

Robust Principal Component Analysis for Background Subtraction: Systematic Evaluation and Comparative Analysis

Charles Guyon, Thierry Bouwmans and El-hadi Zahzah
*Lab. MIA - Univ. La Rochelle
 France*

1. Introduction

The analysis and understanding of video sequences is currently quite an active research field. Many applications such as video surveillance, optical motion capture or those of multimedia need to first be able to detect the objects moving in a scene filmed by a static camera. This requires the basic operation that consists of separating the moving objects called "foreground" from the static information called "background". Many background subtraction methods have been developed (Bouwmans et al. (2010); Bouwmans et al. (2008)). A recent survey (Bouwmans (2009)) shows that subspace learning models are well suited for background subtraction. Principal Component Analysis (PCA) has been used to model the background by significantly reducing the data's dimension. To perform PCA, different Robust Principal Components Analysis (RPCA) models have been recently developed in the literature. 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. However, authors compare their algorithm only with the PCA (Oliver et al. (1999)) or another RPCA model. Furthermore, the evaluation is not made with the datasets and the measures currently used in the field of background subtraction. Considering all of this, we propose to evaluate RPCA models in the field of video-surveillance. Contributions of this chapter can be summarized as follows:

- A survey regarding robust principal component analysis
- An evaluation and comparison on different video surveillance datasets

The rest of this paper is organized as follows: In Section 2, we firstly provide the survey on robust principal component analysis. In Section 3, we evaluate and compare robust principal component analysis in order to achieve background subtraction. Finally, the conclusion is established in Section 4.

2. Robust principal component analysis: A review

In this section, we review the original PCA and five recent RPCA models and their applications in background subtraction:

- Principal Component Analysis (PCA) (Eckart & Young (1936); Oliver et al. (1999))

- RPCA via Robust Subspace Learning (RSL) (Torre & Black (2001); Torre & Black (2003))
- RPCA via Principal Component Pursuit (PCP) (Candes et al. (2009))
- RPCA via Templates for First-Order Conic Solvers (TFOCS¹) (Becker et al. (2011))
- RPCA via Inexact Augmented Lagrange Multiplier (IALM²) (Lin et al. (2009))
- RPCA via Bayesian Framework (BRPCA) (Ding et al. (2011))

2.1 Principal component analysis

Assuming that the video is composed of n frames of size $width \times height$. We arrange this training video in a rectangular matrix $A \in \mathbb{R}^{m \times n}$ (m is the total amount of pixels), each video frame is then vectorized into column of the matrix A , and rows correspond to a specific pixel and its evolution over time. The PCA firstly consists of decomposing the matrix A in the product USV' . where $S \in \mathbb{R}^{n \times n} (diag)$ is a diagonal matrix (singular values), $U \in \mathbb{R}^{m \times n}$ and $V \in \mathbb{R}^{n \times n}$ (singular vectors). Then only the principals components are retained. To solve this decomposition, the following function is minimized (in tensor notation):

$$(S_0, U_0, V_0) = \underset{S, U, V}{\operatorname{argmin}} \sum_{r=1}^{\min(n, m)} \|A - \underset{kk \quad ik \quad jk}{S} \underset{ki \quad kj}{U} \underset{ki \quad kj}{V}\|_F^2, \quad 1 \leq k \leq r \quad \text{subj} \quad \begin{cases} UU = VV = 1 & \text{if } i = j \\ S = 0 & \text{if } i \neq j \end{cases} \quad (1)$$

This imply singular values are straightly sorted and singular vectors are mutually orthogonal ($U_0' U_0 = V_0' V_0 = I_n$). The solutions S_0, U_0 and V_0 of (1) are not unique.

We can define U_1 and V_1 , the set of cardinality $2^{\min(n, m)}$ of all solution;

$$U_1 = U_0 R, \quad V_1 = R V_0, \quad R = \begin{cases} \pm 1 & \text{if } i = j \\ 0 & \text{elsewhere} \end{cases}, \quad m > n \quad (2)$$

We choose k (small) principal components:

$$U = U_1, \quad 1 \leq j \leq k \quad (3)$$

The background is computed as follows:

$$Bg = UU'v \quad (4)$$

where v is the current frame. The foreground detection is made by thresholding the difference between the current frame v and the reconstructed background image (in Iverson notation):

$$Fg = [|v - Bg| < T] \quad (5)$$

where T is a constant threshold.

Results obtained by Oliver et al. (1999) show that the 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. As developped in Bouwmans (2009), this

¹ <http://tfocs.stanford.edu/>

² http://perception.csl.uiuc.edu/matrix-rank/sample_code.html

model presents several limitations. The first limitation of this model 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. The second limitation appears for the background maintenance. Indeed, it is computationally intensive to perform model updating using the batch mode PCA. Moreover without a mechanism of robust analysis, the outliers or foreground objects may be absorbed into the background model. The third limitation is that the application of this model is mostly limited to the gray-scale images since the integration of multi-channel data is not straightforward. It involves much higher dimensional space and causes additional difficulty to manage data in general. Another limitation is that the representation is not multimodal so various illumination changes cannot be handled correctly. In this context, several robust PCA can be used to alleviate these limitations.

2.2 RPCA via Robust Subspace Learning

Torre & Black (2003) proposed a Robust Subspace Learning (RSL) which is a batch robust PCA method that aims at recovering a good low-rank approximation that best fits the majority of the data. RSL solves a nonconvex optimization via alternative minimization based on the idea of soft-detecting and down-weighting the outliers. These reconstruction coefficients can be arbitrarily biased by an outlier. Finally, a binary outlier process is used which either completely rejects or includes a sample. Below we introduce a more general analogue outlier process that has computational advantages and provides a connection to robust M-estimation. The energy function to minimize is then:

$$(S_0, U_0, V_0) = \underset{S, U, V}{\operatorname{argmin}} \sum_{r=1}^{\min(n, m)} \rho(A - \mu \mathbf{1}_n' - S U V) , \quad 1 \leq k \leq r \quad (6)$$

where μ is the mean vector and the ρ - function is the particular class of robust ρ -function (Black & Rangarajan (1996)). They use the Geman-McClure error function $\rho(x, \sigma_p) = \frac{x^2}{x^2 + \sigma_p^2}$ where σ_p is a scale parameter that controls the convexity of the robust function. Similar, the penalty term associate is $(\sqrt{L_{pi}} - 1)^2$. The robustness of De La Torre's algorithm is due to this ρ - function. This is confirmed by the results presented which show that the RSL outperforms the standard PCA on scenes with illumination change and people in various locations.

2.3 RPCA via Principal Component Pursuit

Candes et al. (2009) achieved Robust PCA by the following decomposition:

$$A = L + S \quad (7)$$

where L is a low-rank matrix and S must be sparse matrix. The straightforward formulation is to use L_0 norm to minimize the energy function:

$$\underset{L, S}{\operatorname{argmin}} \operatorname{Rank}(L) + \lambda \|S\|_0 \quad \text{subj } A = L + S \quad (8)$$

where λ is arbitrary balanced parameter. But this problem is NP-hard, typical solution might involve a search with combinatorial complexity. For solve this more easily, the natural way is

to fix the minimization with L_1 norm that provided an approximate convex problem:

$$\underset{L, S}{\operatorname{argmin}} \|L\|_* + \lambda \|S\|_1 \quad \text{subj } A = L + S \quad (9)$$

where $\|\cdot\|_*$ is the nuclear norm (which is the L_1 norm of singular value). Under these minimal assumptions, the PCP solution perfectly recovers the low-rank and the sparse matrices, provided that the rank of the low-rank matrix and the sparsity matrix are bounded by the follow inequality:

$$\operatorname{rank}(L) \leq \frac{\rho_r \max(n, m)}{\mu (\log \min(n, m))^2}, \quad \|S\|_0 \leq \rho_s mn \quad (10)$$

where, ρ_r and ρ_s are positive numerical constants, m and n are the size of the matrix A .

For further consideration, lamda is choose as follow:

$$\lambda = \frac{1}{\sqrt{\max(m, n)}} \quad (11)$$

Results presented show that PCP outperform the RSL in case of varying illuminations and bootstrapping issues.

2.4 RPCA via templates for first-order conic solvers

Becker et al. (2011) used the same idea as Candes et al. (2009) that consists of some matrix A which can be broken into two components $A = L + S$, where L is low-rank and S is sparse. The inequality constrained version of RPCA uses the same objective function, but instead of the constraints $L + S = A$, the constraints are:

$$\underset{L, S}{\operatorname{argmin}} \|L\|_* + \lambda \|S\|_1 \quad \text{subj } \|L + S - A\|_\infty \leq \alpha \quad (12)$$

Practically, the A matrix is composed from datas generated by camera, consequently values are quantified (rounded) on 8 bits and bounded between 0 and 255. Suppose $A_0 \in \mathcal{R}^{m \times n}$ is the ideal data composed with real values, it is more exact to perform exact decomposition onto A_0 . Thus, we can assert $\|A_0 - A\|_\infty < \frac{1}{2}$ with $A_0 = L + S$.

The result show improvements for dynamic backgrounds³.

2.5 RPCA via inexact augmented Lagrange multiplier

Lin et al. (2009) proposed to substitute the constraint equality term by penalty function subject to a minimization under L_2 norm :

$$\underset{L, S}{\operatorname{argmin}} \operatorname{Rank}(L) + \lambda \|S\|_0 + \mu \frac{1}{2} \|L + S - A\|_F^2 \quad (13)$$

This algorithm solves a slightly relaxed version of the original equation. The μ constant lets balance between exact and inexact recovery. Lin et al. (2009) didn't present result on background subtraction.

³ <http://www.salleurl.edu/~ftorre/papers/rpca/rpca.zip>

2.6 RPCA via Bayesian framework

Ding et al. (2011) proposed a hierarchical Bayesian framework that considered for decomposing a matrix (A) into low-rank (L), sparse (S) and noise matrices (E). In addition, the Bayesian framework allows exploitation of additional structure in the matrix. Markov dependency is introduced between consecutive rows in the matrix implicating an appropriate temporal dependency, because moving object are strongly correlated across consecutive frames. A spatial dependency assumption is also added and introduce the same Markov constrain as temporal utilizing the local neighborhood. Indeed, it force the sparse outliers component to be spatially and temporally connected. Thus the decomposition is made as follows:

$$A = L + S + E = U(SB_L)V' + X \circ B_S + E \quad (14)$$

Where L is the low-rank matrix, S is the sparse matrix and E is the noise matrix. Then some assumption about components distribution are done:

- Singular vector (U and V') are drawn from normal distribution.
- Singular value and sparse matrix (S and X) value are drawn from normal-gamma distribution
- Singular sparseness mask (B_L and B_S) from bernoulli-beta process.

Note that L_1 minimization is done by l_0 minimization (number of non-zero values fixed for the sparseness mask), afterwards a l_2 minimization is performed on non-zero values.

The matrix A is assumed noisy, with unknown and possibly non-stationary noise statistics. The Bayesian framework infers an approximate representation for the noise statistics while simultaneously inferring the low-rank and sparse-outlier contributions: the model is robust to a broad range of noise levels, without having to change model hyperparameter settings. The properties of this Markov process are also inferred based on the observed matrix, while simultaneously denoising and recovering the low-rank and sparse components. Ding et al. (2011) applied it to background modelling and the result obtain show more robustness to noisy background, slow changing foreground and bootstrapping issue than the RPCA via convex optimization (Wright et al. (2009)).

3. Comparison

In this section, we present the evaluation of the five RPCA models (RSL, PCP, TFOCS, IALM, Bayesian) and the basic average algorithm (SUB) on three different datasets used in video-surveillance: the Wallflower dataset provided by Toyama et al. (1999), the dataset of Li et al. (2004) and dataset of Sheikh & Shah (2005). Qualitative and quantitative results are provided for each dataset.

3.1 Wallflower dataset ⁴

We have chosen this particular dataset provided by Toyama et al. Toyama et al. (1999) because of how frequent its use is in this field. This frequency is due to its faithful representation of real-life situations typical of scenes susceptible to video surveillance. Moreover, it consists of seven video sequences, with each sequence presenting one of the difficulties a practical task is

⁴ <http://research.microsoft.com/en-us/um/people/jckrumm/wallflower/testimages.htm>

likely to encounter (i.e illumination changes, dynamic backgrounds). The size of the images is 160×120 pixels. A brief description of the Wallflower image sequences can be made as follows:

- **Moved Object (MO):** A person enters into a room, makes a phone call, and leaves. The phone and the chair are left in a different position. This video contains 1747 images.
- **Time of Day (TOD):** The light in a room gradually changes from dark to bright. Then, a person enters the room and sits down. This video contains 5890 images.
- **Light Switch (LS):** A room scene begins with the lights on. Then a person enters the room and turns off the lights for a long period. Later, a person walks in the room and switches on the light. This video contains 2715 images.
- **Waving Trees (WT):** A tree is swaying and a person walks in front of the tree. This video contains 287 images.
- **Camouflage (C):** A person walks in front of a monitor, which has rolling interference bars on the screen. The bars include similar color to the person's clothing. This video contains 353 images.
- **Bootstrapping (B):** The image sequence shows a busy cafeteria and each frame contains people. This video contains 3055 images.
- **Foreground Aperture (FA):** A person with uniformly colored shirt wakes up and begins to move slowly. This video contains 2113 images.

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. Four terms are used in the evaluation:

- True Positive (TP) is the number of foreground pixels that are correctly marked as foreground.
- False Positive (FP) is the number of background pixels that are wrongly marked as foreground.
- True Negative (TN) is the number of background pixels that are correctly marked as background.
- False Negative (FN) is the number of foreground pixels that are wrongly marked as background.

		Algorithm	
		Foreground	Background
Ground Truth	Foreground	TP	FN
	Background	FP	TN

Table 1. Measure for performance evaluation

Table 1 illustrates how to compute these different terms. Then, we computed the following metrics: the detection rate, the precision and the F-measure. Detection rate gives the percentage of corrected pixels classified as background when compared with the total number of background pixels in the ground truth:

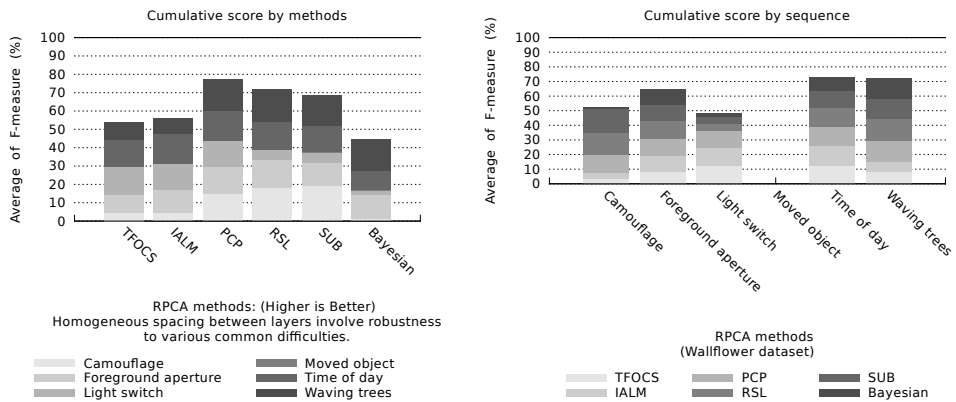


Fig. 1. Performance on the Wallflower dataset. The left (resp. right) figure concern the cumulative score by method (resp. sequence).

$$DR = \frac{TP}{TP + FN} \quad (15)$$

Precision gives the percentage of corrected pixels classified as background as compared at the total pixels classified as background by the method:

$$Precision = \frac{TP}{TP + FP} \quad (16)$$

A good performance is obtained when the detection rate is high without altering the precision. We also computed the F-measure used in (Maddalena & Petrosino (2010)) as follows:

$$F = \frac{2 \times DR \times Precision}{DR + Precision} \quad (17)$$

Table 2 shows the results obtained by the different algorithms on each sequence. For each sequence, the first column shows the original image and the corresponding ground truth.

The second part presents the sparse matrix in the first row and the optimal foreground mask in the second row. The detection rate (DR), Precision (Prec) and F-measure (F) are indicated below each foreground mask. Fig. 1 shows two cumulative histograms of F-measure: The left (resp. right) figure concern the cumulative score by method (resp. sequence). PCP gives the best result followed by RSL, IALM, TFOCS, and Bayesian RPCA. This ranking has to be taken with prrecaution because a poor performance on one sequence influences the overall F-measure and then modifies the rank for just one sequence. For example, the Bayesian obtained a bad score because of the following assumption: the background has necessarily a bigger area than the foreground. It happen in the sequences Camouflage and Light Switch. In the first case, the person hides more than half of the screen space. In the second case, when the person switch on the light all the pixels are affected and the algorithm exchanges the foreground and background. PCP seems to be robust for all critical situations.





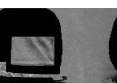








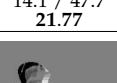

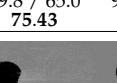
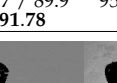
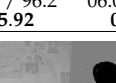
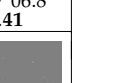






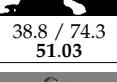
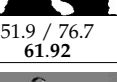
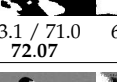

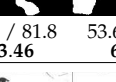
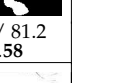
























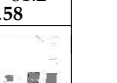






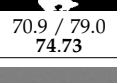
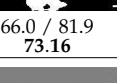
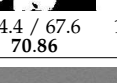
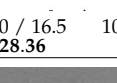
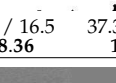
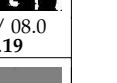
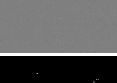



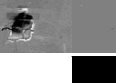











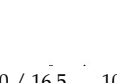
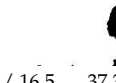
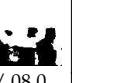

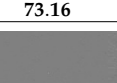
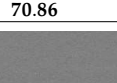
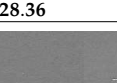
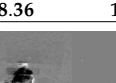
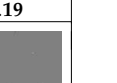

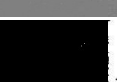


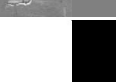



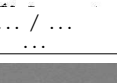
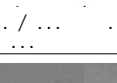
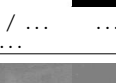

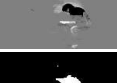





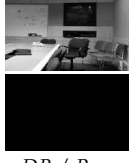
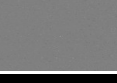



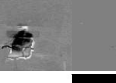

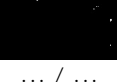
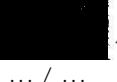
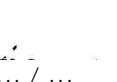
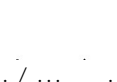














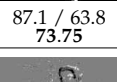
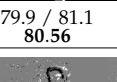
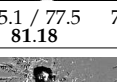
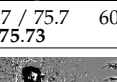
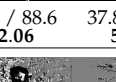
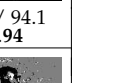



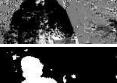
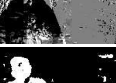


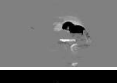










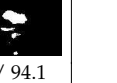












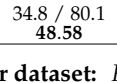
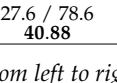
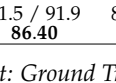
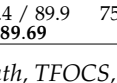
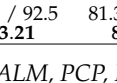
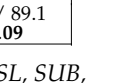
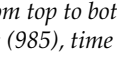
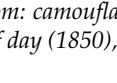
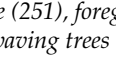
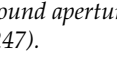
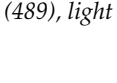
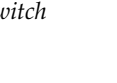




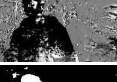
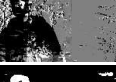







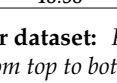
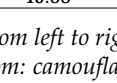
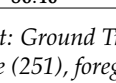
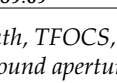
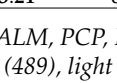
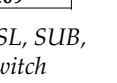
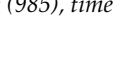
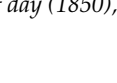
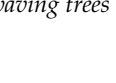
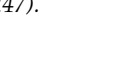














Original / GT	TFOCS	IALM	PCP	RSL	SUB	Bayesian
 DR / Prec F	      14.1 / 47.7 21.77	      13.9 / 53.4 22.02	      89.8 / 65.0 75.43	      93.7 / 89.9 91.78	      95.6 / 96.2 95.92	      06.0 / 06.8 06.41
 DR / Prec F	      38.8 / 74.3 51.03	      51.9 / 76.7 61.92	      73.1 / 71.0 72.07	      69.2 / 80.3 74.37	      51.8 / 81.8 63.46	      53.6 / 81.2 64.58
 DR / Prec F	      70.9 / 79.0 74.73	      66.0 / 81.9 73.16	      74.4 / 67.6 70.86	      100 / 16.5 28.36	      100 / 16.5 28.36	      37.3 / 08.0 13.19
 DR / Prec F	      ... /	      ... /	      ... /	      ... /	      ... /	      ... /
 DR / Prec F	      87.1 / 63.8 73.75	      79.9 / 81.1 80.56	      85.1 / 77.5 81.18	      75.7 / 75.7 75.73	      60.7 / 88.6 72.06	      37.8 / 94.1 53.94
 DR / Prec F	      34.8 / 80.1 48.58	      27.6 / 78.6 40.88	      81.5 / 91.9 86.40	      89.4 / 89.9 89.69	      75.5 / 92.5 83.21	      81.3 / 89.1 85.09

Table 2. **Wallflower dataset:** From left to right: Ground Truth, TFOCS, IALM, PCP, RSL, SUB, Bayesian RPCA. From top to bottom: camouflage (251), foreground aperture (489), light switch (1865), moved object (985), time of day (1850), waving trees (247).








































Original / GT	TFOCS	IALM	PCP	RSL	SUB	Bayesian
 DR / Prec F	  92.3 / 87.2 89.71	  94.4 / 92.3 93.40	  92.5 / 89.3 90.92	  90.5 / 90.7 90.64	  84.9 / 90.6 87.67	  86.9 / 95.2 90.91
 DR / Prec F	  65.2 / 96.0 77.68	  63.4 / 96.7 76.60	  95.3 / 94.6 94.97	  95.6 / 94.9 95.28	  86.1 / 94.1 89.99	  89.1 / 95.2 92.10
 DR / Prec F	  51.7 / 97.5 67.62	  45.4 / 97.1 61.93	  96.0 / 96.1 96.10	  56.4 / 86.6 68.37	  75.0 / 66.7 70.67	  86.2 / 93.7 89.82

Table 3. **Shah dataset:** From left to right: Ground Truth, TFOCS, IALM, PCP, RSL, SUB, Bayesian RPCA. From top to bottom: level crossing (309), level crossing (395), level crossing (462)


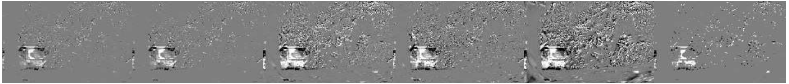

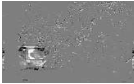

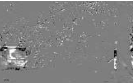

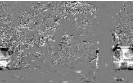

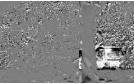

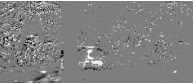














Original / GT	TFOCS	IALM	PCP	RSL	SUB	Bayesian
 DR / Prec F	  67.6 / 66.0 66.82	  70.2 / 65.8 68.00	  58.8 / 81.0 68.20	  60.3 / 81.1 69.19	  57.5 / 61.6 59.52	  51.9 / 67.5 58.73
 DR / Prec F	  86.7 / 79.9 83.22	  87.4 / 85.3 86.37	  87.3 / 83.5 87.42	  92.3 / 83.7 87.82	  82.6 / 81.4 82.03	  79.8 / 80.3 80.06

Table 4. **Li dataset:** From left to right: Ground Truth, TFOCS, IALM, PCP, RSL, SUB, Bayesian RPCA. From top to bottom: campus (1650), campus (1812)

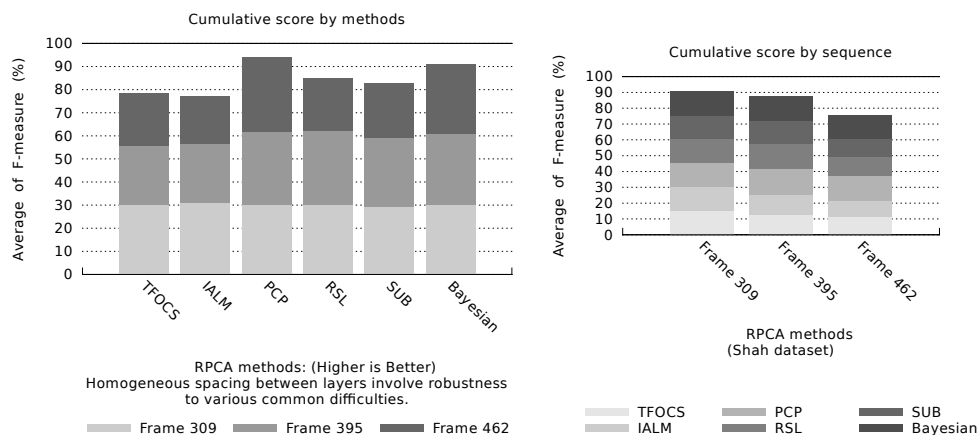


Fig. 2. Performance on the Shah dataset. The left (resp. right) figure concern the cumulative score by method (resp. sequence).

3.2 Shah's dataset ⁵

This sequence involved a camera mounted on a tall tripod and comes from Sheikh & Shah (2005). It contains 500 images and the corresponding GT. The wind caused the tripod to sway back and forth causing nominal motion in the scene. Table 3 shows the results obtained by the different algorithms on three images of the sequence: Frame 309 that contains a walking person, frame 395 when a car arrived the scene and frame 462 when the same car left the scene. For each frame, the first column shows the original image and the corresponding ground truth. The second part presents the sparse matrix in the first row and the optimal foreground mask in the second row. The detection rate (DR), Precision (Prec) and F-measure (F) are indicated below each foreground mask. Fig. 2 shows two cumulative histograms of F-measure: as in previous performance evaluation. PCP gives the best result followed by Bayesian RPCA, RSL, TFOCS and IALM. We can notice that the Bayesian give better performance on this dataset because none of moving object are bigger than the background area.

3.3 Li's dataset ⁶

This dataset provided by Li et al. (2004) consists of nine video sequences, which each sequence presenting dynamic backgrounds or illumination changes. The size of the images is 176*144 pixels. Among this dataset, we have chosen seven sequences that are the following ones:

- **Campus:** Persons walk and vehicles pass on a road in front of waving trees. This sequence contains 1439 images.
- **Water Surface:** A person arrives in front of the sea. There are many waves. This sequence contains 633 images.

⁵ http://www.cs.cmu.edu/~yaser/new_backgroundsubtraction.htm

⁶ http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html

- **Curtain:** A person presents a course in a meeting room with a moving curtain. This sequence contains 23893 images.
- **Escalator:** This image sequence shows a busy hall where an escalator is used by people. This sequence contains 4787 images.
- **Airport:** This image sequence shows a busy hall of an airport and each frame contains people. This sequence contains 3584 images.
- **Shopping Mall:** This image sequence shows a busy shopping center and each frame contains people. This sequence contains 1286 images.
- **Restaurant:** This sequence comes from the wallflower dataset and shows a busy cafeteria. This video contains 3055 images.

The sequences Campus, Water Surface and Curtain present dynamic backgrounds whereas the sequences Restaurant, Airport, Shopping Mall show bootstrapping issues. For each sequence, the ground truth is provided for twenty images when algorithms have to show their robustness. Table 4 shows the results obtained by the different algorithms on the sequence campus. Table 5 presents the results on the dynamic background on the sequences Water Surface, Curtain, Escalator, whereas table 6 presents the result on bootstrapping issues on the sequences Airport, Shopping mall, Restaurant. For each table, the first column shows the original image and the corresponding ground truth. The second part presents the sparse matrix in the first row and the optimal foreground mask in the second row. The detection rate (DR), Precision (Prec) and F-measure (F) are indicated below each foreground mask.

Fig. 3 and 4 shows the two cumulative histograms of F-measure respectively for dynamic background and bootstrapping issues. In each case, Bayesian RPCA gives best results followed by PCP, TFOCS, IALM and RSL.

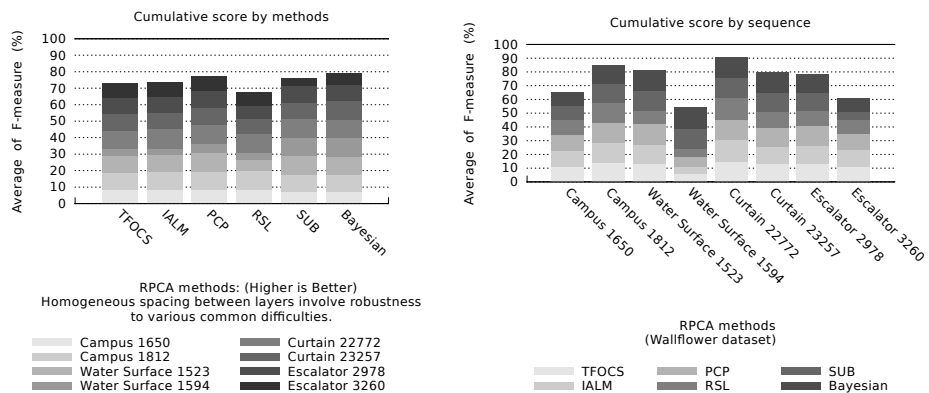


Fig. 3. Performance on the Li dataset (Dynamic backgrounds).

3.4 Implementation and time issues

Regarding the code, we have used the following implementation in *MATLAB*: RSL provided by F. De La Torre⁷, PCP provided by C. Qiu⁸, IALM provided by M. Chen and A. Ganesh⁹,



















































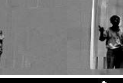















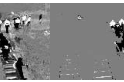

















Original / GT	TFOCS	IALM	PCP	RSL	SUB	Bayesian
  DR / Prec F	  72.9 / 89.5 80.38	  72.5 / 96.3 82.77	  86.8 / 93.0 89.84	  51.1 / 61.4 55.84	  85.8 / 94.6 90.02	  85.4 / 93.2 89.15
  DR / Prec F	  70.2 / 24.5 36.39	  34.9 / 28.4 31.34	  42.4 / 47.9 45.00	  26.6 / 48.5 34.42	  78.6 / 99.1 87.68	  88.6 / 94.9 91.66
  DR / Prec F	  90.0 / 89.7 89.89	  90.3 / 95.8 93.01	  86.2 / 96.2 90.96	  85.6 / 97.2 91.04	  85.4 / 97.5 91.08	  89.5 / 91.1 90.34
  DR / Prec F	  76.2 / 79.7 77.98	  69.8 / 84.8 76.59	  81.1 / 84.8 82.95	  78.2 / 64.5 70.73	  86.3 / 74.9 80.22	  83.5 / 90.9 87.09
  DR / Prec F	  75.1 / 84.2 79.45	  80.7 / 80.3 80.54	  80.3 / 87.5 83.76	  86.5 / 53.7 66.31	  74.6 / 90.1 81.64	  72.3 / 86.4 78.77
  DR / Prec F	  65.8 / 70.4 68.09	  72.1 / 68.7 70.40	  69.6 / 69.0 69.35	  73.8 / 58.0 64.96	  36.2 / 35.5 35.91	  52.1 / 64.1 57.54

Table 5. **Li dataset:** From left to right: Ground Truth, TFOCS, IALM, PCP, RSL, SUB, Bayesian RPCA. From top to bottom: water surface (1523), water surface (1594), curtain (22772), curtain (23257), escalator (2978), escalator (3260)

Original / GT	TFOCS	IALM	PCP	RSL	SUB	Bayesian
 DR / Prec F	 68.1 / 80.3 73.75	 69.3 / 79.9 74.26	 70.7 / 86.2 77.74	 66.8 / 63.7 65.26	 76.1 / 76.4 76.25	 71.9 / 86.3 78.48
 DR / Prec F	 66.5 / 75.1 70.60	 66.3 / 76.2 70.93	 64.7 / 70.8 67.63	 67.7 / 52.3 59.07	 63.9 / 59.0 61.38	 64.9 / 68.7 66.83
 DR / Prec F	 57.9 / 91.3 70.94	 46.9 / 94.6 62.74	 80.6 / 84.6 82.60	 55.3 / 88.7 68.17	 78.7 / 80.3 79.55	 80.1 / 81.7 80.94
 DR / Prec F	 49.1 / 73.6 58.96	 36.2 / 85.5 50.89	 69.3 / 84.1 76.07	 65.6 / 53.0 58.64	 66.7 / 86.4 75.31	 68.5 / 83.0 75.13
 DR / Prec F	 83.4 / 80.8 82.21	 77.4 / 91.8 83.88	 80.2 / 90.9 85.25	 71.1 / 61.0 65.70	 68.3 / 98.2 80.58	 82.1 / 88.5 85.23
 DR / Prec F	 74.2 / 70.1 72.17	 72.0 / 75.5 73.73	 67.2 / 83.3 74.4	 69.9 / 68.8 69.38	 67.6 / 76.0 71.58	 83.4 / 65.6 73.52

Table 6. (Perception) Li dataset: From left to right: Ground Truth, TFOCS, IALM, PCP, RSL, SUB, Bayesian RPCA. From top to bottom: Airport (2926), Airport (3409), shopping mall (1862), shopping mall (1980) Restaurant (1842), Restaurant (2832).

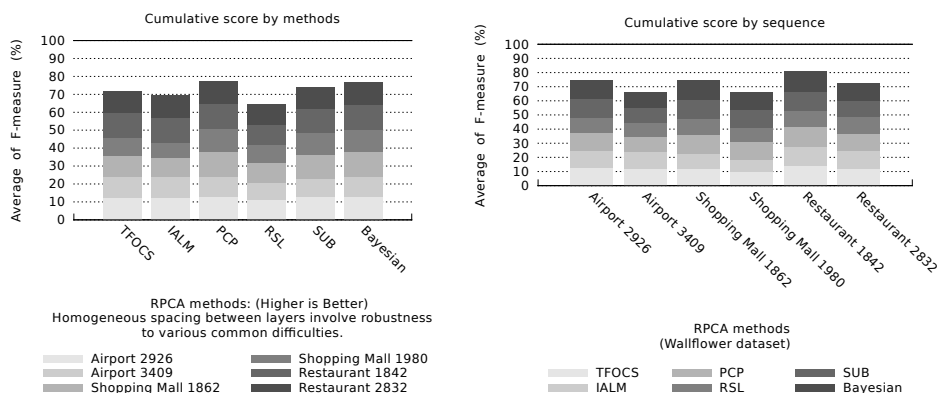


Fig. 4. Performance on the Li dataset (Bootstrap issues).

TFOCS provided by S. Becker¹⁰ and Bayesian provided by X. Ding¹¹. Additionally, a 5×5 median filter is postprocessed in order to suppress peak noise. The thresholding value is automatically choose for maximize the F-measure.

For time issues, the current implementations are faraway to achieve real-time. Indeed, the computing of the backgrounds take few hours for a training sequence with 200 frames for each algorithm. This time can be reduced by *C/Cuda* implementation as suggested in (Mu et al. (2011);Anderson et al. (2011)).

4. Conclusion

In this chapter, we started off by providing a survey on five Robust Principal Component Analysis models recently developed: Robust Subspace Learning, Principal Component Pursuit, Templates for First-Order Conic Solvers, Recursive Projected Compressive Sensing, Bayesian RPCA. We then presented a systematic evaluation and comparative analysis on different datasets used in video-surveillance. PCP demonstrates more robustness on all datasets by providing best global score. The Bayesian RPCA offers also good performance but presents a drawback related to the assumption : the background has necessarily a bigger area than the foreground. For the IALM, its performance is still acceptable.

Futur research directions may concern the evaluation of accelerate hardware implementation of robust PCA (Mu et al. (2011);Anderson et al. (2011)) and robust Independent Components Analysis (Yamazaki et al. (2006)), Incremental Non-negative Matrix Factorization (Bucak et al. (2007)) and Incremental Rank Tensor (Li et al. (2008)).

⁷ <http://www.salleurl.edu/~ftorre/papers/rpca/rpca.zip>

⁸ http://home.engineering.iastate.edu/~chenlu/ReProCS/ReProCS_code.zip

⁹ http://perception.csl.uiuc.edu/matrix-rank/Files/inexact_alm_rpca.zip

¹⁰ <http://tfocs.stanford.edu/code>

¹¹ http://www.ece.duke.edu/~lihan/brpca_code/BRPCA.zip

5. References

- Anderson, M., Ballard, G., Demme, J. & Keutzer, K. (2011). Communication-avoiding qr decomposition for gpu, *IEEE International Parallel and Distributed Processing Symposium, IPDPS 2011*.
- Becker, S., Candes, E. & Grant, M. (2011). Tfocs: Flexible first-order methods for rank minimization, *Low-rank Matrix Optimization Symposium, SIAM Conf. on Optimization*.
- Black, M. & Rangarajan, A. (1996). On the unification of line processes, outlier rejection, and robust statistics with applications in early vision., *International Journal of Computer Vision* pp. 57–92.
- Bouwman, T. (2009). Subspace learning for background modeling: A survey, *Recent Patents on Computer Science* 2(3): 223–234.
- Bouwman, T., Baf, F. E. & Vachon, B. (2008). Background modeling using mixture of gaussians for foreground detection - a survey, *Recent Patents on Computer Science* 1(3): 219–237.
- Bouwman, T., Baf, F. E. & Vachon, B. (2010). Statistical background modeling for foreground detection: A survey, *Handbook of Pattern Recognition and Computer Vision, World Scientific Publishing* 4(2): 181–189.
- Bucak, S., Gunsul, B. & Gursoy, O. (2007). Incremental non-negative matrix factorization for dynamic background modeling, *International Workshop on Pattern Recognition in Information Systems*.
- Candes, E., Li, X., Ma, Y. & Wright, J. (2009). Robust principal component analysis?, *Preprint*.
- Ding, X., He, L. & Carin, L. (2011). Bayesian robust principal component analysis, *IEEE Transaction on Image Processing*.
- Eckart, C. & Young, G. (1936). The approximation of one matrix by another of lower rank, *Psychometrika* 1: 211–218.
- Li, L., Huang, W., Gu, I. & Tian, Q. (2004). Statistical modeling of complex backgrounds for foreground object detection, *IEEE Transaction on Image Processing* pp. 1459–1472.
- Li, X., Hu, W., Zhang, Z. & Zhang, X. (2008). Robust foreground segmentation based on two effective background models, *Multimedia Information Retrieval (MIR)* pp. 223–228.
- Lin, Z., Chen, M., Wu, L. & Ma, Y. (2009). The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices, *UIUC Technical Report*.
- Maddalena, L. & Petrosino, A. (2010). A fuzzy spatial coherence-based approach to background foreground separation for moving object detection, *Neural Computing and Applications, NCA* pp. 1–8.
- Mu, Y., Dong, J., Yuan, X. & Yan, S. (2011). Accelerated low-rank visual recovery by random projection, *International Conference on Computer Vision, CVPR 2011*.
- Oliver, N., Rosario, B. & Pentland, A. (1999). A bayesian computer vision system for modeling human interactions, *International Conference on Vision Systems, ICVS 1998*.
- Sheikh, Y. & Shah, M. (2005). Bayesian modeling of dynamic scenes for object detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27: 1778–1792.
- Torre, F. D. L. & Black, M. (2001). Robust principal component analysis for computer vision, *International Conference on Computer Vision, ICCV 2001*.
- Torre, F. D. L. & Black, M. (2003). A framework for robust subspace learning, *International Journal on Computer Vision* pp. 117–142.
- Toyama, K., Krumm, J., Brumitt, B. & Meyers, B. (1999). Wallflower: Principles and practice of background maintenance, *International Conference on Computer Vision* pp. 255–261.

- Wright, J., Peng, Y., Ma, Y., Ganesh, A. & Rao, S. (2009). Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization, *Neural Information Processing Systems, NIPS 2009* .
- Yamazaki, M., Xu, G. & Chen, Y. (2006). Detection of moving objects by independent component analysis, *Asian Conference on Computer Vision (ACCV) 2006* pp. 467–478.



Principal Component Analysis

Edited by Dr. Parinya Sanguansat

ISBN 978-953-51-0195-6

Hard cover, 300 pages

Publisher InTech

Published online 02, March, 2012

Published in print edition March, 2012

This book is aimed at raising awareness of researchers, scientists and engineers on the benefits of Principal Component Analysis (PCA) in data analysis. In this book, the reader will find the applications of PCA in fields such as image processing, biometric, face recognition and speech processing. It also includes the core concepts and the state-of-the-art methods in data analysis and feature extraction.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Charles Guyon, Thierry Bouwmans and El-hadi Zahzah (2012). Robust Principal Component Analysis for Background Subtraction: Systematic Evaluation and Comparative Analysis, Principal Component Analysis, Dr. Parinya Sanguansat (Ed.), ISBN: 978-953-51-0195-6, InTech, Available from:

<http://www.intechopen.com/books/principal-component-analysis/robust-principal-component-analysis-for-background-subtraction-systematic-evaluation-and-comparative>

INTeCH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2012 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.