Gaussian Random Vector Fields in Trajectory Modelling

Miguel Barão*, Jorge S. Marques**

** University of Évora and INESC-ID Lisboa, Portugal*
*** Institute for Systems and Robotics (ISR/IST), LARSyS,*
Instituto Superior Técnico, University of Lisboa, Portugal

Abstract

This paper proposes the use of Gaussian random vector fields as a generative model to describe a set of observed trajectories in a 2-dimensional space. The observed trajectories are sequences of points in space sampled from continuous trajectories that are assumed to have been generated by an underlying velocity field. Given the observed velocities connecting the trajectory points, a vector field is obtained by conditioning a Gaussian random vector field. Some results obtained in simulation are presented.

Keywords: Random fields, trajectory modelling, Pedestrian surveillance.

1 Introduction

This paper deals with the estimation of a 2-dimensional vector field describing a set of observed trajectories. One of the possible applications is to estimate models for moving people, cars, animals, etc. The models can then be used in surveillance problems to detect abnormal behavior when new observations (trajectories) do not fit well into the previously estimated models, considered as "normal".

This kind of problems has been tackled before using a: 1) a parametric approach where a model with a small number of parameters, e.g. linear dynamical system, is fit to the data; 2) a nonparametric approach where a grid with a large number of nodes is defined and vectors estimated at those nodes, then the vector field is obtaind by interpolation of those nodes [Nascimento et al., 2014, Nascimento et al., 2015, Ferreira et al., 2013]; 3) Using gaussian process regression flow [Kim et al., 2011].

Here we propose the use of random vector fields to model and estimate the underlying vector field generating the observed trajectory data. The use of the random vector fields provides some advantages over the nonparametric approach. The random vector field approach replaces the interpolation by conditioning the random field by the available data. The random vector field works as a prior and by working uniquely under a probabilistic setting, all uncertainties are taken into account automatically which is not the case when using interpolation.

The main contributions are the random vector field proposal and the issues related with the computational complexity of the algorithm, particularly the replacement of the data by a fixed size statistic that may allow online application of the framework.

The paper is organized as follows: section 2 provides some background on random fields, section 3 formulates the problem, 3.1 provides a simulation example, section 3.2 proposes ways to deal with complexity and finally section 4 draws conclusions.

2 Background

A *random field* is a generalization of a stochastic process where the 1-dimensional "time" parameter is replaced by n-dimensional space. In the most general setup a random field is defined as a measurable function

$$T : \mathscr{M} \times \Omega \to \mathscr{N} \tag{1}$$

 ISBN 978-0-9934207-2-6

where $\mathcal{M}$ and $\mathcal{N}$ are manifolds and Ω is a realization space. In this work, we will be dealing with vector valued random fields in a 2-dimensional image space $T : \mathbb{R}^2 \times \Omega \to \mathbb{R}^2$. In this case, for every realization $\omega \in \Omega$ of the random vector field we get a real vector field $T^\omega : \mathbb{R}^2 \to \mathbb{R}^2$ where $T^\omega \triangleq T(\cdot, \omega)$. Similarly, for every point $x \in \mathbb{R}^2$ in the image space we get a random vector $T_x : \Omega \to \mathbb{R}^2$ where $T_x \triangleq T(x, \cdot)$.

A particular example of a random vector field is one where a Gaussian assumption is made. In this case, the random vector field is completely specified by its mean and covariance functions $\mu(\cdot)$ and $K(\cdot, \cdot)$. The mean function $\mu : \mathbb{R}^2 \to \mathbb{R}^2$ assigns a 2D vector to each point in the 2D image space, while the covariance function (kernel) is a function $K : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ such that, given two points in the image space, returns the covariance matrix that relates the two random vectors at those two points. For example, for any pair of points $x_1, x_2 \in \mathbb{R}^2$, the random vectors $T(x_1)$ and $T(x_2)$ are jointly characterized by a multivariable Gaussian distribution

$$\begin{bmatrix} T(x_1) \\ T(x_2) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu(x_1) \\ \mu(x_2) \end{bmatrix}, \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) \\ K(x_2, x_1) & K(x_2, x_2) \end{bmatrix} \right), \tag{2}$$

where the mean vector and covariance matrix have dimensions 4×1 and 4×4 respectively.

The previous example generalizes to any finite number of points $(x_1, \ldots, x_n)$. The multivariable Gaussian distributions obtained this way can be though of as the marginal distributions from an underlying Gaussian random vector field, provided the conditions of the Kolmogorov extension theorem are satisfied [Billingsley, 1995].

Considering again the joint distribution (2), if the vector $T(x_1) = V$ is observed then the conditional distribution $p(T(x_2) \mid T(x_1) = V)$ characterizes the prediction $T(x_2)$ of the field at the point x_2, which is again Gaussian distributed $\mathcal{N}(\mu^*, K^*)$ with mean and covariance given by

$$\mu^* = \mu_2 + K_{21} K_{11}^{-1} (V - \mu_1), \tag{3}$$

$$K^* = K_{22} - K_{21} K_{11}^{-1} K_{12}, \tag{4}$$

where $\mu_i \triangleq \mu(x_i)$ and $K_{ij} \triangleq K(x_i, x_j)$. Again, this generalizes to any finite number of points partitioned into two sets containing observed and unknown vectors. The prediction of the unknown vectors can be performed using the same equations (3)-(4), where the subscript 1 refers to the observed data and the subscript 2 refers to the predictive part of the mean and covariance matrix.

For a more in depth introduction to gaussian processes refer to [Rasmussen and Williams, 2006].

3 Problem Formulation

In this paper, a set of observed trajectories is used to estimate a generative model that best fits the data. The observed trajectories are represented by sequences of points in a 2-dimensional Euclidean space sampled at regular time intervals.

It is assumed that the trajectories $\{x_t\}$ were generated by flowing along an unknown vector field $T(x)$. The additive variable w_t represents unknown additive perturbations affecting the velocity. Using a normalized time interval $\Delta t = 1$ between samples, gives the generative model

$$x_t = x_{t-1} + T(x_{t-1}) + w_t. \tag{5}$$

Given an observed trajectory $(x_0, x_1, \ldots, x_L)$, the computed velocities are calculated by the difference $v_t = x_{t+1} - x_t$, yielding a set of L position/velocity pairs $\{(x_0, v_0), \ldots, (x_{L-1}, v_{L-1})\}$. In what follows, the collection of positions and velocities are represented in matrix form as $\mathbf{X}$ and $\mathbf{T}$ of size $L \times 2$, where each row represents a particular time instant and the two columns represent the horizontal and vertical axis of the 2-dimensional image space.

To predict the velocities at arbitrary points using the random vector field technique, the desired coordinates $\mathbf{G}$ are appended to the trajectory points $\mathbf{X}$, and the velocities $\mathbf{T}^*$ to be predicted are appended to $\mathbf{T}$ to get the augmented matrices

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{G} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{T} \\ \mathbf{T}^* \end{bmatrix}. \tag{6}$$

 ISBN 978-0-9934207-2-6

Then, given a joint probability distribution $p(\mathbf{T}, \mathbf{T}^*)$, the predicted velocities are then obtained by taking the conditional distribution $p(\mathbf{T}^* \mid \mathbf{T})$, as described in section 2. The following assumptions are made to the joint distribution $p(\mathbf{T}, \mathbf{T}^*)$:

1. The random vector field has zero mean everywhere, $\mu(x) = \begin{bmatrix} 0 & 0 \end{bmatrix}$.
2. Given any two points x_i and x_j, the covariance matrix between their respective velocity vectors is isotropic in $\mathbb{R}^2$ and therefore the covariance matrix is given by $k_{ij}\mathbf{I}_{2\times 2}$, where $k_{ij} \triangleq k(x_i, x_j)$ is a scalar function that depends only on the chosen points.

   This allows covariances to be greatly simplified by using the reduced covariance matrix

   $$\mathbf{K} = \begin{bmatrix} k_{11} & \cdots & k_{1n} \\ \vdots & & \vdots \\ k_{n1} & \cdots & k_{nn} \end{bmatrix} \tag{7}$$

   instead of the full matrix, which is given by the Kronecker product $\mathbf{K} \otimes \mathbf{I}_{2\times 2}$.
3. The kernel function $k(\cdot, \cdot)$ used to define the covariance is a positive decreasing function depending on the distance between the two points.

The previous three assumptions impose a stationarity condition in space. Examples of such functions are the Ornstein-Uhlenbeck, squared exponential and triangular functions

$$k_{ou}(x_1, x_2) \triangleq \exp(-\alpha \| x_1 - x_2 \|), \tag{8}$$

$$k_{se}(x_1, x_2) \triangleq \exp(-\alpha \| x_1 - x_2 \|^2), \tag{9}$$

$$k_{tri}(x_1, x_2) \triangleq \max(1 - \alpha \| x_1 - x_2 \|, 0), \tag{10}$$

with parameter α adjusting the spatial dependency between the points.

Given the observed trajectories $\mathbf{X}$ and the grid coordinates $\mathbf{G}$, the covariance matrix is computed in partitioned form as

$$\mathbf{K} = \begin{bmatrix} \mathbf{K}_{xx} & \mathbf{K}_{xg} \\ \mathbf{K}_{gx} & \mathbf{K}_{gg} \end{bmatrix} \tag{11}$$

where the subscripts x and g denote respectively the part of the observed data and the points of the grid where prediction is to take place.

The velocity vectors can now be predicted using (3)-(4):

$$\mu^* = \mathbf{K}_{gx}\mathbf{K}_{xx}^{-1}\mathbf{V} \tag{12}$$

$$\mathbf{K}_{gg}^* = \mathbf{K}_{gg} - \mathbf{K}_{gx}\mathbf{K}_{xx}^{-1}\mathbf{K}_{xg} \tag{13}$$

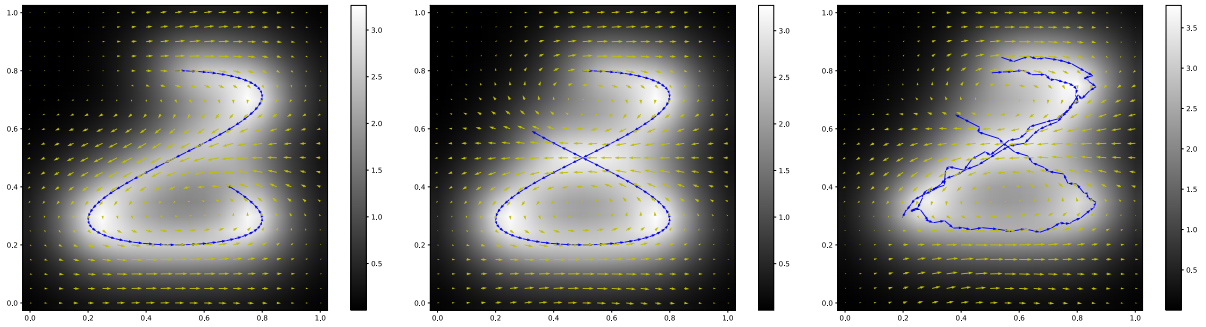
where the zero mean was dropped from the equations.

3.1 Example

To illustrate the algorithm a trajectory was generated and the prediction was performed on a regularly spaced 21×21 grid using a squared-exponential kernel. Figure 1 shows the observed trajectory in blue. The predicted velocities at the grid are jointly gaussian with mean μ^* and covariance matrix $\mathbf{K}_{gg}^*$. The figure shows the marginals distributions for individual points, with the mean represented by green arrows and the variances obtained from the diagonal of $\mathbf{K}_{gg}^*$ represented by the background gray level in log-scale, darker meaning higher variance/uncertainty in the prediction.

It can be seen that the nodes near the trajectory have much lower uncertainty in the predicted velocity than nodes in areas where no nearby data exists. The darker areas tend to give a result close to the prior, assigning near zero velocity.

 ISBN 978-0-9934207-2-6



(a) Single trajectory of length 70. (b) Single trajectory of length 100. (c) Two noisy trajectories with different lengths.

Figure 1: Observed trajectories (blue) and random field prediction at a regularly space grid (yellow). The gray background represents the uncertainty associated with the prediction.

Although a crossing exists in the trajectory, which is impossible in a deterministic dynamical system, the solution found can be interpreted as two separate upper and lower regions with circular motion. The crossing is then explained by the stochastic nature of the problem where a perturbation can produce the jump from one region to the other.

3.2 Dealing with complexity

A practical problem of directly applying the equations (12)-(13) is dealing with variable size and always increasing amount of data. A suboptimal solution to this problem is to use the predicted vectors at the grid as a fixed size statistic that describes the past observed data. As new observations are obtained the statistic can then be updated and the data discarded.

To implement a fixed size statistic, a fixed size grid is used. The algorithm now works in two steps: in the first step the grid vectors are estimated and in the second step the vector field is predicted from the grid, which now acts as a new "virtual" data, instead of the original trajectory.

As a further reduction in complexity, the grid nodes with high uncertainty can be omited and the prediction can be performed using only smaller but relevant information. Figure 2 shows three fields generated from subsets of nodes from solution in figure 1.

This solution is clearly suboptimal since information is being retained in areas where an already good description exists and new contributions are small. A possibly better approach would be to keep the nodes that lead to the largest information gain. This line of research is still ongoing.

4 Conclusions

This paper deals with the use Gaussian vector random fields to build models describing a set of trajectories observed in 2-dimensional space. The use of the random vector field framework has the main advantage that all the uncertainties are being taken into account. The random vector field can be seen as providing a prior, that when constrained on the data provides prediction for the rest of the space. Regions where no data is observed are then closer to the prior and have higher associated uncertainty.

Constraining on all available data has the drawback that, in an online setting, the complexity is always increasing. Here we propose to replace the data by a fixed size statistic that is then updated online. The size of this statistic is experimented with by selecting only the nodes with lower uncertainty. While this is not an optimal solution, it provides a first step in the pursuit of a sparse solution that keeps the complexity of the algorithm sufficiently low to be able to run it online.

 ISBN 978-0-9934207-2-6

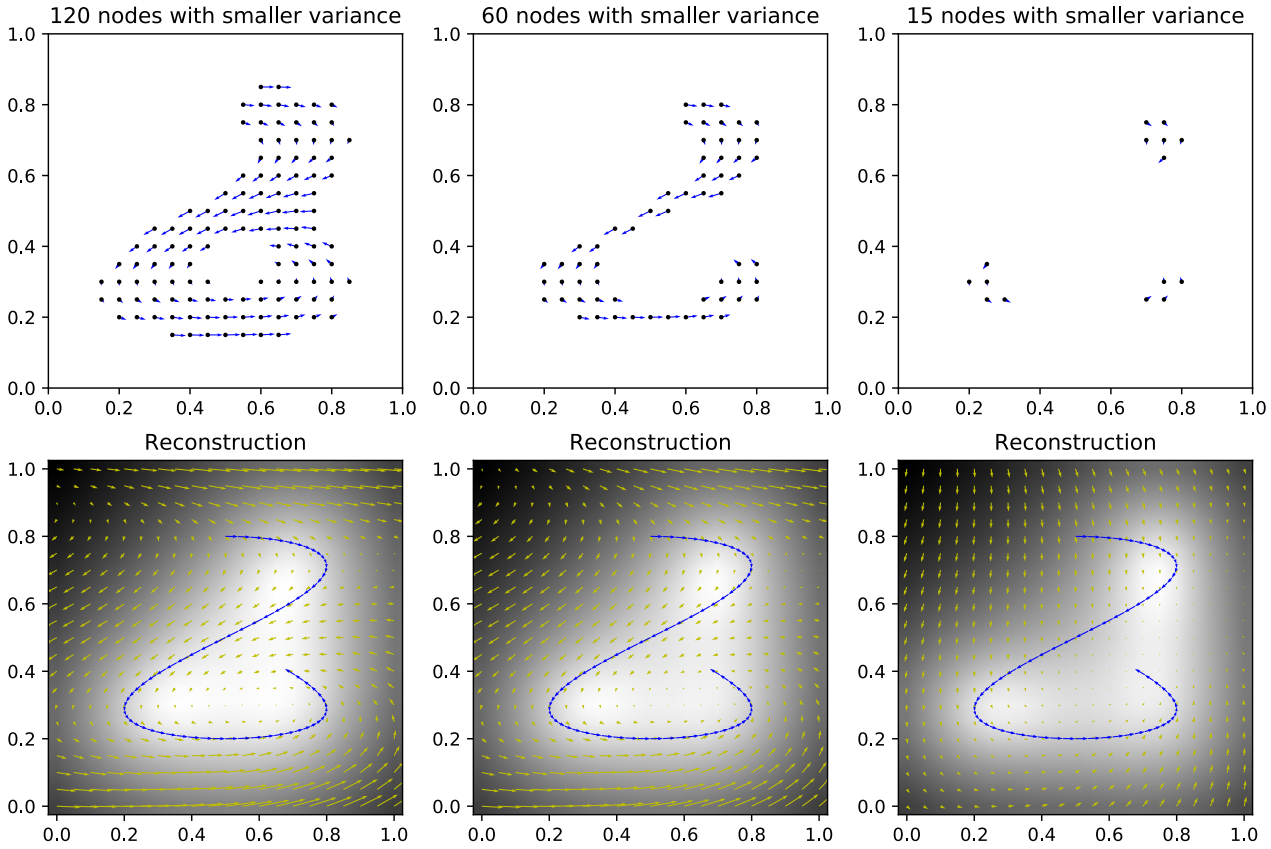


Figure 2: Observed trajectory (blue) and random field prediction at a regularly space grid (yellow). The gray background represents the uncertainty associated with the prediction.

 ISBN 978-0-9934207-2-6

Acknowledgments

This work was partially supported by FCT (Fundação para a Ciência e Tecnologia), grants PTDC/EEIPRO/0426/2014, UID/CEC/50021/2013 and UID/EEA/50009/2013.

References

[Billingsley, 1995] Billingsley, P. (1995). *Probability and Measure*. John Wiley and Sons.

[Ferreira et al., 2013] Ferreira, N., Klosowski, J. T., Scheidegger, C. E., and Silva, C. T. (2013). Vector fields k-means: Clustering trajectories by fitting multiple vector fields. In B. Preim, P. R. and Theisel, H., editors, *Eurographics Conference on Visualization*, volume 32.

[Kim et al., 2011] Kim, K., Lee, D., and Essa, I. (2011). Gaussian process regression flow for analysis of motion trajectories. In *IEEE International Conference on Computer Vision*.

[Nascimento et al., 2014] Nascimento, J. C., Barão, M., Marques, J. S., and Lemos, J. M. (2014). Information geometric algorithm for estimating switching probabilities in space-varying HMM. *IEEE Transactions on Image Processing*, 23(12):5263–5273.

[Nascimento et al., 2015] Nascimento, J. C., Barão, M., Marques, J. S., and Lemos, J. M. (2015). An information geometric framework for the optimization on discrete probability spaces: Application to human trajectory classification. *Neurocomputing*, 150:155–162.

[Rasmussen and Williams, 2006] Rasmussen, C. E. and Williams, C. (2006). *Gaussian Processes for Machine Learning*. MIT Press.