This is the peer reviewd version of the followng article:

Learning People Trajectories using Semi-directional Statistics / Calderara, Simone; Prati, Andrea; Cucchiara, Rita. - STAMPA. - (2009), pp. 213-218. (Intervento presentato al convegno Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance tenutosi a Genoa, Italy nel 2-4 September 2009) [10.1109/AVSS.2009.34].

IEEE
Terms of use:

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

12/10/2024 04:14

(Article begins on next page)

Learning People Trajectories using Semi-directional Statistics

Simone Calderara[1], Andrea Prati[2], Rita Cucchiara[1]
[1]*D.I.I. - Univ. of Modena and Reggio Emilia - Via Vignolese, 905 - Modena - Italy*
[2]*Di.S.M.I. - Univ. of Modena and Reggio Emilia - Via Amendola, 2 - Reggio E. - Italy*

***Abstract*—This paper proposes a system for people trajectory shape analysis by exploiting a statistical approach which accounts for sequences of both directional (the directions of the trajectory) and linear (the speeds) data. A semi-directional distribution (AWLG - Approximated Wrapped and Linear Gaussian) is used with a mixture to find main directions and speeds. A variational version of the mutual information criterion is proposed to prove the statistical dependency of the data. Then, in order to compare data sequences, we define an inexact method with a Kullback-Leibler-based distance measure and employ a global alignment technique is to handle sequences of different lengths and with local shifts or deformations. A comprehensive analysis of variable dependency and parameter estimation techniques are reported and evaluated on both synthetic and real data sets.**

***Keywords*-Semi-directional statistics; trajectory analysis; video-surveillance**

I. Introduction and Related Works

Recently, one of the most addressed topics in video surveillance research is the extraction and analysis of features for behavior understanding. Among the possible features, trajectories represent a rich source of information which can be robustly extracted from single or multiple fixed cameras [1]. Morris and Trivedi in [2] proposed a recent survey on state-of-art techniques for modeling, comparing and classifying trajectories for video surveillance purposes.

In most of the approaches, the trajectories are modeled as a sequence of $< x, y >$ coordinates on the ground plane and different statistic-based inferences are exploited. For instance, in [3] the EM framework is adopted to learn a set of motion models' parameters; Panozzo *et al.*in [4] suitably modified the iterative *Altruistic Vector Quantization* algorithm to robustly cluster trajectories with representative prototypes; in [5], Junejo et al. applied graph cuts to cluster trajectories using the Hausdorff distance; in [6] a system for learning statistical motion patterns is presented using a two-stage fuzzy k-means: first, trajectories are clustered in the spatial domain, then each cluster is subclustered in the temporal domain and new trajectories are classified using Bayes rule; finally anomaly detection is performed by thresholding resulting probability. Accordingly, also in this work we aim to propose an unsupervised classification of trajectories in order to find clusters of similar frequent trajectories, and a second step of classification against the clusters to detect possible abnormal or infrequent trajectories. The main novelty is the analysis of *trajectory shape* instead of *trajectory points*. In this manner we search for similarity of movements of people in large open space, invariant to the spatial position and thus the starting point. For a human observer, people that are walking straight have the same behavior even if their path is more at right or at left side of a corridor, or people staying together and moving in circle have another similar behavior.

To describe trajectory shape, the sequence of directions between two consecutive tracking samples can be adopted [7]), and the obtained representation is independent on the starting point. The use of a sequence of directions to model the trajectories requires the adoption of the so-called *circular* or *directional statistics* [8] since modelling periodic variables, such as angles, by using linear statistics is inappropriate. Since often pure shape is not sufficiently discriminative in surveillance scenarios, (e.g. the same path covered by a walk or by a run has a different meaning in terms of behavior) in this work we propose to add the speed to the shape description to provide a more complete analysis of the trajectory. The introduction of the speed, which is not periodic, requires to account for the different nature of these features, the angle θ, directional, and the speed v, linear. Using a statistical model, the resulting bivariate joint probability $p(\theta, v)$ can be easily modelled as the product $p(\theta) \cdot p(v)$ if and only if the two variables result to be independent for the considered application. If they are not, the joint probability must be modelled by using a directional (univariate) pdf for θ and a linear (univariate) pdf for v. The estimation of the covariance matrix for this bivariate joint pdf can be quite challenging since the dependency between θ and v must be modelled properly. When a directional or periodic variable is combined with a linear one the term *semi-directional* is often used.

The use of a Gaussian pdf for the linear variable v is straightforward, while the choice of the pdf for θ is less obvious. One of the most used (due to the properties it shares with the Gaussian) is the *von Mises* (vM) distribution [9], also known as *circular normal*. However, in the case of semi-directional statistics, the use of a *wrapped Gaussian* [8] distribution is preferable because, due to the closeness to its linear counterpart, it is possible to adopt a linear approximation of the variance parameter even for circular variables. The linear variance approximation allows the employment of the Gaussian maximum likelihood estimator to calculate, with a feasible precision, the covariance matrix

in the case of joint linear and periodic multivariate variables.

Nevertheless, parameter estimation in the case of wrapped Gaussian is not easy. For this reason, Bahlmann [10] proposed to adopt a multivariate semi-directional distribution in handwriting recognition by using an *approximated wrapped Gaussian* (AWG) pdf for the directional variable (the tangent slope of a written segment) and the use of a linear Gaussian for the linear variable, by defining a *semi-wrapped Gaussian* distribution which we will refer to hereinafter as *AWLG* (Approximated Wrapped and Linear Gaussian). Eventually, both directional and linear data can be modelled with multi-modal distributions, for example using parametric mixtures of the corresponding pdfs.

Starting from these premises, this paper presents a new proposal to use semi-directional statistics (specifically a mixture of AWLG) to model and analyze people trajectory shapes in order to classify paths shape and motion models. The AWLG model results to be the more appropriate since we present results of a method based on mutual information [11] for testing the dependency between the directional and linear variables. Since exact mutual information is hard to compute in the case of mixtures of pdfs, a variational approximation of it is derived. Finally, an approach for comparing sequences of semi-directional data is provided: it exploits the global alignment of sequences of symbols with a distance based on Kullback-Leibler divergence. Finally, a complete system for the classification of people trajectories is proposed and experiments on both synthetic and real data are provided to demonstrate its accuracy. Some hours of unconstrained acquisition of people walking around in an open space are evaluated.

II. Semi-directional Distributions

Among the possible pdfs from directional statistics to model angles, the *wrapped Gaussian* (WG) and its approximation, named *approximated wrapped Gaussian* (AWG), are two of the most used since they derive from the Gaussian and can be thus easily treated. Both these distributions start from the wrapping of a Gaussian $\mathcal{N}(x)$ obtained tiling different pdfs shifted by multiple of 2π and with the variable change $\theta = x \bmod 2\pi$. In the case the linear pdf is a univariate Gaussian distribution, the corresponding wrapped Gaussian (WG) can be written as:

$$WG\left(\theta|\theta_0,\sigma\right) = \sum_{w=-\infty}^{+\infty} \mathcal{N}\left(\theta - w2\pi|\theta_0,\sigma\right) \tag{1}$$

$WG\left(\theta|\theta_0,\sigma\right)$ is unimodal with a single local maximum and symmetric about θ_0.

Due to the presence of the summation in equation (1), WG parameters are hardly estimated, especially when a mixture of WG must be considered and possibly estimated with a EM algorithm. Bahlmann [10] proposed an approximation of the wrapped Gaussian by only one, but the most meaningful wrap of it. He demonstrated that, in the presence of sufficiently small variances ($\sigma \lesssim 1$), the sum of wrapped distributions can be approximated with a single periodic distribution since the overlap of neighboring Gaussian wraps is negligible.

The proposed AWG pdf results as:

$$AWG\left(\theta|\theta_0,\sigma\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{\left((\theta-\theta_0) \bmod 2\pi\right)^2}{2\sigma^2}} \tag{2}$$

where the operation "mod" represents the remainder of the division and resulting angle must be in the interval $(-\pi,\pi]$. Being the input angle θ is in the interval $(-\pi,\pi]$, also the "mod" operator maps in the same interval.

The mixture of AWG (MoAWG) is, as usual, obtained as a weighted sum of AWG pdfs:

$$MoAWG\left(\theta|\boldsymbol{\pi},\boldsymbol{\theta_0},\boldsymbol{\sigma}\right) = \sum_{k=1}^{K} \pi_k \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{\left(\left(\theta-\theta_{0,k}\right) \bmod 2\pi\right)^2}{2\sigma_k^2}} \tag{3}$$

The formulation of the EM algorithm for MoAWG is not simple since the mod operator is not derivable. However, thanks to the use of a mathematical artifice, the following update equations can be derived:

$$\theta_{0,k} = \arctan\left(\frac{\sum_{i=1}^{n} \gamma_{ik} \sin\theta_i}{\sum_{i=1}^{n} \gamma_{ik} \cos\theta_i}\right) \tag{4}$$

$$\sigma_k^2 = \frac{\sum_{i=1}^{n} \gamma_{ik} \left(\left(\theta_i - \theta_{0,k}\right) \bmod 2\pi\right)^2}{\sum_{i=1}^{n} \gamma_{ik}} \tag{5}$$

Please note that the update equation for the responsabilities is omitted because it is the same than other mixtures. The complete derivation of these equations can be obtained as a special case of the EM algorithm for MoAWLG reported in the supplementary material.

When dealing with semi-directional data, a joint pdf must be found in order to account, if existing, also for the mutual dependency between directional and linear variables. The pdf is modelled with a AWLG, defined as:

$$AWLG\left(X|\mu,\Sigma\right) = \frac{1}{\sqrt{2\pi \cdot |\Sigma|}} e^{-\frac{1}{2}(X-\mu)^T\Sigma^{-1}(X-\mu)} \tag{6}$$

Here $X = \begin{bmatrix}\theta \\ v\end{bmatrix}$ is the observation, $\mu = \begin{bmatrix}\theta_0 \\ v_0\end{bmatrix}$ is the mean, $X - \mu = \begin{bmatrix}(\theta - \theta_0) \bmod 2\pi \\ v - v_0\end{bmatrix}$ is the "difference" between them, $\Sigma = \begin{bmatrix}\sigma_{\theta,\theta} & \sigma_{\theta,v} \\ \sigma_{v,\theta} & \sigma_{v,v}\end{bmatrix}$ is the covariance matrix and $|\Sigma|$ denotes its determinant. This expression is similar to a bivariate Gaussian pdf, but one of the two dimensions is represented by a AWG.

The corresponding mixture can be written as:

$$MoAWLG\left(X|\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\Sigma}\right)=\sum_{k=1}^{K}\pi_k AWLG\left(X|\mu_k,\Sigma_k\right) \tag{7}$$

At the best of authors' knowledge, a complete EM for MoAWLG has not been proposed yet. The followin equations can be derived:

$$\theta_{0,k}=\arctan\left(\frac{\sum\limits_{i=1}^{n}\gamma_{ik}\sin\theta_i}{\sum\limits_{i=1}^{n}\gamma_{ik}\cos\theta_i}\right) v_{0,k}=\frac{\sum\limits_{i=1}^{n}\gamma_{ik}v_i}{\sum\limits_{i=1}^{n}\gamma_{ik}}$$
$$\Sigma_k=\frac{\sum\limits_{i=1}^{n}\gamma_{ik}\left(X_i-\mu_k\right)^T\left(X_i-\mu_k\right)}{\sum\limits_{i=1}^{n}\gamma_{ik}} \tag{8}$$

For the sake of readability and page limits, the complete derivation is reported in the supplementary material.

III. Verifying Independence

In summary, two possible approaches can be used to model people trajectories as sequences of semi-directional data. The first approach should be chosen whenever the direction θ and the speed v are independent and the probability $p\left(\theta,v\right)$ is computed as the product of a circular pdf by a linear pdf. For the considerations reported in Section II, a mixture of AWG (MoAWG) distributions can be used for θ, while a standard mixture of Gaussians for v. Conversely, in the case of dependent variables, the MoAWLG described above can be used for modelling $p\left(\theta,v\right)$.

In order to evaluate the degree of dependency of θ and v, the *mutual information* I of $p\left(\theta,v\right)$ must be computed [11]:

$$\begin{aligned}I\left[\theta,v\right]&\equiv KL\left(p\left(\theta,v\right)\|p\left(\theta\right)p\left(v\right)\right)\\&=-\int\int p\left(\theta,v\right)\ln\left(\frac{p\left(\theta\right)p\left(v\right)}{p\left(\theta,v\right)}\right)dxdy\end{aligned} \tag{9}$$

where KL is the Kullback-Leibler divergence.

Unfortunately, in our case $p\left(\theta,v\right)$ is a mixture, as well as the numerator of the log in eq. (9) is the product of two mixtures, thus the KL divergence is not analytically tractable, nor does any efficient computational algorithm exist. However, many approximated approaches have been proposed [12]. Among the approaches proposed, the most suitable (since it preserves the majority of the KL properties) is based on the variational computation of the lower bound of the log likelihood. In fact, given two generic mixtures $f(x)=\sum\limits_a\pi_a f_a(x|P_a)$ and $g(x)=\sum\limits_b\omega_b g_b(x|P_b)$, with f_a and g_b the unimodal distributions and P_a and P_b their corresponding parameters, the KL divergence can be written as:

$$\begin{aligned}KL(f|g)&=\int f(x)\ln f(x)dx-\int f(x)\ln g(x)dx\\&=\mathbb{E}_{f(x)}\left[\ln f(x)\right]-\mathbb{E}_{f(x)}\left[\ln g(x)\right]\end{aligned} \tag{10}$$

Being $\mathcal{L}_f\left(g\right)\equiv\mathbb{E}_{f(x)}\left[\ln g(x)\right]=\mathbb{E}_{f(x)}\left[\ln\sum\limits_b\omega_b g_b(x)\right]$ the log likelihood, we can use a variational approach to find a (more tractable) lower bound. Introducing the variational parameters $\phi_{b|a}>0$ such that $\sum\limits_b\phi_{b|a}=1$, we obtain the following equations:

$$\begin{aligned}\mathcal{L}_f\left(g\right)&\equiv\mathbb{E}_{f(x)}\left[\ln\sum_b\omega_b g_b(x)\right]\\&\geq\mathbb{E}_{f(x)}\left[\sum_b\phi_{b|a}\ln\frac{\omega_b g_b(x)}{\phi_{b|a}}\right]\equiv\mathcal{L}_f\left(g,\phi\right)\end{aligned} \tag{11}$$

Since the variational log likelihood $\mathcal{L}_f\left(g,\phi\right)$ is a lower bound of the log likelihood $\mathcal{L}_f\left(g\right)$, the best bound can be found by maximizing it with respect to ϕ. Thus, the following equation [12] can be derived as:

$$\hat{\phi}_{b|a}=\frac{\omega_b e^{-KL(f_a|g_b)}}{\sum\limits_{b'}\pi_{b'}e^{-KL(f_a|g_{b'})}} \tag{12}$$

Applying the same procedure for $\mathcal{L}_f\left(f\right)$, the KL divergence of eq. (10) can be approximated by the following variational KL divergence:

$$KL_{var}\left(f|g\right)=\sum_a\pi_a\ln\frac{\sum\limits_{a'}\pi_{a'}e^{-KL(f_a|f_{a'})}}{\sum\limits_b\omega_b e^{-KL(f_a|g_b)}} \tag{13}$$

Please note that the KL values reported in eq. (13) are now KL divergences between single unimodal distributions (either f_a or g_b), which can be treated analytically. In our specific case, the joint probability (denominator in eq. (9) and $f(x)$ in eq. (10)) is a MoAWLG and, thus, f_a is a AWLG, while the numerator of eq. (9) is the product of a MoAWG and a MoG, which is easily demonstrated to correspond to a mixture again with diagonal covariance, and thus g_b results to be the product of a AWG by a Gaussian. In other words, both the AWLG and the product of AWG by Gaussian can be assimilated to bivariate Gaussians (with full or diagonal covariance, respectively); hence, the KL divergences of eq. (13) can be brought back to the KL between multivariate normal distributions.

An analytical formulation of the KL divergence between normal distributions has been reported in [12]. Given $AWLG\left(X|\mu_1,\Sigma_1\right)$ and $AWG\left(\theta|\theta_{0,2},\sigma_\theta\right)\cdot\mathcal{N}\left(v|v_{0,2},\sigma_v\right)$, their KL divergence is equal to:

$$\begin{aligned}KL\left(AWLG|AWG\cdot\mathcal{N}\right)&=\frac{1}{2}\ln\frac{|\Sigma_2|}{|\Sigma_1|}-\frac{N_1}{2}+\\&+\frac{1}{2}\mathrm{tr}\left(\Sigma_2^{-1}\Sigma_1\right)+\frac{1}{2}\left(\left(\mu_1-\mu_2\right)^T\Sigma_2^{-1}\left(\mu_1-\mu_2\right)\right)\end{aligned} \tag{14}$$

where N_1 is the dimension of Σ_1, $\mu_1 = \begin{bmatrix}\theta_{0,1}\\ v_{0,1}\end{bmatrix}$, $\mu_2 = \begin{bmatrix}\theta_{0,2}\\ v_{0,2}\end{bmatrix}$ and $\Sigma_2 = \begin{bmatrix}\sigma_\theta & 0\\ 0 & \sigma_v\end{bmatrix}$.

IV. Comparing Sequences

A sequence T_j of n_j semi-directional data (θ and v) can be written as: $T_j = \left\{\langle\theta_{1,j}, v_{1,j}\rangle, \cdots, \langle\theta_{n_j,j}, v_{n_j,j}\rangle\right\}$. In order to cluster or classify these sequences, a method for comparing them by means of a suitable similarity or distance measure is mandatory.

In order to compute the distance, we first need to decide which statistical model is more appropriate for the data at hand. Thus, the first step is to evaluate the dependency between θ_i and v_i in our specific people video surveillance application. Supposing that a training set of people trajectories is available and that it is sufficiently representative of the scenario, the formulas (9) and (13) reported in Section III can be employed for this purpose.

Once the dependency has been evaluated, the EM algorithm of the corresponding distribution(s) can be run on each trajectory of the training set to learn the distribution's parameters. In the case of dependent variables, EM algorithm for MoAWLG (see supplementary material) should be used, whilst in the case of independent variables, two separate EM algorithms for MoAWG and MoG should be employed for θ_i and v_i, respectively.

After obtaining the main components of the associated distribution(s), it is possible to encode each data $\langle\theta_i, v_i\rangle$ with a symbol S_i with a Maximum-A-Posteriori (MAP) approach, which associates a single value to the most similar component of the mixture distribution representing the data. With this MAP approach each trajectory T is encoded with a sequence of symbols $\overline{T} = \{S_1, S_2, ..., S_n\}$. In the case of dependent variables, S_i corresponds to a AWLG component ($S_i \leftrightarrow \langle(\mu_r, \Sigma_r)\rangle$ as in eq. (6)), where μ_r and Σ_r are the parameters of the rth components of the MoAWLG that maximizes the posteriori. Conversely, in the case of independent variables, two separate MAP evaluations must be carried out, one to find the r_1^{th} component of the MoAWG that maximizes the posteriori on the angle θ, and one for the r_2^{th} MoG component for the speed v ($S_i \leftrightarrow \langle(\theta_{0,r_1}, \sigma_{\theta,r_1}), (v_{0,r_2}, \sigma_{v,r_2})\rangle$).

This representation allows to compare sequences if a suitable measure $\Omega\left(\overline{T}_i, \overline{T}_j\right)$ is defined. Indeed, exact matching between sequences of data or symbols which correspond to a statistical distribution is not suitable due to acquisition noise, uncertainty and temporal shifts. Additionally, trajectories can have different lengths and thus the similarity between the elements of the sequences must be carried out after the alignment and warping of the sequences.

Among the various techniques, we can borrow from bioinformatics the method for comparing DNA sequences in order to find the best inexact matching between them, also accounting for gaps. Therefore, we propose to adopt the *global alignment*, specifically the well-known Needleman-Wunsch algorithm [13]. A global alignment (over the entire sequence) is preferable (instead of a local one) because the former preserves both global and local shape characteristics. Alternative inexact matching techniques, such as *Dynamic Time Warping* (DTW) or *Longest Common SubSequence* (LCSS), have been applied for comparing trajectories, in different contexts, such as trajectory analysis [14], speech [15] or handwriting recognition [16]. Moreover, Porikli [17] proposes to use a HMM-based similarity measure where each trajectory is modelled with a HMM and compared using the cross likelihood. Though this is an interesting approach, a large amount of data is needed to avoid overfitting in the HMM training phase.

In order to use the global alignment, a cost of the alignment between two symbols must be defined. Considering the two cases listed above, the cost $C(S_i, S_j)$ can be written as:

$$C(S_i, S_j) = \begin{cases} KL_1 & \text{if dependent} \\ KL_2 \cdot KL_3 & \text{if independent} \end{cases} \tag{15}$$

with:

$$\begin{aligned} KL_1 &= KL\left(AWLG\left(X|\mu_{r_i}, \Sigma_{r_i}\right)|AWLG\left(X|\mu_{r_j}, \Sigma_{r_j}\right)\right) \\ KL_2 &= KL\left(AWG\left(\theta|\theta_{0,r_i}, \sigma_{\theta,r_i}\right)|AWG\left(\theta|\theta_{0,r_j}, \sigma_{\theta,r_j}\right)\right) \\ KL_3 &= KL\left(\mathcal{N}\left(v|v_{0,r_i}, \sigma_{v,r_i}\right)|\mathcal{N}\left(v|v_{0,r_j}, \sigma_{v,r_j}\right)\right) \end{aligned} \tag{16}$$

where r_i and r_j are the selected components for symbols S_i and S_j, respectively.

The KL divergences reported in eq. (16) can be computed with the formula in eq. (14) considering that the AWG can be assimilated to a univariate Gaussian and the formula can be computed with σ instead of the covariance matrix Σ.

To verify the proposed approach as a classifier, we build a complete system for trajectory classification. During an initial off-line training step, the training set of trajectories is analyzed and clustered into groups of similarity. The choice of the clustering algorithm is not critical: we need only a method without a fixed number of clusters and capable to create clusters of different cardinality, even containing a single sample. This allows us to not model the classes and their number a-priori, achieving a higher flexibility. Our tests proved that a satisfactory algorithm for clustering trajectory shapes is the *k-medoids* algorithm [18]. This is a suitable modification of the well-known k-means algorithm which has the appreciable characteristic to compute, as prototype of the cluster, the medoid instead of the mean, i.e. the element that minimizes the sum of intra-class distances. In other words, at each iteration the prototype of each cluster is given by the member of the cluster at the minimum average distance from all the other members. Moreover, in order to overcome the limitation of k-medoids (as well as k-means)

clustering to have a fixed number k of clusters, we use its iterative version [7]. The algorithm iteratively merges similar clusters until convergence to obtain the "best" number $\widetilde{k}$ of clusters.

Then, during the on-line phase, whenever a new trajectory T_{new} is collected, its statistical model is computed and compared to the cluster medoids. Based on this comparison, the trajectory can be assigned to the class with the maximum a-posteriori probability. However, more sophisticated classification strategies, such as the *learn-and-predict* approach proposed in [7], could be employed with a minimum effort.

V. Experimental results

In order to verify the accuracy of our approach, we performed extensive experiments with both synthetic and real data. The use of synthetic data is particularly useful since it enables the collection and evaluation of a great amount of data together with their automatic ground truths. In order to analyze trajectories, we specifically developed a trajectory generator in Matlab that produces trajectories similar to the ones collected with a real video-surveillance system. The generator allows the user to select graphically the sequence of angles composing the class of trajectories and the corresponding speeds (by making the length of the segment almost proportional to the speed) and then creates a set of trajectories with added noise to the angles, the speeds, and their occurrences. Additionally, the simulator allows to shuffle the data to remove dependency between angles and speeds, in order to provide both dependent and independent data. Real data are obtained by a multi-camera video surveillance system working on our campus, which exploit multiple cameras with overlapped views to get more robust and longer trajectories [1]. Trajectories are projected on the ground plane and long trajectories are collected thanks to the multiple camera setup.

Two sets of synthetic trajectories (one with dependent and one with independent data) have been generated with the Matlab simulator in order to evaluate both the solutions (with dependent and independent variables). We also evaluated the average mutual information on real data. The average value for real data is high enough to mean that some correlation exists between the angles and the speed in the considered scenario. This is not true in general, but it heavily depends on the context of application and the collected data.

The robustness of the proposed sequences' comparison algorithm has been tested performing two different kinds of experimental campaigns. The first campaign evaluates the performance of our approach in some specific situations, where common approaches tend to fail. First, the robustness against small fluctuations around the zero value of θ (row "Periodicity" of Table I) has been evaluated by generating trajectories that are composed by a unique straight almost-zero direction with added noise. In this case, the system is

	Type	Acc. w AWLG (dep)	Acc. w AWG/G (indep)
Periodicity on θ	dep	100%	100%
Sequence on θ	dep	100%	90.3%
Sequence on v	dep	100%	92.5%
Noise on θ	dep	100%	87.5%
Noise on v	dep	93.5%	87.3%
Classification *(230 tr./11 cl.)*	dep	99.6%	97.8%
Classification *(100 tr./10 cl.)*	indep	90.0%	95.0%
Classification *(356 tr./18 cl.)*	real	93.4%	90.3%

Table I
Experimental results in terms of accuracy in the classification. *tr.*=number of trajectories, *cl.*=number of classes.

able to cluster together all the trajectories thanks to the use of circular (i.e. wrapped) statistics to model angular data.

Subsequently, we tested the capability of the system to handle sequences of either the same principal directions or speeds, but given in different order (rows "Sequence" of Table I). In this case, both the proposed statistical measure and the alignment technique concur to filter out the noise and to correctly cluster this kind of data. Then, a specific test was performed to verify the robustness against severe noise on either angular or speed values (rows "Noise").

The second test campaign evaluates the accuracy of the proposed approach performing sequences' classification on a large amount of data. Synthetic and real data are used for testing, and two synthetic sets are provided with either dependent or independent data (rows 6 and 7 of Table I). The real test (row 8) is composed by 356 trajectories collected by the system previously mentioned and manually ground-truthed. The set of trajectories has been divided randomly in 200 trajectories in the training set and the remaining in the testing set. Examples of the obtained classes (superimposed to a bird-eye view of the multiple camera scenario) are shown in Fig. 1. Please note that trajectories of the same color belong to the same class.

By observing the results in Table I, it is evident that both the proposed solutions achieve high accuracies in most of the tests. It is also worth mentioning that in the classification test with independent-generated data (row 7) the independent model better describes the sequences of data. Eventually, the results on the real data set confirm the slightly-dependent nature of them, in accordance with the consideration reported above on the mutual information. It is remarkable that most of the classification errors are due to not perfect convergence of the EM algorithms and can be potentially avoided by a more sophisticated parameter estimation technique.

In conclusion, in this paper a semi-directional distribution is proposed to jointly model angular and linear data in

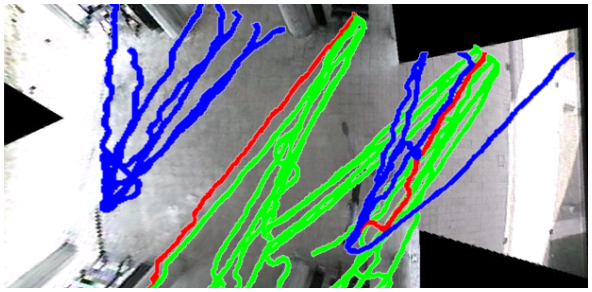

Figure 1. Examples of the achieved clustering results.

order to classify people trajectories for video surveillance applications. Two different models are proposed to account the cases where the input data are either statistically dependent or independent. It is remarkable to see how linear and angular variables can be treated and compared jointly making the proposed framework particularly useful also in other contexts, where sequences of multidimensional and hybrid data must be handled.

REFERENCES

[1] S. Calderara, R. Cucchiara, and A. Prati, "Bayesian-competitive Consistent Labeling for People Surveillance," *IEEE Trans. on PAMI*, vol. 30, no. 2, pp. 354–360, Feb. 2008.

[2] B. Morris and M. Trivedi, "A survey of vision-based trajectory learning and analysis for surveillance," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 8, pp. 1114–1127, Aug. 2008.

[3] M. Bennewitz, W. Burgard, and S. Thrun, "Using em to learn motion behaviors of persons with mobile robots," in *IEEE/RSJ International Conference on Intelligent Robots and System*, vol. 1, 2002, pp. 502– 507.

[4] A. Mecocci and M. Pannozzo, "A completely autonomous system that learns anomalous movements in advanced video-surveillance applications," in *IEEE International Conference on Image Processing*, vol. 2, 2005, pp. 586–589.

[5] I. Junejo, O. Javed, and M. Shah, "Multi feature path modeling for video surveillance," in *Proceedings of the 17th International Conference on Pattern Recognition*, vol. 2, Aug. 2004, pp. 716– 719.

[6] W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, and S. Maybank, "A system for learning statistical motion patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1450–1464, September 2006.

[7] A. Prati, S. Calderara, and R. Cucchiara, "Using circular statistics for trajectory analysis," in *Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR 2008)*, Jun. 2008.

[8] K. Mardia, *Statistics of directional data*, 1972.

[9] R. Fisher, "Dispersion on a sphere," *Proc. Roy. Soc. London Ser. A.*, vol. 217, pp. 295–305, 1953.

[10] C. Bahlmann, "Directional features in online handwriting recognition," *Pattern Recognition*, vol. 39, pp. 115–125, 2006.

[11] C. B. Bell", "Mutual information and maximal correlation as measures of dependence," *Ann. Math. Statist.*, vol. 33, pp. 587–595, 1962.

[12] J. R. Hershey and P. A. Olsen, "Approximating the kullback leibler divergence between gaussian mixture models," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 4, 2007, pp. IV–317–IV–320.

[13] S. Needleman and C. Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins," *Journ. of Molec.Biol.*, vol. 48, no. 3, pp. 443–453, 1970.

[14] D. Buzan, S. Sclaroff, and K. G., "Extraction and clustering of motion trajectories in video," in *ICPR*, vol. 2, 2004.

[15] T. Dutta, "Dynamic time warping based approach to text-dependent speaker identification using spectrograms," in *Proc. of Image and Signal Processing Congress CISP*, 2008, pp. 2: 354–360.

[16] Y. Qiao and M. Yasuhara, "Affine invariant dynamic time warping and its application to online rotated handwriting recognition," in *ICPR*, 2006, pp. 2: 905–908.

[17] F. Porikli and T. Haga, "Event detection by eigenvector decomposition using object and frame features," in *Proc. of CVPR Workshop*, vol. 7, 2004, pp. 114–121.

[18] A. Reynolds, G. Richards, and V. Rayward-Smith, *The Application of K-Medoids and PAM to the Clustering of Rules*. Springer Berlin / Heidelberg, vol. 3177/2004, pp. 173–178.