Jibobot: A Personal Assistant Robot with Social Motions

Eun-Sol Kim, Jiseob Kim, Donghyun Kwak, Kyoung-Woon On and Byoung-Tak Zhang
Department Computer Science and Engineering
Seoul, 08826, South Korea
Email: {eskim, jkim, dhkwak, kwon, btzhang}@bi.snu.ac.kr

***Abstract*—For a personal assistant robot to be attractive, it is crucial that it generates social behaviors that are shared with human users. However, the robot body is different from the human body and mapping human behavior into robot behavior is a challenge. Here we develop a reinforcement learning method for personal robots to generate human-adaptive social behaviors from continuous feedback. We first build a basic motion generation model using a Gaussian Process Dynamic Model (GPDM) that is trained on a human activity dataset. Then, we continuously personalize the generated motions with reward feedback through a Bayesian optimization method for reward regression. The method is evaluated using a personal robot platform to generate seven kinds of robot motions using training data from 700 human motions. Analysis of our results demonstrates the potential for continuous teaching of robot social behaviors using in-situ human feedback.**

We aim at generating social behavior for a personal assistant robot. Fifteen status classes are defined and initial robot motions are learned from human activity data. An initial robot motion represents a status class, so fifteen initial robot motions are generated. The status classes consist of two groups, six emotional statuses and nine informative statuses. Emotional statuses are designed to respond to the human emotion and the following representative emotional categories are selected: happy, sad, angry, bored, surprised and interested. Informative statuses aim to transfer information to humans through robot motions. We selected nine situations: looking around, power on, power off, sleep, warning, notification, alarm, loading, listening and thinking. In this paper, seven status classes are concerned for the range of this problem: happy, sad, angry, bored, looking around, power on and power off.

From the initial robot motion, the system tries to find an optimal robot motion which gets a positive reaction from the human. To do this, the system gets continuous feedback (rewards) about the motion from a human. Using these real-valued rewards, the system finds a function which estimates a reward value of a motion. From the regressed function, the system can generate an optimal robot motion which gets a maximum reward value.

The overall system consists of three parts (Figure 1).

First, we collected human activity data. To find basic robot motions representative to each status class, we used a strategy to mimic human activity. For each status class, human subjects did motions related to the class. The human behavior is recorded using a Kinect depth camera, and then the position coordinates of the joints are extracted from the camera images.

In the second part, basic robot motions for each status class are generated by Gaussian Process Dynamic Model (GPDM). In this step, a low-dimensional sequence for a status class is generated and the sequence is used as the position value sequence of the motion.

In the third step, the system gets human feedback about the robot motions in the form of continuous real-values. Starting with the basic robot motion, the system gets a reward value from a human and then shows a new robot motion to be evaluated. After a small number of iterations, getting a reward from a human and showing a new robot motion, the system finds a reward regression function to find the maximum reward and the corresponding motion.

A. Basic Motion Generation

We suggest GPDM [2], [3] as a method to generate basic robot motions of each status class. The reasons GPDM is used are 1) it compresses nonlinear time series data into a low-dimensional latent space, smoothly; 2) it does not require a large number of datasets for training; 3) the dimension size of the latent space can be chosen without changing the algorithm.

A GPDM is composed of a low-dimensional latent space with associated dynamics, and a map from the latent space to an observation space. The model parameters, related to the dynamics and the mapping, are marginalized out in closed-form using Gaussian Process priors [1]. Therefore, smooth low-dimensional representations of the nonlinear time series data can be obtained with GPDM.

In our problem setting, the GPDM is used to find basic robot motions (low-dimensional latent space) trained on the human activity dataset (nonlinear time series data).

B. Human-Adaptive Robot Motions

After finding basic robot motions modeled after human activities by GPDM, the system keeps finding better robot motions which can get positive responses from a human. In this paper, we suggest a Bayesian optimization method as an exploration method to find better robot motions.

The main idea of finding an optimal robot motion is getting rewards about the current motion from a human and changing the robot motion until the optimal robot motion is discovered. The system gets real-valued rewards from a human about

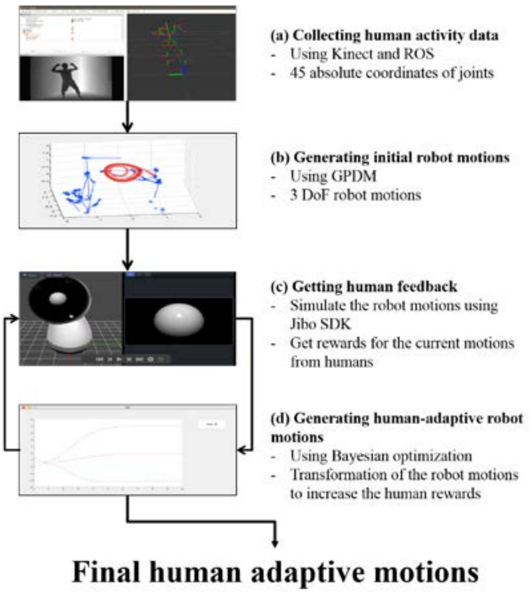

Fig. 1. The overall framework for generating human-adaptive robot motions. (a) Human activity data is collected using Kinect and ROS. Each human motion sequence consists of 45 absolute coordinates of 15 joints and contains 300-400 frames. (b) The human activity data is compressed into low-dimensional latent space by GPDM. Here, we assume the robot platform has three degrees-of-freedom, so the latent space is three- dimsional. In this step, a three-dimensional sequence is generated and is used as the basic motion of the robot. (c)-(d) Human evaluates the motion in terms of a real-value. The system gets the rewards over and over and finds a mapping function from the robot motions to the human rewards. After estimating the reward functions, the system finds the maximum value of the reward and the correspond robot motion which is most attractive to the human-user.

how the motion is relevant or representative of the given classes. After getting a reward, the motion is transformed and the system again requests a human to input a reward. This procedure is repeated until the system finds the optimal robot motions corresponding to the highest reward value.

C. Experimental Results and Conclusion

In Figure 2, a sequence of inspection points and corresponding robot motions are described through Bayesian optimization. As the procedure proceeds, we can see that various motions are explored such as changing motors or velocities.

In this work, we showed a new reinforcement learning method to generate human-adaptive robot motions. The generated robot motions are learned from continuous human feedback, therefore the motions can be regarded as the robot motions having the maximum human rewards. The suggested method could have novelty as the robot motions are evolving in a direction getting more positive human rewards automatically. Also, this method easily applies to other robot platforms which have different robotic structures.

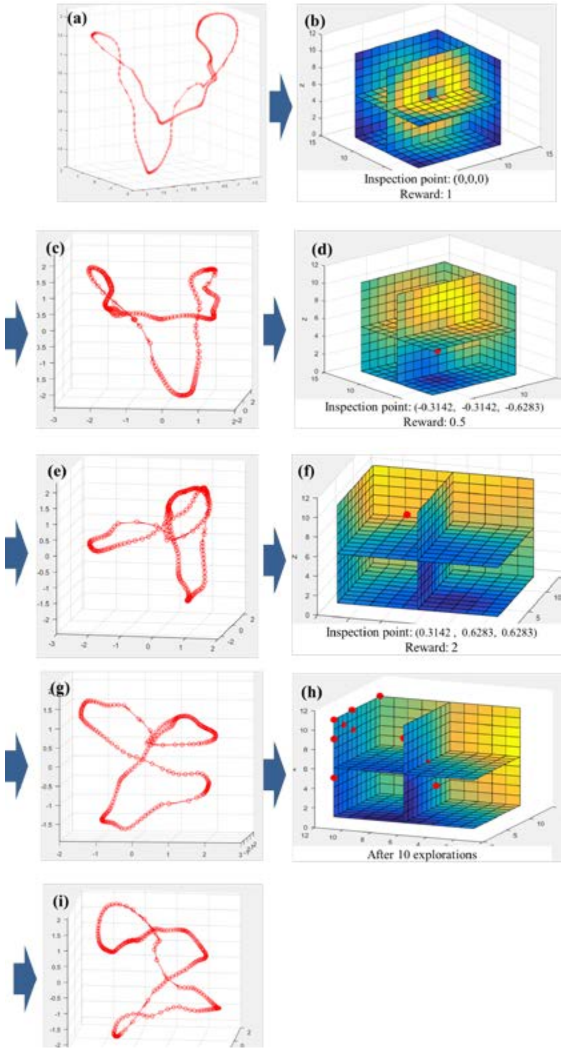

Fig. 2. A process about reward regression and corresponding robot motions. Human evaluates the robot motions after seeing the transformed robot motions.

REFERENCES

[1] C. E. Rasmussen. Gaussian processes for machine learning. 2006.

[2] J. Wang, A. Hertzmann, and D. M. Blei. Gaussian process dynamical models. In *Advances in neural information processing systems*, pages 1441–1448, 2005.

[3] J. M. Wang, D. J. Fleet, and A. Hertzmann. Gaussian process dynamical models for human motion. *IEEE transactions on pattern analysis and machine intelligence*, 30(2):283–298, 2008.