

Kinesthetic teaching of humanoid motion based on whole-body compliance control with interaction-aware balancing

Christian Ott¹, Bernd Henze¹, and Dongheui Lee²

Abstract—In this work we present a framework for kinesthetic teaching and iterative refinement of whole body motions. For detection of external forces we apply a momentum based disturbance observer known from manipulator control to the floating-base model of a humanoid robot. These external forces are used as a trigger for implementing a compliant behavior at the interaction point and are integrated into a predictive balancing algorithm. For representation of the motion data, a hidden Markov model is used, which allows for an iterative update of the discrete motion states as well as a smooth generation of continuous motion data. Finally, we present an application of these algorithms on the humanoid robot TORO.

I. INTRODUCTION

Generating whole body motions for humanoid robots requires the coordination of a large number of degrees of freedom, while complying with unilateral contact constraints. Moreover, if such robots are to be used by non-expert users a demand for intuitive programming interfaces appears.

Bouyarmane and Kheddar [1] presented an optimization based motion planning framework for generating statically stable whole body motions involving multiple state dependent contact constraints. More specific optimization criteria were applied for generating dynamically stable walking [2], [3] and running motions [4]. While optimization based approaches can utilize precise model information for generating dynamically consistent trajectories compatible with unilateral contact constraint, these methods are mainly used for offline motion generation and need a stabilizing feedback controller for online execution.

Generating whole body motions from observation / imitation of human motion was considered in several works in robotics [5] and computer graphics [6]. Recently, a whole body motion programming approach based on a sequence of key frames was presented in [7].

The concept of motion primitives [8] has been introduced in order to reduce the dimensionality of the action space. It is expected that out of a limited set of generic primitive motions more complex actions can be synthesized by combining and adapting individual motions. Such motion primitives can be obtained from offline optimization, human motion observation and imitation, or kinesthetic teaching.

Early works on kinesthetic teaching considered a passive robot behavior [9], [10] by deactivating the controlled motion

or setting very low servo gains. However, these approaches might lead to unsynchronized motions because the teacher moves motors one by one rather than demonstrating natural coordinated movements. Recently, the combination between active impedance control and physical teaching became more popular. In a recent user study on kinesthetic teaching [11], the role of kinematic redundancy was identified as one of the main difficulties for the user.

In [12] kinesthetic teaching of the upper body motion of a humanoid robot was considered. An interaction control approach for the upper body motion was combined with a lower body balancing algorithm based on the reaction null space approach [13]. External forces from kinesthetic teaching or from the task execution were considered as disturbances for the balancer, which employed two strategies, namely ankle- and hip-strategy.

Ibanez et al. [14] proposed an extension of the ZMP preview controller for bipedal balancing under the influence of external forces. In their approach a disturbance model was integrated into the controller for allowing a prediction of the external disturbance forces.

In our previous works [15], [16] we proposed a framework for iterative refinement of learned motion primitives. Therein, a specialized compliance controller was parameterized according to the uncertainty of the learned trajectory. In [15], [16] we applied this approach to a fixed based humanoid manipulator system. In the present paper, we extend this approach towards kinesthetic teaching of a bipedal humanoid robot. External forces from the human teaching are detected by a momentum based disturbance observer known from manipulator control [17]. These external forces are integrated into the bipedal balancing algorithm acting on the horizontal motion of the center of mass (COM). Based on the external force detection, we parameterize the interaction controller differently depending on the detected contact point and its null space. The proposed controller can help for a human subject to teach a full body humanoid robot in a synchronized and easy way.

II. COMPLIANCE CONTROL FOR KINESTHETIC TEACHING

In this section we present the control methods for physical interaction with the robot. During kinesthetic teaching, human interaction forces should allow to correct the motion without affecting the balancing task of the robot. Section II-A discusses the problem of detecting external forces. The result of the force detection can be utilized for influencing the compliant motion control from section II-B. Moreover, in section II-C these forces are incorporated into a balancing

¹Ch. Ott and B. Henze are with the Institute of Robotics and Mechatronics, German Aerospace Center (DLR e.V.), 82234 Wessling, Germany christian.ott at dlr.de, bernd.henze at dlr.de

²D. Lee is with the Department of Electrical Engineering and Information Technology, Technical University of Munich (TUM), 80333 Munich, Germany dhlee at tum.de

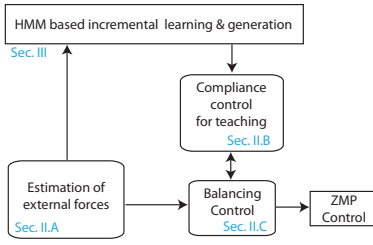


Fig. 1. Overview of the kinesthetic teaching approach.

approach by means of a predictive disturbance model. Figure 1 gives an overview of the main elements in our approach.

A. Contact force detection

The dynamical model of a humanoid robot can be described by a floating base robot with n degrees of freedom (DOF) $\mathbf{q} \in \mathbb{R}^n$, which is interacting with its environment through contact wrenches \mathbf{F}_r and \mathbf{F}_l at the feet and through a vector of generalized external forces $\boldsymbol{\tau}_{\text{ext}}$. Let \mathbf{v} be the body twist describing the generalized velocity of the base link, then the dynamics can be written as

$$\mathbf{M} \begin{pmatrix} \dot{\mathbf{v}} \\ \dot{\mathbf{q}} \end{pmatrix} + \mathbf{C} \begin{pmatrix} \mathbf{v} \\ \dot{\mathbf{q}} \end{pmatrix} + \mathbf{g} = \mathbf{S}\boldsymbol{\tau} + \sum_{i=r/l} \mathbf{J}_i^T \mathbf{F}_i + \boldsymbol{\tau}_{\text{ext}} \quad (1)$$

where \mathbf{M} , \mathbf{C} , and \mathbf{g} are the configuration dependent inertia matrix, centrifugal and Coriolis matrix, and gravity terms. The joint torques $\boldsymbol{\tau}$ act via the matrix $\mathbf{S} = [\mathbf{0}_{6 \times 6} \quad \mathbf{I}_{n \times n}]^T$ only on the actuated degrees of freedom, while the configuration dependent Jacobian matrices \mathbf{J}_r and \mathbf{J}_l describe how the contact wrenches of the feet act onto the system.

In order to estimate the generalized external force $\boldsymbol{\tau}_{\text{ext}}$ in (1) we apply a momentum based disturbance observer as it was proposed in [17]. At first, the change of the generalized momentum $\mathbf{P} = \mathbf{M} \begin{pmatrix} \mathbf{v} \\ \dot{\mathbf{q}} \end{pmatrix}$ is calculated by

$$\dot{\mathbf{P}} = \mathbf{S}\boldsymbol{\tau} + \sum_{i=r/l} \mathbf{J}_i^T \mathbf{F}_i + \boldsymbol{\tau}_{\text{ext}} + \mathbf{C}^T \begin{pmatrix} \mathbf{v} \\ \dot{\mathbf{q}} \end{pmatrix} - \mathbf{g}, \quad (2)$$

wherein the passivity property $\dot{\mathbf{M}} = \mathbf{C} + \mathbf{C}^T$ has been used. It is assumed that the values of $\boldsymbol{\tau}$, \mathbf{F}_i , \mathbf{v} , $\dot{\mathbf{q}}$ and \mathbf{g} can be measured or directly derived from measurement in contrast to $\boldsymbol{\tau}_{\text{ext}}$, which is the unknown external torque to be reconstructed by the observer. Based on an estimate $\boldsymbol{\tau}_{\text{ext}}^{\text{obsv}}$ we can compute an estimate of $\dot{\mathbf{P}}$:

$$\dot{\hat{\mathbf{P}}} = \mathbf{S}\boldsymbol{\tau} + \sum_{i=r/l} \mathbf{J}_i^T \mathbf{F}_i + \boldsymbol{\tau}_{\text{ext}}^{\text{obsv}} + \mathbf{C}^T \begin{pmatrix} \mathbf{v} \\ \dot{\mathbf{q}} \end{pmatrix} - \mathbf{g}, \quad (3)$$

The derivative $\dot{\hat{\mathbf{P}}}$ can be integrated over time and compared to the measured generalized momentum \mathbf{P} . The resulting error is feed back by using a diagonal gain matrix $\mathbf{K} > 0$:

$$\boldsymbol{\tau}_{\text{ext}}^{\text{obsv}} = \mathbf{K}(\mathbf{P} - \hat{\mathbf{P}}) \quad (4)$$

$$= \mathbf{K} \left[\mathbf{M} \begin{pmatrix} \mathbf{v} \\ \dot{\mathbf{q}} \end{pmatrix} - \int_0^t \dot{\hat{\mathbf{P}}}(t) dt - \hat{\mathbf{P}}(0) \right]. \quad (5)$$

Using (2) and (3) it can be easily verified that the observed torque $\boldsymbol{\tau}_{\text{ext}}^{\text{obsv}}$ resulting from (5) complies with the first order dynamics $\dot{\boldsymbol{\tau}}_{\text{ext}}^{\text{obsv}} = \mathbf{K}(\boldsymbol{\tau}_{\text{ext}} - \boldsymbol{\tau}_{\text{ext}}^{\text{obsv}})$ and thus represents a filtered estimate of the generalized external torque $\boldsymbol{\tau}_{\text{ext}}$ [17].

For applying this disturbance observer to the floating-base model of a humanoid robot the complete body twist related to the base link is needed. While the angular velocity is often available from an inertial measurement unit integrated in the robot, the measurement of the translational velocity is more problematic. An estimation based on kinematic measurements assumes a well-defined contact state of the lower limbs. A precise ego-motion detection that is independent on the contact state usually requires additional instrumentation like onboard vision or other external sensors [18]. The experiments reported in section IV are based on a kinematic estimation.

From the distribution of generalized forces within the vector $\boldsymbol{\tau}_{\text{ext}}^{\text{obsv}}$ one can also estimate on which link of the kinematic structure the external force is acting [17]. If we for instance assume that the external forces act on both hands, the external forces can be retrieved from

$$\boldsymbol{\tau}_{\text{ext}}^{\text{obsv}} = \sum_{i=r/l} \mathbf{J}_{\text{ext},i}^T \mathbf{F}_{\text{ext},i} = \begin{bmatrix} \mathbf{J}_{\text{ext},l}^T & \mathbf{J}_{\text{ext},r}^T \end{bmatrix} \begin{bmatrix} \mathbf{F}_{\text{ext},l} \\ \mathbf{F}_{\text{ext},r} \end{bmatrix} \quad (6)$$

by using a pseudo inverse of $\begin{bmatrix} \mathbf{J}_{\text{ext},l}^T & \mathbf{J}_{\text{ext},r}^T \end{bmatrix}$.

B. Compliance control

In our previous work [15], we proposed a compliance controller, which allows for kinesthetic refinement of the robot motion within a tube around the nominal trajectory. The size of this tube was determined by the uncertainty of the commanded motion and the approach in [15] did not use information about the point of interaction. When teaching the arm motion via the end-effector kinesthetically, the human teaching force might act as a disturbance for other parts of the robot. This can lead to an undesired motion in the null space if the parameters of the refinement tube are not set carefully. Therefore, we extend this approach in the following by proposing an alternative teaching mode in which we separate the control behavior for the compliant teaching at the interaction point from the control action in the null space.

The basic control structure resembles a PD+ like tracking controller [19] in which the proportional control action is computed based on a nonlinear stiffness term. The shape of the stiffness term $f(\cdot, \chi)$, shown in Fig. 2, includes a high stiffness around the equilibrium for precise tracking in free motion and the refinement tube area with constant force and thus zero stiffness. For kinesthetic teaching in joint space without utilization of the information about the point of interaction, the control action from [15] is given by

$$\boldsymbol{\tau} = \boldsymbol{\tau}_{\text{ff}} - \mathbf{s}(\tilde{\mathbf{q}}) - \mathbf{D}\dot{\tilde{\mathbf{q}}}, \quad (7)$$

where $\boldsymbol{\tau}_{\text{ff}}$ is a feed forward torque related to the the desired motion $\mathbf{q}_d(t)$, \mathbf{D} is a positive definite damping matrix, and

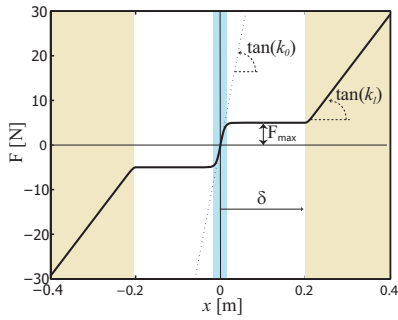


Fig. 2. Nonlinear stiffness term in the compliance controller. The parameters F_{max} , k_0 , k_l , δ are summarized in the vector χ .

the stiffness term $\mathbf{s}(\tilde{\mathbf{q}})$ is defined by $s_i(\tilde{q}_i) = f(\tilde{q}_i, \chi_i)$ for $\tilde{\mathbf{q}} = \mathbf{q} - \mathbf{q}_d$.

As an alternative teaching mode, which allows to avoid undesired disturbance of the null space motion, we utilize

$$\boldsymbol{\tau} = \boldsymbol{\tau}_{ff} - \mathbf{J}^T (\mathbf{s}_c(\tilde{\mathbf{x}}) + \mathbf{D}_x \dot{\tilde{\mathbf{x}}}) - \mathbf{N} (\mathbf{s}(\tilde{\mathbf{q}}) + \mathbf{D}\dot{\tilde{\mathbf{q}}}) , \quad (8)$$

where \mathbf{x} and \mathbf{J} are the Cartesian coordinates and the Jacobian matrix related to the interaction point and $\mathbf{N} = \mathbf{I} - \mathbf{J}^T \mathbf{J}^{+T}$ is a null space projection matrix. The stiffness term $\mathbf{s}_c(\tilde{\mathbf{x}})$ is given by $s_{c,i}(\tilde{x}_i) = f(\tilde{x}_i, \chi_{c,i})$ for $\tilde{\mathbf{x}} = \mathbf{x}(\mathbf{q}) - \mathbf{x}(\mathbf{q}_d)$. By splitting the control action into one part for the Cartesian motion and a second part for the null space motion it is possible to allow teaching of the task motion (i.e. by choosing a large refinement tube δ_c) while avoiding undesired disturbance of the null space motion during kinesthetic teaching (i.e. by setting a small or zero tube parameter δ).

The approach described so far implements a basic compliant behavior for teaching and allows several extensions. If interaction forces are acquired during teaching, these forces can be implemented during motion generation by simple addition or via hybrid impedance force control. For a bipedal humanoid robot, additionally the balancing problem must be considered. In [20] we presented a force based balancing algorithm based on an underlying torque controller. This approach allows to react in a compliant way to external forces. In the present work, we want to compensate for external forces resulting from the teaching. While the extension of [20] in this direction is considered as a future problem, we describe in the next section an approach based on a simplified model, which allows to consider a predictive disturbance model related to external forces. Such predicted force information could come from the algorithms described in section III, but in the present work, we use a simple disturbance model based on the observed external forces from section II-A. The balancing approach from the following section will generate a desired trajectory for the zero moment point (ZMP), which can be realized by a position based ZMP controller. While the horizontal motion of the COM is then required for the realization of the ZMP, the vertical COM motion and the hip orientation can still be controlled based on a compliant relation with the stiffness function from Fig. 2 in order to allow for kinesthetic teaching. In case of inverse

kinematics based ZMP control, the compliant relation for the vertical hip motion h_z can be implemented by

$$\dot{h}_{z,cmd} = \dot{h}_{z,d} + k_v (F_z - f(h_z - h_{z,d}, \chi)) , \quad (9)$$

with a positive gain k_v that controls the convergence to the steady state stiffness using the vertical force F_z measured by the force/torque sensors in the feet. Obviously, an analogous approach can be implemented for the hip orientation.

C. Interaction-aware balancing

Ibanez et al. [14] proposed a preview balancing controller, which is able to take into account disturbances in form of an external force acting at the end-effectors of a humanoid robot. In this section we will consider general wrenches which are detected by the disturbance observer from section II-A. These external forces are considered as task related forces or interaction forces from human upper body teaching and thus shall be compensated by the balancing algorithm. Furthermore we will add a terminal constraint to the optimization in order to ensure the asymptotic stability of the controller and make use of the so called Linear Inverted Pendulum Model (LIPM) [21] instead of the Cart-Table Model [22] applied by [14].

The LIPM can be derived from (1) by reducing the system of multiple rigid bodies to a single point mass, which is equivalent to the total mass m of the system and located at its COM $\mathbf{c} = [c_x \ c_y]^T$. Furthermore, it is assumed that $\boldsymbol{\tau}_{ext}$ from (1) is related to a torque $\boldsymbol{\tau}^{LIP} = [\tau_x^{LIP} \ \tau_y^{LIP}]^T$ acting at the ZMP $\mathbf{p} = [p_x \ p_y]^T$. Neglecting the changes in the height of the mass above the ground c_z the dynamics of the LIPM can be described by

$$\ddot{\mathbf{c}} = \omega^2 (\mathbf{c} - \mathbf{p}) + \frac{1}{m \cdot c_z} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \boldsymbol{\tau}^{LIP} \quad (10)$$

with $\omega = \sqrt{g/c_z}$ and g as gravity acceleration. Equation (10) is discretized in time using the forward differencing method:

$$\underbrace{\begin{pmatrix} c_x \\ c_y \\ \dot{c}_x \\ \dot{c}_y \end{pmatrix}}_{\mathbf{z}_{k+1}} = \underbrace{\begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ \omega^2 T & 0 & 1 & 0 \\ 0 & \omega^2 T & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} c_x \\ c_y \\ \dot{c}_x \\ \dot{c}_y \end{pmatrix}}_{\mathbf{z}_k} + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -\omega^2 T & 0 \\ 0 & -\omega^2 T \end{bmatrix}}_{\mathbf{B}} \mathbf{p}_k + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & T/(mc_z) \\ -T/(mc_z) & 0 \end{bmatrix}}_{\mathbf{E}} \boldsymbol{\tau}_k, \quad (11)$$

where T denotes the length of the time step.

The future states of the system required for the prediction can be extrapolated by

$$\underbrace{\begin{pmatrix} \mathbf{z}_{k+1} \\ \vdots \\ \mathbf{z}_{k+N} \end{pmatrix}}_{\hat{\mathbf{z}}} = \underbrace{\begin{bmatrix} \mathbf{A} \\ \vdots \\ \mathbf{A}^N \end{bmatrix}}_{\hat{\mathbf{A}}} \mathbf{z}_k + \underbrace{\begin{bmatrix} \mathbf{B} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{A}^{N-1} \mathbf{B} & \cdots & \mathbf{B} \end{bmatrix}}_{\hat{\mathbf{B}}} \underbrace{\begin{pmatrix} \mathbf{p}_k \\ \vdots \\ \mathbf{p}_{k+N-1} \end{pmatrix}}_{\hat{\mathbf{p}}} + \underbrace{\begin{bmatrix} \mathbf{E} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{A}^{N-1} \mathbf{E} & \cdots & \mathbf{E} \end{bmatrix}}_{\hat{\mathbf{E}}} \underbrace{\begin{pmatrix} \boldsymbol{\tau}_k^{LIP} \\ \vdots \\ \boldsymbol{\tau}_{k+N-1}^{LIP} \end{pmatrix}}_{\hat{\boldsymbol{\tau}}^{LIP}}. \quad (12)$$

The aim of the controller is to minimize the cost function J , which we choose as

$$J = \frac{1}{2} \hat{\mathbf{z}}^T \hat{\mathbf{R}} \hat{\mathbf{z}} + \frac{1}{2} \hat{\mathbf{p}}^T \hat{\mathbf{Q}} \hat{\mathbf{p}} + \frac{1}{2} \left(\hat{\mathbf{B}} \hat{\mathbf{p}}_k + \hat{\mathbf{E}} \hat{\boldsymbol{\tau}}_k^{LIP} \right)^T \hat{\mathbf{P}} \left(\hat{\mathbf{B}} \hat{\mathbf{p}}_k + \hat{\mathbf{E}} \hat{\boldsymbol{\tau}}_k^{LIP} \right), \quad (13)$$

wherein $\hat{\mathbf{R}} > 0$, $\hat{\mathbf{Q}} > 0$ and $\hat{\mathbf{P}} > 0$ are diagonal weighting matrices. An error in the state of the system is weighted by $\hat{\mathbf{R}}$ while the consumption of input signal is taken into account by $\hat{\mathbf{Q}}$. The additional term depending on $\left(\hat{\mathbf{B}} \hat{\mathbf{p}}_k + \hat{\mathbf{E}} \hat{\boldsymbol{\tau}}_k^{LIP} \right)$ can be interpreted as a soft constraint, which demands that the robot shifts its ZMP in a way that external disturbances are compensated without affecting the COM. Hence, the matrix $\hat{\mathbf{P}}$ can be used for influencing the disturbance response of the closed loop system. The optimal position of the ZMP $\hat{\mathbf{p}}^*$ is provided by the optimization

$$\hat{\mathbf{p}}^* = \underset{\substack{\hat{\mathbf{z}} = \hat{\mathbf{A}} \hat{\mathbf{z}}_k + \hat{\mathbf{B}} \hat{\mathbf{p}} + \hat{\mathbf{E}} \hat{\boldsymbol{\tau}}^{LIP} \\ \hat{\mathbf{z}}_{k+N} = \mathbf{z}_{\text{end}}}}{\text{argmin}} J(\hat{\mathbf{p}}) \quad (14)$$

under the constraint of the system dynamics (12) and a terminal constraint $\mathbf{z}_{k+N} = \mathbf{z}_{\text{end}}$ to enforce closed loop stability. Besides this we assume that the torque $\boldsymbol{\tau}^{LIP}$ is constant over the full length of the horizon, which holds for slowly changing torques compared to the length of the horizon. Note that the vector $\hat{\mathbf{p}}^*$ specifies the feedforward control inputs from the current point in time until the end of the prediction. In order to get the control law one has to close the feedback loop by extracting the current input \mathbf{p}_k^* from $\hat{\mathbf{p}}^*$ and by recomputing it at each time step:

$$\mathbf{p}_k^* = \begin{pmatrix} \mathbf{I}_{1 \times 2} & \mathbf{0}_{1 \times N-2} \end{pmatrix} \cdot \hat{\mathbf{p}}^* \quad (15)$$

Note that in the absence of additional inequality constraints the choice of the cost function and of the terminal constraint results in a linear control law of the form

$$\hat{\mathbf{p}}^* = \mathbf{K}_x \mathbf{z}_k + \mathbf{K}_{\text{end}} \mathbf{z}_{\text{end}} + \mathbf{K}_\tau \boldsymbol{\tau}^{LIP} \quad (16)$$

with \mathbf{K}_x , \mathbf{K}_{end} and \mathbf{K}_τ as constant gain matrices, which allows the optimization to be performed offline.

The asymptotic stability of the controller can be shown by applying Theorem 6.1 from [23], which requires a terminal constraint for the optimization process and a feasible solution. Since there is always a solution to (14) in combination with the invoked terminal constraint, the presented controller is asymptotically stable.

In order to obtain the torque $\boldsymbol{\tau}_k^{LIP}$ one can use the result of the disturbance observer from section II-A. The torque $\boldsymbol{\tau}^{LIP}$ acting at the ZMP has to be computed from the wrenches acting at the arms with

$$\boldsymbol{\tau}_k^{LIP} = \begin{bmatrix} \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} \end{bmatrix} \sum_{i=l/r} \mathbf{A}d_i^T \mathbf{F}_{\text{ext},i}, \quad (17)$$

wherein $\mathbf{A}d_i$ denotes the corresponding adjoint matrix. Note that the assumption of $\boldsymbol{\tau}^{LIP}$ being constant over the full length of the horizon holds if e.g. the wrenches and the lever arms are constant, too.

Note that in the derivation of the balancing controller we assumed a constant height of the COM, while (9) can lead to a vertical motion of the COM resulting in a disturbance of (10). For typical human teaching actions the vertical COM velocity is rather small and this disturbance should be handled by the robustness properties of the balancing controller.

III. MOTION REPRESENTATION AND ITERATIVE LEARNING FOR KINESTHETIC TEACHING

A. Motion representation

HMM-based representation of motion primitives can represent spatiotemporal variabilities in a stochastic way, and recognize online motion of different speed due to their sequence-based nature [24], [25]. Furthermore, recently some approaches have been proposed to overcome the generation of stepwise sequences due to the discrete nature of states in the HMM, by learning the correlation between temporal and spatial data explicitly [16], [15], [26]. In this work, we follow the HMM-representation and basic algorithms in [16].

Each motion primitive is represented as a left-right type HMM with N states. HMM parameters consist of the initial state probability (π_i , the probability for the initial state to be i -th state), the state transition probability (a_{ij} , the probability to transit from state i to state j), and the observation probability distribution ($b_i(\mathbf{o})$, the probability density function for the output of a vector \mathbf{o} at state i). Herein, a Gaussian distribution is used: $b_i(\mathbf{o}) = \mathcal{N}(\mathbf{o} | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$. The vector $\boldsymbol{\mu}_i$ and matrix $\boldsymbol{\Sigma}_i$ are the mean vector and the covariance matrix for the Gaussian in state i ,

$$\boldsymbol{\mu}_i = \begin{bmatrix} {}^t \boldsymbol{\mu}_i \\ {}^s \boldsymbol{\mu}_i \end{bmatrix}, \quad \boldsymbol{\Sigma}_i = \begin{bmatrix} {}^{tt} \boldsymbol{\Sigma}_i & {}^{ts} \boldsymbol{\Sigma}_i \\ {}^{st} \boldsymbol{\Sigma}_i & {}^{ss} \boldsymbol{\Sigma}_i \end{bmatrix}$$

where the symbol ${}^{ts} \boldsymbol{\Sigma}_i$ is the covariance vector between temporal and spatial data of the Gaussian at state i .

The spatial data can be in joint coordinates and/or Cartesian coordinates. In the experiments, Cartesian movements of the end-effectors, force/torque at the end-effectors, and joint configurations are used. As the temporal data, a normalized time variable is introduced in each state. This variable is 0 when entering a state, 1 when leaving the state, and linearly interpolated during the stay at the state. This relative temporal representation allows to handle data sequences of different speed, even whose speed is time-varying within a motion primitive. Note that a preprocessing of observation (e.g., scaling in time) for learning and recognition is not necessary.

B. Iterative Learning

The incremental learning method herein is a variation of the EM algorithm for multiple observations [27]. In our approach no additional information (e.g., memory of all previous training data, posterior probabilities or the number of all previous training data) rather than the current model parameters and new incoming data is required.

Before the training procedure starts, the initial HMM parameters are set randomly. When the first training data ${}^s\mathcal{O}^1 = \{{}^s\mathbf{o}(t)\}$ enters, HMM parameters are trained via an EM algorithm [28]. Here, ${}^s\mathcal{O}^e$ ($e = 1 \cdots E$) denotes the e -th sequence of *spatial* data of training data at each training step. From the second motion sequence, the training data consists of two motion sequences ($E = 2$): a new incoming observation and a generated motion pattern from a motion primitive (Section III-C).

A forgetting factor is used which leads to exponential forgetting of previous data. This allows to correct the models efficiently and to avoid insensitivity to new incoming data for a large training set. The weighting factor for each sequence (${}^s\mathcal{O}^e$) is given as w^e . At initial update of HMM, w^1 is equal to 1 since there is one motion sequence. Afterwards, for the new incoming motion sequence, w^e becomes the forgetting factor $w^e = \eta$. For the generated motion sequence, the weighting factor becomes $1 - \eta$, so that $\sum_{e=1}^E w^e = 1$. In the M-step, new parameters $\bar{\lambda} = \{\bar{\pi}, \bar{\mathbf{a}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\Sigma}}\}$ for the HMM are estimated by using the old HMM parameters $\lambda = \{\pi, \mathbf{a}, \boldsymbol{\mu}, \boldsymbol{\Sigma}\}$ and the two training data. For the detailed update rules, see [16].

So far, a time series of spatial data is used for training. Thereafter $\lambda = \{\pi, \mathbf{a}, {}^s\boldsymbol{\mu}, {}^{ss}\boldsymbol{\Sigma}\}$ has been learned. Then, the corresponding state sequence for the spatial data sequence ${}^s\mathcal{O}$ is calculated via the Viterbi algorithm [28]. From this state sequence, the relative temporal sequence ${}^t\mathcal{O} = \{{}^t\mathbf{o}(t)\}$ within each state is calculated; ${}^t\mathbf{o}(t)$ is 0 when entering a state, 1 when leaving the state. Therefore, the mean of the temporal data of each state is always the same, which is 0.5. Then, from ${}^t\mathcal{O}$ and ${}^s\mathcal{O}$, the covariance ${}^{ts}\boldsymbol{\Sigma}$ for each state is calculated.

C. Motion Generation - Trajectory and Tube

The motion generation is a stochastic process, consisting of state transition generation and motion output generation from the state transition. First, a general state sequence $\mathcal{Q} = \{q(t)\}$ is generated from the initial state probability distribution π and the state transition probability \mathbf{a} . From the state sequence \mathcal{Q} , the relative temporal sequence ${}^t\mathcal{O}$ within each state is calculated. Given the state sequence \mathcal{Q} and relative temporal sequence ${}^t\mathcal{O}$, a sequence of spatial data is generated using Gaussian regression [29]. For each time step t , a mixture of Gaussians in the corresponding state $q(t)$ is considered. Given ${}^t\mathbf{o}(t)$, the conditional expectation ${}^s\mathbf{o}(t)$ is

$${}^s\mathbf{o}(t) = \sum_{i=1}^N \zeta_i(t) \left\{ {}^s\boldsymbol{\mu}_i + \frac{{}^{st}\boldsymbol{\Sigma}_i}{{}^{tt}\boldsymbol{\Sigma}_i} ({}^t\mathbf{o}(t) - {}^t\boldsymbol{\mu}_i) \right\} \quad (18)$$

by considering the responsibility $\zeta_i(t)$ of each state for ${}^t\mathbf{o}(t)$. The conditional variance of ${}^s\mathbf{o}(t)$ is estimated as

$${}^{s|t}\boldsymbol{\Sigma} = \sum_{i=1}^N \zeta_i(t)^2 \left\{ {}^{ss}\boldsymbol{\Sigma}_i - \frac{{}^{st}\boldsymbol{\Sigma}_i {}^{ts}\boldsymbol{\Sigma}_i}{{}^{tt}\boldsymbol{\Sigma}_i} \right\}. \quad (19)$$

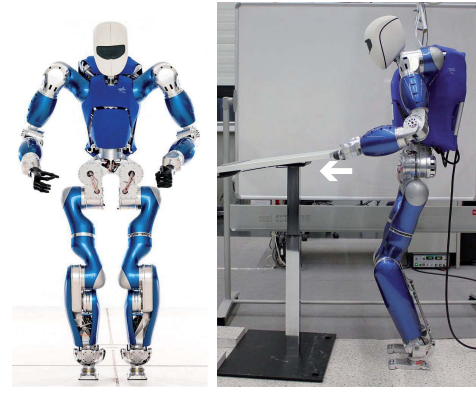


Fig. 3. Left: Humanoid robot TORO (TORque controller RObot). Right: Experimental setup for testing the balancing during physical interaction.

IV. EXPERIMENTS

In order to test the effectiveness of the proposed algorithms, we performed a series of experiments using the humanoid robot TORO (Fig. 3). In the current version TORO has 25 degrees of freedom in the legs (6 DOF), arms (6 DOF), and torso (1 DOF) with a total weight of about 75kg and a total height of about 160cm. All the joints of this robot (except for the fingers which are not used in the reported experiments) are based on the drive units from the DLR-KUKA-Lightweight-Arm-III [30] and thus provide joint torque sensing, which is needed for the disturbance observer from section II-A.

A. Evaluation of the balancing controller

For evaluating the balancing controller, we performed an experiment in which external forces are generated onto the robot. For this experiment, the upper body was controlled by a Cartesian impedance controller with a translational stiffness of 1000N/m implemented in a sampling time of $T = 1ms$. In order to ensure a deterministic force generation, TORO was placed in front of a table and programmed to push against it by moving both arms 9cm to the front, hold it for 5 seconds, and then move back (see Fig. 3). This experiment was done once with the balance-controller taking into account the external forces resulting from the contact between the arms and the desk ($\hat{\mathbf{K}}_\tau \neq \mathbf{0}$) and once without taking into account the disturbances ($\hat{\mathbf{K}}_\tau = \mathbf{0}$). The controller design was computed without end point constraint resulting in the gains listed in Tab. I. Figure 4 shows the corresponding results for the position of the COM, the ZMP and the sum of the forces at the end-effectors only in the direction of x since the other directions do not matter for this particular scenario. One can see that the deviation of the ZMP is higher for $\hat{\mathbf{K}}_\tau \neq \mathbf{0}$ than for $\hat{\mathbf{K}}_\tau = \mathbf{0}$, which is consistent with the observation that the contact force is higher too, since the force resulting from the position of the ZMP relative to the COM must be at an equilibrium with the contact force. The most important fact is that the COM moves about 1.5cm to the front if $\hat{\mathbf{K}}_\tau \neq \mathbf{0}$ instead of about 3 cm to the back if $\hat{\mathbf{K}}_\tau = \mathbf{0}$. That means that by taking the disturbances into account the influence to the COM can be reduced largely.

Parameter	Value
N	300
\hat{P}	$\text{diag}(\dots \text{diag}(0.3, 0.3, 0.3, 0.3) \dots)$
\hat{Q}	$\text{diag}(\dots \text{diag}(1, 1) \dots)$
\hat{R}	$\text{diag}(\dots \text{diag}(8.25, 8.25, 0.0075, 0.0075) \dots)$
Gains	Value
K_x	$\begin{bmatrix} 2.4187 & 0 & 0.4106 & 0 \\ 0 & 2.4187 & 0 & 0.4106 \end{bmatrix}$
K_τ	$\begin{bmatrix} 0 & 0 & 0.0013 & 0 \\ 0 & -0.0013 & 0 & 0 \end{bmatrix}$

TABLE I
PARAMETRIZATION OF THE BALANCING CONTROLLER

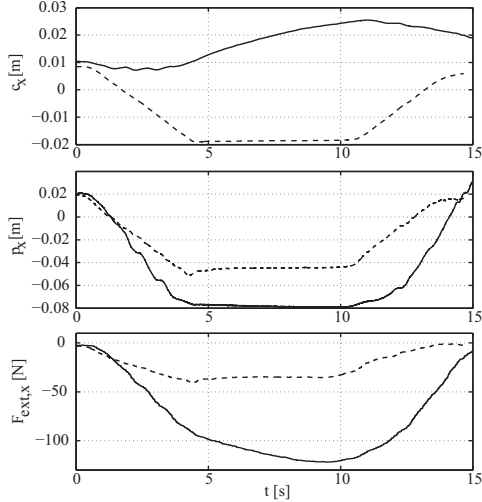


Fig. 4. Experimental results for the balancing controller - dashed line: without taking into account disturbances ($\hat{K}_\tau = 0$), solid line: with taking into account disturbances ($\hat{K}_\tau \neq 0$).

B. Application - Learning a bi-manual lifting task

For kinesthetic teaching, the robot uses the proposed compliance controller from Section II-B based on the interaction force detection. During the kinesthetic demonstrations, the position, orientation (quaternion), force and torque of the robot's both hands as well as the joint configuration are measured. Here, force and torque at the hands are estimated from Section II-A, since the robot is not equipped with force/torque sensors at the wrists, but joint torque sensors. A human demonstrator showed how to lift a box via physical interaction as shown in Fig. 7 (top). By the iterative learning algorithm in Section III-B, the robot learned and performed the lifting task, shown in Fig. 7 (middle). During the task execution, an impedance controller and a force control action are used together (see [31] for imitation learning of force data). Demonstration trajectories and the learned trajectory of a right hand motion are depicted in Fig. 5. Three dimensional trajectories of both hands are given in Fig. 6. In order to show the generalization of the learned behavior, we put a new box whose width is half size of the original box for lifting. Since force/torque data are learned together with movement data, the robot performed the lifting task successfully, as shown in Fig. 7 (bottom).

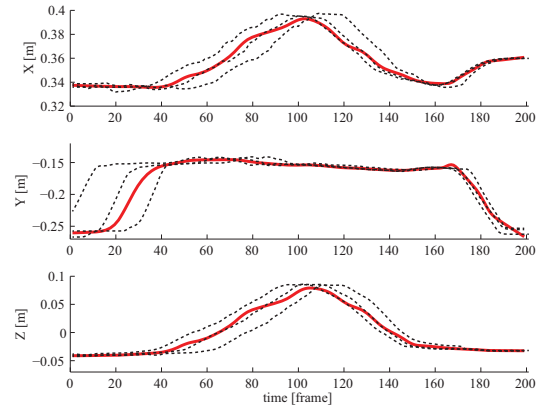


Fig. 5. Demonstrated trajectories (black dashed) during kinesthetic teaching and the learned trajectory (red solid), sampling: 1 frame per 60ms. Three motion sequences are demonstrated via kinesthetic teaching for a box lifting task. The 3D translational motion of the right hand is depicted.

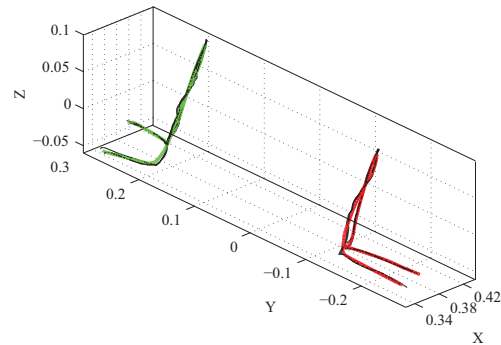


Fig. 6. 3D translational motion of the both hands. Demonstrated trajectories (black) during kinesthetic teaching and the learned trajectory of right hand (red) and the left hand (green).

V. SUMMARY

In this paper we describe a framework for kinesthetic teaching of humanoid robots. For detecting external forces, we utilize a momentum based disturbance observer, which is widely used in manipulator control. The detection of external forces allows to estimate the point of interaction and thus can be used for triggering different impedance behaviors for kinesthetic teaching in which undesired disturbances in the null space can be reduced. Moreover, the detection of external forces allows to improve the balancing behavior. We propose a new cost function for design of a ZMP based balancing controller, which allows to influence the disturbance response of the controller. We implemented and evaluated the interaction-aware balancer and its application in a kinesthetic teaching task using the robot TORO.

While our current approach is general enough to be implemented on many humanoid robots, we would like to mention two possible limitations. First, the disturbance observer from section II-A requires a measurement of the joint torques, which can be difficult in case of highly geared drive units without explicit torque sensors. In such a case, a disturbance detection based on artificial skin might be more effective. The second limitation in our current implementation lies in the fact that the predictive balancing algorithm utilizes a position based ZMP controller. While this allows to apply

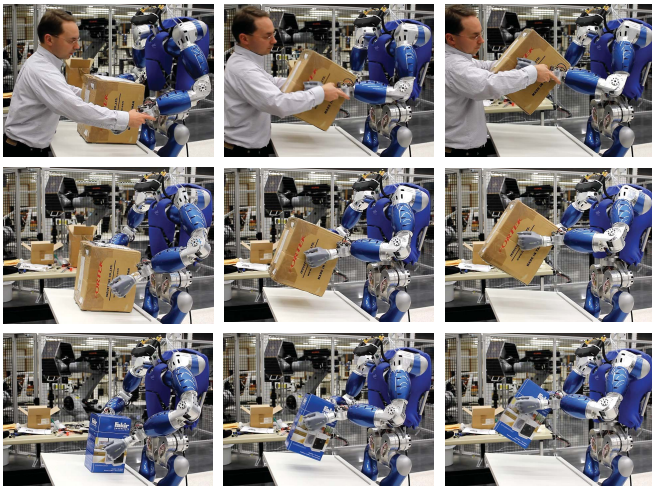


Fig. 7. Snapshots of demonstration (top), learned lifting motion (middle) and generalized behaviors with a new object (bottom).

the approach to any position controlled humanoid, it implies a separate treatment of the upper and lower body motion control. In the future we aim at incorporating such predicted force data also in the torque based balancer from [20]. Moreover, we plan to utilize also learned force data from the human demonstrations in the predictive balancing algorithm.

ACKNOWLEDGMENT

This work has been partially funded by the European Commissions Seventh Framework Programme as part of the project SAPHARI (Grant no. 287513). The first and second author acknowledge support by the Initiative and Networking Fund of the Helmholtz Association through a Helmholtz Young Investigators Group (Grant no. VH-NG-808).

REFERENCES

- [1] K. Bouyarmane and A. Kheddar, "Using a multi-objective controller to synthesize simulated humanoid robot motion with changing contact configurations," in *IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, 2011, pp. 4414–4419.
- [2] A. Werner, R. Lampariello, and C. Ott, "Optimization-based generation and experimental validation of optimal walking trajectories for biped robots," in *IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, 2012.
- [3] T. Buschmann, S. Lohmeier, H. Ulbrich, and F. Pfeiffer, "Optimization based gait pattern generation for a biped robot," in *IEEE-RAS International Conference on Humanoid Robots*, 2005, p. 98–103.
- [4] G. Schultz and K. Mombaur, "Modeling and optimal control of human-like running," *IEEE/ASME Transactions on Mechatronics*, vol. 15, p. 783792, 2010.
- [5] W. Suleiman, E. Yoshida, F. Kanehiro, J.-P. Laumond, and A. Monin, "On human motion imitation by humanoid robot," in *ICRA*, 2008, pp. 2697–2704.
- [6] K. Yamane, J. Kuffner, and J. Hodgins, "Synthesizing animations of human manipulation tasks," *ACM Transactions on Graphics (Proceedings of SIGGRAPH)*, vol. 23, no. 3, pp. 532–539, 2004.
- [7] S. Nakaoka, S. Kajita, and K. Yokoi, "Intuitive and flexible user interface for creating whole body motions of biped humanoid robots," in *IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, 2010, pp. 1675–1682.
- [8] S. Schaal, "Is imitation learning the route to humanoid robots?" *Trends in Cognitive Sciences*, vol. 3, no. 6, pp. 233–242, 1999.
- [9] S. Calinon and A. Billard, "Active teaching in robot programming by demonstration," in *IEEE International Conference on Robot and Human Interactive Communication*, 2007, pp. 702–707.
- [10] T. Inamura, N. Kojo, and M. Inaba, "Situation recognition and behavior induction based on geometric symbol representation of multimodal sensorimotor patterns," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2006, p. 51475152.
- [11] S. Wrede, C. Emmerich, R. Gruenberg, A. Nordmann, A. Swadzba, and J. Steil, "A user study on kinesthetic teaching of redundant robots in task and configuration space," *Journal of Human-Robot Interaction*, vol. 2, no. 1, pp. 56–81, 2013.
- [12] P. Kormushev, D. N. Nenchev, S. Calinon, and D. Caldwell, "Upper-body kinesthetic teaching of a free-standing humanoid robot," in *IEEE Int. Conf. on Robotics and Automation*, 2011.
- [13] D. N. Nenchev and A. Nishio, "Ankle and hip strategies for balance recovery of a biped subjected to an impact," *Robotica*, vol. 26, no. 5, pp. 643–653, 2008.
- [14] A. Ibanez, P. Bidaud, and V. Padois, "Unified preview control for humanoid postural stability and upper-limb interaction adaptation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 1801–1808.
- [15] D. Lee and C. Ott, "Incremental kinesthetic teaching of motion primitives using the refinement tube," *Autonomous Robots*, vol. 31, no. 2, pp. 115–131, 2011.
- [16] —, "Incremental motion primitive learning by physical coaching using impedance control," in *IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, 2010.
- [17] A. De Luca, A. Albu-Schäffer, S. Haddadin, and G. Hirzinger, "Collision detection and safe reaction with the dlr-iii lightweight manipulator arm," in *IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, 2006, pp. 1623–1630.
- [18] K. Schmid and H. Hirschmüller, "Stereo vision and imu based real-time ego-motion and depth image computation on a handheld device," in *IEEE Int. Conf. on Robotics and Automation*, 2013.
- [19] B. Paden and R. Panja, "Globally asymptotically stable 'pd+' controller for robot manipulators," *International Journal of Control*, vol. 47, no. 6, pp. 1697–1712, 1988.
- [20] C. Ott, M. A. Roa, and G. Hirzinger, "Posture and balance control for biped robots based on contact force optimization," in *IEEE-RAS International Conference on Humanoid Robots*, 2011, pp. 26–33.
- [21] S. Kajita, F. Kanehiro, K. Kaneko, K. Yokoi, and H. Hirukawa, "The 3d linear inverted pendulum mode: A simple modeling for a biped walking pattern generation," in *IEEE Int. Conf. on Robotics and Automation*, 2001, pp. 239–246.
- [22] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa, "Biped walking pattern generation by using preview control of zero-moment point," in *IEEE Int. Conf. on Robotics and Automation*, 2003, pp. 1620–1626.
- [23] J. M. Maciejowski, *Predictive Control with Constraints*. Prentice Hall, 2002.
- [24] T. Inamura, Y. Nakamura, and I. Toshima, "Embodied symbol emergence based on mimesis theory," *Int. Journal of Robotics Research*, vol. 23, no. 4, pp. 363–377, 2004.
- [25] D. Lee, C. Ott, and Y. Nakamura, "Mimetic communication model with compliant physical contact in human-humanoid interaction," *Int. Journal of Robotics Research*, vol. 29, no. 13, p. 16841704, 2010.
- [26] S. Calinon, F. D'halluin, E. Sauser, D. Caldwell, and A. G. Billard, "Learning and reproduction of gestures by imitation: An approach based on hidden markov model and gaussian mixture regression," *IEEE Robotics and Automation Magazine*, vol. 17, no. 2, pp. 44–54, 2010.
- [27] J. A. Blimes, "A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models," University of Berkeley, Tech. Rep. ICSI-TR-97-021, 1997.
- [28] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77(2), pp. 257–286, 1989.
- [29] D. A. Cohn, Z. Ghahramani, and M. I. Jordan, "Active learning with statistical models," *Journal of Artificial Intelligence Research*, vol. 4, p. 129145, 1996.
- [30] A. Albu-Schaeffer, S. Haddadin, C. Ott, A. Stemmer, T. Wimboeck, and G. Hirzinger, "The dlr lightweight robot - design and control concepts in human environments," *Industrial Robot: An International Journal*, vol. 34, no. 5, pp. 376–385, 2007.
- [31] A. Schmidts, D. Lee, and A. Peer, "Imitation learning of human grasping skills from motion and force data," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2011, pp. 1002–1007.