



Anomaly detection based on sparse coding with two kinds of dictionaries

Shifeng Li¹  · Chunxiao Liu¹ · Yuqiang Yang¹

Received: 16 June 2017 / Revised: 26 November 2017 / Accepted: 14 January 2018 / Published online: 27 January 2018
© Springer-Verlag London Ltd., part of Springer Nature 2018

Abstract

In this paper, we propose a novel method to detect anomaly from videos based on sparse reconstruction. Different from the traditional methods, two kinds of dictionaries are employed for anomaly detection with one representing the global dictionary and the other indicating the online one. The global dictionary is first trained on training samples, and then used for the local online dictionary learning and anomaly detection. A novel updating scheme is proposed in the local online dictionary learning for an accurate anomaly detection. Experiments on the public databases show that our method can effectively detect abnormal events in complex scenes.

Keywords Sparse coding · Anomaly detection · Two kinds of dictionaries

1 Introduction

With an increasing of videos in the real world, video-based surveillance systems are being increasingly used for traffic violations, accidents, crime, terrorism, vandalism and other suspicious activities. Since manual analysis of such large volumes of data is prohibitively costly, it is necessary to develop effective algorithms to aid in the automatic or semi-automatic interpretation and analysis of video data.

Conventional methods [1–3] intend to detect testing samples with lower probability as anomaly by fitting a probability model over the training data. We propose to detect anomalies in videos based on templates in [4] and [5], where the templates represent the maximum and minimum motion distribution. Recently sparse coding scheme is applied to anomaly detection [6–9] and shows great potential. The sparse-based approaches can detect abnormal events effectively under the assumption that the normal events can be constructed from the normal basis. However, the proposed sparse coding methods mainly use only one kind of dictionary either obtained from the training data or a number of former frames of the testing video, which can not make

fully use of the training and testing samples. Motivated by the class-specific dictionary and shared dictionary proposed in [10], we propose a novel method based on sparse coding using two kinds of dictionaries with one obtained from the global training samples and the other gained from the testing video. Our method not only considers the normal events in the training data, but also takes into account the online information of the testing videos, which can improve the detection result. Our contributions are as follows:

- One novel framework using two kinds of dictionaries for anomaly detection is proposed in this paper, which can make fully use of the training and testing samples.
- We present one scheme to update the online dictionary using the trained global dictionary and the selected testing samples, which can improve the detection result.

The rest of this paper is organized as follows. Section 2 provides a brief overview of previous works on anomaly detection. The detailed explanation of our method is provided in Sect. 3. Section 4 demonstrates the effectiveness of the proposed algorithm in the published datasets, followed by conclusions in Sect. 5.

✉ Shifeng Li
lishifeng2007@gmail.com

¹ The Department of Electronic Engineering, BoHai University, JinZhou, China

2 Related works

Many methods are proposed in anomaly detection. According to whether or not abnormal samples are needed for training, the method can be classified into two types: one type needing abnormal samples and the other only needing normal samples.

The first type of anomaly detection method requires the labels for both normal and abnormal samples in the training process. This kind of method also predefines a special abnormal behavior for detection, such as fight, human fall and traffic violation. Esen et al. [11] proposed a novel motion feature named Motion Co-occurrence Feature (MCF) for fight detection in surveillance videos. Vishwakarma et al. [12] first used an adaptive background subtraction to detect the moving objects and then designed a fall model to confirm the fall action. Fu et al. [13] proposed to detect traffic abnormal events based on vehicle motion trajectories using their pairwise similarity. These methods are constrained due to the quantity limitation of abnormal samples.

The other type mainly focuses on normal events, which is the most popular technique employed by researchers. For example, trajectory based method is usually employed in anomaly detection [14]. However, object tracking is not reliable in densely crowded scenes and it is far likely to lead to unsatisfactory anomaly detection results. Model based methods usually design a local normal model based on low level feature for anomaly detection. For example, Kim and Grauman used a space-time Markov Random Field (MRF) model to detect abnormal activities in videos [1]. Kratz and Nishino employed hidden Markov models (HMM) for anomaly detection [2]. Mehran et al. introduced a novel method to detect and localize abnormal behaviors in crowded videos using social force model (SFM) [3]. These methods work well when applied to the crowded scenes due to the dense motion patterns extracted from the videos. Sparse coding based approaches are developed in [6,9,15,16], which can detect abnormal events effectively under the assumption that the normal events can be constructed from the normal basis. This method usually employs a kind of low level feature of the normal training samples to train a normal global dictionary and uses the trained dictionary to compute the reconstruction error for anomaly detection. Deep learning method is also employed in anomaly detection [10,17,18]. For example, Sakurada and Yairi [10] demonstrated that the auto encoders can be useful as nonlinear techniques to detect subtle anomalies. Sabokrou et al. [17] employed a sparse auto encoder to capture the local and global descriptors for the video properties. Medel and Savakis [18] developed a generative model based on a composite Convolutional Long Short-Term Memory (Conv-LSTM) for anomaly detection. A review modelling representation of video feature for anomaly detection is proposed in [19]. Deep learning methods are popular

in video-based tasks, owing to its ability to produce good representations with raw input. But this method requires a high computer hardware and the training process needs much more training data and time for one representation. In this paper, we propose a novel method based on two kinds of dictionaries, which can make fully use of training and testing data to make the detection results more accurate.

3 Our work

In this paper, we propose a novel method based on sparse coding. Different from the traditional sparse coding methods, we use two different dictionaries in our method, with one representing the global motion distribution and the other standing for the local patterns. First, we obtain a global dictionary from training samples using a kind of effective low-level motion feature. Second, in the testing process, we divide the testing video into several parts every some frames and estimate the anomaly using the global dictionary for the first part. Then, we use the most trusted low-level motion features obtained from the estimated anomaly detection to obtain a local online real-time dictionary, and update it using the global and local dictionaries in the following parts. Last, the abnormal events are detected by the global and online dictionaries, respectively, and the final detection results are the combination of the detected anomalies. The top algorithm flowchart is shown in Fig. 1.

3.1 Low-level motion feature

The low-level motion feature is important in anomaly detection due to its effectiveness and practicability [6,15,20]. Similarly to [20], we also develop a simple but effective motion feature for anomaly detection. Different from the traditional well-known descriptor, such as HOG [21], our new descriptor not only considers the appearance, but also the motion. The low-level feature is also used in [4,5], but the low-level feature used in these two methods does not consider the adjacent cubes.

To compute the representation of each spatio-temporal volume extracted on a regular grid, we define a descriptor based on three-dimensional gradients computed using the luminance values of pixels. The gradient is first computed using finite difference approximations and represented as G_x , G_y , G_t at each pixel (x, y, t) in the t th frame I as follows:

$$\begin{aligned} G_x &= I(x+1, y, t) - I(x-1, y, t) \\ G_y &= I(x, y+1, t) - I(x, y-1, t) \\ G_t &= I(x, y, t+1) - I(x, y, t-1) \end{aligned} \quad (1)$$

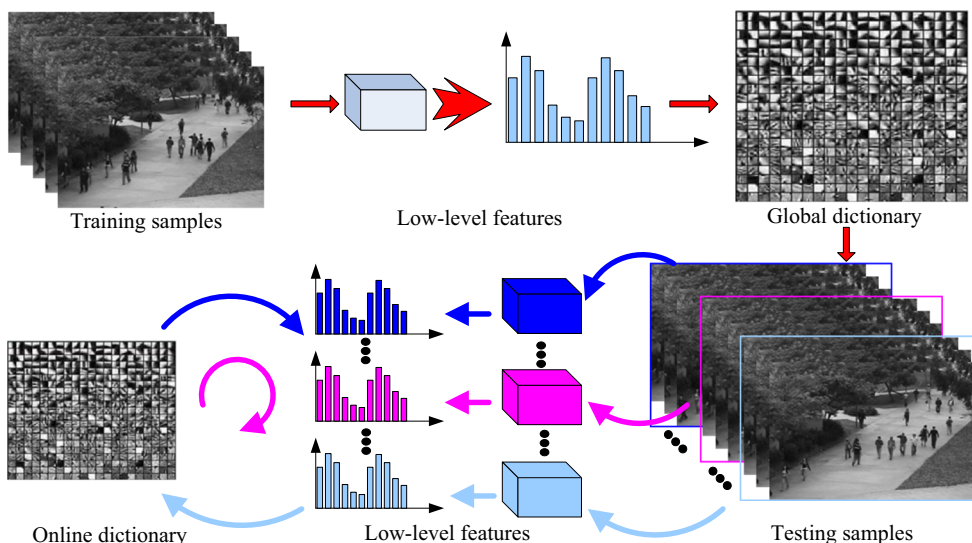


Fig. 1 The top flowchart of our method

Algorithm 1 Online updating framework.

Require: The global dictionary D_{glb} and the testing video $X = [X_1, X_2, \dots, X_N]$

Ensure: Anomaly detection result E_{fnl}

- 1: Set $D_0 = D_{glb}$
- 2: **for** $i=1$ to N **do**
- 3: Calculate the sparse coefficients α_i and α_i^{glb} by l_1 minimization with the updated D_{i-1} and global dictionary D_{glb} using Eq. (3).
- 4: Compute SRC E_i and E_i^{glb} based on the updated D_{i-1} and global dictionary D_{glb} with Eq. (4).
- 5: Calculate the final detection result using $E_i^{fnl} = \max(E_i, E_i^{glb})$ based on X_i .
- 6: Update dictionary D_{i-1} from the current video part X_i using Algorithm 2.
- 7: **end for**
- 8: Combine video abnormal detection results $E_{fnl} = [E_1^{fnl}, E_2^{fnl}, \dots, E_N^{fnl}]$.

- The computation of histogram for one block is different. We first divide the video into $10 \times 10 \times 3$ cubes, and then calculate every histogram in each cube using Eq. (2) and finally stack all histograms within one block in some order. While the method [20] computes the histogram in each subregion.
- We consider the low-level motion feature selectively. In our method, the motion cubes whose summation of gradients G_x and G_y are greater than one threshold are computed for motion feature. While the method [20] computes the histogram densely.

Then, we compute the magnitude M_{3D} and the angles ϕ and θ using the computed gradients as follows:

$$\begin{aligned}
 M_{3D} &= \sqrt{G_x^2 + G_y^2 + G_t^2} \\
 \phi &= \tan^{-1}(G_t / \sqrt{G_x^2 + G_y^2}) \\
 \theta &= \tan^{-1}(G_x / G_y)
 \end{aligned}
 \tag{2}$$

The feature descriptor for each cube includes the histograms of calculated angles ϕ and θ which are quantified into 8 bins with the magnitude M_{3D} as weight over space and temporal regions. We first divide the video into $10 \times 10 \times 3$ cubes and then combine the neighborhood 2×2 cubes as one block. The dimension for each cube is $8 + 8 = 16$, and the final descriptor for one block is $2 \times 2 \times 16 = 64$. There are two differences between our method and [20]:

3.2 Online updating scheme framework

From the training samples, we can easily obtain the global dictionary using sparse coding scheme. We argue that it is useful to update the dictionary during the testing process for a more accuracy anomaly detection result. To achieve this goal, we design a simple but effective scheme to update the online dictionary. Let X be one testing video, and suppose the testing video is divided into N parts. Therefore, we can denote the testing video as $X = [X_1, X_2, \dots, X_N]$ with X_i representing the low-level motion feature set of the i th part. For the i th divided part, we first compute the sparse coefficients α_i and α_i^{glb} based on the last updated dictionary D_{i-1} and the global dictionary D_{glb} , and then calculate the sparse reconstruction cost (SRC) E_i and E_i^{glb} . Last, we obtain the final detection result E_i^{fnl} using max pooling operation. This process is repeated for N times and the final abnormal detection result is achieved from the combination of each E_i^{fnl} . This online updating scheme framework is shown in Algorithm 1.

3.3 Online updating scheme for one video part

In this section, we introduce the core online updating scheme of step 6 in Algorithm 1 for D_{i-1} with one divided video part $X_i = [x_i^1, x_i^2, \dots, x_i^n]$, where x_i^j stands for the j th cube's low level feature in the i th part X_i and n is the total number of cubes for part X_i . This scheme first calculates the sparse coefficients α_i using the last updated dictionary D_{i-1} with Eq. (3) and computes the current SRC E_i with Eq. (4).

$$\alpha_i = \arg \min_{\alpha} \frac{1}{2} \|X_i - D_{i-1}\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (3)$$

$$E_i = \frac{1}{2} \|X_i - D_{i-1}\alpha_i\| + \lambda \|\alpha_i\|_1 \quad (4)$$

where α is the sparse coefficients and λ is a regularization parameter. Then, the method selects the samples from X_i to update the dictionary D_{i-1} using an adaptive threshold ρ based on the computed E_i .

$$\rho = m(E_i) - C(E_i) \quad (5)$$

where $m(E_i)$ is the mean of E_i and $C(E_i)$ is the variance of E_i . Finally, we update D_{i-1} using the selected samples based on Eq. (6) according to [22].

$$D_i = \prod_C [D_{i-1} - \frac{\eta}{i} \nabla_D l(X_i, D_{i-1})] \quad (6)$$

where $l(X_i, D_{i-1}) = \frac{1}{2} \sum_{j=1, \dots, n_i} \|X_i^j - D_{i-1}\alpha_i^j\|_2^2$, η is the learning rate, \prod_C is the orthogonal projection onto \mathcal{C} . For the details, please refer to [22]. The details of updating process using one video part X_i are shown in Algorithm 2.

Algorithm 2 Online updating scheme in step 6 of Algorithm 1.

Require: Last updated dictionary D_{i-1} and current testing video part $X_i = [x_i^1, x_i^2, \dots, x_i^n]$

Ensure: Updated dictionary D_i

- 1: **for** $i=1$ to n **do**
 - 2: Pursuit the sparse coefficient α_i by l_1 minimization with the last D_{i-1} using Eq. (3) based on X_i .
 - 3: Compute the SRC E_i using Eq. (4).
 - 4: Select samples X_i' from X_i by an adaptive threshold ρ , which is computed with Eq. (5).
 - 5: Use X_i' to update D_{i-1} with Eq. (6).
 - 6: **end for**
-

3.4 Anomaly detection

As briefly mentioned in previous section (refer to Algorithm 1), given the newly observed frames X' and the global dictionary D_{glob} , we can compute the detection result E_{fnl} as the

abnormality. Intuitively, A high SRC value implies a high reconstruction cost and a high probability of being an abnormal sample. Similarly to [15], the detected unusual events can be further located with $E_{abs} = E_{fnl} > \varepsilon$, where ε is a user defined threshold that controls the sensitivity of the algorithm to unusual events.

4 Experiments

In this section, we show the experimental performance of the proposed anomaly detection algorithm, both qualitatively and quantitatively on the published datasets UCSD Ped1 and UMN [23].

4.1 Parameters setting

In our experiments, the parameters include the regularization parameter λ , dictionary size K in Eq. (3) and the learning rate η in Eq. (6). In our implementation, we set $\eta = 10^4$, $\lambda = 1.2/\sqrt{m}$ and $K = 256$, where m is the patch size. The parameter settings are the same as in [22]. We find that the results are not sensitive to the parameters.

4.2 Experiments on UCSD ped1

The UCSD Ped1 dataset contains the training and testing two sets. The training set contains 34 short clips for learning of normal patterns, and there is a subset of 10 clips in testing set provided with pixel-level binary masks, which identify the regions whether containing abnormal events or not. Each clip has 200 frames, with a 158×238 resolution. We first resize each video frame into 120×160 , and then split the video frames into $10 \times 10 \times 3$ cubes with no overlap along the temporal direction to obtain a 64 dimension motion feature using the method of Sec. 3.1. The whole testing video is divided into 5 parts with each part size $n = 40$.

4.2.1 Necessity of the updated dictionary

Intuitively, the ability of the global dictionary is restricted due to the limitation of training samples, it is necessary to update global dictionary according to the online video frames. In our case, we find that the accuracy of the sparse coding is improved with the updating steps in Algorithm 2. Fig. 2a shows an example for this case. In this comparison, we first random select one testing video, and then compute the SRC based on one online dictionary and the global dictionary respectively, and last plot the random selected 1000 cubes in the comparison. In Fig. 2a, the red line stands for the SRC with global dictionary and the blue is the SRC with the online updated scheme. It is obvious that the SRC value of the online method is smaller than the global dictionary based

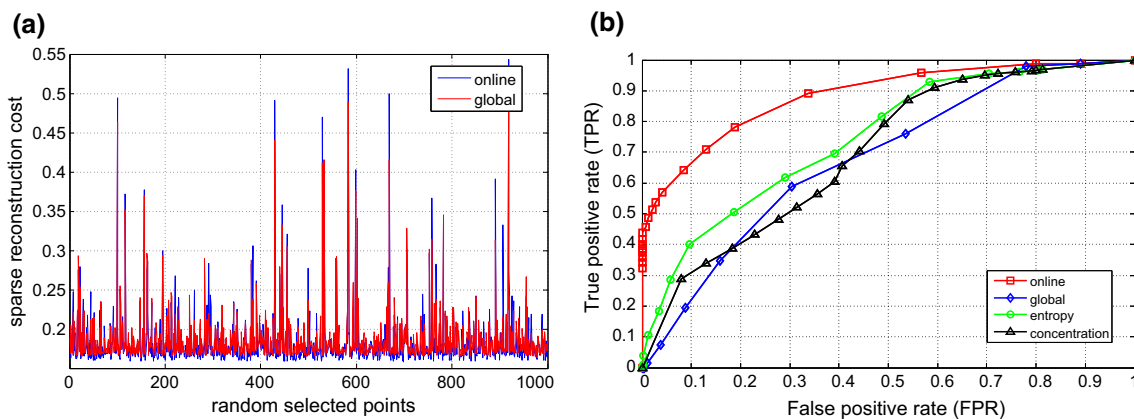


Fig. 2 **a** Comparison of sparse reconstruction cost between the method with updated dictionary and the one without. **b** Comparison of different measurements

Table 1 Area under ROC of different measurements

Method	Area under ROC
Online dictionary	0.9500
global dictionary	0.7094
Concentration function similarity	0.7298
Entropy	0.7647

method in most cases, except the cubes which are possibly abnormal. It suggests that the online updated scheme can not only reduce the SRC of the normal points, but also enlarge the SRC of the abnormal cubes, which can greatly benefit the subsequent anomaly detection.

4.2.2 Comparison with different methods

Similarly to [6], we use receiver operating characteristic (ROC) curve where the x-axis is false positive rate (FPR) and y-axis is true positive rate (TPR) for the comparison. In Fig. 2b, the ROC curves by frame-level measurement are shown to compare SRC using online updating scheme

to three other measurements, which are (1) SRC measurement using global dictionary. (2) Entropy metric: $S_E = -\sum_i p_i \log(p_i)$, where $p_i = |\alpha(i)| / \|\alpha\|_1$. Here $\alpha(i)$ is the i th value of sparse coefficient α , $|\alpha(i)|$ and $\|\alpha\|_1$ is the absolute value and l_1 norm of α , respectively. Intuitively, the sparse coefficients will lead to a small entropy value. (3) Concentration function (CF) which is similar to [24], $S_s = T_k(\alpha) / \|\alpha\|_1$, where $T_k(\alpha)$ is the sum of the k largest positive coefficients of α (the greater the S_s the more likely a normal testing sample).

The red line stands for the SRC with online updating method, the green represents entropy measurements with global dictionary, the black line is the CF measurement with global dictionary, and the blue indicates the SRC measurement with global dictionary. Table 1 shows the area under ROC curve. We can see that the method using online updating scheme outperforms other methods.

Figure 3a shows the comparison with different methods including sparse method [6], MDT [25], SFM [3], MPPCA [25], Adam [26], Sparse 150 [16], Grid Template [4] and MAP method [5]. Our method outperforms the methods MPPCA+SF, MPPCA, SF, MDT and Sparse 150, and has

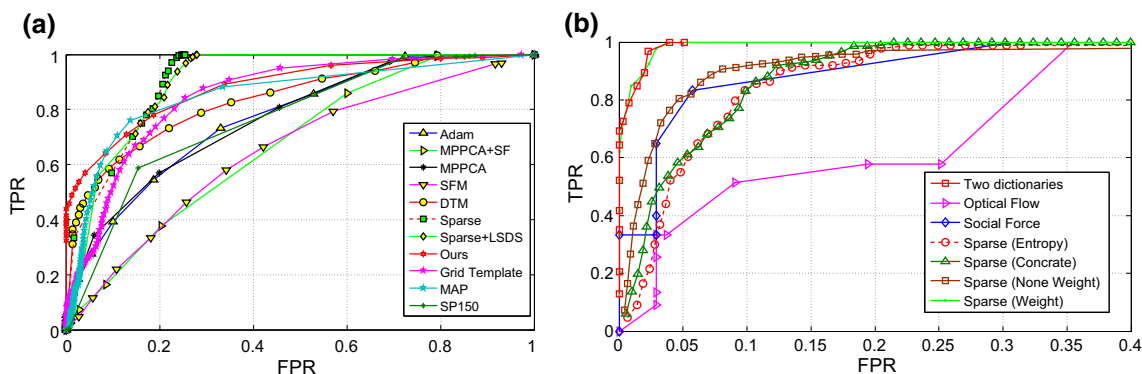


Fig. 3 **a** Comparison of different methods on UCSD Ped1. **b** Comparison of different measurements in UMN dataset

Table 2 Running time comparison on the UCSD Ped1

Method	Second/frame	Platform	CPU	Memory
[25]	25	–	3.0GHz	2G
[6]	3.8	–	2.6 GHz	2G
[27]	5~10	Matlab	–	–
[16]	0.00697	Matlab 2012	3.4GHz	8G
Ours	0.0714	Matlab 2012a	3.4GHz	6G

a similar performance to MAP [5] and Grid Template [4] methods, and is comparable to sparse method [6].

4.2.3 Time cost of our method

In this section, we provide the time cost of our method on UCSD Ped1 dataset. The code is implemented using matlab 2012a on the desktop computer with 3.4Hz double-core CPU and 6G memory. 20 videos are randomly selected from UCSD Ped1 dataset to run the code and the final time cost is computed as the mean. The comparison of running time is shown in Table 2.

4.3 Experiments on UMN dataset

The UMN dataset consists of 3 different scenes of crowded escape events, and the total frame number is 7740 with a 320×240 resolution. Analogously to the experiments on UCSD Ped1 dataset, we first resize each video frame into a 120×160 resolution and then extract the low-level motion feature from the first 400 frames of each scene to train the global dictionary, and leave the others for testing.

In our implementation, the testing video is divided into several subparts with each part including 20 frames. Cong et al. [6] employed different measurements including SRC, entropy and concentration to detect abnormal events based on sparse representation. We compare our method with the sparse method proposed by [6] in Fig. 3b. Our method is comparable to method [6] and better than the other methods. It is necessary to note that our method detects abnormal events with the same feature in all experiments while the method [6] detects abnormal events with different features for different videos. For example, they used spatial feature in UMN dataset while employed spatio-temporal basis in UCSD Ped1 dataset.

5 Conclusion

In this paper, we propose an algorithm to automatically detect abnormal events from video sequences based on sparse coding with two kinds of dictionaries. A global dictionary is first

trained with the training data, and a online dictionary is also trained with the most trusted online data obtained from the global dictionary, and updated in the following sequences. The max-pooling method is employed to combine the detection result obtained from the global and online dictionary. Experimental results on two public datasets demonstrate the effectiveness of our algorithm. Our method detects anomalies from each cube one by one, does not consider the continuity between each cubes. It is the fact that the anomaly usually happens in continuous cubes, so it is necessary to consider anomaly continuity in the future.

Acknowledgements This work was supported by the Natural Science Foundation of China (NSFC), No. 61402049.

References

- Kim, J., Grauman, K.: A space-time MRF for detecting abnormal activities with incremental updates. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), pp. 2921–2928, 20–25 June 2009, Miami, Florida, USA (2009)
- Kratz, L., Nishino, K.: Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), pp. 1446–1453, 20–25 June 2009, Miami, Florida, USA (2009)
- Mehran, R., Oyama, A., Shah, M.: Abnormal crowd behavior detection using social force model. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), pp. 935–942, 20–25 June 2009, Miami, Florida, USA (2009)
- Li, S., Yang, Y., Liu, C.: Anomaly detection based on two global grid motion templates. *Signal Process. Image Commun.* (2017)
- Li, S., Liu, C., Yang, Y.: Anomaly detection based on maximum a posteriori. *Pattern Recognit. Lett.* (2017)
- Cong, Y., Yuan, J., Liu, J.: Sparse reconstruction cost for abnormal event detection. In: The 24th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011, pp. 3449–3456, 20–25 June 2011, Colorado Springs, CO, USA (2011)
- Xu, J., Denman, S., Sridharan, S., Fookes, C., Rana, R.: Dynamic texture reconstruction from sparse codes for unusual event detection in crowded scenes. In: Proceedings of the 2011 Joint ACM Workshop on Modeling and Representing Events, J-MRE '11, pp. 25–30 (2011)
- Adler, A., Elad, M., Hel-Or, Y., Rivlin, E.: Sparse coding with anomaly detection. In: IEEE International Workshop on Machine Learning for Signal Processing (MLSP), 2013, pp. 1–6 (2013)
- Mo, X., Monga, V., Bala, R., Fan, Z.: Adaptive sparse representations for video anomaly detection. *IEEE Trans. Circuits Syst. Video Technol.* **24**(4), 631 (2014)
- Sakurada, M., Yairi, T.: Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, Gold Coast, Australia, QLD, Australia, December 2, 2014, p. 4. (2014). <https://doi.org/10.1145/2689746.2689747>. <http://doi.acm.org/10.1145/2689746.2689747>
- Esen, E., Arabaci, M.A., Soysal, M.: Fight detection in surveillance videos. In: International Workshop on Content-Based Multimedia Indexing, pp. 131–135 (2013)
- Vishwakarma, V., Mandal, C., Sural, S.: Automatic detection of human fall in video. In: Pattern Recognition and Machine

- Intelligence, Second International Conference, PReMI 2007, pp. 616–623, Kolkata, India, December 18–22, 2007, Proceedings (2007)
13. Fu, Z., Hu, W., Tan, T.: Similarity based vehicle trajectory clustering and anomaly detection. In: IEEE International Conference on Image Processing, pp. II–602–5 (2005)
 14. Basharat, A., Gritai, A., Shah, M.: Learning object motion patterns for anomaly detection and improved object detection. In: 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24–26 June 2008, Anchorage, Alaska, USA (2008)
 15. Zhao, B., Li, F., Xing, E.P.: Online detection of unusual events in videos via dynamic sparse coding. In: The 24th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011, pp. 3313–3320, 20–25 June 2011, Colorado Springs, CO, USA (2011)
 16. Lu, C., Shi, J., Jia, J.: Abnormal event detection at 150 FPS in MATLAB. In: IEEE International Conference on Computer Vision, ICCV 2013, Sydney, Australia, December 1–8, 2013, pp. 2720–2727 (2013)
 17. Sabokrou, M., Fathy, M., Hoseini, M., Klette, R.: Real-time anomaly detection and localization in crowded scenes. In: Computer Vision and Pattern Recognition Workshops, pp. 56–62 (2016)
 18. Medel, J.R., Savakis, A.E.: Anomaly detection in video using predictive convolutional long short-term memory networks. CoRR (2016). <http://arxiv.org/abs/1612.00390> [abs/1612.00390]
 19. Chong, Y.S., Tay, Y.H.: CoRR (2015). [arxiv:1505.00523](https://arxiv.org/abs/1505.00523)
 20. Bertini, M., Bimbo, A.D., Seidenari, L.: Multi-scale and real-time non-parametric approach for anomaly detection and localization. *Comput. Vision Image Underst.* **116**(3), 320 (2012)
 21. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005, vol. 1 (2005), pp. 886–893. <https://doi.org/10.1109/CVPR.2005.177>
 22. Mairal, J., Bach, F.R., Ponce, J., Sapiro, G.: Anomaly detection in crowded scenes. In: Proceedings of the 26th Annual International Conference on Machine Learning, ICML 2009, Montreal, Quebec, Canada, June 14–18, 2009, pp. 689–696 (2009)
 23. UMN, http://mha.cs.umn.edu/proj_events.shtml
 24. Wright, J., Yang, A., Ganesh, A., Sastry, S., Ma, Y.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210 (2009). <https://doi.org/10.1109/TPAMI.2008.79>
 25. Mahadevan, V., Li, W., Bhalodia, V., Vasconcelos, N.: Anomaly detection in crowded scenes. In: The Twenty-Third IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2010, San Francisco, CA, USA, 13–18 June 2010, pp. 1975–1981 (2010)
 26. Adam, A., Rivlin, E., Shimshoni, I., Reinitz, D.: Robust real-time unusual event detection using multiple fixed-location monitors. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(3), 555 (2008)
 27. Antic, B., Ommer, B.: Video parsing for abnormality detection. In: IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, November 6–13, 2011, pp. 2415–2422 (2011)