Predicting Whole Body Motion Trajectories using Conditional Neural Movement Primitives

Mehmet Hakan Kurtoglu[1], Yunus Seker[1], Evren Samur[2] and Emre Ugur[1]
[1]Computer Engineering Department, Bogazici University, Istanbul, Turkey.
[2]Department of Mechanical Engineering, Bogazici University, Istanbul, Turkey.

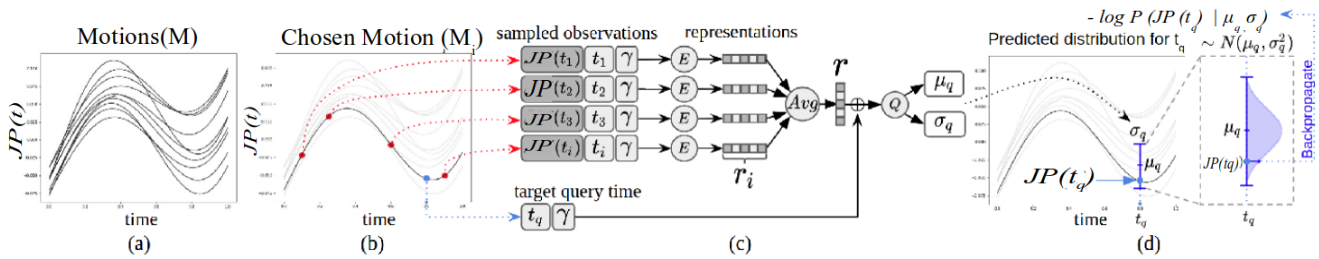


Fig. 1: CNMP framework for learning full body motion. Given multiple trajectories for a single or multiple behaviors, the system learns a robust representation that encodes the distribution of the trajectories. While our framework handles multiple body parts simultaneously, the figure shows a single dimension JP that corresponds to a joint or a part. The figure is adapted from [1]

I. INTRODUCTION

The purpose of this study is to demonstrate a new approach to whole body motion prediction problem using a state-of-the-art robot learning from demonstration framework, the Conditional Neural Movement Primitives (CNMPs)[1]. Previous experiments on robotic learning have shown the capability of CNMPs to learn complex multi-modal sensorimotor relations within the environment, and predict sensorimotor trajectories satisfying external goals such as obstacle avoidance tasks. CNMPs' power of learning a priori knowledge of data and predicting conditional distributions is not only limited to robotic learning. Conditional Neural Movement Primitives(CNMPs) were designed as a learning from demonstration framework for robotic movement learning and generation. It's built on top of Conditional Neural Processes(CNP) architecture which has a deep neural structure trained via gradient descent and defines conditional distributions over functions given a set of observations[2].

Predicting whole body motion trajectories requires novel ways to represent whole-body motion with respecting the coupling of body joints and dimensions, recognizing multiple trajectories for the same behavior, and encoding different behaviors in the same model. Recent state-of-the-art approaches on human motion prediction employed RNN based models which produced promising results[3], [4].

CNMPs can be trained with a single or a number of trajectories of 3D positions of body parts. It learns the underlying trajectory distribution of the motion and predicts whole body motion trajectories over target observation points. Our preliminary results show that our model can learn multiple behaviors simultaneously and successfully generate the motion trajectory of all body parts given a desired body configuration at desired time points.

II. METHOD

CNMPs framework is built on top of CNP neural model[2] which processes the training data to extract prior knowledge by sampling random observations and predicts conditional distributions over target queries. CNMPs learn underlying distribution of the provided multi-variate data, and allow conditioning this distribution on any target time points. Fig. 1 shows the training procedure of CNMPs on a hypothetical 1D trajectory scenario. At each iteration, a) a uniformly-random motion M_i is selected from motion set M. b) n random observation points and a random target query t_q is sampled uniformly from M_i. c) Each observation point goes through the parameter sharing encoder network E which produces latent space representations r_i, in this stage CNMPs allow concatenating external parameters γ. Note that external parameters are not used in this study. The representations obtained for different time points are merged into a single general representation r using a symmetric operator which is generally a mean operation. r is then concatenated with target query time t_q and fed into query network Q which produces distribution mean and variance, namely (μ_q, θ_q). d) CNMPs predict corresponding trajectory distribution of the queried time-step by inferences based on sampled observations. The network is trained via stochastic gradient descent with the negative conditional log probability function.

After training phase, CNMPs can be conditioned on single or multiple time-steps to produce trajectory distributions to satisfy the given conditions. Moreover, whole movement trajectories can be generated by querying CNMPs with all time-steps. Please refer to [1] for details.

CNMPs are exploited to model whole body motion. Trajectories of 3D positions of the body parts are used as inputs to the system. In order to generate motions, our system is conditioned with desired 3D body configuration. Motions were gathered from KIT Whole Body Human Motion database. The time is scaled between [0,1] to ensure time invariance and 3D locations of body parts were normalized to ensure space invariance. 14 markers that cover whole body are selected out of 50 markers. Fig 2 shows sequential 3D body configurations of turning left motion from training data, observed at initial, middle and final time-steps.

III. EXPERIMENTS

In the first analysis, we evaluated the performance of CNMPs in learning single behavior from three different but similar trajectories of walking forward. After our model learns these trajectories, our model is conditioned on observation points with the corresponding body configurations at start, middle and final time-steps. Constrained with these body configurations at different time-steps, the model successfully generated the body configurations for all time-steps. Fig. 3 shows the real and predicted positions of head, shoulder, and knee parts of the body. As shown, the complex trajectories of body parts were successfully learned and generated.

In the second analysis, we investigated whether our system can learn multiple behaviors simultaneously. For this, we trained the system with two different behaviors, namely forward walking with a left turn and forward walking with a right turn. Similar to the first experiment, after the system is trained, the model is conditioned with body configurations that are desired to be observed at specific (initial, middle and final) time-steps. Fig. 3 shows the real and predicted positions of head, shoulder, and knee parts of the body. As shown, the complex trajectories of body parts of different behaviors were successfully learned and generated. Results for both experiments with whole trajectory predictions can be reached here.

IV. CONCLUSIONS

In this paper, we proposed to extend and use CNMPs that were originally designed for robotic manipulation tasks in whole body human motion modelling and learning. Preliminary results show how remarkably well CNMPs can learn to encode complex trajectories of a single or multiple whole body behaviors from only a few motions, extract knowledge about underlying trajectory distributions and generate close to perfect full trajectories of body parts given desired body configurations at desired time points.

Although we only experimented with forward walking and walking with turns, we believe CNMPs can learn more varieties of whole body motions which will be studied in followup works. Future work will include using trajectories of a bigger set of body joints, learning trajectories that involve more complex motions and comparing our results with recent state-of-the-art-approaches [3], [5] .

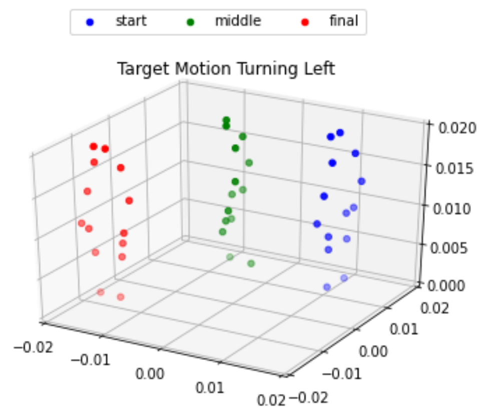
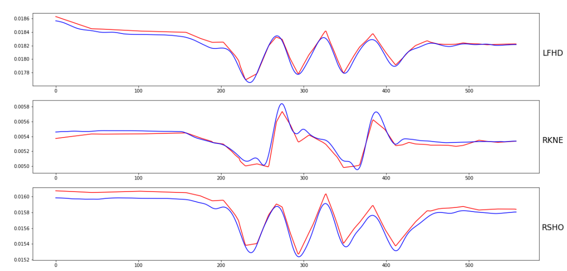
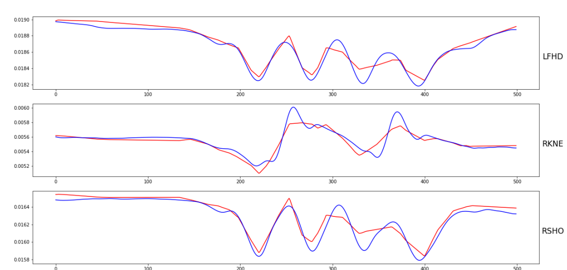


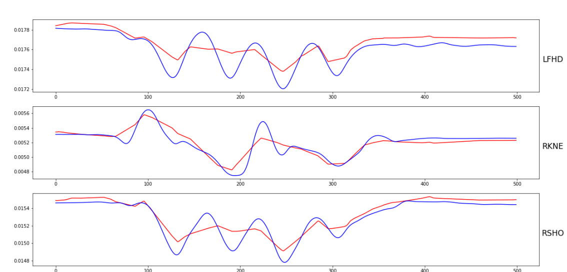
Fig. 2: Snapshots from a training motion. Different colours indicate sequential joint locations in 3D space.



(a) Walking straight



(b) Left Turn



(c) Right Turn

Fig. 3: Each row shows Z-coordinate values of right annotated joints taken from different(head, knee, shoulders from top to below) parts of body over timesteps. Blue lines are target trajectories and red lines are predictions of CNMPs.

ACKNOWLEDGMENT

This work was supported by the Scientific and Technological Research Council of Turkey (TUBITAK, 118E923).

REFERENCES

[1] Seker, M. Y., Imre, M., Piater, J., Ugur, E. "Conditional Neural Movement Primitives", RSS, 2019.

[2] Garnelo, M., D. Rosenbaum, C. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. Rezende and S. M. A. Eslami, "Conditional Neural Processes", ICML, 2018

[3] Tang, Yongyi and Ma, Lin and Liu, Wei and Zheng, Wei-Shi "Long-Term Human Motion Prediction by Modeling Motion Context and Enhancing Motion Dynamics." Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, 2018, doi:10.24963/ijcai.2018/130.

[4] Fragkiadaki, K., Levine, S., Felsen, P., Malik, J. "Recurrent Network Models for Human Dynamics." 2015 IEEE International Conference on Computer Vision (ICCV), 2015, doi:10.1109/iccv.2015.494.

[5] Kratzer, P., Toussaint, M., Mainprice, J. (2018). "Towards Combining Motion Optimization and Data Driven Dynamical Models for Human Motion Prediction". 2018 IEEE-RAS 18th International Conference on Humanoid Robots (Humanoids). https://doi.org/10.1109/humanoids.2018.8624991