Background Recovery by Fixed-rank Robust Principal Component Analysis

Wee Kheng Leow, Yuan Cheng, Li Zhang, Terence Sim, and Lewis Foo

Department of Computer Science, National University of Singapore
Computing 1, 13 Computing Drive, Singapore 117417
`leowwk cyuan zhangli tsim lewis@comp.nus.edu.sg`

Abstract. Background recovery is a very important theme in computer vision applications. Recent research shows that robust principal component analysis (RPCA) is a promising approach for solving problems such as noise removal, video background modeling, and removal of shadows and specularity. RPCA utilizes the fact that the background is common in multiple views of a scene, and attempts to decompose the data matrix constructed from input images into a low-rank matrix and a sparse matrix. This is possible if the sparse matrix is sufficiently sparse, which may not be true in computer vision applications. Moreover, algorithmic parameters need to be fine tuned to yield accurate results. This paper proposes a fixed-rank RPCA algorithm for solving background recovering problems whose low-rank matrices have known ranks. Comprehensive tests show that, by fixing the rank of the low-rank matrix to a known value, the fixed-rank algorithm produces more reliable and accurate results than existing low-rank RPCA algorithm.

Keywords: Background recovery, reflection removal, robust PCA

1 Introduction

Background recovery is a very important recurring theme in computer vision applications. Traditionally, different approaches have been developed to solve different varieties of the problem. Recent research in **robust principal component analysis** (RPCA) offers a promising alternative approach for solving problems such as noise removal, video background modeling, and removal of shadows and specularity [2, 12]. RPCA utilizes the fact that multiple views of a scene contain consistent information about the common background. It constructs a **data matrix** from multiple views and decomposes it into a **low-rank matrix** that contains the background and a **sparse matrix** that captures non-background components. It has been proved that exact solution of RPCA problem is available if the data matrix is composed of a sufficiently low-rank matrix and a sufficiently sparse matrix [2, 3, 9, 13]. Various algorithms have been proposed for solving RPCA problem [6, 9, 12]. In particular, the methods based on **augmented Lagrange multiplier** (ALM) have been shown to be among the most efficient and accurate methods [9].

In computer vision applications, the non-background components may not be sparse. Moreover, algorithmic parameters need to be fine tuned to yield accurate results [1]. These difficulties are especially pronounced for reflection removal problem, and no work on applying RPCA to reflection removal has been reported so far. Fortunately, these application problems can be framed as one of recovering a **fixed-rank** matrix from the data matrix because the rank of the low-rank matrix is known. This paper proposes a **fixed-rank** RPCA algorithm based on ALM (FrALM) for solving background recovering problems. Comprehensive tests on reflection removal and video background modeling show that FrALM produces more accurate results than does low-rank ALM method (LrALM). Moreover, FrALM can produce optimal or near optimal results over a much wider range of parameter values than does LrALM, making it more reliable for solving computer vision problems whose low-rank matrices have known ranks.

2 Existing RPCA Methods

Robust PCA is a term given to a long line of work that aims to render PCA robust to gross corruption and outliers. Various methods have been proposed including influence function [4], multivariate trimming [7], alternating minimization [8], and random sampling [5]. These methods are either inefficient, having non-polynomial time complexity, or do not guarantee optimal solutions [12].

A recent approach directly decomposes a corrupted data matrix into a low-rank matrix and a sparse matrix. The corruption is assumed to be sparse, but the noise amplitude can be large. Various methods have been proposed such as iterative thresholding [9], proximal gradient [12], accelerated proximal gradient [6], and augmented Lagrange multiplier method (ALM) [9]. In particular, ALM has been shown to be among the most efficient and accurate methods [9]. These methods require tuning of algorithmic parameters [1]. On the other hand, [1] applies Bayesian approach to estimate the algorithmic parameters along with the matrices based on prior distributions of inverse variances.

In our applications, the rank of the low-rank matrix is known. So, we adopt the ALM approach but fix the rank of the low-rank matrix, which provides more specific constraint than do prior distributions. This approach allows our algorithm to converge efficiently and accurately, as for the low-rank ALM method of [9], and is simpler and more efficient than the Bayesian method of [1].

Other methods have been proposed to solve related but different problems. For example, [11] solves low-rank matrix factorization and [10] computes a fixed-rank representation for sparse subspace clustering. They are not directly applicable to our application problem, which is a matrix decomposition problem.

3 Fixed-rank RPCA

Given an $m \times n$ data matrix $\mathbf{D}$, PCA seeks to recover a low-rank matrix $\mathbf{A}$ from data matrix $\mathbf{D}$ such that the discrepancy or error $\mathbf{E} = \mathbf{D} - \mathbf{A}$ is minimized:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{E}\|_F, \text{ subject to } \operatorname{rank}(\mathbf{A}) \leq r, \ \mathbf{D} = \mathbf{A} + \mathbf{E} \tag{1}$$

where $r \ll \min(m, n)$ is the target rank of $\mathbf{A}$ and $\|\cdot\|_F$ is the Frobenius norm. Eq. 1 can be solved by SVD but the solution will be vastly inaccurate if the error entries in $\mathbf{E}$ are arbitrarily large. Under the conditions that $\mathbf{A}$ is low-rank and $\mathbf{E}$ is sufficiently sparse, Wright et al. [12] show that $\mathbf{A}$ can be exactly recovered by solving the following convex optimization problem:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1, \text{ subject to } \mathbf{D} = \mathbf{A} + \mathbf{E} \qquad (2)$$

where $\|\cdot\|_*$ denotes the nuclear norm and $\|\cdot\|_1$ denotes the 1-norm,

Lin et al. [9] reformulate Eq. 2 using augmented Lagrange multiplier method. Their method (LrALM) uses a matrix $\mathbf{Y}$ and parameter μ to merge the constraint into the objective function, leading to the following revised problem:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1 + \langle \mathbf{Y}, \mathbf{D} - \mathbf{A} - \mathbf{E} \rangle + \frac{\mu}{2} \|\mathbf{D} - \mathbf{A} - \mathbf{E}\|_F^2 \qquad (3)$$

where $\langle \mathbf{U}, \mathbf{V} \rangle$ is the sum of the product of corresponding elements in $\mathbf{U}$ and $\mathbf{V}$, and λ and μ are parameters that need to be specified. An iterative algorithm is applied to determine the $\mathbf{A}$ and $\mathbf{E}$ that minimize Eq. 3.

For reflection removal, a set of reflection images are arranged as column matrices in $\mathbf{D}$. If the images are well aligned such that the transmitted parts are identical, then $\mathbf{A}$ captures the transmitted parts and has a rank of 1. If the reflection is localized, $\mathbf{E}$ is sparse; otherwise, $\mathbf{E}$ is not sparse. Similar characteristics are observed in background modeling of video taken with a stationary camera.

When $\mathbf{E}$ is not sparse, LrALM may not recover accurate results unless the parameter λ is carefully chosen (Section 4). If λ is too large, the trivial solution of $\mathbf{E} = \mathbf{0}$ is obtained, and $\mathbf{A} = \mathbf{D}$, which has a rank larger than the desired low rank. On the other hand, if λ is too small, $\mathbf{E} = \mathbf{D}$ and $\mathbf{A} = \mathbf{0}$, which has a rank of 0. So, the value of λ directly influences the rank of $\mathbf{A}$ recovered by LrALM. Although Zhou et al. [13] prove theoretically that the optimal λ can be set to $1/\sqrt{\max(m,n)}$, this is true only if $\mathbf{A}$ is low-rank and $\mathbf{E}$ is sufficiently sparse. The parameter μ can also affect the accuracy of the recovered $\mathbf{A}$ by influencing the rank of $\mathbf{A}$ (see discussion below).

To overcome the above difficulties, we frame the background recovery problem as one of recovering a low-rank matrix $\mathbf{A}$ with a known rank r:

$$\min_{\mathbf{A},\mathbf{E}} \|\mathbf{E}\|_F, \text{ subject to } \text{rank}(\mathbf{A}) = \text{known } r, \ \mathbf{D} = \mathbf{A} + \mathbf{E}. \qquad (4)$$

To solve Eq. 4 robustly, we reformulate it in the same manner as the ALM approach (Eq. 3), with the additional constraint of $\text{rank}(\mathbf{A}) = r$. With $\mathbf{A}$'s rank fixed, it may seem that the term $\|\mathbf{A}\|_*$ in Eq. 3 is redundant. Nevertheless, we choose to keep $\|\mathbf{A}\|_*$ in Eq. 3 and solve for $\mathbf{A}$ using ALM approach so that the convergence and optimality properties proved by Lin et al. [9] are preserved.

Our algorithm (FrALM) adopts the exact ALM approach to solve fixed-rank RPCA problem. It is similar to the low-rank ALM algorithm (LrALM) proposed by Lin et al. [9], except that FrALM fixes the rank of $\mathbf{A}$.

FrALM

Input: D, r, λ

1. $\mathbf{A} = \mathbf{0}$, $\mathbf{E} = \mathbf{0}$.
2. $\mathbf{Y} = \operatorname{sgn}(\mathbf{D})/J(\operatorname{sgn}(\mathbf{D}))$, $\mu > 0$, $\rho > 1$.
3. Repeat until convergence:
4.   Repeat until convergence:
5. $\mathbf{U}, \mathbf{S}, \mathbf{V} = \operatorname{svd}(\mathbf{D} - \mathbf{E} + \mathbf{Y}/\mu)$.
6.   If $\operatorname{rank}(T_{1/\mu}(\mathbf{S})) < r$, $\mathbf{A} = \mathbf{U}\, T_{1/\mu}(\mathbf{S})\mathbf{V}^\top$; otherwise, $\mathbf{A} = \mathbf{U}\, \mathbf{S}_r \mathbf{V}^\top$.
7. $\mathbf{E} = T_{\lambda/\mu}(\mathbf{D} - \mathbf{A} + \mathbf{Y}/\mu)$.
8. $\mathbf{Y} = \mathbf{Y} + \mu(\mathbf{D} - \mathbf{A} - \mathbf{E})$, $\mu = \rho\mu$.

Output: A, **E**.

In line 2, $\operatorname{sgn}(\cdot)$ computes the sign of each matrix element, and $J(\cdot)$ computes a scaling factor

$$J(\mathbf{X}) = \max\left(\|\mathbf{X}\|_2, \lambda^{-1}\|\mathbf{X}\|_\infty\right) \tag{5}$$

as recommended in [9]. The function T_ϵ in line 7 is a soft thresholding function:

$$T_\epsilon(x) = \begin{cases} x - \epsilon, & \text{if } x > \epsilon, \\ x + \epsilon, & \text{if } x < -\epsilon, \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

The main difference between FrALM and LrALM lies in Step 6. $\mathbf{S}_r$ is the diagonal matrix of singular values whose diagonal elements above r are set to 0. FrALM fixes $\mathbf{A}$'s rank to the desired rank r if a rank-r matrix is recovered. Otherwise, it behaves in the same manner as LrALM. On the other hand, LrALM allows $\mathbf{A}$'s rank to increase beyond r if μ is too large.

FrALM is algorithmically equivalent to LrALM with a sufficiently small μ that restricts the rank of $\mathbf{A}$ to r. Therefore, the convergence proof of LrALM given in [9] applies to FrALM. Consequently, FrALM can converge as efficiently as LrALM does (Fig. 2(a)). The advantage of FrALM over LrALM is that the user does not have to specify the exact μ that fixes the rank of $\mathbf{A}$ to r.

4   Experiments and Discussions

7 test sets were used to evaluate the performance of FrALM and LrALM on the tasks of reflection removal and video background modeling (Fig. 3, 4). Sets 1 to 3 contained synthetically generated reflection images, and Sets 4 and 5 contained real reflection images. Sets 1 and 4 were corrupted by local reflections whereas Sets 2, 3, and 5 were corrupted by global reflections. Sets 6 and 7 contained video frames of a single moving human and busy traffic junction, respectively.

All the test images were color images of size 200×150 pixels; so $m = 200 \times 150 \times 3 = 90000$. The number of images n in Sets 1 to 7 were respectively, 38, 38, 38, 31, 46, 210, and 250 respectively. Ground truth background images were available for Sets 1 to 6 but not for Set 7. The images were captured with a stationary camera. So the desired rank of the low-rank matrix is 1.
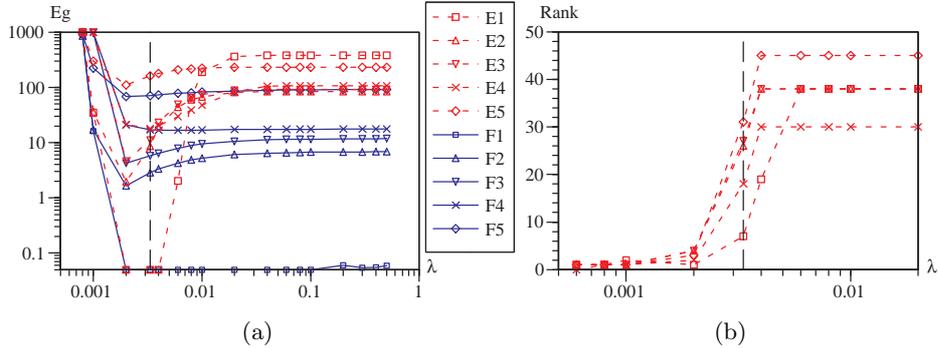
Fig. 1. Performance comparison for reflection removal. (a) Error E_g vs. λ. Error curves above $E_g = 1000$ for very small λ are cropped to reduce clutter. (b) Rank of low-rank matrix recovered by LrALM. (Red dashed lines) LrALM results. Vertical dash lines denote theoretical optimal $\lambda^* = 0.00\dot{3}$. (Blue solid lines) FrALM results.

LrALM and FrALM were tested on the test sets over a range of λ from 0.0001 to 0.5, including the theoretical optimal λ of $\sqrt{m} = 0.00\dot{3}$, denoted as λ^*, as proved in [13]. The parameters ρ and initial μ were set to the default values of 6 and $0.5/\sigma_1$, where σ_1 is the largest singular value of the initial $\mathbf{Y}$, as for LrALM. The algorithms' accuracy was measured in terms of the mean squared error E_g between the ground truth $\mathbf{G}$ and the recovered $\mathbf{A}$:

$$E_g = \frac{1}{mn}\|\mathbf{G} - \mathbf{A}\|_F^2. \tag{7}$$

Test results show that FrALM converges as efficiently as LrALM (Fig. 2(a)). Since the desired rank of $\mathbf{A}$ is 1, λ has to be sufficiently small for LrALM to produce accurate results (Fig. 1(a)). For Sets 2, 3, and 5 with non-sparse $\mathbf{E}$, the empirical optimal λ (0.002) is smaller than the theoretical λ^* (0.00$\dot{3}$), contrary to the theory of [13]. At this lower λ, the ranks of the optimal $\mathbf{A}$ recovered by LrALM are still larger than the known value of 1 (Fig. 1(b)). This shows that LrALM has accumulated higher-rank components in $\mathbf{A}$ and thus over-estimated $\mathbf{A}$. In contrast, FrALM constrains the rank of $\mathbf{A}$ to 1, removing the over-estimation. Consequently, FrALM yields more accurate results than does LrALM, and it returns optimal or near optimal results over a wide range of λ (Fig. 1(a)). We have also verified empirically that reducing the rank of $\mathbf{A}$ to 1 after it is returned by LrALM can reduce over-estimation and improve LrALM's accuracy. However, this post-processing is insufficient for removing the over-estimation entirely and LrALM's error is still larger than that of FrALM.

To investigate the stability of FrALM, we ran it on the test cases at a range of fixed ranks r, with λ set to the empirical optimal of 0.002. FrALM's results were plotted together with LrALM's results obtained in previous tests (Fig. 2(b)). When r is slightly larger than 1, FrALM's error increases only slightly. when r is larger than the rank of $\mathbf{A}$ recovered (line 6 of algorithm), FrALM reduces to LrALM, and its error simply approaches that of LrALM.
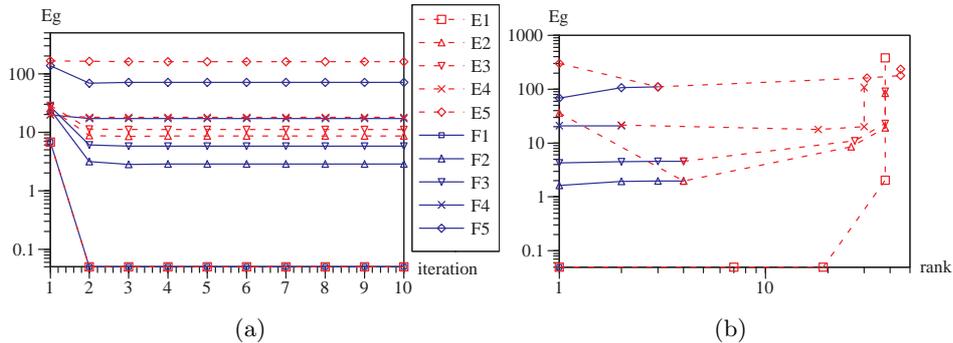
Fig. 2. Performance comparison for reflection removal. (a) Convergence curves with $\lambda = 0.00\dot{3}$. (b) Error vs. rank. Error curves above $E_g = 1000$ for very small λ are cropped to reduce clutter.

Figure 3 displays sample results for reflection removal obtained at the theoretical λ^*. LrALM's results are good for Sets 1 and 4 whose $\mathbf{E}$ is sparse. For Sets 2, 3, and 5, $\mathbf{E}$ is not sparse and LrALM's results have visually noticeable errors (when the images are viewed at higher zoom factors). In contrast, FrALM obtains good results for all test sets.

Figure 4 shows sample results for video background modeling obtained at the theoretical λ^*. In the video frames where the human and vehicles are moving continuously, LrALM can recover the stationary background well (Fig. 4(1c, 2c)). When the vehicles are moving slowly, $\mathbf{E}$ is not sparse, and LrALM shows signs of inaccuracy (Fig. 4(3c)). When the vehicles stop at the traffic junction for an extended period of time, LrALM regards them as part of the low-rank matrix $\mathbf{A}$ and fails to remove them from $\mathbf{A}$ (Fig. 4(4c)). In contrast, FrALM produces much better overall results than does LrALM (Fig. 4(d)).

5   Conclusions

A fixed-rank RPCA algorithm, FrALM, based on exact augmented Lagrange multiplier method is proposed in this paper. By fixing the rank of the low-rank matrix to be recovered, FrALM removes over-estimation of the low-rank matrix and produces more accurate results than does low-rank ALM method (LrALM). Moreover, FrALM returns optimal or near optimal results over a wide range of λ values, whereas LrALM's accuracy is sensitive to λ. If FrALM is fixed to a desired rank that is larger than the actual rank, then FrALM just reduces to LrALM. These properties make FrALM more reliable and accurate than LrALM for solving computer vision problems whose low-rank matrices have known ranks.

References

1. S. D. Babacan, M. Luessi, R. Molina, and A. K. Katsaggelos. Sparse bayesian methods for low-rank matrix estimation. *IEEE Trans. Signal Processing*, 60(8):3964–
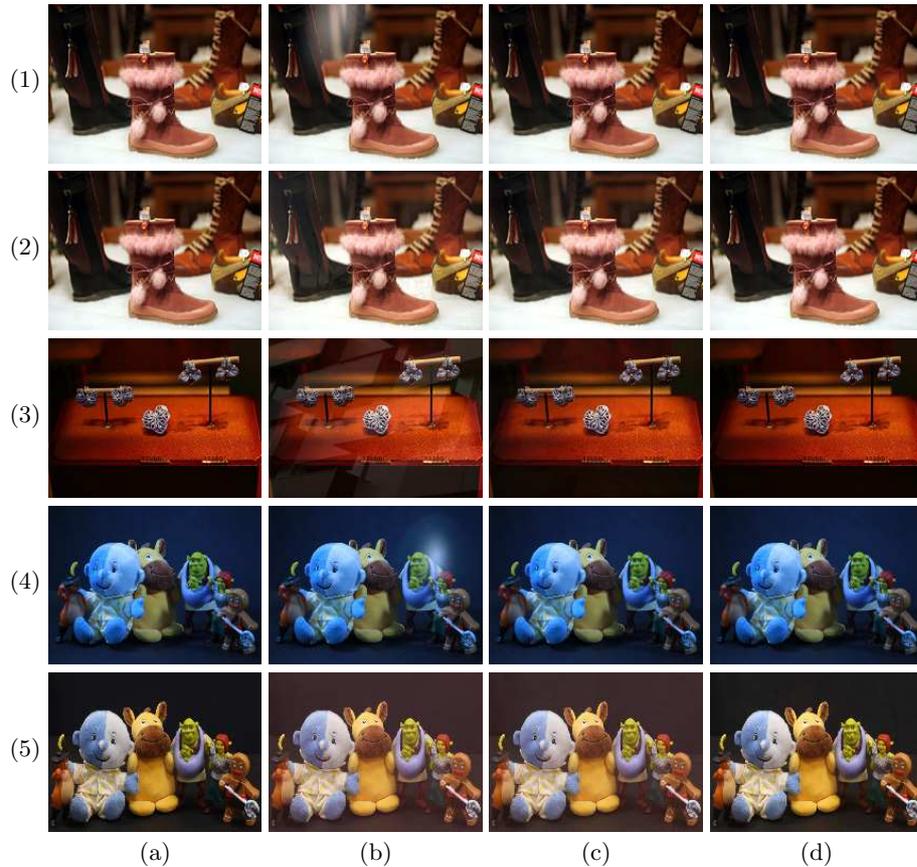
Fig. 3. Sample test results for reflection removal. (a) Ground truth background. (b) Sample input images. (c) LrALM's results. (d) FrALM's results. (1) Set 1: synthetic local reflection. (2) Set 2: synthetic global reflection (light background). (3) Set 3: synthetic global reflection (dark background). (4) Set 4: real local reflection. (5) Set 5: real global reflection. $\lambda = 0.00\dot{3}$ for these test results.

3977, 2012.

2. E. J. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of ACM*, 58(3):11, 2011.

3. E. J. Candès and Y. Plan. Matrix completion with noise. In *Proc. IEEE*, pages 925–936, 2010.

4. F. De la Torre and M. Black. A framework for robust subspace learning. *Int. Journal of Computer Vision*, 54(1–3):117–142, 2003.

5. M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of ACM*, 24(6):381–385, 1981.

6. A. Ganesh, Z. Lin, J. Wright, L. Wu, M. Chen, and Y. Ma. Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. In *CAMSAP*, 2009.
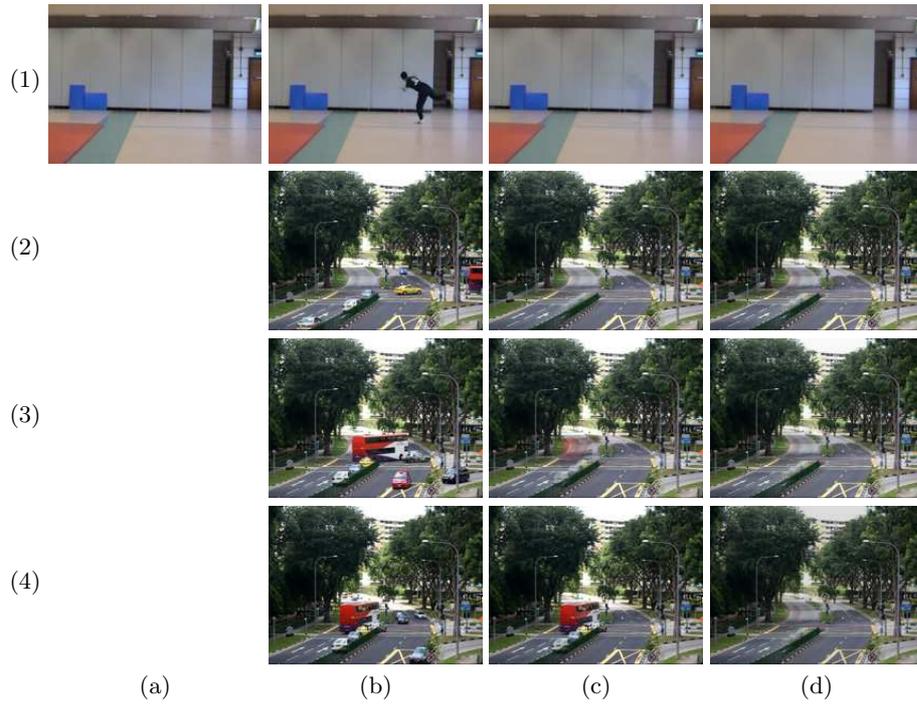
Fig. 4. Sample test results for human and traffic video. (a) Ground truth background. (b) Sample video frames. (c) LrALM's results. (d) FrALM's results. (1) Human motion video. (2–4) Traffic video; ground truth is not available. $\lambda = 0.003$ for these test results.

7. R. Gnanadesikan and J. Kettenring. A framework for robust subspace learning. *Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data (check journal title)*, 28(1):81–124, 1972.
8. Q. Ke and T. Kanade. Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming. In *Proc. CVPR*, pages 739–746, 2005.
9. Z. Lin, M. Chen, L. Wu, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. Technical Report UILU-ENG-09-2215, UIUC, 2009. arXiv preprint arXiv:1009.5055.
10. R. Liu, Z. Lin, F. De la Torre, and Z. Su. Fixed-rank representation for unsupervised visual learning. In *Proc. CVPR*, pages 598–605, 2012.
11. N. Wang, T. Yao, J. Wang, and D.-Y. Yeung. A probabilistic approach to robust matrix factorization. In *Proc. ECCV*, pages 126–139, 2012.
12. J. Wright, Y. Peng, Y. Ma, A. Ganesh, and S. Rao. Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization. In *Proc. NIPS*, pages 2080–2088, 2009.
13. Z. Zhou, X. Li, J. Wright, E. J. Candès, and Y. Ma. Stable principal component pursuit. In *Proc. Int. Symp. Information Theory*, pages 1518–1522, 2010.