Learning Daily Task Motions for Humanoid Robot with Stochastic Model of Human Demonstration

Hideaki Ito
The University of Tokyo, JSK Lab.

I. Introduction

In this research, we have been focusing on how the robot can learn its motor skills for what we are doing in everyday life by observing human demonstrations. Previously, we have developed two frameworks; one is to execute human-robot cooperative tasks and the other is to store the motion information, which is mainly the trajectory of the hands. These frameworks made reproduction of the motions available. In both systems, in order to acquire the motion trajectories for imitatation or remembering, the robot observes human demonstrations. However, these pure imitations had some limitations in terms of speed and generalization. The memory also reaches the capacity limit if the robot continues to store the whole trajectory data in each task. To overcome these limitations, in this paper, we will introduce a generalizing method considering a stochastic approach.

II. Learning Stochastic Models

In this section, we introduce the stochastic model to express human demonstrations in Sec. II-A, and describe how to reproduce a certain trajectory using the learned model in Sec. II-B.

A. Task Generalization using HMM

We use Hidden Markov Model (HMM) to generalize task motions and learn stochastic models from human demonstrations. HMM is a model assumed to be a Markov process with unobservable states, and so that, can deal with time series data, which is trajectory data of hands (or robot grippers) in our case. The parameters in HMM is computed by EM algorithm. With the transition probability, the output probability of i th state can be acquired with the form of distribution $\mathcal{N}(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$.

B. Trajectory Reproduction using LQR

The trajectory can be reproduced by solving Linear Quadratic Regulator (LQR) problem, which estimates an optimized controller input $\boldsymbol{u}_t$ for each descrete timestep in the linear dynamical system,

$$\boldsymbol{\zeta}_{t+1} = \boldsymbol{A}_t\boldsymbol{\zeta}_t + \boldsymbol{B}_t\boldsymbol{u}_t, \tag{1}$$

which can be transformed into the matrix form, by considering a linear uncontrained problem and expressing all future states $\boldsymbol{\zeta}_t$ with the initial state $\boldsymbol{\zeta}_1$;

$$\boldsymbol{\zeta} = \boldsymbol{S}_{\zeta}\boldsymbol{\zeta}_1 + \boldsymbol{S}_{\boldsymbol{u}}\boldsymbol{u}. \tag{2}$$

Now, under the linear unconstrained system represented in Eq. (2), LQR problem is expressed as the minimization of the cost,

$$\begin{aligned} c &= \sum_{t=1}^{T}\left((\boldsymbol{\mu}_t - \boldsymbol{\zeta}_t)^\top \boldsymbol{Q}_t(\boldsymbol{\mu}_t - \boldsymbol{\zeta}_t) + \boldsymbol{u}_t^\top \boldsymbol{R}_t\boldsymbol{u}_t\right) \\ &= (\boldsymbol{\mu}_s - \boldsymbol{\zeta}_s)^\top \boldsymbol{Q}_s(\boldsymbol{\mu}_s - \boldsymbol{\zeta}_s) + \boldsymbol{u}_s^\top \boldsymbol{R}_s\boldsymbol{u}_s. \end{aligned} \tag{3}$$

The computation of the above optimization problem leads to a distribution $\mathcal{N}(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\zeta}})$ in task space with paramerters (See [1] for more details),

$$\hat{\boldsymbol{\zeta}} = \boldsymbol{S}_{\zeta}\boldsymbol{\zeta}_1 + \boldsymbol{S}_{\boldsymbol{u}}\hat{\boldsymbol{u}} \tag{4}$$

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{\zeta}} = \boldsymbol{S}_{\boldsymbol{u}}\left(\boldsymbol{S}_{\boldsymbol{u}}^\top \boldsymbol{Q}_s\boldsymbol{S}_{\boldsymbol{u}} + \boldsymbol{R}_s\right)^{-1}\boldsymbol{S}_{\boldsymbol{u}}^\top. \tag{5}$$

$$\text{with} \quad \hat{\boldsymbol{u}} = \hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}}\boldsymbol{S}_{\boldsymbol{u}}^\top \boldsymbol{Q}_s\left(\boldsymbol{\mu}_s - \boldsymbol{S}_{\zeta}\boldsymbol{\zeta}_1\right)$$

$$\hat{\boldsymbol{\Sigma}}_{\boldsymbol{u}} = \left(\boldsymbol{S}_{\boldsymbol{u}}^\top \boldsymbol{Q}_s\boldsymbol{S}_{\boldsymbol{u}} + \boldsymbol{R}_s\right)^{-1}$$

Assuming $\boldsymbol{Q}_t$ and $\boldsymbol{\mu}_t$ to be the HMM parameters $\boldsymbol{\Sigma}_{s_t}^{-1}$ and $\boldsymbol{\mu}_{s_t}$ in the stepwise references, the trajectory can be estimated and reproduced though the whole steps by Eq. (4) and Eq. (5) ($\boldsymbol{R}_t$, which is a regularization term to consider the smoothness of the reproduced trajectory, is set as $\rho\boldsymbol{I}$).

C. Trajectory Reproduction with Multi Coordinates

When the task is something like approaching things or pushing buttons, it is not enough only to consider the sequence referenced from the absolute coordinate system (world coordinate system). To take both the task coordinates (ex. robot grippers) and the target coordinates (ex. objects to be reached) into account, we can simply replace the first term in Eq. (3) with

$$\sum_{i=1}^{2}(\boldsymbol{\mu}_{s,i} - \boldsymbol{\zeta}_s)^\top \boldsymbol{Q}_{s,i}(\boldsymbol{\mu}_{s,i} - \boldsymbol{\zeta}_s).$$

The replaced minimization problem can be solved in the same way as before, yeilding a new distribution $\mathcal{N}(\hat{\boldsymbol{\zeta}}_{multi}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\zeta}_{multi}})$. It can be proven that $\mathcal{N}(\hat{\boldsymbol{\zeta}}_{multi}, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\zeta}_{multi}})$ corresponds to the product of the two Gaussians $\mathcal{N}(\hat{\boldsymbol{\zeta}}_1, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\zeta}_1})$ and $\mathcal{N}(\hat{\boldsymbol{\zeta}}_2, \hat{\boldsymbol{\Sigma}}_{\boldsymbol{\zeta}_2})$, which is the distribution yeilded by solving Eq. (3) separately in each coordinate system (See [2] for more details).

III. Experiments and Results

We have carried out some simulations and experiments to confirm whether our learning process works.

A. Opening a Garbage Bag

First motion is to open a garbage bag by inflating with air. Since this motion requires enough speed to bring air in the bag, the motion trajectory has to be predefined with the speed configuration in order to apply it to the real robot which has a speed limitation on each joint. Therefore, it is suitable to use our learning process, which can reproduce the whole trajectory beforehand with the reference to the human demonstrations, also taking smoothness into account.

For human demonstration data, we used the time series data of the positions of both right and left hands referenced from the coordinate system defined on the face position.

Fig. 1 shows the predicted HMM model and the output trajectories of the opening-a-garbage-bag motion.

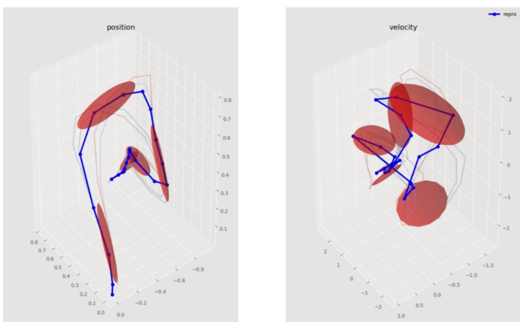

Fig. 1. Reconstructed trajectory of opening a garbage bag using HMM parameters and solving LQR. Left and right figure shows position and velocity information respectively. Blue thick lines represent the reproduced trajectory. Red ellipsoids represent Gaussians which are estimated by HMM from human demonstrations, which is shown with thin colored lines ($N = 4$). Note that the scale of each axis is normalized in positional space so that the scale can be ignored when estimating HMM parameters.

Fig. 2 shows the snapshots of the real robot motion applying the reproduced trajectory above. The speed seemed to be fast enough to accomplish the task, comparing with the human. However, as a result, the bag didn't expand enough. The main reason seemed to be that the robot didn't use its wrists for now and so that it couldn't bring enough air into the bag. We are now working with that problem as one of the future works.

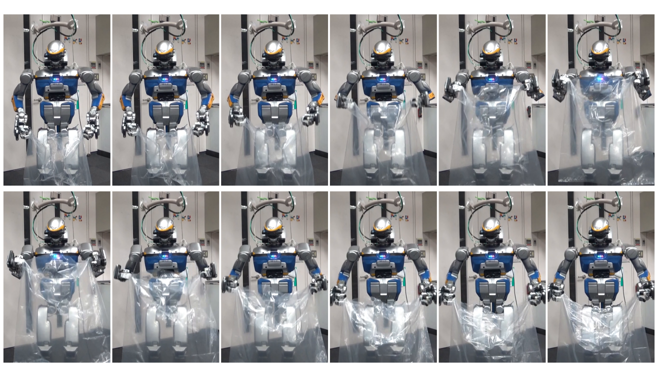

Fig. 2. Snapshots of executing reproduced opening-a-garbage-bag trajectory with a humanoid robot. The result wasn't so efficient (the bag didn't expand enough), although the speed of the motion seemed to be fast enough. One reason might be the non-use of the wrists.

B. Pushing buttons

While the previous task, opening a garbage bag, has no target to approach or to touch, the target coordinates has to be considered when pushing buttons. We brought out an experimental environment, which is shown in the left figure in Fig. 3, and recorded some trajectories which reach the different buttons. As described in Sec. II-C, we applied the multi coordinates version of HMM learning, whose two base coordinate systems are one set on the initial position and the other on the position of the target, which is the button here.

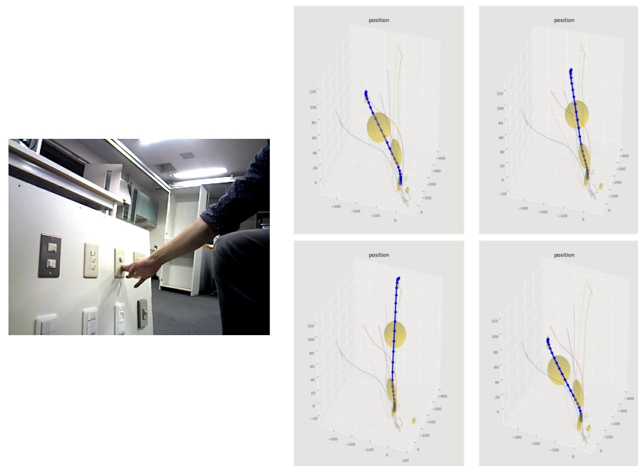

Fig. 3. Left: Experimental setup for collecting human demonstration data of pushing buttons. Buttons are aligned in the same plane. Right: Reproduced trajectory considering multi coordinate system when learning HMM. Yellow ellipsoids represent Gaussians which are the product of two Gaussians considering the initial coordinate system and the button coordinate system.

The right figures in Fig. 3 show the reproduced trajectories whose initial positions are different from each other. Looking at the final position in each trajectory, it surely tend to reach the same position, which means the robot can reach each button correcty. We'll apply the trajectory into the real robot and confirm the accuracy.

IV. Conclusions

We have constructed the system to learn motions with stochastic models and reproduce them from the learned model.

For the future work, we come back to human-robot cooperation tasks, which the human and the robot needs to work synchronously. Now, the robot can acquire the motion in probablistic manner with the system in this paper. The robot can also reproduce each motion from the stochastically learned model information. Here, it is the timing that has to be considered to accomplish the synchronous tasks; when to start and finish the motion.

References

[1] Emmanuel Pignat and Sylvain Calinon. Learning adaptive dressing assistance from human demonstration. *Robotics and Autonomous Systems*, Vol. 93, pp. 61 – 75, 2017.

[2] S. Calinon, F. Guenter, and A. Billard. On learning, representing, and generalizing a task in a humanoid robot. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 37, No. 2, pp. 286–298, April 2007.