Generating Goal-directed Visuomotor Plans with Supervised Learning using a Predictive Coding Deep Visuomotor Recurrent Neural Network

Takazumi Matsumoto(P)[1], Minkyu Choi[1], Minju Jung[2] and Jun Tani[1]

[1]Cognitive Neurorobotics Research Unit, Okinawa Institute of Science and Technology
[2]School of Electrical Engineering, Korea Advanced Institute of Science and Technology
E-mail: tani1216jp@gmail.com

Abstract— The ability to plan and visualize object manipulation in advance is vital for both humans and robots to smoothly reach a desired goal state. In this work, we demonstrate how our predictive coding based deep visuomotor recurrent neural network (P-DVMRNN) can generate plans for a robot to manipulate objects based on a visual goal. A Tokyo Robotics Torbo Arm robot and a basic USB camera were used to record visuo-proprioceptive sequences of object manipulation. Although limitations in resolution resulted in lower success rates when plans were executed with the robot, our model is able to generate long predictions from novel start and goal states based on the learned patterns.

Keywords— Predictive coding, Recurrent neural networks, Visuomotor planning

1 Introduction

The use of neural networks in robotics has become popular in recent years [1], as unlike the traditional method of programming complex models by hand, a neural network can self-determine optimal model parameters. This work employs a hierarchical recurrent neural network (RNN) structure, utilizing Long-Short Term Memory (LSTM) and Convolutional LSTM (ConvLSTM) neural networks [2]. The RNNs can then generate output in a *closed loop* manner, without any input by using the previous predicted output. In order to mimic biological visuomotor coupling we employ a dual hierarchical visuomotor structure which learns visuo-proprioceptual sequences in an end-to-end manner [3] as shown in Fig. 1.

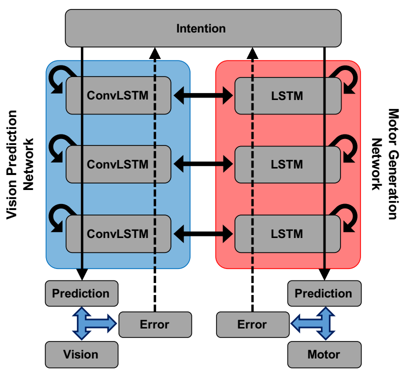

Figure 1: Overall network architecture of P-DVMRNN

The *predictive coding* model [4] used in this work takes *intention* at the top level, that is, instead of attempting to learn a mapping from all possible motor sequences to sensory sequences, the model embeds information learned from the high dimensional visuo-proprioceptive space to a low dimensional intention space. As shown in Fig. 2, predictive coding can produce visuo-proprioceptive sequences from a single intention state. Not only can learned sequences be regenerated given the original intention state, but novel patterns can be generated by altering the intention state [5].

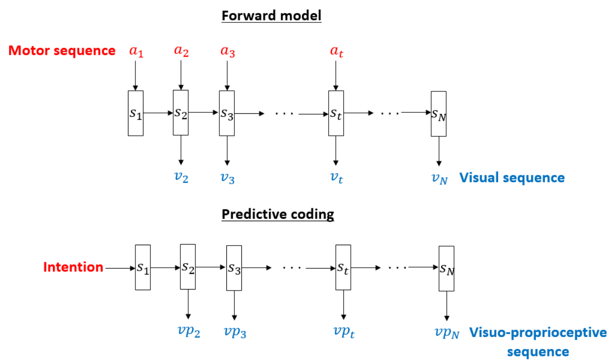

Figure 2: Comparison of the forward and predictive coding models. The predictive coding model is trained to generate both visual and motor sequences

2 Method

During training, our model learns the intention states for each of the training sequences, represented by initial states (IS) of the internal LSTM units as well as the connectivity weights, using back-propagation through time (BPTT) [6] in minimizing prediction error.

When a new input is given, its IS values can be inferred using an error regression (ER) scheme, which minimizes prediction error by changing IS but does not change the network connectivity weights. If the network has sufficient generalization from its training sequences, it is expected that given a novel input an IS value that is between other similar states can be inferred.

For a novel input, as we have no information of the actual IS, it must be searched with the IS initially set to a random value. A prediction is made using the randomly set IS, and given the output using a random IS value is unlikely to match the target sequence, a prediction error is produced. As shown in Eq. 1, this error will be minimized to optimize an IS during ER.

$$\begin{aligned}E_{pred} =&||V_{out,1} - V_{target,1}||^2 + ||V_{out,\hat{T}} - V_{target,T}||^2 \\ &+ \sum_{j=1}^{J} KL(M^j_{out,1}||M^j_{target,1}) \\ &+ \sum_{j=1}^{J} KL(M^j_{out,\hat{T}}||M^j_{target,T})\end{aligned} \tag{1}$$

$V_{out,t}, V_{target,t}, M^j_{out,t}, M^j_{target,t}$ are the visual predicted output at step t, the visual target at step t, the j^{th} joint angle predicted output at step t and the j^{th} joint angle target at timestep t respectively. $\hat{T}$ is the timestep at which which the goal state is reached in the plan. KL refers to Kullback–Leibler divergence.

3 Experimental Results

For this experiment, the experimenter first tutored the robot by completing the task for a set of randomly generated positions. After the joint angles were recorded, the robot recreated the trajectories and captured video of the motions. As this data was generated by a human, it will naturally have noise, gaps and fluctuations. If our model is able to generalize the training trajectories, it should still be able to generate accurate predictions to reach the goal state.

The object used was a plastic cylinder with a diameter of $5cm$ and a height of $10cm$. The target is a circle with a diameter of $12cm$, and both the object and target were randomly placed. The task for the robot was to 1) grasp the object and 2) place it in the target circle. 100 training sequences and 50 test sequences were used. During testing, the task was deemed successfully completed if the robot grasped the object and placed it upright within the target circle.

As shown in Fig. 3, our model is able to generate both visual and proprioceptual predictions for the entire sequence of picking up the object and placing it in the goal circle, although the visual prediction shows some noise. The results are summarized in Tab. 1.

Table 1: Success rate for object manipulation task, with varying degrees of permitted positioning error

	Success rate	With error in grasping 1 pixel	2 pixels	3 pixels
Closed loop Prediction	48%	48%	64%	71%
Open loop Prediction	74%	74%	88%	93%

The unsuccessful cases were often due to deviations in the positioning of the end effector when grasping the object. Due to the low resolution of the visual input and the geometry of the object in relation to the end effector, any error greater than $2cm$ (less than 2 pixels) typically resulted in a failure to grasp the object. In the cases where the robot successfully grasped the object, the object was placed within the goal circle 94% of the time.

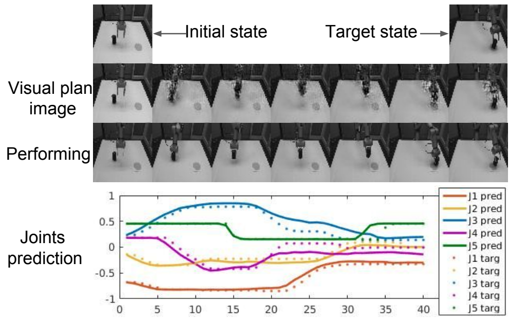

Figure 3: Example of plan generation. The top row shows images of the initial and target visual states, the second row shows the generated image sequence (sampled), and the third row shows the image sequence captured when the robot followed the plan, for comparison. The bottom row shows generated joint angles (solid lines) and ground truth joint angles from the experimenter (dotted lines)

4 Conclusion

In this work, we proposed a new architecture for goal-directed action planning using a predictive coding type deep dynamic neural network. We demonstrated that our model is able to generate visuomotor plans for a novel initial and goal position by inferring a new initial state. These visuomotor plans can be either long (closed loop) or one-step (open loop) predictions, with the success rate of the generated plans being limited by the low image resolution and high precision required to grasp the object. In future work we plan to improve the visual input, model parameters, as well as our robot hardware for better results with long predictions of object manipulation tasks.

References

[1] S. Levine, C. Finn, T. Darrell, and P. Abbeel, "End-to-end training of deep visuomotor policies," *Journal of Machine Learning Research*, no. 17, pp. 1–40, 2016.

[2] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: a machine learning approach for precipitation nowcasting," in *Proc. of NIPS'15*, vol. 1, pp. 802–810, 2015.

[3] J. Hwang, J. Kim, A. Ahmadi, M. Choi, and J. Tani, "Dealing with large-scale spatio-temporal patterns in imitative interaction between a robot and a human by using the predictive coding framework," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, no. 99, pp. 1–14, 2018.

[4] R. P. N. Rao and D. H. Ballard, "Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects," *Nature Neuroscience*, vol. 2, pp. 79–87, 1999.

[5] Y. Yamashita and J. Tani, "Emergence of functional hierarchy in a multiple timescale neural network model: A humanoid robot experiment," *PLoS Computational Biology*, vol. 4, no. 11, 2008.

[6] P. J. Werbos, "Backpropagation through time: what it does and how to do it," in *Proc. of the IEEE*, vol. 78, pp. 1550–1560, 1990.