Fast Dual-Graph Regularized Background Foreground Separation

Longxiu Huang
Department of CMSE, Department of Mathematics
Michigan State University
East Lansing, MI, USA
Email: huangl3@msu.edu

Jing Qin
Department of Mathematics
University of Kentucky
Lexington, KY, USA
Email: jing.qin@uky.edu

***Abstract*—Foreground-background separation is a crucial task in various applications such as computer vision, robotics, and surveillance. Robust Principal Component Analysis (RPCA) is a popular method for this task, which considers the static background as the low-rank component and the moving objects in the foreground as the sparse component. To enhance the performance of RPCA, graph regularization is typically used to incorporate the sophisticated geometry of the background and temporal correlation. However, handling the graph Laplacians can be challenging due to the substantial number of data points. In this study, we propose a novel dual-graph regularized foreground-background separation model based on Sobolev smoothness. Our model is solved using a fast numerical algorithm based on the matrix CUR decomposition. Experimental results on real datasets demonstrate that our proposed algorithm achieves state-of-the-art computational efficiency.**

***Index Terms*—Robust principal component analysis, CUR decomposition, graph regularization, background foreground separation, motion detection.**

I. Introduction

Background and foreground separation (FBS) has been one of the most fundamental problems in many applications, including computer vision, medical imaging, security and surveillance [1]. It can identify the moving object of interest, which can be further used for motion analysis. More recently, due to the rapid development of human-robot interaction devices, it is of high demand to develop fast or even real-time FBS algorithms for detecting human motion and further intention interpretation.

In general, the goal of FBS is to decompose the video sequence into a component of static or dynamic background and a component of foreground. In the case of grayscale videos, due to the repetitive presence at every video frame, the static background can typically expressed as a low-rank matrix after reshaping the video as a matrix whose columns correspond to the video frames. In the meanwhile, the foreground component can be described by a sparse matrix where nonzero entries correspond to the the moving object of interest.

The research of Qin is supported by the NSF grant DMS-1941197. Huang was partially supported by AMS Simons Travel Grant. This work was supported in part through data provided by Intelligent Robotic Arms Lab at the University of Kentucky, and computational resources and services provided by the Institute for Cyber-Enabled Research at the Michigan State University.

Based on the low-rank structure of the background, robust principle component analysis (RPCA) has been one of the most popular methods in FBS [2], [3], [4], wherein convex relaxed formulas for RPCA were proposed and studied. Unfortunately, these earlier approaches only achieved sublinear convergence and thus are computationally intensive [5]. Later, a number of non-convex approaches were studied to solve nonconvex variants of RPCA directly. In particular, [6] proposed an alternating projection based non-convex algorithm and an accelerated version was studied in [7]. Moreover, a gradient descent based method was proposed in [8], which was recently modified for accelerating with ill-conditioned problems [9]. All of the aforementioned non-convex methods offered linear convergence with a complexity of at least $\mathcal{O}(rn^2)$ flops per iteration. In recent developments, CUR decompositions [10], [11] have been harnessed to further expedite RPCA methods [12], [13] and the complexity can be reduced to $\mathcal{O}(r^2 n \log^2(n))$.

Considering the spatial and temporal correlations of a video sequence, graph regularization has been exploited and applied to FBS. For example, spatial and temporal graph Laplacians on the background to be recovered are used in the objective function to preserve local geometries together with low-rank regularization and sparsity of foreground in [14], [15], [16], [17]. In particular, an adaptive low-rankness regularizer is adopted in [16], [17] with enhanced performance. However, as the resolution of video improves, the matrix size of spatial/temporal graph Laplcians grows rapidly, which brings great computational challenges for solving the graph regularized models, e.g., data storage and matrix calculations. To address these issues, we enforce the low-rank structure implicitly using the matrix CUR decomposition. In addition, Sobolev smoothness has shown the great potential in recovering spatiotemporal graph signals [18]. In light of this work, we propose a novel FBS model with Sobolev smoothness based spatial and temporal graph regularizations and develop a fast algorithm based. Specifically, we apply the alternating direction method of multipliers (ADMM) and update rows/columns of the low-rank component based on the matrix CUR decomposition.

The rest of this paper is organized as follows. In Section II, we provide a brief introduction of foreground/background

separation. In Section III, we propose a novel a fast FBS algorithm based on the matrix CUR decomposition. Numerical experiments on real data sets and the performance comparisons are presented in Section IV. Finally, conclusions of this research and future work are presented in Section V.

II. Preliminaries

In this section, we briefly review the matrix CUR decomposition. Given a matrix, classical dimension reduction techniques such as Principal Component Analysis (PCA) transform data from a high-dimensional space into a low-dimensional one while retaining the intrinsic dimension; however, those approximations may lose interpretability in some applications [19]. One way for addressing this issue is to utilize the self-expressiveness of data, i.e., data points are generally well-represented via linear combinations of the other data points in the same set rather than in some abstract bases, e.g., singular vectors. In particular, CUR decomposition is one such efficient self-expressive matrix approximation method, which can maintain the interpretability of the original data during the dimension reduction. Specifically, given a matrix $L \in \mathbb{R}^{n\times n}$ with rank r, CUR matrix decomposition aims to decompose L into terms involving only some of its columns and rows. If we choose some columns/rows of L that can span the column/row space of L, then we can retrieve L itself from these submatrices, which is guaranteed by the following theorem.

Theorem 1. *Let $\mathcal{I}, \mathcal{J} \subseteq [n]$ with $|\mathcal{I}|, |\mathcal{J}| \geq r$ be the respective row and column index sets, and denote the submatrices $C = L_{:,\mathcal{J}}$, $U = L_{\mathcal{I},\mathcal{J}}$ and $R = L_{\mathcal{I},:}$. If $\operatorname{rank}(U) = \operatorname{rank}(L)$, then $L = CU^{\dagger}R$, where $(\cdot)^{\dagger}$ denotes the Moore-Penrose pseudoinverse.*

Theorem 1 could be referred to [20] for a history and proof. From this theorem, one can see that the success of CUR decomposition highly relies on whether the rank of the mixing submatrix U equals that of L. In fact, there are various sampling strategies [21], [22], [23], [24] that can guarantee the condition with high probability. In particularly, [24] pointed out that when a given matrix with rank r has μ-coherence, sampling $|\mathcal{I}| = \mathcal{O}(\mu r \log(n))$ rows and $|\mathcal{J}| = \mathcal{O}(\mu r \log(n))$ columns uniformly with replacement can guarantee that $L = CU^{\dagger}R$ in high probability. Note that when $\mu \sim \mathcal{O}(1)$, U is an $\mathcal{O}(r\log(n)) \times \mathcal{O}(r\log(n))$ matrix under uniform sampling. By Theorem 1, the main computational cost is incurred by calculating the pseudo-inverse of U, which requires only $\mathcal{O}(r^3\log^2(n))$ flops. In contrast, computing the SVD requires $\mathcal{O}(rn^2)$ flops. This confirms the computational efficiency of the CUR decomposition with larger n and smaller r.

Assume that the background is static in a video. The goal of FBS is to decompose a video in a matrix form D into a low-rank component L and a sparse component S. To solve this problem, RPCA [2] proposes the model

$$\min_{L,S} \|L\|_* + \lambda\|S\|_1$$

with the constraint $L + S = D$ where $\lambda > 0$ is a regularization parameter and $\|\cdot\|_*$ is the matrix nuclear norm. Recently, graph regularized models with known rank of L have been explored with the form

$$\min_{\operatorname{rank} L \leq r, S} \lambda_1 R(L) + \lambda_2\|S\|_1 + \frac{1}{2}\|D - L - S\|_F^2$$

where $\lambda_1, \lambda_2 > 0$ are parameters and $R(L)$ is typically a graph regularization of L. Some related works include [14], [15] and one most recent robust variant [16].

III. Proposed Method

A. Construction of Two Graph Laplacians

Given a video, we first convert it to a matrix $D \in \mathbb{R}^{n\times m}$, whose rows and columns correspond to the respective number of spatial and the temporal samples. Next we aim to generate a weighted graph in the temporal domain $G_t = (V_t, E_t, A_t)$ where $V_t = \{\mathbf{v}_i^t\}_{i=1}^m$ is a set of temporal samples, E_t is an edge set and $A_t \in \mathbb{R}^{m\times m}$ is the adjacency matrix whose entries correspond to the weights or the similarity for any two temporal samples. To fulfill this task, we create an adjacency matrix A_t whose (i,j)-th entry is defined as

$$(A_t)_{i,j} = \exp\left(-\frac{\|\mathbf{v}_i^t - \mathbf{v}_j^t\|_2^2}{h_t^2}\right), \quad i,j \in \{1,\ldots,m\}.$$

Here h_t is a positive filtering parameter. Let W_t be the degree diagonal matrix of G_t with $(W_t)_{i,i} = \sum_{j=1}^m (A_t)_{i,j}$. Next we follow the ideas in [25], [26] to define a symmetrically normalized temporal graph Laplacian $\Psi_t \in \mathbb{R}^{m\times m}$ defined as $\Psi_t = I_m - W_t^{-1/2}A_tW_t^{-1/2}$, which can handle irregular graph structures.

Similarly, we use spatial samples to generate a weighted graph $G_s = (V_s, E_s, A_s)$ in the spatial domain where $V_s = \{\mathbf{v}_i^s\}_{i=1}^n$ is the spatial sample set, E_s is the edge set and $A_s \in \mathbb{R}^{n\times n}$ is the associated spatial adjacency matrix. Note that to exploit non-local geometries in the spatial domain, we utilize the patch based similarity for designing A_s. To be specific, the (i,j)-th entry of A_s is defined as

$$(A_s)_{i,j} = \exp\left(-\frac{\|\mathcal{N}(\mathbf{v}_i^s) - \mathcal{N}(\mathbf{v}_j^s)\|_F^2}{h_s^2}\right), i,j \in \{1,\ldots,n\}$$

where $\mathcal{N}(\mathbf{v}_i^s) \in \mathbb{R}^{p^2\times m}$ is a vectorized version of the video patch centered at the i-th pixel and h_s is a positive spatial filtering parameter. The k-nearest neighbors in terms of location are used to ease the computational burden when calculating A_s. In our experiments, we use four nearest neighboring pixels in the spatial domain to calculate patchwise similarities and then assemble them to obtain the spatial adjacency matrix A_s. In the temporal domain, we also use four nearest neighboring temporal samples along the time, i.e., preceding and receding two temporal samples, to compute A_t. Note that if the given video is noisy, some local smoothing techniques or more robust similarity metrics could be incorporated into this procedure. Finally, we define the spatial graph Laplacian in a symmetrically normalized form to be

$$\Psi_s = I_n - W_s^{-1/2}A_sW_s^{-1/2}.$$

Similar to W_t, W_s is the degree diagonal matrix corresponding to G_s which can be obtained by using A_s. It is worth noting that all graph Laplacians are stored as sparse matrices to substantially reduce the data storage memory.

B. Proposed Algorithm

To preserve high-order spatial and temporal correlations for the background, we propose the following robust background foreground separation model

$$\min_{L,S} \|D-L-S\|_1 + \zeta\|S\|_1 + \gamma_1 \operatorname{tr}(L^T\Phi_s L) + \gamma_2 \operatorname{tr}(L\Phi_t L^T) \quad \text{s.t.} \quad \operatorname{rank}(L) \le r.$$

Here $\Phi_s = (\Psi_s + \varepsilon_s I)^{\alpha_s}$ and $\Phi_t = (\Psi_t + \varepsilon_t I)^{\alpha_t}$ are the modified spatial and temporal graph Laplacians based on the Sobolev smoothness [18] with $\alpha_s > 0, \alpha_t > 0$ being the order of smoothness and $\varepsilon_s, \varepsilon_t \ge 0$. We rewrite the above model as

$$\min_{L,S} \|P\|_1 + \zeta\|S\|_1 + \gamma_1 \operatorname{tr}(L^T\Phi_s L) + \gamma_2 \operatorname{tr}(L\Phi_t L^T) \quad \text{s.t.} \quad \operatorname{rank}(L) \le r,\ D-L-S=P.$$

Next we introduce a matrix space with low rank structure

$$\Pi = \{X \in \mathbb{R}^{n\times m} : \operatorname{rank}(X) \le r\},$$

and define its corresponding indicator function as:

$$\delta_\Pi(X) = \begin{cases} 0, & X \in \Pi; \\ +\infty, & \text{otherwise.} \end{cases}$$

Then we define the augmented Lagrangian function as

$$\mathcal{L}(S,L,P,\widehat{P}) = \delta_\Pi(L) + \|P\|_1 + \zeta\|S\|_1 + \frac{\gamma_1}{2}\operatorname{tr}(L^T\Phi_s L) + \frac{\gamma_2}{2}\operatorname{tr}(L\Phi_t L^T) + \frac{\rho}{2}\|D-L-S-P+\widehat{P}\|_F^2. \tag{1}$$

By applying the ADMM framework, we get the algorithm

$$\begin{cases} L \leftarrow \underset{\operatorname{rank}(L)\le r}{\operatorname{argmin}} \frac{\gamma_1}{2}\operatorname{tr}(L^T\Phi_s L) + \frac{\gamma_2}{2}\operatorname{tr}(L\Phi_t L^T) \\ \quad + \frac{\rho}{2}\|D-L-S-P+\widehat{P}\|_F^2 := f(L) \\ S \leftarrow \underset{S}{\operatorname{argmin}}\, \zeta\|S\|_1 + \frac{\rho}{2}\|D-L-S-P+\widehat{P}\|_F^2 \\ P \leftarrow \underset{P}{\operatorname{argmin}} \|P\|_1 + \frac{\rho}{2}\|D-L-S-P+\widehat{P}\|_F^2 \\ \widehat{P} \leftarrow \widehat{P} + (D-L-S-P) \end{cases}$$

In the context of background-foreground separation, it is worth noting that the background remains static throughout the application. As a result, we can assume that the rank of the background matrix, denoted as L, is equal to 1, i.e., $\operatorname{rank}(L) = 1$. Based on this assumption, we can approach the L-subproblem by assuming that L can be factorized as $L = CU^\dagger R$. Here, C and R represent the column and row submatrices of L, respectively, with column and row index sets denoted as J and I. Furthermore, U corresponds to the submatrix of L formed by the intersection of columns J and rows I, denoted as $L(I,J)$. The gradient of f in the L-subproblem is

$$\nabla f(L) = \gamma_1\Phi_s L + \gamma_2 L\Phi_t + \rho(L+S+P-D-\widehat{P}). \tag{2}$$

Algorithm 1 Fast algorithm for dual-graph regularized FBS

Input: observed data matrix D; rank r; spatial/temporal graph Laplacians Ψ_s, Ψ_t; initial shrinkage value ζ_0; graph Laplacian smoothness order α_s, α_t; graph Laplacian smoothness parameter ε_s, ε_t; regularization parameters γ_1, γ_2 for spatial/temporal graph Laplacians; penalty parameter ρ; decay rate β; step size η; step size η; maximum iteration number T; row/column sampling index sets $\mathcal{I}, \mathcal{J}$.
$P_0 = 0$; $\widehat{P}_0 = 0$; $S_0 = \operatorname{shrink}(D, \zeta_0)$
$\mathcal{I}^c = [m] \setminus \mathcal{I}$, $\mathcal{J}^c = [n] \setminus \mathcal{J}$
$C_0 = D_{:,\mathcal{J}}$, $R_0 = D_{\mathcal{I},:}$
for $k = 1, \cdots, T-1$ **do**
  Find the gradient $G_k = \nabla f(L^{(k)})$ given by (2)
  $C_{k+1}(\mathcal{I}^c,:) = C_k(\mathcal{I}^c,:) - \eta G_k(\mathcal{I}^c,\mathcal{J})$,
  $R_{k+1}(:,\mathcal{J}^c) = R_k(:,\mathcal{J}^c) - \eta G_k(\mathcal{I},\mathcal{J}^c)$
  $U_{k+1} = \mathcal{H}_r(U_k - \eta G_k(\mathcal{I},\mathcal{J}))$; //$\mathcal{H}_r(\cdot)$ denotes the best rank r approximation to the argument.
  $C_{k+1}(\mathcal{I},:) = U_{k+1}$ and $R_{k+1}(:,\mathcal{J}) = U_{k+1}$
  $L_{k+1} = C_{k+1}U_{k+1}^\dagger R_{k+1}$
  $\zeta_{k+1} = \beta^k\zeta_0$
  $S_{k+1} = \operatorname{shrink}(D - L_{k+1} - P_k + \widehat{P}_k, \zeta_{k+1}/\rho)$
  $P_{k+1} = \operatorname{shrink}(D - L_{k+1} - S_{k+1} + \widehat{P}_k, 1/\rho)$
  $\widehat{P}_{k+1} = \widehat{P}_k + D - L_{k+1} - S_{k+1} - P_{k+1}$
Output: C_k, U_k, R_k: CUR decomposition of L.

We update the C, R, U separately:

$$\begin{cases} C(I^c,:) \leftarrow C(I^c,:) - \eta\nabla f(I^c,J) \\ R(:,J^c) \leftarrow R(:,J^c) - \eta\nabla f(I,J^c) \\ U \leftarrow \mathcal{H}_r(U - \eta\nabla f(I,J)) \\ C(I,:) \leftarrow U, R(:,J) \leftarrow U \end{cases} \tag{3}$$

with $I^c = [n] \setminus I$, $J^c = [m] \setminus J$ and $\eta > 0$ being the step size for the gradient descent.

Both S-subproblem and P-subproblem can be cast as finding the proximal operator of ℓ_1-norm, which is the shrink operator defined by

$$\operatorname{shrink}(X,\mu)_{ij} = \operatorname{sign}(x_{ij})\max(|x_{ij}| - \mu, 0).$$

Then S is updated via

$$S \leftarrow \operatorname{shrink}(D - L - P + \widehat{P}, \zeta/\rho),$$

and P is updated via

$$P \leftarrow \operatorname{shrink}(D - L - S + \widehat{P}, 1/\rho).$$

IV. Numerical Experiments

In this section, we compare the empirical performance of our method (i.e. Alg. 1) for the FBS problem with the other related algorithms, including robust dual-graph moving object detection (RDMOD) [16], [17], robust PCA on graphs (RPCAG) [27] and PCA using graph total variation (PCAGTV) [28] on two real videos. All color videos are converted to gray-scale and the intensities are rescaled to $[0,1]$. Since our method involves the randomly selection of indices I, J to form matrices R and C,

we run our methods for 20 times and record the corresponding reconstruction quality and runtime. The selection of window size and filtering parameters in the graph Laplacians can be found in [16], and all the other parameters are tuned to optimize the performance.

For performance evaluations, we adopt the various comparison metrics for the background recovery quality and the foreground detection accuracy. For the background, we use the following two metrics:

(a) Peak Signal-to-Noise Ratio (PSNR) in dB is defined by

$$\text{PSNR} = 20\log_{10}(I_{\max}/\|\widehat{L} - L\|_{\text{F}}),$$

where $I_{\max}$ is the maximum intensity, and $\widehat{L}$ is an estimate of the ground truth of the background L;

(b) Structural Similarity Index Measure (SSIM) [29].

For the foreground, we use the following three comparison metrics from sensitivity analysis. Assume TP is the number of the foreground pixels that are accurately labeled as foreground, FP is the number of the background pixels that are inaccurately labeled as foreground, and FN is the number of foreground pixels that are incorrectly labeled as background:

(a) Precision (Pr) is defined as $\text{Pr} = TP/(TP + FP)$;
(b) Recall (Re) is defined as $\text{Re} = TP/(TP + FN)$;
(c) F-measure (Fm) is defined as $\text{Fm} = 2\text{Re}/\text{Pr}$.

All numerical experiments are conducted on Matlab R2021b and executed on a laptop with Intel CPU i7-11800H (2.30 GHz) and 16 GB RAM.

A. Experiment 1

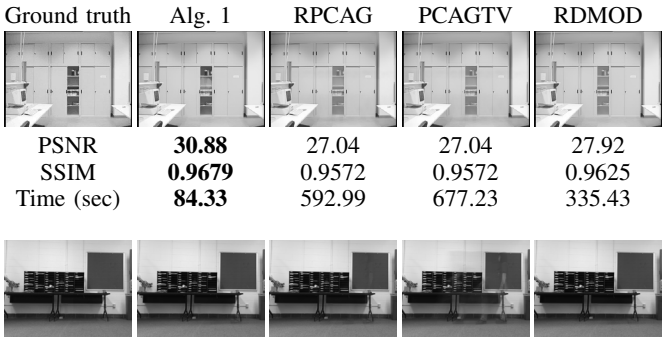
The first video, denoted by Lab, is downloaded from http://web.eee.sztaki.hu/∼bcsaba/FgShBenchmark.htm. The ground truth of the background is shown in Fig. 1. This dataset has 205 frames in total and each frame is an image of size 240×320. This dataset is challenging for background recovery since the closest door in the video is closed for the first 84 frames and then is opened by one person afterwards.

The average runtime and the reconstruction qualities in terms of PSNR and SSIM for each result are summarized in Fig 1. Furthermore, one visual result for each method is shown in Figure 1. We can observe that the PSNR and SSIM values are higher in our results.

In addition, upon closer examination of the visual results, it becomes evident that the contrast of the cabinet door is superior to that of the other doors. In terms of runtime, our approach demonstrates significantly faster performance with a running time of only approximately one-fourth that of RDMOD and merely one-eighth that of PCAGTV.

B. Experiment 2

In the second experiment, we use the walking video from [16], [17], whose original color video can be downloaded from http://bxie.engr.uky.edu/dataset. After taking a preprocessing step, the test video consists of 65 gray-scale image frames, each with 150×200 pixels. The average runtimes and recovered backgrounds for all the methods are shown in the second row of Fig. 1. The quantitative comparison for the recovered backgrounds and foregrounds are listed in Table I. Note that our result has the highest F-measure value indicating a good balance in precision and recall. Our method outperforms RPCAG and PCAGTV in almost all quality measurements and runtime. Compared to RDMOD, our approach significantly reduces runtime while achieving comparable or better performance in terms of recovered background image quality and detection accuracy.



	Alg. 1	RPCAG	PCAGTV	RDMOD
PSNR	**30.88**	27.04	27.04	27.92
SSIM	**0.9679**	0.9572	0.9572	0.9625
Time (sec)	**84.33**	592.99	677.23	335.43

	Alg. 1	RPCAG	PCAGTV	RDMOD
PSNR	41.54	39.98	28.58	**41.84**
SSIM	0.9945	0.992	0.9525	**0.9947**
Time (sec)	**12.15**	149.52	77.39	38.68

Fig. 1. Video background subtraction results. Column 1 contains the ground truth background, the other four columns are the reconstructed background outputted by Alg. 1, RPCAG, PCAGTV and RDMOD respectively.

TABLE I
QUANTITATIVE COMPARISON IN EXPERIMENT 2

	PSNR	SSIM	Pr	Re	Fm
Alg. 1	41.50	0.9944	0.9044	0.7565	**0.8239**
RPCAG	39.98	0.9920	0.9008	**0.7569**	0.8226
PCAGTV	28.95	0.9548	0.6292	0.2464	0.3541
RDMOD	**41.84**	**0.9947**	**0.9150**	0.7488	0.8236

V. CONCLUSIONS AND FUTURE WORK

Foreground-background separation is one of the most fundamental tasks in many applications, such as social security and computer vision. Due to the static presence of the background, a video can be typically decomposed into a sum of low-rank background and a sparse foreground. Many graph regularization involving graph Laplacians have been proposed to enhance the spatiotemporal geometry in the RPCA framework. However, they suffer from the computational bottleneck when the graph Laplacian is large. In this work, we propose a Sobolev smoothness enhanced FBS model with spatial and temporal graph regularizations, and an efficient algorithm based on CUR matrix decomposition and ADMM algorithmic framework. Numerical results have shown our proposed method performs well on realistic data sets. In the future, we will extend this framework to develop tensor-based fast color video FBS algorithms and consider shadow removal from the foreground under complex lightening situations.

REFERENCES

[1] Thierry Bouwmans, Andrews Sobral, Sajid Javed, Soon Ki Jung, and El-Hadi Zahzah, "Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset," *Computer Science Review*, vol. 23, pp. 1–71, 2017.

[2] Emmanuel J Candès, Xiaodong Li, Yi Ma, and John Wright, "Robust principal component analysis?," *J. ACM*, vol. 58, no. 3, pp. 1–37, 2011.

[3] Huan Xu, Constantine Caramanis, and Sujay Sanghavi, "Robust PCA via outlier pursuit," in *Adv. NIPS*, 2010, pp. 2496–2504.

[4] Venkat Chandrasekaran, Sujay Sanghavi, Pablo A Parrilo, and Alan S Willsky, "Rank-sparsity incoherence for matrix decomposition," *SIAM J. Optim.*, vol. 21, no. 2, pp. 572–596, 2011.

[5] Shiqian Ma and Necdet Serhat Aybat, "Efficient optimization algorithms for robust principal component analysis and its variants," *Proc. IEEE*, vol. 106, no. 8, pp. 1411–1426, 2018.

[6] Praneeth Netrapalli, UN Niranjan, Sujay Sanghavi, Animashree Anandkumar, and Prateek Jain, "Non-convex robust PCA," in *Adv. NIPS*, 2014, pp. 1107–1115.

[7] HanQin Cai, Jian-Feng Cai, and Ke Wei, "Accelerated alternating projections for robust principal component analysis," *J. Mach. Learn. Res.*, vol. 20, no. 1, pp. 685–717, 2019.

[8] Xinyang Yi, Dohyung Park, Yudong Chen, and Constantine Caramanis, "Fast algorithms for robust PCA via gradient descent," *Advances in neural information processing systems*, vol. 29, 2016.

[9] Tian Tong, Cong Ma, and Yuejie Chi, "Accelerating ill-conditioned low-rank matrix estimation via scaled gradient descent," *The Journal of Machine Learning Research*, vol. 22, no. 1, pp. 6639–6701, 2021.

[10] Sergei A. Goreĭnov, Eugene E. Tyrtyshnikov, and Nickolai L. Zamarashkin, "A theory of pseudoskeleton approximations," *Linear algebra and its applications*, vol. 261, no. 1-3, pp. 1–21, 1997.

[11] Petros Drineas, Michael W Mahoney, and S Muthukrishnan, "Relative-error CUR matrix decompositions," *SIAM Journal on Matrix Analysis and Applications*, vol. 30, no. 2, pp. 844–881, 2008.

[12] HanQin Cai, Keaton Hamm, Longxiu Huang, and Deanna Needell, "Robust cur decomposition: Theory and imaging applications," *SIAM Journal on Imaging Sciences*, vol. 14, no. 4, pp. 1472–1503, 2021.

[13] HanQin Cai, Keaton Hamm, Longxiu Huang, Jiaqi Li, and Tao Wang, "Rapid robust principal component analysis: CUR accelerated inexact low rank estimation," *IEEE Signal Processing Letters*, vol. 28, pp. 116–120, 2020.

[14] Sajid Javed, Seon Ho Oh, Andrews Sobral, Thierry Bouwmans, and Soon Ki Jung, "Background subtraction via superpixel-based online matrix decomposition with structured foreground constraints," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2015, pp. 90–98.

[15] Sajid Javed, Arif Mahmood, Thierry Bouwmans, and Soon Ki Jung, "Spatiotemporal low-rank modeling for complex scene background initialization," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 6, pp. 1315–1329, 2016.

[16] Jing Qin, Ruilong Shen, Ruihan Zhu, and Biyun Xie, "Robust dual-graph regularized moving object detection," in *2022 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2022, pp. 487–492.

[17] Jing Qin and Biyun Xie, "Human motion detection based on dual-graph and weighted nuclear norm regularizations," *arXiv preprint arXiv:2304.04879*, 2023.

[18] Jhony H Giraldo, Arif Mahmood, Belmar Garcia-Garcia, Dorina Thanou, and Thierry Bouwmans, "Reconstruction of time-varying graph signals via sobolev smoothness," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 8, pp. 201–214, 2022.

[19] Michael W Mahoney and Petros Drineas, "CUR matrix decompositions for improved data analysis," *Proc. Nat. Acad. Sci.*, vol. 106, no. 3, pp. 697–702, 2009.

[20] Keaton Hamm and Longxiu Huang, "Perspectives on CUR decompositions," *Appl. Comput. Harmon. Anal.*, vol. 48, no. 3, pp. 1088–1099, 2020.

[21] Keaton Hamm and Longxiu Huang, "Stability of sampling for CUR decompositions," *Found. Data Sci.*, vol. 2, no. 2, pp. 83–99, 2020.

[22] Sergey Voronin and Per-Gunnar Martinsson, "Efficient algorithms for CUR and interpolative matrix decompositions," *Advances in Computational Mathematics*, vol. 43, no. 3, pp. 495–516, 2017.

[23] Saifon Chaturantabut and Danny C Sorensen, "Nonlinear model reduction via discrete empirical interpolation," *SIAM Journal on Scientific Computing*, vol. 32, no. 5, pp. 2737–2764, 2010.

[24] Jiawei Chiu and Laurent Demanet, "Sublinear randomized algorithms for skeleton decompositions," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 3, pp. 1361–1383, 2013.

[25] Jing Qin, Harlin Lee, Jocelyn T Chi, Yifei Lou, Jocelyn Chanussot, and Andrea L Bertozzi, "Fast blind hyperspectral unmixing based on graph laplacian," in *2019 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. IEEE, 2019, pp. 1–5.

[26] Jing Qin, Harlin Lee, Jocelyn T Chi, Lucas Drumetz, Jocelyn Chanussot, Yifei Lou, and Andrea L Bertozzi, "Blind hyperspectral unmixing based on graph total variation regularization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 3338–3351, 2020.

[27] Nauman Shahid, Vassilis Kalofolias, Xavier Bresson, Michael Bronstein, and Pierre Vandergheynst, "Robust principal component analysis on graphs," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 2812–2820.

[28] Nauman Shahid, Nathanael Perraudin, Vassilis Kalofolias, Benjamin Ricaud, and Pierre Vandergheynst, "PCA using graph total variation," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Ieee, 2016, pp. 4668–4672.

[29] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.