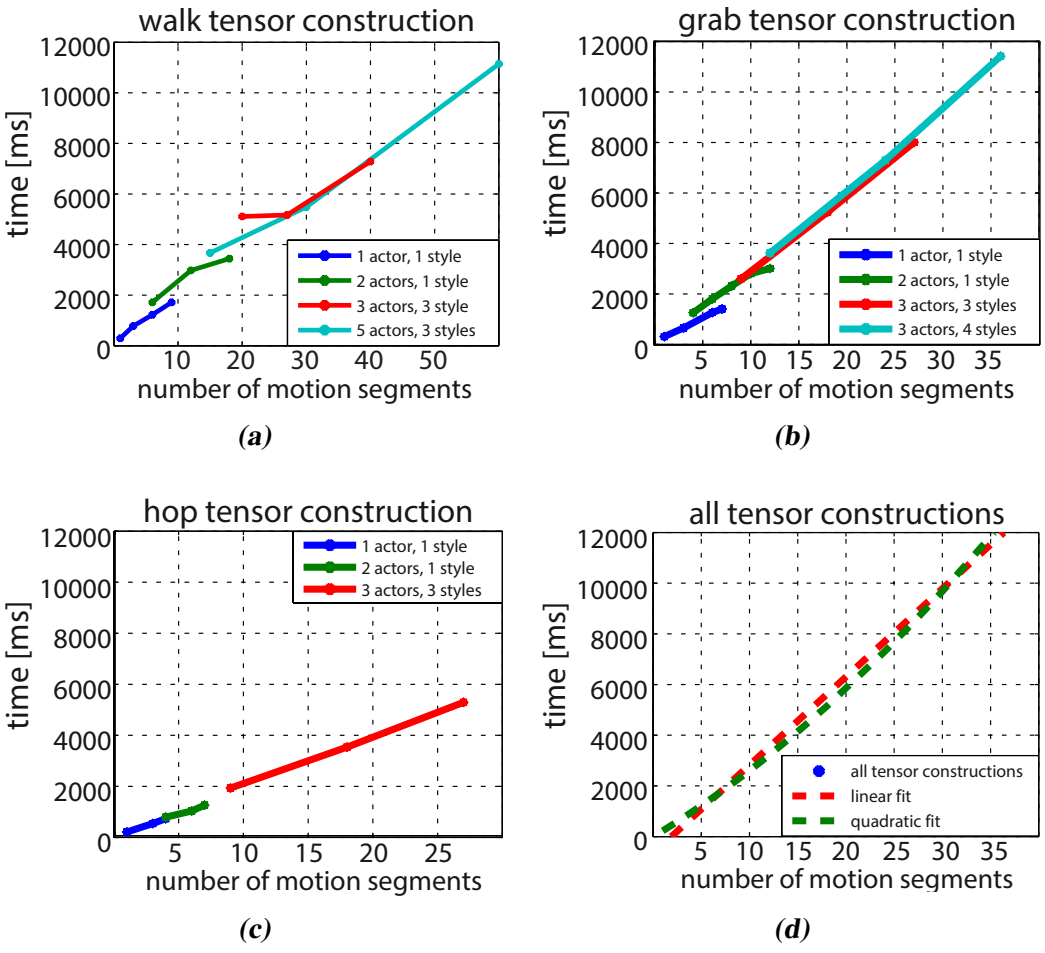


Figure 2.6: *Scatter plot of the timings, when querying the AMC dataset to construct motion tensors. We observed a linear relation between the number of fetched motion segments, but no dependency in the number of actors, motion classes or styles.*

2.1.6.2 Automatic Generation of Statistical Models

Another set of techniques that can benefit from the connection to a relational mocap database is the automatic construction of statistical models. Such models are used for analysis tasks such as motion classification and segmentation or motion synthesis tasks where motion sequences should be developed that fulfill certain user specified constraints. To show the effectiveness of our approach, we show the automatic construction of motion tensors, that have been shown to be useful for motion synthesis [KTMW08, MLC10]. To this end, we fetched data from the database for various actors in the same number of repetitions for multiple motion classes that belong to various styles of a motion. Krüger et al. [KTMW08] introduced so called *natural modes*, that belong to different actors, repetitions or motion styles. In the original works these example motions were selected by hand carefully. By using a data retrieval function, which is written in procedural language for the PostgreSQL (PL/pgSQL), we fetch the individual motions for construction of the multi-modal model. The function takes as input *actor name*, *motion class*, and *number of repetitions* and retrieves related data from the database. Using this approach the construction of each tensor model, as described by Krüger et al. [KTMW08], needed less than ten seconds. Larger sets of motions, including up to 5 actors and 4 motion classes could be retrieved in about 12 seconds. The actual motions for tensor construction were taken from the HDM05 motion capture database. For the *walk*-tensor examples motions from the motion classes: *walkRightCircle4StepsRstart*, *walk4StepsRstart*, *walkLeftCircle4StepsRstart* were used. For the *grab*-tensor the classes *grabHighR*, *grabMiddleR*, *grabLowR*, *grabFloorR* were retrieved. And for the *hop*-tensor the classes *hopRLeg2hops*, *hopBothLegs2hops*, *hopLLeg2hops* were used. The annotations from the classes were taken from the so called *cut-files* subset which is described in the documentation [MRC$^+$07] of the HDM05 data set. Overall we observed that retrieval times depend linear on the number of fetched motion segments instead of the number of actors or motion classes (See Fig. 2.6). Thus, large data sets can be the basis for an efficient construction of statistical models and therefore for a bunch of new applications in motion analysis and synthesis.

2.2 A Relational Database for Quadruped Motion Data

Capturing high quality 3D motions of animals is also gaining considerable attention for quite some time. The basic aim of animals mocap data sets is to provide research community with high quality data and to develop data-driven application using such data sets. However, just like the human motion data, no centralized access to the quadruped motion data is available. In this chapter, we propose a relational database schema to organize freely available quadruped motion data at a single point and to provide a centralized data access. Several example of retrieving contents of interest using standard SQL queries are also presented. Additional benefits of the proposed relational database are also outlined such as easier management of mocap data of different quadruped into a single database, simple mechanism of defining hierarchical annotations and faster data search and retrieval through standard SQL statements.

2.2.1 Background

Several free and commercial motion capture data sets are available for human motion as discussed in the Chapter 1. Different types of motion capturing systems have been used to record human motions including low-cost consumer devices such as Microsfot Kinect, Nintendo Wii Remote as well as high-cost professional recording systems such as Vicon MX, Synertial, Shadow, OptiTrack etc. Due to the availability of freely available mocap data sets, several data-driven methods for full-body human motion reconstruction have been proposed as discussed in Section 2.1.2.

Unlike human mocap data sets, only few data sets of animals or quadruped motions are available. Currently, no large collections of quadruped mocap data sets are freely available. There exist some small image collections for horses, tigers, and dogs such as Weizmann horse data sets[1], INRIA horses[2], and CALVIN TigDog[3]. All of these data sets are publicly available, however, the recorded data are image sequences and are intended to be used in computer vision applications only . Similarly, some commercial mocap data collections are also available by companies such as Equine mechanics[4] and Horse Loco-

[1] http://jamie.shotton.org/work/data.html
[2] http://calvin.inf.ed.ac.uk/datasets/inria-horses/
[3] http://calvin.inf.ed.ac.uk/datasets/tigdog/
[4] http://www.equinemechanics.com/mocap

motion[5]. These collections include mocap data sets recorded with professional motion capture systems, however, these collections are either not freely available or can not be used for research purposes.

Recording of quadruped motions has several restrictions. It is not possible to record quadruped motions with low-cost devices. Professional motion capturing systems are required for recording in indoor environments. In an indoor environment, treadmills can be used to record animal movements but the motions recorded on treadmills can differ significantly from the motions recorded in natural outdoor environments. Recording in outdoor environments is also not a trivial task, because it requires a large area which must be covered by the sensors. For example, the video cameras should completely cover whole area and the wireless sensors should be able to connect with the docking stations without any significant frame loss. Similarly, covering a large area with motion capture systems is very expensive and might not work as desired. For example, retro-reflective markers may not be unambiguously distinguished and tracked by the system.

Many methods of reconstruction of quadruped animations have been presented. One of the well-known motion reconstruction models is physically based models. Marsland and Lapeer [ML05] have proposed a physics-based model of animating trotting horse in real-time. They have used open dynamics engine to model horses as collections of connected bodies and to simulate gravity. P-controllers have been used to minimize errors. Raibert and Hodgins [RH91] have proposed a control algorithm to animate dynamic legged locomotion such as biped robot, quadruped robot, and a kangroo. They have shown that the animated characters can maintain their balance while walking and running at various speeds using various gaits. However, the proposed algorithm is not generic and requires human intervention to create new creatures. Wilhelms and Van Gelder [WVG97] have presented an anatomically based approach to modeling and animating animals. Simmons et al. [SWVG02] have also presented a method of automatic reconstruction of complete jointed creatures using an anatomically based canonical model of similar structures. The model can extract the geometry and animation hierarchy as well as the appearance and motion of the creature.

Another popular method of representing non-rigid motions is using deforming mesh sequence [XZY+07], which is a surface mesh of fixed connectivity and changing vertex positions. Several methods of mesh based quadruped animation have been proposed such as skeleton fitting to a mesh [WP02, SY07, BP07, DATTS08], skeletal deformation [LCF00], defining and exploring shape space [KMP07, YYPM11], arbitrary deform-

[5] http://www.horselocomotion.com/

ing mesh sequences [KG06, XZY+07, KG08, CH12]. Controller-based methods such as one proposed by Coros et al. [CKJ+11] use a representation based on gait graphs, dual leg frame model, and flexible spine model in the synthesizing process.

Lack of a large and publicly available quadruped mocap data set has motivated some of the research groups to work in this direction. One of such projects is the Generic Motion Models based on Quadrupedal Data (GeMMQuad) [Com15] project which focuses on the development of a generic quadrupedal motion model parametrized by spine flexibility. The long-term goal of this project is to develop generic data-driven and biomechanics based methods that can be used from capturing to analyzing, and synthesizing quadrupedal motions. One of the important objectives of this project is construction of a generic and extensible quadruped motion capture database of different quadrupeds including horses, sheep, and dogs. Both indoor and outdoor environments will be used to record different types of quadruped gaits such as walk, trot, canter, and gallop. The motion data will be recorded with different sensor devices including optical motion capture system, accelerometers, image sensors etc.

An important goal of the GeMMQuad project is to publish quadruped mocap data sets publicly in a freely accessible relational database. To achieve this objective, a flexible relational database schema is required. In this context, we present a relational database schema for the GeMMQuad mocap data set. The design of the database allows easier data organization of different sensor modalities. The database also provides a simple way of searching and retrieving contents by means of standard SQL statements. The flexibility of the schema allows storage of mocap data of different animals in the same database. Thus, it is not required to create and manage separate databases for each type of animal. The major advantages of the proposed database architecture are same as the advantages of a relational database for human motion data as discussed in Section 2.1.1. Since the recordings of GeMMQuad mocap data are still in progress, therefore no real mocap data sets are available for testing. The sole purpose of the proposed schema is to provide a conceptual architecture of the relation database for quadruped motion data.

2.2.2 Database Architecture

The Entity-Relationship model of the proposed database architecture for the quadruped motion data is shown in Figure 2.7. The core schema is divided into four different categories:

1. **Controller Entity:** The entry point of the proposed schema. It's basic purpose is

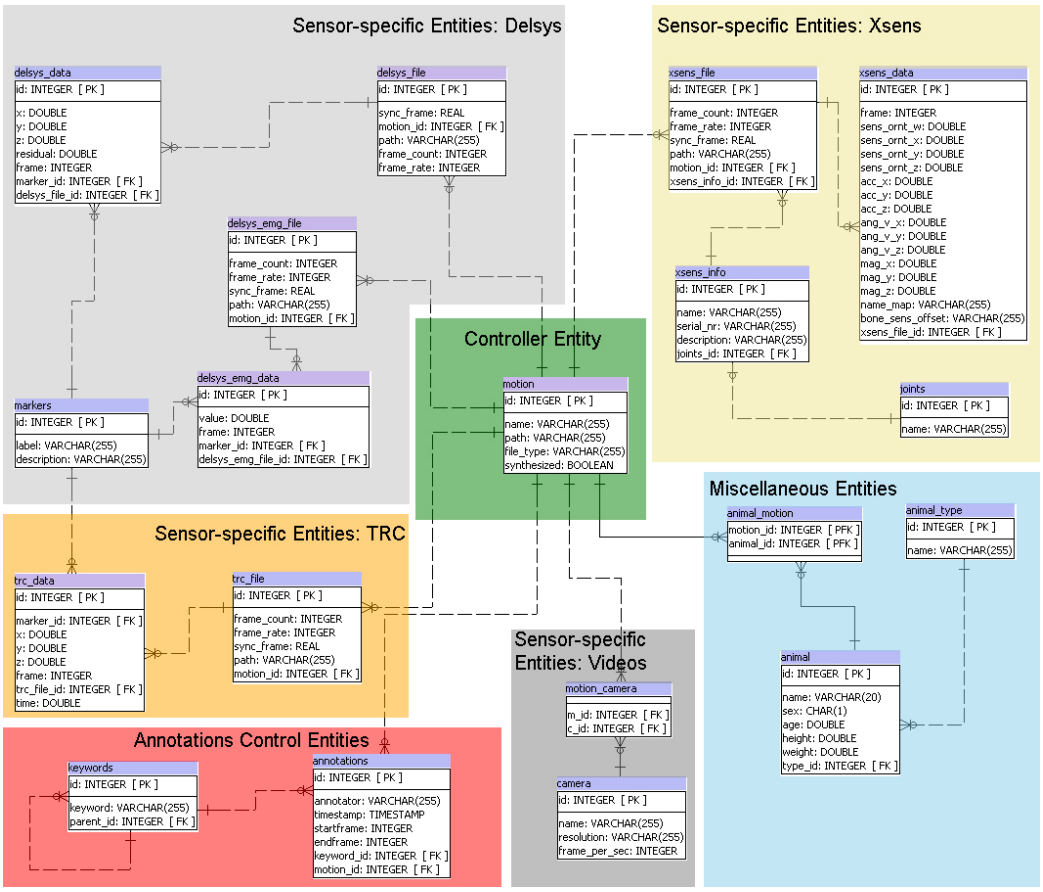


Figure 2.7: *Entity Relationship model of the proposed database for the quadruped motion data. The core schema is divided into four different categories, each of which handles an aspect of the proposed schema.*

controlling flow of information.

2. **Sensor-specific Entities:** These entities are used to handle sensor specific data for each sensor.

3. **Annotations Control Entities:** These entities are used to define and organize annotations.

4. **Miscellaneous Entities:** The purpose of these entities is to organize addition supporting information.

We will briefly explain each of these categories in the following subsections.

2.2.2.1 Controller Entity

The purpose of the controller entity is to control the logical flow of information. It is the entry point of the current schema and most of the data retrieval queries require this entity to search and retrieve information of interest. In our proposed database architecture, the *motion* entity acts as the controller entity (Figure 2.7). The schema of the controller entity is rather simple containing general attributes of a file such as 1) *name:* stores actual name of the file, 2) *file_type:* stores type of the file e.g. trc, delsys, xsens etc., 3) *path:* path of the of data file, and 4) *synthesized:* a boolean flag indicating if the motion data belongs to an actual recording or is synthesized through some procedure.

Table 2.6: *Sensor-specific entities, their attributes and description of each attribute.*

Entities	Attributes	Description
delsys_file, delsys_emg_file, trc_file, xsens_file	frame_count, frame_rate	Total frames and frame rate
	sync_frame	Synchronization frame
	path	Physical path on HDD
delsys_data	x, y, z	Sensor acceleration (x, y, z)
	residual	Residual
	frame	Frame number
delsys_emg_data	value	EMG value
	frame	Frame number
trc_data	x, y, z	3D coordinates
	residual	Residual
	frame	Frame number
xsens_data	sens_ornt_w, sens_ornt_x, sens_ornt_y, sens_ornt_z	Sensor orientation (w, x, y, z)
	acc_x, acc_y, acc_z	Sensor acceleration (x, y, z)
	ang_v_x, ang_v_y, ang_v_z	Angular velocities (x, y, z)
	mag_x, mag_y, mag_z	3D vector of magnetometer readings
	name_map	Name map
	bone_sens_offset	Bone sensor offset
	frame	Frame number

2.2.2.2 Sensor-specific Entities

The basic purpose of sensor-specific entities is to store sensor specific data in a organized manner. Similar to the sensor-specific entities discussed in Section 2.1.3.2, the data of each sensor are split into several sensor-specific entities. General properties of each recording are stored in separate entities and actual data is stored in separate entities. Supporting entities are used to handle additional information. For example to store TRC data, the general properties are stored into the *trc_file* table and the actual data is stored into the *trc_data* table. The supporting information about markers is handled in the *markers* table. Four different types of sensor-specific entities are shown in the proposed schema as discussed below.

Delsys Data: The Delsys Inc. [Del15] manufactures and markets several types of electromyography (EMG) instruments, accelerometers, and biomechanical sensors. In the current schema, the accelerometer data and the EMG data provided by the Delsys sensors are handled separately. The main reason of handling data in separate entities is due to the difference in data types and frame rates of each sensor. The EMG sensor has a higher frame rate (around 2000 Hz) than the frame rate of the accelerometer sensor (around 148 Hz). The EMG data is stored into the *delsys_emg_data* entity whereas the general properties of the recording are stored into the *delsys_emg_file* entity. The markers data, which is supporting information, is stored into the *makers* entity. The accelerometer data is stored into the *delsys_data* entity, whereas the general properties of the recording are stored into the *delsys_file* entity. The *makers* entity is used as a supporting entity to handle marker positions of each sensor. The attributes of each entity explained above are presented in Table 2.6.

TRC Data: The Track Row Column (TRC) file format is a non-skeleton based file format which is used to specify the positions of the markers attached on the subject's body during motion capture trials [Ope15]. This file format was created by the Motion Analysis Corporation. Each data row in TRC file format contains a time value and positional values of 3D coordinates of each marker. In the proposed schema, the data in TRC file format is handled into the *trc_data* entity while the general properties of the recording are stored into the *trc_file* entity. The *markers* table is used as a supporting entity to handle makers' details. The attributes of TRC data entities are described in Table 2.6.

Video Data: In the proposed schema, video streams are not directly saved into the database. However, sensor specific properties of each sensor are saved in the *camera* table. The attributes of the table include *camera name*, *resolution*, and *frame rate*. The *camera* table has a *many-to-many* relationship with the *motion* entity which is maintained through the *motion_camera* table.

Inertial Measurements Data (Xsens): Xsens's MTx [Xse14] is an inertial measurement unit (IMU), which provides 3D information about an object's orientation, acceleration, and angular velocities. Following entities are created to store Xsens data: Two mandatory entities *xsens_file* and *xsens_data*, two supporting entities *xsens_info* and *joints*. The *xsens_file* table has same attributes as the *delsys_file* table. The *xsens_data* entity stores orientation, acceleration, angular velocities, and magnetic data of each joint in each frame. The *xsens_info* entity stores general properties of the sensor such as *sensor name*, *serial number*, and *description*. The *joints* entity has a *one-to-one* relationship with the *xsens_info* table. This relationship helps to identify which sensor has been attached to which joint. The attributes of these entities are explained in Table 2.6.

2.2.2.3 Annotations Control Entities

The annotations control entities in the proposed schema serve as a way of managing hierarchical annotations by means of keywords. They provide a simple yet effective way of retrieving contents based on a specific keyword or a set of keywords. The schema of annotation control entities is borrowed from the one discussed in Section 2.1.3.3. The *annotations* entity stores the general attributes of an annotation such as *start frame*, *end frame*, *timestamp* etc. The *keywords* entity has a *one-to-many* relationship with the *annotations* entity. This entity also has a *self-relation* to maintain a *hierarchical relationship* between different keywords. For details on the attributes of annotations control entities, please refer to Table 2.2 and Figure 2.7.

2.2.2.4 Miscellaneous Entities

The miscellaneous entities are general purpose entities used to store information related to the subjects. In the proposed schema, the *animal* entity is used to store basic information of each animal such as *name*, *sex* , *age*, *height*, and *weight*. The *animal_type* entity is used to handle different types of animals and it has *one-to-many* relationship with the *animal*

entity. A *many-to-many* relationship between the *motion* entity and the *animal* entity is handled using the *animal_motion* table.

2.2.3 Basic Database Operations

2.2.3.1 Processing and Storing Quadruped Mocap Data into Database

PostgreSQL, which is a well-known and commonly used RDBMS, has been used to create a relation database according to the proposed schema. The selected version of PostgreSQL is 9.0 - 64 bit.

2.2.3.2 Retrieving Collections

Collections can be retrieved using standard SQL statements. In this subsection, we will present some example queries to show how contents of interest can be retrieved from the proposed database. The data retrieval queries are performed on dummy data since no real data is yet available.

Retrieving All Subjects: This is a simple example of retrieving data of all *animals*. Each trial recorded during motion capture sessions is performed by one or more animals, which we call subjects. The query below shows retrieval of information about all subjects.

```
select * from animal
```

Retrieving All Subjects of a Particular Type: This is another simple example of retrieving all *subjects* of a particular type. Here, we are interested to see all subjects of type *sheep*.

```
select a.* from animal a, animal_type at
 where a.type_id = at.id
  and at.name=''sheep''
```

Retrieving Motion Data Using Annotations: This example shows how to retrieve motion IDs of all motions which are annotated as *'canter'* or *'gallop'*.

```
select m.* from motion m,
 annotations a, keywords k
```

```
 where m.id=a.motion_id
  and a.keyword_id = k.id
  and (k.keyword=``canter'' or k.keyword=``gallop'')
```

Retrieving Camera Information: In this example, we show how to retrieve information of the cameras used in the recordings of a particular gait type, i.e. *'canter'*.

```
select c.* from camera c,
 motion_camera cm, motion m,
 annotations a, keywords k
 where c.id = cm.c_id
  and cm.m_id = m.id
  and m.id=a.motion_id
  and a.keyword_id = k.id
  and (k.keyword=``canter'')
```

Retrieving Sensor Specific Data: Here we extract TRC data of all *'trot'* events performed by the subject *'Gaia'*. Furthermore, we are only interested into the data of the *'right shoulder'* of the sujbect. This is a very complex query involving multiple joins between eight entities. Retrieval of this sort of data from flat files is a very challenging task.

```
select td.* from trc_data td,
 trc_file tf, motion m,
 annotations a, keywords k,
 animal_motion am, animal an
 where td.trc_file_id = tf.id
  and tf.motion_id = m.id
  and m.id=a.motion_id
  and a.keyword_id = k.id
  and m.id = am.motion_id
  and am.animal_id = an.id
  and an.name = ``Gaia''
  and (k.keyword=``trot'')
  and td.marker_id = (select id from markers
```

```
    where label = ''ShoulderR'')
```

3

Motion Reconstruction Using Very Few Accelerometers and Ground Contacts

Due to the rapid development in sensor technology, the recording of human motion sequences is making its way out of controlled studio environments. Accelerometers are available in a broad range of devices that can be used practically everywhere. In order to show that the reconstruction of full body motions is possible with standard wearable devices, we introduce a prototype that is capable of doing so on the basis of a very sparse sensor setup: We make use of accelerometers placed on both wrists and lower trunk, and ground contact information only. In this setting human motion reconstruction is a difficult and challenging task due to sparse spatial distribution of sensors and the noisy nature of input data. That is why, we first identify ground contacts from the lower trunk sensor signals and then in a second step combine these results with a fast database look-up that allows a data-driven motion reconstruction. We show the effectiveness of our approach in an extensive set of experiments on both simulated and real data. Our results show

appealing reconstructed motions in a variety of human motion scenarios.

3.1 Background

Practical motion capturing with low-cost sensor has been a topic of intense research in the last decade. One major line of research was on reducing the costs and especially the efforts of recording motions by minimizing the number of sensors respective markers set. A considerable yet effective configuration to achieve sparseness is by attaching sensors to the end-effectors only, i.e. to hands and feet [CH05, KCHO10, TZK+11, KL13]. In the reconstruction process, a database of whole-body mocap data is used as a knowledge base for synthesizing motion of missing segments and regularizing the reconstruction process.

Kim and Lee [KL13] have shown that for surprisingly many motions, very well reconstruction is possible from the data of wrist trajectories only. Although the restriction to data of wrists trajectories is of little practical relevance, the results presented in [KL13] naturally imply the question whether motion reconstruction from other and practically obtainable sensor data attached to the wrists and the lower trunk —specifically from accelerometers and gyroscopes— allow similarly good full body motion reconstruction?

In general, the data of time series of wrist acceleration contains less information than the data of time series of wrist position, as has been used in [KL13]. However, having position data and velocity data for a single key frame in an idealized setting (disregarding noise and drifts due to double integration) the information content of the entire time series of acceleration data is not less than the one of the position data.

One can easily come up with settings of motions where one cannot deduce position and velocity data from the time series of accelerations: The wrist accelerations when riding a bicycle with constant velocity cannot be distinguished from the one having another velocity or even from a T-pose, whereas from the positional data these motions can be distinguished. However, when testing input sequences of the HDM05 database [MRC+07] we found that for each of the sequences for Walking, Locomotion on the spot, Locomotion, Locomotion with weights, Hopping and jumping, and Climbing stairs, there exist key frames for which the position and velocity information of the wrists can be statistically deduced from the accelerometer information with a high probability.

Although these results are theoretically encouraging and imply that motion reconstruction on the basis of wrist accelerometer data is possible for many motions in principle, a robust and practical algorithm along with this line of investigation is rather challenging. In this context, we propose a two-step reconstruction method. In the first step, we es-

timate ground contacts from the lower trunk sensor. In the second step, we use ground contact information along with accelerations from both wrists and lower trunk senors to reconstruct motions. Our main contributions presented in this paper are as follows:

- We show that motion reconstruction based on ground contact information together with very few additional accelerometers (such as two wrist attached ones) is possible in general.
- We present a novel ground contact detection algorithm for various locomotion activities from the data of a third accelerometer/gyroscope sensor attached to the lower trunk. Hence in this case the overall reconstruction method requires three accelerometers only.
- We show that it is very advantageous in the case of motion reconstruction from two hand attached and one lower trunk attached accelerometers to estimate ground contacts from the lower trunk sensor first and then to proceed similar to the method proposed by Tautges et al. [TZK+11]. The average reconstruction error using our method is reduced compared to Tautges [TZK+11] method, in which a method of motion reconstruction from four accelerometers attached to the wrists and the ankles is presented, but no ground contact information had been deduced from the accelerometer readings.

  Using the ground contact information, the average processing time is also reduced significantly in general (by about a factor on 3) closing the gap to real time applications significantly.

Our focus on pure accelerometer data for wrist attached sensors is not only motivated by previous research, but mainly by its practical significance: For integration into smart watches or sensor wristlets the integration of accelerometers only without additional gyroscopes allows the creation of smaller devices. For a lower trunk sensor such a consideration is of smaller importance.

3.2 Related Work

Human motion capturing (mocap) has become a standard technique for data-driven animation. Many mocap systems were developed based on optical, mechanical, magnetic, inertial sensors and markers. All these systems have strengths and weak-

nesses in terms of accuracy, capturing volume and operation effort. For an overview see [Mai96, MHK06, Wik14].

The most widely used are optical systems [Pha14b, Opt14, Vic14, Qua14], that make use of a set of high resolution calibrated cameras to capture the 3D position of markers attached to the actors's body by triangulation from the camera images. However, such systems are highly expensive as they require multiple high quality calibrated image sensors, special gears to wear, and special recording environment.

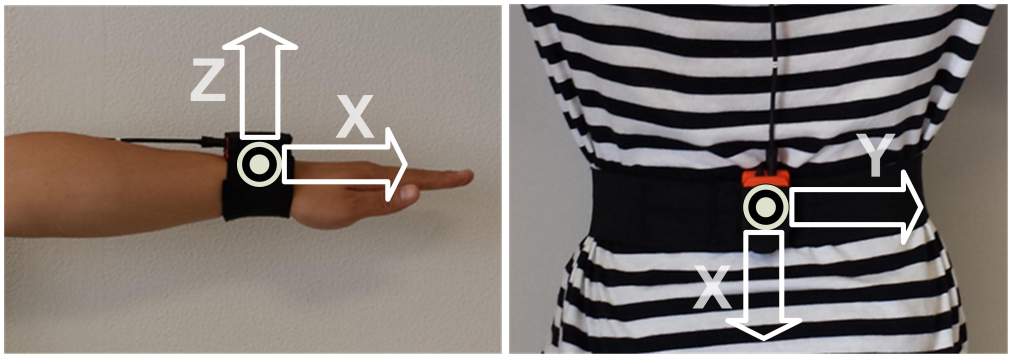
An alternative to marker-based optical mocap systems is low-cost inertial sensors [Xse14, RLS09, Ine14]. These sensors are small in size and easily attachable to an actor's body. They do not require any special recording environment or special gears to wear. However, inertial information from these sensors is very noisy and prone to drifts. To handle sensor drifts, several approaches have been proposed by the research community. Schepers et al. [SRV10] have used magnetic field sensors to avoid drifts, while Vlasic et al. [VAV+07] employed ultrasonic distance sensors for this purpose. Kelly et. al. [KCOH13] proposed a motion graph based memory efficient technique that uses a single overhead camera for global position and orientation data along with accelerometers.

Synthesis of visually appealing and natural looking human motions is a challenging task. Several approaches have been proposed in this context. Well established are data-driven methods, which use a mocap database as the knowledge base [LS99, LS01, AF02, PB02, RSH+05, SH08] of how people move. Typically, the knowledge base is searched during the reconstruction process to find best matches according to the low dimensional input. However, the quality of synthesized motion decreases significantly for such motions which are absent from the knowledge base. The problem of absent motions can be fixed to an extent by creating new motion through interpolation [KG04, MK05, SH07]. However, interpolation methods do not allow creation of new poses that are not present in the knowledge base.

Several data retrieval methods have been proposed to identify motions similar to a query in large motion capture databases. Chai and Hodgins [CH05] used low-dimensional inputs to search k-nearest poses in a pre-computed neighbor graph. Krüger et al. [KTWZ10] have developed a lazy neighborhood graph (LNG) method for fast motion matching in neighborhoods that are found in a kd-tree index structure. An online capable extension (OLNG), that operates on a window of frames, is introduced by Tautges et al. [TZK+11].

Methods for motion reconstruction combine fast retrieval and motion synthesis techniques. Here the challenges are to find similar motions on the basis of low-dimensional

inputs and to synthesize poses that might not be included in the database. Chai and Hodgins [CH05] build a local linear model on the basis of the retrieved nearest neighbors. A best match is obtained by energy optimization. Liu et al. [LWC+11] extended this approach by constructing a series of local linear dynamic models online to approximate nonlinear dynamic behaviors of human movement. Tautges et al. [TZK+11] constructed pose spaces in an online fashion. They came up with a kernel based energy minimization to synthesize the output poses. With this method they showed that appealing motions can be reconstructed on the basis of four accelerometers (attached on the wrists and the ankles) only. Kim and Lee [KL13], have shown that wrist and ankle trajectories can be used to reconstruct various whole body motions. The control inputs consist of two wrists trajectories and the ankles trajectories are computed from the wrists trajectories by solving a shortest path problem in a directed acyclic graph. Yin and Pai [YP03] proposed a method which uses foot pressure distributions to reconstruct full-body motions. However, their method requires a knowledge base which contains pre-captured pressure and motion data.



(a) *(b)*

Figure 3.1: *Sensors placement: two accelerometers are attached to both wrists and one is attached to the center of lower trunk using straps.*

3.3 Estimation of Ground Contacts from a Lower Trunk Sensor

The lower trunk of a human skeleton is a natural point at which the ground contacts of the body can be detected. A suitable sensor attached to the lower trunk sensor (see Figure 3.1) can detect such impacts in terms of accelerations and angular velocities. In this section

we present a novel method of estimating ground contacts from the lower trunk sensor. An overview of our proposed algorithm, which consists of five phases, is presented in Figure 3.2. In the following subsections, the proposed algorithm in explained in detail.

3.3.1 Pre-processing

In the pre-processing phase, input signals from the lower trunk sensor are down-sampled from 100Hz to 25Hz in order to minimize processing time. The measured accelerations from an accelerometer are known to be very noisy and using raw data is not suitable. We used moving average method with a window size of 5 frames to suppress noise and to smooth the down sampled signal. We then compute the magnitude of 3D accelerations and subtract gravity from it. Let $\boldsymbol{a}_x$, $\boldsymbol{a}_y$, and $\boldsymbol{a}_z$ be 3D accelerations from the lower trunk sensor and g be the acceleration due to gravity, then the modified signal, $\widehat{\boldsymbol{a}}$, is given as

$$\widehat{\boldsymbol{a}} = ||a_{xyz}|| - ||g|| \tag{3.1}$$

$$\widehat{\boldsymbol{a}} = \sqrt{\boldsymbol{a}_x{}^2 + \boldsymbol{a}_y{}^2 + \boldsymbol{a}_z{}^2} - ||g|| \tag{3.2}$$

3.3.2 Time-Frequency Reassignment

In the second phase of our proposed algorithm, we analyze pre-processed magnitude of the input signal in frequency domain to estimate time-frequency representations of the signal. In this context, we use the Time-Frequency reassignment method proposed by [FF09]. A brief description of this method along with its usage in the context of our proposed algorithm is given below.

Spectrograms, which are time-frequency representations of a signal, are a well known way of analyzing time-varying signals such as accelerations of a non-rigid body. The spectrogram of a signal can be estimated by squared magnitude of the short-time Fourier transform (STFT) of the signal. To compute the STFT of the input signal, a moving window of fixed size is used. A trade-off of the fixed sized moving window, T, is that the STFT can not resolve any event happening faster than T or slower than 1/T. An STFT with a larger window size resolves frequency features better compromising on time features. On the other hand, an STFT with shorter window size resolves time features better compromising on frequency features.

A solution to address these trade-offs is to use the phase spectrum of the signal to

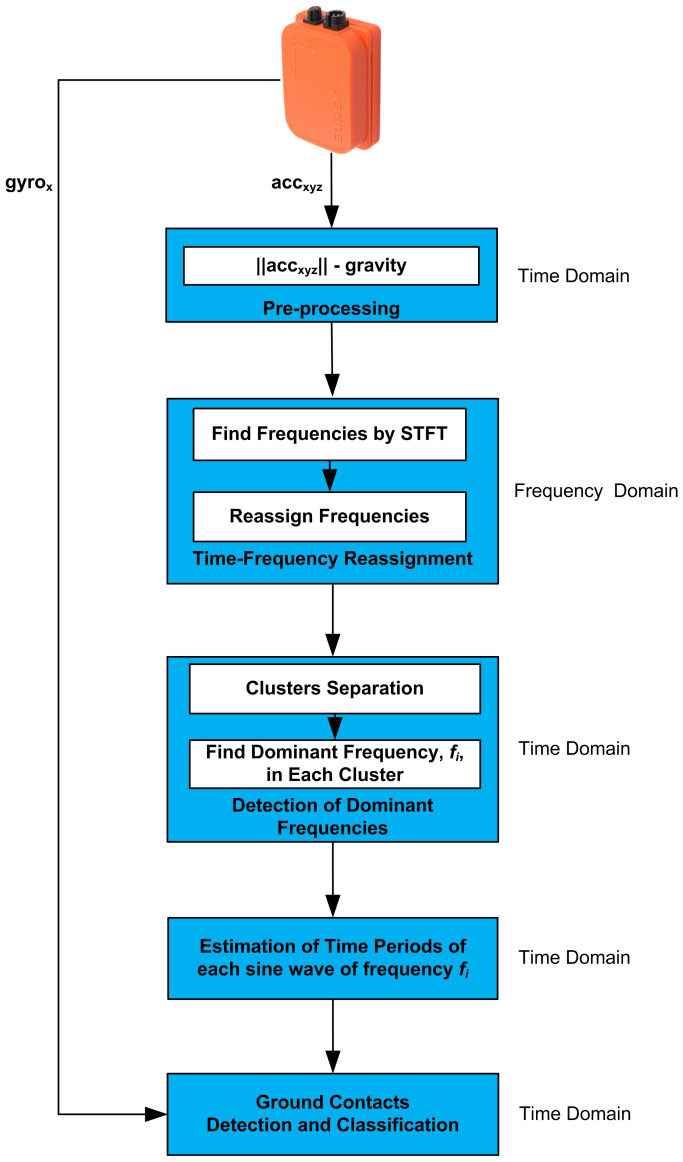


Figure 3.2: *Overview of the algorithm to estimate ground contacts from the lower trunk sensor. The estimation algorithm analyzes input signal in frequency domain in order to find dominant frequencies. Each dominant frequency is used to estimate the time period of its sine wave and angular changes around x-axis, which is in the direction of gravity, in this time period are analyzed to correctly detect and classify ground contacts.*

sharpen amplitude spectrum as proposed by [FF09]. The proposed scheme locates a signal at the correct frequency or time with a higher resolution than the inherent FFT resolution

trade-offs. The idea is to calculate instantaneous frequency (IF) in each FFT frequency band by taking first time derivative of the STFT phase at frequency ω and time T. An IF, simply speaking, is the instantaneous frequency of the dominant component at a particular time and frequency. Similarly the local group delay (GD), also known as time delay, in each FFT frequency band is calculated by taking first frequency derivative of the STFT phase at frequency ω and time T.

$$IF = \frac{\partial}{\partial(T)}[arg(STFT(\omega, T))] \tag{3.3}$$

$$GD = \frac{\partial}{\partial(\omega)}[arg(STFT(\omega, T))] \tag{3.4}$$

where the function *arg* returns the phase angle of the transform. These local estimates of instantaneous frequency and group delay are used to correct the time and frequency coordinates of the spectral data, and map them back onto the true regions of support of the input signal.

Due to the noisy nature of accelerations, a simple spectral analysis of the input signal in frequency domain using STFT does not yield any significant information. For this reason, we have used the Time-Frequency reassignment method presented by [FF09]. Their method is specific to audio and signal processing domain where signals have higher frequencies. In our case, the signal frequency is ultra low e.g., the frequency of a normal human walk is around 1.5-2Hz. However, by setting correct input parameters such as window size, window overlap size, discrete Fourier transform analysis length (fftn) etc., the algorithm was successfully applied to signals with ultra-low frequencies. We have used a window size of 25 frames (i.e. 1 second@25Hz sampling rate), a window overlap size of 24 frames, and an fftn length of 512 frames.

From the time-frequency reassignment method, we get various frequencies in time domain along with their intensity on a decibel scale. Since the reassignment method finds the correct time and frequency coordinates of the spectral data, so the estimated frequencies are grouped together around different time periods as shown in Figure 3.3. We call these groups of frequencies *clusters of frequencies*. It is important to note that for each acceleration cycle, we get exactly one cluster of frequencies and clusters are separated by clear boundaries. Figure 3.3 shows results of an input signal when passed through the Time-Frequency reassignment method. The input signal, which is magnitude of 3D accelerations (gravity removed) from the lower trunk sensor, is shown in the top image. Clusters of frequencies are detected around different time periods as shown in the

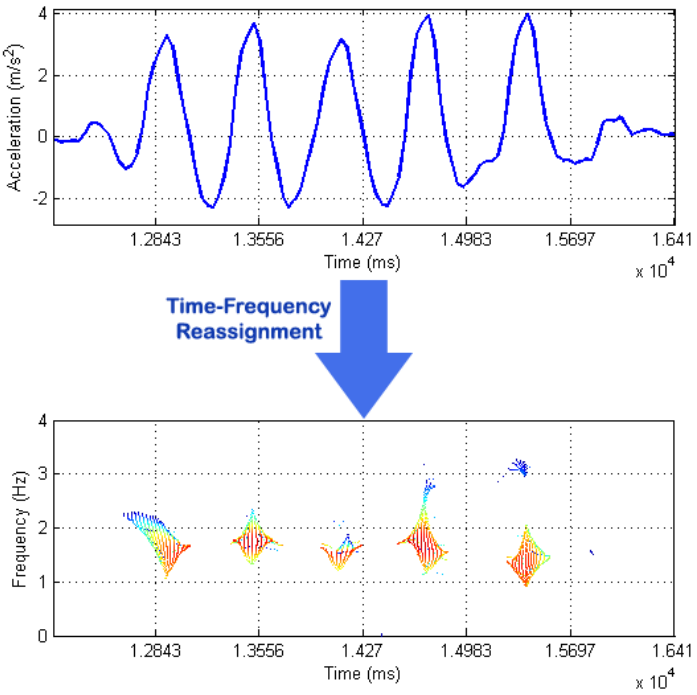


Figure 3.3: *An example of Time-Frequency reassignment of a signal. The top image shows the input signal (magnitude of acceleration - gravity) from the lower trunk sensor. The bottom image shows groups of frequencies (points) computed by the Time-Frequency reassignment method. The color of each point indicates the intensity of each frequency (blue to red ⇒ low to high).*

bottom image. The color of each point in the scatter diagram presents the intensity of the detected frequency on a decibel scale (blue to red ⇒ low to high). For the sake of clarity, we have excluded the third axis from the diagram which depict intensity of each frequency.

3.3.3 Detection of Dominant Frequencies

In the detection of dominant frequencies phase, we start with clusters separation by using two thresholds: one to define the minimum cluster size and the other one to define the minimum distance between two clusters. We set the minimum size of a cluster to 100 frequency points detected by the time frequency reassignment method. Here, one point refers to one isolated dominant frequency at a specific time frame. A cluster with less than 100 frequency points is discarded by considering it as noise. Such clusters usually appear due to small accelerations e.g., during turn around (see Figure 3.3 around 1.58×10^4ms).

Similarly, we define minimum distance between two clusters as 50ms. This means that two clusters must be at least 50ms apart from each other to be considered as separate clusters. Any clusters which are not 50ms apart from each other are considered as a single cluster. The values of these two thresholds are chosen by experimentation. After clusters separation, a single dominant frequency in each cluster is chosen. This is achieved by selecting a frequency with highest intensity in each cluster.

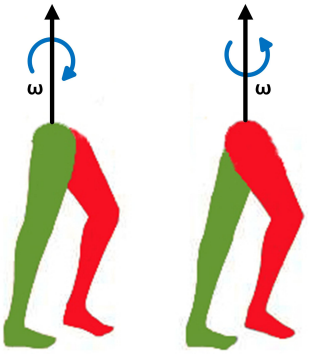


Figure 3.4: *Ground contacts detection from normal human walk. Angular velocity around the x-axis (which is in the direction of gravity) is negative when left foot swings and positive when right foot swings.*

3.3.4 Estimation of Time Periods of Dominant Signals

In this phase, we estimate time periods of the dominant frequencies chosen in the previous phase. The time periods give important information about when a movement is started (on a time scale) and when one cycle of this movement is completed. For example, assume that a dominant frequency of 2Hz is detected. This means that *2* cycles of movements were carried out in one second. If the movement is started at t seconds then, the cycle of first step will finish in $t+1/2$ seconds and the cycle of second step will finish in $t+1$ seconds. We have used the same indirect relation between time periods of dominant signals and walk cycles to detect walk steps. Let T be the desired time period of the signal and f be the dominant frequency of the signal computer in previous step, then T is computed as:

$$T = \frac{1}{f} \tag{3.5}$$

3.3.5 Ground Contacts Detection and Classification

The final step of our proposed algorithm is detection and classification of ground contacts. From the estimated time periods of each dominant frequency, we exactly know the time periods of each acceleration. However, we still do not know which leg generated which acceleration. In this context, we analyze angular velocity in x-axis, which is in the direction of gravity, to correctly classify an acceleration. It is important to note that in our proposed algorithm we only consider such human movements in which the lower back is straight and x-axis always remain in the direction of gravity. Figure 3.4 shows how the angular velocity around x-axis behaves when left foot and right foot are accelerated. The angular velocity around x-axis is positive when right foot is accelerated while left foot is on stance. Similarly, the angular velocity around x-axis is negative when left foot is accelerated and right foot is on stance. Let T be the time period of a dominant signal estimated by equation (3.5), ω_T be the average angular velocity around x-axis in T, then left foot ground contact, g_{left}, and right foot ground contact, g_{right}, are computed by the following equations:

$$g_{\text{left}} = \begin{cases} 1 & \text{if } \omega_T > 0 \\ 0 & \text{if } \omega_T < 0 \end{cases} \tag{3.6}$$

$$g_{\text{right}} = \begin{cases} 1 & \text{if } \omega_T < 0 \\ 0 & \text{if } \omega_T > 0 \end{cases} \tag{3.7}$$

Here a value of *1* indicates that the respective foot has a ground contact and a value of *0* indicates that the respective foot has no ground contact.

Estimated ground contacts are classified into one of three categories: a) *ground contact*, b) *not a ground contact*, and c) *unreliable*. It is important to note that our proposed algorithm can classify several frames into *unreliable*. When a dominant frequency is detected, the algorithm estimates ground contacts of all frames in *only* one cycle of the dominant frequency. We call these frames as *key frames* where the reliability of the classified ground contacts is maximum. As soon as one cycle is completed, the algorithm waits for detection of next dominant frequency. All frames during this wait state are not classified and treated as *unreliable* frames (e.g., see Figure 3.5 around $1.335 - 1.352 \times 10^4 \text{ms}$). In case of stationary human body during a wait state, the algorithm simply finds zero-acceleration time periods and classify them into ground contacts (e.g., see Figure 3.5 after

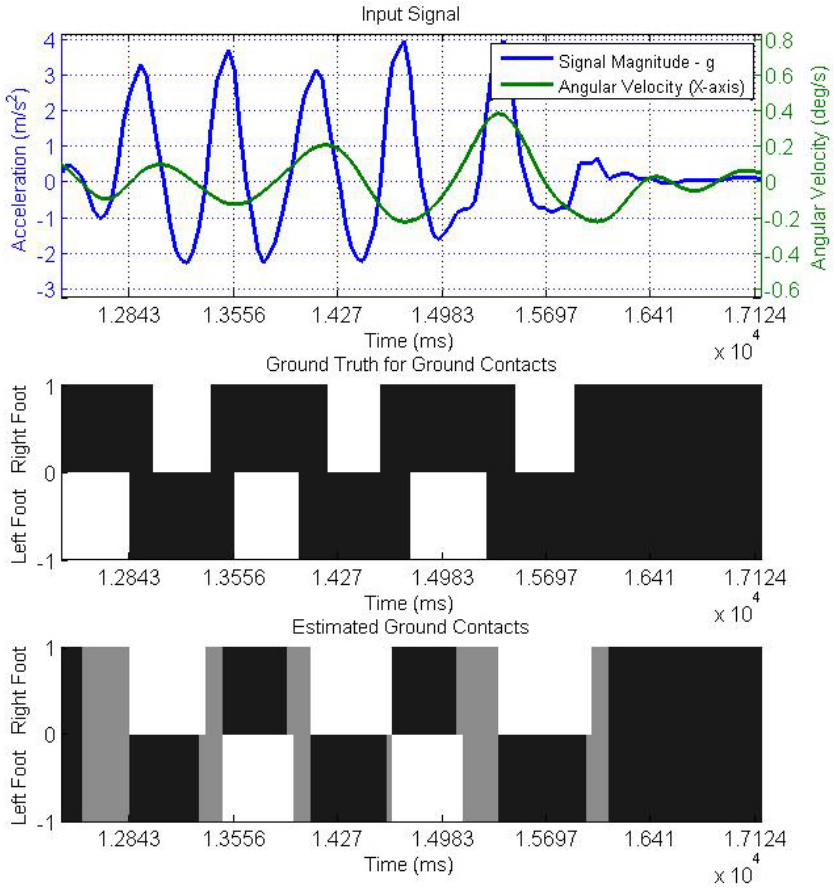


Figure 3.5: *Results of estimation of ground contacts from the lower trunk sensor. The input signal (the acceleration and the normalized angular velocity), the ground truth for ground contacts, and estimated ground contacts are shown in top, middle, and bottom images respectively. The black patches represent 'ground contact', the white patches represent 'not a ground contact', and the gray patches represent 'unreliable' frames.*

1.641×10^4ms). The algorithm, however, is not capable of distinguishing between single stance and double stance in static postures and it classifies both feet on ground in such cases. The idea of *key frames* means that it is possible to reconstruct visually appealing motions with only few reliable ground contacts. In the results section (Section 3.5), we will further prove this idea with experimental results. Analysis of the acceleration signals in the frequency domain to find *key frames* has several advantages over a simple peaks detection approach such as easier and less error prone decomposition of signals, better time period estimates, and hence more accurate ground contacts estimates.

Figure 3.5 shows the results of estimation of ground contacts from the accelerations of the lower trunk sensor. The input signal (the acceleration and the normalized angular velocity), the ground truth for ground contacts, and estimated ground contacts are shown in top, middle, and bottom images respectively. The black patches represent *ground contact*, the white patches represent *not a ground contact*, and the gray patches represent *unreliable* frames. An estimated right foot *ground contacts* are seen around $1.2843 - 1.335 \times 10^4 \mathrm{ms}$. Left foot is classified as *not on ground* in this time period. An *unreliable* state is estimated around $1.335 - 1.352 \times 10^4 \mathrm{ms}$.

3.4 Motion Reconstruction from Accelerometers and Ground Contact Information

In this section we discuss in detail our approach to reconstruct motion sequences on the basis of accelerometer readings that are constrained with ground contact information.

3.4.1 Control Signals and Knowledge Base

Our motion synthesis technique can be controlled by input signals that can come from various sensor setups. We consider two different types of sensor signals in this work. First, accelerometer readings that measure 3D accelerations in a sensor coordinate frame. We assume accelerations to be measured in SI units, thus, it is measured in $\mathrm{m/s^2}$. Second, ground contact information: For each foot, the ground contacts signal has three values, a *true* — indicating a ground contact, a *false* — indicating a no ground contact, and a *null* — indicating an unreliable state. In the reconstruction process, the ground contact information is only considered for such frames where ground contacts are reliable (i.e. either true or false) and ignored otherwise.

We denote accelerometer readings at time t with α_t and the ground contact state vector with β_t. The vector α_t is of dimension $3 \times N$, where N denotes the number of employed sensors. In contrast the vector β_t of ground contacts is two dimensional.

Throughout this work we distinguish between real accelerometer data, captured by XSens MTx [Xse14] devices and virtual data. The virtual sensors are placed on the limbs in the same way as the real sensors in our simulation. Positions of these sensors are computed using forward kinematics. Accelerations in the global coordinate frame are obtained by computing the second time derivatives. After adding accelerations due to

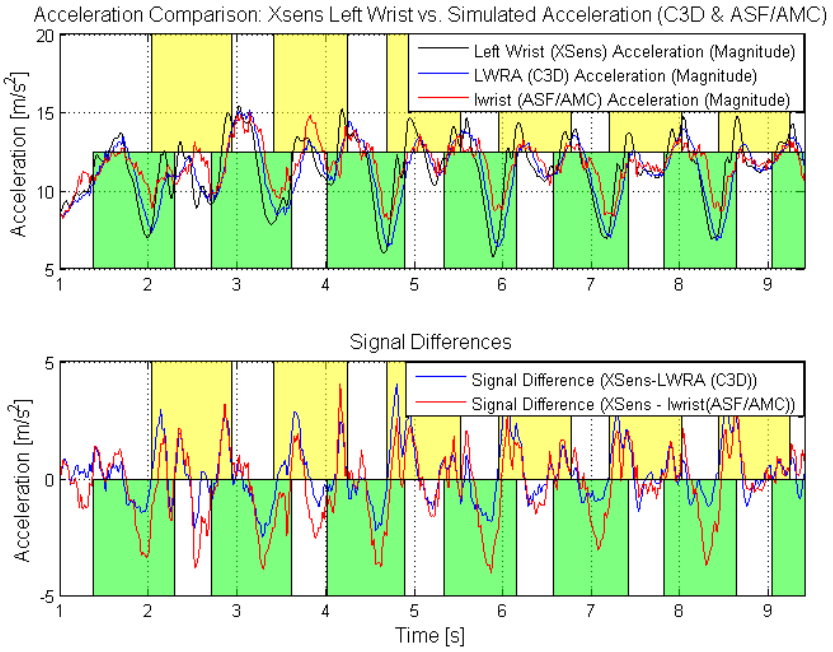


Figure 3.6: *Comparison between the results of measured MTx data and simulated accelerometer data obtained from C3D marker data and a skeletal joint position. The magnitudes of simulated accelerometer data on a left wrist marker, the left wrist joint, and a MTx sensor attached to the left wrist marker are given for the motion HDM08_it_03_01 are given. Using hand annotations ground contacts of the left foot are marked in yellow, of the right foot in green.*

gravity we obtain the simulated signal by transforming the global accelerations to the virtual sensors local system.

As shown by Tautges et al. [TZK+11], the results of virtual sensors and real sensor readings coincide quite well, and reconstruction results can be reliably estimated by virtual sensor data. However, for purposes like ground contact estimation the data of the virtual sensors might differ significantly. In Figure 3.6 we give a comparison between the results of measured MTx data and virtual accelerometer data obtained from C3D marker data and a skeletal segment position.

For real sensors data, ground contacts are estimated using the method proposed in section 3.3. We employ a motion capture database as a knowledge base for the reconstruction process. In the knowledge base, 86.42% of the frames are manually annotated for ground contacts by several annotators. A high quality of the annotations is guaranteed by performing a voting on the annotations of six annotators. The remaining frames are

annotated by a heuristic algorithm, which identifies ground contacts by detecting local minima of the foot velocity and proximity to the ground. We assume the ground contact information in the knowledge is accurate and reliable.

3.4.2 Ground Contact Constrained Motion Retrieval

Due to the ambiguity in acceleration data the search for similar poses in a huge motion database on the basis of pure accelerometer readings is a hard problem. Tautges et al. [TZK+11] replaced the frame-wise knn-search with a search for a sequence of similar readings within a window of successive frames to resolve for ambiguity. Although their procedure works well in the setup with four motion sensors, it does not give the desired results if applied to a more sparse sensor setup directly. To tackle this problem, we propose to constrain the windowed nearest neighbor search on the basis of the ground contact information.

In our sparse sensors setup, we do not have any direct control inputs from lower extremities. However, we rely on the control input from the lower trunk sensor in order to detect ground contacts. The ground contact information is used to perform windowed nearest neighbor search. In the reconstruction process, we additionally use accelerations from both wrists and from the lower trunk.

Let us assume that the control input consists of a continuous stream of sensor accelerations $(\ldots, \boldsymbol{\alpha}^{t-2}, \boldsymbol{\alpha}^{t-1}, \boldsymbol{\alpha}^{t}, \ldots)$ and ground contacts $(\ldots, \boldsymbol{\beta}^{t-2}, \boldsymbol{\beta}^{t-1}, \boldsymbol{\beta}^{t}, \ldots)$, where $\boldsymbol{\alpha}^t$ and $\boldsymbol{\beta}^t$ denote the accelerations and ground contacts at time t, and our goal is to find the poses that match the actual pose and ground contacts at time t.

In order to perform a windowed nearest neighbor search, we consider the current M sensor readings $(\boldsymbol{\alpha}^{t-M+1}, \ldots, \boldsymbol{\alpha}^{t})$ and ground contacts $(\boldsymbol{\beta}^{t-M+1}, \ldots, \boldsymbol{\beta}^{t})$ for a fixed number $M \in \mathbb{N}$. Let K be the number of locally nearest neighbors, and let S^t be the set of indices representing the K nearest neighbors of $\boldsymbol{\alpha}^t$, we search for K nearest neighbors using kd-tree based nearest neighbors search method as proposed by Krüger et al. [KTWZ10]. These K nearest neighbors are used to build the nodes of the OLNG as proposed by Tautges et al. [TZK+11]. The nodes can be represented by an $M \times K$ array, where the m^{th} column, $m \in [1 : M]$, corresponds to the sorted set S_{gc}^{t-m+1} (see Figure 3.7a for an illustration). Our goal is now to identify a subset of S_{gc}^t that is reliable enough to build a local statistical model of poses and matches the ground contacts at time t.

In the update process, instead of selecting all nodes with minimal costs as in OLNG,

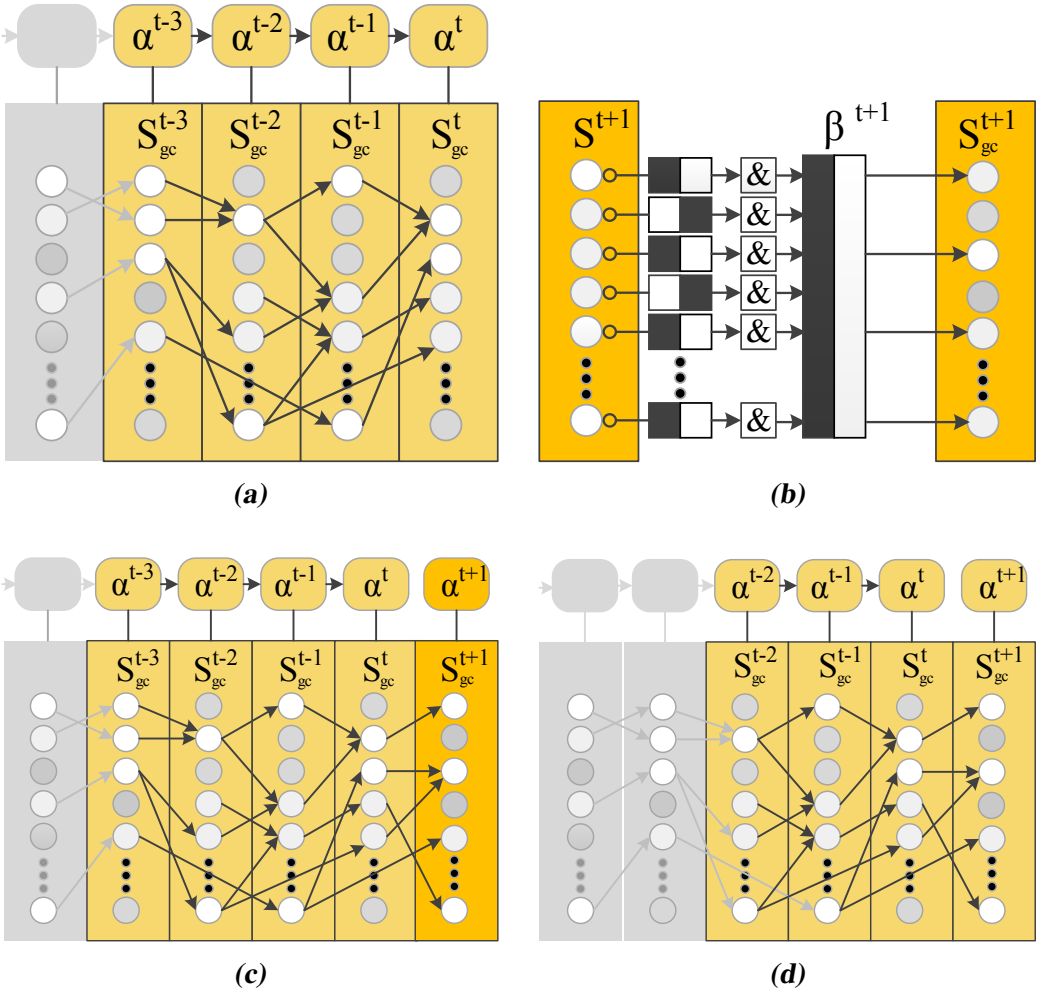


Figure 3.7: *A toy example of the implementation of the Ground Contact Constrained Motion Retrieval with $M = 4$ and K nearest neighbors. Each vertical column corresponds to the K nearest neighbors (each neighbor indicated by a circle) of a sensor reading $\boldsymbol{\alpha}^{t-m+1}$ and ground contacts $\boldsymbol{\beta}^{t-m+1}$, $m \in [1 : M]$. The dark yellow column represents the K nearest neighbors that have been currently retrieved. The white circles represent the nodes which are covered by the path while the gray circles represent the nodes which have been discarded. In (b), ground contacts are composed of left foot ground contacts (left column patches) and right foot ground contacts (right column patches). The black patches represent 'ground contact', the white patches represent 'not a ground contact'.*

we use ground contact information to restrict nodes selection. We name this approach as Ground Contact Constrained Motion Retrieval (GCCMR). Keeping in mind the concept

of key frames we used in estimation of ground contacts from the lower trunk sensor (see Section 3.3), ground contacts are only compared if the ground contact information of the input frame is reliable. For reliable ground contacts, we compare them and only matching nodes are considered regardless of their cost. All nodes with non-matching ground contacts are discarded in this phase, regardless of their costs. On the other hand, if the ground contact information of the input frame is not reliable, we do not match ground contacts and consider costs of the nodes, same as OLNG, and nodes with lowest costs are selected. All high cost nodes are hence discarded in this phase.

For inputs with reliable ground contacts, the update process is explained as follows. Since we have ground contact information in the knowledge base, which are directly comparable with the control input $(\ldots, \boldsymbol{\beta}^{t-2}, \boldsymbol{\beta}^{t-1}, \boldsymbol{\beta}^{t}, \ldots)$, we can make a ground contacts comparison when adding new nodes to the graph. Suppose that the graph has been constructed for the readings $(\boldsymbol{\alpha}^{t-M+1}, \ldots, \boldsymbol{\alpha}^{t})$ and that for each node in S^{t}_{gc} we have identified the path with lowest costs leading there. Now a new reading $\boldsymbol{\alpha}^{t+1}$ arrives, the K nearest neighbors are retrieved (using the kd-tree) and stored in S^{t+1}. For each new node, $\boldsymbol{\beta}^{t+1}$, if the ground contact information of $\boldsymbol{\alpha}^{t+1}$ is reliable, it is first compared with the ground contacts, see Figure 3.7b. If the ground contact is not matched for either one of the two feet, the new node is removed from S^{t+1}. In this way, the nodes that have similar wrist accelerometer readings but mismatching ground contacts are not added to the GCCMR, see Figure 3.7d. Thus, S^{t+1}_{gc} has less elements than S^{t+1} which results in faster computation times for path construction and corresponding cost updates.

3.4.3 Motion Reconstruction

We formulate the pose synthesis per frame as an energy optimization problem following one of the standard approaches in data-driven motion synthesis. We use the same energy function as introduced by Tautges et al. [TZK+11]:

$$\boldsymbol{q}_{\text{best}} = \underset{\boldsymbol{q}}{\operatorname{argmin}}(w_{\text{prior}} \cdot E_{\text{prior}}(\boldsymbol{q}) + w_{\text{contr}} \cdot E_{\text{contr}}(\boldsymbol{q})). \tag{3.8}$$

This energy function consists of two components: a data-driven *prior term* and a *control term* ensuring that the resulting motion corresponds to the actual sensor readings. The prior model consists of three different components: a) a *pose prior*, which characterizes the probability of a pose with respect to the distribution in pose space, b) a *motion prior*, which computes the likelihood of a pose considering the temporal evolution of a motion.

c) a *smoothness prior* to reduce jerkiness.

Table 3.1: *Results of ground contacts estimation from the lower trunk sensor. Two different aspects of the algorithm are tested: accuracy of walk steps classification and quantitative frame by frame detection of ground contacts for each foot.*

Motion Type	Walk Steps			Right Foot Ground Contacts			Left Foot Ground Contacts		
	Total Steps	Detection Rate (%)	Classif. Rate (%)	Total Frames	Total Classif.	Classif. Rate (%)	Total Frames	Total Classif.	Classif. Rate (%)
Normal Walk	120	90.0	83.0	741	666	89.9	744	615	82.7
Circular Walk	90	80.0	60.0	414	321	77.5	432	267	61.8
Sideways Walk	60	90.0	50.0	408	219	53.7	417	162	38.8
Walk & Wave	84	85.7	71.5	402	327	81.3	399	327	82.0
Walk & Greet	72	91.7	83.3	510	312	61.2	495	396	80.0
Walk Forwar & Backward	60	93.3	86.7	420	183	81.6	414	201	79.2
March with Wide Arms	60	93.3	90.0	432	384	88.9	432	375	86.8
Average	–	89.1	74.9	–	–	76.3	–	–	73.0

3.5 Results

The proposed methodology has been extensively tested under different criteria. We have carried out a number of experiments on both simulated sensors data and real sensors data. In this section, we will explain our findings in detail.

Our presented approach, which requires accelerations from both wrists and a lower trunk, and ground contacts as input, is directly comparable with the approach proposed by Tautges et al. [TZK+11], which requires four control inputs — accelerations of both wrists and both ankles. For the sake of comparison, the experimental setup has been kept same in all experiments.

We have grouped the results into three subsections, which are directly related to our three contributions mentioned in Section 3.1. In the first subsection, we discuss results of ground contacts estimation from the lower trunk sensor. In the second subsection, we present motion reconstruction results of several experiments, in which virtual accelerations are used as input. This includes effects of size and diversity of knowledge base on reconstruction, effect of window size on reconstruction, and how reconstruction er-

ror changes with different sensor setups. In the third subsection, the results from real accelerations and ground contacts are discussed.

3.5.1 Ground Contacts Estimation from Lower Trunk Sensor

We have tested our proposed algorithm with several different types of motions. All motions are performed by three different actors; two males and one female. For ground truths we have used Sensor Insole by Moticon [Mot14]. It is a fully integrated sensor which is placed as an insole inside a shoe. It is a wireless sensor consisting of 13 pressure sensors and an accelerometer, which can measure pressure data and acceleration of the foot. The pressure data from this sensor is used as a ground truth for a quantitative comparison.

The proposed algorithm has been tested in two different aspects. Firstly, we are interested to see how accurately the algorithm can detect walk steps. In this case, we are only interested to find a single key-frame of ground contact for each step. Secondly, for how many frames the algorithm can accurately detect ground contacts for each foot. To achieve this, a frame to frame comparison with the ground truth data is carried out.

The results of both aspects of testing are given in Table 3.1. The results under column *walk steps* correspond to the first aspect. The detection rate of foot ground contacts (under column *Detection Rate*) is around 89% on average. However, the classification of ground contact into left or right foot ground contact is much lower, around 75%. This is due to the fact that the algorithm relies on the angular velocities around the x-axis, which are known to be very noisy and hence have direct impact on classification. Also, the algorithm assumes that the angular velocities around the x-axis is always in the direction of gravity, but for certain motions such as *circular walk* this is not the case. The classification results are degraded for motions like *sideways walk* because the rotation of the hips is not around x-axis.

The results of the second aspect, finding number of frames with accurate ground contacts, are presented in last six columns of Table 3.1. The average classification rate of right foot ground contacts is 76.3% and the average classification rate of left foot ground contacts is 73.0%. The algorithm performed best in *normal walk* and worst in *sideways walk*, as expected. Another interesting observation is that the classification rate of both feet does not match with each other in any motion. This is due to the fact that in the detection of dominant frequency phase of our algorithm, there are many candidate frequencies. We choose one which has highest intensity, however, it is possible that the time period of the chosen frequency do not exactly match with the acceleration. Such an offset, which

can be of several milliseconds, effects the overall classification rate.

In summary, our proposed algorithm is better in detecting ground contacts (around 89%) and degrades when classifying ground contacts into left or right foot ground contacts. The numbers do not look impressive, however, we still see visually appealing reconstruction results. Our concept of using key frames implies preferring reliability over quantity. This means that even a lower number of accurately classified ground contacts can greatly improve overall reconstruction. The reconstruction results of our proposed algorithm prove that visually appealing reconstruction is possible with key frames.

3.5.2 Motion Reconstruction from Virtual Accelerations and Ground Contacts

We have tested our proposed algorithm in several experimental setups using virtual wrists and lower trunk accelerations, and ground contacts as input data. In this subsection, we will discuss our findings both quantitatively and qualitatively.

3.5.2.1 Experimental setup and test data

The knowledge base of the following experiments is composed of motion clips taken from the publicly available motion database HDM05 [MRC+07]. The motions in the database were performed by five different actors, referred to by their initials (`bd`, `bk`, `dg`, `mm`, `tr`). Accordingly, we use the following naming pattern to denote different knowledge bases:

`HDM_{actor}_{part}-{scene}_{take}_{framerate}`.

The asterisks are used as wild-cards to represent any possible value of that field. "M" is used to represent mirrored motions (switching the left and right side of the actor). For example, the knowledge base `HDM_bd_01-**_**_25M` represents all motion clips from Part 1 of the HDM05 database performed by the actor `bd`, together with their mirrored copies, down sampled to 25Hz.

In all of the following experiments the accelerometer data was obtained by simulating virtual sensors using a set of test motions also obtained from the HDM05 database. The ground contacts were retrieved from the annotations of the same test motions. The actual test motion itself was never included in the knowledge base. Three different scenarios were defined by whether the actor of the considered test motion was included in the knowledge base or not.

- DB_1: All motions of five actors contained in the HDM05 database together with mirrored copies.
- DB_2: A subset of DB_1, contains motions of the actor to be reconstructed.
- DB_3: A subset of DB_1 without motions of the actor to be reconstructed.

3.5.2.2 Size and diversity of the knowledge base.

In this experiment, we analyzed how the size and diversity of the knowledge base influences the motion reconstruction quality using our approach. Five sets of test motions (one for each actor) were created. Each motion contains six motion clips from Part 1 of the HDM05. We chose motion clips of `HDM_**_01-**_02` for every actor and every motion class as our test motions. We composed a set of 18 different knowledge bases, which are of different size and diversity and reflecting all the three scenarios. We then reconstructed all test motions of every actor using the 18 different knowledge bases reflecting the previously described scenarios.

The reconstructed motions were compared to the original test motions using the Root Square Mean (RMS) error of joint positions. The average reconstruction errors are shown in Figure 3.8. The columns give the name of the knowledge base, the scenarios, the size of the knowledge base (with average numbers of frames), and the color-coded averaged RMS errors for each actor of the reconstructed motions.

The results further prove the following three facts:

- When the motions of the actor to be reconstructed are contained in the knowledge base, the reconstruction quality will be significantly improved (scenario 1 and 2). This can directly be seen by the prominent blue diagonals in Figure 3.8 (a) and 3.8 (c) representing scenario 2. In Figure 3.8 (a) and (c), all diagonal entries present scenario DB_2 where motions of the actor to be reconstructed are included. Similarly, all non-diagonal entries present scenario DB_3 where motions of the actor to be reconstructed are excluded.
- When knowledge base becomes larger and less homogeneous, the reconstruction quality only slightly decreases. This can be seen by comparing rows corresponding to same scenarios in Figure 3.8 (b) and 3.8 (d).
- Increasing the diversity has a higher influence on the results than a mere inflation of the knowledge base, see the results of actor `mm`.

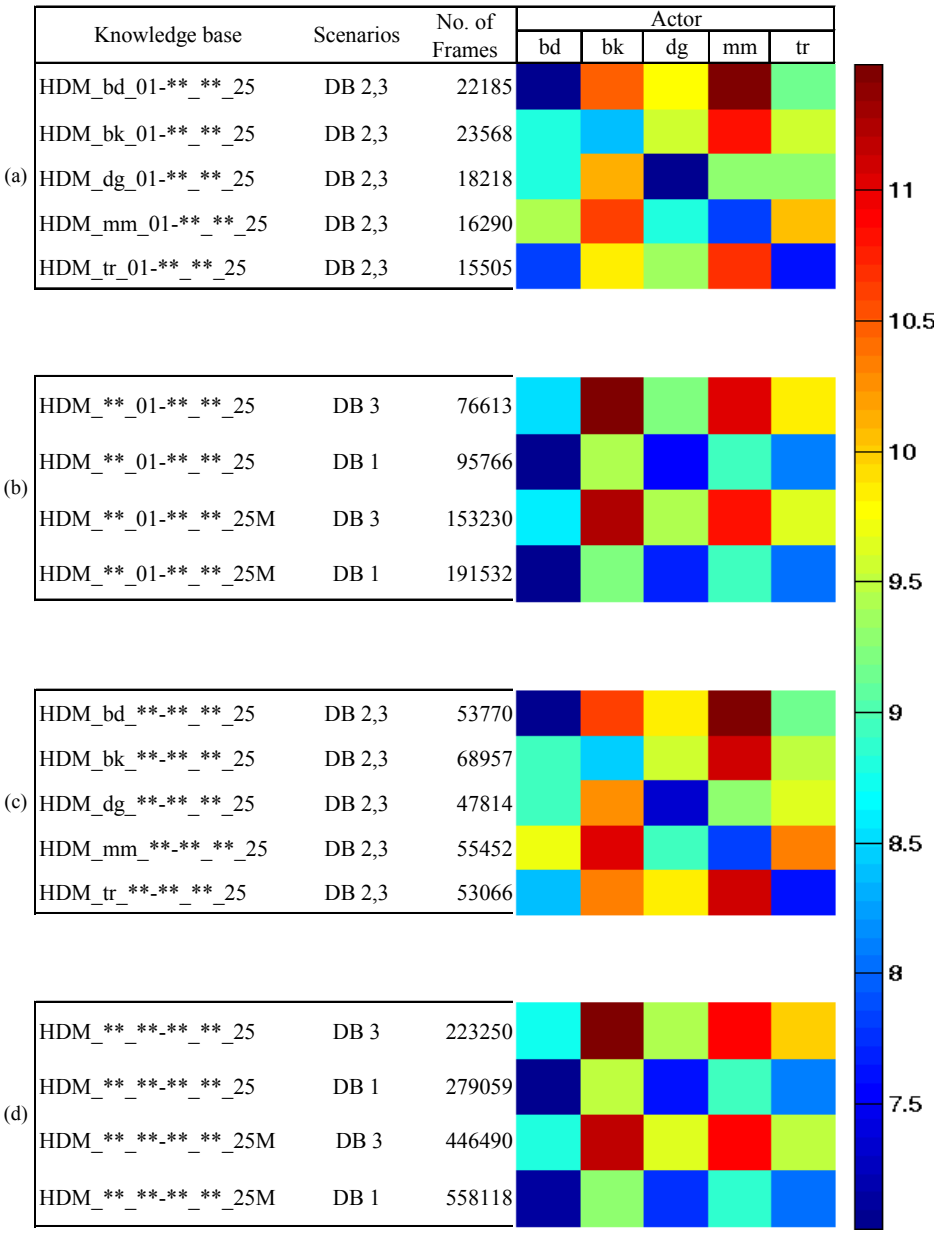


Figure 3.8: *Average reconstruction error for different sets of knowledge base and different actors to be reconstructed.*

A comparison with one of the existing methods is presented in Figure 3.9. We have chosen all motions of actor *dg* as test cases. It is observable that the average reconstruction error of our proposed method is much smaller in most cases such as motions 3, 7, 9, 12, 16, and 18. In some cases (such as motion#3, motion#12), the average reconstruction error of

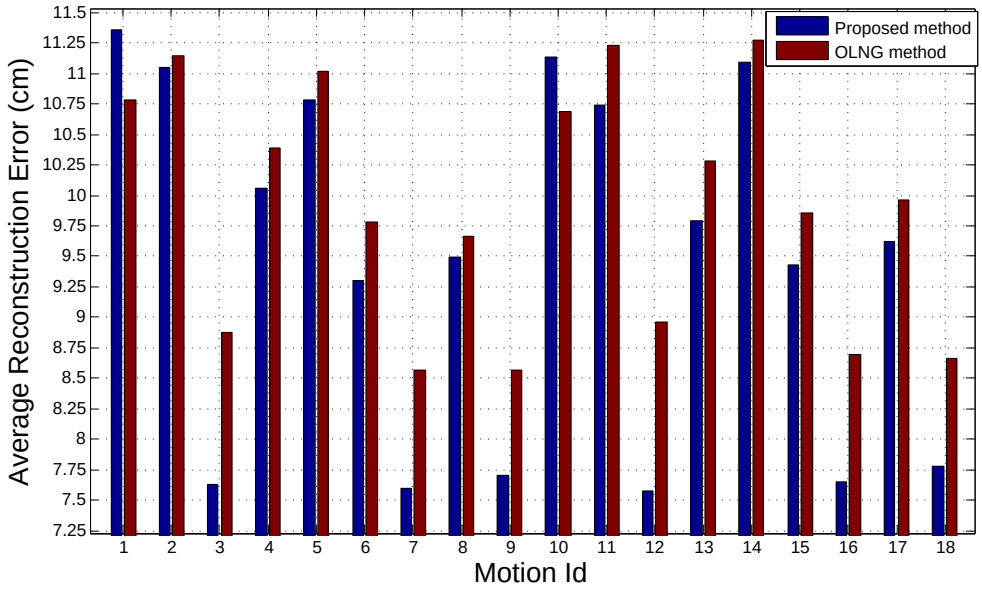


Figure 3.9: *Comparison of average reconstruction error between proposed method and the OLNG method. The average reconstruction error is much smaller in motions 3, 7, 9, 12, 16, and 18. The average reconstruction error difference is up-to 1.39cm in certain motions such as motions 3 and 12.*

our proposed method is up to *1.39cm* less than OLNG method. From these results, we can safely conclude that our proposed method, which uses only three input sensors and ground contacts, either outperforms the OLNG method proposed by Tautges et al. [TZK+11], which uses four input sensors, or produces almost same results.

Figure 3.10 shows a comparison of average reconstruction error between the proposed method, the original OLNG method, and the OLNG method with ground contacts. Our proposed method produces lower average reconstruction error than the original OLNG method in all three scenarios. The OLNG method with ground contacts (i.e. 4 sensors + ground contacts) performs better than our proposed method (with 3 sensors + ground contacts) in DB_1 and DB_3. However, it is important to note that we have extended the original OLNG method to incorporate ground contacts for the sake of comparison only. This variant of the OLNG algorithm is new and it does not exist in the original article. Moreover, this further proves our idea that the use of ground contacts in full body motion reconstruction produces better results by minimizing average reconstruction error.

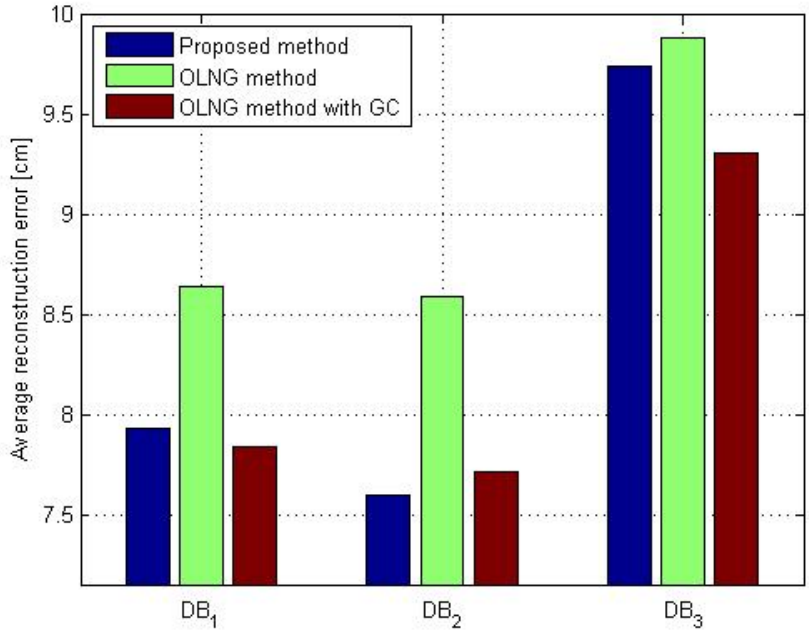
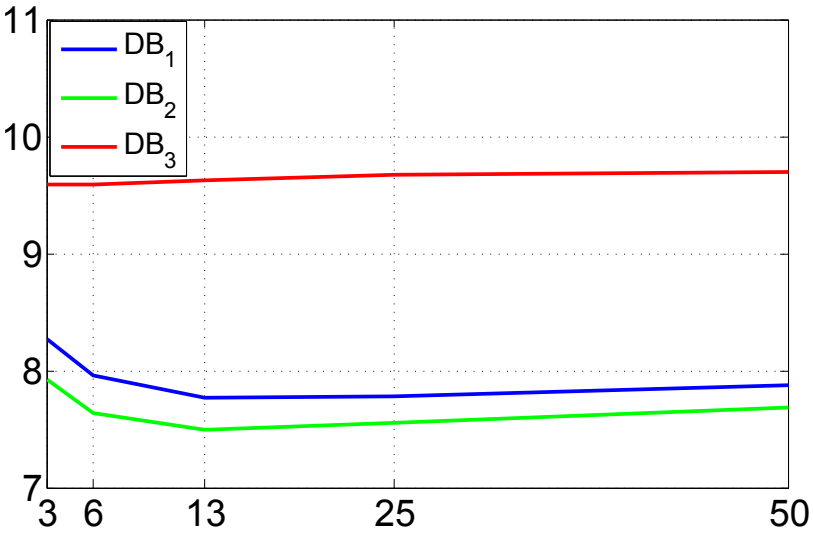


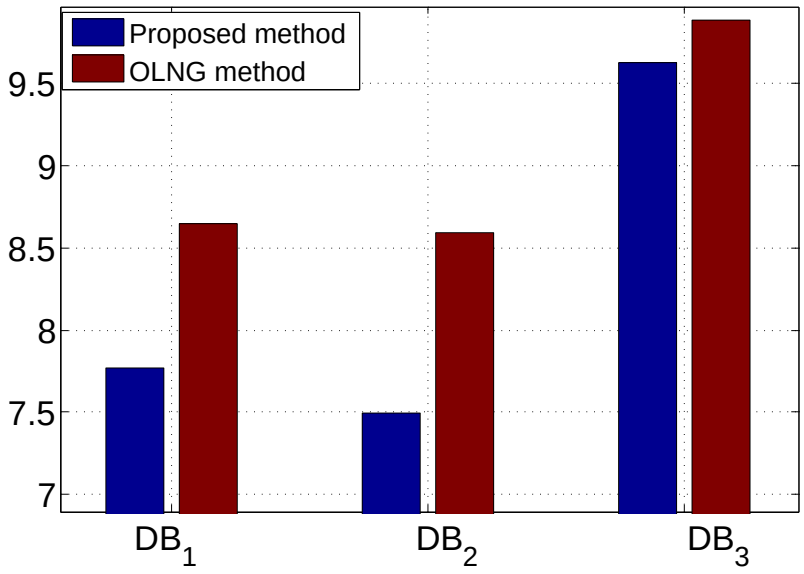
Figure 3.10: *A comparison of average reconstruction error between the proposed method, the original OLNG method, and the OLNG method with ground contacts. Our proposed method produces lower average reconstruction error than the original OLNG method in all three scenarios. The extended OLNG method with ground contacts (i.e. 4 sensors + ground contacts) performs better than the proposed method (3 sensors + ground contacts) in DB_1 and DB_3.*

3.5.2.3 Window size

As mentioned in section 3.4.2, the window size, M, is used to retrieve paths in the updating process. It is an important parameter which will affect the motion reconstruction quality of our method. Thus, we evaluate the average reconstruction error by incrementally increasing the value of M from 3 to 50 using different knowledge bases (DB_1, DB_2, DB_3). Figure 3.11a shows the reconstruction results using M different window sizes. As can be seen, the minimum reconstruction error is obtained when using the window size of 13 frames, which corresponds to half a second at a sampling rate of 25 frames per second. A comparison of the reconstruction error of proposed method and the OLNG method is shown in Figure 3.11b, where $M = 13$. It can be seen that in all the three scenarios, our method produces better results as compared to the OLNG method.



(a)

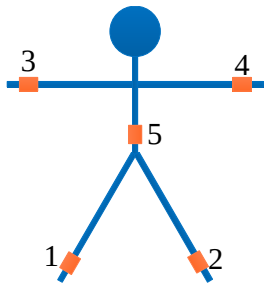


(b)

Figure 3.11: *a) Influence of the window size M of the proposed method. b) presents a comparison of average reconstruction error between proposed method (blue) and the OLNG method (red) under different scenarios, where $M = 13$.*

3.5.2.4 Tests with different sensor setups

Since ground contacts serve as an additional control input in our approach, it is important to evaluate how the number and placing of sensors affect the reconstruction quality when



Sensors setup	Number of ankle sensors	Used sensors
A	0	[3]
B	1	[2 3]
C	0	[3 4]
D	1	[2 3 4]
E	0	[3 4 5]
F	1	[2 3 5]
G	1	[2 3 4 5]
H	2	[1 2 3 4]

Table 3.2: *Different sensor setups*

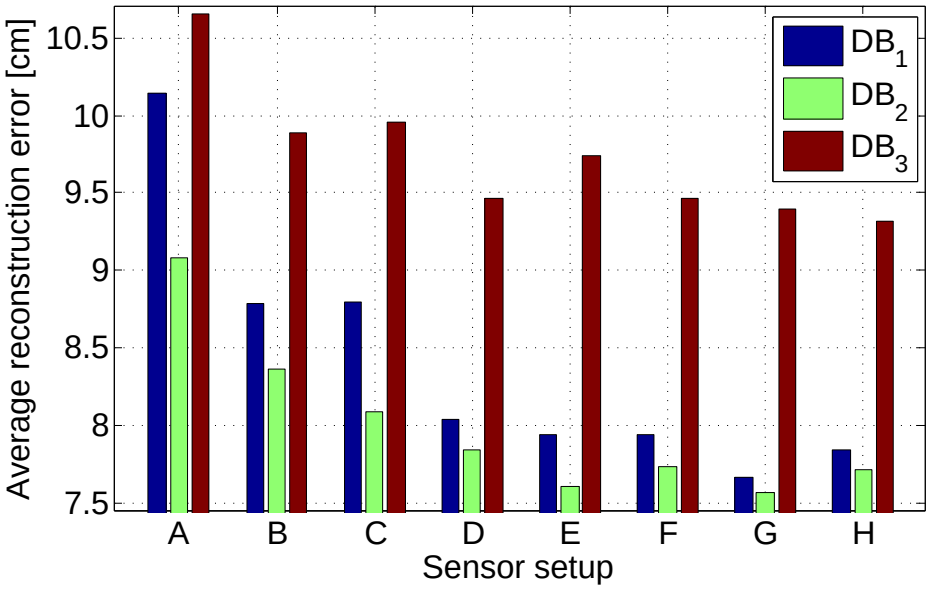


Figure 3.12: *The average reconstruction error on different sensors setup.*

using the ground contacts. In the experiment we reduce the number of ankle sensors to 1 or even to 0 (e.g., A–G in Table 3.2). Figure 3.12 shows the average reconstruction error for these different sensor setups. As can be seen from the figure, additional sensors improve the reconstruction quality as less information needs to be inferred from the knowledge base, especially when the ground contacts are presented. It is observed that when more than two sensors are used, the reconstruction quality is substantially improved (see sensor setup D–H), particularly when the trunk sensor is used (comparing D with F, and G with H). This is because the trunk sensor could help to discriminate same motions of the hands and feet (like walking and walking with a bending trunk). On the other hand, when

the sensors become more sparse (e.g., less than three), more errors will be introduced as expected. Notice that setup H corresponds to the standard setup used by Tautges et al. [TZK+11] *enhanced with ground contact information*, whereas setup E corresponds to our standard setup. Thus, it can be concluded that when using the three sensors and ground contacts, the best sensor setup for three sensors is E. Same results were observed by Kim et al. [KSL13] for three sensors setup without ground contacts.

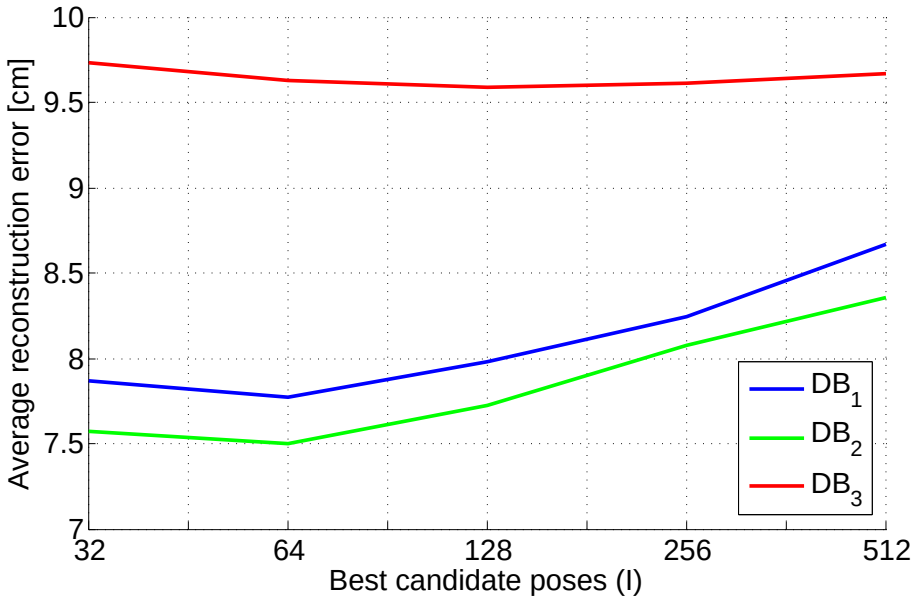


Figure 3.13: *Influence of the size of chosen best candidate poses on reconstruction. Lowest average reconstruction error is seen for a size of 64 best candidate poses.*

3.5.2.5 Best pose candidates size

The best pose candidates are a few selected poses from the nearest neighbor search set, which are further processed to choose the best pose. The number of selected poses has a direct effect on both reconstruction errors and processing time. A smaller number minimizes processing time compromising on reconstruction quality i.e. produces higher reconstruction errors. Similarly, a larger number minimizes reconstruction errors but requires more processing time. Hence, it is important to choose an optimal size of best pose candidates which takes reasonable processing time and keeps reconstruction error to minimum.

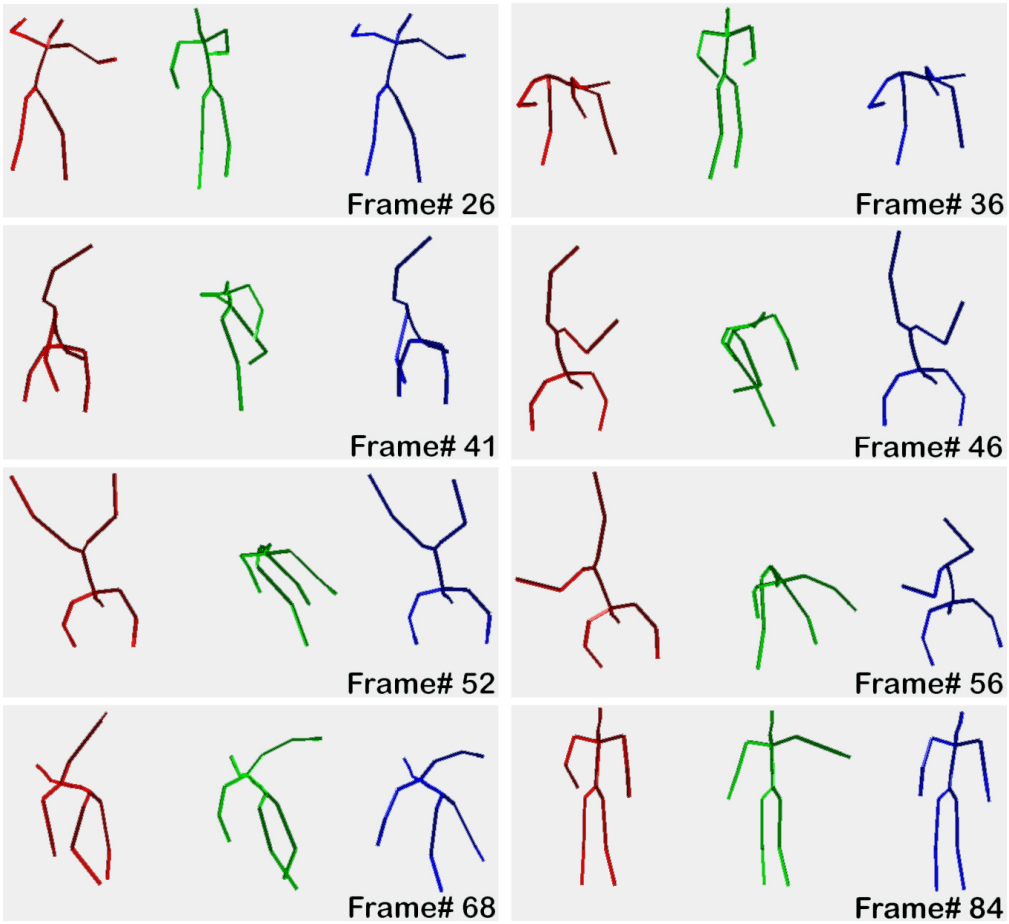


Figure 3.14: *Results of reconstruction of a cartwheel motion showing 8 selected frames. The red skeleton represents the ground truth, the green skeleton represents the reconstructed motion using the OLNG method, and the blue skeleton represents the reconstructed motion using the proposed method. Visually appealing reconstruction results can be observed especially in 2nd and 3rd rows where the OLNG method fails.*

In order to choose an optimal number, we have performed several experiments with different sizes of best pose candidates. In this context, we have chosen all motions of actor *dg* as a test case and performed experiments with DB_1, DB_2 and DB_3. The results are presented in Figure 3.13. From the graph, it is observable that the average reconstruction error is minimum for a size of 64 for all three knowledge bases. The average reconstruction error increases for all other sizes.

Figure 3.15: *Examples of reconstructed motions showing 3 frames for each motion. The red skeleton represents the ground truth, the green skeleton represents the reconstructed motion using the OLNG method, and the blue skeleton represents the reconstructed motion using the proposed method. Different motions include (from top to bottom): 1) jumping on right leg, 2) jumping on both legs, 3) climbing stairs, 4) jumping jacks. The reconstruction results of our proposed method are comparatively better than the OLNG method in several motions such as jumping on right leg, climbing stairs etc.*

3.5.2.6 Qualitative Comparison

In order to qualitatively evaluate the motions reconstructed using the proposed method, we compare them with the original motions, which are considered as ground truth. In this context we choose, several different types of motions, both simple and complex, in order to show the strength of our algorithm. Figure 3.14 shows the reconstruction results of a special motion — cartwheel motion — which could not be well reconstructed by the exist-

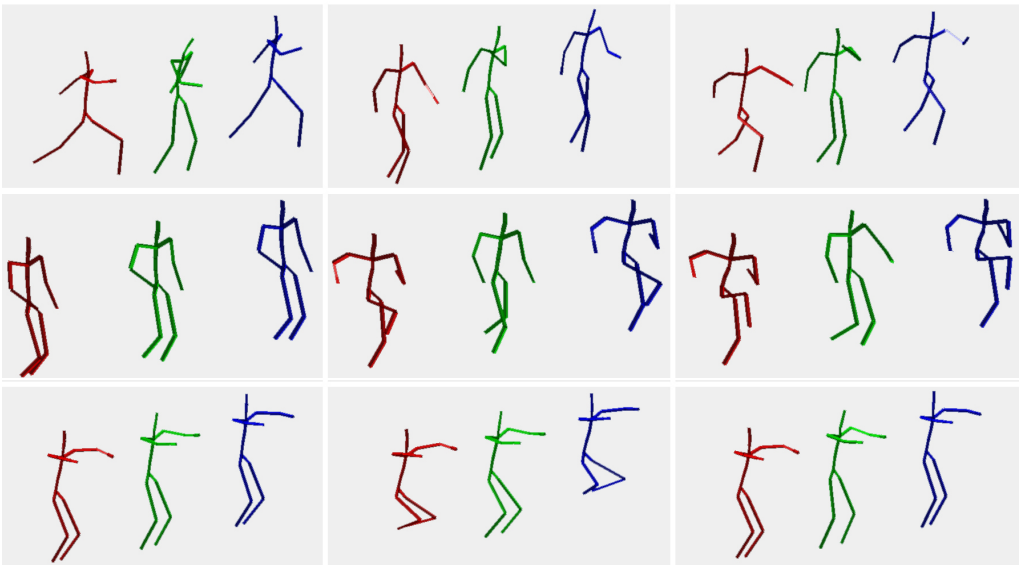


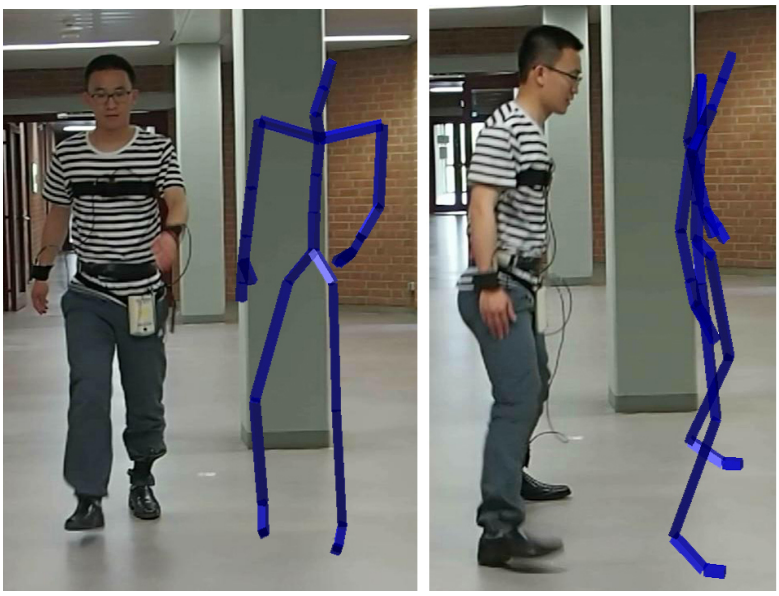
Figure 3.16: *Further examples of reconstructed motions showing 3 frames for each motion. The red skeleton represents the ground truth, the green skeleton represents the reconstructed motion using the OLNG method, and the blue skeleton represents the reconstructed motion using the proposed method. Different motions include (from top to bottom): 1) skiing exercise, 2) elbow-to-knee exercise, and 3) squats. The reconstruction results of our proposed method are comparatively better than the OLNG method in several motions such as skiing, elbow-to-knee exercise etc.*

ing methods because it involves drastic acceleration changes and noise. The red skeleton represents the ground truth, the green skeleton represents the reconstructed motion using the OLNG method, and the blue skeleton represents the reconstructed motion using the proposed method. Eight selected frames of the motion are shown. Visually appealing reconstruction results can be seen especially in 2nd and 3rd rows where the OLNG method fails. It is observable that our proposed method can successfully reconstruct complex motions such as cartwheel as well. Figure 3.15 shows the snapshots of several reconstructed motions including (from top to bottom): 1) jumping on right leg, 2) jumping on both legs, 3) climbing stairs, 4) jumping jacks. Figure 3.16 shows further snapshots of several reconstructed motions including (from top to bottom): 1) skiing exercise, 2) elbow-to-knee exercise, and 3) squats. The red skeleton represents the ground truth, the green skeleton represents the reconstructed motion using the OLNG method, and the blue skeleton represents the reconstructed motion using the proposed method. Three different frames of each motion are shown. The reconstruction results of our proposed method are comparatively

better than the OLNG method in several motions such as jumping on right leg, climbing stairs, skiing, elbow-to-knee exercise etc. All motions presented in Figures 3.14, 3.15, and 3.16 are reconstructed using known ground contacts. Our proposed ground contacts estimation method cannot reliably estimate ground contacts for such motions.

3.5.3 Motion Reconstruction From Real Accelerations and Ground Contacts

In this subsection we will present results of our testing with real wrists and lower trunk accelerations, and ground contact data. We have performed several experiments in an indoor environment with three actors. The ground contacts are estimated from the lower trunk sensor using our proposed algorithm explained in section 3.3.



(a) ***(b)***

Figure 3.17: *Examples of motion reconstruction from real wrist accelerations and ground contact data. a) Normal walk, b) Sideways walk.*

A number of different locomotive activities such as variations of walking, waving, greeting, running etc were perform by different actors (see Table 3.1). Since all experiments are performed in a non-lab environment without any motion capture system, so a quantitative comparison is not possible. Therefore, we only provide qualitative comparison of reconstruction results as shown in Figure 3.17 and Figure 3.18 and in the accom-

(a) *(b)*

Figure 3.18: *Further examples of motion reconstruction from real wrist accelerations and ground contact data. a) Walk and wave right hand, b) Run.*

panying video. For each motion, an image of a selected frame and reconstructed motion of that frame are shown. Here, we provide examples of 4 different variants of locomotion with different accelerations: Figure 3.17a) normal walk, Figure 3.17b) sideways walk, Figure 3.18a) walk and wave with right hand, and Figure 3.18b) running. From these examples, one can clearly see that it is possible to reconstruct visually appealing motions with our proposed method.

3.5.3.1 Runtime

Our proposed method is computationally less costly as compared to the OLNG method presented by Tautges et al. [TZK+11]. This is due to the fact that our proposed algorithm effectively removes many nearest neighbors based on ground contact information. Also, our proposed algorithm requires only 64 best candidate poses, which reduces processing time significantly. Table 3.3 shows runtime comparison between the two methods. In the pre-processing step, the proposed method is used to estimate ground contacts for all motions presented in Table 3.1. The method took 3573ms to process 6660 frames ($\approx$ @1864 frames per second). Substantial difference in energy minimization step can be seen in our proposed method where for each frame, our proposed method is *4.75* times faster than the OLNG method. Overall, our proposed algorithm is *2.79* times faster than the OLNG method. Following configuration was used for our proposed method: DB_1

Table 3.3: *A comparison of run times for different components of reconstruction pipeline between OLNG Method and our proposed method. Substantial difference in energy minimization step can be seen in our proposed method.*

		OLNG Method	Proposed Method
Pre-processing	Ground contacts estimation (trunk sensor)	–	3 573 ms
	kd-tree construction	1 390 ms	1 390 ms
Motion reconstruction	NN search (using kd-tree)	51 ms	51 ms
	Update phase	12 ms	28 ms
	Energy minimization	380 ms	80 ms

with $N \approx 6 \cdot 10^5$, $K = 4096$, $I = 64$. The algorithm was executed on a Core i7 @ 3.07 GHz system on a single thread of Matlab.

4

Estimation of Gender, Height, and Age from Recordings of One Step by a Single Inertial Sensor

A number of previous works have shown that information about a subject is encoded in its sparse kinematic data such as the one revealed by so called point light walkers. In this chapter, we are investigating whether it is possible to extend the concept of point light walkers from recordings of one step by an inertial sensor attached at a single body location. For this purpose, we recorded accelerations and angular velocities of 26 subjects using integrated measurement units attached at four locations (chest, lower back, right wrist and left ankle) when performing standardized gait tasks. The collected data where segmented into individual walking steps. We trained random forest classifiers in order to estimate soft biometrics (gender, height, and age). We applied two different validation methods to the process, 10-fold cross validation and subject-wise cross validation. For all three classification tasks we achieve high accuracy values for all four sensor locations. From these results we can conclude that the data of a single walking step (6D: accelera-

tions and angular velocities) allow for a robust estimation of gender, height, and age of a person.

4.1 Background

Sparse representation of human motions has been investigated for some decades now. It is well-known that representation of human motion by point light displays and similar concepts (e. g. point light walker [Joh73, Tro02]) contain detailed information on several aspects of motions and their initiators.

Over the years, the possibilities to identify certain parameters characterizing given motions have been explored. On the one hand, it is possible to discover information about the displayed motions as such. In the field of action recognition, it has been shown that estimation of poses and skeletons from video and motion capture data allows for recognition and analysis of human movement (Lv et al. [LN07], Junejo et al. [JDLP11], Barnachon et al. [BBBG14], Oshin et al. [OGB14]). The survey of vision-based human motion capture by Moeslund et al. [MHK06] discusses the advances and application of motion-capture related techniques for tracking, pose estimation and recognition of movement. Recognition of motion patterns from video data can be achieved by machine learning approaches exploiting local space-time features (e. g. for SVM-based methods, Schüldt et al. [SLC04]). On the other hand, information on kinematics properties of living beings or animated objects can be detected by analyzing representations of motions. This can be done using motion capture data from passive or active devices as well as contact forces measurements (Venture et al. [VAN08], Kirk et al. [KOF05]).

More recently, the market for wearable devices has virtually exploded (Liew et al. [LWS+15], Son et al. [SLQ+14]). The sheer number of devices [TLZF12] reflects that there are numerous methods to capture and analyze human motion in a relatively new field of application associated with ubiquitous computing. Even though information acquired by such devices may be less accurate than information acquired by modern motion capture systems (Le Masurier et al. [LMTL03], Foster et al. [FLFM+05]), it has been shown that reconstruction of motion from extremely sparse sensor setups is possible in practice (Tautges et al. [TZK+11], Riaz et al. [RTKW15]). This indicates that data collected using tri-axial accelerometers are suitable for classification tasks, e. g. associated with social actions (Hung et al. [HEK13]), general everyday activities (Parkka et al. [PEK+06], Jean-Baptiste et al. [JNP+14], Dijkstra et al. [DKZ10]) or repetitive physical exercises (Morris et al. [MSGK14]).

We investigated if data from a single wearable sensor can reveal similar information about the moving subject as motion capture data in the sense of the above-quoted ([Joh73, Tro02]). We focus on classification of gender, age and height defining exemplary inertial properties of moving subjects. Our experiments show that it is indeed possible to classify and thereby estimate such properties. Our method is able to process representations of single steps recorded by one accelerometer (as opposed to longer data sequences, Neugebauer et al. [NHB12]). In sum, our method is able to recover soft biometrics information with high accuracy consistently over various sensor positions. Since the classification task depends on the chosen feature sets, we further investigated this by evaluating the role of different possible feature sets in the classification.

Modern machine learning techniques like decision trees can target pattern recognition and prediction tasks based on many different representations of motion (Brand et al. [BOP97], Bao et al. [BI04], Kwapisz et al. [KWM11]). We used random forests, a learning method based on the construction of multiple decision trees which can be used for classification as well as regression tasks. While learning predictive models by using decision trees on their own may result in over-fitting to a training set (Phan et al. [Pha14a]), random forests are less prone to this problem. For an overview of random forests, refer to the works of Breimann [Bre01] or Liaw and Wiener [LW02].

Table 4.1: *Characteristics of the study population including age, sex, and height. For validation, two types of models were used: k-fold cross validation and subject-wise cross validation.*

Variable	Characteristics
Total Population	26
Age (y, mean, ±SD)	48.1±12.7
Female Participants	14
Male Participants	12
Height (cm, ±SD)	174±10.2

4.2 Materials and Methods

4.2.1 Participants' Consents

All participants were informed in detail about the purpose of the study, the nature of the experiments, types of data to be recorded, and data privacy policy. The subjects were aware that they were taking part in experiments where a number of biometric and kinematics properties were monitored. The main focus of the study was communicated to the subjects during their progress over the course of the training by the specialists of Gokhale Method Institute [Met15]. Each willing participant was asked to fill in the data collection form with personal details including full name, sex, age, and height.

4.2.2 Population Characteristics and Sampling

The participants were selected during a gait and posture training program conducted in July of 2014 by the specialists of Gokhale Method Institute. They use special gait and posture training methods to help regain the structural integrity of the body. The training program consisted of six 90-minute training sessions. The study population consisted of a total of 26 adults with a male to female ratio of 12:14 and an average age of 48.1 years ($\sigma = \pm 12.7$). The average height of the participants was recorded at 174cm ($\sigma = \pm 10.2$). The characteristics of the study population are shown in Table 4.1.

A *k-fold* cross validation model (chosen value of k=10) was used to compute classification accuracy of the classifier. In *k-fold* cross validation, original sample data are randomly partitioned into k equally sized sub-samples or folds. Out of the k folds, *k-1* folds are used for training and the *left-out* fold is used for validation. The cross validation process is repeated k times and each of the k folds is used exactly once for validation. For sampling, the stratified sampling method [SS03] is used to divide the population into training and test data sets.

A *subject-wise* cross validation model was also employed to compute classification accuracy of each participant against others. *Subject-wise* cross validation is a special variant of *leave-one-out* cross validation in which instead of leaving one sample out for validation, all samples of one participant are left out for validation. For n participants (n=26, in our case), all samples of *n-1* participants are used for training and all samples of the *left-out* participant are used for testing. The cross validation process is repeated n times in order to validate each participant exactly once against the rest. Unlike *10-fold*

cross validation, the number of samples in each fold are not equal in *subject-wise* cross validation. This is due to the difference in the step length of each subject. Subjects with shorter step length have more steps that the others.

Table 4.2: *Standardized gait tasks. Experiments were performed on different surfaces with and without shoes as shown here. For each participant, 9 different recording sessions were carried out in total.*

		4x10-meter Straight Walk		
		Hard Surface		Soft Surface
		Shoes On	Barefoot	Barefoot
Recordings	Before 1st Session	✓	✓	✓
	After 3rd Session	✓	✓	✓
	After 6th Session	✓	✓	✓

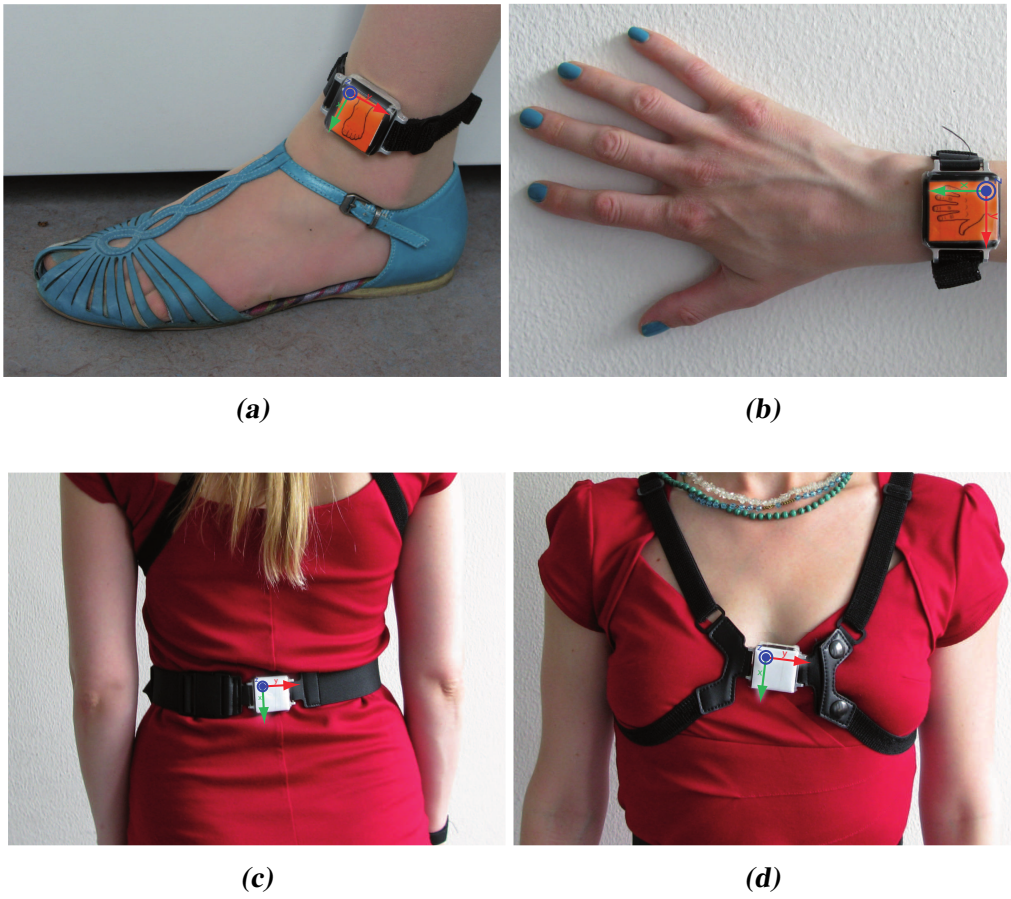
Table 4.3: *Technical specifications of the APDM Opal IMU.*

	Accelerometer	Gyroscope	Magnetometer
Axes	3 axes	3 axes	3 axes
Range	$\pm$ 2g or $\pm$ 6g	$\pm$ 2000 deg/s	$\pm$ 6 Gauss
Noise	0.0012 m/s^2/$\sqrt{Hz}$	0.05 deg/s/$\sqrt{Hz}$	0.5 mGauss/$\sqrt{Hz}$
Sample Rate	1280 Hz	1280 Hz	1280 Hz
Output Rate	20 to 128 Hz	20 to 128 Hz	20 to 128 Hz
Bandwidth	50 Hz	50 Hz	50 Hz
Resolution	14 bits	14 bits	14 bits

4.2.3 Standardized Gait Tasks

The gait task consisted of a 10-meter straight walk from a starting point, turn-around, and walk back to the starting point. Participants were asked to walk in their natural manner and to repeat the gait task two times resulting in a *4x10-meter* walk. Three different types of experiments were performed: 1) *walking on hard surface (concrete floor) with shoes on*, 2) *walking on hard surface (concrete floor) with bare feet*, and 3) *walking on a soft surface (exercise mattress) with bare feet*. Data were recorded during three different stages of the

training course: a) *at the start of the training* (before 1st session), b) *in the middle of the training* (after 3rd session), and c) *at the end of the training* (after 6th session). Hence, for each participant, 9 different recording sessions were carried out in total (see Table 4.2).



(a) ***(b)*** ***(c)*** ***(d)***

Figure 4.1: *Placement of four APDM Opal IMUs on different body parts. The sensors were placed on four different locations: a) left ankle, b) right wrist, c) lower back, and d) chest.*

4.2.4 Sensor Placement and Data Collection

A set of four APDM Opal wireless inertial measurement units [Opa15] were used to record accelerations and angular velocities. An APDM Opal IMU consists of a triad of three accelerometers and three gyroscopes. The technical specifications of the sensor are given in Table 4.3. The sensors were tightly attached to different body parts using adjustable elastic straps. We were particularly interested in the inertial measurements of

four different body parts: a) *chest*, b) *lower back*, c) *right wrist*, and d) *left ankle*. The sensor placement at each body part is shown in Figure 4.1.

4.2.5 Pre-processing

The output sampling rate of an APDM Opal IMU sensor is adjustable between 20 and 128 Hz. In our experiments, an output sampling rate of 128 Hz was chosen. Due to the noisy nature of the acceleration measurements, raw data were pre-processed to suppress noise. To this end, we used moving average method with a window size of 9 frames to smooth the raw signal and suppress noise.

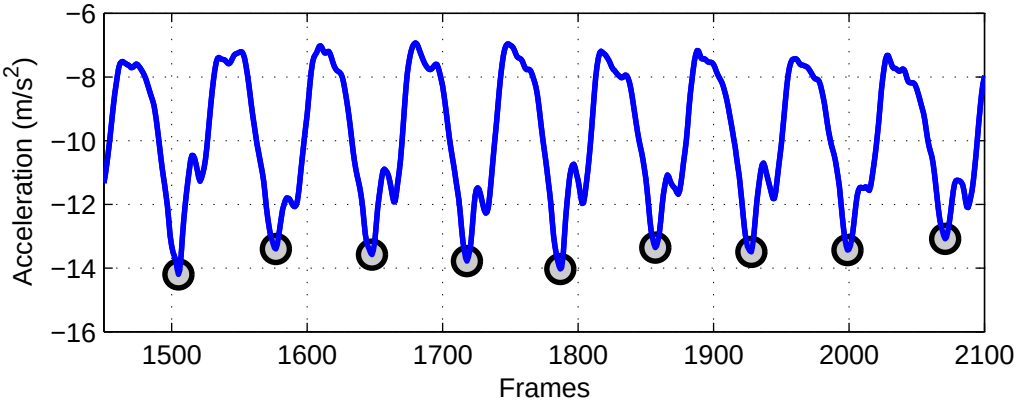


Figure 4.2: *The pre-processed input signal from the x-axis of the accelerometer attached to the lower back and an extracted single step are shown. Detected valleys are highlighted with ◯.*

4.2.6 Signal Decomposition

The input signal consists of a long sequence of steps which is segmented into single steps in order to extract features. A simple approach to decompose a long sequence of steps into single steps is by means of peaks and valleys detection [LZD+12, DNBB10, Zij04]. In this approach, peaks are detected by finding local maxima whereas valleys are detected by finding local minima. Detection of false peaks is minimized by using two threshold: Δ_d and Δ_h. The Δ_d is used to define the minimum distance between two peaks and the Δ_h is used to define the minimum height of the peak. We have used the same approach to detect peaks and valleys from the input signal. The values of the two thresholds are chosen by

experimentation. The valleys are then used to cut the input signal into individual steps. Peaks and valleys are only detected in the x-axis of the acceleration signal and are used to decompose y and z axes of acceleration and all axes of gyroscope. This approach makes sure that the length of the individual step is consistent in all axes of acceleration and gyroscope. Figure 4.2 presents the pre-processed input signal from the x-axis of the IMU's accelerometer attached to the lower back. The detected valleys, highlighted with circles (◯), are also shown.

Table 4.4: *Description of the extracted features for each step from the accelerometer (A) and/or the gyroscope (G). For each step, 56 features from the time and frequency domains are computed.*

Feature Name	Sensor	Axis	Total	Description
Step Length	A	x	1	Total number of frames
Step Duration (s)	A	x	1	Step duration in seconds
Average	A, G	x, y, z	6	Mean value of the step
Standard Deviation	A, G	x, y, z	6	σ of the step
Minimum	A, G	x, y, z	6	Global minimum of the step
Maximum	A, G	x, y, z	6	Global maximum of the step
Root Mean Square	A, G	x, y, z	6	RMS value of the step
Entropy	A, G	x, y, z	6	Uncertainty measure of the step, s_i.: $-\sum_{i=1}^{n}(p_i)log_2(p_i)$ where $p_i = \frac{\frac{s_i}{max(s_i)}}{\sum_{i=1}^{n}\frac{s_i}{max(s_i)}}$
Signal Energy	A, G	x, y, z	6	Energy of the step: $\sum_{n=1}^{N}\|x[n]\|^2$
Signal Energy Ratio	A, G	x, y, z	6	Ratio between the energy of the single step and the energy of the whole signal
Amplitude	A, G	x, y, z	6	Maximum amplitude of the frequency spectrum of the signal of the single step calculated using FFT

4.2.7 Extraction of Features

All single steps detected from the signal decomposition are further processed to extract different features from the time and frequency domains. Table 4.4 presents a complete list of features extracted from different components of accelerations and angular velocities. For each single step, the feature set consists of *56* features in total. Statistical features include: *step length*, *step duration*, *average*, *standard deviation*, *global minimum*, *global maximum*, *root mean square*, and *entropy*. Energy features include *energy of the step*

and *ratio of the step energy to the energy of the whole signal*. The maximum *amplitude* of the frequency spectrum of the signal of single step is calculated using fast Fourier transform (FFT). The *step length* and the *step duration* are only computed for the x-axis of the accelerations as they remain same in all other axes. All of the remaining features are computed for all 3D accelerations and 3D angular velocities. Figure 4.3 presents a decomposed signal depicting a single step between the vertical dash-dot lines (- .). Some of the extracted features are also shown including: 1) Square (□): global minimum, 2) Diamond (◇): global maximum, 3) Solid line (–): mean, 4) Horizontal dash-dot line (- .): standard deviation, 5) Dashed line (- -): root mean square, 6) Between vertical dash-dot lines (- .): length and duration of the step.

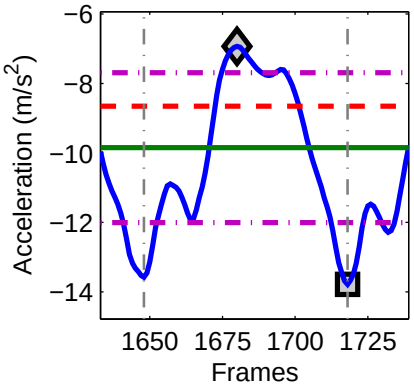


Figure 4.3: *A decomposed signal depicting a single step is shown between the vertical dash-dot lines (- .). Some of the extracted features from the single step are: 1) Square (□): global minimum, 2) Diamond (◇): global maximum, 3) Solid line (–): mean, 4) Horizontal dash-dot line (- .): standard deviation, 5) Dashed line (- -): root mean square, 6) Between vertical dash-dot lines (- .): length and duration.*

4.2.8 Classification of Features

Training and validation data were prepared for each sensor using the features extracted in the previous step. Three types of group classification tasks were performed: a) *gender classification*, b) *height classification*, and c) *age classification*. Furthermore, training and validation data was also prepared for classification within participant subgroups for *height* and *age* classification. In Table 4.5, characteristics of the population within different classification tasks are presented.

Table 4.5: *Characteristics of the population within different group and subgroup classification tasks.*

Task	Classes	N	Age (y, mean, ±SD)
Group Classification Tasks			
Gender Classification	Male	12	43.75±14.50
	Female	14	51.79±11.15
Age Classification	Age < 40	9	34.11±03.62
	$40 <$ Age < 50	6	46.67±02.58
	Age ≥ 50	11	60.67±07.48
Height Classification	Height $\leq$ 170cm	8	55.62±11.29
	170cm $<$ Height $<$ 180cm	10	44.70±11.31
	Height $\geq$ 180cm	8	44.75±13.81
Subgroup Classification Tasks			
Age Classification			
Male Group	Age ≤ 40	6	32.67±02.94
	Age > 40	6	54.83±09.87
Female Group	Age ≤ 50	6	41.83±06.08
	Age > 50	8	59.25±07.48
Height Classification			
Male Group	Height $\leq$ 180cm	7	38.43±11.27
	Height $>$ 180cm	5	51.20±13.85
Female Group	Height $\leq$ 170cm	8	55.62±11.29
	Height $>$ 170cm	6	46.67±09.48

As classifier, Random Forest [LW02] was chosen and trained on the training data set with the following values of parameters: number of trees = 400, maximum number of features for best split = 7, maximum depth = $None$. Two types of validation strategies were employed: *stratified 10-fold* cross validation and *subject-wise* cross validation. The *10-fold* cross validation was employed for all group and subgroup classification tasks whereas the *subject-wise* cross validation was employed to group classification tasks only.

For each sensor in a classification task, the classifier was trained and validated for three different sets of features: 1) *3D accelerations* (29 features), 2) *3D angular velocities* (29 features), and 3) *6D accelerations and angular velocities* (56 features). The *10-fold* cross validation was employed to all three sets of features whereas the *subject-wise* cross

validation was employed to third set of features (56 features) only. Finally, classification rate, specificity, sensitivity, positive predictive value (PPV) for each set of features were calculated as explained in [Umb10]. Same approach was used for all group and subgroup classification tasks. The classification rate c or classification accuracy is given by the formula in equation 4.1

$$c = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{4.1}$$

where TP, TN are the numbers of true positives and true negatives respectively and FP, FN are the numbers of false positives and false negatives respectively.

4.2.9 Statistics

Pearson's Chi-Square test from Matlab's statistics and machine learning toolbox was used to evaluate the goodness of fit for all classification tasks.

4.3 Results

In the following sections, we present the results of our investigations of the recorded gait data. Our classification results prove a number of hypotheses regarding biometric and biographic characteristics of the human subjects. Specifically the gender, the body height and the age of participants could be classified well. Each of classification tasks was solved by training random forest classifiers as introduced in the previous section.

4.3.1 Gender Classification

Our goal was to show that classification tasks regarding the gender of trial subject can be performed sufficiently well by using the proposed sensors attached to each of the given locations.

H_0 **: The gender can be identified by motion recordings of any of the employed sensors**

The results presented in Figure 4.4 show that the statement holds true for each of the four sensors individually. For each sensor there are three different images visualizing the results of the binary classification, namely for the investigation of accelerations, of angular velocities as well as of both combined. The confusion matrices encode the following

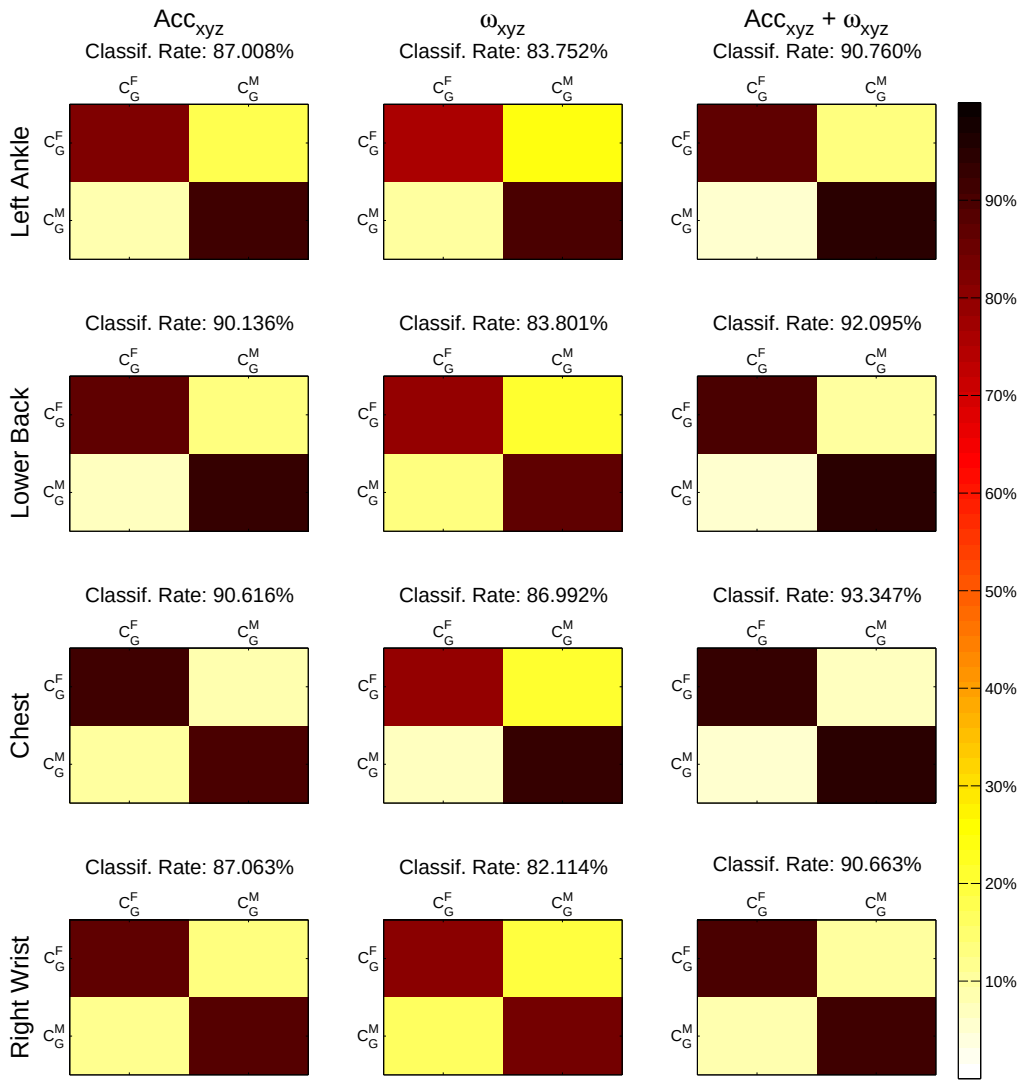


Figure 4.4: *Confusion matrices of gender classification. Each column presents sensor positions (left to right): left ankle, lower back, chest, and right wrist. Each row present feature sets used for classification (top to bottom): 3D accelerations (29 features), 3D angular velocities (29 features), and 6D accelerations and angular velocities (56 features). Classes:* C_G^F = $gender\ female$, C_G^M = $gender\ male$.

information: Each column represents the instances in one of the predicted classes, while each row represents the instances in the actual class (female/male).

For application of acceleration only, the classification rates are higher than 87% for each of the sensors. Classification results based on angular velocities show lower classi-

Table 4.6: *Classification results obtained by using 10-fold cross validation for different classification categories: Gender, Height, and Age. The results show balanced correct classification rates, sensitivity, specificity, and positive predictive value (PPV) for all classification categories (p<0.0001 in all cases; Pearson's Chi-Square test).*

Classification Task	Body Part	Sensor	Class. Rate	Sens.	Spec.	PPV
Gender Classification	Chest	A_{xyz}, G_{xyz}	93.35	92.23	94.22	92.59
	Lower Back	A_{xyz}, G_{xyz}	92.09	89.54	94.10	92.25
	Right Wrist	A_{xyz}, G_{xyz}	90.66	90.25	91.01	89.20
	Left Ankle	A_{xyz}, G_{xyz}	90.59	86.55	93.74	91.49
Body Height Classification	Chest	A_{xyz}, G_{xyz}	91.01	92.79	94.65	91.73
	Lower Back	A_{xyz}, G_{xyz}	90.99	92.91	92.78	89.23
	Right Wrist	A_{xyz}, G_{xyz}	87.15	88.56	91.92	87.12
	Left Ankle	A_{xyz}, G_{xyz}	88.44	93.13	89.85	85.02
Age Classification	Chest	A_{xyz}, G_{xyz}	90.05	94.67	90.85	86.82
	Lower Back	A_{xyz}, G_{xyz}	90.60	93.82	94.52	91.69
	Right Wrist	A_{xyz}, G_{xyz}	85.36	90.58	88.76	83.97
	Left Ankle	A_{xyz}, G_{xyz}	86.78	92.93	87.46	83.63

fication rate but still above 82%. The classification based on the combined features performs better than each of the individual feature sets, namely above 90%. More precisely, the results for the combined features are (listed by sensor in descending order of rates): chest (93.35%), lower back (92.09%), right wrist (90.66%), left ankle (90.59%). Table 4.6 presents 10-fold cross validation results of gender classification including correct classification accuracy, sensitivity, specificity, and positive predictive value (p<0.0001 in all cases; Pearson's Chi-Square test).

4.3.2 Body Height Classification

Another goal was body height classification from only accelerations, angular velocities and a combination of both.

H_1 **: The body height can be identified by motion recordings of any of the employed sensors**

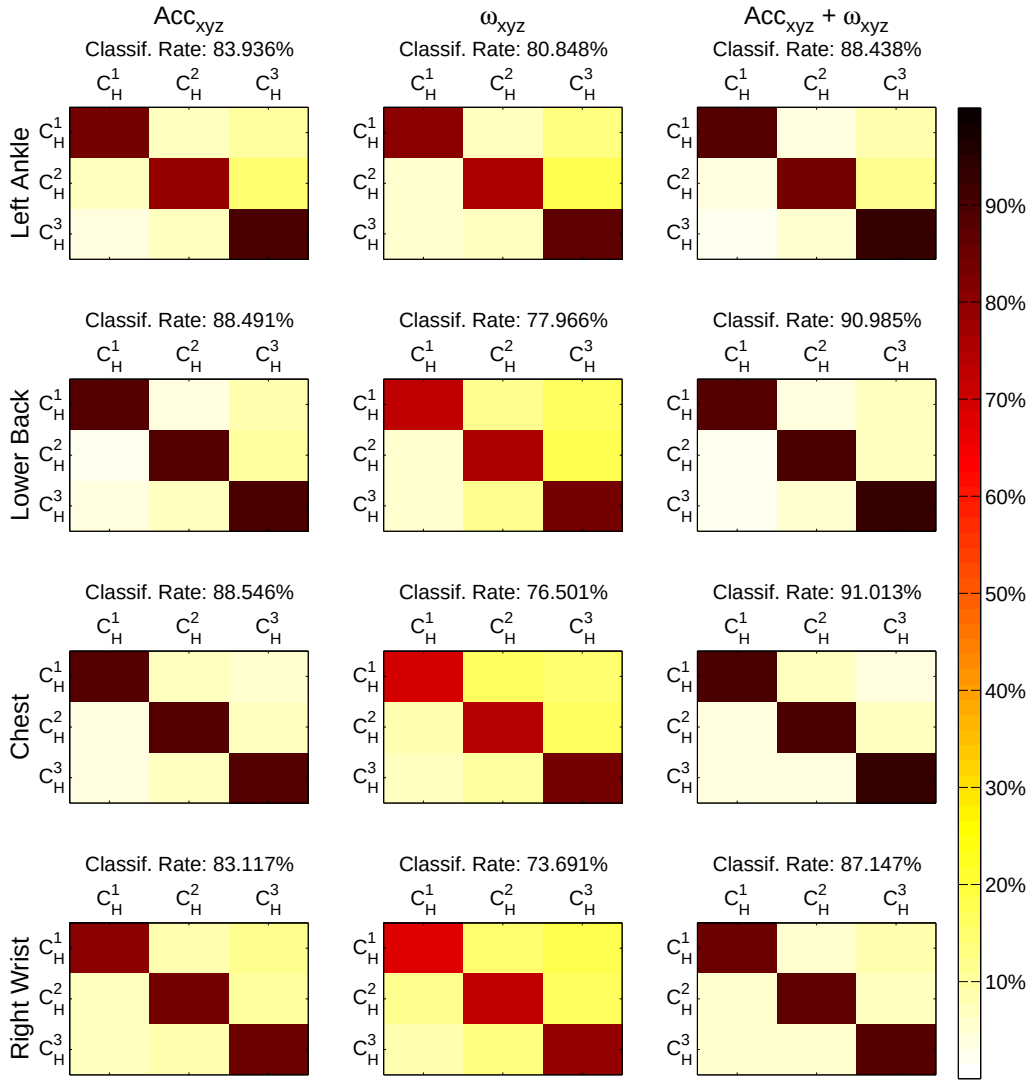


Figure 4.5: *Confusion matrices of body height classification. Each column presents sensor positions (left to right): left ankle, lower back, chest, and right wrist. Each row present feature sets used for classification (top to bottom): 3D accelerations (29 features), 3D angular velocities (29 features), and 6D accelerations and angular velocities (56 features).* $C_H^1 : height \leq 170cm$, $C_H^2 : height = 171 - 179cm$, $C_H^3 : height \geq 180cm$.

The results of the ternary classification for each individual sensor are given in Figure 4.5. Here, the classification estimated the assignment to three classes ($C_H^1 : height \leq 170cm$, $C_H^2 : 170cm < height < 180cm$, $C_H^3 : height \geq 180cm$). The results are divided into three cases: use of accelerations only, use of angular velocities only and both

types of features combined. For application of only acceleration-based features, the classification rates are higher than 83% in each of the cases. Similar to the gender classification task, results based on angular velocities have lower classification rates, between 73% and 80%. The classification based on the combined features performs better than the other two, namely above 87%. More precisely, the results for the combined features are (listed by sensor in descending order of rates): chest (91.01%), lower back (90.99%), left ankle (88.44%), right wrist (87.15%). Table 4.6 presents 10-fold cross validation results of body height classification including correct classification accuracy, sensitivity, specificity, and positive predictive value ($p<0.0001$ in all cases; Pearson's Chi-Square test).

4.3.3 Age Classification

Another goal was age group classification from only accelerations, angular velocities and their combination.

H_2 : The age group of individuals can be identified by motion recordings of any of the employed sensors

The results of the ternary classification for each individual sensor are given in Figure 4.6. Here, the classification estimated the assignment to three classes according to three age groups ($C_A^1 : age < 40$, C_A^2: $40 \leq age < 50$, C_A^3: $age \geq 50$) of participants. Again, the results are divided into three cases: use of accelerations only, use of angular velocities only and both types of features combined. For application of only acceleration-based features, the classification rates range from 81.64% to 88.07% depending on the sensor. As in both previous classification tasks, results based on angular velocities show lower classification rates, between 74% and 80.51%. The classification based on the combined features performs better than in each of the individual feature types, namely above 85%. More precisely, age classification results for the combined features are (listed by sensor in descending order of rates): lower back (90.60%), chest (90.05%), left ankle (86.78%), right wrist (85.36%). Table 4.6 presents 10-fold cross validation results of age classification including correct classification accuracy, sensitivity, specificity, and positive predictive value ($p<0.0001$ in all cases; Pearson's Chi-Square test).

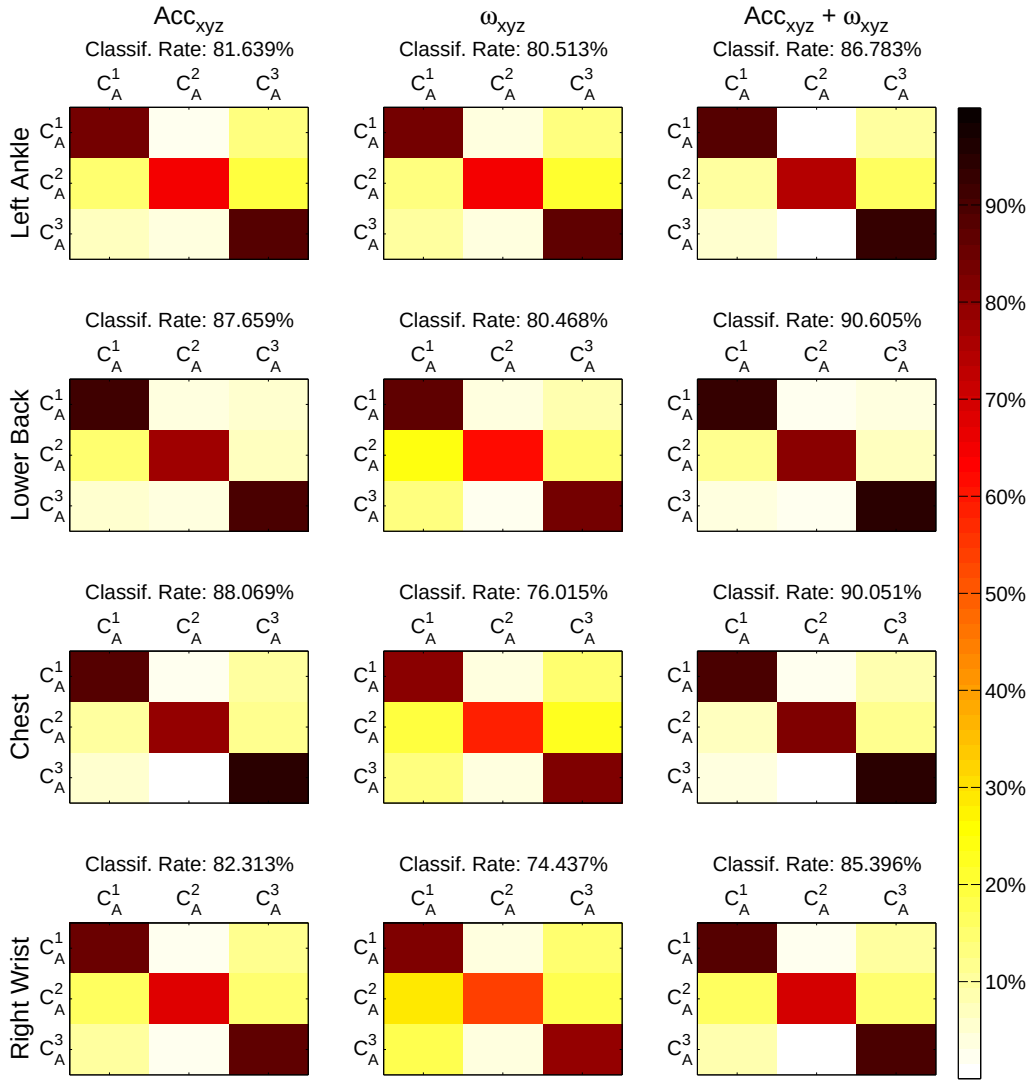


Figure 4.6: *Confusion matrices of age classification. Each column presents sensor positions (left to right): left ankle, lower back, chest, and right wrist. Each row present feature sets used for classification (top to bottom): 3D accelerations (29 features), 3D angular velocities (29 features), and 6D accelerations and angular velocities (56 features).* C_A^1 *:* $age < 40$*,* C_A^2*:* $age < 50$ *,* C_A^3*:* $age \geq 50$*.*

4.3.4 Contribution of Individual Features to Classification Results

The contribution of each of the employed features in all three classification tasks was homogenous in the sense that there is not one outstanding feature with major contribution

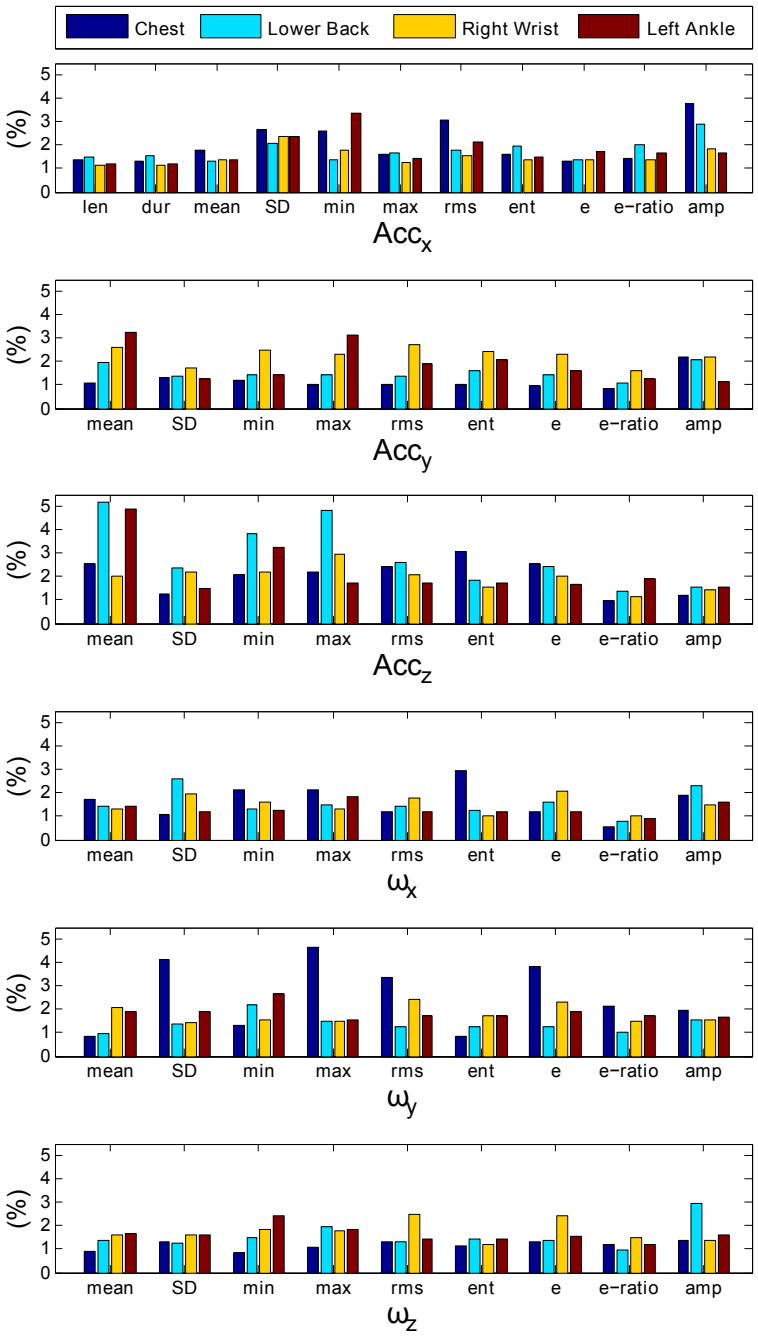


Figure 4.7: *Bar graphs of features importance computed during gender classification. The graphs present a comparison of importance of each feature (in %) with respect to different sensor positions. In general, all features are significantly contributing in the classification task.*

Table 4.7: *Features importance computed during gender classification. Top 5 contributing features are highlighted with bold text. All values are in percentage.*

	Sen.	Len	Dur	Mean	SD	Min	Max	RMS	Ent	E	E Ratio	Amp
Chest	A_x	1.38	1.33	1.80	2.67	2.63	1.59	3.05	1.63	1.29	1.43	**3.77**
	A_y	–	–	1.10	1.33	1.19	1.01	1.00	1.02	0.97	0.84	2.17
	A_z	–	–	2.53	1.28	2.10	2.17	2.45	3.08	2.54	0.95	1.19
	G_x	–	–	1.75	1.07	2.14	2.11	1.21	2.97	1.17	0.54	1.90
	G_y	–	–	0.87	**4.10**	1.31	**4.63**	**3.34**	0.86	**3.85**	2.16	1.99
	G_z	–	–	0.93	1.32	0.88	1.09	1.33	1.13	1.29	1.19	1.38
Lower Back	A_x	1.48	1.53	1.34	2.10	1.36	1.64	1.80	1.95	1.36	2.01	**2.91**
	A_y	–	–	1.94	1.40	1.45	1.43	1.38	1.61	1.45	1.09	2.08
	A_z	–	–	**5.15**	2.34	**3.85**	**4.83**	2.62	1.82	2.45	1.39	1.54
	G_x	–	–	1.41	2.59	1.31	1.50	1.45	1.27	1.60	0.79	2.33
	G_y	–	–	0.96	1.40	2.17	1.49	1.24	1.24	1.27	1.01	1.55
	G_z	–	–	1.38	1.26	1.48	1.98	1.29	1.41	1.40	0.97	**2.93**
Right Wrist	A_x	1.13	1.14	1.39	2.39	1.79	1.27	1.56	1.37	1.40	1.37	1.85
	A_y	–	–	**2.62**	1.72	**2.48**	2.30	**2.70**	2.41	2.32	1.61	2.17
	A_z	–	–	2.00	2.21	2.18	**2.96**	2.10	1.54	2.05	1.15	1.40
	G_x	–	–	1.29	1.93	1.62	1.30	1.79	1.04	2.08	1.02	1.52
	G_y	–	–	2.10	1.41	1.57	1.48	2.42	1.70	2.32	1.50	1.55
	G_z	–	–	1.64	1.61	1.83	1.79	**2.47**	1.19	2.43	1.47	1.38
Left Ankle	A_x	1.22	1.19	1.36	2.37	**3.35**	1.45	2.11	1.48	1.71	1.68	1.70
	A_y	–	–	**3.24**	1.27	1.42	**3.13**	1.89	2.08	1.59	1.29	1.13
	A_z	–	–	**4.87**	1.50	**3.26**	1.71	1.71	1.74	1.69	1.91	1.56
	G_x	–	–	1.41	1.23	1.27	1.84	1.19	1.22	1.17	0.89	1.58
	G_y	–	–	1.90	1.91	2.68	1.57	1.71	1.74	1.89	1.72	1.67
	G_z	–	–	1.66	1.61	2.42	1.84	1.43	1.45	1.57	1.22	1.61

to the classification results. In all experiments we made the following observation: In sum, accelerations contributed more to overall results than angular velocities. However, the combination of the two feature types did better than accelerations or angular velocities

individually.

In detail, the classification results related to sensors at different locations can depend on quite different feature sets. In the following, we will give an overview of the most important contributors for each of the locations.

4.3.4.1 Gender Classification

For the sensor placed at the chest, angular velocities (around the y-axis, i.e. transverse axis) contributed most, especially standard deviation, max, energy, and RMS. These are related to rotation of the upper body around a horizontal axis over the course of the motion. Note that this is not a contradiction to our other claims. In sum, the accelerations contributed more than angular velocities, but some features stand out. Furthermore, the amplitude of the accelerations along the x-axis, i.e. the cranio-caudal axis is of high importance. For the lower back, the most important features are associated with acceleration of the z-axis. This corresponds to changes in the velocity of the hip movement within the sagittal plane, i.e. front-to-back. In addition, angular velocities associated with the z-axis, i.e. rotation around the anteroposterior axis (swinging of hips), contribute significantly to the results. Furthermore, the amplitude of the accelerations along the x-axis, i.e. the cranio-caudal axis is also of high importance. For the right wrist, features associated with acceleration along the y- and z-axis are top contributors. Particularly, mean, minimum and maximum acceleration values associated with dorso-ventral as well as lateral movement of the hand play a more important part in the classification. Also, the RMS of angular velocities associated to the z-axis is important. This is also linked to the swinging of the hand in lateral direction.

For the left ankle, the contribution of accelerations along each axis is generally higher compared to the contribution of other single features. Figure 4.7 shows bar graphs of features importance computed during gender classification. The graphs present a comparison of importance of each feature (in percentage) with respect to different sensor positions. In general, all features are significantly contributing in the classification task. An overview of contribution percentages where the most important features are highlighted is given in Table 4.7.

4.3.4.2 Body Height Classification

For the sensor placed at the chest, accelerations along the z-axis contributed most, especially mean, max and RMS. These are associated with motion of the upper body in

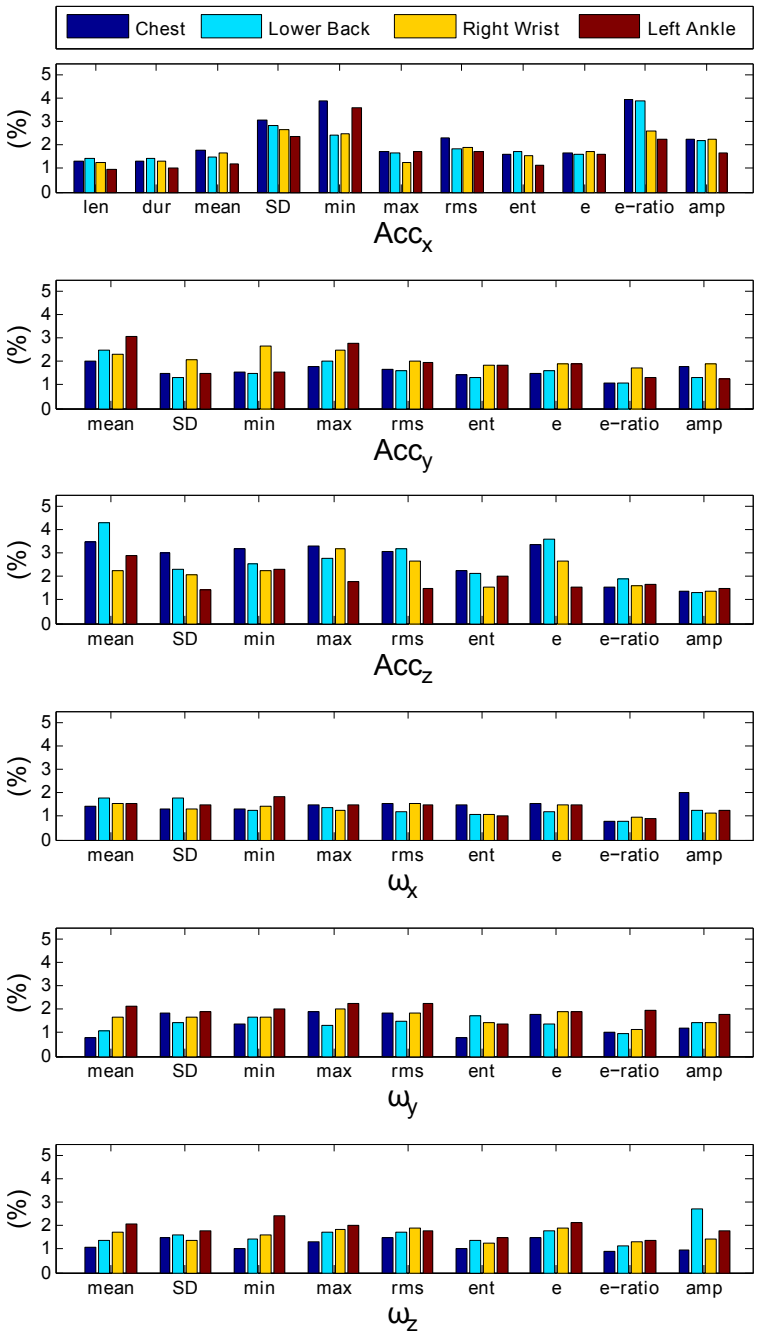


Figure 4.8: *Bar graphs of features importance computed during body height classification. The graphs present a comparison of importance of each feature (in %) with respect to different sensor positions. In general, all features are significantly contributing in the classification task.*

Table 4.8: *Features importance computed during body height classification. Top 5 contributing features are highlighted with bold text. All values are in percentage.*

	Sen.	Len	Dur	Mean	SD	Min	Max	RMS	Ent	E	E Ratio	Amp
Chest	A_x	1.32	1.34	1.79	3.09	**3.91**	1.74	2.33	1.60	1.69	**3.93**	2.22
	A_y	–	–	1.99	1.49	1.54	1.77	1.65	1.43	1.48	1.08	1.80
	A_z	–	–	**3.49**	3.00	3.18	**3.30**	3.05	2.25	**3.35**	1.56	1.37
	G_x	–	–	1.41	1.32	1.31	1.50	1.52	1.50	1.57	0.76	2.03
	G_y	–	–	0.82	1.82	1.35	1.90	1.86	0.79	1.77	1.05	1.20
	G_z	–	–	1.09	1.51	1.02	1.31	1.49	1.00	1.51	0.90	0.95
Lower Back	A_x	1.42	1.43	1.47	**2.82**	2.44	1.65	1.83	1.73	1.61	**3.89**	2.22
	A_y	–	–	2.48	1.31	1.47	2.02	1.58	1.34	1.61	1.08	1.30
	A_z	–	–	**4.29**	2.33	2.56	2.79	**3.20**	2.12	**3.59**	1.88	1.29
	G_x	–	–	1.79	1.79	1.27	1.35	1.22	1.11	1.21	0.81	1.27
	G_y	–	–	1.08	1.46	1.68	1.31	1.50	1.70	1.38	0.97	1.45
	G_z	–	–	1.38	1.63	1.42	1.72	1.70	1.37	1.77	1.12	2.73
Right Wrist	A_x	1.26	1.33	1.68	**2.66**	2.50	1.28	1.91	1.55	1.72	2.60	2.24
	A_y	–	–	2.32	2.06	**2.65**	2.50	1.99	1.83	1.91	1.72	1.88
	A_z	–	–	2.27	2.07	2.24	**3.18**	**2.65**	1.53	**2.65**	1.62	1.36
	G_x	–	–	1.53	1.31	1.41	1.25	1.52	1.10	1.50	0.97	1.17
	G_y	–	–	1.67	1.65	1.68	1.99	1.83	1.42	1.90	1.13	1.44
	G_z	–	–	1.70	1.39	1.62	1.82	1.89	1.28	1.92	1.33	1.41
Left Ankle	A_x	0.99	1.00	1.21	2.36	**3.59**	1.72	1.72	1.14	1.58	2.24	1.68
	A_y	–	–	**3.05**	1.47	1.53	**2.79**	1.98	1.82	1.88	1.34	1.24
	A_z	–	–	**2.92**	1.42	2.34	1.78	1.48	1.99	1.58	1.69	1.50
	G_x	–	–	1.54	1.50	1.86	1.48	1.48	1.03	1.49	0.93	1.27
	G_y	–	–	2.11	1.91	2.00	2.25	2.25	1.39	1.91	1.93	1.79
	G_z	–	–	2.06	1.78	**2.44**	2.00	1.79	1.47	2.14	1.39	1.78

dorso-ventral direction. Furthermore, the minima and signal energy ratio of accelerations along the x-axis, i.e. the cranio-caudal axis are of importance.

For the lower back, the most important features are associated with acceleration of the

z-axis. This corresponds to changes in the velocity of the movement of the hips within the sagittal plane, i.e. front-to-back. In addition, features describing accelerations in the x-axis contribute significantly to the results. These are linked to movement of the hips along the cranio-caudal axis (up and down). For the right wrist, features associated with acceleration along each of the three axes contribute significantly. Particularly, maximum, RMS and energy values associated with dorso-ventral movement of the hand play a more important part. For the left ankle, also the contribution of accelerations along each axis is generally high. Additionally, angular velocities associated with rotation of the feet from side to side (around the z-axis) are significant contributors. Figure 4.8 shows bar graphs of feature contribution computed during body height classification. The graphs present a comparison of importance of each feature (in percentage) with respect to different sensor positions. In general, all features are significantly contributing in the classification task. An overview of contribution percentages where the most important features are highlighted is given in Table 4.8.

4.3.4.3 Age Classification

For the sensor placed at the chest, the importance of the features is similarly distributed as in the height classification results: accelerations along the z-axis contributed most, especially mean, maximum, RMS, and energy. These are associated with motion of the upper body in dorso-ventral direction. Furthermore, the minimum acceleration along the x-axis, i.e. the cranio-caudal axis is important. For the lower back, the most important features are associated especially with acceleration of the z-axis. This is similar to the results found in the height classification scenario and corresponds to changes in the velocity of the movement of the hips within the sagittal plane, i.e. front-to-back. For the right wrist, features associated with acceleration along each of the three axes contribute significantly. Additionally, the minimum angular velocity associated with rotation around the z-axis, i.e. swinging laterally, is important. For the left ankle, the contribution of features associated with lateral acceleration is high. Additionally, angular velocities associated with swinging of the feet from side to side (around the z-axis) as well as rolling over from heel to toes (rotation around the y-axis) are significant contributors. Figure 4.9 shows bar graphs of features importance computed during age classification. The graphs present a comparison of importance of each feature (in percentage) with respect to different sensor positions. In general, all features are significantly contributing in the classification task. An overview of contribution percentages where the most important features are highlighted is given in

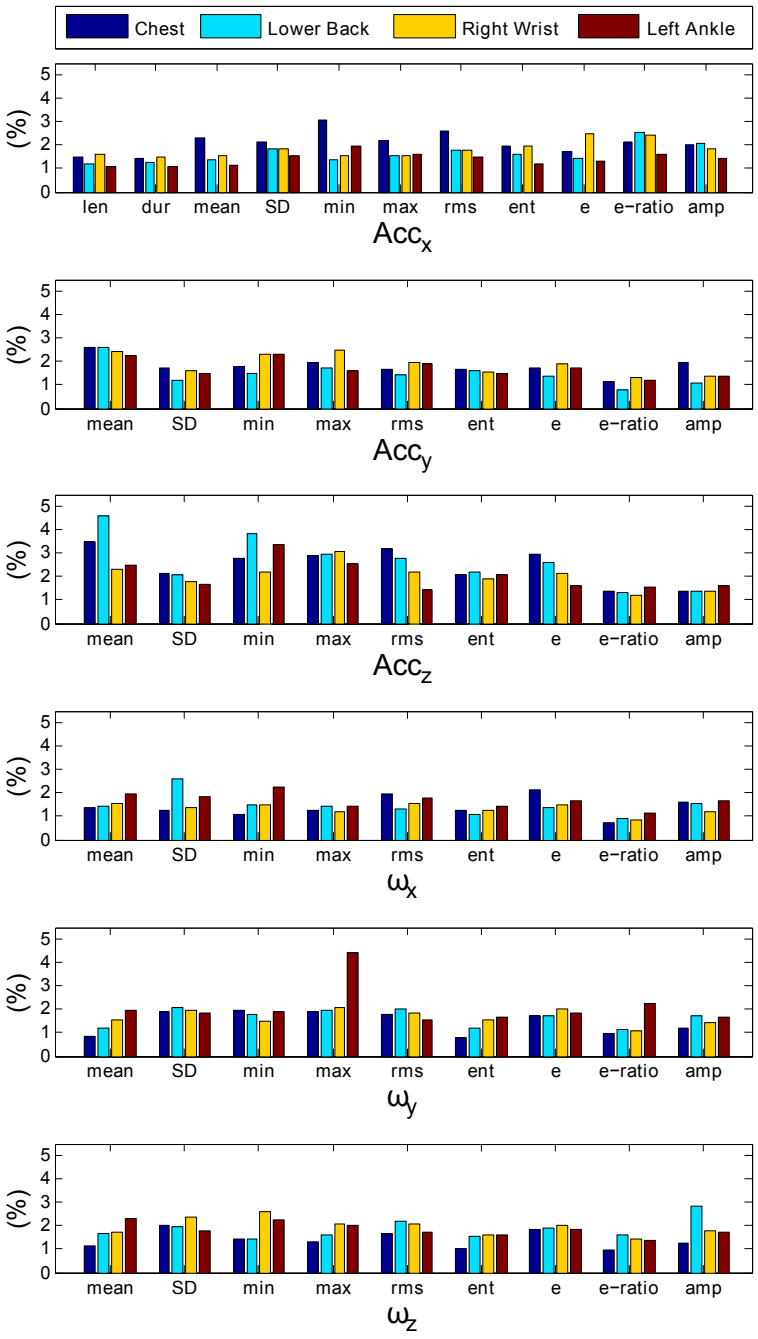


Figure 4.9: *Bar graphs of features importance computed during age classification. The graphs present a comparison of importance of each feature (in %) with respect to different sensor positions. In general, all features are significantly contributing in the classification task.*

Table 4.9: *Features importance computed during age classification. Top 5 contributing features are highlighted with bold text. All values are in percentage.*

	Sen.	Len	Dur	Mean	SD	Min	Max	RMS	Ent	E	E Ratio	Amp
Chest	A_x	1.49	1.43	2.30	2.16	**3.10**	2.17	2.61	1.93	1.70	2.15	2.04
	A_y	–	–	2.62	1.74	1.80	1.96	1.69	1.67	1.74	1.13	1.97
	A_z	–	–	**3.48**	2.11	2.77	**2.88**	**3.18**	2.08	**2.98**	1.38	1.38
	G_x	–	–	1.38	1.26	1.11	1.25	1.95	1.26	2.12	0.72	1.61
	G_y	–	–	0.85	1.87	1.97	1.89	1.79	0.78	1.71	0.98	1.21
	G_z	–	–	1.15	2.02	1.42	1.31	1.64	1.00	1.86	0.98	1.28
Lower Back	A_x	1.19	1.27	1.39	1.85	1.36	1.56	1.79	1.59	1.41	2.55	2.09
	A_y	–	–	2.61	1.20	1.48	1.70	1.45	1.62	1.39	0.82	1.08
	A_z	–	–	**4.60**	2.10	**3.82**	**2.98**	**2.75**	2.21	**2.61**	1.31	1.37
	G_x	–	–	1.42	2.61	1.52	1.42	1.32	1.10	1.39	0.88	1.57
	G_y	–	–	1.18	2.08	1.81	1.93	2.04	1.22	1.75	1.14	1.74
	G_z	–	–	1.65	1.94	1.41	1.59	2.21	1.55	1.91	1.63	2.84
Right Wrist	A_x	1.61	1.55	1.60	1.88	1.54	1.53	1.78	1.96	**2.49**	2.45	1.88
	A_y	–	–	**2.52**	1.60	2.30	**2.46**	2.01	1.64	1.83	1.35	1.42
	A_z	–	–	2.24	1.72	2.18	**3.13**	2.17	1.90	2.09	1.21	1.35
	G_x	–	–	1.56	1.38	1.48	1.16	1.53	1.20	1.49	0.90	1.18
	G_y	–	–	1.61	1.94	1.51	2.00	1.83	1.49	1.96	1.09	1.48
	G_z	–	–	1.78	2.41	**2.56**	2.06	2.08	1.62	2.09	1.50	1.76
Left Ankle	A_x	1.05	1.07	1.13	1.56	1.96	1.59	1.52	1.22	1.34	1.61	1.42
	A_y	–	–	2.25	1.49	2.29	1.60	1.92	1.51	1.70	1.21	1.39
	A_z	–	–	**2.46**	1.64	**3.35**	**2.54**	1.44	2.05	1.61	1.57	1.60
	G_x	–	–	1.94	1.83	2.23	1.44	1.76	1.43	1.67	1.11	1.65
	G_y	–	–	1.95	1.86	1.91	**4.44**	1.56	1.68	1.86	2.23	1.69
	G_z	–	–	**2.30**	1.80	2.26	2.02	1.75	1.58	1.83	1.36	1.75

Table 4.9.

Table 4.10: *Results of body height and age classifications within participant subgroups using 10-fold cross validation. The results show balanced correct classification rates, sensitivity, specificity, and positive predictive value (PPV) for all classification categories ($p<0.0001$ in all cases; Pearson's Chi-Square test).*

Classification Task	Body Part	Sensor	Class. Rate	Sens.	Spec.	PPV
Body Height Classification						
Male Group	Chest	A_{xyz}, G_{xyz}	95.66	97.33	93.35	95.32
	Lower Back	A_{xyz}, G_{xyz}	94.61	96.13	92.52	94.62
	Right Wrist	A_{xyz}, G_{xyz}	93.91	97.16	89.51	92.62
	Left Ankle	A_{xyz}, G_{xyz}	93.25	94.88	91.20	93.16
Female Group	Chest	A_{xyz}, G_{xyz}	92.56	94.20	90.49	92.61
	Lower Back	A_{xyz}, G_{xyz}	94.47	96.70	91.65	93.60
	Right Wrist	A_{xyz}, G_{xyz}	90.44	93.59	86.58	89.52
	Left Ankle	A_{xyz}, G_{xyz}	93.04	94.64	90.94	93.18
Age Classification						
Male Group	Chest	A_{xyz}, G_{xyz}	94.08	93.85	94.31	94.58
	Lower Back	A_{xyz}, G_{xyz}	94.20	93.99	94.43	94.64
	Right Wrist	A_{xyz}, G_{xyz}	94.50	95.23	93.71	94.17
	Left Ankle	A_{xyz}, G_{xyz}	92.90	92.54	93.26	93.16
Female Group	Chest	A_{xyz}, G_{xyz}	94.12	92.11	95.95	95.41
	Lower Back	A_{xyz}, G_{xyz}	95.94	96.53	95.41	95.01
	Right Wrist	A_{xyz}, G_{xyz}	89.72	90.04	89.42	88.81
	Left Ankle	A_{xyz}, G_{xyz}	92.61	89.41	95.36	94.29

4.3.5 Classification Results Based on Restriction to Subgroups

Since the correlation between body height and gender is very high (on average, men are taller than women), we performed a gait-based classification task on each of the groups of female and male participants in order to present height classification results that are independent of this particular phenomenon. Moreover, we also performed age classification on data of each subgroup (female vs male) separately. The number of subjects present in the study did not allow for ternary classification of subgroups (see Table 4.5 for the pop-

ulation characteristics). Therefore there were two different classes in the height-related experiment: C_H^1 = body height of subject is less than or equal to t_h cm, C_H^2 = body height of subject is greater than t_h cm ($t_h = 180$ for male, $t_h = 170$ for female subjects). In the age-related experiment, assigned classes were: C_A^1 = subject is less than or equal to t_a years old, C_A^2 = subject is greater than t_a years old ($t_a = 40$ for male, $t_a = 50$ for female subjects).

Table 4.10 shows an overview of the results. It is quite clear that the results are very good in all cases with the classification rate higher than 90% in all but one case (89.72%). The results also present balanced sensitivity, specificity, and positive predictive value (PPV) for all cases ($p<0.0001$ in all cases; Pearson's Chi-Square test). The classification results are computed by using 10-fold cross validation model.

Table 4.11: *Subject-wise classification results of different classification categories: Gender, Height, and Age. The results show balanced correct classification rates, sensitivity, specificity, and positive predictive value (PPV) for all classification categories ($p<0.0001$ in all cases; Pearson's Chi-Square test).*

Classification Task	Body Part	Sensor	Class. Rate	Sens.	Spec.	PPV
Gender Classification	Chest	A_{xyz}, G_{xyz}	88.23	88.83	87.60	88.21
	Lower Back	A_{xyz}, G_{xyz}	87.26	87.21	87.29	84.35
	Right Wrist	A_{xyz}, G_{xyz}	80.65	75.92	86.37	87.09
	Left Ankle	A_{xyz}, G_{xyz}	78.06	81.63	74.74	75.10
Body Height Classification	Chest	A_{xyz}, G_{xyz}	84.21	83.15	92.36	83.60
	Lower Back	A_{xyz}, G_{xyz}	84.95	85.14	92.23	84.94
	Right Wrist	A_{xyz}, G_{xyz}	75.32	75.06	87.65	75.11
	Left Ankle	A_{xyz}, G_{xyz}	71.58	71.00	85.65	71.25
Age Classification	Chest	A_{xyz}, G_{xyz}	71.16	71.83	85.86	74.48
	Lower Back	A_{xyz}, G_{xyz}	74.71	74.93	87.02	76.29
	Right Wrist	A_{xyz}, G_{xyz}	64.37	64.14	82.11	64.32
	Left Ankle	A_{xyz}, G_{xyz}	65.78	64.87	82.82	65.16

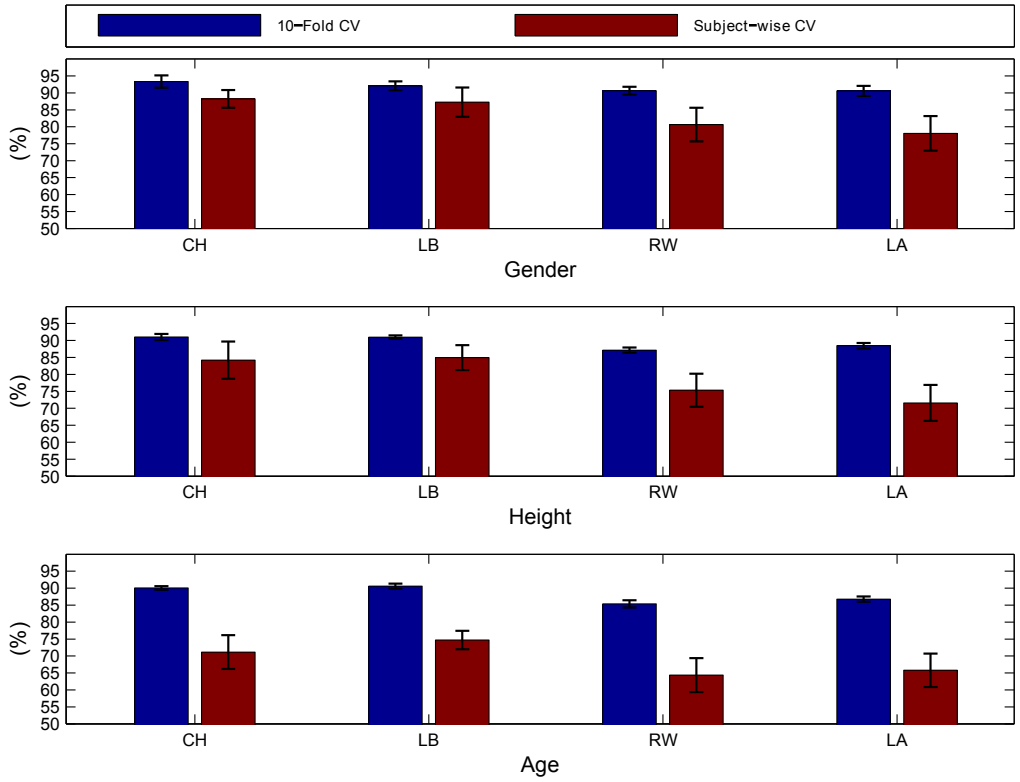


Figure 4.10: *A comparison of results of group classification tasks (gender, height, and age) using 10-fold cross validation and subject-wise cross validation. Sensor positions include: chest (CH), lower back (LB), right wrist (RW), and left ankle (LA). The 10-fold cross validation model outperforms the subject-wise cross validation model in all cases.*

4.3.6 Subject-Wise Cross Validation

In order to show that our results are not caused by over-fitting the classification to specific subjects rather than learning the properties we are looking for (gender, height, age), a *subject-wise* cross validation model was also employed (as explained in Section 4.2.2). Table 4.11 present classification results of *subject-wise* cross validation for all three group classification tasks: gender, height, and age. The feature set contained all features of 6D accelerations and angular velocities (56 in total). For each sensor position, sensitivity, specificity, and PPV were also computed. A comparison of classification results of group classification tasks using *10-fold* cross validation and *subject-wise* cross validation for chest (CH), lower back (LB), right wrist (RW), and left ankle (LA) are presented in Figure 4.10. It is clearly observable that *10-fold* cross validation outperforms *subject-wise*

cross validation in all cases.

In case of gender classification using chest and lower back sensors, the classification rates are 5.12% and 4.83% lower than *10-fold* cross validation. For right wrist and left ankle sensors, the classification rates are 10.01% and 12.53% lower than *10-fold* cross validation. In case of height classification using chest and lower back sensors, the classification rates are 6.8% and 6.04% lower than *10-fold* cross validation. For right wrist and left ankle sensors, the classification rates are 11.83% and 16.86% lower than *10-fold* cross validation.

For the age classification task, a sharp decline in the classification rates is observable in *subject-wise* cross validation. For chest and lower back sensors, the classification rates are 18.89% and 15.89% lower than *10-fold* cross validation. For right wrist and left ankle, the classification rates are 20.99% and 21% lower than *10-fold* cross validation. The main reason of such a sharp decline is because of the unbalanced population in classes C_A^1, C_A^2, and C_A^3 with a subject ratio of 9:6:11.

4.3.7 Body Height Regression

On the level of subject-wise cross validation, it is also possible to consider height and age as regression tasks. Since the population in the recorded data sets is unbalanced with respect to height and age, so the population was divided into balanced classes and it was treated as a classification task. However, it is also interesting to see the results of regression on an unbalance population. In this context, body height regression using random forest regressor was performed for each sensor position. The computed results are presented in Figure 4.11 and Table 4.12. The bar graphs show root mean square (RMS) error of each subject computed during regression. The average RMS errors for each sensor position (in ascending order) are as follows: chest – 8.43cm, left ankle – 8.59cm, right wrist – 9.0cm, and lower back – 9.08cm. For all sensor positions, lower average RMS error is observed for the height group where height is between 170-180cm. The observed values (in ascending order) are as follows: lower back – 4.26cm, chest – 4.37cm, left ankle – 4.68cm, and right wrist – 4.81cm. On the other hand, higher average RMS error is observed for other height groups: between 15.61-17.37cm for the height group where height is less than 170cm and between 11-13.60cm for the height group where height is greater than 180cm. The basic reason of higher average RMS error is sparseness of samples for these two height groups. Only 6 subjects (approx. 23% of all samples) have body height of less than 170cm and 5 subjects (20% of all samples) have a body height

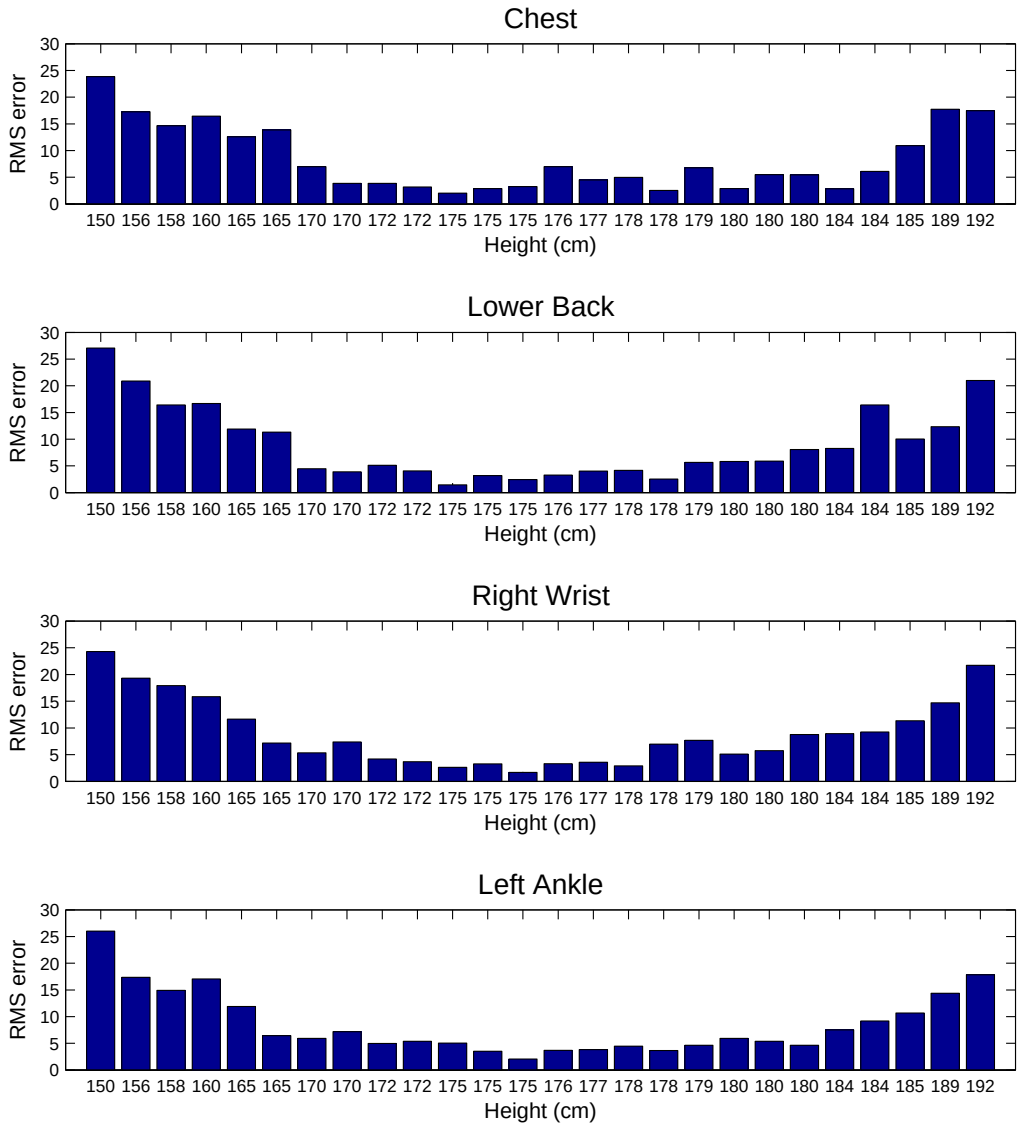


Figure 4.11: *Bar graphs showing root mean square (RMS) error computed during body height regression. For all sensor positions, lower average RMS error (between 4.37-4.81cm) is observed for the height group between 170-180cm.*

of above 180cm. The observed clusters of body heights based on the average RMS error heavily coincide with the chosen classification classes used in body height classification (see Table 4.5). From the observed trends of average RMS error, it can be concluded that a larger data set with balanced and less sparse population of different heights can minimize overall average RMS error of body height regression.

Table 4.12: *Results of body height regression for different sensor positions. Groups of subjects are identified based on the visual analysis of clusters of average RMS errors for each sensor. For all sensor positions, minimum average RMS errors are seen for the group of subjects whose heights are between 170-180cm.*

Regression Task	Body Part	Subjects Group	N	Avg. RMS Error (cm)
Body Height Regression	Chest	Height $<$ 170	6	16.44
		170cm $\leq$ Height $\leq$ 180cm	15	4.37
		Height $>$ 180cm	5	11.00
		All Heights	26	8.43
	Lower Back	Height $<$ 170cm	6	17.37
		170cm $\leq$ Height $\leq$ 180cm	15	4.26
		Height $>$ 180cm	5	13.60
		All Heights	26	9.08
	Right Wrist	Height $<$ 170cm	6	16.02
		170cm $\leq$ Height $\leq$ 180cm	15	4.81
		Height $>$ 180cm	5	13.18
		All Heights	26	9.00
	Left Ankle	Height $<$ 170cm	6	15.61
		170cm $\leq$ Height $\leq$ 180cm	15	4.68
		Height $>$ 180cm	5	11.91
		All Heights	26	8.59

4.3.8 Age Regression

Age regression using random-forest regressor was also performed for each sensor position. The computed results are presented in Figure 4.12 and Table 4.13. In Figure 4.12, the bar graphs show root mean square (RMS) error of each subject computed during regression. The average RMS errors for each sensor position (in ascending order) are: lower back – 11.37, left ankle – 11.54, chest – 12.24, right wrist – 12.43. Similar to body height regression, lower average RMS errors are observed in the center and higher average RMS errors are observed otherwise. Typically, the age group between 34-59 years have lower average RMS error for all sensor positions. The basic reason of lower average RMS error

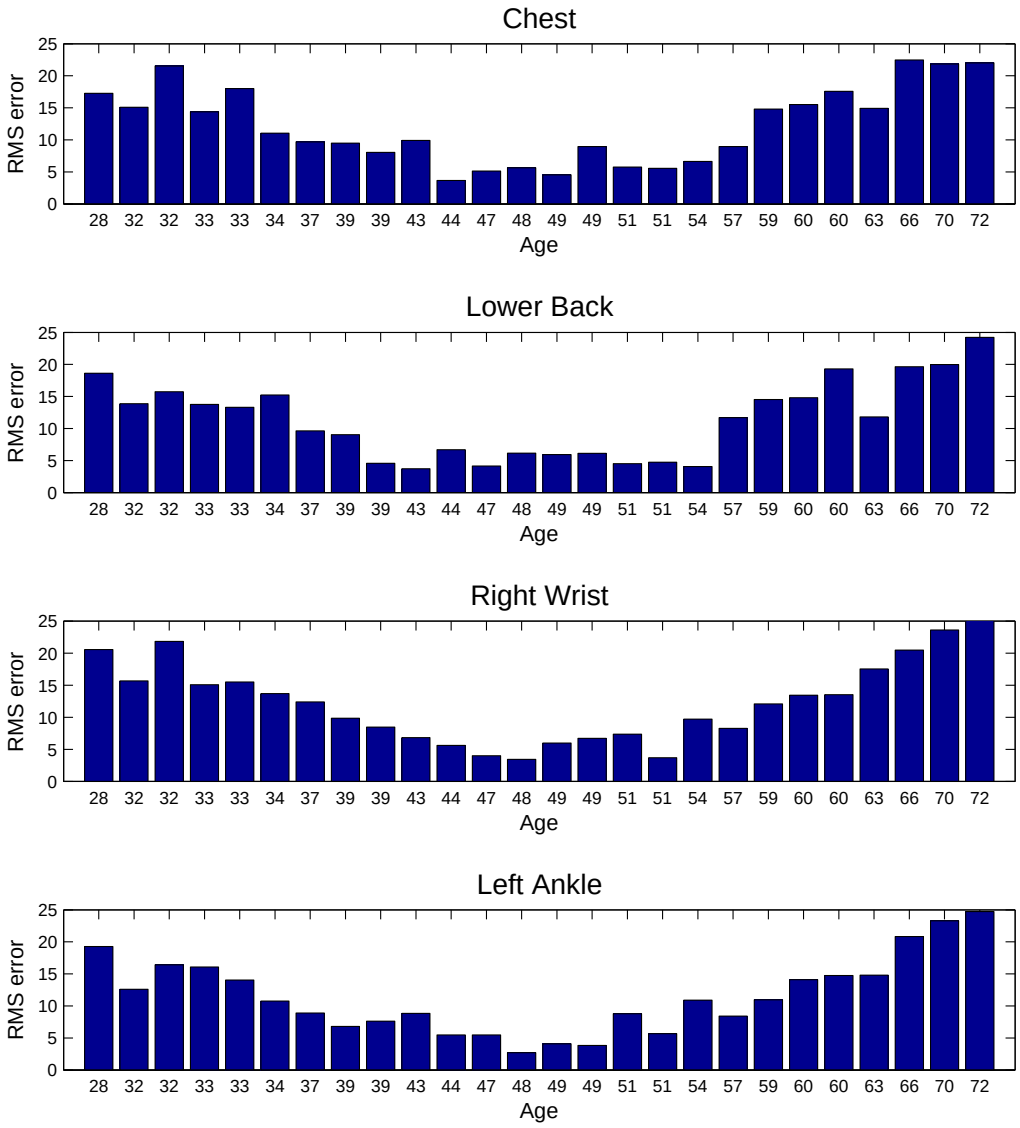


Figure 4.12: *Bar graphs showing root mean square (RMS) error computed during age regression. For all sensor positions, lower average RMS error is observed for the age group between 34-59 years.*

in this age group is due to balanced and enough sample data. This age group contains approximately 58% of the whole samples (15 subjects belong to this group). Maximum average RMS error is observed for the age group where age is greater than 59 years. This group contains only 5 subjects (20% of whole samples) whereas it covers a longer age period between 60-72 years of age (13 years). From these trends, it can be concluded that

Table 4.13: *Results of age regression for different sensor positions. Groups of subjects are identified based on the visual analysis of clusters of average RMS errors for each sensor. For all sensor positions, minimum average RMS errors are seen for the group of subjects whose ages are between 34-59 years.*

Regression Task	Body Part	Subjects Group	N	Avg. RMS Error (years)
Age Regression	Chest	Age<34	5	17.25
		$34 \leq$ Age ≤ 59	15	7.85
		Age > 57	6	19.05
		All Ages	26	12.24
	Lower Back	Age < 37	5	15.05
		$34 \leq$ Age ≤ 59	15	7.38
		Age > 59	6	18.29
		All Ages	26	11.38
	Right Wrist	Age < 37	5	17.72
		$34 \leq$ Age ≤ 59	15	7.86
		Age > 59	6	19.43
		All Ages	26	12.43
	Left Ankle	Age < 34	5	15.68
		$34 \leq$ Age ≤ 59	15	7.28
		Age > 59	6	18.75
		All Ages	26	11.54

a larger data set with balanced and less sparse population covering different age periods can minimize overall average RMS error in age regression.

4.4 Discussion

4.4.1 Summary of Findings

The general problem we tackled is estimation of soft biometrics information from one single step recorded by one inertial sensor. We did so by solving different classification tasks based on motion data of human walking steps represented by accelerations and

angular velocities. Data were recorded by one sensor placed at various locations on the human body, i.e. chest, lower back, right wrist and left ankle. The results show that these classification tasks can be solved well by using accelerometers and/or gyroscopes at any of the given locations. The classification rates were highest for sensors located at the lower back and chest in each of the experiments but still convincingly high when the sensor is attached to wrist or ankle.

Our analysis of the feature sets used in each of the experiments has made clear that there is not one feature mainly responsible for any of the distinctions necessary for a classification. However, the feature importance in each of the classifications gave pointers as to what combination of features produces best results. The most important findings were that angular velocities did not perform better than accelerations.

4.4.2 Comparison with Existing Research

It is not surprising that information about the gender can be recovered by analysis of chest or lower back movement. Effects of marker placement and viewpoint selection for recording locomotion are discussed extensively in the works of Troje [Tro02], as was the high relevance of hip movement for gender classification by human observers. However, we have presented new findings that accelerations associated to wrist and ankle movement alone allow for classification of gender also. To our knowledge, we are also the first to show that classification of height and age groups is possible from non-visual features. This is as yet done by solely relying on image- or video-based features. Makihara et al. [MOIY11] introduce a paper on gait-based age estimation by Gaussian process regression on silhouette-based features of bodies (contrary to face-based age estimation as presented by Stewart et al. [SPZ13]). Their investigation was based on standard resolution video data. They have constructed a whole-generation database of over 1000 individuals, their age ranging from 2 to 94.

Our initial situation is clearly different from this in terms of sensor modalities. The use of commercial smart phones and wearables is an attractive chance to monitor biometrics properties nowadays. Mobile phones and smart devices are a convenient platform for recording information in an every-day setup. Our experiments have shown that information recorded by a single sensor such as a smart device suffices for the estimation of basic soft biometrics properties. Particularly, the wrist was an important subject for tests because smart devices are commonly worn at that location.

Estimating biometrics properties based on motion data makes sense in a number of

different scenarios. In some of them, the focus may be on hard biometrics properties in order to facilitate online identity checks and close security gaps. A number of previous works have shown that identification and authentication problems can be solved by classification of motion data acquired by mobile devices. Derawi and Bours [DB13] show that recognition of specific users can be done in real-time based on data collected by mobile phones. Their method can correctly identify enrolled users based on learning templates of different walking trials.

On the other hand, the attention may be directed to soft biometrics properties. Monitoring health or preventing and curing injury are use cases that represent this idea. Previous works have shown that accelerometers are well-suited for detection and recognition of events and activities. In their paper on sensory motor performance, Albert et al.[AKHJ12] discuss a new method to classify different types of falls in order to rapidly assess the cause and necessary emergency response. They present very good results classifying accelerometer data acquired by commercial mobile phones which were attached to the lower backs of test subjects. In their comparative evaluation of five machine learning classifiers, support vector machines performed best achieving accuracy values near 98%. Classification by decision trees only performed second best in their experiments at $94 - 98\%$ accuracy for fall detection and at $98 - 99\%$ accuracy for fall type classification. In their paper on gait pattern classification, Von Tscharner et al.[TEM13] even conclude that a combination of PCA, SVM and ICA is most reliable dealing with high intra- and inter-subject variability. However, in their survey on mobile gait classification, Schneider et al. [SMA$^+$13] make an attempt to settle the disagreement about suitable classification algorithms. In their study, they conclude that Random Forest is best suited for classification of gait-related properties. In our setup, we decided to use random forest in order to produce comparable results. One additional benefit of this choice is that there is a low number of parameters that have to be chosen. Also, the random forest method enables computing the significance and importance of each feature in overall classification. This helped us to investigate and perform a comparative study of features importance for each sensor position in different classification tasks.

4.4.3 Limitations

Since our database is much smaller than the one introduced by Makihara et al. [MOIY11] and the variety of biometrics features was also smaller (e. g. age covered only three decades), our experiments can only serve as proof of concept for now. Testing classifiers

of non-image-based features on a larger database comprising wider ranges of biometric properties is a direction for future work.

Another limitation of our database is that it only consists of data belonging to patients with complaints of back pain. It will be worthy to perform further experiments to record data of participants without back pain (control group). Classification tasks can then be performed for the patients group, the control group, and a combination of both.

One noteworthy limitation we had to face in our experiments was the possible uncertainty in sensor placement. Irrespective of how carefully each involved sensor is placed, the accuracy of placement depends on physical characteristics of test subjects which may vary between individuals to some extent.

5

Conclusion

In this doctoral research work, we have explored various different aspects of human motion analysis. We have proposed a relational database approach for better data management of human motion capture data sets, which are required in a range of data-driven applications. A number of data search and retrieval example has been presented, which show how complex data using annotations can be effectively fetched from the database in minimum time. Two different applications of the proposed database are also presented. The first application is an extended motion capture player which can play motion sequences by searching the given input from the relational database. The second application is automatic generation of statistical models using the proposed relational database. Furthermore, a similar concept of relation database is also proposed to manage quadruped mocap data sets. The proposed database can handle motions of different types of animals as well as a range of sensor modalities such as accelerometers, motion capture data, videos etc. Several examples of data retrieval using standard SQL statements are also shown.

Motion synthesis is another important aspect in human motion analysis. In this respect, we have proposed a full body motion reconstruction algorithm using a very sparse configuration of sensors. We use a trunk mounted sensor to estimate ground contacts.

In the reconstruction process, we use low dimensional inputs from two wrists mounted sensors along with the estimated ground contacts. Extensive testing of the proposed algorithm is carried out with a number of different types of motions and the results show that realistic and visually appealing motion reconstruction is possible with such a sparse configuration of sensors.

Analysis of human motion using inertial sensors in order to estimate soft biometrics is the last aspect of the research conducted in this doctoral study. In this context, we have proposed a novel method of estimating gender, height, and age of a human from the inertial measurements of a single step. Here, we extract several time domain and frequency domain features using our feature extraction module. Classification and regression tasks are performed using random forest classifier and random forest regressor in order to estimate human soft biometrics. The results of extensive testing have shown that inertial measurements of a single step can reveal the gender, height, and age of a human with higher accuracy using our proposed algorithm.

5.1 Future work

The relational databases for human motion and quadruped motion data will be extended to include more sensor modalities. The mocap data from other publicly available data sets will also be added so that all freely available motion capture data sets can be accessed from a single point and in a single data format. Fast similarity search is another direction of future work, where one can develop effective methods of searching similar poses in the relational database. Additionally, development of more complex data-driven methods using the presented databases are planned for future research.

Our proposed motion reconstruction algorithm has several limitations with estimation of ground contacts from the lower trunk sensor. A natural direction of future work is to develop a better algorithm which can estimate ground contacts for a broader range of motions. We will also work in the direction of estimation of ground contacts from the wrists mounted sensors. Since the low dimensional inputs of the wrists sensors are used in the reconstruction process, so estimation of ground contacts from the same sensors will further minimize the number of sensors required for full body motion reconstruction.

Our empirical finding on estimation of human soft biometrics can be extended further in several directions. One direction is additional refinement of the proposed algorithm especially for estimating age in order to minimize average RMS error. Typically, a good estimate of the direction of gravity should improve the results – at sensors position with

less change in orientation (chest, lower back) the classification rates had been better than at the ones with higher change (wrist, ankle). In future work we will try to adopt a model based estimate of body-part orientation using techniques similar to the ones used in [RTKW15] to come up with such estimates. A natural question arising from these empirical findings is that whether humans have unique gait patterns and whether this uniqueness can be reliably estimated from the inertial measurements. In future, we will work in this direction as well to develop an algorithm which can estimate uniqueness in the gait patterns of humans.

The proposed algorithm can possibly be used to estimate and classify the type of the surface on which a person has walked. In the current setup, we have performed experiments on two types of surfaces only: hard surface (concrete floor) and soft surface (exercise mattress). In future, new data will be collected from various other surfaces such as grassy surfaces, sandy surfaces, rough terrains (e.g. covered with pebbles) etc. The proposed algorithm will be further refined and developed in order to reliably estimate the ground surface based on the inertial data of the gait.

Another possible direction of the future work is estimating the moods or emotions of a person from the inertial measurements of his walk. Michalak et al. [MTF+09], have shown that the mood states can have causal effects on the way people walk. The gait patterns of a sad and depressed walker are characterized by reduced walking speed, arm swing, larger lateral swaying movements of the upper body, vertical head movements, and a more slumped posture. However, they have used a standard motion capture system to record gait data, which makes it less practical in non-lab environments. In future, our focus of investigation will be the question of estimation of various moods of a person from the analysis of the inertial data of his gait.

Bibliography

[Acc15] ACCLAIM ASF/AMC: *Acclaim ASF/AMC*. http://research.cs.wisc.edu/graphics/Courses/cs-838-1999/Jeff/ASF-AMC.html, 2015. – Accessed Jul 2nd, 2015

[ACG09] AWAD, Charly ; COURTY, Nicolas ; GIBET, Sylvie: A Database Architecture for Real-Time Motion Retrieval. In: *Content-Based Multimedia Indexing, 2009. CBMI'09. Seventh International Workshop on* IEEE, 2009, S. 225–230

[AF02] ARIKAN, Okan ; FORSYTH, David A.: Interactive motion generation from examples. In: *ACM Transactions on Graphics (TOG)* Bd. 21 ACM, 2002, S. 483–490

[AFO03] ARIKAN, Okan ; FORSYTH, David A. ; O'BRIEN, James F.: Motion synthesis from annotations. In: *ACM Transactions on Graphics (TOG)* 22 (2003), Nr. 3, S. 402–408

[AKHJ12] ALBERT, Mark V. ; KORDING, Konrad ; HERRMANN, Megan ; JAYARAMAN, Arun: Fall Classification by Machine Learning Using Mobile Phones. In: *PLoS ONE* 7 (2012), 05, Nr. 5, e36556. http://dx.doi.org/10.1371/journal.pone.0036556. – DOI 10.1371/journal.pone.0036556

[Ale84] ALEXANDER, R M.: The gaits of bipedal and quadrupedal animals. In: *The International Journal of Robotics Research* 3 (1984), Nr. 2, S. 49–59

[BBBG14] BARNACHON, Mathieu ; BOUAKAZ, Saïda ; BOUFAMA, Boubakeur ; GUILLOU, Erwan: Ongoing human action recognition with motion capture. In: *Pattern Recognition* 47 (2014), Nr. 1, 238 - 247. http:

//dx.doi.org/http://dx.doi.org/10.1016/j.patcog.2013.06.020. – DOI http://dx.doi.org/10.1016/j.patcog.2013.06.020. – ISSN 0031–3203

[BI04] BAO, Ling ; INTILLE, StephenS.: Activity Recognition from User-Annotated Acceleration Data. In: FERSCHA, Alois (Hrsg.) ; MATTERN, Friedemann (Hrsg.): *Pervasive Computing* Bd. 3001. Springer Berlin Heidelberg, 2004. – ISBN 978–3–540–21835–7, S. 1–17

[Bio15] BIOVISION: *Biovision BVH*. http://research.cs.wisc.edu/graphics/Courses/cs-838-1999/Jeff/BVH.html, 2015. – Accessed Jul 2nd, 2015

[BOP97] BRAND, M. ; OLIVER, N. ; PENTLAND, A.: Coupled Hidden Markov Models for Complex Action Recognition. In: *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97)*. Washington, DC, USA : IEEE Computer Society, 1997 (CVPR '97). – ISBN 0–8186–7822–4, 994–

[BP07] BARAN, Ilya ; POPOVIĆ, Jovan: Automatic rigging and animation of 3D characters. In: *ACM Transactions on Graphics (TOG)* Bd. 26 ACM, 2007, S. 72

[Bre01] BREIMAN, Leo: Random Forests. In: *Machine Learning* 45 (2001), Nr. 1, 5-32. http://dx.doi.org/10.1023/A:1010933404324. – DOI 10.1023/A:1010933404324. – ISSN 0885–6125

[BSC05] BASU, Suddha ; SHANBHAG, Shrinath ; CHANDRAN, Sharat: Search and transitioning for motion captured sequences. In: *Proceedings of the ACM symposium on Virtual reality software and technology* ACM, 2005, S. 220–223

[BWK+13] BERNARD, Jürgen ; WILHELM, Nils ; KRÜGER, Björn ; MAY, Thorsten ; SCHRECK, Tobias ; KOHLHAMMER, Jörn: MotionExplorer: Exploratory Search in Human Motion Capture Data Based on Hierarchical Aggregation. In: *IEEE Trans. on Visualization and Computer Graphics (Proc. VAST)* (2013)

[BWKW14] BAUMANN, Jan ; WESSEL, Raoul ; KRÜGER, Björn ; WEBER, Andreas: Action Graph: A Versatile Data Structure for Action Recognition. In: *GRAPP 2014 - International Conference on Computer Graphics Theory and Applications*, SCITEPRESS, Januar 2014

[BZ13] BARBIČ, Jernej ; ZHAO, Yili: *ASF/AMC Motion Capture Player*. `http://mocap.cs.cmu.edu/tools.php`, 2013. – Accessed Jul 7th, 2013

[CH05] CHAI, Jinxiang ; HODGINS, Jessica K.: Performance animation from low-dimensional control signals. In: *ACM Transactions on Graphics (TOG)* 24 (2005), Nr. 3, S. 686–696

[CH12] CASHMAN, Thomas J. ; HORMANN, Kai: A continuous, editable representation for deforming mesh sequences with separate signals for time, pose and shape. In: *Computer Graphics Forum* Bd. 31 Wiley Online Library, 2012, S. 735–744

[CKJ+11] COROS, Stelian ; KARPATHY, Andrej ; JONES, Ben ; REVERET, Lionel ; VAN DE PANNE, Michiel: Locomotion skills for simulated quadrupeds. In: *ACM Transactions on Graphics (TOG)* Bd. 30 ACM, 2011, S. 59

[CMU03] CMU: *CMU Motion Capture Database*. `http://mocap.cs.cmu.edu/`, 2003. – Accessed Mar 23rd, 2014

[Com15] COMPUTER GRAPHICS, UNIVERSITY OF BONN: *Generic Motion Models based on Quadrupedal Data*. `http://cg.cs.uni-bonn.de/en/projects/gemmquad/`, 2015. – Accessed Jul 5th, 2015

[DATTS08] DE AGUIAR, Edilson ; THEOBALT, Christian ; THRUN, Sebastian ; SEIDEL, Hans-Peter: Automatic Conversion of Mesh Animations into Skeleton-based Animations. In: *Computer Graphics Forum* Bd. 27 Wiley Online Library, 2008, S. 389–397

[DB13] DERAWI, Mohammad ; BOURS, Patrick: Gait and activity recognition using commercial phones. In: *Computers & Security* 39, Part B (2013), Nr. 0, 137 - 144. `http://dx.doi.org/http://dx.doi.org/10.1016/j.cose.2013.07.004`. – DOI http://dx.doi.org/10.1016/j.cose.2013.07.004. – ISSN 0167–4048

[Del15] DELSYS INC.: *Delsys, Inc. — EMG, EKG, Accelerometers, Foot Switches, Goniometers*. http://www.delsys.com/about-delsys/, 2015. – Accessed Jul 2nd, 2015

[DKZ10] DIJKSTRA, Baukje ; KAMSMA, Yvo ; ZIJLSTRA, Wiebren: Detection of gait and postures using a miniaturised triaxial accelerometer-based system: accuracy in community-dwelling older adults. In: *Age and ageing* 39 (2010), Nr. 2, S. 259–262

[DNBB10] DERAWI, M.O. ; NICKEL, C. ; BOURS, P. ; BUSCH, C.: Unobtrusive User-Authentication on Mobile Phones Using Biometric Gait Recognition. In: *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on*, 2010, S. 306–311

[FF05] FORBES, Kate ; FIUME, Eugene: An efficient search algorithm for motion data using weighted PCA. In: *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation* ACM, 2005, S. 67–76

[FF09] FITZ, Kelly R. ; FULOP, Sean A.: A unified theory of time-frequency reassignment. In: *arXiv preprint arXiv:0903.3080* (2009)

[FLFM+05] FOSTER, Randal C. ; LANNINGHAM-FOSTER, Lorraine M. ; MANOHAR, Chinmay ; MCCRADY, Shelly K. ; NYSSE, Lana J. ; KAUFMAN, Kenton R. ; PADGETT, Denny J. ; LEVINE, James A.: Precision and accuracy of an ankle-worn accelerometer-based pedometer in step counting and energy expenditure. In: *Preventive Medicine* 41 (2005), Nr. 3–4, 778 - 783. http://dx.doi.org/http://dx.doi.org/10.1016/j.ypmed.2005.07.006. – DOI http://dx.doi.org/10.1016/j.ypmed.2005.07.006. – ISSN 0091–7435

[Geo15] GEORGIA TECH: *Human Identification at Distance Database*. http://www.cc.gatech.edu/cpl/projects/hid/, 2015. – Accessed Jul 2nd, 2015

[GF15] GRUENVOGEL, Stefan ; FUHRMANN, Arnulph: *Cologne motion capture database*. http://mocap.web.fh-koeln.de/index.php, 2015. – Accessed Jul 2nd, 2015

[GFB11] GUERRA-FILHO, G. ; BISWAS, A.: The human motion database: A cognitive and parametric sampling of human motion. In: *Automatic Face Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on*, 2011, S. 103–110

[HEK13] HUNG, Hayley ; ENGLEBIENNE, Gwenn ; KOOLS, Jeroen: Classifying Social Actions with a Single Accelerometer. In: *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. New York, NY, USA : ACM, 2013 (UbiComp '13). – ISBN 978–1–4503–1770–2, 207–210

[HKH07] HE, Hui ; KIGUCHI, Kazuo ; HORIKAWA, Etsuo: A study on lower-limb muscle activities during daily lower-limb motions. In: *International Journal of Bioelectromagnetism* 9 (2007), Nr. 2, S. 79–84

[HKL06] HWANG, Bon-Woo ; KIM, Sungmin ; LEE, Seong-Whan: A full-body gesture database for automatic gesture recognition. In: *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on* IEEE, 2006, S. 243–248

[Ine14] INERTIATECHNOLOGY: *Motion Tracking — Inertia Technology*. `http://inertia-technology.com/category/products/motion-tracking`, 2014. – Accessed May 11th, 2014

[JDLP11] JUNEJO, Imran N. ; DEXTER, Emilie ; LAPTEV, Ivan ; PEREZ, Patrick: View-Independent Action Recognition from Temporal Self-Similarities. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (2011), jan, Nr. 1, 172–185. `http://dx.doi.org/10.1109/TPAMI.2010.68`. – DOI 10.1109/TPAMI.2010.68. – ISSN 0162–8828

[JNP+14] JEAN-BAPTISTE, Emilie M. D. ; NABIEI, Roozbeh ; PAREKH, Manish ; FRINGI, Evangelia ; DROZDOWSKA, Bogna ; BABER, Chris ; JANCOVIC, Peter ; ROTSHEIN, Pia ; RUSSELL, Martin J.: Intelligent Assistive System Using Real-Time Action Recognition for Stroke Survivors. In: *2014 IEEE International Conference on Healthcare Informatics, ICHI 2014, Verona, Italy, September 15-17, 2014*, 2014, 39–44

[Joh73] JOHANSSON, Gunnar: Visual perception of biological motion and a model for its analysis. In: *Perception & Psychophysics* 14 (1973), Nr.

2, 201-211. http://dx.doi.org/10.3758/BF03212378. – DOI 10.3758/BF03212378. – ISSN 0031–5117

[KCHO10] KELLY, Philip ; CONAIRE, Ciarán ¦ HODGINS, Jessica ; O'CONNER, Noel E.: Human motion reconstruction using wearable accelerometers (Poster). In: *ACM SIGGRAPH / Eurographics Symposium on Computer Animation (SCA)*, 2010

[KCOH13] KELLY, Philip ; CONAIRE, Ciarán ¦ O'CONNER, Noel E. ; HODGINS, Jessica: Motion Synthesis for Sports Using Unobtrusive Lightweight Body-Worn and Environment Sensing. In: *Computer Graphics Forum* Bd. 32 Wiley Online Library, 2013, S. 48–60

[KG04] KOVAR, Lucas ; GLEICHER, Michael: Automated extraction and parameterization of motions in large data sets. In: *ACM Transactions on Graphics* 23 (2004), Nr. 3, 559–568. http://doi.acm.org/10.1145/1015706.1015760. – ISSN 0730–0301. – SIGGRAPH 2004

[KG06] KIRCHER, Scott ; GARLAND, Michael: Editing arbitrarily deforming surface animations. In: *ACM Transactions on Graphics (TOG)* 25 (2006), Nr. 3, S. 1098–1107

[KG08] KIRCHER, Scott ; GARLAND, Michael: Free-form motion processing. In: *ACM Transactions on Graphics (TOG)* 27 (2008), Nr. 2, S. 12

[KL13] KIM, Hyejin ; LEE, Sung-Hee: Reconstructing whole-body motions with wrist trajectories. In: *Graphical Models* 75 (2013), Nr. 6, S. 328–345

[KMP07] KILIAN, Martin ; MITRA, Niloy J. ; POTTMANN, Helmut: Geometric modeling in shape space. In: *ACM Transactions on Graphics (TOG)* 26 (2007), Nr. 3, S. 64

[KOF05] KIRK, Adam G. ; O'BRIEN, James F. ; FORSYTH, David A.: Skeletal Parameter Estimation from Optical Motion Capture Data. In: *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) 2005*, 2005, 782–788

[KSL13] KIM, Jongmin ; SEOL, Yeongho ; LEE, Jehee: Human motion reconstruction from sparse 3D motion sensors using kernel CCA-based regression. In: *Computer Animation and Virtual Worlds* 24 (2013), Nr. 6, S. 565–576

[KTMW08] KRÜGER, Björn ; TAUTGES, Jochen ; MÜLLER, M. ; WEBER, Andreas: Multi-Mode Tensor Representation of Motion Data. In: *Journal of Virtual Reality and Broadcasting* 5 (2008), Juli, Nr. 5. – ISSN 1860–2037

[KTWZ10] KRÜGER, Björn ; TAUTGES, Jochen ; WEBER, Andreas ; ZINKE, Arno: Fast local and global similarity searches in large motion capture databases. In: *Proceedings of the 2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* Eurographics Association, 2010, S. 1–10

[KWM11] KWAPISZ, Jennifer R. ; WEISS, Gary M. ; MOORE, Samuel A.: Activity Recognition Using Cell Phone Accelerometers. In: *SIGKDD Explor. Newsl.* 12 (2011), mar, Nr. 2, 74–82. http://dx.doi.org/10.1145/1964897.1964918. – DOI 10.1145/1964897.1964918. – ISSN 1931–0145

[LCF00] LEWIS, John P. ; CORDNER, Matt ; FONG, Nickson: Pose space deformation: a unified approach to shape interpolation and skeleton-driven deformation. In: *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* ACM Press/Addison-Wesley Publishing Co., 2000, S. 165–172

[Len90] LENZ, J.E.: A review of magnetic sensors. In: *Proceedings of the IEEE* 78 (1990), Jun, Nr. 6, S. 973–989. – ISSN 0018–9219

[Lin06] LIN, Yi: Efficient human motion retrieval in large databases. In: *Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia* ACM, 2006, S. 31–37

[Lin15] LIN, Edgar Chia-Han: A Research on 3D Motion Database Management and Query System Based on Kinect. Version: 2015. http://dx.doi.org/10.1007/978-94-017-9558-6_4. In: PARK, James J. (Jong H. (Hrsg.) ; PAN, Yi (Hrsg.) ; KIM, Cheonshik (Hrsg.) ; YANG, Yun (Hrsg.): *Future Information Technology - II* Bd. 329. Springer Netherlands, 2015. – ISBN 978–94–017–9557–9, 29-35

[LM15] LIANG, Qinghua ; MIAO, Zhenjiang: An editable interface for motion data retrieval based on Labanotation. In: *Multimedia Expo Workshops (ICMEW), 2015 IEEE International Conference on*, 2015, S. 1–6

[LMTL03] LE MASURIER, Guy C. ; TUDOR-LOCKE, Catrine: Comparison of pedometer and accelerometer accuracy under controlled conditions. In: *Medicine and science in sports and exercise* 35 (2003), May, Nr. 5, 867-871. http://dx.doi.org/10.1249/01.mss.0000064996.63632.10. – DOI 10.1249/01.mss.0000064996.63632.10. – ISSN 0195–9131

[LN07] LV, Fengjun ; NEVATIA, Ramakant: Single View Human Action Recognition using Key Pose Matching and Viterbi Path Searching. In: *2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), 18-23 June 2007, Minneapolis, Minnesota, USA*, 2007

[LS99] LEE, Jehee ; SHIN, Sung Y.: A hierarchical approach to interactive motion editing for human-like figures. In: *Proceedings of the 26th annual conference on Computer graphics and interactive techniques* ACM Press/Addison-Wesley Publishing Co., 1999, S. 39–48

[LS01] LEE, Jehee ; SHIN, Sung Y.: A Coordinate-Invariant Approach to Multiresolution Motion Analysis. In: *Graphical Models* 63 (2001), Nr. 2, 87 - 105. http://dx.doi.org/http://dx.doi.org/10.1006/gmod.2001.0548. – DOI http://dx.doi.org/10.1006/gmod.2001.0548. – ISSN 1524–0703

[LW02] LIAW, Andy ; WIENER, Matthew: Classification and Regression by randomForest. In: *R News* 2 (2002), Nr. 3, 18-22. http://CRAN.R-project.org/doc/Rnews/

[LWC+11] LIU, Huajun ; WEI, Xiaolin ; CHAI, Jinxiang ; HA, Inwoo ; RHEE, Taehyun: Realtime human motion control with a small number of inertial sensors. In: *Symposium on Interactive 3D Graphics and Games* ACM, 2011, S. 133–140

[LWS+15] LIEW, Chee S. ; WAH, Teh Y. ; SHUJA, Junaid ; DAGHIGHI, Babak u. a.: Mining Personal Data Using Smartphones and Wearable Devices: A Survey. In: *Sensors* 15 (2015), Nr. 2, S. 4430–4469

[LZD+12] LI, Fan ; ZHAO, Chunshui ; DING, Guanzhong ; GONG, Jian ; LIU, Chenxing ; ZHAO, Feng: A reliable and accurate indoor localization method using phone inertial sensors. In: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* ACM, 2012, S. 421–430

[LZWM05] LIU, Guodong ; ZHANG, Jingdan ; WANG, Wei ; MCMILLAN, Leonard: A system for analyzing and indexing human-motion databases. In: *Proceedings of the 2005 ACM SIGMOD international conference on Management of data* ACM, 2005, S. 924–926

[Mai96] MAIOCCHI, Roberto: *3-D character animation using motion capture*. Upper Saddle River, NJ, USA : Prentice-Hall, Inc., 1996. – 10–39 S. – ISBN 0–13–518309–X

[MC12] MIN, Jianyuan ; CHAI, Jinxiang: Motion Graphs++: A Compact Generative Model for Semantic Motion Analysis and Synthesis. In: *ACM Trans. Graph.* 31 (2012), November, Nr. 6, 153:1–153:12. `http://dx.doi.org/10.1145/2366145.2366172`. – DOI 10.1145/2366145.2366172. – ISSN 0730–0301

[Met15] METHOD, Gokhale: *Gokhale Method — Gokhale Method Institute*. `http://www.gokhalemethod.com/`, 2015. – Accessed Feb 27th, 2015

[MHK06] MOESLUND, Thomas B. ; HILTON, Adrian ; KRÜGER, Volker: A survey of advances in vision-based human motion capture and analysis. In: *Computer vision and image understanding* 104 (2006), Nr. 2, S. 90–126

[MK05] MUKAI, Tomohiko ; KURIYAMA, Shigeru: Geostatistical Motion Interpolation. In: *ACM SIGGRAPH 2005 Papers*. New York, NY, USA : ACM, 2005 (SIGGRAPH '05), 1062–1070

[ML05] MARSLAND, SA ; LAPEER, RJ: Physics-based animation of a trotting horse in a virtual environment. In: *Information Visualisation, 2005. Proceedings. Ninth International Conference on* IEEE, 2005, S. 398–403

[MLC10] MIN, Jianyuan ; LIU, Huajun ; CHAI, Jinxiang: Synthesis and Editing of Personalized Stylistic Human Motion. In: *Proceedings of the 2010 ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*. New York, NY, USA : ACM, 2010 (I3D '10). – ISBN 978–1–60558–939–8, 39–46

[MM01] MEREDITH, Maddock ; MADDOCK, S: Motion capture file formats explained. In: *Department of Computer Science, University of Sheffield* 211 (2001), S. 241–244

[MOIY11] MAKIHARA, Y. ; OKUMURA, M. ; IWAMA, H. ; YAGI, Y.: Gait-based Age Estimation using a Whole-generation Gait Database. In: *Proc. of the International Joint Conference on Biometrics (IJCB2011)*. Washington D.C., USA, Oct. 2011, S. 1–6

[Mot14] MOTICON: *OpenGo science - Moticon*. `http://www.moticon.com`, 2014. – Accessed May 11th, 2014

[Mot15] MOTION LAB SYSTEMS: *C3D.ORG*. `https://www.c3d.org/`, 2015. – Accessed Jul 2nd, 2015

[MPP06] MA, Yingliang ; PATERSON, Helena M. ; POLLICK, Frank E.: A motion capture library for the study of identity, gender, and emotion perception from biological motion. In: *Behavior research methods* 38 (2006), Nr. 1, S. 134–141

[MRC05] MÜLLER, Meinard ; RÖDER, Tido ; CLAUSEN, Michael: Efficient content-based retrieval of motion capture data. In: *ACM Transactions on Graphics (TOG)* Bd. 24 ACM, 2005, S. 677–685

[MRC+07] MÜLLER, M. ; RÖDER, T. ; CLAUSEN, M. ; EBERHARDT, B. ; KRÜGER, B. ; WEBER, A.: Documentation Mocap Database HDM05 / Universität Bonn. 2007 (CG-2007-2). – Forschungsbericht. – ISSN 1610–8892

[MSGK14] MORRIS, Dan ; SAPONAS, T. S. ; GUILLORY, Andrew ; KELNER, Ilya: RecoFit: Using a Wearable Sensor to Find, Recognize, and Count Repetitive Exercises. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA : ACM, 2014 (CHI '14). – ISBN 978–1–4503–2473–1, 3225–3234

[MST03] MORI, T ; SHIMOSAKA, M ; TSUJIOKA, K: *ICS action database*. 2003

[MTD+15] MANDERY, C. ; TERLEMEZ, Ö. ; DO, M. ; VAHRENKAMP, N. ; ASFOUR, T.: The KIT Whole-Body Human Motion Database. In: *International Conference on Advanced Robotics (ICAR)*, 2015

[MTF+09] MICHALAK, Johannes ; TROJE, Nikolaus F. ; FISCHER, Julia ; VOLLMAR, Patrick ; HEIDENREICH, Thomas ; SCHULTE, Dietmar: Embodiment of sadness and depression—gait patterns associated with dysphoric mood. In: *Psychosomatic medicine* 71 (2009), Nr. 5, S. 580–587

[NHB12] NEUGEBAUER, Jennifer M. ; HAWKINS, David A. ; BECKETT, Laurel: Estimating youth locomotion ground reaction forces using an accelerometer-based activity monitor. In: *PloS one* 7 (2012), Nr. 10, S. e48182

[OCK+13] OFLI, Ferda ; CHAUDHRY, Rizwan ; KURILLO, Gregorij ; VIDAL, René ; BAJCSY, Ruzena: Berkeley mhad: A comprehensive multimodal human action database. In: *Applications of Computer Vision (WACV), 2013 IEEE Workshop on* IEEE, 2013, S. 53–60

[OGB14] OSHIN, Olusegun ; GILBERT, Andrew ; BOWDEN, Richard: Capturing relative motion and finding modes for action recognition in the wild. In: *Computer Vision and Image Understanding* 125 (2014), Nr. 0, 155 - 171. http://dx.doi.org/http://dx.doi.org/10.1016/j.cviu.2014.04.005. – DOI http://dx.doi.org/10.1016/j.cviu.2014.04.005. – ISSN 1077–3142

[Opa15] OPAL, APDM: *Wireless, Wearable, Synchronized Inertial Measurement Units (IMUs) — APDM, Inc.* http://www.apdm.com/wearable-sensors/, 2015. – Accessed Mar 1st, 2015

[Ope15] OPENSIM: *Marker (.trc) Files - OpenSim Documentation - Confluence*. http://simtk-confluence.stanford.edu:8080/display/OpenSim/Marker+(.trc)+Files, 2015. – Accessed Jul 2nd, 2015

[Opt14] OPTITRACK: *Motion Capture Systems by OptiTrack*. http://www.naturalpoint.com/optitrack/, 2014. – Accessed May 11th, 2014

[PB02] PULLEN, Katherine ; BREGLER, Christoph: Motion capture assisted animation: Texturing and synthesis. In: *ACM Transactions on Graphics (TOG)* Bd. 21 ACM, 2002, S. 501–508

[PEK+06] PARKKA, J. ; ERMES, M. ; KORPIPAA, P. ; MANTYJARVI, J. ; PELTOLA, J. ; KORHONEN, I.: Activity classification using realistic data from wearable sensors. In: *Information Technology in Biomedicine, IEEE Transactions on* 10 (2006), Jan, Nr. 1, S. 119–128. http://dx.doi.org/10.1109/TITB.2005.856863. – DOI 10.1109/TITB.2005.856863. – ISSN 1089–7771

[Pha14a] PHAN, Thomas: Improving Activity Recognition via Automatic Decision Tree Pruning. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. New York, NY, USA : ACM, 2014 (UbiComp '14 Adjunct). – ISBN 978–1–4503–3047–3, 827–832

[Pha14b] PHASESPACE: *PhaseSpace Motion Capture*. `http://www.phasespace.com`, 2014. – Accessed May 11th, 2014

[Pos13a] POSTGRESQL: *Documentation: 9.0: EXPLAIN*. `http://www.postgresql.org/docs/9.0/static/sql-explain.html`, 2013. – Accessed Jul 7th, 2013

[Pos13b] POSTGRESQL: *Documentation: 9.0: Index Types*. `http://www.postgresql.org/docs/9.0/static/indexes-types.html`, 2013. – Accessed Apr 15th, 2013

[Qua14] QUALISYS: *Motion capture - Mocap - Qualisys Motion Capture Systems*. `http://www.qualisys.com`, 2014. – Accessed May 11th, 2014

[RF03] RAMANAN, Deva ; FORSYTH, David A.: Automatic Annotation of Everyday Movements. In: *Neural Information Processing Systems*, 2003

[RH91] RAIBERT, Marc H. ; HODGINS, Jessica K.: Animation of dynamic legged locomotion. In: *ACM SIGGRAPH Computer Graphics* Bd. 25 ACM, 1991, S. 349–358

[RKW15] RIAZ, Qaiser ; KRÜGER, Björn ; WEBER, Andreas: A Relational Database for Human Motion Data. In: *Computational Science and Its Applications – ICCSA 2015* Bd. 9159. Springer International Publishing, 2015. – ISBN 978–3–319–21412–2, S. 234–249

[RLS09] ROETENBERG, Daniel ; LUINGE, Henk ; SLYCKE, Per: Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors. In: *Xsens Motion Technologies BV, Tech. Rep* (2009)

[RSH+05] REN, Liu ; SHAKHNAROVICH, Gregory ; HODGINS, Jessica K. ; PFISTER, Hanspeter ; VIOLA, Paul: Learning Silhouette Features for Control of

Human Motion. In: *ACM Trans. Graph.* 24 (2005), Oktober, Nr. 4, 1303–1331. http://dx.doi.org/10.1145/1095878.1095882. – DOI 10.1145/1095878.1095882. – ISSN 0730–0301

[RTKW15] RIAZ, Qaiser ; TAO, Guanhong ; KRÜGER, Björn ; WEBER, Andreas: Motion reconstruction using very few accelerometers and ground contacts. In: *Graphical Models* 79 (2015), S. 23–38

[RW03] ROBBINS, K L. ; WU, Q: Development of a computer tool for anthropometric analyses. In: *Proceedings of the International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS'03), pages 347–353, Las Vegas*, CSREA Press, 2003

[SBB10] SIGAL, Leonid ; BALAN, Alexandru O. ; BLACK, Michael J.: Humaneva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. In: *International Journal of Computer Vision* 87 (2010), Nr. 1-2, S. 4–27. – ISSN 0920–5691

[SH07] SAFONOVA, Alla ; HODGINS, Jessica K.: Construction and Optimal Search of Interpolated Motion Graphs. In: *ACM SIGGRAPH 2007 Papers*. New York, NY, USA : ACM, 2007 (SIGGRAPH '07)

[SH08] SLYPER, Ronit ; HODGINS, Jessica K.: Action capture with accelerometers. In: *Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* Eurographics Association, 2008, S. 193–199

[SLC04] SCHULDT, Christian ; LAPTEV, Ivan ; CAPUTO, Barbara: Recognizing Human Actions: A Local SVM Approach. In: *Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 3 - Volume 03*. Washington, DC, USA : IEEE Computer Society, 2004 (ICPR '04). – ISBN 0–7695–2128–2, 32–36

[SLQ+14] SON, Donghee ; LEE, Jongha ; QIAO, Shutao ; GHAFFARI, Roozbeh ; KIM, Jaemin ; LEE, Ji E. ; SONG, Changyeong ; KIM, Seok J. ; LEE, Dong J. ; JUN, Samuel W. u. a.: Multifunctional wearable devices for diagnosis and therapy of movement disorders. In: *Nature nanotechnology* 9 (2014), Nr. 5, S. 397–404

[SMA+13] SCHNEIDER, Oliver S. ; MACLEAN, Karon E. ; ALTUN, Kerem ; KARUEI, Idin ; WU, Michael M.: Real-time Gait Classification for Persuasive Smartphone Apps: Structuring the Literature and Pushing the Limits. In: *Proceedings of the 2013 International Conference on Intelligent User Interfaces*. New York, NY, USA : ACM, 2013 (IUI '13). – ISBN 978–1–4503–1965–2, 161–172

[SPZ13] STEWART, D. ; PASS, A. ; ZHANG, J.: Gender classification via lips: static and dynamic features. In: *IET Biometrics* 2 (2013)

[SRV10] SCHEPERS, H. M. ; ROETENBERG, Daniel ; VELTINK, Peter H.: Ambulatory human motion tracking by fusion of inertial and magnetic sensing with adaptive actuation. In: *Med Biol Eng Comput* 48 (2010), Nr. 1, S. 27–37

[SS03] SÄRNDAL, C.E. ; SWENSSON, B.: *Model Assisted Survey Sampling*. Springer, 2003. – 100–101 S.

[SWVG02] SIMMONS, Maryann ; WILHELMS, Jane ; VAN GELDER, Allen: Model-based reconstruction for creature animation. In: *Proceedings of the 2002 ACM SIGGRAPH/Eurographics symposium on Computer animation* ACM, 2002, S. 139–146

[SY07] SCHAEFER, Scott ; YUKSEL, Can: Example-based skeleton extraction. In: *Symposium on Geometry Processing* Citeseer, 2007, S. 153–162

[TEM13] TSCHARNER, Vinzenz von ; ENDERS, Hendrik ; MAURER, Christian: Subspace Identification and Classification of Healthy Human Gait. In: *PLoS ONE* 8 (2013), 07, Nr. 7, e65063. `http://dx.doi.org/10.1371/journal.pone.0065063`. – DOI 10.1371/journal.pone.0065063

[TLZF12] TAO, Weijun ; LIU, Tao ; ZHENG, Rencheng ; FENG, Hutian: Gait Analysis Using Wearable Sensors. In: *Sensors* 12 (2012), Nr. 2, 2255–2283. `http://dx.doi.org/10.3390/s120202255`. – DOI 10.3390/s120202255. – ISSN 1424–8220

[Tro02] TROJE, Nikolaus F.: Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. In: *Journal of vision* 2 (2002), Nr. 5, S. 2

[TZK+11] TAUTGES, Jochen ; ZINKE, Arno ; KRÜGER, Björn ; BAUMANN, Jan ; WEBER, Andreas ; HELTEN, Thomas ; MÜLLER, Meinard ; SEIDEL, Hans-Peter ; EBERHARDT, Bernd: Motion reconstruction using sparse accelerometer data. In: *ACM Transactions on Graphics (TOG)* 30 (2011), Nr. 3, S. 18

[Umb10] UMBAUGH, Scott E.: *Digital image processing and analysis: human and computer vision applications with CVIPtools*. CRC press, 2010. – 373–376 S.

[VAN08] VENTURE, Gentiane ; AYUSAWA, Ko ; NAKAMURA, Yoshihiko: Motion capture based identification of the human body inertial parameters. In: *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, 2008, 4575–4578

[VAV+07] VLASIC, Daniel ; ADELSBERGER, Rolf ; VANNUCCI, Giovanni ; BARNWELL, John ; GROSS, Markus ; MATUSIK, Wojciech ; POPOVIĆ, Jovan: Practical motion capture in everyday surroundings. In: *in Proc. SIGGRAPH 2007, ACM*, 2007

[Vic14] VICON: *Motion Capture Systems from Vicon*. `http://www.vicon.com`, 2014. – Accessed May 11th, 2014

[VKK14] VÖGELE, Anna ; KRÜGER, Björn ; KLEIN, Reinhard: Efficient Unsupervised Temporal Segmentation of Human Motion. In: *2014 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, 2014

[WF02] WELCH, Greg ; FOXLIN, Eric: Motion Tracking: No Silver Bullet, but a Respectable Arsenal. In: *IEEE Comput. Graph. Appl.* 22 (2002), November, Nr. 6, 24–38. `http://dx.doi.org/10.1109/MCG.2002.1046626`. – DOI 10.1109/MCG.2002.1046626. – ISSN 0272–1716

[Wik14] WIKIPEDIA: *Motion capture*. `http://en.wikipedia.org/wiki/Motion_capture`, 2014. – Accessed May 11th, 2014

[WLZ+11] WANG, Pengjie ; LAU, Rynson W. ; ZHANG, Mingmin ; WANG, Jiang ; SONG, Haiyu ; PAN, Zhigeng: A real-time database architecture for motion capture data. In: *Proceedings of the 19th ACM international conference on Multimedia* ACM, 2011, S. 1337–1340

[WP02] WADE, Lawson ; PARENT, Richard E.: Automated generation of control skeletons for use in animation. In: *The Visual Computer* 18 (2002), Nr. 2, S. 97–110

[WVG97] WILHELMS, Jane ; VAN GELDER, Allen: Anatomically based modeling. In: *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* ACM Press/Addison-Wesley Publishing Co., 1997, S. 173–180

[WVZ+15] WILHELM, Nils ; VÖGELE, Anna ; ZSOLDOS, Rebeka ; LICKA, Theresia ; KRÜGER, Björn ; BERNARD, Jürgen: FuryExplorer: Visual-Interactive Exploration of Horse Motion Capture Data. In: *Visualization and Data Analysis (VDA 2015)*, 2015

[Xse14] XSENS: *Home — Xsens 3D motion tracking*. `http://www.xsens.com`, 2014. – Accessed May 11th, 2014

[XZY+07] XU, Weiwei ; ZHOU, Kun ; YU, Yizhou ; TAN, Qifeng ; PENG, Qunsheng ; GUO, Baining: Gradient domain editing of deforming mesh sequences. In: *ACM Transactions on Graphics (TOG)* Bd. 26 ACM, 2007, S. 84

[YP03] YIN, KangKang ; PAI, Dinesh K.: FootSee: An Interactive Animation System. In: *Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Aire-la-Ville, Switzerland, Switzerland : Eurographics Association, 2003 (SCA '03). – ISBN 1–58113–659–5, 329–338

[YYPM11] YANG, Yong-Liang ; YANG, Yi-Jun ; POTTMANN, Helmut ; MITRA, Niloy J.: Shape space exploration of constrained meshes. In: *ACM Trans. Graph.* 30 (2011), Nr. 6, S. 124

[Zij04] ZIJLSTRA, Wiebren: Assessment of spatio-temporal parameters during unconstrained walking. In: *European Journal of Applied Physiology* 92 (2004), Nr. 1-2, 39-44. `http://dx.doi.org/10.1007/s00421-004-1041-5`. – DOI 10.1007/s00421–004–1041–5. – ISSN 1439–6319

[ZTH13] ZHOU, Feng ; TORRE, F. D. ; HODGINS, J. K.: Hierarchical Aligned Cluster Analysis for Temporal Clustering of Human Motion.

In: *IEEE Trans. on Pattern Analysis and Machine Intelligence* 35 (2013), Nr. 3, S. 582–596. http://dx.doi.org/http://doi.ieeecomputersociety.org/10.1109/TPAMI.2012.137. – DOI http://doi.ieeecomputersociety.org/10.1109/TPAMI.2012.137. – ISSN 0162–8828