



**HAL**  
open science

# Foreground Detection by Robust PCA solved via a Linearized Alternating Direction Method

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah

► **To cite this version:**

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah. Foreground Detection by Robust PCA solved via a Linearized Alternating Direction Method. International Conference on Image Analysis and Recognition, ICIAR 2012, Jun 2012, Portugal. pp.115-122. hal-00809463

**HAL Id: hal-00809463**

**<https://hal.science/hal-00809463>**

Submitted on 9 Apr 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Foreground Detection by Robust PCA solved via a Linearized Alternating Direction Method

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah

Laboratoire MIA, Univ. La Rochelle, France  
tbouwman@univ-lr.fr

**Abstract.** Robust Principal Components Analysis (RPCA) shows a nice framework to separate moving objects from the background. The background sequence is then modeled by a low rank subspace that can gradually change over time, while the moving foreground objects constitute the correlated sparse outliers. RPCA problem can be exactly solved via convex optimization that minimizes a combination of the nuclear norm and the  $l_1$ -norm. To solve this convex program, an Alternating Direction Method (ADM) is commonly used. However, the subproblems in ADM are easily solvable only when the linear mappings in the constraints are identities. This assumption is rarely verified in real application such as foreground detection. In this paper, we propose to use a Linearized Alternating Direction Method (LADM) with adaptive penalty to achieve RPCA for foreground detection. LADM alleviates the constraints of the original ADM with a faster convergence speed. Experimental results on different datasets show the pertinence of the proposed approach.

**Keywords:** Foreground Detection, Robust PCA, Linearized ADM

## 1 Introduction

The detection of moving objects is the basic low-level operations in video analysis. This detection is usually done using foreground detection. This basic operation consists of separating the moving objects called "foreground" from the static information called "background" [1][2]. In 1999, Oliver et al. [3] are the first authors that model the background by Principal Component Analysis (PCA). PCA provides a robust model of the probability distribution function of the background, but not of the moving objects while they do not have a significant contribution to the model. The main limitation of this model [2] is that the size of the foreground object must be small and don't appear in the same location during a long period in the training sequence. Recent research on robust PCA [4] [5] can be used to alleviate these limitations. For example, Candes et al. [5] proposed a convex optimization to address the robust PCA problem. The observation matrix is assumed represented as:  $A = L + S$  where  $L$  is a low-rank matrix and  $S$  must be sparse matrix with a small fraction of nonzero entries. This research seeks to solve for  $L$  with the following optimization problem:

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1 \quad \text{subj } A = L + S \quad (1)$$

where  $\|\cdot\|_*$  and  $\|\cdot\|_1$  are the nuclear norm and  $l_1$ -norm, respectively, and  $\lambda > 0$  is an arbitrary balanced parameter. Under these minimal assumptions, this approach perfectly recovers the low-rank and the sparse matrices. The optimization in Equation (1) can be solved as a general convex optimization problem by any off-the-shelf interior point solver [6]. However, although interior point methods normally take very few iterations to converge, they have difficulty in handling large matrices because the complexity of computing the step direction is  $O(n^6)$ , where  $n$  is the dimension of the matrix. So the generic interior point solvers are too limited for many real applications. The interior point solvers do not scale well for large matrices because they keep the second-order information. To overcome the scalability issue, only the first-order information can be used. Cai et al. [7] showed that this technique, called Singular Value Thresholding (SVT), can be used to minimize the nuclear norm. As the matrix recovery problem in Equation (1) needs minimizing a combination of both the  $l_1$ -norm and the nuclear norm, Wright et al. [8] adopted an iterative thresholding technique to solve it and obtained similar convergence and scalability properties than interior point methods. However, the iterative thresholding scheme converges extremely slowly. To alleviate this slow convergence, Lin et al. [9] proposed the accelerated proximal gradient (APG) algorithm and the gradient-ascent algorithm applied to the dual of the problem in Equation (1). However, these algorithms are all the same to slow for real application. More recently, Lin et al. [10] proposed two algorithms based on augmented Lagrange multipliers (ALM). The first algorithm is called exact ALM (EALM) method that has a Q-linear convergence speed, while the APG is in theory only sub-linear. The second algorithm is an improvement of the EALM that is called inexact ALM (IALM) method, which converges practically as fast as the exact ALM, but the required number of partial SVDs is significantly less. The IALM is at least five times faster than APG, and its precision is also higher [10]. However, the direct application of ALM treats the Equation (1) as a generic minimization problem and ignores its separable structure emerging in both the objective function and the constraint [10]. Hence, the variables  $S$  and  $L$  are minimized simultaneously. Yuan and Yang [11] proposed to alleviate this ignorance by the Alternating Direction Method (ADM) which minimizes the variables  $L$  and  $S$  serially. However, the subproblems in ADM are easily solvable only when the linear mappings in the constraints are identities. This assumption is rarely verified in real application. Recently, Lin et al. [12] addressed this problem by Linearized Alternating Direction Method (LADM) with adaptive penalty. They proved the global convergence of LADM and applied it to solve low-rank representation (LRR). Furthermore, this method reduces the complexity  $O(n^3)$  of the original ADM based method to  $O(rn^2)$ , where  $r$  and  $n$  are the rank and size of the representation matrix, respectively. These properties allow us to use it for foreground detection. The rest of this paper is organized as follows. In Section 2, we present how the background and the foreground can be separated via RPCA. Then, the proposed foreground detection method with the LADM algorithm is presented. Finally, performance evaluation and comparison with four other algorithms are given in Section 3.

## 2 Foreground detection by RPCA solved via LADM

Denote the training video sequences  $D = \{I_1, \dots, I_N\}$  where  $I_t$  is the frame at time  $t$  and  $N$  is the number of training frames. Let each pixel  $(x, y)$  be characterized by its intensity in the grey scale. The decomposition involves the following model:

$$D = L + S \quad (2)$$

where  $L$  and  $S$  are the low-rank component and sparse component of  $D$ , respectively. The matrix  $L$  contains the background and the matrix  $S$  contains mostly zero columns, with several non-zero ones corresponding to the foreground. The matrices  $L$  and  $S$  can be recovered by the convex program based on the Alternating Direction Method (ADM). For this work on foreground detection, we propose to use the Linearized Alternating Direction Method (LADM) with adaptive penalty [12]. The proposed algorithm to solve RPCA via LADM is theoretically the same problem to solve the low-rank representation (LRR) problem [12]. The LRR problem can be expressed by:

$$\min_{Z, E} \|Z\|_* + \mu \|E\|_{2,1}, s.t. X = XZ + E \quad (3)$$

By analogy with equation (2), the matrices  $X$ ,  $E$ , and  $XZ$  correspond respectively to  $D$ ,  $S$  and  $L$ . By further representing  $Z$  as its skinny SVD and utilizing an advanced functionality of the PROPACK<sup>1</sup> package, the complexity of solving LRR by LADM becomes only  $O(rn^2)$ , as there is no full sized matrix-matrix multiplications, where  $r$  is the rank of the optimal  $Z$ . Theoretically, the problem is initially formulated as a linear constrained convex program [12].

$$\min_{x, y} f(x) + g(y), s.t. \mathcal{A}(x) + \mathcal{B}(y) = c \quad (4)$$

Solving this problem by ADM, one operates on the following augmented Lagrangian function:

$$\mathcal{L}_A(x, y, \lambda) = f(x) + g(y) + \langle \lambda, \mathcal{A}(x) + \mathcal{B}(y) - c \rangle + \frac{\beta}{2} \|\mathcal{A}(x) + \mathcal{B}(y) - c\|^2 \quad (5)$$

where  $\lambda$  is the Lagrange multiplier,  $\langle \cdot, \cdot \rangle$  is the inner product, and  $\beta > 0$  is the penalty parameter. ADM [11] decomposes the minimization of  $\mathcal{L}_A$  w.r.t.  $(x, y)$  into two subproblems that minimize w.r.t.  $x$  and  $y$ , respectively. By linearizing ADM and adding a proximal term, Lin et al. [12] obtain a linearized ADM that accelerates the convergence with an adaptive rule for updating the penalty parameter as follows:

$$x_{k+1} = \operatorname{argmin}_x f(x) + \frac{\beta \eta_A}{2} \|x - x_k + \mathcal{A}^*(\lambda_k + \beta(\mathcal{A}(x_k) + \mathcal{B}(y_k) - c)) / (\beta \eta_A)\|^2 \quad (6)$$

where  $\mathcal{A}^*$  is the adjoint of  $\mathcal{A}$  and  $\eta_A > 0$  is a parameter used in the stopping criterion. Similarly,

$$y_{k+1} = \operatorname{argmin}_y g(y) + \frac{\beta \eta_B}{2} \|y - y_k + \mathcal{B}^*(\lambda_k + \beta(\mathcal{A}(x_{k+1}) + \mathcal{B}(y_k) - c)) / (\beta \eta_B)\|^2 \quad (7)$$

<sup>1</sup> <http://soi.stanford.edu/rmunk/PROPACK/>

$$\lambda_{k+1} = \lambda_k + \beta[\mathcal{A}(x_{k+1}) + \mathcal{B}(y_{k+1} - c)] \quad (8)$$

The adaptative updating strategy for the penalty parameter  $\beta$  is as follows:

$$\beta_{k+1} = \min(\beta_{max}, \rho\beta_k), \quad (9)$$

where  $\beta_{max}$  is an upper bound of  $\{\beta_k\}$ . the value of  $\rho$  is defined as

$$\rho = \begin{cases} \rho_0, & \text{if } \beta_k \max(\sqrt{\eta_A}\|x_{k+1} - x_k\|, \sqrt{\eta_B}\|y_{k+1} - y_k\|)\|c\| < \varepsilon_1; \\ 1, & \text{otherwise.} \end{cases} \quad (10)$$

The algorithm runs while the following two stopping criteria are not satisfied:

$$\|\mathcal{A}(x_{k+1}) + \mathcal{B}(y_{k+1}) - c\|/\|c\| < \varepsilon_1 \quad (11)$$

$$\beta_k \max(\sqrt{\eta_A}\|x_{k+1} - x_k\|, \sqrt{\eta_B}\|y_{k+1} - y_k\|)\|c\| < \varepsilon_1 \quad (12)$$

Then, the proposed algorithm for foreground detection is as follows.

#### Algorithm for Foreground Detection via RPCA-LADM

**Require:** Training sequence with  $N$  images that is contained in  $D$  and parameter  $\mu > 0$ .

**Ensure:** Background in  $L$  and foreground in  $S$ .

##### **LRR Algorithm via LADM**

**Initialize:** Set  $X$  to  $D$ , set  $E_0$ ,  $Z_0$  and  $\Lambda_0$  to zero matrices, where  $Z_0$  is represented as  $(U_0, \Sigma_0, V_0) \leftarrow (0, 0, 0)$ . Set  $\varepsilon_1 > 0, \varepsilon_2 > 0, \beta_{max} \gg \beta_0 > 0, \eta_X > \sigma_{max}^2(X), r = 5$ , and  $k \leftarrow 0$ .

**while** stopping criteria (11) and (12) are not satisfied **do**

**Step1:** Update  $E_{k+1} = \operatorname{argmin}_E \mu \|E\|_{2,1} + \frac{\beta_k}{2} \|(XU_k)\Sigma_k V_k^T - X + V_k/\beta_k\|^2$ .

**Step2:** Update the skinny SVD  $(U_{k+1}, \Sigma_{k+1}, V_{k+1})$  of  $Z_{k+1}$ .

**Step3:** Update the predicted rank  $r$ .

if  $r' < r$ , then  $r = \min(r' + 1, n)$ ; otherwise,  $r = \min(r' + \operatorname{round}(0.05n), n)$ .

**Step4:** Update  $\Lambda_{k+1} = \Lambda_k + \beta_k((XU_{k+1})\Sigma_{k+1}V_{k+1}^T + E_{k+1} - X)$ .

**Step5:** Update  $\beta_{k+1}$  by (9) and (10).

**Step6:**  $k \leftarrow k + 1$ .

**end while**

$L \leftarrow XZ, S \leftarrow E$ .

**Foreground Detection** Threshold the matrix  $S$  to obtain the foreground.

For the initialization, the matrix  $X$  is set to  $D$ . The output are the matrices  $L$  and  $S$ . The foreground detection mask is obtained by thresholding the matrix  $S$ . Fig. 1 shows the original frame 309 of the sequence from [13] and its decomposition into the low-rank matrix  $L$  and sparse matrix  $S$ . We can see that  $L$  corresponds to the background whereas  $S$  corresponds to the foreground. The fourth image shows the foreground mask obtained by thresholding the matrix  $S$  and the fifth image is the ground truth image.



Fig. 1: Original image (309), low-rank matrix L (background), sparse matrix S (foreground), foreground mask, ground truth.

### 3 Experimental Results

We compared the proposed approach with the PCA [3], RSL [4], RPCA-EALM [10], RPCA-IALM [10]. The experiments were conducted qualitatively and quantitatively on the Wallflower dataset [14] and I2R dataset [15]. The algorithms were implemented in batch mode with matlab.

#### 3.1 Wallflower dataset<sup>2</sup>

This dataset consists of seven video sequences, with each sequence presenting one of the difficulties a practical task is likely to encounter. The images are  $160 \times 120$  pixels. For each sequence, the ground truth is provided for one image when the algorithm has to show its robustness to a specific change in the scene. Thus, the performance is evaluated against hand-segmented ground truth. The figure 2 shows the qualitative results. For the quantitative evaluation, we used metrics based on the true negative (TN), true positive (TP), false negative (FN), false positive (FP) detections. Then, we computed the detection rate, the precision and the F-measure. A good performance is obtained when the detection rate is high without altering the precision. This can be measured by the F-measure [15]. A good performance is then reached when the F-measure is closed to 1 (100 in percentage). Table 1 shows the F-measure in percentage for each sequence. The F-measure value of MO sequence can't be computed due to the absence of true positives in its ground-truth. RPCA-LADM outperforms globally RSL, RPCA-EALM and RPCA-IALM. For the four sequences (MO, C, B, FA), the proposed method gives better results than RSL, RPCA-EALM and RPCA-IALM. For the sequence WT, TD and LS, results are still acceptable. As these encouraging results are evaluated by using one ground-truth image, we have evaluated the proposed method on a dataset with more ground-truth images in the following sub-section.

#### 3.2 I2R dataset<sup>3</sup>

This dataset provided by [8] consists of nine video sequences, which each sequence presenting dynamic backgrounds or illumination changes. The size of the images is  $176 \times 144$  pixels. For each sequence, the ground truth is provided for 20 images.

<sup>2</sup> <http://research.microsoft.com/en-us/um/people/jckrumm/wallflower/testimages.htm>

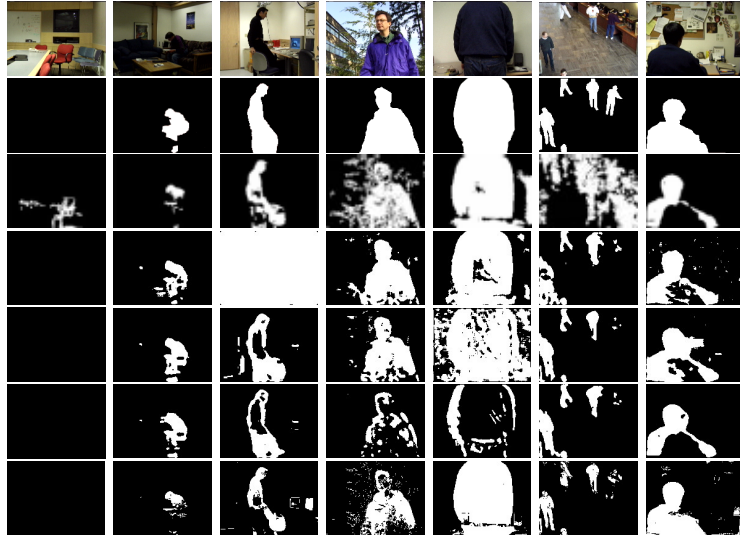


Fig. 2: Experimental results on the Wallflower dataset. From top to bottom: original image, ground truth, PCA, RSL, RPCA-EALM, RPCA-IALM, RPCA-LADM. From left to right: MO (985), TD (1850), LS (1865), WT (247), C (251), B (2832), FA (449).

Table 1: F-measure on the Wallflower dataset

	RSL	RPCA-EALM	RPCA-IALM	RPCA-LADM
TD	75.73	81.18	80.56	55.07
LS	28.36	70.86	73.16	59.92
WT	89.69	86.40	40.88	76.43
C	91.78	75.43	22.02	<b>96.74</b>
B	69.38	74.4	73.73	<b>76.07</b>
FA	74.37	72.07	61.92	<b>91.06</b>

Table 2: F-measure on the I2R dataset

	RSL	RPCA-IALM	RPCA-LADM
Airport	65.26	74.26	<b>75.51</b>
Water surface	34.42	31.34	<b>86.86</b>
Curtain	70.73	76.59	<b>82.25</b>

Among this dataset, we have chosen to show results on three representative sequences that are the following ones: airport, water surface and curtain. We ran RPCA-IALM and RPCA-LADM on these sequences. We skipped RPCA-EALM since it would take excessive amounts of time due to full SVD calculations. Since all these video clips have more than 1000 frames, we took a part of each clip with 200 frames. Fig. 3 shows the qualitative results. For example, we can see that the RPCA-LADM allows to detect the complete silhouette in the sequence called "water surface". Table 2 shows the average F-measure in percentage that is obtained on 20 ground truth images for each sequence. We can see that RPCA-LADM outperforms RSL and RPCA-IALM. Furthermore, we have compared the CPU times required by RPCA-EALM, RPCA-IALM and RPCA-LADM in term of both the number of SVDs and CPU times. In the case of the sequence "Airport" in the resolution  $176 \times 144$  with 200 training images, the CPU time of each algorithm is respectively 43min, 13min, and 5min20s and the SVDs times is respectively 550 SVDs, 38 SVDs, and 6 SVDs. On these problems of extremely low ranks, the partial SVD technique used in RPCA-IALM and RPCA-LADM becomes quite effective. Even so, the CPU times required by RPCA-IALM are still about two times of those required by RPCA-LADM. Furthermore, the speed can be improved by a GPU implementation.

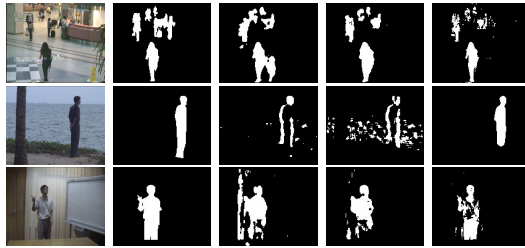


Fig. 3: Experimental results on the I2R dataset. From left to right: original image, ground truth, RSL, RPCA-IALM, RPCA-LADM. From top to bottom: airport (2926), water surface (1594), curtain (23257).

## 4 Conclusion

In this paper, we have presented a foreground detection method based on RPCA that is optimized via a linearized alternating direction method. This method allows us to alleviate the constraints of the identities. Furthermore, experiments on video surveillance datasets show that this approach is more robust than RSL, RPCA-EALM, and RPCA-IALM in the presence of dynamic backgrounds and illumination changes. In terms of computational efficiency, the proposed approach

<sup>3</sup> <http://perception.i2r.a-star.edu.sg/>



has exhibited a significant speed advantage over RPCA-IALM. Furthermore, its speed can be improved by a GPU implementation. Although the main drawback of the proposed method is its batch aspect, future research consists in developing an incremental version to update the model at every frame and to achieve real-time requirements.

## 5 Acknowledgements

The authors would like to thank Zhouchen Lin (Visual Computing Group, Microsoft Research Asia) who has kindly provided the batch algorithm of LADM.

## References

1. T. Bouwmans, "Recent advanced statistical background modeling for foreground detection: A systematic survey," *RPACS*, vol. 4, no. 3, pp. 147–176, Nov. 2011.
2. T. Bouwmans, "Subspace learning for background modeling: A survey," *RPACS*, vol. 2, no. 3, pp. 223–234, Nov. 2009.
3. N. Oliver, B. Rosario, and A. Pentland, "A bayesian computer vision system for modeling human interactions," *ICVS 1999*, Jan. 1999.
4. F. De La Torre and M. Black, "A framework for robust subspace learning," *International Journal on Computer Vision*, pp. 117–142, 2003.
5. E. Candes, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?," *International Journal of ACM*, vol. 58, no. 3, May 2011.
6. V. Chandrasekharan, S. Sanghavi, P. Parillo, and A. Wilsky, "Rank-sparsity incoherence for matrix decomposition," *Preprint*, 2009.
7. J. Cai, E. Candes, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *International Journal of ACM*, May 2008.
8. 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," *NIPS 2009*, Dec. 2009.
9. Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, and Y. Ma, "Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix," *UIUC Technical Report*, Aug. 2009.
10. Z. Lin, M. Chen, L. Wu, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *UIUC Technical Report*, Nov. 2009.
11. X. Yuan and J. Yang, "Sparse and low-rank matrix decomposition via alternating direction methods," *Optimization Online*, Nov. 2009.
12. Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," *NIPS 2011*, Dec. 2011.
13. Y. Sheikh and M. Shah, "Bayesian modeling of dynamic scenes for object detection," *IEEE T-PAMI*, vol. 27, pp. 1778–1792, 2005.
14. K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: Principles and practice of background maintenance," *ICCV*, pp. 255–261, Sept. 1999.
15. L. Maddalena and A. Petrosino, "A fuzzy spatial coherence-based approach to background foreground separation for moving object detection," *Neural Computing and Applications*, pp. 1–8, 2010.