

Unifying Representations and Large-Scale Whole-Body Motion Databases for Studying Human Motion

Christian Mandery, Ömer Terlemez, Martin Do, Nikolaus Vahrenkamp, and Tamim Asfour, *Member, IEEE*

Abstract—Large-scale human motion databases are key for research questions ranging from human motion analysis and synthesis, biomechanics of human motion, data-driven learning of motion primitives, and rehabilitation robotics to the design of humanoid robots and wearable robots such as exoskeletons. In this paper we present a large-scale database of whole-body human motion with methods and tools, which allows a unifying representation of captured human motion, and efficient search in the database, as well as the transfer of subject-specific motions to robots with different embodiments. To this end, captured subject-specific motion is normalized regarding the subject’s height and weight by using a reference kinematics and dynamics model of the human body, the master motor map (MMM). In contrast with previous approaches and human motion databases, the motion data in our database consider not only the motions of the human subject but the position and motion of objects with which the subject is interacting as well. In addition to the description of the MMM reference model, we present procedures and techniques for the systematic recording, labeling, and organization of human motion capture data, object motions as well as the subject–object relations. To allow efficient search for certain motion types in the database, motion recordings are manually annotated with motion description tags organized in a tree structure. We demonstrate the transfer of human motion to humanoid robots and provide several examples of motion analysis using the database.

Index Terms—Humanoid robots, models of the human body, whole-body human motion databases.

I. INTRODUCTION

UNDERSTANDING human motion and its transfer to robots represents a promising way toward intuitive programming of robot systems with different body morphologies. Such understanding can only be gained by observing humans performing actions, when each action can yield a variety of different motions depending on the given situation, the constraints of the current task, the involved objects, and the properties of

Manuscript received August 01, 2015; revised February 18, 2016; accepted May 12, 2016. This paper was recommended for publication by Editor A. Kheddar and Guest Editors D. Kulic and E. Demircan upon evaluation of the reviewers’ comments. This work was supported by the European Union Seventh Framework Programme and the European Union H2020 Programme under Grant 611909 (KoroiBot), Grant 643666 (I-SUPPORT), and Grant 270273 (Xperience). (Christian Mandery and Ömer Terlemez contributed equally to this work.)

The authors are with the Institute for Anthropomatics and Robotics, High Performance Humanoid Technologies Lab (H²T), Karlsruhe Institute of Technology, Karlsruhe 76131, Germany (e-mail: mandery@kit.edu; terlemez@kit.edu; martin.do@kit.edu; vahrenkamp@kit.edu; asfour@kit.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRO.2016.2572685

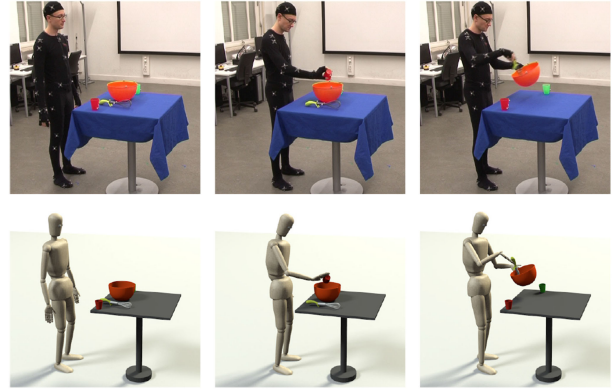


Fig. 1. Preparing the dough: Key frames of a manipulation task available in the motion database that includes four environmental objects. The MMM framework is used to convert the raw capture data to the unifying MMM reference format.

the human subject. Thus, it is important to collect sufficient amounts of motion data consisting of multiple demonstrations of actions performed by different subjects and under different conditions.

For this purpose, a great amount of research efforts have been dedicated to the field of human motion capture, leading to commercial systems that feature an outstanding performance in terms of fast and precise tracking of human motions. The currently available systems differ in capture techniques (e.g., stereo-vision or marker-based approaches with LEDs or reflecting markers) and in the resulting outcome in terms of data formats and human motion modeling. Furthermore, a large number of approaches for action and activity recognition exist, expecting input data specific to their own internal representation. Finally, any target platform for the reproduction of human motion, e.g., visualization models for animation purposes or motion transfer to humanoid robots, expects human motion capture data in terms of its own kinematic model.

In order to unify the representation of human motion capture data, we proposed the master motor map (MMM) approach in previous work (see [16]–[18]). With the MMM framework, it is possible to map and consistently represent human motions coming from varying motion capture systems to the unifying MMM reference model.

The unifying representation of human motion decouples the capture process from the motion analysis and reproduction process and hence leads to a consistent way of storing human motion capture data originating from varying sources. We established

the Karlsruhe Institute of Technology (KIT) whole-body human motion database [19] as a rich motion data corpus that not only focuses on mere whole-body motions but whole-body actions as well. For that reason, the motion data in the database considers human as well as object motions as exemplified in Fig. 1. The raw motion data entries are enriched with additional descriptions and labels. Beside the captured motion in its raw format (e.g., marker motions of the capture system), information about the subject anthropometric measurements and the setup of the scene including environmental elements and objects are provided. The motions are annotated with motion description tags that allow efficient search for certain motion types through structured queries.

The paper is organized as follows. Section II provides an overview of related work in the area of representing human motion data, human motion databases, and modeling the human body. In Section III an overview on the MMM framework, including the MMM reference model and the conversion framework, is given. Section IV describes the motion capture process with details on the motion capture system and the experimental setup. The large-scale human motion database is described in Section V. The motion transfer process to the MMM format is explained in Section VI using a marker-based motion capture setup. In this section we also show how the reproduction of MMM motions on a humanoid robot can be realized and we discuss some exemplary motion analysis approaches that have been realized with the MMM framework. The content of the database is discussed by means of a quantitative analysis and several exemplary motions in Section VII. The work is summarized and notes on future work are given in Section VIII.

This paper builds upon our previous work presented in [18] and [19] and extends it by showing how human motions are transferred to humanoid robots like ARMAR-4 [20] and NAO [21]. It also presents example applications of our framework for the detection of human whole-body support contacts and the calculation of motion dynamics. In addition, we show how additional sensors, e.g., force and IMU measurements, can be used to enrich the captured motion data and we provide an extended description of the content of our motion database with regard to motions involving objects and multisubject motions.

II. RELATED WORK

Numerous motion databases for different purposes such as motion analysis, recognition, classification, animation, and synthesis have been introduced in recent years. In many cases such as in [8], [13], [14], [22], and [23], the acquisition of motions and the maintenance of the corresponding databases have been discontinued after a short time. Thus, only a few approaches have resulted in large-scale motion databases. One of the largest and most prominent motion databases is the Carnegie Mellon University (CMU) Graphics Laboratory Motion Capture Database [1]. With a focus on human character animation and simulation in computer graphics, the database contains more than 2600 recordings with 144 different human subjects performing a variety of motions. The motion data is freely accessible. In addition to the data, tools are provided that allow for viewing

and editing the files. However, inconsistencies in the motion descriptions and the simple structure and organization of the data prevent the uniform processing and analysis of the motions. In [15], a database is introduced that contains partially commercial and free motion capture recordings. This collection comprises around 5500 motion recordings. The focus of this data collection has been set on animation purposes (e.g., game development); thus, the motion data features mostly dance and martial arts actions. The motion files are very coarsely structured and essential information about the subjects and the capturing procedure, e.g., the marker placement is missing, which makes this data challenging to use for research purposes. The HDM05 motion database has been created for motion analysis, synthesis, and classification [2]. It provides around 50 min of motion data, stored in 1457 motion clips in a limited number of roughly 100 motion classes. These motion clips are created by the manual slicing of motion recordings. Compared with the CMU database, HDM05 uses a more stringent structuring of the data. However, only five subjects have been captured for the recordings and some of the motion classes contain only data from even fewer subjects. The human motion database in [3] uses a controlled sampling approach to determine the structure of the database and the motion data to be collected in order to provide a dataset that facilitates the quantitative evaluation of algorithms addressing motion cognition problems. Thus, the database consists of five different datasets: the praxicon dataset with recordings of about 350 actions from a single subject, a cross-validation dataset that describes a subset of 70 actions for 50 different subjects, a generalization dataset that contains recordings with varying motion parameters, and two additional datasets providing recordings for composed motions and interaction between two subjects. The database in [4] has been built for the study of gender, identity, and emotional affect from biological motion. It contains a total number of 4080 recordings, consisting of a small set of actions recorded from 30 subjects playing four different emotional states (neutral, angry, happy, and sad). The Edinburgh Computer Graphics and Visualization (CGVU) Interaction Database [5] provides recordings of numerous different manipulation and object interaction tasks that have been captured using magnetic motion capture and RGB-D sensors as described in [24], which solves the problem of occlusions that frequently arises when purely visual approaches are being used. Similar to our approach, objects are also tracked in this dataset and object mesh models are available. The HuMoD Database presented in [6] provides motion capture data and anthropometric measurements with a focus on lower limbs for the study of human motion dynamics with muscle-driven actuation. Two subjects (one male, one female) performed a total number of 13 repetitive motions like walking, running, or squats, respectively, while a given period of time was assigned for each motion. Focusing on the evaluation of vision-based human pose recognition methods, the HumanEva Database [9] has been created to provide appropriate datasets. It comprises six actions recorded from four subjects wearing loose clothing, which leads to large inaccuracies and uncertainties in the resulting motion data. The data is not freely accessible. Human3.6M [10] provides another large-scale dataset for the evaluation of

TABLE I
SELECTION OF LARGE-SCALE MOTION DATABASES AVAILABLE TO THE SCIENTIFIC COMMUNITY, INCLUDING OUR DATABASE

Database Name		Description	Motion Types	Data Format	# subj.	# rec.
CMU Graphics Lab Motion Capture Database	[1]	Very well known and commonly used whole-body motion database	Mainly locomotion tasks, motions showing human interactions, and motions of physical activities and sports	C3D, ASF/AMC, Videos	109	2605
HDM05	[2]	Approx. 50 min recording of whole-body motion (approx. 3 h before manual segmentation)	Mainly locomotion and sports motions	C3D, ASF/AMC, Videos (some)	5	1457
Human Motion Database	[3]	Systematic sampling methodology to determine the motion data collected, whole-body motion	Five different datasets (see text) in the praxicon dataset: 350 different actions demonstrated by one actor for “concrete” verbs extracted from a lexical database	N/A	five datasets (see text)	
DB for study of gender, identity and emotion	[4]	Small set of motions played for four different emotional states (neutral, angry, happy, sad)	Walking, arm movements, and sequences of arm movements separated by walking	CSM (3ds Max Character Studio format), PTD	30	4080
Edinburgh CGVU Interaction Database	[5]	Whole-body motion captured using magnetic motion capture and RGB-D sensors	Numerous different manipulation and object interaction motions	FBX, BVH	N/A	N/A
HuMoD Database	[6]	Motion capture with focus on lower body (few marker on upper body), measurements from EMG and two force plates, anthropometric measurements	Walking and running at different speeds, avoiding an obstacle, squats, kicks, and jumps	MAT (MATLAB format)	2	26
HMDB51	[7]	Video database to evaluate action recognition algorithms (no MoCap), clips extracted from different sources (e.g., YouTube) and manually annotated	51 defined actions, such as chew, clap, golf, or kick	Videos	N/A	6766
CMU Motion of Body (MoBo) Database	[8]	11 s long clips of human walking on a treadmill recorded with six cameras at 30 FPS	Walking in four different styles: slow, fast, inclined, and with a ball	Images (PPM/JPG)	25	100
HumanEva Dataset	[9]	Dataset for the evaluation of tracking and human pose estimation algorithms	Six common actions: Walk, jog, throw/catch, gesture, box, and combo	N/A	4	56
Human3.6M	[10]	Combines high-resolution camera recordings with motion capture data and time-of-flight data for the evaluation of human pose estimation methods	15 different poses/motions, e.g., walking, waiting, posing, and sitting	N/A	11	N/A
IEMOCAP	[11]	12 h of data for emotion recognition, motion capture markers only attached to face, head, and hands, includes speech and transcriptions	Scripted and improvised scenes played in ten different emotional states	N/A	10	N/A
Human Identification at a Distance Database	[12]	Recordings of walking subjects to study the identification of subjects from gaits	Walking motions	MB (Maya), MOT (text), Videos	20	N/A
Korea University Gesture (KUG) Database	[13]	Whole-body motion capture recordings of human gestures	54 different gestures, classified into normal, abnormal, and command gestures	HTR, Videos	20	N/A
NUS Motion Capture Database	[14]	Whole-body human motion capture recordings by the National University of Singapore	Mainly locomotion tasks, and motions of dance and martial arts	C3D, BVH, FBX, and more	8	N/A
Mocapdata.com	[15]	Operated by Eyes, JAPAN Co. Ltd., some data only available commercially	Mainly motions tailored to animation purposes, e.g., artistic motions, sports and dancing	C3D, BVH, and more (varies depending on motion)	N/A	5604
KIT Whole-Body Human Motion Database		Whole-body motion capture recordings including objects, systematic structure, and normalized motion representation	Locomotion in constrained and unconstrained environments, (loco-)manipulation, multisubject interaction, push recovery (see Section VII)	C3D, MMM, Videos, for some: accel., force	53 ^a	6701 ^a

^aAs of 2016, February 2nd (database still growing).

human pose recognition methods. It provides a total number of 3.6 million human poses from 11 professional actors that have been recorded using four high-resolution cameras. Additionally, reconstructed joint angles from a motion capture system, time-of-flight range data, and 3-D laser scans of the actors are provided. Further, databases that focus on the evaluation of vi-

sion algorithms are presented in [7], [11], [12], and [25], where only video camera footage of human actions has been collected. Due to the lack of motion capture data, the extraction of relevant information from such databases for motion analysis, cognition, and synthesis problems can be difficult. Table I provides a brief overview of the most important large-scale motion databases

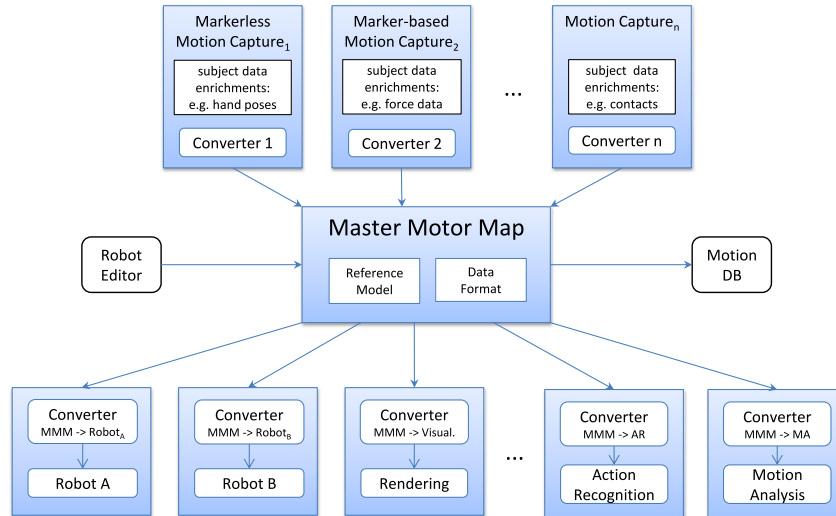


Fig. 2. MMM framework decouples the human motion capture process from motion analysis and reproduction components by providing a reference model of the human body and a unifying data format with corresponding interfaces. This allows unification of a wide variety of data sources and to make motion capture data available in a transparent way. As shown above, the MMM framework defines a reference model of the human body accompanied with an extensible data format. Any motion capture system can be supported as long as a converter is implemented that converts the raw data to the MMM format. In addition, complementary data, such as force or contact data, can be stored and processed. The MMM data can then be converted to different embodiments, such as humanoid robots. Furthermore, the MMM data can be postprocessed and analyzed, e.g., for action recognition or task specific motion conversion. Finally, the motions can be uploaded to the KIT motion database, which provides a large-scale dataset of human whole-body motions.

employed in the scientific community. For further details, [3] provides a survey of many existing motion capture databases.

In order to make use of the motion data stored in these databases, specifications and ideally tools have to be provided, which allow the transfer of human motion from the model of the captured human subject to the model that is used for further motion data processing. Solutions that are tailored for the transfer of one specific model to another have been proposed in [26]–[28], where correspondences between both models are exploited in order to optimize the original motion in such a way that it fits the target model. Various unifying representations, such as [29]–[31], which generalize the motion from the human subject by translating it to a parameterizable model, have emerged during the development of model-based human tracking algorithms. However, most of these representations are very rudimentary and do not allow the encoding of rich information about a motion. A unifying representation that is capable of acting as an intermediate model for the transfer and the comprehensive encoding of human motion has been introduced by the MMM in [16]. The MMM comprises a kinematic and dynamic reference model of the human body with a total number of 104 degrees of freedom (DoF), which can be processed by any component that uses input motion in the representation of the MMM data format. The properties of the segments and joints are inspired by [32] and [33]. Similar to existing motion analysis toolboxes such as [34]–[36], the MMM has been extended to a framework with a multitude of methods and tools that enable a wide range of processing capabilities ranging from the analysis of human motion to the reproduction on humanoid robots. In combination with the KIT whole-body human motion database, we present a comprehensive source and powerful tools for the understanding and the synthesis of human motion.

III. MASTER MOTOR MAP FRAMEWORK

With the MMM, we provide a unifying framework for representing whole-body human motion data in conjunction with tools and methods for motion analysis and reproduction. For this purpose, the MMM framework provides a reference model of the human body together with extendable data structures. As shown in Fig. 2, several interfaces offer unified access for motion data conversion. The input layer is used to transfer human motion capture data to the MMM reference model by converting the raw input to the unifying MMM data format. The data can be exported to the motion database and/or further processed via the output layer of the MMM framework. The output converters can be used to transfer the whole-body human motions to different embodiments, such as humanoid robots, as shown in Section VI-C. In addition, motions can be processed in order to apply filters that may be used to improve the quality or to apply task-specific constraints.

A. Master Motor Map Reference Model

The MMM reference model consists of a rigid body system with a well-defined kinematic configuration enriched with dynamic properties. The kinematic and dynamic specifications of the whole-body reference model are based on biomechanical studies, such as [32], [33], [37], [38], and [39]. The hand specifications are derived from the analysis reported in [40] and [41].

The reference model is provided in a normalized form regarding size and weight. In order to scale the model according to subject’s height, total weight, and potentially available custom segment lengths, the MMM framework provides several ready-to-use model processor components. The model processor

system allows us to customize the model adaption by parameterizing the provided model processors or by implementing custom model processors, e.g., when precise dynamic measurements are available using an approach such as [42].

Although the model provides a detailed representation of human kinematics and dynamics, the full complexity might not be needed for all applications. For example, when studying walking motions, it is unlikely that the movements of the fingers are of interest. Hence, the MMM framework allows the specification of an active set of joints that are considered for analysis, conversion, or transfer of motions. Additionally, representable whole-body motions are limited to healthy subjects with a normal range of motion. Due to the approach of generalization, motions performed by people with extremely deviant relative segment lengths, extreme physical flexibility, or missing limbs are not provided for. While adaptations to the reference model can be made to cover some of these cases, they invalidate the generalization approach and are therefore advised only in individual cases.

1) *Kinematic Model*: The kinematics of the MMM reference model consists of 104 DoF: 6 DoF cover the model pose, 23 DoF are assigned to each hand, and the remaining 52 DoF are distributed on arms, legs, head, eyes, and body. The reference model consists of 55 scalable segments that represent the human body. All joint definitions incorporate a specification of lower and upper limits. More details on reference frames, joint limits, and segment dimensions can be found in [18].

The human spine is modeled using three segments, since it has been shown in [39] that such a three segment approach provides an appropriate approximation for marker-based human motion capture and that a 3-DoF substitutional joint is sufficient to describe the kinematic relationship between the pelvis and upper torso.

The model includes two hands with 23 DoF each (see Fig. 4), where the thumb is modeled with a total of 5 DoF, starting with two at the carpometacarpal (CMC) and metacarpophalangeal (MCP) joints each and 1 DoF on the interphalangeal (IP) joint. The index finger and the middle finger are modeled with 4 DoF; 2 DoF are located at the MCP joint and one each on the proximal interphalangeal (PIP) and distal interphalangeal (DIP) joints. The ring finger and little finger are extended with an additional DoF at the CMC joint to enable better hand closure. A description of the joint locations of the left hand can be found in Table II. The joint locations of the right hand follow from symmetry.

Literature shows that a four-segment foot consisting of tibia and fibula, hindfoot, midfoot, and hallux is sufficient to map human foot motions accurately [43]. We therefore added 2 DoFs (see Fig. 3) to the foot model, which connect the hindfoot with the midfoot (LMrot) and the midfoot with the hallux (LFX).

2) *Dynamic Model*: Since obtaining dynamic properties, such as the center of mass (CoM) or the inertia tensor, from living subjects is difficult, we rely on reference values from literature that provide statistical analysis and mean values. Several works such as [44], where human segment parameters are learned from motion capture and force plate recordings, tend to confirm this approach. Each segment of the MMM refer-

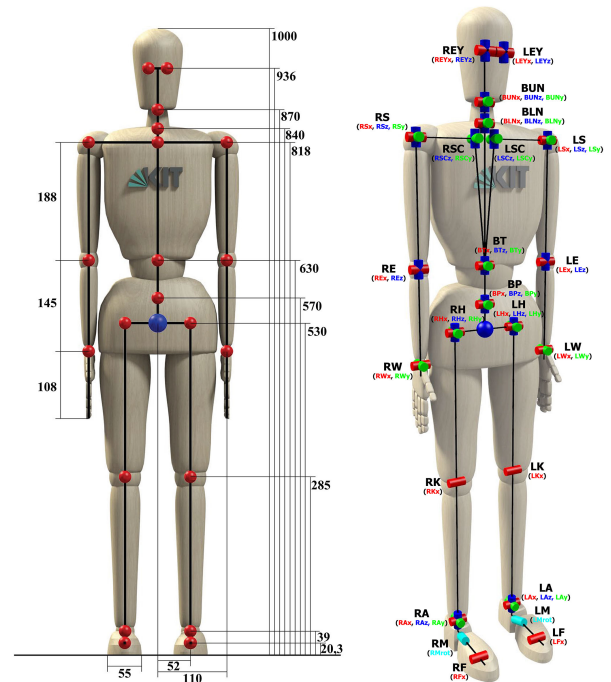


Fig. 3. Kinematics of the MMM reference model. Segment lengths (left side) are in one thousands of the total body height. The rotational order (right side) is given in brackets starting from the root node (blue sphere).

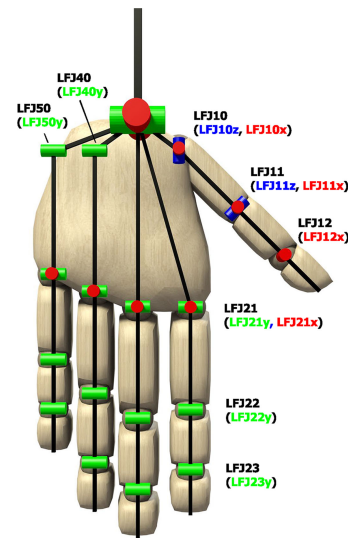


Fig. 4. MMM model of the left hand. Each hand is modeled with 23 DoF, where the thumb consists of 5 DoF and the remaining fingers of 4 DoF. Two additional DoFs at the CMC-joint of the ring-finger and little finger enable a better fist closure. The configuration of the right hand follows from symmetry. Joint names for the thumb and index finger are depicted. The names for the remaining fingers follow from the anatomical index for the fingers (3 for middle finger, 4 for ring-finger, and 5 for little finger).

ence model contains dynamic properties such as the position of the CoM of the segment, its mass, and its inertia tensor. While the position of the CoM and the mass can be linearly scaled according to given subject parameters, the inertia tensor has to be recalculated. This can be done either by deriving it from the appropriate radius of gyration values [17] or, as realized by the MMM framework, by applying the parallel-axis theorem [32].

TABLE II
JOINT LOCATIONS OF THE LEFT HAND IN RELATION TO THE WRIST JOINT, WITH THE Z-AXIS POINTING UPWARD AND THE Y-AXIS POINTING TOWARD THE THUMB

Joint (index)	Finger (index)									
	Thumb (1)		Index Finger (2)		Middle Finger (3)		Ring Finger (4)		Pinky (5)	
	Y	Z	Y	Z	Y	Z	Y	Z	Y	Z
CMC (0)	-0.009740	-0.007884	—	—	—	—	0.010234	-0.007884	0.019971	-0.007884
MCP (1)	-0.009740	-0.034992	-0.012470	-0.048276	0.000000	-0.048168	0.010234	-0.044172	0.019971	-0.039744
PIP/IP (2)	-0.009740	-0.056160	-0.012470	-0.074736	0.000000	-0.076896	0.010234	-0.070524	0.019971	-0.061776
DIP (3)	—	—	-0.012470	-0.090180	0.000000	-0.095256	0.010234	-0.088344	0.019971	-0.074412
tip	-0.009740	-0.073224	-0.012470	-0.100656	0.000000	-0.106920	0.010234	-0.099900	0.019971	-0.084456

Distances are represented in relation to total body height.

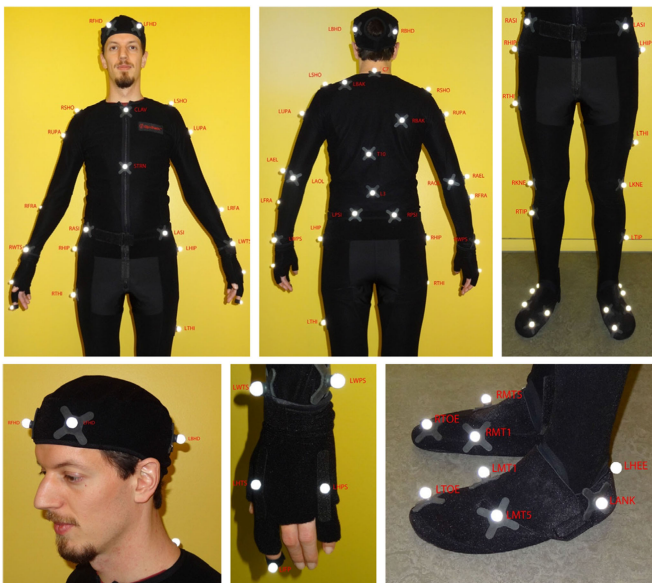


Fig. 5. Reference marker set used for whole-body human motion capture as proposed in the MMM framework.

B. Master Motor Map Software Framework

We provide a reference implementation of the MMM concept that is published under an open source license.¹ The implementation is split in two parts: The *MMMCORE* library contains the data structures and the corresponding I/O methods for robot models and motion data. The *MMMTOLS* package, based on Simox [45], provides converter implementations, extended visualization and inspection tools, as well as the MMM reference model of the human body with its kinematic and dynamic specifications. A detailed overview of the MMM software framework is provided in [18].

IV. ACQUISITION OF MOTION DATA

As shown in Fig. 2, the MMM framework supports any motion capture system as long as a converter is implemented that processes and converts the raw output data of the given capture

system to the unifying MMM data format. In the following, we show how the optical marker-based Vicon MX motion capture system [46] can be used for capturing human whole-body motion data.

A. Motion Capture Setup

The Vicon MX motion capture system is a passive marker-based recording system. Several cameras surrounding the workspace emit infrared light, which is reflected by artificial markers that can be placed on human subjects, objects, and surrounding environmental elements. Recordings of multiple cameras are compared and corresponding reflections are triangulated to create a 3-D point cloud, which is saved in the C3D file format [47], the industry standard for storing motion capture data, and made available in the motion database. Given a predefined reference marker set (see Section IV-B), the points in the 3-D point cloud can be labeled accordingly. Additionally, a time-synchronized video of the motion is recorded and also made available in the database.

B. Whole-Body Marker Set

Fig. 5 shows the reference marker set used for whole-body human motion capture. The marker set consists of 56 markers that are derived from specific anatomical landmarks of the human body. More information about the marker set including close-up pictures is available online.²

C. Object Marker Set

Since including objects in motion capture is essential for the study of human-object interaction, all the necessary tools to handle object interaction are included in the MMM framework. However, some guidelines have to be considered while building marker sets for objects.

To rule out ambiguities, marker placement on objects should not be symmetric and the lines connecting the markers in 3-D space should not be parallel. For optimal robustness, it is best to place markers in such a way that the connecting lines form orthogonal angles. At least three markers have to be visible to

¹<https://gitlab.com/mastermotormap>

²https://motion-database.humanoids.kit.edu/marker_set/



Fig. 6. Exemplary environmental elements and objects with markers attached to them for simultaneous capturing of subject and object motion.

estimate an object pose; however, estimation robustness is increased with a larger number of visible markers. For every used object, a 3-D model is created, which also incorporates the exact positions of the markers placed on the object. These models are saved in the database along with a custom object description. Fig. 6 exemplarily shows several objects for motion capture with markers attached to them. Similar to subject information, every motion in the database is tagged with all objects that are involved in the motion.

D. Additional Data

Complementing the raw motion capture data, additional data is collected to enhance the possibilities for analysis of the captured motion. In the following, we provide several examples of how such additional data streams can be processed and incorporated.

1) *Video Recordings*: For all motions, a video is recorded that is time synchronized to the motion capture data. These videos are manually postprocessed using a tool to anonymize subjects by blurring their faces. The anonymized videos are uploaded to the database and also support the user in navigating through the content.

2) *External and Internal Forces*: For the study of certain motion tasks like push recovery experiments, it is specifically interesting to consider force data recordings. We developed a wireless sensor that allows the measurement of an external force that is exerted on a human subject. In addition, it is interesting to record data from an inertial measurement unit (IMU), especially since almost every bipedal humanoid robot is equipped with an IMU. We use an Xsens MT IMU that is strapped to the chest of the human subject to record 3-axis linear accelerations and 3-axis angular speeds. Fig. 7 shows both the force sensor and the IMU as well as an exemplary plot of measurements for a push recovery motion. These measurements are time synchronized with motion capture data as well as provided through the motion database. More information about our approach to force and IMU measurement can be found in [48].

3) *Subject-Specific Data*: Anthropometric measurements are collected from all captured human subjects. For this purpose, 25 distances between certain anatomical landmarks

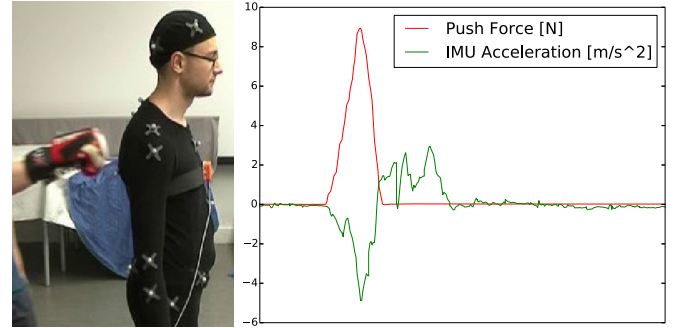


Fig. 7. Wireless force measurement device (based on a boxing glove) and inertial measurement unit (attached to the chest) for push recovery experiments, and an exemplary plot of measured data from such an experiment.

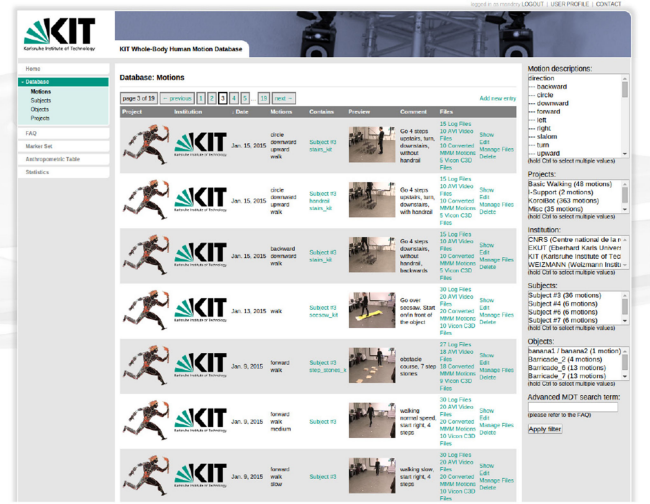


Fig. 8. Web interface of the motion database (screenshot).

are measured manually by using a tape measure. The measured properties are further described in [19] and online.³ We, however, found out that it is difficult to precisely define and reproduce these measurements, which is why we do not recommend using these data, e.g., an exact model parametrization, but rather as an initial starting point for further optimization steps.

V. MOTION DATABASE

The motion database provides two different methods of accessing its content: One the one hand, a web interface provides a convenient method to access the motion database that can be used without client-side software installation. Fig. 8 shows a screenshot of the web page used to list and search available motions. In contrast with other motion databases, like [1] that also provide a web interface, our interface allows registered users to log in and edit the database content directly from their web browser. On the other hand, the database can be accessed directly through an application programming interface (API) based on the object-oriented middleware platform Ice [49]. The API allows direct access to the motion database from several popular platforms and programming languages like C++, .NET,

³https://motion-database.humanoids.kit.edu/anthropometric_table/

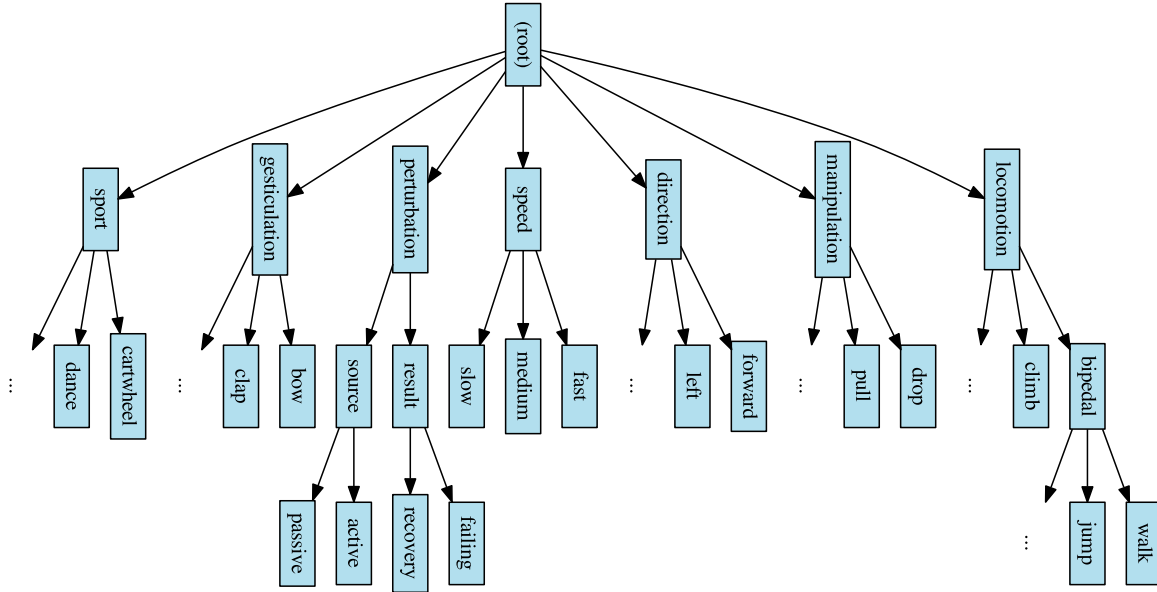


Fig. 9. Structure of the motion description tree (excerpt).

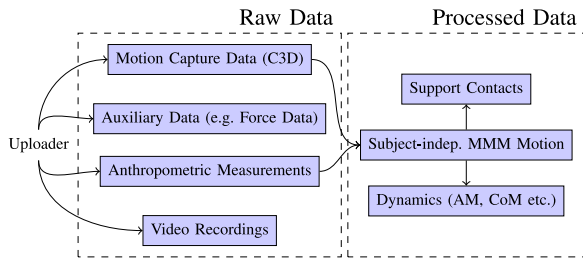


Fig. 10. Data types and data flow in the motion database.

Java, Python, or Ruby, and thus underlines our attempt to facilitate the integration of the motion database in new and existing tools and pipelines. Read access using both access methods is publicly available.

Fig. 10 shows which types of data are available in the motion database and the resulting data flow. Raw data that is uploaded by contributing users consists of motion capture data in the C3D format (see Section IV-A), video recordings, and optional auxiliary data like force or acceleration measurements (see Section IV-D). Additionally, manually measured anthropometric data of recorded subjects are entered into the database (see Section IV-D3).

From the uploaded motion capture data, a subject-independent representation of the motion based on the MMM reference model is created within the motion database. This normalization of motion data is described in Section VI-A. Based on the normalized motion and the associated kinematic and dynamic model of the human body, further properties of the motion like support contacts or trajectories of the CoM and the angular momentum can be obtained, as explained in Section VI-D.

A. Motion Organization

Recorded motions are classified within the so-called motion description tree. The motion description tree is a structure that

we developed for the classification of motions. It consists of a hierarchical declaration of tags describing motion types and additional description nodes for motion properties such as movement type, direction, speed, and stability. These motion properties are manually attached to each new motion during the motion recording sessions before entering it into the database, in contrast with approaches from the computer graphics community [50] that aim to automatically extract motion features that are useful for searching in large-scale motion databases. Fig. 9 shows an excerpt from the present structure of the motion description tree. Creation of the basic initial structure of the tree was aided by the lexical database WordNet [51] to extract hyponymy and hypernymy relations between motion-related terms in the English language. Hyponymy and hypernymy relations define a semantic relationship between words in which a hypernym is a word that has a semantic field broader than that of another word and includes this word. For example, according to WordNet, a hypernym of the verb “bow” is “gesticulation,” which has been reflected in the structure of the motion description tree, as shown in Fig. 9.

The tree itself is managed in the database and thus is extendable by the users. In contrast with a classic categorization approach, motions can be associated with an arbitrary number of nodes in the motion description tree. For example, a motion of a subject that trips while walking downstairs can be categorized using the following tree nodes:

- 1) locomotion \rightarrow bipedal \rightarrow walk
- 2) speed \rightarrow medium
- 3) direction \rightarrow forward
- 4) direction \rightarrow downward
- 5) perturbation \rightarrow result \rightarrow failing
- 6) perturbation \rightarrow source \rightarrow passive

The motion description tree serves as a flexible way to classify motions and allows for a precise classification of compound movements (e.g., “rise from chair and walk two steps”) without

introducing the ambiguities entailed by a free-text description. Motions that are associated with a specific tree node or contained in a specific subtree can be found efficiently in the database. This allows for efficient search for motions with certain properties in the database. For example, the walking motion described would be returned when searching for “locomotion \rightarrow bipedal.” These queries can also be chained using logical operators to create more elaborate search queries. For example, `run OR (walk AND NOT (slow))` returns motions where the subject is either running, with any speed, or walking, but not slow.

VI. MOTION TRANSFER

As shown in Fig. 2, motion capture data has to be converted to the unifying MMM data format for further processing. Once the motions are represented in the MMM data format, they can be processed by the MMM framework and transferred to different embodiments, e.g., humanoid robots. For the whole conversion process (to/from the MMM reference model), the MMM converter framework provides several ready-to-use components that can be applied for motion conversion (see [18] for details). In this section, we show how the conversion from a marker-based motion capture system to the MMM model is realized. Further, we present a reproduction method that can be used to convert whole-body motions to humanoid robots and show its application on ARMAR-4 and the NAO robot.

A. Reconstruction of Human Motion

To be able to extract semantic knowledge from the recorded motions, we first need to transfer these motions to the MMM reference model, i.e., reconstruct joint angle trajectories from raw motion capture data. In the proposed setup (see Section IV-A), the motion capture data consists of marker trajectories in Cartesian space. For every motion capture marker on the human body, a corresponding virtual marker is placed on the reference model. In order to transfer motions to the MMM reference model, the reference model has to be scaled to match the size of the human subject. For this purpose, the MMM framework provides methods to uniformly scale all segments of the reference model according to the known body height. Similar to the approaches presented in [28], [36], and [52], in [53] we introduced the procedure for the motion transfer as an optimization problem, as follows.

Let $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ be an observation of the 3-D positions of the n captured markers and $\mathbf{x} = (p_x, p_y, p_z, \alpha, \beta, \gamma, \theta_1, \dots, \theta_m)$ be the vector describing the pose of the reference model, consisting of the root position and rotation of the model and its m joint angle values. Additionally, let $\mathbf{V}(\mathbf{x}) = (\mathbf{v}_1(\mathbf{x}), \dots, \mathbf{v}_n(\mathbf{x}))$ be the positions of corresponding virtual markers as determined by the forward kinematics of the model. The problem of determining the pose of the MMM reference model for a given marker observation \mathbf{U} is then solved by minimizing

$$f(\mathbf{x}) = \sum_i \|\mathbf{u}_i - \mathbf{v}_i(\mathbf{x})\|^2$$

while maintaining the box constraints for $\theta_1, \dots, \theta_m$ given by the joint limits of the reference model. For every motion frame,

TABLE III
AVERAGE DISTANCES BETWEEN OBSERVED MARKERS AND VIRTUAL MARKERS ON THE MMM REFERENCE MODEL FOR DIFFERENT TYPES OF MOTIONS

Motion type	# Rec.	Avg. all frames	Avg. worst frame
Locom. w./out env. elem.	1200	34.20 mm \pm 11.12 mm	39.98 mm \pm 20.68 mm
Locom. w./ env. elem.	404	36.03 mm \pm 4.10 mm	41.97 mm \pm 12.95 mm
Push Recovery	249	33.56 mm \pm 1.18 mm	37.64 mm \pm 5.79 mm
Loco-Manipulation	138	34.83 mm \pm 8.38 mm	41.19 mm \pm 11.81 mm
Manipulation and Misc.	199	36.88 mm \pm 2.31 mm	42.00 mm \pm 8.37 mm

this optimization problem is solved by using the reimplementa-tion of the Subplex algorithm [54] provided by the NLOpt library [55] for nonlinear optimization.

To facilitate the conversion of motions stored in the database, we developed a toolkit that automatically performs a batch conversion of new motions uploaded to the database, as described above, by using the database API described in Section V. The conversion process results in a motion of the reference model represented in the XML-based MMM data format. The motion file is stored alongside the corresponding raw C3D motion in the database for subsequent processing. For differing applications in which the normalized motion provided by the database might be not suitable, the open nature of the MMM framework also allows the implementation of custom motion converters that perform the joint angle reconstruction by other means.

Table III shows the average distances between the observed markers and the corresponding virtual markers on the MMM reference model for different types of motions available in the motion database. The third column shows the average distance across all markers of all frames of all the considered motions. The fourth column only considers the “worst” frame of each motion, i.e., the frame with the highest average marker distance.

B. Reconstruction of Object Motion

Poses of objects involved in a motion are reconstructed from object markers, similar to the reconstruction of human motions by using a joint-less 6-D pose vector.

C. Transfer to Humanoid Robots

Transferring human motion capture data to humanoid robots can be done in various ways. There are online approaches, in which kinematic transfer is addressed [18], and controller-based approaches which try to ensure that dynamic constraints are maintained [56]. In [57], a learning approach using Gaussian processes is applied in order to perform kinematic mapping on the joint level. A torque-based controller is presented in [58], which allows to consistently represent contact forces with the environment.

In the following, we present an approach for a kinematic transfer of MMM motions to humanoid robots. The starting point for such a retargeting of motions is always the MMM representation of the human motion, reconstructed using the procedure described in Section VI-A. This allows us to abstract the retargeting step from individual subject parameters and the motion acquisition process, and thus enables interchangeability of different acquisition methods within the MMM approach.

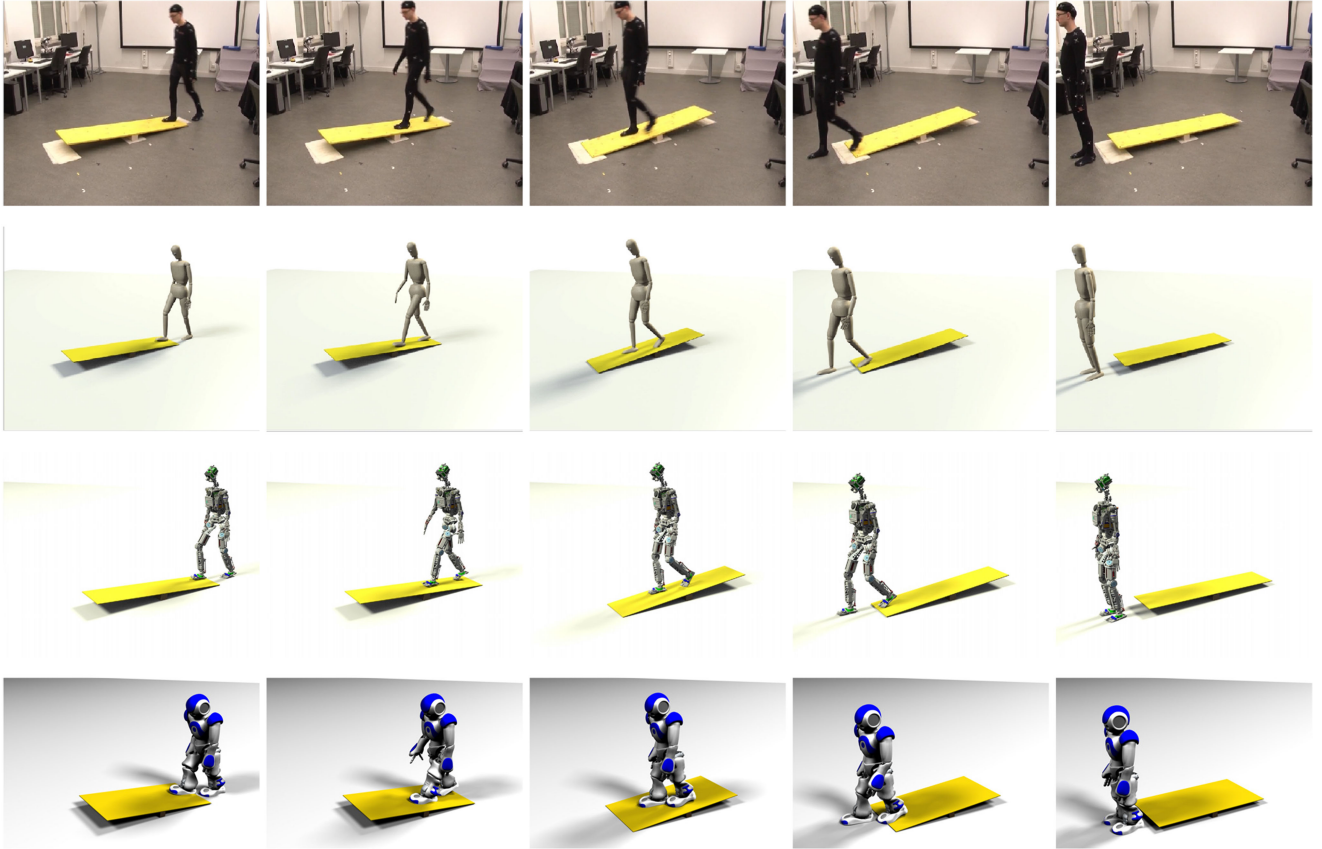


Fig. 11. Top row shows selected frames of a captured walking motion. The MMM framework is used to convert the raw capture data to the unifying MMM reference format as shown in the second row. The two bottom rows show the conversion to the humanoid robot ARMAR-4 and the NAO robot.

Further, an intermediate model for representing human motion in large-scale databases, such as the MMM reference model, provides a unique opportunity for research efforts in the area of robot programming by demonstration, since with such an intermediate model, for n motion acquisition methods and m robot morphologies, there are only $n + m$ possible conversion paths that should be tested instead of $n \times m$. Dynamic aspects, such as stability and maximum torque values, can either be considered in a second filtering step [59] or such considerations can be applied through online controllers that ensure the stability of the robot [56].

The conversion is performed similar to the approach of [26], in which virtual markers are attached to a marionette in order to convert human motion capture data to it. Therefore, the captured motion data is preprocessed to adapt the recorded motion to the morphology of the marionette. In the proposed approach, such a motion adaption step is not needed. The only assumption that is made is that the robot embodiment is similar to the human one.

In a first step, the robot model is enriched by virtual markers that are placed on the surface identically to the virtual markers on the MMM reference model. By doing this, the same kinematic conversion process as described in Section VI-A can be applied for the robot model. The Cartesian marker trajectories of the reference model are used as source and the robot model with its attached markers is used as target. It is important to note that the resulting source point cloud has to be scaled according to the size of the recorded subject and the size of the target robot

platform. Therefore, the point cloud is uniformly scaled based on the ratio between these two sizes. If the subject is interacting with environmental elements, such as a seesaw, these elements also have to be scaled accordingly.

An exemplary motion is shown in Fig. 11. The top row shows the motion recordings, the second row depicts the converted MMM motion, and the bottom rows visualize the converted motion on a model of the humanoid robot ARMAR-4 [20] and the NAO robot [21].

D. Motion Analysis

Beside the possibility of transferring motions to different embodiments, the MMM framework can be used to analyze human motions in various ways. In the following, we give some examples of several motion analysis tools that are provided by the framework.

1) *Supporting Contacts*: In recent work [53], we are using motions of the MMM reference model and environmental objects to determine contacts that are used by the human subject during motion to provide support. It should be noted that the reconstruction of human motion as well as its transfer to humanoids, as described in Sections VI-A and VI-C, use a purely kinematic approach which does not specifically handle contact information and cannot guarantee that such information is preserved during motion transfer. In contrast, related motion retargeting techniques from the computer graphics community

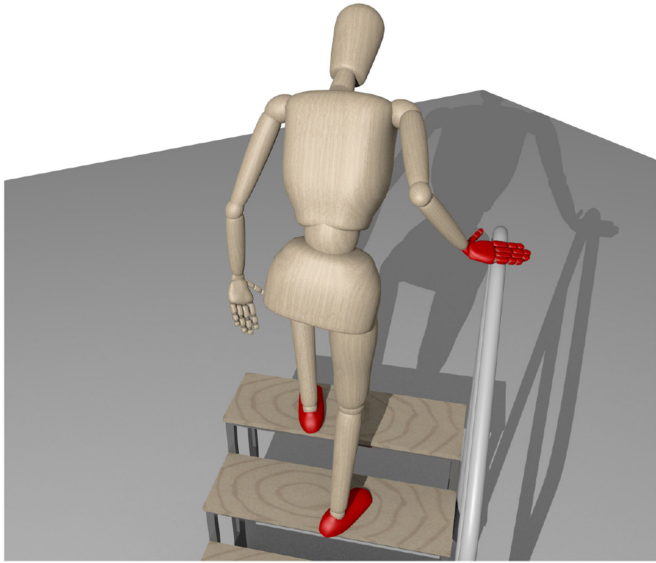


Fig. 12. Example frame showing computed support contact information between the MMM reference model and the environment (supporting segments highlighted in red).

explicitly model various motion features [60] such as spatial relationship between the human character and the environment [61], and can guarantee their preservation through the process by assigning different priorities to the modeled features of the motion [62]. Similar guarantees could be provided with our retargeting procedure by determining contact or distance information from the source data and extending the optimization procedure described in Section VI-A to maintain these additional constraints.

Since contacts are not specifically modeled in the MMM framework, our detection of support contacts is based on a heuristic that considers distances between the models of the human body and the environment computed using Simox [45] as the distances between pairs of closest points from the respective mesh model. Additionally, the speed of relevant segments of the human reference model is considered to distinguish between contacts used for manipulation and contacts used for support. The entirety of the support contacts used by the human defines a whole-body human support pose and our work provides insight in how the human transitions between these support poses during the execution of different loco-manipulation tasks. Building on the heuristics to estimate contacts from the MMM representation, we have shown in [53] that the contacts of the human subject usually can be reliably recovered from the MMM representation. In [53], the analyzed motions have been manually annotated regarding the used support poses (combinations of contacts of feet, hands, elbows, and knees with the floor and environmental elements like stairs). From these support poses, only 2.13% have been recognized incorrectly and 4.53% have not been recognized at all, which appears to be a satisfying result considering that recognizing only a single contact incorrectly leads to an incorrect support pose for that motion frame. Fig. 12 shows the support pose of the MMM reference model for an exemplary motion of a human

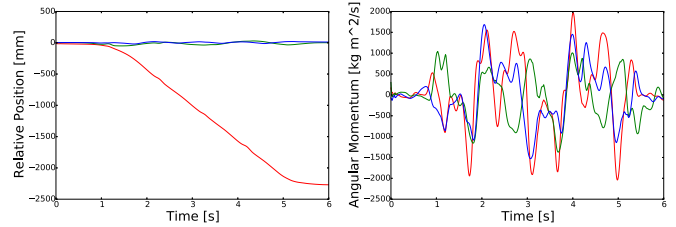


Fig. 13. Computed evolution of the CoM relative to the starting position and the whole-body angular momentum for a walking motion, shown for all three model coordinate axes (X: red, Y: green, Z: blue).

subject going upstairs. The supporting contacts are highlighted in red.

2) *Dynamic Subject Properties:* The dynamic parameters of the MMM reference model allow computing of the evolution of the CoM of the human body during motion. Additionally, the angular momentum can be computed, which has been investigated in recent years as a parameter for balancing bipedal systems, compensating disturbances, or control in general [63]–[65].

The whole-body angular momentum \mathbf{L} of a system consisting of n segments can be computed as

$$\mathbf{L} = \sum_{i=1}^n \left[m_i (\mathbf{r}_i^c \times \mathbf{v}_i^c) + I_i^c \boldsymbol{\omega}_i \right], \quad \mathbf{L} \in \mathbb{R}^3 \quad (1)$$

with

$$\mathbf{r}_i^c = \mathbf{r}_{\text{CoM}_i} - \mathbf{r}_{\text{CoM}} \quad (2)$$

$$\mathbf{v}_i^c = \dot{\mathbf{r}}_{\text{CoM}_i} - \dot{\mathbf{r}}_{\text{CoM}}. \quad (3)$$

Here, m_i denotes the mass of segment i , \mathbf{r}_i^c is its position, \mathbf{v}_i^c is its velocity, and I_i^c is its inertia tensor with respect to the CoM. $\boldsymbol{\omega}_i$ describes the angular velocity of segment i .

Fig. 13 exemplarily shows the CoM and the whole-body angular momentum at the subject's CoM for a walking motion. Since, to our knowledge, there is no definite approach to obtain highly accurate ground truth measurements for the whole-body angular momentum of a human subject, we could not perform a quantitative evaluation of the computed values. However, the shape of the trajectories shows a substantial similarity to the evolution of the whole-body angular momentum during human walking, as described in existing literature like [65].

VII. DATABASE CONTENT

As of February 2, 2016, the motion database consists of 6701 motion-capture recordings of a total run length of 17.21 h. Data from 53 different subjects (37 male, 16 female) have been entered into the database. Table IV provides an overview of the distribution of age, height, and weight for these subjects. A total of 449 motion-capture experiments, defined as a combination of subject and motion type, have been performed in our motion-capture laboratory at KIT. Additionally, 302 experiments have been performed at the Eberhard Karls University Tübingen, the Weizmann Institute of Science, and the LAAS-CNRS. In total, the database comprises 19 880 files that contain the information presented in Fig. 10 with a total size of 79.7 GiB.

TABLE IV
DISTRIBUTION OF AGE, HEIGHT, AND WEIGHT FOR THE 53 SUBJECTS
AVAILABLE IN THE MOTION DATABASE

	Mean/Std. Dev.	Minimum	Maximum
Age	26.2 ± 6.7 years	15 years	55 years
Height	1.75 ± 0.078 m	1.63 m	1.92 m
Weight	70.2 ± 10.7 kg	51 kg	93 kg

TABLE V
TYPES OF LOCOMOTION MOTIONS STORED IN THE DATABASE (EXCERPT)
AND THE CORRESPONDING NUMBER OF RECORDINGS

Motion Type	# Rec.
Walking and running with different speeds	1220
Walking a left/right bend	164
Walking paths (circle/ellipse/figure-of-8 etc.)	1241
Walking backwards	79
Turning on the spot	421
Walking up/down slopes	253
Walking up/down stairs	259
Walking over a beam	284
Walking over a seesaw	66
Walking over step stones	72
Walking on a soft mattress	29
Walking around obstacle objects	88
Walking around a moving obstacle	158
Stepping over an (imaginary) gap	107
Push recovery while walking and standing	476

A. Locomotion Motions

A major category of motions in the database is locomotion on and around objects. Table V shows some motion types that have been collected within this category that are already available in the motion database.

In addition to basic walking motions, more complex locomotion tasks have been recorded involving objects such as seesaws or stairs. The motion of these objects is also tracked as described in Section IV-C.

Examples for locomotion recording involving environmental objects can be seen in Fig. 11.

B. Manipulation Motions with Objects

Another important category of recordings shows manipulation motions that are particularly useful due to tracking of environmental objects. For example, the database provides motions of drinking, shaking, pouring, and throwing of objects. Fig. 1 shows exemplary key frames from a dough preparation task that involves four environmental objects (bowl, whisk, and two cups).

In total, the database currently contains information about 55 objects, which includes models, images, or descriptions of the object. Among these objects are 24 objects that we consider as objects used for manipulation, e.g., a whisk, a cup, or a banana. The remaining 31 objects are environmental elements such as beam, slopes, or stairs that can be used to study human locomotion in constrained environments. A total of 262 motions

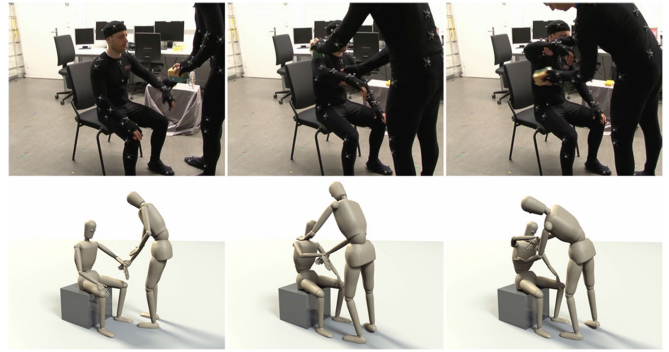


Fig. 14. Conversion of a motion with multiple subjects.

in the database are associated with at least one object and 99 motions include at least two objects.

C. Multisubject Motions

An exemplary motion with two subjects is shown in Fig. 14. For this recording, a washing action was performed and captured. Both subjects were tracked and their motions were converted into two individual MMM reference models. Since MMM data format supports multiple subjects and objects, the MMM framework with its tools can be transparently used for such multisubject recordings.

D. Motions With Force Recordings

Fig. 7 illustrates the measured force and velocity data during a push recovery experiment. The MMM motion is enriched with force and velocity data coming from the force measurement device and the IMU, as described in Section IV-D. The data are analyzed in order to determine the acceleration of the subject and the step length that was applied to capture from the experienced disturbance. More details can be found in [48].

VIII. CONCLUSION

In this paper, we summarized our work on unifying representations of human whole-body motion with the MMM framework and on the KIT Whole-Body Human Motion Database. With the MMM framework available under an open source license, we aim to provide a solid base for human motion studies, for which large datasets of motions are needed in a unified representation that decouples motion recordings from the data analysis and reproduction process. We provide a 104-DoF reference model of the human body that is based on kinematic and dynamic specifications of the human body from biomechanical studies. Additionally, the MMM framework provides well-defined interfaces for data storage and the conversion process, in which several ready-to-use components are used to realize the whole motion transfer pipeline covering motion capturing, motion transfer to the reference model, and finally reproduction of motions on humanoid robots. Several features of the framework, such as processing of objects, dealing with multisubject motions and motion analysis, have been presented and with the support pose

detection used in [53] and the estimation of dynamic parameters for human motion, we showed two exemplary applications of the MMM framework.

Furthermore, we discussed the structure and the content of the KIT Whole-Body Human Motion database, which provides a large-scale repository of human whole-body motion that is publicly available and actively maintained and extended by several contributors. We described the procedures for a systematic recording of motion data with associated complementary data such as video recordings and additional sensor measurements, as well as environmental elements and objects. The availability of highly accurate object trajectories together with the associated object mesh models makes the data especially useful for the analysis of manipulation, locomotion, and loco-manipulation tasks. Additionally, we presented an approach for labeling the content of the database using a motion description tree in such a way that efficient search for certain motion types is readily possible. Access to the database is possible through both a convenient web interface and an API. Since the database is integrated with the MMM framework, in addition to the raw data we can provide a unified representation based on the MMM for all motions. This representation can readily be used with the tools provided by the MMM for visualization and analysis.

It is our hope that the unifying motion representation of our MMM framework and our motion database with its features will provide a valuable contribution not only to our own work in the area of learning from human observation and humanoid robotics, but to a wide community as well.

REFERENCES

- [1] CMU Graphics Lab Motion Capture Database, [Online]. Available: <http://mocap.cs.cmu.edu/>. Accessed on: Feb. 4, 2016.
- [2] M. Müller, T. Röder, M. Clausen, B. Eberhardt, B. Krüger, and A. Weber, "Documentation mocap database hdm05," Univ. Bonn, Bonn, Germany, Tech. Rep. CG-2007-2, 2007.
- [3] G. Guerra-Filho and A. Biswas, "The human motion database: A cognitive and parametric sampling of human motion," *Image Vision Comput.*, vol. 30, no. 3, pp. 251–261, 2012.
- [4] Y. Ma, H. M. Paterson, and F. E. Pollick, "A motion capture library for the study of identity, gender, and emotion perception from biological motion," *Behavior Res. Methods*, vol. 38, no. 1, pp. 134–141, 2006.
- [5] Edinburgh CGVU Interaction Database, [Online]. Available: <http://www.ipab.inf.ed.ac.uk/cgvu/InteractionDatabase/interactiondb.html>. Accessed on: Feb. 4, 2016.
- [6] J. Wojtusch and O. von Stryk, "HuMoD—A versatile and open database for the investigation, modeling and simulation of human motion dynamics on actuation level," in *Proc. IEEE/RAS Int. Conf. Humanoid Robots*, 2015, pp. 74–79.
- [7] H. Kuehne, H. Jhuang, E. Garrote, T. Poggio, and T. Serre, "HMDB: A large video database for human motion recognition," in *Proc. IEEE Int. Conf. Comput. Vision*, 2011, pp. 2556–2563.
- [8] R. Gross and J. Shi, "The CMU motion of body (MoBo) database," Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-RI-TR-01-18, 2001.
- [9] L. Sigal, A. O. Balan, and M. J. Black, "HumanEva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion," *Int. J. Comput. Vision*, vol. 87, no. 1–2, pp. 4–27, 2010.
- [10] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6M: Large scale datasets and predictive methods for 3d human sensing in natural environments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 7, pp. 1325–1339, Jul. 2014.
- [11] C. Busso *et al.*, "IEMOCAP: interactive emotional dyadic motion capture database," *Language Resources Eval.*, vol. 42, no. 4, pp. 335–359, 2008.
- [12] "Human Identification at a Distance," [Online]. Available: <http://www.cc.gatech.edu/cpl/projects/hid/>. Accessed on: Feb. 4, 2016.
- [13] B.-W. Hwang, S. Kim, and S.-W. Lee, "A full-body gesture database for automatic gesture recognition," in *Proc. Int. Conf. Autom. Face Gesture Recog.*, 2006, pp. 243–248.
- [14] "NUS Mocap," [Online]. Available: <http://animation.comp.nus.edu.sg/nusmocap.html>. Accessed on: Feb. 4, 2016.
- [15] "Mocapdata.com," [Online]. Available: <http://www.mocapdata.com/>. Accessed on: Feb. 4, 2016.
- [16] P. Azad, T. Asfour, and R. Dillmann, "Toward an unified representation for imitation of human motion on humanoids," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2007, pp. 2558–2563.
- [17] S. Gärtner, M. Do, C. Simonidis, T. Asfour, W. Seemann, and R. Dillmann, "Generation of human-like motion for humanoid robots based on marker-based motion capture data," in *Proc. 41th Int. Symp. Robot.*, 2010, pp. 1–8.
- [18] O. Terlemez, S. Ulbrich, C. Mandery, M. Do, N. Vahrenkamp, and T. Asfour, "Master motor map—A toolkit for humanoid robots to capture, represent, and reproduce human motions," in *Proc. IEEE/RAS Int. Conf. Humanoid Robots*, 2014, pp. 894–901.
- [19] C. Mandery, O. Terlemez, M. Do, N. Vahrenkamp, and T. Asfour, "The KIT whole-body human motion database," in *Proc. Int. Conf. Adv. Robot.*, 2015, pp. 329–336.
- [20] T. Asfour *et al.*, "ARMAR-4: A 63 dof torque controlled humanoid robot," in *Proc. IEEE/RAS Int. Conf. Humanoid Robots*, 2013, pp. 390–396.
- [21] "The Nao Robot," [Online]. Available: <https://www.aldebaran.com/en/cool-robots/nao/>. Accessed on: Feb. 4, 2016.
- [22] ICS Action Database, [Online]. Available: <http://www.ics.t.u-tokyo.ac.jp/action/>. Accessed on: Feb. 4, 2016.
- [23] LAAS-CNRS Novela Motion Capture Database, [Online]. Available: <http://projects.laas.fr/gepetto/novela/noveladb/>. Accessed on: Feb. 4, 2016.
- [24] P. Sandilands, M. G. Choi, and T. Komura, "Interaction capture using magnetic sensors," *Comput. Animation Virtual Worlds*, vol. 24, no. 6, pp. 527–538, 2013.
- [25] W. Takano, J. Ishikawa, and Y. Nakamura, "Design of a large-scale motion database which enables recovery of 3D motion from 2D images," in *Proc. JSME Robot. Mechatron. Conf.*, 2012, pp. 2A1–Q04.
- [26] K. Yamane, J. Hodgins, and H. Brown, "Controlling a marionette with human motion capture data," in *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 3, 2003, pp. 3834–3841.
- [27] A. Safonova, N. Pollard, and J. K. Hodgins, "Optimizing human motion for the control of a humanoid robot," in *Proc. Int. Symp. Adaptive Motion Animals Mach.*, vol. 78, 2003, pp. 1390–1397.
- [28] T. Minato and H. Ishiguro, "Generating natural posture in an android by mapping human posture in three-dimensional position space," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2007, pp. 609–616.
- [29] D. M. Gavrila and L. S. Davis, "3-D model-based tracking of humans in action: A multi-view approach," in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, 1996, pp. 73–80.
- [30] S. X. Ju, M. J. Black, and Y. Yacoob, "Cardboard people: A parameterized model of articulated image motion," in *Proc. Int. Conf. Autom. Face Gesture Recog.*, 1996, pp. 38–44.
- [31] H. Sidenbladh, M. J. Black, and D. J. Fleet, "Stochastic tracking of 3D human figures using 2D image motion," in *Proc. Eur. Conf. Comput. Vision*. New York: Springer, 2000, pp. 702–718.
- [32] D. A. Winter, *Biomechanics and Motor Control of Human Movement*, 4th ed. Hoboken, NJ, USA: Wiley, 2009.
- [33] P. de Leva, "Adjustments to Zatsiorsky–Seluyanov's segment inertia parameters," *J. Biomechan.*, vol. 29, no. 9, pp. 1223–1230, 1996.
- [34] P.-B. Wieber, F. Billet, L. Boissieux, and R. Pissard-Gibollet, "The HUMANs toolbox, a homogenous framework for motion capture, analysis and simulation," in *Proc. Int. Symp. 3D Anal. Human Movement*, 2006.
- [35] R. Boulic *et al.*, "The HUMANOID environment for interactive animation of multiple deformable human characters," in *Proc. Comput. Graphics Forum*, vol. 14, no. 3, 1995, pp. 337–348.
- [36] S. L. Delp *et al.*, "OpenSim: Open-source software to create and analyze dynamic simulations of movement," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 11, pp. 1940–1950, Nov. 2007.
- [37] W. T. Dempster, W. C. Gabel, and W. J. L. Felts, "The anthropometry of the manual work space for the seated subject," *Am. J. Phys. Anthropology*, vol. 17, no. 4, pp. 289–317, 1959.
- [38] R. Contini, "Body segment parameters, Part II," *Artif. Limbs*, vol. 16, no. 1, pp. 1–19, 1972.

- [39] C. Simonidis, G. Stelzner, and W. Seemann, "A kinematic study of human torso motion," in *Proc. ASME Int. Design Eng. Tech. Conf. Comput. Inform. Eng. Conf.*, vol. 8, 2007, pp. 773–780.
- [40] B. Buchholz, T. J. Armstrong, and S. A. Goldstein, "Anthropometric data for describing the kinematics of the human hand," *Ergonomics*, vol. 35, no. 3, pp. 261–273, 1992.
- [41] B. León, A. Morales, and J. Sancho-Bru, *From Robot to Human Grasping Simulation*. (ser. Cognitive Systems Monographs). New York, NY, USA: Springer, 2014, vol. 19.
- [42] G. Venture, K. Ayusawa, and Y. Nakamura, "Motion capture based identification of the human body inertial parameters," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2008, pp. 4575–4578.
- [43] K. A. Myers, M. Wang, R. M. Marks, and G. F. Harris, "Validation of a multi-segment foot and ankle kinematic model for pediatric gait," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 12, no. 1, pp. 122–130, Mar. 2004.
- [44] G. Venture, K. Ayusawa, and Y. Nakamura, "Real-time identification and visualization of human segment parameters," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2009, pp. 3983–3986.
- [45] N. Vahrenkamp, "Simox: A robotics toolbox for simulation, motion and grasp planning," in *Proc. Int. Conf. Intell. Auton. Syst.*, 2012, pp. 585–594.
- [46] Vicon Motion Capture Systems, [Online] Available at: <http://www.vicon.com/>. Accessed on: Feb. 4, 2016.
- [47] "C3D: The 3D Biomechanics Data Standard," [Online] Available at: <http://www.c3d.org/>. Accessed on: Feb. 4, 2016.
- [48] L. Kaul and T. Asfour, "Human push-recovery: Strategy selection based on push intensity estimation," in *Proc. Int. Symp. Robot.*, 2016, pp. 547–554.
- [49] "ZeroC—The Internet Communications Engine (Ice)," [Online] Available at: <http://www.zeroc.com/ice.html>. Accessed on: Feb. 4, 2016.
- [50] M. Kapadia, I.-k. Chiang, T. Thomas, N. I. Badler, and J. T. Kider Jr., "Efficient motion retrieval in large motion databases," in *Proc. ACM SIGGRAPH Symp. Interactive 3D Graphics Games*, 2013, pp. 19–28.
- [51] G. A. Miller, "WordNet: a lexical database for english," *Commun. ACM*, vol. 38, no. 11, pp. 39–41, 1995.
- [52] T.-W. Lu and J. O'Connor, "Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints," *J. Biomechanics*, vol. 32, no. 2, pp. 129–134, 1999.
- [53] C. Mandery, J. B. Sol, M. Jöchner, and T. Asfour, "Analyzing whole-body pose transitions in multi-contact motions," in *Proc. IEEE/RAS Int. Conf. Humanoid Robots*, 2015, pp. 1020–1027.
- [54] T. H. Rowan, "Functional stability analysis of numerical algorithms," Ph.D. dissertation, Dept. Comput. Sci., Univ. Texas, Austin, TX, USA, 1990.
- [55] S. G. Johnson, "The NLOpt Nonlinear-Optimization Package," [Online] Available at: <http://ab-initio.mit.edu/nlopt/>, 2014.
- [56] K. Yamane and J. Hodgins, "Simultaneous tracking and balancing of humanoid robots for imitating human motion capture data," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2009, pp. 2510–2517.
- [57] A. Shon, K. Grochow, and R. Rao, "Robotic imitation from human motion capture using gaussian processes," in *Proc. IEEE/RAS Int. Conf. Humanoid Robots*, 2005, pp. 129–134.
- [58] Y. Zheng and K. Yamane, "Human motion tracking control with strict contact force constraints for floating-base humanoid robots," in *Proc. IEEE/RAS Int. Conf. Humanoid Robots*, 2013, pp. 34–41.
- [59] L. Boutin, A. Eon, S. Zeghloul, and P. Lacouture, "An auto-adaptable algorithm to generate human-like locomotion for different humanoid robots based on motion capture data," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2010, pp. 1256–1261.
- [60] M. Müller and T. Röder, "Motion templates for automatic classification and retrieval of motion capture data," in *Proc. ACM SIGGRAPH/Eurographics Symp. Computer Animation*, 2006, pp. 137–146.
- [61] R. A. Al-Asqhar, T. Komura, and M. G. Choi, "Relationship descriptors for interactive motion adaptation," in *Proc. ACM SIGGRAPH/Eurographics Symp. Comput. Animation*, 2013, pp. 45–53.
- [62] H. J. Shin, J. Lee, S. Y. Shin, and M. Gleicher, "Computer puppetry: An importance-based approach," *ACM Trans. Graphics*, vol. 20, no. 2, pp. 67–94, 2001.
- [63] A. Hofmann, M. Popovic, and H. Herr, "Exploiting angular momentum to enhance bipedal center-of-mass control," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2009, pp. 4423–4429.
- [64] M. Popovic and A. Englehart, "Angular momentum primitives for human walking: Biomechanics and control," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, vol. 2, 2004, pp. 1685–1691.
- [65] H. Herr and M. Popovic, "Angular momentum in human walking," *J. Exp. Biol.*, vol. 211, no. 4, pp. 467–481, 2008.



Christian Mandery received the Diploma degree in computer science from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2013. He is currently working toward the Ph.D. degree at the Institute for Anthropomatics and Robotics, High Performance Humanoid Technologies Laboratory (H²T), KIT, Karlsruhe, Germany.

His major research interests include representation and analysis of human whole-body motion, learning from human motion, and mapping of human motion to humanoid robots.



Ömer Terlemez received the Diploma degree in computer science from University of Karlsruhe, Karlsruhe, Germany, in 2009. He is currently working toward the Ph.D. degree at the Institute for Anthropomatics and Robotics, High Performance Humanoid Technologies Laboratory (H²T), KIT, Karlsruhe, Germany.

His major research interests include kinematic and dynamic modeling of the human body and transfer of human motion to humanoid robots.



Martin Do received the Diploma and Ph.D. degrees from Karlsruhe Institute of Technology, Karlsruhe, Germany, in 2007 and 2014, respectively.

He is a Postdoctoral Researcher with the Institute for Anthropomatics and Robotics, High Performance Humanoid Technologies Laboratory (H²T), KIT, Karlsruhe, Germany. His major research interests include robotic grasping and learning from human observation with a focus on new representations for grasping and goal-directed imitation of human grasping actions.



Nikolaus Vahrenkamp received the Diploma and Ph.D. degrees from Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2005 and 2011, respectively.

He is a Postdoctoral Researcher with the Institute for Anthropomatics and Robotics, High Performance Humanoid Technologies Laboratory (H²T), KIT, Karlsruhe, Germany. His research interests are motion planning, grasping and mobile manipulation and software development.



Tamim Asfour (M'05) received the diploma degree in electrical engineering in 1994 and the Ph.D. degree in computer science in 2003, both from University of Karlsruhe, Karlsruhe, Germany.

He is a Full Professor with the Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, where he holds the Chair of Humanoid Robotics Systems and is the Head of the High Performance Humanoid Technologies Laboratory (H²T). His major research interest is engineering high-performance humanoid robots. He

is a Developer and a Leader of the development team of the ARMAR humanoid robot family.