Research Article

Histogram of Maximal Optical Flow Projection for Abnormal Events Detection in Crowded Scenes

Ang Li, 1 Zhenjiang Miao, 1 Yigang Cen, 1 Tian Wang, 2 and Viacheslav Voronin 3

1 Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China
2 School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China
3 Department of Radio-Electronic Systems, Don State Technical University, Shakhty 346500, Russia

Correspondence should be addressed to Ang Li; lianghit@126.com

Received 21 May 2015; Revised 29 July 2015; Accepted 11 August 2015

Academic Editor: Shaojie Tang

Copyright © 2015 Ang Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abnormal events detection plays an important role in the video surveillance, which is a challenging subject in the intelligent detection. In this paper, based on a novel motion feature descriptor, that is, the histogram of maximal optical flow projection (HMOFP), we propose an algorithm to detect abnormal events in crowded scenes. Following the extraction of the HMOFP of the training frames, the one-class support vector machine (SVM) classification method is utilized to detect the abnormality of the testing frames. Compared with other methods based on the optical flow, experiments on several benchmark datasets show that our algorithm is effective with satisfying results.

1. Introduction

Nowadays, more and more surveillance cameras have been used in public places. Behavior analysis in crowded scenes [1–5] becomes more and more popular and important for public safety. In order to eliminate the world representation layer which can be a significant source of errors for algorithm modeling, an approach based on modeling directly at the pixel level was described in [6]. In [7, 8], social force model was used in abnormal crowd behavior detection. In [9, 10], a model named social attribute-aware force model was proposed. In this model, in order to improve the algorithm performance for the interaction behavior of the crowd, social characteristics of crowd behavior were taken into account.

In [11], SIFT features were extracted for the Bag of Words (Bow) model with Spatial Pyramid Matching Kernel (SPM). Then a SVM classifier was used for cross-scene abnormal events detection. In [12], based on the fact that the occurrence of abnormal events is rare while the frequently occurring events are normal in general human perception, proximity clustering for abnormal events detection in video sequence was proposed. In [13], when labeled information about normal events was limited and information about abnormal events was not available, projection subspace associated with detectors was discovered by using both labeled and unlabeled segments. In wireless sensor networks, a fact has been observed that instead of being transient, most abnormal events persist over a considerable period time. Thus, a technique for handling data in a segment-based manner was introduced in [14]. Without using any tracking and motion features, a feature extraction and events detection method were presented in [15], where features were extracted from foreground blobs and then confined in SVM based models for real-time events detection.

Unlike most existing approaches used for abnormal events detection, sparse representation based approaches attracted many researchers in the recent years. In [16], a method to detect abnormal events by a sparse subspace clustering was proposed. In [17, 18], a model based on the optical flow was described, which utilized the sparse reconstruction cost (SRC) over the normal dictionary to measure the normalness of the tested samples. As we know, optical flow is the approximated motion vector at each pixel location, which can reflect the relative distances of moving objects. Therefore, it is important and useful in video surveillance and abnormal events detection. Other methods based on the...
histogram of optical flow were described in [19–21]. Also, it was improved and used in this paper.

Although the above approaches could successfully realize abnormal events detection, they were limited in some aspects. Some models were established complicatedly and others cost a long time in the detection process. Based on these, we propose a novel detection model in crowded scenes, which is relatively simple and time-saving in calculation. Similar to the approach introduced in [21], our algorithm is mainly based on a proper processing method in the optical flow field.

The rest of the paper is organized as follows. In Section 2, we present how to acquire the motion features. In Section 3, the theory of one-class SVM is reviewed. In Section 4, the algorithm of abnormal events detection is introduced in detail. Section 5 presents our experiment results. Finally, some conclusions are presented in Section 6.

2. Motion Feature Extraction

Optical flow field is the movement on the surface of grayscale images, which reflects the movement information of two consecutive frames. Optical flow provides the information of direction and amplitude of the moving object in a scene, which can describe the behavior of people very well. Optical flow is derived from the following basic equation:

\[ I_x u + I_y v + I_t = 0, \]

where \( I_x, I_y, \) and \( I_t \) are the partial derivatives of the image grayscale value along the \( x, y, \) and \( t \) dimension, respectively; \( u \) and \( v \) are the horizontal \((x\) dimension) and vertical \((y\) dimension) components of the optical flow. Equation (1) is an ill-posed problem. In [22], Horn and Schunck proposed an algorithm. It is known as the HS algorithm to compute the optical flow by introducing a global constraint of smoothness, which is equal to the additional condition

\[ \min \nabla^2 u + \nabla^2 v, \]

where \( \nabla^2 u \) and \( \nabla^2 v \) are Laplace operators of \( u \) and \( v \), respectively. The problem to get optical flow can be concluded as follows:

\[ \min \iint \left( (I_x u + I_y v + I_t)^2 + \alpha^2 \left( \nabla^2 u + \nabla^2 v \right)^2 \right) dx dy, \]

where \( \alpha \) is the parameter that represents the weights of the regularization term. Then the Euler-Lagrange equation can be acquired, which are solved by utilizing the Gauss-Seidel method. It can get an iterative result to compute the optical flow:

\[ u^{n+1} = \tilde{u} - \frac{I_x (I_x u + I_y v + I_t)}{\alpha^2 + I_x^2 + I_y^2}, \]

\[ v^{n+1} = \tilde{v} - \frac{I_y (I_x u + I_y v + I_t)}{\alpha^2 + I_x^2 + I_y^2}, \]

where \( \tilde{u} \) and \( \tilde{v} \) are weighted average value of \( u \) and \( v \), respectively, which are calculated in a neighborhood around the pixel location. \( n \) denotes the algorithm iteration number.

In this paper, we propose a novel motion feature descriptor, called histogram of maximal optical flow projection (HMOFP). Figure 1 briefly shows the process for computing the HMOFP.

As shown in Figure 2, the optical flow field of frame \( s \) is divided into \( m \) image patches with overlap areas. Each block contains \( B \times B \) pixels. Then we deal with the optical flow in each patch as follows: \( 0 \sim 360^\circ \) are segmented into \( p \) bins. For an image patch, the optical flow vector of each pixel must belong to a bin according to its direction. Thus, each bin may contain several optical flow vectors. We project all optical flow vectors in the same bin onto the angle bisector of this bin. Then the maximal projection vector is selected as the feature descriptor. For example, in Figure 3(a), there are two vectors \( \overrightarrow{\sigma_1} \) and \( \overrightarrow{\sigma_2} \) falling into the first bin. It is easy to know that the projection of \( \overrightarrow{\sigma_2} \) is longer than the projection of \( \overrightarrow{\sigma_1} \). Thus, the length of the projection vector \( ||\overrightarrow{\sigma_2}|| \) is selected as the feature descriptor of the first bin. After computing \( m \) patches, we obtain the feature descriptor vector of each image patch, denoted as \([h_1, h_2, \ldots, h_m]^T_{\text{prom}}\), where \( h_i = [h_{i1}, h_{i2}, \ldots, h_{ip}]^T \).

For the \( i \)th patch, \( h_{ij}, 1 \leq j \leq p, 1 \leq i \leq m \) denotes the maximal amplitude among all projection vectors in the \( j \)th bin. As shown in Figure 3(b), we take the concatenation of the \( m \) feature descriptor vectors, which is named \( H_s \), as the global HMOFP feature of the frame \( s \).

In order to describe a crowd scene well, sufficient crowd movement information is required. On the other hand, for distinguishing two different scenes, detailed comparisons of them are needed and useless information in these two scenes should be eliminated. In the classification process, overlapping block-division can increase the number of significant motion features in two different frames such that these two frames can be more distinguishable. Thus, it is adopted in our algorithm since the optical information can be utilized sufficiently. Moreover, to describe the motion of a crowd,
we need two factors: explicit directions and the moving distance along each direction. The operation of segmenting the 2D space into $p$ bins provides us ample information to describe the directions of moving people. To let the direction in each bin be unique, we select the $p$ angle bisectors as the direction standard. Since there may be far more than one optical flow vector in each bin, in order to enhance the distinction between the normal scene and the abnormal scene, we select the maximal vector projection rather than the sum of all the vector projections on the bisector as the motion feature descriptor. If we ignore the background area, the amplitudes of motion vectors that belong to the normal area are very small in a normal frame and the motion vectors corresponding to the abnormal area are large in an abnormal frame. Usually, the number of normal motion vectors is much more than that of the abnormal area. If we use the sum of all projection vectors on the angle bisector as the feature descriptor of each bin, the accumulation of the massive small motion vectors in the normal frame may confuse the small number of large motion vectors in the abnormal frame; that is, the sum of all projection vectors on the angle bisector in each bin of the normal frame is likely to be close to that of the abnormal frame. Thus, in order to improve the distinguishability between the abnormal and normal frames, we select the maximal projection vector as the feature descriptor of each bin, as it was demonstrated in Figure 3.

3. One-Class SVM

SVM was initiated by Vapnik and Lerner [23]. Since the kernel methods were introduced, SVM has been applied extensively in nonliner classification problems [24–26]. In one-class classification problem, the substance is that the boundary, that is, an appropriate region, needs to be determined in the data space $\mathcal{X}$, which contains most of the samples coming from an unknown probability distribution $D$. This goal can be realized by searching for an optimal decision hyperplane in the feature space, which is known as the Hilbert space $\mathcal{H}$. This hyperplane can maximize the distance between itself and the original point, while only a small part of data falls between them [27]. The relationship between $\mathcal{X}$ and $\mathcal{H}$ is shown in Figure 4.
One-class SVM problem can be presented as an optimization model:

$$
\begin{align*}
\min_{w, \xi, \rho} & \quad \frac{1}{2} \|w\|^2 - \rho + \frac{1}{v} \sum_{i=1}^{l} \xi_i \\
\text{s.t.} & \quad w^T \phi(x_i) \geq \rho - \xi_i, \quad \xi_i \geq 0,
\end{align*}
$$

where $x_i \in \mathcal{X}, i \in [1 \cdots l]$ are training samples in the input data space $\mathcal{X}$ and $\phi : \mathcal{X} \to \mathcal{H}$ can map a vector $x_i$ into the feature space $\mathcal{H}$. $w^T \phi(x_i) - \rho = 0$ is the decision hyperplane. $\xi_i$ is the slack variable for penalizing the outliers. $\nu \in (0,1]$ is the hyperparameter, which is the weight for controlling slack variable and tunes the number of acceptable outliers. $\phi$ is a mapping function, which provides us a way to solve the nonlinear classification problem in the space $\mathcal{X}$ by a linear solution in the space $\mathcal{H}$. By calculating dot product in $\mathcal{H}$, the kernel function is defined as $k(x_i, x_j) = \phi^T(x_i)\phi(x_j)$. The decision function in the space $\mathcal{X}$ with a Lagrangian multiplier $\alpha_i$ is defined as

$$
f(x) = \text{sgn} \left( \sum_{i=1}^{l} \alpha_i k(x_i, x) - \rho \right). \quad (6)$$

In [28], it was introduced that if appropriate parameters were selected, polynomial and sigmoid kernels will result in similar results with Gaussian. We choose Gaussian kernel in our algorithm. This kernel is defined as

$$
k(x_i, x_j) = \exp \left( -\frac{\|x_i - x_j\|^2}{2\sigma^2} \right), \quad (7)$$

where $x_i$ and $x_j$ belong to the space $\mathcal{X}$ and $\sigma$ is the scale factor at which the data should be clustered.

In our method, one-class SVM is utilized as follows. Firstly, the training set is used to establish a model. Then an appropriate boundary in the data space can be determined. The new incoming frames will be clustered by the following rule: if the HMOFP feature of the testing frame falls inside the boundary, it will be clustered as a normal frame. Otherwise, it is abnormal.

4. Abnormal Events Detection

In this section, an algorithm for abnormal events detection in surveillance video is described in detail. Suppose that for a given scene, there is a set of training frames $[f_1, \ldots, f_l]$, which describe the normal behavior of crowded people. The general procedures for the abnormal events detection based on the histogram of maximal optical flow projection (HMOFP) are presented as follows.

Step 1. Calculate the optical flow, that is, $[\text{OP}_1, \ldots, \text{OP}_{l-1}]$, by the HS method at each pixel of the first $l - 1$ frames:

$$
[f_1, \ldots, f_l]_{axgbxd} \xrightarrow{\text{HS}} [\text{OP}_1, \ldots, \text{OP}_{l-1}]_{axgbx(l-1)}, \quad (8)
$$

where $a \times b$ is the size of the frame image and $l$ is the number of the frames in the training set. Our method to compute optical flow is based on the two consecutive frames, which is only effective to the first frame, so in the right side of (8), the maximal subscript is $l - 1$.

Step 2. Extract the motion features of the first $l - 1$ training frames. Then the HMOFP feature vectors of them can be obtained, which is denoted as the set $[H_1, \ldots, H_{l-1}]^T$. Consider

$$
[\text{OP}_1, \ldots, \text{OP}_{l-1}]_{axgbx(l-1)} \xrightarrow{\text{HMOFP}} [H_1, \ldots, H_{l-1}]_{p_{ome}(l-1)}^T. \quad (9)
$$

Step 3. Based on HMOFP, one-class SVM is utilized to calculate the optimal boundary of the set $[H_1, \ldots, H_{l-1}]^T$, which corresponds to the set of support vectors or the optimal hyperplane in the feature space.

Step 4. Detect HMOFP of the testing frames based on the model trained by the motion feature of the first $l - 1$ training frames.

The whole procedure is illustrated in Figure 5.

5. Experimental Results

In this section, based on the UMN dataset [29] and PETS2009 dataset [30], we evaluate our method for abnormal event
5.1. Experiments on the UMN Dataset. There are three different crowded scenes in the UMN dataset, which are named lawn, indoor, and plaza, respectively. In our experiments, we select a part of the normal frames of each scene as the training set and take the rest of the video sequence as the testing set.

5.1.1. Detection in the Lawn Scene. The video sequence of the lawn scene contains 1453 frames in total. The first 480 frames are taken as the training set. As shown in Figure 6, in the lawn scene, the normal event is that individuals walk in different directions. The abnormal event is that individuals suddenly run away. The detection results of the lawn scene are shown in Figure 7. The accuracy of the detection results is 95.5141%.

5.1.2. Detection in the Indoor Scene. The video sequence of the indoor scene contains 4144 frames in total. The first 319 frames are taken as the training set. As shown in Figure 8, in the indoor scene, the normal event is that some people are talking and standing in a relatively fixed location while some others are walking along the road in the hall. The abnormal event is that people run out of the doors suddenly. The detection results of the indoor scene are shown in Figure 9. The accuracy of the detection results is 91.2857%.

5.1.3. Detection in the Plaza Scene. The video sequence of the plaza scene contains 2412 frames in total. The first 550 frames are taken as the training set. As shown in Figure 10, in
the plaza scene, the normal event is that people walk around the center of the square. The abnormal event is that people suddenly run away from the square. The detection results of the plaza scene are shown in Figure 11. The accuracy of the detection results is 94.3352%.

5.2. Experiments on the PETS2009 Dataset. In the following experiments, we can choose some specific scenes we are interested in as the targets in the detection progress. In the PETS2009 dataset, we firstly select the training set and the normal testing set, respectively, in the same scene. Then another video clip in a different scene is taken as the corresponding abnormal testing set. Our experiments and the detection results are shown as follows.

5.2.1. People Scatter Detection. In this part, the training set is the video sequence Time 14-16 (Frame 0 to Frame 222), where people are walking or running towards one direction. The normal testing set includes 41 frames (Frame 48 to Frame 88) of Time 14-17. 41 frames (Frame 337 to Frame 377) of Time 14-33 are labeled as abnormal for testing, in which people are scattered in all directions. The two different scenes are shown in Figure 12. The accuracy of the detection results is 97.5%, as shown in Figure 13.

5.2.2. Crowd Movement Direction Detection. In this part, the training set is the video sequence Time 14-55 (Frame 0 to Frame 399), where people are walking towards all directions. The normal testing set includes 89 frames (Frame 400 to
5.2.3. People Running Detection. In this part, the training set contains 50 frames (Frame 0 to Frame 49) of the video sequence Time 14-31 and 61 frames (Frame 0 to Frame 60) of the video sequence Time 14-17, where people are walking from right to left and from left to right, respectively. The normal testing set includes 104 frames (Frame 0 to Frame 37 and Frame 108 to Frame 173) of Time 14-16. 119 frames (Frame 38 to Frame 107 and Frame 174 to Frame 222) of Time 14-16 are labeled as abnormal for testing, in which people are running towards one direction. The two different scenes are shown in Figure 14. The accuracy of the detection results is 92.6136%, as shown in Figure 15.
Figure 14: Two different scenes in the same location.

![Normal scene](a) ![Abnormal scene](b)

Figure 15: The detection results of the sequence Time 14-17. "1" means normal and "-1" means abnormal.

Figure 16. The accuracy of the detection results is 93.6937%, as shown in Figure 17.

5.2.4. People Splitting Detection. In this part, the training set contains Frames 0 to 40 of the video sequence Time 14-16, where people are walking towards the same direction. The normal testing set includes 64 frames (Frame 0 to Frame 63) of Time 14-31. 66 frames (Frame 64 to Frame 129) of the video sequence Time 14-31 are labeled as abnormal for testing, in which the crowd is splitting. The normal scene and abnormal scene are shown in Figure 18. The accuracy of the detection results is 96.1538%, as shown in Figure 19.

5.2.5. Comparison. We compared our algorithm with the histogram of optical flow orientation (HOFO) method proposed in [21], as shown in Table 1. Most results of our algorithm are better than those of HOFO.

| Method      | Time 14-33 | Time 14-17 | Time 14-16 | Time 14-31 |
|-------------|------------|------------|------------|------------|
| HOFO        | 97.5%      | 90%        | 93.24%     | 94.6154%   |
| HMOFP (ours)| 97.5%      | 92.6136%   | 93.6937%   | 96.1538%   |

Table 1: The comparison of HMOFP with HOFO.

6. Conclusion

In this paper, we proposed an algorithm for abnormal events detection in crowded scenes with global-frame scale. Our method contains two main procedures: first is computing the histogram of maximal optical flow projection (HMOFP) descriptor of the input video sequence. Second, one-class SVM classifier is utilized for nonlinear classification of the
Figure 16: Two different scenes in the same location.

Figure 17: The detection results of the sequence Time 14-16. "1" means normal and "−1" means abnormal.

Figure 18: Two different scenes in the same location.
testing sets. The proposed method has been tested on several surveillance video datasets with good detection accuracy.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**Acknowledgments**

This work is supported by the NSFC (nos. 61273274, 61370127, 61572067, and 61272028), 973 Program (no. 2011CB302203), National Key Technology R&D Program of China (nos. 2012BAH01F03, NSFB4123104, FRFCU 2014JZ004, and Z1311000193143), and Tsinghua-Tencent Joint Lab. for IIT.

**References**

[1] R. Mehran, B. E. Moore, and M. Shah, “A streakline representation of flow in crowded scenes,” in *Computer Vision—ECCV 2010: 11th European Conference on Computer Vision*, Heraklion, Crete, Greece, September 5–11, 2010, Proceedings, Part III, vol. 6313 of Lecture Notes in Computer Science, pp. 439–452, Springer, Berlin, Germany, 2010.

[2] Y. Cong, J. Yuan, and Y. Tang, “Video anomaly search in crowded scenes via spatio-temporal motion context,” *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 10, pp. 1590–1599, 2013.

[3] F. Daniyal and A. Cavallaro, “Abnormal motion detection in crowded scenes using local spatio-temporal analysis,” in *Proceedings of the 36th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’11)*, pp. 1944–1947, May 2011.

[4] W. Li, V. Mahadevan, and N. Vasconcelos, “Anomaly detection and localization in crowded scenes,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 18–32, 2014.

[5] M. Thida, H.-L. Eng, and P. Remagnino, “Laplacian eigenmap with temporal constraints for local abnormality detection in crowded scenes,” *IEEE Transactions on Cybernetics*, vol. 43, no. 6, pp. 2147–2156, 2013.

[6] D. Kosmopoulos and S. P. Chatzis, “Robust visual behavior recognition,” *IEEE Signal Processing Magazine*, vol. 27, no. 5, pp. 34–45, 2010.

[7] R. Mehran, A. Oyama, and M. Shah, “Abnormal crowd behavior detection using social force model,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR ’09)*, pp. 935–942, IEEE, Miami, Fla, USA, June 2009.

[8] S.-H. Yen and C.-H. Wang, “Abnormal event detection using HOSE,” in *Proceedings of the 3rd International Conference on IT Convergence and Security (ICITCS ’13)*, pp. 1–4, IEEE, Macao, China, December 2013.

[9] Y. Zhang, L. Qin, H. Yao, and Q. Huang, “Abnormal crowd behavior detection based on social attribute-aware force model,” in *Proceedings of the 19th IEEE International Conference on Image Processing (ICIP ’12)*, pp. 2689–2692, October 2012.

[10] Y. Zhang, L. Qin, R. Ji, H. Yao, and Q. Huang, “Social attribute-aware force model: exploiting richness of interaction for abnormal crowd detection,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 7, pp. 1231–1245, 2015.

[11] T.-Y. Hung, J. Lu, and Y.-P. Tan, “Cross-scene abnormal event detection,” in *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS ’13)*, pp. 2844–2847, May 2013.

[12] T. Sandhan, A. Sethi, T. Srivastava, and J. Y. Choi, “Unsupervised learning approach for abnormal event detection in surveillance video by revealing infrequent patterns,” in *Proceedings of the 28th International Conference on Image and Vision Computing New Zealand (IVCNZ ’13)*, pp. 494–499, November 2013.

[13] I. Tziakos, A. Cavallaro, and L. Xu, “Local abnormal detection in video using subspace learning,” in *Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS ’10)*, pp. 519–525, 2010.

[14] M. Xie, J. Hu, and S. Guo, “Segment-based anomaly detection with approximated sample covariance matrix in wireless sensor networks.”
networks,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, no. 2, pp. 574–583, 2015.

[15] M. Haque and M. Murshed, “Panic-driven event detection from surveillance video stream without track and motion features,” in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME ’10)*, pp. 173–178, IEEE, Singapore, July 2010.

[16] H. Ren and T. B. Moeslund, “Abnormal event detection using local sparse representation,” in *Proceedings of the 11th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS ’14)*, pp. 125–130, IEEE, Seoul, Republic of Korea, August 2014.

[17] Y. Cong, J. Yuan, and J. Liu, “Sparse reconstruction cost for abnormal event detection,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR ’11)*, pp. 3449–3456, June 2011.

[18] Y. Cong, J. Yuan, and J. Liu, “Abnormal event detection in crowded scenes using sparse representation,” *Pattern Recognition*, vol. 46, no. 7, pp. 1851–1864, 2013.

[19] T. Wang and H. Snoussi, “Histograms of optical flow orientation for visual abnormal events detection,” in *Proceedings of the IEEE 9th International Conference on Advanced Video and Signal-Based Surveillance (AVSS ’12)*, pp. 13–18, September 2012.

[20] T. Wang and H. Snoussi, “Histograms of optical flow orientation for abnormal events detection,” in *Proceedings of the IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS ’13)*, pp. 45–52, January 2013.

[21] T. Wang and H. Snoussi, “Detection of abnormal visual events via global optical flow orientation histogram,” *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 6, pp. 988–998, 2014.

[22] B. K. P. Horn and B. G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, no. 1–3, pp. 185–203, 1981.

[23] V. N. Vapnik and A. Lerner, “Pattern recognition using generalized portrait method,” *Automation and Remote Control*, vol. 24, no. 6, pp. 774–780, 1963.

[24] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “Training algorithm for optimal margin classifiers,” in *Proceedings of the 5th Annual ACM Workshop on Computational Learning Theory*, pp. 144–152, July 1992.

[25] C. Piciarelli, C. Micheloni, and G. L. Foresti, “Trajectory-based anomalous event detection,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 11, pp. 1544–1554, 2008.

[26] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, Cambridge, UK, 2000.

[27] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, “Estimating the support of a high-dimensional distribution,” *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.

[28] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond*, MIT Press, Cambridge, Mass, USA, 2002.

[29] UMN, “Unusual crowd activity dataset of university of minnesota, department of computer science and engineering,” 2006.

[30] PETS, “Performance evaluation of tracking and surveillance (pets) 2009 benchmark data. multisensor sequences containing different crowd activities,” 2009, http://www.cvg.reading.ac.uk/PETS2009/a.html.
Submit your manuscripts at http://www.hindawi.com