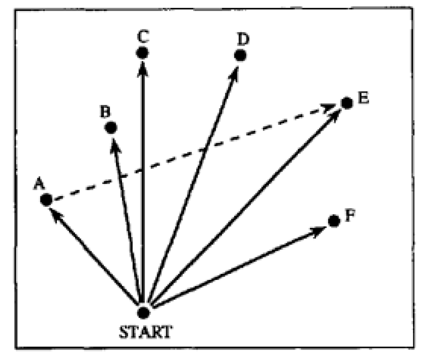
Visuo-Motor Learning using Image Manifolds

Anurag Misra
Indian Institute of Technology Kanpur

INTRODUCTION

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. This is an attempt to learn such maps from different tasks.

'Route' v/s 'Map'

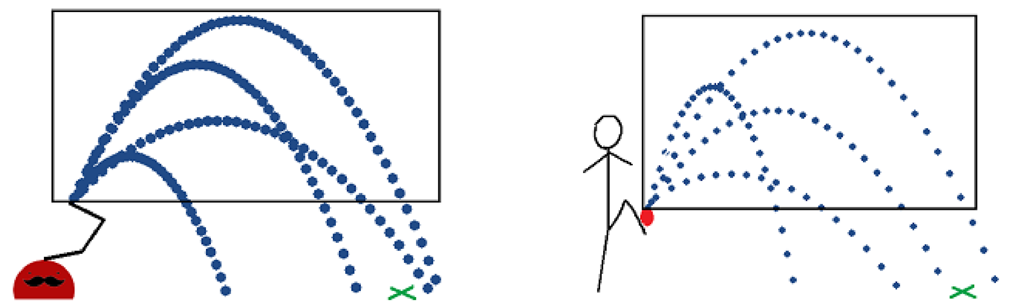


A schematic representation of an experimental layout (based on Rieser, Guth, and Hill's descriptions, 1982).

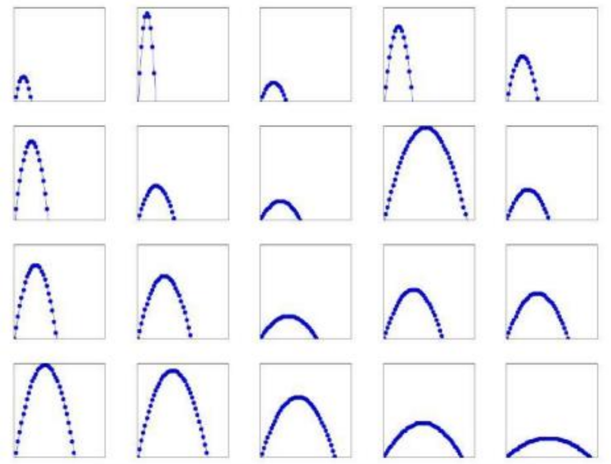
Vision is an important sense to form spatial knowledge. Congenitally blind people poorer at 'maps' but good at 'routes'.
We demonstrate how several actions can be combined into a (map), while separate behavioral modules for each of these remain (routes).

Skill Learning using Maps

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.



Visual input for skills like throwing basketball, kicking, catching etc.

OBJECTIVES

- Learn the generalized map of the actions which can be used for all kinds of activities that involve a visual input and a reaction to that input using any motor action, and find out the basic properties, that if known can help us control the system in the most optimized way
- Develop and describe a generalized algorithm that can be applied to the images of any physical phenomenon and will discover the understanding of the underlying law without any prior domain specific knowledge about the system
- Apply the generalized map of projectile motion to solve the striker-goalkeeper problem i.e. striker has to score given the goal and position of the goalkeeper

ALGORITHM AND APPROACH

Underlying law governing the system

- The basic purpose of this algorithm is to extract the underlying law governing the system
- We generate and supply 1080 images of 100*100 pixels of projectile motion as visual input to the algorithm
- The subspace of throw images varies only the parameters (θ, v), and using this dimensionality of the local tangent space at any point, and by stitching up these tangent spaces we can model the entire subspace as a non-linear manifold of the same intrinsic dimensionality, and mapped to a single joint manifold using ISOMAP

Algorithm 1 Mechanics Law Discovery Algorithm

1: **Input:** Set of high dimensional images $\{I_1, I_2, I_3, ...I_N\}$ and corresponding control parameters

2: **Output:** Value of the control parameters for the set of query images

3: **Step1:** Obtain a low dimensional embedding of the high dimensional image data using ISOMAP dimensionality reduction technique.

4: **Step2:** Train a regression model to acquire a mapping from the low dimensional (curved)manifold to the respective control parameters.

5: **Step3:**For executing a new throw, use a (query) image with desired path. Find a linear interpolation J for this query image $J_i = \sum_{j=1}^{k} w_j I_j$

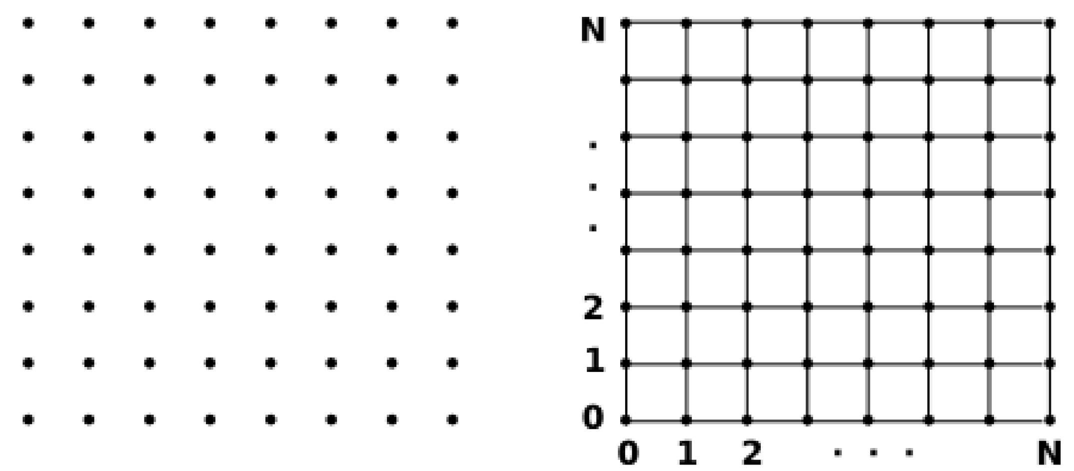
6: **Step4:** Calculate the embedding points for the query image using the weights learnt in Step3. $\hat{Q}_i = \sum_{j=1}^{k} w_j \hat{q}_j$

7: **Step5:** Use the mapping learnt in Step2 to obtain the corresponding parameters for the query image J_i.

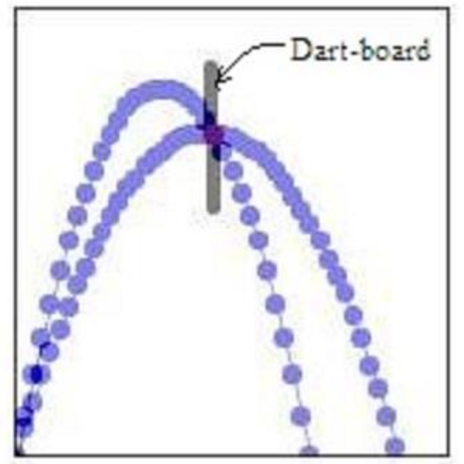
8: **Step6:** Analyse the manifold to depict interesting relations about the physical system.

Solving the Striker-Goalkeeper problem using learnt system

- Assumptions:
  - Goalkeeper still
  - Point of Strike fixed
  - GK position: in front of the goal at some distance
- Approach to solve the problem:
  - Assume the 2D goal to a grid as shown:



- Each point as shown in this grid can be considered as pixel of the (N+1)*(N+1) pixels image/goal
- It is possible for the striker to striker at many of these points considering the position of GK
- Assume there is no goal keeper. For this condition, given the striker position we find the possibly trajectories for each of these points, progressing in a vertical fashion
- For each vertical line in the grid, given the strike position the value of Range R is fixed, which we supply as a parameter
- This approach reduces the problem to 1D, a 1D containing N+1 eligible points for impact. This reduces to Dart Throwing problem:



Dart throwing problem

- Assumptions:
  - Dart board fixed
  - Shoot at fixed point on the dart
  - Point of shoot fixed
- Approach to solve the problem:
  - Dart throwing algorithm

Algorithm 2 Dart Throwing Algorithm

1: **Input:** Image containing the dart.

2: **Output:** Series of trajectories $I_1, I_2..I_p$ that will hit the dart in the most optimal position.

3: **Step1:** Calculate the errors e for all sets of trajectories and select the ones with error $< \epsilon$.

4: **Step2:** Fit a quadratic through the embedding points of the selected trajectories and their corresponding errors such that $\hat{q}^T S \hat{q} + b^T \hat{q} = e$.

5: **Step3:** Use the values of S and b computed in Step 2 to compute the q-coordinates for which the error is minimised i.e. $\hat{q}^T S \hat{q} + b^T \hat{q} = 0$.

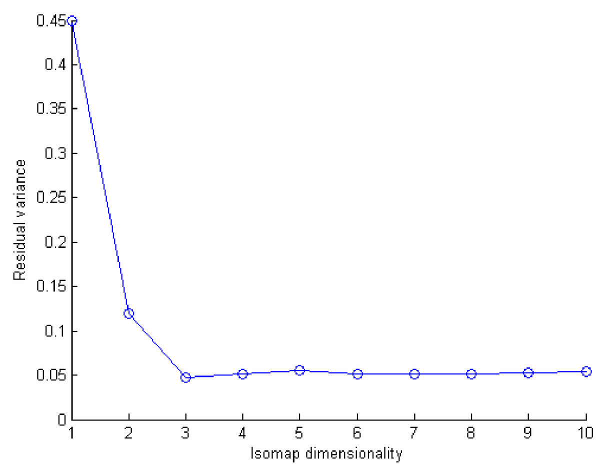
6: **Step4:** Plot the embedding points on the manifold to find the region of accurate trajectories (Figure 3.5).

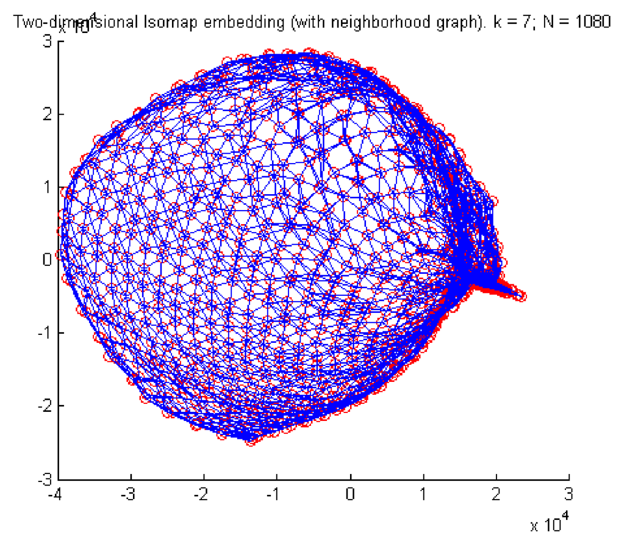
7: **Step5:** Use the mapping learnt in Chapter 2 to generate more throws in the accurate region. Re-calculate errors.

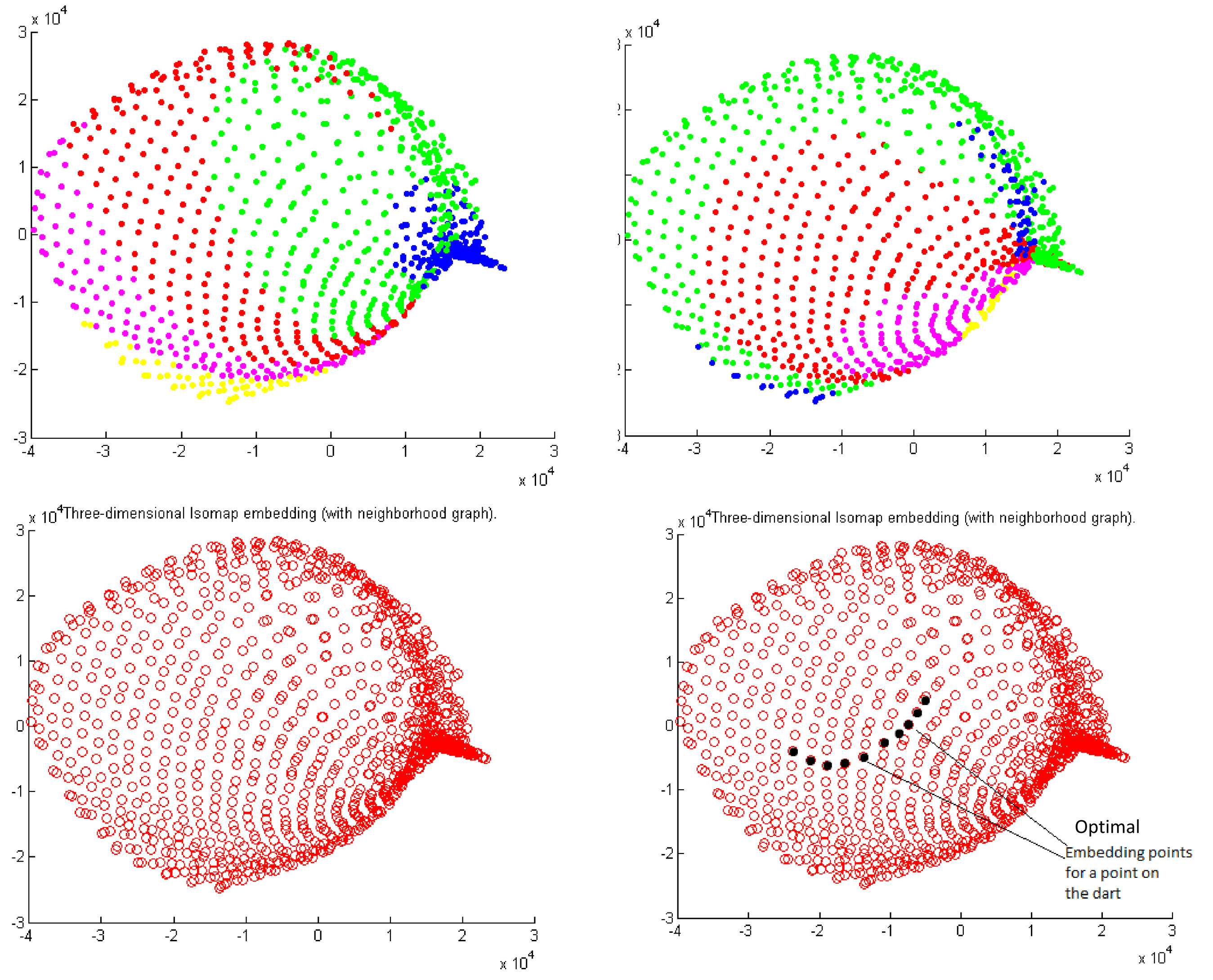
8: **Step6:** Analyse the accurate region on the manifold to select the trajectories with motor parameters that optimise one's effort in throwing.

- After getting all the trajectories for points in the goal, those trajectories which pass through the goalkeeper reach area are eliminated
- This is achieved by obtaining ypos in trajectory for xpos of GK Reach area in the current vertical trajectory plane and checking if that ypos is true for GK Reach Area

RESULTS







CONCLUSIONS

- A generalized model to discover the visuo-motor patterns from the set of images of physical systems or systems involving motor activities like throwing developed and executed
- The model built requires no prior knowledge and hence can be used to learn the patterns for any system where images capture the inherent variability of the system.
- Using the model we solved our goalkeeper-striker problem which can be made an integral of machine football games where skill improves by training

ACKNOWLEDGEMENTS

- I am thankful to Prof. Amitabh Mukherjee for his valuable inputs, constant push, inspiration and guidance
- I am thank to Garuna DevCorp for providing technical help at places I got struck, be it configuring MATLAB Compiler or mex or writing and editing the codes or places I felt lost

REFERENCES

- From visuo-motor to language, Deepali Semwal, Sunakshi Gupta, Amitabha Mukerjee, Department of Computer Science and Engineering, IIT Kanpur
- Resources and code available at http://cocosci.mit.edu/
- Resources available at Google
- A Global Geometric Framework for Nonlinear Dimensionality Reduction Joshua B. Tenenbaum, Vin de Silva, John C. Langford