Pattern Recognition 63 (2017) 653–666

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Extracting contrast-filled vessels in X-ray angiography by graduated RPCA with motion coherency constraint☆

Mingxin Jin[a,1], Rong Li[b,1], Jian Jiang[b], Binjie Qin[a,*]

[a] School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
[b] School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

ARTICLE INFO

Keywords:
Subspace estimation
Robust principal component analysis (RPCA)
Low-rank model
Spatio-temporal motion coherency
Matrix decomposition
Trajectory decomposition
X-ray coronary angiograms

ABSTRACT

X-ray coronary angiography can provide rich dynamic information of cardiac and vascular function. Extracting contrast-filled vessel from the complex dynamic background (caused by the movement of diaphragm, lung, bones, etc.) in X-ray coronary angiograms has great clinical significance in assisting myocardial perfusion evaluation, reconstructing vessel structures for diagnosis and treatment of heart disease. Considering the angiography image sequence is a sum of a low-rank background matrix and a sparse flowing contrast agent matrix, we propose a novel graduated robust principal component analysis (RPCA) with spatio-temporal motion coherency constraint to accurately extract contrast-filled vessel from the X-ray coronary angiograms: (1) We first use a statistically structured RPCA with complex noise model exploiting the complex structural connectivity of vessel regions to identify all candidate foreground contrast-filled vessels; (2) To eliminate the background remained in the candidates, we further introduce trajectory decomposition on the candidate foregrounds to accurately extract contrast-filled vessels using motion coherency regularized RPCA, which imposes total variation minimization on the foreground trajectories to model the spatio-temporal contiguity and smoothness of the foreground trajectories. The graduated RPCA with motion coherency constraint shows to consistently outperform other state-of-the-art methods, in particular on real-world X-ray coronary angiograms that contain a significant amount of complex dynamic background motion. Experimental results on twelve sequences of real X-ray coronary angiograms are evaluated using both qualitative and quantitative methods to demonstrate the obvious advantages of our method over the state-of-the-art alternatives.

1. Introduction

The low-rank and sparse matrix decomposition based on subspace estimation has recently become an important topic in machine learning and computer vision. Based on the fact that big data lie on some low-dimensional subspace, robust principal component analysis (RPCA) [1,2] assumes that an original observed vectorized video frame $\mathbf{D} \in \mathbb{R}^{m \times n}$ (m = total image pixels in each frame, n = number of frames) is composed of low-rank matrix $\mathbf{L}$ representing the static or relatively static background, and a sparse matrix $\mathbf{S}$ consisting of the outlier moving objects. Moving object detection using RPCA is therefore to decompose the matrix $\mathbf{D}$ into low-rank plus sparse matrices. RPCA is popularly used in various computer vision applications, such as change detection, video surveillance and image restoration [3–5]. In biomedical imaging, Gao et al. [6] employed the low-rank and sparse decomposition to reconstruct 4D-CT images with undersampled cardiac data. Otazo et al. [7] applied this decomposition model to reconstruct undersampled dynamic MRI in various problems of clinical interest. Liu et al. [8] proposed low-rank-plus-sparse image decomposition with group-wise image registration to realize low-rank atlas image analysis in the presence of pathologies. Karn et al. [9] used RPCA-based method to recover a low-rank matrix for iris identification by removing possible sparse corruptions. Campbell-Washburn et al. [10] presented an application of the RPCA algorithm to remove RF spike noise from k-space in magnetic resonance imaging.

In this paper, graduated RPCA with motion coherency constraint is proposed to automatically extract contrast-filled vessel from the complex dynamic background, which is caused by the occlusions and movement of diaphragm, lung, bones, etc. in X-ray coronary angiograms (XCA). X-ray coronary angiography is considered the gold standard in clinical decision making and therapy guidance [11] for cardiovascular disease (CD). It records the contrast agent inflow

☆ Fully documented templates are available in the elsarticle package on CTAN.

* Corresponding author.
E-mail address: bjqin@sjtu.edu.cn (B. Qin).

[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.patcog.2016.09.042
Received 31 January 2016; Received in revised form 7 September 2016; Accepted 21 September 2016
Available online 06 October 2016
0031-3203/© 2016 Elsevier Ltd. All rights reserved.

through the coronary vessels and myocardium to represent the complex 3D/4D (3D+time) structure of coronary vessels by 2D X-ray projections. However, X-ray coronary angiography is fundamentally limited in some aspects: the loss of 3D/4D information of the coronary arteries due to the consequences of the projection operation, intensity inhomogeneities due to blood flow inside the arteries, overlap of different other background structures with varying intensity (e.g., spine, ribs or diaphragm), and complex respiratory and cardiac motions. To facilitate diagnosis and treatment of CD, automatically and accurately extracting contrast-filled vessel from complex background and motion disturbances is a prerequisite for further processing, such as selective myocardial perfusion measurement [12–14], visual estimation of myocardium healthiness [15,16], automatic detection of coronary stenosis in XCA [17], coronary motion analysis [18], tomographic reconstruction of coronary arteries [19], and pre/intra-operative image registration [20]. In addition, most vessel segmentation [21,22] problems have been addressed in two different steps, and definitely benefit from the first step of vessel extraction (also called vessel detection or vessel enhancement). Usually, the vessel extraction is used for removing noise and complex background from XCA while emphasizing vessel-like structures for the final step of vessel-like structure classification.

Currently, there are three main types of vessel extraction methods for XCA in various clinical applications: image registration based methods, vessel enhancement or vessel detection, and motion layer separation. Firstly, vessel extraction is implemented by digital subtraction angiography (DSA) [12] or DSA-like synchronized subtraction, which clearly visualize blood vessels and remove interfering background structures by subtracting a pre-contrast mask image from later contrast images. However, the mask and the contrast images in DSA are acquired at different times, and the position of the background tissues around the vessels often changes due to human body motion. Thus, direct subtraction of these images would produce motion artifacts in the resulting DSA images. To reduce the motion artifact, some registration methods are often employed before subtraction [23,24]. In minimizing the dissimilarity between the mask and live images, the registration methods find the correspondence between pixels in mask and live images, and calculate certain warping transformation which is to be applied to the mask image. After subtracting the live images and locally transformed mask image, the motion artifacts can be greatly eliminated. Tang et al. [25] separated the vessel signals and the background signals from the mask images and the contrast images by implementing independent component analysis. This approach still requires the acquisition and registration of corresponding mask and live images. However, the efficiency of the registration-based methods might reduce with a large camera motion and the noisy intensity variation in the two images. Secondly, most vessel enhancement methods proposed as vessel extraction for subsequent vessel segmentation can be classified into two main categories: model-free and model-based methods. The model-free methods usually use linear filtering (such as Gaussian and Gabor filtering) [26,27], nonlinear filtering and Hessian-based [28,29] multi-scale filtering to extract the vessel-like feature from image content. These methods detect vessel arteries at the pixel level without considering the global vessel structures and the structure relationship among image sequence, resulting in their poor performance in detecting various vessels from the complex background with occlusions and artifacts. On the other hand, some model-based methods, such as centerline-based model [30], active contour model [31] and minimal path techniques [32], have been proposed to represent vessels by exploiting the spatial coherence existing in the image. Most model-based algorithms are semi-automatic approaches, requiring user input as an initial model state. However, the model-based methods are usually sensitive to the initial model state and may have limitations for suffering from low contrast, complex background and varied vessel shape within XCA. Though all these methods are required to remove motion artifacts from the extracted vessel, these methods do not benefit from exploiting the different motion characteristics inherent in the different structures in XCA.

In fact, we observe that the different anatomical structures and contrast agent inflow in XCA have different motion patterns by viewing all these structures from the perspective of motion analysis. Specifically, lung is breathing slow and heart is beating fast, while the contrast agent in artery vessels is flowing faster. Therefore, XCA is a superposition of different motion layers, extraction of contrast-filled vessel can be regarded as motion layer separation [33–37]. Several research groups [33,34] proposed a Bayesian probabilistic model or least square computation framework [33] combined with dense motion estimation to separate coronary layers from background structures. However, these methods either required several pre-contrast images as static masks for background estimation [33] or manually selected control points to get sparse motion estimation [34,35]. Though surrogate-driven estimation of respiratory motion [36] can improve motion layer separation, it is still very difficult to accurately solve the ill-posed problem of motion estimation for multi-layers in XCA. Without requiring motion estimation, separating foreground moving objects from quasi-static background has recently attracted great attention by exploiting the low-rank and sparse matrix decomposition in computer vision. Ma et al. [37] decomposed XCA into independently moving layers using classical RPCA to achieve vessel extraction from complex background. However, it is a prerequisite to remove large-scale breathing structures (such as diaphragm) from the original images.

Before we dig into some RPCA-based matrix decomposition methods, it will help us to know that XCA is a natural candidate for low-rank modeling, due to the correlation between the frames of XCA. One of the most challenging tasks in vessel extraction is to estimate a good model for the background variations in XCA: the complex dynamic background is overlap of underlying large-scale diaphragm, lung and bones with slow motion disturbance and varying illumination. This task is complicated by the presence of foreground contrast-filled vessels: due to possible irregular heart beat and foreshortening of 2D projection as well as vessel bifurcations, the contrast agents are perfused with relatively fast and different speeds in an intricate network of large vessels and smaller vessel branches. These contrast-filled vessels are then displayed as foreground anomaly out of the complex background. In such situations, it is natural to model the background variations as approximately low rank. The contrast-filled vessels generally occupy only a fraction of the image pixels and hence can be treated as sparse errors in XCA. Therefore, foreground/background separation for XCA is low-rank and sparse matrix decomposition to solve the following RPCA problem under via principal component pursuit (RPCA-PCP) [1,2]:

$$\min_{\mathbf{L},\mathbf{S}} \| \mathbf{L} \|_* + \lambda \| \mathbf{S} \|_1 \quad \text{s. t.} \quad \mathbf{D} = \mathbf{L} + \mathbf{S}. \tag{1}$$

where $\| \mathbf{L} \|_*$ is the nuclear norm of matrix $\mathbf{L}$ (which is the sum of its singular values), the second term is L_1 norm regularization with $\| \mathbf{S} \|_1$. λ is a regularizing parameter to control the amount of outliers in the RPCA, whereby setting λ to a large value would absorb a lot of foreground into the background while a small value would make the foreground detection method very sensitive to background motion and noises.

However, using classical RPCA-PCP based methods [1,37–40] to extract contrast-filled vessels presents the following limitations: First, the dynamic background in XCA can have complex changes. It mainly contains respiratory and cardiac motion as well as possible patient movement, which can be larger than that of the foreground contrast agents. There also could be complex occlusions and artifacts caused by gas, bones and motion, causing complex intensity variation. All of these factors can cause background remained in the foreground extraction results. It is thus very difficult to directly model the background well.

Second, because the geometry of L_1 norm is diamond shape and its regularization treats each entry pixel independently, these methods did not fully consider complex structural distribution of the contrast-filled vessels with complex noise in the low dose XCA. Though foreground object's prior has been reflected in either block-sparsity measure [41,3] to detect the block containing moving objects or Markov random field (MRF) [42,43] prior to impose smoothness constraint on foreground, the block-sparsity property still does not exploit structured information to fully model sparse foreground moving objects with complex noise. Recent work [3] relies heavily on block-sparsity constrained RPCA on a low resolution video to roughly detect candidate foreground regions with salient motion, and adapt the value of λ based on the saliency measure of detected regions in the subsequent PRCA. However, as we will see in Section 3, the block-sparsity constrained RPCA cannot guarantee accurate estimation of foreground candidates and full removal of background for the final result. Though the MRF-based method can effectively eliminate noise and small background movements, foreground regions tend to be over-smoothed due to the smoothness constraint. Recently, Zhao et al. [44] modeled sparse foreground component and complex noise as a mixture of Gaussians (MoG) to enhance the detection rate of complex foreground [45], but it cannot fully remove background and noise from the foreground moving object. Third, flowing contrast agents in vessels present spatio-temporal contiguity and smoothness (or continuity) with smooth trajectories through the image sequence. This spatio-temporal motion coherency can be used as an appropriate constraint to accurately extract the contrast-filled vessels. However, the spatio-temporal contexts between the moving foreground objects in different frames are usually lost in classical RPCA-based methods. Furthermore, the contrast agents have varying scales and different flowing speeds through vessel network, and might be insensitive to some motion saliency detection methods [3]: it is not even meaningful to have a single global λ in a single RPCA without other motion constraint for vessel extraction from XCA.

To handle all these challenges by merely assuming low-rank model for the complex background with noise disturbances, we need graduated RPCA strategy with spatio-temporal constraint to accurately extract the contrast-filled vessels. In fact, the structured sparsity [46–49] norm has been successfully developed to model the spatial contextual information about nonzero structural patterns of variables in sparse signal recovery. Recently, the structured sparsity constrained RPCA has been applied [49] into the sub-sampled video sequences to extract the foreground structured information, but the subsequent group-sparse RPCA relies on motion-saliency-based foreground support estimation, which is insensitive to some large dynamic background motions and may not detect real small-scale foreground objects. Being different from the two works [3,49] that use motion saliency detection in designing an adaptive λ to segment the various foreground motions, our graduated scheme performs a statistically structured RPCA [44] under the Bayesian framework for full extraction of all candidate foreground structures with noisy disturbances. We further propose motion coherency regularized RPCA on the candidate foreground matrix to eliminate the background motion while accurately extract all contrast-filled vessels. This is based on the fact that the flowing contrast agents are continuously projected to neighboring pixels on multiple frames with the pixels being grouped together as contiguous and sparse trajectories. In a 2D coordinate system with two axes being parallel to the foreground matrix's row and column directions, a specific trajectory of flowing contrast agent takes a sinuous course along the temporal direction (i.e., the row direction or x-axis) of foreground matrix, so that the entries of foreground matrix are not randomly distributed but spatio-temporally clustered within some specific rows and their branches. This observation makes the foreground trajectory matrix satisfy the mathematical definition of total variation (TV), which usually takes spatio-temporal information differences into consideration to segment coherent motion in computer vision [52]. Therefore, the TV norm is assumed to well guarantee motion coherency and spatio-temporal smoothness for contrast-filled vessel extraction.

To the best of our knowledge, we are the first to integrate TV into constrained trajectory decomposition for vessel extraction from XCA. The contribution of this paper are summarized as follows:

(1) First, we propose a novel graduated scheme of RPCA with hierarchical motion coherency constraint to automatically and accurately extract contrast-filled vessels. We first guarantee the extraction of all candidate foreground structures by implementing a statistically structured RPCA on original images, and then exploit spatio-temporal contiguity and smoothness in foreground contrast agent trajectory to design motion coherency regularized RPCA, which decomposes the candidate foreground matrix into foreground/background motion trajectories to accurately extract all contrast-filled vessel structures.

(2) Second, we integrate TV regularization into the RPCA framework to design the motion coherency regularized RPCA on candidate foreground for accurate foreground/background motion trajectory decomposition. During the preparation of this manuscript we became aware of a similar TV usage in RPCA by Cao et al. [53], named TVRPCA, which separates dynamic background from moving object using the spatial continuity of foreground, and detects lingering objects using the temporal continuity of foreground in observed video. However, the focus of that work was on application of single TV regularization for foreground detection in RPCA framework, which did not integrate the structure and trajectory priori information into the background subtraction and may probably miss some small and fast moving objects [53].

The remaining part of this paper is organized as follows: In Section 2, a detailed description of the proposed graduated RPCA with motion coherency constraint is given. In Section 3, we present experimental comparison and evaluation with the proposed and other state-of-the-art methods and Section 4 provides the conclusion.

2. Methods

2.1. Candidate contrast-filled vessel detection

Notations: Given a sequence of XCA, we obtain an original data matrix $\mathbf{D} \in \mathbb{R}^{m\times n}$, which is composed of the vectorized angiographic image frames $\mathbf{D} = \{I_1, \ldots, I_t, \ldots, I_n\}$, in which I_t is the frame at time t and n is the total frames.

To accurately detect all foreground vessels without any omission, we identify all candidate foreground contrast-filled vessels for the refined vessel extraction in subsequent spatio-temporally coherent RPCA. Liu et al. [49] investigated how block-sparse RPCA [3] achieves incomplete detection of foreground candidates. To avoid this incomplete foreground detection and provide more complete candidate vessels than classical block-sparse RPCA methods, we harness the structural information of vessels containing flowing contrast agent within each angiogram. We observe that the candidate foreground regions are not pixel-wisely (and block-wisely) sparse but rather sparsely clustered and correlated within vessel structures, such that candidate vessels constitute contiguous regions. Therefore, we try to introduce structured sparsity constraint [46–49] in RPCA for extracting all candidate contrast-filled vessels. The most natural form of structured sparsity is group sparsity [38] to partition variables into disjoint groups. Given a single image vector $\mathbf{s} \in \mathbb{R}^m$, the variables with indices $\{1, \ldots, m\}$ can be partitioned into a disjoint set of groups, denoted as $\mathcal{G}$, with each group $\mathsf{g} \in \mathcal{G}$ containing a subset of these indices, and η_{g} being a positive scalar weight indexed by group g. A group sparsity inducing norm is defined as $\Omega(\mathbf{s}) = \sum_{\mathsf{g}\in\mathcal{G}} \eta_{\mathsf{g}} \| \mathbf{s}_{\mathsf{g}} \|_q$ for any $q \in (1, \infty]$. As expected, regularizing with $\Omega(\mathbf{s})$ divides $\mathbf{s}$ into disjoint local patches. However, the real structural pattern in $\mathbf{s}$ corresponds to the overlapping projection of the complex network of large vessels and smaller vessel bifurcations as well as complex noise.

Therefore, pre-design of disjoint group structures cannot accurately model the overlapping projection of complex vessel network to extract the real foreground pattern **s** in XCA.

Generally, a sparse matrix **S** is associated with disturbances (e.g., corrupted data values, a moving object in the foreground of motion videos), and sparse components are also called 'noise' in some literatures. To induce more complex and diverse sparsity patterns for contrast-filled vessel extraction, we therefore consider the sparse foreground flowing contrast agents as a mixture of statistical structures and complex noise. Since Gaussian mixture model constructs a universal estimator to any continuous density function in theory, we thus use generative RPCA model under the Bayesian framework to model the complex vessel network and sparse data noise as a MoG distribution. In fact, some rigorous experimental evaluations using both synthetic and real videos indicate [45] that the MoG-based RPCA (MoG-RPCA) [44] can achieve high detection rate (or recall) of foreground moving object. The detection rate is a high quality metric to indicate RPCA's power to extract all foreground pixels when compared with the total number of foreground pixels in the ground truth. Therefore, MoG-RPCA is used to fully extract the candidate contrast-filled vessel in the first step of our graduated RPCA scheme.

2.1.1. Foreground/background modeling

A. *Candidate contrast-filled vessel modeling*: To extract the candidate contrast-filled vessel, each s_{ij} in sparse matrix **S** is assumed to have followed a MoG distribution [44,54]:

$$p(s_{ij}|\boldsymbol{\mu},\boldsymbol{\tau},\mathbf{z}_{ij}) = \prod_{k=1}^{K}\mathcal{N}(s_{ij}|\mu_k,\tau_k^{-1})^{z_{ijk}}p(\mathbf{z}_{ij}|\boldsymbol{\pi}) = \prod_{k=1}^{K}\pi_k^{z_{ijk}} \tag{2}$$

where $\mathbf{z}_{ij}$ is a 1-of-K indicator vector associated with s_{ij}, i.e., $z_{ijk}\in\{0,1\}$, $\sum_{k=1}^{K}z_{ijk}=1$. $\boldsymbol{\pi}=(\pi_1,\ldots\pi_K)$ is the mixing coefficient vector, with π_k denoting the existence probability of the kth Gaussian component, where $\boldsymbol{\pi}$ satisfies $0\le\pi_k\le 1$ and $\sum_{k=1}^{K}\pi_k=1$. $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_K)$ and $\boldsymbol{\tau}=(\tau_1,\ldots,\tau_K)$, where μ_k and τ_k are the mean and precision (inverse variances) of the kth Gaussian component, respectively. Here we let $\mathbf{Z}$ denote a $m\times n\times K$ array with the (i,j,k) th element denoted by z_{ijk} to facilitate the following description.

The parameters μ_k and τ_k are also modeling by the Gaussian-Gamma distribution, and the Dirichlet distribution is chosen to model the mixing coefficient $\boldsymbol{\pi}$ [54]:

$$p(\mu_k,\tau_k) = \mathcal{N}(\mu_k|\mu_{0k},\beta_0^{-1}\tau_k^{-1})\cdot\mathrm{Gam}(\tau_k|\alpha_0,\chi_0)p(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}|\boldsymbol{\eta}_0) \tag{3}$$

where μ_{0k} is the mean of the kth Gaussian component, $\mathrm{Gam}(\cdot)$ denotes the Gamma distribution, the hyperparameters α_0, β_0, and χ_0 are set to be small deterministic value (e.g., 10^{-5}) to obtain broad hyperpriors, and $\mathrm{Dir}(\boldsymbol{\pi}|\boldsymbol{\eta}_0)$ denotes the Dirichlet distribution parameterized by $\boldsymbol{\eta}_0=(\eta_{01},\ldots,\eta_{0K})$.

B. *Low-rank background modeling*: Due to its fast speed and good scalability, the automatic relevance determination (ARD) [55] is adopted in MoG-RPCA for low-rank background modeling. We formulate $\mathbf{L}\in\mathbb{R}^{m\times n}$ with rank $l\le\min(m,n)$ as the product of $\mathbf{U}\in\mathbb{R}^{m\times R}$ and $\mathbf{V}\in\mathbb{R}^{n\times R}$:

$$\mathbf{L}=\mathbf{U}\mathbf{V}^T=\sum_{r=1}^{R}\mathbf{u}_{\cdot r}\mathbf{v}_{\cdot r}^T \tag{4}$$

where $R>l$, and $\mathbf{u}_{\cdot r}(\mathbf{v}_{\cdot r})$ is the rth column of $\mathbf{U}(\mathbf{V})$. To guarantee the low-rank nature of $\mathbf{L}$, our goal is to achieve column sparsity in $\mathbf{U}$ and $\mathbf{V}$, such that some columns in $\mathbf{U}$ and $\mathbf{V}$ will approach zeros. This goal can be achieved by imposing the following priors on $\mathbf{U}$ and $\mathbf{V}$:

$$p(\mathbf{u}_{\cdot r}) = \mathcal{N}(\mathbf{u}_{\cdot r}|\mathbf{0},\gamma_r^{-1}\mathbf{I}_m),\quad p(\mathbf{v}_{\cdot r}) = \mathcal{N}(\mathbf{v}_{\cdot r}|\mathbf{0},\gamma_r^{-1}\mathbf{I}_n) \tag{5}$$

where $\mathbf{I}_m$ denotes the $m\times n$ identity matrix. The conjugate prior on each precision variable is:

$$p(\gamma_r) = \mathrm{Gam}(\gamma_r|a_0,b_0)\propto\gamma_r^{a_0-1}\exp(-b_0\gamma_r) \tag{6}$$

where the parameters a_0 and b_0 are also treated as small deterministic values to obtain broad hyperpriors. It has been validated that this common precision variable γ_r could lead to large precision values of some γ_rs, resulting in a good low-rank estimate of **L** [55].

We combine Eqs. (2)–(6) with condition $\mathbf{D}=\mathbf{L}+\mathbf{S}$ to construct the statistically structured MoG-RPCA for candidate contrast-filled vessel detection. The goal of candidate contrast-filled vessel detection turns to infer the posterior of all involved variables:

$$p(\mathbf{U},\mathbf{V},\mathbf{Z},\boldsymbol{\mu},\boldsymbol{\tau},\boldsymbol{\pi},\boldsymbol{\gamma}|\mathbf{D}) \tag{7}$$

where $\mathbf{Z}=\{\mathbf{z}_{ij}\}$, $\boldsymbol{\mu}=(\mu_1,\ldots,\mu_K)$, $\boldsymbol{\tau}=(\tau_1,\ldots,\tau_K)$, and $\boldsymbol{\gamma}=(\gamma_1,\ldots,\gamma_R)$.

2.1.2. Variational Bayesian inference

To infer the posterior of MoG-RPCA,[2] we seek an approximation distribution $q(\mathbf{x})$ to the true posterior $p(\mathbf{x}|\mathcal{D})$ by minimizing the following Kullback–Leibler (KL) divergence $\mathrm{KL}(q\|p)$ between $q(\mathbf{x})$ and $p(\mathbf{x}|\mathcal{D})$:

$$\min_{q\in C}\mathrm{KL}(q\|p) = -\int q(\mathbf{x})\ln\left\{\frac{p(\mathbf{x}|\mathcal{D})}{q(\mathbf{x})}\right\}\mathrm{d}\mathbf{x} \tag{8}$$

where C denotes the set of probability densities with certain restrictions to make the minimization tractable. Here q is generally set by partitioning the elements of $\mathbf{x}$ into disjoint groups $\{\mathbf{x}_i\}$, and then assuming that it can be factorized as $q(\mathbf{x})=\prod_i q_i(\mathbf{x}_i)$. Under this assumption, the closed-form solution for each group $\{\mathbf{x}_j\}$, with the others fixed, can be attained by

$$q_j^*(\mathbf{x}_j) = \frac{\exp(\mathbb{E}_{i\ne j}[\ln p(\mathbf{x},\mathcal{D})])}{\int\exp(\mathbb{E}_{i\ne j}[\ln p(\mathbf{x},\mathcal{D})])\mathrm{d}\mathbf{x}_j} \tag{9}$$

where $p(\mathbf{x},\mathcal{D})$ is the joint distribution of parameter $\mathbf{x}$ and the observations $\mathcal{D}$, and $\mathbb{E}_{i\ne j}[\cdot]$ denotes the expectation with respect to $\mathbf{x}_i$ s except $\mathbf{x}_j$. The solution to Eq. (8) can then be approached through alternatively optimizing each $q_j(\mathbf{x}_j)$ by Eq. (9). Utilizing the general results above, the closed-form variational inference schemes of MoG-RPCA can then be derived as follows.

The posterior distribution (7) is derived with the following factorized form:

$$q(\mathbf{U},\mathbf{V},\mathbf{Z},\boldsymbol{\mu},\boldsymbol{\tau},\boldsymbol{\pi},\boldsymbol{\gamma}) = \prod_i q(\mathbf{u}_{i\cdot})\prod_j q(\mathbf{v}_{j\cdot})\prod_{ij}q(\mathbf{z}_{ij})\prod_k q(\mu_k,\tau_k)q(\boldsymbol{\pi})\prod_r q(\gamma_r) \tag{10}$$

where $\mathbf{u}_{i\cdot}(\mathbf{v}_{j\cdot})$ is the ith (jth) row of $\mathbf{U}(\mathbf{V})$. We can estimate all the factorized distributions involved in Eq. (10) as follows.

A. *Estimation of candidate foreground component*: To estimate the parameters $\boldsymbol{\mu}$, $\boldsymbol{\tau}$, $\mathbf{Z}$ and $\boldsymbol{\pi}$ involved in the foreground detection, we get the following update equation for each parameter using the prior imposed in Eq. (3) and its conjugate property:

$$q(\mu_k,\tau_k) = \mathcal{N}(\mu_k|m_k,(\beta_k\tau_k)^{-1})\cdot\mathrm{Gam}(\tau_k|\alpha_k,\chi_k) \tag{11}$$

where

[2] http://www4.comp.polyu.edu.hk/~cslzhang/code/MoG-RPCA.rar

$$\beta_k = \beta_0 + \sum_{ij} \mathbb{E}[z_{ijk}],\ m_k = \frac{1}{\beta_k}\left(\beta_0\mu_{0k} + \sum_{ij}\mathbb{E}[z_{ijk}](d_{ij} - \mathbb{E}[\mathbf{u}_{i\cdot}]\mathbb{E}[\mathbf{v}_{j\cdot}]^T)\right),$$

$$\alpha_k = \alpha_0 + \frac{1}{2}\sum_{ij}\mathbb{E}[z_{ijk}],$$

$$\chi_k = \chi_0 + \frac{1}{2}\left\{\sum_{ij}\mathbb{E}[z_{ijk}]\mathbb{E}[(d_{ij} - \mathbf{u}_{i\cdot}\mathbf{v}_{j\cdot}^T)^2] + \beta_0\mu_{0k}^2 - \frac{1}{\beta_k}\left(\sum_{ij}\mathbb{E}[z_{ijk}](d_{ij} - \mathbb{E}[\mathbf{u}_{i\cdot}]\mathbb{E}[\mathbf{v}_{j\cdot}]^T) + \beta_0\mu_0\right)^2\right\}$$

The posterior mixing coefficients $\boldsymbol{\pi}$ is similarly updated using the following equation:

$$q(\boldsymbol{\pi}) = \mathrm{Dir}(\boldsymbol{\pi}|\boldsymbol{\eta}) \tag{12}$$

where $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_K)$, $\eta_k = \eta_{0k} + \sum_{ij}\mathbb{E}[z_{ijk}]$. The variational posterior for the indicators $\mathbf{Z}$ is derived in the following form:

$$q(\mathbf{z}_{ij}) = \prod_k r_{ijk}^{z_{ijk}} \tag{13}$$

where

$$r_{ijk} = \frac{\delta_{ijk}}{\sum_k \delta_{ijk}},$$

$$\delta_{ijk} = \tfrac{1}{2}\mathbb{E}[\ln\tau_k] - \tfrac{1}{2}\ln 2\pi - \tfrac{1}{2}\mathbb{E}[\tau_k]\mathbb{E}[(d_{ij} - \mathbf{u}_{i\cdot}\mathbf{v}_{j\cdot}^T - \mu_k)^2] + \mathbb{E}[\ln\pi_k]$$

B. *Estimation of low-rank background component*: The posterior distribution for each row $\mathbf{u}_{i\cdot}$ of $\mathbf{U}$ (and $\mathbf{v}_{j\cdot}$ of $\mathbf{V}$) involved in low-rank background component can be approximated by

$$q(\mathbf{u}_{i\cdot}) = \mathcal{N}\left(\mathbf{u}_{i\cdot}\middle|\boldsymbol{\mu}_{\mathbf{u}_{i\cdot}}, \sum_{\mathbf{u}_{i\cdot}}\right), \quad q(\mathbf{v}_{j\cdot}) = \mathcal{N}\left(\mathbf{v}_{j\cdot}\middle|\boldsymbol{\mu}_{\mathbf{v}_{j\cdot}}, \sum_{\mathbf{v}_{j\cdot}}\right) \tag{14}$$

where

$$\boldsymbol{\mu}_{\mathbf{u}_{i\cdot}}^T = \sum_{\mathbf{u}_{i\cdot}}\left\{\sum_k \mathbb{E}[\tau_k]\sum_j \mathbb{E}[z_{ijk}](d_{ij} - \mathbb{E}[\mu_k])\mathbb{E}[\mathbf{v}_{j\cdot}]\right\}^T,$$

$$\sum_{\mathbf{u}_{i\cdot}} = \left\{\sum_k \mathbb{E}[\tau_k]\sum_j \mathbb{E}[z_{ijk}]\mathbb{E}[\mathbf{v}_{j\cdot}^T\mathbf{v}_{j\cdot}] + \boldsymbol{\Gamma}\right\}^{-1},$$

$$\boldsymbol{\mu}_{\mathbf{v}_{j\cdot}}^T = \sum_{\mathbf{v}_{j\cdot}}\left\{\sum_k \mathbb{E}[\tau_k]\sum_i \mathbb{E}[z_{ijk}](d_{ij} - \mathbb{E}[\mu_k])\mathbb{E}[\mathbf{v}_{j\cdot}]\right\}^T,$$

$$\sum_{\mathbf{v}_{j\cdot}} = \left\{\sum_k \mathbb{E}[\tau_k]\sum_i \mathbb{E}[z_{ijk}]\mathbb{E}[\mathbf{u}_{i\cdot}^T\mathbf{u}_{i\cdot}] + \boldsymbol{\Gamma}\right\}^{-1},\ \boldsymbol{\Gamma} = \mathrm{diag}(\mathbb{E}[\boldsymbol{\gamma}])$$

2.2. Motion coherency regularized RPCA for trajectory decomposition

Through the statistically structured RPCA, the candidate foregrounds have been extracted from angiographic image sequence completely. However, there still exist a small part of background motion and complex noise disturbance (as shown in the middle two columns in Fig. 1 and the middle column in Fig. 2) in the candidates' foregrounds, because the statistically structured RPCA cannot encode the long term temporal coherency through the image sequences due to the lack of correspondence between nonadjacent frames. For this reason, Cui et al. [50] and Ren et al. [51] propose trajectory decomposition to decompose trajectory matrix into low-rank background matrix and group-sparse foreground matrix. These methods preserve temporal consistency by enforcing the pixels belonging to the same trajectory to have similar labels, and further preserve spatial connectivity [51] by incorporating MRF model into trajectory decomposition. However, they cannot deal with the occlusions in XCA when implementing optical flow based dense point tracker to get trajectory matrix as input data, and cannot model spatio-temporal coherency with simultaneously avoiding over-smoothness effect caused by the MRF modeling on the foreground regions. The spatio-temporal motion coherency is directly represented as spatio-temporal contiguity and smoothness in the image pixels of foreground trajectories. If such trajectory properties can be effectively modeled, there will be highly significant in accurately extracting contrast-filled vessels from background.

After the candidate foreground extraction, the foreground matrix is simply assumed to be a trajectory-level matrix for subsequent foreground/background trajectory decomposition. Due to the sparsely distributed sinuous courses of coronary vessels and small vessels, the foreground trajectories occupy spatio-temporally contiguous regions through image sequences. Therefore, the entries of candidate foreground matrix are sparsely clustered in the column (or y) and the row (or x) directions simultaneously. Thus, the $L_{2,1}$ norm based group sparsity constraint [50,51] $\|\mathbf{S}\|_{2,1}$, which forces exact zero columns or rows, is not appropriate for this application. Nevertheless, the widely used $\boldsymbol{L}_1$ norm, which handles images pixelly, will be appropriate to be combined with some motion coherency constraints. Specifically, for any matrix $\mathbf{S} \in \mathbb{R}^{m\times n}$, the L_1 norm is defined as $\|\mathbf{S}\|_1 = \sum_{i,j}|\mathbf{S}_{i,j}|$.

As for the motion coherency constraint, the TV minimization based algorithm is a very popular approach because of its effectiveness in preserving edge information and promoting piecewise smoothness. In mathematics, the difference operator can be used to measure the sensitivity of change of a discrete function. Based on the difference operator, the TV regularization of matrix $\mathbf{S}$ can be defined as $\|\mathbf{S}\|_{\mathrm{TV}} = \|\nabla_x\mathbf{S}\|_1 + \|\nabla_y\mathbf{S}\|_1 + \|\nabla_z\mathbf{S}\|_1$, where ∇_x, ∇_y and ∇_z correspond to the difference operation performed along the x (or row), y (or column) and z (or frame) directions, respectively. Due to the possible coherency differences between spatial dimension and temporal dimension, it is reasonable to set different weights on ∇_x, ∇_y and ∇_z, such that $\|\mathbf{S}\|_{\mathrm{TV}} = w_x\|\nabla_x\mathbf{S}\|_1 + w_y\|\nabla_y\mathbf{S}\|_1 + w_z\|\nabla_z\mathbf{S}\|_1$. However, testing different weights inside the TV norm is out of the scope of this work. This work simply sets $w_x = w_y = w_z = 1$. The combined regularization effect of TV and L_1-norm minimization encourages the proximate foreground trajectories with similar appearances to group into the same foreground class.

Based on the above-mentioned spatio-temporal constraints, the motion coherency regularized RPCA is carried out on the candidate foregrounds to elaborately extract contrast-filled vessels through the following formulation:

$$\min_{\mathbf{L},\mathbf{S}} \|\mathbf{L}\|_* + \lambda_1\|\mathbf{S}\|_1 + \lambda_2\|\mathbf{S}\|_{\mathrm{TV}} \quad \text{s. t. } \mathbf{D} = \mathbf{L} + \mathbf{S} \tag{15}$$

where matrix $\mathbf{D}$ represents the candidate foregrounds, the matrixes $\mathbf{L}$ and $\mathbf{S}$ represent new low-rank background and sparse foreground trajectories, respectively. λ_1 and λ_2 are regularizing parameters.

In recent years, a lot of efficient numerical algorithms have been proposed to solve the difficult minimization problem in Eq. (15). Among them, alternating direction method (ADM) is one of the most successful algorithms [56]. In theory, ADM can guarantee the convergence of algorithm in most situations. Therefore, we decide to solve the problem in Eq. (15) by ADM algorithm with the solver IALM (Inexact Augmented Lagrangian Multiplier Method) [56]. In order to utilize ADM, we have to rewrite Eq. (15) into another form:

$$\min_{\mathbf{L},\mathbf{S},\mathbf{T}} \|\mathbf{L}\|_* + \lambda_1\|\mathbf{S}\|_1 + \lambda_2\|\mathbf{T}\|_{\mathrm{TV}} \quad \text{s. t. } \mathbf{D} = \mathbf{L} + \mathbf{S}\ \ \mathbf{T} = \mathbf{S} \tag{16}$$

The augmented Lagrangian function of Eq. (16) is given by:

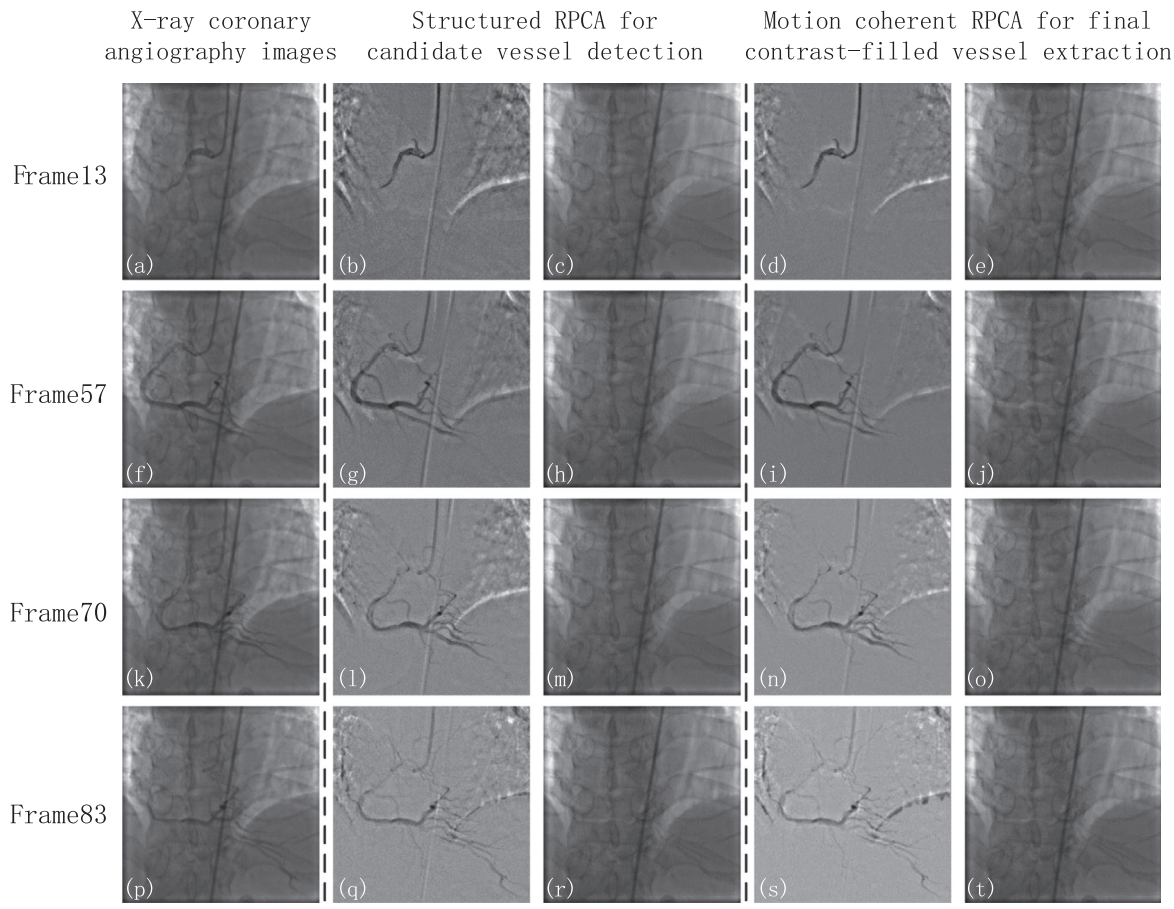

Fig. 1. Extracting contrast-filled vessels from right coronary angiography at four different phases: the left column is original images of different frames; the middle two columns are the foreground candidates and the corresponding background; the right two columns are the final vessel extraction and the corresponding background. (a)–(e) for frame 13 in the first row correspond to the original image **D**, the intermediate foreground/background separation results by statistically structured MoG-RPCA, and the final foreground vessel extraction and the corresponding background by the motion coherency regularized RPCA. Similarly, (f)–(j) in the second row, (k)–(o) in the third row, and (p)–(t) in the last row correspond to the original **D**, the intermediate foreground/background separation results, and the final foreground/background separation results for frame 57, frame 70, and frame 83, respectively.

$$\mathcal{L}(\mathbf{L}, \mathbf{S}, \mathbf{T}, \mathbf{X}, \mathbf{Y}; \mu) = \| \mathbf{L} \|_* + \lambda_1 \| \mathbf{S} \|_1 + \lambda_2 \| \mathbf{T} \|_{\mathrm{TV}} + \frac{\mu}{2} \| \mathbf{D} - \mathbf{L} - \mathbf{S} \|_F^2 + \langle \mathbf{X}, \mathbf{D} - \mathbf{L} - \mathbf{S} \rangle + \frac{\mu}{2} \| \mathbf{T} - \mathbf{S} \|_F^2 + \langle \mathbf{Y}, \mathbf{T} - \mathbf{S} \rangle \tag{17}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are the Lagrangian multipliers, and μ is a positive penalty scalar. We summarize the solutions of the sub-problems based on the ADM strategy in Algorithm 1.

Algorithm 1. ADM for motion coherency regularized RPCA.

Input: Matrix $\mathbf{D} \in \mathbb{R}^{m \times n}$, $\lambda_1 > 0$, $\lambda_2 > 0$

1: Initializing: $\mathbf{L}$, $\mathbf{S}$ and $\mathbf{X}$, $\mathbf{Y}$

2: **while** not converged **do**

3: **L sub-problem**:

$$\mathbf{L}_{k+1} = \arg\min_{\mathbf{L}} \mathcal{L}(\mathbf{L}, \mathbf{S}_k, \mathbf{T}_k, \mathbf{X}_k, \mathbf{Y}_k, \mu) = \arg\min_{\mathbf{L}} \| \mathbf{L} \|_* + \frac{\mu}{2} \| (\mathbf{D} - \mathbf{S}_k + \mu^{-1}\mathbf{X}_k) - \mathbf{L} \|_F^2$$

solved by:

$$(\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}) = \mathrm{svd}(\mathbf{D} - \mathbf{S}_k + \mu^{-1}\mathbf{X}_k) \quad \mathbf{L}_{k+1} = \mathbf{U}\mathcal{S}_{\mu^{-1}}(\mathbf{\Sigma})\mathbf{V}^T$$

4: **S sub-problem**:

$$\mathbf{S}_{k+1} = \arg\min_{\mathbf{S}} \mathcal{L}(\mathbf{L}_{k+1}, \mathbf{S}, \mathbf{T}_k, \mathbf{X}_k, \mathbf{Y}_k, \mu)$$

solved by the operator (20).

5: **T sub-problem**:

$$\mathbf{T}_{k+1} = \arg\min_{\mathbf{T}} \mathcal{L}(\mathbf{L}_{k+1}, \mathbf{S}_{k+1}, \mathbf{T}, \mathbf{X}_k, \mathbf{Y}_k, \mu) = \arg\min_{\mathbf{T}} \lambda_2 \| \mathbf{T} \|_{\mathrm{TV}} + \frac{\mu}{2} \| \mathbf{T} - \mathbf{S}_{k+1} \| + \langle \mathbf{Y}_k, \mathbf{T} - \mathbf{S}_{k+1} \rangle = \arg\min_{\mathbf{T}} \lambda_2 \| \mathbf{T} \|_{\mathrm{TV}} + \frac{\mu}{2} \| \mathbf{T} - \mathbf{S}_{k+1} + \mu^{-1}\mathbf{Y}_k \|_F^2$$

solved by split Bregman method [57].

6: Update Lagrange multiplier $\mathbf{X}$ and $\mathbf{Y}$:

$$\mathbf{X}_{k+1} = \mathbf{X}_k + \mu(\mathbf{D} - \mathbf{L}_{k+1} - \mathbf{S}_{k+1}) \quad \mathbf{Y}_{k+1} = \mathbf{Y}_k + \mu(\mathbf{T}_{k+1} - \mathbf{S}_{k+1})$$

7: **end while**

Output: $(\mathbf{L}_k, \mathbf{S}_k)$

Using the ADM algorithm, we alternately optimize one variable ($\mathbf{L}$, or $\mathbf{S}$, or $\mathbf{T}$) with the other two variables fixed: for $\mathbf{L}$ sub-problem optimization, it has closed-form solution by a soft shrinkage operator $\mathcal{S}_{\mu^{-1}}(y) = \mathrm{sgn}(y)\max(|y| - \mu^{-1}, 0)$ [56] with a threshold μ^{-1} for a scalar y, which operator is extended entry-wisely to vectors and matrices; $\mathbf{S}$ sub-problem is solved by a shrinkage operator with the implementation details in the following section; we adopt split Bregman method [57] to

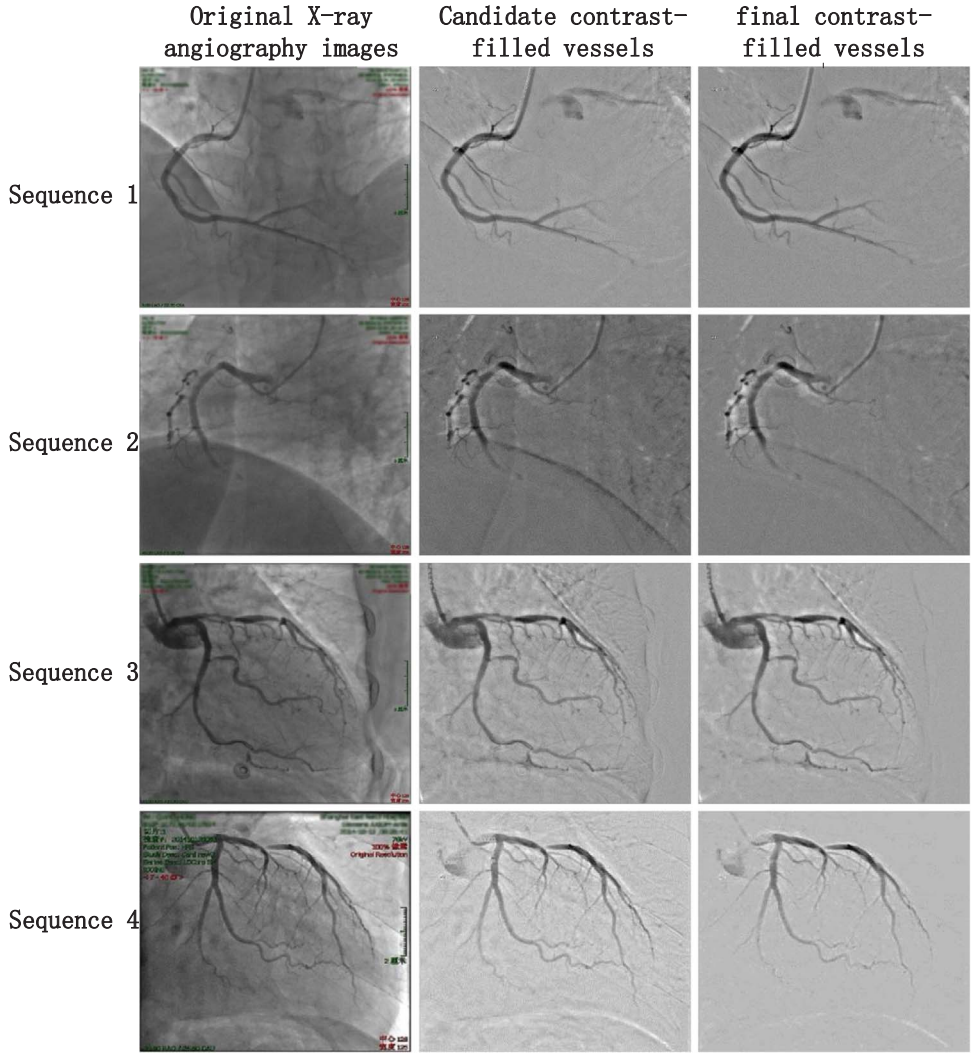

Fig. 2. Vessel extraction results for four different coronary angiography sequences: the original images at the left column; the foreground candidates in the middle column are obtained by the statistically structured MoG-RPCA; the right column corresponds to the final vessel extraction results by the motion coherency regularized RPCA.

solve $\mathbf{T}$ sub-problem optimization. In Algorithm 1, $\lambda_1 = 0.5/\sqrt{\max(m, n)}$ and $\lambda_2 = 0.2/\sqrt{\max(m, n)}$, which can perform well on most experimental data.

For the $\mathbf{S}$ sub-problem optimization in Algorithm 1, the objective function is:

$$\begin{aligned}
&\mathbf{S}_{k+1} = \arg\min_{\mathbf{S}} \mathcal{L}(\mathbf{L}_{k+1}, \mathbf{S}, \mathbf{T}_k, \mathbf{X}_k, \mathbf{Y}_k,\\
&\mu) = \arg\min_{\mathbf{S}} \lambda_1 \| \mathbf{S} \|_{2,1} + \frac{\mu}{2} \| \mathbf{D} - \mathbf{L}_{k+1} - \mathbf{S} \|_F^2 + \langle \mathbf{X}_k,\\
&\mathbf{D} - \mathbf{L}_{k+1} - \mathbf{S}\rangle + \frac{\mu}{2} \| \mathbf{T}_k - \mathbf{S} \|_F^2 + \langle \mathbf{Y}_k,\\
&\mathbf{T}_k - \mathbf{S}\rangle = \arg\min_{\mathbf{S}} \lambda_1 \| \mathbf{S} \|_{2,1} + \mu \| \mathbf{Q} - \mathbf{S} \|_F^2
\end{aligned} \tag{18}$$

where we denote $\mathbf{Q} = \frac{\mathbf{D} - \mathbf{L}_{k+1} + \mathbf{T}_k + \mu^{-1}(\mathbf{X}_k + \mathbf{Y}_k)}{2}$. The objective function can be further expressed as:

$$\mathbf{S}_{k+1} = \arg\min_{\mathbf{S}\in\mathbb{R}^{m\times n}} \sum_{i=1}^{m}\sum_{j=1}^{n} (\lambda_1 |\mathbf{S}_{i,j}| + \mu |\mathbf{Q}_{i,j} - \mathbf{S}_{i,j}|^2) \tag{19}$$

The Eq. (19) can be minimized for each $\mathbf{S}_{i,j}$ separately. The solution $\mathbf{S}_{k+1}$ for $\mathbf{S}$ sub-problem is obtained by the following equation:

$$\mathbf{S}_{k+1} = \mathcal{S}_{\lambda_1(2\mu)^{-1}}(\mathbf{Q}) \tag{20}$$

where the soft shrinkage operator $\mathbf{S}$ is performed element-wisely.

As for the $\mathbf{T}$ sub-problem optimization in Algorithm 1, the objective function is as follows:

$$\begin{aligned}
&\mathbf{T}_{k+1} = \arg\min_{\mathbf{T}} \mathcal{L}(\mathbf{L}_{k+1}, \mathbf{S}_{k+1}, \mathbf{T}, \mathbf{X}_k, \mathbf{Y}_k,\\
&\mu) = \arg\min_{\mathbf{T}} \lambda_2 \| \mathbf{T} \|_{\mathrm{TV}} + \frac{\mu}{2} \| \mathbf{T} - \mathbf{S}_{k+1} \| + \langle \mathbf{Y}_k,\\
&\mathbf{T} - \mathbf{S}_{k+1}\rangle = \arg\min_{\mathbf{T}} \lambda_2 \| \mathbf{T} \|_{\mathrm{TV}} + \frac{\mu}{2} \| \mathbf{T} - (\mathbf{S}_{k+1} - \mu^{-1}\mathbf{Y}_k) \|_F^2\\
&\quad = \arg\min_{\mathbf{T}} \lambda_2 (\|\nabla_x \mathbf{T}\|_1 + \|\nabla_y \mathbf{T}\|_1 + \|\bigtriangledown_z \mathbf{T}\|_1) + \frac{\mu}{2} \| \mathbf{T}\\
&\qquad - (\mathbf{S}_{k+1} - \mu^{-1}\mathbf{Y}_k) \|_F^2
\end{aligned} \tag{21}$$

To apply Bregman splitting method, we first replace $\nabla_x \mathbf{T}$ by d_x, $\nabla_y \mathbf{T}$ by d_y and $\bigtriangledown_z \mathbf{T}$ by d_z to yield the following constrained problem:

$$\begin{aligned}
&\mathbf{T}_{k+1} = \arg\min_{\mathbf{T}} \lambda_2 (\| d_x \|_1 + \| d_y \|_1 + \| d_z \|_1) + \frac{\mu}{2} \| \mathbf{T} - \mathbf{K} \|_F^2 \quad \text{s. t. } d_x\\
&\quad = \nabla_x \mathbf{T}\ d_y = \nabla_y \mathbf{T}\ \mathrm{d}_z = \nabla_z \mathbf{T}
\end{aligned} \tag{22}$$

where $\mathbf{K} = \mathbf{S}_{k+1} - \mu^{-1}\mathbf{Y}_k$. To weakly enforce the constraints in this formulation, we add penalty function terms to obtain the following objective function:

$$\begin{aligned}
\mathbf{T}_{k+1} = \arg\min_{d_x, d_y, d_z, \mathbf{T}} &\lambda_2 (\| d_x \|_1 + \| d_y \|_1 + \| d_z \|_1) + \frac{\mu}{2} \| \mathbf{T} - \mathbf{K} \|_F^2\\
&+ \frac{\lambda_t}{2} \| d_x - \nabla_x \mathbf{T} \|_2^2 + \frac{\lambda_t}{2} \| d_y - \nabla_y \mathbf{T} \|_2^2 + \frac{\lambda_t}{2} \| d_z - \nabla_z \mathbf{T} \|_2^2
\end{aligned} \tag{23}$$

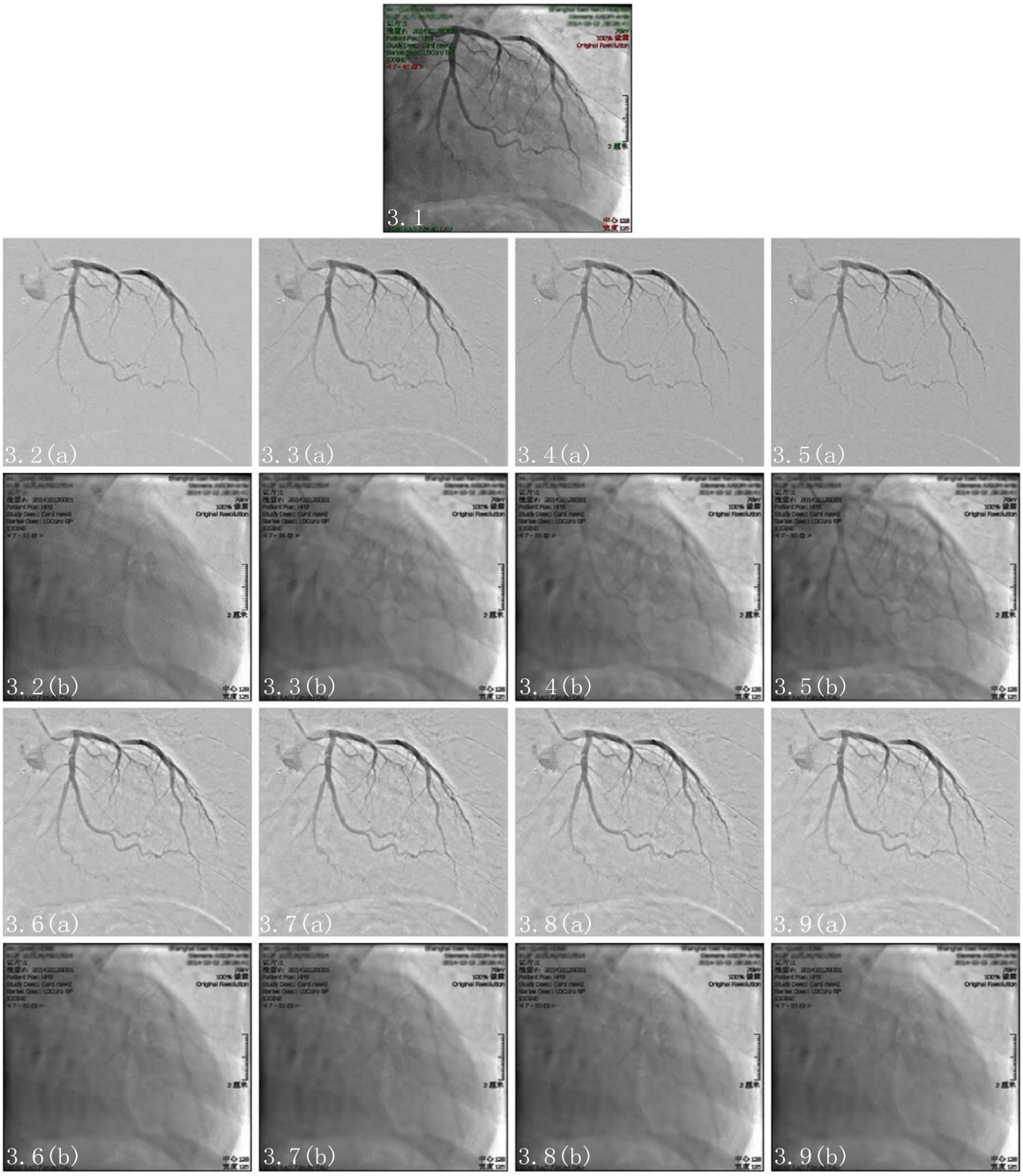

Fig. 3. The image 3.1 is one frame of the left coronary angiographic image sequence. Fig. 3.2(a)–3.2(b) to Fig. 3.9(a)–3.9(b) correspond to the different foreground vessels separated from background by ours, Block-RPCA, DECOLOR, IALM-BLWS, FPCP, GoDec, TFOCS, and PRMF. In the third row, Block-RPCA, DECOLOR, and IALM-BLWS methods all have some foreground vessels remained in the background layer except ours. In the fourth row, the FPCP, GoDec, TFOCS, and PRMF methods all have background residue in the foreground layer.

where λ_t represents the weight of penalty function. Finally, we strictly enforce the constraints by applying the Bregman iteration [57] to get

$$\mathbf{T}_{k+1} = \arg\min_{d_x,d_y,d_z,\mathbf{T}} \lambda_2(\| d_x \|_1 + \| d_y \|_1 + \| d_z \|_1) + \frac{\mu}{2}\| \mathbf{T} - \mathbf{K} \|_F^2 + \frac{\lambda_t}{2}\| d_x - \nabla_x\mathbf{T} - b_x^k \|_2^2 + \frac{\lambda_t}{2}\| d_y - \nabla_y\mathbf{T} - b_y^k \|_2^2 + \frac{\lambda_t}{2}\| d_z - \nabla_z\mathbf{T} - b_z^k \|_2^2 \quad (24)$$

where the proper values of b_x^k , b_y^k and b_z^k are chosen through Bregman iteration.

We can efficiently solve the minimization problem in Eq. (24) by iteratively minimizing it with respect to d and T separately. For the solution of T, this leads to the following subproblem:

$$\mathbf{T}_{k+1} = \arg\min_{\mathbf{T}} \frac{\mu}{2}\| \mathbf{T} - \mathbf{K} \|_F^2 + \frac{\lambda_t}{2}\| d_x^k - \nabla_x\mathbf{T} - b_x^k \|_2^2 + \frac{\lambda_t}{2}\| d_y^k - \nabla_y\mathbf{T} - b_y^k \|_2^2 + \frac{\lambda_t}{2}\| d_z^k - \nabla_z\mathbf{T} - b_z^k \|_2^2 \quad (25)$$

which has the following optimality condition:

$$(\mu I - \lambda_t\Delta)\mathbf{T}_{k+1} = \mu\mathbf{K} + \lambda_t\nabla_x\mathbf{T}(d_x^k - b_x^k) + \lambda_t\nabla_y\mathbf{T}(d_y^k - b_y^k) + \lambda_t\nabla_z\mathbf{T}(d_z^k - b_z^k) \quad (26)$$

In order to achieve optimal efficiency, we must use a fast iterative algorithm to get an approximate solution to the problem in Eq. (26). Due to the problem being strictly diagonally dominant, a natural component-wise solution of Gauss–Seidel method for this problem should be represented as $\mathbf{T}_{i,j}^{k+1} = \mathbf{G}_{i,j}^k$:

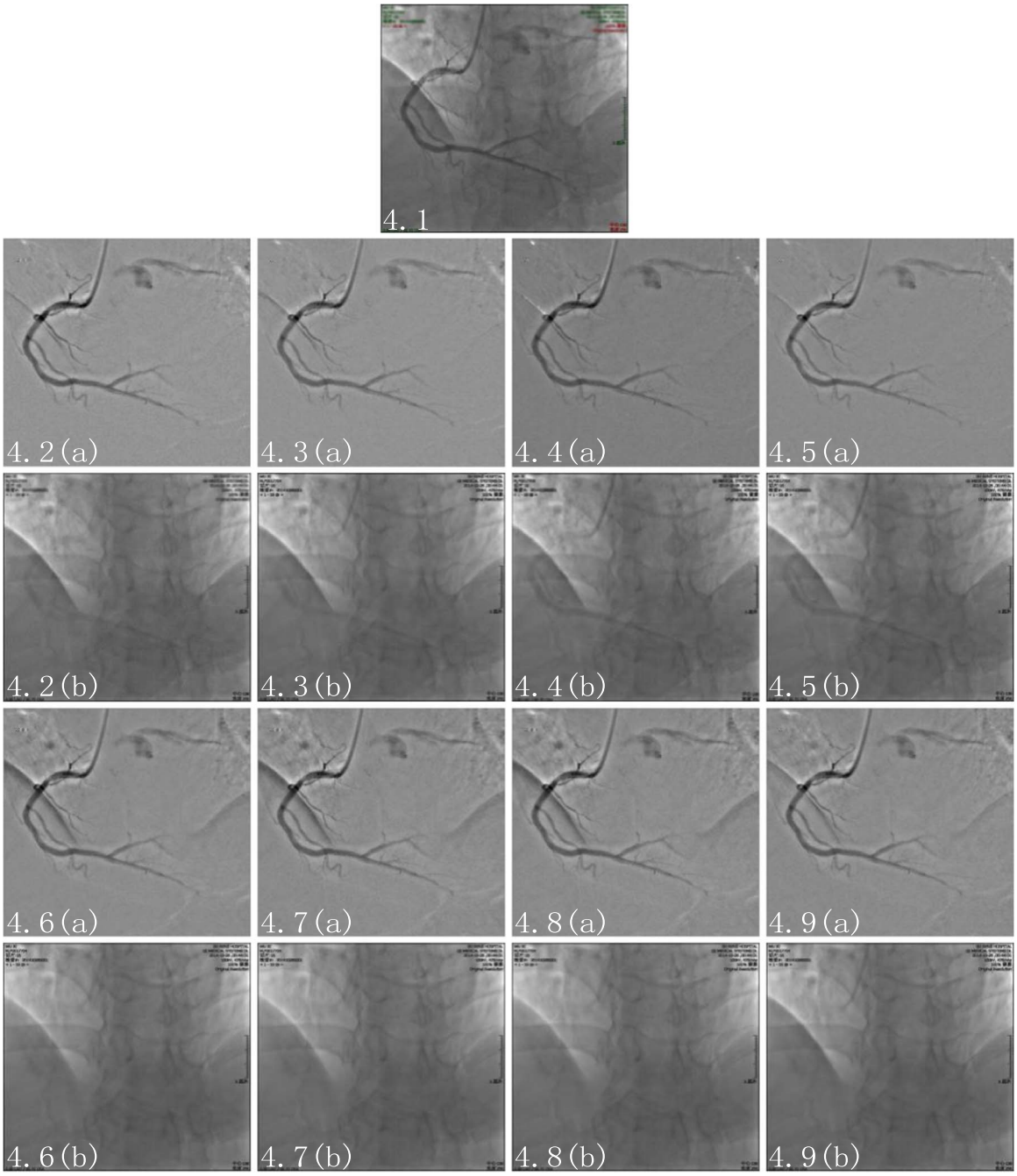

Fig. 4. The image 4.1 is one frame of the right coronary angiographic image sequence. Fig. 4.2(a)–4.2(b) to Fig. 4.9(a)–4.9(b) correspond to the different foreground vessels separated from background by ours, Block-RPCA, DECOLOR, IALM-BLWS, FPCP, GoDec, TFOCS, and PRMF. In the third row, Block-RPCA, DECOLOR, and IALM-BLWS methods all have some foreground vessels remained in the background layer except ours. In the fourth row, the FPCP, GoDec, TFOCS, and PRMF methods all have background residue in the foreground layer.

$$\begin{aligned}\mathbf{G}^k_{i,j,l} = \frac{\lambda_t}{\mu + 6\lambda_t}(&\mathbf{T}^k_{i+1,j,l} + \mathbf{T}^k_{i-1,j,l} + \mathbf{T}^k_{i,j+1,l} + \mathbf{T}^k_{i,j-1,l} + \mathbf{T}^k_{i,j,l+1} + \mathbf{T}^k_{i,j,l-1} \\ &+ d^k_{x,i-1,j,l} - d^k_{x,i,j,l} + d^k_{y,i,j-1,l} - d^k_{y,i,j,l} + d^k_{z,i,j,l-1} - d^k_{z,i,j,l} - b^k_{x,i-1,j,l} \\ &+ b^k_{x,i,j,l} - b^k_{y,i,j-1,l} + b^k_{y,i,j,l} - b^k_{z,i,j,l-1} + b^k_{z,i,j,l}) + \frac{\mu}{\mu + 6\lambda_t}\mathbf{K}_{i,j,l}\end{aligned} \tag{27}$$

where $\mathbf{T}^{k+1}$ has the same meaning as $\mathbf{T}_{k+1}$ in Eqs. (21) and (25), here is for the sake of convenience.

3. Experimental results

In this section, we make real-data experiments using other state-of-arts RPCA-based methods to qualitatively and quantitatively evaluate the performance of our proposed method.[3] All the twelve sequences of XCA used in this work are from Renji Hospital of Shanghai Jiao Tong University. The frame rate of each video is 15 fps and the resolution of each frame is 256 × 256 or 400 × 400 with 8 bits depth. Each sequence represents one cardiac cycle. To make a precise comparative valuations on the performances of contrast-filled vessel extraction for different methods, we should require ground truth at pixel resolution [66,67]. Ideally, XCA images should be labeled a number of times by different experts and the results averaged out. Therefore, a specific contrast-filled vessel was manually segmented separately by three experts for each sequence. After these three binary label images are registered to

[3] http://www.escience.cn/people/bjqin/research.html

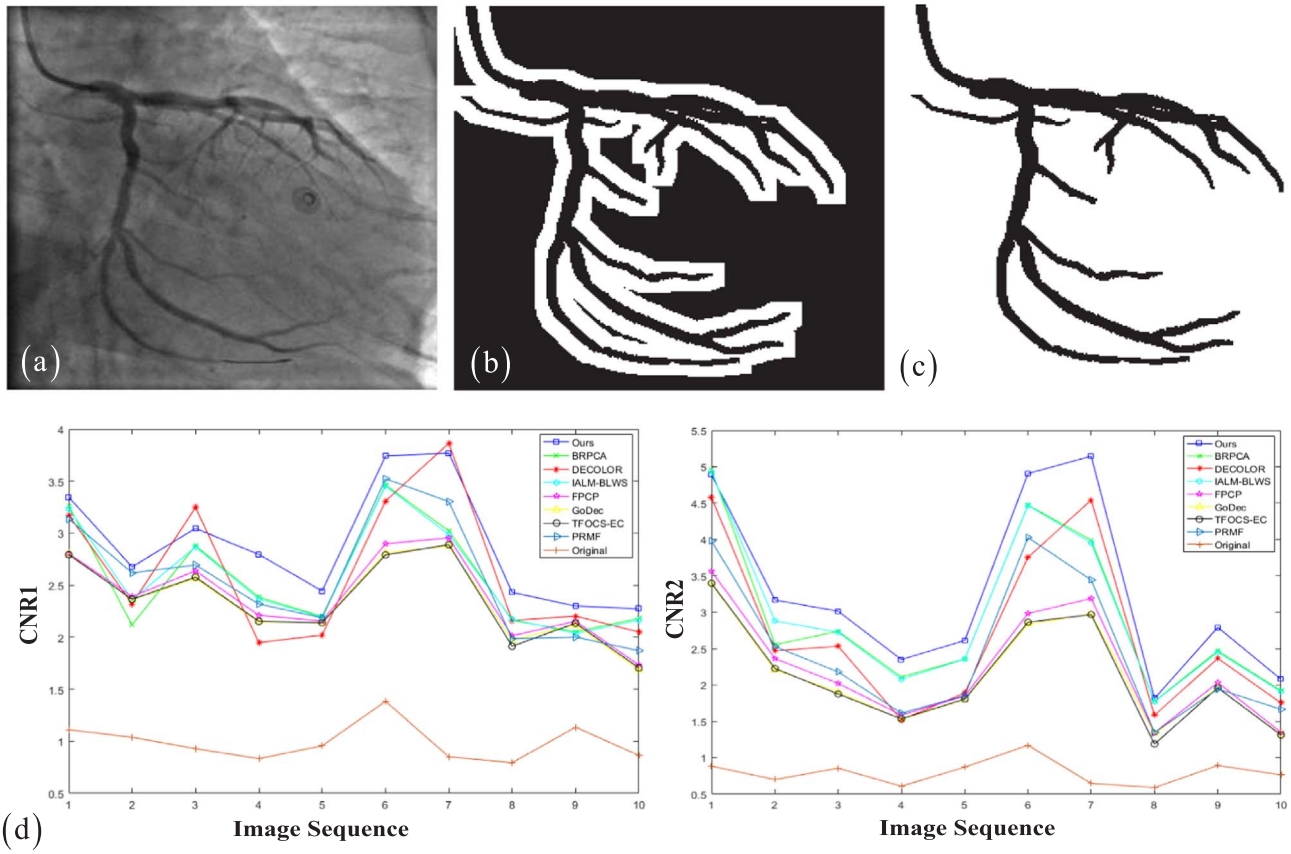

Fig. 5. The average CNR using two types of masks for the XCA sequences. (a) Original XCA image; (b)–(c) Two types of mask with background being defined as the white image region and foreground being the dark area within the white part: (b) is for Mask 1 for the image in (a) and (c) is for Mask 2; (d) CNR1 and CNR2.

get the correspondences of boundary points of the three images by feature-based image registration implemented by ITK software[4] [62], we kept the corresponding boundary points of binary vessels and removed some outlier boundary points by judging the distances from each point of one image to its closest corresponding points in the other two images. The position coordinate values of remained boundary point have been averaged at each point to get final ground truth binary label image of vessel. Besides its importance for ground truth generation in computer vision and image analysis [66,67], introducing an optimization rule to fuse multi-labels for image segmentation has attracted much attention recently in computer vision and image processing [68,69].

We first conduct experiments on several sequences of XCA: left coronary angiography (LCA) and right coronary angiography (RCA) image sequences. The vessel extraction results obtained by the proposed RPCA method are presented in Figs. 1 and 2. In Fig. 1, we select four different phases in the whole right coronary angiographic image sequence with a catheter being inserted into the affected coronary artery and radio-opaque liquid being injected through it: (1) at frame 13, the contrast agent is not distributed in vessels with a catheter being inserted; (2) at frame 57, the contrast agent is filled within vessels completely; (3) at frame 70, the contrast agent is attenuated and some parts of vessels are missing; (4) at frame 83, the contrast agent has completely faded fast with the vessel structures being nearly disappeared. In Fig. 2, we select four coronary angiography images from different coronary angiography sequences. We can find that the vessel detection and the catheter extraction of all the four frames are satisfying, with the complex dynamic background (such as the slow movement of diaphragm, lung and bones) being almost completely removed from the foreground vessel regions.

We also compared the proposed approach with other open source state-of-the-art algorithms: Block-RPCA[5] [3], DECOLOR[6] [42], IALM-BLWS[7] [58], FPCP[8] [59], GoDec[9] [39], TFOCS[10] [60], and PRMF[11] [40]. We download the codes from the authors' websites using the set of default parameters: Block-RPCA uses $\lambda = 1/\sqrt{\max(m,n)}$ in the fist pass RPCA, $\lambda_i = \frac{0.1}{\sqrt{\max(m,n)}} \cdot \frac{\overline{SM}_{\min}}{\overline{SM}_i}$, $\mu_0 = 1.25/\|D\|_2$, $\rho = 1.6$, $J(D) = \max(\|D\|_2, \lambda^{-1}\|D\|_\infty)$ and $\lambda = 0.1/\sqrt{\max(m,n)}$ in the last pass RPCA; DECOLOR method lets $K = \sqrt{n}$, $\beta = 4.5\hat{\sigma}^2$ ($\hat{\sigma}$ is the variance of image) and $\gamma = \beta$ and 5β for simulation and real sequences, respectively; IALM-BLWS sets $\lambda = 1/\sqrt{m}$ with the rank of background matrix being $10\%m$ and the number of corrupted entries in foreground matrix being $10\%m^2$; FPCP makes $\lambda = 0.025$; GoDec sets *rank*=2 with the tolerance $\varepsilon = 10^{-6}$; TFOCS-EC sets $\lambda = 0.01$; PRMF sets *rank*=1, $\lambda_u = \lambda_v = \lambda = 1$, $\tau = 0.01$. In all our experiments, the proposed algorithm simply sets the maximum K as 6 [44] in the MoG-RPCA.

Figs. 3 and 4 show the foreground/background separation results by different methods for LCA and RCA image sequence, respectively. From the qualitative visual evaluation, our method outperforms the other seven methods with clear foreground vessel detection and almost no background. Among the other seven methods, the Block-RPCA, DECOLOR, and IALM-BLWS methods can achieve good foreground vessel detection results (see Fig. 3.3(a)–3.5(a) and Fig. 4.3(a)–4.5(a)). Especially, the foreground vessel detection result obtained by Block-

[4] http://www.itk.org

[5] https://www.ece.nus.edu.sg/stfpage/eleclf
[6] https://fling.seas.upenn.edu/xiaowz/dynamic/wordpress/decolor
[7] http://www.cis.pku.edu.cn/faculty/vision/zlin/zlin.htm
[8] https://sites.google.com/a/istec.net/prodrig/Home/en/pubs
[9] https://sites.google.com/site/godecomposition
[10] http://cvxr.com/tfocs/download/
[11] http://winsty.net/prmf.html

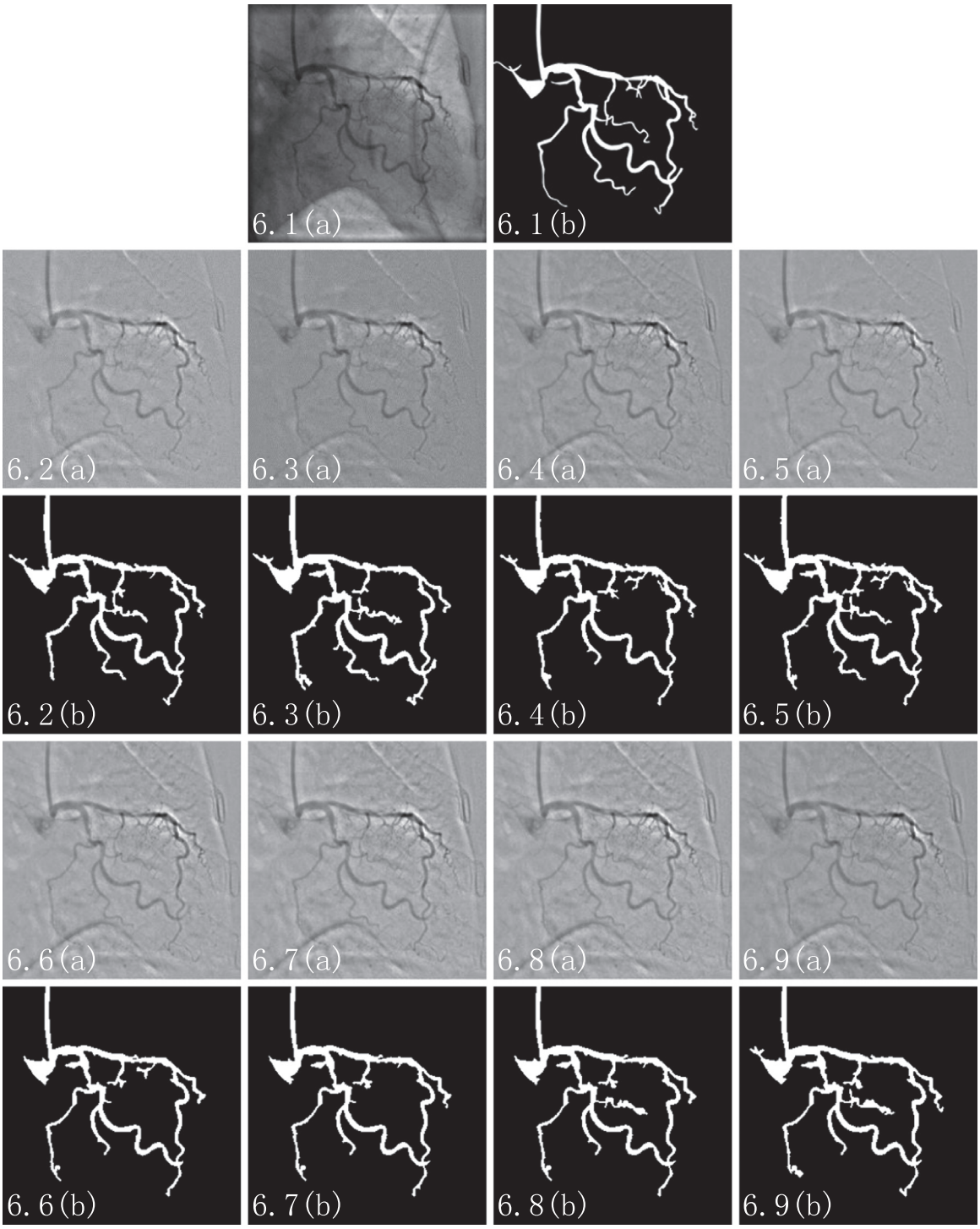

Fig. 6. The image 6.1 is one frame of XCA sequence and its expert-defined foreground vessel. Fig. 6.2(a)–6.2(b) to Fig. 6.9(a)–6.9(b) correspond to the different foreground vessel extractions by different methods and the corresponding segmented vessels by region growing of the seed regions in the vessels.

RPCA is close to our method. Nevertheless, these three methods all have large false negative errors due to their evident foreground vessels remained in the background layer after image decomposition (see Fig. 3.3(b)–3.5(b) and Fig. 4.3(b)–4.5(b)). As for the other four methods FPCP, GoDec, TFOCS-EC, and PRMF, it is obvious to find a lot of background residue in the foreground vessel layer. Due to these obvious false positive errors in the extracted foreground, these four methods achieve poor performance of contrast-filled vessel extraction. Based on this visual evaluation, we find that the proposed method performs the best among all methods.

Only qualitative visual analysis of foreground/background separation in XCA is not convincing, we thus look for appropriate metrics to quantitatively evaluate the performance of vessel extraction. To quantify the visibility of vessels in an image, the contrast-to-noise ratio (CNR) [37] is first used to evaluate the image quality in the experiments of ten XCA sequences. Once the foreground vessel and background of a sequence of XCA are defined, the definition of CNR can be formulated as:

$$\mathrm{CNR} = \frac{|\mu_F - \mu_B|}{\sigma_B} \tag{28}$$

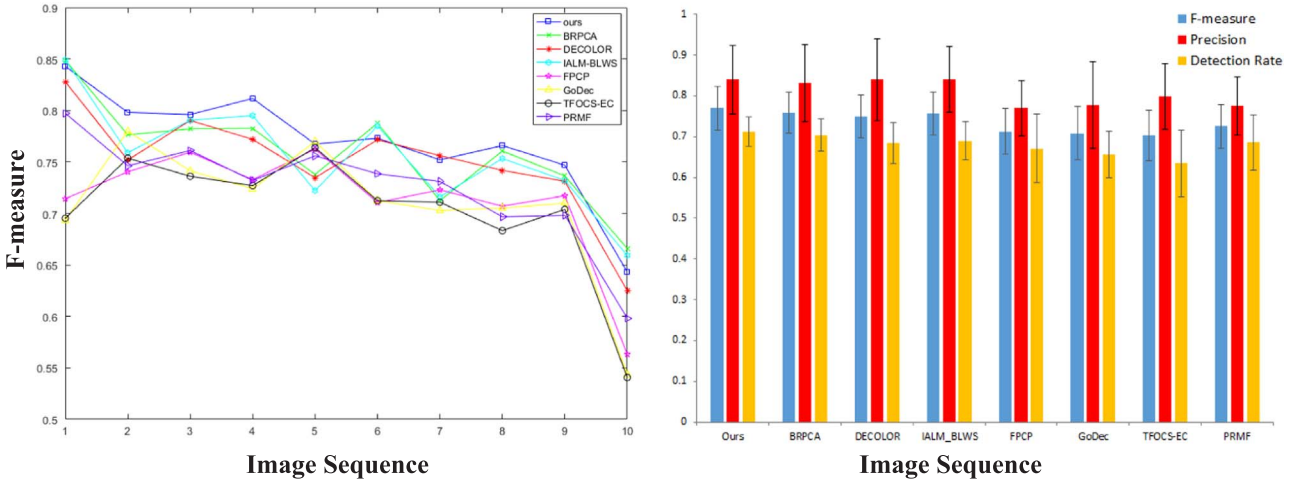

Fig. 7. Different *F*-measures of different image decomposition methods on ten different sequences of XCA. Left: *F*-measures of each sequence for each method; right: mean *F*-measure, detection rate and precision of all image sequences for each method.

Table 1
Performance comparison of eight RPCA methods on the XCA sequences.

Method	Detection Rate	Precision	*F*-measure	CNR1	CNR2	Time
Ours	**0.71258**	0.83955	**0.76976**	**2.88208**	**3.27754**	919.34 s
Block-RPCA	0.70313	0.83132	0.7591	2.57224	2.93497	4 h
DECOLOR	0.68435	0.84038	0.75024	2.62923	2.69954	148.5 s
IALM-BLWS	0.68957	**0.84053**	0.75618	2.58344	2.95259	74.84 s
FPCP	0.67088	0.76977	0.71303	2.39507	2.22859	0.98 s
GoDec	0.656308	0.77700	0.70826	2.34807	2.12574	1.41 s
TFOCS-EC	0.63397	0.79817	0.70273	2.34603	2.11375	304.37 s
PRMF	0.68573	0.77546	0.72556	2.56356	2.45889	42.06 s

where μ_F and μ_B are the mean of foreground and background pixel values, respectively, and σ_B is the standard deviation of the background pixel values. This definition of CNR measures the contrast between the foreground and background pixel intensities in relation to the standard deviation of the background pixel intensities. Larger CNR values imply a better contrast and thus a better performance of foreground/background separation.

We computed two different versions of CNR [37], using two different masks for defining the background in XCA images. For mask 1 in Fig. 5(b), a 7 pixel-wide neighborhood area around the manually segmented vessels (the dark area) is defined as the background (white region surrounding the vessels). This mask can be used to assess the local contrast around vessels in XCA. For mask 2 in Fig. 5(c), the whole regions outside the foreground vessels are considered background to benefit evaluating the removal of the complex dynamic background, such as diaphragm, guiding catheters, etc. The CNR values of the original image sequence and vessel layers resulting from different algorithms are illustrated in Fig. 5(d). The CNR values of our method are higher than those of other methods in general, but not in all sequences. Especially, the proposed method has lower CNR1 than DECOLOR for sequence 3 and 7. The CNR testing indicates that our method has a more robust and better performance than other methods.

We also tested the postprocessing ability of the extracted vessel images. After the expert manually segmented vessels from original image as ground truth (Fig. 6.1(b)), we further segmented foreground vessels by implementing a region-growing algorithm [61,62] on the extracted vessel images for a full quantitative evaluation in Fig. 7. The criteria of detection rate (DR, or called recall), precision (P), and *F*-measure (F) [63] are employed to accurately evaluate the performances of the foreground/background separation methods:

$$\mathrm{DR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \quad \mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \quad \mathrm{F} = \frac{2*\mathrm{DR}*\mathrm{P}}{\mathrm{DR} + \mathrm{P}} \tag{29}$$

where TP is the total number of correctly classified foreground pixels (true positives), FP is the total number of background pixels that are wrongly marked as foreground (false positives), and FN is the total number of foreground pixels that are wrongly marked as background (false negatives). For a certain method, detection rate indicates its power to extract intact foreground with few wholes, precision measures the ratio of true positive foreground vessel to ground truth region and indicates its power to suppress background influence. *F*-measure combines precision and recall to indicate the overall performance of certain extractor. The performance measures shown in Fig. 7 indicate that the proposed method outperforms other methods.

The foreground detection accuracy, background subtraction performance, foreground/background contrast, and computation times, averaged over 50 runs, for all competing methods are reported in Table 1. It is easy to observe that the proposed method achieves more accurate results than the other utilized methods. Considering the computational time, the proposed method runs slower than the other six methods while much faster than the Block-RPCA, FPCA and GoDec methods are the fast methods but have low performances of vessel extraction and foreground/background separation.

4. Discussion and conclusion

Accurately and automatically extracting contrast-filled vessel from XCA has great clinical significance in diagnosis and treatment of cardiovascular disease. To remove the severe disturbance of dynamic background caused by the heart beating and the movement of diaphragm, lung, and bones, we propose a low-rank and sparse matrix decomposition method called graduated RPCA with motion coherency constraint, which fully harnesses the complex structural connectivity

and the spatio-temporal coherency of flowing contrast agents to accurately extract contrast-filled vessels from the complex dynamic background. Specifically, we impose statistically structured constraint on foreground image to detect the candidate contrast-filled vessels, which further are imposed by flowing contrast agent's trajectory group sparsity and TV-based smoothness regularization for fine vessel extraction. This graduated RPCA scheme yields substantially better results than the state-of-the-art RPCA methods. By both qualitative and quantitative evaluation of experiments on twelve real clinical data, the proposed method has demonstrated obvious advantage on the state-of-the-art methods in terms of high accuracy and robustness in extracting contrast-filled vessels from various real XCA.

The automatic contrast-filled vessel extraction from XCA is often greatly hardened by low contrast and the presence of complex background structures as well as human body motion disturbances. The current three main types of vessel extraction methods, i.e., image registration based, image enhancement based and motion layer separation, either are sensitive to the low contrast and/or the complex dynamic background or are very dependent on the accuracy of motion estimation. However, by exploiting the low-rank modeling of complex dynamic background and the sparse-outlier modeling of foreground contrast-filled vessels in XCA, the proposed graduated RPCA based on low-rank and sparse matrix decomposition can accurately and automatically extract the contrast-filled vessel trees with pretty clean background. We believe that some tiny vessel structures in XCA can be extracted well in the context of low-rank and sparse decomposition: though some tiny vessels are not easily detected at a single image due to their faint intensity edges, they can be easily distinguished by observing the tiny structures as moving sparse outliers within the low-rank modeling of complex background in XCA sequence. However, because the perfusion speeds of contrast agents are decreasingly varied according to the flowing blood itself in the artery vessel, small vessel branches, blood capillary and myocardium, the tiny vessel branches in the distal end contain reduced contrast agents which have small motions and relative intensities similar to those of dynamic background. Due to this poor contrast between the tiny vessels and the complex dynamic background, the tiny vessels will be not well kept in the extracted vessels after low-rank and sparse matrix decomposition. However, all tiny vessel structures can be detected by adding structure-aware Gabor filter [64,65] that incorporates the centerline, seed points and directional attributes of the extracted contrast-filled vessels. Although the tiny vessel detection (or segmentation) is hampered by many circumstances and the direct clinical utilization of detected tiny vessels is limited, we consider this issue being a core part of common concern in the image segmentation problem, and believe that the tiny vessel segmentation can indirectly improve various areas of vessel extraction.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (61271320 and 60872102) and Medical Engineering Cross Fund of Shanghai Jiao Tong University (YG2014MS29). The authors would like to thank all authors for opening source codes and ITK software used in the experimental comparison in this work. We are thankful to the anonymous reviewers for their valuable comments that greatly helped to improve this paper.

References

[1] E.J. Cands, X. Li, Y. Ma, J. Wright, Robust principal component analysis?, J. ACM 58 (3) (2011) 11.
[2] T. Bouwmans, E.H. Zahzah, Robust PCA via principal component pursuit: a review for a comparative evaluation in video surveillance, Comput. Vis. Image Underst. 122 (2014) 22–34.
[3] Z. Gao, L. Cheong, Y. Wang, RPCA block-sparse for salient motion detection, IEEE Trans. Pattern Anal. Mach. Intell. 36 (10) (2014) 1975–1987.
[4] C. Dang, H. Radha, RPCA-KFE: key frame extraction for video using robust principal component analysis, IEEE Trans. Image Process. 24 (11) (2015) 3742–3753.
[5] N. He, J.B. Wang, L.L. Zhang, K. Lu, Convex optimization based low-rank matrix decomposition for image restoration, Neurocomputing 172 (2016) 253–261.
[6] H. Gao, J.F. Cai, Z. Shen, H. Zhao, Robust principal component analysis-based four-dimensional computed tomography, Phys. Med. Biol. 56 (11) (2011) 3181–3198.
[7] R. Otazo, E. Cands, D.K. Sodickson, Low-rank plus sparse matrix decomposition for accelerated dynamic MRI with separation of background and dynamic components, Magn. Reson. Med. 73 (2015) 1125–1136.
[8] X. Liu, M. Niethammer, R. Kwitt, N. Singh, M. McCormick, S. Aylward, Low-rank atlas image analysis in the presence of pathologies, IEEE Trans. Med. Imaging 34 (12) (2015) 2583–2591.
[9] P. Kam, X.H. He, S. Yang, X.H. Wu, Iris recognition based on robust principal component analysis, J. Electron. Imaging 23 (6) (2014) 063002.
[10] A.E. Campbell-Washburn, D. Atkinson, Z. Nagy, R.W. Chan, O. Josephs, M.F. Lythgoe, R.J. Ordidge, D.L. Thomas, Using the robust principal component analysis algorithm to remove RF spike artifacts from MR images, Magn. Reson. Med. 75 (2016) 2517–2525.
[11] D.B. Mark, D.S. Berman, M.J. Budoff, J.J. Carr, T.C. Gerber, H.S. Hecht, et al., ACCF/ACR/AHA/NASCI/SAIP/SCAI/SCCT 2010 expert consensus document on coronary computed tomographic angiography: a report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents, J. Am. Coll. Cardiol. 55 (23) (2010) 2663–2699.
[12] T. Ungi, Z. Zimmermann, E. Balazs, A. Lasso, I. Ungi, et al., Vessel masking improves densitometric myocardial perfusion assessment, Int. J. Cardiovasc. Imaging 25 (2009) 229–236.
[13] J. Pu, S. Ding, P. Shan, et al., Comparison of epicardial and myocardial perfusions after primary coronary angioplasty for ST-elevation myocardial infarction in patients under and over 75 years of age, Aging Clin. Exp. Res. 22 (4) (2010) 295–302.
[14] J. Linard, R. Vaillant, Quantitative tool for the assessment of myocardial perfusion during X-ray angiographic procedures, in: Functional Imaging and Modeling of the Heart. Springer Berlin, Heidelberg, 2009, pp. 124–133.
[15] S. Ding, J. Pu, Z. Qiao, et al., TIMI myocardial perfusion frame count: a new method to assess myocardial perfusion and its predictive value for short-term prognosis, Catheter. Cardiovasc. Interv. 75 (5) (2010) 722–732.
[16] S. Petkov, X. Carrillo, P. Radeva, C. Gatta, Diaphragm border detection in coronary X-ray angiographies: new method and applications, Comput. Med. Imaging Graph. 38 (2014) 296–305.
[17] C.B. Compas, T. Syeda-Mahmood, P. McNeillie, et al., Automatic detection of coronary stenosis in X-ray angiography through spatio-temporal tracking, in: 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), 2014, pp. 1299–1302.
[18] M. Panayiotou, A.P. King, R.J. Housden, et al., A statistical method for retrospective cardiac and respiratory motion gating of interventional cardiac x-ray images, Med. Phys. 41 (7) (2014) 071901.
[19] S. Cimen, A. Gooya, M. Grass, et al., Reconstruction of coronary arteries from X-ray angiography: a review, Med. Image Anal. 32 (2016) 46–68.
[20] Y. Khoo, A. Kapoor, Non-iterative rigid 2D/3D point-set registration using semidefinite programming, IEEE Trans. Image Process. 25 (7) (2016) 2956–2970.
[21] C. Kirbas, F. Quek, A review of vessel extraction techniques and algorithms, ACM Comput. Surv. 36 (2) (2004) 81–121.
[22] D. Lesage, E.D. Angelini, I. Bloch, et al., A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes, Med. Image Anal. 13 (6) (2009) 819–845.
[23] Y. Bentoutou, N. Taleb, Automatic extraction of control points for digital subtraction angiography image enhancement, IEEE Trans. Nucl. Sci. 52 (1) (2005) 238–246.
[24] M. Nejati, H. Pourghassem, Multiresolution image registration in digital X-ray angiography with intensity variation modeling, J. Med. Syst. 38 (2) (2014) 1–10.
[25] S. Tang, Y. Wang, Y.W. Chen, Application of ICA to X-ray coronary digital subtraction angiography, Neurocomputing 79 (2012) 168–172.
[26] I. Cruz-Aceves, F. Cervantes-Sanchez, A. Hernandez-Aguirre, et al., A novel Gaussian matched filter based on entropy minimization for automatic segmentation of coronary angiograms, Comput. Electr. Eng. 53 (2016) 263–275.
[27] I. Cruz-Aceves, F. Oloumi, R.M. Rangayyan, et al., Automatic segmentation of coronary arteries using Gabor filters and thresholding based on multiobjective optimization, Biomed. Signal Process. Control 25 (2016) 76–85.
[28] H. Lee, H. Shim, H.J. Chang, Intensity-vesselness Gaussian mixture model (IVGMM) for 2D+ t segmentation of coronary arteries for X-ray angiography image sequences, J. X-ray Sci. Technol. 23 (5) (2015) 579–592.
[29] A. Kerkeni, A. Benabdallah, A. Manzanera, et al., A coronary artery segmentation method based on multiscale analysis and region growing, Computer. Med. Imaging Graph. 48 (2016) 49–61.
[30] Z. Li, Y. Zhang, G. Liu, et al., A robust coronary artery identification and centerline extraction method in angiographies, Biomed. Signal Process. Control 16 (2015) 1–8.
[31] K. Sun, Z. Chen, S. Jiang, Local morphology fitting active contour for automatic vascular segmentation, IEEE Trans. Biomed. Eng. 59 (2) (2012) 464–473.
[32] Y. Chen, J. Yang, Q. Cao, et al., Curve-like structure extraction using minimal path propagation with back-tracing, IEEE Trans. Image Process. 25 (2) (2015) 988–1003.
[33] Y. Zhu, S. Prummer, P. Wang, et al., Dynamic layer separation for coronary DSA and enhancement in fluoroscopic sequences, in: Medical Image Computing and

Computer-Assisted Intervention, MICCAI 2009, 2009, pp. 877–884.
[34] P. Fischer, T. Pohl, T. Kohler, et al., A robust probabilistic model for motion layer separation in X-ray fluoroscopy, in: Information Processing in Medical Imaging, Springer International Publishing, Switzerland, 2015, pp. 288–299.
[35] W. Zhang, H. Ling, S. Prummer, et al., Coronary tree extraction using motion layer separation, in: Medical Image Computing and Computer-Assisted Intervention, CMICCAI 2009, Springer, Berlin, Heidelberg, 2009, pp. 116–123.
[36] P. Fischer, T. Pohl, A. Maier, et al., Surrogate-driven estimation of respiratory motion and layers in X-ray fluoroscopy, in: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Springer International Publishing, Switzerland, 2015, pp. 282–289.
[37] H. Ma, G. Dibildox, J. Banerjee, et al., Layer separation for vessel enhancement in interventional x-ray angiograms using morphological filtering and robust PCA, in: Augmented Environments for Computer-Assisted Interventions, Springer International Publishing, Switzerland, 2015, pp. 104–113.
[38] F. Bach, R. Jenatton, J. Mairal, Structured sparsity through convex optimization, Stat. Sci. 27 (4) (2012) 450–468.
[39] T. Zhou, D. Tao, Godec: Randomized low-rank and sparse matrix decomposition in noisy case, in: Proceedings of the 28th International Conference on Machine Learning, 2011, pp. 33–40.
[40] N. Wang, T. Yao, J. Wang, A probabilistic approach to robust matrix factorization, in: Computer Vision—ECCV 2012, Lecture Notes in Computer Science, vol. 7578, 2012, pp. 126–139.
[41] G. Tang, A. Nehorai, Robust principal component analysis based on low-rank and block-sparse matrix decomposition, in: IEEE 2011 45th Annual Conference on Information Sciences and Systems (CISS), 2011, pp. 1–5.
[42] X. Zhou, C. Yang, W. Yu, Moving object detection by detecting contiguous outliers in the low-rank representation, IEEE Trans. Pattern Anal. Mach. Intell. 35 (3) (2013) 597–610.
[43] Z. Ren, L.T. Chia, D. Rajan, Background subtraction via coherent trajectory decomposition, in: Proceedings of the 21st ACM International Conference on Multimedia, 2013, pp. 545–548.
[44] Q. Zhao, D. Meng, Z. Xu, W. Zuo, L. Zhang, 2014 Robust principal component analysis with complex noise. In: Proceedings of the 31st International Conference on Machine Learning, Beijing, China, 21–26 June 2014. See ⟨http://jmlr.org/proceedings/papers/v32/zhao14.pdf⟩.
[45] T. Bouwmans, A. Sobral, S. Javed, et al., Decomposition into low-rank plus additive matrices for background/foreground separation: a review for a comparative evaluation with a large-scale dataset. arXiv preprint arXiv:1511.01245, 2015.
[46] R. Jenatton, J.Y. Audibert, F. Bach, Structured variable selection with sparsity-inducing norms, arXiv preprint arXiv:0904.3523, 2009.
[47] J. Huang, T. Zhang, D. Metaxas, Learning with structured sparsity, J. Mach. Learn. Res. 12 (2011) 3371–3412.
[48] K. Jia, T.H. Chan, Y. Ma, Robust and practical face recognition via structured sparsity, in: Computer Vision—ECCV 2012, Lecture Notes in Computer Science, vol. 7575, 2012, pp. 331–344.
[49] X. Liu, G. Zhao, J. Yao, C. Qi, Background subtraction based on low-rank and structured sparse decomposition, IEEE Trans. Image Process. 24 (8) (2015) 2502–2514.
[50] X. Cui, J. Huang, S. Zhang, Background subtraction using low rank and group sparsity constraints, in: Computer Vision—ECCV 2012. Springer, Berlin, Heidelberg, 2012, pp. 612–625.
[51] Z. Ren, L-T. Chia, D. Rajan, S. Gao, Background subtraction via coherent trajectory decomposition, in: Proceedings of ACM International Conference on Multimedia, 2013, pp. 545–548.
[52] D. Sun, S. Roth, M.J. Black, A quantitative analysis of current practices in optical flow estimation and the principles behind them, Int. J. Comput. Vis. 106 (2) (2014) 115–137.
[53] X. Cao, L. Yang, X. Guo, Total variation regularized RPCA for irregularly moving object detection under dynamic background, IEEE Trans. Cybern. 46 (4) (2016) 1014–1027.
[54] C.M. Bishop, Approximate inference, in: Pattern Recognition and Machine Learning, 1st ed. Springer-Verlag, New York, NY, USA, 2006.
[55] S.D. Babacan, M. Luessi, R. Molina, A.K. Katsaggelos, Sparse Bayesian methods for low-rank matrix estimation, IEEE Trans. Signal Process. 60 (8) (2012) 3964–3977.
[56] Z.C. Lin, M.M. Chen, L.Q. Wu, Y. Ma, The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices, arXiv preprint arXiv:1009.5055, 2010.
[57] T. Goldstein, S. Osher, The split Bregman method for L_1-regularized problems, SIAM J. Imaging Sci. 2 (2) (2009) 323–343.
[58] Z. Lin, S. Wei, A block Lanczos with warm start technique for accelerating nuclear norm minimization algorithms, arXiv preprint,arXiv:1012.0365
[59] P. Rodriguez, B. Wohlberg, Fast principal component pursuit via alternating minimization, Proceedings of IEEE International Conference on Image Processing (ICIP), Melbourne, Australia, 2013, pp. 69–73.
[60] S. Becker, E.J. Cands, M. Grant, Templates for convex cone problems with applications to sparse signal recovery, Math. Program. Comput. 3 (3) (2011) 165–218.
[61] R. Adams, L. Bischof, Seeded region growing, IEEE Trans. Pattern Anal. Mach. Intell. 16 (6) (1994) 641–647.
[62] L. Ibanez, W. Schroeder, L. Ng, J. Cates, The ITK Software Guide, 2003.
[63] S. Brutzer, B. Hoferlin, G. Heidemann, Evaluation of background subtraction techniques for video surveillance, in Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 1937–1944.
[64] B. Yin, H. Li, B. Sheng, et al., Vessel extraction from non-fluorescein fundus images using orientation-aware detector, Med. Image Anal. 26 (1) (2015) 232–242.
[65] G. Kovcs, A. Hajdu, A self-calibrating approach for the segmentation of retinal vessels by template matching and contour reconstruction, Med. Image Anal. 29 (2016) 24–46.
[66] N. Goyette, P.M. Jodoin, F. Porikli, et al., A novel video dataset for change detection benchmarking, IEEE Trans. Image Process. 23 (11) (2014) 4663–4679.
[67] S. Ruffieux, D. Lalanne, E. Mugellini, et al., Gesture recognition corpora and tools: a scripted ground truthing method, Comput. Vis. Image Underst. 131 (2015) 72–87.
[68] J.E. Iglesias, M.R. Sabuncu, Multi-atlas segmentation of biomedical images: a survey[J], Med. Image Anal. 24 (1) (2015) 205–219.
[69] L. Khelifi, M. Mignotte, A novel fusion approach based on the global consistency criterion to fusing multiple segmentations, IEEE Trans. Syst. Man Cybern.: Syst. PP (99) 1–14, http://dx.doi.org/10.1109/TSMC.2016.2531645

Rong Li received the B.S. degree in Mechanical Engineering from Nanjing University of Science and Technology, in 2014, and He is currently pursuing his Master's Degree in Nanjing University of Science and Technology and Shanghai Jiao Tong University.

Binjie Qin is currently an Associate Professor of Biomedical Engineering with Shanghai Jiao Tong University, Shanghai, China. Over the past years, his research has been focused on biomedical imaging and image analysis for clinical applications. His group has developed image guided surgery system and small animal optical imaging system that are successfully used in clinical applications. His current research interests include biomedical imaging, biomedical image processing, biomedical signal processing, and computer vision.