Medical Image Analysis 38 (2017) 30–49

Contents lists available at ScienceDirect

Medical Image Analysis

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

Low-rank and sparse decomposition based shape model and probabilistic atlas for automatic pathological organ segmentation

Changfa Shi[a,b], Yuanzhi Cheng[b,*], Jinke Wang[b], Yadong Wang[b], Kensaku Mori[c], Shinichi Tamura[d]

[a] Mobile E-business Collaborative Innovation Center of Hunan Province, Hunan University of Commerce, Changsha 410205, China
[b] School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China
[c] Information and Communications Headquarters, Nagoya University, Nagoya 464-8601, Japan
[d] Department of Radiology, Graduate School of Medicine, Osaka University, Suita 565-0871, Japan

ARTICLE INFO

Article history:
Received 18 April 2016
Revised 13 February 2017
Accepted 15 February 2017
Available online 22 February 2017

Keywords:
Pathological organ segmentation
Low-rank and sparse decomposition
Active shape model
Probabilistic atlas
Hierarchical model

ABSTRACT

One major limiting factor that prevents the accurate delineation of human organs has been the presence of severe pathology and pathology affecting organ borders. Overcoming these limitations is exactly what we are concerned in this study. We propose an automatic method for accurate and robust pathological organ segmentation from CT images. The method is grounded in the active shape model (ASM) framework. It leverages techniques from low-rank and sparse decomposition (LRSD) theory to robustly recover a subspace from grossly corrupted data. We first present a population-specific LRSD-based shape prior model, called LRSD-SM, to handle non-Gaussian gross errors caused by weak and misleading appearance cues of large lesions, complex shape variations, and poor adaptation to the finer local details in a unified framework. For the shape model initialization, we introduce a method based on patient-specific LRSD-based probabilistic atlas (PA), called LRSD-PA, to deal with large errors in atlas-to-target registration and low likelihood of the target organ. Furthermore, to make our segmentation framework more efficient and robust against local minima, we develop a hierarchical ASM search strategy. Our method is tested on the SLIVER07 database for liver segmentation competition, and ranks 3rd in all the published state-of-the-art automatic methods. Our method is also evaluated on some pathological organs (pathological liver and right lung) from 95 clinical CT scans and its results are compared with the three closely related methods. The applicability of the proposed method to segmentation of the various pathological organs (including some highly severe cases) is demonstrated with good results on both quantitative and qualitative experimentation; our segmentation algorithm can delineate organ boundaries that reach a level of accuracy comparable with those of human raters.

1. Introduction

The task of segmenting human organs (i.e., the extraction of the target organ's contours from its surroundings) from computed tomography (CT) images plays an essential role in various clinical applications, such as medical diagnosis, surgical planning and treatment evaluation. In current clinical practice, organ segmentation is still typically performed manually by radiologists in a slice-by-slice fashion. Since manual delineation is labor-intensive, and prone to errors due to observer dependence, fully automated organ segmentation techniques are thus extremely valuable in clinical environment.

In the past decade, many sophisticated organ segmentation methods have been developed with varying degrees of success (Heimann et al., 2009; van Rikxoort and van Ginneken, 2013). Of these, model-based methods have proved to be the most effective one, where a prior anatomical knowledge of the target organ is incorporated into the segmentation process. The prior information is mainly represented by two types of statistical models: the statistical shape model (SSM) (Heimann and Meinzer, 2009) which learns the local and global organ shape prior models; and the probabilistic atlas (PA) (Park et al., 2003) which learns the organ's spatial existence probability model. Among various model-based methods so far proposed, active shape model (ASM) (Cootes et al., 1995) has achieved state-of-the-art organ segmentation accuracy (Heimann and Meinzer, 2009), thanks to its ability to combine both low-level image appearance information and high-level shape prior information in a unified framework.

* Corresponding author.
E-mail address: yzcheng@hitwh.edu.cn (Y. Cheng).

http://dx.doi.org/10.1016/j.media.2017.02.008
1361-8415/

Fig. 1. Illustration of challenges in accurate and robust pathological organ segmentation from CT images: a pathological liver (a) and a pathological lung (b) case with large lesions; and (c) existence of imaging artifacts (i.e., metal shadows) in CT images caused by implants.

The aforementioned state-of-the-art organ segmentation approaches work well for certain challenging cases (e.g., large anatomical variations across the population; low contrast and ambiguous boundaries between organs). However, when directly applied to segment pathological organs, which frequently occur in clinical data, these methods are most likely to fail. Because the premises on which these proposed algorithms typically rely on are only applicable to organ segmentation under healthy or minimal pathological conditions rather than under severe pathology. This, however, greatly hampers the usefulness of these computer-based organ segmentation methods in the clinical environment. Specifically, the main challenges of segmenting pathological organs may be outlined as follows:

- *Existence of large pathological lesions*: Fig. 1a shows a 2-D slice image from an abdominal CT scan with large liver tumors, which result in large differences in intensity values from that of normal liver parenchyma. Accurate and robust segmentation of diseased organs with large pathological lesions is especially challenging due to: (1) large deformation of organ shapes caused by big lesions; and (2) complex size and image appearance variability induced by different types of lesions (e.g., hypodense and hyperdense lesions).
- *Presence of pathology affecting organ borders*: Fig. 1b shows another challenging case of a lung CT scan, where the segmentation of lung tissue is heavily challenged by the presence of pathology at organ borders. This is due to the insufficient intensity distinction between the diseased tissue and neighboring tissue with similar appearance (e.g., the fat and muscle), resulting in blurred boundaries between them. And there is a high risk of excluding the peripheral pathological tissue from the final segmentation (i.e., under-segmentation).
- *Existence of other imaging artifacts and noise*: The quality of CT images can be severely deteriorated by imaging artifacts, such as metal shadows caused by implants (Fig. 1c). It is still a challenge to reduce these imaging artifacts in CT images.

Therefore, pathological organs are usually preferred to be delineated manually by clinical experts. However, it is a daunting task for clinicians to manually delineate organs with large lesions, due to pathological organs' highly deformed shape and variable appearance, which goes beyond their prior knowledge of normal anatomy. Thus, it will also suffer from high intra- and inter-operator variability. Although several methods have been proposed specifically for segmenting pathological organs (see Section 2.1 for a detailed review), the results obtained so far remain unsatisfactory under severe pathological conditions. Overcoming the above limitations is exactly what we are concerned within this study.

In this paper, we present an accurate and robust approach in segmenting pathological organs from CT images based on the ASM framework. We first describe a new shape prior model to overcome the limitations of conventional principal component analysis (PCA) (Jolliffe, 2002) based SSM (Cootes et al., 1995) when segmenting pathological organs: (1) sensitivity to non-Gaussian gross errors of input shapes caused by weak and misleading appearance cues of large lesions; (2) having difficulty modeling complex shape variations; and (3) poor adaptation to the finer local details of input shapes. Specifically, inspired by the recently popular low-rank and sparse decomposition (LRSD) theory (Chandrasekaran et al., 2011), also known as Robust PCA (RPCA) (Candès et al., 2011), in the signal processing and computer vision communities, a population-specific LRSD-based shape prior model, called LRSD-SM, is developed. It can tackle the above-mentioned limitations of SSM in a unified framework under severe pathological conditions. LRSD was specifically proposed to overcome the problems of classical PCA when recovering a subspace from grossly corrupted data. Given a data matrix **D**, LRSD decomposes it into two parts: a low-rank component **A** representing the underlying subspace, and a sparse component **E** corresponding to the sparse gross errors. LRSD thus can robustly recover the underlying subspace in the presence of gross errors or outliers, due to the explicit modeling of these errors by the sparse component **E**.

An important prerequisite for ASM-based methods is the shape model initialization. PA-based methods have been employed to automatically initialize the shape model (Okada et al., 2008). However, conventional PA-based methods suffer from problems in the following two situations: (1) The target organ exhibits large anatomical variations in both shape and size, and the PA built with all the training data cannot precisely account for it; (2) The target organ contains large lesions (e.g., tumors), which show very different intensity values from that of normal organ, resulting in large errors in atlas-to-target registration and low likelihood of the target organ. In this paper, we introduce a new PA-based method to circumvent the above-mentioned problems. Specifically, a patient-specific LRSD-based PA method formulated in a statistical Bayesian framework, called LRSD-PA, is developed. It can derive more patient-specific initial shapes and largely eliminate the negative effects of large lesions.

Lastly, in order to make our proposed segmentation framework more efficient and robust against local minima, we develop a hierarchical ASM search strategy, where the shape model deformation proceeds in a global-to-local fashion. To illustrate the applicability of our proposed framework to various pathological organs, we have tested it in two challenging clinical applications involving pathological liver and right lung segmentation from CT images. Extensive evaluations on 95 clinical CT scans of pathological cases (including some highly severe ones) show that our proposed LRSD-SM and LRSD-PA methods achieve significantly better accuracy and robustness than the conventional methods in the presence of severe pathology. Also, the performance of our proposed hierarchical framework is comparable with that of state-of-the-art methods and human raters.

In summary, the main contributions of this paper are as follows:

(1) LRSD-SM is proposed to tackle the limitations of PCA-based SSM in a unified framework under severe pathological conditions (Section 3.2);
(2) An LRSD-PA based shape initialization method is introduced to derive more patient-specific initial shapes and largely eliminate the negative effects of large lesions (Section 3.3);
(3) A hierarchical ASM search strategy is developed to make the segmentation framework more efficient and robust against local minima (Section 3.4);
(4) The proposed segmentation framework is successfully applied to segment the challenging pathological liver and right lung tissue from a total of 95 clinical CT scans (Section 5).

2. Related work

In this section, we briefly review closely related work, and explain how previous work differs from ours.

2.1. Pathological organ segmentation

In the following, we briefly review a few organ segmentation methods that are specifically tailored for the challenging datasets containing pathological organs.

Umetsu et al. (2014) proposed a liver CT segmentation method specifically for the challenging cases with unusual shapes and large lesions. They used sparse representation theory to select the most relevant training datasets for an input case to construct the probabilistic atlas (PA). To make the method robust against large lesions, they also introduced a novel term based on a set of "lesion bases" to account for the differences between pathological cases and normal liver parenchyma. However, their method failed when neighboring organs (e.g., the gallbladder) show similar intensity values to the peripheral lesions, resulting in over-segmentation of the liver tissue in those areas. Dakua et al. (2016) proposed a semi-automated pathological liver segmentation method in low contrast CT images, where the stochastic resonance algorithm in the discrete Gabor transform domain was first employed to enhance the low contrast images, followed by dynamic cellular automaton and level sets based image segmentation. However, an extensive evaluation of their method on more clinical data is needed to fully demonstrate its benefit.

A thorough review of different algorithms for pathological lung segmentation in CT images has been recently presented by Mansoor et al. (2015), including a critical analysis of their feasibility and shortcomings in the case of the most common lung abnormalities. van Rikxoort and van Ginneken (2013) also presented a comprehensive survey on lung CT segmentation methods that are specifically designed for cases containing pathological abnormalities. Wang et al. (2009), van Rikxoort et al. (2009), and Mansoor et al. (2014) developed two-stage approaches for pathological lung segmentation in CT images. In the first stage, a conventional approach was used to perform initial lung segmentation, followed by a pathology recognition system that automatically detected the failure of initial segmentation and the presence of pathology. Then, if necessary, a refinement stage with a more sophisticated algorithm was triggered to derive better results for pathological cases. Their extensive experiments showed that these methods performed much better than the conventional approach at a relatively low increase in computational cost. However, a common problem with these two-stage approaches is that large amounts of training data from healthy subjects is required for the pathology recognition system to generalize well and achieve high accuracy. Sluimer et al. (2005) proposed a PA-based pathological lung segmentation method in CT scans. It avoided the drawbacks of using a single normal reference lung as the atlas, and the initial segmentation was then refined by voxel classification on the border volume. Their method achieved high accuracy when the lung tissue was affected by minimal to moderate pathology, but failed for severely pathological cases. Prasad et al. (2008) used an adaptive thresholding technique guided by the curvature of ribs to segment the lung tissue with diseases. Although their method showed significant improvement over the conventional thresholding-based techniques, it may fail when large high density lesions lie inside the lung parenchyma. Sofka et al. (2011) and Sun et al. (2012) proposed active shape model (ASM) based segmentation methods for patients with lung diseases in CT data. Although their methods delivered promising results, their heuristics-based model initialization method depends heavily on the quality of automatic landmarks detection near lung boundaries (e.g., rib cage and carina), and it is not robust against false detections. Nakagomi et al. (2013) presented an extended graph cuts framework for lung segmentation in CT images, which incorporated multiple shapes and prior information on neighbor structures of the lung. However, their method was mainly targeted at one particular type of abnormality, i.e., cases with pleural effusion. Hua et al. (2011) proposed a graph search based method for lung segmentation in CT images. Their method achieved better accuracy than the intensity-only methods for pathological cases. However, the construction of the graph was restricted to a narrow band around the pre-segmented lung surface obtained from a conventional intensity-based method, which usually failed under severe pathological conditions.

2.2. Low-rank and sparse decomposition for medical applications

Recently, the low-rank and sparse decomposition (LRSD) theory has found many applications in the field of medical imaging computing (Zhou et al., 2014). Deng et al. (2014) proposed an LRSD-based local morphological analysis method by defining the local deformations as the norm of the decomposed sparse component, and applied it to the diagnosis of cirrhosis livers. Hamy et al. (2014) proposed an LRSD-based robust data decomposition registration method for respiratory motion correction in dynamic contrast-enhanced MRI (DCE-MRI), where the registration of DCE-MR time-series was based on the decomposed motion related components, thus confoundings from contrast agent induced local intensity changes are mostly eliminated. Zweng et al. (2015) proposed an LRSD-based automatic guide-wire detection method in fluoroscopic video sequences for neurointerventions. Ma et al. (2015) proposed an LRSD-based layer separation method for vessel enhancement in interventional X-ray angiograms. And both of the two aforementioned methods used the same assumption as in the LRSD-based foreground detection method, that is the nearly static background can be modeled by a low-rank component, while the moving foreground objects constitute sparse outliers (Bouwmans and Zahzah, 2014). However, one of the main drawbacks of these above-mentioned methods is that only the gross errors are explicitly modeled with the sparse component in their implementation of the low-rank image decomposition. Therefore, some small dense Gaussian noise will still be present in the decomposed low-rank component, which can degrade these methods' overall performance.

2.3. Relation to previous work

Our proposed patient-specific LRSD-PA method is inspired by recent papers by Liu et al. (2014) and Liu et al. (2015), who proposed a common framework for atlas-based tissue segmentation and unbiased atlas building in brain MR images by using low-rank and sparse image decomposition to handle patients with large

pathology and deformations. In this paper, we extend their methods in several important ways: (1) In their methods, all the training data was directly employed to build an atlas, and a healthy brain atlas (i.e., the SRI24 atlas (Rohlfing et al., 2010)) instead of the constructed atlas was used as the template for image-to-atlas registrations in tissue segmentation. Whereas we construct new patient-specific liver and lung atlases (i.e., LRSD-PA) based on clustered training data containing large pathology and deformations. Our method not only can derive a more accurate atlas for the target image, but also is much more efficient in both computation time and memory cost. Moreover, we utilize the constructed atlases as the template for the initial shape extraction. (2) Their methods fail when the target image differs considerably from the training data, especially for soft tissue organs (e.g., the liver) with highly complex anatomical variability (as will shown in Section 5.3). Because when the target image has distinctive shape and appearance,its landmarks are mostly assigned to the decomposed sparse component rather than the low-rank component, leading to the failure of registration between its low-rank component and atlas. To resolve this problem, more similar type datasets are needed as suggested by the authors. Whereas we propose a patient-specific LRSD-PA to deal with this problem, where the training data is partitioned into multiple clusters of similar type cases, and the training data for a specific target image comes from the selected best cluster based on the similarity between it and the cluster-specific PA. (3) In the experiments, they only used tissue-class standard deviation (TCSD) as the evaluation metric to assess the accuracy of tissue label alignment between atlas and the target image,rather than comparisons between final tissue labels and ground truth labels, whereas we evaluate our method on both liver and lung datasets with corresponding ground truth tissue labels. (4) In their implementation of the low-rank and sparse image decomposition, only the gross errors are modeled with the sparse component, and the important weight parameters, which determine the amount of information allocated in low-rank and sparse components, were heuristically and manually chosen.

Whereas in our method, small dense Gaussian noise is also explicitly modeled with a separate component, and the weight parameters are set automatically.

3. Methods

Here we propose a novel ASM-based pathological organ segmentation framework in detail. Fig. 2 shows the flowchart of the proposed segmentation framework, including training and testing phases. For the purpose of illustration, we will use the liver tissue as an example of pathological target organ.

In the training phase, we construct the models employed in the hierarchical ASM search: the proposed population-specific LRSD-based shape prior model (LRSD-SM) (Section 3.2); and the appearance model based on normalized gradient profiles (Cootes and Taylor, 2001). In the testing phase, for a given input CT image to be segmented, we utilize the proposed patient-specific LRSD-based PA (LRSD-PA) to obtain an initial organ shape (Section 3.3). Subsequently, during the hierarchical ASM search, the initial shape is iteratively deformed and refined by using the trained appearance and shape prior models (i.e., LRSD-SM) respectively, in a global-to-local fashion (Section 3.4). Essentially, the proposed segmentation framework is based on a research line of model initialization via a multi-atlas segmentation (MAS) approach (Iglesias and Sabuncu, 2015), followed by an ASM refinement step.

3.1. Main notations

Throughout this paper, matrix and its entries are denoted by capital boldface and small symbols, respectively (e.g., x_{ij} is the (i, j)-th entry of the matrix $\mathbf{X}$). A variety of matrix norms are used: $\|\mathbf{X}\|_* = \sum_k \sigma_k$ the nuclear norm, where σ_k denotes the k-th singular value of $\mathbf{X}$; $\|\mathbf{X}\|_0$ the ℓ_0-norm (i.e., the number of non-zero entries in $\mathbf{X}$); $\|\mathbf{X}\|_1 = \sum_{i,j} |x_{ij}|$ the ℓ_1-norm; $\|\mathbf{X}\|_F = \sqrt{\sum_{i,j} x_{ij}^2}$ the Frobenius norm; $\|\mathbf{X}\|_2 = \max_k \sigma_k$ the spectral norm; and $\|\mathbf{X}\|_\infty = \max_{i,j} |x_{ij}|$ the infinity norm.

3.2. Low-rank and sparse decomposition based shape model

Here we describe the procedure to construct our proposed LRSD-SM from training data and optimize the objective function of the resulting model.

3.2.1. Problem formulation

Assume that we have a set of K spatially aligned corresponding triangular mesh liver training shapes $\{\mathcal{M}_i \mid i = 1, 2, \ldots, K\}$, and each shape $\mathcal{M}_i$ is represented by a shape vector $\mathbf{d}_i = (\mathbf{v}_{i1}^T, \mathbf{v}_{i2}^T, \ldots, \mathbf{v}_{in_p}^T)^T$ by concatenating the coordinates of all its vertices, where $\mathbf{v}_{ij} = (x_{ij}, y_{ij}, z_{ij})$ denotes the jth point of shape $\mathcal{M}_i$, and n_p is the number of landmark points. Using these pre-aligned training shapes, we can build a liver shape repository $\mathbf{D_0}$ by stacking all the training shape vectors column-wisely: $\mathbf{D_0} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_K] \in \mathbb{R}^{3N \times K}$, where N is the number of vertices of a shape. Given a pre-aligned input shape to be refined $\mathbf{y} \in \mathbb{R}^{3N}$, the data matrix $\mathbf{D}$ can thus be defined as the concatenation of the shape repository and input shape: $\mathbf{D} = [\mathbf{D_0}, \mathbf{y}] \in \mathbb{R}^{3N \times (K+1)}$.

In the classical statistical shape model (SSM) (Cootes et al., 1995), principal components analysis (PCA) is applied to the anatomical landmarks to learn the principal low-dimensional subspace of plausible liver shapes. Given a pre-aligned input shape to be refined $\mathbf{y} \in \mathbb{R}^{3N}$, it performs a linear basis expansion given r principal modes of shape variation, that is, approximating $\mathbf{y}$ by a linear combination of r principal components. Mathematically, it seeks the best rank-r estimate of $\mathbf{D}$ in a least-squares sense, leading to the following minimization problem:

$$\hat{\mathbf{A}} = \arg\min_{\mathbf{A}} \|\mathbf{D} - \mathbf{A}\|_F \quad \text{s.t.} \quad \text{rank}(\mathbf{A}) \le r, \ \mathbf{D} = \mathbf{A} + \mathbf{Z}. \tag{1}$$

It can be solved analytically via the singular value decomposition (SVD) of the data matrix $\mathbf{D}$. Then the refined input shape $\hat{\mathbf{y}}$ is defined as the last column of the low-rank matrix $\hat{\mathbf{A}}$.

The formulation of classical SSM in Eq. (1) says that the liver shapes are linearly correlated with each other and form a low-rank matrix, and the parts that cannot be fitted in the least squares sense are the shape errors, which implies that it only accounts for small dense Gaussian noise due to the use of ℓ_2-norm. Besides, it can only achieve optimality under the condition that the input data follows a Gaussian distribution. Therefore, classical SSM will have poor performance in cases where the input shapes are corrupted by gross errors or outliers. Unfortunately, such gross segmentation errors are prevalent in the ASM-based local search strategy due to weak and misleading appearance cues and limited search coverage.

In real applications, the gross errors may be arbitrarily large, but the part corrupted by these errors can be considered sparse compared with the whole liver shape. To overcome the shortcomings of classical SSM, we cast the liver shape refinement procedure as a low-rank and sparse decomposition (LRSD) problem (Chandrasekaran et al., 2011), also known as Robust PCA (RPCA) (Candès et al., 2011). Specifically, it decomposes the data matrix $\mathbf{D}$ into two parts: (1) a low-rank component $\mathbf{A}$ corresponding to the global liver shape structure; and (2) a sparse component $\mathbf{E}$ corresponding to the sparse gross errors. Therefore, the LRSD model fits our shape refinement problem quite well. Mathematically, it leads to the following minimization problem:

$$(\hat{\mathbf{A}}, \hat{\mathbf{E}}) = \arg\min_{\mathbf{A},\mathbf{E}} \text{rank}(\mathbf{A}) + \lambda \|\mathbf{E}\|_0 \quad \text{s.t.} \quad \mathbf{D} = \mathbf{A} + \mathbf{E}, \tag{2}$$

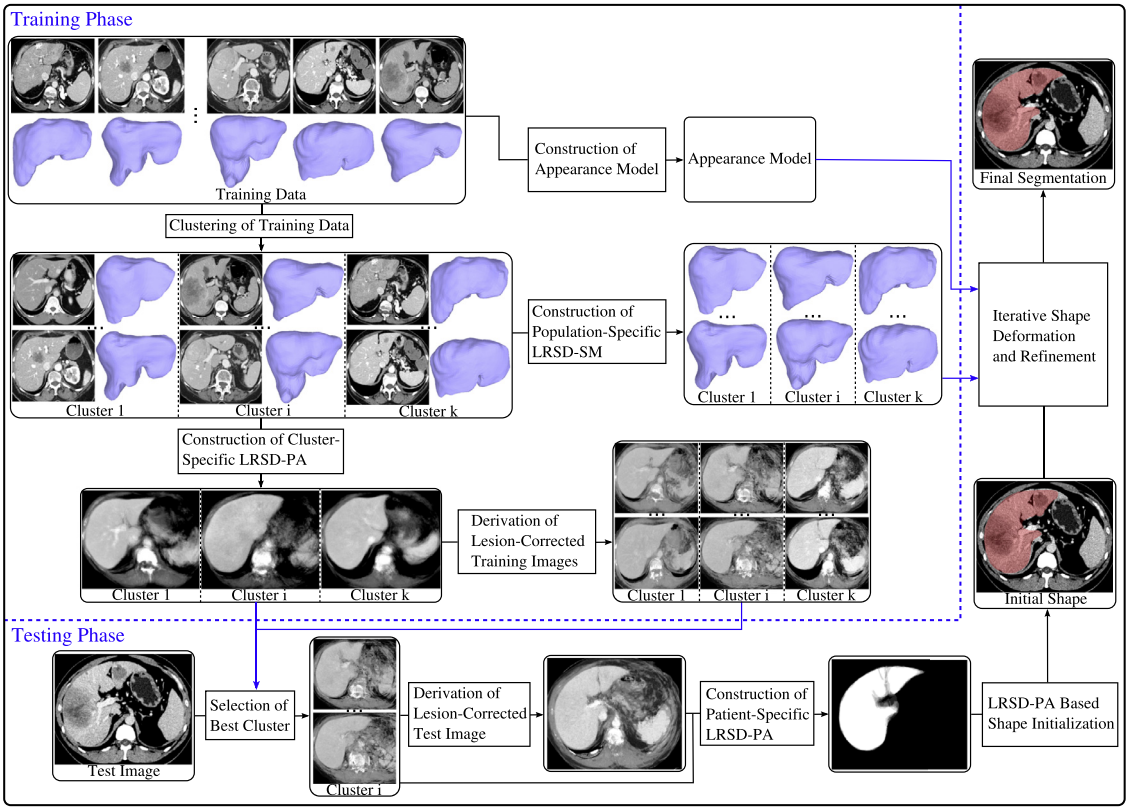


Fig. 2. Flowchart of the proposed framework for pathological organ segmentation, including training and testing phases. Here the liver tissue is used as an example of pathological target organ. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where λ controls the trade-off between the low-rank component $\mathbf{A}$ and the sparse component $\mathbf{E}$. Then the refined input shape $\hat{\mathbf{y}}$ is defined as the last column of the low-rank matrix $\hat{\mathbf{A}}$.

One remaining limitation of the LRSD model in Eq. (2) is its assumption that the input liver shape only has sparse gross errors because of the ℓ_0-norm nature. However, in real cases the input shape is often also corrupted by small dense Gaussian noise, affecting every vertex of the input shape. Therefore, we propose a novel LRSD-based shape model (i.e., LRSD-SM) by extending the LRSD model in Eq. (2) to consider the presence of both sparse gross errors and entry-wise Gaussian noise in the input shape, resulting in the following minimization problem:

$$\begin{aligned}(\hat{\mathbf{A}}, \hat{\mathbf{E}}) &= \arg\min_{\mathbf{A},\mathbf{E}} \ \operatorname{rank}(\mathbf{A}) + \lambda\|\mathbf{E}\|_0 \\ \text{s.t.} \quad &\|\mathbf{D} - \mathbf{A} - \mathbf{E}\|_F \le \delta,\end{aligned} \tag{3}$$

where $\delta > 0$ defines the upper bound of the considered Gaussian noise.

The optimization problem in Eq. (3) is in general intractable and known to be NP-hard due to the non-convexity of the matrix rank and ℓ_0-norm. Recently, it has been proved that solving the following relaxed dual convex optimization problem, named Stable Principal Component Pursuit (SPCP) (Zhou et al., 2010), can achieve the same recovery accuracy:

$$\begin{aligned}(\hat{\mathbf{A}}, \hat{\mathbf{E}}, \hat{\mathbf{Z}}) &= \arg\min_{\mathbf{A},\mathbf{E},\mathbf{Z}} \ \|\mathbf{A}\|_* + \lambda\|\mathbf{E}\|_1 + \gamma\|\mathbf{Z}\|_F^2 \\ \text{s.t.} \quad &\mathbf{D} = \mathbf{A} + \mathbf{E} + \mathbf{Z},\end{aligned} \tag{4}$$

where λ and γ control the trade-off among the three optimized components, and the nuclear norm and ℓ_1-norm are the convex surrogates of the rank and ℓ_0-norm, respectively. Because of the sparsity-inducing property of the ℓ_1-norm, the solution of this optimization problem $\hat{\mathbf{E}}$ will be the sparsest, meaning most entries will be zero. Under some mild conditions on the rank upper bound and noise sparsity level, it has been proven that SPCP can stably recover both $\mathbf{A}$ and $\mathbf{E}$ with high probability and with an error bound proportional to the noise level (Zhou et al., 2010).

Eq. (4) is the optimization problem of our proposed LRSD-SM, through which the refined input shape $\hat{\mathbf{y}}$ lies in the last column of the low-rank matrix $\hat{\mathbf{A}}$, while the segmentation errors are extracted in $\hat{\mathbf{E}}$ and $\hat{\mathbf{Z}}$. The Frobenius norm is employed to model the small dense Gaussian noise, while $\mathbf{E}$ is used to explicitly model the sparse gross errors. Therefore, our proposed LRSD-SM is robust to handle both large sparse shape errors and small dense Gaussian noise. Fig. 3 shows an example of refining an input liver shape $\mathbf{y}$ by using our proposed LRSD-SM.

3.2.2. Optimization framework

Since the introduction of the RPCA model in Candès et al. (2011), various optimization algorithms have been proposed to solve the low-rank and sparse decomposition problem and stable recovery of both $\mathbf{A}$ and $\mathbf{E}$ can be guaranteed (Candès et al., 2011; Zhou et al., 2010). In Zhou et al. (2010), the SPCP problem is solved by using the accelerated proximal gradient (APG) algorithm (Ganesh et al., 2009), which is essentially a variant of the fast iterative shrinkage-thresholding algorithm (FISTA) (Beck and Teboulle, 2009) coupled with a fast continuation technique. In consideration of both efficiency and scalability (in our application), we present the use of alternating direction method of multipliers (ADMM) method (Lin et al., 2011; Boyd et al., 2010). The ADMM method is a popular first-order optimization method that improves the classical augmented Lagrangian method by exploiting the separable structure of the objective function and variable splitting. It has been shown that, compared with the APG algorithm, the ADMM method has the following main advantages (Lin et al., 2011): (1) it performs

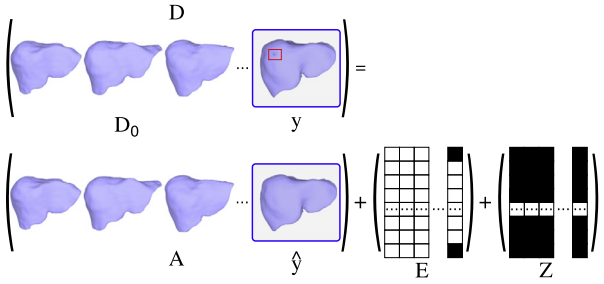


Fig. 3. An example of refining an input liver shape **y** (indicated by the blue rectangle in **D**) as $\hat{\mathbf{y}}$ using our low-rank and sparse decomposition based shape model (LRSD-SM). Filled-in blocks in **E** and **Z** indicate nonzero entries. Therefore, the last column of **E** which corresponds to the input shape **y** is sparse in this case. It demonstrates that the data matrix **D** can be decomposed into a low-rank component **A** corresponding to the global liver shape structure, a sparse component **E** corresponding to the sparse gross errors, and component **Z** corresponding to the small dense Gaussian noise. And **y** contains sparse gross errors (indicated by a red rectangle in **y**). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

more than five times faster, converges with many fewer iterations, and needs less memory space; and (2) it can derive an accurate and exact solution rather than just an approximation.

We follow a research line (Lin et al., 2011) in which the standard ADMM method is employed to solve the LRSD problem. In that work, the standard ADMM method can only handle objective functions containing at most two separable blocks of variables. In our application, however, it cannot be used straightforwardly to solve the SPCP problem in Eq. (4) which involves three blocks of variables (i.e., **A, E**, and **Z**). Therefore, we extend the standard ADMM method to solve the SPCP problem in Eq. (4) with three convex separable objective functions.

Algorithm 1 summarizes our derived ADMM method for solving the SPCP problem of the proposed LRSD-SM in Eq. (4). The detailed mathematical derivation procedure is given in Appendix A. Suppose that the size of the data matrix **D** is $m \times n$ with $m \geq n$. The computational cost of the employed ADMM method is mainly dominated by performing SVD in the **A** subproblem (Eq. (A.7)). Therefore, the computational complexity of ADMM is $O(mn^2)$ per iteration. Generally, the convergence properties of the ADMM method with at most two blocks have been well established (Lin et al., 2011; Boyd et al., 2010). However, for the ADMM method with three blocks, it is still very hard to guarantee its global convergence under general conditions. Fortunately, Hong and Luo (2013) have proved that under some restrictive conditions on the dynamically changing penalty parameter sequence $\{\mu_k\}$, the convergence properties of the derived ADMM method with three blocks can be guaranteed. In our experiments, we also found that the algorithm always converges. Empirically, the ADMM method with three blocks has been applied successfully in existing applications (Peng et al., 2012; Tao and Yuan, 2011; Liu et al., 2013).

Algorithm 1 The ADMM algorithm.

Input: Data matrix $\mathbf{D}$, weighting parameters λ and γ.
Initialize: $\mathbf{A}_0 = \mathbf{E}_0 = \mathbf{Z}_0 = 0, \mathbf{Y}_0 = \frac{\mathbf{D}}{\max(\|\mathbf{D}\|_2, \lambda^{-1}\|\mathbf{D}\|_\infty)}, \mu_0 = \frac{1.25}{\|\mathbf{D}\|_2}$, $\mu_{\max} = 10^9$, $\rho = 1.2$, $\varepsilon = 10^{-7}$, and $k = 0$.

while *not converged* **do**

`// Update` $\mathbf{A}$.
$(\mathbf{U}, \mathbf{\Sigma}, \mathbf{V}) \leftarrow \mathtt{svd}(\mathbf{D} - \mathbf{E}_k - \mathbf{Z}_k + \mu_k^{-1}\mathbf{Y}_k)$,
$\mathbf{A}_{k+1} \leftarrow \mathbf{U}\mathcal{S}_{\mu_k^{-1}}(\mathbf{\Sigma})\mathbf{V}^T$.

`// Update` $\mathbf{E}$.
$\mathbf{E}_{k+1} \leftarrow \mathcal{S}_{\frac{\lambda}{\mu_k}}(\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{Z}_k + \mu_k^{-1}\mathbf{Y}_k)$.

`// Update` $\mathbf{Z}$.
$\mathbf{Z}_{k+1} \leftarrow \frac{\mu_k}{\mu_k + 2\gamma}(\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1} + \mu_k^{-1}\mathbf{Y}_k)$.

`// Check the stopping criteria.`
if $\frac{\|\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1} - \mathbf{Z}_{k+1}\|_F}{\|\mathbf{D}\|_F} < \varepsilon$ **then**
break.
end if

`// Update` $\mathbf{Y}$.
$\mathbf{Y}_{k+1} \leftarrow \mathbf{Y}_k + \mu_k(\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1} - \mathbf{Z}_{k+1})$.

`// Update` μ.
$\mu_{k+1} \leftarrow \mathtt{min}(\rho\mu_k, \mu_{\max})$.

$k \leftarrow k + 1$.

end while

Output: $\mathbf{A} = \mathbf{A}_{k+1}, \mathbf{E} = \mathbf{E}_{k+1}, \mathbf{Z} = \mathbf{Z}_{k+1}$.

3.3. Initial shape extraction using patient-specific low-rank and sparse decomposition based probabilistic atlas

In order to derive more patient-specific initial shapes for our ASM-based segmentation framework, we propose an LRSD-based probabilistic atlas (PA) (i.e., LRSD-PA) for shape model initialization. It has been shown that in atlas-based segmentation, selecting a subset of training data anatomically most similar to the target image can produce a better segmentation result than that of using the full training data (Aljabar et al., 2009; Iglesias and Sabuncu, 2015). This is especially true for soft tissue organs (e.g., the liver) that exhibit high anatomical variability (Oda et al., 2011; Chu et al., 2013). Also, the computational cost will decrease drastically when only a small subset of training data is used to learn the PA instead of the full training data. Therefore, we partition the training data into multiple clusters of similar type cases and separately construct a cluster-specific PA for each cluster. In the initial shape extraction procedure, we first select the best cluster based on the similarity between the cluster-specific PA and the target image, and generate a patient-specific PA from all the training data within the selected cluster. Then we use a statistical Bayesian framework to derive the initial shape, and the patient-specific PA provides the prior probability of organ existence.

3.3.1. Clustering of training data using spectral clustering

To partition the training data into multiple clusters of similar type cases, we employ the spectral clustering algorithm (von Luxburg, 2007). It consists of two major steps: (1) aligning training data: Before performing clustering of training data, we need to align them together in terms of intensity images and establish correspondence between liver shapes. Because in the clustering step, we employ a similarity measure between all the training data based on both intensity images and the corresponding shapes; and (2) clustering using the spectral clustering algorithm. The detailed procedures for these two steps are available in the Supplementary Material.

3.3.2. Construction of cluster-specific LRSD-based PA

After subdividing the training data into multiple clusters, we separately construct a cluster-specific PA for each cluster. However, most of the training data employed in this study contains pathological abnormalities. To largely eliminate the undesired effects of these abnormalities on the final constructed PA, we propose an unbiased PA construction method based on low-rank and sparse decomposition, named LRSD-PA. Its formulation is based on the same

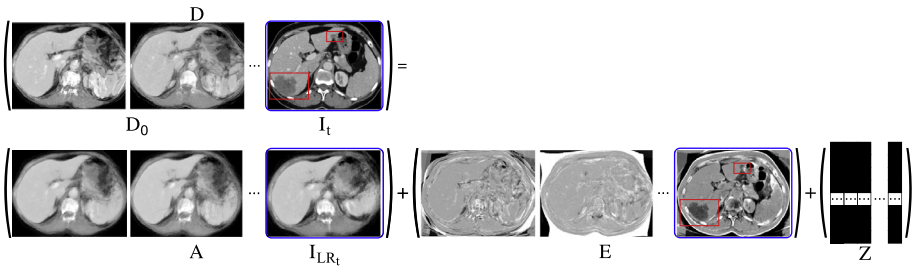


Fig. 4. An example of deriving the lesion-corrected image I_{LR_t} for a target liver image I_t (indicated by the blue rectangle in **D**) using our low-rank and sparse decomposition based probabilistic atlas (LRSD-PA). Filled-in blocks in **Z** indicate nonzero entries. All the columns of **E** are sparse compared with the whole image in this case. It demonstrates that the data matrix **D** can be decomposed into a low-rank component **A** corresponding to the lesion-corrected liver images, a sparse component **E** corresponding to the sparse pathological abnormalities, and component **Z** corresponding to the small dense Gaussian noise. And the target image I_t contains sparse pathological abnormalities (indicated by red rectangles in I_t and **E**). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

two facts as in our proposed LRSD-SM: (1) the aligned liver images are linearly correlated with each other and form a low-rank matrix; and (2) the parts that cannot be modeled by the low-rank part are gross errors or outliers (e.g., pathological abnormalities), which also can be considered sparse compared with the whole image. Also, through LRSD-PA, we can derive lesion-corrected training data for each cluster. Specifically, the vectorized voxel intensity values of each aligned liver image form a column vector of the data matrix **D**, which is decomposed into three parts according to Eq. (4): (1) a low-rank component **A** corresponding to the lesion-corrected liver images; (2) a sparse component **E** corresponding to the sparse pathological abnormalities; and (3) component **Z** corresponding to the small dense Gaussian noise. Then the lesion-corrected liver images in the final low-rank component **A** are employed to construct cluster-specific PA. The iterative procedure for constructing cluster-specific PA using our proposed LRSD-PA and deriving lesion-corrected training data for each cluster is summarized in Algorithm 2. Throughout this paper, the B-splines based free-form deformation (FFD) model (Rueckert et al., 1999) is employed to perform all the image registration. Note that in order to increase the robustness against large lesions, we firstly perform registrations between input images and the chosen initial template to align them to a common coordinate system as in Liu's method (Liu et al., 2014; Liu et al., 2015). Otherwise, if we directly apply LRSD on unaligned input liver images with large anatomical variability to compute the low-rank images, followed by a registration to the low-rank images, then the derived low-rank images will deviate considerably from their corresponding input images, resulting in blurry and anatomically unrepresentative atlases.

3.3.3. Construction of patient-specific LRSD-based PA

Due to high anatomical variability of the liver tissue, given a target image I_t to be segmented, we construct a patient-specific LRSD-based PA for it to derive a more accurate initial liver shape. We first select the best training data cluster based on the similarity between the cluster-specific PA and the target image. Specifically, the cluster c is selected when the similarity between the target image I_t and the atlas intensity image $\bar{I}_c'$ that is warped to the space of I_t is the largest. The normalized cross correlation (NCC) is employed as the similarity measure, that is $NCC(\bar{I}_c', I_t)$ is the largest. A patient-specific PA can then be constructed using all the lesion-corrected training data within the selected cluster c. However, most of the testing data in this study contains lesions. To largely remove the negative effects of these lesions on the constructed patient-specific PA and on the final likelihood image, we use the proposed LRSD-PA to derive the lesion-corrected target image I_{LR_t}, as shown in Algorithm 3. Similar to Algorithm 2, we firstly align all the selected training images to the target image I_t before applying LRSD on data matrix **D**. An example of deriving the

Algorithm 2 Cluster-specific probabilistic atlas construction and lesion-corrected training data derivation procedure.

Input: Training data for class c: $\{I_i \mid i = 1, \ldots, N_c\}$,and the maximum number of iterations: $N_{\max}$.

Choose the training sample with the least amount of lesionas the initial average template $\bar{I}^0$.

for $i = 1$ **to** N_c **do**
  Find the nonrigid transformation T_i^0 that aligns I_i to the template $\bar{I}^0$:$I_i^1 \leftarrow T_i^0(I_i)$.
end for

for $j = 1$ **to** $N_{\max}$ **do**
  `// Compute low-rank parts via LRSD on data matrix` **D**.
  $[I_{LR_1}^j, I_{LR_2}^j, \ldots, I_{LR_{N_c}}^j] \leftarrow$ `LRSD(` $\mathbf{D} = [I_1^j, I_2^j, \ldots, I_{N_c}^j]$ `)`.
  `// Derive the new average template using low-rank parts.`
  $\bar{I}^j \leftarrow \frac{1}{N_c} \sum_{k=1}^{N_c} I_{LR_k}^j$.
  for $i = 1$ **to** N_c **and** $j < N_{\max}$ **do**
    `// The low-rank parts` $I_{LR_i}^j$ `are used for registration.`
    Find the nonrigid transformation T_i^j that aligns $I_{LR_i}^j$ to the template $\bar{I}^j$.
    `// Warp` I_i^j `to the space of the template` $\bar{I}^j$ `using` T_i^j.
    $I_i^{(j+1)} \leftarrow T_i^j(I_i^j)$.
  end for
end for

`// Derive the cluster-specific probability atlas.`
$\bar{I}_c \leftarrow \bar{I}^{N_{\max}}$.

`// Warp the training data to the derived probability atlas.`
for $i = 1$ **to** N_c **do**
  Find the nonrigid transformation T_i that aligns I_i to the probability atlas $\bar{I}_c$:$I_i' \leftarrow T_i(I_i)$.
end for

`// Compute low-rank parts via LRSD on data matrix` $\mathbf{D}'$.
$[I'_{LR_1}, I'_{LR_2}, \ldots, I'_{LR_{N_c}}] \leftarrow$ `LRSD(` $\mathbf{D}' = [I_1', I_2', \ldots, I_{N_c}']$ `)`.

`// Derive the lesion-corrected training data for class` c.
for $i = 1$ **to** N_c **do**
  $\hat{I}_i \leftarrow T_i^{-1}(I'_{LR_i})$.
end for

Output: The constructed probability atlas for class c: $\bar{I}_c$,and the lesion-corrected training data for class c: $\{\hat{I}_i \mid i = 1, \ldots, N_c\}$.

Algorithm 3 Lesion-corrected target image derivation procedure.

Input: The lesion-corrected training data within the selected cluster c: $\{\hat{I}_i \mid i = 1, \ldots, N_c\}$, the target image to be segmented: I_t, and the maximum number of iterations: $N_{\max}$.

for $i = 1$ **to** N_c **do**
  Find the nonrigid transformation T_i^0 that aligns $\hat{I}_i$ to the target image I_t: $\hat{I}_i^1 \leftarrow T_i^0(\hat{I}_i)$.
end for
for $j = 1$ **to** $N_{\max}$ **do**
  `// Compute the low-rank target image via LRSD.`
  $I_{LR_t}^j \leftarrow \mathrm{LRSD}(\mathbf{D} = [I_t, \hat{I}_1^j, \hat{I}_2^j, \ldots, \hat{I}_{N_c}^j])$.
  for $i = 1$ **to** N_c **and** $j < N_{\max}$ **do**
    `// The low-rank target image` $I_{LR_t}^j$ `is the new template.`
    Find the nonrigid transformation T_i^j that aligns $\hat{I}_i$ to $I_{LR_t}^j$: $\hat{I}_i^{(j+1)} \leftarrow T_i^j(\hat{I}_i)$.
  end for
end for

`// Derive the lesion-corrected target image.`
$I_{LR_t} \leftarrow I_{LR_t}^{N_{\max}}$.

Output: The lesion-corrected target image I_{LR_t}.

lesion-corrected image for a target liver image I_t by using LRSD-PA is shown in Fig. 4.

After the lesion-corrected target image I_{LR_t} is derived, it is employed as the new target image instead of I_t, and lesions in the original target image I_t are largely removed in I_{LR_t}. Then all the lesion-corrected training data within the selected cluster $c\{(\hat{I}_i, L_i) \mid i = 1, \ldots, N_c\}$ (L_i is the corresponding label image) is nonrigidly warped to the space of the new target image I_{LR_t}. A patient-specific PA $(\bar{I}_t, \bar{L}_t)$ can thus be constructed as the weighted sum of the warped data $\{(\hat{I}_i^{'}, L_i^{'}) \mid i = 1, \ldots, N_c\}$:

$$\bar{I}_t = \frac{\sum_{i=1}^{N_c} \omega_i \hat{I}_i^{'}}{\sum_{i=1}^{N_c} \omega_i}, \quad \bar{L}_t = \frac{\sum_{i=1}^{N_c} \omega_i L_i^{'}}{\sum_{i=1}^{N_c} \omega_i}, \tag{5}$$

where the weights ω_i are defined as the similarity between the warped data $\hat{I}_i^{'}$ and the lesion-corrected target image I_{LR_t}: $\omega_i = NCC(\hat{I}_i^{'}, I_{LR_t})$.

3.3.4. Initial organ shape extraction

Once the patient-specific PA $(\bar{I}_t, \bar{L}_t)$ is constructed, we employ it to derive the initial organ shape based on a statistical Bayesian framework. Assuming that the liver intensity values in a CT volume follow the Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$, where μ is the mean intensity value and σ is the standard deviation. The conditional probability $p(I_{LR_t}(\mathbf{x})|l)$, which equals the target likelihood image $I_{LH_t}(\mathbf{x})$, can be defined as:

$$p(I_{LR_t}(\mathbf{x})|l) = I_{LH_t}(\mathbf{x}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(I_{LR_t}(\mathbf{x}) - \mu)^2}{2\sigma^2}\right). \tag{6}$$

The Gaussian parameters (i.e., μ and σ) are estimated in a region of interest (ROI) via the expectation-maximization (EM) algorithm (Dempster et al., 1977). And the ROI is defined as the region where the atlas label image $\bar{L}_t$ is larger than a threshold value $T_{PA} = 0.9$. Based on a statistical Bayesian framework, we can define the total liver likelihood image $I_{TLH_t}(\mathbf{x})$ as the posterior probability of liver tissue existence in the lesion-corrected target image I_{LR_t}: $p(l|I_{LR_t}(\mathbf{x}))$. It is computed via the Bayes' theorem:

$$I_{TLH_t}(\mathbf{x}) = p(l|I_{LR_t}(\mathbf{x})) = \frac{p(I_{LR_t}(\mathbf{x})|l)\, p(l)}{Z} = \frac{I_{LH_t}(\mathbf{x})\, \bar{L}_t(\mathbf{x})}{Z}, \tag{7}$$

where the prior probability of liver tissue existence $p(l)$ is given by the atlas label image $\bar{L}_t(\mathbf{x})$, and Z is a normalizing constant. The initial liver shape is then extracted by thresholding $I_{TLH_t}(\mathbf{x})$ using Otsu's method (Otsu, 1979), followed by performing opening operator, removing unconnected components, and filling small holes via closing operator. In our proposed ASM-based pathological organ segmentation framework, the derived initial organ shape is employed as the shape initialization.

3.4. Hierarchical pathological organ segmentation framework

It is known that the classical PCA-based SSM has poor adaptation to the finer local details of input shapes (Heimann and Meinzer, 2009; Zhang et al., 2012). This is because in the final constructed SSM, these important local details will be smoothed out even when they are present in the training samples but not statistically significant. It is especially the case when modeling soft tissue organs (e.g., the liver) that exhibit high shape variability. To alleviate this problem, we propose to partition the training samples into different subgroups that belong to the same population using the identical spectral clustering procedure as described in Section 3.3.1, and then we separately construct a population-specific LRSD-SM shape prior model for each subgroup. Since the shape subspace spanned by samples from a subgroup is much more compact than that of using all training samples, the resulting shape models can thus recover the finer local details more accurately and handle complex shape variations (Shi and Shen, 2008; Foruzan et al., 2014).

After the construction of population-specific LRSD-SM, we now integrate it into our proposed ASM-based pathological organ segmentation framework. In order to make the segmentation framework more efficient and robust against local minima, we propose a hierarchical ASM search strategy, where the shape model deformation proceeds in a global-to-local fashion. The proposed ASM search strategy alternatively iterates between two procedures until convergence as in Cootes et al. (1995): firstly the initialized liver shape is deformed by searching for a better position for each landmark point according to the trained appearance model; and secondly, the intermediate deformed shape is refined by projecting onto the learned shape space with the LRSD-SM model, through which both large gross errors and small Gaussian noise caused by misleading appearance cues will be largely eliminated. Specifically, we first generate multi-resolution image pyramids (Burt, 1981) for all training images, and match population-specific LRSD-SM with different number of shape subgroups to images at different resolutions. Then we separately construct appearance models based on normalized gradient profiles (Cootes and Taylor, 2001) for images at different resolutions. For a given input CT image to be segmented, we also construct a multi-resolution image pyramid for it. At the start of hierarchical ASM search (i.e., level $l = 0$ of the image pyramid), the LRSD-SM based on all training samples is used to guide the initial segmentation. Then as the level l of the image pyramid increases, the LRSD-SM that only employs training samples within a selected shape subgroup c (the deformed shape belongs to) is utilized to further refine the segmentation results. The shape subgroup c is selected from a total of $(l+1)$ shape subgroups according to the shape similarity between the deformed shape and each subgroup. The shape model is thus deformed in a global-to-local fashion, and more accurate recovery

Algorithm 4 Hierarchical pathological organ segmentation framework.

Input: An input CT image: $\mathcal{V}$,the number of resolution levels: $L_{\max}$,and the number of iterations per resolution level: $N'_{\max}$.
$\mathcal{M}_0 \leftarrow$ LRSD-PA based shape model initialization.
$\{\mathcal{V}_0, \mathcal{V}_1, \ldots, \mathcal{V}_{(L_{\max}-1)} = \mathcal{V}\} \leftarrow$ Construction of a multi-resolution image pyramid for input image $\mathcal{V}$.

```
for l = 0 to (L_max − 1) do
  for i = 0 to (N'_max − 1) do
    for all model points v_j ∈ M_l do
      // Shape deformation procedure.
      t_j ← Search for a better position for v_j in V_l using the appearance model trained at level l.
      v_j ← t_j.
    end for
    // Shape refinement procedure.
    M_l ← Refine the intermediate deformed shape usingLRSD-SM that employs training sampleswithin the selected shape subgroup c from (l + 1) shape subgroupsanatomically most similar to the deformed shape.
  end for
end for
```

Output: Organ segmentation result $\mathcal{M}_{(L_{\max}-1)}$.

of local detailed shape information can be achieved. The hierarchical ASM search begins at the coarsest resolution input image (i.e., level $l = 0$ of the image pyramid) using the models trained at that level, and then switches to a finer resolution input image after convergence or a predefined number of iterations (Cootes and Taylor, 2001). This procedure is repeated until convergence criterion is met on the original input image in the pyramid. Note that all the training images used to build the appearance models and the input images are the originally acquired CT scans rather than the lesion-corrected low-rank images. Algorithm 4 shows the details of our hierarchical pathological organ segmentation framework.

4. Experiments

In order to evaluate the performance of our proposed method, and to illustrate the applicability of our method to various pathological organs, we have tested it in two challenging clinical applications involving pathological liver and right lung segmentation from a total of 95 clinical CT scans. These two applications differ in the target organ's shape and lesion pattern variability.

4.1. Datasets

4.1.1. Liver datasets

For the application of pathological liver segmentation, we have used a set of 60 portal phase contrast-enhanced abdominal CT images. 30 CT images were from the public database SLIVER07 [1] (Heimann et al., 2009). It consists of 20 training data (SLIVER07-Train) with gold standard, and 10 testing data (SLIVER07-Test) without gold standard. Most cases in this database were pathological, including tumors, metastases and cysts of different sizes (Heimann et al., 2009). Another 15 CT images with gold standard were from the public database 3Dircadb1.[2] Those 15 cases in this database contained pathology, mainly including metastases, hepatocellular carcinoma (HCC) and hemangiomas. The CT scans in these two databases were acquired using a variety of different CT scanners and acquisition protocols. The remaining 15 CT images were from a locally acquired database Non-Public. All 15 cases were from different patients, and the corresponding expert segmentations were manually labeled by our clinical partner. The CT scans were acquired using a 64-detector row and dual-source CT scanner (SOMATOM Definition Flash; Siemens Healthcare, Forchheim, Germany), and all cases were pathological, mainly including metastases, HCC, cysts and hemangiomas.

In this study, the SLIVER07-Train database was used for training liver models and determining the parameter settings for our proposed method. The SLIVER07-Test database was only used for the liver segmentation competition and to compare our method with state-of-the-art methods. Also we combined the 3Dircadb1 database (15 cases) and Non-Public database (15 cases) to form a new database, called Liver-Pathology, which was used for testing the performance of our proposed method. As in Sun et al. (2012) and Umetsu et al. (2014), we used the longest lesion diameter in the axial image plane as a measure to roughly quantify the size of the pathology. The average and maximum diameter in the SLIVER07-Test and Liver-Pathology database were 39.3 mm and 97.7 mm, 44.6 mm and 110.2 mm, respectively. Table 1 lists the details of the liver CT scans in different databases.

4.1.2. Lung datasets

For the application of pathological right lung segmentation, we have employed a set of 35 thoracic CT images. 20 CT images with gold standard were from the public database EMPIRE10 [3] (Murphy et al., 2011). Among them, 10 cases in this database had right lung pathology, mainly including pulmonary nodules, cavities, consolidations, and ground-glass opacity (GGO). The CT scans were acquired using a variety of different CT scanners and acquisition protocols (Murphy et al., 2011). The remaining 15 CT images were from a locally acquired database Lung-Pathology. All 15 cases were from different patients, and the corresponding expert segmentations were provided by our clinical partner. The CT scans were acquired using a 16-detector row CT scanner (Brilliance 16; Philips Medical Systems, Best, The Netherlands), and all cases had diseased right lung, mainly including pulmonary nodules, consolidations, cavities and pulmonary fibrosis. To roughly quantify the severity of the pathology for the right lung tissue, we measured the percentage of voxels within the manual reference segmentations with values larger than −500 HU as in Sluimer et al. (2005). The average and maximum percentage of unhealthy tissue in the Lung-Pathology database were 22.6% and 40.6%, respectively.

In this study, the EMPIRE10 database was used for training right lung models. And the Lung-Pathology database was used for testing the performance of our proposed method. Table 2 gives the details of the lung CT scans in different databases.

4.2. Evaluation metrics and statistical tests

In order to quantitatively assess the accuracy of our proposed method, we compared the segmentation result with the manual reference segmentation according to two volume and surface based metrics: Jaccard index (*JI*) (Jaccard, 1901) and average symmetric surface distance (*ASD*) (Heimann et al., 2009). The *JI* and *ASD* are given in percent and millimeters, respectively. For *JI* (*ASD*), the larger (smaller) the value is, the better the segmentation result.

In all the comparison experiments performed in this study, the paired t-test was conducted to assess whether the difference in segmentation accuracy between our method and other methods was statistically significant, with a significance level set at $p < 0.05$. All statistical hypothesis tests are based on both *JI* and *ASD*

[1] http://www.sliver07.org.
[2] http://www.ircad.fr/softwares/3Dircadb/3Dircadb1/index.php?lng=en.
[3] http://empire10.isi.uu.nl.

Table 1
Specifications of the liver CT scans in different databases.

Database	Number of scans	In-plane matrix size	In-plane resolution [mm]	Number of slices	Slice thickness [mm]
SLIVER07-Train	20	512 × 512	0.58–0.81	64–394	0.7–5.0
SLIVER07-Test	10	512 × 512	0.54–0.87	73–502	0.5–3.0
Liver-Pathology	30	512 × 512	0.56–0.87	74–217	1.0–4.0

Table 2
Specifications of the lung CT scans in different databases.

Database	Number of scans	In-plane matrix size	In-plane resolution [mm]	Number of slices	Slice thickness [mm]
EMPIRE10	20	(276–464)×(158–382)	0.60–0.97	127–537	0.7–2.5
Lung-Pathology	15	512 × 512	0.61–0.83	261–361	1.0

metrics, and the null hypothesis is that no difference exists between our method and other (the compared) methods in mean values of the same evaluation metric. The *t*-statistic, which follows a *t*-distribution under the null hypothesis, is used as the test statistic to calculate the *p*-value and to decide whether to reject the null hypothesis.

4.3. Implementation details

Before processing all the training and testing images, we firstly employed the 3-D anisotropic diffusion filter (Perona and Malik, 1990) to reduce image noise while retaining important edge information. For establishing point correspondences between training shapes, we used the spherical harmonics (SPHARM) method (Styner et al., 2006) and the Minimum Description Length (MDL) method implemented by Heimann et al. (2006) as described in the Supplementary Material. In the testing phase, the establishment of point correspondences only needs to be performed once for the derived initial shapes. We also employed the above-mentioned methods to establish point correspondences between the initial shapes and the mean shape. Since our LRSD-PA based shape initialization method can largely eliminate the negative effects of large lesions, the derived initial shapes thus can robustly align with the mean shape in the presence of pathologies.

In Algorithm 2, the training sample with the least amount of lesion is chosen as the initial average template $\bar{I}^0$ in the following operations: firstly we roughly extract the normal liver parenchyma for each training sample by calculating the liver likelihood image $I_{LH}(\mathbf{x})$ (Eq. (6)) from the masked liver region image via their gold standard, followed by thresholding $I_{LH}(\mathbf{x})$ using Otsu's method (Otsu, 1979); and secondly, we perform an exclusive OR (XOR) operation between the extracted normal liver parenchyma and their gold standard, then the XORed result with the least volume among all training samples is chosen as the initial average template $\bar{I}^0$. The masked training images via liver masks are used to obtain all the nonrigid transformations in Algorithm 2, the resulting transformations are then propagated to the original training images, followed by applying LRSD to the aligned original images. While in Algorithm 3, no liver masks are used. The B-splines based FFD model (Rueckert et al., 1999) implemented in the open source toolbox elastix[4] (Klein et al., 2010) was used to perform all the image registration in this study. The normalized mutual information (NMI) (Studholme et al., 1999) was employed as the similarity measure. For the nonrigid registrations performed in the model initialization module (including Algorithms 2 and 3), the following parameters are used: three levels of a coarse-to-fine B-spline grid resolutions with a control point spacing of 80, 40 and 20 mm, a maximum of 600 iterations of optimization per resolution level, and 32 histogram bins for calculating NMI.

A leave-one-out cross-validation (LOOCV) on the training set (i.e., the SLIVER07-Train database) was used to tune the parameter settings of the proposed hierarchical segmentation framework offline. The objective function for optimizing the parameters is to minimize the *ASD* segmentation errors. The set of parameter values that achieved the best segmentation accuracy in the LOOCV was chosen for test evaluation. The settings of all parameters, which remained the same throughout the experiments, are presented in Table 3.

In the experiments, we compared our proposed LRSD-SM method with two other shape prior models: classical PCA-based SSM (Cootes et al., 1995) and sparse shape composition (SSC) (Zhang et al., 2012). SSC is a sparse representation-based shape prior modeling method. It implicitly incorporates shape priors on-the-fly by refining an input shape as a sparse linear combination of training shapes in a shape repository. We refer the reader to Zhang et al. (2012) for more details. Also, we compared our proposed LRSD-PA based shape initialization method with the conventional PA-based method as in Okada et al. (2008). In our implementation, to avoid bias towards the selected reference image, we employed an iterative procedure to construct the PA. To allow for a fair comparison between different methods, the same hierarchical segmentation framework was utilized, except for the module to be compared (i.e., shape prior modeling or shape initialization). We also fine-tuned the parameter values of all compared methods to obtain the best results via a LOOCV on the training set (the SLIVER07-Train database). The criteria for choosing the best parameter values are the same as in our method.

We implemented all methods in C++[5] on a Linux system, while using ITK (Ibáñez et al., 2005) and VTK (Schroeder et al., 2006) to perform basic image processing operations and 3-D data visualization, respectively. All experiments in this study were conducted on a desktop PC with a 2.67 GHz Intel quad-core processor and 8 GB memory. For the segmentation of each testing image, our method took an average time of about 145 min, and the vast majority of it was spent performing the image registrations in the model initialization module. On average, the time to perform one image registration was about 5 min.

5. Results

Here we first describe the challenging application of pathological liver segmentation in detail; then we present the main results of pathological right lung segmentation, to demonstrate the applicability of our method to segmentation of the various pathological organs.

[4] http://elastix.isi.uu.nl.

[5] The source code is available at https://github.com/ivanshih/LRSD-SM-PA.

Table 3
Parameter values of the proposed hierarchical segmentation framework.

Step	Parameter	Value	Description
Probabilistic Atlas	k	3	Number of target training data clusters for LRSD-PA
	λ	λ_0	Parameter in Eq. (4) for LRSD-PA. $\lambda_0 = 1/\sqrt{\max(m,n)}$ is the default value for λ as suggested in Candès et al. (2011), where (m, n) is the dimension of the data matrix.
	γ	γ_0	Parameter in Eq. (4) for LRSD-PA. $\gamma_0 = 1/(2(\sqrt{m}+\sqrt{n}))$ is the default value for γ as suggested in Candès and Plan (2010), assuming that the standard deviation of the Gaussian noise is $\sigma = 1$.
	$N_{\max}$	3	Maximum number of iterations in Algorithms 2 and 3
Shape Model	n_p	2562	Number of landmark points
	λ	$1.8\lambda_0$	Parameter in Eq. (4) for LRSD-SM
	γ	$50\gamma_0$	Parameter in Eq. (4) for LRSD-SM
Appearance Model	k'	9	Number of sample points in a normalized gradient profile on either side of the landmark point
Search Algorithm	$L_{\max}$	3	Number of resolution levels in the hierarchical ASM search strategy
	$N'_{\max}$	10	Number of iterations per resolution level
	n_s	6	Number of new positions to search on either side of the landmark point during each iteration

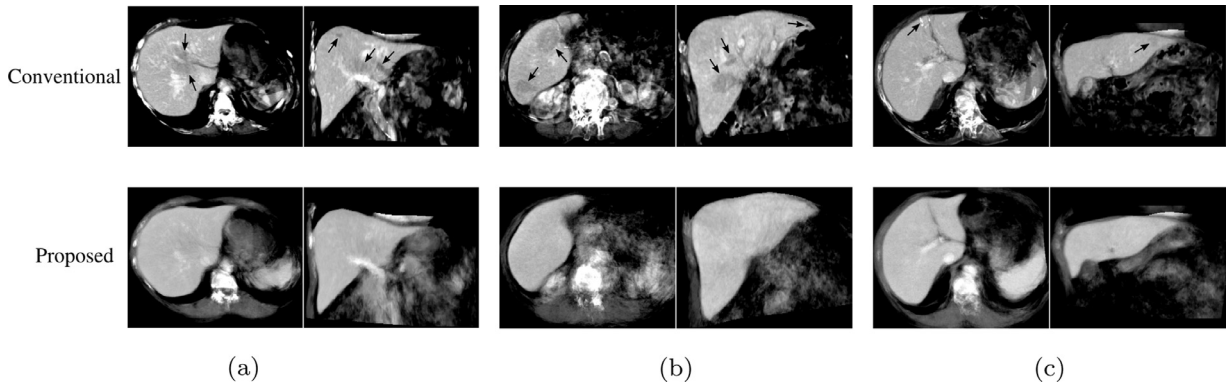


Fig. 5. Comparison of the two different probabilistic atlases (PAs) generated by the conventional PA model (1st row) and the proposed LRSD-PA model (2nd row) for each of three clusters. Each column shows the axial and coronal slices of the PA for each cluster. For the results of the conventional model, regions with heterogeneous appearance induced by the large lesions are highlighted by black arrows.

5.1. Pathological liver segmentation from CT images

5.1.1. Qualitative results

Fig. 5 displays the comparison of the probabilistic atlas (PA) construction for three clusters by using the conventional PA model and the proposed LRSD-PA model. It clearly shows that the PAs of the conventional method have large regions with very different appearance from the normal liver tissue (highlighted by black arrows), suggesting that the conventional method is strongly affected by the present of severe pathology in the training data. By contrast, the liver tissue in our results appears much more homogeneous and sharper, mainly because the use of the lesion-corrected training data leads to much more accurate pairwise image registration. It demonstrates that the negative effects of large lesions are largely eliminated in our method.

Fig. 6 presents the results of the derived lesion-corrected liver images for four challenging pathological cases via our proposed LRSD-PA method, including liver with hypodense lesion, hyperdense lesion, hypo- and hyperdense lesions, and metal shadows caused by implants, respectively. It can be seen that for all these challenging cases, our method successfully separates the sparse lesions or imaging artifacts with different patterns from the derived lesion-corrected images. Also, the appearance of the resulting liver tissue is much more homogeneous. Therefore, in the subsequent processing steps, the negative effects of these large lesions will be largely eliminated.

Fig. 7 depicts the comparative results of initial liver shape extraction by using conventional PA-based method and our proposed LRSD-PA based method in three challenging pathological cases, including liver with hypodense lesion, hyperdense lesion, and hypo- and hyperdense lesions, respectively. It shows that the initial shapes of the conventional method have low overlaps with the ground truth. And the situation becomes even worse as the size of the lesions increase. Because in conventional method, large lesions will lead to large errors in atlas-to-target registration, as can be seen from the conventional PAs of the first and third cases. Also the likelihood of the liver tissue becomes very low in regions with large lesions. By contrast, the initialization of our method is more accurate and much closer to the ground truth, and the superiority of our method becomes even more evident in regions with large lesions. Because in our method, the input images are the derived lesion-corrected pathological images, the negative effects of these lesions that confound the conventional method are thus largely eliminated.

Fig. 8 illustrates the comparative results of liver segmentation by using shape prior models based on SSM, SSC, and our proposed LRSD-SM method in three challenging pathological cases, including liver with hypodense lesion, and hypo- and hyperdense lesions. For the first two cases, all methods can restore a large portion of the liver tissue with lesions, since liver shapes without these regions are not considered to be valid according to these shape models. Both SSM and SSC based methods, however, fail to recover some finer local shape details, illustrating the difficulty in fully capturing the large shape variability of the liver tissue. Due to the small size of training samples, PCA-based SSM tends to over-constrain the shape deformations, thus it generally lacks the flexibility to adapt accurately to these local details. In SSC, the limited number of training samples are directly included in the shape repository to refine an input shape, thus these important local details will be smoothed out in the constructed large shape space. Also, it is known that the ℓ_1-norm based sparse representation method employed in SSC lacks the grouping effect

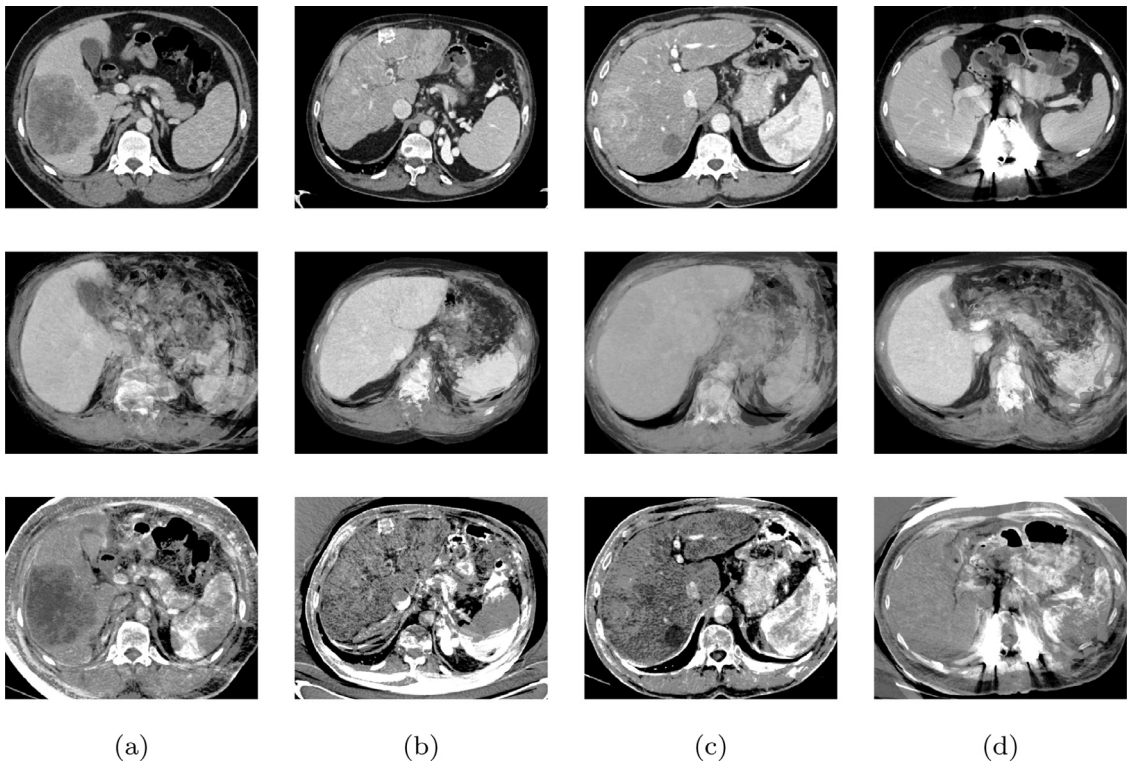


Fig. 6. Results of the derived lesion-corrected liver images for four challenging pathological cases via our proposed LRSD-PA method, including liver with (a) hypodense lesion, (b) hyperdense lesion, (c) hypo- and hyperdense lesions, and (d) metal shadows caused by implants. The first row shows the pathological input images. The second and third rows are their corresponding lesion-corrected images and sparse lesions, respectively.

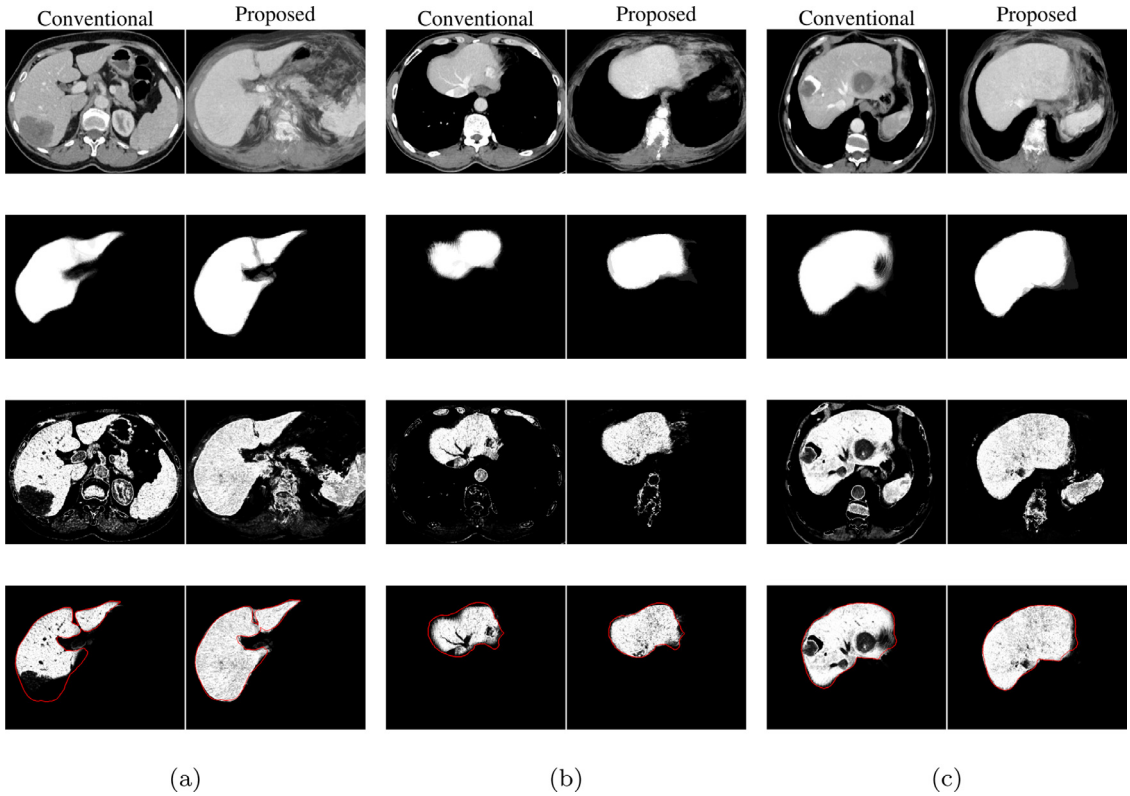


Fig. 7. Comparative results of initial liver shape extraction by using conventional PA-based method and our proposed LRSD-PA based method in three challenging pathological cases, including liver with (a) hypodense lesion, (b) hyperdense lesion, and (c) hypo- and hyperdense lesions. Every column shows one case, representing the conventional and proposed methods, respectively. In each column, the followings are shown sequentially from top to bottom: input image, PA, likelihood image, and the initial shape (i.e., the combination of PA and likelihood image) with the ground truth indicated by red contours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

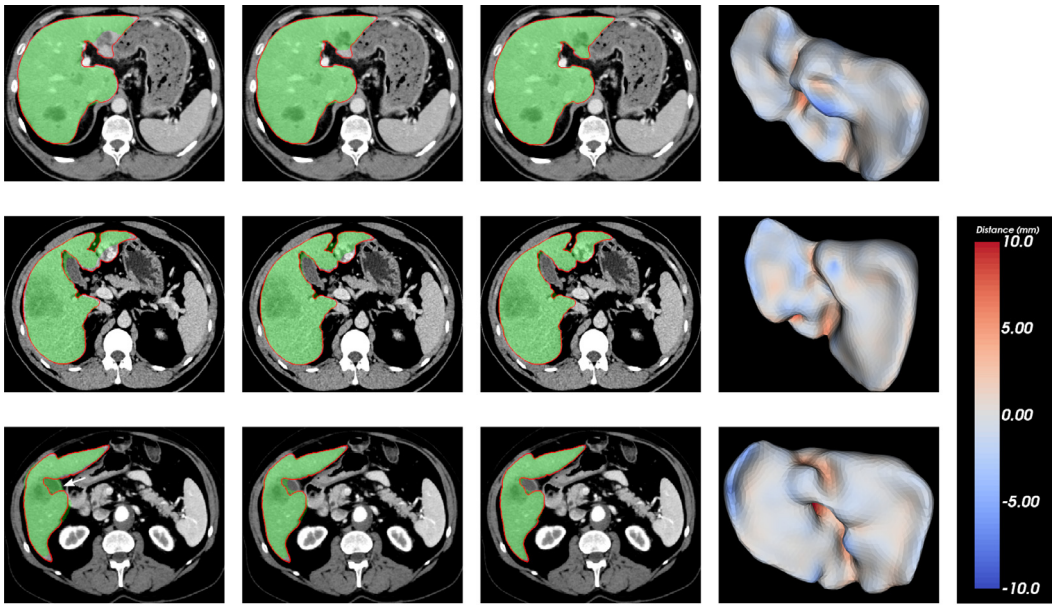


Fig. 8. Comparative results of liver segmentation by using shape prior models based on (a) SSM, (b) SSC, and (c) our proposed LRSD-SM method in three challenging pathological cases, including liver with hypodense lesion (1st and 3rd rows), and hypo- and hyperdense lesions (2nd row). Each row shows one case. The segmentation results are drawn as green regions, and the ground truth is indicated by red contours. (d) The 3-D visualization of average symmetric surface distance (ASD) errors of our method. The red and blue regions indicate over- and under-segmentation, respectively. (e) The distance to color bar. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(Zou and Hastie, 2005), that is when a group of highly correlated liver shapes exists in the shape repository, SSC then tends to select only one from the group, rather than simultaneously selecting all the correlated shapes as a group. On the other hand, our method successfully recovers these finer local details and the delineation is much more accurate. The reason is that, although the proposed model is based on clustered training data with less training samples, it partitions the shape space into multiple subspaces that have much more compact shape distributions. Moreover, the proposed population-specific model is built from training samples within the selected shape subgroup anatomically most similar to the target shape, therefore, it is more specific and detailed than that based on all the training data. Also, in the hierarchical ASM search procedure, our proposed population-specific model is deformed in a global-to-local fashion, thus allowing for more flexibility during shape deformation.

The third case is a particularly challenging one, where the neighboring gallbladder (indicated by a white arrow) exhibits almost similar intensity values to the peripheral hypodense lesion, thus further complicating the problem of accurate delineation. SSM based method, as expected, readily leaks into the neighboring gallbladder and over-segments the live tissue. It is mainly because in the SSM based method, these over-segmented regions roughly fall within the principal subspace spanned by training samples. Also, a previous proposed pathological liver segmentation method by Umetsu et al. (2014) has failed in such situation (see Fig. 16 in their paper), resulting in over-segmentation of the liver tissue in those areas as in SSM. While both SSC and our method successfully separate the neighboring gallbladder from the liver tissue to a large extent, because these over-segmented regions in SSM are explicitly modeled and recovered as the sparse gross errors. Moreover, our method achieves much more accurate result than SSC, thanks to the high flexibility of our proposed population-specific model which is deformed in a global-to-local fashion. Therefore, our method is highly robust against both over- and under-segmentation errors in the presence of severe pathology.

Table 4
Quantitative comparative results of the liver segmentation with three different shape prior modeling methods on the Liver-Pathology database.

Method	JI [%]	ASD [mm]
SSM	$89.43 \pm 2.61^*$	$1.99 \pm 0.92^*$
SSC	$91.68 \pm 1.90^*$	$1.47 \pm 0.42^*$
Our LRSD-SM	$\mathbf{92.66 \pm 1.55}$	$\mathbf{1.27 \pm 0.36}$

For each metric, the mean and standard deviation of the overall datasets are given. Bold values are the best result in that column. * indicates a statistically significant difference between the marked result and the corresponding one of our method at a significance level of 0.05.

5.1.2. Quantitative results

Table 4[6] presents the quantitative comparative results of liver segmentation by using three different shape prior modeling methods on the Liver-Pathology database. It shows that our proposed LRSD-SM achieves the highest accuracy with the smallest variances, and consistently outperforms the other two compared methods, suggesting the robustness of our method across diverse test data. Also, our method yields statistically significant improvements in all metrics over both SSM (paired t-test: $t = 7.03$, $p = 9.97 \times 10^{-8}$ for *JI*; $t = -4.85$, $p = 3.86 \times 10^{-5}$ for *ASD*; $n = 30$) and SSC (paired t-test: $t = 6.84$, $p = 1.64 \times 10^{-7}$ for *JI*; $t = -6.94$, $p = 1.25 \times 10^{-7}$ for *ASD*; $n = 30$). Specifically, the mean *JI* of our method is 92.66%, which is 3.23% and 0.98% higher than SSM and SSC based method, respectively. Regarding *ASD* metric, our method obtains a mean value of 1.27 mm, which achieves an improvement of 13% on average compared to SSC and up to 36% over SSM. These results demonstrate that our proposed method is significantly more accurate and robust than both SSM and SSC in the presence of severe pathology.

[6] The corresponding box plots are available in the Supplementary Material.

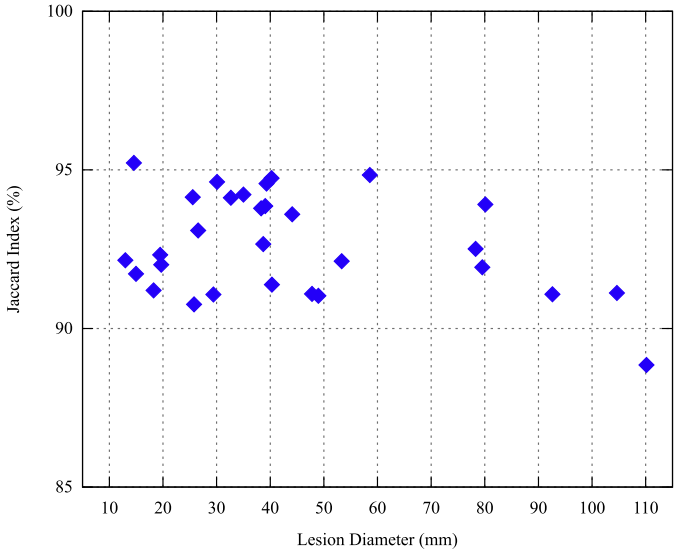


Fig. 9. Influence of liver lesion diameter size on the segmentation accuracy.

Table 5
Quantitative comparative results of the final liver segmentation by using two different model initialization methods on the Liver-Pathology database.

Method	JI [%]	ASD [mm]
Conventional PA	$90.76 \pm 1.67^*$	$1.65 \pm 0.43^*$
Our LRSD-PA	$\mathbf{92.66 \pm 1.55}$	$\mathbf{1.27 \pm 0.36}$

For each metric, the mean and standard deviation of the overall datasets are given. Bold values are the best result in that column. * indicates a statistically significant difference between the marked result and the corresponding one of our method at a significance level of 0.05.

To study the robustness and insensitivity of our method with respect to pathological liver with different lesion diameter size, Fig. 9 shows the relationship between *JI* metric and liver lesion diameter. We can see that the performance of our method remains relatively stable as the size of lesion diameter increases. For all the 30 pathological cases from Liver-Pathology database, our method achieves relatively high accuracy. The *JI* metric of all cases is above 90%, except for the case with the largest lesion, whose diameter is larger than 110 mm. Also the tumor volume is around 50% of the whole liver volume. However, even for this particularly challenging case, with a *JI* of 88.85%, the performance is still considered to be satisfactory. Therefore, our method is quite robust and to a large extent insensitive to the changes relating to the size and pattern of liver lesions.

To investigate the effectiveness of our proposed LRSD-PA based shape initialization method, Table 5[7] provides the comparative results of final liver segmentation by using two different model initialization methods on the Liver-Pathology database. It shows that our method achieves statistically significantly better accuracy (paired t-test: $t = 11.17,\ p = 5.09 \times 10^{-12}$ for *JI*; $t = -8.68,\ p = 1.46 \times 10^{-9}$ for *ASD*; $n = 30$) than the conventional PA-based method according to all metrics. Specifically, the mean *JI* and *ASD* of final segmentation results for the conventional method are 90.76% and 1.65 mm, respectively. And our method outperforms it by 1.9% and up to 23% on average in terms of *JI* and *ASD*, respectively. Therefore, in the presence of severe pathology, our proposed method can derive more patient-specific initial shapes for the hierarchical segmentation framework, which then leads to significantly more accurate final segmentation results than conventional PA-based method.

[7] The corresponding box plots are available in the Supplementary Material.

Table 6
Quantitative evaluation results of the liver shape refinement with and without the Frobenius norm term from the input liver shapes with synthetic noise.

Metrics	Input shape	Without Frobenius norm	With Frobenius norm
JI [%]	$74.17 \pm 5.61^*$	$78.96 \pm 3.82^*$	$\mathbf{79.82 \pm 3.66}$
ASD [mm]	$6.09 \pm 1.52^*$	$4.38 \pm 0.85^*$	$\mathbf{4.12 \pm 0.71}$

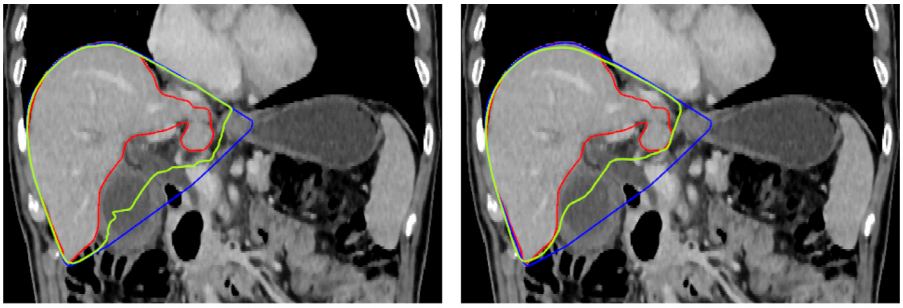
For each metric, the mean and standard deviation of the overall datasets are given. Bold values are the best result in that row. * indicates the statistically significant differences between input shape and results with the Frobenius norm term, and between results without and with the Frobenius norm term at a significance level of 0.05.

5.2. Comparison of the shape refinement with and without the Frobenius norm

One of the important steps in our algorithm is the use of the Frobenius norm term in Eqs. (3) and (4). This term can explicitly model the small dense Gaussian noise. In order to show the usefulness of this term on the final shape refinement, we tested the performance of our shape model LRSD-SM with the Frobenius norm term by adding synthetic noise to the input liver shapes on the Liver-Pathology database. More specifically, we calculated the convex hulls for all the triangular mesh shapes of the gold standard segmentations, and utilized them as the input shapes for the refinement test, which include both gross errors and Gaussian noise. Note that since the 3-D anisotropic diffusion filter (Perona and Malik, 1990) was firstly employed to denoise all the training and testing images before processing, we only tested the influence of this term on the proposed shape model LRSD-SM instead of the proposed probabilistic atlas model LRSD-PA. Subsequently, we performed the comparison of the shape refinement with and without the Frobenius norm term.

Fig. 10a shows the liver shape refinement result without the Frobenius norm term, namely, by using Eq. (2) and only modeling the sparse gross errors. It can be seen that most of the gross errors presented in the input shape are eliminated, and the refinement results are close to the gold standard. However, there are some unsatisfactory results with a slightly serrated contour. When a Frobenius norm penalty is added to Eqs. (3) and (4), the refinement results are much improved as shown in Fig. 10b. Specifically, the recovered contour is much smoother and much closer to the gold standard than that shown in Fig. 10a, especially in the concave regions. This is due to the explicit modeling of Gaussian noise through the Frobenius norm term.

Table 6 tabulates the quantitative evaluation results of the liver shape refinement with and without the Frobenius norm term from the input liver shapes with synthetic noise. We found the statistically significant differences between input shape and result with the Frobenius norm term (paired t-test: $t = 11.31,\ p = 3.80 \times 10^{-12}$ for *JI*; $t = -11.46,\ p = 2.73 \times 10^{-12}$ for *ASD*; $n = 30$), and between results without and with the Frobenius norm term (paired t-test: $t = 4.46,\ p = 1.13 \times 10^{-4}$ for *JI*; $t = -7.83,\ p = 1.25 \times 10^{-8}$ for *ASD*; $n = 30$). These results demonstrate that the incorporation of the Frobenius norm term is beneficial to the overall performance, and our proposed LRSD-SM is robust against both large sparse shape errors and small dense Gaussian noise, thus leading to more accurate shape refinement results. Note that in this experiment, we added very large synthetic noise to the input liver shapes, resulting in a mean *ASD* and *JI* of 6.09 mm and 74.17%, respectively. This introduced relatively low accuracy of the refinement results as shown in Table 6. While in our segmentation of pathological liver tissue, the input liver shapes obtained by the LRSD-PA based shape



(a) (b)

Fig. 10. Visual comparisons of the liver shape refinement results with and without the Frobenius norm term. Refinement results were obtained from the LRSD-based shape models (a) without the Frobenius norm term and (b) with the term, respectively. Synthetic noise was added to the input liver shapes on the Liver-Pathology database. All images are of 2-D coronal slices. The red and blue contours show the gold standard and the input shape, respectively. The shape refinement results are shown in green contours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 7
Quantitative comparative initial shape results obtained from the LRSD-based atlas models with and without the clustering step on the Liver-Pathology database.

Metrics	Without clustering	With clustering
JI [%]	$85.68 \pm 3.99^{*}$	$\mathbf{86.97 \pm 3.32}$
ASD [mm]	$2.55 \pm 1.27^{*}$	$\mathbf{2.23 \pm 0.96}$

For each metric, the mean and standard deviation of the overall datasets are given. Bold values are the best result in that row. * indicates the statistically significant differences between the results without and with the clustering step at a significance level of 0.05.

initialization method are much closer to the gold standard (results are tabulated later in Table 7).

5.3. Comparison of the LRSD-based atlas models with and without the clustering of training data

A clear drawback of the Liu's LRSD-based atlas model (Liu et al., 2014; Liu et al., 2015) is that it fails when the target image differs considerably from the training data, especially for soft tissue organs (e.g., the liver) with large anatomical variations. To alleviate this difficulty, we firstly partition the training data into multiple clusters of similar type cases before applying LRSD, instead of straightforwardly applying LRSD presented in Liu's method (Liu et al., 2014; Liu et al., 2015). Next, we compare the performance of LRSD-based atlas models working with and without the clustering step for demonstrating the usefulness of the clustering of training data.

Fig. 11 shows the comparisons of the derived lesion-corrected liver images for one challenging case that differs considerably from the training data with and without the clustering step. For this challenging case, the liver parenchyma adjacent to the gallbladder (the quadrate lobe) is much smaller than that of the training cases. It can be seen that the LRSD-based atlas model without the clustering step fails to obtain satisfactory results. In particular, the appearance in regions with atypical shapes (highlighted by the white rectangles), especially in the gallbladder (indicated by black arrows), deviates considerably from the original input images. By contrast, the derived lesion-corrected images of our method (i.e., with the clustering step) are more faithful to the input images in all regions. Moreover, our method is demonstrated to offer a substantial reduction in average computation time and memory cost of about 73% and 67%, respectively, when compared to that without the clustering step (the number of clusters $k = 3$). Table 7 lists the quantitative comparative initial shape results with and without the clustering step on the Liver-Pathology database. It can be seen that our method obtains higher accuracy with smaller variances, and statistically significant differences are found between the results without and with the clustering step (paired t-test: $t = 2.94$, $p = 6.32 \times 10^{-3}$ for *JI*; $t = -2.79$, $p = 9.32 \times 10^{-3}$ for *ASD*; $n = 30$). These results indicate that the clustering of training data before applying LRSD significantly improves the results over a direct application of LRSD.

In terms of clinical application, that is to generally achieve a clinical accuracy to within 2 mm (Nicolau et al., 2004; Kokudo et al., 2002; Nogueira et al., 2009), LRSD-based atlas model method does not yield a satisfactory segmentation yet (see Table 7). Therefore, an ASM refinement step is developed in our algorithm to further improve the segmentation results. Before the ASM refinement step, a mean *ASD* of 2.23 ± 0.96 mm is found for our method. After the refinement step, the mean *ASD* is improved to 1.27 ± 0.36 mm. We observe similar results when the *JI* is used as the metric for evaluation. More specifically, the mean *JI* is improved from 86.97 ± 3.32% to 92.66 ± 1.55%. Therefore, the proposed ASM refinement step is absolutely necessary in order to yield accurate and satisfactory results for clinical applications.

5.4. Comparison with state-of-the-art methods

To evaluate the performance of our proposed hierarchical segmentation framework in the broader context of recent research, we compared it with published state-of-the-art automatic methods that have participated in the SLIVER07 competition (Heimann et al., 2009). We also validated our proposed method on the SLIVER07-Test database which was specifically tailored for liver segmentation competition. Since the manual reference segmentation of the database is unavailable, we submitted our segmentation results to the competition organizers, who then independently evaluated the accuracy for us. The following five volume and surface based metrics were employed for quantitative comparison: volumetric overlap error (*VOE*), signed relative volume difference (*SRVD*), average symmetric surface distance (*ASD*), root mean square symmetric surface distance (*RMSD*), and maximum symmetric surface distance (*MSD*). For a more detailed definition of these metrics, we refer the reader to Heimann et al. (2009).

Table 8 summarizes the quantitative results of all compared methods that have achieved top 10 results. Among all compared methods, our method ranks in the third place. Our method achieves a total score of 76, and it is only one point below the best score of 77 reached by Wimmer's and Kainmüller's method. The top two ranked methods by Wimmer et al. and Kainmüller et al. are also based on the ASM framework. Table 9 lists the differences between their main components and ours. Compared with these two top-ranked methods, our method offers the following advan-

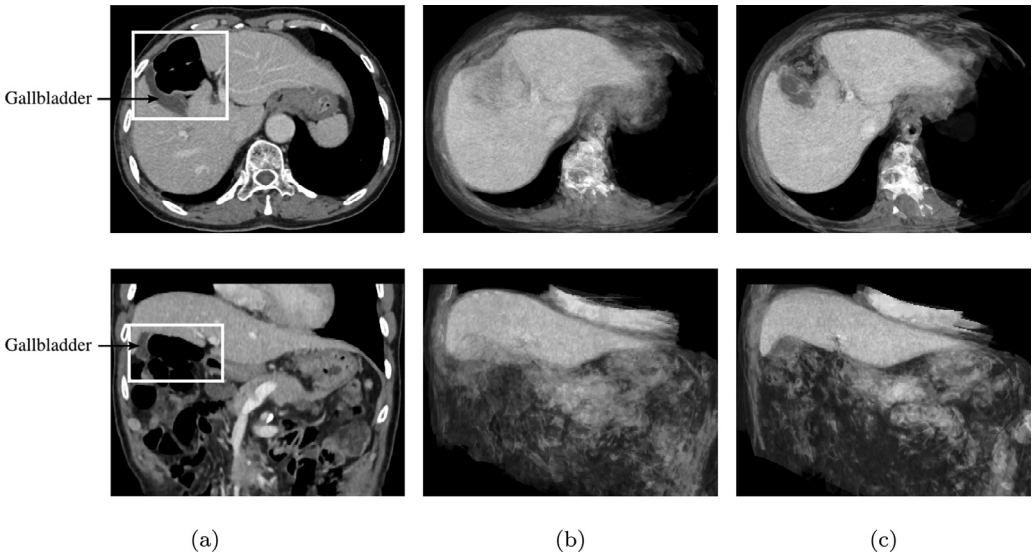


Fig. 11. Visual comparisons of the derived lesion-corrected liver images for one challenging case that differs considerably from the training data with and without the clustering step. (a) The pathological input images. Results were obtained from the LRSD-based atlas models (b) without the clustering step and (c) with the step, respectively. The first and second rows are of 2-D axial and coronal slices, respectively. The regions with atypical shapes are highlighted by the white rectangles.

Table 8
Comparison of the proposed method to the state-of-the-art methods in the SLIVER07-Test database.

Method	Training Samples	VOE [%]	SRVD [%]	ASD [mm]	RMSD [mm]	MSD [mm]	Total Score
Wimmer et al., 2009	20	6.5 ± 0.9	1.0±2.7	1.0 ± 0.2	2.0 ± 0.3	**18.3 ± 4.7**	**77 ± 4**
Kainmüller et al., 2007	112	**6.1 ± 2.1**	−2.9±2.9	**0.9 ± 0.3**	1.9 ± 0.8	18.7 ± 8.5	77 ± 9
Our method	20	7.1 ± 1.1	**−0.5±2.8**	1.1 ± 0.2	2.0 ± 0.5	19.4 ± 7.1	76 ± 6
Linguraru et al., 2012	27	6.4 ± 1.4	2.3±2.3	1.0 ± 0.3	**1.9 ± 0.6**	20.7 ± 4.0	76 ± 6
Wimmer et al., 2008	20	6.5 ± 1.2	0.6±2.0	1.1 ± 0.3	2.5 ± 0.9	24.1 ± 9.0	74 ± 6
Huang et al., 2012	20	7.6 ± 1.7	−1.3±3.4	1.3 ± 0.4	2.3 ± 0.9	22.1 ± 9.7	72 ± 8
Gauriau et al., 2013	20	7.2 ± 1.8	2.6±3.0	1.3 ± 0.5	2.6 ± 0.8	23.1 ± 5.0	72 ± 10
Kirschner, 2013	47	7.5 ± 1.2	−0.9±3.7	1.3 ± 0.3	2.7 ± 0.7	25.4 ± 6.3	70 ± 7
Badakhshannoory and Saeedi, 2010	20	8.5 ± 1.3	−2.4±2.9	1.3 ± 0.3	2.6 ± 0.7	22.4 ± 7.3	70 ± 6
Heimann et al., 2007	35	7.7 ± 1.9	1.7±3.2	1.4 ± 0.4	3.2 ± 1.3	30.1 ± 10.2	67 ± 11

Table shows the quantitative results of all the existing methods that have achieved top 10 results. For each metric, the mean and standard deviation of the overall datasets are given. Bold values are the best result in that column.

Table 9
The differences in main components between the top two ranked ASM-based liver segmentation methods and our method.

Method	Shape representation	Shape correspondence	Shape initialization	Shape model	Appearance model	Search algorithm
Wimmer et al., 2009	Level set function (signed distance map)	N/R	Largest object detection based	Parzen density estimation based	Classifier-based boundary and region models	Multi-scale level set evolution
Kainmüller et al., 2007	Triangular mesh	Manual + surface parameterization	Right lung lobe based	SSM	Heuristic intensity model	Constrained free-form deformation
Our method	Triangular mesh	SPHARM + MDL	LRSD-PA based	LRSD-SM	SAM (normalized gradient profiles)	Multi-resolution ASM

N/R and SAM stand for Not Required and statistical appearance model, respectively.

tages: (1) Our method achieves the best *SRVD*, which directly assesses volumetric information important for applications such as liver surgery planning (Heimann et al., 2009); (2) Our shape model can explicitly handle gross errors caused by weak and misleading appearance cues of large lesions; (3) Our shape initialization method is not based on any heuristics, and can largely eliminate the negative effects of large lesions; and (4) We use a very small training set of 20 samples, which is much less than that used in Kainmüller's method (112); and when the training size drops to 43, the performance of their method decreases sharply, and only a total score of 73 is achieved. Specifically, the mean *VOE* and *ASD* of our method are 7.1% and 1.1 mm, respectively (Team HIT-liver[8] in SLIVER07 competition). Moreover, the score of our proposed framework compares favorably to that of a second independent human rater's manual segmentation with a score of 75 (Team Lara in SLIVER07 competition). These results demonstrate that the performance of our method is comparable with that of state-of-the-art methods and human raters, thus it can be deployed for accurate and robust liver segmentation in the presence of severe pathology.

[8] http://www.sliver07.org/showresult.php?rank=32&submission=2015-10-17-0655.

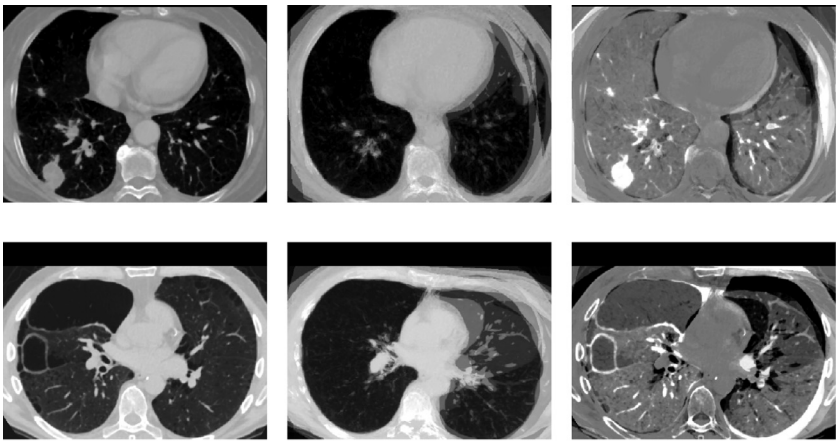


Fig. 12. Results of the derived lesion-corrected right lung images for two challenging pathological cases via our proposed LRSD-PA method, including right lung with hyperdense lesion (1st row) and hypodense lesion (2nd row). The first column shows the pathological input images. The second and third columns are their corresponding lesion-corrected images and sparse lesions, respectively.

Table 10
Quantitative comparative results of the final right lung segmentation by using two different model initialization methods on the Lung-Pathology database.

Method	JI [%]	ASD [mm]
Conventional PA	$94.80 \pm 2.02^{*}$	$0.91 \pm 0.33^{*}$
Our LRSD-PA	$\mathbf{96.13 \pm 1.54}$	$\mathbf{0.63 \pm 0.25}$

For each metric, the mean and standard deviation of the overall datasets are given. Bold values are the best result in that column. * indicates a statistically significant difference between the marked result and the corresponding one of our method at a significance level of 0.05.

5.5. Pathological right lung segmentation from CT images

To illustrate the applicability of our method to various pathological organ segmentation, we also evaluated it on pathological right lung cases.

Fig. 12 shows the results of the derived lesion-corrected right lung images for two challenging pathological cases via our proposed LRSD-PA method, including right lung with hyperdense and hypodense lesions, respectively. We observe that all these lesions with different patterns are successfully isolated from the derived lesion-corrected images, and the resulting right lung tissue has a much more homogeneous appearance. Fig. 13 presents the comparative results of initial right lung shape extraction by using conventional PA-based method and our proposed LRSD-PA based method in two challenging pathological cases, including right lung with hyperdense and hypodense lesions, respectively. It shows that the initial shapes of our method have much higher overlaps with the ground truth than the conventional method, especially in regions with large lesions. This is because the derived lesion-corrected pathological images are employed as the input images in our method, which leads to more accurate atlas-to-target registration and higher likelihood of the lung tissue. Therefore, the negative effects of these large lesions that complicate the conventional method are largely eliminated in our method.

Table 10 lists the final right lung segmentation results by using our proposed LRSD-PA based method for shape initialization on the Lung-Pathology database, when compared to the conventional PA-based initialization method. For the final segmentation results, our method yields mean *JI* and *ASD* of 96.13% and 0.63 mm, respectively. Also, our method brings significant improvements (paired *t*-test: $t = 3.26,\ p = 5.70 \times 10^{-3}$ for *JI*; $t = -3.59,\ p = 2.96 \times 10^{-3}$ for *ASD*; $n = 15$) over the conventional method in all metrics. Specifically, our method outperforms it by 1.33% and up to 30% on average in terms of *JI* and *ASD*, respectively.

Therefore, for the lung tissue showing different lesion patterns, our proposed method also achieves significantly more accurate and robust segmentation results than the conventional method. It can thus be deployed for accurate and robust pathological lung segmentation. Moreover, all these experimental results demonstrate the strong consistency and generality of our method not only across organ type, but also across pathology type with considerably different characteristics.

6. Conclusion

In this paper, we have presented a novel method for accurate and robust pathological organ segmentation from CT images. This method is grounded on the ASM framework. Our method improves the ASM framework by the proposal of: (1) a new LRSD-SM shape prior model; (2) a new LRSD-PA based shape initialization method; and (3) a hierarchical ASM search strategy. To illustrate the applicability of our proposed segmentation framework to various pathological organs, we have tested it in two challenging clinical applications involving pathological liver and right lung segmentation from CT images. Extensive evaluations were conducted on a total of 95 clinical CT scans of pathological cases (including some highly severe ones), and high accuracy and robustness have been achieved. Both qualitative and quantitative experimental results demonstrate that:

(1) The proposed LRSD-SM method achieves significantly better accuracy and robustness than both the conventional SSM model and the SSC model;
(2) The proposed LRSD-PA based shape initialization method derives more patient-specific initial shapes for the hierarchical segmentation framework and achieves significantly better results than the conventional PA-based method;
(3) The performance of the proposed segmentation framework is comparable with that of state-of-the-art methods and human raters.

Therefore, our proposed method significantly improves the performance of the ASM framework in the presence of severe pathol-

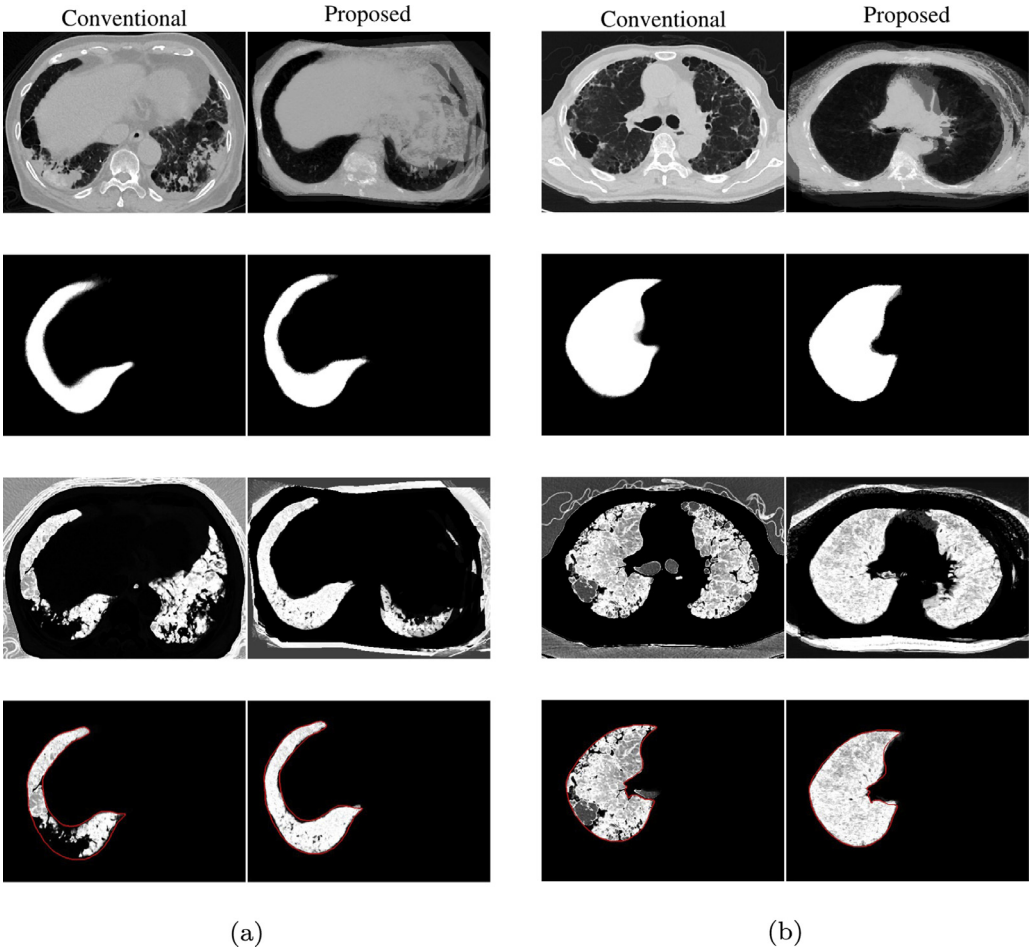


Fig. 13. Comparative results of initial right lung shape extraction by using conventional PA-based method and our proposed LRSD-PA based method in two challenging pathological cases, including right lung with (a) hyperdense lesion and (b) hypodense lesion. Every column shows one case, representing the conventional and proposed methods, respectively. In each column, the followings are shown sequentially from top to bottom: input image, PA, likelihood image, and the initial shape (i.e., the combination of PA and likelihood image) with the ground truth indicated by red contours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

ogy, and it can be deployed for accurate and robust pathological organ segmentation.

Although very encouraging results have been obtained by our proposed segmentation framework, there are still several ways to further enhance it:

(1) We employed the simple appearance model based on normalized gradient profiles in our framework, because we want to focus on the shape prior model (i.e., LRSD-SM). Using more sophisticated appearance models such as classifiers or clustering techniques would further improve the segmentation accuracy as suggested in Heimann and Meinzer (2009);
(2) In our current implementation, the computational time of our method is a bit high, mainly due to the computationally demanding nature of image registrations performed in the model initialization module. And they account for the vast majority of processing time in the testing phase (more than 90%). With recent advances in graphics processing unit (GPU) technology, the computational costs of our framework would be substantially reduced by parallelizing the image registration algorithms on GPU as suggested in Eklund et al. (2013).

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments and helpful suggestions that greatly improved the paper's quality. This work was supported in part by the National Natural Science Foundation of China under Grant No. 61571158, and the Nature Science Foundation of Heilongjiang Province of China (No. F2015005 and No. QC2016090), and the Kakenhi by MEXT and JSPS (26108006) and JSPS Bilateral International Collaboration Grants.

Appendix A. The Solutions to the SPCP Problem of LRSD-SM via the ADMM Method

We first derive the augmented Lagrangian for the SPCP problem in Eq. (4):

$$L_\mu(\mathbf{A},\mathbf{E},\mathbf{Z},\mathbf{Y}) = \|\mathbf{A}\|_* + \lambda\|\mathbf{E}\|_1 + \gamma\|\mathbf{Z}\|_F^2 + \langle \mathbf{Y}, \mathbf{D}-\mathbf{A}-\mathbf{E}-\mathbf{Z}\rangle + \frac{\mu}{2}\|\mathbf{D}-\mathbf{A}-\mathbf{E}-\mathbf{Z}\|_F^2, \quad (A.1)$$

where $\mathbf{Y}\in\mathbb{R}^{3N\times(K+1)}$ is the Lagrange multiplier, $\mu > 0$ is the penalty parameter, and $\langle \mathbf{P},\mathbf{Q}\rangle = tr(\mathbf{P}^T\mathbf{Q})$ denotes Euclidean inner product between two matrices, where $\mathbf{P}^T$ is the transpose of a matrix, and $tr(\cdot)$ is the trace of a matrix. Instead of directly solving the problem in Eq. (A.1) as in the classical augmented Lagrangian

method, ADMM divides it into three subproblems, and minimizes the variables **A, E** and **Z** sequentially with others fixed, leading to the following iterations:

$$\mathbf{A}_{k+1} = \arg\min_{\mathbf{A}} \ L_{\mu_k}(\mathbf{A}, \mathbf{E_k}, \mathbf{Z_k}, \mathbf{Y}_k), \tag{A.2}$$

$$\mathbf{E}_{k+1} = \arg\min_{\mathbf{E}} \ L_{\mu_k}(\mathbf{A_{k+1}}, \mathbf{E}, \mathbf{Z}_k, \mathbf{Y}_k), \tag{A.3}$$

$$\mathbf{Z}_{k+1} = \arg\min_{\mathbf{Z}} \ L_{\mu_k}(\mathbf{A_{k+1}}, \mathbf{E_{k+1}}, \mathbf{Z}, \mathbf{Y}_k), \tag{A.4}$$

$$\mathbf{Y}_{k+1} = \mathbf{Y}_k + \mu_k(\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1} - \mathbf{Z}_{k+1}). \tag{A.5}$$

All three minimization subproblems admit simple closed-form solutions.

(i) **A** *subproblem*:

Theorem 1 (Cai et al., 2010)**.** *Given a matrix* $\mathbf{W}$ *and* $\tau > 0$*, the optimal solution to the following minimization problem is given by:*

$$\mathcal{D}_\tau(\mathbf{W}) = \arg\min_{\mathbf{X}} \ \tau\|\mathbf{X}\|_* + \frac{1}{2}\|\mathbf{X} - \mathbf{W}\|_F^2, \tag{A.6}$$

where $\mathcal{D}_\tau$ *is the singular value thresholding (SVT) operator (Cai et al., 2010). It is defined as:* $\mathcal{D}_\tau(\mathbf{W}) = \mathbf{U}\mathcal{S}_\tau(\mathbf{\Sigma})\mathbf{V}^T$*, where* $\mathbf{U\Sigma V}^T = \mathbf{W}$ *is the singular value decomposition (SVD) of* $\mathbf{W}$*, and* $\mathcal{S}_\tau(\sigma_{ij}) = \max(|\sigma_{ij}| - \tau, 0) \cdot \mathrm{sgn}(\sigma_{ij})$ *is the shrinkage operator applied on* $\mathbf{\Sigma}$ *element-wisely, where sgn(·) is the sign function.*

Given other variables fixed, the closed-form solution for **A** subproblem in Eq. (A.2) can thus be derived, according to Theorem 1, as following:

$$\begin{aligned}\mathbf{A}_{k+1} &= \arg\min_{\mathbf{A}} \ L_{\mu_k}(\mathbf{A}, \mathbf{E_k}, \mathbf{Z_k}, \mathbf{Y}_k)\\ &= \arg\min_{\mathbf{A}} \ \|\mathbf{A}\|_* + \frac{\mu_k}{2}\|\mathbf{D} - \mathbf{A} - \mathbf{E}_k - \mathbf{Z}_k + \mu_k^{-1}\mathbf{Y}_k\|_F^2\\ &= \mathcal{D}_{\mu_k^{-1}}(\mathbf{D} - \mathbf{E}_k - \mathbf{Z}_k + \mu_k^{-1}\mathbf{Y}_k).\end{aligned} \tag{A.7}$$

(ii) **E** *subproblem*:

Theorem 2 (Hale et al., 2008)**.** *Given a matrix* $\mathbf{W}$ *and* $\tau > 0$*, the optimal solution to the following minimization problem is given by:*

$$\mathcal{S}_\tau(\mathbf{W}) = \arg\min_{\mathbf{X}} \ \tau\|\mathbf{X}\|_1 + \frac{1}{2}\|\mathbf{X} - \mathbf{W}\|_F^2, \tag{A.8}$$

where $\mathcal{S}_\tau$ *is the shrinkage operator.*

Similarly, the closed-form solution for **E** subproblem in Eq. (A.3) can be derived, according to Theorem 2, as following:

$$\begin{aligned}\mathbf{E}_{k+1} &= \arg\min_{\mathbf{E}} \ L_{\mu_k}(\mathbf{A_{k+1}}, \mathbf{E}, \mathbf{Z}_k, \mathbf{Y}_k)\\ &= \arg\min_{\mathbf{E}} \ \lambda\|\mathbf{E}\|_1 + \frac{\mu_k}{2}\|\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E} - \mathbf{Z}_k + \mu_k^{-1}\mathbf{Y}_k\|_F^2\\ &= \mathcal{S}_{\frac{\lambda}{\mu_k}}(\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{Z}_k + \mu_k^{-1}\mathbf{Y}_k).\end{aligned} \tag{A.9}$$

(iii) **Z** *subproblem*: With other variables fixed, **Z** subproblem in Eq. (A.4) leads to a quadratic optimization problem, which has the following closed-form solution:

$$\begin{aligned}\mathbf{Z}_{k+1} &= \arg\min_{\mathbf{Z}} L_{\mu_k}(\mathbf{A_{k+1}}, \mathbf{E_{k+1}}, \mathbf{Z}, \mathbf{Y}_k)\\ &= \arg\min_{\mathbf{Z}} \ \gamma \ \|\mathbf{Z}\|_F^2 + \frac{\mu_k}{2} \ \|\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1} - \mathbf{Z} + \mu_k^{-1}\mathbf{Y}_k\|_F^2\\ &= \frac{\mu_k}{\mu_k + 2\gamma}\left(\mathbf{D} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1} + \mu_k^{-1}\mathbf{Y}_k\right).\end{aligned} \tag{A.10}$$

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.media.2017.02.008.

References

Aljabar, P., Heckemann, R.A., Hammers, A., Hajnal, J.V., Rueckert, D., 2009. Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. NeuroImage 46 (3), 726–738.

Badakhshannoory, H., Saeedi, P., 2010. Automatic liver segmentation from CT scans using multi-layer segmentation and principal component analysis. In: Proc. International Symposium on Visual Computing (ISVC'10). Las Vegas, NV, USA, pp. 342–350.

Beck, A., Teboulle, M., 2009. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. 2 (1), 183–202.

Bouwmans, T., Zahzah, E.H., 2014. Robust PCA via principal component pursuit: a review for a comparative evaluation in video surveillance. Comput. Vision Image Understanding 122, 22–34.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2010. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. 3 (1), 1–122.

Burt, P.J., 1981. Fast filter transform for image processing. Comput.GraphicsImage Process. 16 (1), 20–51.

Cai, J., Candès, E.J., Shen, Z., 2010. A singular value thresholding algorithm for matrix completion. SIAM J. Optim. 20 (4), 1956–1982.

Candès, E.J., Li, X., Ma, Y., Wright, J., 2011. Robust principal component analysis? J. ACM 58 (3), 11.

Candès, E.J., Plan, Y., 2010. Matrix completion with noise. Proc. IEEE 98 (6), 925–936.

Chandrasekaran, V., Sanghavi, S., Parrilo, P.A., Willsky, A.S., 2011. Rank-sparsity incoherence for matrix decomposition. SIAM J. Optim. 21 (2), 572–596.

Chu, C., Oda, M., Kitasaka, T., Misawa, K., Fujiwara, M., Hayashi, Y., Wolz, R., Rueckert, D., Mori, K., 2013. Multi-organ segmentation from 3D abdominal CT images using patient-specific weighted-probabilistic atlas. Proceedings of the SPIE Medical Imaging. Orlando, Florida, USA. 86693Y–1–86693Y–7.

Cootes, T.F., Taylor, C.J., 2001. Statistical Models of Appearance for Computer Vision. Technical Report. University of Manchester.

Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J., 1995. Active shape models-their training and application. Comput. Vision Image Understanding 61 (1), 38–59.

Dakua, S.P., Abi-Nahed, J., Al-Ansari, A., 2016. Pathological liver segmentation using stochastic resonance and cellular automata. J. Vis. Commun. Image Represent. 34, 89–102.

Dempster, A.P., Laird, N.M., Rubin, D.B., 1977. Maximum likelihood from incomplete data via the EM algorithm. J. R. Stat. Soc. 39 (1), 1–38.

Deng, J., Han, X., Chen, Y., Xu, G., Sato, Y., Hori, M., Tomiyama, N., 2014. Sparse and low-rank matrix decomposition for local morphological analysis to diagnose cirrhosis. IEICE Trans. Inf. Syst. E97-D (12), 3210–3221.

Eklund, A., Dufort, P., Forsberg, D., LaConte, S., 2013. Medical image processing on the GPU - past, present and future. Med. Image Anal. 17 (8), 1073–1094.

Foruzan, A.H., Chen, Y., Hori, M., Sato, Y., Tomiyama, N., 2014. Capturing large shape variations of liver using population-based statistical shape models. Int. J. Comput. Assist. Radiol. Surg. 9 (6), 967–977.

Gauriau, R., Cuingnet, R., Prevost, R., Mory, B., Ardon, R., Lesage, D., Bloch, I., 2013. A generic, robust and fully-automatic workflow for 3D CT liver segmentation. In: Proc. MICCAI Workshop on Abdominal Imaging: Computational and Clinical Applications. Nagoya, Japan, pp. 241–250.

Ganesh, A., Lin, Z., Wright, J., Wu, L., Chen, M., Ma, Y., 2009. Fast algorithms for recovering a corrupted low-rank matrix. In: Proceedings of the Third International Workshop Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'09). Aruba, Dutch Antilles, The Netherlands, pp. 213–216.

Hale, E.T., Yin, W., Zhang, Y., 2008. Fixed-point continuation for ℓ_1-minimization: methodology and convergence. SIAM J. Optim. 19 (3), 1107–1130.

Hamy, V., Dikaios, N., Punwani, S., Melbourne, A., Latifoltojar, A., Makanyanga, J., Chouhan, M., Helbren, E., Menys, A., Taylor, S., Atkinson, D., 2014. Respiratory motion correction in dynamic MRI using robust data decomposition registration - application to DCE-MRI. Med. Image Anal. 18 (2), 301–313.

Heimann, T., Meinzer, H.P., Wolf, I., 2007. A statistical deformable model for the segmentation of liver CT volumes. In: Proc. MICCAI Workshop 3-D Segmentation in the Clinic: A Grand Challenge, pp. 161–166.

Heimann, T., van Ginneken, B., Styner, M., Arzhaeva, Y., Aurich, V., Bauer, C., Beck, A., Becker, C., Beichel, R., Bekes, G., Bello, F., Binnig, G.K., Bischof, H., Bornik, A., Cashman, P., Chi, Y., Cordova, A., Dawant, B.M., Fidrich, M., Furst, J.D., Furukawa, D., Grenacher, L., Hornegger, J., Kainmüller, D., Kitney, R., Kobatake, H., Lamecker, H., Lange, T., Lee, J., Lennon, B., Li, R., Li, S., Meinzer, H.-P., Németh, G., Raicu, D.S., Rau, A., van Rikxoort, E.M., Rousson, M., Ruskó, L., Saddi, K.A., Schmidt, G., Seghers, D., Shimizu, A., Slagmolen, P., Sorantin, E., Soza, G., Susomboon, R., Waite, J.M., Wimmer, A., Wolf, I., 2009. Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans. Med. Imaging 28 (8), 1251–1265.

Heimann, T., Meinzer, H.-P., 2009. Statistical shape models for 3D medical image segmentation: a review. Med. Image Anal. 13 (4), 543–563.

Heimann, T., Oguz, I., Wolf, I., Styner, M., Meinzer, H.-P., 2006. Implementing the automatic generation of 3D statistical shape models with ITK. In: Proceedings of Insight Journal - ISC/NA-MIC Workshop on Open Science at MICCAI 2006. Copenhagen, Denmark.

Hong, M., Luo, Z.-Q., 2013. On the linear convergence of the alternating direction method of multipliers. CoRR abs/1208.3922v3, 1–37.
Huang, C., Jia, F., Li, Y., Zhang, X., Luo, H., Fang, C., Fan, Y., 2012. Automatic liver segmentation based on shape constrained differeomorphic demons atlas registration. In: Proc. International Conference on Electronics, Communications and Control (ICECC'12), pp. 126–129.
Hua, P., Song, Q., Sonka, M., Hoffman, E.A., Reinhardt, J.M., 2011. Segmentation of pathological and diseased lung tissue in CT images using a graph-search algorithm. In: Proceedings of the International Symposium on Biomedical Imaging: From Nano to Macro (ISBI'11). Chicago, Illinois, USA, pp. 2072–2075.
Ibáñez, L., Schroeder, W., Ng, L., Cates, J., 2005. The ITK Software Guide, second edition Kitware, Inc., Clifton Park, NY.
Iglesias, J.E., Sabuncu, M.R., 2015. Multi-atlas segmentation of biomedical images: a survey. Med. Image Anal. 24 (1), 205–219.
Jaccard, P., 1901. Étude comparative de la distribution florale dans une portion des alpes et des jura. Bulletin de la Société Vaudoise des Sciences Naturelles 37, 547–579.
Jolliffe, I., 2002. Principal Component Analysis, second edition Springer Series in Statistics, Springer, Berlin, Germany.
Kainmüller, D., Lange, T., Lamecker, H., 2007. Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In: Proc. MICCAI Workshop 3-D Segmentation in the Clinic: A Grand Challenge, pp. 109–116.
Kirschner, M., 2013. The Probabilistic Active Shape Model: From Model Construction to Flexible Medical Image Segmentation. Technische Universität Darmstadt, Darmstadt Ph.D. thesis. URL: http://tuprints.ulb.tu-darmstadt.de/3519/.
Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W., 2010. elastix: a toolbox for intensity-based medical image registration. IEEE Trans. Med. Imaging 29 (1), 196–205.
Kokudo, N., Miki, Y., Sugai, S., Yanagisawa, A., Kato, Y., Sakamoto, Y., Yamamoto, J., Yamaguchi, T., Muto, T., Makuuchi, M., 2002. Genetic and histological assessment of surgical margins in resected liver metastases from Colorectal Carcinoma: minimum surgical margins for successful resection. Arch. Surg. 137 (7), 833–840.
Lin, Z., Liu, R., Su, Z., 2011. Linearized alternating direction method with adaptive penalty for low-rank representation. In: Proceedings of the Advances in Neural Information Processing Systems 24 (NIPS'11). Granada, Spain, pp. 612–620.
Liu, G., Lin, Z., Yan, S., Sun, J., Yu, Y., Ma, Y., 2013. Robust recovery of subspace structures by low-rank representation. IEEE Trans. Pattern Anal. Mach. Intell. 35 (1), 171–184.
Liu, X., Niethammer, M., Kwitt, R., McCormick, M., Aylward, S., 2014. Low-rank to the rescue - atlas-based analyses in the presence of pathologies. In: Proceedings of the Medical Image Computing and Computer-Assisted Interventation (MICCAI'14). Boston, MA, USA, pp. 97–104.
Liu, X., Niethammer, M., Kwitt, R., Singh, N., McCormick, M., Aylward, S., 2015. Low-rank atlas image analyses in the presence of pathologies. IEEE Trans. Med. Imaging 34 (12), 2583–2591.
Linguraru, M.G., Richbourg, W.J., Liu, J., Watt, J.M., Pamulapati, V., Wang, S., Summers, R.M., 2012. Tumor burden analysis on computed tomography by automated liver and tumor segmentation. In: IEEE Trans. Med. Imag, 31, pp. 1965–1976.
Ma, H., Dibildox, G., Banerjee, J., Niessen, W., Schultz, C., Regar, E., van Walsum, T., 2015. Layer separation for vessel enhancement in interventional x-ray angiograms using morphological filtering and robust PCA. In: Proceedings of the MICCAI Workshop on Augmented Environments for Computer-Assisted Interventions (AE-CAI). Munich, Germany, pp. 104–113.
Mansoor, A., Bagci, U., Foster, B., Xu, Z., Papadakis, G.Z., Folio, L.R., Udupa, J.K., Mollura, D.J., 2015. Segmentation and image analysis of abnormal lungs at CT: current approaches, challenges, and future trends. RadioGraphics 35 (4), 1056–1076.
Mansoor, A., Bagci, U., Xu, Z., Foster, B., Olivier, K.N., Elinoff, J.M., Suffredini, A.F., Udupa, J.K., Mollura, D.J., 2014. A generic approach to pathological lung segmentation. IEEE Trans. Med. Imaging 33 (12), 2293–2310.
Murphy, K., van Ginneken, B., Reinhardt, J.M., Kabus, S., Ding, K., Deng, X., Cao, K., Du, K., Christensen, G.E., Garcia, V., Vercauteren, T., Ayache, N., Commowick, O., Malandain, G., Glocker, B., Paragios, N., Navab, N., Gorbunova, V., Sporring, J., de Bruijne, M., Han, X., Heinrich, M.P., Schnabel, J.A., Jenkinson, M., Lorenz, C., Modat, M., McClelland, J., Ourselin, S., Muenzing, S.E.A., Viergever, M.A., Nigris, D.D., Collins, D.L., Arbel, T., Peroni, M., Li, R., Sharp, G.C., Schmidt-Richberg, A., Ehrhardt, J., Werner, R., Smeets, D., Loeckx, D., Song, G., Tustison, N.J., Avants, B.B., Gee, J.C., Staring, M., Klein, S., Stoel, B.C., Urschler, M., Werlberger, M., Vandemeulebroucke, J., Rit, S., Sarrut, D., Pluim, J.P.W., 2011. Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge. IEEE Trans. Med. Imaging 30 (11), 1901–1920.
Nakagomi, K., Shimizu, A., Kobatake, H., Yakami, M., Fujimoto, K., Togashi, K., 2013. Multi-shape graph cuts with neighbor prior constraints and its application to lung segmentation from a chest CT volume. Med. Image Anal. 17 (1), 62–77.
Nicolau, S., Pennec, X., Soler, L., Ayache, N., 2004. An accuracy certified augmented reality system for therapy guidance. In: Proceedings of the 8th European Conference on Computer Vision (ECCV'04). Prague, Czech Republic, pp. 79–91.
Nogueira, J.F., Stamm, A.C., Lyra, M., 2009. Novel compact laptop-based image-guidance system: preliminary study. Laryngoscope 119 (3), 576–579.
Oda, M., Nakaoka, T., Kitasaka, T., Furukawa, K., Misawa, K., Fujiwara, M., Mori, K., 2011. Organ segmentation from 3d abdominal CT images based on atlas selection and graph cut. In: Proceedings of the MICCAI Workshop on Abdominal Imaging: Computational and Clinical Applications. Toronto, ON, Canada, pp. 181–188.
Okada, T., Shimada, R., Hori, M., Nakamoto, M., Chen, Y.-W., Nakamura, H., Sato, Y., 2008. Automated segmentation of the liver from 3D CT images using probabilistic atlas and multilevel statistical shape model. Acad. Radiol. 15 (11), 1390–1403.
Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man Cybern. 9 (1), 62–66.
Park, H., Bland, P.H., Meyer, C.R., 2003. Construction of an abdominal probabilistic atlas and its application in segmentation. IEEE Trans. Med. Imaging 22 (4), 483–492.
Peng, Y., Ganesh, A., Wright, J., Xu, W., Ma, Y., 2012. RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images. IEEE Trans. Pattern Anal. Mach. Intell. 34 (11), 2233–2246.
Perona, P., Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. 12 (7), 629–639.
Prasad, M.N., Brown, M.S., Ahmad, S., Abtin, F., Allen, J., da Costa, I., Kim, H.J., McNitt-Gray, M.F., 2008. Automatic segmentation of lung parenchyma in the presence of diseases based on curvature of ribs. Acad. Radiol. 15 (9), 1173–1180.
van Rikxoort, E.M., van Ginneken, B., 2013. Automated segmentation of pulmonary structures in thoracic computed tomography scans: a review. Phys. Med. Biol. 58 (17), R187–R220.
van Rikxoort, E.M., de Hoop, B., Viergever, M.A., Prokop, M., van Ginneken, B., 2009. Automatic lung segmentation from thoracic computed tomography scans using a hybrid approach with error detection. Med. Phys. 36 (7), 2934–2947.
von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17 (4), 395–416.
Rohlfing, T., Zahr, N.M., Sullivan, E.V., Pfefferbaum, A., 2010. The SRI24 multichannel atlas of normal adult human brain structure. Hum. Brain Mapp. 31 (5), 798–819.
Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J., 1999. Non–rigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging 18 (8), 712–721.
Schroeder, W., Martin, K., Lorensen, B., 2006. Visualization Toolkit: An Object-Oriented Approach to 3D Graphics, fourth edition Kitware, Inc., Clifton Park, NY.
Shi, Y., Shen, D., 2008. Hierarchical shape statistical model for segmentation of lung fields in chest radiographs. In: Proceedings of the Medical Image Computing and Computer-Assisted Interventation (MICCAI'08). New York, NY, USA, pp. 417–424.
Sluimer, I., Prokop, M., van Ginneken, B., 2005. Toward automated segmentation of the pathological lung in CT. IEEE Trans. Med. Imaging 24 (8), 1025–1038.
Sofka, M., Wetzl, J., Birkbeck, N., Zhang, J., Kohlberger, T., Kaftan, J., Declerck, J., Zhou, S.K., 2011. Multi-stage learning for robust lung segmentation in challenging CT volumes. In: Proceedings of the Medical Image Computing and Computer-Assisted Interventation (MICCAI'11). Toronto, ON, Canada, pp. 667–674.
Studholme, C., Hill, D.L.G., Hawkes, D.J., 1999. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognit. 32 (1), 71–86.
Styner, M., Oguz, I., Xu, S., Brechbühler, C., Pantazis, D., Levitt, J.J., Shenton, M.E., Gerig, G., 2006. Framework for the statistical shape analysis of brain structures using SPHARM-PDM. In: Proceedings of Insight Journal - ISC/NA-MIC Workshop on Open Science at MICCAI 2006. Copenhagen, Denmark.
Sun, S., Bauer, C., Beichel, R., 2012. Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach. IEEE Trans. Med. Imaging 31 (2), 449–460.
Tao, M., Yuan, X., 2011. Recovering low-rank and sparse components of matrices from incomplete and noisy observations. SIAM J. Optim. 21 (1), 57–81.
Umetsu, S., Shimizu, A., Watanabe, H., Kobatake, H., Nawano, S., 2014. An automated segmentation algorithm for CT volumes of livers with atypical shapes and large pathological lesions. IEICE Trans. Inf. Syst. E97-D (4), 951–963.
Wang, J., Li, F., Li, Q., 2009. Automated segmentation of lungs with severe interstitial lung disease in CT. Med. Phys. 36 (10), 4592–4599.
Wimmer, A., Hornegger, J., Soza, G., 2008. Implicit active shape model employing boundary classifier. In: Proc. International Conference on Pattern Recognition (ICPR'08). Tampa, Florida, USA, pp. 1–4.
Wimmer, A., Soza, G., Hornegger, J., 2009. A generic probabilistic active shape model for organ segmentation. In: Proc. Medical Image Computing and Computer-Assisted Interventation (MICCAI'09). London, UK, pp. 26–33.
Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D.N., Zhou, X.S., 2012. Towards robust and effective shape modeling: sparse shape composition. Med. Image Anal. 16 (1), 265–277.
Zhou, X., Yang, C., Zhao, H., Yu, W., 2014. Low-rank modeling and its applications in image analysis. ACM Comput. Surv. 47 (2), 36:1–36:33.
Zhou, Z., Li, X., Wright, J., Candès, E.J., Ma, Y., 2010. Stable principal component pursuit. In: Proceedings of the IEEE International Symposium on Information Theory (ISIT'10). Austin, Texas, USA, pp. 1518–1522.
Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. 67 (2), 301–320.
Zweng, M., Fallavollita, P., Demirci, S., Kowarschik, M., Navab, N., Mateus, D., 2015. Automatic guide-wire detection for neurointerventions using low-rank sparse matrix decomposition and denoising. In: Proceedings of the MICCAI Workshop on Augmented Environments for Computer-Assisted Interventions (AE-CAI). Munich, Germany, pp. 114–123.