On the Applications of Robust PCA in Image and Video Processing

— **Source link**

Thierry Bouwmans, Sajid Javed, Hongyang Zhang, Zhouchen Lin ...+1 more authors

Institutions: University of La Rochelle, University of Warwick, Carnegie Mellon University, Peking University ...+1 more institutions

Topics: Image processing, Video denoising, Video processing, Structure from motion and Robust principal component analysis

Related papers:

- Robust principal component analysis

- Decomposition into Low-rank plus Additive Matrices for Background/Foreground Separation: A Review for a Comparative Evaluation with a Large-Scale Dataset

- Robust Subspace Learning: Robust PCA, Robust Subspace Tracking, and Robust Subspace Recovery

- Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices via Convex Optimization

- A Singular Value Thresholding Algorithm for Matrix Completion

On the Applications of Robust PCA in Image and Video Processing

Thierry Bouwmans, Sajid Javed, Hongyang Zhang, Zhouchen Lin, Ricardo Otazo

On the Applications of Robust PCA in Image and Video Processing

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

Abstract—Robust PCA (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 3D 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 like dynamic Magnetic Resonance Imaging (MRI) for acceleration of data acquisition, background suppression and learning of inter-frame motion fields, and (3) imaging for 3D computer vision with additional depth information like in Structure from Motion (SfM) and 3D 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.

Index Terms—Robust PCA, Image processing, Video processing, 3D Computer Vision, Medical Imaging.

I. INTRODUCTION

PRincipal component analysis was introduced by Karl Pearson in 1901 and was first widely used in statistics. But 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 Principal Component Analysis (PCA) methods was first addressed in statistics by replacing the standard estimation of the covariance matrix with a robust estimator [36],[128] or by using projection pursuit techniques [56],[129]. On the other hand, in neural networks

T. Bouwmans is with Laboratoire MIA, Univ. La Rochelle, 17000 La Rochelle, France. e-mail: tbouwman@univ-lr.fr.

Sajid Javed is with Department of Computer Science, University of Warwick, UK. e-mail: s.javed.1@warwick.ac.uk.

Hongyang Zhang is with Machine Learning Department, Carnegie Mellon University, USA. e-mail: hongyanz@cs.cmu.edu.

Zhouchen Lin is with Key Laboratory of Machine Perception (Ministry of Education), School of Electronics Engineering and Computer Science, Peking University, P.R. China. He is also with Cooperative Medianet Innovation Center, Shanghai Jiao Tong University, P.R. China. email: zlin@pku.edu.cn.

Ricardo Otazo is with Center for Advance Imaging Innovation and Research, New-York University, USA email: ricardo.otazo@nyumc.org

PCA was robustified by designing a neural network that relied on self-organizing rules based on statistical physics [298]. But 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 which appeared on arXiv in 2009, Candès et al. [37] addressed the robustness by decomposition into low-rank plus sparse matrices (also called **L+S** decomposition), and practically 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 **A** is assumed represented as:

$$\mathbf{A} = \mathbf{L} + \mathbf{S} \qquad (1)$$

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

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

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

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

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, $\lambda = \frac{1}{\sqrt{max(m,n)}}$. 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 robust PCA based either on robust estimators and projection pursuit, and the **L+S** decomposition model is that the first approaches assume outlying data points in which entire row or column of the data matrix is corrupted whilst 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 **L+S** decomposition model in computer vision applications like 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 **L** matrix. But, applying directly this original RPCA formulation in applications in which there is also information of interest in the sparse matrix **S** (such as background/foreground separation) results that outliers contains both the information of interest (moving objects that is considered as sparse) and the noise. Thus, most of the time, the stable RPCA formulation [344] is preferred for this kind of computer vision applications. The stable formulation PCP (also called **L+S+E** decomposition) assumes that the observation matrix **A** is represented as follows:

$$\mathbf{A} = \mathbf{L} + \mathbf{S} + \mathbf{E} \qquad (4)$$

where **E** is a noise term (say i.i.d. noise on each entry of the matrix) and $||\mathbf{E}||_F < \delta$ for some $\delta > 0$. To recover **L** and **S**, Zhou et al. [344] proposed to solve the following optimization problem, as a relaxed version to PCP:

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

where $||.||_F$ is the Frobenius norm and $\lambda = \frac{1}{\sqrt{n}}$. 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 **L**, **S** and **E**. A general formulation of the optimization problem suitable for a RPCA formulation applied to a computer vision application can be written as follows:

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

where $T(\cdot)$ and $\Pi(\cdot)$ are linear operators applied on **L** and **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 **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 like in satellite communication [150], seismology [67],[53], speech enhancement [125][290], Synthetic-Aperture Radar (SAR) imaging [127],[307],[240],[159],[155],[329],[23],[24],[95] and direction-of-arrival tracking [167],[61], (2) computer vision applications like in image processing, video processing and 3D computer vision as developed in Bouwmans et al. [28], (3) computer science applications such as the detection traffic anomalies [191][216], and (4) astronomy for auroral substorm detection [304] and exoplanet detection [101],[217]. 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

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

2000s [268],[269], the work of Candès et al. [37] is the main reason why there has been a resurgence of interest in robust PCA and extensions in computer vision since 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 [170], linear time SVD algorithms [303], limited memory SVD algorithms [178], symmetric low-rank product-Gauss-Newton algorithms [179], Block Lanczos with Warm Start (BLWS) algorithms [172], and randomized SVD algorithms [83][335][145]. 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 [274]), and some of them specifically focused on online dynamic robust PCA [182],[274] with performance guarantees [181],[183],[201],[315] and memory efficient algorithms [200]. This line of research allows its application in computer vision such as background/foreground separation which requires incremental and real-time computations. Furthermore, robust PCA often outperforms previous state-of-the-art methods in several computer vision applications [38],[104],[169] with rigorous theoretical analysis [19],[37],[318]. Indeed, as this decomposition is non-parametric 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) combination of 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 **L+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 **L+S** decomposition and **L+S+E** decomposition to enforce its adequacy to the concerned task.

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 [105],[149],[261],[280],[281], texture image denoising [166], hyperspectral image denoising [50],[100],[285], image completion and inpainting [39],[299], image composition for high-dynamic range imaging [21], image decomposition for intrinsic image computation [151],[313] and for structural image decomposition [43], image alignment and rectification [219],[231],[259],[293],[328], image stitching and mosaicking [163], image colorization [306], multi-focus image [277],[278],[325],[326],[327], pansharpening [322], change detection [51], face recognition [185],[289],[320], partial-duplicate image search [302], image saliency detection [147],[160],[161],[222],[228] and image analysis [343],[173].
- **Medical imaging:** RPCA has become a powerful tool to increase the performance of data acquisition [89],[90],[210],[211],[270], image reconstruction [215] and image analysis of brain images [14],[89],[212],[213],[250],[13], cardiac images [48],[49],[90],[210],[211],[270],[215],[296], vessels images [143] and retina images [86]. A key initial application was to reduce the number of measurements in dynamic imaging (space + time), which resulted in increased imaging speed for MRI [90],[210],[211],[270] and radiation dose reduction for CT [89]. 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 datasets, such as functional and diffusion MRI of the brain [14],[212],[213]. In an other way, RPCA can also detect changes in the retina [86] and also aligned image for speckle reduction of retinal OCT images [18].
- **Imaging for 3D computer vision:** This application requires mechanical measurement of the camera positions or manual alignment of partial 3D views of a scene. Thus, RPCA can also be used to reduce outliers and noise in algorithms such as Structure from Motion (SfM) [177],[291],[9],[8], 3D motion recovery [283], and 3D 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],[208],[223], background initialization [255],[258], moving target detection [241], motion saliency detection [47], [300], [332], motion estimation [238], visual object tracking [168][276], action recognition [126], key frame extraction [60], video object segmentation [130],[153],[197],[317],[319], video coding [45],[46],[110],[331], video restoration and denoising [142],[334],[109],[318],[176], video inpainting [142], hyperspectral video processing [96],[42], and video stabilization [68].

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 III-A whilst the specific case of medical imaging is then investigated in Section III-B. We review the 3D computer vision applications in Section III-C. Section IV review 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 $\mathbf{L+S}$ decomposition presents a suitable framework for these different tasks. In addition, the spatial aspects in images is exploited in the $\mathbf{L+S}$ decomposition to enforce its use to the concerned task. In the following sub-sections, we review these different tasks categorized in low-level imaging, medical imaging and 3D computer vision.

A. Low-level imaging

In low-level processing tasks, RPCA via $\mathbf{L+S}$ 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 like 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 like in intrinsic image computation, (3) only the low-rank aspect is of interest like in image alignement, image stitching, image colorization and pan-sharpening, and (4) only the sparse aspect is of interest like in multi-focus 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. [149] 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. [149] calculated the deformation fields $\mathbf{V}_k^j$ from a fixed frame I_k^{samp} to other ones. Define

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

Fig. 1: Image Restoration. From top to bottom: Ground truth image, distorted and blurred image, Sobolev gradient-Laplacian method [184], Centroid method [195] deblurred [248], Two-stage reconstruction method [206] and RPCA algorithm [149] (Images from Lau et al. [149]).

(a)   (b) PSNR: 14.16dB   (c) PSNR: 26.78dB   (d) PSNR: 26.65dB

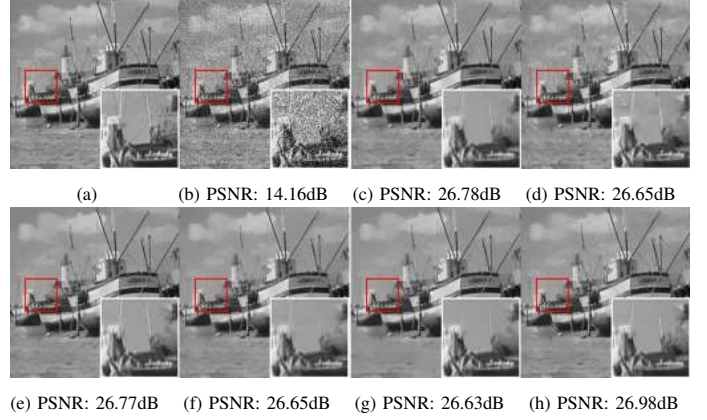(e) PSNR: 26.77dB   (f) PSNR: 26.65dB   (g) PSNR: 26.63dB   (h) PSNR: 26.98dB

Fig. 2: Image Denoising. From top to bottom: (a) Ground-truth image, (b) noiy image, (c) BM3D [59], (d) EPLL [345], (e) LSSC [190], (f) NCSR [64], (g) SAIST [63], and (h) RPCA algorithm called WNNM [105] (Images from Gu et al. [105]).

where vec($\mathbf{V}$) indicates the vectorization of $\mathbf{V}$ and $\mathbf{V}_{k,1} :=$ Re($\mathbf{V}_k$) and $\mathbf{V}_{k,2} :=$ 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:

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

where $\|\mathbf{L}\|_* = \sum_i \sigma_i(\mathbf{L})$ is the nuclear norm, i.e, 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. [149] then warped I_k^{samp} by a post-processing 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-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 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 non-local 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. [105] 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. [105] searched for its non-local similar patches by block matching methods. Then they stacked those non-local similar patches into a matrix $\mathbf{Y}_j$ and decomposed it as $\mathbf{X}_j + \mathbf{N}_j$, where the subscript j indicates the j-th 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 the noise. With this observation, Gu et al. [105] 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 i-th largest singular value of matrix $\mathbf{X}$ and $\mathbf{w} = [\mathbf{w}_1, ..., \mathbf{w}_n]^T$ is the nonnegative weight vector. To set an appropriate weight vector $\mathbf{w}$, Gu et al. [105] 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 preserves much better edges against the following state-of-art methods BM3D [59], EPLL [345], LSSC [190], NCSR [64] and SAIST [63] 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 non-convex 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. [281] 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 non-convex surrogates with few iterations. Experiments show that p, q-PCP achieves the best image recovery performance. Liang [166] 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 UV (ultra-violet) 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. But, 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. [285] addressed hyperspectral data denoising in the RPCA formulation by taking advantage of hyperspectral unmixing and modeling it probabilistically. Let $\mathbf{X}$ be an observed 3D 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 2D 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 $\mathbf{S} \in \mathbb{R}^{n_p \times n_b}$. rank($\mathbf{S}$) is always full with rank($\mathbf{A}$) $\approx$ 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 endmember matrix and abundance matrix, respectively. n_e is the number of endmembers and rank($\mathbf{L}$) is no larger than n_e (i.e rank($\mathbf{L}$) $\leq n_e$). Because n_e is far smaller than n_p and n_b, rank($\mathbf{L}$) $\ll$ rank($\mathbf{A}$). Thus, $\mathbf{L}$ is effectively a low-rank matrix and we have $\mathbf{A} = \mathbf{L} + \mathbf{S}$. Experimental results show that RPCA algorithms outperforms the standard approach based on wavelet as can be seen in Fig. 3 on the Washington DC Mall dataset which contains hyperspectral images of size 1208×307 pixels. Each has 191 spectral channels and a subimage of size $256 \times 256 \times 191$ is cropped from this dataset. 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 like 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 [333], 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
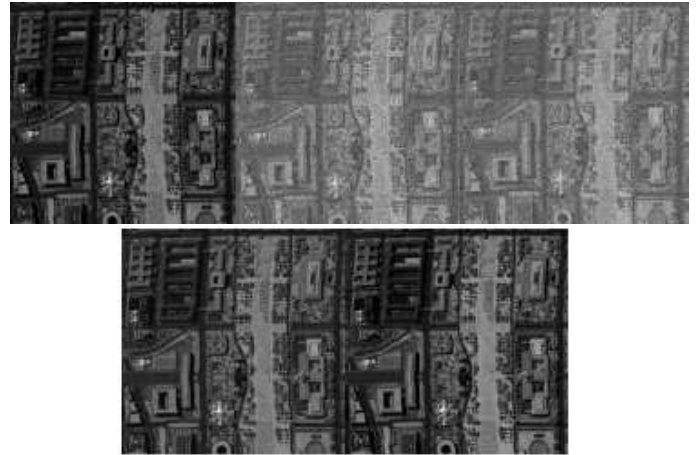

Fig. 3: Hyperspectral image denoising results of band 100 for Washington DC Mall dataset[1]. From top to bottom: Original band, noisy band, wavelet based result, GoDec [339], and RPCA algorithm [285] (Images from Wei et al. [285]).

dynamic objects which typically results in ghosting artifacts.

Bhardwaj and Raman [21] addressed the above-mentioned issues by the robust PCA framework. Specifically, they first modelled 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 $\mathbf{A}$:

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

which is solved by the Alternating Direction Method of Multipliers (ADMM) [171]. 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 imaged 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 multi-exposure images and multi-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 lightings 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

---

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

Fig. 4: Image Composition (HDR). From left to right: Tone mapped [87], and multi-exposed images obtained with the RPCA based approach [21]. The differences are noticeable within the regions delimited by a circle. (Images from Bhardwaj and Raman [21]).

(a) RPCA   (b) RCPCA   (c) RPCA   (d) RCPCA

Fig. 5: Intrinsic Image Computation on two subsets of the Extended Yale B database. From left to right (by group of 6 images=: (a) RPCA results on the subset 18, (b) rank-constrained RPCA results on the subset 18, (c) RPCA results on the subset 22, and (b) rank-constrained RPCA results on the subset 22. Note that the reflectance image remains the same for the rank-constrained RPCA (Images from Yu [313]).

utilizes the fact that the intrinsic image, which reflects the light reflectance properties of a face, is common for the face images taken under different lightings. 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 the sparse matrices 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 [313] proposed the rank-constrained PCA (RCPCA) model, by explicitly enforcing the rank of $\mathbf{L}$ to be one:

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

The above model can also be easily solved by ADMM [171]. Experiments on the *MIT intrinsic image* dataset and the *Yaleface* dataset (see Fig. 6) show that the proposed fixed rank model in Equation (11) enjoys a lower local mean squared error than the prior methods for intrinsic image computation. Similarly, Leow et al. [151] used a different norm, the Frobenius norm, for the matrix $\mathbf{S}$ in Equation (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 pre-processing 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. [219] 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, the mathematical model of RASL is similar to robust PCA but with a characterization of geometric deformation τ:

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

where $\mathbf{A} \circ \tau$ refers to applying frame-wise geometric deformation τ to each frame. For efficient solution, Peng et al. [219] converted the non-convex problem to a computable convex optimization by iteratively linearizing τ locally and updating with the increment $\Delta\tau$ of τ:

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

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 [171], and the solution of Equation 13 converges to solution of Equation 12. An improved optimization method for RASL can be found in [231], where $\Delta\tau_k$ is cancelled first. Such a local linearization algorithm can be viewed as a Gauss-Newton method for minimizing the composition of a non-smooth convex function with a smooth, nonlinear mapping. The convergence behavior of such algorithms has been established in the literature [219]. There are many generalizations of RASL. For example, Wu et al. [293] proposed a method for Online Robust Image Alignment (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 datasets, the heuristic basis updating scheme using thresholding and replacement reduces the robustness of image alignment. Motivated by online robust PCA, Song et al. [259] 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 [293] and SGD [259] 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. [337] employed an online image alignment method via subspace learning from Image Gradient
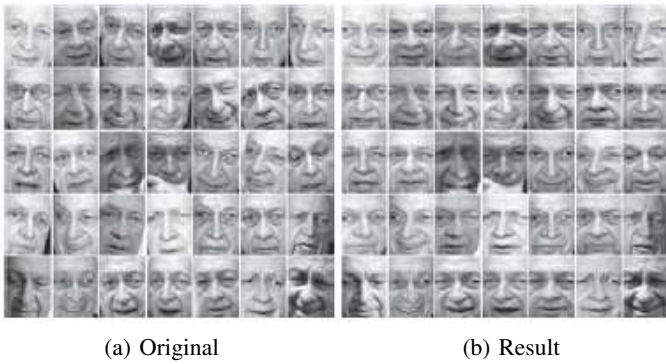
(a) Original   (b) Result

Fig. 6: Image Alignement. From left to right: (a) Original images from the LFW dataset [124], (b) Result obtained by Online Robust Image Alignment (ORIA) [293] (Images from Wu [293]).

(a)

(b)   (c)   (d)

Fig. 7: Image Stiching. First row: (a) Input images from the Shanghai dataset. Second row: (b) Brown and Lowe [33], (c) Gao et al. [91], and (d) SIASM [163] (Images from Li and Monga [163]).

Orientations (IGO). Fig. 6 shows alignment of faces from the Labeled Faces in the Wild (LFW) dataset [124].

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) [328] 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. [330] 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 [323] and text detection [324].

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 [163] 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
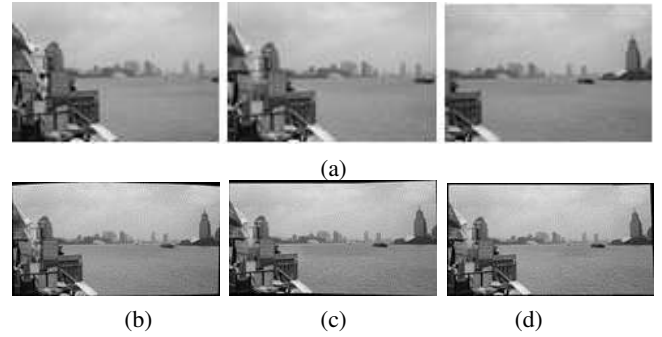
utilizes the sparse components. Their model is

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

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 [171]. With a few postprocessing steps the multiple images can be stitched together. Experiments on the *Shanghai* dataset (See Fig. 7) and the *Windows* dataset show that the proposed method creates much less ghosting artifacts and blur than the prior methods.

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 global image but holds true for local similar patches of the images. With this observation, Yao and Kwok [306] 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 trade-off parameter. For each image patch, denote by $\mathbf{G}$ the matrix that contains the k most similar patches including itself. Yao and Kwok [306] proposed

to minimize the local low-rank approximation problem:

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

(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, $\mathbf{L}$ 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) Multi-Focus Image Fusion: Robust PCA has important applications in multi-focus image fusion as well. Existing imaging devices, such as auto-focus cameras, have limited focus range: only objects around a particular depth are clear while other objects are blurry. Multi-focus 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. [277],[278] formulated the problem of multi-focus 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. [325],[326],[327] proposed to use the Pulse-Coupled Neural Network (PCNN) to record the feature comparison results. The advantage is that the biological characteristics 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 multi-spectral (LRMS) images to generate the synthetic high-resolution multi-spectral (HRMS) images with high spatial and spectral resolutions.

Yang et al. [322] 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 multi-spectral images naturally imply the inherent low-rank property of the matrix formed by stacking HRMS images together. To be more specific, denote by $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_n]$ the matrix by stacking the bands of n LRMS images, each being a column of $\mathbf{A}$, and let $\mathbf{L} = [\mathbf{L}_1, \mathbf{L}_2, ..., \mathbf{L}_n]$ be the matrix of stacking the n HRMS images. Yang et al. [322] decomposed $\mathbf{A}$ as the sum of $\mathbf{L}$ and two offset matrices $\mathbf{S}_1$ and $\mathbf{S}_2$:

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

(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 the HRMS images, both of which should be sparse. The 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:

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

(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 [171], 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 pre-step 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 dataset [193]. For example, Luan et al. [185] used 15 images for an individual that are stacked in the observed matrix A. The first row of Fig. 8 shows 8 images of 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 Gaussian model (RPCA-MOG [196]) and
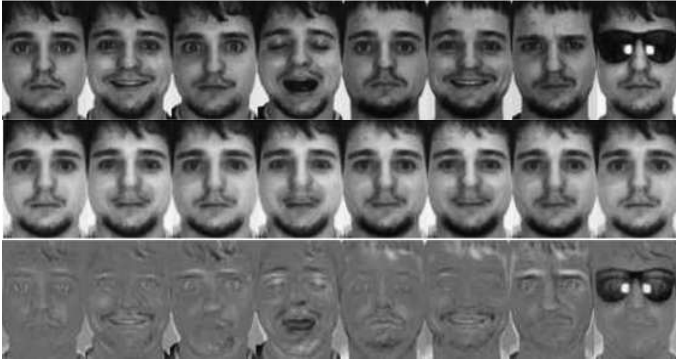
Fig. 8: Face Modeling: Removal of facial occlusion using RPCA. First row: Original face images with facial expression and occlusion. Second row: Low-rank images. Third row: Sparse error images (Images Luan et al. [185]).

Laplacian model (RegL1ALM [338]) 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. [289] 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:

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

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 $\widehat{\mathbf{x}}$.

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. [185] 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. [185] combined *sparsity descriptor*
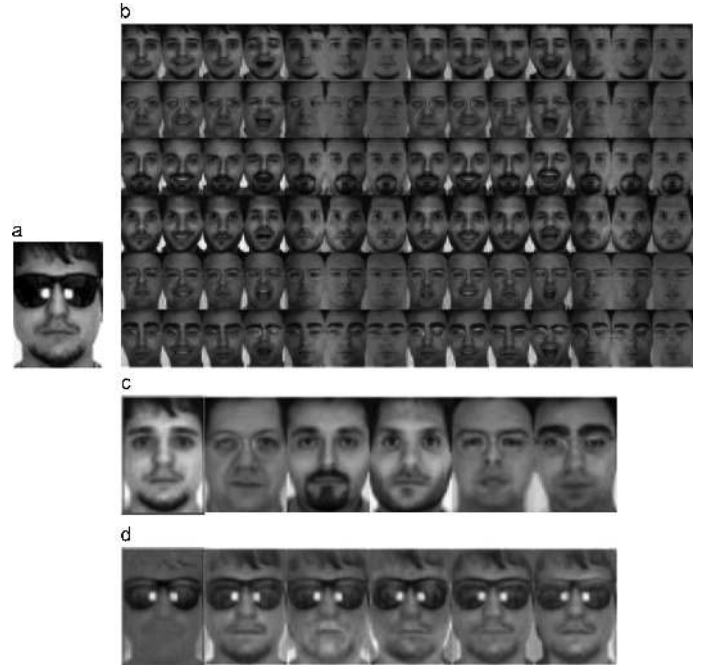

Fig. 9: Face Recognition: Low-rank and sparse error images of a given test image. (a) Test image. (b) Training images of 6 individuals. (c) Low-rank images of the test image under 6 individuals. (d) Corresponding sparse error images (Images from Luan et al. [185]).

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 datasets (Yale Face Database[2], Extended Yale Face Database B[3] and AR face database[4] [193]). As an illustration, Fig. 9 shows the decomposition of a test face image under different subjects using RPCA.

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. [302] introduced the global geometric consistency to detect the

[2]http://vision.ucsd.edu/content/yale-face-database

[3]http://vision.ucsd.edu/content/extended-yale-face-database-b-b

[4]http://www2.ece.ohio-state.edu/ aleix/ARdatabase.html

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 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. [147] proposed a method called Multi-Task 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 $\mathbf{S}$:

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

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. [147] then defined the score function $S(P_i)$ for the i-th 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 Equation (20) is actually called Low-Rank Representation (LRR) [175],[174]. It also has wide applications in image processing, such as motion segmentation [175],[174], image segmentation [52], and image tag completion and refinement [123]. More thorough investigations on LRR can be found in [173].

Instead of working on all the image, Li and Haupt [161],[160] 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 [222],[228]. Thus, anomaly detection is viewed as a matrix decomposition problem with the minimum volume constraint for the multi-modular background and sparsity constraint for the anomaly image pixels.

In summary, the RPCA formulation provides better or similar performances than previous state-of-art methods over these twelve 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
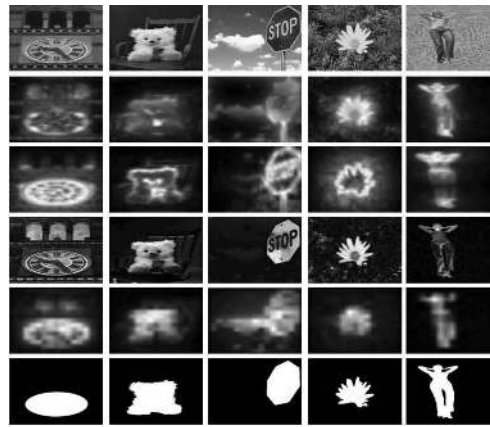

Fig. 10: Saliency Detection: Comparison on images from MSRA dataset [1]. The rows for top to bottom are: original images and saliency maps produced by GBVS [116], CSD [99], FT [1], and the RPCA algorithm (MTSP) [147], respectively. The last row is the ground truth. (Images from Lang et al. [147]).

the sum of a low-rank clean image and a sparse perturbations like in background suppression in accelerated dynamic [211] and in change detection [86]. In the application of joint image reconstruction and registration, only the low-rank aspect is of interest as it concerned image alignment [215].

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],[146], signal relaxation [242], 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 spatio-temporal 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 surprinsingly, 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 [271]. For example, the application of compressed sensing to dynamic MRI [189],[209] 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 **S** the frame-by-frame innovation. **L+S** reconstruction of undersampled dynamic MRI data is performed by solving [211]:

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

here, T is a linear sparsifying transform for **S**, **E** is the encoding operator and **d** is the undersampled k-t data. **L** and **S** are defined as space-time matrices, where each column is a temporal frame, and **d** is defined as a stretched-out single column vector. We assume that **S** has a sparse representation in some transformed domain (e.g. temporal frequency domain, temporal finite differences), hence the idea of minimizing $\left\| T(\mathbf{S}) \right\|_1$ and not $\left\| \mathbf{S} \right\|_1$ itself. For a single-coil acquisition, the encoding operator **E** performs a frame-by-frame undersampled spatial Fourier transform. For acquisition with multiple receiver coils, **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 **L+S** model, which presents improved performance over enforcing a coil-by-coil **L+S** model due to the exploitation of inter-coil correlations, as demonstrated previously for the combination of compressed sensing and parallel imaging based on joint multicoil sparsity [165],[209]. **L+S** reconstruction aims to simultaneously (a) remove aliasing artifacts in the space-time domain (or equivalently to estimate the value of nonsampled points in k-t space) and (b) separate the resulting spatiotemporal low-rank and sparse components. Fig. 11 shows the application of **L+S** reconstruction for 4D contrast-enhanced liver MRI (3D+time), where **L+S** presented improved spatiotemporal resolution with respect to compressed sensing (sparsity alone) and the automatic background suppression in the **S** component improved the visualization of contrast enhancement. **L+S** compares favorably to CS, which suffers from spatiotemporal blurring. Moreover, the **S** component, in which the background has been suppressed, offers improved visualization of contrast-enhancement.

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 spatio-temporal sparsity [144],[214],[272] (including the **L+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, **L+S** reconstruction introduces temporal blurring, leading to non-diagnostic information, or even worse, information that can lead to a false diagnosis. Using ideas from computer vision RPCA techniques such as TILT [328] and RASL [219], the **L+S** model can be modified to include a inter-frame motion operator **W** that describes the deformation between consecutive frames, this is, $\mathbf{M} = \mathbf{W}(\mathbf{L} + \mathbf{S})$. Optical flow [122] can be used to estimate
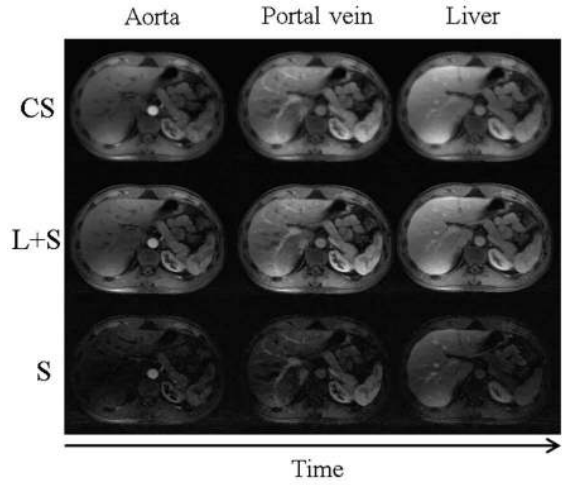

Fig. 11: CS (sparsity-only) and **L+S** reconstruction of 4D dynamic contrast-enhanced abdominal data acquired with golden-angle radial sampling (8 spokes/frame, undersampling factor is 48 and temporal resolution is 0.94 seconds per 3D volume) corresponding to a representative slice and three contrast-enhancement phases (aorta, portal vein, liver).

motion between consecutive frames. For frames $\mathbf{M}_{t-1}$ and $\mathbf{M}t$, the optical flow constraint is:

$$0 = \mathbf{M}_{t-1} - \mathbf{M}_t + \frac{\partial \mathbf{M}_t}{\partial x} w_{x,t} + \frac{\partial \mathbf{M}_t}{\partial y} w_{y,t},$$
$$0 = \mathbf{M}_t + \nabla \mathbf{M}_t \mathbf{W}_t, \tag{22}$$

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 [265], 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 $\mathbf{M}_t$:

$$\mathbf{W}_t = \frac{\nabla \mathbf{M}_t \cdot \mathbf{M}_t}{\left\| \nabla \mathbf{M}_t \right\|^2 + \left\| \mathbf{M}_t \right\|^2}. \tag{23}$$

Motion-guided **L+S** reconstruction [215] aims to estimate **L**, **S** and **W** from undersampled data only. The dependency between **L+S** and **W** makes the optimization problem non-convex and alternating optimization was employed to update **L** and **S** with fixed **W**, and vice-versa, update **W** with fixed $\mathbf{M} + \mathbf{L} + \mathbf{S}$, as follows:

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

$$\mathbf{W}_{k+1} = \frac{\nabla(\mathbf{L}_{k+1} + \mathbf{S}_{k+1}) \cdot \mathbf{m}_{k+1}}{\left\| \nabla(\mathbf{L}_{k+1} + \mathbf{S}_{k+1}) \right\|^2 + \left\| \mathbf{m}_{k+1} \right\|^{2"}}, \tag{25}$$

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
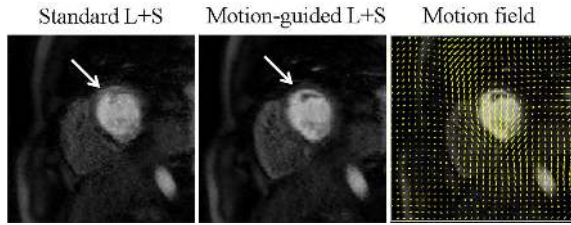
Fig. 12: Standard and motion-guided **L+S** reconstruction of 8-fold accelerated free-breathing cardiac perfusion MRI for a representative contrast phase and slice. The arrows indicate temporal blurring artifacts in standard **L+S** caused by mis-alignment among frames, which are significantly removed by motion-guided **L+S**. In addition, motion-guided **L+S** enables access to motion fields between consecutive frames.

L+S exploits an inherent self-consistency between the **L+S** model and image registration, this is, the rank of **L** will be lowest and the sparsity of **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 **L+S** to free-breathing 8-fold accelerated cardiac perfusion MRI data, where in addition to improved reconstruction, motion-fields that describe inter-frame motion are estimated as 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 [229]. Fu et al. [86] presented a change detection method based on RPCA for retinal fundus images. After alignment and illumination correction, each considered couple of temporal images is expanded into an image serial through linear interpolation between the grey image and the normalized one to progressively decrease the intensity variation between two frames. Then, the linear interpolation images between the grey 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 $i = 1, ..., N$. Vectorizing these frames and concatenating them together, one can obtain an image matrix $\mathbf{A}$ of size $M \times N$. The matrix $\mathbf{A}$ is then decomposed as **L+S**. The decomposition can thus be done by solving the robust PCA problem (10). The sparse component contains the changes between the two images, that are lesion. Extensive experimental results are made on clinical medical 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. 3D Computer Vision

In 3D computer vision, several tasks need to avoid outliers to obtain a reliable 3D reconstruction like in Struture from Motion and 3D motion recovery in which the information of interest is in the low-rank matrix **L**.
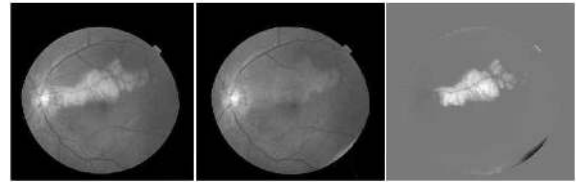

Fig. 13: From left to right: Original image, the low-rank component, and the sparse component (Images from Fu et al. [86]).

1) Structure from Motion: Structure from Motion (SfM) refers to the process of automatically generating a 3D structure of an object by its tracked 2D image frames. Practically, the goal is to recover both 3D structure, namely 3D coordinates of scene points, and motion parameters, namely attitude (rotation) and position of the cameras, starting from image point correspondences. Then, finding the full 3D reconstruction of this object can be posed as a low-rank matrix recovery problem [9],[177],[291]. Suppose that the object is rigid, 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 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 whilst the first three rows of $\mathbf{Q} \in \mathbb{R}^{4 \times M}$ represent the relative 3D positions for each feature point in the reconstructed object. But, when there exists 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 the noise. Then, recovering the full 3D structure of the object can be a low-rank matrix recovery problem in the RPCA formulation. Practically, Liu et al. [177] employed an ℓ_1- filtering approach to solve the decomposition while Wu et al. [291] used the augmented Lagrange multiplier (ALM) method [170]. Experiments on 43344 tracked points over 1001 frames show that this approach provides the best compromise between time and accuracy in comparison with RSL [268] and the original RPCA-ALM [291]. In an other 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 the completion of the missing entries. Then, a modified version of GoDec [339] 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 as compared to previous state-of-the-art methods.

2) 3D Reconstruction: In robotics, the optical sensor begins by capturing points of objects that exist in robots field of vision but the acquired 3D 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 3D point clouds. Let's assume that the captured 3D point cloud $\mathbf{A}$ consists of m points represented as a vector $\mathbf{v} = [\mathbf{x}, \mathbf{y}, \mathbf{z}]$ in a 3D 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 decrease due to the removal of the outliers, so the number of the remaining vertices are m_r where $m_r < m$. Because the acquired 3D 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 3D mesh is obtained which has a exploitable form to be used by other applications or processing tasks.

3) 3D Motion Recovery: Skeleton tracking is a useful and popular application of Microsoft Kinect but it cannot provide accurate reconstructions for complex motions like 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. [283] developed a 3D 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 3D 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 (21 in [283], 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 [283] 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 like 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 like in key frame extraction and UHD super resolution video, and (3) only the low-rank aspect is if interest like 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 [244],[251],[260], fuzzy modeling [15],[16],[17],[27] and conventional subspace learning model either reconstructive [65],[66],[162],[239],[207],[253],[282],[321] and discriminative [84],[85],[192]. However, RPCA methods immediately provided a very promising solution towards 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 [69],[72],[73],[74],[112],[113],[114],[115],[264],[297].

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 dataset [273] 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 rank($\mathbf{L}$) influences the number of modes of the background that can be represented by rank($\mathbf{L}$): if rank($\mathbf{L}$) is too high, the model will incorporate the moving objects into its representation whereas if rank($\mathbf{L}$) is too low, the model tends to be uni-modal and then the multi-modality 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 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. [344].

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 dataset [102] and the extension of background/foreground separation for moving cameras in the following sub-sections. More details can be found in Bouwmans et al. [31],[32], and Namrata et al. [274].
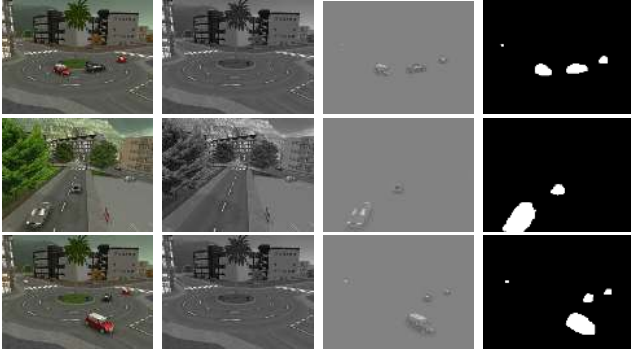
Fig. 14: Background/foreground separation. From left to right: Original images, low-rank matrix **L** (background), sparse matrix **S** (foreground), and foreground mask (Sequences from BMC 2012 dataset [273]).

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]:

$$\min_{\mathbf{L},\mathbf{S},\mathbf{E}} \underbrace{||T(\mathbf{L})||_{norm_1}^{p_1} + \lambda_1||\Pi(\mathbf{S})||_{norm_2}^{p_2} + \lambda_2||\mathbf{E}||_{norm_3}^{p_3}}_{\text{Decomposition}}$$
$$+ \underbrace{\lambda_3||\mathbf{L}||_{2,1} + \delta_1||grad(\mathbf{S})||_1 + \delta_2 TV(\mathbf{S}) + \delta_3\Omega(\mathbf{S})}_{\text{Application}},$$
$$\text{s.t.} \quad \mathbf{A} = \mathbf{L} + \mathbf{S} + \mathbf{E}, \quad \text{or} \quad \mathbf{A} = \mathbf{W} \circ (\mathbf{L} + \mathbf{S} + \mathbf{E}), \quad \text{or}$$
$$\mathbf{A} \circ \tau = \mathbf{L} + \mathbf{S} + \mathbf{E},$$
$$(26)$$

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 **L**, **S**, and **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 [62]. Other loss functions can be used such as ℓ_σ-loss function [305], Least Squares (LS) loss function ($|| \cdot ||_F^2$), Huber loss function [7], M-estimator based loss functions [121], and the generalized fused Lasso loss function [294],[295]. $norm_2$ is usually taken to force spatial homogeneous fitting in the matrix **S**, that is for example the norm $\ell_{2,1}$ with $p_2 = 1$ [69],[72],[73],[74],[112],[113],[115],[114],[264]. 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 **L** like in incPCP-TI [236], [252] to tackle translational and rotational camera jitter.
- The function $\Pi(\cdot)$ is a linear operator processed on **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 are set to zero. $\Pi(\cdot)$

can be computed with optical flow [208] and with salient motion detection [256].

- The term $\lambda_3||\mathbf{L}||_{2,1}$ ensures that the recovered **L** has exact zero columns corresponding to the outliers.
- $||grad(\mathbf{S})||_1$, $TV(\mathbf{S})$, and $\Omega(\mathbf{S})$ are the gradient [113],[114],[115],[287], the total variation [41],[111],[113],[114],[288] and the static/dynamic tree structured sparsity norm [70],[71],[74],[180],[267] applied on the matrix **S** to enforce the spatial and/or temporal constraints, respectively.
- A weighting matrix **W** [256],[301],[308] or a transformation τ [68],[69],[70],[71],[72],[73],[74],[119],[120],[219] can also be used as a constraint in (26) to enforce the recovery of the background that appears at 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],[158],[164],[314],[336] were developed to make the iterations as few as possible as well as SVD algorithms [83],[145],[335] were designed to make the iterations as efficient as possible, batch algorithms can not reach the requirement of real-time computation for background/foreground separation. Thus, to update the model when a new data arrives, several online/incremental algorithms can be found and they can be classified in the following categories [274]: (1) dynamic RPCA algorithms such as the Recursive Projected Compressive Sensing (ReProCS) algorithm and its variants [107],[107],[108],[106][224],[226] provided with performance guarantees, (2) incremental PCP algorithms such as incPCP and its variants [232],[233],[234],[235],[236],[252], (3) online deomposition algorithms [140], (4) subspace tracking algorithms such as the Grassmannian Robust Adaptive Subspace Tracking Algorithm (GRASTA) [117], the ℓ_p-norm Robust Online Subspace Tracking algorithm (pROST) [243], the Grassmannian Online Subspace Updates with Structured-Sparsity algorithm (GOSUS) [297], Fast Adaptive Robust Subspace Tracking algorithm (FARST) [3], and (5) life-long 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 [188],[218],[279],[221],[187],[186], sub-matrices computation [220] and GPU implementations [6],[232].

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 [257], by camera trap [97],[98],[246] and by aerial camera [79],[80],[81],[82]. Furthermore, dedicated methods also exist for infrared cameras [241] and RGB-D cameras [267],[132]. 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. [131] used a 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. [136] 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 $\mathbf{A}$ [68],[70],[72],[74],[120],[219],[259] or to the low-rank matrix $\mathbf{L}$ [236],[252], (2) compensation based methods in which the motion due to the camera is compensated in the pre-processing step like in Tian et al. [267] and Javed et al. [136], and (3) endogenous convolution based methods in which convolutional sparse representations model the effects of non-linear transformations such as translation and rotation, thereby simplifying or eliminating the alignment pre-processing task [286].

- **Illumination changes:** To be robust to illumination changes, Javed et al. [136] incorporated spectral graph regularization in the RPCA framework while Newson et al. [204] used a weighted cluster graph. In the case of time-lapse videos and low-frame rate videos, Shakeri and Zhang [246] 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 [340],[341] 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. [139],[141] used Markov Random Field (MRF) in OR-PCA. In RPCA based on Salient Motion Detection (SMD-RPCA), Chen et al. [47] defined a saliency clue over the sparse matrix $\mathbf{S}$ to filter out the dynamic backgrounds globally. In another work, Wu et al. [292] employed a Multi-Component 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 images iis over-segmented into 80 super-pixels using the Simple Linear Iterative Clustering (SLIC)

[2] to take into accoun the spatial constraint.

- **Intermittent motion of foreground objects:** In MAGRPCA, Javed et al. [136] 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. [47] leveraged the previously detected salient motion to guide the update of the current low-rank prior. Newson et al. [204] used a weighted cluster graph.

- **Ghost suppression:** Rodriguez and Wohlberg [237] 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. [70], [71] 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 [92],[93], 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 [70],[71].

- **Shadows:** Li et al. [154] 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 dataset such as the CD.net 2012/2014 dataset[5] [102][103] or the BMC 2012 dataset[6] [273] 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 dataset [102], and a more detailed analysis can be found in Namrata et al. [274]. This dataset contains almost 31 video sequences which are divided into six different video categories comprising 'Baseline', 'Dynamic Backgrounds' (DB), '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 hundred to thousand number of frames. We compared a total of 25 existing methods comprising 15 batch algorithms and 10 online algorithms. The implementation of all these algorithms is also available in

---

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

the LRSLibrary. These methods are classified into three main categories:

- **Provable methods:** Principal Component Pursuit (PCP) [37], non-convex Alternating Projections based RPCA (AltProj) [203], Near Optimal RMC (NO-RMC)[54], RPCA via Gradient Descent (RPCA-GD) [309], Recursive Projected Compressive Sensing (ReProCS-provable) [201],[315], and Modified-PCP [316].
- **Heuristic methods:** Recursive Projected Compressive Sensing (ReProCS) [106], Grassmannian Robust Adaptive Subspace Tracking Algorithm (GRASTA) [118], Three Term Decomposition (3TD) [208], Two-Pass RPCA (2PRPCA) [92], Go Decomposition (GoDec) [339], Online RPCA (OR-PCA) [139], [134], ℓ_p Robust Online Subspace Tracking (pROST) [243] and Probabilistic Robust Matrix Factorization (PRMF) [284] .
- **Heuristic methods with application specific constraints:** incremental Principal Component Pursuit (incPCP) [235], Motion-assisted Spatiotemporal Clustering of Low-rank (MSCL) [138], Detecting Contiguous Outliers in the LOw-rank Representation (DECOLOR) [342], Low-rank Structured-Sparse Decomposition (LSD) [180], Total Variation RPCA (TVRPCA) [41], Spatiotemporal RPCA (SRPCA) [137], Robust Motion Assisted Matrix Restoration (RMAMR) [301], Generalized Fussed Lasso [294], Grassmannian Online Subspace Updates with Structured-sparsity (GOSUS) [297], Contiguous Outliers Representation via Online Low-rank Approximation (COROLA) [245], and Online Mixture of Gaussians for Matrix Factorization with Total Variation (OMoGMF+TV) [310], respectively.

For qualitative evaluation, visual results were reported using 15 challenging sequences which contained two sequences namely 'highway' and 'office' from 'Baseline' category, three sequences 'canoe', 'boats', and 'overpass' from DB category, two sequences 'traffic' and 'badminton' from 'Camera Jitter' category, three sequences 'winterDriveway', 'sofa', and 'streetLight' from IOM category, three sequences 'backdoor', 'copyMachine' and 'cubicle' from 'Shadows' category, and two sequences 'library' and 'lakeside' from '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 4GB of RAM. For quantitative evaluation, the used metrics come from the CD.net 2012 dataset [102] 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 as compared at 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 $F_1 = \frac{2 \times Recall \times Precision}{Recall + Precision}$. The F-measure characterizes the performance of classification in precision-

recall space. The aim is to maximize F_1-measure closed to one. Table I 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, ReProCS had the best performance with F_1 scores of 74-78%. On average for all datasets, only two of the methods that use extra constraints that are MSCL and GOSUS were better and only by a little by achieving 83 and 81% scores, respectively. For computational time, ReProCS, 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 in computation time in seconds. Practically speaking, these results testify the fact that a 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.

5) Extension to moving cameras: Background-foreground separation is also needed in video taken by moving cameras such as PTZ cameras and handheld cameras [194]. This issue is actually less investigated than the static case. Unlike strategies [71],[119],[219],[236],[252] 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. [88],[198] 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:

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

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 [114]. The low-rank component is obtained via the optimal low-rank matrix estimator (OptShrink [199]) 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. [266] 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.
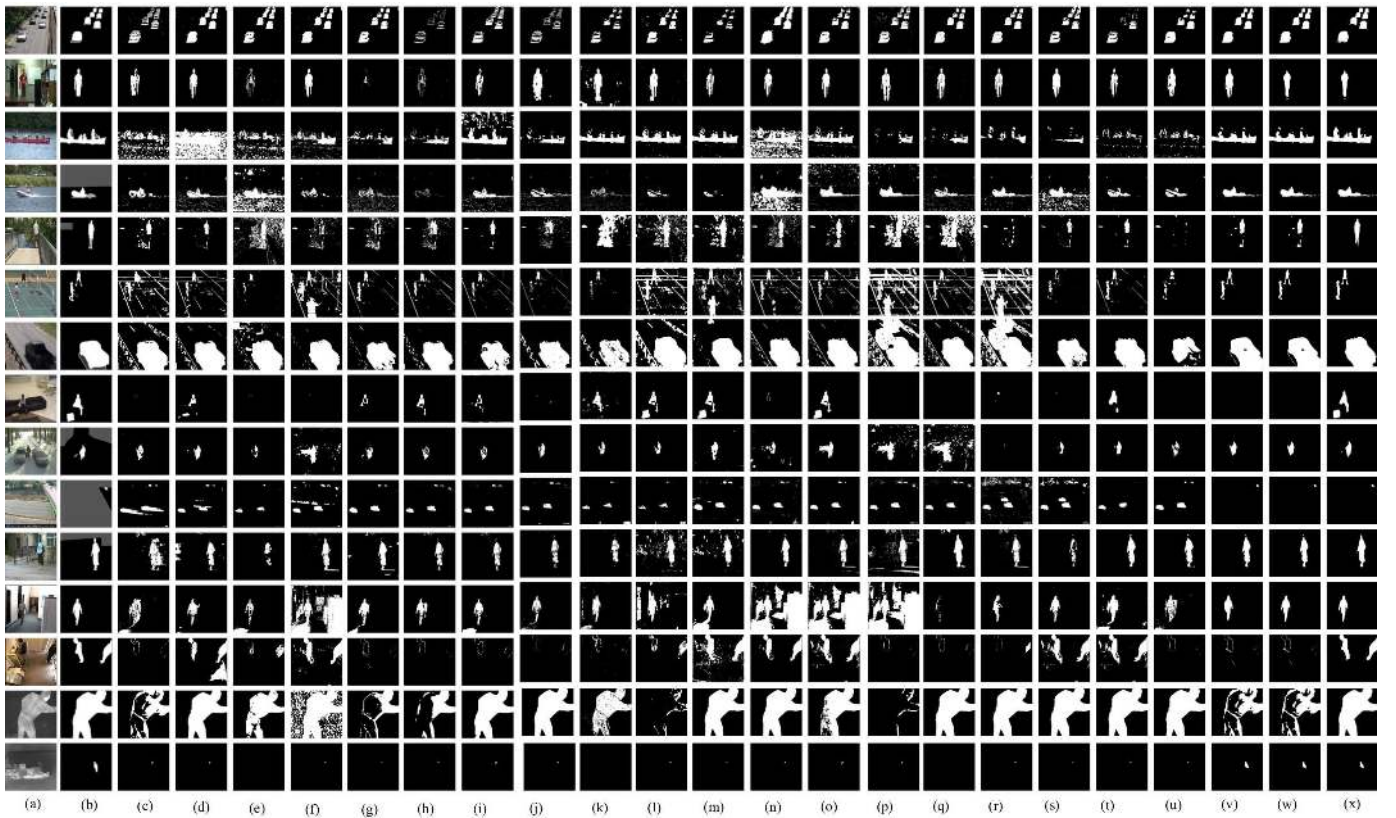
Fig. 15: Comparison of the qualitative results of the 15 input images from ChangeDetection.net dataset. From left to right: (a) set of 15 input images. (b) the ground truth of foreground objects. (c) background subtraction estimated by RPCA via PCP method. (d) GoDec. (e) RPMF. (f) RPCA-GD. (g) 3TD. (h) 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. (x) MSCL. From to bottom: Rows (1)-(2): Sequences 'Highway' and 'office' from category 'Baseline'. Rows (3)-(5): Sequences 'canoe', 'boat', and 'overpass' from category DB. Rows (6)-(7): Sequences 'badminton' and 'traffic' from category 'Camera Jitter'. Rows (8)-(10): Sequences 'sofa', 'winter Driveway', and 'streetLight' from category IOM. Rows (11)-(13): Sequences 'BackDoor', 'cubicle', and 'copyMachine' from category 'Shadow'. Rows (14)-(15): Sequences 'library' and 'lakeside' from category 'Thermal' (Images from Namrata et al. [274]).

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. [266] designed an algorithm based on the Robust Subspace Recovery (RoSuRe [22]) which is able to project a data onto a union of subspaces of lower dimensions. The moving-camera RoSuRe (mcRoSuRe) provides good detection results while at the same time avoiding the need for previous video synchronization. For moving and panning cameras, Chau and Rodriguez [44] 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 [236], 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 like

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. [300] used the low-rank and sparse decomposition on video slices along $X-T$ and $Y-T$ 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 $\mathbf{S}_{cube\mathbf{XT}}$ ($\mathbf{S}_{cube\mathbf{YT}}$) along $\mathbf{X} - \mathbf{Y} - \mathbf{T}$. The initial saliency map cube is obtained by computing $norm(\mathbf{S}_{cube\mathbf{XT}} * \mathbf{S}_{cube\mathbf{YT}})$ where $*$ is the element-wise 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 sub-video. In addition, a spatial information refinement preserve

TABLE I: 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 $8,000$ 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) [203]	**0.78**	*0.71*	0.60	**0.76**	0.69	0.58	*0.68*	2.38
NO-RMC (batch) Fig. 15 (o) [54]	0.71	0.64	0.64	0.66	**0.71**	0.50	0.64	2.85
RPCA-GD (batch) Fig. 15 (f) [309]	0.74	0.62	0.68	*0.75*	0.66	0.49	0.65	2.46
ReProCS-provable (online) [201], [315]	0.77	**0.77**	*0.69*	0.71	**0.74**	**0.70**	**0.73**	0.74
Mod-PCP (online) [316]	*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) [106]	0.80	0.76	0.72	0.75	0.77	**0.69**	*0.74*	**0.61**
GRASTA (online) Fig. 15 (k) [118]	0.66	0.35	0.43	0.52	0.42	0.35	0.45	1.16
3TD (batch) Fig. 15 (g) [208]	*0.88*	0.75	0.72	0.68	*0.78*	0.55	0.72	2.17
2PRPCA (batch) Fig. 15 (u) [92]	**0.92**	*0.79*	*0.81*	*0.80*	0.76	*0.65*	**0.78**	1.63
GoDec (batch) Fig. 15 (d) [339]	0.77	0.58	0.48	0.51	0.62	0.38	0.55	1.56
OR-PCA (online) Fig. 15 (t) [139]	0.86	0.75	0.70	0.74	0.76	0.56	0.72	**0.22**
pROST (online) Fig. 15 (h) [243]	0.79	0.59	0.79	0.70	0.58	0.48	0.65	2.03
PRMF (batch & online) Fig. 15 (e) [284]	**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) [235]	0.81	0.71	0.78	0.74	0.70	*0.75*	0.74	0.41
MSCL (batch) Fig. 15 (x) [138]	0.87	*0.85*	**0.83**	0.82	**0.82**	**0.80**	**0.83**	1.68
DECOLOR (batch) Fig. 15 (v) [342]	**0.92**	0.70	0.68	*0.83*	0.70	0.59	0.73	1.88
LSD (batch) Fig. 15 (p) [180]	**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) [137]	0.82	0.84	0.78	0.77	0.79	**0.80**	0.80	0.59
RMAMR (batch) Fig. 15 (j) [301]	0.89	0.82	0.75	0.73	0.75	0.66	0.76	1.32
GFL (batch) Fig. 15 (w) [294]	0.83	0.74	0.78	0.8 2	0.76	0.59	0.75	2.40
GOSUS (online) Fig. 15 (q) [297]	*0.90*	0.79	*0.82*	**0.84**	*0.80*	0.74	*0.81*	0.89
COROLA (online) Fig. 15 (s) [245]	0.85	**0.86**	*0.82*	0.78	*0.80*	0.71	0.80	*0.39*
OMoG+TV (online) Fig. 15 (r) [310]	0.85	0.76	0.78	0.68	0.70	0.71	0.74	**0.19**

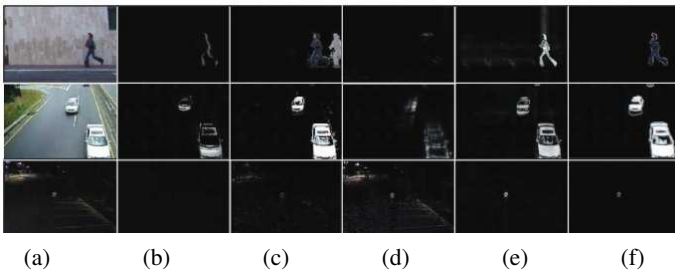Fig. 16: Motion Saliency Detection. From left to right: (a) Original images, (b) consecutive frame difference, (c) MOG [260], (d)Temporal Spectrum Residual (TSR) [57], (e) Raw saliency map [57], and (f) final result obtained by the RPCA algorithm [57] (Images from Xu et al. [300]).

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) [57]as well as background subtraction algorithms like Consecutive Frame Difference (CFD) and Mixture of Gaussians (MoG) [260].

C. Motion Estimation

Motion estimation concerns the process of determining motion vectors for the transformation from one $2D$ image to another which is usually done from adjacent frames in a video sequence. Ros et al. [238] addressed this problem with a modified formulation of RPCA in the special case of camera-pose recovery and visual odometry. Practically, Ros et al. [238] 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 time instant t_i. This formulation is suitable for the stereo visual odometry problem with a rigid 3D transformation [94]. 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. [238] exploited the rank constraints present in rigid 3D 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:

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

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 dataset[7] [94] 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

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

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. But, object detection can be improved by using a tracking feedback. Thus, Lin et al. [168] introduced tracking feedback in the RPCA formulation as follows:

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

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. $\mathbf{P_G}(\mathbf{X})$ is the orthogonal projection of the matrix $\mathbf{X}$ onto the linear space of matrice supported by $\mathbf{G}$, and $\mathbf{P_{G^{\perp}}}(\mathbf{X})$ 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 non-convex, 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 [342] on the I2R dataset [156].

Shan and Chao [247] designed an improved ℓ_1-tracker in a particle filter framework using RPCA and random projection. Practically, three target 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 dataset $\mathbf{A}_0$ 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 dataset 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 a next N^+ tracking results are available in $\mathbf{A}^+$, they are used to update the data matrix $\mathbf{A}_0$. So, $\begin{bmatrix} \mathbf{A}_0 \mathbf{A}^+ \end{bmatrix}$ is cleaned by RPCA as follows:

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

  where $\begin{bmatrix} \mathbf{L_0 L^+} \end{bmatrix}$ denotes the new cleaned matrix, and $\begin{bmatrix} \mathbf{S_0 S^+} \end{bmatrix}$ is the new sparse error matrix. The j-th column of matrix $\mathbf{A}_0$ is then replaced by the i-th 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 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 as compared to several state-of-the-art algorithms. In another work, Elnakeeb and Mitra [78] 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 [78] 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. [126] designed a motion representation method for action recognition by extracting refined low-rank features of RPCA. After extensive experiments, Huang et al. [126] 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. [126] employed the Edges Distribution Histogram (EDH) and 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 EDH or AEDH feature. Experiments on the KTH action dataset[8] [148] 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,

---

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

indexing, and prints from video. Most state-of-the-art methods work directly with the input video dataset, without considering the underlying low-rank structure of the dataset. 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. [60] developed a Key Frame Extraction (KFE) algorithm based on RPCA which decomposes the input video data into a low-rank component which reveals the information across the elements of the dataset, and a set of sparse components each of which containing distinct information about each element. Then, Dang et al. [60] combined the two information types into a single ℓ_1-norm based non-convex 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.

G. Video Object Segmentation

Video segmentation concerns the partition of a video into several semantically consistent spatio-temporal regions. It is a fundamental computer vision problem in several applications like video analytics, summarization and indexing. But, its computational complexity and inherent difficulties such as the large intra-category variations and the large inter-category similarities make this task very challenging. For streaming video segmentation, Li et al. [152],[153] employed a Sub-Optimal 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:

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

where r is the desired rank and $r \ll n$. Then, Li et al. [152],[153] integrated the discriminative replication prior based on internal video statistics into SOLD based on the observation that 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. [45] 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 Go Decomposition (GoDec) algorithm [339] which is a randomized algorithm for low-rank and sparse matrix decomposition in noisy case is 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, up to 3 dB PSNR gain, especially at relatively low bit rate. But, LRSD cannot handle high-resolution or long-time videos due to its high memory requirement. To remedy to these limitations, Chen et al. [46] designed an incremental LRSD (ILRSD) algorithm that can effectively handle large-scale video sequences without much performance loss. Guo et al. [110] 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 [339] is also used for the decomposition. Then, the compressed motion are 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 had a larger proportion. To address this problem, Zhao et al. [331] 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 [42],[96] 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 the matrix $\mathbf{A}$. But, the memory requirement of this problem is typically more challenging than in the color case. Let each frame of a dataset be a $n_r \times n_c \times n_b$ ($128 \times 320 \times 129$) 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, pre-processing 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 (non-reduced) 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, hairs and dust. Ji et al. [142] grouped similar patches in the spatio-temporal 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 spatio-temporal 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 $\mathbf{A} = (\mathbf{p}_1, \mathbf{p}_2, ...\mathbf{p}_3)$. As the 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 two-dimensional samples to one-dimensional vectors and the input matrix should be approximatly low rank matrix. To remedy this limitation, Zhao et al. [334] used an extended RPCA algorithm called Low Rank Approximations of Matrices (GLRAM) to obtain better performance than RPCA. As Ji et al. [142], Guo and Vaswani [109] 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 are first used to initialize the low-rank layer, the sparse layer, and the small residual which is small and bounded. After, ReProCS [106] 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 [58] 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 becomes easier if the denoiser is applied to each layer separately or to only the layer of interest. Fig. 17 shows an examples of videos denoising and enhancement, respectively. For video denoising, we compare RPCA-VBM3D [109] with VBM3D [58]. For video enhancement, we show the comparison between the RPCA algorithm called ReProCS [106] to 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 is a challenging problem because finding important or informative parts of the original video requires to understand 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 [230] 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 the 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 [230] 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 6 human subjects (384 images in total, so $\mathbf{A} \in \mathbb{R}^{32,256 \times 384}$. Experiments [230] show that this sampling algorithm is robust in the presence of corrupted data.

L. UHD Super Resolution 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 super-resolution (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 over-complete dictionary. In order to fully utilize the spatio-temporal information, Ebadi et al. [75] employed a multi-frame video SR approach that is aided by a low-rank plus sparse decomposition of the video sequence. First, Ebadi et al. [75] defined Group of Pictures (GOP) structure and saught a rank-1 low-rank part that recovers the shared spatio-temporal information among the frames in the GOP. Then, the low-rank frames and the sparse frames are super-resolved 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,

| original | noisy | RPCA-VBM3D (PSNR=30dB) | VBM3D (PSNR=25dB) |
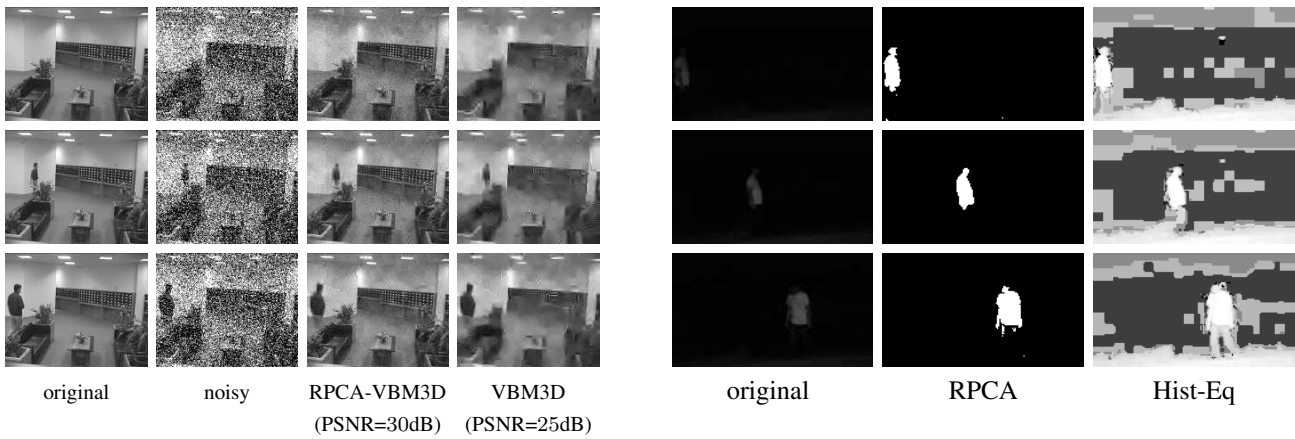
(a) Denoising a very noisy video. PSNR shown in parenthesis

| original | RPCA | Hist-Eq |

(b) Video enhancement: "Seeing" in the dark

Fig. 17: Video denoising and enhancement: (a) Denoising a very noisy video with Gaussian noise of standard deviation $\sigma = 70$ and hence PSNR is 11dB. From left to right: Original videos, noisy videos, RPCA-VM3D [109] results, VBM3D [58] results. Note that VBM3D gives a much more blurred denoised image. Peak Signal to Noise Ratio (PSNR) is noted below each figure too. (b) Video enhancement. From left to right: Original videos, RPCA algorithm (ReProCS [106]) results and histogram equalization (Hist-Eq) results (Images from Namrata et al. [274]).

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. But, dynamic RPCA [182],[274] has received significant attention much more recently, reducing these limitations with performance guarantees [183],[181],[201],[315] and memory efficient algorithms [200], and thus allowing to consider its uses in very challenging applications such as background/foreground separation in videos taken 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 3D 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 datasets 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 [205], it works only for one-dimensional RPCA. On the other hand, ReProCS [225] is a nearly memory optimal solution to dynamic RPCA, but it requires more than one pass through the data.

An open question is how can the RPCA formulation 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 dataset into one of K different low-dimensional subspaces. This can be viewed as a generalization of PCA which tries to represent a given dataset 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 [76][77] and low-rank representation [175],[174], where each sample is represented by other samples and the representation matrix is regularized by either sparsity [76][77], low-rankness [175],[174], or both [76]. Other works also concern scalable subspace clustering [275] which can be solved using algorithms [275],[312],[311] that are provably correct when subspaces are sufficiently separated and data are well distributed within each subspace. A complete review can be found in [173]. Then, given that subspace clusters have been computed for a given dataset, 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 [249].

Another open question is whether one can solve the phaseless RPCA or **L**+**S** problem. Indeed, one can only acquire magnitude-only measurements in applications like ptychography, sub-diffraction 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 under-sampled phaseless measurements.

Finally, this article does not review the literature on the recent works on RPCA for tensor data. Interested readers may refer to the works in [55],[157],[262],[263] for application to background/foreground separation, to the works in [5],[133],[135],[202],[254],[227],[257] for online/incremental tensor algorithms, and to the works of Lin et al. [173] for some recent results. All of 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. *CVPR 2009*, pages 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 Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2274–2281, 2012.

[3] J. Ahn. Fast adaptive robust subspace tracking for online background subtraction. *International Conference on Pattern Recognition, ICPR 2014*, 2014.

[4] Z. An. Video background modeling based on optimization algorithms of robust PCA. *Thesis*, February 2014.

[5] A. Anandkumar, P. Jain, Y. Shi, and U. Niranjan. Tensor vs matrix methods: Robust tensor decomposition under block sparse perturbations. *Preprint*, 2015.

[6] M. Anderson, G. Ballard, J. Demme, and K. Keutzer. Communication-avoiding QR decomposition for GPUs. *IEEE International Parallel and Distributed Processing Symposium, IPDPS 2011*, 2011.

[7] A. Aravkin, S. Becker, V. Cevher, and P. Olsen. A variational approach to stable principal component pursuit. *Conference on Uncertainty in Artificial Intelligence, UAI 2014*, July 2014.

[8] F. Arrigoni, L. Magri, B. Rossi, P. Fragneto, and A. Fusiello. Robust absolute rotation estimation via low-rank and sparse matrix decomposition. *3DV 2014*, 2014.

[9] F. Arrigoni, B. Rossi, and A. Fusiello. Robust and efficient camera motion synchronization via matrix decomposition. *International Conference on Image Processing, ICIAP 2015*, September 2015.

[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. *International Conference on Robotics, ICR 2017*, pages 11–19, 2017.

[11] D. Atkinson, D. Burnstein, and R. Edelman. First-pass cardiac perfusion: evaluation with ultrafast MR imaging. *Radiology*, 174(3):757–62, 1990.

[12] D. Atkinson and R. Edelman. Cineangiography of the heart in a single breath hold with a segmented turboflash sequence. *Radiology*, 178(2):357–60, 1991.

[13] S. Baete, J. Chen, Y. Lin, R X. Wang, R. Otazo, and F. Boada. Low rank plus sparse decomposition of odfs for improved detection of group-level differences and variable correlations in white matter. *Preprint*, 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. *ISMRM 2016*, page 1045, 2016.

[15] F. El Baf, T. Bouwmans, and B. Vachon. A fuzzy approach for background subtraction. *IEEE International Conference on Image Processing, ICIP 2008*, pages 2648–2651, October 2008.

[16] F. El Baf, T. Bouwmans, and B. Vachon. Fuzzy integral for moving object detection. *IEEE International Conference on Fuzzy Systems, FUZZ-IEEE 2008*, pages 1729–1736, June 2008.

[17] F. El Baf, T. Bouwmans, and B. Vachon. Type-2 fuzzy mixture of Gaussians model: Application to background modeling. *International Symposium on Visual Computing, ISVC 2008*, pages 772–781, December 2008.

[18] A. Baghaie, R. De Souza, and Z. Yu. Sparse and low rank decomposition based batch image alignment for speckle reduction of retinal OCT images. *IEEE International Symposium on Biomedical Imaging, ISBI 2015*, pages 226–230, 2015.

[19] M. Balcan, Y. Liang, D. Woodruff, and H. Zhang. Matrix completion and related problems via strong duality. In *Innovations in Theoretical Computer Science*, volume 94, pages 2955–2963, 2018.

[20] M. Balcan and H. Zhang. Noise-tolerant life-long matrix completion via adaptive sampling. In *Advances in Neural Information Processing Systems*, pages 2955–2963, 2016.

[21] A. Bhardwaj and S. Raman. Robust PCA-based solution to image composition using augmented lagrange multiplier (alm). *Visual Computer*, March 2015.

[22] X. Bian and H. Krim. Bi-sparsity pursuit for robust subspace recovery. *IEEE International Conference on Image Processing, ICIP 2015*, September 2015.

[23] F. Biondi. Low rank plus sparse decomposition of synthetic aperture radar data for maritime surveillance. *International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing, CoSeRa 2016*, pages 75–79, 2016.

[24] F. Biondi. Low-Rank Plus Sparse Decomposition and Localized Radon Transform for Ship-Wake Detection in Synthetic Aperture Radar Images. *IEEE Geoscience and Remote Sensing Letters*, 15(1):117–121, January 2018.

[25] T. Bouwmans. Subspace learning for background modeling: A survey. *RPCS*, 2(3):223–234, November 2009.

[26] T. Bouwmans. Recent advanced statistical background modeling for foreground detection: A systematic survey. *RPCS*, 4(3):147–176, November 2011.

[27] T. Bouwmans. Background Subtraction For Visual Surveillance: A Fuzzy Approach. *Chapter 5, Handbook on Soft Computing for Video Surveillance, Taylor and Francis Group, S.K. Pal, A. Petrosino, L. Maddalena*, pages 103–139, March 2012.

[28] T. Bouwmans, N. Aybat, and E. Zahzah. Handbook on robust low-rank and sparse matrix decomposition: Applications in image and video processing. *CRC Press, Taylor and Francis Group*, May 2016.

[29] T. Bouwmans, F. El Baf, and B. Vachon. Background modeling using mixture of Gaussians for foreground detection - a survey. *RPCS*, 1(3):219–237, November 2008.

[30] T. Bouwmans, F. Porikli, B. Horferlin, and A. Vacavant. Handbook on background modeling and foreground detection for video surveillance. *CRC Press, Taylor and Francis Group*, July 2014.

[31] T. Bouwmans, A. Sobral, S. Javed, S. Jung, and E. Zahzah. Decomposition into low-rank plus additive matrices for background/foreground separation: A review for a comparative evaluation with a large-scale dataset. *Computer Science Review,*, November 2016.

[32] T. Bouwmans and E. Zahzah. Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance. *Special Issue on Background Models Challenge, Computer Vision and Image Understanding, CVIU 2014*, 2014.

[33] M. Brown and D. Lowe. Automatic Panoramic Image Stitching using Invariant Features. *International Journal on Computer Vision*, 74(1):59–73, 2007.

[34] A. Buades, B. Coll, and J-M. Morel. A non-local algorithm for image denoising. *IEEE Conference on Computer Vision and Pattern Recognition*, 2:60–65, 2005.

[35] H. Cai, J. Cai, and K. Wei. Accelerated alternating projections for robust principal component analysis. *Preprint*, 2017.

[36] N. Campell. Robust procedures in multivariate analysis i: Robust covariance estimations. *Applied Statistics*, 1(29):231–237, 1980.

[37] E. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *International Journal of ACM*, 58(3), May 2011.

[38] E. Candès and M. Soltanolkotabi. Discussion of latent variable graphical model selection via convex optimization. *Annals of Statistics*, 40(4), 2012.

[39] J. Cao, J. Zhou, X. Liu, W. Wang, P. Tao, and J. Wang. Low-rank image completion with entropy features. *Machine Vision and Applications*, 28(1):129–139, February 2017.

[40] X. Cao, Y. Chen, Q. Zhao, D. Meng, Y. Wang, D. Wang, and Z. Xu. Low-rank matrix factorization under general mixture noise distributions. *International Conference on Computer Vision, ICCV 2015*, December 2015.

[41] X. Cao, L. Yang, and X. Guo. Total variation regularized RPCA for irregularly moving object detection under dynamic background. *IEEE Transactions on Cybernetics*, 46(4):1014–1027, April 2016.

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

[43] X. Chang, S. Liu, P. Zhao, and X. Li. Convergent predictioncorrection-based ADMM for multi-block separable convex programming. *Journal of Computational and Applied Mathematics*, 2017.

[44] G. Chau and P. Rodrguez. Panning and jitter invariant incremental principal component pursuit for video background modeling. *International Workshop RSL-CV 2017 in conjunction with ICCV 2017*, October 2017.

[45] C. Chen, J. Cai, W. Lin, and G. Shi. Surveillance video coding via low-rank and sparse decomposition. *ACM international conference on Multimedia*, pages 713–716, 2012.

[46] C. Chen, J. Cai, W. Lin, and G. Shi. Incremental low-rank and sparse decomposition for compressing videos captured by fixed cameras. *Journal of Visual Communication and Image Representation*, December 2014.

[47] 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 Recognition*, October 2016.

[48] J. Chen, S. Liu, and M. Huang. Low-Rank and Sparse Decomposition Model for Accelerating Dynamic MRI Reconstruction. *Journal of Healthcare Engineering*, 17, 2017.

[49] L. Chen, B. Yang, and X. Wang. Dynamic Magnetic Resonance Imaging via Nonconvex Low-Rank Matrix Approximation. *Mathematical Problems in Engineering*, 2017, 2017.

[50] Y. Chen, Y. Guo, Y. Wang, D. Wang, C. Peng, and G. He. Denoising of Hyperspectral Images using Nonconvex Low Rank Matrix Approximation. *IEEE Transactions on Geoscience and Remote Sensing*, 55(9):5366–5380, September 2017.

[51] Z. Chen and B. Wang. Spectrally-Spatially Regularized Low-Rank and Sparse Decomposition: A Novel Method for Change Detection in Multi-temporal Hyperspectral Images. *MDPI Remote Sensing*, October 2017.

[52] B. Cheng, G. Liu, J. Wang, Z. Huang, and S. Yan. Multi-task low-rank affinity pursuit for image segmentation. *International Conference on Computer Vision*, pages 2439–2446, 2011.

[53] J. Cheng, K. Chen, and M. Sacchi. Robust Principal Component Analysis (RPCA) for seismic data denoising. *GeoConvention 2015*, 2015.

[54] Y. Cherapanamjeri, K. Gupta, and P. Jain. Nearly-optimal robust matrix completion. *Preprint*, June 2016.

[55] 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. *Wiley*, September 2009.

[56] C. Croux and A. Ruiz-Gazen. A fast algorithm for robust principal components based on projection pursuit. *Computational Statistics, COMPSTAT 1996*, 1996.

[57] X. Cui, Q. Liu, and D Metaxas. Temporal spectral residual: Fast motion saliency detection. *ACM MM*, pages 617–620, 2009.

[58] K. Dabov, A. Foi, and K. Egiazarian. Video denoising by sparse 3D transform-domain collaborative filtering. *European Signal Processing Conference*, pages 145–149, October 2007.

[59] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007.

[60] C. Dang, A. Moghadam, and H. Radha. RPCA-KFE: key frame extraction for consumer video based robust principal component analysis. *Preprint*, May 2014.

[61] A. Das. A Bayesian Sparse-Plus-Low-Rank Matrix Decomposition Method for Direction-of-Arrival Tracking. *IEEE Sensors Journal*, 17(15):4894–4902, 2017.

[62] Y. Deng, Q. Dai, R. Liu, and Z. Zhang. Low-rank structure learning via log-sum heuristic recovery. *Preprint*, 2012.

[63] W. Dong, G. Shi, and X. Li. Nonlocal image restoration with bilateral variance estimation: a low-rank approach. *IEEE Transactions on Image Processing*, 22(2):700–711, 2013.

[64] W. Dong, L. Zhang, and G. Shi. Centralized sparse representation for image restoration. *ICCV 2011*, 2011.

[65] Y. Dong and G. DeSouza. Adaptive learning of multi-subspace for foreground detection under illumination changes. *Computer Vision and Image Understanding*, 2010.

[66] Y. Dong, T. Han, and G. DeSouza. Illumination invariant foreground detection using multi-subspace learning. *Journal International of Knowledge-Based and Intelligent Engineering Systems*, 14(1):31–41, 2010.

[67] L. Duarte, E. Nadalin, K. Filho, R. Zanetti, J. Romano, and M. Tygel. Seismic wave separation by means of robust principal component analysis. *European Signal Processing Conference, EUSIPCO 2012*, 2012.

[68] S. Erfanian Ebadi and E. Izquierdo. Approximated RPCA for fast and efficient recovery of corrupted and linearly correlated images and video frames. *IEEE International Conference on Systems Signals and Image Processing, IWSSIP 2015*, September 2015.

[69] S. Erfanian Ebadi and E. Izquierdo. Efficient background subtraction with low-rank and sparse matrix decomposition. *IEEE International Conference on Image Processing, ICIP 2015*, September 2015.

[70] S. Erfanian Ebadi and E. Izquierdo. Foreground segmentation via dynamic tree-structured sparse RPCA. *European Conference on Computer Vision, ECCV 2016*, 2016.

[71] S. Erfanian Ebadi and E. Izquierdo. Foreground segmentation with tree-structured sparse RPCA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[72] S. Erfanian Ebadi, V. Guerra One, and E. Izquierdo. Approximated robust principal component analysis for improved general scene background subtraction. *IEEE Transactions Image Processing*, 2015.

[73] S. Erfanian Ebadi, V. Guerra One, and E. Izquierdo. Efficient background subtraction with low-rank and sparse matrix decomposition. *Workshop on Signal Processing with Adaptive Sparse Structured Representations, SPARS 2015*, July 2015.

[74] S. Erfanian Ebadi, V. Guerra One, and E. Izquierdo. Dynamic tree structured sparse rpca via column subset selection for background modeling and foreground detection. *IEEE International Conference on Image Processing, ICIP 2016*, September 2016.

[75] S. Erfanian Ebadi, V. Ones, and E. Izquierdo. UHD video super-resolution using low-rank and sparse decomposition. *International Workshop on RSL-CV in conjunction with ICCV 2017e*, October 2017.

[76] E. Elhamifar and R. Vidal. Sparse subspace clustering. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2790–2797, 2009.

[77] E. Elhamifar and R. Vidal. Sparse subspace clustering: algorithm, theory, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(11):2765–2781, 2013.

[78] A. Elnakeeb and U. Mitra. Low-rank, sparse and line constrained estimation: Applications to target tracking and convergence. *IEEE ISIT 2017*, pages 2283–2287, 2017.

[79] A. ElTantawy and M. Shehata. Moving object detection from moving platforms using Lagrange multiplier. *IEEE International Conference on Image Processing, ICIP 2015*, 2015.

[80] A. ElTantawy and M. Shehata. UT-MARO: unscented transformation and matrix rank optimization for moving objects detection in aerial imagery. *International Symposium on Visual Computing, ISVC 2015*, pages 275–284, 2015.

[81] A. ElTantawy and M. Shehata. A novel method for segmenting moving objects in aerial imagery using matrix recovery and physical spring model. *ICPR 2016*, pages 3898–3903, 2016.

[82] A. ElTantawy and M. Shehata. MARO: matrix rank optimizationnfor the detection of small-size moving objects from aerial camera platforms. *Signal, Image and Video Processing*, 2017.

[83] N. Erichson, S. Voronin, S. Brunton, and J. Kutz. Randomized Matrix Decompositions using R. *Preprint*, 2016.

[84] D. Farcas and T. Bouwmans. Background modeling via a supervised subspace learning. *International Conference on Image, Video Processing and Computer Vision, IVPCV 2010*, pages 1–7, July 2010.

[85] D. Farcas, C. Marghes, and T. Bouwmans. Background subtraction via incremental maximum margin criterion: A discriminative approach. *Machine Vision and Applications*, 23(6):1083–1101, October 2012.

[86] Y. Fu, W. Wang, and C. Wang. Image change detection method based on RPCA and low-rank decomposition. *Chinese Control Conference, CCC 2016*, pages 9412–9417, 2016.

[87] N. Gallo and N. Gelfand. Artifact-free high dynamic range imaging. *IEEE International Conference on Intelligent Computer Communication and Processing, ICCP 2009*, 2009.

[88] C. Gao, B. Moore, and R. Nadakuditi. Augmented robust PCA for foreground-background separation on noisy, moving camera video. *Preprint*, September 2017.

[89] H. Gao, J. Cai, Z. Shen, and H. Zhao. Robust principal component analysis-based four-dimensional computed tomography. *Physics in Medicine and Biology*, 56(11):3181–98, 2011.

[90] H. Gao, S. Rapacchi, D. Wang, J. Moriarty, C. Meehan, J. Sayre, G. Laub, P. Finn, and P. Hu. Compressed sensing using prior rank, intensity and sparsity model (PRISM): Applications in cardiac cine MRI. *ISMRM 2012*, page 2242, 2012.

[91] J. Gao, S. Kim, and M. Brown. Constructing image panoramas using dual-homography warping. *IEEE International Conference on Computer Vision and Pattern Recognition, CVPR 2011*, pages 49–56, 2011.

[92] Z. Gao, L. Cheong, and M. Shan. Block-sparse RPCA for consistent foreground detection. *ECCV 2012*, 2012.

[93] Z. Gao, L. Cheong, and M. Shan. Block-sparse RPCA for salient motion detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(10):1975–1987, October 2014.

[94] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. *IEEE Conference on Computer Vision, CVPR 2012*, June 2012.

[95] J. Geng, J. Fan, and H. Wang. Weighted Fusion-Based Representation Classifiers for Marine Floating Raft Detection of SAR Images. *IEEE Geoscience and Remote Sensing Letters*, 14(3):444–448, March 2017.

[96] T. Gerhart. Convex optimization techniques and their application in hyperspectral video processing. *Thesis*, December 2013.

[97] 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. *Preprint*, 2017.

[98] J. Giraldo-Zuluaga, A. Gomez, A. Salazar, and A. Diaz-Pulido. Camera-trap images segmentation using multi-layer robust principal component analysis. *Preprint*, 2017.

[99] S. Goferman, L. Zelnikmanor, and A. Tal. Context aware saliency detection. *CVPR 2010*, pages 2376–2383, 2010.

[100] 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. *IEEE International Conference on Image Processing, ICIP 2014*, September 2014.

[101] C. Gomez 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. *Astronomy and Astrophysics*, 2016.

[102] N. Goyette, P. Jodoin, F. Porikli, J. Konrad, and P. Ishwar. changedetection.net: A new change detection benchmark dataset. *IEEE Workshop on Change Detection, CVPR 2012*, June 2012.

[103] N. Goyette, P. Jodoin, F. Porikli, J. Konrad, and P. Ishwar. A novel video dataset for change detection benchmarking. *IEEE Transactions on Image Processing*, 2014.

[104] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, and L. Zhang. Weighted Nuclear Norm Minimization and Its Applications to Low Level Vision. *International Journal of Computer Vision*, 121(2):183–208, January 2017.

[105] S. Gu, L. Zhang, W. Zuo, and X. Feng. Weighted nuclear norm minimization with application to image denoising. *Preprint*, March 2014.

[106] H. Guo, C. Qiu, and N. Vaswani. An online algorithm for separating sparse and low-dimensional signal sequences from their sum. *IEEE Transactions on Signal Processing*, 62(16):4284–4297, 2014.

[107] H. Guo, C. Qiu, and N. Vaswani. Practical ReProCS for separating sparse and low-dimensional signal sequences from their sum - part 1. *ICASSP 2014*, May 2014.

[108] H. Guo, C. Qiu, and N. Vaswani. Practical ReProCS for separating sparse and low-dimensional signal sequences from their sum - part 3. *GlobalSIP 2014*, 2014.

[109] H. Guo and N. Vaswani. Video denoising and enhancement via dynamic video layering. *Preprint*, October 2017.

[110] X. Guo, S. Li, and X. Cao. Motion matters: A novel framework for compressing surveillance videos. *ACM International Conference on Multimedia*, October 2013.

[111] X. Guo, X. Wang, L. Yang, X. Cao, and Y. Ma. Robust foreground detection using smoothness and arbitrariness constraints. *European Conference on Computer Vision, ECCV 2014*, September 2014.

[112] C. Guyon, T. Bouwmans, and E. Zahzah. Foreground detection based on low-rank and block-sparse matrix decomposition. *IEEE International Conference on Image Processing, ICIP 2012*, September 2012.

[113] C. Guyon, T. Bouwmans, and E. Zahzah. Foreground detection via robust low rank matrix decomposition including spatio-temporal constraint. *International Workshop on Background Model Challenges, ACCV 2012*, November 2012.

[114] C. Guyon, T. Bouwmans, and E. Zahzah. Foreground detection via robust low rank matrix factorization including spatial constraint with iterative reweighted regression. *International Conference on Pattern Recognition, ICPR 2012*, November 2012.

[115] C. Guyon, T. Bouwmans, and E. Zahzah. Moving object detection via robust low rank matrix decomposition with IRLS scheme. *International Symposium on Visual Computing, ISVC 2012*, pages 665–674, July 2012.

[116] J. Harel, C. Koch, and P. Perona. Graph-based visual saliency. *Neural Information Processing Systems, NIPS 2006*, pages 545–552, 2006.

[117] J. He, L. Balzano, and J. Luiz. Online robust subspace tracking from partial information. *IT 2011*, September 2011.

[118] J. He, L. Balzano, and A. Szlam. Incremental gradient on the grassmannian for online foreground and background separation in subsampled video. *International on Conference on Computer Vision and Pattern Recognition, CVPR 2012*, June 2012.

[119] J. He, D. Zhang, L. Balzano, and T. Tao. Iterative grassmannian optimization for robust image alignment. *Image and Vision Computing*, June 2013.

[120] J. He, D. Zhang, L. Balzano, and T. Tao. Iterative online subspace learning for robust image alignment. *IEEE Conference on Automatic Face and Gesture Recognition, FG 2013*, 2013.

[121] R. He, T. Tan, and L. Wang. Recovery of corrupted low-rank matrix by implicit regularizers. *IEEE Transaction on Pattern Analysis and Machine Intelligence, PAMI 2013*, September 2013.

[122] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[123] Y. Hou and Z. Lin. Image tag completion and refinement by subspace clustering and matrix completion. *Visual Communications and Image Processing*, pages 1–4, 2015.

[124] G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. *ECCV Workshop on Faces in Real-Life Images*, 2008.

[125] P. Huang, S. Chen, P. Smaragdis, and M. Hasegawa-Johnson. Singing-voice separation from monaural using robust principal component analysis. *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, pages 57–60, 2012.

[126] 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 Journal for Science and Engineering*, 40(2):1427–1441, March 2015.

[127] 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 Transactions on Geoscience and Remote Sensing*, 56(3):1774–1787, March 2018.

[128] P. Huber. Robust statistics. *Wiley, New York*, 1981.

[129] M. Hubert, P. Rousseeuw, and S. Verboven. A fast method for robust principal components with applications to chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 60:101–111, 2002.

[130] W. Jang, C. Lee, and C. Kim. Primary object segmentation in videos via alternate convex optimization of foreground and background distributions. *International Conference on Computer Vision and Pattern Recogntion, CVPR 2016*, June 2016.

[131] S. Javed, T. Bouwmans, and S. Jung. Combining ARF and OR-PCA background subtraction of noisy videos. *International Conference in Image Analysis and Applications, ICIAP 2015*, September 2015.

[132] S. Javed, T. Bouwmans, and S. Jung. Depth extended online RPCA with spatiotemporal constraints for robust background subtraction. *Korea-Japan Workshop on Frontiers of Computer Vision, FCV 2015*, January 2015.

[133] S. Javed, T. Bouwmans, and S. Jung. Stochastic decomposition into low rank and sparse tensor for robust background subtraction. *ICDP 2015*, July 2015.

[134] S. Javed, T. Bouwmans, and S. Jung. Improving OR-PCA via Smoothed Spatially-Consistent Low-rank Modeling for Background Subtraction. *ACM Symposium On Applied Computing, SAC 2017*, 2017.

[135] S. Javed, T. Bouwmans, and S. Jung. SBMI-LTD: Stationary Background Model Initialization based on Low-rank Tensor Decomposition. *ACM Symposium On Applied Computing, SAC 2017*, 2017.

[136] S. Javed, A. Mahmood, T. Bouwmans, and S. Jung. Motion-Aware Graph Regularized RPCA for Background Modeling of Complex Scenes. *Scene Background Modeling Contest, International Conference on Pattern Recognition, ICPR 2016*, December 2016.

[137] S. Javed, A. Mahmood, T. Bouwmans, and S. Jung. Spatiotemporal Low-rank Modeling for Complex Scene Background Initialization. *IEEE Transactions on Circuits and Systems for Video Technology*, December 2016.

[138] S. Javed, A. Mahmood, T. Bouwmans, and S. Jung. Background-Foreground Modeling Based on Spatio-temporal Sparse Subspace Clustering. *IEEE Transactions on Image Processing*, 26(12):5840–5854, December 2017.

[139] S. Javed, S. Oh, T. Bouwmans, and S. Jung. Robust background subtraction to global illumination changes via multiple features based OR-PCA with MRF. *Journal of Electronic Imaging*, 2015.

[140] S. Javed, S. Oh, A. Sobral, T. Bouwmans, and S. Jung. Background subtraction via superpixel-based online matrix decomposition with structured foreground constraints. *Workshop on Robust Subspace Learning and Computer Vision, ICCV 2015*, December 2015.

[141] S. Javed, A. Sobral, S. Oh, T. Bouwmans, and S. Jung. OR-PCA with MRF for robust foreground detection in highly dynamic backgrounds. *Asian conference on computer vision, ACCV 2014*, 2014.

[142] H. Ji, S. Huang, Z. Shen, and Y. Xu. Robust video restoration by joint sparse and low rank matrix approximation. *SIAM Journal on Imaging Sciences*, 4(4):1122–1142, January 2011.

[143] 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 Recognition*, 63:653–666, March 2017.

[144] H. Jung and J. Ye. Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: what we can learn from video compression techniques. *International Journal of Imaging Systems and Technology*, 20:81–98, 2010.

[145] M. Kaloorazi and R. Lamare. Subspace-Orbit Randomized Decomposition for Low-rank Matrix Approximation. *Preprint*, 2018.

[146] F. Korosec, R. Frayne, T. Grist, and C. Mistretta. Time-resolved contrast-enhanced 3D MR angiography. *Magnetic Resonance in Medicine*, 36(3):345–51, 1996.

[147] C. Lang, C. Liu, J. Yu, and S. Yan. Saliency detection by multitask sparsity pursuit. *IEEE Transactions on Image Processing*, 21(3):1327–1338, 2012.

[148] I. Laptev and T. Lindeberg. Velocity adaptation of space-time interest points. *IEEE International Conference on Pattern Recognition*, 2004.

[149] C. Lau, Y. Lai, and L. Lui. Restoration of atmospheric turbulence-distorted images via RPCA and quasi conformal maps. *Preprint*, April 2017.

[150] T. Lee. Analysis of DS-CDMA Receivers using the Robust PCA against Multi-Tone Jamming for Land Mobile Satellite Channels. *Master Thesis, Gwangju Institute of Science and Technology, South Korea*, 2017.

[151] W. Leow, Y. Cheng, L. Zhang, T. Sim, and L. Foo. Background recovery by fixed-rank robust principal component analysis. *International Conference on Computer Analysis of Images and Patterns*, pages 54–61, 2013.

[152] C. Li, L. Lin, W. Zuo, W. Wang, and J. Tang. SOLD: Sub-optimal low-rank decomposition for efficient video segmentation. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2015*, pages 5519–5527, 2015.

[153] C. Li, L. Lin, W. Zuo, W. Wang, and J. Tang. An Approach to Streaming Video Segmentation with Sub-Optimal Low-rank Decomposition. *IEEE Transactions on Image Processing*, 25(5):1947–1960, May 2016.

[154] H. Li, Z. Mia, Y. Li, Y. Xu, and Y. Zhang. Moving object detection via box constrained RPCA. *Journal of PLA University of Science and Technology*, 17(5):403–407, October 2016.

[155] J. Li, Y. Huang, G. Liao, and J. Xu. Moving Target Detection via Efficient ATI-GoDec Approach for Multichannel SAR System. *IEEE Geoscience and Remote Sensing Letters*, 13(9):1320–1324, 2016.

[156] L. Li, W. Huang, I. Gu, and Q. Tian. Statistical modeling of complex backgrounds for foreground object detection. *IEEE Transaction on Image Processing*, pages 1459–1472, 2004.

[157] L. Li, P. Wang, Q. Hu, and S. Cai. Efficient background modeling based on sparse representation and outlier iterative removal. *IEEE Transactions on Circuits and Systems for Video Technology*, December 2014.

[158] Q. Li and Z. Wang. Riemannian submanifold tracking on low-rank algebraic variety. *AAAI Conference on Artificial Intelligence, AAAI 2017*, 2017.

[159] Q. Li, H. Yan, L. Wu, and R. Wang. Robust PCA for Ground Moving Target Indication in Wide-Area Surveillance Radar System. *Journal of the Operations Research Society of China*, 1(1):135–153, March 2013.

[160] X. Li and J. Haupt. Identifying Outliers in Large Matrices via Randomized Adaptive Compressive Sampling. *IEEE Transactions on Signal Processing*, 63(7):1792–1807, April 2015.

[161] X. Li and J. Haupt. Outlier identification via randomized adaptive compressive sampling. *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2015*, 2015.

[162] Y. Li. On incremental and robust subspace learning. *Pattern Recognition*, 37(7):1509–1518, 2004.

[163] Y. Li and V. Monga. SIASM: Sparsity-Based Image Alignment and Stitching Method for Robust Image Mosaicking. *IEEE International Conference on Image Processing, ICIP 2016*, 2016.

[164] Y. Li and W. Yu. A fast implementation of singular value thresholding algorithm using recycling rank revealing randomized singular value decomposition. *Preprint*, April 2017.

[165] D. Liang, B. Liu, J. Wang, and L. Ying. Accelerating SENSE using compressed sensing. *Magnetic Resonance in Medicine*, 62(6):1574–84, 2009.

[166] X. Liang. Removing mixed noise in low rank textures by convex optimization. *Computational Visual Media*, 2(3):267–276, 2016.

[167] B. Lin, J. Liu, M. Xie, and J. Zhu. Direction-of-Arrival Tracking via Low-Rank Plus Sparse Matrix Decomposition. *IEEE Antennas and Wireless Propagation Letters*, 14:1302–1305, 2015.

[168] L. Lin, W. Lin, and S. Huang. Group object detection and tracking by combining RPCA and fractal analysis. *Soft Computing*, pages 1–12, September 2016.

[169] Z. Lin. A review on low-rank models in data analysis. *Big Data and Information Analytics*, 1(2/3):139–161, 2016.

[170] Z. Lin, M. Chen, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. *UIUC Technical Report*, November 2009.

[171] Z. Lin, R. Liu, and Z. Su. Linearized alternating direction method with adaptive penalty for low-rank representation. *Advances in Neural Information Processing Systems*, pages 612–620, 2011.

[172] Z. Lin and S. Wei. A block Lanczos with warm start technique for accelerating nuclear norm minimization algorithms. *Preprint*, 2010.

[173] Z. Lin and H. Zhang. Low-rank models in visual analysis: Theories, algorithms, and applications. *Academic Press*, June 2017.

[174] G. Liu, Z. Lin, S. Yan, J. Sun, and Y. Ma. Robust recovery of subspace structures by low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):171–184, 2013.

[175] G. Liu, Z. Lin, and Y. Yu. Robust subspace segmentation by low-rank representation. *International Conference on Machine Learning*, pages 663–670, 2010.

[176] H. Liu and L. Chau. Deepsea video descattering. *Multimedia Tools and Applications*, pages 1–11, December 2017.

[177] R. Liu, Z. Lin, and Z. Su. Exactly recovering low-rank matrix in linear time via l_1 filter. *Preprint*, August 2011.

[178] X. Liu, Z. Wen, and Y. Zhang. Limited memory block Krylov subspace optimization for computing dominant singular value decomposition. *Preprint*, 2012.

[179] X. Liu, Z. Wen, and Y. Zhang. An efficient Gauss-Newton algorithm for symmetric low-rank product matrix approximations. *Technical Report*, June 2014.

[180] X. Liu, G. Zhao, J. Yao, and C. Qi. Background subtraction based on low-rank model and structured sparse decomposition. *IEEE Transactions on Image Processing*, 2015.

[181] B. Lois and N. Vaswani. A correctness result for online robust PCA. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2015*, 2015.

[182] B. Lois and N. Vaswani. Online Matrix Completion and Online Robust PCA. *IEEE International Symposium on Information Theory, ISIT 2015*, 2015.

[183] B. Lois, N. Vaswani, and C. Qiu. Performance guarantees for undersampled recursive sparse recovery in large but structured noise. *GlobalSIP 2013*, pages 1061–1064, December 2013.

[184] Y. Lou, H. Sung, S. Soatto, and A. L. Bertozzi. Video stabilization of atmospheric turbulence distortions. *Inverse Problems and Imaging*, 7(3):839–861, 2013.

[185] 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 Recognition*, 47(2):495–508, 2014.

[186] H. Luong, N. Deligiannis, S. Forchhammer, and A. Kaup. Compressive online decomposition of dynamic signals via $n-l_1$ minimization with clustered priors. *IEEE Statistical Signal Processing Workshop, SSP 2018*, June 2018.

[187] H. Luong, N. Deligiannis, S. Forchhammer, and A. Kaup. Online decomposition of compressive streaming data using $n-l_1$ cluster-weighted minimization. *Data Compression Conference*, 2018.

[188] H. Luong, N. Deligiannis, J. Seiler, S. Forchhammer, and A. Kaup. Incorporating prior information in compressive online robust principal component analysis. *Preprint*, 2017.

[189] M. Lustig, J. Santos, D. Donoho, and J. Pauly. k-t SPARSE: High frame rate dynamic MRI exploiting spatio-temporal sparsity. *ISMRM 2006*, page 2420, 2006.

[190] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman. From learning models of natural image patches to whole image restoration. *ICCV 2009*, 2009.

[191] M. Mardani, G. Mateos, and G. Giannakis. Recovery of low-rank plus compressed sparse matrices with application to unveiling traffic anomalies. *IEEE Transactions on Information Theory*, pages 5186–5205, 2013.

[192] C. Marghes and T. Bouwmans. Background modeling via incremental maximum margin criterion. *International Workshop on Subspace Methods, ACCV 2010 Workshop Subspace 2010*, November 2010.

[193] A. Martinez. The AR face database. *CVC Technical Report 24*, 4, 1998.

[194] Y. Mehran and T. Bouwmans. New Trends on Moving Object Detection in Video Images Captured by a Moving Camera: A Survey. *Computer Science Review*, 28:157–177, May 2018.

[195] E. Meinhardt-Llopis and M. Micheli. Implementation of the centroid method for the correction of turbulence. *IEEE Transactions on Image Processing*, 4:187–195, 2014.

[196] D. Meng and F. De La Torre. Robust matrix factorization with unknown noise. *International Conference on Computer Vision, ICCV 2013*, December 2013.

[197] S. Minaee and Y. Wang. Masked signal decomposition using subspace representation and its applications. *Preprint*, 2017.

[198] B. Moore, C. Gao, and R. Nadakuditi. Panoramic robust PCA for foreground-background separation on noisy, free-motion camera video. *Preprint*, December 2017.

[199] R. Nadakuditi. OptShrink: an algorithm for improved low-rank signal matrix denoising by optimal, data-driven singular value shrinkage. *IEEE Transactions of Information Theory*, 60(5):3002–3018, May 2013.

[200] P. Narayanamurthy and N. Vaswani. Medrop: Memory-efficient dynamic robust PCA. *Preprint*, 2017.

[201] P. Narayanamurthy and N. Vaswani. New results for provable dynamic robust PCA. *Preprint*, May 2017.

[202] C. Navasca and X. Wang. Adaptive low rank approximation of tensors. *Workshop on Robust Subspace Learning and Computer Vision, ICCV 2015*, 2015.

[203] P. Netrapalli, U. Niranjan, S. Sanghavi, A. Anandkumar, and P. Jain. Non-convex robust PCA. *Preprint*, October 2014.

[204] A. Newson, M. Tepper, and G. Sapiro. Low-rank spatio-temporal video segmentation. *British Machine Vision Conference, BMVC 2015*, 2015.

[205] U. Niranjan and Y. Shi. Streaming robust PCA. *Preprint*, 2016.

[206] T. Pace O. Oreifej, G. Shu and M. Shah. A two-stage reconstruction approach for seeing through waterblurring from a single image. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2011*, pages 1153–1160, 2011.

[207] N. Oliver, B. Rosario, and A. Pentland. A bayesian computer vision system for modeling human interactions. *ICVS 1999*, January 1999.

[208] O. Oreifej, X. Li, and M. Shah. Simultaneous video stabilization and moving object detection in turbulence. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI 2012*, 2012.

[209] E. Otazo, D. Kim, L. Axel, and D. Sodickson. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. *Magnetic Resonance in Medicine*, 64(3):767–76, 2010.

[210] R. Otazo, E. Candes, and D. Sodickson. Low-rank and sparse matrix decomposition for accelerated DCE-MRI with background and contrast separation. *ISMRM Workshop on Data Sampling and Image Reconstruction, Sedona*, page 7, 2013.

[211] R. Otazo, E. Candes, and D. Sodickson. Low-rank and sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components. *Magnetic Resonance in Medicine*, 73(3):1125–1136, 2015.

[212] 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. *Organization for Human Brain Mapping, OHBM 2015*, page 1690, 2015.

[213] R. Otazo, A. Franco, A. Yoshimoto, J. Chen, C. Marmar, and F. Boada. Robust resting state fMRI using robust principal component analysis (rpca). *Organization for Human Brain Mapping, OHBM 2016*, page 4114, 2016.

[214] R. Otazo, D. Kim, L. Axel, and D. Sodickson. Combination of compressed sensing and parallel imaging with respiratory motion correction for highly-accelerated first-pass cardiac perfusion MRI. *ISMRM 2011*, page 63, 2011.

[215] R. Otazo, T. Koesters, E. Candes, and D. Sodickson. Motion-guided low-rank plus sparse (L+S) reconstruction for free-breathing dynamic MRI. *ISMRM 2014*, page 742, 2014.

[216] R. Paffenroth, K. Kay, and L. Servi. Robust PCA for Anomaly Detection in Cyber Networks. *Preprint*, January 2018.

[217] B. Pairet, L. Jacques, C. Gomez Gonzalez, and O. Absil. Low Rank and Group-A verage Sparsity Driven Convex Optimization for Direct Exoplanets Imaging. *International Traveling Workshop on Interactions between Sparse models and Technology, iTWIST 2016*, 2016.

[218] P. Pan, J. Feng, L. Chen, and Y. Yang. Online compressed robust PCA. *International Joint Conference on Neural Networks, IJCNN 2017*, pages 1041–1048, 2017.

[219] 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 Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2233–2246, 2012.

[220] G. Pope, M. Baumann, C. Studery, and G. Durisi. Real-time principal component pursuit. *Asilomar Conference on Signals, Systems, Computation*, November 2011.

[221] S. Prativadibhayankaram, H. Luong, T. Le, and A. Kaup. Compressive online robust principal component analysis with optical flow for video foreground-background separation. *Preprint*, October 2017.

[222] H. Qi, B. Ayhan, C. Kwan, W. Wang, S. Li, and S. Vance. Identify anomaly component by sparsity and low rank. *IEEE Workshop on Hyperspectral Image and Signal Processing, Whispers 2015*, 2015.

[223] H. Qin, Y. Peng, and X. Li. Foreground extraction of underwater videos via sparse and low-rank matrix decomposition. *Workshop on Computer Vision for Analysis of Underwater Imagery , ICPR 2014*, 2014.

[224] C. Qiu and N. Vaswani. Real-time robust principal components pursuit. *International Conference on Communication Control and Computing*, 2010.

[225] C. Qiu and N. Vaswani. ReProCS: a missing link between recursive robust PCA and recursive sparse recovery in large but correlated noise. *Preprint*, 2011.

[226] C. Qiu and N. Vaswani. Support predicted modified-CS for recursive robust principal components' pursuit. *IEEE International Symposium on Information Theory, ISIT 2011*, 2011.

[227] C. Qiu, X. Wu, and H. Xu. Recursive projected sparse matrix recovery (ReProCSMR) with application in real-time video layer separation. *IEEE International Conference on Image Processing, ICIP 2014*, pages 1332–1336, October 2014.

[228] 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. *IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2016*, 2016.

[229] R. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image change detection algorithms: a systematic survey. *IEEE Transactions on Image Processing*, 14(3):294–307, March 2005.

[230] M. Rahmani and G. Atia. Robust and Scalable Column/Row Sampling from Corrupted Big Data. *International Workshop on RSL-CV in conjunction with ICCV 2017*, October 2017.

[231] X. Ren and Z. Lin. Linearized alternating direction method with adaptive penalty and warm starts for fast solving transform invariant low-rank textures. *International journal of computer vision*, 104(1):1–14, 2013.

[232] P. Rodriguez. Real-time incremental principal component pursuit for video background modeling on the TK1. *GPU Technical Conference, GTC 2015*, March 2015.

[233] P. Rodriguez and B. Wohlberg. A Matlab implementation of a fast incremental principal component pursuit algorithm for video background modeling. *IEEE International Conference on Image Processing, ICIP 2014*, October 2014.

[234] P. Rodriguez and B. Wohlberg. Video background modeling under impulse noise. *IEEE International Conference on Image Processing, ICIP 2014*, October 2014.

[235] P. Rodriguez and B. Wohlberg. Incremental principal component pursuit for video background modeling. *Springer Journal of Mathematical Imaging and Vision*, 2015.

[236] P. Rodriguez and B. Wohlberg. Translational and rotational jitter invariant incremental principal component pursuit for video background modeling. *IEEE International Conference on Image Processing, ICIP 2015*, 2015.

[237] P. Rodriguez and B. Wohlberg. Ghosting suppression for incremental principal component pursuit algorithms. *IEEE Global Conference on Signal and Information Processing, GlobalSIP 2016*, December 2016.

[238] G. Ros, J. Alvarez, and J. Guerrero. Motion estimation via robust decomposition with constrained rank. *Preprint*, October 2014.

[239] J. Rymel, J. Renno, D. Greenhill, J. Orwell, and G. Jones. Adaptive eigen-backgrounds for object detection. *ICIP 2004*, pages 1847–1850, October 2004.

[240] 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 2016*, 16(9), 2016.

[241] 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. *MDPI Sensors 2016*, 2016.

[242] P. Schmitt, M. Griswold, P. Jakob, M. Kotas, V. Gulani, M. Flentje, and A. Haase. Inversion recovery TrueFISP: quantification of T(1), T(2), and spin density. *Magnetic Resonance in Medicine*, 51(4):661–67, 2004.

[243] F. Seidel, C. Hage, and M. Kleinsteuber. pROST - a smoothed Lp-norm robust online subspace tracking method for realtime background subtraction in video. *Special Issue on Background Modeling for Foreground Detection in Real-World Dynamic Scenes, Machine Vision and Applications*, 2013.

[244] M. Shah, J. Deng, and B. Woodford. Video background modeling: recent approaches, issues and our proposed techniques. *Machine Vision and Applications*, 25(5):1105–1119, July 2014.

[245] M. Shakeri and H. Zhang. COROLA: a sequential solution to moving object detection using low-rank approximation. *Preprint*, May 2015.

[246] M. Shakeri and H. Zhang. Moving object detection in time-lapse or motion trigger image sequences using low-rank and invariant sparse decomposition. *ICCV 2017*, October 2017.

[247] D. Shan and Z. Chao. Improved l_1-tracker using robust PCA and random projection. *Machine Vision and Applications*, 2016.

[248] Q. Shan, J. Jia, and A. Agarwala. High-quality motion deblurring from a single image. *ACM Transactions on Graphics*, 27(3), 2008.

[249] J. Shen, P. Li, and H. Xu. Online Low-Rank Subspace Clustering by Basis Dictionary Pursuit. *International Conference on Machine Learning*, pages 622–631, 2016.

[250] F. Shi, J. Cheng, L. Wang, P. Yap, and D. Shen. LRTV: MR Image Super-Resolution with Low-Rank and Total Variation Regularizations. *IEEE Transactions on Medical Imaging*, 34(12):2459–2466, 2015.

[251] A. Shimada, D. Arita, and R. Taniguchi. Dynamic control of adaptive mixture of Gaussians background model. *IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2006*, 2006.

[252] G. Silva and P. Rodriguez. Jitter invariant incremental principal component pursuit for video background modeling on the tk1. *Asilomar Conference on Signals, Systems, and Computers, ACSSC 2015*, November 2015.

[253] D. Skocaj and A. Leonardis. Incremental and robust learning of subspace representations. *Image and Vision Computing, IVC 2006*, pages 1–12, 2006.

[254] A. Sobral, C. Baker, T. Bouwmans, and E. Zahzah. Incremental and multi-feature tensor subspace learning applied for background modeling and subtraction. *International Conference on Image Analysis and Recognition, ICIAR 2014*, October 2014.

[255] A. Sobral, T. Bouwmans, and E. Zahzah. Comparison of matrix completion algorithms for background initialization in videos. *SBMI 2015 Workshop in conjunction with ICIAP 2015*, September 2015.

[256] A. Sobral, T. Bouwmans, and E. Zahzah. Double-constrained RPCA based on saliency maps for foreground detection in automated maritime surveillance. *ISBC 2015 Workshop conjunction with AVSS 2015*, 2015.

[257] A. Sobral, S. Javed, S. Jung, T. Bouwmans, and E. Zahzah. Online stochastic tensor decomposition for background subtraction in multi-spectral video sequences. *Workshop on Robust Subspace Learning and Computer Vision, ICCV 2015*, December 2015.

[258] A. Sobral and E. Zahzah. Matrix and tensor completion algorithms for background model initialization: A comparative evaluation. *Special Issue on Scene Background Modeling and Initialization, Pattern Recognition Letters*, 2016.

[259] W. Song, J. Zhu, Y. Li, and C. Chen. Image alignment by online robust pca via stochastic gradient descent. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(7):1241–1250, 2016.

[260] C. Stauffer and E. Grimson. Adaptive background mixture models for real-time tracking. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 1999*, pages 246–252, 1999.

[261] C. Tan, J. Chen, and L. Chau. Edge-preserving rain removal for light field images based on RPCA. *International Conference on Digital Signal Processing, DSP 2017*, pages 1–5, 2017.

[262] 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*, January 2013.

[263] H. Tan, B. Cheng, J. Feng, L. Liu, and W. Wang. Mixture augmented lagrange multiplier method for tensor recovery and its applications. *Special Issue on Green Intelligent Transport System*, 2013.

[264] G. Tang and A. Nehorai. Robust principal component analysis based on low-rank and block-sparse matrix decomposition. *CISS 2011*, 2011.

[265] J. Thirion. Image matching as a diffusion process: an analogy with Maxwell demons. *Medical Image Analysis*, 2(3):243–260, 1998.

[266] 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 Transactions on Circuits and Systems I: Regular Papers*, 2017.

[267] D. Tian, H. Mansour, and A. Vetro. Depth-weighted group-wise principal component analysis for foreground/background separation. *IEEE International Conference on Image Processing, ICIP 2015*, September 2015.

[268] F. De La Torre and M. Black. A robust principal component analysis for computer vision. *International Conference on Computer Vision*, 2001.

[269] F. De La Torre and M. Black. A framework for robust subspace learning. *International Journal on Computer Vision*, pages 117–142, 2003.

[270] 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 Transactions on Medicine Imaging*, 33(8):1689–701, 2014.

[271] J. Tsao and S. Kozerke. MRI temporal acceleration techniques. *Magnetic Resonance in Medicine*, 36(3):543–560, 2012.

[272] M. Usman, D. Atkinson, F. Odille, C. Kolbitsch, G. Vaillant, T. Schaeffter, P. Batchelor, and C. Prieto. Motion corrected compressed sensing for free-breathing dynamic cardiac MRI. *Magnetic Resonance in Medicine*, 70(2):504–16, 2013.

[273] A. Vacavant, T. Chateau, A. Wilhelm, and L. Lequievre. A benchmark dataset for foreground/background extraction. *International Workshop on Background Models Challenge, ACCV 2012*, November 2012.

[274] N. Vaswani, T. Bouwmans, S. Javed, and P. Narayanamurth. Robust PCA and Robust Subspace Tracking. *IEEE Signal Processing Magazine*, July 2018.

[275] R. Vidal. Scalable Subspace Clustering. *Invited Talk*, 2016.

[276] M. Wan, G. Gu, W. Qian, K. Ren, and Q. Chen. Total variation regulation based low-rank and sparse matrix representation model for infrared moving target tracking. *MDPI Remote Sensing*, 2018.

[277] T. Wan, Z. Qin, C. Zhu, and R. Liao. A robust fusion scheme for multifocus images using sparse features. *ICASSP 2013*, pages 1957–1961, 2013.

[278] T. Wan, C. Zhu, and Z. Qin. Multifocus image fusion based on robust principal component analysis. *Pattern Recognition Letters*, 34(9):1001–1008, July 2013.

[279] C. Wang, C. Li, and J. Wang. Two modified augmented Lagrange multiplier algorithms for Toeplitz matrix compressive recovery. *Computers and Mathematics with Applications*, 2017.

[280] H. Wang, Y. Cen, Z. He, R. Zhao, Y. Cen, and F. Zhang. Robust Generalized Low-Rank Decomposition of Multimatrices for Image Recovery. *IEEE Transactions on Multimedia*, 19(5):969–983, May 2017.

[281] J. Wang, M. Wan, X. Hu, and S. Yan. Image denoising with a unified schatten-p norm and l_q norm regularization. *Journal of Optimization Theory and Applications*, April 2014.

[282] L. Wang, L. Wang, M. Wen, Q. Zhuo, and W. Wang. Background subtraction using incremental subspace learning. *ICIP 2007*, 5:45–48, 2007.

[283] M. Wang, K. Li, F. Wu, Y. Lai, and J. Yang. 3-D Motion Recovery via Low Rank Matrix Analysis. *VCIP 2016*, November 2016.

[284] N. Wang, T. Yao, J. Wang, and D. Yeung. A probabilistic approach to robust matrix factorization. *European Conference on Computer Vision, ECCV 2012*, 2012.

[285] W. Wei, L. Zhang, Y. Zhang, C. Wang, and C. Tian. Hyperspectral image denoising from an incomplete observation. *International Conference on Orange Technologies, ICOT 2015*, pages 177–180, 2013.

[286] B. Wohlberg. Endogenous convolutional sparse representations for translation invariant image subspace models. *IEEE International Conference on Image Processing, ICIP 2014*, 2014.

[287] B. Wohlberg, R. Chartrand, and J. Theiler. Local principal component analysis for nonlinear datasets. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, March 2012.

[288] H. Woo and H. Park. Robust asymmetric nonnegative matrix factorization. *Computational and Applied Mathematics Reports, University of California, USA*, 2014.

[289] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009.

[290] C. Wu, H. Hsu, S. Wang, J. Hung, Y. Lai, H. Wang, and Y. Tsao. Wavelet speech enhancement based on robust principal component analysis. *InterSpeech 2017*, 2017.

[291] L. Wu, Y. Wang, Y. Liu, and Y. Wang. Robust structure from motion with affine camera via low-rank matrix recovery. *China Information Sciences*, 56(11):1–10, November 2015.

[292] 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*, 2018.

[293] Y. Wu, B. Shen, and H. Ling. Online robust image alignment via iterative convex optimization. *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1808–1814, 2012.

[294] B. Xin, Y. Kawahara, Y. Wang, L. Hu, and W. Gao. Efficient generalized fused Lasso and its applications. *ACM Transactions on Intelligent Systems and Technology, TIST 2016*, 7(4), May 2016.

[295] B. Xin, Y. Tian, Y. Wang, and W. Gao. Background subtraction via generalized fused Lasso foreground modeling. *IEEE CVPR 2015*, June 2015.

[296] F. Xu, J. Han, Y. Wang, M. Chen, Y. Chen, G. He, and Y. Hu. Dynamic Magnetic Resonance Imaging via Nonconvex Low-Rank Matrix Approximation. *IEEE Access*, 5:1958–1966, 2017.

[297] J. Xu, V. Ithapu, L. Mukherjee, J. Rehg, and V. Singh. GOSUS: grassmannian online subspace updates with structured-sparsity. *International Conference on Computer Vision, ICCV 2013*, September 2013.

[298] L. Xu and A. Yuille. Robust principal component analysis by self-organizing rules based on statistical physics approach. *IEEE Transaction on Neural Networks*, pages 131–143, January 1995.

[299] H. Xue, S. Zhang, and D. Cai. Depth Image Inpainting: Improving Low Rank Matrix Completion with Low Gradient Regularization. *IEEE Transactions on Image Processing*, 26(9):4311–4320, September 2017.

[300] Y. Xue, X. Gu, and X. Cao. Motion saliency detection using low-rank and sparse decomposition. *International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2012*, March 2012.

[301] J. Yang, X. Sun, X. Ye, and K. Li. Background extraction from video sequences via motion-assisted matrix completion. *IEEE International Conference on Image Processing, ICIP 2014*, October 2014.

[302] L. Yang, Y. Lin, Z. Lin, and H. Zha. Low rank global geometric consistency for partial-duplicate image search. *International Conference on Pattern Recognition*, pages 3939–3944, 2014.

[303] M. Yang and Z. An. Video background modeling using low-rank matrix recovery. *Journal of Nanjing University of Posts and Telecommunications*, April 2013.

[304] 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 Transactions on Neural Networks and Learning Systems*, 27(1):32–46, 2016.

[305] Y. Yang, Y. Feng, and J. Suykens. A nonconvex relaxation approach to robust matrix completion. *Preprint*, 2014.

[306] Q. Yao and J. Kwok. Colorization by patch-based local low-rank matrix completion. *AAAI Conference on Artificial Intelligence*, 2015.

[307] M. Yasin, M. Cetin, and A. Khwaja. SAR imaging of moving targets by subaperture based low-rank and sparse decomposition. *SIU 2017*, 2017.

[308] 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 Transactions on Circuits and Systems for Video Technology*, 2015.

[309] X. Yi, D. Park, Y. Chen, and C. Caramanis. Fast algorithms for robust PCA via gradient descent. *Preprint*, 2016.

[310] H. Yong, D. Meng, W. Zuo, and L. Zhang. Robust online matrix factorization for dynamic background subtraction. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2017.

[311] C. You, C. Li, D. Robinson, and R. Vidal. Oracle Based Active Set Algorithm for Scalable Elastic Net Subspace Clustering. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, 2016.

[312] C. You, D. Robinson, and R. Vidal. Scalable Sparse Subspace Clustering by Orthogonal Matching Pursuit. *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, 2016.

[313] J. Yu. Rank-Constrained PCA for Intrinsic Images Decomposition. *IEEE International Conference on Image Processing, ICIP 2016*, 2016.

[314] N. Zarmehi and F. Marvasti. Recovery of sparse and low rank components of matrices using iterative method with adaptive thresholding. *Preprint*, 2017.

[315] J. Zhan, B. Lois, H. Guo, and N. Vaswani. Online (and Offline) Robust PCA: Novel Algorithms and Performance Guarantees. *Journal of Machine Learning Research*, 2016.

[316] J. Zhan and N. Vaswani. Robust PCA with partial subspace knowledge. *Preprint*, 2014.

[317] H. Zhang, Z. Lin, and C. Zhang. Completing low-rank matrices with corrupted samples from few coefficients in general basis. *IEEE Transactions on Information Theory*, 62(8):4748–4768, 2016.

[318] H. Zhang, Z. Lin, C. Zhang, and E. Chang. Exact recoverability of robust pca via outlier pursuit with tight recovery bounds. In *AAAI*, pages 3143–3149, 2015.

[319] H. Zhang, Z. Lin, C. Zhang, and J. Gao. Robust latent low rank representation for subspace clustering. *Neurocomputing*, 145:369–373, 2014.

[320] H. Zhang, Z. Lin, C. Zhang, and J. Gao. Relations among some low-rank subspace recovery models. *Neural computation*, 27:1915–1950, 2015.

[321] J. Zhang and Y. Zhuang. Adaptive weight selection for incremental eigen-background modeling. *ICME 2007*, pages 851–854, July 2007.

[322] K. Zhang, M. Wang, and S. Yang. Learning low-rank decomposition for pansharpening with spatial-spectral offsets. *Submitted to IEEE Transactions on Neural Networks and Learning Systems*, 2016.

[323] X. Zhang, Z. Lin, F. Sun, and Y. Ma. Rectification of optical characters as transform invariant low-rank textures. *IEEE International Conference on Document Analysis and Recognition*, pages 393–397, 2013.

[324] X. Zhang, Z. Lin, F. Sun, and Y. Ma. Transform invariant text extraction. *The Visual Computer*, 30(4):401–415, 2014.

[325] Y. Zhang, L. Chen, Z. Zhao, and J. Jia. Multi-focus image fusion with sparse feature based pulse coupled neural network. *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, 12(2), 2014.

[326] Y. Zhang, L. Chen, Z. Zhao, and J. Jia. A novel pulse coupled neural network based method for multi focus image fusion. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 7(3):361–370, 2014.

[327] 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-International Journal for Light and Electron Optics*, 125(17):5002–5006, 2014.

[328] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma. TILT: transform invariant low-rank textures. *International Journal of Computer Vision*, page 124, 2012.

[329] Z. Zhang, H. Lei, and Z. Lv. Vehicle Layover Removal in Circular SAR Images via ROSL. *IEEE Geoscience and Remote Sensing Letters*, 12(12):2413–2417, 2015.

[330] Z. Zhang, X. Liang, and Y. Ma. Unwrapping low-rank textures on generalized cylindrical surfaces. *IEEE International Conference on Computer Vision*, pages 1347–1354, 2011.

[331] L. Zhao, X. Zhang, Y. Tian, R. Wang, and T. Huang. A background proportion adaptive Lagrange multiplier selection method for surveillance video on high HEVC. *ICME 2013*, July 2013.

[332] M. Zhao, L. Jiao, W. Ma, H. Liu, and S. Yang. Classification and saliency detection by semi-supervised low-rank representation. *Pattern Recognition*, 2015.

[333] Q. Zhao, D. Meng, Z. Xu, W. Zuo, and L. Zhang. Robust principal component analysis with complex noise. *International Conference on Machine Learning, ICML 2014*, 2014.

[334] X. Zhao, G. An, Y. Cen, H. Wang, and R. Zhao. Robust generalized low rank approximations of matrices for video denoising. *IEEE International Conference on Signal Processing, ICSP 2016*, pages 815–818, 2016.

[335] W. Zhen and Y. Min. New methods for solving the nuclear norm with random matrix and the application in Robust Principal Component Analysis. *Chinese Control and Decision Conference, CCDC 2017*, pages 1323–1328, 2017.

[336] P. Zheng and A. Aravkin. Fast methods for non-smooth non-convex minimization. *Preprint*, 2018.

[337] Q. Zheng, Y. Wang, and P. Heng. Online robust image alignment via subspace learning from gradient orientations. *IEEE International Conference on Computer Vision, ICCV 2017*, 2017.

[338] Y. Zheng, G. Liu, S. Sugimoto, S. Yan, and M. Okutomi. Practical low-rank matrix approximation under robust l1-norm. *International Conference on Computer Vision and Pattern Recognition, CVPR 2012*, 2012.

[339] T. Zhou and D. Tao. GoDec: randomized low-rank and sparse matrix decomposition in noisy case. *ICML 2011*, 2011.

[340] 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. *IJCAI 2013*, April 2013.

[341] T. Zhou and D. Tao. Unmixing incoherent structures of big data by randomized or greedy decomposition. *Preprint*, 2013.

[342] X. Zhou, C. Yang, and W. Yu. Moving object detection by detecting contiguous outliers in the low-rank representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35:597–610, 2013.

[343] X. Zhou, C. Yang, H. Zhao, and W. Yu. Low-rank modeling and its applications in image analysis. *Preprint*, 2014.

[344] Z. Zhou, X. Li, J. Wright, E. Candès, and Y. Ma. Stable principal component pursuit. *IEEE ISIT Proceedings*, pages 1518–1522, June 2010.

[345] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. *International Conference on Computer Vision, ICCV 2011*, pages 479–486, 2011.

Zhouchen Lin is currently a professor with the Key Laboratory of Machine Perception, School of Electronics Engineering and Computer Science, Peking University. His research interests include computer vision, image processing, machine learning, pattern recognition, and numerical optimization. He is an area chair of CVPR 2014/2016, ICCV 2015, NIPS 2015/2018, and AAAI 2019, and a senior program committee member of AAAI 2016/2017/2018 and IJCAI 2016/2018. 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 IAPR Fellow and IEEE Fellow. He also co-authored a book entitled Low-Rank Models in Visual Analysis in 2017.

Thierry Bouwmans is currently an Associate Professor at the University of La Rochelle (France) since 2000. He obtained the diploma of HDR for full professor position in 2014. His research interests consist mainly in the detection of moving objects in challenging environments. He has coauthored two books in CRC Press (background/foreground separation for video surveillance, robust PCA via decomposition in low rank and sparse matrices). His research investigated particularly the use of robust PCA in video surveillance. He is also the main organizers of the RSL-CV workshops hosted at ICCV in 2015 and 2017. He is a reviewer for international journals including IEEE (Trans. on Image Processing, Trans. on Multimedia, Trans. on CSVT, etc.), SPRINGER (IJCV, MVA, etc.) and ELSEVIER (CVIU, PR, PRL, etc.), and top-level conferences such as CVPR, ICPR, ICIP, AVSS, etc..

Sajid Javed is currently a Post-Doctoral research fellow in the Department of Computer Science, University of Warwick, United Kingdom, from October, 2017. Dr. Sajid is working actively on MiCAHiL project under the supervision of Prof. Nasir Rajpoot at TIA Lab. Before joining the TIA Lab, he obtained his Bs.c (hons) degree in Computer Science from University of Hertfordshire, UK, in 2010. He joined the Virtual Reality Laboratory of Kyungpook National University, Republic of Korea, in 2012, where he completed his combined Master's and Doctoral degrees in Computer Science under the supervision of Prof. Soon Ki Jung and co-supervision of Prof. Thierry Bouwmans from MIA Lab, France in 2017. Dr. Sajid has co-authored about 30 publications, this includes several journals and international conferences publications in the area of RPCA for background-foreground modeling. His current research areas are related to medical image analysis using deep learning algorithms.

Ricardo Otazo is an Associate Professor of Radiology at New York University School of Medicine. 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, or GRASP; and low-rank plus sparse, or 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 in 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.

Hongyang Zhang is currently a Ph.D. candidate in Machine Learning Department, Carnegie Mellon University, Pittsburgh, USA. He received the Masters degree in computer science from Peking University, Beijing, China in 2015. His research interests include machine learning, theoretical computer science, optimization, and statistics. He also co-authored a book entitled Low-Rank Models in Visual Analysis in 2017.