Parametric Trajectory Representations for Behaviour Classification

Rowland R. Sillito
r.sillito@ed.ac.uk

Robert B. Fisher
rbf@inf.ed.ac.uk

Inst. Perception Action and Behaviour
School of Informatics
University of Edinburgh, UK

This paper attempts to answer a simple question: what is the best way to represent a set of trajectories with fixed-length vectors?

Background. While trajectories provide a highly efficient way to summarise surveillance footage for behaviour classification, their unbounded dimensionality presents an obstacle when applying machine learning techniques. Several approaches [1, 2, 3] address this problem by applying 1D signal approximation techniques to describe the X and Y coordinate sequences of a given trajectory with a fixed set of parameters. We compare four previously-applied variants of this strategy, according to their ability to separate the classes present in several different trajectory datasets.

Trajectory representations. We compare four different ways to represent trajectories, which each correspond to a different family of basis functions (and technique for determining coefficients) including: Haar wavelets ([2], using the Discrete Wavelet Transform), Fourier basis functions ([1], using the Discrete Fourier Transform), Chebyshev polynomials ([1], using the Chebyshev approximation formula) and Cubic Spline basis functions ([3], using least-squares regression), as illustrated in Figure 1.

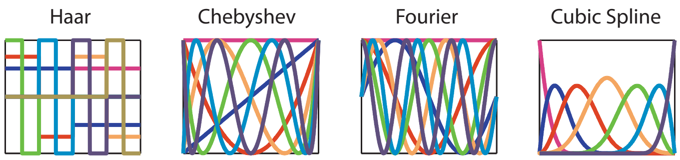

Figure 1: Families of basis functions for curve approximation.

Each trajectory representation method determines two sets of coefficients $\vec{C}^X$ and $\vec{C}^Y$ which define an approximation to the X and Y components of each trajectory in terms of a set of M basis functions $h_1(s), \ldots, h_M(s)$ as follows (where s defines a curve parameter eg. time):

$$X(s) = \sum_{m=1}^{M} h_m(s) \cdot C_m^X \qquad Y(s) = \sum_{m=1}^{M} h_m(s) \cdot C_m^Y$$

We compare the preceding representations in conjunction with two different ways for defining the curve parameter s_n that accompanies the nth coordinate pair x_n, y_n (both normalised to the interval $[0,1]$):

1. Time: $s_n = \dfrac{t_n}{t_N}$

2. Arc-length: $s_n = \dfrac{\sum_{i=2}^{n} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}{\sum_{i=2}^{N} \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}$

Empirical comparison. To determine which of the above strategies is most useful for trajectory classification, we measure the class separability of five different trajectory datasets when described in terms of each different representation for a range of different dimensionalities. These datasets - illustrated in Figure 2 - include pedestrian, vehicle, hand, and pen trajectories, and all contain several distinct classes of motion.

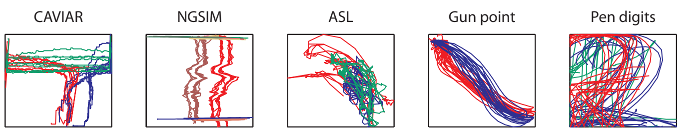

Figure 2: Examples from trajectory datasets used for testing.

Class separability is determined using a non-parametric measure proposed in by Zighed et al. in [5], which operates on the basis of a Relative Neighbourhood Graph [4] spanning a given dataset, with edges weighted by the inverse Euclidean distance between each pair of connected data points. Separability is then defined in terms of the sums of within- and between-class edge weights (E_{within} and $E_{between}$) as follows:

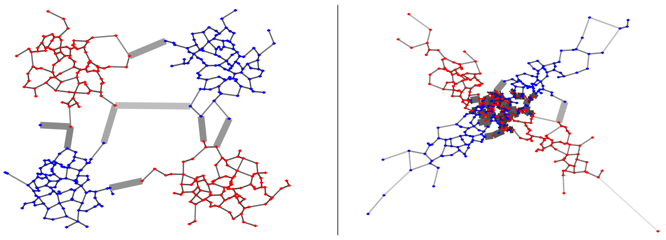

Figure 3: Examples of the Relative Neighbourhood Graph.

$$J_{RNG} = \frac{E_{within}}{E_{within} + E_{between}}$$

For datasets with unequal class sizes, we also measure separability using a version of the leave-one-out nearest-neighbour classification rate that is corrected for differences in class size.

Main findings. Results obtained over a range of dimensionalities indicate that the different representations yield similar levels of class separability. However, a persistent ordering was observed where the Haar representation improved upon the Fourier representation, and the best separability was provided by either the Chebyshev or Spline representations. Figure 4 provides an example of the results obtained for a pen trajectory dataset with 10 classes corresponding to different handwritten digits.

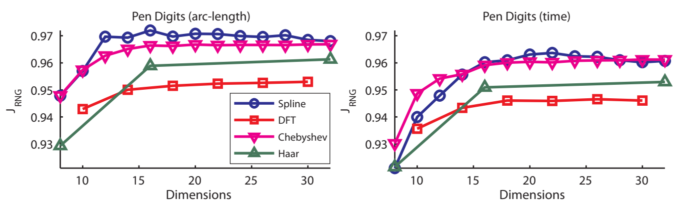

Figure 4: Example result: curves show separability vs dimensionality for each representation (right/left plots ⇒ time/arc-length parametrisation.)

For the datasets considered here, each representation appears to yield better results when used in conjunction with a curve parametrisation strategy based on arc-length, rather than time. However, we illustrate a situation - pertinent to surveillance applications - where the converse is true. These contradictory findings suggest that fixed-length trajectory representations may not necessarily be a good choice for modelling/classifying surveillance footage as they entail an unacceptable compromise between adequately representing the spatial and spatiotemporal characteristics of different trajectories.

[1] A. Naftel and S. Khalid. Classifying spatiotemporal object trajectories using unsupervised learning in the coefficient feature space. *Multimedia Systems*, 12:227–238, 2006.

[2] E. Sahouria and A. Zakhor. Motion indexing of video. In *Proc. IEEE ICIP*, volume 2, pages 526–529, 1997.

[3] R. R. Sillito and R. B. Fisher. Semi-supervised learning for anomalous trajectory detection. In *Proc. BMVC*, pages 1035–1044, 2008.

[4] G. T. Toussaint. The Relative Neighbourhood Graph of a finite planar set. *Pattern Recognition*, 12:261–268, 1980.

[5] D. A. Zighed, S. Lallich, and F. Muhlenbach. A statistical approach to class separability. *Applied Stochastic Models in Business and Industry*, 21:187–197, 2005.