



HAL
open science

Foreground Detection via Robust Low Rank Matrix Decomposition including Spatio-Temporal Constraint

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah

► **To cite this version:**

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah. Foreground Detection via Robust Low Rank Matrix Decomposition including Spatio-Temporal Constraint. International Workshop on Background Model Challenges, ACCV 2012,, Nov 2012, Daejeon, South Korea. pp.315-320. hal-00809473

HAL Id: hal-00809473

<https://hal.science/hal-00809473>

Submitted on 19 Nov 2015

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 via Robust Low Rank Matrix Decomposition including spatio-temporal Constraint

Charles Guyon, Thierry Bouwmans, El-Hadi Zahzah

Laboratoire MIA (Mathematiques, Image et Applications) - University of La Rochelle
thierry.bouwmans@univ-lr.fr

Abstract. Foreground detection is the first step in video surveillance system to detect moving objects. 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. In this paper, we propose to use a low-rank matrix factorization with IRLS scheme (Iteratively reweighted least squares) and to address in the minimization process the spatial connexity and the temporal sparseness of moving objects (e.g. outliers). Experimental results on the BMC 2012 datasets show the pertinence of the proposed approach.

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". Recent research on robust PCA shows qualitative visual results with the background variations appromatively lying in a low dimension subspace, and the sparse part being the moving objects. First, Candes et al. [1] proposed a convex optimization problem 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:

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

where $\|\cdot\|_*$ and $\|\cdot\|_1$ are the nuclear norm (which is the L_1 norm of singular values) and l_1 norm, respectively, and $\lambda > 0$ is an arbitrary balanced parameter. Under these minimal assumptions, this approach called Principal Component Pursuit (PCP) solution perfectly recovers the low-rank and the sparse matrices.

In this paper, we propose a robust low-matrix factorization with IRLS scheme to adress the second limitation. For a data matrix A containing the sequence, we

assume that a part is approximatively low-rank and product of two matrices, and a small part of this matrix is corrupted by the outliers. Furthermore, we directly introduced a spatial term in the l_1 minimization to address the spatial connexity of the pixels. So, our contributions can be summarized as follows: 1) Addition of spatial constraint to minimization process, 2) IRLS alternating scheme for weighted the 2-parameters $\|\cdot\|_{\alpha,\beta}$ for matrix low-rank decomposition. The rest of this paper is organized as follows. The Section 2 focus on IRLS method applied on vector regression problems. In Section 3, we present a robust low-rank matrix factorization which allows us to detect foreground objects in dynamic backgrounds. In Section 4, we present results on the BMC 2012 datasets¹ and the Section 5 provides the conclusion.

2 L_p Minimization with spatial constraint

In most applications, video surveillance data is assumed to be compose of background, foreground and noise. Regression task is a crucial part of the proposed decomposition algorithm. We consider the following minimization problem (2), where A is a dictionary matrix (row order) and b is a row vector, the second term forces the error E to be a connexe shape, through the TV (Total Variation) of the residual must be small, where the matrix ∇_s is a spatial gradient.

$$\operatorname{argmin}_x \|Ax - b\|_{\alpha} + \lambda \|\nabla_s(Ax - b)\|_1 \quad (2)$$

The left part of the problem (3) is convex for $\alpha > 1$ and the usual IRLS (Iteratively reweighted least squares) scheme for solve $\operatorname{argmin}_x \|Ax - b\|_{\alpha}$ is given by

$$\begin{cases} D^{(i)} = \operatorname{diag}((\varepsilon + |b - Ax^{(i)}|)^{\alpha-2}) \\ x^{(i+1)} = (A^t D^{(i)} A)^{-1} A^t D^{(i)} b \end{cases} \quad (3)$$

It was proven that a suitable IRLS method is convergent for $1 \leq \alpha < 3$ [2]. Since if the process is expressed with a residual formulation, we gain more numericaly stability and let us to choose freely $\alpha \in [1, \infty[$ with an adapted step size λ_{opt} on every iteration.

$$\begin{cases} r^{(i)} = b - Ax^{(i)} \\ D = \operatorname{diag}((\varepsilon + |r^{(i)}|)^{\alpha-2}) \\ y^{(i)} = (A' D A)^{-1} A' D r^{(i)} \\ x^{(i+1)} = x^{(i)} + (1 + \lambda_{opt}) y^{(i)} \end{cases} \quad (4)$$

With a fixed λ_{opt} , we should choose λ_{opt} as developed [3].

¹ <http://bmc.univ-bpclermont.fr>

Otherwise, the algorithm is twice iterative, where we try to get an optimal x and an optimal λ at each step.

$$\left\{ \begin{array}{l} c^{(i)} = Ay^{(i)} \\ d^{(i)} = b - A(x^{(i)} + y^{(i)}) \\ \operatorname{argmin}_{\lambda_{opt}} \|c^{(i)}\lambda - d^{(i)}\|_{\alpha} \end{array} \right. \left\{ \begin{array}{l} \lambda^{(0)} = \Lambda(\alpha) \\ s^{(k)} = d - \lambda^{(k)}c \\ E = \operatorname{diag}((\varepsilon + |s^{(k)}|)^{\alpha-2}) \\ z^{(k)} = \frac{c^t E s^{(k)}}{c^t E c} \\ \lambda^{(k+1)} = \lambda^{(k)} + (1 + \Lambda(\alpha))z^{(k)} \end{array} \right. \quad (5)$$

Only few iterations (≈ 10) is enough for acceptable approximation of λ_{opt} of the $\lambda^{(k)}$ sequence. Moreover, the convergence is usually improved by a Aitken process or an other acceleration technique. Note for case $\alpha > 2$, convergence is achieved when $0 < 1 + \lambda < \frac{2}{\alpha-1}$. Additionally, TV is particular case of the following problem:

$$\operatorname{argmin}_x \|Ax - b\|_{\alpha} + \lambda \|Cx - d\|_{\beta} \quad (6)$$

By derivation, the associated IRLS scheme is,

$$\left\{ \begin{array}{l} r_1 = b - Ax^{(i)}, r_2 = d - Cx^{(i)}, e_1 = \varepsilon + |r_1|, e_2 = \varepsilon + |r_2| \\ D_1 = (\sum e_1^{\alpha})^{\frac{1}{\alpha}-1} \operatorname{diag}(e_1^{\alpha-2}), D_2 = \lambda (\sum e_2^{\beta})^{\frac{1}{\beta}-1} \operatorname{diag}(e_2^{\beta-2}) \\ y^{(i)} = (A^t D_1 A + C^t D_2 C)^{-1} (A^t D_1 r_1 + C^t D_2 r_2) \\ x^{(i+1)} = x^{(i)} + (1 + \lambda_{opt}) y^{(i)} \end{array} \right. \quad (7)$$

More generally, we consider the following matrix regression problem with two parameters norm (α, β) and a weighted matrix (W) ,

$$\min_X \|AX - B\|_{\alpha, \beta}^W \quad \text{with} \quad \|M_{ij}\|_{\alpha, \beta}^W = \left(\sum_{i=1}^n \left(\sum_{j=1}^m W_{ij} |M_{ij}|^{\beta} \right)^{\frac{\alpha}{\beta}} \right)^{\frac{1}{\alpha}} \quad (8)$$

The problem is solved in the same manner on matrices with a reweighted regression strategy,

$$\left\{ \begin{array}{l} \text{Until X is stable, repeat on each } k\text{-columns} \\ R \leftarrow B - AX \\ S \leftarrow \varepsilon + |R| \\ D_k \leftarrow \operatorname{diag}(S_{ik}^{\beta-2} \circ (\sum_j (S_{ij}^{\beta} \circ W_{ij}))^{\frac{\alpha}{\beta}-1} \circ W_{ik})_k \\ X_{ik} \leftarrow X_{ik} + (1 + \Lambda(\max(\alpha, \beta))) (A^t D_k A)^{-1} A^t D_k R_{ik} \end{array} \right. \quad (9)$$

3 Foreground Detection via Robust Low-Rank Matrix Factorization and temporal constraint

The training video sequence $A \in \mathbb{R}^{n \times m}$ is stored as a matrix with a particular structure. Columns are spatial frames and rows are values of a fixed pixel over time. For $A = \{I_1, \dots, I_m\}$, I_j denotes a vectorized frame of n pixels at j -time with m is the number of frames. $A_{y+hx,t}$ implies the pixel intensity at coordinate

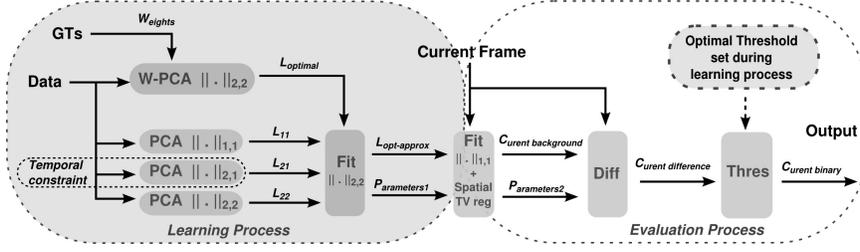


Fig. 1. Overview of the learning and evaluation process. Learning process needs GT (Groundtruth) for better fits the eigenbackground components.

x, y, t . The background modeling process finds an ideal subspace of the video sequence, which describes the best as possible the (dynamic) background as shown in Fig. 2. Then, the decomposition involves the following model:

$$A = L + S = BC + S \quad (10)$$

where B is a *low-rank* matrix corresponding to the background model plus noise and C allows to approximate L by linear combination. S is a *sparse* matrix which corresponds to the foreground component obtained by subtraction.

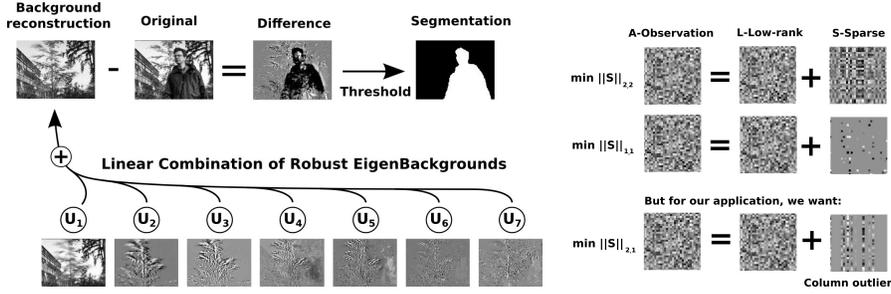


Fig. 2. At left: The common process of background subtraction via PCA (Principal Component Analysis). At final step, an adaptative threshold is used to get a binary result.

Fig. 3. At right: Using the previous decomposition on a low-rank random matrix plus noise, different kind of pattern on residual matrix emerge with the choice of the norm.

The model involves the error reconstruction determined by the following constraints:

$$\min_{B \in \mathbb{R}^{n \times p}, C \in \mathbb{R}^{p \times m}} \|(A - BC) \circ W\|_{\alpha, \beta} + \mu \|BC\|_* \quad (11)$$

where $\|\cdot\|_*$ denote the nuclear norm. The decomposition is split into two parts. Firstly, we track 1-Rank decomposition since the first eigen-vector is strongly

dominant in video surveillance.

$$\begin{aligned} R_1 &= A - B_1 C_1 \\ R &= A - B_1 C_1 - B_r C_r \end{aligned} \quad \left| \quad \begin{aligned} \min_{B_1, C_1} \|R_1\|_{1,1} \\ \min_{B_r, C_r} \|R \circ \phi(R_1)\|_{2,1 \rightarrow 0} \end{aligned} \right. \quad (12)$$

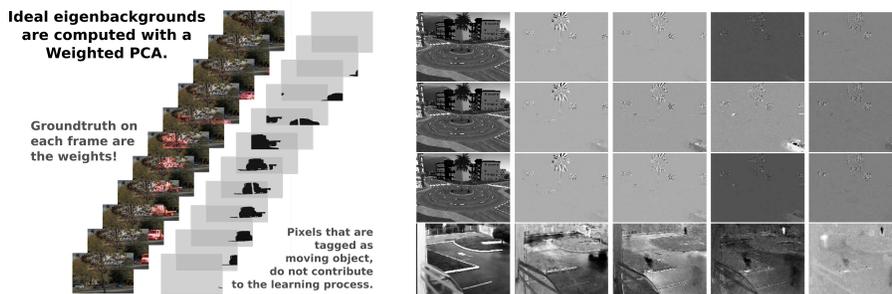


Fig. 4. At Left: Schema of ideal PCA processing. The eigenbackground are computed using a Weighted-PCA with GroundTruth.

Fig. 5. At right: First eigenBackground of the fifths sequence of Rotary (BMC) with the norms $\|\cdot\|_{opt}$, $\|\cdot\|_{1,1}$ and $\|\cdot\|_{2,1}$. Last row shows eigenBackground on real dataset with $\|\cdot\|_{2,1}$.

We use $\|\cdot\|_{2,1 \rightarrow 0}$ instead of usual $\|\cdot\|_{1,1}$ because it forces spatial homogeneous fitting. Besides $\beta = (1 \rightarrow 0)$ means the β parameter decreases during iteration. First, we search a solution of the convex problem $\|\cdot\|_{2,1}$, then use the solution as an initial guess for non-convex problem $\|\cdot\|_{2,(1-\varepsilon)}$. Finally, we find a local minimum of $\|\cdot\|_{2,0}$ and hope that is near of the global minimum of this problem. Furthermore, this norm enforce temporarily sparseness of outliers as shown in Fig.4. In the case where $\alpha = \beta = 2$, the decomposition is usually solved by a SVD (Singular Value Decomposition). Thus, our SVD algorithm can be seen as an iterative regression. The proposed scheme determines alternatively the optimal coefficients, it means searching C for B fixed and searching B for C fixed.

$$\begin{aligned} C^{(k+1)} &= (A^t A)^{-1} A^t B^{(k)} \\ \bar{C}^{(k+1)} &= C^{(k+1)} \sqrt{C^{t(k+1)} C^{(k+1)}}^{-1} \\ B^{(k+1)} &= (A^t A)^{-1} A^t \bar{C}^{(k+1)} \end{aligned} \quad (13)$$

Additionally, this alternating regression framework allows to associate a weighed matrix W which is entrywise multiplied to the error term,

$$\min_{B,C} \|(A - BC) \circ W\|_{\alpha,\beta} \quad (14)$$

The W mask is iteratively computed and aims to enforce the fit exclusively on guessed background region.

We define a function ϕ that have two goals, smooth the error (like spatial median filtering) and transform the error for obtain a suitable weighted mask for regression.

$$W = \phi(|A - BC|) \quad , \quad \phi(E) = e^{-\gamma TV(E)} \quad (15)$$

By including local penalty as a constraint in RPCA, this explicitly increases local coherence of the sparse component as filled/plain shapes (therefore moving object).

4 Experimental Results

Here, we show experimental results on the real dataset of BMC,

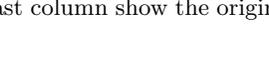
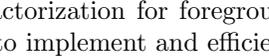
Video	Recall	Precision	F-measure	PSNR	Visual Results
1	0.9139	0.7170	0.8036	38.2425	
2	0.8785	0.8656	0.8720	26.7721	
3	0.9658	0.8120	0.8822	37.7053	
4	0.9550	0.7187	0.8202	39.3699	
5	0.9102	0.5589	0.6925	30.5876	
6	0.9002	0.7727	0.8316	29.9994	
7	0.9116	0.8401	0.8744	26.8350	
8	0.8651	0.6710	0.7558	30.5040	
9	0.9309	0.8239	0.8741	55.1163	

Table 1. Quantitative results with common criterions. Last column show the original, GT and result of the first four real video sequences.

5 Conclusion

In this paper, we have presented a robust matrix factorization for foreground detection. This method is conceptually simple, easy to implement and efficient. Furthermore, experiments on video surveillance datasets show that this approach is more robust than recent RPCA approaches in the presence of dynamic backgrounds and illumination changes. Further research consists in developping an incremental version to update the model at every frame and to achieve real-time requirements.

References

1. Candes, E., Li, X., Ma, Y., Wright, J.: Robust principal component analysis? International Journal of ACM **58** (2011)
2. Osborne, M.R.: Finite algorithms in optimization and data analysis. John Wiley & Sons (1985)
3. Guyon, C., Bouwmans, T., Zahzah, E.: Foreground detection via robust low rank matrix factorization including spatial constraint with iterative reweighted regression. International Conference on Pattern Recognition, ICPR 2012 (2012)