A NEW SPARSE LOW-RANK MATRIX DECOMPOSITION METHOD AND ITS APPLICATION ON TRAIN PASSENGER ABNORMAL ACTION IDENTIFICATION

Y. Zhang, G. Cai*, J. Sun*, Y. Wang*, J. Chen‡*

Abstract: In the article a new sparse low-rank matrix decomposition model is proposed based on the smoothly clipped absolute deviation (SCAD) penalty. In order to overcome the computational hurdle we generalize the alternating direction method of multipliers (ADMM) algorithm to develop an alternative algorithm to solve the model. The algorithm we designed alternatively renew the sparse matrix and low-rank matrix in terms of the closed form of SCAD penalty. Thus, the algorithm reduces the computational complexity while at the same time to keep the computational accuracy. A series of simulations have been designed to demonstrate the performances of the algorithm with comparing with the Augmented Lagrange Multiplier (ALM) algorithm. Ultimately, we apply the model to an on-board video background modeling problem. According to model the on-board video background, we can separate the video background and passenger's actions. Thus, the model can help us to identify the abnormal action of train passengers. The experiments show the background matrix we estimated is not only sparser, but the computational efficiency is also improved.

Key words: *sparse, low-rank matrix, alternative algorithms, SCAD, abnormal action identification*

Received: July 24, 2013 **DOI:** 10.14311/NNW.2015.25.033
Revised and accepted: September 14, 2015

1. Introduction

It is well known that many countries have improved the video inspector ability of public facilities after the 911 terrorist attack. Many exception events (such as fighting, rubbing) happened sometimes in the trains. When passengers have some dangerous behavior, the video monitoring system identifies the abnormal action and

[*]Ye Zhang – Corresponding author, Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing, China, E-mail: ease2003@163.com

[†]Guoqiang Cai, Jian Sun, Yatao Wang, State Key Lab of Rail Traffic Control & Safety, Beijing Jiaotong University, Beijing, China, E-mail: guoqiangcai@163.com, aidnsun@gmail.com, 13120967@bjtu.edu.cn

[‡]Jianxiao Chen, CSR Times Electric Co., Ltd, Shidai Road, Shifeng District, Zhuzhou, 412001, Hunan, P.R.C, E-mail: lisacls_1@163.com

©CTU FTS 2015

alarm security staff to deal with promptly. How to fast identify the abnormal action of passengers is an urgent and researchable problem for the security of an operating train. Generally speaking, a video involves abnormal action has strong kinetic action and relatively stable background. According to the properties, we model the abnormal action identification problem as a matrix decomposition problem in this paper.

A common approach to deal with engineering problems is to decompose the complex system to several simple systems to study. Especially, the complex system can be decomposed by a sparse matrix and a low-rank matrix, when the complex system is expressed by the summation of a sparse matrix and a low-rank matrix. The sparse low-rank matrix decomposition problem has attracted interest in the research community due to the successful applications to collaborative filtering and video background modeling.

Let us suppose matrix D can be expressed by a summation of a low-rank matrix A and a sparse matrix E. A natural question is how to obtain an accurate estimations of A and E based on D. This problem has been modeled by the l_0 regularization form

$$\begin{aligned} &\min_{A,E\in\mathbb{R}^{m\times n}} \operatorname{rank}(A) + \lambda\|E\|_{l_0}, \\ &\text{s.t. } D = A + E, \end{aligned} \tag{1}$$

where $\operatorname{rank}(A)$ is the rank of A, $\|E\|_{l_0}$ is the number of non-zero entries of E, and $\lambda > 0$ is a regularization parameter. Obviously, sparse low-rank matrix decomposition problem is an ill-posed problem. However, with the development of compressive sensing and sparse modeling, [5, 6, 8, 2] found the solvable conditions can be obtained based on rank-sparsity incoherence property. In addition, [4, 10, 13, 14, 15] introduced the principle component pursuit to handle this problem. The formulations (l_1 regularization form) are as follows:

$$\begin{aligned} &\min_{A,E\in\mathbb{R}^{m\times n}} \|A\|_* + \lambda\|E\|_{l_1}, \\ &\text{s.t. } D = A + E, \end{aligned} \tag{2}$$

where $\|\|_*$ is the kernel norm (the summation of all the singular values), and $\|\|_{l_1}$ is the summation of the absolute value of all entries. Moreover, matrix D can not be exactly expressed by the summation of E and A due to the noisy effect in the real system. Hence, [16] proposed the robust principle component pursuit based on the constrain$\|D - A - E\|_F \le \delta$. That is

$$\begin{aligned} &\min_{A,E\in\mathbb{R}^{m\times n}} \|A\|_* + \lambda\|E\|_{l_1}, \\ &\text{s.t. } |D - A - E\|_F \le \delta. \end{aligned} \tag{3}$$

[15] has introduced an iterative thresholding algorithm to solve (2) and (3). However, the convergence speed is slow due to the singular value decomposition procedure at each iterative step. So, it cannot solve large-scale real application problems. Furthermore, [12] proposed APG algorithm to solve the original problem and a kind of gradient ascent algorithm to solve the dual problem which are justified that they can be used to deal with 1000×1000 matrix. These two algorithms enhance convergence speed 50 times than the thresholding algorithm. Moreover, [11] introduced an

approximate ALM algorithm which cannot only improve the convergence speed but also keep the computational accuracy. It becomes the best well-known algorithm for large-scale problem (sparse low-rank matrix decomposition).

Recently, developed SCAD regularization theory indicates the SCAD penalty can obtain more sparse and robust (stable) solution than l_1 [1, 9]. So, we would like to share the same light with SCAD and apply SCAD penalty to design a new model to handle sparse low rank matrix decomposition problem. To obtain a better estimation of D, we utilize the SCAD regularization to replace l_1 regularization in this article, and generalize SCAD regularization to describe from 1-dimensional vector to 2-dimensional matrix. The basic formulation we propose as follow:

$$\begin{aligned} &\min_{A,E\in\mathbb{R}^{m\times n}} \|A\|_{\mathrm{SCAD}} + \lambda\|E\|_{l_a}^a, \\ &\text{s.t. } |D - A - E\|_F \le \delta, \end{aligned} \tag{4}$$

where $\|A\|_{\mathrm{SCAD}} = P_\lambda(\sigma)$, $\sigma = (\sigma_1, ..., \sigma_r)$ are the singular values of A, P_λ function satisfies $P'_\lambda(x) = \lambda\{I(x \le \lambda) + \frac{(b\lambda-\theta)_+}{(b-1)\lambda} I(x > \lambda)\}$, $b > 2$ is a constant, and $\|E\|_{l_a}^a = (\sum_{i=1}^m \sum_{j=1}^n |E_{ij}|^a)^{1/a}$.

The rest of this paper is organized as follows. In Section 2, we introduce an alternative thresholding algorithm based on the ADMM framework [3, 7]. The algorithm utilizes the alternative projection idea to renew low rank matrix and sparse matrix simultaneously. We design a series of simulations to demonstrate the performance of the algorithm with comparing with some baseline algorithms in Section 3. The Section 4 describes how to use the model we proposed to model on-board video background. A real case is studied from an inspector video of subway, which can help staffs identify the abnormal actions of passengers.

2. Iterative thresholding algorithm based on ADMM

2.1 ADMM framework

ADMM is a kind of efficient algorithm framework to handle the distributed convex optimization. It uses the decomposition coordination procedure to construct a decomposed gradient ascent and extended Lagrange multiplier method simultaneously. The basic framework is as follows:

$$\begin{aligned} &\min_{x,z} \;\; f(x) + g(z), \\ &\text{s.t. } Mx + Nz = c, \end{aligned} \tag{5}$$

where $x \in \mathbb{R}^n$, $z \in \mathbb{R}^m$, $M \in \mathbb{R}^{p\times n}$, $N \in \mathbb{R}^{p\times m}$, and f, g are convex functions. Let us consider the extended Lagrange formulation

$$L_\mu(x, z, y) = f(x) + g(z) + y^\top(Mx + Nz - c) + \frac{1}{2}\mu\|Mx + Nz - c\|_2^2, \tag{6}$$

where y is the multiplier of linear equation constrains, and μ is the multiplier of inequality constrains, that is extended Lagrange parameter. The three main steps of ADMM are constructed by

$$\begin{cases} x^{k+1} = \arg\min_x L_\mu(x, z^k, y^k), \\ z^{k+1} = \arg\min_z L_\mu(x^{k+1}, z, y^k), \\ y^{k+1} = y^k + \mu(Mx^{k+1} + Nz^{k+1} - c). \end{cases} \tag{7}$$

Obviously, it includes minimizing x, minimizing z and renewing dual variable y three steps. The convergence theory of ADMM is developed in terms of the convexity of f and g. However, the basic steps can be generalized to nonconvex situations. As we known, the SCAD function is nonconvex function, so we consider a new extended Lagrange multiplier formulation

$$L_\mu(A, E, Y) = \|A\|_{\mathrm{SCAD}} + \lambda\|E\|_{l_a}^a + \langle Y, D - A - E\rangle + \frac{1}{2}\mu\|D - A - E\|_F^2, \tag{8}$$

where μ is a parameter which depends on δ. To minimize $L_\mu(A, E_k, Y_k)$, we iteratively solve three sub problems

$$\begin{cases} A_{k+1} = \arg\min_A \quad L_\mu(A, E_k, Y_k), \\ E_{k+1} = \arg\min_E \quad L_\mu(A_{k+1}, E, Y_k), \\ Y_{k+1} = Y_k + \mu(D - A_{k+1} - E_{k+1}). \end{cases} \tag{9}$$

To simplify the formulation of the three main steps, we reform them as

$$\begin{cases} A_{k+1} = \arg\min_A \quad \frac{1}{2}\mu\|A - (D - E_k + \frac{1}{\mu}Y_k)\|_F^2 + \|A\|_{\mathrm{SCAD}}, \\ E_{k+1} = \arg\min_E \frac{\mu}{2\lambda}\|E - (D - A_{k+1} + \frac{1}{\mu}Y_k)\|_F^2 + \|E\|_{l_a}^a, \\ Y_{k+1} = Y_k + \mu(D - A_{k+1} - E_{k+1}). \end{cases} \tag{10}$$

The crucial point of the above alternative thresholding algorithm is renewing the low rank matrix, sparse matrix and Lagrange multiplier simultaneously until convergence condition holds.

In order to renew the spare and low-rank matrix, we have to solve two subproblems in Eq. (10):

$$\min_X \|X - W\|_F^2 + \lambda\|X\|_{l_a}^a, \tag{11}$$

$$\min_X \|X - W\|_F^2 + \lambda\|X\|_{\mathrm{SCAD}}. \tag{12}$$

Subproblem in Eq. (11) is relatively simple. When $a = 1$, the optimal solution $ST_{\lambda/2}(W)$ of Eq. (11) is the soft thresholding function [8]. Namely, shrinking all entries of E based on $\lambda/2$ and soft thresholding function $ST_{\lambda/2}$, where

$$ST_\lambda(x) = \begin{cases} x - \mathrm{sgn}(x)\lambda/2 & \text{if } |x| > \lambda/2, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Solving subproblem in Eq. (12) is relatively difficult, so we show the next theorem to give some insights to figure it out.

Theorem 1. *Suppose the rank of matrix* $W \in \mathbb{R}^{m\times n}(m \geq n)$ *is* $r \in [1, \min(m,n)]$. *The* $W = U_{m\times r}D_{r\times r}V^T_{r\times n}$ *is the singular value decomposition, where* $D = diag(\sigma_1.\ \sigma_2, ..., \sigma_r)$, $U = (u_1, u_2, ..., u_r)$, $V = (v_1, v_2, ..., v_r), u_i \in \mathbb{R}^m, v_i \in \mathbb{R}^n$. *Then the closed form of optimal solution of (12) is* $X^* = US_\lambda(D)V^\top$, *where* $S_\lambda(D) = (S_\lambda(\sigma_1), ..., S_\lambda(\sigma_r))$ *and*

$$S_\lambda(x) = \begin{cases} \operatorname{sgn}(x)(|x| - \lambda) & \textit{if } |x| \leq 2\lambda, \\ (a-1)x - \operatorname{sgn}(x)a\lambda/(a-2) & \textit{if } 2\lambda < |x| \leq a\lambda, \\ x & \textit{if } |x| > a\lambda. \end{cases} \tag{14}$$

The singular value decomposition of X *is*

$$X = U'D'V'^\top,$$

where $D' = \operatorname{diag}(\sigma'_1, ..., \sigma'_n)$, $U' = [u'_1, ..., u'_n]$, $V' = [v'_1, ..., v'_n]$.

Proof. Based on Eq. (11), we have

$$\|W - X\|_F^2 + \lambda\|X\|_{\text{SCAD}} = \|W\|_F^2 - 2\sum_{i=1}^n \sigma'_i u_i'^\top W v'_i + \sum_{i=1}^n \sigma_i'^2 + \sum_{i=1}^n P_\lambda(\sigma'_i). \tag{15}$$

Let $Q(U', V') = \min_{D'\geq 0} -2\sum_{i=1}^n \sigma'_i u_i'^\top W v'_i + \sum_{i=1}^n \sigma_i'^2 + P_\lambda(\sigma'_i)$. So Eq. (11) is equivalent to

$$\begin{aligned} &\min_{U',V'} \ Q(U', V'), \\ &\text{s.t. } U'^\top U' = I_n, \\ &\quad V'^\top V' = I_n. \end{aligned} \tag{16}$$

Note $t_i = u_i'^\top W v'_i$ and $f_\lambda(\sigma'_i) = -2\sigma'_i t_i + \sigma_i'^2 + P_\lambda(\sigma_i)$, then

$$Q(U', V') = \min_{D'\geq 0} \sum_{i=1} (\sigma'_i). \tag{17}$$

So Eq. (17) is equivalent to

$$\sigma_i^* = \arg\min_{\sigma'_i \geq 0} f_\lambda(\sigma'_i). \tag{18}$$

In terms of [9], the derivative of the objective function is

$$-2t_i + 2\sigma_i^* + P'_\lambda(\sigma_i) = 0, \tag{19}$$

thus the solution satisfy the $\sigma_i = S_\lambda(t_i)$. So, we can complete our proof to get $X^* = US_\lambda(D)V^\top$. □

2.2 Alternative thresholding algorithm

We will implement an alternative thresholding algorithm to obtain the closed form of subproblems Eqs. (11) and (12).

The sketch of the algorithm is in the sequence:

- Initializing $\{Y_0, E_0, A_0\}$, and choosing parameter μ and error precision ε and $k = 0$;
- Renewing low rank matrix A_{K+1}: $D - E_k + \mu^{-1} Y_k \approx U_r D_r V_r^\top$, $A_{k+1} = U_r S_\lambda(D_r) V_r^\top$;
- Renewing sparse matrix E_{k+1}:

$$E_{k+1} = \begin{cases} S_\lambda(D - A_{k+1} + \mu^{-1} Y_k), \\ ST_{\lambda/\mu}(D - A_{k+1} + \mu^{-1} Y_k), \quad a = 1; \end{cases} \tag{20}$$

- Renewing Lagrange multiplier: $Y_{k+1} = Y_k + \mu(D - A_{k+1} - Y_{k+1})$;
- If $\dfrac{\|D - A_{k+1} - E_{k+1}\|_F}{\|D\|_F} < \varepsilon$, then stop the algorithm, otherwise let $k := k + 1$;

It is known that the convergence speed of ADMM depends on parameter μ. To avoid the dependency of μ and enhance the convergence speed, we adaptively used parameter as $\mu_{k+1} = \rho\mu_k, \rho \leq 1$. Next section, we will design some simulations to demonstrate the performance of the algorithm.

3. Simulations

In this section, we compare performance of three algorithms ASS (Adaptive SCAD-Soft Alternative Thresholding Algorithm), ISS (Inadaptive SCAD-Soft Alternative Thresholding Algorithm) and ALM in the context of matrices with noise and without noise. We suppose sr to be a measurement of sparsity (proportion of non-zero elements) and lr to ba a measurement of low rank (the ratio of rank r and dimension m) of sparse low rank matrix D. The simulated matrix D is generated by

- Generate random matrix L and R both in $\mathbb{R}^{m\times r}$ independently, and all the entries of L and R satisfy standard normal distribution independently. Then we take low rank matrix $A = \frac{1}{\sqrt{r}} LR^\top$ which can guarantee the variance of each elements is 1.
- Generate $sr \times m^2$ independent random numbers from $U[0, 1]$ and choose $sr \times m^2$ positions from sparse matrix $E \in \mathbb{R}^{m\times m}$. Then put the random numbers into the positions.
- Generate a random noise matrix $N \in \mathbb{R}^{m\times m}$ and $N_{ij} \sim \mathcal{N}(0, \sigma)$, i.i.d. If $\sigma > 0$, we say $D = A + E + N$ is the noise matrix.

3.1 Simulations without noise

The matrix D can be exactly decomposed as a sparse matrix and a low rank matrix when D does not effect by noise. In this simulation we consider different size matrices such as $m = 500 : 500 : 4000$ and set $lr = 0.01$ and $sr = 0.05$. We repeat each size simulation 20 times and make use of the average of $\mathrm{err}A^*$, $\mathrm{err}E^*$, rank(A^*)/rank(A) and $\|E\|_{l_0}^*/\|E\|_{l_0}$ to compare the performance of the algorithms. The definitions are

$$\mathrm{err}A^* = \frac{\|A - A^*\|_F}{\|A\|_F}, \qquad \mathrm{err}E^* = \frac{\|E - E^*\|_F}{\|E\|_F}. \tag{21}$$

m		500	1000	1500	2000	2500	3000	3500	4000
rank(A*)	ALM	5.8	11	16	21	26	31	36	41
	ASS	**5**	**10**	**15**	**20**	**25**	**30**	**35**	**40**
	ISS	5	10	15	20	25	30	35	40
iterations	ALM	27.5	27.6	28	27.9	28	28	28	28
	ASS	**7**	**7**	**7**	**7**	**7**	**7**	**7**	**7**
	ISS	27.4	27.1	27.3	26.9	26.7	26.7	26.3	26.5
time [s]	ALM	5.42	28.17	72.12	135.63	239.73	375.79	539.10	754.44
	ASS	**1.82**	**7.86**	**19.30**	**42.03**	**77.35**	**121.91**	**205.56**	**288.60**
	ISS	4.76	18.90	42.92	88.71	162.19	254.38	434.71	603.38

Tab. I *Results of simulations without noise.*

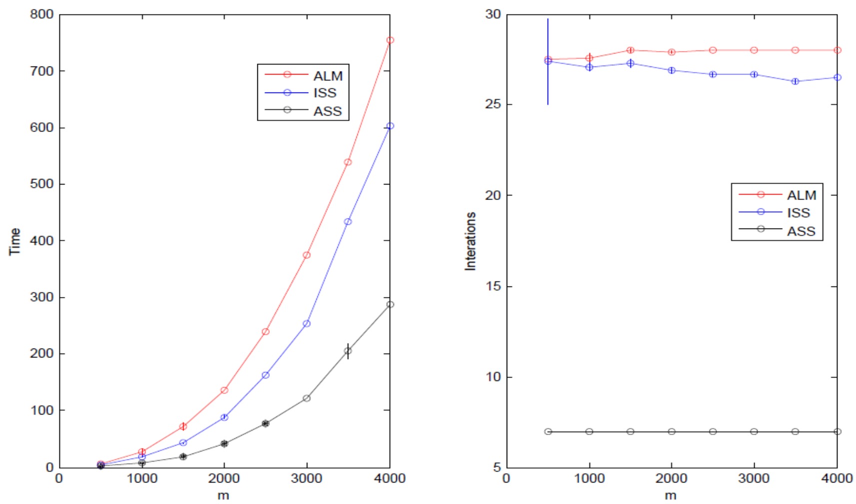

Fig. 1 *Comparison of computational complexity ($lr = 0.01$, $sr = 0.05$).*

Based on the Fig. 1, 2 and Tab. I we can find several facts. First, the relative errors of estimated sparse matrix and low rank matrix in terms of three algorithms are all very small. The estimated rank of ALM is relatively bigger while ISS and ASS can get the corrected rank. Second, from computational speed perspective we can see the computational time of ALM and ISS increase dramatically and iterative number stays at around 27. However, computational time of ASS algorithm increase much slower than others and iterative number stays at 7.

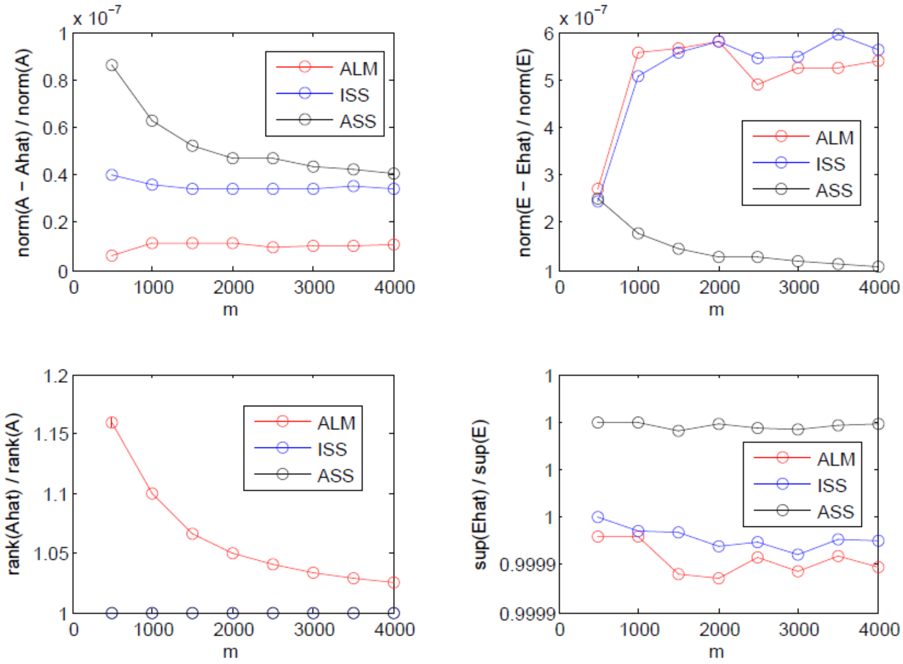

Fig. 2 *Comparison of computational accuracy ($lr = 0.01$, $sr = 0.05$).*

3.2 Simulation with noise

In this subsection we will compare the performance of the three algorithms to apply to the noisy sparse low rank matrix. The same to the simulation 1, the different size matrix D ($m = 500 : 500 : 4000$) can be considered. We set $lr = 0.01$, $sr = 0.05$ and $\sigma = 0.1 : 0.1 : 1$. For each group parameters, the simulations have repeated 10 times and recorded the average of $\mathrm{err}A^*$, $\mathrm{err}E^*$, $\mathrm{rank}(A^*)/\mathrm{rank}(A)$ and $\|E\|^*_{l_0}/\|E\|_{l_0}$. Based on the results we find that the relative error of ALM is about 3 times of ASS and ISS, and the estimated rank is much higher than real rank. Comparing with the iterative numbers to achieve convergence, ALM is about 6 times of ASS and ISS. The convergence time of ALM increases dramatically as number of dimensionality arising. However, ASS and ISS increase slowly.

4. Application to train passenger abnormal action identification based on inspector video of subway

Background modeling and abnormal action identification of videos all can be naturally modeled by sparse low-rank matrix decomposition problems [4]. We suppose each frame of a video is according to each column of D, then the background which involves each frame have very strong similarity. Thus, it can be constructed by low-rank matrix A. The moving objects and background changes can be modeled by the sparse matrix and noise matrix $E + N$. In this section, an inspector video from Chinese subway system is used to demonstrate the performance of the algorithm. The video includes the normal actions (stable background and working passengers), and abnormal actions (fighting, running). The Fig. 3 shows the decomposed performance of ASS algorithm. The next table shows the comparison results of all the algorithms.

Measurements	ALM	ASS	ISS
rank(A*)	134	6	5
iterations	37	9	8
times	108.7414	17.7623	33.2454

Tab. II *Results of chinese subway data.*

From the Tab. II and Fig. 4 we can find that the ASS algorithm is the most efficient algorithm for this data set. The computational complexity (times) improved about 10 times comparing with ALM algorithm. And the iterations is also less than the ALM algorithm. The ASS algorithm was used in subway train passenger

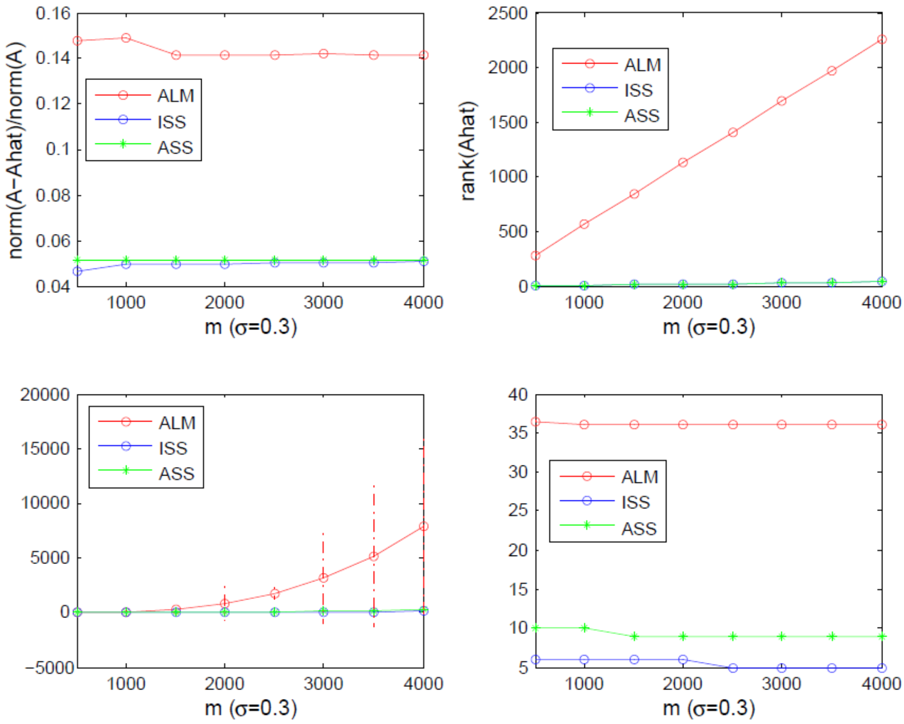

Fig. 3 *Comparison of computational accuracy and complexity* ($lr = 0.01$, $sr = 0.05$, $\sigma = 0.3$).

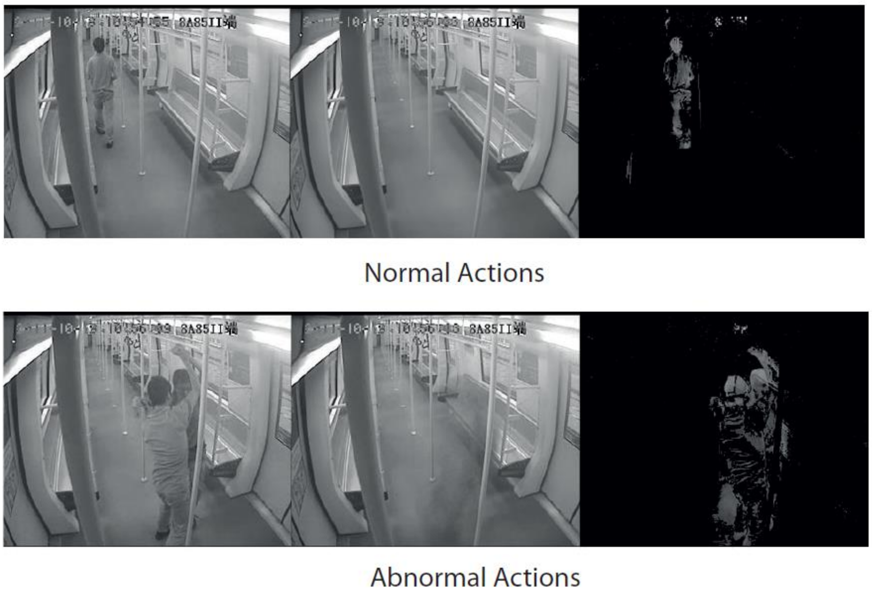

Fig. 4 *Identification of abnormal actions.*

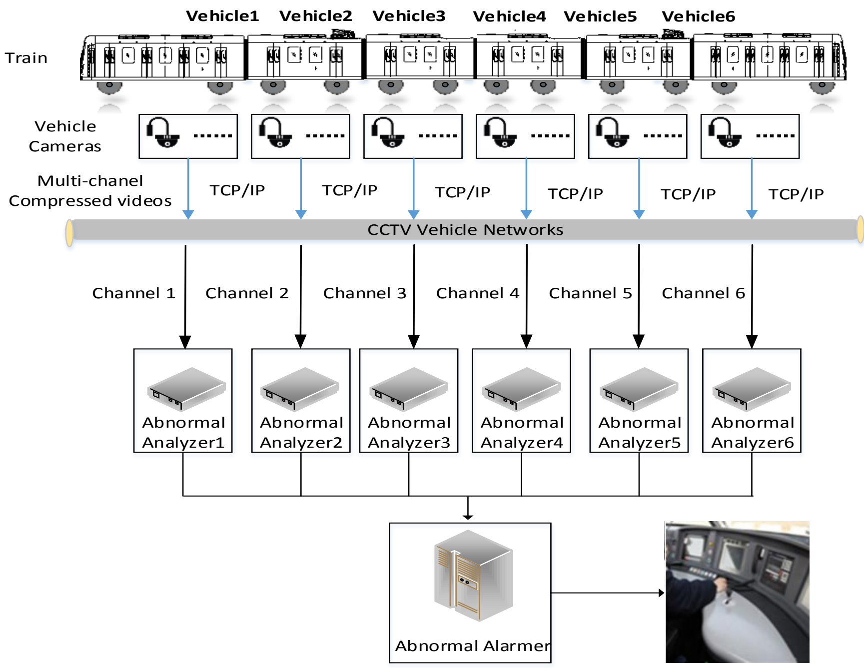

Fig. 5 *Subway train passenger abnormal action identification system.*

abnormal action analyze. The application system is shown in Fig. 5. Real-time vehicle monitoring cameras are connected by a CCTV Vehicle Networks. Their outputs are sent to several Abnormal Analyzers to identify abnormal or dangerous behavior by means of the ASS algorithm. Abnormal Analyzers transfer alarming information to Abnormal Alarm. Eventually, Abnormal Alarm prompts emergent alarm in display terminal for a train driver to deal with the event quickly.

5. Conclusions

In this article, we have introduced a new sparse low-rank matrix decomposition model based on SCAD penalty. With sharing the light of the SCAD penalty (the closed form), we designed an efficient iterative algorithm to solve the model. The performances of the algorithm have been demonstrated by the simulations. Ultimately, we generalize this model directly to abnormal action identification problem based an on on-board inspector video data set. The results showed ASS algorithm has the best potential capacity to fast analyze the massive video data.

Acknowledgement

This paper was supported by the China High Technology Project (2015BAG19B02) (KIK15007531), the Beijing Postdoc Program (2014ZZ-64).

References

[1] ANTONIADIS A., FAN J. Regularization of Wavelet Approximations. *Journal of the American statistical association*. 2001, 96(455), pp. 939/-955, doi: 10.1198/016214501753208942.

[2] BOUWMANS T., ZAHZAH El H. Robust PCA via Principal Component Pursuit: A Review for a Comparative Evaluation in Video Surveillance. *Computer Vision and Image Understanding*. 2014, 122, pp. 22–34, doi: 10.1016/j.cviu.2013.11.009.

[3] BOYD S., PARIKH N., CHU E., PELEATO B., ECKSTEIN J. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. *Foundations and Trends in Machine Learning*. 2010, 3(1), pp. 1–122, doi: 10.1561/2200000016.

[4] CANDES E.J., LI X.D., MA Y., WRIGHT J. Robust Principle Component Analysis? *Journal of the ACM*. 2011, 58(3), doi: 10.1145/1970392.1970395.

[5] CANDES E.J., ROMBERG J., TAO T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Trans on Information Theory*. 2006, 52(2), pp. 489–509, doi: 10.1109/TIT.2005.862083.

[6] CHANDRASEKARAN V., SANGHAVI S., PARRILO P.A., WILLSKY A.S. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*. 2011, 21(2), pp.572-596, doi: 10.1137/090761793.

[7] DAUBECHIES I., DEFRISE M., DE MOL C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications on Pure and Applied Mathematics*. 2004, 57(11), pp. 1413–1457, doi: 10.1002/cpa.20042.

[8] DONOHO D.L. Compressed sensing. *IEEE Trans on Information Theory*. 2006, 52(4), pp. 1289–1306, doi: 10.1109/TIT.2006.871582.

[9] FAN J., LI R. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*. 2011, 94(465), pp. 1348–1360.

[10] JAVED S., SOBRAL A., OH S., BOUWMANS T., JUN S. OR-PCA with MRF for Robust Foreground Detection in Highly dynamic backgrounds. In: *Computer Vision - ACCV* 2014. 12*th Asian Conference on Computer Vision*, Lecture Notes in Computer Science, Singapore. Springer, 2015, pp. 289–299, doi: 10.1007/978-3-319-16811-1_19.

[11] LIN Z.C., CHEN M.M., MA Y. *The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-rank Matrices* [online]. UTUC Technical Report UILU-ENG-09-2214. 2010 [veiewed 2015-12-01]. Available from: http://www.cis.pku.edu.cn/faculty/vision/zlin/Publications/ALM-v5.pdf

[12] Lin Z.C., Ganesh A., Wright J., Wu L., Chen M., Ma Y. *Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix* [online]. UIUC Techinical Report UILU-ENG-09-2214. 2009 [veiewed 2015-12-01]. Available from: http://perception.csl.illinois.edu/matrix-rank/Files/rpca_algorithms.pdf

[13] TANG G., NEHORAI A. Robust principal component analysis based on low-rank and block-sparse matrix decomposition. In: 45*th Annual Conference on Information Sciences and Systems (CISS 2011)*, Baltimore, MD. IEEE, 2011, pp. 1–5, doi: 10.1109/CISS.2011.5766144.

[14] WOHLBERG B., CHARTRAND R., THEILER J. Local Principal Component Pursuit for Nonlinear Datasets. In: *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP* 2012*)*, Kyoto, Japan. IEEE, 2012, pp. 3925–3928, doi: 10.1109/ICASSP.2012.6288776.

[15] Wright J., Ganesh A., Rao S., Peng Y., Ma Y. Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices via Convex Optimization. In: Y. BENGIO, D. SCHUURMANS, J.D. LAFFERTY, C.K.I. WILLIAMS, A. CULOTTA, eds. *Proceedings of the Advances in Neural Information Processing Systems 22 (NIPS 2009)*. Curran Associates, Inc., 2009, pp. 2080–2088, http://papers.nips.cc/paper/3704-robust-principal-component-analysis-exact-recovery-of-corrupted-low-rank-matrices-via-convex-optimization

[16] ZHOU Z., LI X., WRIGHT J., CANDES E.J., MA Y. Stable principal component pursuit. In: *Proceedings of the IEEE International Symposium on Information Theory (ISIT* 2010*)*, Austin, TX, USA. IEEE, 2010, pp. 1518–1522, doi: 10.1109/ISIT.2010.5513535.