



HAL
open science

Comparison of Matrix Completion Algorithms for Background Initialization in Videos

Andrews Sobral, Thierry Bouwmans, El-Hadi Zahzah

► **To cite this version:**

Andrews Sobral, Thierry Bouwmans, El-Hadi Zahzah. Comparison of Matrix Completion Algorithms for Background Initialization in Videos. SBMI 2015 Workshop in conjunction with ICIAP 2015, Sep 2015, Genoa, Italy. pp.510-518, 10.1007/978-3-319-23222-5_62 . hal-01227959

HAL Id: hal-01227959

<https://hal.science/hal-01227959v1>

Submitted on 12 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.

Comparison of Matrix Completion Algorithms for Background Initialization in Videos

Andrews Sobral^{1,2} and Thierry Bouwmans² El-hadi Zahzah¹

¹ Université de La Rochelle, Lab. L3I, 17000, La Rochelle, France

² Université de La Rochelle, Lab. MIA, 17000, La Rochelle, France

Abstract. Background model initialization is commonly the first step of the background subtraction process. In practice, several challenges appear and perturb this process such as dynamic background, bootstrapping, illumination changes, noise image, etc. In this context, this work aims to investigate the background model initialization as a matrix completion problem. Thus, we consider the image sequence (or video) as a partially observed matrix. First, a simple joint motion-detection and frame-selection operation is done. The redundant frames are eliminated, and the moving regions are represented by zeros in our observation matrix. The second stage involves evaluating nine popular matrix completion algorithms with the Scene Background Initialization (SBI) data set, and analyze them with respect to the background model challenges. The experimental results shows the good performance of LRGeomCG [17] method over its direct competitors.

Keywords: matrix completion, background modeling, background initialization

1 Introduction

Background subtraction (BS) is an important step in many computer vision systems to detect moving objects. This basic operation consists of separating the moving objects called “foreground” from the static information called “background” [2, 16]. The BS is commonly used in video surveillance applications to detect persons, vehicles, animals, etc., before operating more complex processes for intrusion detection, tracking, people counting, etc. Typically the BS process includes the following steps: a) background model initialization, b) background model maintenance and c) foreground detection. With a focus on the step (a), the BS initialization consists in creating a background model. In a simple way, this can be done by setting manually a static image that represents the background. The main reason is that it is often assumed that initialization can be achieved by exploiting some clean frames at the beginning of the sequence. Naturally, this assumption is rarely encountered in real-life scenarios, because of continuous clutter presence. In addition, this procedure presents several limitations, because it needs a fixed camera with constant illumination, and the background needs to be static (commonly in indoor environments), and having no moving object in the first frames. In practice, several challenges appear and perturb this process such as noise acquisition, bootstrapping, dynamic factors, etc [11].

The main challenge is to obtain a first background model when more than half of the video frames contain foreground objects. Some authors suggest the initialization of the

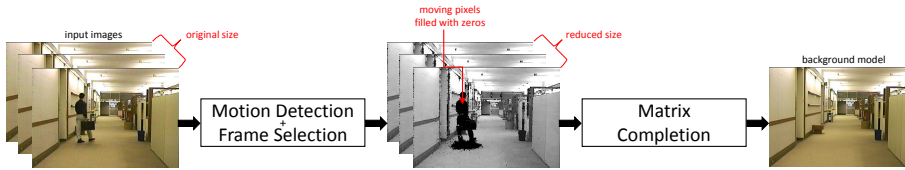


Fig. 1. Block diagram of the proposed approach. Given an input image, a joint motion detection and frame selection operation is applied. Next, a matrix completion algorithm tries to recover the background model from the partially observed matrix. In this paper, the processes described here are conducted in a batch manner.

background model by the arithmetic mean [9] (or weighted mean) of the pixels between successive images. Practically, some algorithms are: (1) batch ones using N training frames (consecutive or not), (2) incremental with known N or (3) progressive ones with unknown N as the process generates partial backgrounds and continues until a complete background image is obtained. Furthermore, initialization algorithms depend on the number of modes and the complexity of their background models. However, BS initialization has also been achieved by many other methodologies [11, 2]. We can cite for example the computation of eigen values and eigen vectors [15], and the recent research on subspace estimation by sparse representation and rank minimization [3]. The background model is recovered by the low-rank subspace that can gradually change over time, while the moving foreground objects constitute the correlated sparse outliers.

In this paper, the initialization of the background model is addressed as a matrix completion problem. The matrix completion aims at recovering a low rank matrix from partial observations of its entries. The image sequence (or video) is represented as a partially observed real-valued matrix. Figure 1 shows the proposed framework. First, a simple joint motion-detection and frame-selection operation is done. The redundant frames are eliminated, and the moving regions are represented with zeros in our observation matrix. This operation is described in the Section 2. The second stage involves evaluating nine popular matrix completion algorithms with the Scene Background Initialization (SBI) data set [12] (see Section 3). This enables to analyze them with respect to the background model challenges. Finally, in Sections 4 and 5, the experimental results are shown as well as conclusions.

Throughout the paper, we use the following notations. Scalars are denoted by lowercase letters, e.g., x ; vectors are denoted by lowercase boldface letters, e.g., \mathbf{x} ; matrices by uppercase boldface, e.g., \mathbf{X} . In this paper, only real-valued data are considered.

2 Joint Motion Detection and Frame Selection

In order to reduce the number of redundant frames, a simple joint motion detection and frame selection operation is applied. First, the color images are converted into its gray-scale representation. So, let a sequence of N gray-scale images (frames) $\mathbf{I}_0 \dots \mathbf{I}_N$ captured from a static camera, that is, $\mathbf{I} \in \mathbb{R}^{m \times n}$ where m and n denotes the frame resolution (rows by columns). The difference between two consecutive frames (motion detection step) is calculated by:

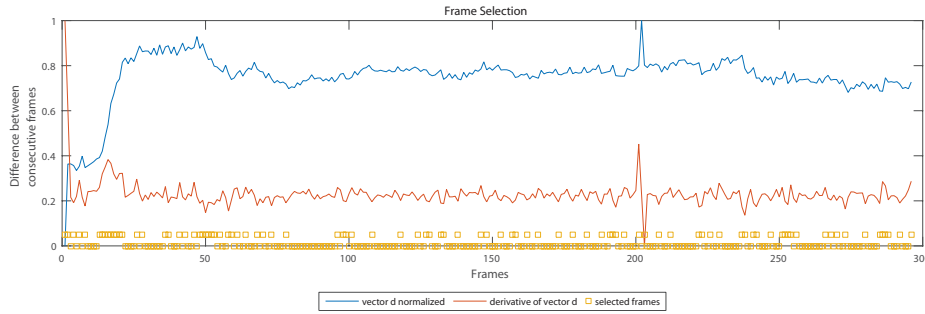


Fig. 2. Illustration of frame selection operation. The normalized vector (in blue) shows the difference between two consecutive frames. The derivative vector draw how much the normalized vector changes (in red), and then it is thresholded and the frames are selected (in orange).

$$\mathbf{D}_t = \sqrt{(\mathbf{I}_t - \mathbf{I}_{t-1})^2} \mid_{t=1, \dots, N}, \quad (1)$$

where $\mathbf{D}_t \in \mathbb{R}^{m \times n}$ denotes the matrix of pixel-wise L_2 -norm differences from frame $t - 1$ to frame t . Next, the sum of all elements of \mathbf{D}_t , for $t = 1, \dots, N$, is stored in a vector $\mathbf{d} \in \mathbb{R}^N$ whose t -th element is given by:

$$d_t = \sum_{i=1}^m \sum_{j=1}^n \mathbf{D}_t(i, j), \quad (2)$$

where $\mathbf{D}_t(i, j)$ is the matrix element located in the row $i \in [1, \dots, m]$ and column $j \in [1, \dots, n]$. Then, the vector \mathbf{d} is normalized between 0 and 1 by:

$$\hat{\mathbf{d}} = \frac{d_t - d_{min}}{d_{max} - d_{min}} \mid_{t=1, \dots, N}, \quad (3)$$

where d_{min} and d_{max} denotes the minimum value and the maximum value of the vector \mathbf{d} . The frame selection step is done by calculating the derivative of $\hat{\mathbf{d}}$ by:

$$\mathbf{d}' = \frac{d}{dt} \hat{\mathbf{d}}, \quad (4)$$

Next, the vector \mathbf{d}' is also normalized by Equation 3 and represented by $\hat{\mathbf{d}}'$. Finally, the index of the more relevant frames are given by thresholding $\hat{\mathbf{d}}'$:

$$\mathbf{y} = \begin{cases} 1 & \text{if } |\hat{\mathbf{d}}' - \hat{\mu}'| > \tau \\ 0 & \text{otherwise} \end{cases}, \quad (5)$$

where $\hat{\mu}'$ denotes the mean value of the vector $\hat{\mathbf{d}}'$, and $\tau \in [0, \dots, 1]$ controls the threshold operator. In this paper, $R \leq N$ represent the set of all frames where $\mathbf{y} = 1$, and the parameter τ was chosen experimentally for each scene: $\tau = 0.025$ for HallAndMonitor, $\tau = 0.05$ for HighwayII, $\tau = 0.10$ for HighwayI, and $\tau = 0.15$ to all other scenes. Figure 1 illustrates our frame selection operation, in this example, with $\tau = 0.025$, only 92 relevant frames are selected from a total of 296 frames (68, 92% of reduction). In the next section, the matrix completion process is described.

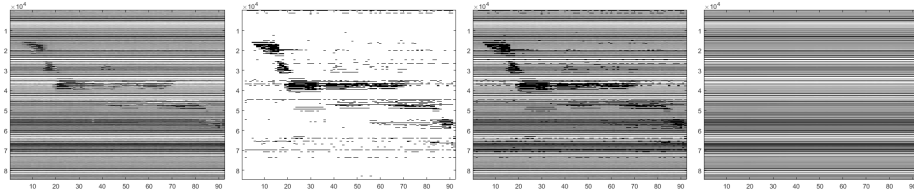


Fig. 3. Illustration of the matrix completion process. From the left to the right: a) the selected frames in vectorized form (our observation matrix), b) the moving regions are represented by non-observed entries (black pixels), c) the moving regions filled with zeros (modified version of the observation matrix), and d) the recovered matrix after the matrix completion process.

3 Matrix Completion

As explained previously, the matrix completion aims to recover a low rank matrix from partial observations of its entries. Considering the general form of low rank matrix completion, the optimization problem is to find a matrix $\mathbf{L} \in \mathbb{R}^{n_1 \times n_2}$ with minimum rank that best approximates the matrix $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$. Candès and Recht [6] show that this problem can be formulated as:

$$\begin{aligned} & \text{minimize} && \text{rank}(\mathbf{A}), \\ & \text{subject to} && P_{\Omega}(\mathbf{A}) = P_{\Omega}(\mathbf{L}), \end{aligned} \quad (6)$$

where $\text{rank}(\mathbf{A})$ is equal to the rank of the matrix \mathbf{A} , and P_{Ω} denotes the sampling operator restricted to the elements of Ω (set of observed entries), i.e., $P_{\Omega}(\mathbf{A})$ has the same values as \mathbf{A} for the entries in Ω and zero values for the entries outside Ω . Later, Candès and Recht [6] propose to replace the $\text{rank}(\cdot)$ function with the nuclear norm $\|\mathbf{A}\|_* = \sum_{i=1}^r \sigma_i$ where $\sigma_1, \sigma_2, \dots, \sigma_r$ are the singular values of \mathbf{A} and r is the rank of \mathbf{A} . The nuclear norm make the problem tractable and Candès and Recht [6] have proved theoretically that the solution can be exactly recovered with a high probability. In addition, Cai et. al [4] propose an algorithm based on soft singular value thresholding (SVT) to solve this convex relaxation problem. However, in real world application the observed entries may be noisy. In order to make the Equation 6 robust to noise, Candès and Plan [5] propose a stable matrix completion approach. The equality constraint is replaced by $\|P_{\Omega}(\mathbf{A} - \mathbf{L})\|_F \leq \epsilon$, where $\|\cdot\|_F$ denotes the Frobenious norm and ϵ is an upper bound on the noise level. Recently, several matrix completion algorithms have been proposed to deal with this challenge, and a complete review can be found in [21].

In this paper, we address the background model initialization as a matrix completion problem. Once frame selection process is done, the moving regions of the R selected frames are determined by:

$$\mathbf{M}_k(i, j) = \begin{cases} 1 & \text{if } 0.5(\mathbf{D}_k(i, j))^2 > \beta \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where $k \in R$, and β is the thresholding parameter (in this paper, $\beta = 1e^{-3}$ for all experiments). Next, the moving regions of each selected frame are filled with zeros

Table 1. List of low-rank matrix completion algorithms evaluated in this paper.

Category	Method	Main techniques	Reference
Rank Minimization	IALM	Augmented Lagrangian	[10, Lin et al. (2010)]
	RMAMR	Augmented Lagrangian	[20, Ye et al. (2015)]
Matrix Factorization	SVP	Hard thresholding	[13, Meka et al. (2009)]
	OptSpace	Grassmannian	[8, Keshavan et al. (2010)]
	LMaFit	Alternating	[19, Wen et al. (2012)]
	ScGrassMC	Grassmannian	[14, Ngo and Saad (2012)]
	LRGeomCG	Riemannian	[17, Vandereycken (2013)]
	GROUSE	Online algorithm	[1, Balzano et al. (2013)]
	ORIMP	Matching pursuit	[18, Wang et al. (2015)]

by $\mathbf{I}_k \circ \overline{\mathbf{M}}_k$, where $\overline{\mathbf{M}}_k$ denotes the complement of \mathbf{M}_k , and \circ denotes the element-wise multiplication of two matrices. For color images, each channel is processed individually, then they are vectorized into a partially observed real-valued matrix $\mathbf{A} = [\text{vec}(I_1) \dots \text{vec}(I_k)]$, where $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2}$, $n_1 = (m \times n)$, and $n_2 = k$. Figure 3 illustrates our matrix completion process. It can be seen that the partially observed matrix can be recovered successfully even with the presence of many missing entries. So, let \mathbf{L} the recovered matrix from the matrix completion process, the background model is estimated by calculating the average value of each row, resulting in a vector $\mathbf{l} \in \mathbb{R}^{n_1 \times 1}$, and then reshaped into a matrix $\mathbf{B} \in \mathbb{R}^{m \times n}$.

4 Experimental results

In order to evaluate the proposed approach, nine matrix completion algorithms have been selected, and they are listed in Table 1. The algorithms were grouped in two categories, as well as its main techniques (following the same definition of Zhou et al. [21]).

In this paper, the Scene Background Initialization (SBI) data set was chosen for the background initialization task. The data set contains seven image sequences and corresponding ground truth backgrounds. It provides also MATLAB scripts for evaluating background initialization results in terms of eight metrics³. Figure 4 show the visual results for the top three best matrix completion algorithms, and Table 2 reports the quantitative results of each algorithm over the data set⁴. The algorithms are ranked as follow: 1) for each algorithm we calculate its rank position for each metric, we call it as *metric rank* (i.e. RMAMR have the first position for the AGE metric in the HallAndMonitor scene), next, 2) we sum the rank position value of each algorithm over the eight metrics, and finally, 3) we calculate the rank position over the sum, and we call it as *scene rank*. For the Global Rank, first we sum the scene rank for each MC algorithm, then

³ Please, refer to <http://sbmi2015.na.icar.cnr.it/> for a complete description of each metric.

⁴ Full experimental evaluation and related source code can be found in the main website: <https://sites.google.com/site/mc4bmi/>



Fig. 4. Visual comparison for the background model initialization. From top to bottom: 1) example of input frame, 2) background model ground truth, and background model results for the top 3 best ranked MC algorithms: 3) LRGeomCG, 4) LMaFit, and 5) RMAMR.

we calculate its rank position over the sum. As we can see, the experimental results show the good performance of LRGeomCG [17] method over its direct competitors. Furthermore, in most cases the matrix completion algorithms outperform the traditional approaches such as Mean [9], Median [7] and MoG [22] as can be seen in the full experimental evaluation available at <https://sites.google.com/site/mc4bmi/>.

5 Conclusion

In this paper, we have evaluated nine recent matrix completion algorithms for the background initialization problem. Given a sequence of images, the key idea is to eliminate the redundant frames, and consider its moving regions as non-observed values. This approach results in a matrix completion problem, and the background model can be recovered even with the presence of missing entries. The experimental results on the SBI data set shows the comparative evaluation of these recent methods, and highlights the good performance of LRGeomCG [17] method over its direct competitors. Finally, MC shows a nice potential for background modeling initialization in video surveillance. Future research may concern to evaluate incremental and real-time approaches of matrix completion in streaming videos.

References

1. L. Balzano and S. J. Wright. On GROUSE and incremental SVD. In *CAMSAP'2013*, pages 1–4, 2013.
2. T. Bouwmans. Traditional and recent approaches in background modeling for foreground detection: An overview. In *Computer Science Review*, 2014.
3. T. Bouwmans and E. Zahzah. Robust PCA via Principal Component Pursuit: A review for a comparative evaluation in video surveillance. In *Special Issue on Background Models Challenge, Computer Vision and Image Understanding*, volume 122, pages 22–34, May 2014.
4. J.-F. Cai, E. J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM J. on Optimization*, 20(4):1956–1982, March 2010.
5. E. J. Candès and Y. Plan. Matrix completion with noise. *CoRR*, abs/0903.3131, 2009.
6. E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *CoRR*, abs/0805.4471, 2008.
7. R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting objects, shadows and ghosts in video streams by exploiting color and motion information. In *ICIAP'2001*, pages 360–365, Sep 2001.
8. R. H. Keshavan, A. Montanari, and S. Oh. Matrix completion from noisy entries. *The Journal of Machine Learning Research*, 99:2057–2078, 2010.
9. A. H. S. Lai and N. H. C. Yung. A fast and accurate scoreboard algorithm for estimating stationary backgrounds in an image sequence. In *IEEE SCS'98*, pages 241–244, 1998.
10. Z. Lin, M. Chen, and Y. Ma. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices. *Mathematical Programming*, 2010.
11. L. Maddalena and A. Petrosino. Background model initialization for static cameras. In *Background Modeling and Foreground Detection for Video Surveillance*. CRC Press, Taylor and Francis Group, 2014.
12. L. Maddalena and A. Petrosino. Towards benchmarking scene background initialization. *CoRR*, abs/1506.04051, 2015.
13. R. Meka, P. Jain, and I. S. Dhillon. Guaranteed rank minimization via singular value projection. *CoRR*, abs/0909.5457, 2009.
14. T. Ngo and Y. Saad. Scaled gradients on grassmann manifolds for matrix completion. In *Advances in Neural Information Processing Systems 25*, pages 1412–1420, 2012.
15. N. M. Oliver, B. Rosario, and A. P. Pentland. A bayesian computer vision system for modeling human interactions. *IEEE PAMI*, 22(8):831–843, 2000.
16. A. Sobral and A. Vacavant. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *CVIU*, 122(0):4–21, 2014.
17. B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013.
18. Z. Wang, M. Lai, Z. Lu, W. Fan, H. Davulcu, and J. Ye. Orthogonal rank-one matrix pursuit for low rank matrix completion. *SIAM J. Scientific Computing*, 37(1), 2015.
19. Z. Wen, W. Yin, and Y. Zhang. Solving a low-rank factorization model for matrix completion by a nonlinear successive over-relaxation algorithm. *Mathematical Programming Computation*, 4(4):333–361, 2012.
20. X. Ye, J. Yang, X. Sun, K. Li, C. Hou, and Y. Wang. Foreground-background separation from video clips via motion-assisted matrix restoration. *IEEE T-CSVT*, PP(99):1–1, 2015.
21. X. Zhou, C. Yang, H. Zhao, and W. Yu. Low-rank modeling and its applications in image analysis. *ACM Computing Surveys (CSUR)*, 47(2):36, 2014.
22. Z. Zivkovic. Improved adaptive gaussian mixture model for background subtraction. In *ICPR 2004*, volume 2, pages 28–31 Vol.2, Aug 2004.