Improving a Switched Vector Field Model for Pedestrian Motion Analysis

Catarina Barata, Jacinto C. Nascimento, and Jorge S. Marques[1]

[1]Instituto de Sistemas e Robótica, Instituto Superior Técnico,
1049-001 Lisboa, Portugal

Abstract. Modeling the trajectories of pedestrians is a key task in video surveillance. However, finding a suitable model to describe the trajectories is challenging, mainly because several of the models tend to have a large number of parameters to be estimated. This paper addresses this issue and provides insights on how to tackle this problem. We model the trajectories using a mixture of vector fields with probabilistic switching mechanism that allows to efficiently change the trajectory motion. Depending on the probabilistic formulation, the motions fields can have a dense or sparse representation, which we believe influences the performance of the model. Moreover, the model has a large set of parameters that need to be estimated using the initialization-dependent EM-algorithm. To overcome the previous issues, an extensive study of the parameters estimation is conducted, namely: (i) *initialization*, and (ii) *priors distribution* that controls the sparsity of the solution. The various models are evaluated in the trajectory prediction task, using a newly proposed method. Experimental results in both synthetic and real examples provide new insights and valuable information how the parameters play an important in the proposed framework.

Keywords: Surveillance, Trajectories, Motion fields, Hidden Markov Models, Expectation Maximization

1 Introduction

Activity recognition, movement prediction (see Fig. 1), and detection of abnormal behaviors are critical tasks in surveillance applications, requiring appropriate models to describe them. Since many of the surveillance settings are based on far-field cameras that do not provide detailed information about the pedestrians in a video scene, several of these models are based on the characterization of the trajectories performed by the pedestrians. In fact, trajectories allow us to identify and collect statistics of typical motions, activities, and interactions in a video scene [1]. Consequently, having a reliable description of the possible trajectories is of major importance. However, trajectory modeling is a challenging problem due to its great spatial and temporal variability.

Various methods have been proposed to efficiently describe trajectories. Amongst these methods, we are interested in the ones that resort to a generative approach,

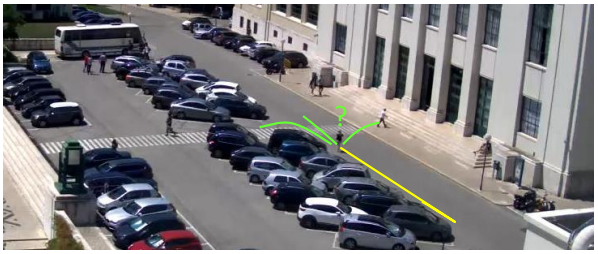

Fig. 1. Pedestrian trajectory: past (yellow) and possible future motions (green).

i.e, the trajectories are assumed to follow a dynamical equation and are governed by known motion patterns. In [2–4] a Gaussian process (GP) regression is used for trajectory modeling under a probabilistic viewpoint. More specifically, in [3] the authors present an unsupervised algorithm in which GP is used to estimate the instantaneous velocity of the pedestrian, while in [2] GP is used to model flow functions. The latter representation also allows for incrementally predicting possible paths and detecting anomalous events from online trajectories. In [4] it is proposed an incremental and unsupervised approach, where the model is updated by receiving new trajectory samples. Dirichlet processes have also been used to model trajectories [6, 8, 7], as well as vector fields estimated using the k-means algorithm [5]. The latter work demonstrated the ability of vector fields to describe other types of trajectories besides pedestrian ones (e.g., hurricanes and cell-phone GPS data).

Our work is related to [9] where vector fields are used to model pedestrian trajectories. This method assumes that a grid with $\sqrt{n} \times \sqrt{n}$ nodes is defined in the image domain. Each node is characterized by the following parameters: (*i*) a set of *motion vector fields* (velocity vectors), (*ii*) a field of isotropic *covariance matrices*, and (*iii*) a switching *probability matrix* that governs the transition between fields in that node. Only one motion field is active at each trajectory point and the transition between them depends on the spatial location of the pedestrian in the grid. This method has been show to efficiently characterize trajectories in the task of activity recognition [9]. However, it has some limitations: (*i*) the parameter estimation is performed using the EM algorithm that strongly depends on the initialization, and may encourage wrong estimates; and (*ii*) the fields were assumed to be dense and estimated in nodes where information was not available. Recently, a new formulation was proposed to deal with the latter issue: imposing a sparsity constraint to the vector fields [10]. However, this work failed to demonstrate the impact of such strategy at describing pedestrian trajectories. Moreover, to the best of our knowledge, none of the vector fields models has ever been used to predict future positions in a trajectory. We believe that this is a critical task in surveillance applications.

This paper provides valuable and comprehensive contributions, namely: (*i*) provides an effective way on how to *initialize* the vector fields in the EM algorithm, (*ii*) evaluates different vector fields formulations, and (*iii*) provides a new methodology for trajectory prediction, using the space-varying dynamical model.

For the first item, we perform a systematic study comprising the following three strategies: (a) *random*, (b) *uniform*, and (c) *clustering* initialization, thus not jeopardizing deficient estimates related to poor initializations. For the second item, we compare the vector fields formulations of [9] and [10]. The proposed prediction strategy is used to evaluate the initialization and type of fields.

2 Switching Dynamical Model

In this work, the trajectory of a pedestrian is assumed to be represented as a sequence of positions $x = (x_1, \ldots, x_L)$, with $x_t \in [0,1]^2$. We also consider that all of the possible motions in a scene may be described using K motion fields $T_k : [0,1]^2 \rightarrow \mathbb{R}^2$, where $k \in \{1, ..., K\}$ is the identifier of the field. Only one of these fields is active at a specific time instant, but it is possible to switch between them. Under these assumptions, each position is generated as follows

$$x_t = x_{t-1} + T_{k_t}(x_{t-1}) + w_t \ , \tag{1}$$

where k_t denotes the active motion field at time t and $w_t \sim N(0, \sigma^2_{k_t} I)$ is a white random perturbation with zero mean and isotropic covariance.

Switching between active fields depends on the pedestrian position, and it is modeled as a first order Markov process

$$P(k_t = j | k_{t-1} = i, x_{t-1}) = B_{ij}(x_{t-1}) \ , \tag{2}$$

where $B(x) = \{B_{ij}(x), i, j \in \{1, \ldots, K\}\}$ is a stochastic space-varying matrix of switching probabilities (transition matrix) that depends on the position x, *i.e.*, switching probabilities at different positions are usually different.

We specify the motion fields T_k and the transition matrix B using a regular grid on the image $\mathcal{G} = \{g_i \in [0,1]^2, i = 1, \ldots, n\}$. This means that each node g_i of the grid is associated with K velocity vectors T^i_k and a transition matrix B^i. Outside of the nodes, the velocity vector $T_k(x)$ and switching matrix $B(x)$ are obtained by bilinear interpolation, as detailed below

$$T_k(x) = \mathcal{T}_k \Phi(x) \ , \tag{3}$$

$$B(x) = \sum_{i=1}^{n} B^i \phi^i(x) \ , \tag{4}$$

where $\mathcal{T}_k$ is the dictionary of node velocities T^i_k associated with the $k-th$ field, B^i is the switching matrix associated with node g_i, $\Phi(x)$ is a $N \times 1$ normalized sparse vector of known interpolation coefficients, and $\phi^i(x)$ is its $i-th$ element.

If we assume a L-sequence trajectory and a known sequence of active fields $k = (k_1, ..., k_L)$, it is easy to compute the joint probability $p(x, k)$

$$
\begin{aligned}
p(x,k|\theta) &= p(x_1,k_1)\prod_{t=2}^{L} p(x_t|k_t,x_{t-1})p(k_t|k_{t-1},x_{t-1}) \\
&= p(x_1)P(k_1)\prod N(x_t|x_{t-1}+T_{k_t}(x_{t-1}),\sigma^2_{k_t}I) \\
&\times B_{k_{t-1},k_t}(x_{t-1}) \ , \qquad (5)
\end{aligned}
$$

where $\theta = (\mathcal{T}, B, \sigma^2)$ denotes the complete set of model parameters. From (5), it is clear that the sequence of active fields $k = (k_1, ..., k_L)$ is modeled as a realization of a first order Markov process, with some initial (known) distribution $P(k_1)$ and a space-varying transition matrix, i.e., $B_{k_{t-1},k_t}(x_{t-1})$. In the following section we address the estimation of the model parameters.

3 MAP Estimation Of The Model

Lets assume that we have a set of S observed trajectories, denoted as $\mathcal{X} = \{x^{(1)}, \ldots, x^{(S)}\}$, where $x^{(j)} = (x_1^{(j)}, ..., x_{L_j}^{(j)})$ is the j-th trajectory with L_j samples. Naturally, the corresponding sequences of active fields $\mathcal{K} = \{k^{(1)}, ..., k^{(S)}\}$ are missing. Given this set, our goal is to estimate the model parameters $\theta = (\mathcal{T}, B, \sigma^2)$ using the MAP formulation

$$
\hat{\theta} = \arg\max_{\theta} \Big[\log p(\mathcal{X}|\theta) + \log p(\theta)\Big] \ , \qquad (6)
$$

where $\log p(\theta) = \sum_{k=1}^{K} \log p(\theta_k)$ is the log prior distribution. Assuming independence between the various parameters, it is possible to further decompose the prior as follows

$$
\log p(\theta) = \sum_{k=1}^{K} \log p(\mathcal{T}_k) + \sum_{k=1}^{K} \log p(\sigma_k^2) + \log p(B) \ . \qquad (7)
$$

The complete likelihood $\log p(\mathcal{X}, \mathcal{K}|\theta)$ may also be easily defined from (5), by computing a sum over all the trajectories. However, its marginalizing over all the admissible labels $\mathcal{K}$ cannot be computed.

The aforementioned limitation may be addressed using the Expectation-Maximization (EM) method to estimate the model parameters. This amounts to iteratively maximizing the auxiliary function

$$
U(\theta, \theta') = E\left\{\log p(\mathcal{X}, \mathcal{K}|\theta)|\mathcal{K}, \theta^{(t)}\right\} + \log p(\theta) \qquad (8)
$$

with respect to θ and taking into account the available estimates θ' from the previous iteration.

Using (5) and replacing into (8) we obtain the following expression (disregarding the constants)

$$U(\theta,\theta') = -\sum_{s=1}^{S}\sum_{t=2}^{L}\sum_{k=1}^{K} w_k^{(s)}(t)\log N(x_t|x_{t-1}+T_{k_t}(x_{t-1}),\sigma_{k_t}^2 I) + \sum_{s=1}^{S}\sum_{t=2}^{L}\sum_{p,q=1}^{K} w_{p,q}^{(s)}(t)\log B_{k_{t-1},k_t}(x_{t-1}) + \sum_{k=1}^{K}\log p(\mathcal{T}_k) + \sum_{k=1}^{K}\log p(\sigma_k) + \log p(B) \ , \quad (9)$$

where $w_p^{(s)}(t) = P(k_t^{(s)} = p|x^{(s)},\theta')$ and $w_{p,q}^{(s)}(t) = P(k_{t-1}^{(s)} = p, k_t^{(s)} = q||x^{(s)},\theta')$. These weights are computed in the E-step, using the forward-backward algorithm proposed in [12]. The optimization with respect to θ is performed in the M-step, and involves a separate maximization with respect to $\mathcal{T}, B$, and σ^2 [9].

As discussed in [10] the prior $\log p(\theta)$ may be used to incorporate *a priori* knowledge about the parameters, as well as to enforce specific properties. In particular, efforts have been made to define a prior distribution $p(\mathcal{T})$ that could lead to *smooth* and *sparse* fields:

$$\log p(\mathcal{T}_k) = \alpha\|\Delta\mathcal{T}_k\|_2^2 + \beta\|\mathcal{T}_k\|_p^p \ , \quad (10)$$

where Δ is an operator that computes all differences between velocities of neighboring nodes and $\|.\|_p$ denotes the pth norm. Here, the first term sets that any neighbor grid nodes $x_{g1}, x_{g2} \in \mathcal{G}$ should have similar estimated velocities, i.e., the difference $\mathcal{T}_{k_t}(x_{g1}) - \mathcal{T}_{k_t}(x_{g2})$ should be small. The second term enforces small velocity values in most of the nodes. These values may be zero if $p = 1$, leading to sparse vector fields where the velocities are only estimated for nodes that are supported by observations [10]. The values chosen for the regularization constants α and β and norm p define the importance of each term, and may influence the estimation of the fields T_k. However, the assessment of their role and relevance has never been performed.

Besides the priors, another criteria that strongly influences the estimation of the model is the initialization of the parameters θ_{init}, since the EM algorithm is very sensitive to initialization [11]. In the following section we propose an experimental framework to evaluate the role of these criteria in the estimation of the motion models.

4 Experimental Setup

The goal of this work is to provide an extensive experimental study on the role of both initialization and the prior $p(\mathcal{T})$ in the estimation of the motion models. To conduct such a study it is necessary to i) identify the types of initialization and priors to be tested; ii) select one or more datasets to conduct the experiments; and iii) define an evaluation metric. In this section we address each of these requirements.

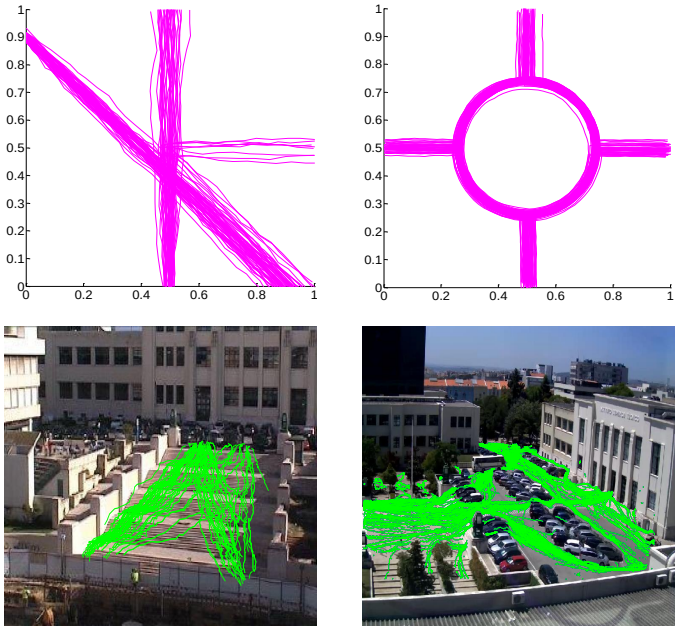

Fig. 2. Experimental datasets: synthetic examples (1st row) and real pedestrian trajectories (2nd row).

i) Initialization: The EM-algorithm, which is used in the estimation of the parameters, strongly depends on initialization. Thus, we will evaluate three types of initialization: i) random, where each node of a field is initialized with a random velocity vector; ii) uniform, where each node of a field is initialized with the same (random) velocity vector; and iii) clustered, where we perform a preliminary clustering of the velocity values and set the number of centroids to be equal to the number of fields K.

ii) Prior: Another goal of this paper is to provide a comprehensive study of the role of the fields' prior (10). This means that for each of the aforementioned initializations, we train several models using different combinations of $\alpha \in \{0.5, 1, 2\}$, $\beta \in \{0, 0.05, 0.1, 0.15\}$, and $p \in \{1, 2\}$, leading to a total of 21 models per initialization type.

iii) Datasets: Our study is conducted using four datasets (two synthetic and two real). The first synthetic case comprises $K = 3$ motion types and 100 randomly generated trajectories, as is exemplified in Fig. 2 (1st row-left). The trajectories were generated according to the following procedure. First, the starting point was randomly chosen between a small region around $[0, 0.9]^T$ and $[0.5, 0]^T$. Conditioned on the selected point, the motion regime can either be a down motion with a slope of 45° or an up motion. When the trajectories reach the center of the image there is a probability of switching between these motions and a third one, oriented form left to right.

Table 1. Dataset description. $\bar{L}$ is the average trajectory length and $\bar{v} = \frac{1}{L}\sum_{t=1}^{L} \|x_t - x_{t-1}\|_2^2$ is the average displacement.

Dataset	# Trajectories	$\bar{L}$	$\bar{v}$ ($\times 10^{-3}$)
Synthetic 1	100	27 ± 5	46 ± 4
Synthetic 2	100	77 ± 40	21 ± 0.4
Stairway	86	57 ± 18	12 ± 2
Campus II	1083	50 ± 48	2 ± 0.9

The second synthetic example mimics a roundabout with four entrance/exist points, as exemplified in Fig. 2 (1st row-right). Three motion regimes ($K = 3$) are considered: a circular and counterclockwise one, and two linear movements that represent, respectively, the entrance and exiting of the roundabout. Switching may occur at any of the entrance/exit points. As before, we randomly generate 100 trajectories.

The two real datasets were acquired at the IST university campus in Lisbon (see Fig. 2 (2nd row)), using static far field cameras. The first dataset (2nd row-left), which we call Stairway, comprises 86 trajectories of pedestrians going up and down a stairway, while the second dataset, called Campus II (2nd row-right), comprises more than 900 trajectories of pedestrians crossing the IST campus. Two cameras were used to acquired the Stairway dataset, while the Campus II dataset was acquired using only one. In the first case, two homographies were performed to i) align the images and ii) correct the distortions cause by the perspective projection. Regarding the Campus II dataset, it was only necessary to correct the distortions. We set the number of motion fields to be estimated to $K = 2$ for the stairway example (up and down), and to $K = 4$ in the Campus II set (roughly north, south, east, and west).

In Table 1 we summarize the properties of each dataset, namely the number of trajectories, the mean trajectory length ($\bar{L}$), and the average displacement between two consecutive trajectory points.

iv) Evaluation - prediction: Finding an appropriate metric to evaluate the role of the initialization and prior is a challenging task. One could assess the error in the estimation of the model parameters, but such metric is suitable only when we know the ground truth values that generated the trajectories. Such information is only available in the case of the synthetic data. Alternatively, we can assess the ability of the estimated models to predict the future of the trajectory, i.e., given a sequence of past positions $(x(1), ..., x(t_0))$ we want to predict the position $\hat{x}(t_0 + \delta)$, δ steps ahead in time using the dynamical model (1).

In our case, this means that we have to predict not only the position, but also the most probable motion field k^* that leads to it (see Fig. 1). To achieve this goal, we apply the forward-backward algorithm [12] used in the E-step of model estimation, namely the forward part, that works as follows. First, lets consider the past, i.e., $t \leq t_0$, and assume the existence of a set of forward variables

$$\alpha_i(t) = Pr\{x(1), ..., x(t), k_t = i|\hat{\theta}\} \ , \tag{11}$$

such that $\alpha(t) = [\alpha_1(t), ..., \alpha_K(t)]$ is a vector that stores all of the forward probabilities of time instant t. For a given past instant, $\alpha(t)$ is estimated as follows:

$$\alpha(t) = D(x_t)B(x_{t-1})^T\alpha(t-1),\ t = 2, ...,\ t_0,\ \alpha(1) = D_1\pi\ , \tag{12}$$

$D(x_t) \in \mathbb{R}^{K\times K}$ is a diagonal matrix with

$$D_{ii} = N\left(x_t; x_{t-1} + T_i(x_{t-1}), \sigma_i^2 I\right),\ i = 1, ..., K\ . \tag{13}$$

For $t > t_0$ the real position $x(t)$ is unknown. Therefore, matrix D can not be computed and the estimation of the forward probabilities $\hat{\alpha}(t)$ is given by:

$$\hat{\alpha}(t) = B(\hat{x}_{t-1})^T\hat{\alpha}(t-1),\ t = t_0 + 1, ..., t_0 + \delta\ , \tag{14}$$

where

$$\hat{x}_t = \hat{x}_{t-1} + T_{k_t^*}(\hat{x}_{t-1}) + 0\ . \tag{15}$$

k_t^* denotes the active field at each instant $t > t_0$, and is defined as: $k^* = \max_k \alpha_k(t)$.

We use the displacement error (DE) as the evaluation metric:

$$DE = \sum_{t_0=1}^{L-\delta} \|x(t_0 + \delta) - \hat{x}(t_0 + \delta)\|_2^2\ , \tag{16}$$

where $\hat{x}(t_0 + \delta)$ is computed by recursively applying (14) to estimate the best field and (15) to predict the position. We set $\delta = 5$, as proposed in other works (e.g., [13]).

5 Results

In Fig. 3 we show the best performances for each dataset (synthetics 1 and 2, Stairway, and Campus II). These results were obtained using leave-one-trajectory-out cross validation and the performances were computed for each pair (*initialization*,*prior*). Here dense means that $\beta = 0$, while l_2 and l_1 correspond to $p = 2, 1$, in (10). Table 2 shows the best overall performances, as well as the combination of model parameters and initialization that led to it. These results demonstrate that the proposed prediction strategy is able to successfully the estimate the position of a pedestrian $\delta = 5$ time steps ahead, achieving a low DE in all datasets.

A more thorough analysis of the results suggest that the initialization method has a significant impact in the synthetic and Stairway datasets, while the performance in the Campus II remains almost unchangeable. The role of the initialization is particularly notorious for synthetic example 2 (the roundabout), where velocity clustering leads to a large increase in the DE. This higher error explained by the existence of a circular movement that is inefficiently modeled by velocity clustering.

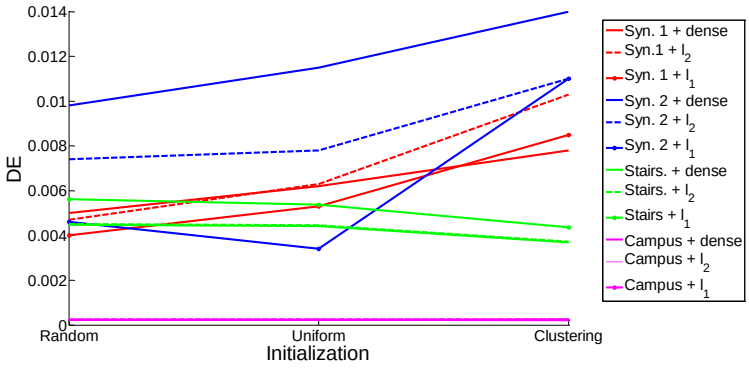

Fig. 3. Experimental results for synthetic examples (red and blue) and real pedestrian trajectories (green and magenta).

Table 2. Best experimental results out of 63 possible configurations for each dataset.

Dataset	DE	**Prior**	**Best Configuration**
Synthetic 1	$5.0 \pm 11.0 \times 10^{-3}$	Dense	Random, $\alpha = 0.5, \beta = 0$
	$4.7 \pm 10.9 \times 10^{-3}$	l_2	Random, $\alpha = 0.5, \beta = 0.1$
	$\mathbf{4.0 \pm 9.0 \times 10^{-3}}$	l_1	**Random**, $\alpha = \mathbf{0.5}, \beta = \mathbf{0.05}$
Synthetic 2	$1.1 \pm 0.80 \times 10^{-2}$	Dense	Random, $\alpha = 1, \beta = 0$
	$7.4 \pm 5.8 \times 10^{-3}$	l_2	Random, $\alpha = 0.5, \beta = 0.15$
	$\mathbf{4.6 \pm 3.0 \times 10^{-3}}$	l_1	**Random**, $\alpha = \mathbf{0.5}, \beta = \mathbf{0.15}$
Stairway	$\mathbf{3.7 \pm 5.8 \times 10^{-3}}$	Dense	**Clustering**, $\alpha = \mathbf{2}, \beta = \mathbf{0}$
	$3.7 \pm 6.4 \times 10^{-3}$	l_2	Clustering, $\alpha = 0.5, \beta = 0.05$
	$4.4 \pm 8.4 \times 10^{-3}$	l_1	Clustering, $\alpha = 1, \beta = 0.05$
Campus II	$\mathbf{2.2 \pm 0.5 \times 10^{-4}}$	Dense	**Clustering**, $\alpha = \mathbf{2}, \beta = \mathbf{0}$
	$2.4 \pm 0.6 \times 10^{-4}$	l_2	Clustering, $\alpha = 0.5, \beta = 0.15$
	$2.4 \pm 0.5 \times 10^{-4}$	l_1	Random, $\alpha = 0.5, \beta = 0.05$

Enforcing sparsity in the estimation of the fields ($p = 1$) leads to improved prediction results for the synthetic data. However, the opposite is observed for the real datasets, where dense fields achieve the best results. Such disparity may be due to the proportionally higher variability that exists in pedestrian velocities, when compared with the synthetic datasets (see Table 1). Setting $p = 2$ in (10) has some impact in the performance, but it is clearly outperformed by the other priors. Regarding the prior regularization constants, α and β, it seems that setting $\alpha = 0.5$ and $\beta = 0.05$ is the preferred combination for the sparse prior in three out of the four datasets. In the case of the dense prior, we observe that setting α to a higher value leads to the best results for the real datasets, while no trend is observed for the synthetic ones.

Conclusions

Efficiently modeling the trajectories of pedestrians is critical for most surveillance setups. However, several of the models proposed to tackle this issue rely on complex methods that require the tunning of multiple parameters. In this paper we addressed this issue, taking as starting point a motion model based on switching motion fields.

Our contributions are two-folded. First we proposed a trajectory prediction strategy for the aforementioned method, which allows a qualitative evaluation of the method. Second, we conducted a thorough study of the model parameters, namely initialization, velocity priors, and constants, in order to obtain insights of their role. Both contributions are original of this paper. Experiments were conducted in four datasets and the obtained results are relevant and promising. Future work should rely on exploring the applicability of the motion fields to other tasks, such as the detection of abnormal trajectories or activities.

6 Acknowledgments

This work was supported by the FCT project and plurianual funding: [PTDC/EEIPRO/0426/2014], [UID/EEA/50009/2013].

References

1. J. Aggarwal and M. Ryoo, "Human activity analysis: A review," ACM Comput. Surv., vol. 43, no. 3, p. 16, Apr. 2011.
2. K. Kim, D. Lee, and I. Essa, "Gaussian process regression flow for analysis of motion trajectories," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2011.
3. David Ellis, Eric Sommerlade and Ian Reid, "Modelling Pedestrian Trajectory Patterns with Gaussian Processes", in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV) Workshops*, 2009.
4. V. Bastani, L. Marcenaro and C. S. Regazzoni, "Online Nonparametric Bayesian Activity Mining and Analysis From Surveillance Video", *IEEE Trans. on Imag. Proc.*, vol. 25, no. 5, pp. 2089-2102, 2016
5. N. Ferreira, J. T. Klosowski, C. E. Scheidegger and C. T. Silva, "Vector Field k-Means: Clustering Trajectories by Fitting Multiple Vector Fields", Eurographics Conference on Visualization, vol. 32 no. 3, pp. 201-210, 2013.
6. V. Bastani, L. Marcenaro, and C. Regazzoni, "Unsupervised trajectory pattern classification using hierarchical Dirichlet process mixture hidden Markov model," in Proc. IEEE Int. Workshop Mach. Learn. Signal Process. (MLSP), pp. 16, 2014
7. W. Hu, X. Li, G. Tian, S. Maybank, and Z. Zhang, "An incremental DPMM-based method for trajectory clustering, modeling, and retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 5, pp. 10511065, 2013.
8. D. Lin, "Online learning of nonparametric mixture models via sequential variational approximation," in Advances in Neural Information Processing Systems, pp. 395403, 2013.

9. J. C. Nascimento, M. A. T. Figueiredo, J. S. Marques. “Activity recognition using mixture of vector fields”, IEEE Trans. on Image Processing, vol. 22, no. 5, pp. 1712 - 1725, 2013.
10. C. Barata, J. C. Nascimento and J. S. Marques “A Sparse Approach to Pedestrian Trajectory Modeling using Multiple Motion Fields”, Int. Conf. on Image processing, 2017
11. M. A. T. Figueiredo, A. K. Jain “Unsupervised learning of finite mixture models”, IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no. 3, pp. 381-396, 2002
12. L. R. Rabiner “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”, in Proceedings of the IEEE, vol. 77, no. 2, pp. 257-286, 1989
13. A. Alahi, V. Ramanathan, K. Goel, A. Robicquet, A. A. Sadeghian, L. Fei-Fei, and S. Savarese, Learning to predict human behaviour in crowded scenes, Group and Crowd Behavior for Computer Vision, pp. 183207, 2017.