

Sit-to-Stand Task on a Humanoid Robot from Human Demonstration

Michael Mistry, Akihiko Murai, Katsu Yamane, and Jessica Hodgins

Abstract—In this work, we perform the challenging task of a humanoid robot standing up from a chair. First we recorded demonstrations of sit-to-stand motions from normal human subjects as well as actors performing stylized standing motions (e.g. imitating an elderly person). Ground contact force information was also collected for these motions, in order to estimate the human’s center of mass trajectory. We then mapped the demonstrated motions to the humanoid robot via an inverse kinematics procedure that attempts to track the human’s kinematics as well as their center-of-mass trajectory. In order to estimate the robot’s center-of-mass position accurately, we additionally used an inertial parameter identification technique that fit mass and center-of-mass link parameters from measured force data. We demonstrate the resulting motions on the Carnegie Mellon/Sarcos hydraulic humanoid robot.

I. INTRODUCTION

As humans we are generally able to execute dynamic and dexterous motions with relative ease: walking, jumping, and dancing are a few examples. The motion of sitting and standing up from a chair is also a seemingly simple and commonplace motion we do everyday, but in reality is quite complex and dynamic, and can become challenging and even dangerous as we grow older. Success of this task requires whole body coordination and balance, careful control and placement of the body’s center-of-mass (COM) as it moves from a large support basin under the chair to much smaller one above the feet, all the while coping with sudden contact state and force changes. As roboticists, we are interested in such a motion and aim to better understand its complexities by synthesizing natural human sitting-to-standing on a humanoid robot platform. To do so, we will first record the motions of humans naturally standing up from chairs, and transfer that motion onto a humanoid robot. However, this transfer process is non-trivial as simple playback of the recorded motion will generally fail, because the human and robot have different kinematic and dynamic properties. Therefore, we will also record the contact forces of the humans as they execute the motion, and use the human’s COM trajectory as a guide for the robot.

In this work, we outline our procedure for generating and implementing natural human-like sit-to-stand motions from human demonstration onto a humanoid robot. Because the center-of-mass trajectory of both the human and robot are an essential component of this procedure, we will first discuss how we estimate the sagittal plane COM of the robot using joint torques and contact forces recorded from static

postures. The inertial parameter identification method can provide a more accurate estimate than the manufacturers’ CAD data. Next, in order to map the recorded human motion to the robot motion, we will discuss our inverse kinematics procedure. The inverse kinematics optimization attempts to match the human’s kinematic motion as closely as possible, while tracking the human’s COM trajectory. Finally we will demonstrate the resulting motion implemented on the Carnegie Mellon/Sarcos hydraulic humanoid robot.

II. ROBOT CENTER OF MASS ESTIMATION

Planning a successful sit-to-stand motion without falling requires accurate knowledge of the robot’s center-of-mass position throughout the trajectory. Not only must the robot’s COM be placed above the feet by the end of the movement, we also wish to track the human demonstrator’s COM trajectory as closely as possible, to realize natural human-like motion. If we know the masses and relative COM locations of each individual link on the robot, we can easily obtain the total COM location of any configuration via forward kinematics. Although the manufacturers of most robots provide estimates of masses and COMs of each link computed by computer aided design (CAD) software, these estimates do not account for several components including (in the case of our robot) hydraulic hoses, electronics, and onboard computers. These extra components contribute approximately 50% additional mass to our robot.

To obtain a more accurate estimate of the robot’s center of mass position, we employ a simplified version of a traditional least-squares optimization inertial parameter identification technique. As shown in [1], [2] (and extended to floating base systems by [3], [4]), the complete dynamics equations of a floating base rigid body robot can be written to be linear with respect to inertial parameters:

$$\mathbf{K}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}})\phi = \mathbf{S}^T \tau + \mathbf{J}_C^T(\mathbf{q})\lambda, \quad (1)$$

where $\mathbf{q} \in \mathbb{R}^{n+6}$ is the robot configuration vector (n joints plus 6 DOF floating base), $\tau \in \mathbb{R}^n$ is the vector of actuation torques, $\mathbf{S} = [\mathbf{0}_{n \times 6} \quad \mathbf{I}_{n \times n}]$ delineates the actuated and unactuated degrees-of-freedom (note that the floating base is unactuated), \mathbf{J}_C is the contact Jacobian, and λ is the vector of contact forces. Finally, $\phi = [\phi_0^T \quad \phi_1^T \quad \dots \quad \phi_n^T]^T$ is the vector of inertial parameters of $n+1$ links (n joints plus the floating base link), and each link has 12 inertial parameters, as follows:

$$\phi_i = [m_i \quad m_i c_{x_i} \quad m_i c_{y_i} \quad m_i c_{z_i} \quad I_{xx_i} \quad I_{xz_i} \quad I_{xzi} \quad I_{yy_i} \quad I_{yz_i} \quad I_{zzi} \quad f_{c_i} \quad f_{v_i}]^T, \quad (2)$$

All authors are affiliated with Disney Research Pittsburgh, 4615 Forbes Avenue, Suite 420, Pittsburgh PA 15213, USA [mmistry, amurai, kyamane, jkh]@disneyresearch.com

K. Yamane and J. Hodgins are additionally associated with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA

where m_i is the mass of link i , the vector $\langle c_{x_i}, c_{y_i}, c_{z_i} \rangle$ is the link's center of mass position, and $\langle I_{xx_i}, I_{xy_i}, I_{xz_i}, I_{yy_i}, I_{yz_i}, I_{zz_i} \rangle$ are the 6 independent components of its inertia tensor. Additionally f_{c_i} and f_{v_i} are coulomb and viscous friction, respectively.

By considering only static postures of the robot (i.e., when $\dot{\mathbf{q}} = \ddot{\mathbf{q}} = \mathbf{0}$), we further simplify (1), by considering only the gravity contribution of \mathbf{K} :

$$\mathbf{K}_g(\mathbf{q})\phi_g = \mathbf{S}^T \tau + \mathbf{J}_C^T(\mathbf{q})\lambda, \quad (3)$$

with each link's parameters defined as:

$$\phi_{g,i} = [m_i \quad m_i c_{x_i} \quad m_i c_{y_i} \quad m_i c_{z_i}]^T \quad (4)$$

By measuring the contact forces and torques while the robot is in several static postures, we can collect several data points and stack them in the following manner:

$$\begin{bmatrix} \mathbf{K}_g(\mathbf{q}_1) \\ \mathbf{K}_g(\mathbf{q}_2) \\ \vdots \\ \mathbf{K}_g(\mathbf{q}_N) \end{bmatrix} \phi_g = \begin{bmatrix} \mathbf{S}^T \tau_1 + \mathbf{J}_{C_1}^T \lambda_1 \\ \mathbf{S}^T \tau_2 + \mathbf{J}_{C_2}^T \lambda_2 \\ \vdots \\ \mathbf{S}^T \tau_N + \mathbf{J}_{C_N}^T \lambda_N \end{bmatrix}, \quad (5)$$

or in more simple notation as:

$$\bar{\mathbf{K}}_g \phi_g = \bar{\mathbf{f}}, \quad (6)$$

where the bar notation refers to an augmented stack of N matrices, and the following vector of total generalized force:

$$\mathbf{f}_i = \mathbf{S}^T \tau_i + \mathbf{J}_{C_i}^T \lambda_i \quad (7)$$

The inertial parameters, which contribute to gravity forces, can then be estimated using ordinary weighted least squares:

$$\hat{\phi}_g = \left(\bar{\mathbf{K}}_g^T \bar{\mathbf{K}}_g \right)^{-1} \bar{\mathbf{K}}_g^T \bar{\mathbf{f}}, \quad (8)$$

which will compute the parameter set $\hat{\phi}_g$ that minimizes $\left\| \bar{\mathbf{K}}_g \hat{\phi}_g - \bar{\mathbf{f}} \right\|_2$.

A. Reduced Robot Model

One major challenge of data-driven inertial parameter estimation approaches is acquiring a rich data set that evenly samples the entire configuration/motion space. Humanoid robots compound this problem, since they are generally high degree of freedom systems and the useable workspace is limited to a set of stable postures. However, we may be able to reduce the complexity by estimating only those parameters relevant to the task. In our sit-to-stand task, the forward progression of the center of mass is most critical for robot stability. Therefore, we decided to further simplify the identification problem by considering a reduced dimensionality model of the robot with motion restricted to the sagittal plane. Therefore for center of mass parameter estimation, we use a 5 link model with 4 actuated DOFS: ankle, knee, hip, and torso (again the base link is unactuated) (Fig. 1 (right)). One or two arm links would also have been reasonable to add to the planar model, however, we decided to neglect the arms due to their small mass relative to the heavier legs and torso. By simplifying the inertial parameter problem to

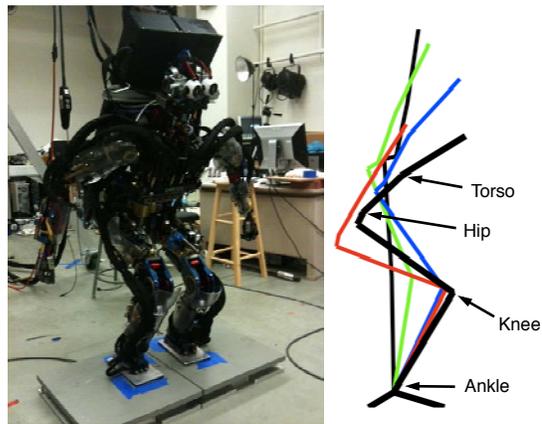


Fig. 1. Left: The Carnegie Mellon/Sarcos humanoid robot standing in a static squat position. Contact forces are measured using force sensing plates shown below the robot. Right: Some sample postures used for data collection.

only reflect the center of mass of a 5 link planar robot, we have reduced the total number of inertial parameters from 468 (the full robot with 39 links) to 15 (5 links with 3 parameters each: mass and x, y positions). Note, however, by restricting estimation to a planar model, we sacrifice the identifiability of individual links. Instead we are limited to only being able to identify groups of parameters in linear combinations with each other [5]. While identification of individual link parameters may be erroneous, the aggregated center of mass estimation of all links can still be accurate.

B. Data Collection and Estimation

To collect data for center of mass estimation, we place the robot into several statically stable standing postures of various ankle, knee, hip, and torso configurations. In each posture, we record joint torques of the robot using onboard load cells and the contact forces and moments of each foot with individual force sensing plates (AMTI Biomechanics Force Platform) (Fig. 1 (left)). We recorded a total of 88 different standing postures including upright straight knee standing and low bending squatting (the right and left sides of the robot were kept symmetrical). The total range of displacement of each joint in the entire data set is 1.47, 1.69, 0.56, and 0.56 radians for hip, knee, ankle, and torso joints respectively. Fig. 1 (right) shows five representative postures from the data set.

To evaluate the quality of our estimation procedure, we separated the data set into 90% for training and 10% for evaluation (9 of 88 postures). For each evaluation posture we compared the predicted center of mass position versus the recorded center-of-pressure position obtained from the AMTI force sensor. RMS error for forward/backwards center of mass position was 0.0196 m in the evaluation set.

III. MOTION GENERATION FROM HUMAN DEMONSTRATION

We wish to understand how a human controls his whole body in an environment with multiple and variable contact

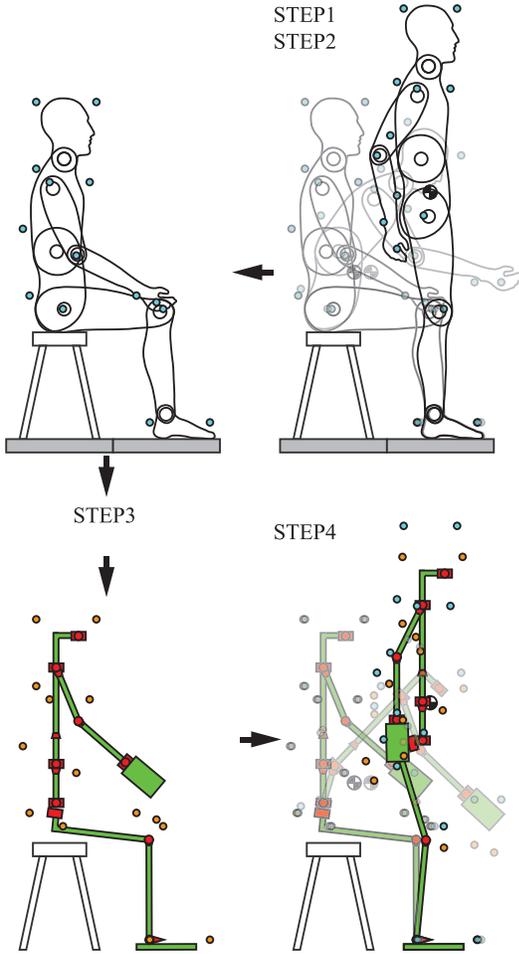


Fig. 2. Procedure for generating the robot motion trajectory from human motion capture data. STEP1: measure the human motion data by the optical motion capture system, STEP2: estimate the COM trajectory from the force plate data, STEP3: overlay the marker data onto the measured humanoid robot posture, and STEP4: solve the inverse kinematics computation with 1) the COM trajectory (hard) and 2) the marker trajectory (soft) as the constraints.

forces, and realize a natural human-like motion on a humanoid robot in a complex environmental contact. Many robotics researches have considered the center of mass (COM) as an essential point for the humanoid robot motion control, and used it for complex tasks such as the control of biped walking or the response to a sudden external force [6], [7]. Indeed, for human sit-to-standing motion, biomechanics researchers measured kinematic variance patterns during repeated standing motions and concluded that precise COM trajectory control in the sagittal plane is a key component to successful standing [8]. Therefore, in this work, our goal is not only record and playback natural human standing motion, but also accurately measure the human's COM trajectory experimentally, and plan the humanoid robot's motion so that its COM will trace that of human. Here the following difficulties should be solved:

- 1) The actual COM position of a human is difficult to obtain, because accurate kinematics and mass properties

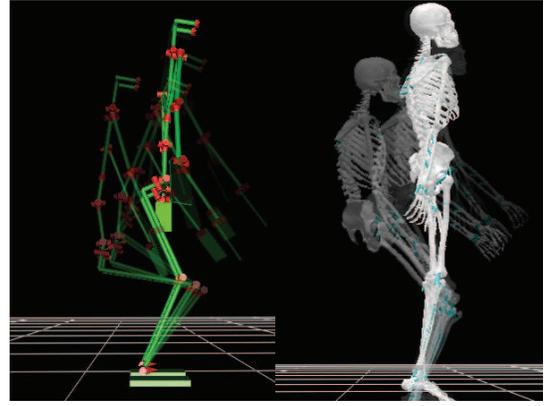


Fig. 3. Left: Kinematic model of humanoid robot. Right: 155 DOF skeletal model used to estimate the human's COM position.

of humans are difficult to estimate.

- 2) The kinematics and mass properties differ between humans and the humanoid robot.
- 3) The accurate dynamics properties of the humanoid robot are unknown and differ from the manufacturer's CAD data.

We generate the humanoid robot's motion in the following steps while addressing these problems. An overview of the process is provided in Fig. 2.

First, using an optical motion capture system (VICON), we measure the experimental human motion data as a set of 3-D marker trajectories $\mathbf{m}(t) \in \mathbb{R}^{3n_m}$, where n_m is the number of markers. Two force sensor plates (AMTI) also measure ground contact forces and are synchronized to the motion capture data spatially and temporally. One plate is positioned beneath the chair and the other is under the subject's feet. From the six axis contact force measurement of each plate, we can compute the position of the center of pressure (COP) $\mathbf{p}_{\text{COPexp}}(t) \in \mathbb{R}^2$.

To estimate the human's COM trajectory, we use a 155 DOF skeletal model [9], [10], (Fig. 3, right) with an average set of human inertia properties. We use this model to map the motion capture markers to human joint angle data $\theta(t) \in \mathbb{R}^{n_{j_h}}$, where n_{j_h} is the number of degrees of freedom of the skeleton model. Then, we can compute a simulated COM and COP trajectory during the human motion ($\mathbf{p}_{\text{COMsim}}(t) \in \mathbb{R}^3$ and $\mathbf{p}_{\text{COPsim}}(t) \in \mathbb{R}^2$) based on the skeletal model's inertia parameters. There may be a difference between simulated and measured COP, because the inertia properties are different between the subject and the average human. We, therefore, manually scale the link lengths and masses of the skeletal model appropriately so that the resulting $\mathbf{p}_{\text{COPsim}}(t)$ matches $\mathbf{p}_{\text{COPexp}}(t)$ within 2cm to solve 1). Lastly, the appropriate COM trajectory $\mathbf{p}_{\text{COMsim}}(t)$ is computed.

Next we generate the whole-body joint trajectory for the humanoid robot. Other researchers have explored mapping human motion capture data to humanoid robot motion, e.g. mapping the joint angle data, or tracing the end effector's trajectory. In this paper, we apply an inverse kinematics



Fig. 4. Top: Frame sequence of the human actor performing an elderly style standing motion. Bottom: The corresponding humanoid robot’s standing motion emulating the actor’s performance.

computation considering the motion capture marker data as a constraint using the method proposed in [11]. We add the simulated human COM trajectory as an additional constraint in the inverse kinematics computation. First, we adjust the marker position on the humanoid robot model. To solve 2) and realize a complicated contact condition like sitting on a chair, we experimentally measure the humanoid robot’s joint angles in a sitting posture. Then we pick the corresponding frame from the human motion capture data, and overlay the measured marker positions of that frame onto the measured humanoid robot’s posture. The relative position between each marker and the corresponding humanoid robot’s link is chosen as the robot’s marker position set. The inertia parameters of humanoid robot that were estimated in the previous section and the robot marker position set gives the Jacobian matrix of COM w.r.t. the joint angles $\mathbf{J}_{\text{COM}}(\theta_r) \in \mathbb{R}^{3 \times n_{j_r}}$ and that of the marker positions w.r.t. the joint angles $\mathbf{J}_m(\theta_r) \in \mathbb{R}^{3n_m \times n_{j_r}}$, where n_{j_r} is the number of degrees of freedom of the humanoid robot. Lastly, we compute the trajectory of the humanoid robot joint angles $\theta_r(t) \in \mathbb{R}^{n_{j_r}}$ iteratively that satisfies $\mathbf{p}_{\text{COMsim}}(t)$ as a hard constraint and $\mathbf{m}(t)$ as a soft constraint. The resulting joint angle trajectory maintains the desired COM trajectory strictly to keep balance

while respecting the measured marker trajectories as a soft constraint to emulate the human motion.

IV. EVALUATION

A. Experiments

We evaluate the proposed method on the Carnegie Mellon/Sarcos humanoid robot. This force controllable robot has a total of 34 hydraulic actuators with a load sensor at each joint, and a 6 axis force sensor at each foot. First, we recorded human motion using the VICON optical motion capture system, while simultaneously measuring ground contact force with AMTI force plates. During motion capture, we asked our demonstrators not to move their feet during the motion or touch their hands to the chair or body. No other instructions such as speed or timing were given. While we recorded healthy subjects executing natural standing motions, we also recorded a professional actor performing stylized standing, including imitating an elderly person. After motion capture, we generate the corresponding robot motion using the estimated robot COM parameters from Sec. II and the proposed inverse kinematics computation from Sec. III. Finally we execute the planned trajectory, at full speed, on the humanoid robot via joint PD control. Fig. 4 shows an

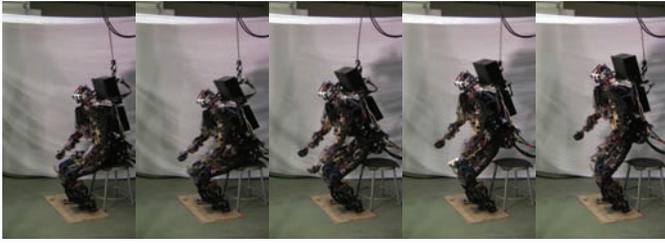


Fig. 5. Sequence of a natural standing motion implemented on the humanoid robot.

example motion sequence. Here the professional actor is performing a stylized elderly stand up motion. The bottom shows the corresponding robot sequence generated by the proposed method, aligned in time with the demonstrated motion. Fig. 5 shows the sequence that results from a demonstration of natural standing. From these frame shots, we see that the humanoid robot successfully realizes the human standing motion, while maintaining balance during the whole motion. In the next subsection, we analyze the data measured during the humanoid robot experiment, and evaluate the efficiency of the proposed method.

B. Results

Fig. 6 shows the COM and COP position in the longitudinal direction, plotted against time in the horizontal axis, for the example of a natural standing motion shown in Fig. 5. The vertical axis represents the COM and COP position [m] with respect to the left foot position, and a positive value represents the forward direction. Each line represents the following value:

- green dashed line: the estimated COM position of the human demonstrator. The motion is captured by the optical motion capture system, and this COM is estimated using a skeletal model whose inertial parameters are from an average human but adjusted based on the recorded COP position (not shown).
- blue solid line: the planned COM position for the generated humanoid robot motion. The humanoid robot motion is generated by the inverse kinematics computation with COM position as a hard constraint. The iterative inverse kinematics optimization produces a very close match to the human’s COM trajectory.
- red solid line: the estimated COM position from the experiment with the humanoid robot, computed by forward kinematics from measured joint angles and the COM parameters estimated from Sec II.
- red dash line: the actual COP position in the experiment with the humanoid robot. The CMU/Sarcos humanoid robot has 6-axis force sensors at each foot link, and the COP position is computed from this contact force data (not AMTI force plates as used in previous sections). Because of this approximation, the COP measurement is inaccurate due to the chair contact, until lift-off at approximately 1 second.

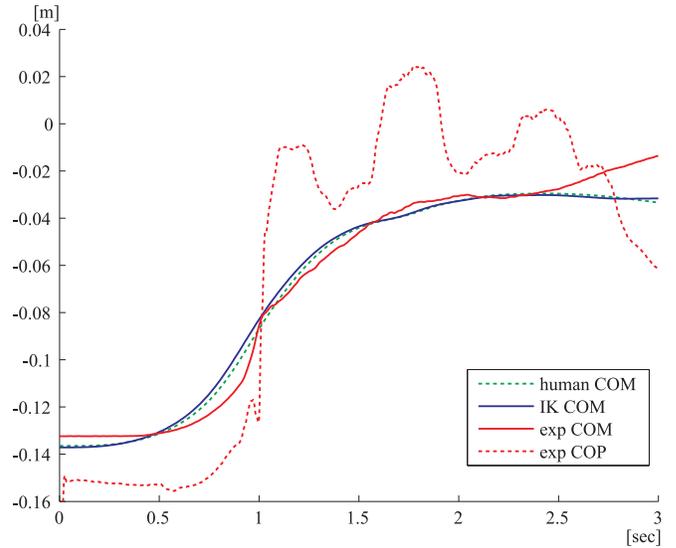


Fig. 6. COM and COP position in the longitudinal direction vs. time, for a natural standing motion. Chair lift-off occurs at approximately 1 second. Green dash line: the human COM position estimated from the experimental human motion data, blue solid line: the planned COM position for the generated humanoid robot motion, red solid line: the estimated COM position of the robot during the experiment, and red dash line: the recorded COP position of the robot during the experiment.

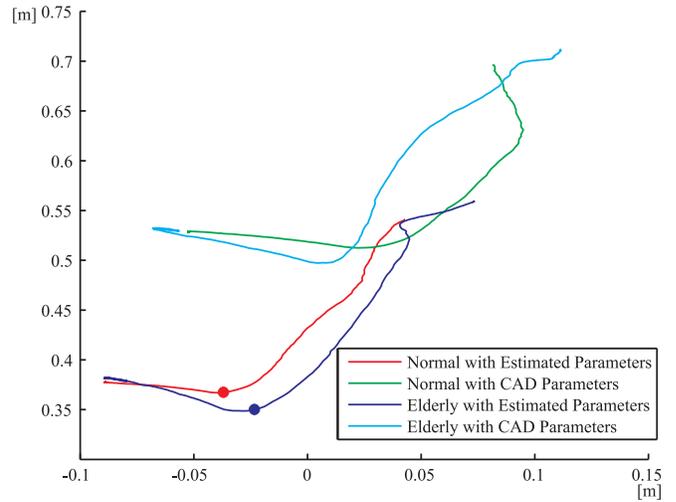


Fig. 7. Robot COM position trajectory in the sagittal plane. Red and dark blue lines: estimated COM trajectories for natural and elderly motions (respectively) using estimated COM parameters from Sec. II. Dots mark the location of chair lift-off. Green and light blue lines: trajectories for same motions if original CAD parameters are used for COM estimation.

From this graph, we see that the solid red line strictly follows the green dashed line and blue solid line, showing that the motion of the humanoid robot realizes the human’s COM trajectory well. The COP position data (red dashed line), also follows the COM trajectory but with some variance, because the generated motion is not quasi-static. Because we generated this motion from an inverse kinematics computation of experimental human motion data, with kinematics constraints, dynamics constraints are not

considered. However, although the resulting motion is not quasi-static, it can still maintain balance throughout the entire motion even though the contact situation changed dramatically during the motion.

In Fig. 7 we compare the resulting COM trajectories, from the robot, for two different motions: natural standing (red) and stylized elderly (dark blue). Trajectories are shown in the sagittal plane (x vs. y position), and traverse towards the right with time. In the elderly case, we see that the COM actually first moves lower before rising, representing a deeper torso bend in order to move the COM closer to the feet and generate more momentum before lift off. This is consistent with the strategy actually used by elderly subjects [12]. At the end of the motion we also see the COM move back, possibly to compensate for the additional generated momentum. Additionally we also plot how the resulting COM trajectories would have looked if CAD parameters are used in the model, instead of our estimated inertial parameters from Sec II. We see that the estimation procedure moves the COM significantly lower and towards the rear. This difference can mainly be attributed to the relatively massive backpack not accounted for in the CAD files.

V. CONCLUSION

In this work, we implemented a sit-to-stand task on a humanoid robot, where the robot was able to successfully stand up from a sitting posture. Our goal was not to engineer a stable robotic motion for this task, but rather to emulate natural human-like movement and speed. Therefore we used demonstrations from healthy subjects naturally standing up as well as actors performing stylized motions. Because the dynamics plays a key roll in the realization of this task, we also recorded the contact forces from our human demonstrators. This contact force information, along with a skeletal model, allowed us to estimate the demonstrator's COM trajectory along the motion. Maintenance of a consistent COM trajectory has been identified by biomechanics researchers as a critical control component of this task. Therefore, it was important for us to not only emulate the demonstrator's kinematic motion with our robot, but also their COM trajectory. However, in order to be able to do so, we require an estimate of the robot's COM position. To obtain as accurate an estimate as possible, we used an inertial parameter estimation technique to identify the mass and center of mass parameters of a 5 link planar robot model. We then used this model to estimate the robot's COM position during the motion.

While we had some success with this technique and we were able to successfully track some of the demonstrators at full speed, an obvious next step is to use more dynamic

constraints such as ZMP trajectory. Additionally, since our robot is force-controllable, we would like to employ force control techniques such as inverse dynamics [13]. Doing so will require full inertial parameter estimation including the inertia tensors and joint friction. We have developed techniques for the estimation of floating base robots [4]. Finally, we would also like to include additional contacts during a standing motion such as arm rests, and include arm push off as part of the motion. Interestingly, while humans often use hand contact as an aid in standing up, for a robot this may present an additional challenge.

ACKNOWLEDGMENTS

The authors acknowledge the help and contributions of Stuart Anderson for the development of the realtime control software as well as the overall calibration of the hardware.

REFERENCES

- [1] H. Mayeda, K. Osuka, and A. Kanagawa, "A new identification method for serial manipulator arms," *Proceedings of the IFAC 9th World Congress*, vol. 2, pp. 74–79, 1984.
- [2] C. Atkeson, C. An, and J. Hollerbach, "Estimation of inertial parameters of manipulator loads and links," *The International Journal of Robotics Research*, Jan 1986.
- [3] K. Ayusawa, G. Venture, and Y. Nakamura, "Identification of the inertial parameters of a humanoid robot using unactuated dynamics of the base link," *8th IEEE-RAS International Conference on Humanoid Robots*, pp. 1 – 7, Nov 2008.
- [4] M. Mistry, S. Schaal, and K. Yamane, "Inertial parameter estimation of floating base humanoid systems using partial force sensing," *9th IEEE-RAS International Conference on Humanoid Robots*, 2009.
- [5] W. Khalil and E. Dombre, *Modeling, Identification, and Control of Robots*. London and Sterling, VA: Kogan Page Science, 2002.
- [6] T. Sugihara and Y. Nakamura, "Whole-body cooperative balancing of humanoid robot using cog jacobian," *2002 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 2575–2580, 2002.
- [7] C. Ott, D. Lee, and Y. Nakamura, "Motion capture based human motion recognition and imitation by direct marker control," *IEEE-RAS International Conference on Humanoid Robots 2008*, pp. 399–405, 2008.
- [8] J. P. Scholz and G. Schöner, "The uncontrolled manifold concept: identifying control variables for a functional task," *Exp Brain Res*, vol. 126, no. 3, pp. 289–306, Jun 1999.
- [9] Y. Nakamura, K. Yamane, and A. Murai, "Macroscopic modeling and identification of the human neuromuscular network," *Proceedings of the 28th IEEE EMBS Annual International Conference*, pp. 99–105, 2006.
- [10] A. Murai, K. Yamane, and Y. Nakamura, "Modeling and identifying the somatic reflex network of the human neuromuscular system," *Proceedings of the 29th IEEE EMBS Annual International Conference, Lyon, FRANCE*, 2007.
- [11] K. Yamane and Y. Nakamura, "Natural motion animation through constraining and deconstraining at will," *IEEE Transaction on Visualization and Computer Graphics*, vol. 9, pp. 352–360, 2003.
- [12] E. Papa and A. Cappozzo, "Sit-to-stand motor strategies investigated in able-bodied young and elderly subjects," *J. Biomechanics*, vol. 33, no. 9, pp. 1113–22, 2000.
- [13] M. Mistry, J. Buchli, and S. Schaal, "Inverse dynamics control of floating base systems using orthogonal decomposition," *Proceedings of the 2010 IEEE Int. Conference on Robotics and Automation*, 2010.