
Neural Embedding for Physical Manipulations

Lingzhi Zhang^{1*} Andong Cao^{2*} Rui Li¹ Jianbo Shi¹
¹University of Pennsylvania ²Yale University
{zlz, lirui613, jshi}@seas.upenn.edu, antonio.cao@yale.edu

Abstract

In common physical manipulations, action and state spaces can be vast and sometimes unknown, and observations are often relatively sparse. How do we learn the full topology of action and state spaces when given only few and sparse observations? Inspired by the properties of grid cells in mammalian brains, we build a generative model that enforces a normalized pairwise distance constraint between the latent space and action space to achieve data-efficient discovery of action space. We demonstrate that our model can effectively alleviate mode collapse issue and actively explore the unknown action space, and outperform other generative models, such as Generative Adversarial Networks and Variational Auto-Encoders.

1 Introduction

In real-world physical manipulations, common tasks like predicting ball collision or rope manipulations can involve vast or continuous action and state spaces. But often we only see relatively sparse observations. Can we learn a physical embedding such that we can safely interpolate and sample unseen and plausible actions? Common approaches are to use generative models to encode a high-dimension space into a low-dimensional parametric distribution. However, generative models such as Generative Adversarial Networks (GANs) [3] and Variational Auto-Encoders (VAEs) [4] suffer from mode collapse, which prevents fully exploring the unknown topological structure of the unseen action and state spaces in our application. Thus, our goal is to solve this problem by proposing a generative model inspired by grid cells [11] and normalized diversification [6].

Grid cells, the grid-like neural circuit in mammalian brains, is known to dynamically map the external environment as the animal navigates the world [11]. Remarkably, this encoding preserves metric distance relationships, such that objects close in the real-world are close in the brain’s intrinsic map [2]. Moreover, with a few observations and actions, the grid cells can rescale the mapping according to changes in the size and shape of the environment [1]. Such mental model allows quick adaptation to new surroundings, efficient localization and path-planning, and imagination of unseen events.

Inspired by the properties of grid cells, we propose a constraint on the latent space of a generative model that achieves data-efficient discovery of output spaces. Similar to the grid cell’s distance-preserving encoding of the world, our proposed model preserves the normalized pairwise distance of samples between the parametric low-dimensional latent space and high dimensional output space. Intuitively, our approach encourages the neural network to actively explore the output space by enforcing the latent space to decode as diverse as possible. In addition, our model is coupled with adversarial learning that enforces the generated samples to be plausible.

Our model has several practical applications: first, during many robotic operations, the constraints on the action space, such as safety or geometric limits specific to the task, are dynamic and largely unknown, whereas a wide range of possible maneuvers remains uncharted. Our proposed method can help the robot efficiently explore the action space and predict the future states. Second, the latent

*Indicates equal contribution.

embedding in our model supports safe interpolation in the output space, thus enabling robots to generate reliable and plausible action proposals. Third, our model can accurately predict future states given current states and sampled actions with a learned inverse dynamic model. Overall, our method can help a robot build a mental physical model of the task at hand.

2 Methodology

In physical manipulation scenarios, there exists many stochastic actions when given a state. For example, there are many potential locations, directions, and forces to push a given object. Thus, we learn a multimodal generative model to predict diverse actions given a current state. On the hand, when given a current state and an action, there is one and only one future state. Thus, we use a forward dynamic model to predict a deterministic future state by inputting a current state and an action.

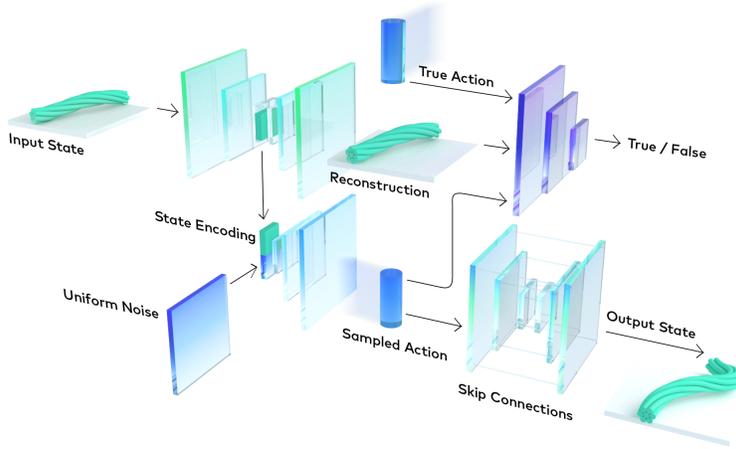


Figure 1: This is an overview of our model architecture.

2.1 A Generative Model to Unfold Action Space

The overall model architecture is shown in Fig.(1). The model first takes an input state and encodes into a feature embedding guided by an auto-encoder. Then the input state feature is concatenated with a sampled random variable and are decoded into a possible action proposal. During training, a normalized pairwise distance constraint is applied to encourage active exploration of action space and a conditional discriminator is used to check all the sampled actions are plausible. Finally, a forward dynamic model is trained to map the input state and predicted action into an output future state.

2.1.1 Active Exploration Via Normalized Diversification

To encourage active exploration of action space, our generative model preserves the normalized pairwise distance of different generated samples in between the latent space and action space. The distance metric $d_z(\cdot, \cdot)$ between any two samples is a Euclidean distance. We denote z as latent variables, a as actions, and i, j as sample indices.

$$d_z(z_i, z_j) = \|z_i - z_j\| \quad , \quad d_{a_t}(a_t^i, a_t^j) = \|a_t^i - a_t^j\| \quad (1)$$

Furthermore, we define the normalized pairwise distance matrices $D_{ij}^z, D_{ij}^{a_t} \in \mathbb{R}^{N \times N}$ as follows,

$$D_{ij}^z = \frac{d_z(z_i, z_j)}{\sum_j d_z(z_i, z_j)} \quad , \quad D_{ij}^{a_t} = \frac{d_{a_t}(a_{t_i}, a_{t_j})}{\sum_j d_{a_t}(a_{t_i}, a_{t_j})} \quad (2)$$

During training, we treat the normalizer in (2) as a constant when back-propagating the gradient to the generator network. This ensures that we optimize the absolute pairwise distance for a sample,

rather than adjusting normalizer to satisfy the loss constraint. The normalized diversity loss function is defined as follows,

$$\mathcal{L}_{div}(s_t, a_t, z) = \frac{1}{N^2 - N} \sum_{i=1}^N \sum_{i \neq j}^N \max(0, \alpha D_{ij}^z - D_{ij}^{a_t}) \quad (3)$$

where α is a hyperparameter. The diagonal elements of the distance matrix are all zeros and thus not included.

Unlike most generative models, our generative model parameterizes the latent space as a uniform distribution $U(0, 1)$ instead of Gaussian distribution because of two reasons. First, the uniform distribution is bounded so that the sampled latent variables will never be too far away from each other, which might induce extremely large pairwise distance and thus might lead to exploding gradients when optimizing with respect to the loss. Second, while Gaussian distribution puts a Gaussian bump as a strong prior on the data distribution, uniform distribution has the relative flexibility to map to diverse modes by cutting itself into many tiles and mapping each tile into a different mode in the output space.

2.1.2 Safe Mapping Via Adversarial Training

While the normalized diversity loss encourages the model to actively explore in the action space, the adversarial checks all the predicted actions are plausible. Our adversarial training framework is based on conditional GAN [9]. The generator sampled a possible action based on a given input state and a random latent variable. The discriminator takes both real and generated actions as inputs and predicts whether these actions are real or fake conditioned on the corresponding input state. During implementation, the discriminator takes the concatenation of the action and input state feature embedding as inputs, and hinge loss [5] [12] is used to alternatively update the generator and discriminator,

$$\mathcal{L}_D(s_t, a_t, z) = \mathbb{E}_{a_t \sim q_{data}(a_t)}[\min(0, 1 - D(a_t|s_t))] + \mathbb{E}_{z \sim p(z)}[\min(0, 1 + D(G(s_t, z)|s_t))] \quad (4)$$

$$\mathcal{L}_G(s_t, z) = -\mathbb{E}_{z \sim p(z)}[D(G(s_t, z)|s_t)] \quad , \quad \mathcal{L}_{adv} = \mathcal{L}_D + \mathcal{L}_G \quad (5)$$

To stabilize training, spectral normalization [10] is applied to scale down the weight matrices in the discriminator by their largest singular values, which effectively restricts the Lipschitz constant of the network. After training converges to an equilibrium, the generator is able to sample diverse and plausible actions given a current state.

2.2 A Forward Dynamic Model to Predict the Future State

In the second part of our method, we train a forward dynamic model to predict future states given a pair of current state and action. We consider predicting a future state as a deterministic process and thus directly optimize the L2 loss between the predicted and ground truth future states.

$$\mathcal{L}_{recon}(s_{t+1}^*, s_{t+1}) = \|s_{t+1}^* - s_{t+1}\| \quad (6)$$

where s_{t+1}^* and s_{t+1} indicate predicted and the ground truth future states. The states could be high-dimensional images or some low-dimensional representation depending on different applications.

3 Experimental Results

To evaluate the effectiveness of our model, we conduct experiments using both a synthetic dataset and two simulated physical manipulation applications - hitting a capsule and pushing a rope.

In the synthetic experiment, we model a simple physical manipulation: pushing a ball away from the center of a table. In Fig.(2), blue and green dots indicate the force (action) and the corresponding

| Model | Synthetic | | Rope | | Roller | |
|--------|------------------|------------------|-------------------|------------------|------------------|------------------|
| | FD ↓ | JSD ↓ | FD ↓ | JSD ↓ | FD ↓ | JSD ↓ |
| VAE[4] | 21.55±2.20 | 0.05±0.00 | 12.36±1.04 | 0.67±0.01 | 10.14±0.00 | 0.66±0.01 |
| GAN[3] | 26.83±1.94 | 0.16±0.03 | 16.48±10.45 | 0.66±0.01 | 13.05±6.79 | 0.66±0.01 |
| Ours | 3.48±0.75 | 0.02±0.00 | 11.08±4.46 | 0.54±0.10 | 9.66±4.90 | 0.50±0.08 |

Table 1: A table shows quantitative results between different generative models and ours. **FD** indicates Fréchet Distance [7], and **JSD** indicates Jensen-Shannon Divergence [8].

future location (state) respectively. We assume there exists unknown geometric constraint in the action sampling space and non-linear transformation from the action space to the state space given fixed input state. All models are trained with only 600 sparse actions and are used to sample 10,000 actions when converged. The predicted states are computed by feeding predicted actions into the trained forward dynamic model. As shown in Fig.(2), VAE [4] seems to put a Gaussian structure in the predicted action space and thus does not capture the correct topology of the true action space. GAN [3] is able to capture the rough topological structure but collapses to some certain modes. Finally, compared to VAE [4] and GAN [3], our model can clearly better approximate the true action and state spaces by sampling in the embedding space.

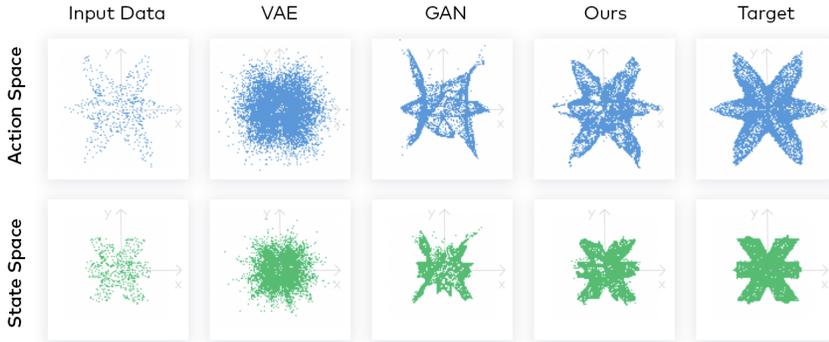


Figure 2: Comparison of generative models' ability to discover the unknown action and state spaces.

In the simulation environment, we conduct experiments on two scenarios: hitting a capsule and pushing a rope. The actions are the forces acting on the object and states are the images. In Fig.(3), we demonstrate that our model can sample diverse and plausible actions given an input state, where the arrow vector indicates the force magnitude and direction. Furthermore, we also show that our model can generate not only diverse but plausible results, which outperforms the other two generative model frameworks, as shown in Table.(1).



Figure 3: Qualitative results of diverse action sampling on rope and roller manipulations.

4 Conclusion and Future Work

In this work, we propose a generative model that can actively unfold the unknown action and state spaces with only sparse observations. With the proposed distance preserving constraint and adversarial training, our learned physical embedding can better generalize to unseen and diverse action space in comparison with commonly used generative models in both synthetic experiments and common physical manipulation scenarios. In the future work, we envision to integrate our method into a learning system that can learn from a few observations and infer the invisible object physical properties and Newtonian physics model between objects.

Acknowledgement

We gratefully appreciate support through Honda Research Institute Curious Minded Machine program.

References

- [1] C. Barry, R. Hayman, N. Burgess, and K. J. Jeffery. Experience-dependent rescaling of entorhinal grids. *Nature neuroscience*, 10(6):682, 2007.
- [2] R. A. Epstein, E. Z. Patai, J. B. Julian, and H. J. Spiers. The cognitive map in humans: spatial navigation and beyond. *Nature neuroscience*, 20(11):1504, 2017.
- [3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.
- [4] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.
- [5] J. H. Lim and J. C. Ye. Geometric gan. *arXiv preprint arXiv:1705.02894*, 2017.
- [6] S. Liu, X. Zhang, J. Wangni, and J. Shi. Normalized diversification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 10306–10315, 2019.
- [7] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet. Are gans created equal? a large-scale study. In *Advances in neural information processing systems*, pages 700–709, 2018.
- [8] C. D. Manning, C. D. Manning, and H. Schütze. *Foundations of statistical natural language processing*. MIT press, 1999.
- [9] M. Mirza and S. Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.
- [10] T. Miyato, T. Kataoka, M. Koyama, and Y. Yoshida. Spectral normalization for generative adversarial networks. *arXiv preprint arXiv:1802.05957*, 2018.
- [11] M.-B. Moser, D. C. Rowland, and E. I. Moser. Place cells, grid cells, and memory. *Cold Spring Harbor perspectives in biology*, 7(2):a021808, 2015.
- [12] D. Tran, R. Ranganath, and D. M. Blei. Deep and hierarchical implicit models. *arXiv preprint arXiv:1702.08896*, 7, 2017.