LOW RANK SPARSITY PRIOR FOR ROBUST VIDEO ANOMALY DETECTION

Xuan Mo[†] *Vishal Monga*[†] *Raja Bala*[⋆] *Zhigang Fan*[⋆] *Aaron Burry*[⋆]

[†]Department of Electrical Engineering, Pennsylvania State University, USA
[⋆]Xerox Research Center, Webster, NY, USA

ABSTRACT

Recently, sparsity based classification has been applied to video anomaly detection. A linear model is assumed over video features (e.g. trajectories) such that the feature representation of a new event is written as a sparse linear combination of existing feature representations in the dictionary. Sparsity based video anomaly detection shows promise but open challenges remain in that existing methods assume object specific and class specific event dictionaries making them applicable mostly in highly structured scenarios. Second, using conventional sparsity models on matrices/vectors, the computational burden is often high. In this work, we advocate a more general and practical sparsity model using a low-rank structure on the matrix of sparse coefficients. We find that enforcing a low-rank structure can ease the rigidity of traditional row-sparse constraints on sparse coefficient vectors/matrices. Because low-rank matrices are of course not always sparse, an additional l_1 regularization term is added. Further, if rank is substituted by its convex nuclear norm alternative, then significant computational benefits can be obtained over existing methods in sparsity based video anomaly detection. Experimental evaluation on benchmark video datasets reveal, our method is competitive with state-of-the art while providing robustness benefits under occlusion.

1. INTRODUCTION

Vast amounts of video footage are collected and analyzed for traffic violations, accidents, crime, terrorism, vandalism, and other suspicious activities. An active area of research within this domain is video anomaly detection, which refers to the problem of *automatically* finding patterns in data that do not conform to expected behavior, and that may warrant special attention or action. Video anomaly detection involves *encoding* an event followed by a decision rule often facilitated by a model. Approaches include the use of finite state machines [1], Markov chain models [2], hierarchical Bayesian models [3], decision trees [4], Hidden Markov Model (HMM) [5], infinite hidden Markov model [6] and Support Vector Machines (SVM) [7]. An excellent overview of video anomaly detection techniques can be found in [2].

RESEARCH WAS SUPPORTED BY A GRANT FROM THE XEROX RESEARCH CENTER IN WEBSTER, NY. THIS PAPER CONTAINS HYPERLINKS TO EXTERNAL VIDEOS.

Relation to Prior Work: Sparse reconstruction techniques [8], [9], [10] have recently been employed in video anomaly detection. The fundamental premise underlying these methods is that any new feature representation of a normal/anomalous event can be approximately modeled as a (sparse) linear combination of pre-labeled feature representations (of previously observed events) in a training dictionary. Li *et al.* [8] use object trajectories as event descriptors while Zhao *et al.* [9] use features extracted from spatio-temporal volumes. The approaches in [8], [9] were motivated by sparsity based face recognition proposed by Wright *et al.* [11] which demonstrated sparse representations could exhibit robustness to significant amounts of noise and occlusion. Extensions of vector sparsity to multi-task scenarios, e.g. multi-object anomaly detection have been developed [10] which employ a row-sparse structure on the sparse coefficient matrix.

Motivation and Contribution: Sparsity based video anomaly detection shows promise but open challenges remain in that existing methods assume object specific and class specific event dictionaries making them applicable mostly in highly structured scenarios. Second, using conventional sparsity models on matrices/vectors, the computational burden is often high. In this work, we advocate a more general and practical sparsity model using a low-rank structure on the matrix of sparse coefficients. We find that enforcing a low-rank structure can ease the rigidity of traditional row-sparse constraints on sparse coefficient vectors/matrices. A significant practical benefit with the proposed method is that it is not necessary to assign class labels to the normal trajectories, and therefore the manual effort in building the training dictionary is much reduced. All the normal trajectories are collected together as a big dictionary, and there is no need to group training trajectories (or events) into different classes as is done in [8], [9], [10]. Because low-rank matrices are of course not always sparse, an additional l_1 regularization term is added. Further, if rank is substituted by its convex nuclear norm alternative, then significant computational benefits can be obtained over existing methods in sparsity based video anomaly detection. Experiments on benchmark video datasets reveal that our method is competitive with state-of-the art while providing robustness benefits under occlusion.

978-1-4799-2893-4/14/$31.00 ©2014 IEEE

2. SPARSITY-BASED ANOMALY DETECTION

In the rest of the paper, we will focus on object trajectories as the *event encoding* or video feature of choice. Let each trajectory representation lie in $\mathbb{R}^N$, and t denote the number of training samples (i.e. example trajectory representations) from each of k different classes, i.e. behavior patterns in a video which may be normal or anomalous. The t training samples (trajectory representations) from the i-th class are arranged as the columns of a matrix $\mathbf{A}_i \in \mathbb{R}^{N\times t}$. The dictionary $\mathbf{A} \in \mathbb{R}^{N\times T}$ ($T = kt$) of training samples from all classes is then formed as follows: $\mathbf{A} = [\mathbf{A}_1\ \mathbf{A}_2\ \ldots \mathbf{A}_k]$.

Given a sufficient number of training samples from the m-th trajectory class, a test image $\mathbf{y} \in \mathbb{R}^N$ from the same class is conjectured to approximately lie in the linear span of those training samples. Any trajectory feature vector is synthesized by a linear combination of the set of all training trajectory samples as follows:

$$\mathbf{y} \approx \mathbf{A}\boldsymbol{\alpha} = [\mathbf{A}_1\ \mathbf{A}_2\ \ldots \mathbf{A}_k] \begin{bmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \\ \vdots \\ \boldsymbol{\alpha}_k \end{bmatrix}, \qquad (1)$$

where each $\boldsymbol{\alpha}_i \in \mathbb{R}^t$. Typically for an example trajectory $\mathbf{y}$, only one of the $\boldsymbol{\alpha}_i$'s will be active (corresponding to the class/event from which $\mathbf{y}$ is generated). Thus the coefficient vector $\boldsymbol{\alpha} \in \mathbb{R}^T$ is modeled as *sparse* and is recovered by solving the following optimization problem:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \text{ subject to } \|\mathbf{y} - \mathbf{A}\boldsymbol{\alpha}\|_2 < \varepsilon, \qquad (2)$$

where the objective is to minimize the number of non-zero elements in $\boldsymbol{\alpha}$. The residual error between the test trajectory and each class behavior pattern is computed to find the class to which the test trajectory belongs:

$$r_i(\mathbf{y}) = \|\mathbf{y} - \mathbf{A}_i\hat{\boldsymbol{\alpha}}_i\|_2 \quad i = 1, 2, \ldots, K \qquad (3)$$

Fig. 1 shows an example of event classification using the sparsity model. The training dictionary consists of 2 classes, each with 4 different trajectories. The test trajectory can be well represented by the linear combination of trajectory no. 1 and trajectory no. 3 from class 1 (see Fig. 1). This is infact tantamount to saying that the coefficient vector $\boldsymbol{\alpha}$ is indeed sparse - in this example, two of eight entries being active.

3. LOW RANK SPARSITY PRIOR FOR VIDEO ANOMALY DETECTION

3.1. Motivation: Low Rank Sparsity Prior

The aforementioned set-up in (1)-(3) assumes that training is available from both normal and anomolous events and hence anomaly detection reduces to a classification problem. In the absence of training from anomalous events (the more practical scenario), outlier rejection measures [9], [10] on the sparse vector $\boldsymbol{\alpha}$ may be used to detect anomalies. A careful preparation of the dictionary $\mathbf{A}$ is neverthless needed often with training examples that are manually labeled to belong to particular event classes. Multi-object or multi-view anomaly detection leads to a sparse coefficient matrix (and not vector), in those cases training dictionaries are often labeled not only per class but also per object [8], [10] leading to a row-sparse structure on the coefficient matrix. Such elaborate preparation of the dictionary is sometimes unrealistic and invariably burdensome requiring a pre-analysis of video footage prior to anomaly detection. Fig. 2(a) (the video is available at: `http://youtu.be/M6_PJigg5CY`) shows an example video frame of a structured scenario (detection of stop sign violations) where preparation of a dictionary clearly separated into class-specific sub-dictionaries is possible. In many other settings however, multiple trajectories are simultaneously extracted and a clear separation into normal event classes is difficult. An example of a video frame from such a scenario is shown in Fig. 2(b) (the video is available at: `http://youtu.be/jEzLkWF65Io`).

We therefore seek a more general and practical sparsity prior which can deal with unstructured scenarios. Note that, row-sparse matrices that have been used in [9], [10] for anomaly detection are also low-rank. Inspired by this observation and known connections between low-rank and sparse matrices [12], we propose using a low rank sparsity prior.

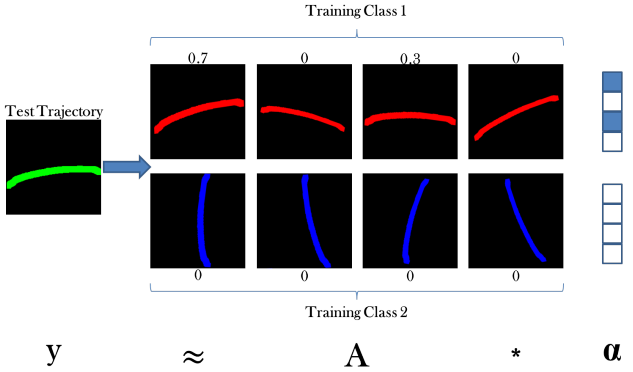

Fig. 1. An example illustration of trajectory classification using a sparse reconstruction model.

(a) (b)

Fig. 2. (a) structured scenario (b) unstructured scenario.

3.2. Low Rank Sparsity Prior for Anomaly Detection

In the proposed framework, there is no need to group training trajectories into different normal event classes. All observed training trajectories corresponding to normal events are collected together as a big dictionary: $\mathbf{A} \in \mathbb{R}^{N \times T}$. We also collect M test trajectories extracted from the video into a matrix $\mathbf{Y} = \{\mathbf{y}_i\} \in \mathbb{R}^{N \times M}, i = 1, \dots, M$.

Under a linear model $\mathbf{Y} \approx \mathbf{AS}$, and given sufficient training, the coefficient matrix $\mathbf{S} \in \mathbb{R}^{T \times M}$ is expected to be sparse. Making a departure from the typical $\|\|_{\text{row},0}$ norm, we propose to use a low-rank structure to measure the sparsity of $\mathbf{S}$. A couple of examples of simultaneously sparse and low rank matrices are:

$$\mathbf{S}_1 = \begin{pmatrix} 1&1&1&1&1&1&1&1&1&1\\ 1&1&1&1&1&1&1&1&1&1\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0 \end{pmatrix} \mathbf{S}_2 = \begin{pmatrix} 0&0&1&1&1&1&1&1&1&0\\ 0&0&0&1&1&1&1&1&1&1\\ 1&0&0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0&0&0\\ 0&1&0&0&0&0&0&0&0&0 \end{pmatrix}$$

Then, we propose to replace (2) by:

$$\begin{array}{ll} \text{minimize} & \text{rank}(\mathbf{S}) \\ \text{subject to} & \|\mathbf{Y} - \mathbf{AS}\|_F \le \varepsilon, \end{array} \tag{4}$$

A convex relaxation of (4) can be obtained via subsituting rank($\mathbf{X}$) by $\|\mathbf{X}\|_* = \sum_i \sigma_i(\mathbf{X})$ (where $\|\|_*$ denotes nuclear norm and $\sigma_i(\mathbf{X})$ is the i-th singular value of $\mathbf{X}$) [13]. This results in a convex optimization problem:

$$\begin{array}{ll} \text{minimize} & \|\mathbf{S}\|_* \\ \text{subject to} & \|\mathbf{Y} - \mathbf{AS}\|_F \le \varepsilon. \end{array} \tag{5}$$

While low-rank and sparse matrix structures often simultaneously exist (as is expected here as well), in general the two are not the same, and low-rank does not imply sparsity. To encourage sparse matrices which are simultaneously low-rank, we further add a l_1 regularization term to the cost function and convexity still holds:

$$\begin{array}{ll} \text{minimize} & \|\mathbf{S}\|_* + \lambda \|\mathbf{S}\|_1 \\ \text{subject to} & \|\mathbf{Y} - \mathbf{AS}\|_F \le \varepsilon. \end{array} \tag{6}$$

Anomaly Detection: Once we get the optimal coefficient matrix $\hat{\mathbf{S}}$, the recovered trajectory can be computed using columns of $\hat{\mathbf{S}} = \{\hat{\mathbf{s}}_1, ..., \hat{\mathbf{s}}_M\}$:

$$\hat{\mathbf{y}}_i = \mathbf{A}\hat{\mathbf{s}}_i, \tag{7}$$

where those test trajectories which are very similar to the recovered trajectories can be regarded as normal trajectories.

$$\frac{\|\mathbf{y}_i - \hat{\mathbf{y}}_i\|_2}{\|\mathbf{y}_i\|_2} < \tau \rightarrow \mathbf{y}_i \text{ is } \textbf{normal}. \tag{8}$$

Computational Complexity:[1] The problem in (6) can in fact be cast as a semidefinite program (SDP) [14]. This semidefinite program can then be solved using a "custom" interior point method [15], [16] and has an *average* complexity of $O(N^2TM)$ where N, T, M are as stated above.

On the other hand, (2) (minimizing the l_0 norm) is well-known to be an NP-hard problem. Thus the l_1 norm is often used as an effective approximation to l_0. Several fast l_1-minimization algorithms have been published [17]. The Homotopy method is amongst the most popular algorithms and has a computational complexity of its j-th iteration as $O(jN^2 + jNT)$ [18]. Let J denote the number of iterations, the total complexity becomes: $O(\sum_{j=1}^{J}(jN^2 + jNT)) = O(J^2N^2 + J^2NT)$. Here, J depends on the number of non-zero elements in $\boldsymbol{\alpha}$, so $O(J) = O(N)$. Therefore, the computational complexity of evaluating one test trajectory (event representation) using (2) is $O(N^4 + N^3T)$. Since there are M trajectories, the total computational complexity is $O(N^4M + N^3TM)$. The proposed method hence has much lower complexity, this is experimentally confirmed in Section 4.

[1] A full comparison of complexity entails memory, number of cannonical operations (adds/multiplies) and their interaction with the implementation architecture. Such analysis is beyond the scope of this article and will be considered in future work. We present a comparison of *popular* algorithms in terms of the number of fundamental operations (multiplies).

4. EXPERIMENTAL RESULTS

The datasets used in our experiments are: 1.) The well-known Public Data Set for Traffic Video (PDTV) [19] and 2.) The Xerox Stop Sign data set (while this video data set is proprietary, an example video clip is available at: `http://youtu.be/M6_PJigg5CY`). Figs. 3 (a) and (b) show an example anomaly in PDTV data, where a car fails to yield to oncoming car while turning left. Figs. 3 (c) and (d) show an example anomaly in Xerox Stop Sign data, where a driver backs his car in front of stop sign.

In all subsequent experiments, object trajectories are used to represent events. We use well-known techniques to extract trajectories [20] as a collection of coordinate pairs $[x(t), y(t)]$. We approximate a raw trajectory using a basic B-spline function [21] with 50 knots (50 x-coordinates and 50 y-coordinates) and these knots are finally used to form the trajectory feature vector.

4.1. Comparison against a State of the Art Trajectory-based Video Anomaly Detection Technique

Our proposed algorithm is called low rank sparsity prior (abbreviated to LRSP). We compare LRSP against a widely cited method by Piciarelli *et al.* [7] which is based on trajectory extraction and one class SVMs. For the experiment involving the PDTV data set, we obtain a training dictionary consisting of 319 normal event trajectories (No training corresponding to anomalous events was used). 117 normal trajectories and 24 anomalous trajectories are used as independent test data. The confusion matrices of LRSP are compared with Piciarelli *et al* [7] in Table 1. The benefits of LRSP are readily apparent.

For the Xerox Stop Sign data set, the training dictionary comprises 72 normal trajectories. We specifically conduct this experiment to compare the performance of the two methods when object occlusion is involved, i.e. *occluded trajec-*

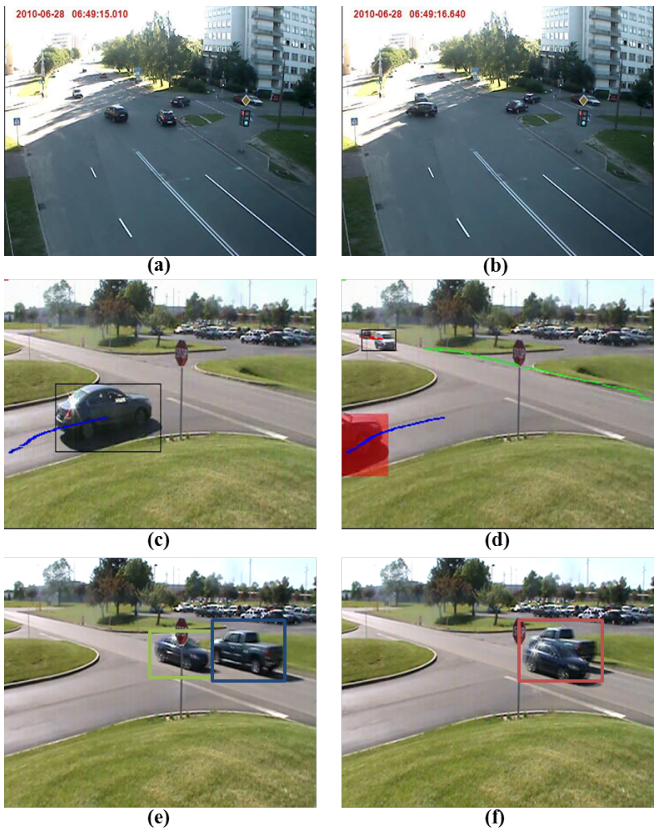

Fig. 3. (a) (b): Example anomaly in PDTV data, (c) (d): Example anomaly in Xerox Stop Sign data, (e) (f): Example frames that show object occlusion.

Table 1. Confusion matrices of PDTV data

	LRSP		Piciarelli *et al.* [7]	
	Normal	Anomaly	Normal	Anomaly
Normal	78.6%	37.5%	70.9%	41.7%
Anomaly	21.4%	62.5%	29.1%	58.3%

tories are used as our test data. Figs. 3 (e) and (f) show an example where a car is occluded by another car (the video is available at: `http://youtu.be/4Azh2yZjA4o`). An independent set of 13 *normal but occluded* trajectories and 6 *anomalous but occluded trajectories* are used to test our approach. The confusion matrices of both methods - LRSP and the method in Picarelli *et al.* are reported in Table 2. In this case LRSP is vastly better. This can be reasoned as follows: The optimization problem in (6) is well-conditioned. A set of occluded trajectories $\mathbf{Y}_o$ can be thought of as a perturbation on non-occluded trajectories $\mathbf{Y}$ and if $\|\mathbf{Y}_o - \mathbf{Y}\|_2$ is small enough, then by perturbation theory the solution $\hat{\mathbf{S}}$ under occlusion should only change slightly. This robustness of LRSP is a major practical benefit in real-world surveillance videos for example where noise and occlusion are typical.

4.2. Performance Variation with Regular Parameter λ

In our optimization problem in (6), there is parameter λ which controls the relative importance of $\|\cdot\|_*$ and $\|\cdot\|_1$ terms. In

Table 2. Confusion matrices of Stop Sign occluded data

	LRSP		Piciarelli *et al.* [7]	
	Normal	Anomaly	Normal	Anomaly
Normal	76.9%	33.3%	61.5%	50.0%
Anomaly	23.1%	66.7%	38.5%	50.0%

Table 3. Confusion matrices of Xerox Stop Sign data

	LRSP		ESP	
Run time	37 seconds		159 seconds	
	Normal	Anomaly	Normal	Anomaly
Normal	88.2%	25.0%	91.2%	25.0%
Anomaly	11.8%	75.0%	8.8%	75.0%

Fig. 4, we plot the detection rate curves against the value of λ for PDTV data. Fig. 4 reveals that $\lambda \in [0.25, 0.75]$ leads to good performance. Both excessively low and high values of λ lead to a loss in performance. In particular, when λ is large, the cost function reduces largely to the $\|\cdot\|_1$ matrix norm and the performance drop is very significant. This emphasizes the value of the low-rank term which allows greater generality over row-sparsity and can capture sparse matrix structures arising in real-world scenarios. Note the results of LRSP in Tables 1 - 3 are reported using the "best" λ.

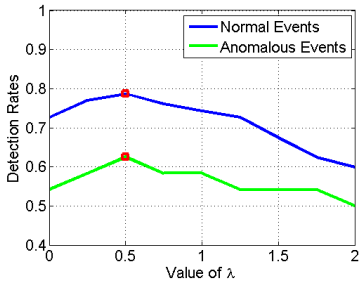

Fig. 4. Detection rates curves with respect to the value of λ

4.3. Computational Benefits and Trade Off

We now compare our proposed method against existing sparsity models as described in Section 2 (abbreviated to E-SP[2]) [8–10]. Since ESP can only detect video anomaly in structured scenarios, the Xerox Stop Sign data set is used to test both LRSP and ESP.

For the Xerox Stop Sign data set, the training dictionary contains 9 normal event classes (containing 8 trajectories each) and 1 anomalous trajectory class (containing 4 trajectories). An independent set of 34 normal trajectories and 8 anomalous trajectories are used to test our approach. Table 3 shows the confusion matrices and run time (total run time of running all the 42 test trajectories) of LRSP and ESP. We can see that the proposed LRSP method runs much faster than the ESP with a small loss in detection rates. This is expected because ESP has the benefit of pre-labeled event classes. In structured scenarios, the performance of ESP in fact serves as the practical upper bound for LRSP.

[2]Although methods in [8], [9] and [10] use varying event representations, their underlying sparsity model is the same. We use the abbreviation "ESP" to represent these three techniques.

5. REFERENCES

[1] C.-H. Chuang, *et al.*, "Carried object detection using ratio histogram and its application to suspicious event analysis," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 19, no. 6, pp. 911–916, Jun. 2009.

[2] V. Saligrama, J. Konrad, and P. Jodoin, "Video anomaly identification," *IEEE Signal Processing Magazine*, vol. 27, no. 5, pp. 18–33, Sept. 2010.

[3] X. Wang, X. Ma, and W. Grimson, "Unsupervised activity perception in crowded and complicated scenes using hierarchical bayesian models," *IEEE Trans. on Pattern Analysis and Machine Int.*, vol. 31, no. 3, pp. 539–555, Mar. 2009.

[4] C. Simon, J. Meessen, and C. De Vleeschouwer, "Visual event recognition using decision trees," *Multimedia Tools Appl.*, vol. 50, no. 1, pp. 95–121, Oct. 2010.

[5] N. Vaswani, A. Roy-Chowdhury, and R. Chellappa, "Shape activity: a continuous-state hmm for moving/deforming shapes with application to abnormal activity detection," *IEEE Trans. on Image Processing*, vol. 14, no. 10, pp. 1603–1616, Oct. 2005.

[6] I. Pruteanu-Malinici and L. Carin, "Infinite hidden markov models for unusual-event detection in video," *IEEE Trans. on Image Processing*, vol. 17, no. 5, pp. 811–822, May. 2008.

[7] C. Piciarelli, C. Micheloni, and G. Foresti, "Trajectory-based anomalous event detection," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 18, no. 11, pp. 1544–1554, Nov. 2008.

[8] C. Li, Z. Han, Q. Ye, and J. Jiao, "Abnormal behavior detection via sparse reconstruction analysis of trajectory," in *Proc. IEEE Int. Conf. Image and Graphics*, Aug. 2011, pp. 807–810.

[9] B. Zhao, L. Fei-Fei, and E. Xing, "Online detection of unusual events in videos via dynamic sparse coding," in *Proc. IEEE Conf. Computer Vision Pattern Recognition*, Jun. 2011, pp. 3313–3320.

[10] X. Mo, V. Monga, R. Bala, and Z. Fan, "Adaptive sparse representations for video anomaly detection," *IEEE Trans. on Circuits and Systems for Video Technology*, to be published. [Online]. Available: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6587741

[11] J. Wright, *et al.*, "Robust face recognition via sparse representation," *IEEE Trans. on Pattern Analysis and Machine Int.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[12] Z. Zhang, H. Zha, and H. Simon, "Low-rank approximations with sparse factors i: Basic algorithms and error analysis," *SIAM Journal on Matrix Analysis and Applications*, vol. 23, no. 3, pp. 706–727, 2002.

[13] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM*, vol. 58, no. 3, pp. 11:1–11:37, Jun. 2011.

[14] X. Mo and V. Monga, "Low rank sparsity prior for robust video anomaly detection," The Pennsylvania State University, Tech. Rep., 2013. [Online]. Available: http://signal.ee.psu.edu/LRSP_CC.pdf

[15] V. Chandrasekaran, S. Sanghavi, P. Parrilo, and A. Willsky, "Rank-sparsity incoherence for matrix decomposition," *SIAM Journal on Optimization*, vol. 21, no. 2, pp. 572–596, 2011.

[16] Z. Liu and L. Vandenberghe, "Interior-point method for nuclear norm approximation with application to system identification," *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 3, pp. 1235–1256, 2010.

[17] A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Fast l1-minimization algorithms and an application in robust face recognition: A review," EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2010-13, Feb 2010. [Online]. Available: http://www.eecs.berkeley.edu/Pubs/TechRpts/2010/EECS-2010-13.html

[18] D. Malioutov, M. Cetin, and A. Willsky, "Homotopy continuation for sparse signal representation," in *Proc. IEEE Int. on Conf. Acoustics, Speech, and Signal Processing*, vol. 5, 2005, pp. 733–736.

[19] "Public dataset of traffic video (pdtv)." [Online]. Available: http://www.tft.lth.se/video/co_operation/data_exchange/

[20] C. Stauffer and W. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Trans. on Pattern Analysis and Machine Int.*, vol. 22, no. 8, pp. 747–757, Aug. 2000.

[21] G. Knott, *Interpolating Cubic Splines*. Progress in Computer Science and Applied Logic, V. 18, 2000.