HUMoD - A versatile and open database for the investigation, modeling and simulation of human motion dynamics on actuation level

Janis Wojtusch, Member, IEEE and Oskar von Stryk, Member, IEEE

Abstract—Collecting high-quality biomechanical measurement data for investigating human motion dynamics with muscle driven actuation, i.e., joint trajectories, ground reaction forces and muscle activities, usually requires experienced examiners, expensive measurement equipment and sophisticated processing. Several published databases offer biomechanical measurement data of various human motions. However, most of these databases are not primarily intended for modeling, simulation and validation of human motion dynamics with muscle driven actuation and do not include ground reaction forces together with muscle activities and detailed anthropometric parameters of the subjects.

In this paper, the HUMoD Database, a versatile and quite unique combination of comprehensive biomechanical measurement data and anthropometric parameters with a focus on lower limbs, is introduced. The provided datasets allow to create and validate biomechanical models of the human locomotor system on actuation level and to investigate and simulate human motion dynamics including muscle driven actuation. Besides investigations in biomechanics, the database can be of value especially for the design and development of musculoskeletal humanoid robots and for better understanding and benchmarking human-like robot locomotion. The database contains raw and processed biomechanical measurement data from a three-dimensional motion capture system, an instrumented treadmill and an electromyographical measurement system for eight different motion tasks performed by a healthy female and male subject as well as anthropometric parameters for both subjects. The biomechanical measurement data, anthropometric parameters and source code of the applied computational scripts are open and can be obtained free of charge from the HUMoD Database website.

I. INTRODUCTION

Research on human motion dynamics of the lower limbs with muscle driven actuation is largely based on biomechanical measurement data, i.e., kinematic motion data that describes the spatial motion of body segments, ground reaction force data and muscle activity data [1]. Measured ground reaction forces and kinematic motion data can be used to mathematically derive center of pressure trajectories, joint center estimates or joint trajectories. In combination with anthropometric parameters, i.e., body segment lengths, masses, centers of mass and moments of inertia, this data allows to create biomechanical models of the human locomotor system and to investigate and simulate the dynamics of the measured human motion. Measured muscle activities can be used to parametrize and validate neuromuscular models and investigate muscle interaction. One example for the application of biomechanical measurement data is the computation of joint torques \( \tau \), e.g., [2], [3], or muscle forces \( m \), e.g., [4], [5], from joint trajectories \( q, \dot{q}, \ddot{q} \) and ground reaction forces \( g \) by using inverse dynamics simulation in combination with an appropriate dynamics model \( M(q), C(q, \dot{q}), G(q), J(q) \) given by

\[
\tau = M(q) \ddot{q} + C(q, \dot{q}) + G(q) + J(q)g,
\]

\[
m = f(\tau).
\]

The computed biomechanical torques \( \tau \) and forces \( m \) can be applied to design and validate drive trains for musculoskeletal humanoid robots, e.g., [6], or assistive devices, e.g., [7]. Another example are optimization based investigations of human motion dynamics where optimal control approaches, e.g., [8], [9], or inverse optimal control approaches, e.g., [10], [11] are used to find optimal joint torques \( \tau^* \), optimal muscle activations \( \alpha^* \) or optimization criteria \( \Phi[i] \) for human-like locomotion of a humanoid robot from biomechanical measurement data with

\[
\min_{\alpha} \Phi[q, \dot{q}, \ddot{q}, \tau, \alpha]
\]

subject to

\[
\ddot{q} = M^{-1}(q) (\tau - C(q, \dot{q}) - G(q) - J(q)g),
\]

\[
\tau = g(m),
\]

\[
m = h(\alpha).
\]

Validated with measured ground reaction forces and muscle activity data, the obtained optimization results facilitate the understanding of fundamental principles and concepts in humanoid motion.

The collection of high-quality biomechanical measurement data usually requires careful preparation, experienced examiners and a gait laboratory with expensive measurement equipment. Subsequent to the measurement, the collected data has to be further processed before it can be used for investigation, modeling, simulation or validation. In order to reduce this preparative effort or in case of no access to a well-equipped gait laboratory, it is possible to resort to one of several published databases with biomechanical measurement data of various motion tasks performed by human subjects. The Motion Capture Database\(^1\) and the Multimodal Human Action Database [12] provide kinematic motion data acquired with marker-based motion capture systems for locomotion, human interaction, interaction with the

\(^1\)http://mocap.cs.cmu.edu

This work was funded by the German Research Foundation DFG under grant STR533/8-1.

Janis Wojtusch and Oskar von Stryk are with the Department of Computer Science, Simulation, Systems Optimization and Robotics Group, Technische Universität Darmstadt, Hochschulstraße 10, 64289 Darmstadt, Germany.

wojtusch | stryk@sim.tu-darmstadt.de
environment or physical activities. In addition to kinematic motion data, the Motion Capture Database HDM05 [13] includes an individual kinematic body model for each subject that is generated from motion capture data. This kinematic body model can be used to estimate joint centers or joint trajectories. The Whole-Body Human Motion Database\(^2\) contains kinematic motion data for several activities of daily living as well as information about age, gender, body mass, body height and a set of anthropometric lengths for each subject. Some datasets also provide force plate measurements and activity-related information like geometric data of stairs or handrails. The OpenSim Motion and Simulation Data website\(^3\) is a collection of small databases containing various biomechanical measurement data for normal and pathological locomotion provided by different researchers and prepared for the modeling and simulation software OpenSim [14]. The type and amount of included data differs among the individual databases, some of which are related to specific publications. Most of these published databases are not primarily intended for the investigation, modeling, simulation and validation of human motion dynamics with muscle driven actuation and do not include ground reaction forces together with muscle activities and detailed anthropometric parameters of the subjects.

In this paper, the HuMoD Database, a versatile and open database providing kinematic motion data, ground reaction forces and muscle activities with a focus on lower limbs for eight different motion tasks performed by a healthy female and male subject, is introduced. The database contains a quite unique combination of raw and processed data from a three-dimensional motion capture system, an instrumented treadmill and an electromyographical measurement system as well as anthropometric parameters for both subjects.

Details on the measurement protocol are given in Section II. The applied motion protocol, measurement setup and data processing are described. The structure and content of the HuMoD Database is presented in Section III. A description of the provided data as well as the included computational scripts is given. A concluding discussion of the HuMoD Database is provided in Section IV.

### II. MEASUREMENT PROTOCOL

A healthy female subject (27 yrs, 161 cm, 57 kg) and male subject (32 yrs, 179 cm, 85 kg) performed thirteen trials without shoes dressed in underwear. The subjects were given time to become familiar with the measurement setup and equipment before the measurements and to rest between the trials. The measurement procedure was reviewed and approved by the ethical review committee of Friedrich-Schiller-Universität Jena, Germany. The subjects provided informed consent in accordance with the policies of the ethical review committee.

#### A. MOTION PROTOCOL

The subjects performed eight motion tasks, partially at different speeds or under changed conditions resulting in thirteen trials. The motion tasks cover locomotion, interaction with an object and physical activity representing a sample of typical repetitive tasks and goal-oriented tasks useful for biomechanics and humanoid robotics research. These include walking, running, squatting and jumping as well as avoiding obstacles and kicking a ball. Locomotion trials comprise straight walking at 1.0 \(\text{m/s}\), 1.5 \(\text{m/s}\) and 2.0 \(\text{m/s}\), straight running at 2.0 \(\text{m/s}\), 3.0 \(\text{m/s}\) and 4.0 \(\text{m/s}\), sideways walking at 1.0 \(\text{m/s}\) and a transition between standing and straight running with accelerating at 0.1 \(\text{m/s}^2\) from 0.0 \(\text{m/s}\) to 4.0 \(\text{m/s}\), holding 4.0 \(\text{m/s}\) for 20s and decelerating at -0.1 \(\text{m/s}^2\) from 4.0 \(\text{m/s}\) to 0.0 \(\text{m/s}\). Trials including interaction with an object are avoiding a box obstacle (41×20×15 cm) that was placed lengthwise and crosswise.

#### TABLE I

**SUMMARY OF THE DATASETS IN THE HuMoD DATABASE.**

<table>
<thead>
<tr>
<th>#</th>
<th>Description</th>
<th>Time</th>
<th>Events of Subject A</th>
<th>Events of Subject B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Straight walking at 1.0 (\text{m/s})</td>
<td>60 s</td>
<td>60 left steps, 60 right steps</td>
<td>54 left steps, 53 right steps</td>
</tr>
<tr>
<td>1.2</td>
<td>Straight walking at 1.5 (\text{m/s})</td>
<td>60 s</td>
<td>71 left steps, 71 right steps</td>
<td>60 left steps, 60 right steps</td>
</tr>
<tr>
<td>1.3</td>
<td>Straight walking at 2.0 (\text{m/s})</td>
<td>60 s</td>
<td>81 left steps, 81 right steps</td>
<td>70 left steps, 71 right steps</td>
</tr>
<tr>
<td>2.1</td>
<td>Straight running at 2.0 (\text{m/s})</td>
<td>60 s</td>
<td>89 left steps, 89 right steps</td>
<td>73 left steps, 73 right steps</td>
</tr>
<tr>
<td>2.2</td>
<td>Straight running at 3.0 (\text{m/s})</td>
<td>60 s</td>
<td>96 left steps, 97 right steps</td>
<td>79 left steps, 78 right steps</td>
</tr>
<tr>
<td>2.3</td>
<td>Straight running at 4.0 (\text{m/s})</td>
<td>60 s</td>
<td>105 left steps, 105 right steps</td>
<td>86 left steps, 85 right steps</td>
</tr>
<tr>
<td>3</td>
<td>Sideways walking at 0.5 (\text{m/s})</td>
<td>60 s</td>
<td>61 left steps, 60 right steps</td>
<td>47 left steps, 46 right steps</td>
</tr>
<tr>
<td>4</td>
<td>Transition between 0.0 (\text{m/s}) and 4.0 (\text{m/s})</td>
<td>112 s</td>
<td>150 left steps, 149 right steps</td>
<td>130 left steps, 129 right steps</td>
</tr>
<tr>
<td>5.1</td>
<td>Avoiding a long box obstacle at 1.0 (\text{m/s})</td>
<td>120 s</td>
<td>6 obstacles, 119 left steps, 119 right steps</td>
<td>7 obstacles, 105 left steps, 105 right steps</td>
</tr>
<tr>
<td>5.2</td>
<td>Avoiding a wide box obstacle at 1.0 (\text{m/s})</td>
<td>120 s</td>
<td>6 obstacles, 120 left steps, 121 right steps</td>
<td>7 obstacles, 106 left steps, 106 right steps</td>
</tr>
<tr>
<td>6</td>
<td>Squats with stopped treadmill</td>
<td>40 s</td>
<td>16 squats</td>
<td>10 squats</td>
</tr>
<tr>
<td>7</td>
<td>Kicks with stopped treadmill</td>
<td>100 s</td>
<td>8 kicks</td>
<td>9 kicks</td>
</tr>
<tr>
<td>8</td>
<td>Jumps with stopped treadmill</td>
<td>20 s</td>
<td>35 jumps</td>
<td>35 jumps</td>
</tr>
</tbody>
</table>

\(^2\)https://motion-database.humanoids.kit.edu

\(^3\)https://simtk.org/project/xml/downloads.xml?group_id=777
on the treadmill at 1.0 m/s and kicking a soft football (20 cm, 160 g). Physical activity trials contain continuous squats and continuous jumps with arms akimbo. Squats, kicks and jumps were performed with stopped treadmill. During the first and last 10 s of each trial the force plates remained unloaded. Before and after performing the particular motion, the subject stood still on the treadmill for at least 10 s. This idle time was increased to 20 s after fast motion tasks. Details of the single trials are summarized in Table I.

B. MEASUREMENT SETUP

All measurements were performed at the Locomotion Lab of André Seyfarth at Technische Universität Darmstadt, Germany. The spatial motion of upper and lower limbs was recorded at 500 Hz with a three-dimensional motion capture system consisting of four Oqus 310+ cameras and six Oqus 300+ cameras (Qualysis, Sweden). A set of thirty-five reflective markers with a diameter of 19 mm mounted on thin cardboard was placed on the skin at anatomical landmarks. One additional reflective marker was placed on top of the underpants above the pubic symphysis landmark [15]. The electrical activity of fourteen selected muscles in the legs was recorded at 2000 Hz with the electromyographical measurement system Bagnoli-16 Desktop (Delsys, USA). The measured signals were internally filtered to a bandwidth between 20 Hz and 450 Hz. The set of fourteen surface electrodes was placed according to SENIAM guidelines [16]. Figure 1 illustrates the locations of the thirty-six reflective markers for motion capture and fourteen surface electrodes for electromyographical measurement. All trials were performed on the instrumented treadmill ADAL3D-WR (Tecmachine, France). The belt of the treadmill runs over two force plates with four single-axis force sensors (Kistler, Switzerland) that were used to measure the vertical ground reaction forces $F_y$ of the left and right foot. The two force plates are mounted on top of four multi-axis force sensors (Kistler, Switzerland) that measured the lateral forces $F_x$ and $F_z$. All forces were recorded at 1000 Hz. Figure 2 shows the instrumented treadmill and the single- and multi-axis force sensors.

C. DATA PROCESSING

The raw data measured with the motion capture system, instrumented treadmill and electromyographical measurement system was processed with the numerical computing software MATLAB (MathWorks, USA) in order to compensate measurement errors and provide additional information for the investigation, modeling, simulation and validation of human motion dynamics.

1) KINEMATIC MOTION DATA: Raw kinematic motion and ground reaction force data were synchronized by compensating temporal offset and drift as well as transforming the global reference frame of the motion capture system into the global reference frame of the instrumented treadmill considering the ISB recommendations for biomechanical reference frame notation [17]. Figure 2 illustrates the applied global reference frame where the origin is located at the center of the rectangle spanned by the left and right force plates projected to the top of the belt surface. Infrequent gaps in raw kinematic motion data of up to 300 ms resulting from temporarily covered reflective markers were filled by applying fitted polynomial approximations. The measured spatial positions of the reflective markers were shifted to the approximated skin surface and used to estimate the joint centers of fifteen joints in arms, trunk, pelvis and legs [15], [18]–[20]. A detailed description of this procedure is provided on the HuMoD Database website. The preliminary joint center estimates are affected by soft tissue artifacts resulting from relative skin to bone motion that have a high impact on subsequent kinematics and dynamics computations [21], [22]. An evident indicator for soft tissue artifacts is a variation in body segment lengths which connect the joint centers and are assumed to be rigid. Figure 3 shows the relative variation of the left thigh length for the motion tasks straight walking at 2.0 m/s, straight running at 4.0 m/s and
normally distributed random variables \( w \). Process and measurement noise are modeled as zero-mean markers \( x \) markers. Joint trajectories \( \tilde{x} \) markers. The joint trajectories \( x \) markers with beginning at the first time instant. Third, a backward recursion (smoothing) uses information of future samples to further improve the joint trajectory estimates \( \hat{x}^{(t)} \) with beginning at the last time instant. The recursive equations of the three update steps are

\[
\begin{align*}
\hat{x}^{(t)} &= f(\hat{x}^{(t-1)}), \\
\tilde{x}^{(t)} &= \tilde{x}^{(t)} + K^{(t)}(z^{(t)} - h(\tilde{x}^{(t)})), \\
x^{(t)} &= \hat{x}^{(t)} + J^{(t)}(x^{(t+1)} - f(\tilde{x}^{(t)}))
\end{align*}
\]

with the adaptive filter gain \( K^{(t)} \) and smoother gain \( J^{(t)} \). The parameters of process and measurement noise were derived from measurement error ratings of the three-dimensional motion capture system. The joint trajectories are given as Tait–Bryan angles in \( x'y'z' \) convention.

The applied kinematics model, shown in Figure 4, consists of thirteen rigid body segments and twelve rotatory joints with different degrees of freedom. Each leg is modeled with three body segments representing thigh, shank and foot. These body segments are connected by the ankle joint (AJ) with three degrees of freedom, the knee joint (KJ) with one degree of freedom and the hip joint (HJ) with three degrees of freedom. Each arm is modeled with two body segments representing upper and lower arm as well as the elbow joint (EJ) with one degree of freedom and the shoulder joint (SJ) with three degrees of freedom. Head, trunk and pelvis are individual body segments that are connected by the lower neck joint (LNJ) corresponding to the C7/T1 joint and the lower lumbar joint (LLJ) corresponding to the L5/S1 joint each with three degrees of freedom. The joint center estimates also include estimates for the upper lumbar joint (ULJ) corresponding to the T12/L1 joint and two toe joints (TJ). These joint center estimates were not considered in the kinematics model because of missing anthropometric parameters for upper and lower torso and separate toes. The

![Figure 3](image-url) Relative left thigh length variation of subject A for the motion tasks straight walking at 2.0 m/s (red), straight running at 4.0 m/s (blue) and squats with stopped treadmill (green) as a result of soft tissue artifacts in preliminary joint center estimates. Faint lines represent individual thigh length variation for each cycle, while strong lines represent average thigh length variation over all cycles of the particular motion task. The thigh length variation is normalized in time to gait or squat cycles.

In order to compensate the influence of soft tissue artifacts and other unpreventable measurement errors, an extended Kalman smoother in combination with a subject-specific kinematics model with twenty-eight degrees of freedom (DOF) was applied to compute the joint trajectories including joint positions \( q \), marker velocities \( \dot{q} \), marker accelerations \( \ddot{q} \) and joint jerk \( \dddot{q} \) as well as smoothed joint center estimates [22], [23]. The extended Kalman smoother estimates the joint trajectories \( x = [q_1, \dot{q}_1, \ddot{q}_1, \dddot{q}_1, \ldots, q_{28}, \dot{q}_{28}, \ddot{q}_{28}, \dddot{q}_{28}]^T \) by combining noisy measurements of the reflective markers \( z \) with prior system knowledge and minimizes the estimation error statistically. The prior system knowledge is represented by a process model \( f(x) \) that characterizes the expected time evolution of the joint trajectories \( x \) and a measurement model \( h(x) \) that describes the non-linear relation between the joint trajectories \( x \) and noisy measurements of the reflective markers \( z \) which is given by the applied kinematics model. Process and measurement noise are modeled as zero-mean normally distributed random variables \( w, v \). The process model is based on the assumption that the joint jerk \( \dddot{q}_i \) is constant which results in a linear process model

\[
f_i(x_i) = \begin{bmatrix} \Delta t & \Delta t^2 & \frac{\Delta t^3}{2} \\ 0 & 1 & \Delta t \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x_i
\]

with \( i = 1, \ldots, 28 \) and the sample time \( \Delta t \). The extended Kalman smoother applies three successive update steps. First, the process model \( f(x^{(t-1)}) \) is used for a predictive time update \( \hat{x}^{(t)} \). Second, a forward recursion (filtering) estimates the joint trajectories \( \hat{x}^{(t)} \) from all measurements \( x^{(t)} \) integrating information of past samples with beginning at the first time instant. Third, a backward recursion (smoothing) uses information of future samples to further improve the joint trajectory estimates \( \hat{x}^{(t)} \) with beginning at the last time instant. The recursive equations of the three update steps are

\[
\begin{align*}
\hat{x}^{(t)} &= f(\hat{x}^{(t-1)}), \\
\tilde{x}^{(t)} &= \hat{x}^{(t)} + K^{(t)}(z^{(t)} - h(\tilde{x}^{(t)})), \\
x^{(t)} &= \hat{x}^{(t)} + J^{(t)}(x^{(t+1)} - f(\tilde{x}^{(t)}))
\end{align*}
\]
kinematic model was implemented with the multibody systems library MBSLIB [24] that provides efficient algorithms to compute the forward kinematics simulation \( h(x) \) and the Jacobian matrix \( \frac{\partial h(x)}{\partial x} \) using the automatic differentiation.

In the absence of a deterministic model for the unpreventable measurement errors, the application of the extended Kalman smoother is the best choice to reduce the influences of instrumental errors, modeling errors and soft tissue artifacts. On one side, deterministic soft tissue artifacts and modeling errors are treated inadequately as stochastic noise. In the case of identical soft tissue artifacts on all markers of one body segment for example, it is impossible to distinguish between body segment motion and soft tissue artifacts [22]. On the other side, all available data and system knowledge is used to eliminate the variation in body segment lengths and filter stochastic measurement noise. This ensures smooth joint trajectories and consistent kinematic motion data.

2) GROUND REACTION FORCES: The raw ground reaction forces were filtered using a sixth order zero-lag low-pass filter with a cut-off frequency of 50 Hz. Individual force sensor offset and drift were compensated with a linear regression based on the measurements of the unloaded force plates during the first and last 10 s of each trial. In order to decompose the measured lateral ground reaction forces \( F_x \) and \( F_z \) and the measured vertical ground reaction force \( F_y \) for the locomotion trials in the event of mixed force plate contact during double support phase, parametrized transition functions determined using a multiple regression analysis were applied [25]. The shape of the transition functions, that approximate the ground reaction force decrease of the foot leaving the ground, is specified by force characteristics, the duration of the double support phase and the forward velocity of the subject. This parameterization ensures a smooth transition of the separated forces and allows consideration of step-to-step variability in human locomotion [26]. The ground reaction forces were used to estimate the center of pressure and detect individual events like left and right steps, squats or kicks. The number and type of detected events in each trial are summarized in Table I.

3) MUSCLE ACTIVITIES: The raw muscle activities were rectified and filtered using a root-mean square filter with a window size of 300 ms [27]. In addition, the filtered muscle activities were normalized to the maximum activity level over all trials of the subject. Each dataset provides filtered as well as normalized electrical activities for fourteen selected muscles in the legs. All processed data is combined in a single data file per dataset with individual data structures for each measurement system. The processed kinematic motion data includes the synchronized spatial positions of the thirty-six reflective markers, joint center estimates for fifteen joints as well as joint trajectories for twelve joints in arms, trunk, pelvis and legs. The processed ground reaction forces comprise synchronized, filtered and decomposed ground reaction forces and center of pressure estimates. The processed muscle activities contain filtered as well as normalized electrical activities for fourteen selected muscles in the legs. All processed data is combined in a single data file per dataset with individual data structures for each type of measurement and additional information about the trial. Detailed anthropometric parameters including age, gender, body mass, body height, origin as well as body segment lengths, masses, centers of mass and moments of inertia are provided in separate data files for each subject. Positions and lengths are given in mm, while all other data
is given in base and derived SI units. In addition to the biomechanical measurement data, the source code of the applied computational scripts for processing the raw measurement data is made available in a revision control system. A written documentation provides a reference guide on the structure and content of the data files and the computational scripts as well as information about the subjects, motion protocol, measurement setup and data processing. All data files, documents and computational scripts are open and can be obtained free of charge from the HUMOD Database website:

http://www.sim.informatik.tu-darmstadt.de/humod/

IV. CONCLUSION

The HUMOD Database offers versatile raw and processed biomechanical measurement data for the investigation, modeling and simulation of human motion dynamics focusing on lower limbs. The quite unique combination of kinematic motion data, ground reaction forces and muscle activity data with detailed anthropometric parameters allows to create and validate biomechanical models of the human locomotor system and to investigate and simulate the dynamics of the measured human motion including muscle driven actuation. Results of these investigations and simulations can provide helpful information in the design of musculoskeletal humanoid robots or assistive devices. The open, easy to access and free of charge availability of the database enables researchers to use comprehensive and high-quality biomechanical measurement data of several motion tasks. The provided raw data allows to derive additional biomechanical information or validate the given processed data. The open-source release of the computational scripts in a revision control system is supposed to improve understandability, transparency and quality of the provided data and applied data processing.

ACKNOWLEDGMENT

The authors would like to thank Martin Grimmer for his valuable support during the measurements.

REFERENCES


