2015 IEEE International Conference on Robotics and Automation (ICRA)

Learning Movement Primitives for Force Interaction Tasks

Jens Kober[1], Michael Gienger[2], and Jochen J. Steil[3]

***Abstract*— Kinesthetic teaching is a promising approach to acquire robot skills in an intuitive way. This paper focuses on learning skills that do not solely rely on kinematics but also need to take into account interaction forces. We present three novel concepts towards learning such force interaction skills. Firstly, we determine segments from a small number of continuous kinesthetic demonstrations using contact information. Secondly, we associate each segment with a movement primitive, and determine its composition, i.e., the control variables and reference frames that allow to reproduce the demonstrated task. Lastly, we propose a concept to determine the transitions between the primitives during reproduction. The proposed methods are evaluated on a box pulling and flipping task, and show very good generalization abilities for objects with different geometries, and situations with different object arrangements.**

I. Introduction

Compliant manipulation is a hot topic in robotics research. It has received significant attention in the scientific community, both driven by its potential in various fields of application, and by the development of affordable and reliable robots with compliant control abilities [1]. In particular, torque-controlled robots offer advantages. Their backdriveability allows kinesthetic teaching, while recording both kinematic movement data and interaction forces with a wrist force-torque (FT) sensor. Other than with robots that are controlled actively compliant, the force measurements are undisturbed if taught appropriately (moving the links before the FT sensor). In addition, there is no need for mapping demonstration data between different embodiments (correspondence problem). Finally, it is intuitive to kinesthetically train a robot to perform interaction tasks, so that extensive programming can be avoided [2].

This opens up a wide range of interesting questions. In this paper, we focus on imitation learning of sequential movement tasks that can be represented by a sequence of elementary movements (movement primitives). The primitives are composed of a set of control variables, and can contain kinematic and force modalities. We propose an approach to simultaneously determine the composition (control variables) and parameters (attractor goals) in order to reconstruct and explain the observation data best.

The presented concepts are evaluated in a "pull and flip" task with a Barrett WAM robot, where the learned sequence generalizes to unseen initial positions and unseen box sizes, see Figs. 1 and 2 for an illustration. The underlying methods have the potential for other tasks, such as screwing, joining of parts, or simple assembly tasks in the industrial domain.

[1]J. Kober is with Delft Center for Systems and Control, TU Delft, 2628 CD Delft, The Netherlands, j.kober@tudelft.nl

[2]M. Gienger is with Honda Research Institute Europe, 63073 Offenbach, Germany, michael.gienger@honda-ri.de

[3]J. J. Steil is with CoR-Lab, Bielefeld University, 33615 Bielefeld, Germany, jsteil@cor-lab.uni-bielefeld.de

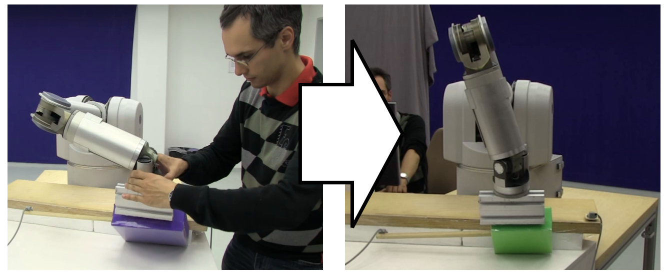

Fig. 1. Kinesthetic learning of a "pull and flip" task

A. Related Work

The approach presented in this paper approach covers the whole process of getting from human demonstrations to a robotic skill, with a particular focus on determining the movement primitive (MP) composition. Related papers often treat some sub-aspects of this process in isolation.

Decomposing the demonstrations into simpler sub-movements, also called segmentation has been studied extensively. Most previous work focuses on kinematic demonstrations. This problem is also highly relevant in the computer graphics domain, e.g., [3] employs principal component analysis to segment motion capture data. Similarly, non-negative matrix factorization has been employed in the robotics domain [4]. Alternatively probabilistic models can be utilized, e.g., [5] has employed hidden Markov models. Skill learning from visual observations via semantic event chains [6] does not directly yield a representation suitable for skill reproduction. Our approach is more closely related to [7] where correlative features between the hand and objects, which indicate whether an object is being manipulated actively, are used to detect segmentation points. Also from a biological point of view such contact events are highly relevant for dextrous manipulation [8]. We employ MPs based on linear attractors without additional modulation [9] that are comparatively easy to detect but might require more MPs to represent a movement than more complex MP representations.

Various representation and teaching methods for movements that need to consider force or compliance have been proposed. The weights of dynamic movement primitives (DMPs) [10] can be adapted in order to achieve a desired force profile [11] or additional modulation terms can be introduced [12]. Reinforcement learning has been employed to adapt finely discretized force trajectories to the desired task [13]. Force demonstrations can be supplied separately

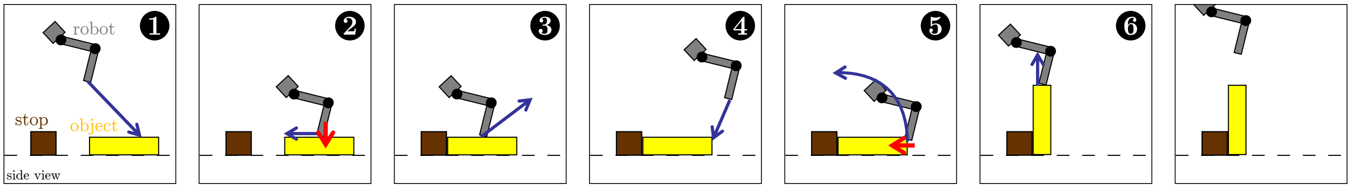

Fig. 2. This figure illustrates an example of the "pull and flip" task. The robot moves towards the object ❶, pulls the object towards the stop ❷, moves behind the object ❸–❹, flips the object up ❺, and finally moves away from the object ❻. Blue arrows indicate kinematic movements, the red ones applied forces.

using a haptic device as in [14] while [15] describes an interface for adapting stiffness online through human robot interaction. Similar to [16], [17] we simultaneously teach positions and forces by kinesthetic teach-in and employ linear dynamical systems as representation. Smooth transitions between consecutive DMPs have been studied in [18].

The sequence of MPs can be specified by hand [9], determined by motion planning [19], learned by reinforcement learning [20], [21] or extracted from demonstrations [16], [22]. Here we only consider a linear sequence and extract termination conditions from the demonstrations.

In most approaches, the composition of the learned MPs is defined by hand. The authors of [23] employ an Inverse Optimal Control approach. Other recent work applies statistical methods and prior task knowledge to find the appropriate task space composition [24].

There are only very few papers that propose approaches that address several of these topics simultaneously: Based on inter- and intra-demonstration variance [17] segments motions as well as assigns a reference frame and control modality. The extractedMPs are learned by SEDS [25]. How transitions between the MPs are handled is not discussed. The system is evaluated in a vegetable grating task.

In contrast, [16] considers pre-segmented MPs. The employed DMPs effectively encode the mean of the demonstrations while the reference frame and the control modality are specified by hand. Once a MP has finished its movement or when a deviation from the demonstrations is detected, the MP is executed for which the observed starting sensor values match the current ones best. This framework is illustrated by the DARPA ARM drilling task.

In [20] only kinematic control is considered. The demonstrations are segmented by change point detection. The individual MPs are modeled as controllers functioning as attractors that act on automatically selected relevant dimensions. The robot learns to navigate a maze via reinforcement learning where the MPs are considered to be options [26].

II. Movement Primitive Learning

In this section we will discuss the teaching procedure (Sect. II-A), as well as the four steps of our approach: The raw demonstrations are first aligned (Sect. II-B) and then segmented into candidate MPs (Sect. II-C). Our main focus lies on determining the MP composition (Sect. II-D). Finally we present a first proof of concept to determine the transitions between the recovered MPs (Sect. II-E). All steps are illustrated in Fig. 3 for a simplified example.

A. Kinesthetic Demonstrations

As method for demonstrations we are employing kinesthetic teach-in. By directly moving the robot through the task in gravity compensation mode, we do not have to solve the correspondence problem. In contrast to motion capture it is straightforward to record all the sensory information (especially the occurring forces) during demonstrations. For each time step of an observation, we record a vector $\mathbf{q} = (q_1 \dots q_d \; f_x \dots m_z \; \mathbf{o}_1^{\mathrm{T}} \dots \mathbf{o}_k^{\mathrm{T}})^{\mathrm{T}}$ comprising d joint angles, six interaction forces and torques at the FT sensor, and transformation vectors $\mathbf{o}$ of the k objects in the scene.

In a second step, we project these data on a vector of n pre-defined task variables $\mathbf{x} = (x_1 \dots x_n)^{\mathrm{T}} = \mathbf{f}(\mathbf{q})$. The task variables capture the positions, orientations and interaction forces of the end effector in the world coordinate system, or in one of the absolute or relative coordinate frames of the objects that are part of the scene (see Fig. 3). They solely depend on the configuration vector $\mathbf{q}$ of the scene. Task-level forces are compensated for the gravity forces of the robot links distal to the sensor. One out of the m kinesthetic demonstration comprises T_m time steps $\mathbf{X} = (\mathbf{x}_1^{\mathrm{T}}, \dots, \mathbf{x}_{T_m}^{\mathrm{T}})$, so that after the kinesthetic teaching, we have m task-level demonstrations $\mathbf{X}_{1\dots m}$.

B. Alignment

The multiple demonstrations $\mathbf{X}_{1\dots m}$ usually differ slightly in timing and movement velocity (see Fig. 3). In order to align those we employ Dynamic Time Warping (DTW) [27]. DTW aligns the demonstrations in time according to a similarity measure. For the tasks we are interested in, there is often a significant variation in the position dimensions, e.g., caused by different starting positions of an object, which needs to be reflected in the weights of the individual data dimensions in the similarity measure. We therefore assign high weights to force norms and velocity norms as well as to a binary feature that indicates whether the robot is in contact with the environment or not (see Sect. II-C).

C. Segmentation

Intuitively, getting into contact with the environment or an object as well as coming to a halt, mark transitions in a sequence of movements (see Fig. 3). For switching between position and force control it is also very important to detect contact changes. The aligned demonstrations are segmented jointly.

Potential segmentation points are added if there is a contact event or when the movement starts or stops. The latter

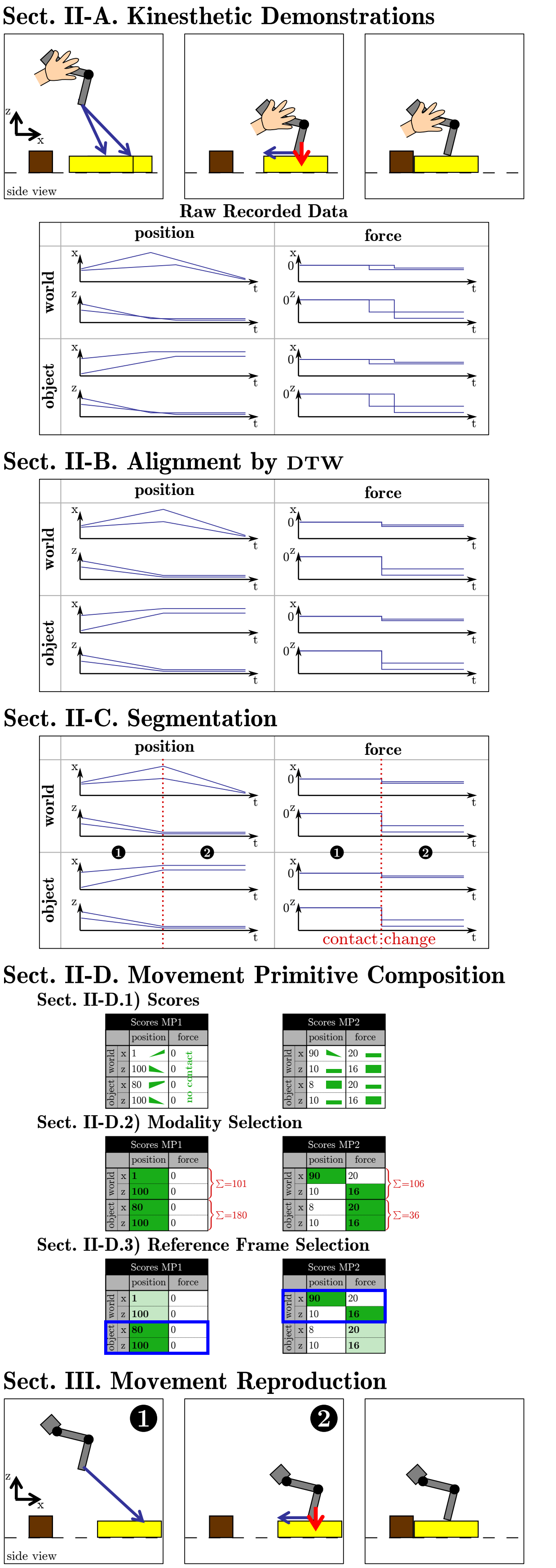

Fig. 3. This figure illustrates the overall process of our approach. Please refer to the corresponding sections for details.

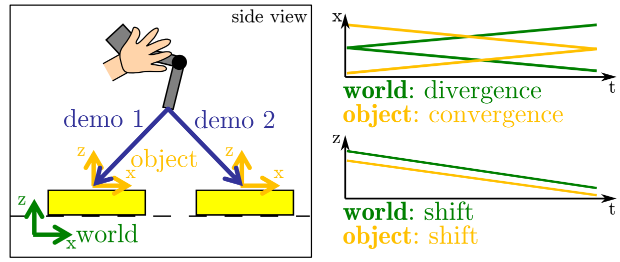

Fig. 4. This figure illustrates the basic idea of our approach. Demonstrations in different reference frames exhibit different convergence behavior: If the end-effector starts at a fixed position and moves towards the object, the demonstrations diverge in x-direction in world coordinates but converge in object coordinates. In z-direction they behave identical in both reference frames.

zero velocity crossings are especially relevant for free-space movements. The former can be detected in a variety of ways: considering correlations between hand and object movements, or sensor information related to touch. In the presented experiments we employed a threshold on the force norm.

Segmentation points stemming from the zero velocity crossing and contact events often occur together but are not perfectly aligned. Therefore, we merge segmentation points that differ only in a few frames in a post-processing step.

D. Movement Primitive Composition

Each of the detected segments is treated as a separate MP. Semantically identical MPs could be merged before the reproduction based on the similarity of their composition and goals. The core of our method lies in figuring out the most plausible composition of the MP. That is, we want to a) assign each MP a reference frame in which it acts and b) assign each individual coordinate within this reference frame a control modality, i.e., force or position control. Our MP representation is based on linear attractors [9], i.e., first order dynamical systems with a velocity limit, and without additional modulation. This representation results in movements that converge to a point. The basic idea of our approach is to use the convergence behavior of the demonstrations in the candidate reference frames to determine the best match.

For example if the end-effector always starts at the same position and is then moving towards an object that can be placed at arbitrary locations, we will observe a convergence behavior in object coordinates while the demonstrations are diverging in world coordinates (see Fig. 4). A converging behavior is characterized by a decrease in inter-demonstration variance over time, a diverging behavior by an increase in inter-demonstration variance.

Similarly if the demonstrations diverge in both reference frames, the one with the lower end variance is preferable.

In between converging and diverging behaviors we have the case where the inter-demonstration variance stays almost constant. This corresponds to a directional movement, i.e., a shift with constant velocity. This type of behavior occurs frequently when holding a position or force in the considered dimension. Also here we favor the frame with less variance.

To sum up, we prefer a converging behavior over a shifting behavior over a diverging behavior.

In our discussion so far we have concentrated on kinematic tasks, but we can also include the choice of modality by comparing the convergence behavior of the position values to those of the force values. For doing so we have to take into account the different scaling of the demonstrated values[1]. Similarly, if comparing reference frames with different types of coordinates we have to keep in mind the scaling, e.g., when comparing Cartesian coordinates and cylindrical coordinates.

1) Scores: We calculate a score for each candidate frame based on statistics obtained from the demonstrations. The number of candidates is determined by the number of reference frames, the dimensionality of the coordinate systems, and the number of modalities we consider. For example, if we have four reference frames with three dimensions each and differentiate between force and position control, we need to calculate $4 \times 3 \times 2 = 24$ score values (see Fig. 3).

The score is based on the change in variance as well as the end variance. Based on these scores we find the combination of candidates that matches the demonstrations best.

In order to calculate the scores we first fit two linear models to each candidate. These represent the respective mean and standard deviation as a function of the aligned time[2]. Next we determine how well the candidate matches each of the three behaviors explained above. We calculate two factors that can be seen as pseudo-probabilities for each behavior. The first factor expresses whether the candidate shows the desired convergence behavior, e.g., for the converging behavior it depends on how much the standard deviation decreases over time and we have

$$p_{\text{converging\_strongly}} = \operatorname{clip}_{[0,1]}\left(-\alpha \frac{d}{dt} \operatorname{std}\left(\widetilde{\mathbf{x}}_{\text{demo}}\right)\right),$$

where ${}^{d}/_{dt} \operatorname{std}\left(\widetilde{\mathbf{x}}_{\text{demo}}\right)$ is the change in the standard deviation according to the linear model, $\operatorname{clip}_{[\min,\max]}(\cdot)$ ensures that the resulting value is bounded between a maximum and a minimum, and α is a scaling factor that allows to change the cut-off at which we are certain to have a converging behavior. Instead of the linear function with cutoffs, we could also employ a sigmoid function similar to logistic regression. For the shifting behavior we evaluate how constant the standard deviation stays over time and for the diverging behavior how much it increases.

The second factor expresses how favorable the behavior in the demonstrations is. Taking again the example of the converging behavior we want to prefer candidates with low final standard deviations, e.g.,

$$p_{\text{converging\_favorably}} = \operatorname{clip}_{[0,1]}\left(1 - \beta \operatorname{std}\left(\widetilde{x}_{\text{demo,end}}\right)\right),$$

where $\operatorname{std}\left(\widetilde{x}_{\text{demo,end}}\right)$ is the standard deviation at the end of the segment according to the linear model and β a scaling factor that determines how large the final standard deviation can be maximally. For the shifting behavior the equivalent factor depends on the overall standard deviation, for the diverging behavior we employed a combination of low final standard deviation and how similar the movement directions of the demonstrations are. The open parameters are chosen such that they cover the range of behaviors observed in the demonstrations.

[1]A standard deviation of 0.01 m for position tasks is for example roughly equivalent to a standard deviation of 1 N for force tasks in our "pull and flip" skill.

[2]We employ L1 linear regression in order to improve robustness w.r.t. outliers that occur frequently at the start and at the end of a segment due to imprecise alignment and segmentation.

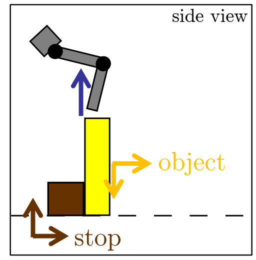

Fig. 5. This figure illustrates an ambiguity in the reference frames. The upward movement can be represented equally well in stop and object coordinates. However, if the flipping movement failed it is more desirable to move in stop coordinates, which cannot be extracted from the demonstrations as they do not contain failures.

The complete score for each behavior is then

$$p_{\text{behavior}} = \eta\, \gamma_{\text{behavior}}\, p_{\text{behavior\_strongly}}\, p_{\text{behavior\_favorably}}.$$

As we will discuss in Sect. II-D.5 we need to assign different priors for the modalities and reference frames to resolve ambiguities, which corresponds to the factor η. The factor γ_{behavior} can be seen as a prior for the individual behaviors. In our experiments we employed $\gamma_{\text{converging}} = 100$, $\gamma_{\text{shifting}} = 10$, and $\gamma_{\text{diverging}} = 1$. As the demonstration data is noisy we never get a perfectly constant standard deviation corresponding to the shifting behavior. Therefore we allow some overlap between the behaviors and assign $\max\left(p_{\text{converging}}, p_{\text{shifting}}, p_{\text{diverging}}\right)$ as the global score for the candidate. We also employ this mechanism to decide whether to use a MP based on a position attractor or a velocity attractor.

2) Modality Selection: First we select a modality per coordinate, i.e., whether the coordinate is force or position controlled, by selecting the modality with the maximum score for each coordinate of each reference frame (see Fig. 3). Now we are only left with the choice of the reference frame.

3) Reference Frame Selection: In this paper, we restrict all dimensions to be of the same reference frame. We calculate the sum of scores over the dimensions of each reference system in order to pick the best reference frame (see Fig. 3).

4) Parameter Selection: Finally, the MPs' parameters for the attractor (i.e., position, velocity, or force) and the velocity limit are obtained via linear regression.

5) Ambiguities & Priors: Obviously, with an increasing number of reference frames and types of control, the possible number of combinations increases very quickly. This introduces ambiguities where the reference frames behave essentially equivalently. An example is illustrated in Fig. 5. If a human were to decide upon the best reference frame, he or she would additionally take into account further criteria

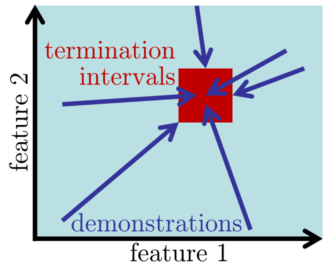

Fig. 6. This figure illustrates how the termination intervals are extracted from demonstrations.

such as generalization properties or error recovery properties. This is a general problem of learning from demonstrations which can be solved by introducing priors, by collecting more training data to disambiguate the choice, or by evaluating the reproductions. We suggest to incorporate some priors, however in a rather generic form for the class of tasks at hand. Firstly, we take into account that it is undesirable to have a force controlled component if we are not in contact with the environment as this could lead to dangerous accelerations. Secondly, if several reference frames have the same scores, we select the candidate that is associated with the object with the slowest dynamics, e.g., we prefer the world coordinate frame over the object frame, and objects with larger masses over those with lower ones.

E. Transitions within the Sequence

To reproduce the learned skill, we need to decide when to change from one MP to the next one. Currently we only consider linear sequences, therefore it is sufficient to detect when we need to switch to the next MP. Our MP representation is time-independent and only the position attractors have an explicit target. Similar to options [26] termination conditions or transition regions can be defined. The current features of the MP are compared to the values of the features at which the demonstrations terminated. In this paper, we employ the dimensions controlled by the current and subsequent MP as our features. For example, a "moving down" MP would be terminated if the height gets below a height limit observed in the demonstrations. Here we have to strike a balance between switching too early and not switching at all, especially when generalizing to unseen regions.

As a proof of concept intervals based on the features which spans the values at which the demonstrations terminated (see Fig. 6) are automatically constructed. Dimensions where the MP only holds the position or force are discarded for better generalization. This approach naturally includes the attractor targets while adding additional safeguards when the transition involves a switch in reference frames.

In a related project we are exploring more advanced models for transitions [22] where we include the possibility to have branching and repetitions. There the possible successor MPs are encoded in a transition graph where each node has an associated classifier that determines when to determinate the current MP and to which MP to switch next. We are currently working on integrating this more advanced transition model with the presented framework.

III. Movement Reproduction

In order to control force and kinematic control variables simultaneously, we employ a hybrid position-force control concept based on a task-level inverse dynamics approach previously described in [9]. The desired joint torque $\mathbf{T}$ is

$$\begin{aligned}\mathbf{T} = &\mathbf{M}\,\mathbf{J}^{\#}\mathbf{S}\left(\mathbf{a}_x - \dot{\mathbf{J}}\dot{\mathbf{q}}\right)\\ &+ \mathbf{J}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{S}\right)\mathbf{a}_f + \mathbf{M}\,\mathbf{J}^{\#}\left(\mathbf{I}-\mathbf{S}\right)\left(\mathbf{a}_d - \dot{\mathbf{J}}\dot{\mathbf{q}}\right)\\ &- \mathbf{M}\left(\mathbf{I}-\mathbf{J}^{\#}\mathbf{J}\right)\boldsymbol{\xi} + \mathbf{g} + \mathbf{h}.\end{aligned} \tag{1}$$

Vector $\mathbf{a}_x$ contains a PID control law based on a vector of augmented kinematic task descriptors (reference frame and control modalities of the MP) and their time derivatives. Vector $\mathbf{a}_f$ comprises the desired forces with an associated task-level damping term $\mathbf{a}_d$. Each of the task elements can be represented in various frames of reference as described in Sect. II-A. The diagonal selection matrix $\mathbf{S}$ enables either kinematic or force components of a task variable in the spirit of a hybrid position force controller [28]. Vector $\boldsymbol{\xi}$ accounts for joint speed damping and joint limit avoidance and is projected into the null space of the movement.

Vectors $\mathbf{g}$ and $\mathbf{h}$ comprise the gravity and Coriolis forces, $\mathbf{M}$ denotes the mass and inertia matrix, $\mathbf{J}$ and $\mathbf{J}^{\#}$ are the Jacobian of the task and its pseudo inverse. The pseudo inverse is scaled with a weighting matrix, which allows to continuously modulate the contribution of each task variables. Our approach allows for an arbitrary number of possibly conflicting tasks. However, here we only partition linearly independent tasks between the control modalities, and leave the conflict resolution between dependent task variables to the task frame selection.

As representation of the individual MPs we chose linear attractors. The attractor dynamics are driving the desired positions, velocities and forces (control modality) of the MP's reference frame, which enter the controller equation (1) as terms $\mathbf{a}_x$ and $\mathbf{a}_f$ (see Fig. 3). For example, a position MP determines the desired velocity $\dot{\mathbf{x}}_{\text{des}}$ based on the difference between the current position and the MP's goal as well as on the velocity limit. The first order dynamical system yields the desired position $\mathbf{x}_{\text{des}}$. Finally, the PID control law is applied to obtain $\mathbf{a}_x$ from $\mathbf{x}_{\text{des}}$ and $\dot{\mathbf{x}}_{\text{des}}$.

IV. Experiments

In this section we will discuss how the proposed approach performed on the "pull and flip" task experimentally. In Sect. IV-A we give a brief overview of the setup, in Sect. IV-C we contrast our method to the criterion from [17], and in Sect. IV-B we show how our approach performed on various demonstration sets.

A. System, Modalities & Reference Frames

As discussed in Sect. III we employ a hybrid position-force control concept. Hence we can control an individual

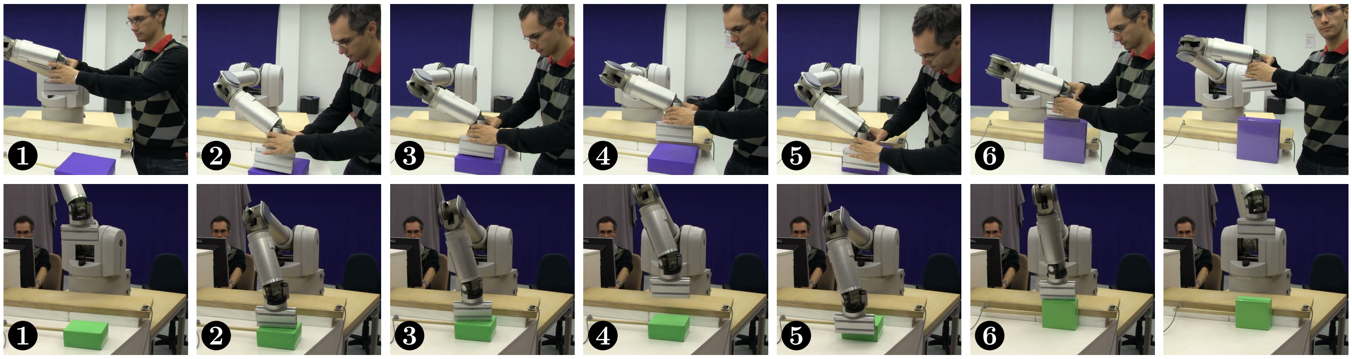

Fig. 8. This figure illustrates one kinesthetic demonstration and the learned "pull and flip" skill. The behavior was learned with demonstrations with the large object only. It generalizes to a different object size and unseen initial positions.

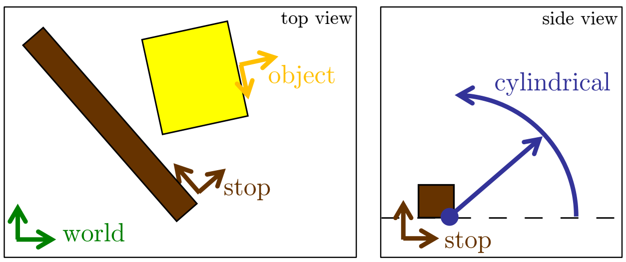

Fig. 7. This figure illustrates the coordinate frames employed in the experiments. There are Cartesian coordinates in the world reference frame, relative to the object, and relative to the stop, as well as cylindrical coordinated relative to the stop.

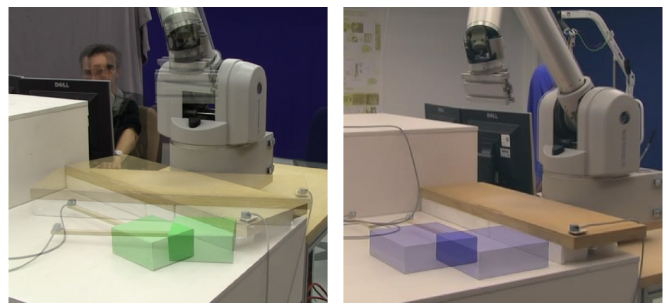

Fig. 9. This figure illustrates the generalization capabilities of the learned "pull and flip" skill. The skill generalizes to unseen positions and orientations of the stop, unseen initial positions of the object, as well as different object sizes.

coordinate either in position or force mode. For our scenario we defined four candidate reference frames, see Fig. 7. The first one is the end-effector in world coordinates, the second one is the end-effector relative to the object, the third one relative to the stop. Please note that all relative reference frames are not only shifted but also rotated with the reference object. The former reference frames are all in Cartesian coordinates. Finally, we have a reference frame relative to the stop in cylindrical coordinates. Its axis is aligned with the long side of the stop. The cylindrical coordinates work very well for many kinds of rotatory movements (e.g., opening doors) and are very robust w.r.t. inaccuracies in the rotation axis if is combined with force control.

The position and orientation of the stop and the object are tracked using a Polhemus Liberty electromagnetic motion tracker system. The robot distorts the magnetic field significantly and there can be errors up to approximately $0.07\,\mathrm{m}$. The robot is a torque controlled Barrett WAM with a Barrett FT sensor at the wrist. The metal end-effector is coated in order to increase friction between it and the object while the underside of the object has a low friction material.

B. Results

We collected nine demonstrations each with a larger object ($0.2\,\mathrm{m} \times 0.2\,\mathrm{m} \times 0.065\,\mathrm{m}$) and a smaller object ($0.12\,\mathrm{m} \times 0.15\,\mathrm{m} \times 0.065\,\mathrm{m}$) including various initial positions of the object and varied the orientation and position of the stop ($0.25\,\mathrm{m} \times 0.8\,\mathrm{m} \times 0.08\,\mathrm{m}$). Our approach was evaluated on each data set individually and on the combined data. The resulting sequence consists of 14 MPs for the skills learned from the large and combined object sizes and 13 MPs for the small one. The MPs match the structure given in Fig. 2 while some MPs were split and we have some additional MPs where we paused briefly during the demonstration.

All learned MPs managed to successfully reproduce the demonstrated behavior and to generalize to unseen initial positions of the object and stop (see Fig. 9 and the attached video). The skill trained with the larger object sizes manages successfully to generalize to the smaller object size as we have demonstrated a movement pushing the object horizontally against the stop in between ❹ and ❺. If this MP is not demonstrated and the object was not very close to the stop after ❷ , flipping up the box ❺ often fails as the robot moves up too quickly and looses contact with the object. Hence, we included this pushing behavior in the demonstrations to compensate for imperfect reproductions of the pull, which the additional benefit of generalizing to smaller boxes. For the skill trained with the smaller box, moving behind the object ❹ fails with the larger object as the learned position is relative to the stop and too close for the larger object. The skill learned from the combined data set copes nicely with both object by pulling the object ❷ until they hit the stop irrespective of their size.

For all experiments the movement towards the object ❶ was expressed in the reference frame relative to the object, and

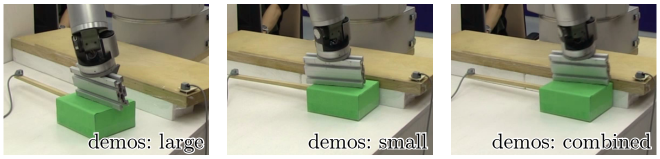

Fig. 10. This figure illustrates the recovered pulling MPs ❷. When training with the demonstration set containing only one box size, we get position attractors. Please notice the gap in the left snapshot. When training with the combined demonstration, we get a force MP.

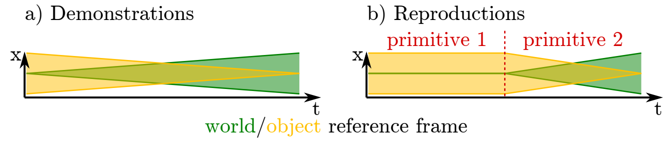

Fig. 11. Movement in x-direction of the toy example (Fig. 4 or ❶ in Fig. 2). a) Demonstrations b) Reproductions resulting from the criterion in [17].

the flipping movement ❺ in cylindrical coordinates relative to the stop. For most other MPs the Cartesian or cylindrical reference frames relative to the stop where chosen. All pulling movements ❷ employ a downward force and all flipping movements ❺ a force in radial direction.

However, there are some noticeable differences between the MPs resulting from the three training sets. For the pulling movement ❷ of the skills resulting from the single object sizes the approach chose position attractors relative to the stop, which results in the generalization properties discussed above (see Fig. 10). For the combined training set we got a MP that increases the force in horizontal direction until the object hits the stop. This force-based pulling is less robust as the downward force needs to be strong enough to avoid slippage of the end-effector, which accelerates rapidly once it has lost contact. This is also the only MP where the extracted transitions failed consistently.

For moving behind the object ❹, the two data sets with the single box size resulted in MPs that are moving to an absolute position relative to the stop, for the combined set we get a directional movement where again the burden of stopping at the correct point is placed on the transition mechanism. Given the simplistic nature of the transition mechanism it worked surprisingly well. As we are taking the whole interval of demonstrated end points, transitions sometimes occurred too early (e.g., for moving behind the object in the skill learned from the combined set) or were sometimes not detected at all when starting far from the demonstrated initial conditions. Ignoring the pulling and moving behind MP of the combined data set, the transitions failed less than 10%.

C. Comparison: Alternative Criterion

Of the papers discussed in Sect. I-A, [17] is the closest to the focus of our approach. The paper describes a criterion for selecting the best reference frame, the best modality, and for segmenting the demonstration. It is defined as

$$C = \mathrm{var}_{\mathrm{intra}} - \mathrm{var}_{\mathrm{inter}},$$

where $\mathrm{var}_{\mathrm{intra}}$ is the variance over a time window (i.e., how much the variable changes over the course of a single demonstration) and $\mathrm{var}_{\mathrm{inter}}$ is the variance over trials (i.e., how similar the demonstrations where). We applied this criterion to our data in order to evaluate the respective benefits and shortcomings of the methods. The open scaling parameters and window sizes where tuned by hand.

In contrast to their paper, the MPs we want to learn based on the segmentation and detected composition are a lot less expressive, hence, a coarser segmentation would be sufficient for the approach in [17]. Contrary to this observation, the proposed criterion tends to over-segment our data. Let us take a look at the example from Fig. 4 again. In Fig. 11a we have plotted the converging and diverging behavior of the x-coordinate. The intra-demonstration variance $\mathrm{var}_{\mathrm{intra}}$ is going to be constant over the whole time and have the same value for both reference frames. Therefore, C depends only on the variance over the trials $\mathrm{var}_{\mathrm{inter}}$. The one related to the object reference frame is decreasing, the other one increasing. Initially the criterion would choose the world reference frame (as it has a smaller variance) and then switch to the object reference frame after 50% of the time. As a result we would get two MPs, where initially the position is held or would diverge randomly, and then the end-effector would move rapidly towards the object. This example illustrates the problems we encountered with the purely statistics-based segmentation and deciding upon the composition of a MP by comparing variances without taking into account the convergence behavior of the movement. Fig. 12 illustrates the same problem on real data.

Furthermore, introducing segmentation points both for switches in the reference frames and the modalities separately resulted in rather small MPs. Therefore the resulting skill will need to rely more on the transition mechanism than on the MP representation. The approach often picks force MPs even if the robot is moving in free-space, which would also have to be discarded in an additional step. We manged to find parameters that resulted in reasonable MPs for parts of the sequence but could not get it to run on the whole sequence.

V. Conclusion

We have proposed a number of concepts that allow to acquire force interaction skills for robots from a small number of kinesthetic demonstrations. The demonstrations are decomposed into our notion of attractor-based Movement Primitives (MPs) using contact information. The main novelty of this paper is to acquire a task-level representation for the MPs which gives it the ability to generalize to different situations, such as to objects with different geometrical or physical properties, and to different arrangements of the objects in the work space. We propose an approach that combines statistical and prior information to disambiguate the MPs reference frame, and the controlled variables within the selected frame. To reproduce the demonstrated skills, we apply a data-driven way to determine the transitions between MPs, and a hybrid position/force controller to track the forces and positions coming from the MPs attractor dynamics.

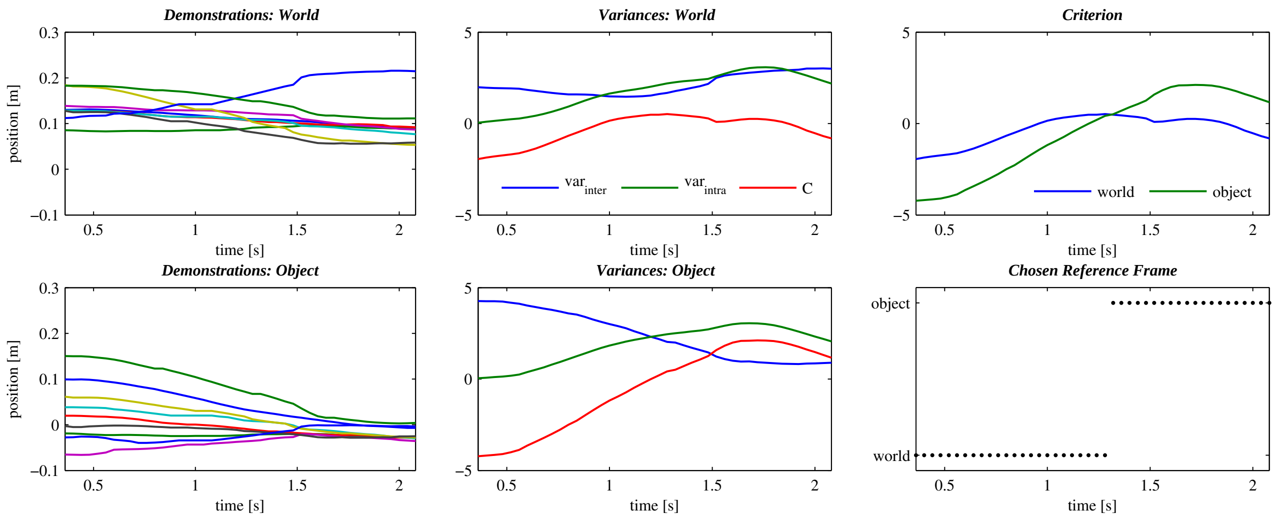

Fig. 12. Results of the criterion of [17] on real world data. Left: raw demonstrations, Middle: variance over trials, sliding variance within a trial, and resulting criterion, Right: selected reference frames.

The concepts have been evaluated in a box pulling and flipping task, and show very good generalization abilities for objects with different sizes, masses, and situations with different object arrangements. The system can generalize the learned skill to unseen situation as shown in our box pulling and flipping experiments.

REFERENCES

[1] T. Brogårdh, "Present and future robot control development—an industrial perspective," *Annual Reviews in Control*, vol. 31, no. 1, pp. 69–79, 2007.

[2] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Springer Handbook of Robotics* (B. Siciliano and O. Khatib, eds.), pp. 1371–1394, Springer Berlin Heidelberg, 2008.

[3] J. Barbič, A. Safonova, J.-Y. Pan, C. Faloutsos, J. K. Hodgins, and N. S. Pollard, "Segmenting motion capture data into distinct behaviors," in *Graphics Interface Conference*, 2004.

[4] O. Mangin and P.-Y. Oudeyer, "Unsupervised learning of simultaneous motor primitives through imitation," in *Frontiers in Computational Neuroscience Conference Abstract: IEEE ICDL-EPIROB*, 2011.

[5] D. Kulić, C. Ott, D. Lee, J. Ishikawa, and Y. Nakamura, "Incremental learning of full body motion primitives and their sequencing through human motion observation," *International Journal of Robotics Research*, vol. 31, no. 3, pp. 330–345, 2012.

[6] E. E. Aksoy, A. Abramov, J. Dörr, K. Ning, B. Dellen, and F. Wörgötter, "Learning the semantics of object-action relations by observation," *International Journal of Robotics Research*, vol. 30, no. 10, pp. 1229–1249, 2011.

[7] M. Mühlig, M. Gienger, and J. J. Steil, "Human-robot interaction for learning and adaptation of object movements," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2010.

[8] J. R. Flanagan, M. C. Bowman, and R. S. Johansson, "Control strategies in object manipulation tasks.," *Current Opinion in Neurobiology*, vol. 16, no. 6, pp. 650–659, 2006.

[9] T. Luksch, M. Gienger, M. Mühlig, and T. Yoshiike, "Adaptive movement sequences and predictive decisions based on hierarchical dynamical systems," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012.

[10] A. J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement imitation with nonlinear dynamical systems in humanoid robots," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2002.

[11] A. Gams, M. Do, A. Ude, T. Asfour, and R. Dillmann, "On-line periodic movement and force-profile learning for adaptation to new surfaces," in *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2010.

[12] A. Gams, B. Nemec, A. Ijspeert, and A. Ude, "Coupling movement primitives: Interaction with the environment and bimanual tasks," *IEEE Transactions on Robotics*, vol. 30, no. 4, pp. 816–830, 2014.

[13] M. Kalakrishnan, L. Righetti, P. Pastor, and S. Schaal, "Learning force control policies for compliant manipulation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011.

[14] P. Kormushev, S. Calinon, and D. G. Caldwell, "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input," *Advanced Robotics*, vol. 25, no. 5, pp. 581–603, 2011.

[15] K. Kronander and A. Billard, "Online learning of varying stiffness through physical human-robot interaction," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2012.

[16] P. Pastor, M. Kalakrishnan, L. Righetti, and S. Schaal, "Towards associative skill memories," in *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2012.

[17] A. L. Păiş, K. Umezawa, Y. Nakamura, and A. Billard, "Learning robot skills through motion segmentation and constraints extraction," in *HRI Workshop on Collaborative Manipulation*, 2013.

[18] B. Nemec and A. Ude, "Action sequencing using dynamic movement primitives," *Robotica*, vol. 30, no. 05, pp. 837–846, 2012.

[19] M. Kallmann, R. Bargmann, and M. Mataric, "Planning the sequencing of movement primitives," in *International Conference on Simulation of Adaptive Behavior (SAB)*, 2004.

[20] G. Konidaris, S. Kuindersma, R. A. Grupen, and A. G. Barto, "Robot learning from demonstration by constructing skill trees," *International Journal of Robotics Research*, vol. 31, no. 3, pp. 360–375, 2012.

[21] C. Daniel, G. Neumann, O. Kroemer, and J. Peters, "Learning sequential motor tasks," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2013.

[22] S. Manschitz, J. Kober, M. Gienger, and J. Peters, "Learning to sequence movement primitives from demonstrations," in *IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*, 2014.

[23] N. Jetchev and M. Toussaint, "Discovering relevant task spaces using inverse feedback control," *Autonomous Robots*, vol. 37, no. 2, pp. 169–189, 2014.

[24] M. Mühlig, M. Gienger, J. J. Steil, and C. Goerick, "Automatic Selection of Task Spaces for Imitation Learning," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2009.

[25] S. Khansari-Zadeh and A. Billard, "Learning stable nonlinear dynamical systems with Gaussian mixture models," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 943–957, 2011.

[26] R. S. Sutton, D. Precup, and S. Singh, "Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning," *Artificial Intelligence*, vol. 112, no. 1, pp. 181–211, 1999.

[27] R. Bellman and R. Kalaba, "On adaptive control processes," *IRE Transactions on Automatic Control*, vol. 4, no. 2, pp. 1–9, 1959.

[28] M. H. Raibert and J. J. Craig, "Hybrid position / force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, vol. 103, no. 2, pp. 126–133, 1981.