INVITED PAPER

On the Applications of Robust PCA in Image and Video Processing

This paper surveys the applications of RPCA in computer vision and biomedical imaging by reviewing representative image processing applications (low-level imaging, biomedical imaging, 3-D computer vision), and video processing applications such as background/foreground separation.

By Thierry Bouwmans, Sajid Javed, Hongyang Zhang, Zhouchen Lin, *Fellow IEEE*, and Ricardo Otazo

ABSTRACT | Robust principal component analysis (RPCA) via decomposition into low-rank plus sparse matrices offers a powerful framework for a large variety of applications such as image processing, video processing, and 3-D computer vision. Indeed, most of the time these applications require to detect sparse outliers from the observed imagery data that can be approximated by a low-rank matrix. Moreover, most of the time experiments show that RPCA with additional spatial and/or temporal constraints often outperforms the state-of-the-art algorithms in these applications. Thus, the aim of this paper is to survey the applications of RPCA in computer vision. In the first part of this paper, we review representative image processing applications as follows: 1) low-level imaging such as image recovery and denoising, image composition, image colorization, image alignment and rectification, multi-focus image, and face recognition; 2) medical imaging such as dynamic magnetic resonance imaging (MRI) for acceleration of data acquisition, background suppression, and learning of interframe motion fields; and 3) imaging for 3-D computer vision with additional depth information such as in structure from motion (SfM) and 3-D motion recovery. In the second part, we present the applications of RPCA in video processing which utilize additional spatial and temporal information compared to image processing. Specifically, we investigate video denoising and restoration, hyperspectral video, and background/foreground separation. Finally, we provide perspectives on possible future research directions and algorithmic frameworks that are suitable for these applications.

KEYWORDS | Image processing; medical imaging; robust principal component analysis (RPCA); video processing; 3-D computer vision

I. INTRODUCTION

Principal component analysis (PCA) was introduced by Karl Pearson in 1901 and was first widely used in statistics. However, its main limitation includes its sensitivity to outliers, its high computation time, and memory requirements, which make the model unsuitable for high-dimensional data as in computer vision applications. The robustness of PCA methods was first addressed in statistics by replacing the standard estimation of the covariance matrix with a robust estimator [36], [133] or by using projection pursuit techniques [57], [134]. On the other hand, in neural networks, PCA was robustified by designing a neural network that relied on self-organizing rules based

Manuscript received January 15, 2018; revised May 7, 2018; accepted June 26, 2018. Date of current version August 2, 2018. The work of Z. Lin was supported by the National Basic Research Program of China (973 Program) under Grant 2015CB352502, by the National Natural Science Foundation of China under Grants 61625301 and 61731018, by Qualcomm, and by Microsoft Research Asia. *(Corresponding author: Thierry Bouwmans.)*

T. Bouwmans is with the Laboratoire MIA, University of La Rochelle, 17000, La Rochelle, France (e-mail: tbouwman@univ-lr.fr.).

S. Javed is with the Department of Computer Science, University of Warwick, Coventry, U.K. (e-mail: s.javed.1@warwick.ac.uk.).

H. Zhang is with the Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: hongyanz@cs.cmu.edu.).

Z. Lin is with the Key Laboratory of Machine Perception (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, Beijing, China, and also with the Cooperative Medianet Innovation Center, Shanghai Jiao Tong University, Shanghai, China (e-mail: zlin@pku.edu.cn.).

R. Otazo is with the Center for Advance Imaging Innovation and Research, New York University, New York, NY, USA (e-mail: ricardo.otazo@nyumc.org).

Digital Object Identifier 10.1109/JPROC.2018.2853589

on statistical physics [309]. However, all these robust methods are still limited to relatively low-dimensional data and thus they are not applicable for computer vision applications with high-dimensional data. In a further work, Candès *et al.* [37] addressed the robustness by decomposition into low-rank plus sparse matrices (also called $\mathbf{L}+\mathbf{S}$ decomposition [42], [128]), and provided by several ways a suitable framework for many signal processing and computer vision applications. Practically, Candès *et al.* [37] proposed a convex optimization to address the robust PCA problem. The observation matrix $\mathbf{A}$ is assumed to be represented as

$$\mathbf{A} = \mathbf{L} + \mathbf{S} \tag{1}$$

where $\mathbf{L}$ is a low-rank matrix and $\mathbf{S}$ must be a sparse matrix with a small fraction of nonzero entries. The straightforward formulation is to use the l_0-norm to minimize the energy function

$$\min_{\mathbf{L},\mathbf{S}} \ \text{rank}(\mathbf{L}) + \lambda\|\mathbf{S}\|_0 \quad \text{subj} \quad \mathbf{A}-\mathbf{L}-\mathbf{S}=0 \tag{2}$$

where $\lambda > 0$ is an arbitrary balanced parameter. However, this problem is NP-hard, and typical solution might involve a search with combinatorial complexity. This research seeks to solve for $\mathbf{L}$ with the following optimization problem:

$$\min_{\mathbf{L},\mathbf{S}} \ \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1 \quad \text{subj} \quad \mathbf{A}-\mathbf{L}-\mathbf{S}=0 \tag{3}$$

where $\|.\|_*$ and $\|.\|_{l_1}$ are the nuclear norm (which is the l_1-norm of singular value) and l_1-norm, respectively, and $\lambda > 0$ is an arbitrary balanced parameter. Usually, $\lambda = (1/\sqrt{\max(m,n)})$. Under these minimal assumptions, this approach called principal component pursuit (PCP) solution perfectly recovers the low-rank and the sparse matrices.

The main difference between the robust PCA (RPCA) based on robust estimators and projection pursuit and the $\mathbf{L}+\mathbf{S}$ decomposition model is that the first approaches assume outlying data points in which the entire row or column of the data matrix is corrupted while the second approach assumes outliers that are uniformly distributed. In addition, the classical approaches on robust PCA focus mostly on estimators with excellent worst case robustness but poor computational profiles because they are NP-hard to compute or they involve combinatorial search making them unsuitable for computer vision applications with high-dimensional data. On the other hand, Candès *et al.* [37] showed results with the $\mathbf{L}+\mathbf{S}$ decomposition model in computer vision applications such as face images and background modeling that demonstrated encouraging performance. This original RPCA formulation suffices in applications (such as image denoising and image alignment) where the information of interest is in the low-rank $\mathbf{L}$ matrix. However, applying directly this original RPCA formulation in applications in which there is also information of interest in the sparse matrix $\mathbf{S}$ (such as background/foreground separation) results in outliers containing both the information of interest (moving objects that are considered as sparse) and the noise. Thus, most of the time, the stable RPCA formulation [355] is preferred for this kind of computer vision applications. The stable formulation PCP (also called $\mathbf{L}+\mathbf{S}+\mathbf{E}$ decomposition) assumes that the observation matrix $\mathbf{A}$ is represented as follows:

$$\mathbf{A} = \mathbf{L} + \mathbf{S} + \mathbf{E} \tag{4}$$

where $\mathbf{E}$ is a noise term [say independent identically distributed (i.i.d.) noise on each entry of the matrix] and $\|\mathbf{E}\|_F < \delta$ for some $\delta > 0$. To recover $\mathbf{L}$ and $\mathbf{S}$, Zhou *et al.* [355] proposed to solve the following optimization problem, as a relaxed version to PCP:

$$\min_{L,S} \ \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1 \quad \text{subj} \quad \|\mathbf{A}-\mathbf{L}-\mathbf{S}\|_F < \delta \tag{5}$$

where $\|.\|_F$ is the Frobenius norm and $\lambda = (1/\sqrt{n})$. Furthermore, to enhance the adequacy of the RPCA formulation for computer vision, spatial and/or temporal additional constraints need to be introduced by using specific regularization terms or function applied on $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{E}$. A general formulation of the optimization problem suitable for a RPCA formulation applied to a computer vision application can be written as follows:

$$\begin{aligned}\min_{\mathbf{L},\mathbf{S},\mathbf{E}} & \underbrace{\|T(\mathbf{L})\|_* + \lambda_1\|\Pi(\mathbf{S})\|_1 + \lambda_2\|\mathbf{E}\|_F}_{\text{Constrained Stable RPCA Decomposition}} \\ & + \underbrace{\delta_1 F(\mathbf{L}) + \delta_2 G(\mathbf{S})}_{\text{Computer Vision Application}} \\ \text{s.t. } & \mathbf{A} = \mathbf{L}+\mathbf{S}+\mathbf{E}, \quad \text{or} \quad \mathbf{A} = \mathbf{W}\circ(\mathbf{L}+\mathbf{S}+\mathbf{E}), \\ & \text{or } \mathbf{A}\circ\tau = \mathbf{L}+\mathbf{S}+\mathbf{E}\end{aligned} \tag{6}$$

where $T(\cdot)$ and $\Pi(\cdot)$ are linear operators applied on $\mathbf{L}$ and $\mathbf{S}$, respectively. They allow to take into account spatial and temporal constraints as well as the functions $F(\cdot)$ and $G(\cdot)$ that are usually suitable norms for the specific constraints met in the application. A weighting matrix $\mathbf{W}$ or a transformation τ can also be used in the constraint of the minimization. In literature, numerous publications used the robust PCA formulation by improving its computational efficiency and its adequacy to the concerned application in 1) signal processing applications such as in satellite communication [155], seismology [68], [54], speech enhancement [130], [301], synthetic aperture radar (SAR) imaging [132], [318], [251], [164], [160], [340], [23], [24], [99], and direction-of-arrival tracking [172], [62]; 2) computer vision applications such as in image processing, video processing, and 3-D computer vision as developed in [28]; 3) computer science applications such as the detection traffic anomalies [198], [223];

and 4) astronomy for auroral substorm detection [315] and exoplanet detection [105], [224]. A full list of publications of RPCA in these different applications is available at the DLAM website.[1]

Thus, even if PCA is a problem that has existed for over a century, and also applied in computer vision since 2000s [279], [280], the work of Candès *et al.* [37] is the main reason why there has been a resurgence of interest in RPCA and extensions in computer vision for the last six years. The other reasons concern two main points. First, new singular value decomposition (SVD) solutions have been developed to make the iterations as efficient as possible and to deal with the fact that the standard SVD solution fails if the data are corrupted by anything other than small noise. For example, approximated SVD solutions exist to avoid full SVD computations in order to reduce computation time such as partial SVD algorithms [175], linear time SVD algorithms [314], limited memory SVD algorithms [183], symmetric low-rank product-Gauss–Newton algorithms [184], block Lanczos with warm start (BLWS) algorithms [177], and randomized SVD algorithms [87], [346], [150]. Moreover, a lot of video data arrive sequentially over time and the subspace in which the data lie can change with time. Motivated by these reasons, there has been an array of papers using online or streaming robust PCA (also called robust subspace tracking [285]), and some of them specifically focused on online dynamic robust PCA [187], [285] with performance guarantees [186], [188], [208], [326] and memory-efficient algorithms [207]. This line of research allows its application in computer vision such as background/foreground separation which requires incremental and real-time computations. Moreover, there are also algorithms for high-dimensional data [238], [239], [240], [241]. Furthermore, RPCA often outperforms previous state-of-the-art methods in several computer vision applications [38], [108], [174] with rigorous theoretical analysis [19], [37], [329]. Indeed, as this decomposition is nonparametric and does not make many assumptions, it is widely applicable to large-scale problems ranging from image processing to video processing.

II. PRELIMINARY OVERVIEW

Many tasks in image and video processing present in the observed data combination of 1) one information of interest and perturbations, or 2) two information of interest and perturbations. Here, perturbations include both notion of noise and errors than occur in computer vision systems. In the first case, information of interest and perturbations present low-rank and sparsity aspects, respectively. Thus, RPCA via $\mathbf{L}+\mathbf{S}$ decomposition offers a suitable framework for these processing. Then, the low-rank component mathematically contains the inliers (information of interest) and the sparse components contains the outliers (noise). In the second case, the first information of interest and the second information of interest present low-rank and sparsity aspects, respectively. Thus, the stable RPCA formulation is required to avoid the matrix S to contain both the second information of interest and the perturbations. Furthermore, the spatial aspects present in images, and the temporal constraints in video sequences can be used in the $\mathbf{L}+\mathbf{S}$ decomposition and $\mathbf{L}+\mathbf{S}+\mathbf{E}$ decomposition to enforce its adequacy to the concerned task.

[1] https://sites.google.com/site/robustdlam/

A. Image Processing

RPCA framework was applied with a great success in the following imaging applications.

- Low-level imaging and analysis: image restoration and denoising [109], [154], [272], [291], [292], texture image denoising [171], hyperspectral image denoising [51], [104], [296], image completion and inpainting [39], [310], image composition for high-dynamic range imaging [21], image decomposition for intrinsic image computation [156], [324] and for structural image decomposition [44], image alignment and rectification [226], [242], [270], [304], [339], image stitching and mosaicking [168], image colorization [317], multifocus image [288], [289], [336], [337], [338], pansharpening [333], change detection [52], face recognition [190], [300], [331], partial-duplicate image search [313], image saliency detection [152], [165], [166], [230], [236], and image analysis [354], [178].
- Medical imaging: RPCA has become a powerful tool to increase the performance of data acquisition [93], [94], [217], [218], [281], image reconstruction [222] and image analysis of brain images [13], [14], [93], [219], [220], [261], cardiac images [49], [50], [94], [217], [218], [222], [281], [307], vessels images [148], and retina images [90]. A key initial application was to reduce the number of measurements in dynamic imaging (space + time), which resulted in increased imaging speed for MRI [94], [217], [218], [281] and radiation dose reduction for CT [93]. In addition, the separation of the background in the low-rank component performed automated background suppression for angiography and contrast-enhanced studies. RPCA can also perform a robust separation of common and individual information when analyzing a group of clinical data sets, such as functional and diffusion MRI of the brain [14], [219], [220]. In another way, RPCA can also detect changes in the retina [90] and also aligned image for speckle reduction of retinal OCT images [18].
- Imaging for 3-D computer vision: This application requires mechanical measurement of the camera positions or manual alignment of partial 3-D views of a scene. Thus, RPCA can also be used to reduce outliers and noise in algorithms such as SfM [9], [8], [182], [302], 3-D motion recovery [294], and 3-D reconstruction [10].

B. Video Processing

This application is the most investigated one. Indeed, numerous authors used RPCA problem formulations in applications such as background/foreground separation [4], [215], [231], background initialization [266], [269], moving target detection [252], motion saliency detection [48], [311], [343], motion estimation [249], visual object tracking [173], [287], action recognition [131], key frame extraction [61], video object segmentation [135], [158], [204], [328], [330], video coding [46], [47], [114], [342], video restoration and denoising [147], [113], [181], [329], [345], video inpainting [147], hyperspectral video processing [43], [100], and video stabilization [72].

In the following sections, we introduce how the RPCA formulation is employed in these applications. Particularly, we indicate how the observed image and video data are stacked in the input matrix $\mathbf{A}$, and the signification of the low-rank $\mathbf{L}$ and sparse $\mathbf{S}$ matrices. Furthermore, several authors have added specific constraints in the RPCA formulation to make it suitably designed for the target applications. The rest of this paper is organized as follows. Section III reviews the applications of RPCA in image processing. Particularly, low-level imaging is surveyed in Section III-A while the specific case of medical imaging is then investigated in Section III-B. We review the 3-D computer vision applications in Section III-C. Section IV reviews the applications of RPCA in video processing. Finally, we present the conclusion with future research directions.

III. IMAGE PROCESSING

In image processing, several tasks can be formulated into low-rank and/or sparsity aspects. Thus, the $\mathbf{L}+\mathbf{S}$ decomposition presents a suitable framework for these different tasks. In addition, the spatial aspects in images are exploited in the $\mathbf{L}+\mathbf{S}$ decomposition to enforce its use to the concerned task. In the following sections, we review these different tasks categorized in low-level imaging, medical imaging, and 3-D computer vision.

A. Low-Level Imaging

In low-level processing tasks, RPCA via $\mathbf{L}+\mathbf{S}$ decomposition is of interest in tasks in which 1) the observed image can be viewed as the sum of a low-rank clean image and a sparse perturbations as in image restoration and denoising, hyperspectral image denoising, and image composition; 2) the observed image can be viewed as the sum of a low-rank image and a sparse image as in intrinsic image computation; 3) only the low-rank aspect is of interest as in image alignment, image stitching, image colorization, and pan sharpening; and 4) only the sparse aspect is of interest as in multifocus image fusion.

1) Image Restoration and Denoising: Image restoration is one of the most fundamental problems in image processing and computer vision, especially in the current days with the growing number of cameras and closed circuit monitors. Its goal is to restore a clear image from degraded images. There are two main kinds of degradations: geometric distortion and blur. Lau *et al.* [154] addressed the degradation issues by first optimizing a mathematical model to subsample sharp and mildly distorted video frames, and then applying a two-step stabilization to stabilize the subsampled video with Beltrami coefficients, replacing blurry images with sharp ones by optical flow and robust PCA. In particular, for every frame I_k^{samp}, Lau *et al.* [154] calculated the deformation fields $\mathbf{V}_k^j$ from a fixed frame I_k^{samp} to other ones. Define

$$\begin{aligned}\mathbf{V}_k &:= (\text{vec}(\mathbf{V}_k^1), \text{vec}(\mathbf{V}_k^2), \ldots, \text{vec}(\mathbf{V}_k^n)) \\ &= \mathbf{V}_{k,1} + i\mathbf{V}_{k,2} \end{aligned} \tag{7}$$

where $\text{vec}(\mathbf{V})$ indicates the vectorization of $\mathbf{V}$ and $\mathbf{V}_{k,1} := \text{Re}(\mathbf{V}_k)$ and $\mathbf{V}_{k,2} := \text{Im}(\mathbf{V}_k)$ contain the horizontal and vertical displacement vectors, respectively. They applied robust PCA to decompose each of $\{\mathbf{V}_{k,p} \mid p = 1, 2\}$ into low-rank and sparse terms

$$\begin{aligned}\mathbf{V}_{k,p} &= \mathbf{L}_{k,p}^* + \mathbf{S}_{k,p}^* \\ (\mathbf{L}_{k,p}^*, \mathbf{S}_{k,p}^*) &= \underset{\mathbf{L}+\mathbf{S}=\mathbf{V}_{k,p}}{\text{argmin}} \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1, \quad \text{for } p = 1, 2 \end{aligned} \tag{8}$$

where $\|\mathbf{L}\|_* = \sum_i \sigma_i(\mathbf{L})$ is the nuclear norm, i.e., the sum of singular values $\|\mathbf{S}\|_1 = \sum_{ij} |\mathbf{S}_{i,j}|$ is the ℓ_1-norm, and the sparse part $\mathbf{S}_{k,p}^*$ corresponds to the outlier. Lau *et al.* [154] then warped I_k^{samp} by a postprocessing of low-rank part $\mathbf{L}_k$ for each k to obtain the stabilized frames. Experiments on both synthetic and real experiments demonstrate the effectiveness of the proposed method in alleviating distortions and blur, restoring image details, and enhancing visual quality against several state-of-the-art methods, as can be seen in Fig 1.

Image denoising is a problem closely related to image restoration, where the degradation is caused by noise. The goal of image denoising is to effectively identify and remove noise from the ground-truth image. To this end, many classic image denoising algorithms assume a specific statistical model of the noise, and apply the

Fig. 1. *Image restoration. From top to bottom: the ground-truth image, the distorted and blurred image, the Sobolev gradient-Laplacian method [189], the centroid method [202] deblurred [259], the two-stage reconstruction method [213] and the RPCA algorithm [154]. (Images from [154].)*

maximum-likelihood estimator to do the inference. However, the assumed statistical model, e.g., the Gaussian white noise, cannot always hold true in practice. This observation motivates some new ideas to the problem of image denoising. The seminal work of [34] first proposed the nonlocal self-similarity-based methods for image denoising. The idea is that the repeated local patterns across a natural image may help reconstruct the similar local patches. Inspired by this idea, Gu *et al.* [109] combined with the new technique of weighted nuclear norm minimization to perform image denoising. In particular, for a local patch $\mathbf{y}_i$ in an image, Gu *et al.* [109] searched for its nonlocal similar patches by block matching methods. Then, they stacked those nonlocal similar patches into a matrix $\mathbf{Y}_j$ and decomposed it as $\mathbf{X}_j + \mathbf{N}_j$, where the subscript j indicates the jth class of patches. Intuitively, matrix $\mathbf{X}_j$ should be of low rank as it is stacked by the similar local patches while $\mathbf{N}_j$ corresponds to noise. With this observation, Gu *et al.* [109] proposed to minimize the following objective function with weighted nuclear norm regularization:

$$\min_{\mathbf{X}_j} \lambda \|\mathbf{Y}_j - \mathbf{X}_j\|_F^2 + \|\mathbf{X}_j\|_{w,*} \tag{9}$$

where $\|\cdot\|_{w,*}$ is the weighted nuclear norm defined as $\|\mathbf{X}\|_{w,*} = \sum_i \mathbf{w}_i \sigma_i(\mathbf{X})$, in which $\sigma_i(\mathbf{X})$ is the ith largest singular value of matrix $\mathbf{X}$ and $\mathbf{w} = [\mathbf{w}_1, \ldots, \mathbf{w}_n]^T$ is the nonnegative weight vector. To set an appropriate weight vector $\mathbf{w}$, Gu *et al.* [109] chose $\mathbf{w}_i$ to be inversely propositional to $\sigma(\mathbf{Y}_i)$, thus encouraging low-rank solutions more effectively than the usual nuclear norm. Extensive experiments show that the proposed method called weighted nuclear norm minimization (WNNM) is able to recover more details, generate much less artifacts, and preserve much better edges against the following state-of-the-art methods: BM3D [60], EPLL [356], LSSC [197], NCSR [65], and SAIST [64], as can be seen in Fig. 2. The main drawback to the above approach is that iteratively reweighted algorithms can only approximate either the low-rank component or the sparse one with a nonconvex surrogate. One important reason for this is that it is difficult to solve a problem whose objective function contains two or more nonsmooth terms. In this context, Wang *et al.* [292] employed a Schatten-p norm and ℓ_q-norm regularized principal component pursuit (p, q-PCP) to approximate the low rank and sparse functions with nonconvex surrogates with few iterations. Experiments show that p, q-PCP achieves the best image recovery performance. Liang [171] considered the restoration of a low-rank texture contaminated by both Gaussian and salt-and-pepper noise. The algorithm formulates texture image denoising in terms of solving a low-rank matrix optimization problem.

2) Hyperspectral Image Denoising: Traditional RGB images capture light in the red, green, and blue portions of the visible light spectrum. Each band represents the amount of energy emitted at a particular wave length. Images having more than three bands are referred to as multispectral or hyperspectral images. These images can involve light that is outside the visible spectrum, such as infrared (IR) and ultraviolet (UV) light. Hyperspectral images have a higher spectral resolution compared to multispectral images while being limited to a narrow spectral bandwidth. By imaging the light that is absorbed and reflected in high detail within a certain region of the electromagnetic spectrum, it is possible to identify particular materials present in the image. Thus, hyperspectral images contain rich spectral information which facilitates lots of computer vision tasks. However, hyperspectral data are easily affected by different factors such as noise, missing data, etc., which degrades the image quality and makes hyperspectral data incomplete. Wei *et al.* [296] addressed hyperspectral data denoising in the RPCA formulation by taking advantage of hyperspectral unmixing and modeling it probabilistically. Let $\mathbf{X}$ be an observed 3-D hyperspectral image with $\mathbf{X} \in \mathbb{R}^{n_r \times n_c \times n_b}$ where n_r, n_c, and n_b are the height, width, and the number of bands, respectively. For convenience, $\mathbf{X}$ is rearranged in a 2-D matrix $\mathbf{A}$ by reshaping the image of each band as a vector of $\mathbf{A} \in \mathbb{R}^{n_p \times n_b}$ with $n_p = n_r \times n_c$ which corresponds to the number of pixels. Suppose that a noisy hyperspectral image $\mathbf{A}$ can be decomposed into a noiseless/clean hyperspectral image $\mathbf{L} \in \mathbb{R}^{n_p \times n_b}$ and a noise image $\mathbf{S} \in \mathbb{R}^{n_p \times n_b}$. $\text{rank}(\mathbf{S})$ is always full with $\text{rank}(\mathbf{A}) \approx \text{rank}(\mathbf{S}) = \min(n_p, n_b)$. $\mathbf{L}$ can be represented as a multiplication of a matrix $\mathbf{L}_1 \in \mathbb{R}^{n_p \times n_e}$ with a matrix $\mathbf{L}_2 \in \mathbb{R}^{n_e \times n_b}$, called the endmember matrix and the abundance matrix, respectively. n_e is the number of endmembers and $\text{rank}(\mathbf{L})$ is no larger than n_e (i.e., $\text{rank}(\mathbf{L}) \leq n_e$). Because n_e is far smaller than n_p and n_b, $\text{rank}(\mathbf{L}) \ll \text{rank}(\mathbf{A})$. Thus, $\mathbf{L}$ is effectively a low-rank matrix and we have $\mathbf{A} = \mathbf{L} + \mathbf{S}$. Experimental results show

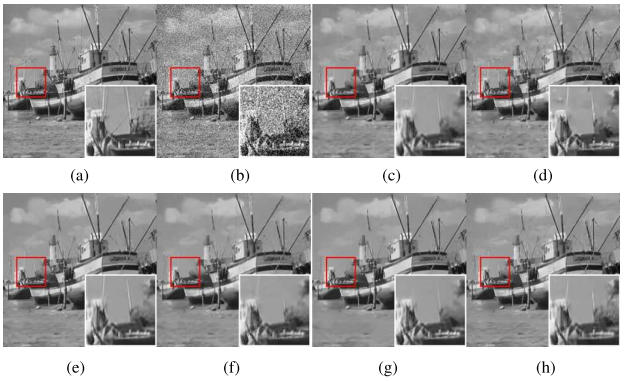


Fig. 2. *Image denoising. From top to bottom: (a) ground-truth image; (b) noisy image, PSNR: 14.16 dB; (c) BM3D, PSNR: 26.78 dB [60]; (d) EPLL, PSNR: 26.65 dB [356]; (e) LSSC, PSNR: 26.77 dB [197]; (f) NCSR, PSNR: 26.65 dB [65]; (g) SAIST, PSNR: 26.63 dB [64]; and (h) RPCA algorithm called WNNM, PSNR: 26.98dB [109]. (Images from [109].)*

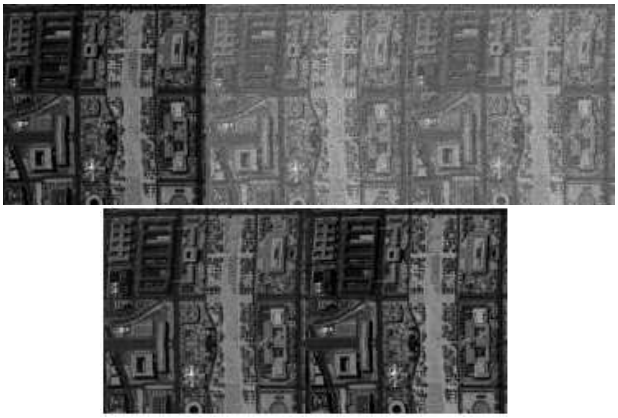


Fig. 3. *Hyperspectral image denoising results of band 100 for Washington DC Mall data set. From top to bottom: original band, noisy band, the wavelet-based result, GoDec [350], and RPCA algorithm [296]. (Images from [296].)*

that RPCA algorithms outperform the standard approach based on wavelet, as can be seen in Fig. 3 on the Washington DC Mall data set[2] which contains hyperspectral images of size 1208×307 pixels. Each image has 191 spectral channels and a subimage of size $256 \times 256 \times 191$ is cropped from this data set. Even if this RPCA-based method outperforms the state-of-the-art methods, real noise in hyperspectral date often exhibits very complex statistical distributions, rather than simply being sparse. So the noise cannot be easily described by a simple norm such as the ℓ_1-norm. From the probabilistic perspective, the low-rank part $\mathbf{L}$ and the noise part $\mathbf{S}$ can be modeled more directly and flexibly with a generative model using a mixture of Gaussians model as in the MOG-RPCA model [344], or using a mixture of exponential power (MoEP) distributions as in the penalized MoEP (PMoEP) model [40]. Experiments show that this probabilistic method can denoise noisy incomplete hyperspectral data more effectively when compared with previous denoising methods.

3) Image Composition: Image composition is the problem of combining multiple images captured by a camera or multiple cameras to generate a desired image of a scene. A typical example is a high-contrast scene captured by a low-dynamic range (LDR) camera. It has many important applications in computational photography, such as high dynamic range (HDR) imaging and flash/no-flash imaging. Classic techniques for this problem suffer from issues caused by defocus blur and dynamic objects which typically results in ghosting artifacts.

Bhardwaj and Raman [21] addressed the aforementioned issues by the robust PCA framework. Specifically, they first modeled the camera response function by a gamma correction function to linearize the intensity values. This operation is applied to all n images and a matrix $\mathbf{A}$ is constructed by stacking each image as a column of $\mathbf{A}$. They then applied robust PCA to $\mathbf{A}$

$$\min_{\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1, \quad \text{s.t. } \mathbf{A} = \mathbf{L} + \mathbf{S} \tag{10}$$

which is solved by the alternating direction method of multipliers (ADMM) [176]. Next, they used the inverse of the gamma correction to the columns of $\mathbf{L}$ to obtain the high-contrast LDR images which are free from defocus blur and ghosting artifacts. Finally, they fused the obtained high-contrast LDR images into a high-quality HDR image by an existing method. The motivation here is that the static part of the scene presented in all images should be of low rank ($\mathbf{L}$) as they are similar to each other. This RPCA technique penalizes the lower singular values while retaining the higher singular value. Experiments on multiple-exposure images and multiple-aperture images show that the proposed method can capture better contrast details and have less defocus blur and specularities, as can be seen in Fig. 4.

4) Intrinsic Image Computation: Intrinsic image computation aims at separating a given input image into its material-related properties, such as reflectance or albedo, and its light-related properties, e.g., shading and shadows. It is probably one of the most important preprocessing steps for photometric computer vision and image based modeling.

The seminal work by Candès *et al.* [37] first proposed to apply robust PCA to compute the intrinsic image of face images. The idea is simple: by stacking multiple facial images from the same person taken under different lighting as the columns of a matrix $\mathbf{A}$, $\mathbf{A}$ should be decomposed as $\mathbf{L} + \mathbf{S}$ with a low-rank matrix $\mathbf{L}$ and a sparse matrix $\mathbf{S}$. This idea utilizes the fact that the intrinsic image, which reflects the light reflectance properties of a face, is common for the face images taken under a different lighting. The decomposition can thus be done by solving the robust PCA problem (10).

However, inappropriate choice of the regularization parameter λ between the low-rank and sparse matrices

Fig. 4. *Image composition (HDR). From left to right: Tone mapped [91], and multiexposed images obtained with the RPCA-based approach [21]. The differences are noticeable within the regions delimited by a circle. (Images from [21].)*

[2]http://lesun.weebly.com/hyperspectral-data-set.html

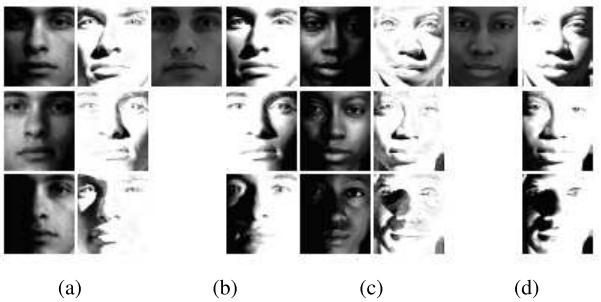


Fig. 5. *Intrinsic image computation on two subsets of the Extended Yale B database. From left to right (by group of six images=: (a) RPCA results on subset 18; (b) rank-constrained RPCA results on subset 18; (c) RPCA results on subset 22; and (d) rank-constrained RPCA results on subset 22. Note that the reflectance image remains the same for the rank-constrained RPCA. (Images from [324].)*

in the classic robust PCA problem often results in an $\mathbf{L}$ with a rank greater than one, while for intrinsic image computation, the rank of $\mathbf{L}$ should be one, as there should be only one intrinsic image. To resolve this issue, Yu [324] proposed the rank-constrained PCA (RCPCA) model, by explicitly enforcing the rank of $\mathbf{L}$ to be one

$$\min_{\mathbf{L},\mathbf{S}} \|\mathbf{S}\|_1, \quad \text{s.t. } \text{rank}(\mathbf{L}) = 1, \qquad \mathbf{A} = \mathbf{L} + \mathbf{S}. \tag{11}$$

The above model can also be easily solved by ADMM [176]. Experiments on the MIT intrinsic image data set and the Yaleface data set (see Fig. 6) show that the proposed fixed rank model in (11) enjoys a lower local mean squared error than the prior methods for intrinsic image computation. Similarly, Leow *et al.* [156] used a different norm, the Frobenius norm, for the matrix $\mathbf{S}$ in (11) and reasonably good intrinsic images were obtained.

5) Image Alignment and Rectification: Image alignment refers to the problem of transforming different images into the same coordinate system. It is a critical preprocessing step for multiple applications, such as background modeling, where the frames of a video are assumed to be aligned in order to obtain a low-rank background. Practically, robust and efficient image alignment remains a challenging task, due to the massiveness of images, great illumination variations between images, partial occlusion, and corruption. Peng *et al.* [226] first proposed robust alignment by sparse and low rank (RASL) to solve the problem based on the assumption that a batch of aligned images should form a low-rank matrix $\mathbf{L}$. The sparse component $\mathbf{S}$ models local differences among images. Let $\mathbf{A}$ be the matrix which stacks each frame as its one column, then the mathematical model of RASL is similar to robust PCA but with a characterization of geometric deformation τ

$$\min_{\tau,\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1, \quad \text{s.t. } \mathbf{A} \circ \tau = \mathbf{L} + \mathbf{S} \tag{12}$$

where $\mathbf{A} \circ \tau$ refers to applying framewise geometric deformation τ to each frame. For efficient solution, Peng *et al.* [226] converted the nonconvex problem to a computable convex optimization by iteratively linearizing τ locally and updating with the increment $\Delta\tau$ of τ

$$\begin{cases} \min_{\Delta\tau_k,\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \lambda\|\mathbf{S}\|_1, \text{ s.t. } \mathbf{A} \circ \tau_k + \mathbf{J}\Delta\tau_k = \mathbf{L} + \mathbf{S} \\ \tau_{k+1} \leftarrow \tau_k + \Delta\tau_k \\ k \leftarrow k + 1. \end{cases} \tag{13}$$

Here $\mathbf{J}$ is the Jacobian of $\mathbf{A} \circ \tau$ w.r.t. the parameters of deformation τ. The above convex optimization problem can also be efficiently solved by ADMM [176], and the solution of (13) converges to solution of (12). An improved optimization method for RASL can be found in [242], where $\Delta\tau_k$ is canceled first. Such a local linearization algorithm can be viewed as a Gauss–Newton method for minimizing the composition of a nonsmooth convex function with a smooth, nonlinear mapping. The convergence behavior of such algorithms has been established in the literature [226]. There are many generalizations of RASL. For example, Wu *et al.* [304] proposed a method for ORIA by employing a fixed-rank model along with a basis update scheme and by assuming that the aligned image without corruption is a linear composition of well-aligned basis. Although quite efficient on large data sets, the heuristic basis updating scheme using thresholding and replacement reduces the robustness of image alignment. Motivated by online robust PCA, Song *et al.* [270] took advantage of closed-form solutions and a stochastic gradient descent (SGD) updating scheme, which have better convergence performance. However, as well as RASL, ORIA [304] and SGD [270] all assume that large errors such as occlusion and corruption among the images are sparse and separable with respect to intensity, which may fail in aligning images with severe intensity distortions. To address this limitation, Zheng *et al.* [348] employed an online image alignment method via subspace learning from image gradient orientations (IGOs). Fig. 6 shows alignment of faces from the Labeled Faces in the Wild (LFW) data set [129].

Image rectification is a similar task as image alignment, both of which are to deform (one or more) images into a "standard" coordinate system. However, the difference is that, instead of transforming multiple images into the same coordinate system as in the alignment problem, image rectification has only access to one image, which is more challenging. Transform invariant low-rank textures (TILT) [339] provides a possible solution to this problem. The intuition is as follows: an image, viewed as a matrix, should be of approximately low rank if it is in its regular status, e.g., being symmetric or periodic. Interestingly, TILT has the same mathematical model (12) as RASL, and the solution methods of TILT and RASL are also identical. The only difference is on the interpretation

of matrix $\mathbf{D}$. In TILT $\mathbf{D}$ consists of an image patch of a single image, while in RASL $\mathbf{D}$ consists of a collection of images, stacked in columns. Therefore, RASL and TILT are complementary to each other in that they try to capture temporal and spatial correlation among image(s), respectively.

There are some other generalizations and improvements of TILT. For example, Zhang *et al.* [341] considered the parameterized transformations of TILT, in particular, generalized cylindrical transformations, which can be conveniently applied to unwrapping textures from buildings. Zhang *et al.* also applied TILT to text rectification [334] and text detection [335].

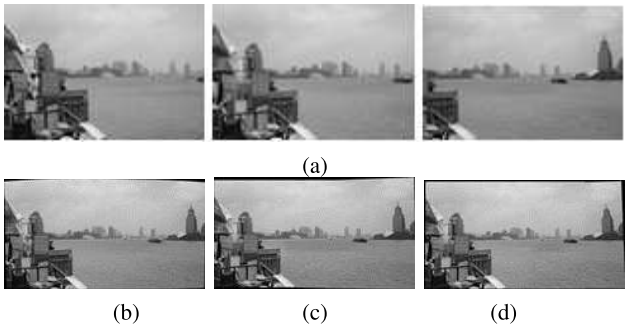
6) Image Stitching: Image stitching refers to the problem of aligning and stitching multiple images. It has many applications in computer vision and computer graphics, such as video stabilization and augmented reality. Despite significant progress on this problem, many methods have limited robustness to occlusions and local object motion in different captures. In order to remove this obstacle, Li and Monga [168] formulated the alignment problem as a low-rank and sparse matrix decomposition problem with incomplete observations, and the stitching problem as a multiple labeling problem that utilizes the sparse components. Their model is

$$\min_{\tau,\mathbf{L},\mathbf{S}} \|\mathbf{S}\|_1, \quad \text{s.t. } \mathbf{A}\circ\tau = \mathcal{P}_\Omega(\mathbf{L}+\mathbf{S}), \quad \mathsf{rank}(\mathbf{L}) \le r \quad (14)$$

where $\mathbf{A}$ is constructed by stacking each image as one column, τ models the geometric transformation on each image, and $\mathcal{P}_\Omega$ is the standard projection operator on the observed set Ω due to the fact that each image is a partial observation of the underlying mosaics in terms of pixels values. Thus, the columns of the output matrix $\mathbf{L}$ are the desired aligned images. Problem (14) can also be solved by ADMM [176]. With a few postprocessing steps, the multiple images can be stitched together. Experiments on the Shanghai data set (see Fig. 7) and the Windows data set show that the proposed method creates much less ghosting artifacts and blur than the prior methods.

(a) (b)

Fig. 6. *Image alignment. From left to right: (a) original images from the LFW data set [129]; and (b) result obtained by ORIA [304]. (Images from Wu [304].)*



(a) (b) (c) (d)

Fig. 7. *Image stitching: (a) input images from the Shanghai data set; (b) Brown and Lowe [33]; (c) Gao et al. [95]; and (d) SIASM [168]. (Images from [168].)*

7) Image Colorization: Image colorization is the problem of recovering the original color of a monochrome image from only a few user-provided color pixels. A strategy to solve this problem is by matrix completion, which assumes that the underlying color image is of low rank. Though it is shown that many images can be modeled by low-rank matrices, the low-rank assumption is typically untrue for the coloring of a global image but holds true for local similar patches of the images. With this observation, Yao and Kwok [317] achieved image colorization by Patch-based Local Low-Rank (PaLLR) matrix completion. In particular, instead of assuming the whole $m\times n$ image to be low rank, they first extracted groups of similar image patches, each of which has its own low-rank structure. The extraction of similar patches is by the similarity measure

$$\|P_{i,j} - P_{i',j'}\|_F^2 + \beta\left(\frac{1}{m^2}(i-i')^2 + \frac{1}{n^2}(j-j')^2\right) \quad (15)$$

between the patch $P_{i,j}$ at position (i,j) and the patch $P_{i',j'}$ at position (i',j'), where $\beta>0$ is a tradeoff parameter. For each image patch, denote by $\mathbf{G}$ the matrix that contains the k most similar patches including itself. Yao and Kwok [317] proposed to minimize the local low-rank approximation problem

$$\min_{\mathbf{L}} \frac{1}{2} \underbrace{\|\mathbf{LT}-\mathbf{G}\|_F^2}_{\text{consistency with gray values}} + \underbrace{\frac{\lambda}{2}\|\Omega\odot(\mathbf{L}-\mathbf{O})\|_F^2}_{\text{sparse labeled errors}} + \underbrace{\mu|\mathbf{L}|_*}_{\text{local low-rank}} \quad (16)$$

where $\odot$ is the dot product which is the sum of the products of the corresponding entries and returns a single number, $\mathbf{T}$ is the color-to-monochrome linear transform, and $\mathbf{L}$ is the target colorization of $\mathbf{G}$. $\mathbf{O}$ and Ω indicate the values and positions of the color pixels in the k patches, respectively. The effectiveness of PaLLR is guaranteed by the observation that the singular value spectrum of a

typical patch group decays quickly, i.e., the patch is low rank. Finally, the color of a pixel is obtained by averaging the color values in patches that contain the pixel.

8) Multifocus Image Fusion: Robust PCA has important applications in multifocus image fusion as well. Existing imaging devices, such as autofocus cameras, have limited focus range: only objects around a particular depth are clear while other objects are blurry. Multifocus image fusion aims at resolving this issue: it creates a single image in which all scene objects appear sharp. It has many applications in digital photography and computer vision.

Wan *et al.* [288], [289] formulated the problem of multifocus image fusion as choosing the most significant features from a sparse matrix which is obtained by robust PCA to form a composite space of features. They then integrated the local sparse features that represent the salient regions of the input images to construct the desired fused images. Their method consists of five steps. 1) Stack the images with different focuses as columns of matrix $\mathbf{A}$. 2) Perform robust PCA decomposition (10) on matrix $\mathbf{A}$ so as to obtain the low-rank matrix $\mathbf{L}$ and the sparse matrix $\mathbf{S}$. Unstack each column of $\mathbf{S}$ into multiple matrices, each of which corresponds to one source image. 3) Divide the resultant matrices into small blocks. Choose the blocks with lager standard deviations to construct the fused image, with a sliding window technique. 4) Record the feature comparison results. 5) Apply a consistency verification process to refine the decision map by a majority filter. In a further work, Zhang *et al.* [336], [337], [338] proposed to use the pulse-coupled neural network (PCNN) to record the feature comparison results. The advantage is that the biological characteristic of PCNN is able to take full use of the local features obtained from sparse matrices and improve the accuracy of determining in-focus objects.

9) Pan Sharpening: With the development of optical sensors, more and more remote sensing images are collected, with numerous applications in environment monitoring, battlefield reconnaissance, etc. Unfortunately, due to the uncontrolled environments and some physical limitations of sensors, images from a single sensor typically have low spatial and spectral resolution. The technique of pan sharpening is designed to resolve the issue: it fuses the panchromatic (PAN) image with the low-resolution multispectral (LRMS) images to generate the synthetic high-resolution multispectral (HRMS) images with high spatial and spectral resolutions.

Yang *et al.* [333] proposed to apply low-rank decomposition to the problem of pan sharpening with spatial–spectral offsets. The idea is that the spatial redundancy and spectral correlation among the multispectral images naturally imply the inherent low-rank property of the matrix formed by stacking HRMS images together. To be more specific, denote by $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_n]$ the matrix by stacking the bands of n LRMS images, each being a column of $\mathbf{A}$, and let $\mathbf{L} = [\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_n]$ be the matrix of stacking the n HRMS images. Yang *et al.* [333] decomposed $\mathbf{A}$ as the sum of $\mathbf{L}$ and two offset matrices $\mathbf{S}_1$ and $\mathbf{S}_2$

$$\mathbf{A} = \mathbf{L} + \mathbf{S}_1 + \mathbf{S}_2. \tag{17}$$

The spatial offset matrix $\mathbf{S}_1$ counteracts the spatial details in HRMS images while the spectral offset matrix $\mathbf{S}_2$ contains the information of spectral changes between the LRMS and HRMS images, both of which should be sparse. Matrix $\mathbf{L}$ should be of low rank due to the spatial and spectral correlations among the HRMS images. Besides, the PAN image $\mathbf{P}$ can be viewed as the spectral degradation of HRMS images. Therefore, the PAN image can be represented by the HRMS image: $\mathbf{P} = \mathbf{LW}$ for some representation coefficient matrix $\mathbf{W}$. So the pan-sharpening problem can be formulated as the optimization problem

$$\begin{aligned} \min_{\mathbf{L},\mathbf{S}_1,\mathbf{S}_2,\mathbf{W}} \ & \|\mathbf{L}\|_* + \alpha\|\mathbf{S}_1\|_1 + \beta\|\mathbf{S}_2\|_1 \\ \text{s.t. } & \mathbf{A} = \mathbf{L} + \mathbf{S}_1 + \mathbf{S}_2, \quad \mathbf{LW} = \mathbf{P} \end{aligned} \tag{18}$$

where α and β are the regularization parameters. With additional physical constraints on $\mathbf{S}_1$ and $\mathbf{S}_2$ and solving problem (18) by ADMM [176], extensive experiments show that the calculated spatial and spectral offsets $\mathbf{S}_1$ and $\mathbf{S}_2$ are able to approach the reference differences well, implying that the fused images by the two offsets are of high quality.

10) Face Modeling and Recognition: Robust face modeling under uncontrolled conditions is crucial for the face recognition systems, and it is a prestep before face recognition. Common objects, such as sunglasses and scarves, cause facial features partially occluded. Fig. 8 shows an example with face images of size 84×60 pixels the AR data set [200]. For example, Luan *et al.* [190] used 15 images for an individual that are stacked in the observed matrix A. The first row of Fig. 8 shows eight images of

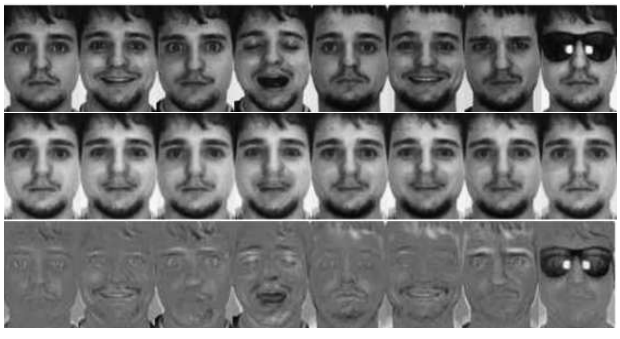


Fig. 8. ***Face modeling: Removal of facial occlusion using RPCA. First row: Original face images with facial expression and occlusion. Second row: Low-rank images. Third row: Sparse error images. (Images from [190].)***

the same individual with varied facial expression and contiguous occlusion (sunglasses). The low-rank components $\mathbf{L}$ among different images look very similar, even if in the presence of expressions and occlusion. The sparse errors $\mathbf{S}$ depict the difference between original and corresponding low-rank face images. In the case of sunglasses occlusion, nothing but a pair of sunglasses can be seen from the error image. In a probabilistic approach, Cao *et al.* [40] modeled the low-rank part $\mathbf{L}$ and the noise part $\mathbf{S}$ with a generative model using a mixture of exponential power (MoEP) distributions. This model called penalized MoEP (PMoEP) outperforms both the Gaussian model (RPCA-MOG [203]) and the Laplacian model (RegL1ALM [349]) on the Extended Yale B database.

Robust face recognition, i.e., automatically recognizing human faces from views with varying expression and illumination as well as disguise and occlusion, is one of the most important problems in computer vision and pattern recognition. The basic problem in face recognition is to use labeled training data from k distinct classes to correctly identify the class of a new test sample. The challenge is that the disguise and occlusion may heavily degrade the performance of the traditional methods.

To robustify the existing methods, Wright *et al.* [300] proposed to use sparse representation to perform face recognition. The intuition is based on a simple observation that the test samples are approximately representable by a linear combination of those training samples from the same class. Therefore, the representation should be sparse, involving only a small fraction of samples in the same class, and is robust to outliers/disguise/occlusions. Pursuing the sparsest representation naturally discriminates between various classes by the following convex optimization problem:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x}\|_1, \quad \text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{y}. \tag{19}$$

Here, as usual $\mathbf{A}$ is the matrix formed by stacking each training sample as a column of the matrix and $\mathbf{y}$ is a column vector corresponding to the test image. Finally, the given test image is assigned to the class with the smallest reconstruction error by the representation coefficient $\hat{\mathbf{x}}$.

However, the sparsest representation (19) is not robust to large contiguous occlusion such as scarf and sunglasses. To mitigate the issue, rather than minimizing the sparse representation model (19), Luan *et al.* [190] proposed an approach for robust face recognition by exploiting the sparse term obtained by robust PCA (10). In particular, they first constructed a matrix of normalized (training and testing) samples by stacking all facial images as the columns of the matrix. Their algorithm then applies robust PCA to the constructed matrix. Focusing on the sparse term obtained by robust PCA, Luan *et al.* [190] combined sparsity descriptor and smoothness descriptor to characterize the similarity between a testing image and any given class. The testing image is finally assigned to the class with the highest similarity. Practically, RPCA is employed for removal of facial specularities and shadows, and for removal of facial occlusion. Experiments show that the associated sparse term by robust PCA exhibits more discriminative information for face identification, being more robust to varying illumination and pixel corruption on both synthetic and real data sets (Yale Face Database,[3] Extended Yale Face Database B,[4] and AR face database[5] [200]). As an illustration, Fig. 9 shows the decomposition of a test face image under different subjects using RPCA.

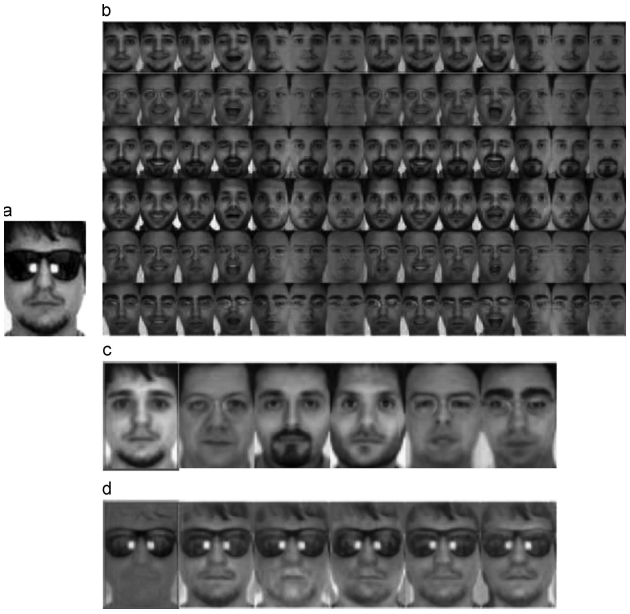


Fig. 9. *Face recognition: Low-rank and sparse error images of a given test image. (a) Test image. (b) Training images of six individuals. (c) Low-rank images of the test image under six individuals. (d) Corresponding sparse error images. (Images from [190].)*

11) Partial-Duplicate Image Search: Partial-duplicate image search refers to the problem of searching images from a database containing the same contents as the query image. The challenge is that the retrieved images might be modified versions of the query image, such as the changes in color, scale, rotation, and contrast, having partial occlusions and different viewpoints, etc. Traditional methods, e.g., the bag of visual words, suffer from low retrieval precision and recall as they only consider the local features and the feature quantization error may easily lead to false matches among images. To remedy these issues, Yang *et al.* [313] introduced the global geometric consistency to detect the false matches by a low-rank model. They noticed that the rank of the squared Euclidean distance matrix between the feature points is at most 4 when the matchings are correct. In contrast, when there are false matches between feature points, the stacked

[3]http://vision.ucsd.edu/content/yale-face-database
[4]http://vision.ucsd.edu/content/extended-yale-face-database-b-b
[5]http://www2.ece.ohio-state.edu/aleix/ARdatabase.html

squared distance matrix should be of higher rank. Applying robust PCA to the stacked squared Euclidean distance matrices, false matches can be detected effectively.

12) Saliency Detection: Saliency detection in still image is a crucial step for improving visual experience, which has many applications such as image cropping, image collection browsing, video compression, etc. The goal of image saliency detection is to find the image regions in which one or more features differ significantly from their surroundings. In other words, if we use other regions to "predict" the selected salient region, the representation error should be large. Based on this observation, Lang *et al.* [152] proposed a method called multitask sparsity pursuit (MTSP) which decomposes the feature matrix $\mathbf{A}$ of image patches into a highly correlated low-rank component $\mathbf{AZ}$ and a sparse salient component $\mathbf{S}$

$$(\mathbf{Z}^*, \mathbf{S}^*) = \underset{\mathbf{Z},\mathbf{S}}{\operatorname{argmin}} \|\mathbf{Z}\|_* + \lambda\|\mathbf{S}\|_{2,1}, \quad \text{s.t. } \mathbf{A} = \mathbf{AZ} + \mathbf{S}. \tag{20}$$

The idea is that by breaking an image into patches with the extracted features $\mathbf{A}$ (stacked by columns as usual), the salient regions should correspond to those with large sparse noise in $\mathbf{S}$. Lang *et al.* [152] then defined the score function $S(P_i)$ for the ith patch P_i by $S(P_i) = \|\mathbf{S}^*_{:i}\|_2$. The salient regions are then identified by a threshold which is set to discard small $S(P_i)$'s. Fig. 10 shows that MTSP obtained competitive results against state-of-the-art methods even if a standard approach named FT [1] offers the best overall performance.

Note that the model in (20) is actually called low-rank representation (LRR) [180], [179]. It also has wide applications in image processing, such as motion segmentation [180], [179], image segmentation [53], and image tag completion and refinement [127]. More thorough investigations on LRR can be found in [178].

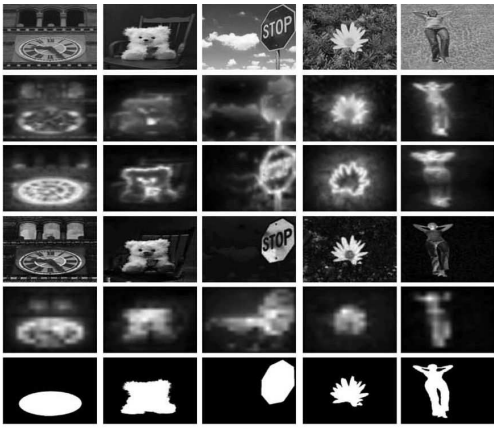


Fig. 10. *Saliency detection: Comparison on images from MSRA data set [1]. The rows for top to bottom are: original images and saliency maps produced by GBVS [120], CSD [103], FT [1], and the RPCA algorithm (MTSP) [152], respectively. The last row is the ground truth. (Images from [152].)*

Instead of working on all the image, Li and Haupt [165], [166] estimate the saliency map directly from compressive samples in applications where the goal is to identify regions of anomalous behavior rather to image the entire scene. Furthermore, saliency detection is also addressed as anomaly detection in spectral images [230], [236]. Thus, anomaly detection is viewed as a matrix decomposition problem with the minimum volume constraint for the multimodular background and sparsity constraint for the anomaly image pixels.

In summary, the RPCA formulation provides better or similar performances than previous state-of-the-art methods over these 12 low-level processing tasks.

B. Medical Imaging

In medical imaging, the $\mathbf{L} + \mathbf{S}$ decomposition was used for applications in which the observed image can be considered as the sum of a low-rank clean image and a sparse perturbations as in background suppression in accelerated dynamic [218] and in change detection [90]. In the application of joint image reconstruction and registration, only the low-rank aspect is of interest as it concerned image alignment [222].

1) Accelerated Dynamic MRI With Automated Background Suppression: Dynamic MRI techniques acquire a time series of images that encode physiological information of clinical interest, such as organ motion [12], contrast agent uptake [11], [151], signal relaxation [253], among others. The acquisition of each time point needs to be short relative to the dynamic process to obtain an instantaneous snapshot. However, MRI hardware is not fast enough to sample k-space (Fourier space of the image) for each time point at the Nyquist/Shannon rate, particularly if the required spatial and temporal resolution is high and/or volumetric coverage is large. As a consequence, spatial resolution and/or volumetric coverage are usually sacrificed for temporal resolution. Dynamic MRI has a real need for speed.

Given the extensive spatiotemporal correlations in the series of images of dynamic MRI, acquiring fully sampled images at each time point is a wasteful process since the information that is common to all frames is sampled over and over again. Not surprisingly, a number of methods have been developed to acquire undersampled k-space data at each time point and exploit spatiotemporal correlations in order to reconstruct a time series of images without aliasing artifacts [282]. For example, the application of compressed sensing to dynamic MRI [196], [216] exploits temporal sparsity along with incoherent sampling to reduce the number of measurements needed at each time point without information loss. RPCA or low-rank plus sparse ($\mathbf{L}+\mathbf{S}$) decomposition can be applied in the context of compressed sensing to replace the pulse sparsity model by a $\mathbf{L} + \mathbf{S}$ model, where $\mathbf{L}$ would represent the common background among all frames and $\mathbf{S}$ the frame-by-frame innovation. $\mathbf{L}+\mathbf{S}$ reconstruction of undersampled dynamic

MRI data is performed by solving [218]

$$[\mathbf{L},\mathbf{S}] = \underset{\mathbf{L},\mathbf{S}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{E}(\mathbf{L}+\mathbf{S})-\mathbf{d}\|_2^2 + \lambda_{\mathbf{L}}\|\mathbf{L}\|_* + \lambda_{\mathbf{S}}\|T(\mathbf{S})\|_1. \quad (21)$$

Here T is a linear sparsifying transform for $\mathbf{S}$, $\mathbf{E}$ is the encoding operator, and $\mathbf{d}$ is the undersampled $k-t$ data. $\mathbf{L}$ and $\mathbf{S}$ are defined as space-time matrices, where each column is a temporal frame, and $\mathbf{d}$ is defined as a stretched out single column vector. We assume that $\mathbf{S}$ has a sparse representation in some transformed domain (e.g., temporal frequency domain, temporal finite differences), hence the idea of minimizing $\|T(\mathbf{S})\|_1$ and not $\|\mathbf{S}\|_1$ itself. For a single-coil acquisition, the encoding operator $\mathbf{E}$ performs a frame-by-frame undersampled spatial Fourier transform. For acquisition with multiple receiver coils, $\mathbf{E}$ is given by the frame-by-frame multicoil encoding operator, which performs a multiplication by coil sensitivities followed by an undersampled Fourier transform. The multicoil reconstruction case enforces a joint multicoil $\mathbf{L}+\mathbf{S}$ model, which presents improved performance over enforcing a coil-by-coil $\mathbf{L}+\mathbf{S}$ model due to the exploitation of intercoil correlations, as demonstrated previously for the combination of compressed sensing and parallel imaging based on joint multicoil sparsity [170], [216]. $\mathbf{L}+\mathbf{S}$ reconstruction aims to simultaneously 1) remove aliasing artifacts in the space-time domain (or equivalently to estimate the value of nonsampled points in $k-t$ space); and 2) separate the resulting spatiotemporal low-rank and sparse components. Fig. 11 shows the application of $\mathbf{L}+\mathbf{S}$ reconstruction for 4-D contrast-enhanced liver MRI (3-D + time), where $\mathbf{L}+\mathbf{S}$ presented improved spatiotemporal resolution with respect to compressed sensing (sparsity alone) and the automatic background suppression in the $\mathbf{S}$ component improved the visualization of contrast enhancement. $\mathbf{L}+\mathbf{S}$ compares favorably to CS, which suffers from spatiotemporal blurring. Moreover, the $\mathbf{S}$ component, in which the background has been suppressed, offers improved visualization of contrast enhancement.

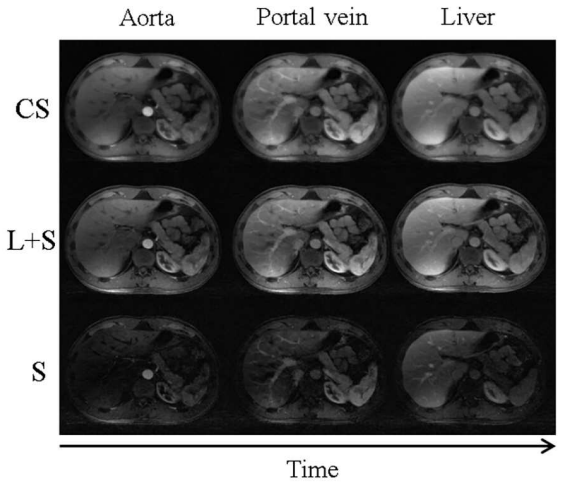


Fig. 11. *CS (sparsity-only) and $\mathrm{L}+\mathrm{S}$ reconstruction of 4-D dynamic contrast-enhanced abdominal data acquired with golden-angle radial sampling (eight spokes per frame, undersampling factor is 48 and temporal resolution is 0.94 s per 3-D volume) corresponding to a representative slice and three contrast-enhancement phases (aorta, portal vein, liver).*

2) Joint Image Reconstruction and Registration: The superposition of organ motion with the physiological process of interest (e.g., contrast enhancement) introduces significant challenges for reconstruction of undersampled data based on spatiotemporal sparsity [149], [221], [283] (including the $\mathbf{L}+\mathbf{S}$ reconstruction approach). Organ motion causes misalignment among temporal frames, which reduces the degree of temporal correlations; consequently, the low-rank and sparsity assumptions break down. Under these conditions, $\mathbf{L}+\mathbf{S}$ reconstruction introduces temporal blurring, leading to nondiagnostic information, or even worse, information that can lead to a false diagnosis. Using ideas from computer vision RPCA techniques such as TILT [339] and RASL [226], the $\mathbf{L}+\mathbf{S}$ model can be modified to include an interframe motion operator $\mathbf{W}$ that describes the deformation between consecutive frames, this is, $\mathbf{M} = \mathbf{W}(\mathbf{L}+\mathbf{S})$. Optical flow [126] can be used to estimate motion between consecutive frames. For frames $\mathbf{M}_{t-1}$ and $\mathbf{M}t$, the optical flow constraint is

$$\begin{aligned} 0 &= \mathbf{M}_{t-1} - \mathbf{M}_t + \frac{\partial \mathbf{M}_t}{\partial x} w_{x,t} + \frac{\partial \mathbf{M}_t}{\partial y} w_{y,t} \\ 0 &= \mathbf{M}_t + \nabla \mathbf{M}_t \mathbf{W}_t \end{aligned} \quad (22)$$

where $\mathbf{W}_t = \begin{pmatrix} w_{x,t} \\ w_{y,t} \end{pmatrix}$ is the unknown motion field for the frame $\mathbf{M}_t$. This linear system is undetermined since there are two unknowns and only one equation. We follow the solution proposed by Thirion [276], also known as the demons method, which corresponds to a second-order gradient descent on the sum of squares difference between $\mathbf{M}_{t-1}$ and $\mathbf{M}_t$

$$\mathbf{W}_t = \frac{\nabla \mathbf{M}_t \cdot \mathbf{M}_t}{\|\nabla \mathbf{M}_t\|^2 + \|\mathbf{M}_t\|^2}. \quad (23)$$

Motion-guided $\mathbf{L}+\mathbf{S}$ reconstruction [222] aims to estimate $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{W}$ from undersampled data only. The dependency between $\mathbf{L}+\mathbf{S}$ and $\mathbf{W}$ makes the optimization problem nonconvex and alternating optimization was employed to update $\mathbf{L}$ and $\mathbf{S}$ with fixed $\mathbf{W}$, and *vice versa*, update $\mathbf{W}$ with fixed $\mathbf{M}+\mathbf{L}+\mathbf{S}$, as follows:

$$[\mathbf{L}_{k+1},\mathbf{S}_{k+1}] = \underset{\mathbf{L},\mathbf{S}}{\operatorname{argmin}} \frac{1}{2}\|\mathbf{E}\mathbf{W}_k(\mathbf{L}+\mathbf{S})-\mathbf{d}\|_2^2 + \lambda_{\mathbf{L}}\|\mathbf{L}\|_* + \lambda_{\mathbf{S}}\|\mathbf{T}\mathbf{S}\|_1 \quad (24)$$

$$\mathbf{W}_{k+1} = \frac{\nabla(\mathbf{L}_{k+1}+\mathbf{S}_{k+1})\cdot \mathbf{m}_{k+1}}{\|\nabla(\mathbf{L}_{k+1}+\mathbf{S}_{k+1})\|^2 + \|\mathbf{m}_{k+1}\|^{2''}} \quad (25)$$

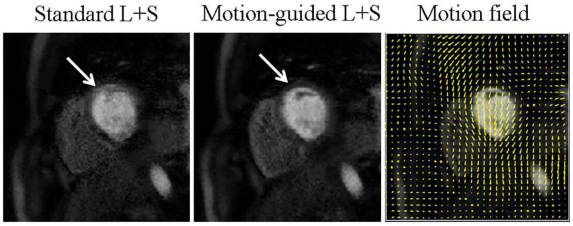


Fig. 12. *Standard and motion-guided* $\mathrm{L}+\mathrm{S}$ *reconstruction of eightfold accelerated free-breathing cardiac perfusion MRI for a representative contrast phase and slice. The arrows indicate temporal blurring artifacts in standard* $\mathrm{L}+\mathrm{S}$ *caused by misalignment among frames, which are significantly removed by motion-guided* $\mathrm{L}+\mathrm{S}$*. In addition, motion-guided* $\mathrm{L}+\mathrm{S}$ *enables access to motion fields between consecutive frames.*

where $\mathbf{m}_{k+1}$ is a vector that concatenates the differences between consecutive frames from $\mathbf{L}_{k+1}+\mathbf{S}_{k+1}$. Here, the first step reconstructs and registers the dynamic image using the previous update on the motion fields, and the second step updates the motion fields based on the current dynamic image update using the demons algorithm. Motion-guided $\mathbf{L}+\mathbf{S}$ exploits an inherent self-consistency between the $\mathbf{L}+\mathbf{S}$ model and image registration, that is, the rank of $\mathbf{L}$ will be lowest and the sparsity of $\mathbf{S}$ will be highest when temporal frames are registered, and *vice versa*, to perform image reconstruction and registration jointly. Fig. 12 shows the application of motion-guided $\mathbf{L}+\mathbf{S}$ to free-breathing eightfold accelerated cardiac perfusion MRI data, where in addition to improved reconstruction, motion fields that describe interframe motion are estimated as an additional piece of information.

3) Change Detection: Change detection between at least two images of the same scene at different time is of widespread interest in many applications including medical imaging, remote sensing, and so on [237]. Fu *et al.* [90] presented a change detection method based on RPCA for retinal fundus images. After alignment and illumination correction, each couple of temporal images considered is expanded into an image serial through linear interpolation between the gray image and the normalized one to progressively decrease the intensity variation between two frames. Then, the linear interpolation images between the gray image and the normalized one are used for the RPCA decomposition to obtain the change mask. Suppose that the given interpolated longitudinal retinal fundus images are of N frames of size $M=m\times n$, and each frame $\mathbf{A}_i$ with $i=1,\ldots,N$. Vectorizing these frames and concatenating them together, one can obtain an image matrix $\mathbf{A}$ of size $M\times N$. Matrix $\mathbf{A}$ is then decomposed as $\mathbf{L}+\mathbf{S}$. The decomposition can thus be done by solving the robust PCA problem (10). The sparse component contains the changes between the two images, which are lesion. Extensive experimental results made on real clinical medical cases show that this method is of lower complexity and higher effectiveness compared to the conventional change detection image, and it is more robust to noise and the registration error. Fig. 13 shows an example of results obtained on retina fundus images.

C. Three-Dimensional Computer Vision

In 3-D computer vision, several tasks need to avoid outliers to obtain a reliable 3-D reconstruction as in SfM and 3-D motion recovery in which the information of interest is in the low-rank matrix $\mathbf{L}$.

1) Structure From Motion: SfM refers to the process of automatically generating a 3-D structure of an object by its tracked 2-D image frames. Practically, the goal is to recover both 3-D structure, namely 3-D coordinates of scene points, and motion parameters, namely attitude (rotation) and position of the cameras, starting from image point correspondences. Then, finding the full 3-D reconstruction of this object can be posed as a low-rank matrix recovery problem [9], [182], [302]. Suppose that the object is rigid, and there are N frames and M tracked points $\mathbf{L}_0=[X,Y]_{2M\times N}$, and the intrinsic camera parameters are fixed, then the trajectories of the feature points all lie in a linear subspace of $\mathbb{R}^{2M\times N}$ with $\mathrm{rank}(\mathbf{L}_0)\leq 4$. $\mathbf{L}_0$ can be factorized as $\mathbf{L}_0=\mathbf{PQ}$ where $\mathbf{P}\in\mathbb{R}^{2M\times 4}$ contains the rotations and translations while the first three rows of $\mathbf{Q}\in\mathbb{R}^{4\times M}$ represent the relative 3-D positions for each feature point in the reconstructed object. However, when there exist errors due to occlusion, missing data or outliers, the feature matrix is no longer of rank 4 and can be viewed as $\mathbf{A}_0=\mathbf{L}_0+\mathbf{S}_0$ where $\mathbf{S}_0$ corresponds to noise. Then, recovering the full 3-D structure of the object can be a low-rank matrix recovery problem in the RPCA formulation. Liu *et al.* [182] employed an ℓ_1-filtering approach to solve the decomposition while Wu *et al.* [302] used the augmented Lagrange multiplier (ALM) method [175]. Experiments on 43 344 tracked points over 1001 frames show that this approach provides the best compromise between time and accuracy in comparison with RSL [279] and the original RPCA-ALM [302]. In another work, Arrigoni *et al.* [8], [9] employed the RPCA and the robust matrix completion (RMC) formulations that are robust to outliers and missing data, respectively. Thus, $\mathbf{A}$ is decomposed into $\mathbf{L}+\mathbf{S}_1+\mathbf{S}_2+\mathbf{E}$ where $\mathbf{S}_1$ is a sparse matrix over a sampling set Ω representing the outliers in the measurements, and $\mathbf{S}_2$ has a support on Ω^C and it is an approximation of $\mathbf{P}_{\Omega^C}(\mathbf{L})$, representing

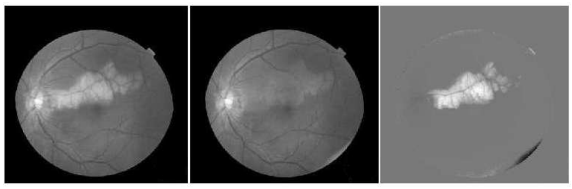


Fig. 13. *From left to right: the original image, the low-rank component, and the sparse component. (Images from [90].)*

the completion of the missing entries. Then, a modified version of GoDec [350], called R-GoDec in [9] and dubbed R-GoDec in [8], is used to solve this decomposition. Extensive experiments show that this method outperforms in accuracy when compared to previous state-of-the-art methods.

2) Three-Dimensional Reconstruction: In robotics, the optical sensor begins by capturing points of objects that exist in robots field of vision but the acquired 3-D point clouds are usually noisy and they also have misalignment. To remedy these problems, Arvanitis *et al.* [10] employed RPCA for removing outliers and noise of 3-D point clouds. Let us assume that the captured 3-D point cloud $\mathbf{A}$ consists of m points represented as a vector $\mathbf{v} = [\mathbf{x}, \mathbf{y}, \mathbf{z}]$ in a 3-D coordinate space $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^{m\times 1}$ and $\mathbf{v} \in \mathbb{R}^{m\times 3}$. Then, some of these points are considered as outliers and $\mathbf{A}$ is considered to be equal to $\mathbf{L} + \mathbf{S}$. $\mathbf{L}$ is a low-rank matrix representing the space of real data while $\mathbf{S}$ is a sparse matrix representing the space where outliers lie. Once RPCA is applied, the number of vertices decreases due to the removal of the outliers, so the number of the remaining vertices is m_r where $m_r < m$. Because the acquired 3-D point cloud is unorganized in $\mathbf{L}$, meaning that the connectivity of its points is unknown, Arvanitis *et al.* [10] used a triangulated model based on the k nearest neighbors (k-NN) algorithm. The triangulation process allows to specify the neighbors of each point so that the bilateral filtering method can be used efficiently as the denoising technique. At the end, a smoothed 3-D mesh is obtained which has an exploitable form to be used by other applications or processing tasks.

3) Three-Dimensional Motion Recovery: Skeleton tracking is a useful and popular application of Microsoft Kinect but it cannot provide accurate reconstructions for complex motions such as in the presence of occlusion. Indeed, the human skeleton is represented by a collection of joints, which are easily influenced by noises and have drifting problems. To address this issue, Wang *et al.* [294] developed a 3-D motion recovery based on the time coherence in a skeleton. Thus, this approach used a low-rank matrix analysis to correct invalid or corrupted motions. Let the captured skeleton sequence be stored in an observation matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$ which is obtained by stacking the 3-D positions of all the joints together, where $m = 3 \times n_f$ with n_f being the number of frames of the input skeleton sequence, and n is the number of joints (see [294, eq. (21)], ignoring the finger joints). Then, $\mathbf{A}$ is decomposed into $\mathbf{L} + \mathbf{S}$. $\mathbf{L}$ contains the clean motions and $\mathbf{S}$ contains the noise. Experiments [294] with Microsoft Kinect V2.0 show that this method accurately recovers high-quality skeletons from the invalid corrupted motion data in high efficiency.

IV. VIDEO PROCESSING

As well as in image processing, video processing tasks present either or both low-rank and sparsity aspects but with the temporal information in addition of the spatial information. Thus, both spatial and temporal aspects present in video sequences can be exploited in the $\mathbf{L} + \mathbf{S}$ decomposition to enforce its adequacy to the concerned task. In practice, RPCA via $\mathbf{L} + \mathbf{S}$ decomposition is suitable for video processing tasks in which 1) the observed video can be viewed as the sum of a low-rank clean video without perturbation and a sparse perturbations as in video restoration and denoising, background/foreground separation, motion saliency detection, video object segmentation, and hyperspectral video processing; 2) the observed video can be viewed as the sum of a low-rank video and a sparse video as in key frame extraction and UHD super resolution video; and 3) only the low-rank aspect is of interest as in background initialization, motion estimation, action recognition, and video summarization.

A. Background/Foreground Separation

Background/foreground separation in a video taken by a static camera is a crucial step for detecting moving objects in the video surveillance systems [25], [26], [29]. Before the work of Candès in 2009, this task was usually addressed by statistical modeling [255], [262], [271], fuzzy modeling [15], [16], [17], [27], and conventional subspace learning model either reconstructive [66], [67], [167], [250], [214], [264], [293], [332] and discriminative [88], [89], [199]. However, RPCA methods immediately provided a very promising solution toward moving object detection. However, because of the well-known challenges such as dynamic backgrounds, illumination conditions, color saturation, shadows, etc., the state-of-the-art RPCA methods do not often provide accurate segmentation [73], [76], [77], [78], [116], [117], [118], [119], [275], [308].

In RPCA, the background sequence is modeled by the low-rank subspace that can gradually change over time, while the moving foreground objects constitute the correlated sparse outliers. Thus, $\mathbf{A}$ contains the observed video in which the frames are stacked into column vectors and further decomposed as $\mathbf{L} + \mathbf{S}$. The decomposition is then solved via the minimization problem (10). Fig. 14 shows original frames of synthetic sequences from the BMC 2012 data set [284] and its decomposition into the low-rank matrix $\mathbf{L}$ and sparse matrix $\mathbf{S}$. We can see that $\mathbf{L}$ corresponds to the background whereas S corresponds to the foreground. The fourth image shows the foreground mask obtained by thresholding the matrix $\mathbf{S}$. The $\text{rank}(\mathbf{L})$ influences the number of modes of the background that can be represented by $\text{rank}(\mathbf{L})$: if $\text{rank}(\mathbf{L})$ is too high, the model will incorporate the moving objects into its representation whereas if $\text{rank}(\mathbf{L})$ is too low, the model tends to be unimodal and then the multimodality which appears in dynamic backgrounds will not be captured. The quality of the background/foreground separation is directly related to the assumption of the low rank and sparsity of the background and foreground, respectively. However, as the

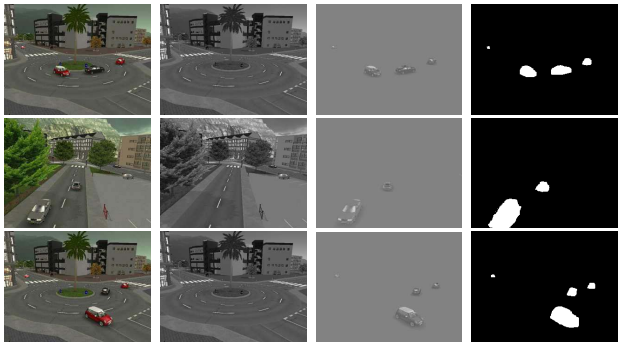


Fig. 14. ***Background/foreground separation. From left to right: original images, low-rank matrix*** L ***(background), sparse matrix*** S ***(foreground), and foreground mask. (Sequences from BMC 2012 data set [284].)***

matrix $\mathbf{S}$ could contain both the moving objects and noise, the stable decomposition $\mathbf{A} = \mathbf{L} + \mathbf{S} + \mathbf{E}$ (with $\mathbf{E}$ is the noise) is more suitable to separate moving objects from noise such as those proposed by Zhou *et al.* [355].

This application of RPCA is the most investigated one in the literature [31] because it is the most representative, challenging, and demanding application as it needs to take into account both spatial and temporal constraints with incremental and real-time constraints [30], [32]. We summarize the main solutions, a comparative evaluation on the CD.net 2012 data set [106] and the extension of background/foreground separation for moving cameras in the following sections. More details can be found in [31], [32], and [285].

1) Adding Spatial and Temporal Constraints: In the literature, spatial and/or temporal constraints are mainly added in the minimization problem. A general formulation can be expressed as follows [31]:

$$\begin{aligned}\min_{\mathbf{L},\mathbf{S},\mathbf{E}} \; & \underbrace{\|T(\mathbf{L})\|_{\text{norm}_1}^{p_1} + \lambda_1 \|\Pi(\mathbf{S})\|_{\text{norm}_2}^{p_2} + \lambda_2 \|\mathbf{E}\|_{\text{norm}_3}^{p_3}}_{\text{Decomposition}} \\ & + \underbrace{\lambda_3 \|\mathbf{L}\|_{2,1} + \delta_1 \|grad(\mathbf{S})\|_1 + \delta_2 TV(\mathbf{S}) + \delta_3 \Omega(\mathbf{S})}_{\text{Application}}, \\ \text{s.t.} \quad & \mathbf{A} = \mathbf{L} + \mathbf{S} + \mathbf{E}, \quad \text{or} \quad \mathbf{A} = \mathbf{W} \circ (\mathbf{L} + \mathbf{S} + \mathbf{E}), \quad \text{or} \\ & \mathbf{A} \circ \tau = \mathbf{L} + \mathbf{S} + \mathbf{E} \end{aligned} \tag{26}$$

where p_1, p_2, and p_3 are power in the set $\{1, 2\}$. λ_1, λ_2, λ_3, δ_1, δ_2, and δ_3 are regularization parameters. norm_1, norm_2, and norm_3 are norms which are used in the loss functions to enforce the low-rankness, sparsity, and noise constraints on $\mathbf{L}$, $\mathbf{S}$, and $\mathbf{E}$, respectively. norm_1 is taken to provide the following loss functions: ℓ_0-loss function ($\|\cdot\|_0$), ℓ_1-loss function ($\|\cdot\|_1$), ℓ_2-loss function ($\|\cdot\|_2$), nuclear norm function, Frobenius loss function, and log-sum heuristic function [63]. Other loss functions can be used such as ℓ_σ-loss function [316], least squares (LS) loss function ($\|\cdot\|_F^2$), Huber loss function [7], M-estimator based loss functions [125], and the generalized fused Lasso loss function [305], [306]. norm_2 is usually taken to force spatial homogeneous fitting in the matrix $\mathbf{S}$, that is, for example, the norm $\ell_{2,1}$ with $p_2 = 1$ [73], [76], [77], [78], [116], [117], [119], [118], [275]. It is important to note that the first part of (26) concerns mainly the decomposition into low-rank plus sparse and noise matrices and second part concerns mainly the application of background/foreground separation. The terms associated with background/foreground separation can be described as follows.

- The function $T(\cdot)$ is a set of invertible and independent transformations processed on $\mathbf{L}$ as in incPCP-TI [247], [263] to tackle translational and rotational camera jitter.
- The function $\Pi(\cdot)$ is a linear operator processed on $\mathbf{S}$ to enforce spatial and/or temporal constraints. $\Pi(\cdot)$ weights its entries according to their confidence of correspondence to a moving object such that the most probable elements are unchanged and the least probable elements are set to zero. $\Pi(\cdot)$ can be computed with optical flow [215] and with salient motion detection [267].
- The term $\lambda_3\|\mathbf{L}\|_{2,1}$ ensures that the recovered $\mathbf{L}$ has exact zero columns corresponding to the outliers.
- $\|grad(\mathbf{S})\|_1$, $TV(\mathbf{S})$, and $\Omega(\mathbf{S})$ are the gradient [117], [118], [119], [298], the total variation [41], [115], [117], [118], [299], and the static/dynamic tree structured sparsity norm [74], [75], [78], [185], [278] applied on the matrix $\mathbf{S}$ to enforce the spatial and/or temporal constraints, respectively.
- A weighting matrix $\mathbf{W}$ [267], [312], [319] or a transformation τ [72], [73], [74], [75], [76], [77], [78], [123], [124], [226] can also be used as a constraint in (26) to enforce the recovery of the background that appears in only a few frames and to eliminate the influence of light conditions, camouflages, and dynamic backgrounds, and to model potential global motion that the foreground region undergoes, respectively.

2) Online/Incremental and Real-Time Algorithms: Even if fast solvers [35], [163], [169], [325], [347] were developed to make the iterations as few as possible as well as SVD algorithms [87], [150], [346] were designed to make the iterations as efficient as possible, batch algorithms cannot reach the requirement of real-time computation for background/foreground separation. Thus, to update the model when new data arrive, several online/incremental algorithms can be found and they can be classified in the following categories [285]: 1) dynamic RPCA algorithms such as the recursive projected compressive sensing (ReProCS) algorithm and its variants [111], [111], [112], [110], [232], [234] provided with performance guarantees; 2) incremental PCP algorithms such as incPCP and its variants [243], [244], [245], [246], [247], [263]; 3) online decomposition algorithms [145]; 4) subspace

tracking algorithms such as the Grassmannian robust adaptive subspace tracking algorithm (GRASTA) [121], the ℓ_p-norm robust online subspace tracking (pROST) algorithm [254], the Grassmannian online subspace updates with structured-sparsity (GOSUS) algorithm [308], and the fast adaptive robust subspace tracking (FARST) algorithm [3]; and 5) the weighted low-rank approximation algorithm [69], [70], [71], and the (6) lifelong learning algorithm [20]. As it is expected that background/foreground separation also needs to be achieved in real time, several strategies have been developed which are generally based on compressive sensing algorithms [195], [193], [191], [228], [225], [290], [229], [194], [192], submatrices computation [227], and GPU implementations [6], [243].

3) Dealing With the Challenges: Several challenges appear in a video because of the type and locations of the camera, and its environments. Thus, several authors designed RPCA formulation for videos taken by a fixed color CCD camera (in most of the cases), but also by hyperspectral camera [268], by camera trap [101], [102], [257], and by aerial camera [83], [84], [85], [86]. Furthermore, dedicated methods also exist for infrared cameras [252] and RGB-D cameras [278], [137]. For the environments which present dynamic backgrounds, illumination changes, camera jitter, etc., many modified RPCA approaches have been designed according to the following very popular background modeling challenges.

- Noisy images: To cope with noisy videos in the presence of rainy or snowy conditions, Javed *et al.* [136] used real-time active random field (ARF) constraints using a probabilistic spatial neighborhood system. After that, online robust PCA (OR-PCA) is used to separate the low-rank and sparse component from denoised frames. In addition, a color transfer function is employed between the source and input image for handling global illumination conditions, which is a very useful technique for surveillance agents to handle the nighttime videos.
- Bootstrapping: In clutter scenes, where background is always occluded by heavy foreground objects, Javed *et al.* [141] developed a motion-aware graphs regularized RPCA (MAGRPCA).
- Camera motion: Several strategies are used in literature to deal with camera motion: 1) transformation-based methods in which a transformation $\tau(\cdot)$ is applied to the data matrix A [72], [74], [76], [78], [124], [226], [270] or to the low-rank matrix L [247], [263]; 2) compensation-based methods in which the motion due to the camera is compensated in the preprocessing step such as in [141] and [278]; and 3) endogenous-convolution-based methods in which convolutional sparse representations model the effects of nonlinear transformations such as translation and rotation, thereby simplifying or eliminating the alignment preprocessing task [297].
- Illumination changes: To be robust to illumination changes, Javed *et al.* [141] incorporated spectral graph regularization in the RPCA framework, while Newson *et al.* [211] used a weighted cluster graph. In the case of time-lapse videos and low-frame rate videos, Shakeri and Zhang [257] proposed a low-rank and invariant sparse decomposition (LISD) method where a prior illumination map is incorporated into the main objective function.
- Dynamic backgrounds: Zhou and Tao [351], [352] tracked multiple sparse object flows (motions) in video by using a shifted subspaces tracking (SST) strategy in order to segment the motions and recover their trajectories by exploring the low-rank property of background and the shifted subspace property of each motion. Thus, SST allows the model to separate the motions of the moving objects of interest and the motions of background objects such as trees and waves. Javed *et al.* [144], [146] used Markov random field (MRF) in OR-PCA. In RPCA based on salient motion detection (SMD-RPCA), Chen *et al.* [48] defined a saliency clue over the sparse matrix S to filter out the dynamic backgrounds globally. In another work, Wu *et al.* [303] employed a multicomponent group sparse RPCA, in which the observed matrix is decomposed into a low-rank static background L, a group sparse foreground S_1, and a dynamic background S_2. Moreover, each image is oversegmented into 80 superpixels using the simple linear iterative clustering (SLIC) [2] to take into account the spatial constraint.
- Intermittent motion of foreground objects: In MAGRPCA, Javed *et al.* [141] used an optical flow algorithm between consecutive frames to generate the binary mask of motion. This motion mask allows to remove the motionless video frames and create a matrix comprising only dynamic video clips. Thus, MAGRPCA incorporates the motion message and encodes the manifold constraints and is very efficient because motionless frames are removed in order to handle large outliers in the background model. In SMD-RPCA, Chen *et al.* [48] leveraged the previously detected salient motion to guide the update of the current low-rank prior. Newson *et al.* [211] used a weighted cluster graph.
- Ghost suppression: Rodriguez and Wohlberg [248] proposed an algorithm called gs-incPCP, which can suppress the ghost by using two simultaneous background estimates based on observations over the previous N_1 and N_2 frames with $N_1 \ll N_2$ in order to identify and diminish the ghosting effect. Ebadi *et al.* [74], [75] proposed a tandem algorithm, which involves an initialization step before the optimization takes place. It is different from algorithms that require a two-pass optimization [96], [97], where the optimization is twice performed to refine results. Introducing a prior knowledge of

the spatial distribution of the outliers to the model, Ebadi *et al.* further proposed methods for faster convergence [74], [75].

- Shadows: Li *et al.* [159] designed a box constraint RPCA (BC-RPCA) to separate the moving objects and the shadows. So BC-RPCA models the input video as three parts which are low-rank background, sparse foreground, and moving shadows. Experiments on several scenes show that BC-RPCA works well on shadow and varying lighting condition challenges.

All these aforementioned key limitations need to be addressed in the RPCA formulation for background/foreground separation. Furthermore, the evaluation needs to be conducted with a large-scale data set such as the CD.net 2012/2014 data set[6] [106], [107] or the BMC 2012 datai set[7] [284] to allow full and fair comparisons.

4) Comparative Evaluation: In this part, we show the performance of the current state-of-the-art RPCA-based methods for background/foreground separation using the CD.net 2012 data set [106], and a more detailed analysis can be found in [285]. This data set contains almost 31 video sequences which are divided into six different video categories comprising "baseline," "dynamic backgrounds" (DBs), "intermittent object motion" (IOM), "thermal," "camera jitter," and "shadows" presenting the different challenges previously enumerated. The resolution of the videos also varies from 320×240 to 480×720 with hundreds to thousands of frames. We compared a total of 25 existing methods comprising 15 batch algorithms and ten online algorithms. The implementation of all these algorithms is also available in the LRSLibrary. These methods are classified into three main categories.

- Provable methods: PCP [37], nonconvex alternating-projection-based RPCA (AltProj) [210], near-optimal RMC (NO-RMC) [55], RPCA via gradient descent (RPCA-GD) [320], recursive projected compressive sensing (ReProCS-provable) [208], [326], and modified-PCP [327].
- Heuristic methods: ReProCS [110], GRASTA [122], three-term decomposition (3TD) [215], two-pass RPCA (2PRPCA) [96], go decomposition (GoDec) [350], OR-PCA [144], [139], ℓ_p robust online subspace tracking (pROST) [254], and probabilistic robust matrix factorization (PRMF) [295].
- Heuristic methods with application specific constraints: incremental PCP (incPCP) [246], motion-assisted spatiotemporal clustering of low rank (MSCL) [143], detecting contiguous outliers in the low-rank representation (DECOLOR) [353], low-rank structured-sparse decomposition (LSD) [185], total variation RPCA (TVRPCA) [41], spatiotemporal RPCA (SRPCA) [142], robust motion-assisted matrix restoration (RMAMR) [312], generalized fussed lasso [305], Grassmannian online subspace updates with structured sparsity (GOSUS) [308], contiguous outliers representation via online low-rank approximation (COROLA) [256], and online mixture of Gaussians for matrix factorization with total variation (OMoGMF+TV) [321], respectively.

For qualitative evaluation, visual results were reported using 15 challenging sequences that contained two sequences, namely, "highway" and "office" from the "baseline" category; three sequences "canoe," "boats," and "overpass" from the DB category; two sequences "traffic" and "badminton" from the "camera jitter" category; three sequences "winterDriveway," "sofa," and "streetLight" from the IOM category; three sequences "backdoor," "copyMachine," and "cubicle" from the "shadows" category; and two sequences "library" and "lakeside" from the "thermal" category. Fig. 15 provides qualitative results and comparisons of 22 current state-of-the-art RPCA-based methods on 15 sequences. The execution times required by all of the algorithms were compared on a machine with a 3:0-GHz Intel core i5 processor and 4 GB of RAM. For quantitative evaluation, the used metrics come from the CD.net 2012 data set [106] such as the recall, the precision, and the F_1-measure score. Recall gives the percentage of corrected pixels classified as background when compared with the total number of background pixels in the ground truth. Precision gives the percentage of corrected pixels classified as background when compared to the total pixels classified as background by the algorithm. A good performance is obtained when the detection rate, also known as recall, is high without altering the precision. Based on these metrics, the F_1-measure is computed as $F_1 = (2 \times Recall \times Precision/Recall + Precision)$. The F-measure characterizes the performance of classification in precision-recall space. The aim is to maximize F_1-measure closed to one. Table 1 shows the quantitative results in terms of average F_1 measure score as well as the computational time in seconds for all of the compared algorithms applied on the large video sequence known as *boats* from DB category. On average, among all algorithms that do not use extra constraints, PRMF, 2PRPCA, ReProCS-provable, and ReProCS had the best performance with F_1 scores of 74%–78%. On average, for all data sets, only two of the methods that use extra constraints, MSCL and GOSUS, were better and only by a little by achieving 83% and 81% scores, respectively. For computational time, ReProCS and ReProCS-provable are the fastest methods in provable methods category, while from the heuristic methods category, OR-PCA and GRASTA are even faster but have worse performance. COROLA and OMoGMF+TV in heuristic methods with additional constraints category are top performing methods in terms of computation time in seconds. Practically speaking, these results show that an RPCA method for background/foreground should take into account both spatial and temporal constraints as well as it should be incremental to be effectively usable in real applications.

[6]http://changedetection.net/
[7]http://bmc.iut-auvergne.com/

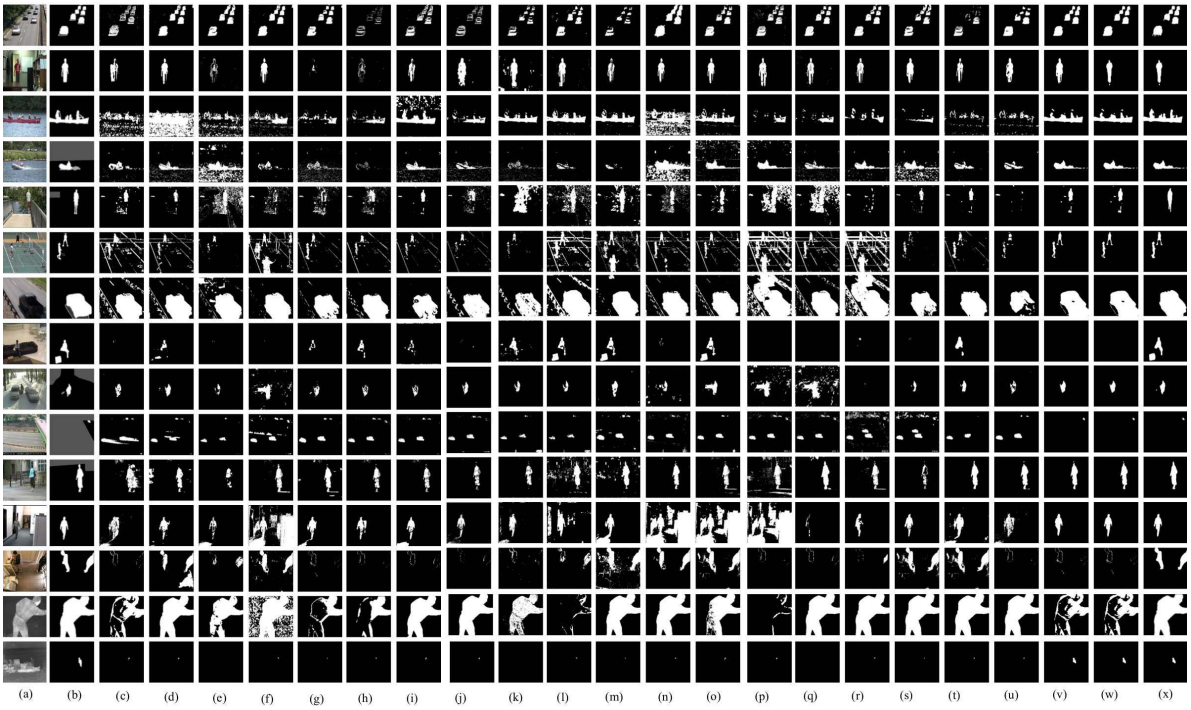


Fig. 15. ***Comparison of the qualitative results of the 15 input images from ChangeDetection.net data set. From left to right: (a) the set of 15 input images; (b) the ground truth of the foreground objects; (c) the background subtraction estimated by RPCA via the PCP method; (d) GoDec; (e) RPMF; (f) RPCA-GD; (g) 3TD; (h) pROST; (i) incPCP; (j) RMAMR; (k) GRASTA; (l) ReProCS; (m) TVRPCA; (n) SRPCA; (o) NO-RMC; (p) LSD; (q) GOSUS; (r) OMoGMF+TV; (s) COROLLA; (t) OR-PCA; (u) 2PRPCA; (v) DECOLOR; (w) GFL; and (x) MSCL. From to bottom: rows (1)–(2) sequences "highway" and "office" from the "baseline" category; rows (3)–(5) sequences "canoe," "boat," and "overpass" from the DB category; rows (6)–(7) sequences "badminton" and "traffic" from the "camera jitter" category; rows (8)–(10) sequences "sofa," "winter Driveway," and "streetLight" from the IOM category; rows (11)–(13) sequences "BackDoor," "cubicle," and "copyMachine" from the "shadow" category; rows (14)–(15) sequences "library" and "lakeside" from the "thermal" category. (Images from [285].)***

5) Extension to Moving Cameras: Background/foreground separation is also needed in video taken by moving cameras such as PTZ cameras and handheld cameras [201]. This issue is actually less investigated than the static case. Unlike strategies [75], [123], [226], [247], and [263] for small camera jitter which used affine transformation model that describes the motion of the frames in the quasi-static cameras case, Gao *et al.* [92], [205] produced a panoramic low-rank component that spans the entire field of view, automatically stitching together corrupted data from partially overlapping scenes. Practically, the algorithm proceeds by registering the frames of the raw video to a common reference perspective and then it minimizes a modified RPCA cost function that accounts for the partially overlapping views of registered frames and includes TV regularization to decouple the foreground from noise and sparse corruption. The augmented RPCA problem formulation is then expressed as follows:

$$\min_{\mathbf{L},\mathbf{S}_1,\mathbf{S}_2} \frac{1}{2}\|\mathbf{P}_M(\mathbf{A}-\mathbf{L}-\mathbf{S}_1-\mathbf{S}_2)\|_F^2 + \lambda_1\|\mathbf{L}\|_* + \lambda_2\|\mathbf{S}_1\|_1 + \lambda_3 TV(\mathbf{S}_2) \quad (27)$$

where $\mathbf{L}$, $\mathbf{S}_1$, and $\mathbf{S}_2$ represent the background (low-rank component), the sparse corruptions (sparse component), and the foreground (smoothly varying matrix), respectively. $TV(.)$ is the total variation regularizer [118]. The low-rank component is obtained via the optimal low-rank matrix estimator (OptShrink [206]) that requires no parameter tuning. Experiments show that this algorithm is robust to both dense and sparse corruptions of the raw video and yields superior background/foreground separations compared to the original RPCA [37] and total variation regularized RPCA [41]. For slowly moving cameras in the case of anomaly detection in videos, Thomaz *et al.* [277] employed an algorithm that computes the union of subspaces that best represents all the frames from a reference video as a low-rank projection plus a sparse residue. The intrinsic structure of the sparse decomposition is used in order to detect the anomalies without requiring previous video synchronization. Because the original RPCA is able to project the data onto a single subspace only, Thomaz *et al.* [277] designed an algorithm based on the robust subspace recovery (RoSuRe [22]) which is able to project data onto a union of subspaces of lower dimensions. The moving-camera RoSuRe (mcRoSuRe) provides good detection results while at the

Table 1 Average F_1 Score of Provable Methods, Heuristic Methods, and Heuristic Methods With Specific Constrains for Background/Foreground Separation. Time Is Shown for a Video Having 320×240 Resolution of 8000 Frames. The Best and Second Best Performing Methods Are Shown in Bold and Bold Italic, Respectively

Provable Methods	Baseline	DB	Camera Jitter	Shadow	Thermal	IOM	Average	Time (secs/frame)
PCP (batch) Fig. 15 (c) [37]	0.75	0.69	0.62	0.73	0.65	0.48	0.65	4.19
AltProj (batch) [210]	**0.78**	***0.71***	0.60	**0.76**	0.69	0.58	***0.68***	2.38
NO-RMC (batch) Fig. 15 (o) [55]	0.71	0.64	0.64	0.66	***0.71***	0.50	0.64	2.85
RPCA-GD (batch) Fig. 15 (f) [320]	0.74	0.62	0.68	***0.75***	0.66	0.49	0.65	2.46
ReProCS-provable (online) [208], [326]	0.77	**0.77**	***0.69***	0.71	**0.74**	**0.70**	**0.73**	0.74
Mod-PCP (online) [327]	***0.75***	0.64	**0.70**	0.65	0.69	**0.70**	***0.68***	**0.44**
Heuristic Methods	**Baseline**	**DB**	**Camera Jitter**	**Shadow**	**Thermal**	**IOM**	**Average**	**Time**
ReProCS (online) Fig. 15 (l) [110]	0.80	0.76	0.72	0.75	0.77	**0.69**	***0.74***	***0.61***
GRASTA (online) Fig. 15 (k) [122]	0.66	0.35	0.43	0.52	0.42	0.35	0.45	1.16
3TD (batch) Fig. 15 (g) [215]	***0.88***	0.75	0.72	0.68	***0.78***	0.55	0.72	2.17
2PRPCA (batch) Fig. 15 (u) [96]	**0.92**	**0.79**	***0.81***	***0.80***	0.76	***0.65***	**0.78**	1.63
GoDec (batch) Fig. 15 (d) [350]	0.77	0.58	0.48	0.51	0.62	0.38	0.55	1.56
OR-PCA (online) Fig. 15 (t) [144]	0.86	0.75	0.70	0.74	0.76	0.56	0.72	**0.22**
pROST (online) Fig. 15 (h) [254]	0.79	0.59	0.79	0.70	0.58	0.48	0.65	2.03
PRMF (batch & online) Fig. 15 (e) [295]	**0.92**	***0.77***	**0.85**	**0.88**	**0.83**	0.48	**0.78**	2.40
Heuristic Methods with Specific Constraints	**Baseline**	**DB**	**Camera Jitter**	**Shadow**	**Thermal**	**IOM**	**Average**	**Time**
incPCP (online) Fig. 15 (i) [246]	0.81	0.71	0.78	0.74	0.70	***0.75***	0.74	0.41
MSCL (batch) Fig. 15 (x) [143]	0.87	***0.85***	**0.83**	0.82	**0.82**	**0.80**	**0.83**	1.68
DECOLOR (batch) Fig. 15 (v) [353]	**0.92**	0.70	0.68	***0.83***	0.70	0.59	0.73	1.88
LSD (batch) Fig. 15 (p) [185]	**0.92**	0.71	0.78	0.81	0.75	0.67	0.77	1.43
TVRPCA (batch) Fig. 15 (m) [41]	0.84	0.55	0.63	0.71	0.69	0.57	0.66	1.48
SRPCA (batch) Fig. 15 (n) [142]	0.82	0.84	0.78	0.77	0.79	**0.80**	0.80	0.59
RMAMR (batch) Fig. 15 (j) [312]	0.89	0.82	0.75	0.73	0.75	0.66	0.76	1.32
GFL (batch) Fig. 15 (w) [305]	0.83	0.74	0.78	0.8 2	0.76	0.59	0.75	2.40
GOSUS (online) Fig. 15 (q) [308]	***0.90***	0.79	***0.82***	**0.84**	***0.80***	0.74	***0.81***	0.89
COROLA (online) Fig. 15 (s) [256]	0.85	**0.86**	***0.82***	0.78	***0.80***	0.71	0.80	***0.39***
OMoG+TV (online) Fig. 15 (r) [321]	0.85	0.76	0.78	0.68	0.70	0.71	0.74	**0.19**

same time avoiding the need for previous video synchronization. For moving and panning cameras, Chau and Rodriguez [45] designed an incremental PCP algorithm called incPCP-PTI which continuously aligns the low-rank component to the current reference frame of the camera. Based on the translational and rotational jitter invariant algorithm incPCP-TI [247], incPCP-PTI continuously estimates the alignment transformation $T(\cdot)$ in order to align the previous low-rank representation with the observed current frame. Furthermore, instead of using iterative hard threshold as in incPCP-TI, the low-rank approximation problem is solved in the reference frame by applying an adaptive threshold to the residual. Further research might focus on other types of distortions such as perspective changes, zooming in/out of the camera, and the reduction of the time for high frame rate real-time applications.

B. Motion Saliency Detection

Motion saliency detection is crucial for video processing tasks, such as video segmentation, object recognition, and adaptive compression. Different from image saliency, moving objects catch human being's attention much easier than static ones. Xu *et al.* [311] used the low-rank and sparse decomposition on video slices along $\mathbf{X}-\mathbf{T}$ and $\mathbf{Y}-\mathbf{T}$ planes to achieve the separation of foreground moving objects from backgrounds. Naturally, the low-rank component $\mathbf{L}$ corresponds to the background and the sparse component $\mathbf{S}$ captures the motion objects in the foreground. Then, the motion matrices, i.e., abs($\mathbf{S}$) obtained from the $\mathbf{X}-\mathbf{T}(\mathbf{Y}-\mathbf{T})$ slices are integrated together as $\mathbf{S}_{\text{cube}\mathbf{XT}}$ ($\mathbf{S}_{\text{cube}\mathbf{YT}}$) along $\mathbf{X}-\mathbf{Y}-\mathbf{T}$. The initial saliency map cube is obtained by computing $\text{norm}(\mathbf{S}_{\text{cube}\mathbf{XT}} * \mathbf{S}_{\text{cube}\mathbf{YT}})$ where $*$ is the elementwise product operator, and norm() represents normalization processing. The size of $\mathbf{T}$ equals the size of the video, and it can also be defined as the size of a subvideo. In addition, a spatial information refinement preserves the completeness of the detected motion objects. From Fig. 16, we can see that the RPCA algorithm outperforms a standard approach called temporal spectrum residual (TSR) [58] as well as background subtraction algorithms such as consecutive frame difference (CFD) and mixture of Gaussians (MoG) [271].

Fig. 16. ***Motion saliency detection. From left to right: (a) original images; (b) consecutive frame difference; (c) MoG [271]; (d) TSR [58]; (e) raw saliency map [58]; and (f) final result obtained by the RPCA algorithm [58]. (Images from [311].)***

C. Motion Estimation

Motion estimation concerns the process of determining motion vectors for the transformation from one 2-D image to another, which is usually done from adjacent frames in a video sequence. Ros *et al.* [249] addressed this problem

with a modified formulation of RPCA in the special case of camera-pose recovery and visual odometry. Practically, Ros *et al.* [249] considered the estimation of motion models $\mathbf{M}_i{}_{i=1}^{N}$ between pair of stereo frames $\mathbf{F}_i, \mathbf{F}_{i+1}$ along a given sequence of N frames $\{\mathbf{F}_i\}_{i=1}^{N}$. Each frame $\mathbf{F}_i = (\mathbf{V}_i^l, \mathbf{V}_i^r)$ consists of two images taken from the left and right cameras at the same instant t_i. This formulation is suitable for the stereo visual odometry problem with a rigid 3-D transformation [98]. When estimating the transformation $\mathbf{M}_i$, one should account for the presence of noise and outliers in the observations in order to avoid a biased solution. Thus, Ros *et al.* [249] exploited the rank constraints present in rigid 3-D motions to identify outliers. Practically, the information resultant from the low-rank and sparse decomposition is used to make a binary decision on each tuple of point matches (column) about its pertinence to the outlier set. Despite the impossibility of performing an exact recovery of every element of the observation matrix, the resultant information is enough to make this set of binary decisions. Thus, a robust decomposition with constrained rank (RD-CR) is employed and is formulated as follows:

$$\min_{\mathbf{L},\mathbf{S}} \frac{1}{2}\|\mathbf{A}-\mathbf{L}-\mathbf{S}\|_F^2 + \lambda\|\mathbf{S}\|_1, \qquad \text{s.t.} \quad \text{rank}(\mathbf{L}) \leq r. \quad (28)$$

This formulation enables solving problems in harder conditions, i.e., higher ranks and greater proportions of outliers. However, in motion estimation problems, the rank is still too high to achieve an exact estimation of $\mathbf{L}$ and $\mathbf{S}$. For this reason, the problem is addressed by using the residual matrix $\mathbf{S}$ to infer which columns (point matches) are outliers. From the results, this approach is competitive against state-of-the-art methods on the KITTI data set[8] [98] in terms of accuracy and is more efficient in terms of computation.

D. Tracking

Tracking in computer vision refers to a problem which allows to track an object from a temporal sequence, and then allows to estimate the trajectory of an object in the image plane when it moves around a scene. Object detection and tracking are two independent processes in video sequences. However, object detection can be improved by using a tracking feedback. Thus, Lin *et al.* [173] introduced tracking feedback in the RPCA formulation as follows:

$$\begin{aligned}\min_{\mathbf{L},\mathbf{G_{i,j}}\in\{\mathbf{0},\mathbf{1}\}} \; & \frac{1}{2}\|\mathbf{P}_{\mathbf{G}^\perp}(\mathbf{A}-\mathbf{L})\|_F^2 + \lambda_1\|\mathbf{L}\|_* \\ & + \lambda_2\|f(\mathbf{G})\|_1 + \lambda_3\|t(\mathbf{G})\|_1 + \gamma\|\mathbf{B}\cdot\text{vec}(\mathbf{G})\|_1, \\ \text{s.t. } & \text{rank}(\mathbf{L}) \leq r \end{aligned} \quad (29)$$

where $\mathbf{G} \in \{0,1\}^{n\times m}$ denotes the foreground support and its value is 0 if (i,j) is background and 1 if (i,j) is foreground. $\mathbf{P}_{\mathbf{G}}(\mathbf{X})$ is the orthogonal projection of the matrix $\mathbf{X}$ onto the linear space of matrice supported by $\mathbf{G}$, and $\mathbf{P}_{\mathbf{G}^\perp}(\mathbf{X})$ is the complementary projection. $f(\mathbf{G})$ is the fractal dimension of the object support $\mathbf{G}$ and $\mathbf{B}$, and it is the node-edge incidence matrix. $t(\mathbf{G})$ is the object tracking process of support $\mathbf{G}$. As the objective function of (29) is nonconvex, an alternating method is to separate the energy minimization over $\mathbf{L}$ and $\mathbf{G}$ into two steps, respectively. $\mathbf{L}$-step is a convex optimization problem using the RPCA algorithm, and $\mathbf{G}$-step can be solved by a graph cut algorithm. This algorithm called group object detection and tracking (GODT) outperforms DECOLOR [353] on the I2R data set [161].

Shan and Chao [258] designed an improved ℓ_1-tracker in a particle filter framework using RPCA and random projection. Practically, three targiet templates and several background templates are employed into a template set.

- The target templates are obtained as follows: 1) a fixed template obtained from a manually selected target in the first frame; 2) a dynamic template updated via RPCA which builds a stable appearance model for long-time tracking; and 3) a dynamic template which is frequently reinitialized based on the stable template and is updated rapidly to represent the fast appearance change of the target. First, a data set $\mathbf{A}_0$ is constructed based on the tracking results in the former N frames. For the similarity of the tracking results, a low-rank matrix is recovered from the data set by removing the gross corruption and even outlier. Each column of $\mathbf{A}_0$ is a reduced dimensional feature vector from one normalized tracking result. When next N^+ tracking results are available in $\mathbf{A}^+$, they are used to update the data matrix $\mathbf{A}_0$. So, $[\mathbf{A}_0\mathbf{A}^+]$ is cleaned by RPCA as follows:

$$\begin{aligned}\min_{[\mathbf{L_0}\mathbf{L}^+],[\mathbf{S_0}\mathbf{S}^+]} \; & \left\|\left[\mathbf{L}_0\mathbf{L}^+\right]\right\|_* + \lambda\left\|\left[\mathbf{S}_0\mathbf{S}^+\right]\right\|_1 \\ \text{s.t. } & \left[\mathbf{A}_0\mathbf{A}^+\right] = \left[\mathbf{L}_0\mathbf{L}^+\right] + \left[\mathbf{S}_0\mathbf{S}^+\right] \end{aligned} \quad (30)$$

  where $[\mathbf{L}_0\mathbf{L}^+]$ denotes the new cleaned matrix, and $[\mathbf{S}_0\mathbf{S}^+]$ is the new sparse error matrix. The jth column of matrix $\mathbf{A}_0$ is then replaced by the ith column of matrix $\mathbf{A}^+$ to be used when the next N^+ tracking results arrive.
- The background templates consist of several background image patches cropped from the background regions of the former frames in order to strengthen the algorithms ability of distinguishing the background and the foreground. These templates combined with the three target templates are then used to represent the candidate image patches sparsely.

Finally, the candidate with the minimum distance to its linear combination corresponding to only the target templates is selected as the tracking target. Experiments show that this RPCA-based ℓ_1-tracker outperforms in certain critical situations when compared to several state-of-the-art algorithms. In another work, Elnakeeb and Mitra [82]

[8] http://www.cvlibs.net/datasets/kitti/eval-odometry.php

considered the incorporation of a line constraint for structured estimation. Practically, multiple forms of structure on matrices are extended from low rank and sparsity. The line constraint is introduced via a rotation that yields a secondary low-rank condition. Then, Elnakeeb and Mitra [82] applied this method to single object tracking in video wherein the trajectory can be parameterized as a line. Noticeable performance improvement is obtained over previous background subtraction methods that do not exploit the line structure.

E. Action Recognition

Motion representation is an important task in human action recognition, and most traditional methods usually require intermediate processing steps such as actor segmentation, body tracking, and interest point detection, making these methods sensitive to errors due to these processing steps. To remedy this limitation, Huang *et al.* [131] designed a motion representation method for action recognition by extracting refined low-rank features of RPCA. After extensive experiments, Huang *et al.* [131] determined the optimal λ for extracting the discriminative information of motion. Then, the RPCA algorithm is applied on the all action image sequences with the appropriate parameter λ to obtain the low-rank images and sparse error images. The low-rank images of all the action image sequences are very similar and represent the discriminative information of motion, while the sparse error images are different and represent the individual differences of each action image. Thus, the low-rank images are kept to perform action recognition, and the sparse error images are discarded. To represent the characteristic of the obtained low-rank images, Huang *et al.* [131] employed the edges distribution histogram (EDH) and the accumulative edges distribution histogram (AEDH) to encode the statistical distribution of the low-rank images into a feature vector. Finally, the support vector machine (SVM) is applied to recognize human actions represented by the EDH or AEDH feature. Experiments on the KTH action data set[9] [153] show that this algorithm outperforms previous approaches with an average accuracy of 96.16%.

F. Key Frame Extraction

Key frame extraction concerns the problem of selecting a subset of the most informative frames from a video to summarize its content such as in video summarization, search, indexing, and prints from video. Most state-of-the-art methods work directly with the input video data set, without considering the underlying low-rank structure of the data set. Other methods exploit the low-rank component only, but they ignored the other key information in the video. On the other hand, Dang *et al.* [61] developed the key frame extraction (KFE) algorithm based on RPCA that decomposes the input video data into a low-rank component which reveals the information across the elements of the data set, and a set of sparse components each of which contains distinct information about each element. Then, Dang *et al.* [61] combined the two information types into a single ℓ_1-norm-based nonconvex optimization problem to extract the desired number of key frames. Extensive experiments on a variety of consumer and other types of videos show that RPCA-KFE with the ground truth and with related state-of-the-art algorithms clearly illustrates its viability.

[9]http://www.nada.kth.se/cvap/actions/

G. Video Object Segmentation

Video segmentation concerns the partition of a video into several semantically consistent spatiotemporal regions. It is a fundamental computer vision problem in several applications such as video analytics, summarization, and indexing. However, its computational complexity and inherent difficulties such as the large intracategory variations and the large intercategory similarities make this task very challenging. For streaming video segmentation, Li *et al.* [157], [158] employed the suboptimal low-rank decomposition (SOLD) algorithm which tracks the low-rank representation by exploiting the low-rank structure of low-level supervoxel features. Since the supervoxel feature matrix is often noisy or grossly corrupted, the low-rank representation can be formulated as follows:

$$\mathbf{A} = \mathbf{AL} + \mathbf{S} + \mathbf{E}, \quad \text{s.t. rank}(\mathbf{Z}) \leq r \tag{31}$$

where r is the desired rank and $r \ll n$. Then, Li *et al.* [157], [158] integrated the discriminative replication prior based on internal video statistics into SOLD based on the observation of small-size video patterns within the same object. An inference algorithm is employed to perform streaming video segmentation in both unsupervised and interactive scenarios. Extensive experiments show that SOLD outperforms other video segmentation approaches in both accuracy and efficiency.

H. Video Coding

Video coding aims to generate a content representation format for storage or transmission. Due to the growing needs for public security, traffic surveillance, and remote healthcare monitoring, efficient compression and fast transmission of large amount of surveillance videos are required in practice. Surveillance videos are usually with a static or gradually changing background. The state-of-the-art block-based codec, H.264/AVC, is not sufficiently efficient for encoding surveillance videos since it cannot exploit the strong background temporal redundancy in a global manner. First, Chen *et al.* [46] applied the RPCA formulation called low-rank and sparse decomposition (LRSD) to decompose a surveillance video into the low-rank component (background) and the sparse component (moving objects). Then, the GoDec

algorithm [350], which is a randomized algorithm for low-rank and sparse matrix decomposition in noisy case, was employed to separate the components of $\mathbf{A}$, so that $\mathbf{A} = \mathbf{L} + \mathbf{S} + \mathbf{E}$, where $\mathbf{L}$ is a rank-r matrix. Then, different coding methods for the two different components were designed. The frames of the background are representing by very few independent frames based on their linear dependency, which significantly removes the temporal redundancy. Experimental results show that LRSD significantly outperforms H.264/AVC, with up to 3-dB peak signal-to-noise ratio (PSNR) gain, especially at relatively low bit rate. However, LRSD cannot handle high-resolution or long-time videos due to its high memory requirement. To remedy to these limitations, Chen *et al.* [47] designed an incremental LRSD (ILRSD) algorithm that can effectively handle large-scale video sequences without much performance loss. Guo *et al.* [114] employed a dictionary approach based on a small number of observed frame. With the trained background dictionary, every frame is separated into the background and moving object via the RPCA formulation. As in LRSD, GoDec [350] is also used for the decomposition. Then, the compressed motion is stored together with the reconstruction coefficient of the background corresponding to the background dictionary. The decoding is carried out on the encoded frame in an inverse procedure. This algorithm outperforms H.264/AVC codec in terms of both file size and PSNR for surveillance videos.

For surveillance video coding, the rate-distortion analysis shows that a larger penalty λ needs to be used if the background in a coding unit has a larger proportion. To address this problem, Zhao *et al.* [342] performed an analysis on the relationship between the optimal penalty and the background proportion, and then designed a penalty selection model to obtain the optimal coding performance for surveillance video.

I. Hyperspectral Video Processing

Chang and Gerhart [43], [100] employed the RPCA decomposition for the detection of gaseous chemical plumes in hyperspectral video data. These video sequences are typically very large in size due to the fact that the images themselves are of high resolutions. An algorithm which decomposes a hyperspectral video sequence into a low-rank and sparse representation $\mathbf{A} = \mathbf{L} + \mathbf{S}$ is then used and applied to the detection of chemical plumes. As the problem is the same as background/foreground separation, the input frames are stacked as columns in matrix $\mathbf{A}$. However, the memory requirement of this problem is typically more challenging than in the color case. Let each frame of a data set be an $n_r \times n_c \times n_b$ $(128 \times 320 \times 129)$ data cube, then by concatenating along the spectral dimension, it produces a vector of length $n_r \times n_c \times n_b$ $(5, 283, 840)$. The data matrix $\mathbf{A}$ with N frames is of size $n_r \times n_c \times n_b \times N$ $(5, 283, 840 \times 100)$. In practice, preprocessing techniques are used to make the task computationally feasible. For example, one can select a subset of the spectral bands based on noise or performing dimension reduction on each frame of the video sequence. Experiments show that the low-rank approximation captures the background very well. After the plume is released, the sparse component captures the movement of the plume through each band of the video sequence. Applying this method to the original (nonreduced) video sequence results in the background matrix approximating stationary signals and the sparse component showing moving signals and noise.

J. Video Restoration and Denoising

Video restoration concerns the recovery of the original one from the degraded video data. It is one of the fundamental problems in video processing, especially in the current days. Indeed, old films which need to be restored present noise contamination, image blurring, and missing data. Second, with the prevalence of webcams and camera phones, the problem of video restoration has become even more important than before. Practically, there are two main kinds of restoration: video denoising in the presence of random-valued noise in the data acquisition and transmission due to faulty sensor or transmission, and video inpainting for archived film to repair videos corrupted by line scratches, hair, and dust. Ji *et al.* [147] grouped similar patches in the spatiotemporal domain and formulated the video restoration problem as a joint sparse and low-rank matrix approximation problem. First, for each reference patch p, similar patches are found in the spatiotemporal domain by using a patch matching algorithm. Assume that m match patches are found and denoted as $\{\mathbf{p}_i\}_{i=1}^{m}$. If each patch $\mathbf{p}_i$ is represented by a vector $\mathbf{p}_i \in \mathbb{R}^{n \times n}$ by concatenating all columns of the patch into a column vector, the resulting patch stack is then a matrix $\mathbf{A} \in \mathbb{R}^{n^2 \times m}$ with $\mathbf{A} = (\mathbf{p}_1, \mathbf{p}_2, \ldots \mathbf{p}_3)$. As matrix $\mathbf{A}$ can be corrupted by noise and/or outliers, $\mathbf{A}$ is then decomposed with the stable RPCA formulation $\mathbf{A} = \mathbf{L} + \mathbf{S} + \mathbf{E}$, where $\mathbf{L}$ is the original patch matrix for recovery, $\mathbf{S}$ is the matrix of outliers, and $\mathbf{E}$ is the random image noise

$$\min_{\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \lambda \|\mathbf{S}\|_1 + \frac{1}{2\mu} \|\mathbf{A} - \mathbf{L} - \mathbf{S}\|_F^2 \qquad (32)$$

with μ defined with an empirical parameter. Experiments show that this method compares favorably against many existing algorithms on both video denoising and video inpainting. This method can effectively remove the noise, but must transform 2-D samples to 1-D vectors, and the input matrix should be approximately a low-rank matrix. To remedy this limitation, Zhao *et al.* [345] used an extended RPCA algorithm called low-rank approximations of matrices (GLRAM) to obtain better performance than RPCA. As Ji *et al.* [147], Guo and Vaswani [113] also considered that many noisy or corrupted videos can be split into three parts, but they used the notion of layers instead of patches. Thus, PCP is used first to initialize

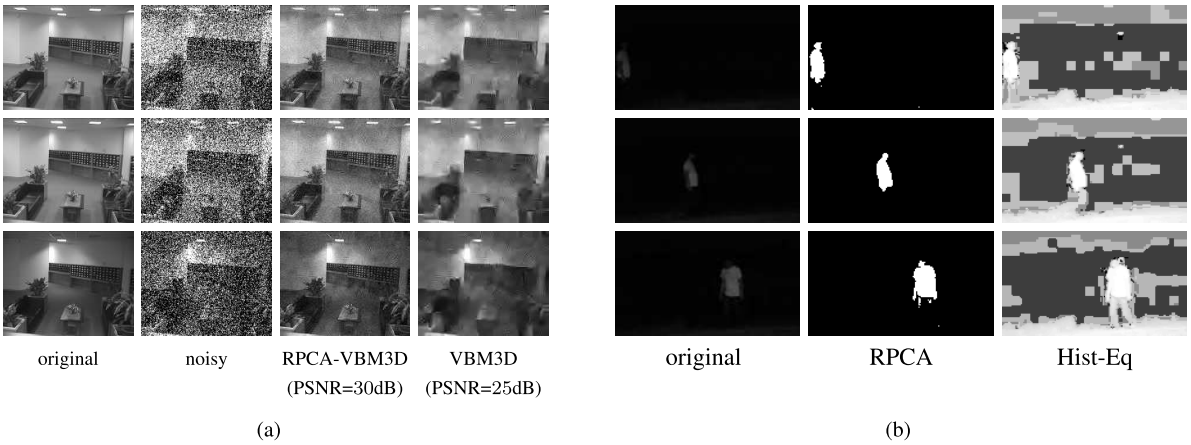


Fig. 17. *Video denoising and enhancement. (a) Denoising a very noisy video with Gaussian noise of standard deviation $\sigma = 70$ and hence PSNR is 11 dB. From left to right: original videos, noisy videos, RPCA-VM3D [113] results, and VBM3D [59] results. Note that VBM3D gives a much more blurred denoised image. PSNR is noted below each figure as well. (b) Video enhancement. From left to right: original videos, the RPCA algorithm (ReProCS [110]) results, and the histogram equalization (Hist-Eq) results. (Images from [285].)*

the low-rank layer, the sparse layer, and the residual which is small and bounded. Then, ReProCS [110] is used overtime to quickly separate the layers in videos with large-sized sparse components and/or significantly changing background images. This video-layering step is followed by VBM3D [59] on each of the two layers. Thus, VBM3D exploits the specific characteristics of each layer and is able to find more matched blocks to filter over, resulting in better denoising performance. Practically, very noisy videos become easier if the denoiser is applied to each layer separately or to only the layer of interest. Fig. 17 shows examples of video denoising and enhancement, respectively. For video denoising, we compare RPCA-VBM3D [113] with VBM3D [59]. For video enhancement, we show the comparison between the RPCA algorithm called ReProCS [110] and the histogram equalization which is the standard approach for such low light data. In each case, the RPCA algorithms outperform the classical state-of-the-art method. The code for this experiment is downloadable from http://www.ece.iastate.edu/∼hanguo/ReLD_Denoising.zip.

K. Video Summarization

Video summarization is a quick way to overview its content, and it is a challenging problem because finding important or informative parts of the original video requires understanding its content. Furthermore, the content of videos is very diverse, ranging from home videos to documentaries, which makes video summarization much more difficult as prior knowledge is almost unavailable. To tackle this problem, Ramani and Atia [241] employed a scalable column/row subspace pursuit algorithm based on the RPCA formulation that enables sampling in challenging scenarios in which the data are highly structured. The idea consists of searching for a set of columns whose low-rank component can cancel out the low-rank component of all the columns. Thus, informative columns are employed for video summarization. For face sampling, Ramani and Atia [241] tested this algorithm on the Yale Face Database B which consists of face images from 38 human subjects. For each subject, there are 64 images with different illuminations. $\mathbf{A}$ containing the vectorized image is built with the images of six human subjects (384 images in total, so $\mathbf{A} \in \mathbb{R}^{32,256\times 384}$. Experiments [241] show that this sampling algorithm is robust in the presence of corrupted data.

L. UHD Superresolution Video

The recovery of high-resolution (HR) images and videos from low-resolution (LR) content is a topic of great interest in digital image processing. The global superresolution (SR) problem assumes that the LR image is a noisy, low-pass filtered, and downsampled version of the HR image. Recent approaches are sparsity-based techniques which assume that image patches can be well represented as a sparse linear combination of elements from an appropriately chosen overcomplete dictionary. In order to fully utilize the spatiotemporal information, Ebadi *et al.* [79] employed a multiframe video SR approach that is aided by a low-rank plus sparse decomposition of the video sequence. First, Ebadi *et al.* [79] defined the group of pictures (GOP) structure and sought a rank-1 low-rank part that recovers the shared spatiotemporal information among the frames in the GOP. Then, the low-rank frames and the sparse frames were superresolved separately. This algorithm results in significant time reduction as well as surpassing state-of-the-art performance, both qualitatively and quantitatively.

V. CONCLUSION

The RPCA formulation has been successfully applied in the last seven to ten years in computer vision applications, outperforming previous state-of-the-art techniques. This success is due to its robustness to outliers and its flexibility to be applied in different types of outliers due to its ability

to allow specific additional constraints such as spatial and temporal ones. In the early times, its memory and time requirements limited its applications in online and/or real-time applications. However, dynamic RPCA [187], [285] has received significant attention recently, reducing these limitations with performance guarantees [188], [186], [208], [326] and memory-efficient algorithms [207], and thus allowing to consider its uses in very challenging applications such as background/foreground separation in videos taken with static or moving cameras.

However, there are still many important issues which need to be solved to allow the RPCA formulation to be fully and broadly employed in image and video processing and 3-D computer vision. The first issue concerns the guarantee for dynamic RPCA under even weaker assumptions. Second, even if robust matrix completion and undersampled robust PCA have been well studied, their dynamic extensions have received almost no attention. It is an important question for very long image or video data sets where a changing subspace assumption is a more appropriate one. Third, simple and provable RPCA or dynamic RPCA solutions that are streaming are required in several computer vision applications. Even if a streaming RPCA solution has been developed in recent work [212], it works only for 1-D RPCA. On the other hand, ReProCS [233] is a nearly memory optimal solution to dynamic RPCA, but it requires more than one pass through the data.

An open question is how the RPCA formulation can be successfully adapted to solve other more general computer vision problems. One such problem is subspace clustering which involves clustering a given image or video data set into one of K different low-dimensional subspaces. This can be viewed as a generalization of PCA which tries to represent a given data set using a single low-dimensional subspace. There has been a lot of work on the subspace clustering problem, developed in the frameworks of both sparse representation [80], [81] and low-rank representation [179], [180], where each sample is represented by other samples and the representation matrix is regularized by either sparsity [80], [81], low-rankness [179], [180], or both [80]. Other works also concern scalable subspace clustering [286] which can be solved using algorithms [286], [322], [323] that are provably correct when subspaces are sufficiently separated and data are well distributed within each subspace. A complete review can be found in [178]. Then, given that subspace clusters have been computed for a given data set, if more data vectors come in sequentially, how can one incrementally solve the clustering problem, i.e., either classify the new vector into one of the K subspaces, or decide that it belongs to a new subspace? There has been sporadic work on this problem. For example, Shen *et al.* proposed an online version of low-rank subspace clustering [260].

Another open question is whether one can solve the phaseless RPCA or $\mathbf{L} + \mathbf{S}$ problem. Indeed, one can only acquire magnitude-only measurements in applications such as ptychography, subdiffraction imaging, or astronomy. If the unknown image sequence is well modeled in the RPCA formulation, the main question is how this model can be exploited to recover it from undersampled phaseless measurements.

Finally, this paper does not review the literature on the recent works on RPCA for tensor data. Interested readers may refer to the works in [56], [162], [273], and [274] for application to background/foreground separation; the works in [5], [138], [140], [209], [235], [265], and [268] for online/incremental tensor algorithms; and the works of Lin *et al.* [178] for some recent results. All of the above are active research topics with many open questions. ■

REFERENCES

[1] R. Achanta, S. Hemami, F. Estrada, and S. Susstrun, "Frequency tuned salient region detection," in *Proc. CVPR*, pp. 1597–1604, 2009.

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC: Super-pixels compared to state of the art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2281, Nov. 2012.

[3] J.-H. Ahn, "Fast adaptive robust subspace tracking for online background subtraction," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, 2014, pp. 2555–2559.

[4] Z. An, "Video background modeling based on optimization algorithms of robust PCA," M.S. thesis, Shang Xueba Univ., China, Feb. 2014.

[5] A. Anandkumar, P. Jain, Y. Shi, andi U. Niranjan, "Tensor vs matrix methods: Robust tensor decomposition under block sparse perturbations," in *Proc. Int. Conf. Artif. Intell. Statist. (AISTATS)*, Cadiz, Spain, 2016, pp. 268–276.

[6] M. Anderson, G. Ballard, J. Demme, and K. Keutzer, "Communication-avoiding QR decomposition for GPUs," in *Proc. IEEE Int. Parallel Distrib. Process. Symp. (IPDPS)*, May 2011, pp. 48–58.

[7] A. Aravkin, S. Becker, V. Cevher, and P. Olsen, "A variational approach to stable principal component pursuit," in *Proc. Conf. Uncertain. Artif. Intell. (UAI)*, Jul. 2014.

[8] F. Arrigoni, L. Magri, B. Rossi, P. Fragneto, and A. Fusiello, "Robust absolute rotation estimation via low-rank and sparse matrix decomposition," in *Proc. 3DV*, Dec. 2014, pp. 491–498.

[9] F. Arrigoni, B. Rossi, and A. Fusiello, "Robust and efficient camera motion synchronization via matrix decomposition," in *Proc. Int. Conf. Image Process. (ICIAP)*, Sep. 2015, pp. 444–455.

[10] G. Arvanitis, A. Lalos, K. Moustakas, and N. Fakotakis, "Real-time removing of outliers and noise in 3D point clouds applied in robotic applications," in *Proc. Int. Conf. Robot. (ICR)*, 2017, pp. 11–19.

[11] D. Atkinson, D. Burnstein, and R. Edelman, "First-pass cardiac perfusion: Evaluation with ultrafast MR imaging," *Radiology*, vol. 174, no. 3, pp. 757–762, 1990.

[12] D. Atkinson and R. Edelman, "Cineangiography of the heart in a single breath hold with a segmented turboflash sequence," *Radiology*, vol. 178, no. 2, pp. 357–360, 1991.

[13] S. H. Baete, J. Chen, Y. C. Lin, R. X. Wang, R. Otazo, and F. E. Boada, "Low rank plus sparse decomposition of ODFs for improved detection of group-level differences and variable correlations in white matter," *NeuroImage*, vol. 174, pp. 138–152, Jul. 2018.

[14] S. Baete, J. Chen, R. Otazo, and F. Boada, "Low rank plus sparse decomposition of ODF distributions for improved detection of group differences in diffusion spectrum imaging," in *Proc. ISMRM*, 2016, p. 1045.

[15] F. El Baf, T. Bouwmans, and B. Vachon, "A fuzzy approach for background subtraction," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2008, pp. 2648–2651.

[16] F. El Baf, T. Bouwmans, and B. Vachon, "Fuzzy integral for moving object detection," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Jun. 2008, pp. 1729–1736.

[17] F. El Baf, T. Bouwmans, and B. Vachon, "Type-2 fuzzy mixture of Gaussians model: Application to background modeling," in *Proc. Int. Symp. Vis. Comput. (ISVC)*, Dec. 2008, pp. 772–781.

[18] A. Baghaie, R. M. D'Souza, and Z. Yu, "Sparse and low rank decomposition based batch image alignment for speckle reduction of retinal OCT images," in *Proc. IEEE Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 226–230.

[19] M. Balcan, Y. Liang, D. Woodruff, and H. Zhang, "Matrix completion and related problems via strong duality," *Innov. Theor. Comput. Sci.*, vol. 94, pp. 2955–2963, Apr. 2018.

[20] M. Balcan and H. Zhang, "Noise-tolerant life-long matrix completion via adaptive sampling," in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 2955–2963.

[21] A. Bhardwaj and S. Raman, "Robust PCA-based solution to image composition using augmented Lagrange multiplier (ALM)," *Vis. Comput.*, vol. 32, no. 5, pp. 591–600, Mar. 2015.

[22] X. Bian and H. Krim, "Bi-sparsity pursuit for robust subspace recovery," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 3535–3539.

[23] F. Biondi, "Low rank plus sparse decomposition of synthetic aperture radar data for maritime surveillance," in *Proc. Int. Workshop Compress. Sens. Theory Appl. Sonar Remote Sens. (CoSeRa)*, 2016, pp. 75–79.

[24] F. Biondi, "Low-rank plus sparse decomposition and localized radon transform for ship-wake detection in synthetic aperture radar images," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 1, pp. 117–121, Jan. 2018.

[25] T. Bouwmans, "Subspace learning for background modeling: A survey," *Recent Patents Comput. Sci.*, vol. 2, no. 3, pp. 223–234, Nov. 2009.

[26] T. Bouwmans, "Recent advanced statistical background modeling for foreground detection: A systematic survey," *Recent Patents Comput. Sci.*, vol. 4, no. 3, pp. 147–176, 2011.

[27] T. Bouwmans, "Background subtraction for visual surveillance: A fuzzy approach," in *Handbook on Soft Computing for Video Surveillance*, S. K. Pal, A. Petrosino, and L. Maddalena, Eds. New York, NY, USA: Taylor & Francis, 2012, pp. 103–139, ch. 5. i

[28] T. Bouwmans, N. Aybat, and E. Zahzah, *Handbook on Robust Low-Rank and Sparse Matrix Decomposition: Applications in Image and Video Processing*. Boca Raton, FL, USA: CRC Press, 2016.

[29] T. Bouwmans, F. El Baf, and B. Vachon, "Background modeling using mixture of Gaussians for foreground detection—A survey," *Recent Patents Comput. Sci.*, vol. 1, no. 3, pp. 219–237, Nov. 2008.

[30] T. Bouwmans, F. Porikli, B. Horferlin, and A. Vacavant, *Handbook on 'Background Modeling and Foreground Detection for Video Surveillance'* Boca Raton, FL, USA: CRC Press, Jul. 2014.

[31] T. Bouwmans, A. Sobral, S. Javed, S. Ki Jung, and El-Z. Zahzah, "Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset," *Comput. Sci. Rev.*, vol. 23, pp. 1–17, Nov. 2016.

[32] T. Bouwmans and E. Zahzah, "Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance," in *Proc. Special Issue Background Models Challenge Comput. Vis. Image Understanding (CVIU)*, 2014, pp. 22–34.

[33] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *Int. J. Comput. Vis.*, vol. 74, no. 1, pp. 59–73, Aug. 2007.

[34] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2005, pp. 60–65.

[35] H. Cai, J. Cai, and K. Wei. (Jan. 2018). "Accelerated alternating projections for robust principal component analysis." [Online]. Available: https://arxiv.org/pdf/1711.05519.pdf

[36] N. A. Campell, "Robust procedures in multivariate analysis I: Robust covariance estimation," *Appl. Stat.*, vol. 1, no. 29, pp. 231–237, 1980.

[37] E. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Int. J. ACM*, vol. 58, no. 3, May 2011, Art. no. 11.

[38] E. Candès and M. Soltanolkotabi, "Discussion: Latent variable graphical model selection via convex optimization," *Ann. Stat.*, vol. 40, no. 4, pp. 1968–1972, 2012.

[39] J. Cao, J. Zhou, X. Liu, W. Wang, P. Tao, and J. Wang, "Low-rank image completion with entropy features," *Mach. Vis. Appl.*, vol. 28, nos. 1–2, pp. 129–139, Feb. 2017.

[40] X. Cao *et al.*, "Low-rank matrix factorization under general mixture noise distributions," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1493–1501.

[41] X. Cao, L. Yang, and X. Guo, "Total variation regularized RPCA for irregularly moving object detection under dynamic background," *IEEE Trans. Cybern.*, vol. 46, no. 4, pp. 1014–1027, Apr. 2016.

[42] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky, "Rank sparsity incoherence for matrix decomposition," *SIAM J. Optim.*, vol. 21, no. 2, pp. 572–596, 2011.

[43] J. Chang and T. Gerhart, "Applications of low rank and sparse matrix decompositions in hyperspectral video processing," in *Handbook on Robust Low-Rank and Sparse Matrix Decomposition: Applications in Image and Video Processing*. Hoboken, NJ, USA: CRC Press, May 2016, ch. 16.

[44] X. Chang, S. Liu, P. Zhao, and X. Li, "Convergent prediction–correction-based ADMM for multi-block separable convex programming," *J. Comput. Appl. Math.*, vol. 335, pp. 270–288, Jun. 2018.

[45] G. Chau and P. Rodríguez, "Panning and jitter invariant incremental principal component pursuit for video background modeling," in *Proc. Int. Workshop RSL-CV Conjunction (ICCV)*, Oct. 2017, pp. 1844–1852.

[46] C. Chen, J. Cai, W. Lin, and G. Shi, "Surveillance video coding via low-rank and sparse decomposition," in *Proc. ACM Int. Conf. Multimedia*, 2012, pp. 713–716.

[47] C. Chen, J. Cai, W. Lin, and G. Shi, "Incremental low-rank and sparse decomposition for compressing videos captured by fixed cameras," *J. Vis. Commun. Image Represent.*, vol. 26, pp. 338–348, Dec. 2014.

[48] C. Chen, S. Li, H. Qin, and A. Hao, "Robust salient motion detection in non-stationary videos via novel integrated strategies of spatio-temporal coherency clues and low-rank analysis," *Pattern Recognit.*, vol. 52, pp. 410–432, Apr. 2016.

[49] J. Chen, S. Liu, and M. Huang, "Low-rank and sparse decomposition model for accelerating dynamic MRI reconstruction," *J. Healthcare Eng.*, vol. 2017, Aug. 2017, Art. no. 9856058.

[50] L. Chen, B. Yang, and X. Wang, "Dynamic magnetic resonance imaging via nonconvex low-rank matrix approximation," *Math. Problems Eng.*, 2017.

[51] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He, "Denoising of hyperspectral images using nonconvex low rank matrix approximation," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5366–5380, Sep. 2017.

[52] Z. Chen and B. Wang, "Spectrally-spatially regularized low-rank and sparse decomposition: A novel method for change detection in multitemporal hyperspectral images," *MDPI Remote Sens.*, vol. 9, no. 10, p. 1044, Oct. 2017.

[53] B. Cheng, G. Liu, J. Wang, Z. Huang, and S. Yan, "Multi-task low-rank affinity pursuit for image segmentation," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 2439–2446.

[54] J. Cheng, K. Chen, and M. Sacchi, "Robust principal component analysis (RPCA) for seismic data denoising," in *Proc. GeoConvention*, 2015.

[55] Y. Cherapanamjeri, K. Gupta, and P. Jain, "Nearly-optimal robust matrix completion," in *Proc. Int. Conf. Mach. Learn. (PMLR)*, Sydney, NSW, Australia, 2017.

[56] A. Cichocki, R. Zdunek, A. Phan, and S. Amari, *Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-Way Data Analysis and Blind Source Separation*. Hoboken, NJ, USA: Wiley, 2009.

[57] C. Croux and A. Ruiz-Gazen, "A fast algorithm for robust principal components based on projection pursuit' computational statistics," in *COMPSTAT*. 1996.

[58] X. Cui, Q. Liu, and D. Metaxas, "Temporal spectral residual: Fast motion saliency detection," in *Proc. ACM MM*, 2009, pp. 617–620.

[59] K. Dabov, A. Foi, and K. Egiazarian, "Video denoising by sparse 3D transform-domain collaborative filtering," in *Proc. Eur. Conf. Signal Process.*, Sep. 2007, pp. 145–149.

[60] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[61] C. Dang, A. Moghadam, and H. Radha, "RPCA-KFE: Key frame extraction for consumer video based robust principal component analysis," *IEEE Trans. Image Process.*, vol. 24, no. 11, 2015.

[62] A. Das, "A Bayesian sparse-plus-low-rank matrix decomposition method for direction-of-arrival tracking," *IEEE Sensors J.*, vol. 17, no. 15, pp. 4894–4902, Aug. 2017.

[63] Y. Deng, Q. Dai, R. Liu, and Z. Zhang, "Low-rank structure learning via nonconvex heuristic recovery," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 3, pp. 383–396, Mar. 2013.

[64] W. Dong, G. Shi, and X. Li, "Nonlocal image restoration with bilateral variance estimation: A low-rank approach," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 700–711, Feb. 2013.

[65] W. Dong, L. Zhang, and G. Shi, "Centralized sparse representation for image restoration," in *Proc. ICCV*, Nov. 2011, pp. 1259–1266.

[66] Y. Dong and G. DeSouza, "Adaptive learning of multi-subspace for foreground detection under illumination changes," *Comput. Vis. Image Understanding*, vol. 115, no. 1, pp. 31–49, 2010.

[67] Y. Dong, T. Han, and G. DeSouza, "Illumination invariant foreground detection using multi-subspace learning," *J. Int. Knowl.-Based Intell. Eng. Syst.*, vol. 14, no. 1, pp. 31–41, 2010.

[68] L. Duarte, E. Nadalin, K. Filho, R. Zanetti, J. Romano, and M. Tygel, "Seismic wave separation by means of robust principal component analysis," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, 2012, pp. 1494–1498.

[69] A. Dutta, B. Gong, X. Li, and M. Shah, "Weighted singular value thresholding and its application to background estimation," *J. Mach. Learn. Res.*, vol. 17, pp. 1–22, 2017.

[70] A. Dutta and X. Li, "A fast algorithm for a weighted low rank approximation," in *Proc. Int. Conf. Mach. Vis. Appl.*, May 2017, pp. 93–96.

[71] A. Dutta, X. Li, and P. Richtarik, "A batch-incremental video background estimation model using weighted low-rank approximation of matrices," in *Proc. Workshop RSL-CV Conjunction (ICCV)*, Oct. 2017, pp. 1835–1843.

[72] S. E. Ebadi and E. Izquierdo, "Approximated RPCA for fast and efficient recovery of corrupted and linearly correlated images and video frames," in *Proc. IEEE Int. Conf. Syst. Signals Image Process. (IWSSIP)*, Sep. 2015, pp. 49–52.

[73] S. E. Ebadi and E. Izquierdo, "Efficient background subtraction with low-rank and sparse matrix decomposition," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 4863–4867.

[74] S. E. Ebadi and E. Izquierdo, "Foreground segmentation via dynamic tree-structured sparse RPCA," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 314–329.

[75] S. E. Ebadi and E. Izquierdo, "Foreground segmentation with tree-structured sparse RPCA," *IEEE Trans. Pattern Anal. Mach. Intell.*, Aug. 2017.

[76] S. E. Ebadi, V. G. One, and E. Izquierdo, "Approximated robust principal component analysis for improved general scene background subtraction," *IEEE Trans. Image Process.*, to be published. [Online]. Available: https://arxiv.org/abs/1603.05875

[77] S. E. Ebadi, V. G. One, and E. Izquierdo, "Efficient background subtraction with low-rank and sparse matrix decomposition," in *Proc. Workshop Signal Process. Adapt. Sparse Struct. Represent. (SPARS)*, Jul. 2015, pp. 4863–4867.

[78] S. E. Ebadi, V. G. One, and E. Izquierdo, "Dynamic tree-structured sparse RPCA via column subset selection for background modeling and foreground detection," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3972–3976.

[79] S. E. Ebadi, V. Ones, and E. Izquierdo, "UHD video super-resolution using low-rank and sparse decomposition," in *Proc. Int. Workshop RSL-CV Conjunction (ICCV)*, Oct. 2017, pp. 1889–1897.

[80] E. Elhamifar and R. Vidal, "Sparse subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 2790–2797.

[81] E. Elhamifar and R. Vidal, "Sparse subspace clustering: Algorithm, theory, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2765–2781, Nov. 2013.

[82] A. Elnakeeb and U. Mitra, "Low-rank, sparse and line constrained estimation: Applications to target

tracking and convergence," in *Proc. IEEE ISIT*, Jun. 2017, pp. 2283–2287.
[83] A. ElTantawy and M. Shehata, "Moving object detection from moving platforms using Lagrange multiplier," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 2586–2590.
[84] A. ElTantawy and M. Shehata, "UT-MARO: Unscented transformation and matrix rank optimization for moving objects detection in aerial imagery," in *Proc. Int. Symp. Vis. Comput. (ISVC)*, 2015, pp. 275–284.
[85] A. ElTantawy and M. Shehata, "A novel method for segmenting moving objects in aerial imagery using matrix recovery and physical spring model," in *Proc. ICPR*, 2016, pp. 3898–3903.
[86] A. ElTantawy and M. Shehata, "MARO: Matrix rank optimization for the detection of small-size moving objects from aerial camera platforms," *Signal Image Video Process.*, vol. 12, no. 4, pp. 641–649, 2018.
[87] N. Erichson, S. Voronin, S. Brunton, and J. Kutz. (Apr. 2018). "Randomized Matrix Decompositions using R." [Online]. Available: https://arxiv.org/abs/1603.05875
[88] D. Farcas and T. Bouwmans, "Background modeling via a supervised subspace learning," in *Proc. Int. Conf. Image Video Process. Comput. Vis. (IVPCV)*, Jul. 2010, pp. 1–7.
[89] D. Farcas, C. Marghes, and T. Bouwmans, "Background subtraction via incremental maximum margin criterion: A discriminative approach," *Mach. Vis. Appl.*, vol. 23, no. 6, pp. 1083–1101, Oct. 2012.
[90] Y. Fu, W. Wang, and C. Wang, "Image change detection method based on RPCA and low-rank decomposition," in *Proc. Chin. Control Conf. (CCC)*, 2016, pp. 9412–9417.
[91] N. Gallo and N. Gelfand, "Artifact-free high dynamic range imaging," in *Proc. IEEE Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Jul. 2009, pp. 9412–9417.
[92] C. Gao, B. Moore, and R. Nadakuditi, "Augmented robust PCA for foreground-background separation on noisy, moving camera video," in *Proc. IEEE GlobalSIP*, 2017, pp. 1240–1244.
[93] H. Gao, J. Cai, Z. Shen, and H. Zhao, "Robust principal component analysis-based four-dimensional computed tomography," *Phys. Med. Biol.*, vol. 56, no. 11, pp. 3181–3198, 2011.
[94] H. Gao *et al.*, "Compressed sensing using prior rank, intensity and sparsity model (PRISM): Applications in cardiac cine MRI," in *Proc. ISMRM*, p. 2242, 2012.
[95] J. Gao, S. Kim, and M. Brown, "Constructing image panoramas using dual-homography warping," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 49–56.
[96] Z. Gao, L. Cheong, and M. Shan, "Block-sparse RPCA for consistent foreground detection," in *Proc. ECCV*, 2012, pp. 690–703.
[97] Z. Gao, L. Cheong, and M. Shan, "Block-sparse RPCA for salient motion detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 10, pp. 1975–1987, Oct. 2014.
[98] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *Proc. IEEE Conf. Comput. Vis. (CVPR)*, Jun. 2012, pp. 3354–3361.
[99] J. Geng, J. Fan, and H. Wang, "Weighted fusion-based representation classifiers for marine floating raft detection of SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 3, pp. 444–448, Mar. 2017.
[100] T. Gerhart, "Convex optimization techniques and their application in hyperspectral video processing, " M.S. thesis, California State Univ., Long Beach, CA, USA, Dec. 2013.
[101] J. Giraldo-Zuluaga, A. Gomez, A. Salazar, and A. Diaz-Pulido, "Automatic recognition of mammal genera on camera-trap images using multi-layer robust principal component analysis and mixture neural networks," in *Proc. IEEE Int. Conf. Tools Artif. Intell. (ICTAI)*, 2017, pp. 53–60.
[102] J. Giraldo-Zuluaga, A. Gomez, A. Salazar, and A. Diaz-Pulido, "Camera-trap images segmentation using multi-layer robust principal component analysis," *Vis. Comput.*, pp. 1–13, Dec. 2017.
[103] S. Goferman, L. Zelnikmanor, and A. Tal, "Context aware saliency detection," in *Proc. CVPR*, 2010, pp. 2376–2383.
[104] A. Gogna, A. Shukla, H. Agarwal, and A. Majumda, "Split Bregman algorithms for sparse/joint-sparse and low-rank signal recovery: Application in compressive hyperspectral imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2014, pp. 1302–1306.
[105] C. G. Gonzalez, O. Absil, P. Absil, M. Van Droogenbroeck, D. Mawet, and J. Surdej, "Low-rank plus sparse decomposition for exoplanet detection in direct-imaging ADI sequences. The LLSG algorithm," *Astron. Astrophys.*, p. A54, 2016.
[106] N. Goyette, P. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "Changedetection.net: A new change detection benchmark dataset," in *Proc. IEEE Workshop Change Detection (CVPR)*, Jun. 2012, pp. 1–8.
[107] N. Goyette, P. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "A novel video dataset for change detection benchmarking," *IEEE Trans. Image Process.*, vol. 23, no. 11, pp. 4663–4679, Nov. 2014.
[108] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang, "Weighted nuclear norm minimization and its applications to low level vision," *Int. J. Comput. Vis.*, vol. 121, no. 2, pp. 183–208, Jan. 2017.
[109] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2862–2869.
[110] H. Guo, C. Qiu, and N. Vaswani, "An online algorithm for separating sparse and low-dimensional signal sequences from their sum," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4284–4297, Aug. 2014.
[111] H. Guo, C. Qiu, and N. Vaswani, "Practical ReProCS for separating sparse and low-dimensional signal sequences from their sum—Part 1," in *Proc. ICASSP*, May 2014, pp. 4161–4165.
[112] H. Guo, C. Qiu, and N. Vaswani, "Practical ReProCS for separating sparse and low-dimensional signal sequences from their sum—Part 2," in *Proc. GlobalSIP*, 2014, pp. 369–373.
[113] H. Guo and N. Vaswani, "Video denoising and enhancement via dynamic video layering," to be published.
[114] X. Guo, S. Li, and X. Cao, "Motion matters: A novel framework for compressing surveillance videos," in *Proc. ACM Int. Conf. Multimedia*, Oct. 2013, pp. 549–552.
[115] X. Guo, X. Wang, L. Yang, X. Cao, and Y. Ma, "Robust foreground detection using smoothness and arbitrariness constraints," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, Sep. 2014, pp. 535–550.
[116] C. Guyon, T. Bouwmans, and E. Zahzah, "Foreground detection based on low-rank and block-sparse matrix decomposition," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2012, pp. 1225–1228.
[117] C. Guyon, T. Bouwmans, and E. Zahzah, "Foreground detection via robust low rank matrix decomposition including spatio-temporal constraint," in *Proc. Int. Workshop Background Model Challenges (ACCV)*, Nov. 2012, pp. 315–320.
[118] C. Guyon, T. Bouwmans, and E. Zahzah, "Foreground detection via robust low rank matrix factorization including spatial constraint with iterative reweighted regression," in *Proc. Int. Conf. Pattern Recognit. (ICPR)*, Nov. 2012, pp. 2805–2808.
[119] C. Guyon, T. Bouwmans, and E. Zahzah, "Moving object detection via robust low rank matrix decomposition with IRLS scheme," in *Proc. Int. Symp. Vis. Comput. (ISVC)*, Jul. 2012, pp. 665–674.
[120] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in *Proc. Neural Inf. Process. Syst. (NIPS)*, 2006, pp. 545–552.
[121] J. He, L. Balzano, and J. Luiz, "Online robust subspace tracking from partial information," in *Proc. IT*, Sep. 2011.
[122] J. He, L. Balzano, and A. Szlam, "Incremental gradient on the grassmannian for online foreground and background separation in subsampled video," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2012, pp. 1568–1575.
[123] J. He, D. Zhang, L. Balzano, and T. Tao, "Iterative Grassmannian optimization for robust image alignment," *Image Vis. Comput.*, vol. 32, no. 10, pp. 800–813, Jun. 2013.
[124] J. He, D. Zhang, L. Balzano, and T. Tao, "Iterative online subspace learning for robust image alignment," in *Proc. IEEE Conf. Autom. Face Gesture Recognit. (FG)*, Apr. 2013, pp. 1–8.
[125] R. He, T. Tan, and L. Wang, "Robust recovery of corrupted low-rankmatrix by implicit regularizers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 4, pp. 770–783, Sep. 2013.
[126] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, 1981.
[127] Y. Hou and Z. Lin, "Image tag completion and refinement by subspace clustering and matrix completion," in *Proc. Vis. Commun. Image Process.*, 2015, pp. 1–4.
[128] D. Hsu, S. M. Kakade, and T. Zhang, "Robust matrix decomposition with sparse corruptions," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7221–7234, Nov. 2011.
[129] G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database forstudying face recognition in unconstrained environments," in *Proc. ECCV Workshop Faces Real-Life Images*, 2008.
[130] P. Huang, S. Chen, P. Smaragdis, and M. Hasegawa-Johnson, "Singing-voice separation from monaural recordings using robust principal component analysis," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 57–60.
[131] S. Huang, J. Ye, T. Wang, L. Jiang, X. Wu, and Y. Li, "Extracting refined low-rank features of robust PCA for human action recognition," *Arabian J. Sci. Eng.*, vol. 40, no. 2, pp. 1427–1441, Mar. 2015.
[132] Y. Huang, G. Liao, J. Xu, J. Li, and D. Yang, "GMTI and parameter estimation for MIMO SAR system via fast interferometry RPCA method," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1774–1787, Mar. 2018.
[133] P. Huber, *Robust Statistics*. New York, NY, USA: Wiley, 1981.
[134] M. Hubert, P. J. Rousseeuw, and K. V. Branden, "A fast method for robust principal components with applications to chemometrics," *Chemometrics Intell. Lab. Syst.*, vol. 60, no. 1, pp. 101–111, 2002.
[135] W.-D. Jang, C. Lee, and C.-S. Kim, "Primary object segmentation in videos via alternate convex optimization of foreground and background distributions," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 696–704.
[136] S. Javed, T. Bouwmans, and S. Jung, "Combining ARF and OR-PCA for robust background subtraction of noisy videos," in *Proc. Int. Conf. Image Anal. Appl. (ICIAP)*, Sep. 2015, pp. 340–351.
[137] S. Javed, T. Bouwmans, and S. Jung, "Depth extended online RPCA with spatiotemporal constraints for robust background subtraction," in *Proc. Korea-Japan Workshop Frontiers Comput. Vis. (FCV)*, Jan. 2015, pp. 1–6.
[138] S. Javed, T. Bouwmans, and S. Jung, "Stochastic decomposition into low rank and sparse tensor for robust background subtraction," in *Proc. ICDP*, Jul. 2015, pp. 1–6.
[139] S. Javed, T. Bouwmans, and S. Jung, "Improving OR-PCA via smoothed spatially-consistent low-rank modeling for background subtraction," in *Proc. ACM Symp. Appl. Comput.*, 2017, pp. 89–94.
[140] S. Javed, T. Bouwmans, and S. Jung, "SBMI-LTD:

Stationary background model initialization based on low-rank tensor decomposition," in *Proc. ACM Symp. Appl. Comput. (SAC)*, 2017, pp. 95–200.
[141] S. Javed, A. Mahmood, T. Bouwmans, and S. K. Jung, "Motion-aware graph regularized RPCA for background modeling of complex scenes," in *Proc. Int. Conf. Pattern Recognit. Scene Background Modeling Contest (ICPR)*, Dec. 2016, pp. 120–125.
[142] S. Javed, A. Mahmood, T. Bouwmans, and S. K. Jung, "Spatiotemporal low-rank modeling for complex scene background initialization," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1315–1329, Dec. 2016.
[143] S. Javed, A. Mahmood, T. Bouwmans, and S. Jung, "Background-foreground modeling based on spatio-temporal sparse subspace clustering," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5840–5854, Dec. 2017.
[144] S. Javed, S. Oh, T. Bouwmans, and S. Jung, "Robust background subtraction to global illumination changes via multiple features-based online robust principal components analysis with Markov random field," *J. Electron. Imag.*, vol. 24, no. 4, p. 043011, 2015.
[145] S. Javed, S. Oh, A. Sobral, T. Bouwmans, and S. Jung, "Background subtraction via superpixel-based online matrix decomposition with structured foreground constraints," in *Proc. Workshop Robust Subspace Learn. Comput. Vis. (ICCV)*, Dec. 2015, pp. 90–98.
[146] S. Javed, A. Sobral, S. Oh, T. Bouwmans, and S. Jung, "OR-PCA with MRF for robust foreground detection in highly dynamic backgrounds," in *Proc. Asian Conf. Comput. Vis. (ACCV)*, 2014, pp. 284–299.
[147] H. Ji, S. Huang, Z. Shen, and Y. Xu, "Robust video restoration by joint sparse and low rank matrix approximation," *SIAM J. Imag. Sci.*, vol. 4, no. 4, pp. 1122–1142, 2011.
[148] M. Jin, R. Li, J. Jiang, and B. Qin, "Extracting contrast-filled vessels in X-ray angiography by graduated RPCA with motion coherency constraint," *Pattern Recognit.*, vol. 63, pp. 653–666, Mar. 2017.
[149] H. Jung and J. Ye, "Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: What we can learn from video compression techniques," *Int. J. Imag. Syst. Technol.*, vol. 20, no. 2, pp. 81–98, 2010.
[150] M. Kaloorazi and R. Lamare, "Subspace-orbit randomized decomposition for low-rank matrix approximation," in *Proc. EUSIPCO*, Roma, Italy, Sep. 2018.
[151] F. Korosec, R. Frayne, T. Grist, and C. Mistretta, "Time-resolved contrast-enhanced 3D MR angiography," *Magn. Reson. Med.*, vol. 36, no. 3, pp. 345–351, 1996.
[152] C. Lang, G. Liu, J. Yu, and S. Yan, "Saliency detection by multitask sparsity pursuit," *IEEE Trans. Image Process.*, vol. 21, no. 3, pp. 1327–1338, Mar. 2012.
[153] I. Laptev and T. Lindeberg, "Velocity adaptation of space-time interest points," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Aug. 2004, pp. 52–56.
[154] C. Lau, Y. Lai, and L. Lui. (Sep. 2017). "Restoration of atmospheric turbulence-distorted images via RPCA and quasiconformal maps." [Online]. Available: https://arxiv.org/pdf/1704.03140.pdf
[155] T. Lee, "Analysis of DS-CDMA receivers using the robust PCA against multi-tone jamming for land mobile satellite channels, " M.S. thesis, School Elect. Eng. Comput. Sci., Gwangju Inst. Sci. Technol., Gwangju, South Korea, 2017.
[156] W. Leow, Y. Cheng, L. Zhang, T. Sim, and L. Foo, "Background recovery by fixed-rank robust principal component analysis," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, 2013, pp. 54–61.
[157] C. Li, L. Lin, W. Zuo, W. Wang, and J. Tang, "SOLD: Sub-optimal low-rank decomposition for efficient video segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 5519–5527.
[158] C. Li, L. Lin, W. Zuo, W. Wang, and J. Tang, "An approach to streaming video segmentation with sub-optimal low-rank decomposition," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 1947–1960, May 2016.
[159] H. Li, Z. Mia, Y. Li, Y. Xu, and Y. Zhang, "Moving object detection via box constrained RPCA," *J. PLA Univ. Sci. Technol.*, vol. 17, no. 5, pp. 403–407, Oct. 2016.
[160] J. Li, Y. Huang, G. Liao, and J. Xu, "Moving target detection via efficient ATI-GoDec approach for multichannel SAR system," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 9, pp. 1320–1324, Sep. 2016.
[161] L. Li, W. Huang, I. Y.-H. Gu, and Q. Tian, "Statistical modeling of complex backgrounds for foreground object detection," *IEEE Trans. Image Process.*, vol. 13, no. 11, pp. 1459–1472, Nov. 2004.
[162] L. Li, P. Wang, Q. Hu, and S. Cai, "Efficient background modeling based on sparse representation and outlier iterative removal," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 2, pp. 278–289, Feb. 2014.
[163] Q. Li and Z. Wang, "Riemannian submanifold tracking on low-rank algebraic variety," in *Proc. AAAI Conf. Artif. Intell. (AAAI)*, 2017, pp. 2196–2202.
[164] Q. Li, H. Yan, L. Wu, and R. Wang, "Robust PCA for ground moving target indication in wide-area surveillance radar system," *J. Oper. Res. Soc. China*, vol. 1, no. 1, pp. 135–153, Mar. 2013.
[165] X. Li and J. Haupt, "Identifying outliers in large matrices via randomized adaptive compressive sampling," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1792–1807, Apr. 2015.
[166] X. Li and J. Haupt, "Outlier identification via randomized adaptive compressive sampling," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2015, pp. 3302–3306.
[167] Y. Li, "On incremental and robust subspace learning," *Pattern Recognit.*, vol. 37, no. 7, pp. 1509–1518, Jul. 2004.
[168] Y. Li and V. Monga, "SIASM: Sparsity-based image alignment and stitching method for robust image mosaicking," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Apr. 2016, pp. 3302–3306.
[169] Y. Li and W. Yu. (Apr. 2017). "A fast implementation of singular value thresholding algorithm using recycling rank revealing randomized singular value decomposition." [Online]. Available: https://arxiv.org/abs/1704.05528
[170] D. Liang, B. Liu, J. Wang, and L. Ying, "Accelerating SENSE using compressed sensing," *Magn. Reson. Med.*, vol. 62, no. 6, pp. 1574–1584, 2009.
[171] X. Liang, "Removing mixed noise in low rank textures by convex optimization," *Comput. Vis. Media*, vol. 2, no. 3, pp. 267–276, 2016.
[172] B. Lin, J. Liu, M. Xie, and J. Zhu, "Direction-of-arrival tracking via low-rank plus sparse matrix decomposition," *IEEE Antennas Wireless Propag. Lett.*, vol. 14, pp. 1302–1305, 2015.
[173] L. Lin, W. Lin, and S. Huang, "Group object detection and tracking by combining RPCA and fractal analysis," *Soft Comput.*, vol. 22, no. 1, pp. 231–242, Sep. 2016.
[174] Z. Lin, "A review on low-rank models in data analysis," *Big Data Inf. Anal.*, vol. 1, nos. 2–3, pp. 139–161, 2016.
[175] Z. Lin, M. Chen, and Y. Ma, "The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices," UIUC, Champaign, IL, USA, Tech. Rep., Nov. 2009.
[176] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proc. Adv. Neural Inf. Process. Syst.*, 2011, pp. 612–620.
[177] Z. Lin and S. Wei, "A block Lanczos with warm start technique for accelerating nuclear norm minimization algorithms," Microsoft, Redmond, WA, USA, Tech. Rep. #MSR-TR-2010-162, 2010.
[178] Z. Lin and H. Zhang, *Low-Rank Models in Visual Analysis: Theories, Algorithms, and Applications*. Orlando, FL, USA: Academic, 2017.
[179] G. Liu, Z. Lin, S. Yan, J. Sun, and Y. Ma, "Robust recovery of subspace structures by low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 171–184, Jan. 2013.
[180] G. Liu, Z. Lin, and Y. Yu, "Robust subspace segmentation by low-rank representation," in *Proc. Int. Conf. Mach. Learn.*, 2010, pp. 663–670.
[181] H. Liu and L. Chau, "Deepsea video descattering," *Multimedia Tools Appl.*, pp. 1–11, Dec. 2017.
[182] R. Liu, Z. Lin, and Z. Su, "Exactly recovering low-rank matrix in linear time via l_1 filter," *Neurocomputing*, May 2018.
[183] X. Liu, Z. Wen, and Y. Zhang, "Limited memory block Krylov subspace optimization for computing dominant singular value decomposition," *SIAM J. Sci. Comput.*, vol. 35, no. 3, pp. A1641–A1668, 2013.
[184] X. Liu, Z. Wen, and Y. Zhang, "An efficient Gauss–Newton algorithm for symmetric low-rank product matrix approximations," *SIAM J. Optim.*, vol. 25, no. 3, pp. 1571–1608, Jun. 2014.
[185] X. Liu, G. Zhao, J. Yao, and C. Qi, "Background subtraction based on low-rank and structured sparse decomposition," *IEEE Trans. Image Process.*, vol. 24, no. 8, pp. 2502–2514, Aug. 2015.
[186] B. Lois and N. Vaswani, "A correctness result for online robust PCA," in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, 2015, pp. 3791–3795.
[187] B. Lois and N. Vaswani, "Online matrix completion and online robust PCA," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jun. 2015, pp. 1826–1830.
[188] B. Lois, N. Vaswani, and C. Qiu, "Performance guarantees for undersampled recursive sparse recovery in large but structured noise," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2013, pp. 1061–1064.
[189] Y. Lou, H. Sung, S. Soatto, and A. L. Bertozzi, "Video stabilization of atmospheric turbulence distortion," *Inverse Problems Imag.*, vol. 7, no. 3, pp. 839–861, 2013.
[190] X. Luan, B. Fang, L. Liu, W. Yang, and J. Qian, "Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion," *Pattern Recognit.*, vol. 47, no. 2, pp. 495–508, 2014.
[191] H. Luong, N. Deligiannis, S. Forchhammer, and A. Kaup, "Compressive online decomposition of dynamic signals via n-ℓ_1 minimization with clustered priors," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Jun. 2018.
[192] H. Luong, N. Deligiannis, S. Forchhammer, and A. Kaup, "Compressive online decomposition of dynamic signals via n-ℓ_1 minimization with clustered priors," in *Proc. IEEE Stat. Signal Process. Workshop (SSP)*, Jun. 2018.
[193] H. Luong, N. Deligiannis, S. Forchhammer, and A. Kaup, "Online decomposition of compressive streaming data using n-ℓ_1 cluster-weighted minimization," in *Proc. Data Compress. Conf.*, Mar. 2018.
[194] H. Luong, N. Deligiannis, S. Forchhammer, and A. Kaup, "Online decomposition of compressive streaming data using n-ℓ_1 cluster-weighted minimization," in *Proc. Data Compress. Conf.*, 2018.
[195] H. Luong, N. Deligiannis, J. Seiler, S. Forchhammer, and A. Kaup. (May 2017). "Incorporating prior information in compressive online robust principal component analysis." [Online]. Available: https://arxiv.org/pdf/1701.06852.pdf
[196] M. Lustig, J. Santos, D. Donoho, and J. Pauly, "k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity," in *Proc. ISMRM*, 2006, p. 2420.
[197] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "From learning models of natural image patches to whole image restoration," in *Proc. ICCV*, 2009, pp. 479–486.
[198] M. Mardani, G. Mateos, and G. Giannakis, "Recovery of low-rank plus compressed sparse matrices with application to unveiling traffic anomalies," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5186–5205, Aug. 2013.
[199] C. Marghes and T. Bouwmans, "Background modeling via incremental maximum margin

criterion," in *Proc. Int. Workshop Subspace Methods (ACCV) Workshop Subspace*, Nov. 2010, pp. 394–403.

[200] A. Martinez, "The AR face database," CVC, Tech. Rep. 24-4, 1998.

[201] Y. Mehran and T. Bouwmans, "New trends on moving object detection in video images captured by a moving camera: A survey," *Comput. Sci. Rev.*, vol. 28, pp. 157–177, May 2018.

[202] E. Meinhardt-Llopis and M. Micheli, "Implementation of the centroid method for the correction of turbulence," *Image Process.*, vol. 4, pp. 187–195, Jul. 2014.

[203] D. Meng and F. De La Torre, "Robust matrix factorization with unknown noise," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, Dec. 2013, pp. 1337–1344.

[204] S. Minaee and Y. Wang, "Masked signal decomposition using subspace representation and its applications," *IEEE Trans. Image Precess*, submitted for publication.

[205] B. Moore, C. Gao, and R. Nadakuditi, "Panoramic robust PCA for foreground-background separation on noisy, free-motion camera video," to be published.

[206] R. Nadakuditi, "OptShrink: An algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage," *IEEE Trans. Inf. Theory*, vol. 60, no. 5, pp. 3002–3018, May 2013.

[207] P. Narayanamurthy and N. Vaswani, "MEDRoP: Memory-efficient dynamic robust PCA," in *Proc. ITA Workshop*, 2018.

[208] P. Narayanamurthy and N. Vaswani, "New results for provable dynamic robust PCA," *CoRR*, 2017.

[209] C. Navasca and X. Wang, "Adaptive low rank approximation for tensors," in *Proc. Workshop Robust Subspace Learn. Comput. Vis. (ICCV)*, Dec. 2015, pp. 939–945.

[210] P. Netrapalli, U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain, "Non-convex robust PCA," in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 1107–1115.

[211] A. Newson, M. Tepper, and G. Sapiro, "Low-rank spatio-temporal video segmentation," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, 2015, pp. 103–111.

[212] U. Niranjan and Y. Shi, "Streaming robust PCA," in *Proc. NIPS Workshop Optim. Mach. Learn.*, 2017.

[213] T. O. Pace Oreifej, G. Shu, and M. Shah, "A two-stage reconstruction approach for seeing through waterblurring from a single image," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2011, pp. 1153–1160.

[214] N. Oliver, B. Rosario, and A. Pentland, "A Bayesian computer vision system for modeling human interactions," in *Proc. ICVS*, Jan. 1999, pp. 255–272.

[215] O. Oreifej, X. Li, and M. Shah, "Simultaneous video stabilization and moving object detection in turbulence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 2, pp. 450–462, Feb. 2012.

[216] E. Otazo, D. Kim, L. Axel, and D. Sodickson, "Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI," *Magn. Reson. Med.*, vol. 64, no. 3, pp. 767–776, 2010.

[217] R. Otazo, E. Candes, and D. Sodickson, "Low-rank and sparse matrix decomposition for accelerated DCE-MRI with background and contrast separation," in *Proc. ISMRM Workshop Data Sampling Image Reconstruct.*, Sedona, AZ, USA, 2013, p. 7.

[218] R. Otazo, E. Candes, and D. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components," *Magn. Reson. Med.*, vol. 73, no. 3, pp. 1125–1136, 2015.

[219] R. Otazo, A. Franco, J. Chen, C. Marmar, and F. Boada, "Low-rank plus sparse $(L + S)$ decomposition for separation of subsampled physiological noise in fMRI," in *Proc. Org. Human Brain Mapping (OHBM)*, 2015, p. 1690.

[220] R. Otazo, A. Franco, A. Yoshimoto, J. Chen, C. Marmar, and F. Boada, "Robust resting state fMRI using robust principal component analysis (RPCA)," in *Proc. Org. Human Brain Mapping (OHBM)*, p. 4114, 2016.

[221] R. Otazo, D. Kim, L. Axel, and D. K. Sodickson, "Combination of compressed sensing and parallel imaging with respiratory motion correction for highly-accelerated first-pass cardiac perfusion MRI," *Magn. Reson. Med.*, vol. 64, no. 4, pp. 1148–1154, 2011.

[222] R. Otazo, T. Koesters, E. Candès, and D. K. Sodickson, "Motion-guided low-rank plus sparse (L+S) reconstruction for free-breathing dynamic MRI," in *Proc. ISMRM*, 2014, p. 742.

[223] R. Paffenroth, K. M. Kay, and L. D. Servi. (Jan. 2018). "Robust PCA for anomaly detection in cyber networks." [Online]. Available: http://export.arxiv.ora/pdf/1801.01571

[224] B. Pairet, L. Jacques, C. Gomez Gonzalez, and O. Absil, "Low rank and group-average sparsity driven convex optimization for direct exoplanets imaging," in *Proc. Int. Traveling Workshop Interact. Between Sparse Models Technol. (iTWIST)*, 2016, pp. 24–26.

[225] P. Pan, J. Feng, L. Chen, and Y. Yang, "Online compressed robust PCA," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, 2017, pp. 1041–1048.

[226] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2233–2246, Nov. 2012.

[227] G. Pope, M. Baumann, C. Studery, and G. Durisi, "Real-time principal component pursuit," in *Proc. Asilomar Conf. Signals Syst. Comput.*, Nov. 2011, pp. 1433–1437.

[228] A. Prativadibhayankaram, H. Luong, T. Le, and A. Kaup, "Compressive online video background-foreground separation using multiple prior information and optical flow," *J. Imag.*, vol. 4, no. 7, p. 90, 2018. [Online]. Available: https://doi.org/10.3390/iimaaina4070090

[229] S. Prativadibhayankaram, H. Luong, T. Le, and A. Kaup, "Compressive online robust principal component analysis with optical flow for video foreground-background separation," in *Proc. 8th Int. Symp. Inf. Commun. Technol.*, 2017, pp. 385–392.

[230] H. Qi, B. Ayhan, C. Kwan, W. Wang, S. Li, and S. Vance, "Identify anomaly component by sparsity and low rank," in *Proc. IEEE Workshop Hyperspectral Image Signal Process. (WHISPERS)*, Jun. 2015, pp. 1–4.

[231] H. Qin, Y. Peng, and X. Li, "Foreground extraction of underwater videos via sparse and low-rank matrix decomposition," in *Proc. Workshop Comput. Vis. Anal. Underwater Imagery (ICPR)*, 2014, pp. 65–72.

[232] C. Qiu and N. Vaswani, "Real-time robust principal components pursuit," in *Proc. Int. Conf. Commun. Control Comput.*, 2010, pp. 65–72.

[233] C. Qiu and N. Vaswani. (2017). "ReProCS: A missing link between recursive robust PCA and recursive sparse recovery in large but correlated noise." [Online]. Available: https://arxiv.ora/pdf/1106.3286.pdf

[234] C. Qiu and N. Vaswani, "Support predicted modified-CS for recursive robust principal components' pursuit," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2011, pp. 668–672.

[235] C. Qiu, X. Wu, and H. Xu, "Recursive projected sparse matrix recovery (ReProCSMR) with application in real-time video layer separation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 1332–1336.

[236] Y. Qu, R. Guo, W. Wang, H. Qi, B. Ayhan, and S. Vance, "Anomaly detection in hyperspectral images through spectral unmixing and low rank decomposition," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 1855–1858.

[237] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: A systematic survey," *IEEE Trans. Image Process.*, vol. 14, no. 3, pp. 294–307, Mar. 2005.

[238] M. Rahmani and G. Atia, "Randomized subspace learning approach for high dimensional low rank plus sparse matrix decomposition," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, 2015, pp. 1796–1800.

[239] M. Rahmani and G. Atia, "A subspace learning approach for high dimensional matrix decomposition with efficient column/row sampling," *J. Mach. Learn. Res.*, 2016.

[240] M. Rahmani and G. Atia, "High dimensional decomposition of coherent/structured matrices via sequential column/row sampling," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Mar. 2017, pp. 6419–6423.

[241] M. Rahmani and G. Atia, "Robust and scalable column/row sampling from corrupted big data," in *Proc. Int. Workshop RSL-CV Conjunction (ICCV)*, Oct. 2017.

[242] X. Ren and Z. Lin, "Linearized alternating direction method with adaptive penalty and warm starts for fast solving transform invariant low-rank textures," *Int. J. Comput. Vis.*, vol. 104, no. 1, pp. 1–14, 2013.

[243] P. Rodriguez, "Real-time incremental principal component pursuit for video background modeling on the TK1," in *Proc. GPU Techn. Conf.*, Mar. 2015.

[244] P. Rodriguez and B. Wohlberg, "A MATLAB implementation of a fast incremental principal component pursuit algorithm for video background modeling," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 3414–3416.

[245] P. Rodriguez and B. Wohlberg, "Video background modeling under impulse noise," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 1041–1045.

[246] P. Rodriguez and B. Wohlberg, "Incremental principal component pursuit for video background modeling," *J. Math. Imag. Vis.*, vol. 55, no. 1, pp. 1–18, 2015.

[247] P. Rodriguez and B. Wohlberg, "Translational and rotational jitter invariant incremental principal component pursuit for video background modeling," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 537–541.

[248] P. Rodriguez and B. Wohlberg, "Ghosting suppression for incremental principal component pursuit algorithms," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Dec. 2016, pp. 537–541.

[249] G. Ros, J. M. Álvarez, and J. Guerrero, "Motion estimation via robust decomposition with constrained rank," *IEEE Trans. Intell. Veh.*, vol. 1, no. 4, pp. 346–357, Oct. 2014.

[250] J. Rymel, J. Renno, D. Greenhill, J. Orwell, and G. A. Jones, "Adaptive eigen-backgrounds for object detection," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2004, pp. 1847–1850.

[251] D. Sabushimike, S. Na, J. Kim, N. Bui, K. Seo, and G. Kim, "Low-rank matrix recovery approach for clutter rejection in real-time IR-UWB radar-based moving target detection," *Sensors*, vol. 16, no. 9, p. 1409, 2016.

[252] D. Sabushimike, S. Na, J. Kim, N. Bui, K. Seo, and G. Kim, "Low-rank matrix recovery approach for clutter rejection in real-time IR-UWB radar-based moving target detection," *Sensors*, vol. 16, no. 9, p. 1409, 2016.

[253] P. Schmitt *et al.*, "Inversion recovery TrueFISP: Quantification of T1, T2, and spin density," *Magn. Reson. Med.*, vol. 51, no. 4, pp. 661–667, 2004.

[254] F. Seidel, C. Hage, and M. Kleinsteuber, "pROST—A smoothed l_p-norm robust online subspace tracking method for realtime background subtraction in video," in *Mach. Vis. Appl.*, 2013.

[255] M. Shah, J. D. Deng, and B. J. Woodford, "Video background modeling: Recent approaches, issues and our proposed techniques," *Mach. Vis. Appl.*, vol. 25, no. 5, pp. 1105–1119, 2014.

[256] M. Shakeri and H. Zhang, "COROLA: A sequential solution to moving object detection using low-rank approximation," *Comput. Vis. Image Understanding*, vol. 146, pp. 27–39, May 2016.

[257] M. Shakeri and H. Zhang, "Moving object detection in time-lapse or motion trigger image sequences using low-rank and invariant sparse decomposition," in *Proc. ICCV*, Oct. 2017, pp. 5133–5141.

[258] D. Shan and Z. Chao, "Improved l_1-tracker using robust PCA and random projection," *Mach. Vis.*

Appl., vol. 27, no. 4, pp. 577–583, 2016.
[259] Q. Shan, J. Jia, and A. Agarwala, "High-quality motion deblurring from a single image," *ACM Trans. Graph.*, vol. 27, no. 3, p. 73, 2008.
[260] J. Shen, P. Li, and H. Xu, "Online low-rank subspace clustering by basis dictionary pursuit," in *Proc. Int. Conf. Mach. Learn.*, 2016, pp. 622–631.
[261] F. Shi, J. Cheng, L. Wang, P.-T. Yap, and D. Shen, "LRTV: MR image super-resolution with low-rank and total variation regularizations," *IEEE Trans. Med. Imag.*, vol. 34, no. 12, pp. 2459–2466, Dec. 2015.
[262] A. Shimada, D. Arita, and R. Taniguchi, "Dynamic control of adaptive mixture of Gaussians background model," in *Proc. IEEE Int. Conf. Adv. Video Signal Based Surveillance (AVSS)*, Nov. 2006, p. 5.
[263] G. Silva and P. Rodriguez, "Jitter invariant incremental principal component pursuit for video background modeling on the tk1," in *Proc. Asilomar Conf. Signals Syst. Comput. (ACSSC)*, Nov. 2015, pp. 1403–1407.
[264] D. Skocaj and A. Leonardis, "Incremental and robust learning of subspace representations," in *Proc. Image Vis. Comput. (IVC)*, 2006, pp. 1–12.
[265] A. Sobral, C. Baker, T. Bouwmans, and E. Zahzah, "Incremental and multi-feature tensor subspace learning applied for background modeling and subtraction," in *Proc. Int. Conf. Image Anal. Recognit. (ICIAR)*, Oct. 2014, pp. 94–103.
[266] A. Sobral, T. Bouwmans, and E. Zahzah, "Comparison of matrix completion algorithms for background initialization in videos," in *Proc. SBMI Workshop Conjunction (ICIAP)*, Sep. 2015, pp. 510–518.
[267] A. Sobral, T. Bouwmans, and E. Zahzah, "Double-constrained RPCA based on saliency maps for foreground detection in automated maritime surveillance," in *Proc. ISBC Workshop Conjunction (AVSS)*, 2015, pp. 1–6.
[268] A. Sobral, S. Javed, S. Jung, T. Bouwmans, and E. Zahzah, "Online stochastic tensor decomposition for background subtraction in multispectral video sequences," in *Proc. Workshop Robust Subspace Learn. Comput. Vis. (ICCV)*, Dec. 2015, pp. 106–113.
[269] A. Sobral and E. Zahzah, "Matrix and tensor completion algorithms for background model initialization: A comparative evaluation," in *Proc. Pattern Recognition Lett.*, 2016, pp. 22–33.
[270] W. Song, J. Zhu, Y. Li, and C. Chen, "Image alignment by online robust pca via stochastic gradient descent," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 7, pp. 1241–1250, Jul. 2016.
[271] C. Stauffer and E. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 1999, pp. 246–252.
[272] C. Tan, J. Chen, and L. Chau, "Edge-preserving rain removal for light field images based on RPCA," in *Proc. Int. Conf. Digit. Signal Process. (DSP)*, 2017, pp. 1–5.
[273] H. Tan, B. Cheng, J. Feng, G. Feng, W. Wang, and Y. Zhang, "Low-n-rank tensor recovery based on multi-linear augmented Lagrange multiplier method," *Neurocomputing*, vol. 119, pp. 144–152, Jan. 2013.
[274] H. Tan, B. Cheng, J. Feng, L. Liu, and W. Wang, "Mixture augmented Lagrange multiplier method for tensor recovery and its applications," *Green Intell. Transp. Syst.*, vol. 2013, Mar. 2013, Art. no. 914963.
[275] G. Tang and A. Nehorai, "Robust principal component analysis based on low-rank and block-sparse matrix decomposition," in *Proc. CISS*, 2011, pp. 1–5.
[276] J. Thirion, "Image matching as a diffusion process: An analogy with Maxwell demons," *Med. Image Anal.*, vol. 2, no. 3, pp. 243–260, 1998.
[277] L. Thomaz, E. Jardim, A. da Silva, E. da Silva, S. Netto, and H. Krim, "Anomaly detection in moving-camera video sequences using principal subspace analysis," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 3, no. 3, pp. 1003–1015, Mar. 2017.
[278] D. Tian, H. Mansour, and A. Vetro, "Depth-weighted group-wise principal component analysis for foreground/background separation," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 3230–3234.
[279] F. De La Torre and M. Black, "A robust principal component analysis for computer vision," in *Proc. Robust Princ. Compon. Anal. Comput. Vis.*, 2001, pp. 362–369.
[280] F. De La Torre and M. Black, "A framework for robust subspace learning," *Int. J. Comput. Vis.*, vol. 54, nos. 1–3, pp. 117–142, 2003.
[281] B. Tremoulheac, N. Dikaios, D. Atkinson, and S. Arridge, "Dynamic MR image reconstruction-separation from undersampled (k, t)-space via low-rank plus sparse prior," *IEEE Trans. Medicine Imag.*, vol. 33, no. 8, pp. 1689–1701, Aug. 2014.
[282] J. Tsao and S. Kozerke, "MRI temporal acceleration techniques," *Magn. Reson. Med.*, vol. 36, no. 3, pp. 543–560, 2012.
[283] M. Usman *et al.*, "Motion corrected compressed sensing for free-breathing dynamic cardiac MRI," *Magn. Reson. Med.*, vol. 70, no. 2, pp. 504–516, 2013.
[284] A. Vacavant, T. Chateau, A. Wilhelm, and L. Lequievre, "A benchmark dataset for outdoor foreground/background extraction," in *Proc. Int. Workshop Background Models Challenge (ACCV)*, Nov. 2012, pp. 291–300.
[285] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurth, "Robust subspace learning: Robust PCA, robust subspace tracking, and robust subspace recovery," *IEEE Signal Process. Mag.*, vol. 35, no. 4, pp. 32–55, Jul. 2018, doi: 10.1109/MSP.2018.2826566.
[286] R. Vidal, "Scalable Subspace Clustering," *Invited Talk*, 2016.
[287] M. Wan, G. Gu, W. Qian, K. Ren, and Q. Chen, "Total variation regularization term-based low-rank and sparse matrix representation model for infrared moving target tracking," *MDPI Remote Sens.*, vol. 10, p. 105, Mar. 2018.
[288] T. Wan, Z. Qin, C. Zhu, and R. Liao, "A robust fusion scheme for multifocus images using sparse features," in *Proc. ICASSP*, 2013, pp. 1957–1961.
[289] T. Wan, C. Zhu, and Z. Qin, "Multifocus image fusion based on robust principal component analysis," *Pattern Recognit. Lett.*, vol. 34, pp. 1001–1008, Jul. 2013.
[290] C. Wang, C. Li, and J. Wang, "Two modified augmented Lagrange multiplier algorithms for toeplitz matrix compressive recovery," *Comput. Math. Appl.*, vol. 74, no. 8, pp. 1915–1921, 2017.
[291] H. Wang, Y. Cen, Z. He, R. Zhao, Y. Cen, and F. Zhang, "Robust generalized low-rank decomposition of multimatrices for image recovery," *IEEE Trans. Multimedia*, vol. 19, no. 5, pp. 969–983, May 2017.
[292] J. Wang, M. Wan, X. Hu, and S. Yan, "Image denoising with a unified schatten-p norm and l_q norm regularization," *J. Optim. Theory Appl.*, Apr. 2014.
[293] L. Wang, L. Wang, M. Wen, Q. Zhuo, and W. Wang, "Background subtraction using incremental subspace learning," in *Proc. ICIP*, vol. 5, Sep. 2007, pp. 45–48.
[294] M. Wang, K. Li, F. Wu, Y. Lai, and J. Yang, "3-D motion recovery via low rank matrix analysis," in *Proc. VCIP*, Nov. 2016, pp. 1–4.
[295] N. Wang, T. Yao, J. Wang, and D. Yeung, "A probabilistic approach to robust matrix factorization," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012.
[296] W. Wei, L. Zhang, Y. Zhang, C. Wang, and C. Tian, "Hyperspectral image denoising from an incomplete observation," in *Proc. Int. Conf. Orange Technol. (ICOT)*, 2013, pp. 177–180.
[297] B. Wohlberg, "Endogenous convolutional sparse representations for translation invariant image subspace models," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 2859–2863.
[298] B. Wohlberg, R. Chartrand, and J. Theiler, "Local principal component analysis for nonlinear datasets," in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2012.
[299] H. Woo and H. Park, "Robust asymmetric nonnegative matrix factorization," Ph.D. dissertation, Comput. Appl. Math. Rep., Univ. California, Berkeley, CA, USA, 2014.
[300] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Tirans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
[301] C. Wu *et al.*, "Wavelet speech enhancement based on robust principal component analysis," in *Proc. InterSpeech*, 2017, pp. 439–443.
[302] L. Wu, Y. Wang, Y. Liu, and Y. Wang, "Robust structure from motion with affine camera via low-rank matrix recovery," *China Inf. Sci.*, vol. 56, no. 11, pp. 1–10, Nov. 2015.
[303] M. Wu, Y. Sun, R. Hang, Q. Liu, and G. Liu, "Multi-component group sparse RPCA model for motion object detection under complex dynamic background," *Neurocomputing*, Apr. 2018.
[304] Y. Wu, B. Shen, and H. Ling, "Online robust image alignment via iterative convex optimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 1808–1814.
[305] B. Xin, Y. Kawahara, Y. Wang, L. Hu, and W. Gao, "Efficient generalized fused Lasso and its applications," *ACM Tran. Intell. Syst. Technol.*, vol. 7, no. 4, p. 60, May 2016.
[306] B. Xin, Y. Tian, Y. Wang, and W. Gao, "Background subtraction via generalized fused Lasso foreground modeling," in *Proc. IEEE CVPR*, Jun. 2015, pp. 2163–2169.
[307] F. Xu *et al.*, "Dynamic magnetic resonance imaging via nonconvex low-rank matrix approximation," *IEEE Access*, vol. 5, pp. 1958–1966, 2017.
[308] J. Xu, V. Ithapu, L. Mukherjee, J. Rehg, and V. Singh, "GOSUS: Grassmannian online subspace updates with structured-sparsity," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, Sep. 2013, pp. 3376–3383.
[309] L. Xu and A. Yuille, "Robust principal component analysis by self-organizing rules based on statistical physics approach," *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 131–143, Jan. 1995.
[310] H. Xue, S. Zhang, and D. Cai, "Depth image inpainting: Improving low rank matrix completion with low gradient regularization," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4311–4320, Sep. 2017.
[311] Y. Xue, X. Gu, and X. Cao, "Motion saliency detection using low-rank and sparse decomposition," in *Proc. Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Mar. 2012, pp. 131–143.
[312] J. Yang, X. Sun, X. Ye, and K. Li, "Background extraction from video sequences via motion-assisted matrix completion," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2014, pp. 2437–2441.
[313] L. Yang, Y. Lin, Z. Lin, and H. Zha, "Low rank global geometric consistency for partial-duplicate image search," in *Proc. Int. Conf. Pattern Recognit.*, 2014, pp. 3939–3944.
[314] M. Yang and Z. An, "Video background modeling using low-rank matrix recovery," *J. Nanjing Univ. Posts Telecommun.*, vol. 33, no. 2, Apr. 2013.
[315] X. Yang, X. Gao, D. Tao, X. Li, B. Han, and J. Li, "Shape-constrained sparse and low-rank decomposition for auroral substorm detection," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 1, pp. 32–46, Jan. 2016.
[316] Y. Yang, Y. Feng, and J. Suykens. (2014). "A nonconvex relaxation approach to robust matrix completion." [Online]. Available: https://arxiv.org/abs/1309.0302?context=stat.ML
[317] Q. Yao and J. Kwok, "Colorization by patch-based local low-rank matrix completion," in *Proc. AAAI Conf. Artif. Intell.*, 2015.
[318] M. Yasin, M. Cetin, and A. Khwaja, "SAR imaging of moving targets by subaperture based low-rank and sparse decomposition," in *Proc. SIU*, 2017, pp. 1–4.
[319] 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 Trans. Circuits Syst. Video Technol.*, vol. 25, no. 11, pp. 1721–1734, Nov. 2015.
[320] X. Yi, D. Park, Y. Chen, and C. Caramanis, "Fast algorithms for robust PCA via gradient descent,"

in *Proc. Adv. Neural Inf. Process. Syst.*, 2016, pp. 4152–4160.
[321] H. Yong, D. Meng, W. Zuo, and L. Zhang, "Robust online matrix factorization for dynamic background subtraction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 7, pp. 1726–1740, Jul. 2018.
[322] C. You, C. Li, D. Robinson, and R. Vidal, "Oracle based active set algorithm for scalable elastic net subspace clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2016, pp. 3928–3937.
[323] C. You, D. Robinson, and R. Vidal, "Scalable sparse subspace clustering by orthogonal matching pursuit," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 3918–3927.
[324] J. Yu, "Rank-constrained PCA for intrinsic images decomposition," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 3578–3582.
[325] N. Zarmehi and F. Marvasti, "Recovery of sparse and low rank components of matrices using iterative method with adaptive thresholding," *Digit. Signal Process.*, vol. 73, pp. 145–152, Feb. 2018.
[326] J. Zhan, B. Lois, H. Guo, and N. Vaswani, "Online (and offline) robust PCA: Novel algorithms and performance guarantees," *J. Mach. Learn. Res.*, 2016.
[327] J. Zhan and N. Vaswani, "Robust PCA with partial subspace knowledge," in *Proc. ISIT*, 2014, pp. 2192–2196.
[328] H. Zhang, Z. Lin, and C. Zhang, "Completing low-rank matrices with corrupted samples from few coefficients in general basis," *IEEE Trans. Inf. Theory*, vol. 62, no. 8, pp. 4748–4768, Aug. 2016.
[329] H. Zhang, Z. Lin, C. Zhang, and E. Chang, "Exact recoverability of robust PCA via outlier pursuit with tight recovery bounds," in *Proc. AAAI*, 2015, pp. 3143–3149.
[330] H. Zhang, Z. Lin, C. Zhang, and J. Gao, "Robust latent low rank representation for subspace clustering," *Neurocomputing*, vol. 145, pp. 369–373, Dec. 2014.
[331] H. Zhang, Z. Lin, C. Zhang, and J. Gao, "Relations among some low-rank subspace recovery models," *Neural Comput.*, vol. 27, no. 9, pp. 1915–1950, 2015.
[332] J. Zhang and Y. Zhuang, "Adaptive weight selection for incremental eigen-background modeling," in *Proc. ICME*, Jul. 2007, pp. 851–854.
[333] K. Zhang, M. Wang, and S. Yang, "Learning low-rank decomposition for pan-sharpening with spatial-spectral offsets," *IEEE Trans. Neural Netw. Learn. Syst.*, pp. 1–11, Aug. 2017.
[334] X. Zhang, Z. Lin, F. Sun, and Y. Ma, "Rectification of optical characters as transform invariant low-rank textures," in *Proc. IEEE Int. Conf. Document Anal. Recognit.*, Aug. 2013, pp. 393–397.
[335] X. Zhang, Z. Lin, F. Sun, and Y. Ma, "Transform invariant text extraction," *Vis. Comput.*, vol. 30, no. 4, pp. 401–415, 2014.
[336] Y. Zhang, L. Chen, Z. Zhao, and J. Jia, "Multi-focus image fusion with sparse feature based pulse coupled neural network," *TELKOMNIKA Telecommun., Comput., Electron. Control*, vol. 12, no. 2, pp. 357–366, 2014.
[337] Y. Zhang, L. Chen, Z. Zhao, and J. Jia, "A novel pulse coupled neural network based method for multi focus image fusion," *Int. J. Signal Process. Image Process. Pattern Recognit.*, vol. 7, no. 3, pp. 361–370, 2014.
[338] Y. Zhang, L. Chen, Z. Zhao, J. Jia, and J. Liu, "Multi-focus image fusion based on robust principal component analysis and pulse-coupled neural network," *Optik-Int. J. Light Electron Opt.*, vol. 125, no. 17, pp. 5002–5006, 2014.
[339] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma, "TILT: Transform invariant low-rank textures," *Int. J. Comput. Vis.*, vol. 99, no. 1, pp. 1–24, 2012.
[340] Z. Zhang, H. Lei, and Z. Lv, "Vehicle layover removal in circular SAR images via ROSL," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 12, pp. 2413–2417, Dec. 2015.
[341] Z. Zhang, X. Liang, and Y. Ma, "Unwrapping low-rank textures on generalized cylindrical surfaces," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2011, pp. 1347–1354.
[342] L. Zhao, X. Zhang, Y. Tian, R. Wang, and T. Huang, "A background proportion adaptive Lagrange multiplier selection method for surveillance video on HEVC," in *Proc. ICME*, Jul. 2013.
[343] M. Zhao, L. Jiao, W. Ma, H. Liu, and S. Yang, "Classification and saliency detection by semi-supervised low-rank representation," *Pattern Recognit.*, vol. 51, pp. 281–294, Mar. 2015.
[344] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang, "Robust principal component analysis with complex noise," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2014, pp. 55–63.
[345] X. Zhao, G. An, Y. Cen, H. Wang, and R. Zhao, "Robust generalized low rank approximations of matrices for video denoising," in *Proc. IEEE Int. Conf. Signal Process. (ICSP)*, Nov. 2016, pp. 815–818.
[346] W. Zhen and Y. Min, "New methods for solving the nuclear norm with random matrix and the application in robust principal component analysis," in *Proc. Chin. Control Decision Conf. (CCDC)*, 2017, pp. 1323–1328.
[347] P. Zheng and A. Aravkin. (May 2016). "Fast methods for nonsmooth nonconvex minimization." [Online]. Available: https://arxiv.org/pdf/1605.06900.pdf
[348] Q. Zheng, Y. Wang, and P. Heng, "Online robust image alignment via subspace learning from gradient orientations," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 1753–1762.
[349] Y. Zheng, G. Liu, S. Sugimoto, S. Yan, and M. Okutomi, "Practical low-rank matrix approximation under robust $l1$-norm," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2012, pp. 1410–1417.
[350] T. Zhou and D. Tao, "GoDec: Randomized low-rank and sparse matrix decomposition in noisy case," in *Proc. ICML*, 2011, pp. 33–40.
[351] T. Zhou and D. Tao, "Shifted subspaces tracking on sparse outlier for motion segmentation, which reduces motion detection, tracking and segmentation to one efficient matrix factorization model," in *Proc. IJCAI*, Apr. 2013.
[352] T. Zhou and D. Tao. (Sep. 2013). "Unmixing incoherent structures of big data by randomized or greedy decomposition." [Online]. Available: https://arxiv.org/abs/1309.0302
[353] X. Zhou, C. Yang, and W. Yu, "Moving object detection by detecting contiguous outliers in the low-rank representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 597–610, Mar. 2013.
[354] X. Zhou, C. Yang, H. Zhao, and W. Yu, "Low-rank modeling and its applications in image analysis," *ACM Comput. Surv.*, vol. 47, no. 2, p. 36, 2014.
[355] Z. Zhou, X. Li, J. Wright, E. J. Candès, and Y. Ma, "Stable principal component pursuit," in *Proc. IEEE Int. Symp. Inf. Theory*, Austin, TX, USA, Jun. 2010, pp. 1518–1522.
[356] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, 2011, pp. 479–486.

ABOUT THE AUTHORS

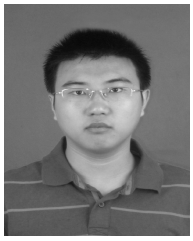
Thierry Bouwmans received the Diploma of HDR for full professor position in 2014.

He has been an Associate Professor at the University of La Rochelle, La Rochelle, France, since 2000. His research interests consist mainly in the detection of moving objects in challenging environments. He has coauthored two books: *Background Modeling and Foreground Detection for Video Surveillance* (Boca Raton, FL, USA: CRC Press, 2014) and *Handbook of Robust Low-Rank and Sparse Matrix Decomposition: Applications in Image Video Processing* (Boca Raton, FL, USA: CRC Press, 2016). His research investigated particularly the use of robust PCA in video surveillance.

Prof. Bouwmans is the main organizer of the Workshop on Robust Subspace Learning and Computer Vision (RSL-CV) hosted at the 2015 and 2017 International Conference on Computer Vision (ICCV). He is a reviewer for international journals including the IEEE Transactions on Image Processing, the IEEE Transactions on Multimedia, the IEEE Transactions on Circuits and Systems for Video Technology, the *International Journal of Computer Vision*, *Machine Vision and Applications*, *Computer Vision and Image Understanding*, *Pattern Recognition*, *Pattern Recognition Letters*, and top-level conferences such as the Conference on Computer Vision and Pattern Recognition (CVPR), the International Conference on Pattern Recognition (ICPR), the International Conference on Image Processing (ICIP), the IEEE International Conference on Advanced Video and Signal-based Surveillance (AVSS), and more.

Sajid Javed received the Bs.c (honors) degree in computer science from the University of Hertfordshire, Hatfield, U.K., in 2010 and the combined M.S. and Ph.D. degrees in computer science from the Virtual Reality Laboratory, Kyungpook National University, Daegu, South Korea, in 2017, under the supervision of Prof. S. K. Jung and co-supervision of Prof. T. Bouwmans.

He has been a Postdoctoral Research Fellow in the Department of Computer Science, University of Warwick, Coventry, U.K., since October 2017. He is working actively on the MiCAHiL project under the supervision of Prof. N. Rajpoot at TIA Lab. He has coauthored about 30 publications, including several articles in journals and international conference publications in the area of RPCA for background/foreground modeling. His current research areas are related to medical image analysis using deep learning algorithms.



Hongyang Zhang received the M.S. degree in computer science from Peking University, Beijing, China, in 2015. He is currently working toward the Ph.D. degree at the Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA, USA.

His research interests include machine learning, theoretical computer science, optimization, and statistics. He also coauthored a book titled *Low-Rank Models in Visual Analysis: Theories, Algorithms, and Applications* (Amsterdam, The Netherlands: Elsevier, 2017).

Zhouchen Lin (Fellow, IEEE) is currently a Professor at the Key Laboratory of Machine Perception, School of Electronics Engineering and Computer Science, Peking University, Beijing, China. He coauthored a book titled *Low-Rank Models in Visual Analysis: Theories, Algorithms, and Applications*(Amsterdam, The Netherlands: Elsevier, 2017). His research interests include computer vision, image processing, machine learning, pattern recognition, and numerical optimization.

Prof. Lin was/is an Area Chair of the 2014/2016/2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), the 2015 International Conference on Computer Vision (ICCV), Advances in Neural Information Processing Systems (NIPS 2015/2018), and the 2019 AAAI Conference on Artificial Intelligence (AAAI), and a Senior Program Committee Member of 2016/2017/2018 AAAI and the 2016/2018 International Joint Conference on Artificial Intelligence (IJCAI). He is an Associate Editor of the IEEE Transactions on Pattern Analysis and Machine Intelligence and the *International Journal of Computer Vision*. He is an International Association for Pattern Recognition (IAPR) Fellow.

Ricardo Otazo is an Associate Professor of Radiology at New York University (NYU) School of Medicine, New York, NY, USA. His research work is focused to the development and clinical translation of MRI and CT techniques using compressed sensing and machine learning. He is one of the early developers of compressed sensing techniques for rapid and high-dimensional MRI, such as k-t SPARSE-SENSE, golden-angle radial sparse parallel (GRASP); and low-rank plus sparse (L+S). He is also working to use compressed sensing and machine learning ideas in the SparseCT method for dose reduction in CT. He is a Project Lead at the Center for Advanced Imaging, Innovation and Research (CAI^{2}R) at NYU, where he leads efforts in acquisition and reconstruction methods for different medical imaging modalities.