



ELSEVIER

Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com)

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Visual abnormal behavior detection based on trajectory sparse reconstruction analysis

Ce Li, Zhenjun Han, Qixiang Ye, Jianbin Jiao*

University of Chinese Academy of Sciences, PR China

ARTICLE INFO

Available online 11 January 2013

Keywords:

Abnormal behavior detection
Trajectory representation
Sparse reconstruction analysis
L1-norm minimization

ABSTRACT

Abnormal behavior detection has been one of the most important research branches in intelligent video content analysis. In this paper, we propose a novel abnormal behavior detection approach by introducing trajectory sparse reconstruction analysis (SRA). Given a video scenario, we collect trajectories of normal behaviors and extract the control point features of cubic B-spline curves to construct a normal dictionary set, which is further divided into Route sets. On the dictionary set, sparse reconstruction coefficients and residuals of a test trajectory to the Route sets can be calculated with SRA. The minimal residual is used to classify the test behavior into a normal behavior or an abnormal one. SRA is solved by L1-norm minimization, leading to that a few of dictionary samples are used when reconstructing a behavior trajectory, which guarantees that the proposed approach is valid even when the dictionary set is very small. Experimental results with comparisons show that the proposed approach improves the state-of-the-art.

Crown Copyright © 2013 Published by Elsevier B.V. All rights reserved.

1. Introduction

Abnormal behavior detection is important to video content understanding, with applications in intelligent video surveillance [1–4] and content-based multi-media retrieval [1,5], etc. Research in abnormal detection has made great progresses in recent years, such as abnormal action detection [5,6], abnormal event detection [7–10], and abnormal crowd detection [11–15]. All these methods can be categorized into two parts, based on visual features of video stream [10–13] or based on trajectory analysis [16–24]. In recent years, trajectory analysis based methods received much attention when performing visual abnormal behavior detection [12–18,30]. Although extensively investigated, trajectory analysis is still an open research topic with challenging problems, including the trajectory length variation [19,22], the trajectory noise [20] and the limited sizes of sample sets [16,21]. Researchers are putting a lot of effort into finding more effective trajectory representation and modeling approaches.

In the early research, trajectory representations include representative sequences corresponding to motion vectors [23], motion vectors with acceleration information [24], etc. Since these representations directly extracted the object positions and velocities in video frames, they lead to variable feature length and bring difficulties to the trajectory analysis. In the later research, fixed-length vectors based on re-sampling and linear interpolation [18,25] are proposed. These vectors can deal with the problem of trajectory length variation, but the interpolation

often brings in redundant and noise information. Naftel and Khalid [26] propose an efficient trajectory representation using function approximation algorithms of Least Square Polynomial, Chebyshev Polynomial and Discrete Fourier Transform (DFT), while the representations in transformation domain increase the complexity of the representation. In Ref. [27], a more efficient trajectory representation is proposed. Haar Wavelet Coefficients and Least-squares Cubic Spline Curves Approximation (LCSA) are adopted as parametric feature vectors. These parametric feature vectors are insensitive to the length of trajectories, providing a general tool. The performance of these representations can be optimized by selecting proper parameters. In this paper, we follow the idea of Ref. [16] to extract trajectory representation, and make a deep discussion about the selection of proper feature parameters based on quantitative experiments.

On the other hand, various methods are investigated or employed to model the trajectories on the representations. Clustering methods, such as Self-Organizing Map Neural Network [24] and hierarchical Fuzzy *K*-means clustering [25], are used to classify trajectories in an un-supervised manner and then build prototypes. Test trajectories will be classified by their distances to the prototypes. The used distances include Euclidean distance [17], Hausdorff distance [19] or Dynamic Time Warping (DTW) [22,28] etc. The disadvantages of these distances lie in that they cannot reflect the statistical nature of behaviors. In other words, classification based on distance measurement does not consider the different importance of features as a probabilistic or discriminative method. In recent years, supervised learning methods, such as Gaussian Mixture Models (GMMs) [16], Bayesian Model [12,21], Hidden Markov Model (HMM) [29], One-Class Support Vector Machine (OC-SVM) [7], Hierarchical Hidden Markov Model [29] and Nonparametric

* Correspondence to: No. 19 A, Yu Quan Road, Shi Jing Shan District, Postcode: 100049 Beijing, PR China. Tel: +86 10 88256968.

E-mail address: jiaojb@ucas.ac.cn (J. Jiao).

Bayesian Model [21], are employed in trajectory analysis. Given a large training set, these methods can reach a good performance. But when facing a small training set, the performance of these methods cannot always be guaranteed. It is known that labeling trajectory samples in video sequences always result in huge workload due to the large amount of video data. Therefore, exploring trajectory analysis approach for a small sample set is significant.

Inspired by the development of sparse reconstruction in face recognition [31] and object tracking [32], we cast the trajectory classification as a sparse reconstruction problem. Intuition behind the sparse reconstruction lies in the fact that the coefficients imply the discriminative information (some coefficients that can compactly express the trajectory are nonzero and the others are almost zero) among different trajectory patterns, which can be used for classification. Sparse Reconstruction Analysis (SRA), solved by L1-norm minimization, is suitable to represent and reconstruct a behavior trajectory with a few of dictionary samples. In theory, given an input test sample $y \in \mathbb{R}^m$, we reconstruct it by a sparse linear combination of an over-complete normal basis set $\Phi = \mathbb{R}^{m \times D}$, where $m \ll D$. Therefore, this method is proposed to quantify the normalness of trajectory via a sparse reconstruction from normal ones. The reconstruction for a special behavior trajectory based on all behavior trajectories is typically sparse. According to the sparse theory [31,34,35], once sparseness is guaranteed, the size of sample set has a little effect on classification performance. This guarantees the classification performance of SRA with a small sample set, which is the main advantage of this work compared with state-of-the-art methods represented by Ref. [16], where a large sample set is required to build the models.

The rest of the paper is organized as follows. In Section 2, we introduce the trajectory representation and the abnormal behavior detection approach, in Section 3, we present the experiments with comparisons, and in Section 4 we conclude the paper with the discussion of the future work.

2. Methodology

In this section, we first present an overview of the proposed abnormal behavior detection approach and then describe the trajectory representation and the sparse reconstruction analysis.

2.1. Overview of the proposed approach

In this paper, based on a predefined trajectory dictionary set for a fixed video scenario, we propose a novel abnormal behavior detection approach using sparse reconstruction analysis. We first collect a set of trajectories of normal behaviors from videos by an object tracking algorithm [32] or a motion detection method [14,29]. By observing their appearances, these trajectories are manually categorized into different sub-sets, called Route sets. For all the collected trajectories, the Least-squares Cubic Spline Curves Approximation (LCSCA) features are extracted for representations and then construction of the dictionary set.

When performing abnormal behavior detection, each test trajectory will also be represented with LCSCA features. Then, we introduce the sparse reconstruction analysis on the normal dictionary set to classify the testing motion trajectories of objects, where our objective is to reconstruct the test trajectory with as few dictionary samples as we can. The L1-norm minimization is used to solve the reconstruction coefficients, on which the reconstruction residuals of each Route set can also be calculated. The minimal reconstruction residual is used to classify the test trajectory into a norm behavior or an abnormal one with an empirically defined threshold. The framework of the proposed approach is shown in Fig. 1.

2.2. Trajectory representation

Since motion trajectories consist of coordinate sequences of different length and are extracted on different frame numbers, we use the control points of LCSCA to extract fixed-length parametric vectors as feature representation. This is achieved by approximating each spatial-temporal trajectory with a uniform cubic B-spline curve [33] parameterized by time (frame number). Cubic B-spline curve [33] can be thought of a method for defining a sequence to approximate the form of trajectory. A spline curve is a sequence of curve segments that are connected together to form a single continuous curve (here trajectory). Further mathematical explanation is shown in the Ref. [33]. Because of the number of control points and weight factors, the representation of the basis is flexible for simple or complicated shape of curves. Given a trajectory sequence in (x, y, t) space, we use B-spline control points to represent both the shape and spatio-temporal profiles of a trajectory $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_{t-1}, y_{t-1}), (x_t, y_t)\}$ in a parametric way $F = \{C_1^x, C_2^x, \dots, C_p^x, C_1^y, C_2^y, \dots, C_p^y\}$, where p is the number of control points and t is the length of trajectory [16], C_p^x is the normalized x coordinate of p th control point, and C_p^y is the normalized y coordinate of p th control point. Fig. 2 shows a normal and an abnormal trajectories, respectively.

The transformation procedure of the trajectory representation is as follows:

- (1) Define the parameter vectors = $\{0, s_2, \dots, s_{t-1}, s_t\}$,

$$s_n = \frac{\sum_{i=2}^n \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}}{\sum_{i=2}^t \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}} \quad (n = 2, 3, \dots, t, s_n \in (0, 1]) \quad (1)$$

where $\sum_{i=2}^n \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}$ is the total distance traversed at a given point (x_n, y_n) ; and define the knot vector

$$\tau = \left\{ \underbrace{0, 0, 0, 0}_{1 \dots 4}, \underbrace{\frac{1}{p-3}, \frac{2}{p-3}, \dots, \frac{p-4}{p-3}}_{5 \dots p}, \underbrace{1, 1, 1, 1}_{p+1 \dots p+4} \right\} \quad (2)$$

where τ denotes a knot vector with $p+4$ elements.

- (2) Calculate the cubic B-spline basis function with the following recursive formulation according to De-Boor algorithm [33]

$$B_{p,1}(s_n) = \begin{cases} 1 & \text{if } \tau_p \leq s_n < \tau_{p+1} \\ 0 & \text{otherwise} \end{cases}, \quad B_{p,m}(s_n) = \frac{s_n - \tau_p}{\tau_{p+m-1} - \tau_p} B_{p,m-1}(s_n) + \frac{\tau_{p+m} - s_n}{\tau_{p+m} - \tau_{p+1}} B_{p+1,m-1}(s_n) \quad (3)$$

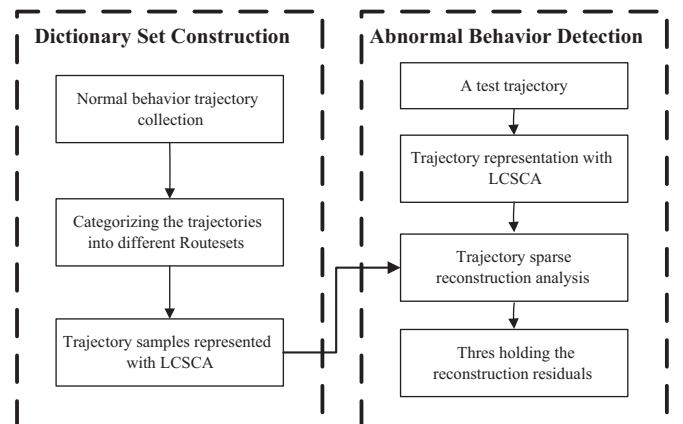


Fig. 1. Framework of the proposed approach.

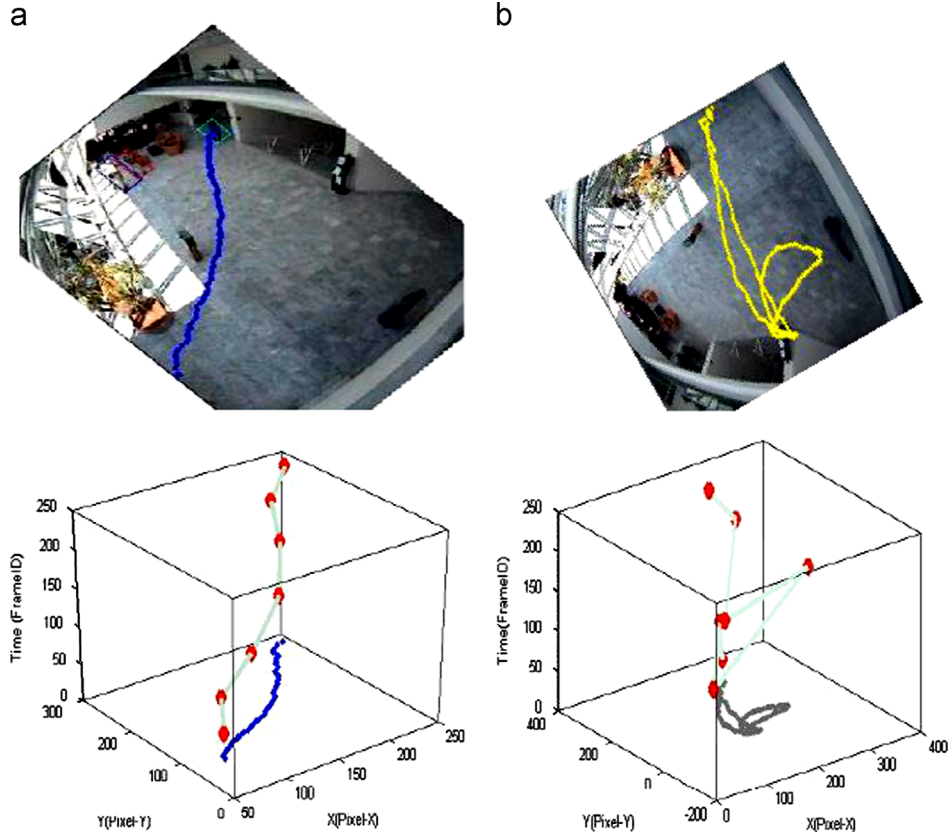


Fig. 2. Trajectory representation when $p=7$. (a) Top: a normal trajectory sample; bottom: the feature, representation with LCSCA. (b) Top: an abnormal trajectory sample; bottom: the feature representation with LCSCA.

where τ_p is the p th knot of above-mentioned τ vector, and m denotes the order of the function (the initialized value of m is 4 for cubic splines).

- (3) Find the control points by (4), which minimizes the sum of squared errors between the original trajectory and its approximation

$$F^{XY} = \Phi^\dagger T^{XY}, \quad (4)$$

where

$$\Phi = \begin{Bmatrix} B_{1,A}(s_1) & \cdots & B_{p,A}(s_1) \\ \vdots & \ddots & \vdots \\ B_{1,A}(s_t) & \cdots & B_{p,A}(s_t) \end{Bmatrix} \text{ and } \Phi^\dagger = \Phi^T \Phi^{-1} \Phi^T$$

2.3. Sparse reconstruction analysis

In this section, we address the problem of trajectory classification. As mentioned in Section 1, it is reasonable to make a hypothesis that the test trajectories can be approximately represented with a linear superposition of the sample set. Supposing that there are J behavior patterns (called Routes) in a surveillance scene, and each Route $A_j = \{a_j^1, a_j^2, \dots, a_j^K\}$ holds K behavior trajectories, we can have the union sample set in the scene as follows:

$$B = \cup \{A_j\} = \{a_1^1, a_1^2, \dots, a_1^K, a_2^1, a_2^2, \dots, a_2^K, \dots, a_J^1, a_J^2, \dots, a_J^K\}, j = 1, \dots, J \quad (5)$$

For a test trajectory represented by above mentioned control point feature F_t , we calculate the sparse linear reconstruction coefficients ψ on set B as (6)

$$B\psi \approx F_t \quad (6)$$

where F_t is the LCSCA representation vector of the test trajectory,

$\psi = \{\psi_j^k\}, j = 1, \dots, J, k = 1, \dots, K$ and ψ_j^k is the reconstruction coefficient corresponding to the k th sample of j th Route in B .

In the real condition of a surveillance scene, the sample set of the training trajectories (in this case, a sample is a representative vector of a normal trajectory) is always quite large, then there may be many redundant samples in the set, which leads to a sparse coefficient vector of the linear superposition. Therefore, a few of samples that are similar with the given input trajectory will be activated, but the whole coefficient vector remains sparse, supposing there are r nonzero coefficients in $\psi = \{\psi_j^k\}$, r is less than 5% of K in the statistical experiment, where $r \ll K$. By this token, the object has an r -sparse representation based on the sample set. The number of the nonzero coefficients is denoted by $\|\psi\|_0$. Minimizing $\|\psi\|_0$ is the principal to obtain a sparse representation, which is, however, an NP-hard problem. Recent development in the theory of compressed sensing [36] shows that the solution of L1-norm minimization subject to a linear system of the samples can be used to find sparse enough representation of the test sample. The resulting optimization problem, similar to the Least Absolute Shrinkage and Selection Operator (LASSO) in statistics [37], which is a constrained optimization problem for estimating regression coefficients to the least squares criterion, penalizes the L1-norm of the coefficients in the linear combination, rather than directly penalizing the number of nonzero coefficients $\|\psi\|_0$. In terms of the set B and the test sample F_t , a sparse representation is computed as follows:

$$\operatorname{argmin} \|\psi\|_1, \text{ s.t. } B\psi = F_t \quad (7)$$

where $\|\cdot\|_1$ denotes the L1-norm. The property of the sparse representation guarantees that the test sample is represented in the most compact way based on the sample set. That is to say, sample-based sparse reconstruction analysis selects a small subset,

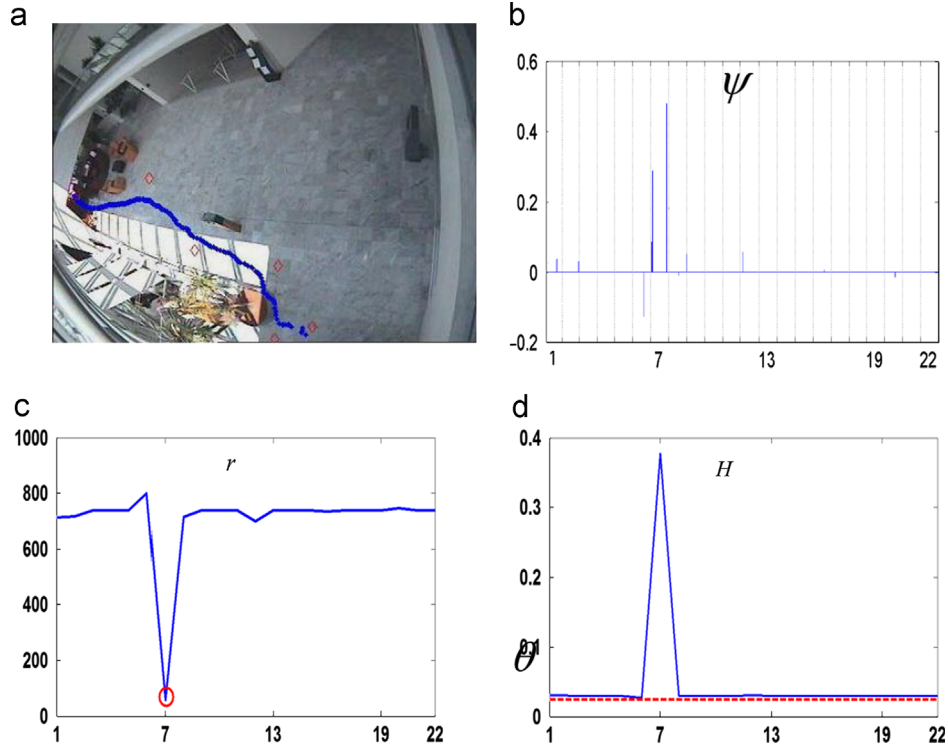


Fig. 3. Illustration of sparse reconstruction coefficients (a) a test trajectory of class 7, (b) the corresponding coefficients, (c) the corresponding residuals, and (d) proportion of the trajectory belongs to normal.

in which the samples are the most representative ones, such as parts of or the whole object.

When we obtain a test sample F_t with its corresponding sparse reconstruction coefficients, we build a classifier by calculating the residuals between the sparse reconstructions with each sample in the set B . Firstly, we define a characteristic function δ_j for each class, which keeps the sparse coefficients corresponding to j th class in B and sets the coefficients to zero of the other classes. Then the sparse reconstruction residuals based on each trajectory class can be defined as

$$r_j(F_t) = \|F_t - B\delta_j(\psi)\|_2, \quad j = 1 \dots J \quad (8)$$

and a threshold θ is used to discriminate the new example as abnormal if the result is negative, as in (9). Finally, if the new example is regarded as a normal trajectory, the trajectory pattern can be classified according to (10). The threshold θ is set as 0.03 empirically. In other words, when there was a H_j less than 0.03, the new example is an outlier of sample set space. For instance, Fig. 3 shows the sparse reconstruction coefficients, the reconstruction residuals based on each class and proportion of the trajectory belongs to class 7, where the horizontal ordinate denotes the identifier of class. Many nonzero coefficients are from class 7 in Fig. 3(b), and the minimum of residuals denotes the classification result in Fig. 3(c). Because all of the H values are more than threshold in Fig. 3(d), the trajectory (shown in Fig. 3(a)) is determined as normal.

$$\text{Detect}(F_t) = \text{sign}(\min_{j=1 \dots J} H_j - \theta), \quad \text{where } H_j = \frac{1/r_j(F_t)}{\sum_{i=1}^J 1/r_j(F_t)} \quad (9)$$

$$\text{Classify}(F_t) = \underset{j}{\text{argmin}} r_j(F_t) \quad (10)$$

3. Experiments

In this section, we carry out experiments with comparisons in order to validate the proposed approach and demonstrate its

advantages. There are two groups of experiment, multiple trajectory representations are carried out with comparisons, and comparisons of the proposed method with GMMs based method [16] are carried out on CAVIAR [38] dataset and NGSIM (Lankershim) [39] dataset.

3.1. Experiments for multiple representation comparisons

In this group of experiment, the used dataset is publicly available CAVIAR [38] dataset, which contains a series of behaviors in the entrance lobby of INRIA lab. There are 11 entry-exit Routes appeared in the dataset. The size of surveillance video frame is 640×460 , and each trajectory consists of a sequence of x -coordinates and y -coordinates, which are the center positions of moving objects in the frames.

Considering the traversal orientations, we obtain totally 22 types of normal trajectories for behavior analysis (each type holding 100 simulated tracks shared by Rowland Sillito¹). Some examples from each of the Routes from dataset are visualized and shown in Fig. 4. In the detection procedure, we choose 21 trajectories to represent the normal behaviors, consisting of people walking directly from one exit to another and 19 trajectories to define abnormal behaviors, consisting of people fighting, falling down, leaving or collecting packages, as test samples.

To choose proper parameters and quantitatively validate the proposed approach, DACC (Detection ACCuracy) and CCR (Correct Classification Rate) are defined as follows:

$$\text{DACC} = \frac{\text{TP} + \text{TN}}{\text{The total of test behaviors}} \quad (11)$$

$$\text{CCR} = \frac{\text{The number of correct route classification}}{\text{The total of normal test behaviors}} \quad (12)$$

¹ Thanks to the kindly help of Rowland Sillito about the dataset, which were used for anomalous trajectory detection in Ref. [16].

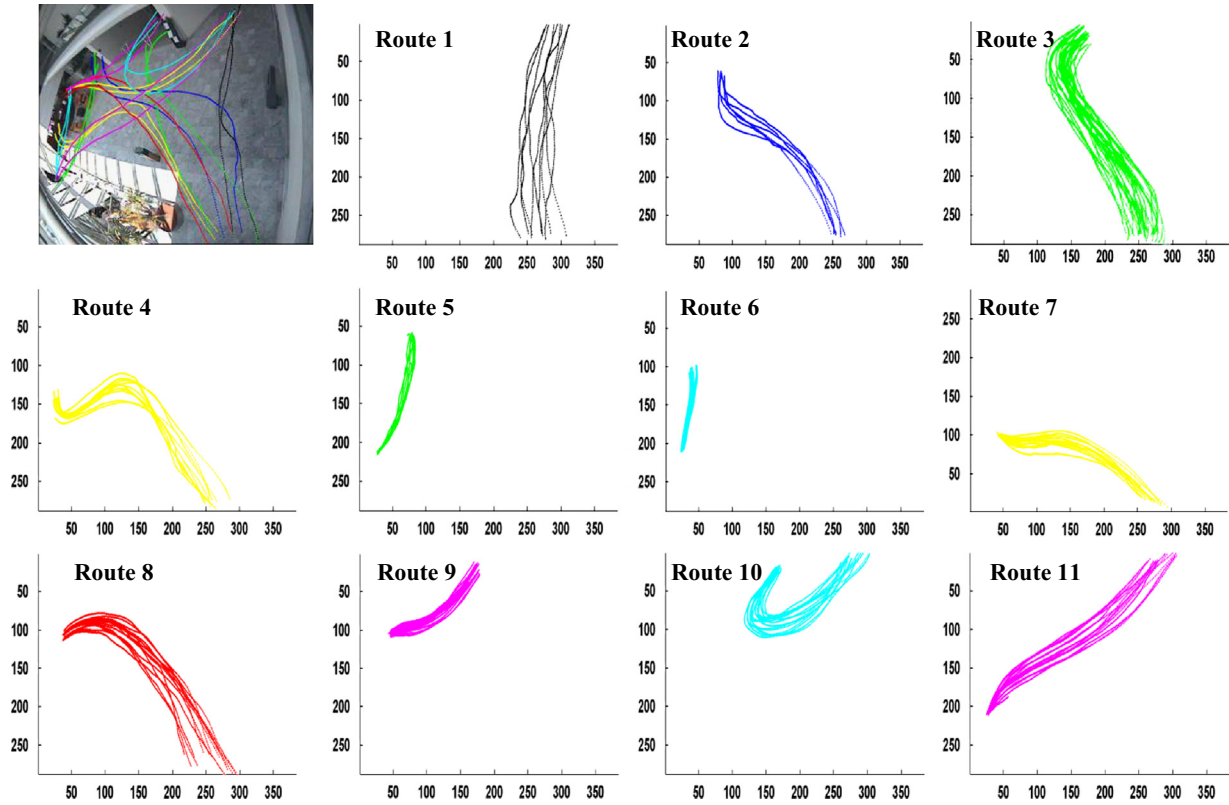


Fig. 4. Route examples of the dataset [38].

where TP (True Positives) is the number of trajectories correctly detected as normal behaviors, and TN (True Negatives) is the number of trajectories correctly detected as abnormal behaviors. DACC is used to evaluate the performance of normal–abnormal behavior detection and CCR is used to evaluate the performance of Route classification when a trajectory has already classified as a normal behavior. In the experiments the higher the two rates are, the better the performance is.

For trajectory representation, we defined two kinds of feature vectors: R1 and R2. R1 is defined as the sequence $F = \{C_1^X, C_2^X, \dots, C_p^X, C_1^Y, C_2^Y, \dots, C_p^Y\}$ calculated in Section 2.2; R2 contains F and the coordinates of the entry and exit positions. To determine the appropriate values of p (the number of control points), we calculate the second derivative of approximation deviation for different p , and find that the second derivative reaches its minimum when $p=5, 7, 9$, or 11 . The approximation deviation is the median of the distances between all control points with their nearest trajectory points. Therefore, we set $p=\{5, 7, 9, 11\}$ for performance comparisons.

The average DACC and CCR of 10 trails based on R1 are shown in Fig. 5. In the CAVIAR lobby scenarios, by adjusting the parameter p , it is observed that when $p=5$, a better DACC=90.42% ($\pm 3.85\%$) and CCR=70.09% ($\pm 6.13\%$) are obtained (as shown in Fig. 5). The figure also shows the variations of DACC and CCR when K increases, where K is the number of the samples in each kind of Route. When K is set as small as 5 (a very small dictionary set), 84.75% DACC and 62.38% CCR can be obtained, showing the effectiveness of the proposed approach even when the dictionary set (sample set) is small.

The results when adding the coordinates of the entry and the exit into R1 (called R2 representation) are shown in Fig. 6. When $p=5$, obtained DACC=83.75% ($\pm 2.88\%$) and CCR=68.15% ($\pm 4.53\%$) are still steadily higher than the others. Besides, a 81.75% DACC and

63.81% CCR can be obtained when $K=5$. Compared with the results reported by R1 (Fig. 5), R2 reports a worse performance. The reason lies in that trajectories sharing the same entry or exit can sometimes been classified as the same behavior with our SRA method.

3.2. Experiments for performance comparisons

In this group of experiment, two different datasets are used to demonstrate the advantages of the proposed method. One is afore mentioned CAVIAR dataset and the other is NGSIM (Lankershim) dataset. The NGSIM (Lankershim) dataset consists of 1213 behavior trajectories in a crossroad scenario. There are four Routes (similarly, considering the traversal orientations, there are eight types of normal trajectories for behavior analysis) appeared in the scenario. The size of surveillance video frame is 720×576 . 800 of the normal examples are used for training and 45 normal examples, along with 12 anomalous examples, for testing.

We compare our proposed SRA approach with the Gauss Mixture Models (GMMs) based method [16] which is the state-of-the-art on visual abnormal trajectory detection. Both methods use the same sample set and test set. In the CAVIAR lobby scenario, if we adjust the number of the samples in each Route K , Fig. 7 shows the variations of DACC with respect to K , where DACC is the average of 10 trails based on R1 with $p=7$. It shows that our approach achieves a better performance than GMMs, especially when K is small. As for the previous dataset, performance measure (DACC) is taken in the NGSIM (Lankershim) scenario. The situation is similar using the same empirical threshold. Fig. 8 illustrates the changes in performance observed: detection performance (DACC) steadily improves with respect to K , which is the number of the samples in each Route. Although the threshold is probably arbitrarily chosen, it shows that our approach achieves a better performance than GMMs, especially

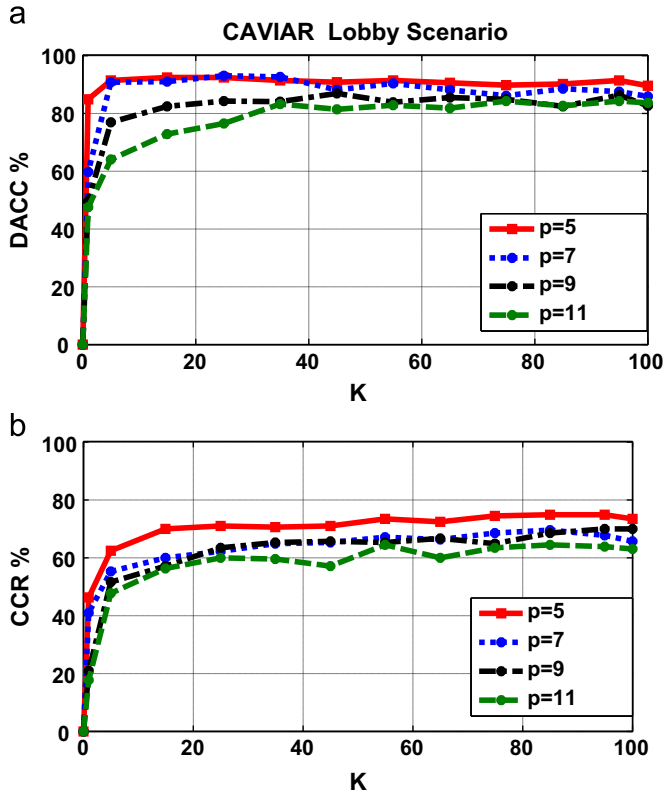


Fig. 5. Results of (a) DACC, and (b) CCR based on R1.

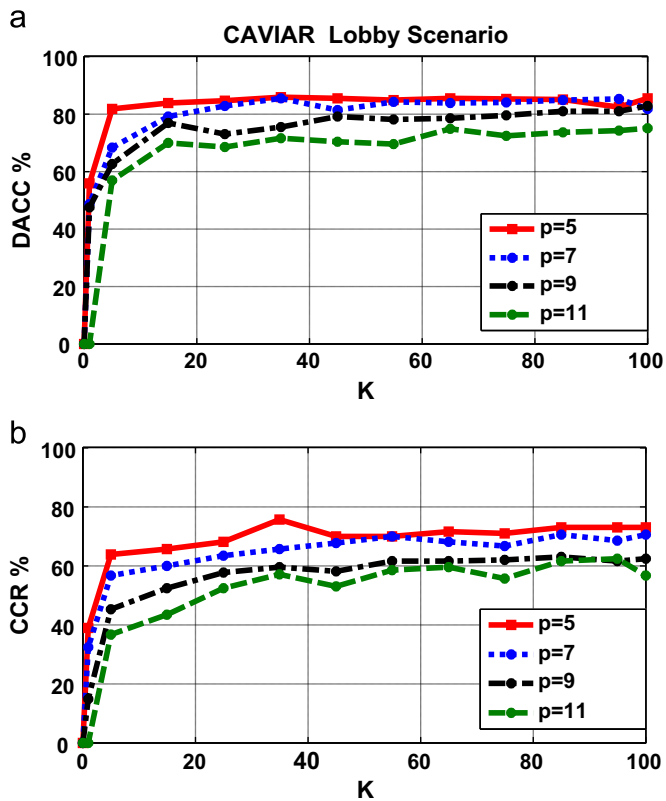


Fig. 6. Results of (a) DACC, and (b) CCR based on R2.

when K is small. This shows that by sparse reconstruction we could compactly represent the trajectories even when the sample set is small while existing GMMs need a larger training set to reach a good performance.

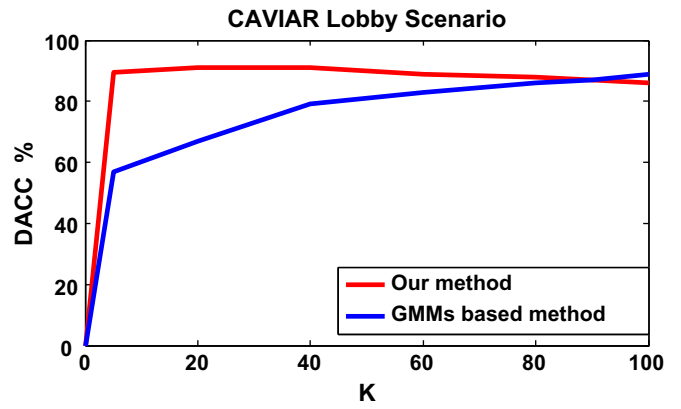


Fig. 7. Comparison of the proposed method with GMMs based method [16] on CAVIAR dataset.

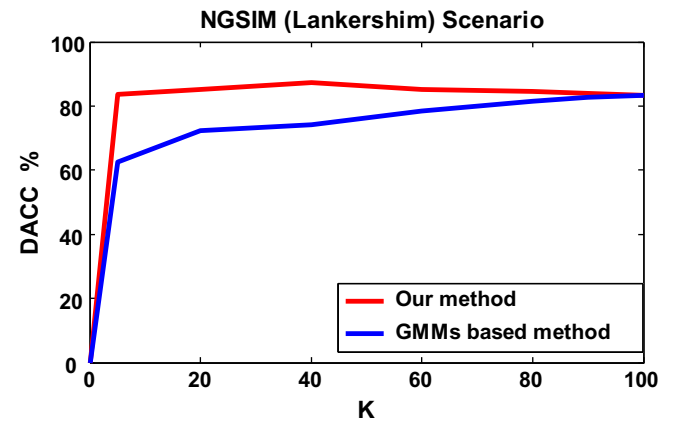


Fig. 8. Comparison of the proposed method with GMMs based method [16] on NGSIM (Lankershim) dataset.

4. Conclusions and future work

In this paper, we propose a novel trajectory-based visual abnormal behavior detection approach with sparse reconstruction analysis. The new concept and technique introduced in this paper include trajectory analysis, trajectory linear reconstruction, the trajectory dictionary set and Route set. Whether a testing sample is abnormal or not is determined by its sparse linear reconstruction coefficients and residuals, through a linear reconstruction of the normal dictionary set. Thanks to the flexibility of the LCSCA representation and the employed SRA method, our approach reports an efficient and robust detection. Experimental results on a real-world dataset show the good performance of our proposed approach on different sizes of samples. The comparison to the state-of-the-art is also provided, which indicates that the proposed approach achieves a better result even though a smaller dictionary set is used. A known disadvantage of the proposed method is that the detection performance affected by the control point parameter, which should be improved in the future work.

Acknowledgment

This work is supported in Part by National Basic Research Program of China (973 Program) with Nos. 2011CB706900, 2010CB731800, and National Science Foundation of China with Nos. 61039003, 61271433 and 61202323.

References

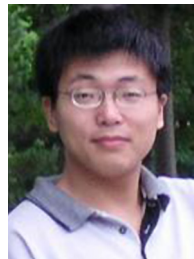
- [1] H. Zhong, J. Shi, and M. Visontai, Detecting unusual activity in video, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2004, pp. 819–826.
- [2] T. Xiang, S. Gong, Video behavior profiling for anomaly detection, IEEE Trans. Patt. Anal. Mach. Intell. (2008) 893–908.
- [3] N. Robertson, and I. Reid, A general method for human activity recognition in video, in: Proceedings of Computer Vision and Image Understanding, 2006, pp. 232–248.
- [4] V. Saligrama, J. Konrad, P. Jodoin, Video anomaly identification, IEEE Signal Process. Mag. (2010) 18–33.
- [5] D. Weinland, R. Ronfard, E. Boyer, A survey of vision-based methods for action representation, segmentation and recognition, Comput. Vis. Image Understand. (2011) 224–241.
- [6] X. Cui, Q. Liu, M. Gao, and D.N. Metaxas, Abnormal detection using interaction energy potentials, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2011, pp. 3161–3167.
- [7] C. Piciarelli, G.L. Foresti, Trajectory-based anomalous event detection, IEEE Trans. Circ. Syst. Video Tech. (2008) 1544–1554.
- [8] A. Adam, E. Rivlin, I. Shimshoni, D. Reinitz, Robust real-time unusual event detection using multiple fixed-location monitors, IEEE Trans. Patt. Anal. Mach. Intell. (2008) 555–560.
- [9] J. Kim, and K. Grauman, Observe locally, infer globally: a space-time MRF for detecting abnormal activities with incremental updates, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2009.
- [10] S. Kwak, H. Byun, Detection of dominant flow and abnormal events in surveillance video, Opt. Eng. (2011).
- [11] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, Anomaly detection in crowded scenes, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2010.
- [12] X. Wang, X. Ma, and E. Grimson, Unsupervised activity perception by hierarchical Bayesian models, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2007.
- [13] S. Wu, B.E. Moore, and M. Shah, Chaotic invariants of Lagrangian particle trajectories for anomaly detection in crowded scenes, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2010, pp. 2054–2060.
- [14] L. Kratz, and K. Nishino, Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2009.
- [15] R. Mehran, A. Oyama, and M. Shah, Abnormal crowd behavior detection using social force model, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2009.
- [16] R.R. Sillito, and R.B. Fisher, Semi-supervised learning for anomalous trajectory detection, in: Proceedings of British Machine Vision Conference, 2008, pp. 1035–1044.
- [17] Z. Fu, W. Hu, and T. Tan, Similarity base vehicle trajectory clustering and anomaly detection, in: Proceedings of International Conference on Image Processing, 2005.
- [18] C. Piciarelli, and G.L. Foresti, Online-trajectory clustering for anomalous events detection in: Proceedings of Pattern Recognition Letters, 2006, pp. 1835–1842.
- [19] X. Wang, K. Tieu, and E. Grimson, Learning semantic scene models by trajectory analysis, in: Proceedings of European Conference on Computer Vision, 2006, pp. 110–123.
- [20] Z. Zhang, K. Huang, T. Tan, and L. Wang, Trajectory series analysis based event rule induction for visual surveillance, in: Proceedings of IEEE Computer Vision and Pattern Recognition, 2007, pp. 1–8.
- [21] X. Wang, and E. Grimson, Trajectory analysis and semantic region modeling using a nonparametric Bayesian model, in: IEEE Conference on Computer Vision and Pattern Recognition, 2008.
- [22] Z. Zhang, K. Huang, and T. Tan, Comparison of similarity measures for trajectory clustering in outdoor surveillance scenes, in: Proceedings of International Conference of Pattern Recognition, 2006, pp. 1135–1138.
- [23] N. Johnson, D. Hogg, Learning the distribution of object trajectories for event recognition, Image Vision Comput. (1996) 609–615.
- [24] J. Owens, A. Hunter, Application of the self-organizing map to trajectory classification, IEEE Int. Workshop Vis. Surveill. (2000) 609–615.
- [25] W. Hu, X. Xiao, Z. Fu, D. Xie, T. Tan, S.J. Maybank, A system for learning statistical motion patterns, IEEE Trans. Patt. Anal. Mach. Intell. (2006) 1450–1464.
- [26] A. Naftel, S. Khalid, Classifying spatiotemporal object trajectories using unsupervised learning in the coefficient feature space, Multimedia Syst. (2008) 227–238.
- [27] R.R. Sillito, and R.B. Fisher, Parametric trajectory representations for behaviour classification, in: Proceedings of British Machine Vision Conference, 2009.
- [28] E. Keogh, and M. Pazzani, Scaling up dynamic time scaling up dynamic time, in: Proceedings of ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2005.
- [29] V. Dizan, F. Thierry, L. Christian, Growing hidden Markov models: an incremental tool for learning and predicting human and vehicle motion, Int. J. Robot. Res. (2009) 11–12.
- [30] A. Zaharescu, and R. Wildes, Anomalous behaviour detection using spatio-temporal oriented energies, subset inclusion histogram comparison and event-driven processing, in: Proceedings of European Conference on Computer Vision, 2010, pp. 563–576.
- [31] J. Wright, A.Y. Yang, A. Ganesh, S.S. Sastry, Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Patt. Anal. Mach. Intell. (2009) 210–227.
- [32] Z.J. Han, J.B. Jiao, B.C. Zhang, Q.X. Ye, J.Z. Liu, Visual object tracking via sample-based adaptive sparse representation (AdaSR), Patt. Recog. (2011) 2170–2183.
- [33] T. Hastie, R. Tibshirani, J. Friedman, The Elements of Statistical Learning, Springer Series in Statistics, Springer, 2001.
- [34] M. Elad, A.M. Bruckstein, A generalized uncertainty principle and sparse representation in pairs of bases, IEEE Trans. Inform. Theory (2002) 2558–2567.
- [35] M. Figueiredo, R. Nowak, S. Wright, Gradient Projection for sparse reconstruction: application to compressed sensing and other inverse problems, IEEE J. Sel. Top. Signal Process. (2007) 586–597.
- [36] D. Donoho, and Y. Tsaig, Fast solution of L1-norm minimization problems when the solution may be sparse, <<http://www.stanford.edu/tsaig/research.html>>, 2006.
- [37] R. Tibshirani, Regression shrinkage and selection via the LASSO, J. Roy. Stat. Soc. B (1996) 267–288.
- [38] CAVIAR test case scenarios. URL <<http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>>.
- [39] NGSIM test case scenarios. URL <<http://ngsim.camsys.com/>>.



Ce Li received her B.S. degree in software engineering from Tianjin University (TJU), Tianjin, in 2008. She received her M.S. degrees from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2012. Since 2012, she has been a Ph.D. candidate of the University of Chinese Academy of Sciences, Beijing, China. Her research interests include visual behavior analysis and intelligent surveillance etc.



Zhenjun Han received his B.S. degree in software engineering from Tianjin University (TJU), Tianjin, in 2006. And he received his M.S. and Ph.D. degrees from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2009 and 2012, respectively. Since 2012, he has been a Postdoctoral researcher of the Graduate University of Chinese Academy of Sciences. His research interests include object tracking, behavior analysis in intelligent surveillance.



Qixiang Ye received his B.S. and M.S. degrees in mechanical & electronic engineering from Harbin Institute of Technology (HIT), Harbin, China, in 1999 and in 2001, respectively. He received his Ph.D. degree from the Institute of Computing Technology, Chinese Academy of Sciences in 2006. Since 2009, he has been an associate professor at the Graduate University of the Chinese Academy of Sciences, Beijing. His research interests include image processing, pattern recognition, intelligent systems, etc. Dr. Ye won the “Sony Outstanding Paper Award” in 2005.



Jianbin Jiao received his B.S., M.S. and Ph.D. degrees in mechanical and electronic engineering from Harbin Institute of Technology (HIT), Harbin, China, in 1989, 1992, and 1995, respectively. From 1997 to 2005, he was an associate professor of HIT. Dr. Jiao had been a visiting professor at Kyushu Institute of Technology in 2009 and a visiting scholar at the University of Nevada, Las Vegas from 2002 to 2005. Since 2006, he has been a professor of the Graduate University of Chinese Academy of Sciences, Beijing. Dr. Jiao is also a “Science-100 professor” of the Chinese Academy of Sciences. His research interests include image processing, pattern recognition, intelligent systems, etc.