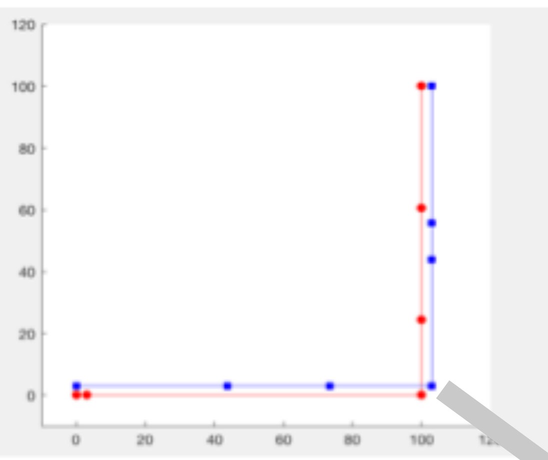
P37 A UNIFORM REPRESENTATION FOR TRAJECTORY LEARNING TASKS

Qingzhe Li, Jessica Lin, Liang Zhao, Huzefa Rangwala
{qli10,jessica,lzhao9}@gmu.edu, rangwala@cs.gmu.edu

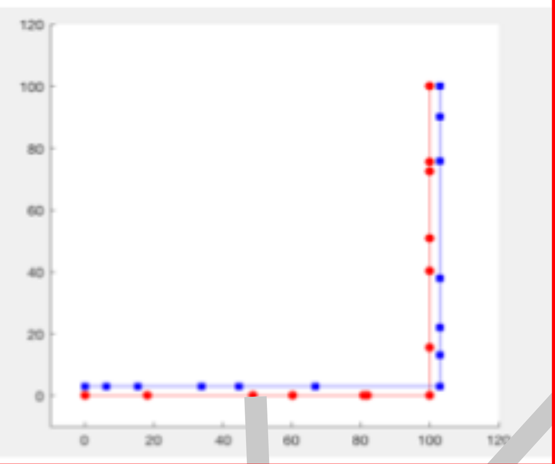
1. Introduction

Motivation : Get Spatial Trajectories Under Control

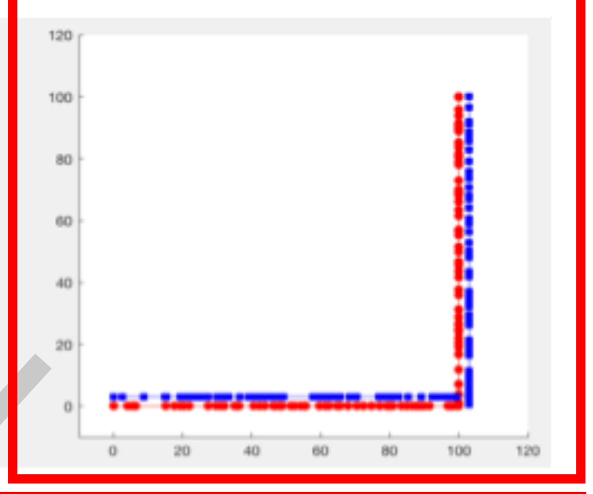
Low sampling rate suffers from:
- Various Speeds
- Uneven Distribution
- Far from the actual distance for most popular trajectory distance measures



Moderate sampling rate suffers from:
- Various Speeds
- Uneven Distribution
- Inaccurate distance for most popular trajectory distance measures



High sampling rate suffers from:
- High battery consumption
- High computational cost
- Uneven Distribution



SIT Representation provides:
- High accuracy
- High quality input data for spatial trajectory algorithms
- Low battery consumption
- Low computational cost

Step Invariant Trajectory (SIT)

SIT Representation handles:
- Low sampling rate
- High sampling rate
- Various Speeds
- Various Sampling rates

2. Methodology

2.1 Notations

- A Spatial Trajectory $T = \{P_0, P_1, \ldots, P_m\}$, where $P_i = (x_i, y_i)$ is the position of the sample point.
- A Step Invariant Trajectory(SIT) $\hat{T}$ is a uniform representation of the trajectory T, where $\hat{T} = \{P_0, P_0^{(1)}, \ldots P_0^{(k0)}, P_1', P_1^{(2)}, \ldots P_1^{(k1)}, \ldots, P_m^{(km)}\}$
- with a constant step distance r for all pairs of consecutive points.
- A Subtrajectory $C^{[s,e]}$ is a subsequence of a step invariant trajectory $\hat{T}$.

2.2 Translating to SIT Representation:

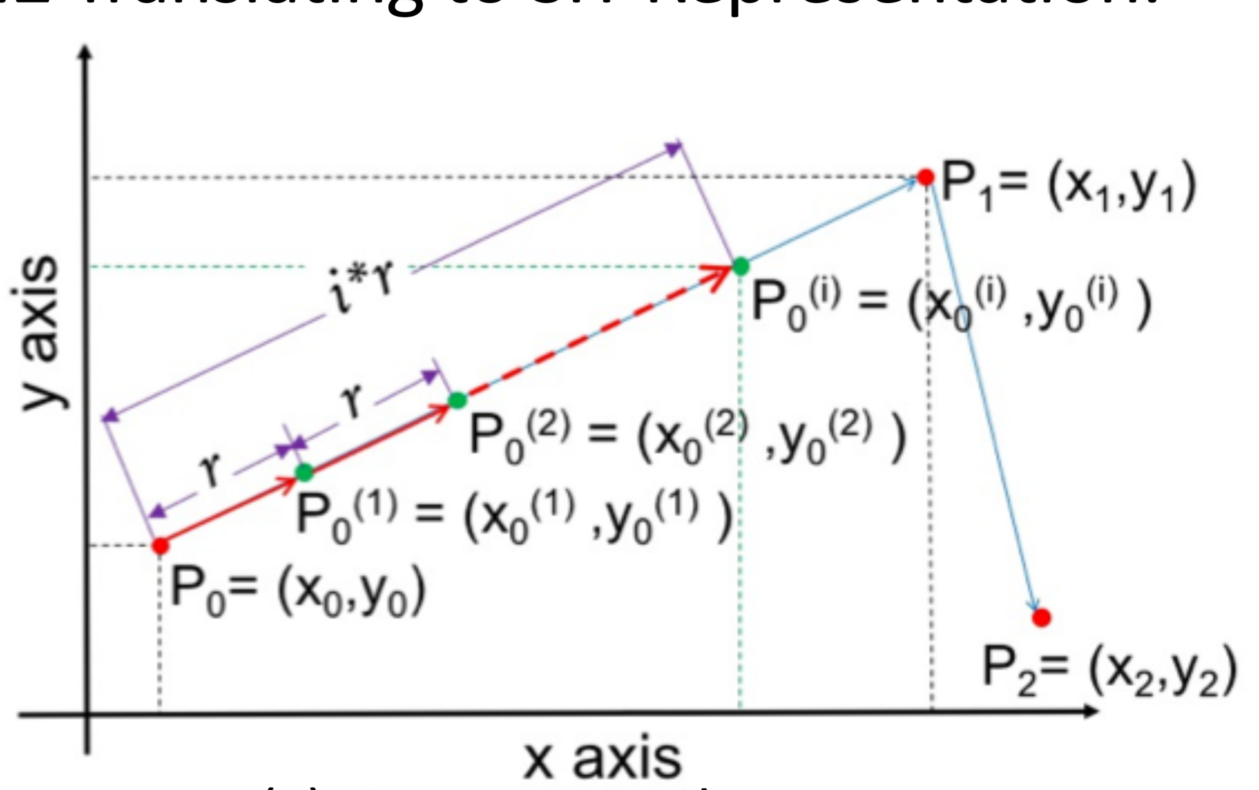


Figure (a) Linear Interpolation

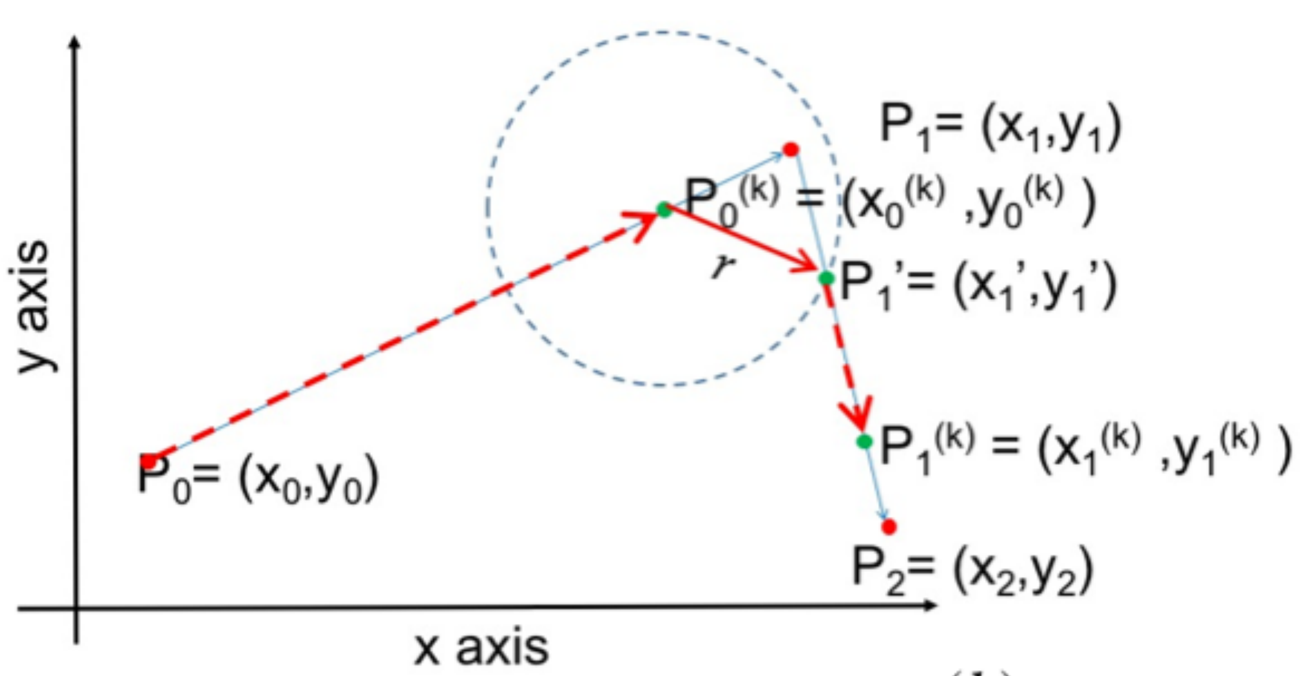


Figure (b) Handling Tails Case 1 $||P_0^{(k)}, P_2|| >= r$

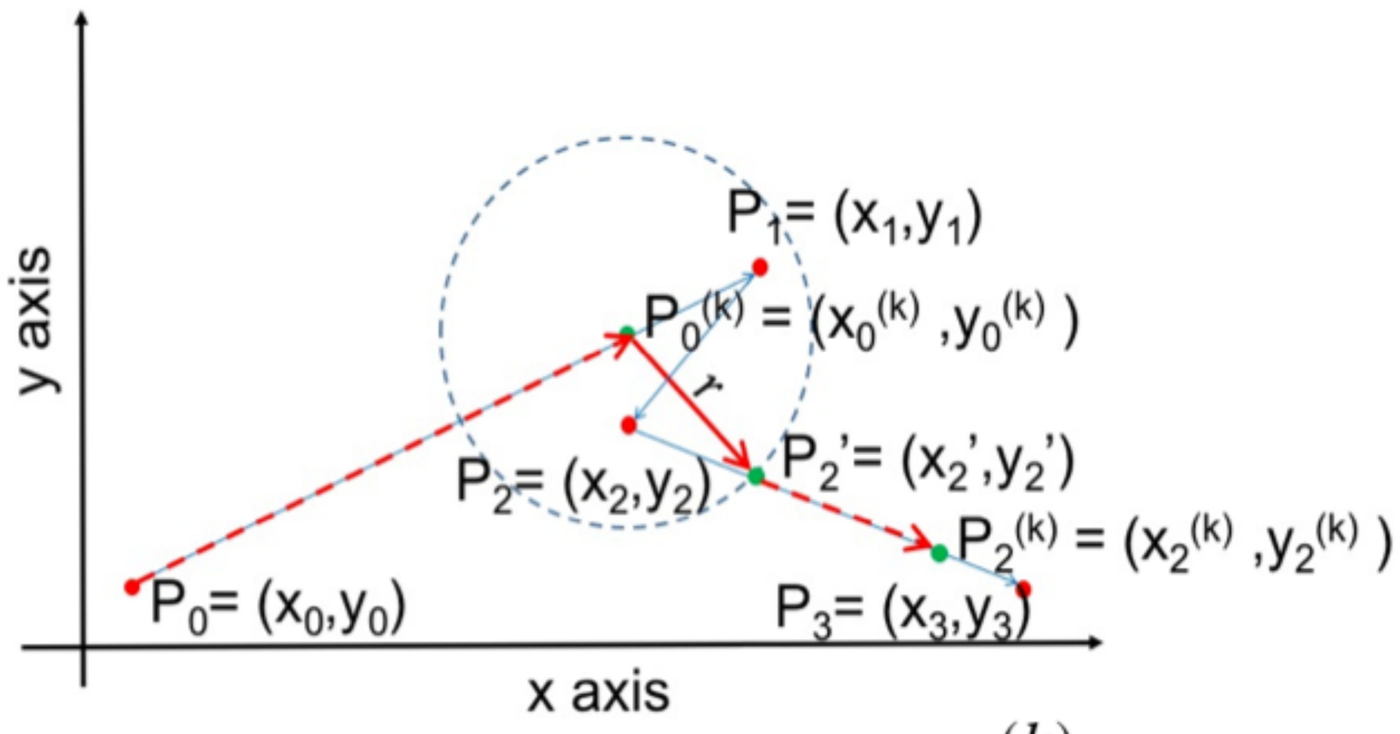


Figure (c) Handling Tails Case 1: $||P_0^{(k)}, P_2|| < r$

2.3 Distance Measure for SIT Representation

- Best Match Euclidean Distance (BMED) by using a sliding window to handle two SITs with different lengths

$$BMED(\hat{T}1, \hat{T}2) = \frac{\min(EuDist(\hat{T}1, C_2^{[s,e]}))}{\sqrt{m}}$$

3. Experiments

Table 3: Translation Time

Representation	CROSS	Speed Up	VMT	Speed Up
AMKS (20 iterations)	286s	91.96	262s	192.64
Tube-Droplet	268s	86.17	184s	135.29
Our SIT ($r = 3$)	**3.11s**	1	**1.36s**	1

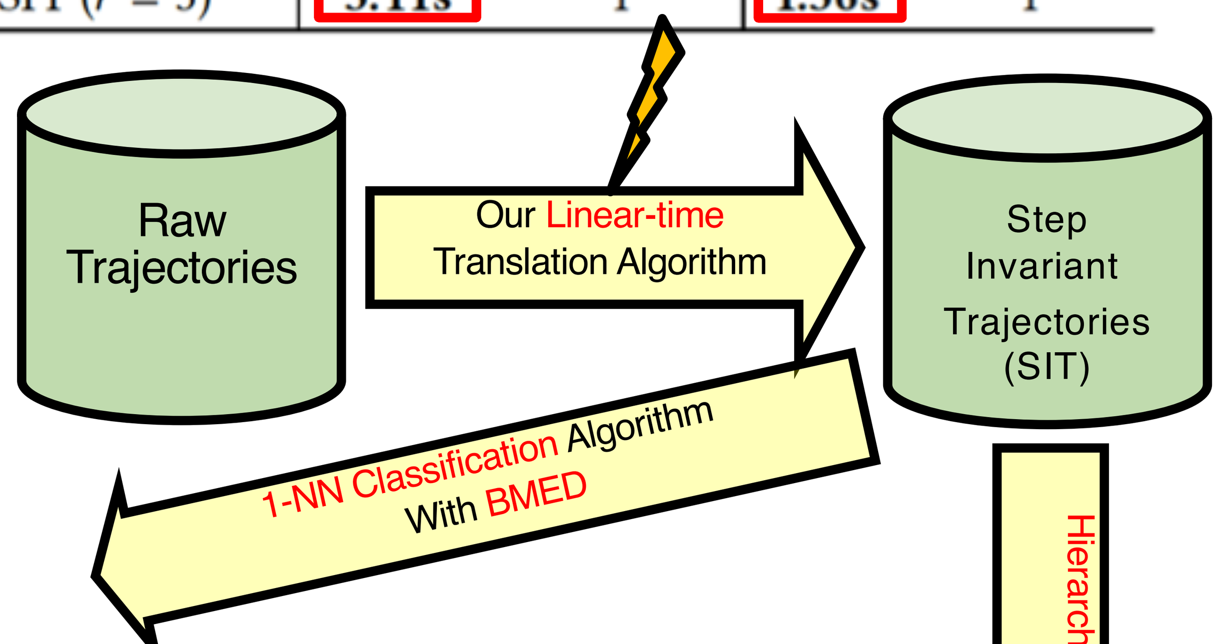
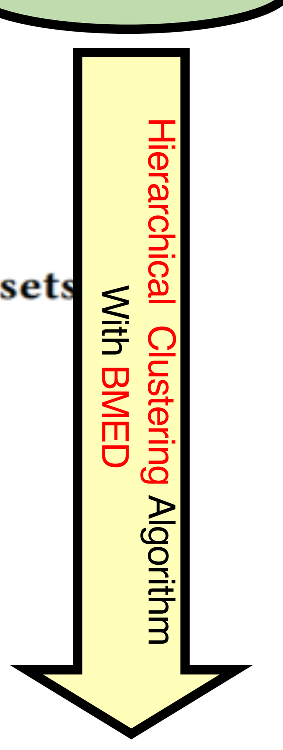
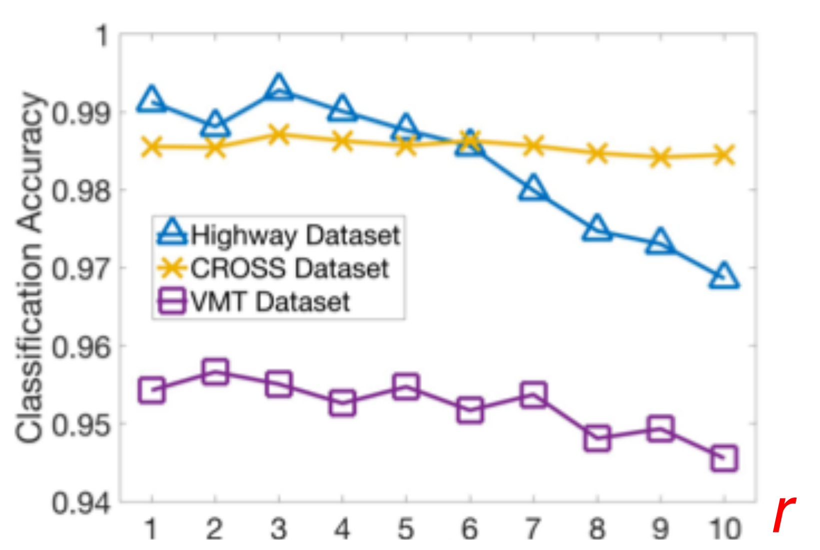


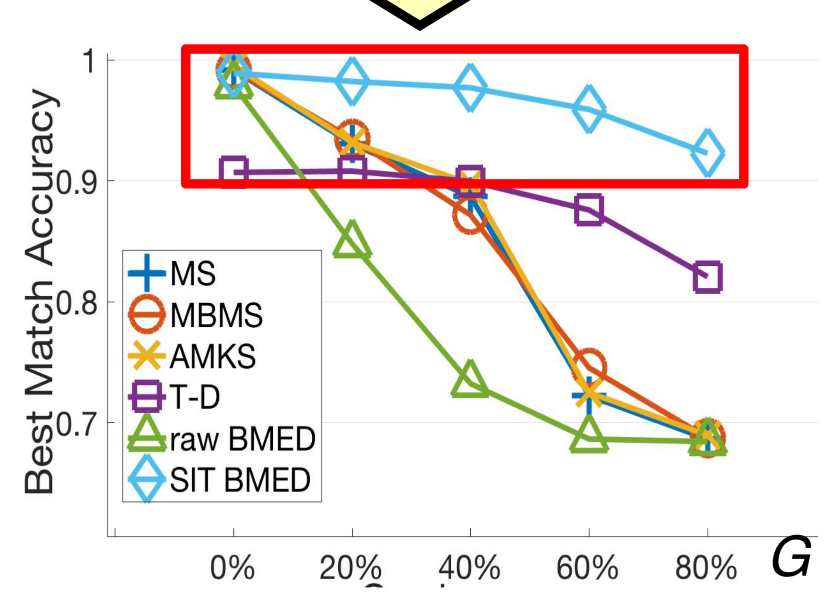
Table 2: Classification Accuracy on Benchmark Datsets

Method	CROSS	VMT
Raw+BMED+1-NN	93.1	92.9
Raw+DTW+1-NN	95.6	94.0
3-Stage Hierarchical [5]	96.8	N/A[1]
tDPMM [3]	98.0	N/A[1]
Tube-Droplet+1-NN[4]	**98.6**	93.6
SIT+BMED+1-NN (Our Method)	**98.6**	**95.4**





Sensitivity Test for r within the same order of magnitude



Clustering result when we removes G points from the original dataset

4. Conclusion

- Effective way to measure trajectory distances
- Low computational cost
- Linear translation time
- Friendly preprocessing step for other spatial trajectory methods
- Improving the performance of other spatial trajectory methods