Abstract

In this work I present an enhanced end-to-end framework for tracking partially observable dynamic environment, using raw sensor input. This study is a further investigation of Deep Tracking, a recently proposed theoretical framework for unsupervised prediction of space occupancy.

Compared to traditional visual tracking methods which heavily relies on feature engineering, the predictive framework in this study demonstrates very good potential without requiring the same level of pre-existing knowledge.

Deep Learning Model

The original research proposes the four layer recurrent neural network. It uses convolutionsal operations followed by a sigmoid nonlinearity as processing step at each layer.

In my enhanced learning model, I added an additional network layer. This layer is set up for the purpose of tracking extra information. The hope is that it can capture different sizes, trajectories, as well as rotations. In order to maintain computation time, I changed sigmoid non-linear operators to ReLU.

Dataset

The data sets used in this study are entirely synthesized. This is done for two reasons. First, the limited time and resource makes it difficult to collaborate with robotics lab and obtain real sensor data. Second, since the Deep Tracking framework is still under theoretical phase, its limited power means simple synthetic sets are most suitable for study purpose.

The data generation procedures for original study is not provided. Thus I implemented my own lua script to generate combinations of geometric shapes in space, and model these shapes to move in different patterns. The data sets are stored as native Torch 7 files, the size of each individual set (with 1 million frames) is 7.31G before compression, 2.9G each after.

<table>
<thead>
<tr>
<th>Variations</th>
<th>Predictable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td>Yes</td>
</tr>
<tr>
<td>Speed</td>
<td>Yes</td>
</tr>
<tr>
<td>Rotation</td>
<td>Less satisfactory</td>
</tr>
<tr>
<td>Trajectory</td>
<td>Less satisfactory</td>
</tr>
</tbody>
</table>

Further Work

My study shows the limitation of the current DeepTracking model. Next steps of work include more quantitative error analysis, or improvement of neural network architecture. The latter is more challenging because of time limit.

References