Is Motion-Capture-Based Biomechanical Simulation Valid for HCI Studies? Study and Implications

Myroslav Bachynskyi[*†], **Antti Oulasvirta**[*], **Gregorio Palmas**[*], **Tino Weinkauf**[*]
[*]Max Planck Institute for Informatics   [†]Saarland University

ABSTRACT

Motion-capture-based biomechanical simulation is a non-invasive analysis method that yields a rich description of posture, joint, and muscle activity in human movement. The method is presently gaining ground in sports, medicine, and industrial ergonomics, but it also bears great potential for studies in HCI where the physical ergonomics of a design is important. To make the method more broadly accessible, we study its predictive validity for movements and users typical to studies in HCI. We discuss the sources of error in biomechanical simulation and present results from two validation studies conducted with a state-of-the-art system. Study I tested aimed movements ranging from multitouch gestures to dancing, finding out that the critical limiting factor is the size of movement. Study II compared muscle activation predictions to surface-EMG recordings in a 3D pointing task. The data shows medium-to-high validity that is, however, constrained by some characteristics of the movement and the user. We draw concrete recommendations to practitioners and discuss challenges to developing the method further.

Author Keywords
Empirical methods; physical ergonomics; validity; biomechanical simulation; optical motion capture.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Motion-capture-based biomechanical simulation is a combination of optical motion capture ("mocap") to record the 3D movements of the human body and the simulation of the biomechanics involved with those movements. The output of optical motion capture is the 3D motion of pointlights, from which performance measures such as speed or accuracy can be derived. Biomechanical simulation augments this with a rich description of human movement, including velocities and angles of limb segments, forces and moments at joints, and most importantly muscle activations [6]. Recently, the

method has gained ground in three disciplines: sports [26], medicine [17], and industrial ergonomics [19]. Mocap-based biomechanical simulation is becoming an option for HCI laboratories as well. The computational intensity is no longer a bottleneck and there is biomechanical simulation software available with human body models.[1] Moreover, optical motion capture equipment[2] matured and became affordable in the recent past. Modern marker-based systems track rigid points on the body with high precision and sampling rates.

However, to our knowledge, mocap-based biomechanical simulation is presently not deployed in HCI research. The method could allow cost-efficient estimation of physical ergonomics. In contrast, the standard ergonomics instruments have turned out to be too expensive, constraining, and specialized to be routinely used in HCI.[3] The mocap-based method studied in this paper does not restrict natural movement (much). This makes it possible to study full-body interactions where EMG recordings and other measurements are impractical.

Although work-related injuries and incidents have declined [2], the method could improve the analysis of user interfaces by better taking into account biomechanical stresses and muscle loads. Designs with cumbersome postures could be avoided, such as the "gorilla arm" when using vertical touch screens. Many times in the history of interaction design, this would have been of benefit. Consider the fate of the light pen: Touted as the ideal input device for information workers, it foundered as it caused strain in shoulder and arm muscles. We believe that the method bears even more potential in the design of novel interfaces. While the ergonomics of desktop-based interfaces have been intensively studied (e.g., [3]), interactions "beyond the desktop" need more attention. For example, tangible computing, tabletops and surfaces, mobile interaction, and various forms of 3D interaction all insist on novel postures and movement types for input. Such interactions can be too demanding for long-term use, or they can be too easy as in the case of exergames. Understanding the physical ergonomics of full-body motion is an important part of designing such interfaces. Furthermore, the performance data from the captured movements and the ergonomics data from the biomechanical simulation can be analyzed in a combined manner to find good trade-offs between them.

---

[1]E.g., OpenSim, AnyBody, LifeModeler, and SantosHuman)
[2]E.g., Vicon, PhaseSpace, OptiTrack
[3]Examples of physical ergonomics instruments rarely used in HCI studies are goniometers, videometry, and EMG.
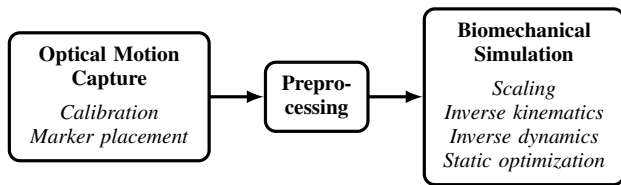
Figure 1: Steps in motion-capture-based biomechanical simulation [24].

This paper critically examines the validity of the method for the needs of HCI, which are different from sports, medicine, and industrial ergonomics. First, in medicine and sports, biomechanical simulation is mostly used with the purpose of deeply understanding one particular subject or a special group. In contrast, HCI research must be able to address large samples with varying age groups and other characteristics. Second, typical analyses in sports and medicine focus on a particular muscle or muscle group, tendon, or joint, such as the knee moment during jumping or landing. In contrast, HCI research is hardly interested in anatomical details, but in the *effects* of motion: uncomfortable postures, stress, load, and fatigue. It is clear that muscles have special importance in HCI. Third, movements in sports and medicine are often gross and, in the case of athletes, very fast. In contrast, HCI focuses on middle- and small-range movements, such as those involving the hand when using input devices. Fourth, sports and medicine are often interested in the lower limbs and analyze gait. In contrast, both industrial ergonomics and HCI tend to have more emphasis on upper extremities. Fifth, industrial ergonomics and HCI have interest in aimed movements to external targets. However, whereas industrial ergonomics has looked at movements to physical objects [18], setups in HCI often involve dynamic displays and dynamic contact forces on surfaces. Such forces are not directly observable using motion capture, but must be recorded with external instruments and integrated into the simulation.

Due to the novelty of mocap-based biomechanical simulation, it has not yet been extensively validated. Validity here refers to the accuracy of the biomechanical simulation results when compared to ground truth. Because the other outcomes of biomechanical simulation are better known, we focus on the validity of muscle-related predictions in user groups and movements that are characteristic to HCI. As we will show, every step of the method pipeline (Figure 1) has different sources of error, which can accumulate. It is not known from previous work whether these errors pose severe problems for HCI studies. What is known from previous work is that the current biomechanical models cannot be considered as "valid in general," but they have to be validated for each application or task [16]. These concerns call for rigorous research to assess validity in HCI tasks.

In our view, the method should be developed for HCI as a *generic* method, which shall be available without strong assumptions about specific equipment or movement ranges. In contrast, previous work outside HCI has developed task-specific simulations, such as a model of workers in reaching and assembly tasks [18], or simulations of prostheses in medicine [17]. Moreover, when mocap-based biomechanical

simulation is used in medicine and sports, the parameters of the human body model are often fine-tuned to each participant using EMG, isokinetic dynamometers and other measurements besides motion capture. Such instrumentations can restrict the natural movement, narrow the scope too much, and are generally too time-consuming for regular HCI studies. For HCI, we insist that the method should be usable without extensive manual modeling, and the scope of biomechanical simulation should be decidable *a posteriori*.

The contribution of this paper is two-fold. First,this is the first work applying and validating a state-of-the-art biomechanical model *with muscles* for HCI tasks. Second, in broader context, this is the first validation of muscle activation predictions of a state-of-the-art model for whole-arm movements in all directions. Two studies are presented that assess the method both technically and against ground truth measurements in typical HCI tasks. As it is impossible to cover all HCI tasks, we carefully selected a broad yet representative set with diverse requirements. **Study I** assesses technical feasibility in five scenarios: a full-body dance game, full-body flight simulator controls, mouse pointing, multitouch gestures, and typing on a keyboard. Our focus is on the technical feasibility, that is, whether the simulation can yield reliable outcomes for such diverse tasks. Moreover, the inspection of mouse and keyboard allows us to compare our predictions to previous work where EMG was used. **Study II** presents a 3D in-air pointing task with 16 subjects and directly compares the muscle activation predictions from the biomechanical simulation against measured ground truth data obtained using surface EMG. Our focus is on the predictive value of muscle activation predictions. The experimental design allows us to gauge inter-subject differences, differences between muscles, as well as to compare against self-reports of fatigue.

BACKGROUND AND RELATED WORK

Biomechanical simulation "reverse engineers" observed motion to explain it in terms of anatomical events. Its input is the movement of pointlights in 3D space. When accompanied by information on how the pointlights map to the human anatomy (mapping and scaling), motion is first explained as rotations of joints (inverse kinematics). Then, given mass distribution of the body, required forces at joints are estimated (inverse dynamics). Finally, given muscle anatomy, plausible muscle activations are estimated (static optimization).

We here compare the outputs to existing measurements in physical ergonomics. We then review previous work in validating the outputs against ground truth.

Outputs: Comparison to Traditional Measurements

Motion capture and biomechanical simulation create a high-dimensional description of a user's movement. The output variables are best understood as descriptors of *physical ergonomics costs*, the anatomical and physiological costs of movement [23, 30]. Within this scope, mocap-based biomechanical simulation can estimate six out of eight measures that would normally require specialized measurement instruments (Table 1).

Instrument	Joint angles	Posture	Kinematics	Forces and moments at joints	Muscular load	Muscular activation	Fatigue	Experience
Goniometers	●	○	●	○	○	○	○	○
Videometry	●	●	●	○	○	○	○	○
Accelerometers	○	●	●	○	○	○	○	○
Force plates	○	○	●	●	○	○	○	○
Dynamometers	○	○	○	●	○	○	○	○
Surface EMG	○	○	○	○	◐	●	◐	○
Self-reports	○	○	○	○	○	○	◐	●
Biom. simulation	●	●	●	●	●	●	○	○

● fully covered   ◐ partially covered   ○ not covered

Table 1: Comparison of biomechanical simulation against traditional instruments for physical ergonomics costs.

Joint angles are indicators of movement constraints and extreme postures and often measured by labor-intensive videometry, or goniometer measurement, which is limited to a few joints at a time and can perturb the movement. *Posture* is the state of the whole kinematic tree. It predicts overloading and musculoskeletal stress. In mocap-based biomechanical simulation, inverse kinematics yields angles and posture.

Kinematics describes angles and the distribution of loads and mass during movement and predicts overloading and repetitive-strain injuries. Force plates, on-limb accelerometers and on-joint friction/bending sensors can be used, but these have limited coverage, are cumbersome to apply, and can influence movements. *Moments and forces at joints* can estimate the overall energy expenditure [14] and also point to arthrokinetic strain and stress. Moments at joints are the sum of muscle forces multiplied by moment arms. Dynamometers are used in sports sciences but are limited to static setups and cover one movement type at a time. Mocap-based biomechanical simulation estimates moments and forces based on the outputs of inverse kinematics and full mass of the participant, assuming standard mass distribution.

Muscular load is the force produced by a muscle for a movement. Direct measurement of muscular forces is intrusive, but surface EMG [5] (sEMG) can be used to estimate it if parameters such as cross-sectional area are known. *Muscle activation* refers to the recruitment of muscle fibers by action potential induced by motor units. sEMG can be used for measurements, but it is limited to preselected muscles. In mocap-based biomechanical simulation, these two are given by static optimization.

Fatigue is the state of a muscle when it cannot produce its maximum force. It is reflected in the sEMG signal. Muscular fatigue can be described by total mechanical energy expenditure of muscle, which can be calculated from muscle activations integrated over time, estimated by static optimization. Presently, the best way to measure muscular fatigue is to infer it from the EMG signal and to use self-reports.

Self-reports are verbal reports of effort, fatigue, and stress; they are typically measured via questionnaires administered after a task. A workload questionnaire widely used in HCI is NASA-TLX that taps to some of these aspects.

Biomechanical Simulation in Human Factors and HCI
To our knowledge, there is no previous research applying or validating biomechanical simulation *with muscles* to core HCI tasks. However, biomechanical simulation has been combined with motion capture in a few earlier tools for ergonomics, and applied to cases in office safety assessment [11] and automobile assembly analysis [7]. These implementations included simplified models that lack muscles or necessitate external EMG recordings [7].

Previous Validation Studies
Although mocap-based biomechanical simulation has gained ground only recently, some steps of the simulation have been known for decades and are more thoroughly understood [30] than muscle activation models.

Error in joint angle prediction has been estimated to be within 1 degree for flexion–extension and abduction–adduction, and within 3 degrees for axial rotation [15]. Mean joint dislocations were smaller than 0.5cm, which should be accurate enough for HCI. Forces and moments have been validated in a study that compared the output of inverse dynamics to *joint moments* calculated with a machine learning algorithm from an EMG signal for the knee [14]. Model fit was high: $R^2 = 0.91 \pm 0.04$.

The only studies looking at the validity of muscle activation predictions of real movement of whole limbs involve lower extremities and consider gait or running [8, 25]. They compare predicted muscle activations against sEMG. Although there are a few validation studies of simulations related to upper extremities, all of them are constrained and simulate only one-dimensional movements of a single joint. The previous validations of lower extremities or separate joints of upper extremities do not generalize to the upper extremity model and cannot be considered as "valid" for HCI tasks [16]. The only study with a comprehensive upper extremity model reported, qualitatively, a lack of agreement in a comparison of predicted muscle activations and recorded EMG of 3 muscles for single specific reaching movement [4]. Please see the *Auxiliary materials* for a summary of previous validation studies. Study II is the first exhaustive validation of upper extremity models for whole-arm aimed movements in all directions and locations.

SOURCES OF ERROR IN BIOMECHANICAL SIMULATION
To motivate the need for a validation study, here we outline the most significant sources of error in the method (Figure 1).

Generally speaking, the estimations of joint and muscle activities should be treated as *hypotheses* made possible by a strong prior: an anatomically correct but generalized full-body model. The two most important sources of error are, first, the precision of motion capture that affects all computations "downstream." State-of-the-art marker-based tracking is accurate to the millimeter level. For example, we used the

PhaseSpace system in the studies which allows $1/5$ mm accuracy at 480 Hz. Second, every subsequent step in the simulation pipeline introduces unique sources of error, some of which are accentuated in HCI studies:

1. Marker Placement and Mapping

Biomechanical simulation depends on a reliable mapping between the body model and pointlights in the 3D mocap data. There are guidelines for marker placements that increase reliability by identifying *anatomical landmarks* that are more matchable (e.g., acromion, elbow) [13]. The mapping, called the *virtual marker set*, is manually defined after the data collection by the experimenter using a GUI.

Sources of error: First, typical optical tracking allows only a limited number of markers on the body (e.g., our system has 38). The experimenter must define which limbs to track and which to leave out. For example, tracking the articulation of hands is limited, if the rest of the upper torso should be tracked as well. Second, the placement of virtual markers on the model can never be perfect, since the virtual body model does not have the exact same geometry, and because markers are placed with an offset from the bones that mark the landmarks. Third, mapping is less accurate in segments that are farther away from anatomical landmarks.

2. Scaling

Every user needs to be scaled to match the anatomy of the generalized human model. *A measurement set* is a set of marker pairs and body parts that are scaled according to the ratio of distances between virtual and physical markers. The model size and weight are adjusted on the basis of the measurement set or from manual measurements. *Automatic Marker Adjustment* is then done based on data from a calibration pose of the user. It adjusts marker positions by means of inverse kinematics, which minimizes the errors between virtual and physical marker positions.

Sources of error: First, scaling assumes that the distribution of mass in a body is a linear function of the model's distribution. Second, automatic marker adjustment can err due to improper weight distribution, causing correctly placed markers to be misplaced.

3. Inverse Kinematics (IK)

IK calculates generalized coordinates that describe a skeletal movement in terms of angles between bones at joints, and translations and rotations of the human model relative to ground. It minimizes the weighted least-squares distance between physical markers and corresponding virtual markers.

Sources of error: First, markers can drift during movement due to non-rigid skin movement. Second, joints are modeled as simple "hinges" and omit for example translation at joints [23].

4. Inverse Dynamics (ID)

ID calculates forces and moments at joints produced by a movement given as a generalized-coordinate sequence. External forces can be added to the simulation at this step — for example, if recorded by force plate, force transducers, or dynamometers.

Sources of error: First, measurements of external forces can be imprecise or temporally or spatially out of synch with the pointlights. Second, mass distribution and anatomy of a given user may differ from that of the model, causing inaccuracies.

5. Static Optimization (SO)

SO resolves the required activations of muscles by minimizing *total* muscle activation as its objective function. It uses two muscle models as constraints: ideal force generators and muscles constrained by force–length–velocity properties.

Sources of error: First, SO assumes that people move "optimally" in terms of minimizing total activation, which cannot be assumed in many HCI tasks. Second, muscle anatomy and strength may differ between the user and the model. Third, movement speed may be an issue: For slow movements, activation patterns could be identified incorrectly, because humans can use a different activation strategy, or use smaller musculature to move.

GENERAL SETUP

The two studies presented in this paper examine HCI-relevant motor control tasks in 3D space. The system we utilize in the studies represents the state-of-the-art. Our recordings are done with high-end, commodity motion capture equipment: the *PhaseSpace* system with *Impulse* cameras tracks a full-body suit and gloves with flexibly attachable marker positions. For simulation, we use *OpenSim* [6], the only comprehensive open source simulator. It supports editing of the musculoskeletal model, scripting, and visual investigation of the results in a GUI. We use the SIMM Full Body Model, which combines measurements from several anatomical studies [10] best representing an average adult male. It contains models of 118 bones, 86 joints, and 285 muscles.

In the validation studies, the following practices were followed:

- *Marker placement* guidelines were followed [13]. Some trackers, like ours, come with a marker suit that has some pre-positioned markers, and not all guidelines can be strictly followed. For studies of aimed movements, it is necessary to add further markers on the *end-effectors* (e.g., index finger).

- *Force sensors* were used when external forces are involved — for example, ground reaction force if a user is jumping — or any other force is exerted on a physical object. We used Micro Load Cell (0–20 kg) with PhidgetBridge mounted under surfaces to-be-touched.

- *Artifact removal*: For preprocessing, we first clean the data from marker occlusions and artifacts. These may arise from shifts in the set of cameras observing a particular marker and are typical when the user moves around or the scene is crowded. Both issues manifest themselves as "jumps" in motion paths that can be reliably identified from the second derivative of coordinate values that are further than 2 SD from the mean. To fill gaps in the remaining data, we have experimented with several interpolations and found linear interpolation to yield sufficient quality.

- *Smoothing* of motion trajectories is necessary for ID and SO to compute on smaller-scale movements. We have tried

standard methods, such as Butterworth, but found Kalman smoothing [21] to produce the best results.

- *Calibration for target registration* is necessary when the study involves physical targets, such as in Study II. We have done calibration by asking users to touch the center of each target with an end-effector equipped with a marker. We perform calibration for every session to avoid the effects of between-session changes in the coordinate system of the mocap sessions.

- *Additional reserve actuators* can be added to the muscles in the model to ensure the existence of a solution in SO [24]. Generally this should be avoided, but it is necessary if the model contains muscles too weak to explain observed movements. We report when this was done.

In all other aspects, we follow the manual of OpenSim [24]. For examples of the setup and outputs, we refer the reader to the supplementary video.

STUDY I: FEASIBILITY IN FIVE HCI TASKS

The first study addresses technical feasibility in five HCI tasks carried out by a subject. The goal in selecting the tasks was to cover a wide range of movements with different velocities and varying number of contributing limbs and muscle groups. We consider mocap-based biomechanical simulation to be feasible for a specific task if all computational steps can be completed without errors. Typical error conditions are abrupt points appearing in the IK output, muscle activations approaching maximum, or large reserve activations appearing in SO. Furthermore, we compare our results against the existing literature on EMG measurements, where available.

Data Collection

A 36-year-old subject (male, 178-cm, 78 kg) volunteered for the study. The subject is right-handed and has no perceptual, neurological, or cognitive deficits. The *PhaseSpace* motion capture system with 12 Impulse cameras at 480 fps was used to record the movement of 43 active markers. In tasks 3, 4, and 5, a force plate of our own construction measured external forces. Interactive software was used in tasks 1, 3, and 5. The tasks were performed in a single three-hour session. This allowed us to use a single calibration and scaling. Details of the tasks follow:

1. **Full-body dance game** involves configural movements of the full human body. It was performed to a song from *Just Dance 2* on the Nintendo Wii.

2. **Plane control** involves steering a plane through continuous aimed movements of the upper part of the human body. Three control schemes were used: The first used a "bird paradigm," the second a steering-wheel paradigm. In the third, the arm was lowered and flexed. The subject had to mimic the motions of a person "flying" in a video as accurately as possible.

3. **Mouse pointing** involves fine-grained movements that deploy muscles from the shoulder down. It consisted of lateral reciprocal aiming movements performed with a commodity mouse. Control-to-display ratio was varied in three conditions: a ratio of 4, 10, and 18 (scale: 1–20). Four

Task	IK	ID	SO
Dance game	●	◒	◒
Flight	●	●	●
Mouse	●	●	●
Typing	◒	◒	◒
Multitouch gestures	○	○	○

● full success   ◒ partial success   ○ failure

Table 2: Completion of simulation steps for five HCI tasks in Study I. See text for details.

target sizes and four distances were used, yielding 1280 selections in all.

4. **Multitouch gestures** on a surface involve fine-grained movements that employ the small muscles of the hand and the arm. We followed an existing gesture set [31]: rotation ($45\,\mathrm{deg}$, $90\,\mathrm{deg}$, and $180\,\mathrm{deg}$), pinch with two fingers ($4\,\mathrm{cm}$ and $10\,\mathrm{cm}$), pull with 2 fingers ($4\,\mathrm{cm}$ and $10\,\mathrm{cm}$), pinch with all fingers ($10\,\mathrm{cm}$), pull with all fingers ($10\,\mathrm{cm}$), drag with index finger (horizontal 4 cm and 10 cm, and vertical 4 cm and 10 cm), drag with four fingers (horizontal and vertical 10 cm), and tap with index finger. Each condition was repeated 50 times.

5. **Typing** involves fast simultaneous movements of multiple end-effectors and recruits small muscles. The participant typed his name as quickly and precisely as possible 50 times on a regular physical Qwerty keyboard.

In all tasks, the subject was trained prior to motion capture to reach a level of performance we considered representative for that task. Sufficient rest was provided throughout.

Results

Table 2 summarizes the success or failure of the biomechanical computations for the specific HCI tasks. Three out of five tasks were completely successful. For these tasks, the method is discriminative and the outputs are sensible as highlighted in Figure 2 using representative muscle activation patterns. In particular, Figure 2a shows a clear activation of upper back, shoulder and biceps muscles, when the dancer moves his arm up, and lower back, shoulder and triceps muscles when he moves his arm down. For the mouse pointing task (Figure 2b), different shoulder muscles (deltoideus anterior and pectoralis major vs. deltoideus posterior and medius) are activated for movements from left to right and vice versa. When typing (Figure 2c), the subject mainly used his middle finger, which is reflected in a higher proportion of activation of muscles controlling that finger.

To follow, we discuss our detailed observations during this study, in particular regarding the cases where the biomechanical simulation could not be applied (cf. Table 2):

Inverse Kinematics
Data from all tasks except multitouch gestures could be processed for IK. IK requires keeping RMS (root mean square) error within 2 cm and largest marker error less than 4 cm. Although such errors are considered to be normal for *full-body* simulation, they were too large for multitouch gestures where the movement size falls within this range. This manifested in
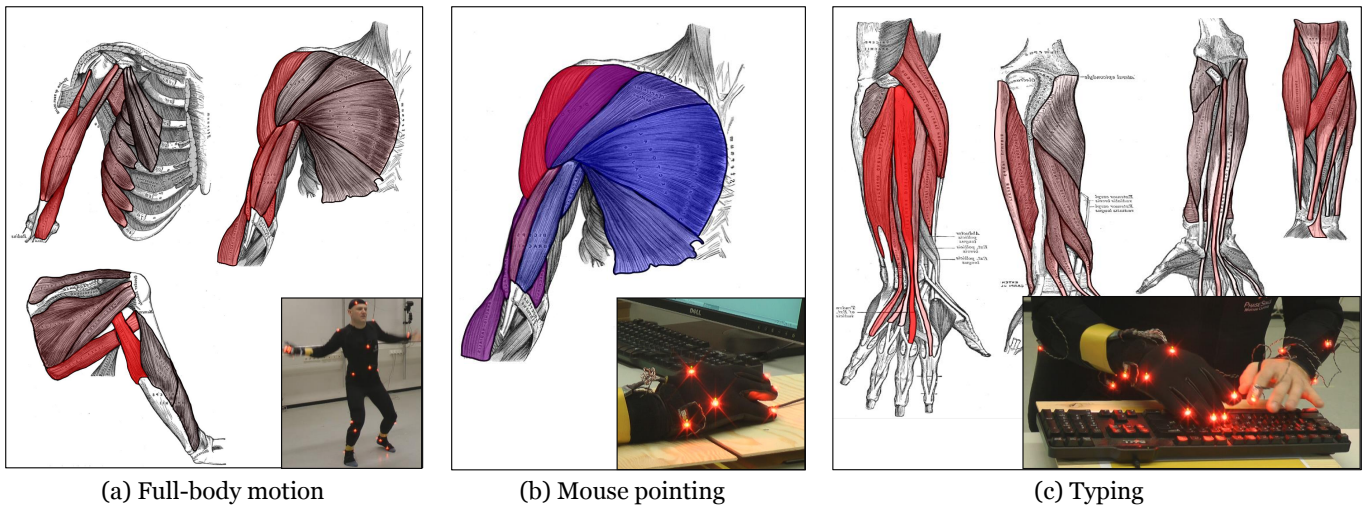
Figure 2: Selected muscle activation predictions for three HCI tasks in Study I. (a) Arm and shoulder muscle activations when moving arm up in a dance game. (b) Shoulder activations for mouse pointing: Activation marked in blue is for movement to left and red for to right. (c) Muscle activations in hand and arm muscles for typing two letters with the middle finger.

fingers being "stuck" in the same pose during the simulation. The other borderline case is typing: in our particular case IK was successful, because the participant used only 3-4 fingers with pronounced up/down movements when typing. Had the participant used the ten finger touch-typing technique, IK would have failed.

Inverse Dynamics
Task that are successful in the IK stage can proceed to ID. The only problem we encountered in this step were tasks where large external forces were applied. For dance, where the user stepped up and down, because we did not have force plates on the ground, we manually estimated ground reaction forces based on observation of movements. This approach improved the validity of full-body results, but the results are not reliable for the lower extremities.

Static Optimization
Our analysis here is limited to selected shorter segments ($<$ 2,000 ms) of the full recordings, because of computational intensity. Computing just 50 frames of SO for the dance task took 15 hours on a desktop computer.

One observed limitation is due to movements that are produced by muscles that are stronger than the corresponding ones of the generalized model. Another was that of motions where limbs are *overextended* or produce very fast abrupt movements. A successful simulation of the "bird" controller in the flight task required adding reserve actuators at the shoulder joints. This issue can be partially addressed by adjusting the *muscle* parameters of the general full-body model to the individual participant. Similarly to ID, SO needs correct external forces, so only part of the outputs from dance can be considered valid.

Agreement with literature
Muscle activation predictions of two cases—mouse pointing and typing—could be checked against an earlier report using EMG to compare the two input devices. Overall, muscle activation predictions agree with previous findings. While we cannot compare *absolute* values, relative activations agree: Trapezius descendens is 1.8 times more activated and deltoids are 1.3 times more activated when working with the mouse than when working with a keyboard [3].

Summary
We conclude that movement size was the determining factor for technical feasibility. Movements with smaller than 4 cm radius were not feasible with current marker setup, however they may be successful with a more complicated setup involving multiple markers on hand [29]. Moreover, abrupt and overextended movements posed a challenge. We also learned that including data from a force sensor (under surface, mousepad, and keyboard) observably improved the realism of ID and SO predictions, even for small-range movements and small external forces.

STUDY II: VALIDATION OF MUSCLE ACTIVATION PREDICTIONS

Study II addresses the *predictive validity* of muscle activations in HCI-relevant motions. Informed by Study I, we decided to focus on gross movements instead of small movements, and chose mid-air pointing gestures as the topic. This topic is relevant for research on interfaces that use computer vision and accelerometers for control.

Surface-EMG was measured for eight muscles of sixteen subjects while performing a 3D pointing task. The participants carried out in-air reciprocal pointing movements among targets in the reachable space of their arms (Figure 3). The experimental design covered the whole reachable space of the arm and allowed us to vary target size and amplitude of movements. Moreover, we had 16 users with varying demographics, which allowed us to learn about potential inter-subject differences and differences between muscles.

EMG was chosen as "ground truth" following existing recommendations [16] in studies of lower limbs [8]. We use EMG amplitude as ground truth. This can be justified because
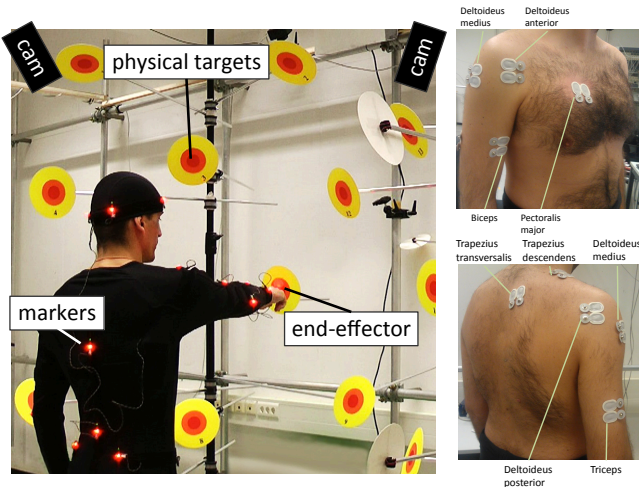
Figure 3: *Left*: Experimental setup for Study II: A 3D space with physical targets is scaled to the reachable area of a user's arm. End effector and physical targets are tracked in addition to the pointlights of the motion tracking suit. *Right*: Electrode placement for a validation study.

for a particular muscle larger EMG amplitude corresponds to larger muscle activation if 1) an EMG segment is recorded during con- or eccentric movement but not both, 2) cross-talk is minimized by precise electrode placement on larger muscles, 3) electrode displacement is minimal (here: for 5 out of 8 muscles), 4) no between-muscle or between-movement comparison is carried out for EMG amplitudes, and 5) the absolute value of EMG is ignored (we use Pearson correlation as the metric).

Method

Sixteen subjects (9 males and 7 females) with ages ranging from 21 to 36 ($M = 25.9$), height from 162 to 178 cm ($M = 170$), and weight from 61 to 79 kg ($M = 70$) were recruited. No subject had musculoskeletal or neural disorders, and every subject took part in some regular physical exercise.

Five targets from a total of 25 physical targets (see Figure 3) were selected for each subject by stratified sampling from five segments of the reachable space of the dominant arm: left upper outer, left lower outer, right upper outer, right lower outer, and central inner. There are three target sizes (20, 40, 80 mm). We recorded three trials for every pairwise combination of the five targets — in total, 30 trials per subject. The order of trials was randomized.

In addition to motion capture, surface EMG was recorded with a Myon 320 [22] and self-adhesive electrodes (Ambu Neuroline 720 00 S/25 with Ag/AgCI and conductive gel) at a sampling rate of 2000 Hz. All subjects confirmed that the EMG electrodes did not restrict their movements. Electrodes were placed on eight muscles: the pectoralis major, deltoideus anterior, deltoideus medius, deltoideus posterior, trapezius descendens, trapezius transversalis, biceps, and triceps (see Figure 3). The skin was prepared for electrode placement following recommendations (SENIAM [9]).

Retrospective self-reports were measured by a questionnaire. Subjects were asked to rate task difficulty and also the
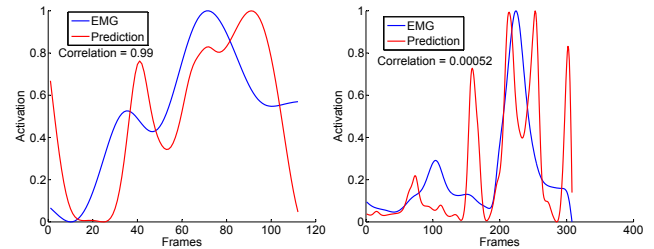

Figure 4: Examples of EMG vs. predicted muscle activation with high (*left*) and low (*right*) correlations.

stress/tension in the muscles of the arm, shoulder, back, and chest with a seven-point Likert scale [1].

Preprocessing and Analysis

All motion capture data was processed through IK and SO. Because of the computational cost of SO, we selected representative movements by choosing a movement for each trial with a movement time closest to the mean and which ended within the effective target.

Following the recommendation of the manufacturer, the DC offset was removed from the EMG data and frequencies below 20 Hz, above 500 Hz, and between 49 and 51 Hz (powerline interference) were filtered. The signal was then full-wave rectified and normalized according to maximum voluntary contraction. Then both EMG and the activations calculated via static optimization were low-pass filtered at a frequency of 4 Hz to create a linear envelope of the signals.

Results

For each movement ($n = 960$), we computed the Pearson correlation coefficients between the time series of the EMG signals and the corresponding SO activations of the studied muscles. The full distribution of the correlation coefficients can be observed in Figure 5a. The median of the correlation coefficients over the full dataset is $r = 0.48$. Examples of high and low correlations are given in Figure 4.

We also segmented the distribution based on different independent variables (i.e. muscle, participant, target size or location) and analyzed distributions within segments. Several observations were made:

- SO predicts better for larger muscles: deltoideus ($r = 0.64$) and trapezius ($r = 0.67$). For smaller and less-recruited muscles, correlations are small or negligible: biceps ($r = 0.27$), triceps ($r = 0.04$), pectoralis major ($r = 0.04$).
- SO better predicts gross movements, i.e. movements toward *large targets* (8 cm radius; $r = 0.59$) than toward small targets (2 cm radius; $r = 0.37$). Perhaps the recruitment of smaller musculature in the finer control of motion is well captured by neither sEMG nor SO.
- Correlations differ significantly among subjects (range: $0.33 < r < 0.62$), with the strongest correlations recorded for the oldest subject (36 years old, male). This is understandable given that the musculature in the model (SIMM-FBM) is based on measurements from adult males.
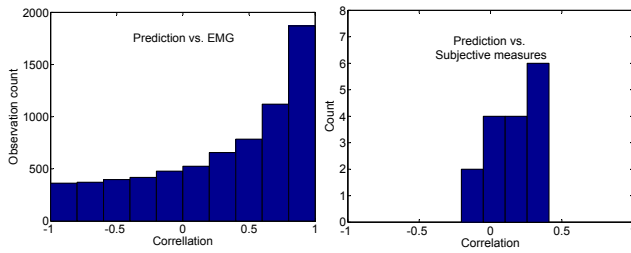
Figure 5: Distributions of correlation coefficients between predicted muscle activations and EMG (*left*) and between prediction and subjective measures (*right*)

- Correlations increase in the course of a session (average slope=0.11%, in total 3.14% over 30 trials), perhaps because muscles are getting fatigued, so that muscle parameters of partially fatigued young people are closer to the muscles of an average adult in the body model.
- Self-reports of stress/tension have only a weak correlation ($0.1 < r < 0.2$) with both recorded and simulated variables (Figure 5).

We did not find effects for the location of movement in the 3D space.

Conclusion

We conclude that the predictions of static optimization were poorer for small and non-contributing muscles, but adequate for the largest muscles and gross movements. In ideal conditions, predictive validity was high: If we consider only the larger muscles of participants whose anatomical parameters are close to the full-body model, and movements between large targets, then the median correlation is 0.81. Generally, middle-aged subjects were better predicted. Finally, self-reporting of fatigue was not predicted by the method.

SUMMARY AND RECOMMENDATIONS

Mocap-based biomechanical simulation consists of many steps and it was previously unclear how sensitive the method is to the accumulation of errors. The general thrust of the findings is that, presently, mocap-based biomechanical simulation is valid for some but not all HCI tasks. In the following, we summarize the lessons learned.

Table 3 provides a list of recommended best practices. Study I (Feasibility) found that movements of less than 4 cm amplitude are presently impossible to process by the standard simulation. The noise level is too high for such signals, even with state-of-the-art tracking equipment and careful marker placement. Very fast and overextended motions are problematic as well and require a temporary increase of the muscle strength during the simulation (reserve actuators). These results imply that movements involving small interactive surfaces or abrupt movements with large forces, as in exergames, are difficult. However, many improvements are possible (see *Future Work*).

The main result of Study II is encouraging, namely that the correlation of muscle activation predictions with ground truth data from surface-EMG is as high as $r = 0.48$ for the 3D aimed movements of sixteen random subjects. Regarding
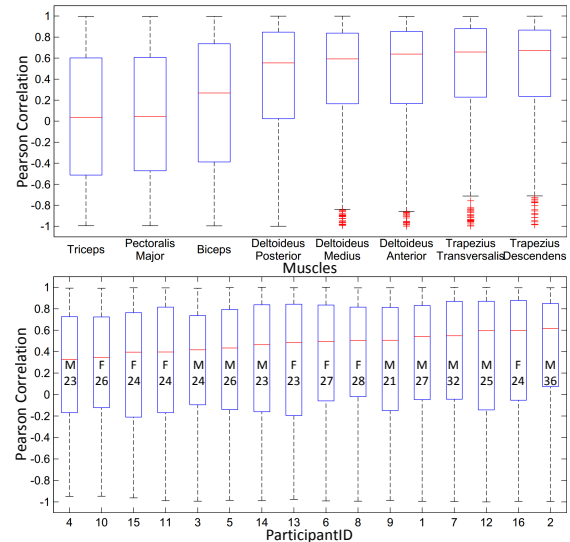

Figure 6: *Top*: **Correlations of muscle activation predictions with EMG recordings. Breakdown by muscle.** *Bottom*: **Breakdown by user.**

the design of a suitable HCI task, we observed the following choices to affect predictive validity:

- Movements to larger targets are better predicted than to smaller ones.
- Larger movements are better predicted than smaller ones, unless they become overextended.
- Faster movements are better predicted than slower ones.

The following observations are also important to consider when assessing the outcomes of biomechanical simulation:

- Activations of larger muscles are better predicted than smaller ones.
- For young participants correlations increase in the course of a session as muscles become more tired.
- The highest correlations are found for subjects who match closely to the full-body model.

For favorably chosen tasks and subjects, correlation against S-EMG was as high as 0.81. This is the case for middle-aged adult males that are the best match with the human body model. But both studies indicate that the highest muscle activations are very well predicted. Moreover, the results from mouse pointing and typing (Study I) are in consensus with findings of previous studies using EMG.

These results should be taken into account when sampling users and designing tasks for experiments.

FUTURE WORK

As there is increasing attention to biomechanical simulation in various disciplines, the involved technologies are likely to be improved by future work in different domains. We identified a number of issues that should be of particular concern to HCI researchers, as detailed in the following paragraphs.

In order to expand the method to fine motor movements like multitouch gestures on a surface, the most critical shortcoming to address is the limited movement size. Generally speak-

System Setup
S1. Follow anatomical guidelines for marker placement.
S2. Use additional markers for end effectors.
S3. Use force plates to record contact forces at surfaces.
Data Preprocessing
D1. Use reserve actuators for overextended movements.
D2. Use Kalman smoothing for more robust IK and ID.
Participants
P1. Middle-aged subjects are better predicted than young.
P2. Males are better predicted than females.
Tasks and Procedure
M1. Movements with large amplitudes are better predicted.
M2. Movements to large targets are better predicted.
M3. Fine movements of less than 4 cm of amplitude are not feasible.
M4. Faster movements are better predicted than slow.
M5. Prediction is better for movements recruiting larger muscles.
M6. Prediction improves when the muscle gets more tired over time.
M7. Over-extended motions are poorly predicted.

Table 3: Recommendations for improving accuracy of muscle activation prediction in HCI studies.

ing, the simulation succeeds if the noise level in its input data is significantly smaller than the actual movements. On the positive side, the motion capture system itself introduces relatively small errors. Some of the culprits are hard to avoid. For example, the locations of anatomical landmarks vary slightly across users, and therefore also the marker placement relative to these landmarks. Furthermore, the markers drift slightly during movements. However, movement size can be partially addressed by adding more markers per joint and fine-tuning the body model to the individual subject.

Related to this issue is the fact that the present body models work best with a middle-aged male. More work is needed in statistical body modeling that can account for statistical variation in body shapes and mass distributions. This way, the method will be applicable to a wider range of subjects. The model also makes simplifying anatomical assumptions, such as joints being "hinges," although for example the thumb has a saddle joint. Improving the accuracy of the model will improve the permissible size of movements. Moreover, more studies are needed to understand limitations with different demographics.

Our validation study showed no real relationship between predicted muscle activations and self-reports of stress and tension. The experience of the user is obviously important for HCI researchers, but it is not novel to find a low correlation between objective and subjective measures in HCI [12]. We suspect that the reason for the low correlation is that our subjects were not tired enough for stress to emerge. At present, we advise collecting said data via questionnaires, but future work should examine whether muscle fatigue can be estimated when more muscle parameters are known [23].

The standard biomechanical simulation considers movements "in a vacuum." We used force sensors to collect ground reaction forces at contact points on surfaces, but our setup is limited to recordings of a single sensor. For HCI investigators dealing with interactions that involve multiple objects (e.g., tangible computing), it will be important to have a way to collect and synchronize data from multiple easily attachable force sensors.

Finally, several practical issues need to be addressed. Although we have shown that data collection and processing is possible and worthwhile, the process is still unwieldy. Computation times of ID, IK, and SO are extensive: 15, 50, and 1800 seconds, respectively, on a desktop computer for an movement of the arm (500 ms). As an alternative to SO there are also two more sophisticated tools in the OpenSim called Residual Reduction Algorithm (RRA) and Computed Muscle Control (CMC) [6]. Together they should be able to provide more accurate results, although their computational cost is much higher. For application in HCI, RRA and CMC also need to be assessed and validated. Faster simulation methods exist [20, 28], but they have not yet been validated for HCI. But even the fastest simulation software cannot make up for the time-consuming setup of markers and other manual inspections and interventions that are required in the preprocessing phase. To support practitioners better, future work should seek to streamline the method. A promising direction is *markerless* motion capture (e.g., [27]). It allows unhindered mobility for the subjects, and the setup effort is reduced to a minimum. However, markerless motion capture has lower accuracy. It remains to be examined for which HCI tasks this lower accuracy is still sufficient.

CONCLUSION

The combination of motion capture and biomechanical simulation allows measuring both performance and physical ergonomics of a user interface within a single session and with relatively low costs for the setup. The benefit is that it allows obtaining a rich description of a user's movement: joint angles and posture, kinematics, forces and moments at joints, as well as muscular load and activation. Computational and manual costs are addressable. It is important to try to tap this rich source of information for HCI. As a first step, we studied the feasibility and validity of this approach for HCI in aimed movement tasks. The positive thrust of the present findings is that the method can already be used for a broad class of HCI studies. Thus, it is a viable alternative to traditional ergonomics instruments. We also identified some limitations. They can largely be addressed by technical improvements, which will expand the range of HCI studies that can benefit from mocap-based biomechanical simulation.

REFERENCES
1. Åhsberg, E., Garnberale, F., and Kjellberg, A. Perceived quality of fatigue during different occupational tasks. development of a questionnaire. *Industrial Ergonomics 20*, 2 (1997), 121–135.
2. Bureau of Labor Statistics. Workplace injury and illness summary – 2012. Washington, DC: United States Department of Labor, 2013.

3. Cooper, A., and Straker, L. Mouse versus keyboard use: A comparison of shoulder muscle load. *Industrial Ergonomics 22*, 45 (1998), 351 – 357.

4. Daly, M. *Adaptations following upper extremity resistance training in older adults*. PhD thesis, Wake Forest University, 2011.

5. De Luca, C. J. The use of surface electromyography in biomechanics. *Applied Biomechanics 13* (1997), 135–163.

6. Delp, S. L., Anderson, F. C., Arnold, A. S., Loan, P., Habib, A., et al. OpenSim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Trans. Biomedical Engineering 54*, 11 (2007), 1940–1950.

7. Du, J., and Duffy, V. G. A methodology for assessing industrial workstations using optical motion capture integrated with digital human models. *Occupational Ergonomics 7*, 1 (2007), 11–25.

8. Hamner, S. R., Seth, A., and Delp, S. L. Muscle contributions to propulsion and support during running. *Journal of biomechanics 43*, 14 (2010), 2709–2716.

9. Hermens, H. J., Freriks, B., Merletti, R., Stegeman, D., Blok, J., Rau, G., Disselhorst-Klug, C., and Hägg, G. *European recommendations for surface electromyography*. Roessingh Research and Development, The Netherlands, 1999.

10. Holzbaur, K. R. S., Murray, W. M., and Delp, S. L. A model of the upper extremity for simulating musculoskeletal surgery and analyzing neuromuscular control. *Ann. of Biomed. Eng. 33* (2005), 829–840.

11. Honglun, H., Shouqian, S., and Yunhe, P. Research on virtual human in ergonomic simulation. *Computers & Industrial Engineering 53*, 2 (2007), 350–356.

12. Hornbæk, K., and Law, E. L.-C. Meta-analysis of correlations among usability measures. In *Proc. SIGCHI* (2007), 617–626.

13. Kontaxis, A., Cutti, A., Johnson, G., and Veeger, H. A framework for the definition of standardized protocols for measuring upper-extremity kinematics. *Clinical Biomechanics 24* (2009), 246–253.

14. Lloyd, D. G., Besier, T. F., et al. An EMG-driven musculoskeletal model to estimate muscle forces and knee joint moments in vivo. *J. Biomech. 36* (2003), 765–776.

15. Lu, T.-W., and O'connor, J. Bone position estimation from skin marker co-ordinates using global optimisation with joint constraints. *J. Biomech. 32*, 2 (1999), 129–134.

16. Lund, M. E., de Zee, M., Andersen, M. S., and Rasmussen, J. On validation of multibody musculoskeletal models. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 226* (2012), 82–94.

17. Machado, M. F., Flores, P., Walter, J., and Fregly, B. Challenges in using OpenSim as a multibody design tool to model, simulate, and analyze prosthetic devices: a knee joint case-study. http://hdl.handle.net/1822/19411, 2012.

18. Mavrikios, D., Karabatsou, V., Pappas, M., and Chryssolouris, G. An efficient approach to human motion modeling for the verification of human-centric product design and manufacturing in virtual environments. *Robotics and Computer-Integrated Manufacturing 23*, 5 (2007), 533 – 543.

19. Miehling, J., Krüger, D., and Wartzack, S. Simulation in human-centered design–past, present and tomorrow. In *Smart Product Engineering*. Springer, 2013, 643–652.

20. Murai, A., Kurosaki, K., Yamane, K., and Nakamura, Y. Musculoskeletal-see-through mirror: Computational modeling and algorithm for whole-body muscle activity visualization in real time. *Prog. Biophys. Mol. Biol. 103*, 2 (2010), 310–317.

21. Musial, J., Verstraete, M., and Gobron, N. Comparing the effectiveness of recent algorithms to fill and smooth incomplete and noisy time series. *Atmos. Chem. Phys. Discuss 11* (2011), 14259–14308.

22. Myon Prophysics, http://www.myon-prophysics.ch/products/hardware/myon-320/.

23. Neumann, D. A., and Rowan, E. E. *Kinesiology of the musculoskeletal system: foundations for physical rehabilitation*. Mosby St. Louis, MO, 2002.

24. OpenSim documentation. http://simtk-confluence.stanford.edu:8080/display/OpenSim.

25. Pronost, N., Sandholm, A., and Thalmann, D. A visualization framework for the analysis of neuromuscular simulations. *The Visual Computer 27* (2011), 109–119.

26. Reinbolt, J. A., Seth, A., and Delp, S. L. Simulation of human movement: applications using opensim. *Procedia IUTAM 2*, 0 (2011), 186 – 198.

27. Stoll, C., Hasler, N., Gall, J., Seidel, H., and Theobalt, C. Fast articulated motion tracking using a sums of Gaussians body model. In *Proc. ICCV* (2011), 951–958.

28. van den Bogert, A. J., Geijtenbeek, T., Even-Zohar, O., Steenbrink, F., and Hardin, E. C. A real-time system for biomechanical analysis of human movement and muscle function. *Medical & biological engineering & computing* (2013), 1–9.

29. Vignais, N., Cocchiarella, D., Kociolek, A., and Keir, P. Dynamic assessment of finger joint loads using kinetic and kinematic measurements. *Digital Human Modeling Symposium* (2013).

30. Winter, D. A. *Biomechanics and motor control of human movement*. Wiley, 2009.

31. Wobbrock, J. O., Morris, M. R., and Wilson, A. D. User-defined gestures for surface computing. In *Proc. SIGCHI* (2009), 1083–1092.