Few-Shot Semi-supervised Learning From Demonstration for Generalisation of Force-Based Motor Skills Across Objects Properties

Marina Y. Aoyama[1], João Moura [1,2], Namiko Saito[1,2] and Sethu Vijayakumar[1,2]

***Abstract*— In many manipulation tasks, force feedback plays an essential role in adapting the motion to the physical properties of the manipulated object. The existing Learning from Demonstration (LfD) approaches require a large number of demonstrations to generalise the learned motor skills to manipulate objects with unknown properties, however, collecting demonstrations is expensive in time and human effort. Therefore, we aim to learn to adapt motion according to object properties from a small number of demonstrations by utilising a large amount of unsupervised data, which is less informative about the task but less expensive to collect. We propose to decouple the haptic representation model from the motion generation model and enable pre-training of the haptic representation model through self-supervised learning on unsupervised haptic data. We validated on the wiping task using wiping tools with different stiffness and surface friction. Our results suggest that pre-training of the haptic model leads to force profiles that are closer to those demonstrated during adaptive wiping using sponges with unseen stiffness and friction. The sim2real transfer of the haptic representation model pre-trained on simulation data in learning downstream tasks on a real robot was also evaluated.**

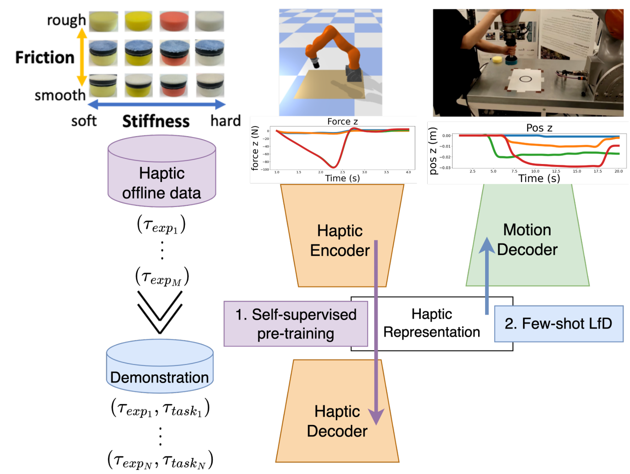


Fig. 1: Proposed few-shot semi-supervised LfD framework. First, the haptic encoder-decoder model is pre-trained using unsupervised data collected in simulation. Then, the motion decoder is trained to map from the haptic representation obtained using the pre-trained haptic encoder to the desired adaptive motion for the target task using demonstration data.

I. INTRODUCTION

One of the important capabilities of fully automated robots is to adapt the motion to the physical properties of the manipulated object using force feedback. In this work, we use a wiping task as an instance of force-based manipulation task because the wiping motion and the force that needs to be applied depend on the stiffness of the wiping tool, the surface friction, etc.

LfD is an intuitive and effective way of transferring such motor skills from human to robot. However, it is often challenging to generalise learnt motor skills to manipulate objects with different physical properties, as collecting demonstrations of various objects is time-consuming and involves extensive human effort.

One promising approach to improve the generalisability and robustness when supervised data is limited is to pre-train model using unsupervised data. However, pre-training of the model has mostly been applied to learning visual and haptic representations. They cannot be directly applied to haptic representation learning through force interactions, as obtaining haptic representations requires a series of force observations rather than a snapshot of the force sensing at each time step. We therefore propose pre-training a haptic representation model using sequences of force observations collected by performing exploratory actions in an unsupervised manner.

The goal of this research is to learn to manipulate objects with different physical properties from a small number of demonstrations. To this end, we aim to learn to identify the properties of the manipulated object using simulation and to learn to generate the motion trajectory for the target task on a real robot accordingly through few-shot LfD.

II. RELATED WORKS

A. *Motor skill transfer across objects*

Several studies proposed the approaches to identify object properties by probing prior to the task and to adapt the motion for the target task accordingly using manually defined rules [1], [2]. Alternatively, it has also been shown that liquids with different properties can be successfully poured by obtaining simulated physical property parameters such as viscosity and surface tension through matching visual observations on the simulator and on the real robot, and then replaying the target task motion optimised for the object in the simulation on the real robot [3]. These studies have shown that even when performing the same task, adapting the motion according to the properties of the object is essential to improve the performance of the task. However, it is unclear how object properties need to be represented in order to

[1]Authors are with the School of Informatics, The University of Edinburgh, Edinburgh, U.K.
[2]Authors are with The Alan Turing Institute, London, U.K.

learn to adapt motions accordingly to the identified object properties, rather than using manually defined rules.

There are some studies on LfD to adapt the target task motion based on the characteristics of the object as identified through interaction with the object [4], [5], [6]. These studies suggest that motor skills can be transferred across objects with different physical properties by identifying the properties of the object and adapting the motion accordingly. However, these approaches require more demonstration data to increase generalisability, as models to identify object properties and models to generate motion are coupled, and can only be trained together. In contrast, our proposed approach to utilise unsupervised data to improve the generalisability of the task across different objects by decoupling the object property identification model trained by self-supervised learning and the motion generation model learnt from demonstrations.

B. *Pre-training haptic representation model*

While the amount of demonstration data tends to be limited as it is expensive to collect, a larger amount of unsupervised or unlabelled data tends to be available as it can be collected without human intervention and is therefore less expensive to collect. Although unsupervised data is less informative about the target task than demonstration data, prior work suggests that self-supervised learning on a large amount of unsupervised data allows to obtain perceptual representation that is useful to improve the generalisability and robustness of the downstream task [7], [8]. These studies learn visual and tactile representations of a snapshot of sensor information at each time step. In contrast, our proposed work considers acquiring haptic representations based on a series of force feedback. This approach allows the identification of object properties rather than only the current interaction information. Our proposed self-supervised representation learning approach is similar to the one employed to identify object properties using a sequence of video frames [9].

III. PROBLEM FORMULATION

The task has two phases: the exploration phase, where the properties of the object are identified by performing pre-defined exploratory actions, and the motion generation phase, where the robot adaptively performs the target task according to the properties of the object identified in the exploration phase.

Two types of data sets, unsupervised haptic data and demonstration data, are available for learning. Unsupervised haptic data is collected autonomously by repeatedly performing pre-defined exploratory actions on objects with different properties either in simulation or on a real robot and recording the force feedback. Demonstration data consists of pairs of force trajectories acquired by performing pre-defined exploratory actions and the desired motion trajectories for the object to perform the target task produced by an expert demonstrator. Unsupervised haptic data is less informative about the target task but cheaper to collect, whereas demonstration data is highly informative about the target task but more expensive to collect due to the time and human effort required. Therefore, the number of available unsupervised data M is much larger than the number of available demonstration data N. The key is to utilise a large amount of unsupervised data, which can be collected at low cost, to enable generalisation of the task across unseen objects with different properties, even when the number of demonstrations is limited.

We do not assume that the robot knows the properties of the real object, both at training and testing time. The objective is to learn a model that generates a motion trajectory for the target task, similar to a human demonstration, which needs to be adapted according to the properties of the object. Every time before executing the target task, the robot performs pre-defined exploratory actions on the manipulated object to identify its properties. It assumes that the environment and the properties of the object do not change during the task.

IV. METHOD

A. *Semi-supervised LfD architecture*

The proposed semi-supervsied LfD framework consists of two steps: pre-training of haptic representation encoder using a large amount of unsupervised haptic data and few-shot LfD for motion generation decoder using demonstration data.

B. *Pre-trained Haptic Encoder*

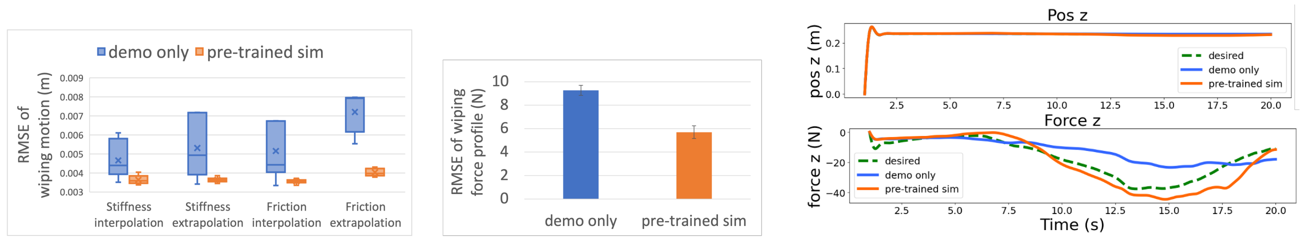
Variational Auto-Encoder (VAE) [10] has been extensively applied to the self-supervised learning of perceptual representation models. First, VAE is adopted to train haptic representation encoder-decoder models on unsupervised haptic data by reconstructing the force trajectories τ_{exp} obtained through pre-defined exploratory actions, thus by minimising a loss function

$$\hat{\theta}, \hat{\phi} = \arg\min_{\theta,\phi} L\left(\tau_{exp}\right) = \beta D_{KL}\left(q_{\phi}\left(z \mid \tau_{exp}\right) \parallel p_{\phi}\left(z\right)\right) + E_{MSE}\left(\bar{\tau}_{exp}, \tau^{*}_{exp}\right) \tag{1}$$

consisting of two loss terms, Kullback–Leibler (KL) divergence between the approximate posterior and prior distributions of the latent variables z and reconstruction loss. ϕ and θ are the the encoder and decoder parameters respectively, and β is the regularisation coefficient. Once the haptic representation encoder has been trained, the weights of the encoder are frozen to be used in the next step.

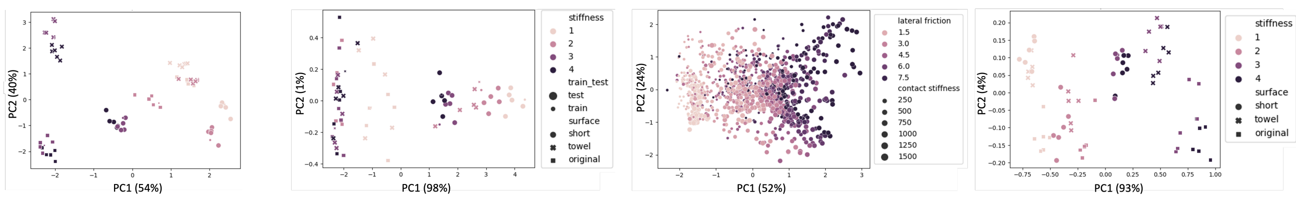
C. *Motion Decoder*

Next, the motion decoder that generates the desired motion for the target task τ_{task} according to the properties of the object is trained on the demonstration data. From the demonstration data, the force trajectory obtained in the exploration phase τ_{exp} is first passed to the pre-trained haptic encoder to obtain the corresponding haptic representation of the object. Next, a mapping from the haptic representation embedding to the desired motion trajectory for the target task is trained by minimising the loss function



(a) RMSE of generated motion (b) RMSE of force profile (c) Position and force trajectory during wiping

Fig. 2: Root Mean Squared Error (RMSE) and plots of motion trajectories and force profiles during wiping when motion is generated by the LfD model, with and without pre-training of the haptic encoder.



(a) Baseline: demo only (Stiffness Extrapolation) (b) Baseline: demo only (Friction Extrapolation) (c) Proposed: pre-trained trained and tested in sim (d) Proposed: pre-trained trained in sim & tested on real

Fig. 3: Principal Component Analysis (PCA) of latent space when the haptic encoder is pre-trained with simulated unsupervised data and when the encoder-decoder model is trained only using demonstration data.

$$\hat{\theta}_{LfD} = \arg\min_{\theta_{LfD}} L\left(\tau_{exp}, \tau_{task}\right) = E_{MSE}\left(\hat{\tau}_{task}, \tau^{*}_{task}\right). \tag{2}$$

where θ_{LfD} is the weights of the motion decoder.

D. *Task Execution*

At test time, the robot first performs pre-defined exploratory actions on the object to be manipulated. Using the haptic encoder and motion decoder trained in the pre-training and LfD steps respectively, the force trajectory obtained from the exploration phase is compressed to represent the properties of the object and then the corresponding desired motion trajectory for the target task is generated. The generated motion trajectory for the target task is sent to the robot, to perform the task by playing back the motion trajectory without using the current sensor feedback (*i.e.,* in an open loop).

V. EXPERIMENTAL SETUP

A. *Wiping task*

The task is to wipe a table placed in front of the robot with a sponge attached to the end-effector. The sponges with different properties (e.g., stiffness and surface friction) are prepared to evaluate the ability of the model to adapt wiping motion according the sponge properties unseen at training time.

B. *Data collection*

In the exploratory phase, the robot sequentially performs two pre-defined exploratory actions: pressing and lateral motion. Each exploratory motion lasts 2 seconds and the force data in x, y and z directions are recorded at 100Hz. The unsupervised data was collected using a high fedality physics simulator, PyBullet by performing the exploratory actions on 1000 objects by varying the simulated contact stiffness, lateral friction and spinning friction parameters. Note that this unsupervised data can be collected either in simulation on a real robot by performing exploratory actions on variety of objects.

For the robot experiment, we use a 7 DoF robotic arm, KUKA iiwa, in position control. To sense the force from the object, a six-axis force-torque sensor was mounted on the wrist. We attached the sponge to the end-effector as shown in Fig. 1. First, the exploration data is collected by performing the pre-defined exploratory actions on the sponge on the robot. Subsequently, demonstration was provided kinaesthetically in gravity compensation mode by following the circular wiping motion pattern. The task was to wipe with as much force as possible in normal direction. Each wiping motion lasts for 20 seconds and the end-effector position in x, y and z is recorded at 100Hz. Demonstration data consists of a list of pairs of force trajectory from the exploration the corresponding motion trajectory from the demonstration for 12 objects (4 stiffness levels * 3 friction levels * 1 trial).

C. *Model training*

First, haptic representation model is pre-trained as described in Section IV-B using simulated unsupervised data. Then, the motion generation model is trained as described in Section IV-C using demonstration data collected on a real robot. A sample of demonstration data was used to test four cases, stiffness interpolation, stiffness extrapolation and friction interpolation, friction extrapolation cases.

D. Evaluation

We evaluated the wiping performance using 1) RMSE of the wiping motion trajectory, and 2) RMSE of the wiping force trajectory demonstrated by human and executed by the robot. We compared the proposed semi-supervised LfD approach which pre-trains the haptic encoder with simulated unsupervised data and the baseline approach which learns the haptic encoder and motion decoder together only using demonstration data.

VI. RESULTS AND DISCUSSION

A. Analysis of motion generation

First, we generated the adaptive wiping motion for objects with unseen properties using the model trained with and without pre-training of the haptic encoder. The demonstrated and generated wiping motions were played back on the robot. We compare the RMSE of force profile during wiping to evaluate the effectiveness of pre-trained haptic encoder in generating motion closer to the demonstrated one and applying force closer to the demonstrator.

We see that in all four cases, stiffness interpolation, stiffness extrapolation, friction interpolation and friction extrapolation, there are statistically significant improvement in reducing RMSE of the motion trajectory and force profile during adaptive wiping as shown in Fig. 2a and Fig. 2b respectively. The plot of force profile during adaptive wiping is shown in Fig. 2c. The results have shown that pre-training of haptic representation model leads to learning of wiping motion closer to the demonstrated one without increasing a number of demonstrations.

B. Analysis of latent space of haptic encoder

Next, we apply the PCA to the latent space of the haptic encoder to evaluate the haptic representation model when trained together with the motion decoder using demonstration data only (*i.e.,* baseline) and when the haptic encoder is pre-trained using unsupervised haptic data collected in simulation (*i.e.,* proposed).

Fig. 3a and Fig. 3b show the examples of PCA applied to the latent space of the haptic encoder model when trained only using demonstration data together with the motion decoder model for the stiffness and surface friction extrapolation cases respectively. In Fig. 3a, the objects with different stiffness and friction are well separated and clustered, however, it performed poorly on generating motions for unseen objects as it does not learn underlying distribution of object properties unlike in the pre-trained model. In Fig. 3b, latent space was not well separated as 8 data (4 stiffness levels $\times$ 2 friction levels) is not enough to learn to distinguish different physical properties.

On the other hand, the latent space of the haptic encoder pre-trained and evaluated on the simulated force trajectories captures the distribution of the stiffness and friction as shown in Fig. 3c. When the haptic encoder is trained purely on the simulated data and evaluated on the real force trajectories, the haptic encoder captures both stiffness and friction of the real objects in the first principal component as shown in Fig. 3d. We observe that the friction of the object was not as clearly captured as the stiffness. One possible reasons for this is because of a larger sim2real gap of the friction model resulted in the mismatch between the simulator and real world observations. The analysis on the latent space suggested that haptic representation model pre-trained with unsupervised data collected in simulation allows to obtain haptic representation of real objects which successfully improved the performance of the downstream task on real robot.

VII. CONCLUSIONS

In this paper, we proposed decoupling the haptic representation model from the motor skill LfD model to utilise a large amount of unsupervised data. The experiment has shown that pre-training the haptic model with unsupervised simulation data enables learning of generalisable motor skills on a real robot from a small number of demonstrations.

One of the limitations of this work is that we did not examine how the sim2real gap can be reduced to obtain a better haptic representation of the real object while maximising the information available from the simulation. Also, while the encoder and decoder structure to generate the desired full motion trajectory was chosen in this study, it remains open to consider how pre-trained haptic representations can be effectively used when adapting the behaviour, for example, using the current sensor feedback.

REFERENCES

[1] A. Straižys, M. Burke, and S. Ramamoorthy, "Surfing on an uncertain edge: Precision cutting of soft tissue using torque-based medium classification," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 4623–4629.

[2] M. C. Gemici and A. Saxena, "Learning haptic representation for manipulating deformable food objects," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2014, pp. 638–645.

[3] T. Lopez-Guevara, R. Pucci, N. K. Taylor, M. U. Gutmann, S. Ramamoorthy, and K. Subr, "Stir to pour: Efficient calibration of liquid properties for pouring actions," *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 5351–5357, 2020.

[4] P. Sundaresan, S. Belkhale, and D. Sadigh, "Learning visuo-haptic skewering strategies for robot-assisted feeding," in *6th Annual Conference on Robot Learning*, 2022.

[5] N. Saito, T. Ogata, S. Funabashi, H. Mori, and S. Sugano, "How to select and use tools? : Active perception of target objects using multimodal deep learning," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2517–2524, 2021.

[6] K. Kawaharazuka, N. Kanazawa, K. Okada, and M. Inaba, "Learning-based wiping behavior of low-rigidity robots considering various surface materials and task definitions," in *2022 IEEE-RAS 21st International Conference on Humanoid Robots (Humanoids)*. IEEE, 2022, pp. 919–924.

[7] J. Pari, N. Muhammad, S. P. Arunachalam, and L. Pinto, "The surprising effectiveness of representation learning for visual imitation," *arXiv preprint arXiv:2112.01511*, 2021.

[8] I. Guzey, B. Evans, S. Chintala, and L. Pinto, "Dexterity from touch: Self-supervised pre-training of tactile representations with robotic play," 2023.

[9] M. Asenov, M. Burke, D. Angelov, T. Davchev, K. Subr, and S. Ramamoorthy, "Vid2param: Modeling of dynamics parameters from video," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 414–421, 2019.

[10] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.