Visuomotor Learning from Image Manifolds

Thesis Supervisor: Dr. Amitabha Mukerjee

Date: 30\textsuperscript{th} June, 2014

Presented By: Sunakshi Gupta
Outline

• Motivation
• Inferring Laws of Mechanics
• Learning Visuomotor Skills
  • Learn to throw: Incremental Approach
  • Goal oriented skills: Dart Throwing
• Conclusions
• References
Motivation

- Children have implicit knowledge of various real life systems involving complex physical equations.

- Such models are built on sensory input and improves with practice.

- Maps for several actions (throw, catch, head, kick) which have overlaps, may be combined.

- Objective of this thesis: to learn such maps for different tasks.
‘Route’ v/s ‘Map’

- **Vision**- an important sense to form spatial knowledge.

- Congenitally blind people poorer at ‘maps’ but good at ‘routes’.

- We show how several actions can be combined into a (map), while separate behavioural modules for each of these remain (routes).

A schematic representation of an experimental layout (based on Rieser, Guth, and Hill's descriptions, 1982).
Skill Learning using Maps: Examples I

- Discover similar patterns in various skills which can be then fine-tuned accordingly to gain expertise in one particular skill.

Common visual parameters in various skills. Black box depicts area of interest for our model.
Skill Learning using Maps: Examples II

Visual input for skills like throwing basketball, kicking, catching etc.
Manifolds & Dimensionality Reduction

- High Resolution Images: high dimensionality but low degree of variability.

- Intrinsic dimensionality: factors to be modified to cause variations in images while remaining locally within their subspace.

- Such a subspace can be modeled as a non-linear manifold.

- Manifold depicts topology in Euclidean space near every point.
Inferring Laws of Mechanics

- Earlier approaches use prior knowledge e.g. [Rajagopalan, Kuipers, 1994], [Schmidt, Lipson, 2009] etc.

- Our approach is generalised and learns these systems using latent parameters of the non linear manifold.
Physical Systems covered

- Ball on Incline
- Box and pulley
- Crank and Piston
- Ball Projectile Motion
Algorithm & Approach

• Obtain a low dimensional embedding of the high dimensional image data using ISOMAP.

• Learn a mapping from this manifold to the respective control parameters.

• Find a linear interpolation J for the query image: \( J_i = \sum w_j I_j \)

• Calculate the embedding points for the query image using the weights learnt: \( Q_i = \sum w_j q_j \)

• Use the mapping learnt in Step2 to obtain the corresponding parameters for the query image \( J_i \).
Ball on Incline

- Examples: slides in parks, inclined ramps

Sample Images
Ball on Incline

- Parameters:
  Slope of incline and position of ball on Incline

2-D Neighborhood Graph

Residual Variance Plot
Ball on Incline

Coloring by slope of incline

Coloring by ball position

Coloring by time to reach the end on incline
Box and Pulley

- Examples: wells, elevators, flag hoisting

Sample Images
Box and Pulley

- **Parameters:**
  - the weight(size) of the red box and its height

![Residual Variance Plot](image1)

![2-D Neighborhood Graph](image2)
Box and Pulley

Coloring by **height** of the Box

Coloring by **size** of the box

Coloring by **acceleration**

Coloring by **velocity**
Box and Pulley

<table>
<thead>
<tr>
<th>Query Set</th>
<th>Acceleration (in m/s²)</th>
<th>Velocity (in m/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight=15</td>
<td>0.24</td>
<td>0.63</td>
</tr>
<tr>
<td>Weight=19</td>
<td>0.39</td>
<td>0.75</td>
</tr>
</tbody>
</table>

The predicted fit for the 2 sets of query images
Crank and Piston

- Examples: cars and locomotives engines
Crank and Piston

- 360 angles each for four crank radii (8, 9, 10, 11)
- Parameters:
  Piston location and crank radius

Neighbourhood Graph  Coloring by **Crank radius**  Coloring by **Piston Location**
Crank and Piston

- Query Images with crank radius : 7
- SAE in piston location : 0.62

Result of Interpolation for query images
Ball projectile Motion

- Examples: throwing basketball, kicking, catching

Sample Images
Ball Projectile Motion

- Parameters: Velocity and Angle of projection
- Use of Hausdorff Distance
  \[ h(A, B) = \max_{a \in A} \ \min_{b \in B} \ |a - b| \]
- Performs better than Euclidean

![Residual Variance Plot](image1)

![2-D Neighborhood Graph](image2)
Ball Projectile Motion

- Parameter space \((v, \theta)\) is homeomorphic to image space.

Mapping between parameter \(\theta\) and image space
Ball Projectile Motion

Coloring by **maximum height**

Coloring by **range** of projectile

Coloring by **velocity** of projection

Coloring by **angle** of projection
Learning Visuomotor Skill

- Incremental Approach in learning throwing skill similar to babies (Prediction Model).

- Application of model learnt in goal-oriented skills like dart throwing (Decision Model).
Learn to throw: Incremental Approach

- Children collect information (visual data) through sensory system.
- Process the perception to build model which improves with practice.

Variation in SAE for different velocity ranges
Learn to throw: Incremental Approach

- With limited exposure, the range of learning is restricted.

Throws with unrestricted velocities (Right) cover larger area and attain more number of heights than the ones having restriction on the velocity range.
Goal oriented skills: Dart Throwing

- Model for holistic throw fine-tuned to improve specific goal oriented skills like dart throwing, basketball, catching, kicking etc.

- **Dart throwing**: decision to find throws which hit the dart at the bull’s eye.

Image showing the dart with the target trajectories
Dart Throwing: Algorithm

- Select trajectories with error $< \varepsilon$.

- Fit quadratic through embedding points and respective errors: $q^T S q + b^T q = e$.

- Use values of $S$ and $b$ to learn points with minimal error: $q^T S q + b^T q = 0$.

- Plot the embedding points on the manifold to find the region of accurate trajectories.
Dart Throwing: Results

Two dimensional embedding showing trajectories with minimal error
# Mapping to Language

<table>
<thead>
<tr>
<th>Schema</th>
<th>Top Hindi Lexemes</th>
<th>Score</th>
<th>Freq</th>
<th>ratio</th>
<th>Top English lexemes</th>
<th>score</th>
<th>Freq</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>[dome]</td>
<td>डोम (dome)</td>
<td>13.5</td>
<td>36</td>
<td>1.76</td>
<td>dome</td>
<td>29.3</td>
<td>29</td>
<td>6.00</td>
</tr>
<tr>
<td></td>
<td>पहले (pehley) [before]</td>
<td>7.67</td>
<td>6</td>
<td></td>
<td>before</td>
<td>4.88</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>[throw]</td>
<td>पहले (pehley) [before]</td>
<td>6.95</td>
<td>6</td>
<td>1.27</td>
<td>threw</td>
<td>14.8</td>
<td>14</td>
<td>1.91</td>
</tr>
<tr>
<td></td>
<td>फेंकी (phenki) [throw]</td>
<td>5.48</td>
<td>10</td>
<td></td>
<td>throwing</td>
<td>7.72</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>[roll]</td>
<td>सरकाया (sarkaayaa) [roll]</td>
<td>16.2</td>
<td>15</td>
<td>2.01</td>
<td>rolled</td>
<td>8.12</td>
<td>7</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>सरकायी (saraayi) [roll]</td>
<td>8.05</td>
<td>7</td>
<td></td>
<td>slid</td>
<td>7.28</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

High confidence lexemes
## Mapping to Language

<table>
<thead>
<tr>
<th>SCHEMA</th>
<th>Top scoring Hindi lexical units / cluster</th>
<th>score</th>
<th>ratio</th>
<th>Top scoring English lexical units / cluster</th>
<th>score</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>[throw]</td>
<td>पहले (pehley) [before]</td>
<td>6.89</td>
<td>1.49</td>
<td>threw</td>
<td>14.86</td>
<td>2.36</td>
</tr>
<tr>
<td></td>
<td>गिरा, गिर, गिरी (gira, gir, giri) [fall]</td>
<td>4.6</td>
<td></td>
<td>Throwing, thrown</td>
<td>6.29</td>
<td></td>
</tr>
<tr>
<td>[roll]</td>
<td>सरकायें, सरकाया (sarkaaya, sarkaayi) [roll]</td>
<td>23.66</td>
<td>7.78</td>
<td>rolled</td>
<td>8.07</td>
<td>1.11</td>
</tr>
<tr>
<td></td>
<td>वहीं (wahin) [there]</td>
<td>3.04</td>
<td></td>
<td>slid</td>
<td>7.27</td>
<td></td>
</tr>
</tbody>
</table>

Verb learning
Conclusions & Future Work

- Generalised model to discover visuo-motor patterns in physical systems.

- Holistic view of throw learnt using visual map and applied to dart throwing.

- Further work requires integration with language. Concept knowledge learnt can be used to bootstrap language.
References

Thank you