The KIT Whole-Body Human Motion Database

Christian Mandery, Ömer Terlemez, Martin Do, Nikolaus Vahrenkamp, Tamim Asfour
Institute for Anthropomatics and Robotics
Karlsruhe Institute of Technology (KIT), Germany
{mandery, asfour}@kit.edu

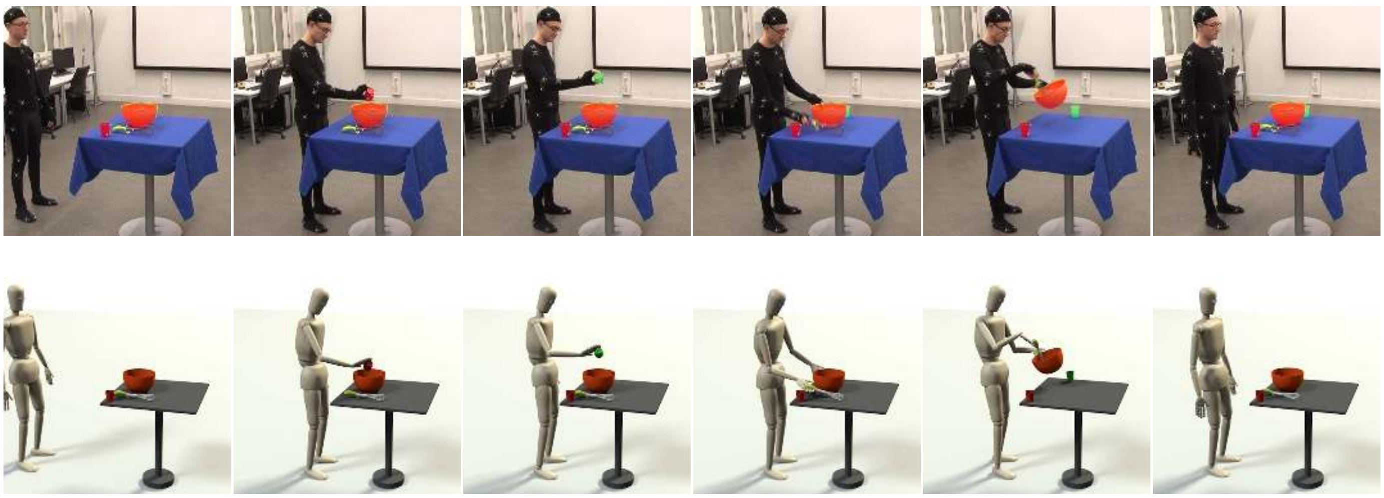

Fig. 1: Preparing the dough: Key frames of a manipulation task available in the motion database which includes four environmental objects.

***Abstract*—We present a large-scale whole-body human motion database consisting of captured raw motion data as well as the corresponding post-processed motions. This database serves as a key element for a wide variety of research questions related e.g. to human motion analysis, imitation learning, action recognition and motion generation in robotics. In contrast to previous approaches, the motion data in our database considers the motions of the observed human subject as well as the objects with which the subject is interacting. The information about human-object relations is crucial for the proper understanding of human actions and their goal-directed reproduction on a robot. To facilitate the creation and processing of human motion data, we propose procedures and techniques for capturing of motion, labeling and organization of the motion capture data based on a Motion Description Tree, as well as for the normalization of human motion to an unified representation based on a reference model of the human body. We provide software tools and interfaces to the database allowing access and efficient search with the proposed motion representation.**

I. INTRODUCTION

Understanding of human motion and its transfer to robots represents a promising way towards intuitive programming of robot systems with different body morphologies. Such understanding can only be gained by observing humans performing actions, where 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 the human subject. Thus, it is important to collect sufficient amount of motion data consisting of multiple demonstrations of actions performed by different subjects and under different conditions. For this purpose, great research efforts have been dedicated to the field of human motion capture, leading to commercial systems which feature an outstanding performance in terms of fast and precise tracking of human motions. Such human motion capture systems allow the collection of a huge amount of motion data which can be stored, structured and made available for further processing. In this context, a large-scale motion database is an essential component which provides the necessary input for motion analysis, motion synthesis and learning from human observation. Most available motion data collections have focused on the mere movements of a single human individual. From the viewpoint of robotics, the resulting data lacks crucial information which is needed to understand, to represent, and to reproduce observed human actions on robots in a goal-directed way.

To address this deficiency, we propose the KIT Whole-Body Human Motion Database as a rich motion data corpus which not only focuses on mere whole-body motions but also whole-body actions. 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

The research leading to these results has received funding from the European Union Seventh Framework Programme under grant agreement no. 611909 (KoroiBot) and grant agreement no. 270273 (Xperience).

Database name	Description	# recordings
CMU Graphics Lab Motion Capture Database [1]	Well-known and commonly used DB, wide range of motions (locomotion, interaction etc.), incoherent structure and organization	2605
HDM05 [2]	70 different actions performed by up to five subjects each (more than three hours of motion data)	1457
Human Motion Database [3]	Systematic sampling methodology, five different data sets (see text) with 350 different actions by one actor in the praxicon data set	five data sets (see text)
DB for study of gender, identity and emotion [4]	Small set of motions (walking, arm movements and combinations) played by 30 subjects for four different emotional states	4080
Eyes, JAPAN Co. Ltd. [5]	Vendor behind mocapdata.com, a commercial motion capture library tailored to animation purposes, low level of organization	933 (comm.) + 4671 (free)
HMDB51 [6]	Large-scale video database to evaluate action recognition algorithms, does not contain motion capture data	7000
KIT Whole-Body Human Motion Database	Wide range of motions from 38 subjects, includes environmental elements and objects, systematic structure and additional normalized motion representation	3704 (growing)

TABLE I: Selection of large-scale motion databases available to the scientific community, including our database.

motion description tags which allow for efficient search for certain motion types through structured queries. Using this motion description tags, a clear and intuitive data structure has been developed, and in combination with a comprehensive user interface, specific motion data can be accessed in an efficient manner. Furthermore, the established database provides a data interface which allows the transfer of motion data from a specific observed human subject to different kinematic and dynamic models of humans and robots. This interface is based on the Master Motor Map (MMM) representation ([7], [8]), which in its core contains a reference model of the human body providing a well-specified kinematic configuration and dynamics, and defines the data formats for capturing, representing and mapping of human motion to a target robot system. This reference model is used to generalize motion recordings from the subject-specific properties and, thus, facilitates the implementation of motion mapping mechanisms to different embodiments.

The paper is organized as follows: Section II provides an overview of related work in the area of human motion databases. In Section III, an overview of the implemented system and the resulting data flow of the underlying software is given. Section IV describes the motion capture process with details on the motion capture system and the experimental setup. The resulting motion recordings are labeled according to Section V and normalized using the MMM interface as described in Section VI. The access to the database and its content are outlined in Section VII and Section VIII. The work is summarized and notes on future work are given in Section IX.

II. RELATED WORK

Table I provides a brief overview of the most important large-scale motion databases employed in the scientific community. One of the largest motion databases is the CMU Graphics Lab Motion Capture Database [1]. It contains a total of 2605 recordings from a wide range of different motion types. Data has been collected from 144 subjects, although some subjects refer to the same actual person. Currently, the CMU database is one of the most commonly used databases in animation, computer graphics and robotics applications. Yet, the lack of structure and organization of motion data in the database as well as the inconsistent motion description between different subjects often makes its use inconvenient. The HDM05 motion database has been created for motion analysis, synthesis and classification [2]. It provides around 50 minutes of motion data, stored in 1457 motion clips of a limited number of roughly 100 motion classes. The motion clips are created by the manual slicing of motion recordings. Compared to the CMU database, HDM05 uses a more stringent structuring of the data. However, only five subjects have been used for the recordings and some of the motion classes only contain data from even less subjects. The Human Motion Database in [3] uses a controlled sampling approach to determine the motion data collected. The database consists of five different data sets: The praxicon data set contains recordings of about 350 actions from a single subject, while a cross-validation data set provides recordings for a subset of 70 actions for 50 different subjects. A generalization data set contains recordings with varying motion parameters, e.g. direction of movement. Two additional data sets provide 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, sad). Eyes, JAPAN Co. Ltd. [5] provides commercial and free motion capture recordings. Their collection contains 933 motion clips for sale and 4671 free recordings, available under a Creative Commons license. Similarly, the data collection by the Motion Capture Society [9] offers some packages of motion capture recordings for download. Both data sets however focus on animation purposes (e.g. game development) and exhibit little structure, making them hardly usable for research purposes. IEMOCAP [10] is a data set created for emotion recognition and the combined analysis of both verbal (speech) and non-verbal communication (gesture). It contains data captured in five sessions from ten specially trained actors which played ten different emotional states like happiness, fear or anger. The ICS Action Database [11] contains recordings for 25 precisely defined actions in a total of five sets. It is special in that every frame recorded is manually annotated for each of the 25 actions on whether the frame is certainly or maybe part of this specific action. However, little information is available about the recorded subjects. The Korea University Gesture database [12] has been created for the analysis of human whole-body

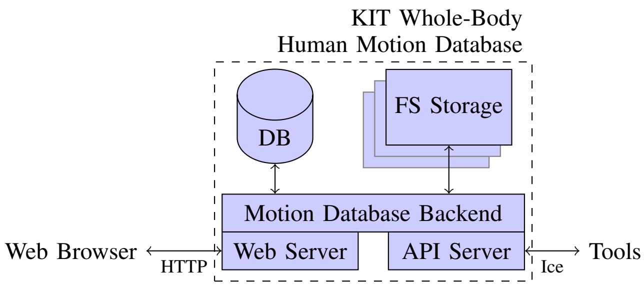

Fig. 2: Organization of the motion database.

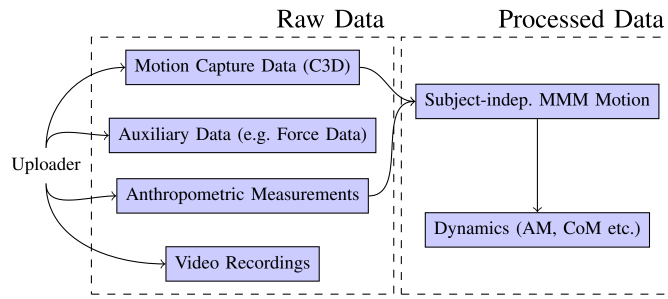

Fig. 3: Data types and data flow in the motion database.

gestures. In total, it contains 54 different gestures classified into "normal", "abnormal" and "command" gestures, captured from 20 subjects. The data set is not publicly available and access must be requested. [13] provides a data set containing position and force data from 18 human-robot and 6 human-human dyads. In this interaction task, both agents attempt to carry a large table together between different positions in the lab. The Motion Capture Database of the National University of Singapore [14] contains recordings of various motions from seven different subjects. The HumanEva database [15] has been created to evaluate algorithms for human pose estimation based on video recordings. It comprises six actions recorded from four subjects wearing everyday loose clothing instead of the usual skin-tight motion capture suit, which causes large inaccuracies in the resulting motion capture data and renders the data less suitable for many applications. Data is only available after signing and faxing a release agreement. The database in [16] provides walking motion recordings for 20 subjects, collected to study the identification of subjects from gaits. The LAAS-CNRS Novela Motion Capture Database [17] provides dance motions of a single subject. In [18], the authors are working on a large-scale motion database which reconstructs 3D motions from 2D images. In addition, there are also motion video databases that do not contain motion capture recordings. HMDB51 [6] is the largest action video database at present that provides about 7000 manually annotated clips. It has been developed to evaluate the performance and robustness of action recognition algorithms. The CMU Motion of Body (MoBo) database [19] contains video recordings of four different walking styles from 25 subjects. [3] provides a good overview over many existing motion capture databases with respect to content, structure and size.

Our general observation is that most existing motion databases have been collected by a single research group in a rather short period of time, often linked to one specific research project. These databases then contain recordings predominantly useful only for the initial research purpose and do not aim to be extendable. With our motion database, we pursue a more universal approach. Our motion database is intended as a long-term proposition to collect high-quality whole-body human motion data that is open to the entire research community and not restricted to a single research project.

III. SYSTEM OVERVIEW

Fig. 2 provides an overview of the organization of the motion database. The motion database provides two different methods of accessing its content which are described in more detail in Section VII. One the one hand, a web server provides a web interface to the user, and on the other hand, the database can be accessed directly through an application programming interface (API) based on the object-oriented middleware platform Ice [20]. Both, the web and the API server, are driven by a common backend which utilizes the object-relational database management system PostgreSQL (DB) and a filesystem-based storage (FS Storage) to manage the database content.

Fig. 3 shows which types of data are available in the motion database and the resulting data flow. Raw data which 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 measurements from force plates or data gloves. Additionally, manually measured anthropometric data of recorded subjects are entered into the database (see Section IV-C).

From the uploaded motion capture data, a normalized subject-independent representation of the motion based on the Master Motor Map (MMM) framework is created within the motion database. This normalization of motion data is described in Section VI. Based on the normalized motion and the associated kinematic and dynamic model of the human body, kinematics and dynamics of the motion like the the angular momentum (AM) or the center of mass (CoM) trajectory can be estimated.

IV. ACQUISITION OF MOTION DATA

A. Human Motion Capture Setup

Recordings of human motion are performed in a motion capture studio equipped with an optical marker-based Vicon MX motion capture system [21]. For the acquisition of motion, artificial reflective markers are placed on human subjects (see Section IV-B) and on the environmental elements and objects. Motion recordings are saved in the C3D file format which is the industry standard for storing motion capture data [22] and represents the input for the normalization steps in our processing pipeline described in Section VI. Additionally, a time-synchronized video of the motion is recorded and also made available through the database.

B. Reference Marker Set

Fig. 4 shows the reference marker set used for whole-body human motion capture. The marker set consists of 56 markers which are derived from specific anatomical landmarks of the

Fig. 4: Marker set used for whole-body human motion capture in our lab.

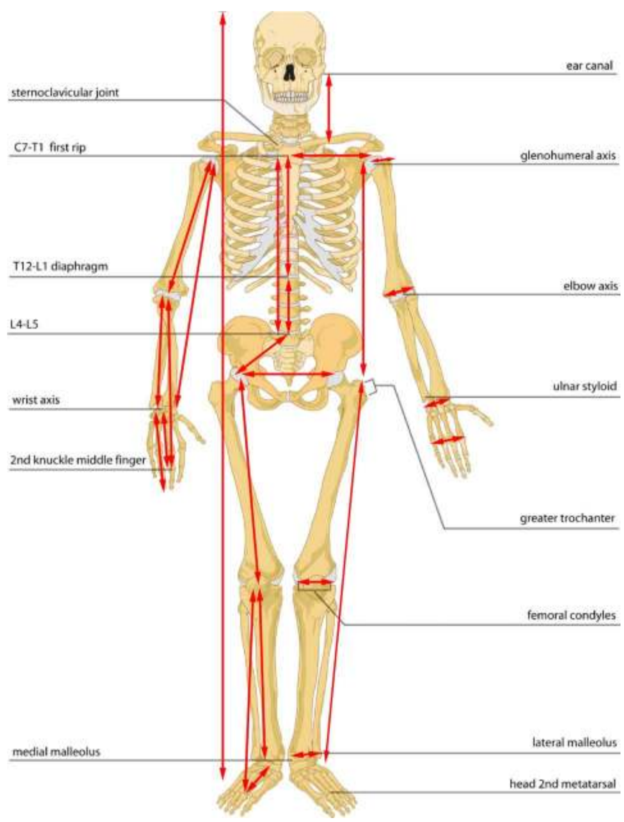

Fig. 5: The anthropometric subject parameters, which are also stored in the database.

human body. More information about the marker set including close-up pictures is available online[1].

[1]https://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.

C. Anthropometric Measurements

To gain a better understanding of the recorded motions, we added the possibility to attach manually measured anthropometric properties to every subject in the database. An overview of the measurable properties is shown in Fig. 5, while detailed information on how to obtain them, including appropriate examples for every property, is available online[2]. Since the values are obtained by manually measuring segment length with a tape measure, they should be considered a good estimate rather than exact values. The anthropometric parameters can be used for plausibility checks when computing kinematic properties of recorded subjects or help to determine parameters which are not computationally extractable from marker trajectories only.

D. Objects

Including objects in motion capture is interesting for research of human-object interaction, e.g. grasping, manipulation and locomotion tasks. We aim to provide data for the study

[2]https://motion-database.humanoids.kit.edu/anthropometric_table/

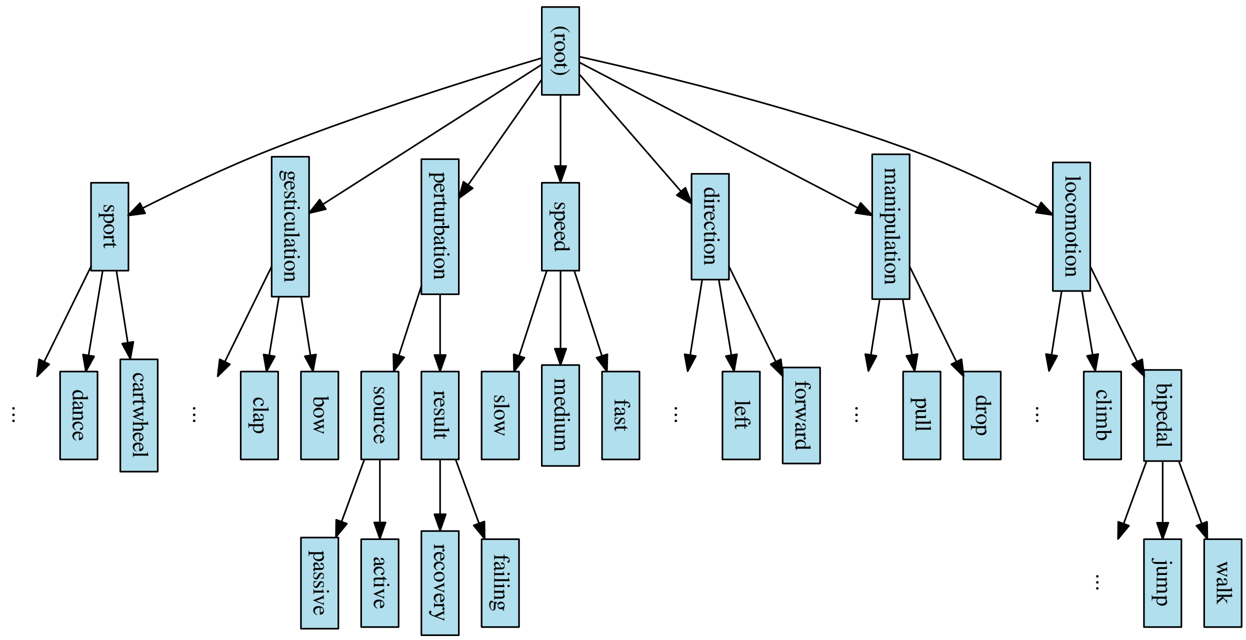

Fig. 7: Structure of the Motion Description Tree (excerpt).

of human actions on objects and environmental elements such as opening doors, stepping stairs, etc. Thus, markers are also attached to the objects in unique defined marker positions.

The marker placement must not be symmetric to rule out ambiguities. At least three markers should always be visible to estimate an object pose, however estimation robustness is increased with a larger number of visible markers. Fig. 6 exemplarily shows several objects with markers attached to them for motion capture. Furthermore, a 3D model of every object is created and saved in the database along with a custom object description. Similarly to subject information, every motion in the database is tagged with all objects that are involved in the motion.

V. MOTION LABELING

Recorded motions are classified within the so-called Motion Description Tree. The Motion Description Tree is a structure which we have 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. Fig. 7 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 [23] 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 which 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. 7.

The tree itself is managed in the database and thus extendable by the users. In contrast to 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 → bipedal → walk
2) speed → medium
3) direction → forward
4) direction → downward
5) perturbation → result → failing
6) perturbation → source → 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 above would be returned when searching for “locomotion → bipedal”. These queries can also be chained using logical operators as explained in Section VII-C to create more elaborate search queries.

VI. MOTION NORMALIZATION

The Master Motor Map (MMM) ([7], [8]) is an open-source framework for the representation, mapping and reproduction of human motion on robots. The fundamental idea of the MMM framework is to offer an unifying reference model of the human body with kinematic (DoF, segment lengths) and dynamic parameters (segment center of mass or inertial tensor). This model has been used in our previous work in the context of imitation learning on humanoid robots (see [24], [25]).

Based on this reference model, we implemented a method which allows the normalization of motions by reconstructing the joint angle movements of the model from the raw marker movements attached to the body of the human subject. This way, motion recordings are generalized from the individual kinematic and dynamic properties. To enable the joint angle reconstruction, virtual markers are placed on the reference model according to the placement of the physical markers on the human subject which are described in Section IV-B. For each frame, an optimization problem is solved in order to minimize the distances between the virtual and the corresponding physical marker positions. Thus, by running a single-step least-squares optimization with a maximal number of N iterations the $\boldsymbol{\theta}(t)$ is calculated as follows

$$\boldsymbol{\theta}_N(t) = \boldsymbol{\theta}_{N-1}(t) + J^{+}_{N-1}\left(\boldsymbol{x} - f(\boldsymbol{\theta}_{N-1}(t))\right),$$

where J denotes a matrix which combines the Jacobian matrices with respect to all marker positions, $\boldsymbol{x}$ denotes the captured marker positions and $f(\boldsymbol{\theta}_{N-1}(t))$ stands for the positions of the virtual markers calculated with direct kinematics function f of the reference model, the MMM model. The optimization procedure is initialized with $\boldsymbol{\theta}_0(t) = \boldsymbol{\theta}(t-1)$ and converges within a few iterations. Details on the optimization, the mapping procedure and extensions for the considerations of possible joint constraints are given in [8].

The MMM framework also provides specialized tools to reconstruct the motion of the human hand and of rigid bodies like the objects involved in a motion. The tools are further described in [8].

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 VII-B. The result of the conversion is a motion of the MMM reference model represented in the XML-based file format, which then is stored alongside the corresponding raw C3D motion in the database for subsequent processing. Based on the dynamic parameters of the MMM model, dynamic motion characteristics such as angular momentum and body center of mass can be calculated automatically for all motions in the database. The video attachment exemplarily shows the calculated motion dynamics of a walking motion in the database. For differing applications where 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.

VII. Accessing the Database

Read access to the database is publicly available. This includes both, raw data like the data described in Section IV and processed data like the normalized motions described in Section VI. To access the database, we provide a web interface and an application programming interface (API).

A. Web Interface

The web interface[3] provides a convenient method to access the motion database which can be used without client-side software installation. In contrast to other motion databases like [1] which also provide a web interface, our interface allows registered users to login and edit the database content directly from their web browser. Files can be downloaded and uploaded both separately or in bulk mode as an archive file. Additionally, detailed descriptions for the marker set and the anthropometric data table outlined in Section IV are available online.

[3]https://motion-database.humanoids.kit.edu/

B. Application Programming Interface

The API allows to directly access the motion database from several popular platforms and programming languages like C++, .NET, Java, Python and Ruby and thus underlines our attempt to facilitate the integration of the motion database in new and existing tools and pipelines. As an example, the tools for motion normalization and motion dynamics calculation described in Section VI have been developed in C++ and Python and access the database content through this API. The technical basis for the database API is the Internet Communitications Engine (Ice) framework [20] which is a Remote Procedure Call (RPC) framework provided by ZeroC for the development of distributed applications.

C. Searching

An important requirement for a large-scale motion database is the integration of efficient search mechanisms to easily locate appropriate motion content. In the KIT motion database, motions can be found by using the classification provided by the Motion Description Tree (see Section V), where search queries can be constructed using logical operators, for example:

- `run AND forward`: Running motions directed forwards.
- `carry AND drop`: Motions where an object is carried and dropped (a specific object may also be included in the search).
- `run OR (walk AND NOT(slow))`: Motions where the subject is either running, with any speed, or walking, but not slow.

Motions can also be searched for environmental objects, subjects, projects or the lab where the recording has been made, and be ordered by various fields. In addition, the API makes the implementation of arbitrary client-side motion selection algorithms feasible.

VIII. Database Content

As of 2015, May 12, the motion database consists of motion data of a total run length of 7.68 hours. 43 different subjects (31 male, 12 female) and 41 different objects have been entered into the database. 289 motion capture experiments, defined as a combination of subject and motion type, have been performed in our motion capture lab at KIT. Additionally, 165 experiments have been performed at the Eberhard Karls University Tübingen (EKUT), the Weizmann Institute of Science and the LAAS-CNRS. In total, the database comprises 9727 files that belong to the file types shown in Fig. 3 with a total size of 32.04 GiB.

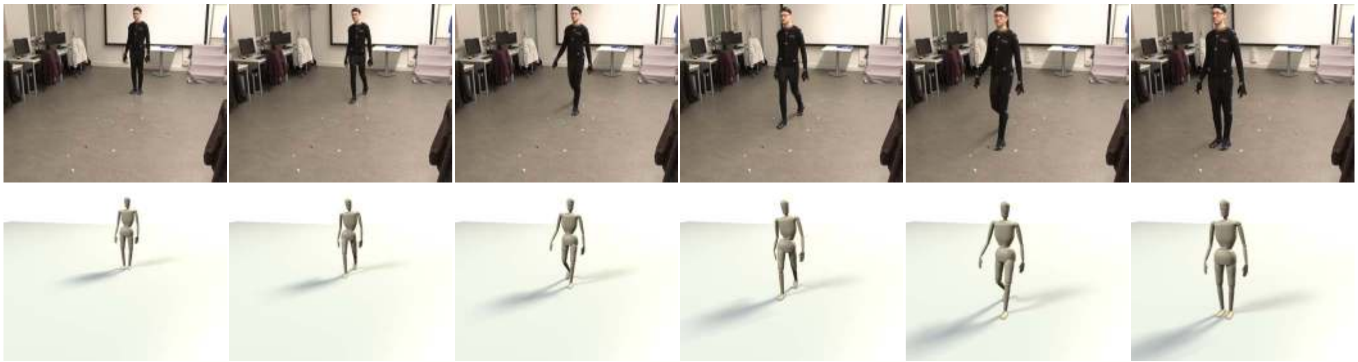

Fig. 8: Walking motion (exemplary key frames).

Motion Type	# rec.
Walking and running with different speeds	1107
Walking a left/right bend	147
Walking paths (circle/ellipse/figure-of-8 etc.)	550
Walking backwards	79
Turning on the spot	332
Walking up/down slopes	232
Walking up/down stairs	206
Walking over a beam	204
Walking over a seesaw	30
Walking over step stones	71
Walking on a soft mattress	29
Walking around obstacle objects	60
Walking around a moving obstacle	158
Stepping over an (imaginary) gap	92
Push recovery while walking and standing	266

TABLE II: Types of locomotion motions stored in the database (excerpt) and the corresponding number of recordings.

The video attachment shows examples for both raw motion data and normalized motions available in the KIT motion database. A major category of motions in the database is locomotion on and around objects as well as push recovery. Table II shows some motion types which have been collected within this category that are already available in the motion database. As an example, Fig. 8 shows exemplary key frames from the video recording and the normalized MMM motion of a simple walking motion available in the database[4]. For more complex locomotion tasks involving objects like a seesaw, the motion of these objects is also tracked as described in Section IV-D. Additionally, for motions offering variation possibilities, these variations have been experimentally explored (e.g. for walking stairs: Walking upwards or downwards, forward or backward, with/without handrail etc.). Examples for locomotion recording involving objects can be seen in Fig. 9.

Another important category of recordings shows manipulation motions which 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 which involves four environmental objects (bowl, whisk and two cups). Furthermore, recordings of human-human interaction, where two subjects are tracked at the same time, are provided. This category of motions includes for example greeting motions (e.g. handshake, high-five), handover motions (e.g. pizza delivery) and combined locomotion tasks (e.g. synchronized marching).

Other categories of motions which are already provided include gesticulation motions, like conversation, pointing or waving motions, and music/sports motions (e.g. playing drums, playing (air) guitar, punching and kicking motions, squats, jumping jacks, dancing). For the purpose of motion normalization by other means than the normalization done with the MMM framework, also static calibration poses (e.g. T-Pose) and dynamic calibration motions are available for some of the subjects.

IX. CONCLUSION

We presented the structure and the content of the publicly available KIT Whole-Body Human Motion Database and described the procedures for a systematic recording of motion data with associated anthropometric subject data as well as environmental elements and objects. We presented an approach for labeling the content of the database in such a way which allows efficient search for certain motion types based on the Motion Description Tree. We showed how the motion data is normalized to a reference model of the human body by using the Master Motor Map framework. Furthermore, we described the software framework which has been developed to allow convenient access and use of the database. In addition to the web interface, the database API allows the easy integration of the database into existing toolchains and frameworks. It is our hope that our motion database with its features will provide a valuable addition not only to our own work in the area of learning from human observation and humanoid robotics, but to the community. Future work will address the collection of more motion data with a focus on whole-body loco-manipulation tasks, e.g. opening doors, and recordings of human push recovery. Improved recording setups may allow the annotation of the database content with additional information like force plate data. Additionally, some improvements to the database interface like auto-suggestion features or a video preview have been proposed. In the long-run, we aim to extend the database further towards automatic analysis of uploaded motions, e.g. automatic generation of linguistic interpretations from uploaded motions, motion clustering and the learning of motion primitives.

REFERENCES

[1] "CMU Graphics Lab Motion Capture Database," http://mocap.cs.cmu.edu/, accessed February 13, 2015.

[4] https://motion-database.humanoids.kit.edu/details/motions/395/

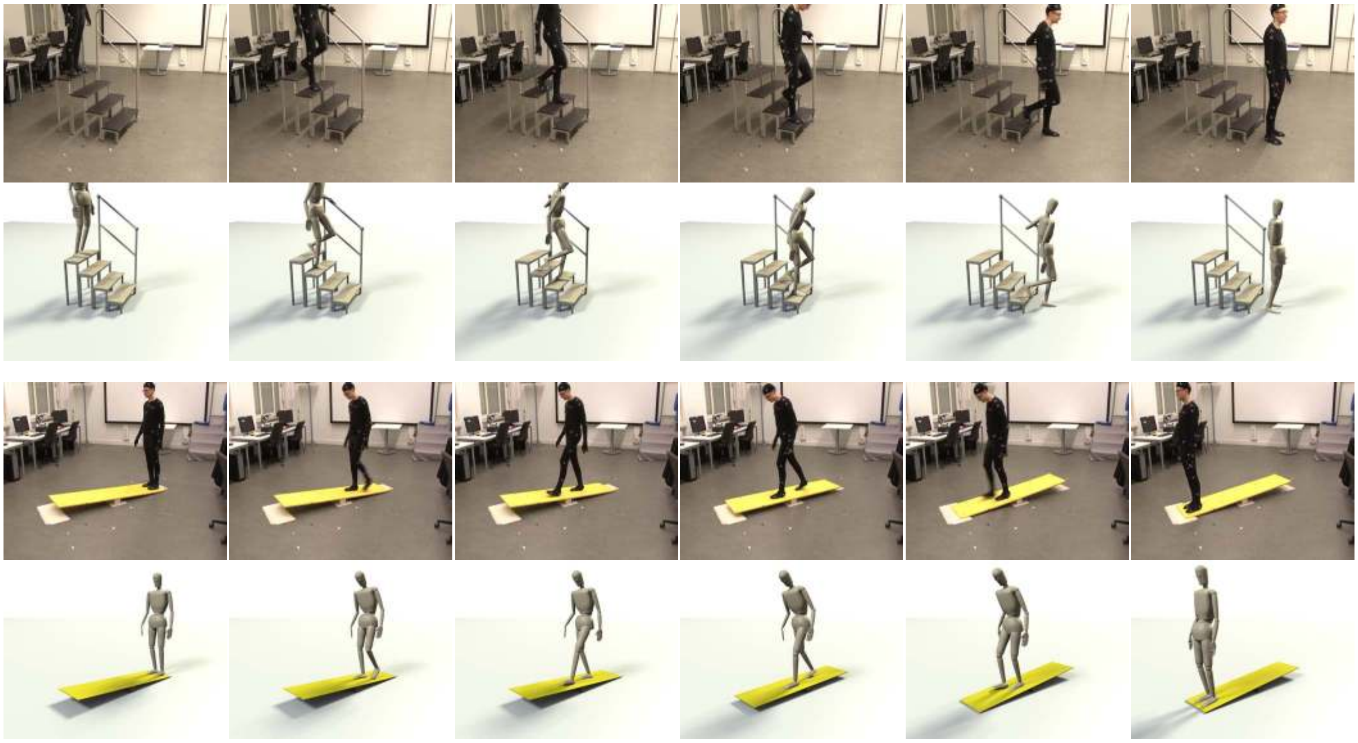

Fig. 9: Two example motions involving environmental objects, namely a flight of stairs and a seesaw (exemplary key frames).

[2] M. Müller, T. Röder, M. Clausen, B. Eberhardt, B. Krüger, and A. Weber, "Documentation mocap database hdm05," 2007.

[3] G. Guerra-Filho and A. Biswas, "The human motion database: A cognitive and parametric sampling of human motion," *Image and Vision Computing*, 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 research methods*, vol. 38, no. 1, pp. 134–141, 2006.

[5] "Mocapdata.com," http://www.mocapdata.com/, accessed February 13, 2015.

[6] H. Kuehne, H. Jhuang, R. Stiefelhagen, and T. Serre, "Hmdb51: A large video database for human motion recognition," in *High Performance Computing in Science and Engineering 12*. Springer, 2013, pp. 571–582.

[7] P. Azad, T. Asfour, and R. Dillmann, "Toward an unified representation for imitation of human motion on humanoids," in *Robotics and Automation, 2007 IEEE International Conference on*. IEEE, 2007, pp. 2558–2563.

[8] 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 *Proceedings, IEEE-RAS International Conference on Humanoid Robotics (Humanoids)*, 2014.

[9] "The Motion Capture (Mocap) Club," http://www.mocapclub.com/, accessed February 13, 2015.

[10] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. S. Narayanan, "IEMOCAP: interactive emotional dyadic motion capture database," *Language resources and evaluation*, vol. 42, no. 4, pp. 335–359, 2008.

[11] "ICS Action Database," http://www.ics.t.u-tokyo.ac.jp/action/, accessed February 13, 2015.

[12] B.-W. Hwang, S. Kim, and S.-W. Lee, "A full-body gesture database for automatic gesture recognition," in *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*. IEEE, 2006, pp. 243–248.

[13] A. Mörtl, M. Lawitzky, A. Kucukyilmaz, M. Sezgin, C. Basdogan, and S. Hirche, "The role of roles: Physical cooperation between humans and robots," *The International Journal of Robotics Research*, vol. 31, no. 13, pp. 1656–1674, 2012.

[14] "NUS Mocap," http://animation.comp.nus.edu.sg/nusmocap.html, accessed February 13, 2015.

[15] 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," *International journal of computer vision*, vol. 87, no. 1-2, pp. 4–27, 2010.

[16] "Human identification at a distance," http://www.cc.gatech.edu/cpl/projects/hid/, accessed February 13, 2015.

[17] "LAAS-CNRS Novela Motion Capture Database," http://projects.laas.fr/gepetto/novela/noveladb/, accessed February 13, 2015.

[18] W. Takano, J. Ishikawa, and Y. Nakamura, "Design of a large-scale motion database which enables recovery of 3d motion from 2d images," in *JSME Robotics and Mechatronics Conference*, 2012.

[19] R. Gross and J. Shi, "The cmu motion of body (mobo) database," Tech. Rep., 2001.

[20] "ZeroC - The Internet Communications Engine (Ice)," http://www.zeroc.com/ice.html.

[21] "Vicon motion capture systems," http://www.vicon.com/.

[22] "C3D: The 3D Biomechanics Data Standard," http://www.c3d.org/.

[23] G. A. Miller, "WordNet: a lexical database for english," *Communications of the ACM*, vol. 38, no. 11, pp. 39–41, 1995.

[24] M. Do, T. Asfour, and R. Dillmann, "Towards a Unifying Grasp Representation for Imitation Learning on Humanoid Robots," in *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 2011, pp. 482–488.

[25] T. Asfour, M. Do, K. Welke, A. Bierbaum, P. Azad, N. Vahrenkamp, S. Gärtner, A. Ude, and R. Dillmann, "From sensorimotor primitives to manipulation and imitation strategies in humanoid robots," *Springer Tracts in Advanced Robotics*, vol. 70, no. STAR, pp. 363–378, 2011.