# Multi-Style Human Motion Prediction and Generation via Meta-Learning

Lingfeng Sun, Masayoshi Tomizuka, and Wei Zhan

## I. INTRODUCTION

The purpose of this study is demonstrate a generation framework which can online generate motions of different styles. Data-driven methods [1]–[4] are widely used in previous works to model kinematic information of different behaviors when analyzing human motion. We focus on the 3D human motion generation task utilizing motion prediction models with meta-learning. We train a style-conditional generative model with an unlabeled stylistic motion dataset. Given historic motion, the model can predict future motion or generate new styled motion based on one-shot imitation. Fig. 1 shows an illustration of the overall framework.

While previous prediction works [5]–[11] focused more on different actions, stylistic variations in actions that represent the personality or mood are also crucial for motion generation. For example, the "angry" **style** in "crossing the road in heavy steps" may be more important than the "walking" **action** in some scenarios. One challenge in previous few-shot learning works on multi-behaviors is that datasets often require labels of the motions [12], but in general human motion is very diversified and hard to interpolate using existing motions. Besides, unlike prediction problems, generation tasks usually have no ground truth to compare with. For the former problem, we assume all unlabeled motion data from similar task space and learn the representation via self supervision. We utilize one-shot imitation [13] to enable fast adaptation during online use. For the latter one, we can naturally evaluate the online motion generation framework with consistent **style** using prediction benchmarks.

## II. METHODS

*1) Formulations:* We first introduce the **Markov Decision Process(MDP)** to model the generation process of the 3D motion sequence similar to [6]. At each time-step $t$, the state $s_t = \{x_i\}_{i=1}^{t\times m}$ is defined as all previous poses starting from the initial one. The action $a_t = \{v_{t\times m}, ..., v_{(t+1)\times m-1}\}$ is defined as the change of pose in the next length-$m$ window. The transition model is deterministic based on accumulated velocity. The reward and initial state distribution vary for different behaviors. This formulation satisfies Markov property and removes the constraint on past sequence length. Based on the MDP formulation, we extend the notion with a probabilistic behavior encoding variable $b \in \mathcal{B}$ represent different styles(e.g., angry, depressed, childlike).

Lingfeng Sun, Masayoshi Tomizuka, and Wei Zhan are with the Department of Mechanical Engineering, University of California, Berkeley. CA 94720, USA (email: lingfengsun, wzhan, tomizuka@berkeley.edu)

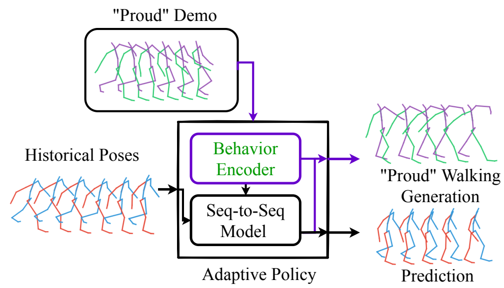


Fig. 1. Given the historical (initial) motion sequence, the model would output motion prediction based on input poses. If given a demonstration of *proud* walking, the model can encode the behavior and generate future motion of *proud* style based on the initial poses.

*2) Memory-based meta learning:* To learn the behavior-conditional policy $\pi : \mathcal{S} \times \mathcal{B} \rightarrow \mathcal{A}$ from heterogeneous multi-task demonstrations as the desired generative model, we need to meta-learn an inference model $q(b|\tau)$ and a policy $\pi(a|s, b)$, with $\tau$ representing a demonstration trajectory. When given a new demonstration $\hat{\tau}$, we can sample $\hat{b} \sim q(b|\hat{\tau})$ and use the policy to generate or predict. We assume no labelling on the motion data and no prior knowledge on the behavior distribution.

*3) Imitation learning with mutual-information regularization:* Consider a trajectory $\tau = \{s_{1:T}, a_{1:T}\}$ generated with the encoded behavior $b$, the trajectory distribution induced by the learned policy $\pi$ parameterized with $\theta$ is:

$$p_\theta(\tau|b) = \eta(s_1)\prod_{t=1}^{T}\pi_\theta(a_t|s_t, b)P(s_{t+1}|s_t, a_t) \qquad (1)$$

We define the behavioral cloning loss as:

$$L_{\mathrm{BC}} = \mathbb{E}_{\tau\sim D_E, b\sim q(b|\tau_E)}[d(\pi_\theta(\tau), \pi_E(\tau))] \qquad (2)$$

However, directly optimizing over the behavior cloning loss on expert dataset $D_E$ could lead to cases where the model simply ignore $b$, i.e, the policy just treats $b$ as an useless augmentation to the states and learn a global policy. We need to make sure that the behavior variable $b$ affect the trajectory generated by the policy. The mutual information between the behavior and the generated trajectory can serve as a useful regularization. By definition, the mutual information between $\tau$ and $b$ under joint distribution is:

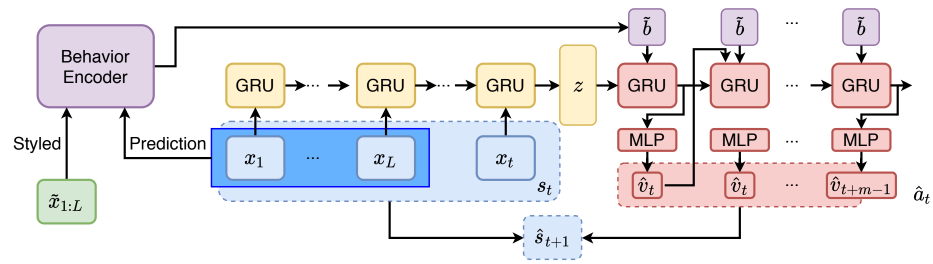


Fig. 2. Policy network structure. The behavior takes styled demonstration $\tilde{x}$ for styled generation task and first $L$ frames of state pose vector for prediction task. The decoder recurrently output the velocities of $m$ steps as action.

TABLE I

MOTION PREDICTION ERROR IN EULER ANGLE ERROR FOR WALKING, EATING, SMOKING, AND DISCUSSION IN THE HUMAN3.6M DATASET FOR HORIZION OF 80, 160, 320 (SHORT-TERM), 1000MS (LONG-TERM).

| Action | Walking | | | | Eating | | | | Smoking | | | | Discussion | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ms | 80 | 160 | 320 | 1000 | 80 | 160 | 320 | 1000 | 80 | 160 | 320 | 1000 | 80 | 160 | 320 | 1000 |
| Res. sup [9] | 0.28 | 0.49 | 0.72 | N/A | 0.23 | 0.39 | 0.62 | N/A | 0.33 | 0.61 | 1.05 | N/A | 0.31 | 0.68 | 1.01 | N/A |
| ConvSeq2Seq [14] | 0.33 | 0.54 | 0.68 | 0.92 | 0.22 | 0.36 | 0.58 | 1.24 | 0.26 | 0.49 | 0.96 | **1.62** | 0.32 | 0.67 | 0.94 | 1.86 |
| Ours | **0.23** | **0.34** | **0.55** | **0.71** | **0.17** | **0.32** | **0.55** | **1.15** | **0.25** | **0.46** | **0.88** | 1.65 | **0.30** | **0.58** | **0.83** | **1.81** |

$$
\begin{aligned}
I(\tau, b) &= D_{KL}(p(\tau, b) \| p(\tau)p(b)) \\
&= \mathbb{E}_{(\tau,b)\sim p(\tau,b)}[\log p(\tau, b) - \log(p(\tau)p(b))] \\
&= \mathbb{E}_{b\sim p(b),\tau\sim p_\theta(\tau|b)}[\log p_\theta(b|\tau) - \log p(b)] \\
&\approx \mathbb{E}_{\tau_E\sim D_E, b\sim q(b|\tau_E)}[\log q_\phi(b|\tau_E) - \log p(b)]
\end{aligned} \tag{3}
$$

The posterior distribution $p_\theta(b|\tau)$ is intractable and we use $q_\phi(b|\tau)$ as a variational approximation, we can conclude the mutual information loss as:

$$
L_{\text{info}} = -\mathbb{E}_{\tau_E\sim D_E, b\sim q(b|\tau_E)} \log q_\phi(b|\tau) \tag{4}
$$

Therefore the overall the meta-training objective is:

$$
\min_{\theta,\phi} L_{\text{BC}} + \lambda L_{\text{info}} \tag{5}
$$

where $\lambda$ is the weight hyper-parameter of loss terms.

*4) Policy model and algorithms:* As shown in Fig. 2, we use a behavior encoder $q$ to encode a behavior latent variable $b$, then the sequence-to-sequence policy encoder takes in the state $s_t = x_1, ..., x_{t\times m}$, the decoder takes in $b$ and outputs fixed $m$-length action $a_t = v_{t\times m}, ...v_{(t+1)\times m-1}$ which is the predicted pose velocities under this certain behavior.

During training, an extensive set of unlabelled motion sequences is given for training. For testing, by inferring a $\tilde{b} \sim q_\phi(b|\tilde{X})$, we can generate same-behavior motion with $\pi_\theta(a|s, \tilde{b})$.

## III. EXPERIMENTS

We first evaluate the prediction performance. We run experiments on the Human3.6M dataset [15] and compare it with previous benchmarks. The results are shown in Table I, data processing methods follows the previous benchmarks. The goal is to confirm that we reach acceptable performance in the prediction task.

Then we evaluate the performance of controlled generation by a stylistic motion dataset [16] containing the same actions and show the one-shot imitation performance on the same walking action. We use the same sequence of *angry walking* behavior as the initial state and use three sequences of *angry*, *proud*, and *depressed* walking as styled demonstrations. We can see the generated sequences with the same initial frames and future poses of different styles in Fig. 3. The model can perform styled generation after one-shot imitation for seen and unseen styles.

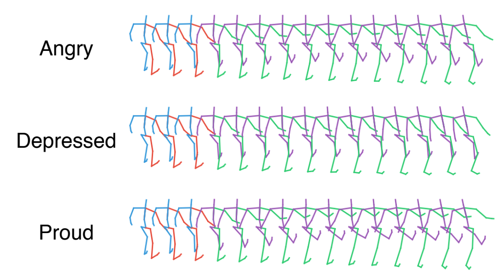


Fig. 3. Qualitative styled generation results. Style affects the future poses. (difference in legs and hands)

## IV. CONCLUSION

We proposed a new learning method for the multi-style motion generation framework. we present an approach to learn an adaptive motion policy from an unlabeled multi-style motion dataset. The model is trained with behavior cloning and regularized with mutual information loss. Experiments show that our proposed method can handle prediction and styled generation tasks via one-shot imitation well. The overall framework provided a high level structure for heterogeneous motion generation and can incorporate any motion prediction model or one-shot learning method.

## REFERENCES


[1] J. Bütepage, M. Black, D. Kragic, and H. Kjellström, "Deep representation learning for human motion prediction and classification," 2017.

[2] S. Levine, J. M. Wang, A. Haraux, Z. Popović, and V. Koltun, "Continuous character control with low-dimensional embeddings," *ACM Transactions on Graphics*, vol. 31, no. 4, p. 28, 2012.

[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*, vol. 30, no. 2, pp. 283–298, 2008.

[4] L. Kovar and M. Gleicher, "Automated extraction and parameterization of motions in large data sets," *ACM Trans. Graph.*, vol. 23, no. 3, p. 559–568, Aug. 2004. [Online]. Available: https://doi.org/10.1145/1015706.1015760

[5] K. Fragkiadaki, S. Levine, P. Felsen, and J. Malik, "Recurrent network models for human dynamics," 2015.

[6] B. Wang, E. Adeli, H.-k. Chiu, D.-A. Huang, and J. C. Niebles, "Imitation learning for human pose prediction," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019.

[7] P. Ghosh, J. Song, E. Aksan, and O. Hilliges, "Learning human motion models for long-term predictions," 2017.

[8] A. Jain, A. R. Zamir, S. Savarese, and A. Saxena, "Structural-rnn: Deep learning on spatio-temporal graphs," 2016.

[9] J. Martinez, M. J. Black, and J. Romero, "On human motion prediction using recurrent neural networks," 2017.

[10] W. Mao, M. Liu, and M. Salzmann, "History repeats itself: Human motion prediction via motion attention," 2020.

[11] E. Aksan, P. Cao, M. Kaufmann, and O. Hilliges, "A spatio-temporal transformer for 3d human motion prediction," 2020.

[12] L.-Y. Gui, Y. Wang, D. Ramanan, and J. M. F. Moura, "Few-shot human motion prediction via meta-learning," in *Proceedings of (ECCV) European Conference on Computer Vision*. Springer, September 2018.

[13] Y. Duan, M. Andrychowicz, B. C. Stadie, J. Ho, J. Schneider, I. Sutskever, P. Abbeel, and W. Zaremba, "One-shot imitation learning," 2017.

[14] C. Li, Z. Zhang, W. S. Lee, and G. H. Lee, "Convolutional sequence to sequence model for human dynamics," 2018.

[15] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu, "Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1325–1339, jul 2014.

[16] S. Xia, C. Wang, J. Chai, and J. Hodgins, "Realtime style transfer for unlabeled heterogeneous human motion," *ACM Trans. Graph.*, vol. 34, no. 4, Jul. 2015.