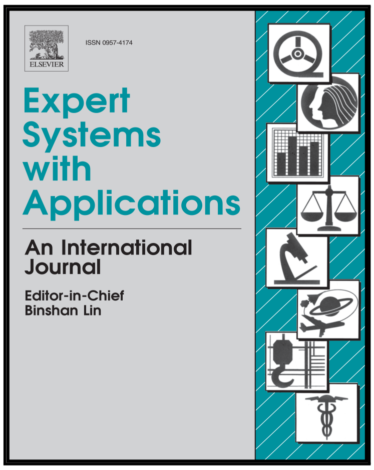
Accepted Manuscript



A Novel Sparse Representation Model for Pedestrian Abnormal Trajectory Understanding

Zhijun Chen , Hao Cai , Yishi Zhang , Chaozhong Wu ,
Mengchao Mu , Zhixiong Li , Miguel Angel Sotelo

PII: S0957-4174(19)30441-5
DOI: https://doi.org/10.1016/j.eswa.2019.06.041
Reference: ESWA 12753

To appear in: *Expert Systems With Applications*

Received date: 17 September 2018
Revised date: 20 June 2019
Accepted date: 20 June 2019

Please cite this article as: Zhijun Chen , Hao Cai , Yishi Zhang , Chaozhong Wu , Mengchao Mu , Zhixiong Li , Miguel Angel Sotelo , A Novel Sparse Representation Model for Pedestrian Abnormal Trajectory Understanding, *Expert Systems With Applications* (2019), doi: https://doi.org/10.1016/j.eswa.2019.06.041

HIGHLIGHTS

- A novel sparse representation method for pedestrian trajectory abnormal analysis.
- Utilizing Lp-regularization ($0 < p < 1$) to get sparser solutions.
- An effective solver for the proposed method with EM algorithm and entropy.

A Novel Sparse Representation Model for Pedestrian Abnormal Trajectory Understanding

Zhijun Chen, Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, P. R. China 430063, Email: chenzj556@whut.edu.cn
Hao Cai, Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, P. R. China 430063, Email: caihao@whut.edu.cn
Yishi Zhang (Corresponding author), School of Management, Huazhong University of Science and Technology, Wuhan, P. R. China 430074, Email: easezh@126.com
Chaozhong Wu, Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, P. R. China 430063, Email: wucz@whut.edu.cn
Mengchao Mu, Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan, P. R. China 430063, Email: 924121889@whut.edu.cn
Zhixiong Li, School of Mechanical, Materials, Mechatronic and Biomedical Engineering, University of Wollongong, Wollongong, NSW 2522, Australia, Email: zhixiong.li@ieee.org
Miguel Angel Sotelo, Computer Engineering Department, University of Alcalá, Alcalá de Henares (Madrid), Spain, Email: miguel.sotelo@uah.es

ABSTRACT
Pedestrian abnormal trajectory understanding based on video surveillance systems can improve public safety. However, manually identifying pedestrian abnormal trajectories is usually prohibitive workload. The objective of this study is to propose an automatic method for understanding pedestrian abnormal trajectory. An improved sparse representation model, namely information entropy constrained trajectory representation method (IECTR), is developed for pedestrian trajectory classification. It aims to reduce the entropy for trajectory representation and to obtain superior analyzing results. In the proposed method, the orthogonal matching pursuit (OMP) is embedded in the expectation maximization (EM) method to iteratively obtain the selection probabilities and the sparse coefficients. In addition, the lower-bound sparser condition of L_p-minimization ($0 < p < 1$) is applied in the proposed method to guarantee salient solutions. In order to validate the performance and effectiveness of the proposed method, classification experiments are conducted using five pedestrian trajectory datasets. The results show that the identification accuracy of the proposed method is superior to the compared methods, including naïve Bayes classifier (NBC), support vector machine (SVM), k-nearest neighbor (kNN), and typical sparse representation-based methods.

Keywords: pedestrian behavior; sparse representation; L_p-norm; information entropy; trajectory understanding

INTRODUCTION

In recent years, with an increasing demand for security issue, traffic safety that is an important part of security issue has attracted more and more attentions. Duo to rapid development of image processing techniques, storage technology, and artificial intelligence, the applications of video analysis to enhance traffic safety become more and more popular. Pedestrian behavior analysis based on video surveillance is one of the important affects for public transportation safety. For example, public transport systems in the central London were attacked during the morning rush hour. Another example, Madrid commuter rail network was attacked and the explosions killed 191 people. If the video surveillance system can automatically detect abnormal pedestrian behaviors in the public areas beforehand, the bombing tragedies of Madrid and London may be effectively prevented.

To date, airports, bus station, and other public places are widely equipped with video surveillance system. However, many of the existing video surveillance system only have the remote monitoring function. Although some alarm functions (e.g. fire alarm) are developed with information technology, automatic analysis and interpretation of pedestrian behaviors are not fully included in the video surveillance system. If the pedestrian abnormal behaviors are analyzed manually, it is prohibitive workload due to the vast amount of traffic video data. Therefore, it is necessary to develop an automatic understanding method for pedestrian abnormal behaviors.

Pedestrian abnormal behavior analysis is one of the hotspots and difficulties. It concerns detection, and tracking of pedestrian and understanding of pedestrian trajectory from video sequences. Due to the advancement of video camera technologies and the availability of more sophisticated computer vision algorithms, the pedestrian detection and tracking technology has become more and more mature (Viola et al., 2003; Yao and Deng, 2012; Wang et al., 2014; He et al., 2017; Xu et al., 2018). Therefore, we only focus on understanding method of pedestrian abnormal behaviors in this study.

In recent years, many methods are proposed to understand pedestrian abnormal behavior based on video surveillance. Based on motion direction of head, hands, and other parts of pedestrian body extracted from video, several methods (such as hierarchical Bayesian network and dynamic Bayesian network) are employed to understand pedestrian abnormal behaviors. However, the pedestrian body features are difficult to extract because they are largely influenced by various environmental factors. The performance of understanding pedestrian abnormal behaviors is seriously affected. An object motion characterization can be described using motion trajectory that contain a sequence of position points. The motion trajectory has been widely employed to identify object motion patterns, action recognition, and so on. Therefore, many methods are proposed to understand pedestrian abnormal behaviors using pedestrian trajectory. For example, cluster method is widely applied. However, the performance and effective of these understand methods are often affected by incompletion and distortion of trajectories. The sparse representation model is a novel idea to understand or analyze object behavior. Sparse representation models typically used to tackle the analysis task can be traced to Wright et al. (2009), of which the application field is face classification. Sparse representation model has been proven to be a powerful classifier even if the object feature is missing and distortion (Wright et al., 2009; Chong et al., 2011; Okyere et al., 2014; Chen et al., 2018).

In this study, a sparse representation-based classification model with L_p-minimization is proposed for pedestrian anomaly trajectory understanding. An information entropy constrained trajectory representation (IECTR) method is then developed to solve the proposed model. The proposed method explicitly pays attention to the entropy of the distribution of selection frequency and modifies the sparse representation model by adding an entropy constraint to reduce the coding costs. In order to validate the performance and effectiveness of the proposed method, five datasets, including Bus station dataset1, Bus station dataset2, Airport dataset, Square dataset, and Railway station dataset, are used in our experimental.

The remainder of the paper is organized as follows: Section II introduces the sparse representation of the trajectories and the proposed method. Then our pedestrian abnormal trajectory understanding framework is proposed in section III. In section IV, verification of the proposed method is given, followed by the conclusion and future work section.

Contributions

The main contributions of this study are threefold:

(1) An improved sparse representation-based classification model with L_p-minimization is proposed for pedestrian abnormal trajectory understanding.
(2) The Information Entropy is introduced in the proposed model to reduce the representation coding costs.
(3) A novel solver based on the EM algorithm and the lower-bound sparser condition of L_p-minimization is proposed to solve the proposed model.

Related Works

Over the last decade, there has been a growing interest within the computer vision and machine learning communities in the problem of understanding pedestrian abnormal behavior based on video surveillance. These studies can be divided into two aspects: Pedestrian characteristics analysis and Trajectory analysis. For pedestrian characteristics analysis, the key is pedestrian behavior representation that needs to extract the characteristics of pedestrian silhouettes, body parts, and so on. For example, motion direction of head, hands, and other parts of pedestrian body are extracted to understand the pedestrian abnormal behavior such as fighting and violence (Datta et al., 2003). A hierarchical Bayesian network is also employed to understand pedestrian abnormal behavior. First, the poses of simultaneously tracked body parts and overall body pose are estimated at the low and high level of the Bayesian network, respectively. Then, a dynamic Bayesian network (DBN) is applied to analyze the evolution of the poses of the multiple body parts (Park and Aggarwal, 2004). Similarly, a coupled hidden Markov model (CHMM) is proposed to understand pedestrian behaviors based on head position, body and motion information (Oliver et al., 1999). Although using pedestrian body features can identify the abnormal behavior, the pedestrian features are difficult to extract and largely influenced by various environmental factors, such as weather, tree, and so on.

For the trajectory analysis methods, the abnormal behaviors are identified during the course of pedestrian motions. For instance, Xiang and Gong (2008) propose an unsupervised learning method where likelihood ratio test (LRT) is used to online understand the abnormal behaviors. The iterative Altruistic Vector Quantization algorithm is used to robustly cluster trajectories in Mecocci and Pannozzo (2005). The abnormal trajectory understanding is based on fitting a spatial Gaussian on each prototype and statistically checking the fit of new trajectory samples. A two-stage fuzzy k-means is proposed to cluster trajectory in Hu et al. (2006). In their study, trajectories are firstly clustered in the spatial domain, and then each cluster is sub-clustered in the temporal domain. Then a threshold is set up to detect abnormal trajectory. Junejo et al. (2004) propose graph cuts to detect abnormal trajectory based on the Hausdorff distance measure.

The sparse representation model is a novel idea to understand or analyze object behavior. Li et al. (2013) propose a sparse reconstruction (representation) method based on trajectory to detect pedestrian abnormal behaviors in video surveillance system. The fundamental underlying assumption of spare representation model is that any new sample can be approximately linear combined by training samples from a dictionary. The key problem of sparse representation model is to solve the L_0-minimization problem. This problem for underdetermined system of linear equations is strongly NP-hard (Amaldi and Kann, 1998). Based on theory in field of compress sensing, we can learn that if the sparse solution of L_0-minimization problem is sufficiently sparse, solving the L_0-minimization problem can be equivalently transformed into solving the L_1-minimization problem (Candes and Tao, 2005). Therefore, L_1-minimization is widely applied as an approximation of L_0-minimization due to its convexity. However, the solutions from L_1-minimization are sometimes not sparse-enough in practice, which may impair the representation and discrimination abilities of the corresponding models. To this end, a tighter relaxation namely L_p-minimization ($0 < p < 1$) is introduced to obtain the sparser solutions (Chen et al., 2010; Zhang et al., 2017). Because the problem of L_p-minimization is non-convex, the existing convex optimization solvers cannot be employed to get the optima within polynomial time. Recent study has proved that the local optimal solution of L_p-minimization can be considered as the global optimal solution under some weak conditions (Fan and Li, 2001; Chartand, 2007; Nikolova, 2006; Chen et al., 2010). Sparse representation based on L_p-minimization is currently being applied in many fields: Chen et al. (2017) propose a sparse representation-based classification model with L_p-minimization ($0 < p < 1$) for vehicle behavior learning. Zhang et al. (2017) propose a large-scale robust semi-supervised classification with an $L_{2,p}$-norm-based framework, where p is smaller than 1. In addition, although getting sparser solutions is better for trajectory learning, the distribution of trajectory selection frequency is neglected in the extant work, which will lead to high entropy for the model and thus may increase the total cost of representation (reconstruction) and at the same time impair the classification ability from the perspective of information theory.

TABLE 1 Summarization of previous studies

Reference(s)	Algorithm	Feature
Datta et al.(2003)	Motion analyze	Pedestrian characteristics
Park and Aggarwal (2004)	DBN	Pedestrian characteristics
Oliver et al.(1999)	CHMM	Pedestrian characteristics
Xiang and Gong(2008)	LRT	Pedestrian trajectory
Mecocci and Pannozzo (2005)	Iterative Altruistic Vector Quantization	Pedestrian trajectory
Hu et al.(2006)	A two-stage fuzzy k-means	Pedestrian trajectory
Junejo et al.(2004)	Cluster	Pedestrian trajectory
Li et al.(2013)	Sparse representation method	Pedestrian trajectory

METHODOLOGY

Sparse Representation for Pedestrian Trajectories

For the sake of convenience, we first show the notations used in this paper as follows. For $\mathbf{x} \in \mathbb{R}^m$, $\|\mathbf{x}\|_p$ denotes L_p-norm, $\|\mathbf{x}\|$ denotes L_2 -norm throughout this paper. Signal representation is a basic problem in the fields where image compression plays an important role. Sparse representation with respect to an over-complete dictionary has been studied with growing recently, especially in the trajectory detection. Let matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ be the training dictionary where each column denotes a training trajectory, a test trajectory $\mathbf{y} \in \mathbb{R}^m$ can be expressed by a linear combination of the training trajectories as following

$$\mathbf{y} = \mathbf{A}\boldsymbol{\sigma} \in \mathbb{R}^m,$$

where $\boldsymbol{\sigma} \in \mathbb{R}^n$ is the weight vector that contains the weight of each training samples in the dictionary (i.e. each column of A) for test trajectory fitting. Figure 1 shows an example of the sparse representation technique used in pedestrian trajectory understanding. In Figure 1, Dictionary A contains 3 subgroups namely A_1, A_2, and A_3, where each subgroup consists of the training trajectories that belong to the same class. The weight vector (coefficient vector) $\sigma = \{0,0,0,a_1,a_2,a_3,a_4,0,0,0\}^T$ corresponding to the test trajectory y implies that y can be represented only by the training trajectories in A_2. The class label of y can thus be considered as the same as those in A_2.

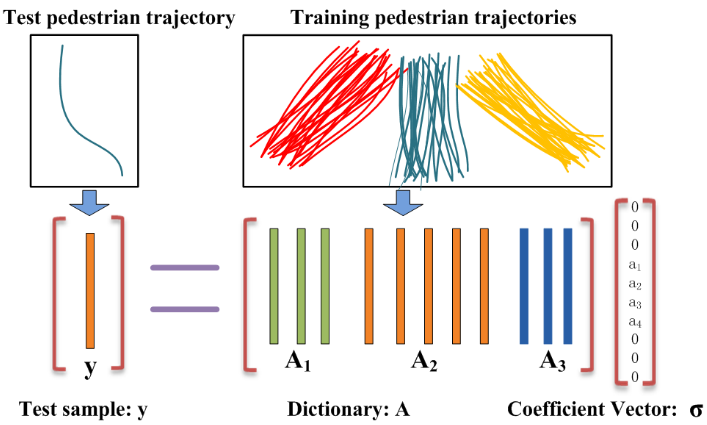


Figure 1 An example of the sparse representation technique applied in pedestrian trajectories understanding.

Sparse representation aims at seeking approximations of $\mathbf{y}$ with the atoms in $\mathbf{A}$ as few as possible. This can be implemented by solving the following problem

$$\min_{\boldsymbol{\sigma}} \|\boldsymbol{\sigma}\|_0 \quad \text{s.t. } \|\mathbf{y}-\mathbf{A}\boldsymbol{\sigma}\|_2 \le \varepsilon \tag{1}$$

Note that the objective function of model (1) is nonconvex. In fact, model (1) is well-known to be strongly NP-hard (Amaldi and Kann, 1998). To this end, model (1) are often transformed to model (2),

$$\min_{\sigma} \|\boldsymbol{\sigma}\|_p \qquad s.t. \|\boldsymbol{y}-\boldsymbol{A\sigma}\|_2 \le \varepsilon$$

(2)

where p is usually set as 1 or 2, and at this time model (2) is in fact a convex approximation of (1). Many methods have been developed to efficiently find solution of model (2) including greedy algorithms and convex relaxation techniques. Taking into account the trade-off between approximation capability and computational complexity, greedy algorithms like matching pursuit (MP) and orthogonal matching pursuit (OMP) (Mo et al., 2014) are widely applied. MP selects one atom that reduces the residual most in each iteration step. OMP differs from MP in that, after one atom is selected every iteration, the support (position indices of nonzero entries) and the projection values are updated by orthogonaling the residual with respect to the selected atoms. This process is repeated until the number of nonzero entries in the solution reaches the given upper limit or the residual is small enough. OMP is thus applied in our method.

In addition, much study (e.g. Candes and Tao, 2005; Chartand, 2007; Fan and Li., 2001; Nikolova, 2006) has been evidenced that using L_p-norm ($0 < p < 1$) in the regularization term can usually find sparser solutions than using L_1-norm. Although finding the exact solutions of the optimization formulation with L_p-norm ($0 < p < 1$) is still NP-hard (Chen et al., 2010), a theorem of the lower bound of L_p-minimization is introduced to guarantee sparser solution under weaker conditions than L_1-approximation of the original model (i.e. model (1)). Inspired by this, we apply the theorem in our method to help find sparser solutions. The theorem is described as follows.

Theorem 1 *(Chen et al., 2010)***:** a lower bound $L = \left(\frac{\lambda p}{2\|A\|\sqrt{f(\sigma)}}\right)^{\frac{1}{1-p}}$, for the absolute value of nonzero entries in a local optimal solution σ_i of model (2), that is

$$\text{for any } i \in N,\ L \le |\sigma_i|, \qquad \text{if } \sigma_i \ne 0$$

it can be written as following,

$$\text{for any } i \in N,\ \sigma_i \in (-L, L) \Rightarrow \sigma_i = 0$$

According to the Theorem 1, if local minimizer σ_i of L_p-minimization satisfies the condition namely $L \le |\sigma_i|$, the corresponding entry in the global minimizer (i.e. the exact optimal solution of model (1)) is 0. Therefore, it is possible to make the solution sparser and closer to the global minimizer by utilizing theorem 1.

Sparse Representation with Information Entropy

Trajectory dictionary often contains rich information. In order to get a better approximation, more training samples will be selected and thus needs a higher coding cost for representation. According to the information theory, the total bit of the representation of the trajectory depends on the product of the information entropy and the number of samples. The nonzero entries of the coefficient matrix $\Sigma = [\boldsymbol{\sigma}_1, \ldots, \boldsymbol{\sigma}_K]^T$ corresponding to the test trajectory matrix $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_K]^T$ can thus be quantized to reduce the entropy. Otherwise, the distribution of nonzero entries in Σ may vary and thus may lead to high coding costs although sparse samples are selected. To this end, an information entropy constrained trajectory representation method (denoted as IECTR) is proposed in this section. Specifically, we first let

$$\mathbf{S} = \begin{bmatrix} s_{11} & \cdots & s_{1n} \\ \cdots & \cdots & \cdots \\ s_{K1} & \cdots & s_{Kn} \end{bmatrix}$$

where $\mathbf{s}_i = [s_{1i}, \ldots, s_{Ki}]^T$ records the positions of the nonzero entries in $\Sigma_i = [\sigma_{1i}, \ldots \sigma_{Ki}]^T$. IECTR contains two constraints, namely the traditional constraint $\|\mathbf{y} - \mathbf{A}\boldsymbol{\sigma}\| \le \varepsilon$ in model (2) and a new one namely $H(\mathbf{S}) \le h$, where h is the upper bound of the information entropy for the indices $\mathbf{S}$. $H(\mathbf{S})$ is defined as

$$H(\mathbf{S}) = -\sum_{\mathbf{s}_i \in \mathbf{S}} p(\mathbf{s}_i) \cdot \log_2 p(\mathbf{s}_i) \tag{3}$$

where $p(\mathbf{s}_i)$ is calculated as

$$p(\mathbf{s}_i) = \frac{\|\mathbf{s}_i\|_1}{\|\mathbf{S}\|_1} = \frac{\|\Sigma_i\|_0}{\|\Sigma\|_0} \tag{4}$$

To make the model consistent with the fast solving algorithm like OMP, the constraint of entropy is taken into the objective function by introducing a Lagrangian multiplier as follows

$$\min \|\boldsymbol{\sigma}\|_p - \lambda \sum_{\Sigma_i \in \Sigma} \frac{\|\Sigma_i\|_0}{\|\Sigma\|_0} \cdot \log_2 \frac{\|\Sigma_i\|_0}{\|\Sigma\|_0} \qquad s.t. \|\boldsymbol{y} - \boldsymbol{A\sigma}\|_2 \le \varepsilon$$

(5)

However, model (5) is hard to be solved not only lies in its NP-hardness but also because Σ in $H(\mathbf{S})$ is made up of the calculated Σ. To tackle this contradiction as well as to make the problem tractable, an OMP-EM mixture method (Hild et al., 2008) are applied to get the near-optimal solutions. The detailed procedure can be described as follows: First, taking $\mathbf{p} = [p(\mathbf{s}_1),..., p(\mathbf{s}_n)]^T$ as a constant to calculate Σ using OMP, then in the next step, using Σ to update **p**, in such a way as to iteratively update **p** and Σ. To start, **p** should be first initialized. In this paper, the similarity between $\mathbf{y}_i$ and the training sample $\mathbf{a}_j$

$$\mathrm{sim}_{ij} = \frac{\mathbf{y}_i \cdot \mathbf{a}_j}{\|\mathbf{y}_i\| \cdot \|\mathbf{a}_j\|} \tag{6}$$

are applied for initialization. Thus, a similarity matrix can be obtained

$$SI = \begin{bmatrix} \mathrm{sim}_{11} & \cdots & \mathrm{sim}_{1n} \\ \cdots & \cdots & \cdots \\ \mathrm{sim}_{K1} & \cdots & \mathrm{sim}_{Kn} \end{bmatrix} \tag{7}$$

and the initial probability can be calculated by

$$p_j = \mathrm{sim}_j / \mathrm{Sim} \tag{8}$$

where $sim_j = \sum_{i=1}^{K} sim_{ij}$ and $\mathrm{Sim} = \sum_{j=1}^{n} \sum_{i=1}^{K} sim_{ij}$.

The pseudo code of IECTR is given in Algorithm 1.

Algorithm 1 Information Entropy Constrained Trajectory Representation (IECTR) method

---

Inputs: Test trajectory set $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_K]^T$, the dictionary **A**

1. **Initialization:** $t = 0$, calculate $\mathbf{p}^0 = [p_1^0,..., p_n^0]$ according to Eqs. (7) and (8)
2. **Make iteration:** Perform the following steps and increase t by 1
3. **Initialization:** $\mathrm{Re} = \varnothing$, $\Sigma^t = \varnothing$, $\mathbf{S}^t = \varnothing$
4. **for** all $\mathbf{y}_k \in \mathbf{Y}$, **do**
5. **Initialization:** $\mathbf{r}_k^0 = \mathbf{y}_k$, $[s_{k1},..., s_{kn}]^T = \mathbf{0}$, $i = 0$
6. **Make iteration:** Perform the following steps and increase i by 1
7. Find $j := \arg\min_{\substack{1 \le j \le n \\ s_{kj} = 0}} \left\| \mathbf{a}_j \cdot \mathbf{a}_j^T \mathbf{r}_k^i / \|\mathbf{a}_j\|_2^2 - \mathbf{r}_k^i \right\|_2^2 - \lambda \cdot p_j^t \log_2 p_j^t$ and set $s_{kj} := 1$
8. Find the minimizer of $\mathbf{r}_k^{i+1} := \|\mathbf{y}_k - \mathbf{A}\boldsymbol{\sigma}_k^i\|_2^2$ s.t. $\mathrm{support}(\boldsymbol{\sigma}_k^i) = [s_{k1},..., s_{kn}]^T$
9. If $\|\mathbf{r}_k^{i+1}\|_2 < \tau$, stop. Otherwise conduct another iteration.
10. **for** each $\sigma_l \in \boldsymbol{\sigma}_k^i$ **do**

11. Calculate L according to Theorem 1
12. **if** $\sigma_l \in (-L, L)$ then
13. $\sigma_l = 0$
14. end **if**
15. end **for**
16. $\Sigma^t = \Sigma^t \cup \{\boldsymbol{\sigma}_k^i\}$, $\mathbf{S}^t = \mathbf{S}^t \cup \{\text{support}(\boldsymbol{\sigma}_k^i)\}$, $\text{Re} = \text{Re} \cup \{\mathbf{r}_k^{i+1}\}$
17. end **for**
18. Update $\mathbf{p}^{t+1}$ using $\mathbf{S}^t$ according to Eq. (4).
19. If $||\text{Re}||_F$ is not changed, stop. Otherwise conduct another iteration.

Outputs: coefficient matrix Σ, indices **S**

Algorithm IECTR consists of four loops, namely EM loop (lines 2 -- 19), testing sample loop (lines 4 -- 17), local minimizer searching loop (lines 6 -- 9), and local minimizer modification part (lines 10 -- 15). The iteration times of the first loop T depends on the convergence of Re. The second loop iterates K times where K is the number of the test samples. And it is easy to find that the worst iteration case of the third loop is no more than $O(n^2)$ and the last part iterates n times. Therefore, the iterative complexity of IECTR can be finally obtained as $O(TKn^2)$.

The Proposed Method for Pedestrian Abnormal Trajectory Learning

Pedestrian Trajectory Classification

The proposed method for pedestrian trajectory learning based on IECTR will be described in this section. Pedestrian trajectory learning includes two aspects: Trajectory classification and abnormal trajectory detection.

Trajectory dictionary requires all training samples in the same dimension, while the length of the original pedestrian trajectories (i.e. the raw data) are usually different due to various pedestrian behaviors. To this end, a frequently used technique called least-squares cubic spline curves approximation (LCSCA) (Li et al., 2013) is employed to normalize all pedestrian trajectories into the same dimension. The control points obtained by LCSCA can discretely represent the shape of complex curves and thus will be used to extract fixed-length vectors as pedestrian trajectory representation.

Using the proposed IETCR method, the sparse coefficient matrix $\mathbf{\Sigma}$ can be obtained. Suppose a test trajectory **y** belongs to the *i*th class, it is more likely that the average residual between **y** and the training trajectories (i.e. the atoms) belong to the same class in dictionary **A** are smaller than that between **y** and those belong to any other class. Thus a class residual error set **R** could be introduced for classification. To identify the class that **y** belongs to, we first denote the class residual error vector set **R** of the test trajectory **y** as

$$\mathbf{R} = \left[R_{C_1}, \ldots, R_{C_k}\right] = \left[\left\|\mathbf{A}\boldsymbol{\sigma}_{C_1} - \mathbf{y}\right\|, \ldots, \left\|\mathbf{A}\boldsymbol{\sigma}_{C_k} - \mathbf{y}\right\|\right] \tag{9}$$

where $\boldsymbol{\sigma}_{C_1}, \ldots, \boldsymbol{\sigma}_{C_k}$ denote the coefficient vectors correspond from $\mathbf{C}_1$ to $\mathbf{C}_k$, respectively. When **R** is obtained, we take the *i*th class of the minimum R_{C_i} as the class of test trajectory. That is, we can simply apply the following equation to identify the class of **y**.

$$Class(\mathbf{y}) = \arg\min_i (R_{C_i}) \qquad i \in (1, \dots, k) \tag{10}$$

Anomaly Detection

Since abnormal training trajectories in the dictionary can be taken as the samples with the class label of 'abnormal', abnormal test trajectories can be classified during the trajectory classification process introduced in previous section. Thus anomaly detection can be actually conducted in the classification process. However, unlike the normal trajectories, classification of abnormal ones may fall into a dilemma since the dictionary may contain statistically insignificant or even no abnormal training samples. Under this circumstance, anomaly detection should be conducted in the context of unsupervised learning. The outlier rejection rule (Wright et al., 2009) is applied in our framework to detect anomaly when the supervised learning framework is invalid. More specifically, Sparsity Concentration Index (SCI) metric is utilized. The SCI of a coefficient vector x is defined as

$$\mathrm{SCI}(\mathbf{x}) = \frac{k \cdot \max_i \|g_i(\mathbf{x})\|_1 / \|\mathbf{x}\|_1 - 1}{k-1} \in [0,1]. \tag{11}$$

where $g_i(\mathbf{x})$ is the vector that contains all nonzero elements corresponding to *i*th class in **x**. We now analyze two extreme cases of the value of SCI: If SCI(**x**) = 1, the test trajectory **y** is represented using only trajectories from a single class, thus it can be classified to that class. On the contrary, SCI(**x**) = 0 means the sparse coefficients are distributed evenly on all classes. In other words, it is more likely that y does not belong to any presented classes. According to the outlier rejection rule, it could be determined as an abnormal trajectory. A threshold $\lambda \in (0,1)$ is selected and **y** is identified to be an abnormal trajectory if $\mathrm{SCI}(\mathbf{y}) < \lambda$.

EXPERIMENTAL RESULTS AND DISSCUSION

In order to validate the performance and effectiveness of the proposed method, five datasets, including Bus station dataset1, Bus station dataset2, Airport dataset, Square dataset, and Railway station dataset, are used in this study. Bus station dataset1 is collected in Brown Hill, Liverpool, UK. Bus station dataset2, Airport dataset, Square dataset, and Railway station dataset that consist of several pedestrian videos are all collected by our lab. In order to obtain pedestrian trajectories that can be used for the purpose of training and testing, the frequently used methods, namely Mask R-CNN (He et al., 2017) and Kalman filter, are employed as the preprocessing methods in our experiments. Mask R-CNN extends Faster R-CNN by adding a branch for predicting segmentation masks on each region of interest. The effectiveness of Mask R-CNN has been guaranteed in several recent works (e.g. He et al., 2017). Figure 2 shows the preprocessing steps: First, Mask R-CNN and Kalman filter are employed to detect and track pedestrian from a set of video clips.. Then, the extracted trajectories are randomly gathered as a training dataset (i.e. a dictionary) and a test dataset, respectively, and finally transformed into the same dimensions by LCSCA.

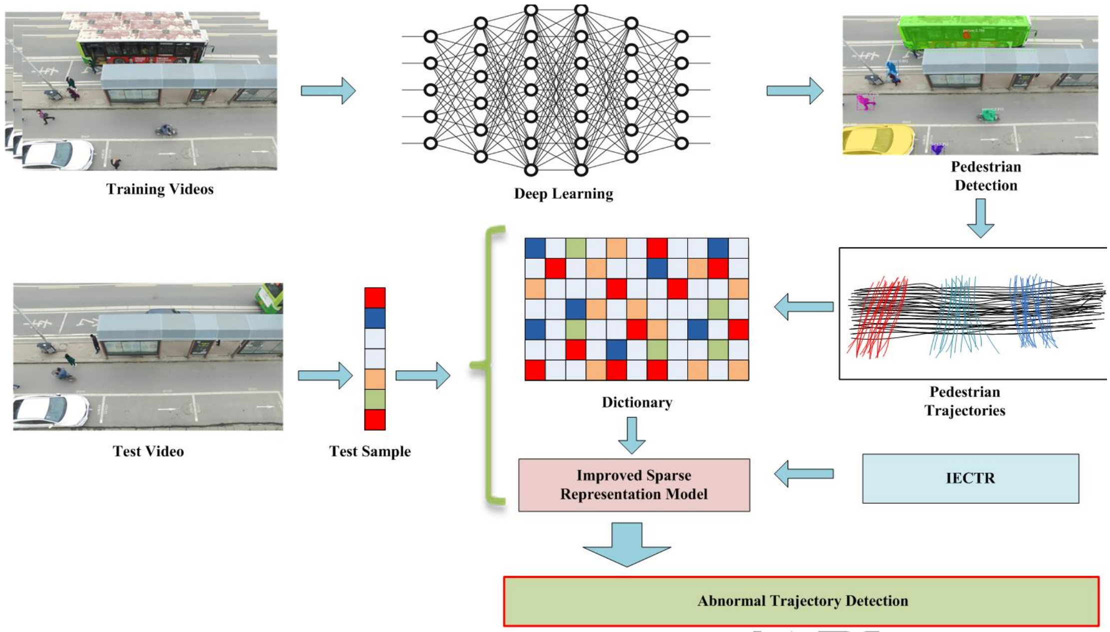


Figure 2 The preprocessing flow for pedestrian trajectory extraction and collection.

For Bus station dataset1, the normal behavior classes and abnormal behavior classes are manually identified and labeled for each trajectory: 2937 training trajectories (including 2850 normal and 87 abnormal trajectories) and 1390 test trajectories with 1300 normal and 90 abnormal trajectories are finally obtained in our experiments. For Bus station dataset2, 5705 training trajectories (including 5550 normal and 155 abnormal trajectories) and 2105 test trajectories with 2000 normal and 105 abnormal trajectories are finally obtained in our experiments. For Airport dataset, 2655 normal and 15 abnormal trajectories are obtained from training video clips, and 1200 normal and 35 abnormal trajectories are used as test samples. For Square dataset, different normal trajectory classes and abnormal trajectory classes are manually identified and labeled for each trajectory: 3500 normal and 82 abnormal trajectories are obtained from training video clips, and 1200 normal and 40 abnormal trajectories are used as test samples. For Railway station dataset, 5350 training trajectories (including 5000 normal and 350 abnormal trajectories) and 900 test trajectories with 870 normal and 80 abnormal trajectories are finally obtained for the experiments. General information of the five datasets is summarized in Table 2.

In this study, several representative classifiers, including naïve Bayes classifier (NBC), support vector machine (SVM), k-nearest neighbor (kNN), sparse representation with l^1-relaxation (namely L_1-minimization, hereafter it is denoted as SR-l^1), sparse representation with l^2-relaxation (SR-l^2), and sparse representation with $l^{0.5}$-relaxation (SR-$l^{0.5}$) are used as the compared methods to verify the performance and effectiveness of the proposed method. All the experiments are conducted on a 2.60 GHz CPU, 8GB RAM personal computer with Windows 7.

NBC (Chen et al., 2015) is a simple probabilistic classifier with an assumption of conditional independence among variables, i.e., the presence (or absence) of a particular variable of a class is unrelated to the presence (or absence) of any other variable. It only requires a small amount of training data to estimate the parameters necessary for classification.

TABLE 2 General information of training and test datasets

	Dataset	Normal trajectory	Abnormal trajectory	Normal class	Abnormal class
Training	Bus station dataset1	2850	87	3	2
	Bus station dataset2	5550	155	7	3
	Airport dataset	2655	15	4	2
	Square dataset	3500	82	3	3
	Railway station dataset	5000	350	4	2
Test	Bus station dataset1	1300	90	3	2
	Bus station dataset2	2000	105	7	3
	Airport dataset	1200	35	4	2
	Square dataset	1200	40	3	3
	Railway station dataset	870	80	4	2

kNN (Aha et al., 1991) is a supervised nonlinear classifier. Given a parameter k and a sample d, the kNN technique considers k training examples closest to *d* in terms of their distance (e.g., Euclidean distance) in the feature space, and then predicts the class of d as the dominant real class among those *k* neighbors by a majority voting manner. A value of $k = 1$ neighbor has been used in our experiments

SVM (Chen et al., 2017) is proved to be very effective in multiple application fields. The fundamental principles of the SVM can be described as finding a best separating hyperplane between two classes of training samples and using it to classify the test samples.

SR-l^1 (Li et al., 2013): Sparse representation model with l^1-relaxation is used for pedestrian trajectory classification.

SR-l^2 (Zhang et al., 2011): Sparse representation model with l^2-relaxation is used for pedestrian trajectory classification.

SR-$l^{0.5}$ (Chen et al., 2010): Sparse representation model with $l^{0.5}$-relaxation is used for pedestrian trajectory classification.

In this study, NBC, kNN and SVM are conducted in the Weka environment. The Gaussian RBF kernel is applied for SVM, and $k = 3$ is set for kNN. SR-l^1, SR-l^2, SR-$l^{0.5}$, and the proposed method are implemented in the Matlab environment. The classification accuracy rate, true positive rate (TPR), and false positive rate (FPR) are applied as the evaluation metrics, where TPR is defined as the ratio of the number of the correctly-identified abnormal trajectories to the total number of abnormal trajectories. FPR is defined as the ratio of the number of the incorrectly-identified abnormal trajectories to the total number of abnormal trajectories. For all normal trajectory classes, we report the classification accuracy of each selected method; and for the abnormal trajectory class, we report TPR and FPR of each selected method.

Figure 3 shows an example of pedestrian detection for Bus station dataset1 (Brown Hill, Liverpool, UK). Figure 4 shows examples of the normal trajectory classes and abnormal trajectory class in Bus station dataset1. In Bus station dataset1, pedestrian normal trajectories are further divided into three classes “N1”, “N2”, “N3”.

Table 3 shows the results of abnormal trajectory detection using NBC, kNN, SVM, SR-l^1, SR-l^2, SR-$l^{0.5}$, and the proposed method on Bus station dataset1, respectively.

Figure 3 An example of pedestrian detection for Bus station dataset1 (Brown Hill, Liverpool, UK)

(a) **(b)**

Figure 4 Examples of the normal trajectory classes and abnormal trajectory class in Bus station dataset1: (a) the normal trajectory classes, (b) the abnormal trajectory class.

TABLE 3 Results of pedestrian trajectory identification on Bus station dataset1

Methods	Accuracy of Normal Trajectory Understanding			Accuracy of Abnormal Trajectory Understanding	
	N1	N2	N3	TPR	FPR
NBC	80.6%	82.8%	80.3%	55.6%	2.2%
SVM	82.2%	84.2%	79.7%	57.8%	2.4%
kNN	82.6%	83.8%	80.7%	62.2%	2.7%
SR-l^1	83.4%	86%	81.6%	81.1%	0.9%
SR-l^2	84.6%	82.2%	77.7%	80%	1.1%
SR-$l^{0.5}$	87.4%	86.6%	81%	83.3%	0.8%
Our method	90.2%	86.6%	83.7%	86.7%	0.2%

As can be seen from Table 3, the accuracies of normal trajectory understanding of the proposed method (N1: 90.2%, N2: 86.6%, and N3: 83.7%) are better than those of other methods (NBC (N1:80.6%, N2:82.8%, and N3: 80.3%), SVM (N1:82.2%, N2:84.2%, and N3: 79.7%), kNN (N1:82.6%, N2:83.8%, and N3: 80.7%), SR-l^1 (N1:83.4%, N2:86%, and N3: 81.6%), SR-l^2 (N1:84.6%, N2:82.2%, and N3: 77.7%), and SR-$l^{0.5}$ (N1:87.4%, N2:86.6%, and N3: 81%)). For abnormal identification, TPR of the proposed method is highest (86.7%) and FPR of the proposed method is smallest (0.2%) among the compared methods (NBC (2.2%), SVM (2.4%), kNN (2.7%), SR-l^1 (0.9%), SR-l^2 (1.1%), and SR-$l^{0.5}$ (0.8%)). Results shown in Table 3 indicate that, the proposed method achieves better performance compared with other methods. In general, the accuracy of the model based on sparse representation is superior to the other methods on Bus station dataset1, especially in abnormal trajectory understanding. For different types of sparse representation models, the performance of the proposed method is better than others, mainly because the proposed method explicitly pays attention to the entropy of the distribution of selection frequency and modifies the traditional sparse representation models by adding the entropy constraint.

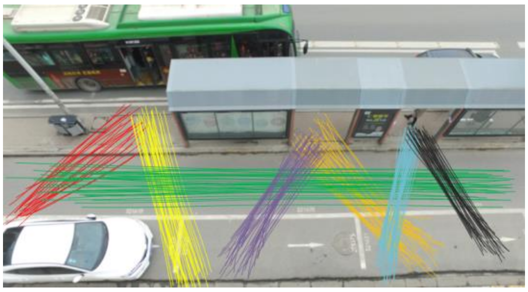
In order to check the efficiency of the proposed method, we compare the execution speed of the proposed method with those of all the selected methods. We select Bus station dataset1 as an example, and report the results in Table 4. As can be seen from Table 4, NBC performs fastest among all the methods due to its simplicity and the strong independence assumption. The execution time of the proposed method is longer than that of kNN, this is possibly due to the fact that the proposed method implemented in Matlab usually run significantly slower than those implemented in Weka which is based on Java environment. However, the proposed method performs significantly better than the selected sparse representation-based methods (SR-l^1, SR-l^2, and SR-$l^{0.5}$) which are all implemented in the Matlab environment, showing that the proposed IECTR is more efficient compared with the ordinary programming solvers.

TABLE 4 Execution time on Bus station dataset1

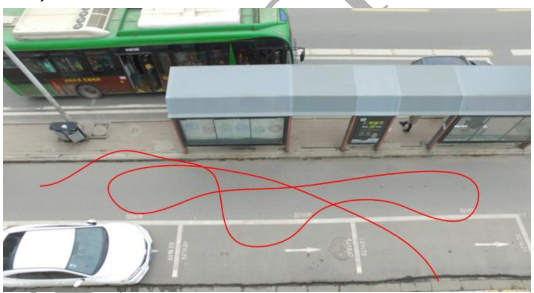
	NBC	SVM	kNN	SR-l^1	SR-l^2	SR-$l^{0.5}$	Our method
Execution time (ms)	2841	15623	7625	32311	27094	48172	13208

Figure 5 shows an example of pedestrian detection for Bus station dataset2 (Wuhan, Hubei, China). Figure 6 shows examples of the normal trajectory classes and abnormal trajectory class in Bus station dataset2. Table 5 shows the results of abnormal trajectory identification using NBC, kNN, SVM, SR-l^1, SR-l^2, SR-$l^{0.5}$, and the proposed method on Bus station dataset2, respectively. In Bus station dataset2, pedestrian normal trajectories are further divided into seven classes. “N1”, “N2”, “N3” , “N4” , “N5” , “N6” , “N7” are used to represent the seven classes, respectively. According to the results shown in Table 5, the accuracies of normal trajectory understanding from the proposed method are the highest in most of cases (N1:88.3%, N2:89.3%, N3: 86.7%, N4:88.7%, N5: 81%, and N7: 83.5%). For abnormal trajectory understanding, TPR of the proposed method is highest (82.9%), and FPR of the proposed method is smallest (0.5%), compared with the selected methods.

Figure 5 An example of pedestrian detection for Bus station dataset2 (Wuhan, Hubei, China)

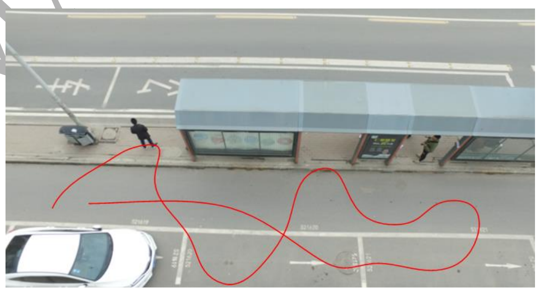


(a)



(b)

(c)



(d)

Figure 6 Examples of the normal and abnormal trajectory classes in Bus station dataset2

TABLE 5 Results of pedestrian trajectory identification on Bus station dataset2

Methods	Accuracy of Normal Trajectory Understanding							Accuracy of Abnormal Trajectory Understanding	
	N1	N2	N3	N4	N5	N6	N7	TPR	FPR
NBC	76.7%	78.3%	81%	81.6%	79.3%	77.6%	79.5%	38.1%	2.9%
SVM	79.3%	77%	77.7%	81.3%	76.6%	79.3%	75.5%	34.3%	3.1%
kNN	81%	82.6%	79.6%	80.7%	77%	81.3%	78%	52.3%	2.3%
SR-l^1	83.6%	85%	82.3%	84.6%	79.6%	82.3%	79.5%	74.3%	1.1%
SR-l^2	82%	79.7%	80.7%	82.3%	78.3%	81.3%	77.5%	72.4%	1.3%
SR-$l^{0.5}$	84.6%	85.3%	84%	85%	80.3%	81.7%	81%	79%	1.1%

Our method	88.3%	89.3%	86.7%	88.7%	81%	81.7%	83.5%	82.9%	0.5%

Figure 7 shows an example of pedestrian detection for Airport dataset (Wuhan, Hubei, China). Figure 8 shows examples of the normal trajectory classes and abnormal trajectory class in Airport dataset. Table 6 shows the results of abnormal trajectory detection using NBC, kNN, SVM, SR-l^1, SR-l^2, SR-$l^{0.5}$, and the proposed method on Airport dataset, respectively. In Airport dataset, pedestrian normal trajectories are further divided into four classes. “N1”, “N2”, “N3”, and “N4” are used to represent the four classes, respectively. The best understanding accuracies are achieved by the proposed method (N1:89.6%, N2:86.3%, N3: 89.3%, and N4: 87.6%). For abnormal trajectory understanding, the accuracy of the proposed method outperforms other selected methods.

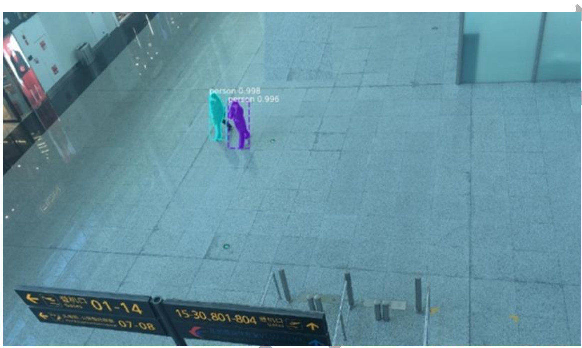
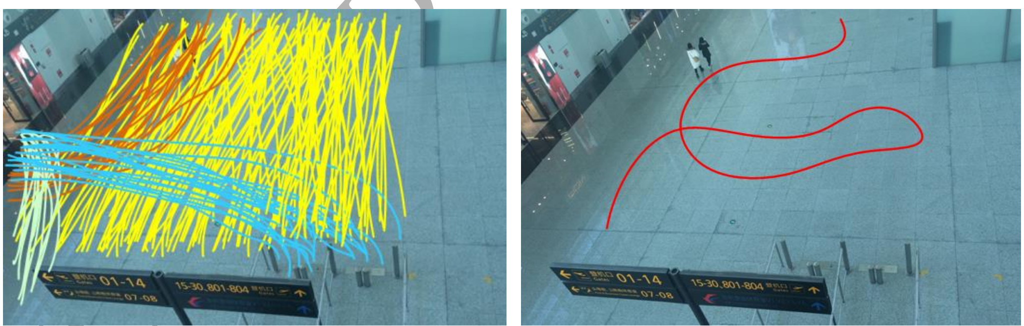


Figure 7 An example of pedestrian detection for Airport dataset (Wuhan, Hubei, China)



(a) **(b)**

Figure 8 Examples of the normal trajectory classes and abnormal trajectory class in Airport dataset: (a) the normal trajectory classes, (b) the abnormal trajectory class.

TABLE 6 Results of pedestrian trajectory identification on Airport dataset

Methods	Accuracy of Normal Trajectory Understanding				Accuracy of Abnormal Trajectory Understanding	
	N1	N2	N3	N4	TPR	FPR
NBC	74.3%	77.6%	74.3%	80.3%	31.4%	1.6%
SVM	73.3%	84.2%	79.3%	77%	31.4%	1.6%
kNN	79.6%	82%	80%	81.3%	42.9%	1.5%

SR-l^1	81.6%	83.6%	81.3%	83%	82.8%	0.4%
SR-l^2	80.3%	82.3%	79.6%	80%	77.1%	0.6%
SR-$l^{0.5}$	85.6%	86.3%	88.3%	85%	85.7%	0.3%
Our method	89.6%	86.3%	89.3%	87.6%	88.6%	0.2%

Figure 9 shows an example of pedestrian detection for Square dataset (Wuhan, Hubei, China). Figure 10 shows examples of the normal trajectory classes and abnormal trajectory class in Square dataset. Table 7 records the results of abnormal trajectory detection using NBC, kNN, SVM, SR-l^1, SR-l^2, SR-$l^{0.5}$, and the proposed method on Square dataset, respectively. In Square dataset, pedestrian normal trajectories are divided into three classes. “N1”, “N2”, and “N3” are used to represent the three classes, respectively.

As can be seen from Table 7, the proposed method (88%) is better than other methods (NBC (77.5%), SVM (80.5%), kNN (82%), SR-l^1 (81.5%), SR-l^2 (79.8%), and SR-$l^{0.5}$ (83.7%)) for normal trajectory class N1. Also, the identification accuracy of N2 of the proposed method (88.3%) is better than other methods (NBC (79.5%), SVM (77.8%), kNN (80.3%), SR-l^1 (82.3%), SR-l^2 (82%), and SR-$l^{0.5}$ (82.8%)). for normal trajectory class N3, our method obtains the highest identification accuracy (89.2%) among all the compared methods. While the lowest accuracy (78.8%) is obtained by NBC. According to values of TPR and FPR, we can learn that the proposed method (TPR: 90% and FPR: 0.2%) outperforms all the compared methods.

Figure 9 An example of pedestrian detection for Square dataset (Wuhan, Hubei, China)

(a) **(b)**

(c) **(d)**

Figure 10 Examples of the normal and abnormal trajectory classes in Square dataset

TABLE 7 Results of pedestrian normal trajectory identification on Square dataset

Methods	Accuracy of Normal Trajectory Understanding			Accuracy of Abnormal Trajectory Understanding	
	N1	N2	N3	TPR	FPR
NBC	77.5%	79.5%	78.8%	47.5%	1.6%
SVM	80.8%	77.8%	81.3%	42.5%	1.7%
kNN	82%	80.3%	79.3%	45%	1.6%
SR-l^1	81.5%	82.8%	82.3%	82.5%	0.4%
SR-l^2	79.8%	82%	80.8%	80%	0.6%
SR-$l^{0.5}$	83.7%	82.8%	80.7%	85%	0.3%
Our method	88%	88.3%	89.2%	90%	0.2%

Figure 11 shows an example of pedestrian detection for Railway station dataset (Wuhan, Hubei, China). Figure 12 shows examples of the normal trajectory classes and abnormal trajectory class in Railway station dataset. Table 8 shows the results of abnormal trajectory detection using NBC, kNN, SVM, SR-l^1, SR-l^2, SR-$l^{0.5}$, and the proposed method on Railway station dataset, respectively. In Railway station dataset, pedestrian normal trajectories are also further divided into four classes. "N1", "N2", "N3", and "N4" are used to represent the four classes, respectively. Table 8 shows that, although the understanding accuracy of the proposed method (83.1%) is lower than that of SR-$l^{0.5}$ (84.6%) in class N4, the best accuracies are achieved by the proposed method in most of cases (N1:82.5%, N2:84.4%, and N3: 87.2%). For abnormal trajectory understanding, the accuracy of the proposed method (TPR: 78.8% and FPR: 0.9%) outperforms all other compared methods.

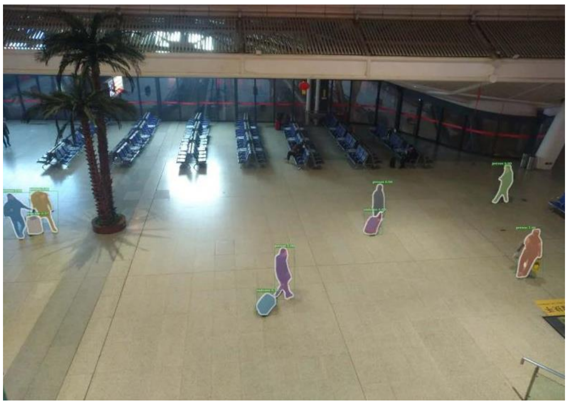
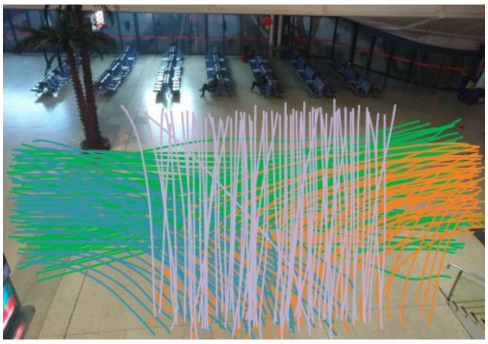


Figure 11 An example of pedestrian detection for Railway station dataset (Wuhan, Hubei, China)



(a) **(b)**

Figure 12 Examples of the normal trajectory classes and abnormal trajectory class in Railway station dataset: (a) the normal trajectory classes, (b) the abnormal trajectory class.

TABLE 8 Results of pedestrian trajectory identification on Railway station dataset

Methods	Accuracy of Normal Trajectory Understanding				Accuracy of Abnormal Trajectory Understanding	
	N1	N2	N3	N4	TPR	FPR
NBC	74.5%	71.9%	76%	76.9%	32.5%	1.4%
SVM	77.5%	74.3%	75.2%	79.2%	31.3%	1.5%
kNN	81%	78.1%	79.6%	83.1%	52.5%	1.2%
SR-l^1	76.5%	76%	82%	81.5%	73.8%	0.9%
SR-l^2	78%	76.8%	78%	80.8%	61.3%	1.2%
SR-$l^{0.5}$	80.5%	81.3%	83.2%	84.6%	75%	1.1%
Our method	82.5%	84.4%	87.2%	83.1%	78.8%	0.9%

CONCLUSIONS

In this study, an information entropy constrained sparse representation model is developed for pedestrian behavior understanding, which can lead to superior identification results. OMP is embedded in EM method to iteratively obtain the selection probabilities and the sparse coefficients of the proposed model. Trajectory similarity is employed for initializing the selection probabilities. In addition, the lower bound theory of nonzero entries in solutions of L_p minimization is introduced to get more salient solutions. In order to validate the performance and effectiveness of the proposed method, five datasets, including Bus station dataset1, Bus station dataset1, Airport dataset, Square dataset, and Railway station dataset, are used in our experiments. Experimental results show that the understanding accuracies achieved by the proposed method are significantly improved on the whole compared with most well-known classifiers namely NBC, kNN, SVM, and typical extant sparse representation methods. We may conclude that the proposed method can be effectively used in pedestrian behavior learning based on video surveillance systems.

This study gives a potential way for the applications of the video surveillance system to provide further services for traffic safety and law enforcement.

Future work about pedestrian behavior learning will include developing an online learning system based on this study in order to adapt to the dynamic environment. In addition, it is noted that most of the extant works that refer to pedestrian trajectory analysis use private datasets for model illustration and validation. Therefore, open-source pedestrian trajectory datasets need to be further established for convincing comparisons among the state-of-the-arts.

ACKNOWLEDGMENT

We thank the editor and the anonymous reviewers for their constructive comments and suggestions. This work was supported in part by the National Natural Science Foundation of China under Grants 61703319, 71702066, 51775396, and U1764262, in part by China Postdoctoral Science Foundation under Grant 2017M612856, in part by the Science and Technology Demonstration Project of Ministry of Transport of China under Grant 2015364X16030, in part by the National Key R&D Program of China under Grants 2017YFB0102500 and 2018YFB1600600, and in part by Australia ARC DECRA under Grant DE190100931. Yishi Zhang served as the corresponding author.

CONFLICTS OF INTEREST

The authors have no conflict of interest to declare.

REFERENCES

Aha, D. Kibler, W, D., and Albert, M. K., 1991. Instance-based learning algorithms, Machine Learning, 6(1): 37-66.

Amaldi, E., and Kann, V., 1998. On the approximability of minimizing nonzero variables or unsatisfied relations in linear systems. Theoretical Computer Science, 209 (1-2): 237-260.

Candes, E. J. and Tao, T., 2005. Decoding by linear programming, IEEE Transactions on Information Theory, 51(12): 4203-4215.

Chen Z., Wu C., Huang Z., Lyu N., Hu Z., Zhong M., Cheng Y., and Ran B., 2017. Dangerous Driving Behavior Detection using Video-extracted Vehicle Trajectory Histograms, Journal of Intelligent Transportation Systems, 21(5): 409-421.

Chen, X., Xu, F., and Ye, Y., 2010. Lower bound theory of nonzero entries in solutions of L_2-L_p minimization, SIAM Journal on Scientific Computing, 32(5): 2832-2852.

Chong, Y., Chen, W., Li, Z., and Lam, W. H. K., 2011. Method for Preceding Vehicle Type Classification Based on Sparse Representation. Transportation Research Record: Journal of the Transportation Research Board, No. 2243, Transportation Research Board of the National Academies, Washington,D.C., pp. 74-80

Chen, Z., Wu, C., Zhang, Y., Huang, Z., Zhong, M. and Lyu, N., 2015. Feature selection with redundancy-complementariness dispersion. Knowledge-Based Systems, 89(C): 203-217.

Chartrand, R. Nonconvex Regularization for Shape Preservation, 2007. In: proceedings of the IEEE International Conference on Image Processing, pp. 293-296.

Chen, Z., Wu, C., Zhang, Y., and Huang, Z., Jiang, J., Lyu, N., and Ran B., 2017. Vehicle Behavior Learning via Sparse Reconstruction with l_2-l_p Minimization and Trajectory Similarity. IEEE Transactions on Intelligent Transportation Systems, 18(2): 236-247.

Chen, Z., Xue J., Wu C., Qin L., and Cheng X., 2018. Classification of vessel motion pattern in inland waterways based on Automatic Identification System. Ocean Engineering, 161: 69-76.

Chartrand, R., 2007. Exact Reconstruction of Sparse Signals via Nonconvex Minimization [J]. IEEE Signal Processing Letters, 14(10): 707-710.

Datta, A., Shah, M., and Vitoria Lobo, N. D., 2003. Person-on-person violence detection in video data. In: proceedings of the 16th IEEE International Conference on Pattern Recognition, pp. 433-438.

Fan J, and Li R., 2001. Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. Journal of the American Statistical Association, 96(456): 1348-1360.

Junejo, I. N., Javed, O., and Shah, M.., 2004. Multi feature path modeling for video surveillance. In: proceedings of the International Conference on Pattern Recognition, pp.716-719.

Hild, K. E., H. T. Attias, and S. S. Nagarajan., 2008. An expectation-maximization method for spatio-temporal blind source separation using an AR-MOG source model. IEEE Transactions on Neural Networks, 19(3): 508-519.

Hu, W., Xiao, X., Fu, Z., Xie, D., Tan, T., and Maybank, S., 2006. A System for Learning Statistical Motion Patterns. IEEE Transactions on Pattern Analysis & Machine Intelligence, 28(9): 1450-1464.

He K, Gkioxari G, Dollar P, and Girshick R., 2017. Mask R-CNN. IEEE Transactions on Pattern Analysis and Machine Intelligence, PP(99):1-1.

Li, C., Han, Z. J., Ye, Q., and Jiao, J., 2013. Visual abnormal behavior detection based on trajectory sparse reconstruction analysis, Neurocomputing, 119(16): 94-100.

Mecocci, A., and Pannozzo, M., 2005. A completely autonomous system that learns anomalous movements in advanced video surveillance applications. In: proceedings of the IEEE International Conference on Image Processing, pp. 586-589.

Mo, X., Monga, V., Bala, R., and Fan, Z, 2014. Adaptive Sparse Representations for Video Anomaly Detection. IEEE Transactions on Circuits & Systems for Video Technology, 24(4): 631-645.

Nikolova, M., 2006.Analysis of the Recovery of Edges in Images and Signals by Minimizing Nonconvex Regularized Least-Squares. Siam Journal on Multiscale Modeling & Simulation, 4(3): 960-991.

Okyere, A. G. Y., and Attoh-Okine, N., 2014. Traffic Sign Recognition Using Sparse Representations and Active Contour Models. Transportation Research Record: Journal of the Transportation Research Board, No. 2463, Transportation Research Board of the National Academies, Washington, D.C., pp. 35-45.

Oliver, N., Rosario, B., and Pentland, A., 1999. A Bayesian Computer Vision System for Modeling Human Interactions. IEEE Transactions on Pattern Analysis & Machine Intelligence, 22(8): 255-272.

Park, S., and Aggarwal, J. K., 2004. A hierarchical Bayesian network for event recognition of human actions and interactions. Multimedia Systems, 10(2): 164-179.

Viola, P. A, Jones, M. J., 2007. Detecting pedestrians using patterns of motion and appearance in videos: Kluwer Academic Publishers, US 7212651 B2[P].

Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S., and Ma, Y., 2009. Robust face recognition via sparse representation. IEEE Transactions on Pattern Analysis & Machine Intelligence, 31(2): 210-227.

Wang, S., Han, Z., Zhu, L., and Chen, Q., 2014. A Novel Approach to Design the Fast Pedestrian Detection for Video Surveillance System. International Journal of Security & Its Applications, 8(1): 93-102.

Xiang, T., and Gong, S., 2008. Incremental and adaptive abnormal behavior detection. Computer Vision & Image Understanding, 111(1): 59-73.

Xu, X., Xie, L., Li, H., and Qin, L. (2018). Learning the route choice behavior of subway passengers from AFC data. Expert Systems with Applications, 95, 324-332.

Yao, W., and Deng, Z., 2012. A robust pedestrian detection approach based on Shapelet feature and Haar detector ensembles. Tsinghua Science & Technology, 17(1): 40-50.

Zhang, L., Luo, M., Li, Z., Nie, F., Zhang, H., Liu, J., and Zheng, Q., 2018. Large-scale robust semisupervised classification, IEEE Transactions on Cybernetics, 99: 1-11.

Zhang, L., Yang, M., and Feng, X., 2011. Sparse representation or collaborative representation: Which helps face recognition? In: proceedings of the IEEE International Conference on Computer Vision, pp. 471-478.

Author Contributions

Yishi Zhang and Zhijun Chen conceived of the presented idea and wrote the manuscript. Yishi Zhang developed and implemented the model. Zhijun Chen and Hao Cai performed the computations and conducted the experiments. Chaozhong Wu and MengChao Mu contributed to pedestrian detection. Zhijun Chen and Hao Cai conducted the work of data preprocessing. Zhixiong Li and Miguel Angel Sotelo checked the results of this work and proofread the manuscript. All authors discussed the results and contributed to the final manuscript.

This work was supported in part by the National Natural Science Foundation of China under Grants 61703319, 71702066, 51775396, and U1764262, in part by China Postdoctoral Science Foundation under Grant 2017M612856, in part by the Science and Technology Demonstration Project of Ministry of Transport of China under Grant 2015364X16030, in part by the National Key R&D Program of China under Grants 2017YFB0102500 and 2018YFB1600600, and in part by Australia ARC DECRA under Grant DE190100931.

This submission has been approved by all the authors and the institutes where the study was conducted. The content of this manuscript has not been published previously, and not been reviewed in any other journals.

The authors have no conflict of interest to declare.