Electronic Letters on Computer Vision and Image Analysis 5(3):148-156, 2005

Trajectory Analysis for Sport and Video Surveillance

Y. L. de Meneses*, P. Roduit* +, F. Luisier, J. Jacot*

* *Laboratoire de Production Microtechnique (LPM-IPR), Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland*
+ *Swarm-Intelligent System Group (SWIS), Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland*
Received 22 December 2004; accepted 27 July 2005

Abstract

In video surveillance and sports analysis applications, object trajectories offer the possibility of extracting rich information on the underlying behavior of the moving targets. To this end we introduce an extension of Point Distribution Models (PDM) to analyze the object motion in their spatial, temporal and spatiotemporal dimensions. These trajectory models represent object paths as an average trajectory and a set of deformation modes, in the spatial, temporal and spatiotemporal domains. Thus any given motion can be expressed in terms of its modes, which in turn can be ascribed to a particular behavior.

The proposed analysis tool has been tested on motion data extracted from a vision system that was tracking radio-guided cars running inside a circuit. This affords an easier interpretation of results, because the shortest lap provides a reference behavior. Besides showing an actual analysis we discuss how to normalize trajectories to have a meaningful analysis.

Key Words: Trajectory Analysis, Tracking, Point Distribution Models, Deformable Templates, Behavior Analysis, Performance, Eigenshapes.

1 Introduction

Object tracking at frame-rate on standard desktop computers has been rendered possible by faster and cheaper hardware (cameras, frame-grabbers, processors). This has sprung up many applications in video surveillance [1], sports analysis [2], human-machine interfaces [3], robotics [4], and ethology [5] [6]. Since tracking data is now relatively easy to acquire [7], it is necessary to process it and to extract meaningful information for higher-level tasks such as behavior analysis or sportsmen performance evaluation.

This project intends to develop new analysis tools to characterize the gestures or motions of sportsmen by means of their trajectories or the trajectories of their extremities. To this end, we intend to use deformable templates to build meaningful models of a set of trajectories. These deformable models should allow to describe variations around an average or reference trajectory. Point Distribution Models (PDMs) [8] are one kind of deformable templates of particular interest because of their statistical meaning and simple mathematical expression. They have been applied to model object shapes. However a trajectory is more than a geometric shape; it comprises also temporal information that can be of key importance in some analyses. This paper proposes

Correspondence to: <pierre.roduit@epfl.ch>

Recommended for acceptance by <Perales F. and Draper B.>
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

an adaptation of Point Distribution Models to analyze trajectories in their spatial, temporal and spatiotemporal dimensions.

This paper is organized as follows. In the next section, Point Distribution Models will be succinctly presented together with the adaptations introduced to model temporal information. Section 3 will describe the experiments we have conducted with trajectories from radio-controlled cars and a discussion will close the paper.

2 Point Distribution Models for trajectories

A trajectory can be represented as an ordered set of points π in space and time. Without loss of generality we shall consider trajectories lying on a plane and hence, given a set of trajectories, the i^{th} point in the k^{th} trajectory can be written as

$$\pi_i^k = \left[x_i^k \; y_i^k \; t_i^k\right]^T, \tag{1}$$

where x_i^k, y_i^k are the spatial coordinates and t_i^k is in the temporal dimension. The k^{th} trajectory, τ_k, made up of N points $\pi_1^k \ldots \pi_N^k$, is then written as

$$\tau_k = \left[x_1^k \ldots x_N^k \; y_1^k \ldots y_N^k \; t_1^k \ldots t_N^k\right]^T. \tag{2}$$

As described in [8], each trajectory in the set can be described as the superposition of an average or reference trajectory and a linear combination of deformation modes:

$$\tau_k = \overline{\tau} + P \cdot B_k, \tag{3}$$

where $P = [p_1 \; p_2 \ldots p_r]$ is the matrix of eigenvectors of the covariance matrix

$$S = \frac{1}{K-1} \sum_{i=1}^{K} (\tau_i - \overline{\tau})(\tau_i - \overline{\tau})^T = P \cdot \Lambda \cdot P^{-1} \tag{4}$$

and K is the number of trajectories available in the training set and $r = \min(3N, K) - 1$ is the number of degrees of freedom in the set.

The computation of matrix P from a set of representative trajectories is known as Principal Component Analysis (PCA) [9] or Karhunen-Loève Transform (KLT). It provide the r vectors $[p_1 \ldots p_r]$ or eigenshapes that define a set of orthogonal deformations. These deformations indicate the directions, relative to the mean shape, in which the trajectory points π_i^k are found to be moving. Generally only the M most important or energetic modes ($M \leq r$) are retained, where their energy is defined by the associated eigenvalue λ_i in matrix Λ.

In equation (3), B_k is a vector of deformation coefficients defined as

$$B_k = \left[b_1^k \; b_2^k \ldots b_r^k\right]^T, \tag{5}$$

that indicates the contribution of each deformation mode p_i toward the actual shape. For a given trajectory τ_k this contribution B_k is computed as

$$B_k = P^{-1}(\tau_k - \overline{\tau}). \tag{6}$$

Vector B_k provides a signature of the trajectory in deformation space. Similar trajectories should have resembling signatures, and since signatures provide a more compact description of trajectories —that is the whole point in using models— they are more convenient for classification tasks.

2.1 Outlier detection

However, if a given trajectory is very different from the ones in the training set, it will require a large amount of deformation to fit the resulting model. Therefore the deformation coefficients B_k can be used to detect outlier trajectories by using Hotelling's T^2 statistic [10], which is a multivariate analogue of the t-distribution. To use this statistic, the deformation modes have to be normalized so that they all have the same unit variance. Mathematically it means that we define a new set of normalized modes and coefficients

$$\tilde{P} = P\Lambda^{-\frac{1}{2}} \quad (7)$$

$$\tilde{B}_k = \Lambda^{-\frac{1}{2}} B_k. \quad (8)$$

where Λ is the diagonal matrix that contains the eigenvalues of the covariance matrix. In this normalized space we can define for each trajectory a scalar value

$$T_k^2 = \tilde{B}_k^T \tilde{B}_k = \sum_{j=1}^{M} (\tilde{b}_j^k)^2 \quad (9)$$

where M is the number of principal modes retained. This scalar T_k^2 is the Mahalanobis distance of the trajectory, and it can be interpreted as the normalized deformation energy of the related trajectory. If deformation coefficients b_j^k in the training set were normally distributed, then $(1-\alpha)\%$ of times the deformation energy would be bounded by the scalar

$$T_{\alpha,M,K}^2 = \frac{M \cdot (K-1)}{K-M} F_{M,K-M;\alpha} \quad (10)$$

where $F_{M,K-M;\alpha}$ stands for the Fisher distribution with M and $K-M$ degrees of freedom and $(1-\alpha)\%$ confidence interval. Therefore a trajectory can be defined as statistically conforming to the set if

$$T_k^2 \leq T_{\alpha,M,K}^2. \quad (11)$$

2.2 Analysis of temporal information

Trajectories can be analyzed in their spatial, temporal or spatiotemporal dimension by projecting them into the corresponding subspaces. If taken only in their spatial dimensions they become geometric shapes as in the original PDM formulation [8]. If they are projected into the time dimension they provide a temporal profile and can also be analyzed with the same methodology.

A spatiotemporal analysis requires measuring spatial and temporal deformation modes combined, but the variance present in the spatial and temporal components is generally different by orders of magnitude, which means that spatial deformation modes can be "masked" by temporal deformation modes. To avoid this problem, instead of applying the PCA on the covariance matrix, the correlation matrix is used, that is, the difference components are normalized with respect to their variance:

$$\tilde{X}_k = \frac{1}{\sigma_x}[x_1^k \dots x_N^k]^T \quad (12)$$

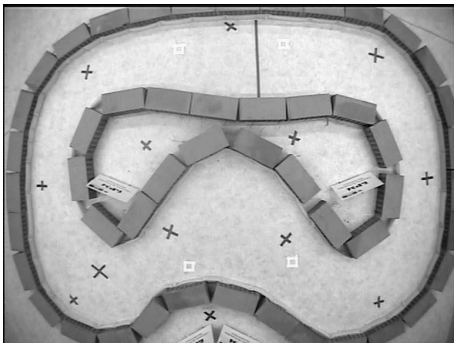
$$\tilde{Y}_k = \frac{1}{\sigma_y}[y_1^k \dots y_N^k]^T \quad (13)$$

$$\tilde{t}_k = \frac{1}{\sigma_t}[t_1^k \dots t_N^k]^T \quad (14)$$

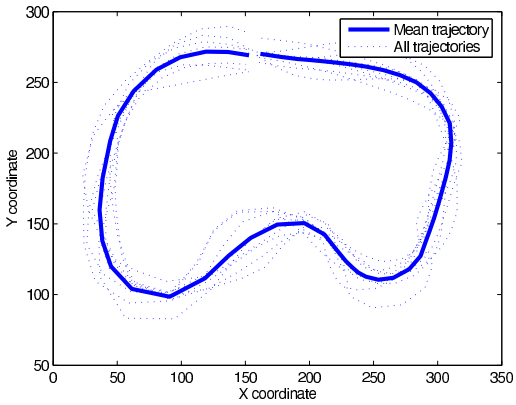
where σ_x stands for the standard deviation of the x component.

3 Experiments

We have applied the analysis described above to a series of trajectories extracted from a vision system that was tracking radio-guided cars running inside a circuit, shown in figure 1(a).



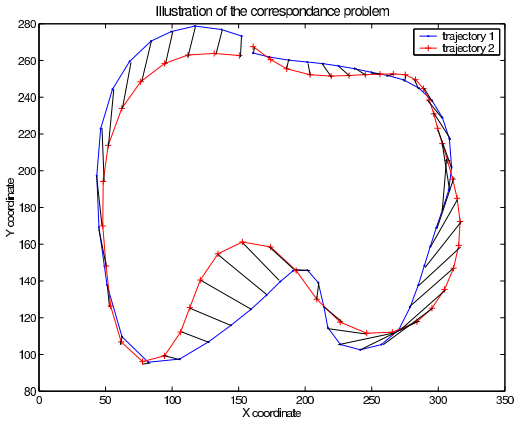
(a)



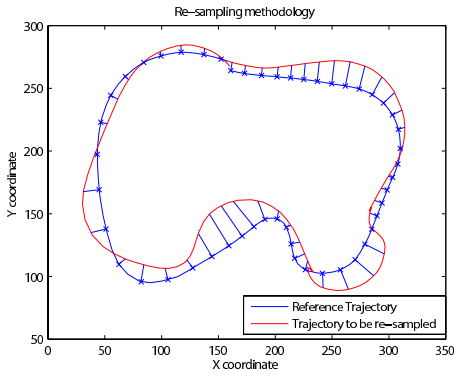
(b)

Figure 1: On the left, the circuit used for the radio-controlled cars, with the cross markings used for camera calibration purposes. On the right the spatial profile of a set of trajectories for one player (9 trajectories). The average trajectory is indicated with a thicker line.

This was a playful demonstrator installed during the public festivities for the EPFL's 150$^{\text{th}}$ anniversary. Spectators had the possibility of racing 2 cars for about 3 minutes, which implied typically 8-9 laps, which will be the trajectories under study (cf. figure 1(b)).



(a)



(b)

Figure 2: On the left the problem of correspondence between the points in two trajectories is shown. On the right, the selected resampling technique is shown: trajectories are resampled along orthogonal lines to the points in a reference trajectory.

Before being analyzed, trajectories have to be preprocessed to meet some requirements imposed by PDMs, such as that they all should have the same number of points. The simplest way is to perform a temporal resam-

pling, whereby each trajectory is fitted by a cubic spline [11] and resampled with a uniform sampling rate on the time axis, but this approach leads to a correspondence problem between points in two trajectories, as shown in figure 2(a).

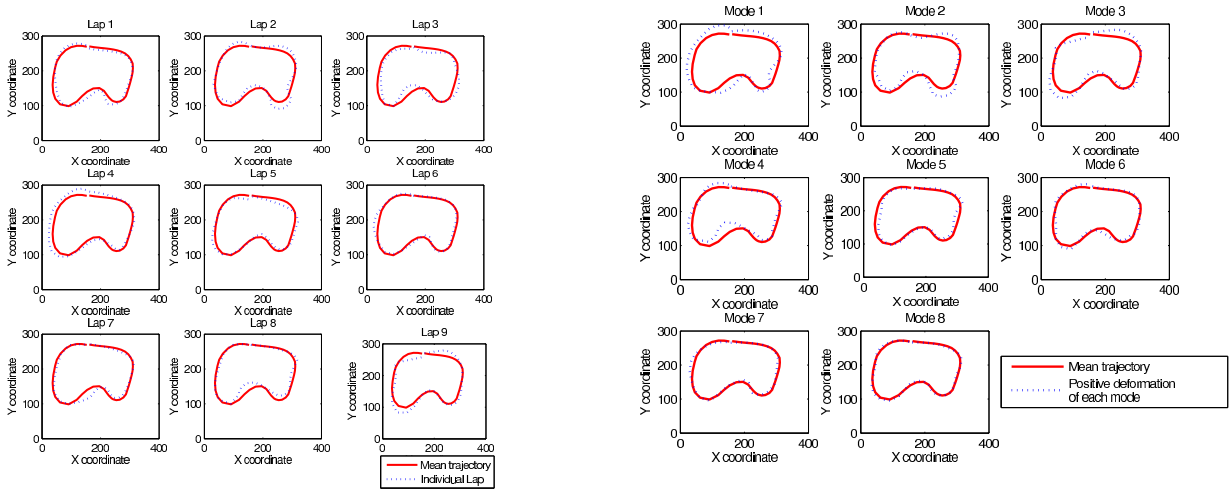
To solve this problem the trajectories are fitted to cubic splines and resampled along orthogonal positions to a reference trajectory (cf. figure 2(b)). The reference trajectory is chosen so that it is smooth enough to afford the resampling of the maximum number of trajectories. Indeed, if the trajectory has many bends, the intersections with the orthogonal lines might not respect the original order or they might not even exist. In any case, sensitivity analysis has shown that, for this data set, the choice of the reference trajectory does not have a noticeable impact on the subsequent modes. Once the trajectories are resampled they can be superposed and an average trajectory and modes can be computed.

3.1 Purely-spatial analysis

To perform a purely-spatial analysis we deal only with the x_i^k and y_i^k coordinates of the trajectory, removing the temporal values t_i^k. Equation (2) becomes

$$\tau_k = \left[x_1^k \dots x_N^k \, y_1^k \dots y_N^k\right]^T \tag{15}$$

and the analysis is performed as in the original PDM formulation [8], providing information about the spatial shape of the trajectories.



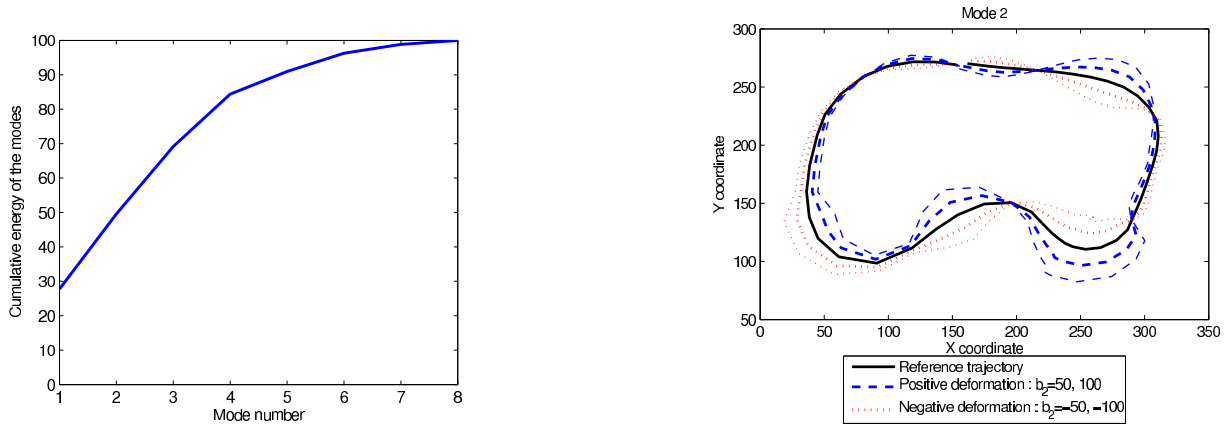
(a) (b)

Figure 3: On the left side each individual lap is plotted together with the average trajectory $\overline{\tau}$. The right image shows the synthetic trajectory corresponding to each mode p_j with its greatest contribution $\max_k(b_j^k)$ found in the set.

Figure 3(a) shows the individual laps together with the average trajectory $\overline{\tau}$. Figure 3(b) represents the different modes of the spatial analysis. Each mode is plotted with a coefficient corresponding to the greatest contribution $\max_k(b_j^k)$ found in the set. The modes are ordered by decreasing amount of shape variation (energy). The cumulative energy is plotted in figure 4(a), where it can be seen that the first 4 modes contain 85%

of the energy in the training set. The last mode can be neglected.

The spatial representation of each mode can be directly linked to the trajectories. Figure 4(b) shows mode 2 with different amounts of deformation by generating synthetic trajectories $\tau = \overline{\tau} + B \cdot P$ with a deformation vector $B = [0\ d\ 0\ 0\ 0\ 0\ 0\ 0]$ where only the second component is nonzero. It can be interpreted that this 2^{nd} mode encodes the variability in the way the bottom curves are negotiated, and most particularly the bottom-right curve. Indeed lap 2, which has a variation in that same curve, shows the greatest contribution from mode 2, as seen in figure 5(a). A similar relationship can be seen between mode 3 and lap 9. However, the most occurring source of variation, mode 1, represents the global variation among trajectories that follow the inner wall or the outer wall of the circuit.



(a) (b)

Figure 4: On the left, the cumulative energy of the deformation modes $Ce(j) = \sum_{i=1}^{j} \lambda_i / \sum_{k=1}^{r} \lambda_k$. It can be seen that that 85% of the total energy is contained in the first 4 modes. The right image shows some synthetic trajectories corresponding to mode 2: $\tau = \overline{\tau} + B \cdot P$ with $B = [0\ d\ 0\ 0\ 0\ 0\ 0\ 0]$ and $d = -100, -50, 0, 50, 100$. It can be interpreted that this 2^{nd} mode seems to encode mainly the variations in the way the bottom-right curve is negotiated.

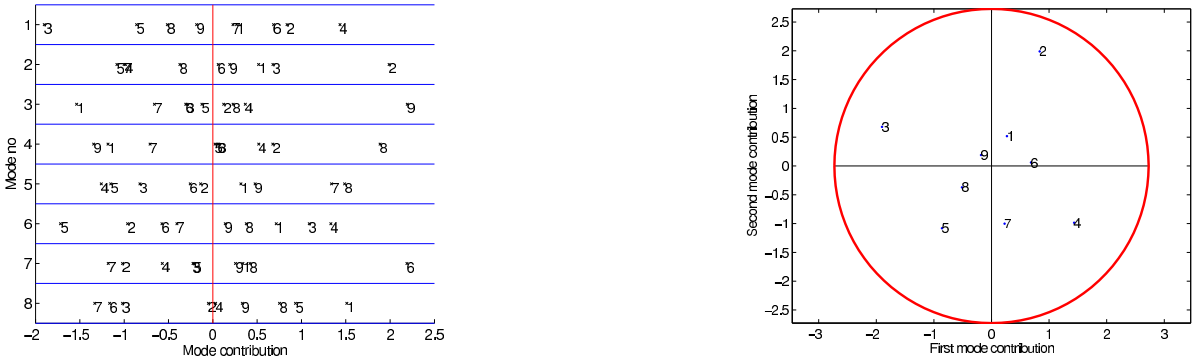
The outlier-detection procedure presented in section 2.1 can be used to investigate trajectories that stand out because they are too different from the original set. This difference can be related to higher efficiency (faster laps) or lower efficiency (slower laps). In figure 5(b) we plot each trajectory in the 2D space defined by the two first components of their normalized deformation vector $\tilde{B}_k$, as given in equation (8). In this 2D space, the upper limit of statistically acceptable deformation with a 90% confidence level is given by $T^2_{0.1,2,9} = 7.45$. This is plotted as a circle of radius $\sqrt{7.45} = 2.73$, because in that case equations (9) and (11) become

$$T_k^2 = (\tilde{b}_1^k)^2 + (\tilde{b}_2^k)^2 < T^2_{0.1,2,9} = 7.45 \tag{16}$$

It can be seen from figure 5(b) that none of the trajectories are statistically outliers, although trajectories 2 and 3 do stand out. Indeed, trajectory 2 happens to be the slowest one.

3.2 Spatiotemporal analysis

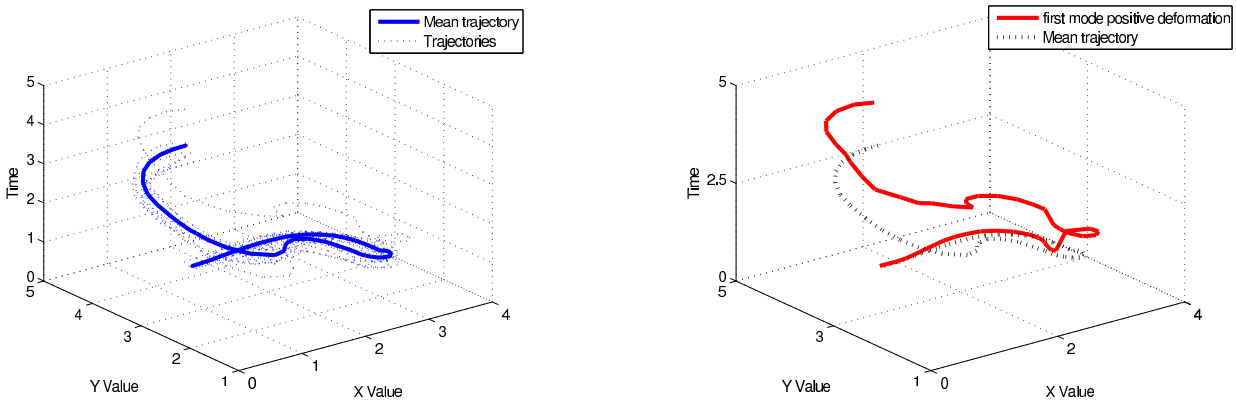
The resampling described in section 3 also involves the temporal component of trajectories. Indeed, a time component can be interpolated for each point on a given trajectory orthogonal to the points on the reference trajectory. This interpolation provides a series of spatiotemporal curves, shown in figure 6(a). After normalizing the trajectory data as explained in section 2.2, they are next decomposed into several deformation modes.



(a) (b)

Figure 5: On the left figure, the contribution of each mode to the different trajectories is plotted. A great contribution from a mode to a trajectory implies that the shape of the mode can be found in the shape of the trajectory. On the right figure the normalized deformation vector $\tilde{B}_k = [\tilde{b}_1^k \ \tilde{b}_2^k]$ for the two first modes is plotted on a circle of radius $\sqrt{T^2_{0.1,2,9}} = 2.73$ (cf. equation 10). In this case, no outlier has been found. Indeed the original variability is so high that none of the trajectories can be considered as an outlier.

Figure 6(b) shows the first mode with the highest amplitude present in the training set. It can be seen that this synthetic curve, which is the highest contribution of the first mode in the trajectory set, corresponds to a slower trajectory —it finishes later than the average trajectory. It has a jump in the time dimension, indicating that the car has been stopped, probably due to a collision.

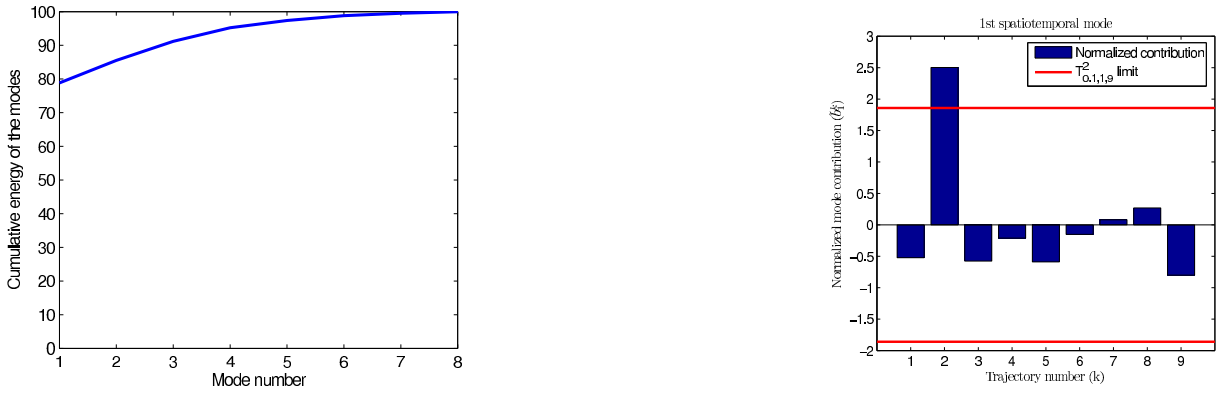


(a) (b)

Figure 6: The left figure shows the trajectories in spatiotemporal dimension superposed together with the mean trajectory. The right figure plots the shape of the maximum contribution of the first mode, showing that it produces a more jumpy, slower trajectory.

In their spatiotemporal domain, the trajectories that have been analyzed seem to have a smaller number of deformation modes. Indeed, the first mode, pictured above, accounts for 80% of the total energy in the set

(see figure 7(a)). From figure 7(b) it can be seen that mode 1 contributes the most, with a positive coefficient, to trajectory or lap 2. So much so that the corresponding coefficient $\tilde{b}_1^2$ is greater than the expected threshold $\sqrt{T^2_{0.1,1,9}}$, as per eq. (10), indicating that the second trajectory is a statistical outlier. It is indeed this trajectory that happens to be the slowest in the set.



(a) (b)

Figure 7: The left figure shows the cumulative energy in the spatiotemporal analysis of the trajectory set. The first deformation mode accounts for 80% of the total energy. The right figure shows the normalized deformation coefficients $\tilde{b}_1^k$ corresponding to the first mode in the spatiotemporal analysis of the trajectory set.

4 Conclusion

This paper presents a method to analyze trajectories in their combined spatial and temporal dimensions. The methodology is based on an extension of Point Distribution Models. This extension is needed to accommodate time-domain information that has a different dynamic range and variance than that of spatial information.

The projection of the spatiotemporal modes on the spatial and temporal dimension shows that they are indeed different from purely spatial and purely temporal deformation modes. This indicates that they do contain additional information compared to shape PDMs. However, compared to spatial PDMs, the interpretation of the resulting modes is much more complicated, because the data are richer. Some relationship can be found with the slowest trajectory in the set, because it has a longer temporal profile. Part of the interpretation problem lies in the fact that the underlying driver's style or behavior is unknown. It would therefore be desirable to apply the methodology on trajectories of known behavior such as those produced by a given stochastic differential equation or a robot, or on a more applied way, on the trajectories of a reduced number of expert pilots.

This project has also shown the importance of preprocessing the data and has provided a solution for the problem at hand, car trajectories in a closed circuit. This implies that the beginning and end of each trajectory was clearly defined (a lap), but this might not be always so easy, particularly in open spaces. The same applies for the time-domain information, because lap times were comparable quantities from lap to lap and they contained the relevant information (to analyze the driver's style.) In other applications, instantaneous velocity might be a more relevant information. Thus the choice of temporal features deserves further research.

4.1 Perspectives

To better correlate the analyzed trajectories with the observed behaviors, we have started a collaborative project with the Swarm-Intelligent Systems group (SWIS), a research team at the EPFL focusing on collective embedded systems involving multi-robot platforms. The point is that mobile robots provide an experimental platform the behavior of which can be programmed and yet provide a natural variability to their trajectories. From the roboticists' point of view, trajectory analysis tools such as the one described in this paper, provide a means of quantifying the robot's behavior and hence predict their performance (in terms of time, energy, work done) for tasks where trajectories play a role.

The first work to be done is to recreate the circuit experiment with mobile robots and to generate a huge number of trajectories, in order to classify the different behaviors. This experiment could demonstrate the validity of these first results.

Acknowledgment

The continuation of this project is supported by the Swiss National Science Foundation grant "Trajectory Analysis and Behavioral Identification in Multi-Robot Systems ", No. 200021–105565.

References

[1] N. Johnson and D. Hogg, "Learning the distribution of object trajectories for event recognition," *Image and Vision Computing*, vol. 14, no. 8, pp. 609–615, 1996.

[2] N. N. Gehrig, V. Lepetit, and P. Fua, "Golf club visual tracking for enhanced swing analysis tools," in *British Machine Vision Conference*, (Norwich, UK), September 2003.

[3] A. F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 257–267, 2001.

[4] R. Vaughan, N. Sumpter, J. Henderson, A. Frost, and S. Cameron, "Experiments in automatic flock control," *Robotics and Autonomous Systems*, vol. 31, pp. 109–117, 2000.

[5] R. Jeanson, S. Blanco, R. Fournier, J. Deneubourg, V. Fourcassie, and G. Theraulaz, "A model of animal movements in a bounded space," *Journal of Theoretical Biology*, vol. 225, pp. 443–451, Dec 2003.

[6] Z. Khan, T. Balch, and F. Dellaert, "Efficient Particle Filter-Based Tracking of Multiple Interacting Targets Using an MRF-based Motion Model," in *Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'03)*, 2003.

[7] T. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231–268, 2001.

[8] T. Cootes, C. Taylor, and D. Cooper, "Active shape-models - their training and applications," *Computer Vision and Image Understanding*, pp. 38–59, 1995.

[9] J. Jackson, "Principal components and factor analysis: part I," *Journal of Quality Technology*, vol. 12, pp. 201–213, October 1980.

[10] D. Montgomery, *Introduction to statistical quality control*. New York: Wiley, 4th ed., 2001.

[11] M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Magazine*, vol. 16, pp. 22–38, November 1999.