



A hierarchical local region-based sparse shape composition for liver segmentation in CT scans



Changfa Shi^a, Yuanzhi Cheng^{a,*}, Fei Liu^b, Yadong Wang^a, Jing Bai^b, Shinichi Tamura^c

^a School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China

^b School of Medicine, Tsinghua University, Beijing 100084, China

^c Center for Advanced Medical Engineering and Informatics, Osaka University, Suita 565-0871, Japan

ARTICLE INFO

Article history:

Received 21 November 2014

Received in revised form

7 August 2015

Accepted 2 September 2015

Available online 14 September 2015

Keywords:

Liver segmentation

Active shape model

Sparse shape composition

Shape segmentation

Hierarchical model

ABSTRACT

Motivated by the goals of improving segmentation of challenging liver cases containing low contrast with neighboring organs and presence of pathologies as well as highly varied shapes between subjects, a novel framework is presented for liver segmentation in portal phase of abdominal CT images. In a first training step, we describe a multilevel local region-based Sparse Shape Composition (SSC) model, called MLR-SSC, to increase the flexibility of shape prior models and capture the detailed local shape information more faithfully. Specifically, the liver shapes are decomposed into multiple regions in a multilevel fashion. Moreover, we build a local shape repository for each region and refine an input shape in a region-by-region manner. In a second testing step, it starts with a blood vessel-based liver shape initialization to derive a more patient-specific initial shape, followed by a hierarchical deformable shape optimization algorithm. It makes the segmentation framework more efficient and robust to local minima. Extensive experiments on 60 clinical CT scans demonstrate that our method achieves much better accuracy and efficiency than two closely related methods in the presence of small training sets. Moreover, our method shows slightly superior performance to three newly published methods. Also, we compare our method with the published semi-automatic methods from the “MICCAI 2007 Grand Challenge” workshop.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Liver segmentation (i.e., the delineation of the liver tissue from its surroundings) from computed tomography (CT) images is essential and forms the basis for various medical procedures, such as liver cancer diagnosis, treatment planning and image-guided surgery. It also facilitates the subsequent segmentation of hepatic vessels and liver tumors by providing a liver mask. In traditional clinical practice, liver tissue is still manually delineated by radiologic technologists in a slice-by-slice manner, which is tedious and time-consuming, and suffers from high intra- and inter-observer variability. Therefore, semi-automatic or automatic liver segmentation techniques are highly desirable, especially in many real-time clinical applications.

However, accurate liver segmentation from CT images is still a challenging task due to (see Fig. 1): (1) large anatomical variation in both shape and size; (2) low contrast and fuzzy boundary between liver tissue and neighboring organs sharing similar

intensity values, such as the heart and the stomach; and (3) the presence of noise, partial volume effects and pathological abnormalities.

In the last few decades, numerous methods have been proposed to perform liver segmentation from CT images. A comprehensive review of different techniques has been presented by Campadelli et al. [1], including detailed analysis of their advantages and drawbacks. A thorough comparison of different sophisticated techniques based on the public datasets of the Segmentation of the Liver Competition 2007 (SLIVER07) has been investigated by Heimann et al. [2]. These methods can be roughly categorized as either image-based or model-based methods according to the amount of shape information employed. The first group is mainly based on pure image information, such as thresholding [3] and region growing [4]. Their major limitation is the tendency to leak into neighboring organs with similar intensity values to the liver tissue. Consequently, in order to obtain accurate and robust segmentation, the incorporation of local and global liver shape prior knowledge is introduced in model-based methods, such as deformable models [5], graph-cuts [6], active shape models (ASMs) [7–14], and probabilistic atlases [15]. Among the myriad

* Corresponding author. Tel./fax: +86 451 8641 3309.

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

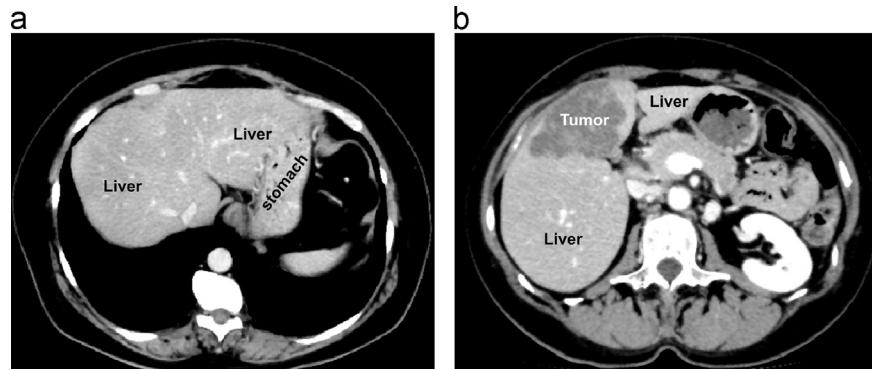


Fig. 1. Examples illustrating challenges in accurate liver segmentation from CT images: (a) low contrast and fuzzy boundary between liver tissue and neighboring organs with similar intensity values, such as the stomach; (b) the presence of pathological abnormalities.

methods recently explored, ASMs have achieved state-of-the-art liver segmentation results [2]. One of the major features of ASM is its capability of combining both low-level appearance information and high-level shape prior information in a unified framework. By using these learned prior knowledge, ASM matches an object-specific model to new images in a top-down fashion, thus more accurate and robust results can be achieved compared to pure low-level methods.

After Cootes' seminal work on ASM [16], various techniques have been proposed to improve it in many aspects, which are thoroughly surveyed in the excellent paper of Heimann and Meinzer [17]. In ASM-based method, statistical shape models (SSMs) using principal component analysis (PCA) [18] are usually employed to learn the shape prior models. However, SSMs suffer from a major limitation: they tend to overly constrain the shape deformations and overfit the training data due to the small size of training data. This results in shape prior models that have low generalization ability, that is, they cannot be adapted accurately to the finer local details of shapes in new images. This is especially the case in the following two situations: (1) modeling 3-D shapes where the number of training samples is always too low, since the manual delineation of 3-D training data is laborious and time-consuming; and (2) modeling soft tissue organs (e.g., liver tissue) that exhibit highly complex local shape variations. The reason is that the dimension of the principal subspace spanned by the eigenvectors of the covariance matrix is at most equal to the number of training samples minus one. Therefore, when a limited number of training samples is available, the principal subspace of plausible shapes cannot capture the full range of shape variability present in the training data. Even if we have enough training data, some local details that are not statistically significant will be smoothed out in the final shape models.

Many techniques have been developed in the literature to increase the flexibility of SSMs, which can be broadly divided into three categories [19,13]: (1) enlargement of training data by synthesizing artificial shapes from existing training samples [19]; (2) hierarchical formulation of SSMs by decomposing the shape into multiple subparts [9] or frequencies [20], and by representing the shape as a multi-resolution geometric model [10]; and (3) incorporation of additional shape variability by introducing a synthetic covariance matrix [21]. Another two limitations of SSMs are their inability to model complex shapes with non-Gaussian shape distribution, and sensitivity to non-Gaussian errors present in input shapes. Several nonlinear extensions of SSMs have been proposed to handle multimodal shape distribution, such as those based on Gaussian mixture models [22], polynomial regression [23], and kernel PCA [24]. To robustify ASMs against outliers, strategies including outlier detection and correction [25],

reduction of outliers' influence [26], and robust point matching [27] have been developed.

Recently, inspired by the popular sparse learning theory [28] in the signal processing and machine learning communities, Zhang et al. [29] proposed Sparse Shape Composition (SSC) shape prior modeling method to tackle the above-mentioned three limitations of SSMs in a unified framework. Without assuming any parametric shape distribution model, SSC implicitly incorporates shape priors on-the-fly by refining the input shape as a sparse linear combination of training shapes in a shape repository. Since all the training shapes are directly incorporated into the shape repository, even local details that are not statistically significant in the training data can be preserved. By explicitly modeling the non-Gaussian errors present in input shapes, SSC is also robust against outliers. However, since SSC still uses global shape information to refine an input shape, the detailed local shape information cannot be accurately reconstructed. Also, the large-scale sparse optimization of SSC has particularly low computational efficiency when both the size of the shape repository and the number of vertices of each shape are large.

In this paper, we introduce a novel framework for accurate and robust liver segmentation in portal phase of abdominal CT images based on ASMs. We first propose a new multilevel local region-based SSC (MLR-SSC) to increase the flexibility of shape prior models and capture the detailed local shape information more faithfully, especially for modeling highly varied liver shapes when a limited number of training samples is available. Specifically, we decompose the liver shapes into multiple regions in a multilevel fashion, such that each region has homogeneous shape variation. Subsequently, we build a local shape repository for each region and refine an input shape in a region-by-region manner. Since local region-based shape statistics often lie in a more compact subspace than that of global shape, it can describe the detailed local shape information more accurately. Moreover, because the number of vertices of each region becomes smaller, the computational efficiency of the sparse optimization will be improved.

It is well-known that shape initialization is critical for ASM-based method due to its local search strategy. Therefore, the initial shape needs to be close to the object of interest, otherwise, the shape deformation will have a low convergence rate or even lead to a complete segmentation failure. Many organ detection methods have been proposed to automatically initialize the shape model, such as those based on evolutionary algorithm [8], generalized Hough transform [11], and machine learning [30]. However, all these approaches typically use the mean shape of training data to initialize the shape model. This may cause shape deformation stuck in local minima due to large anatomical variations, especially for soft tissue organs like the liver. Therefore, more advanced appearance models such as classifiers and clustering

techniques [17] are usually employed for them to achieve good results. In particular, Ling et al. [31] used marginal space learning (MSL) [32] with 3-D Haar features and steerable features to detect livers. Instead of an exhaustive search in the full parameter space, MSL gradually increase the dimensionality of the marginal space by sequentially estimating position, orientation and scale parameters, and coefficients corresponding to the first three shape components. However, to perform well in practice, MSL needs a large amount of training data to learn the detectors, which is difficult to derive from 3-D medical imaging. Inspired by recent blood vessel-based liver segmentation methods [33], we propose a new blood vessel-based liver shape initialization method to derive a more patient-specific initial shape. The key idea behind this method is that hepatic vessels (i.e., portal vein and hepatic vein) of each liver tissue provide unique prior information in terms of its shape and size.

Finally, in order to make the liver segmentation framework more efficient and more robust to local minima, we develop a hierarchical deformable shape optimization strategy, where the shape model is deformed in a coarse-to-fine manner. Extensive experiments on 60 clinical CT scans demonstrate that our proposed MLR-SSC model achieves much better accuracy and efficiency than both the conventional shape model and the original SSC model in the presence of small training sets. Moreover, our method shows slightly superior performance to state-of-the-art model-based methods. Also, we compared our method with the published semi-automatic methods from the “MICCAI 2007 Grand Challenge” workshop.

The main contributions of this paper can be summarized as follows: (1) MLR-SSC is proposed to increase the flexibility of shape prior models and capture the detailed local shape information more faithfully in the presence of small training sets (Section 3.1); (2) An efficient and scalable optimization algorithm (i.e., LARS-Homotopy method [34]) is employed to solve the sparse optimization problem in MLR-SSC (Section 3.1.2); (3) A blood vessel-based liver shape initialization method is proposed to derive a more patient-specific initial shape (Section 3.2.1); (4) A hierarchical optimization strategy is developed to make the segmentation framework more efficient and more robust to local minima (Section 3.2.2) and is successfully applied to segment liver tissue from CT images (Section 5).

2. Background and related work

In this section, we briefly review the closely related work [29] on Sparse Shape Composition (SSC) model, and explain how previous work differs from ours.

SSC [29] is a sparse representation-based shape prior modeling method that implicitly incorporates the shape priors on-the-fly to preserve local detail information. Based on two sparsity properties of the input shape (i.e., the input shape can be approximately represented by a sparse linear combination of existing training shapes in a shape repository; the input shape may include sparse gross errors), SSC cast the shape modeling as the following sparse optimization problem [29]:

$$(\hat{\mathbf{x}}, \hat{\mathbf{e}}) = \underset{\mathbf{x}, \mathbf{e}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{D}\mathbf{x} - \mathbf{e}\|_2^2 + \lambda_1 \|\mathbf{x}\|_1 + \lambda_2 \|\mathbf{e}\|_1, \quad (1)$$

where $\mathbf{y} \in \mathbb{R}^{3N}$ (N is the number of vertices of a shape) is a pre-aligned input shape to be refined, $\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_K] \in \mathbb{R}^{3N \times K}$ is the shape repository (also called dictionary) that consists of K pre-aligned training shapes \mathbf{d}_i , $\mathbf{x} \in \mathbb{R}^K$ denotes the coefficients of linear combination, $\mathbf{e} \in \mathbb{R}^{3N}$ represents the sparse gross errors, $\|\mathbf{v}\|_1 = \sum_i |v_i|$ and $\|\mathbf{v}\|_2$ are the ℓ_1 -norm that induces sparsity and Euclidean norm of \mathbf{v} respectively, λ_1 and λ_2 are positive regularization

parameters that control the sparsity of \mathbf{x} and \mathbf{e} respectively. Then the input shape \mathbf{y} is refined as $\hat{\mathbf{D}}\hat{\mathbf{x}}$ and transformed to its original coordinate frame.

SSC has been successfully applied to segment liver tissue, and shows better performance than several state-of-the-art methods [29]. However, the shape variation over the surface of the liver shape model differs considerably as shown in Fig. 3 (the tips of left and right lobes have more variation than other parts of the liver surface). Since SSC uses global shape information to refine an input shape, training shapes whose high variation parts are similar to that of the input shape are more likely to be chosen (i.e., with nonzero coefficients) to sparsely represent the input shape. Therefore, the parts of the surface with low but important shape variation will tend to be disregarded, which means that the detailed local shape information cannot be accurately reconstructed by the SSC method. Also in SSC, the interior point method is employed to solve the sparse optimization problem, whose computational complexity is $O(N^2 K)$ per iteration [35]. Therefore, it will suffer from a particularly high computational cost when both the size of the shape repository K and the number of vertices of each shape N are large, which limits its wide applicability to large-scale problems and real-time clinical applications.

To alleviate these problems, we propose a novel multilevel local region-based SSC (MLR-SSC) to improve the accuracy and reduce the computational cost. Specifically, we decompose the liver shapes into multiple regions in a multilevel fashion, such that each region has homogeneous shape variation. Then we build a local shape repository for each region and refine an input shape in a region-by-region manner.

We note that recently Zhang et al. [35] proposed a similar mesh partitioning based SSC as our proposed MLR-SSC model. However, our method differs from theirs in four main aspects: (1) They employed appearance features to partition the shape into multiple parts, while we decompose the shape into regions with homogeneous shape variation, as the appearance features are more suitable for learning the appearance model as in [31] instead of the shape model; (2) They only used one shape division level and partitioned the shape into 30 regions, while we decompose the shape in a multilevel fashion with only 16 regions, since the larger the number of sub-region, the more discontinuities in locations where sub-regions intersect; (3) They used the interior point methods to solve the sparse optimization problem, while we employ a much more efficient and scalable optimization algorithm (i.e., LARS-Homotopy method [34]); (4) They applied the sparse coefficients \mathbf{x}_j derived from each sub-region to the full shape space \mathbf{D} to refine the entire input shape, while we employ a region-by-region refinement strategy, which is more accurate and efficient.

After this paper was submitted for publication, we became aware of a recent paper by Shao et al. [36], who independently proposed a hierarchical ASM-based segmentation framework by incorporating a multilevel local sparse shape composition and a local sparse appearance model. While both methods try to solve the problems of high anatomical shape variability and boundary ambiguity, the main differences between our method and the method proposed in [36] are: (1) Their method's specific application is 2-D lung field segmentation from chest radiographs, while we target the more challenging 3-D liver segmentation in CT images, where the liver tissue exhibits much more complex shape variation and fuzzier boundary than the lung tissue. Also the sizes of the lesions in chest radiographs (i.e., pulmonary lung nodules) are much smaller compared to liver tumors in CT images; (2) The objective function of their employed SSC model is (Eq. (6) in their paper)

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_1, \quad (2)$$

which only accounts for small dense Gaussian noise as in PCA-based SSMs due to ℓ_2 -norm nature. While in our proposed MLR-SSC (as shown in Eq. (6)), we also explicitly model the sparse gross errors with the error term \mathbf{e} . Therefore, our method is robust to both sparse gross errors and small Gaussian noise. Also it has been shown that SSC with the explicit error term \mathbf{e} as in Eq. (1) can achieve much more accurate segmentation results than that of Eq. (2) (see Table 3 of [29]); (3) They used affinity propagation clustering as in [35] to perform 2-D contour division, while we adapt the Lloyd's clustering [37] based variational shape approximation (VSA) method [38] for 3-D surface division. In their vertex-based method, they only penalize the proximity of one vertex to another within the same segment, which can result in spatially discontinuous segments. On the contrary, we adopt a face-based mesh segmentation strategy, where face continuity within the resulting divisions is implicitly imposed as a hard constraint; (4) In their hierarchical segmentation framework, a single-resolution technique (i.e., only use the original input image) is implemented, while we develop a multi-resolution optimization strategy (i.e., use Gaussian pyramids of the original input image), which has proved to be more efficient and robust to local minima [39].

3. Methods

In this section, we describe our proposed ASM-based hierarchical liver segmentation framework. The main workflow of the segmentation framework is illustrated in Fig. 2, which consists of offline training and runtime testing stages.

In offline training stage, two main models used in ASM search are learned: (1) our proposed MLR-SSC as the shape prior model (Section 3.1); and (2) an appearance model based on normalized gradient profiles [39].

In runtime testing stage, given a testing image, a blood vessel-based method is employed to derive an initial liver shape (Section 3.2.1). The initial shape is then iteratively deformed and regularized according to the learned appearance and shape prior models, respectively. To make our method more efficient and robust, we adopt a hierarchical optimization strategy, where the shape model is deformed in a coarse-to-fine manner (Section 3.2.2).

3.1. Multilevel local region-based sparse shape composition

Here, we describe how to learn our proposed MLR-SSC from training images. But we need to build a multilevel shape division beforehand.

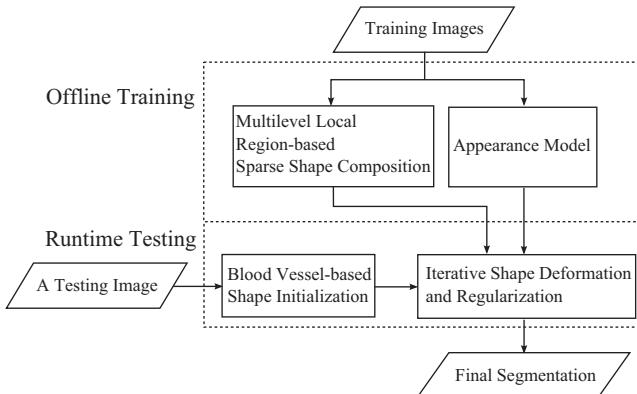


Fig. 2. The main workflow of our method for liver segmentation, which consists of offline training and runtime testing stages.

Assume that we have a set of n_s corresponding triangular liver training shapes $\{\mathcal{M}_i | i = 1, 2, \dots, n_s\}$, and each shape \mathcal{M}_i is represented by a shape vector $\mathbf{d}_i = (\mathbf{v}_{i1}^T, \mathbf{v}_{i2}^T, \dots, \mathbf{v}_{in_p}^T)^T$, where $\mathbf{v}_{ij} = (x_{ij}, y_{ij}, z_{ij})$ denotes the j th point of shape \mathcal{M}_i , and n_p is the number of landmark points. After spatially aligning all training shapes into a common coordinate frame using the generalized Procrustes analysis (GPA) [40], the mean liver shape $\bar{\mathcal{M}}$ can be denoted by a shape vector $\bar{\mathbf{d}} = (\bar{\mu}_1^T, \bar{\mu}_2^T, \dots, \bar{\mu}_{n_p}^T)^T = \sum_{i=1}^{n_s} \mathbf{d}_i / n_s$.

3.1.1. Multilevel shape division

In order to capture more local shape statistics in the shape prior model and decrease the computational complexity of the sparse optimization, the liver shapes are decomposed into multiple regions in a multilevel fashion, such that each region has homogeneous shape variation. The shape variation of i th landmark point is defined as the standard deviation of its position on the surface across all training shapes

$$\sigma_i = \sqrt{\frac{1}{n_s - 1} \sum_{k=1}^{n_s} \|\mathbf{v}_{ki} - \bar{\mu}_i\|_2^2}. \quad (3)$$

For each triangle T_i of the surface, its shape variation can be defined as the average shape variation of its three vertices: $\delta_i = \frac{1}{3} \sum_{j \in T_i} \sigma_j$. Fig. 3 illustrates the shape variation over the surface of the liver shape model. We can see that the tips of left and right lobes have more variation than other parts of the liver surface. Because shape correspondence have been established for all training shapes, the multilevel shape division only needs to perform once for the mean liver surface, then the shape division information is propagated to each training liver surface.

We adapt the variational shape approximation (VSA) method [38] to decompose the mean liver shape into multiple homogeneous regions in a multilevel fashion. Specifically, at a given level l of the multilevel shape division, Lloyd's clustering [37] is employed to partition the mean liver shape $\bar{\mathcal{M}}$ into n_l regions $\mathcal{R} = \{R_1, \dots, R_{n_l}\}$, which are best approximated by n_l proxies $\mathcal{P} = \{P_1, \dots, P_{n_l}\}$ that minimize a total approximation error. The main steps of the multilevel shape division is showed in Algorithm 1.

Algorithm 1. Multilevel shape division procedure.

Input: $\bar{\mathcal{M}}$: the mean liver shape, $\{\delta_i\}$: shape variation for each triangle T_i , n_l : the number of regions at level l of the multilevel shape division.

Step 1. Initialization:

Generate a random seed triangle T_i for each region R_i .

Assign each T_i to the proxy P_i : $P_i = (m_i) \leftarrow \delta_i$.

Step 2. Region partitioning:

$R_i \leftarrow \emptyset$; Add each seed triangle T_i to the region R_i : $R_i \leftarrow R_i \cup T_i$.

Initialize a global min-priority queue: $\mathcal{Q} \leftarrow \emptyset$.

Insert each seed triangle T_i 's three neighboring triangles T_j into \mathcal{Q} with: a priority $Priority_{T_j} \leftarrow$ approximation error

$$\mathcal{L}(T_j, P_i) = |\delta_j - m_i|, \text{ and a tag } Tag_{T_j} \leftarrow \text{label } i \text{ of the proxy } P_i.$$

while $\mathcal{Q} \neq \emptyset$ do

$T_j \leftarrow$ Pop the triangle with the smallest approximation error from \mathcal{Q} .

if T_j not labeled then

Add T_j to the region indicated by its tag Tag_{T_j} :

$$R_{Tag_{T_j}} \leftarrow R_{Tag_{T_j}} \cup T_j.$$

Insert T_j 's unlabeled adjacent triangles into \mathcal{Q} with the same tag Tag_{T_j} .

end if

end while

Step 3. Proxy fitting:

Update each proxy $P_i = (m_i)$ as the average shape variation of their associated new region R_i : $m_i \leftarrow \frac{1}{|R_i|} \sum_{T_j \in R_i} \delta_j$.

Update the seed triangle T_i for each new region R_i as the one with minimal approximation error in R_i :

$$T_i \leftarrow \operatorname{argmin}_{T_j \in R_i} \mathcal{L}(T_j, P_i).$$

Step 4. Steps 2 and 3 are repeated with a predefined number of iterations.

Output: The shape division results of the mean liver shape at level l with n_l regions: $\mathcal{R} = \{R_1, \dots, R_{n_l}\}$.

By applying the above procedure to each level l , we obtain a multilevel shape division $\{\bar{\mathcal{M}}_j | l=0, 1, \dots, L-1; j=1, 2, \dots, n_l\}$, where $\bar{\mathcal{M}}_j$ denotes the j th region at level l of the mean liver shape,

L is the number of shape division levels, and n_l is the number of regions at level l ($n_l = 2^l$ in our implementation). So the mean liver shape is the zero level of the liver shape division $\bar{\mathcal{M}}_{01} = \bar{\mathcal{M}}$ with $n_0 = 1$. Note that to avoid inconsistency among neighboring regions along the boundaries, all regions at each level l are generated from scratch instead of subdividing regions of previous level. An example of the multilevel shape division of the mean liver shape with $L=3$ is showed in Fig. 4. After propagating the shape division information to each training liver surface, we derive the multilevel shape division for each training shape $\{\mathcal{M}_j^i | i=1, 2, \dots, n_s\}$.

3.1.2. Multilevel local region-based SSC

Once the multilevel shape division for each training shape is derived, we can build a multilevel local shape repository for each

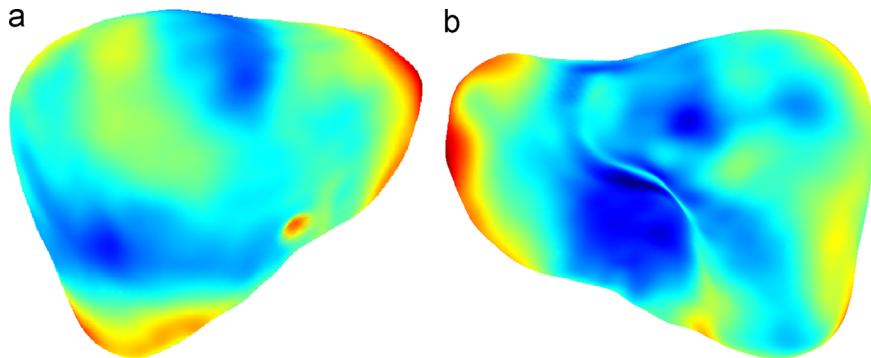


Fig. 3. Color-coded shape variation on the surface of the liver shape model. The variation ranges from small (dark blue) to large (bright red). (a) The anterior view. (b) The posterior view. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

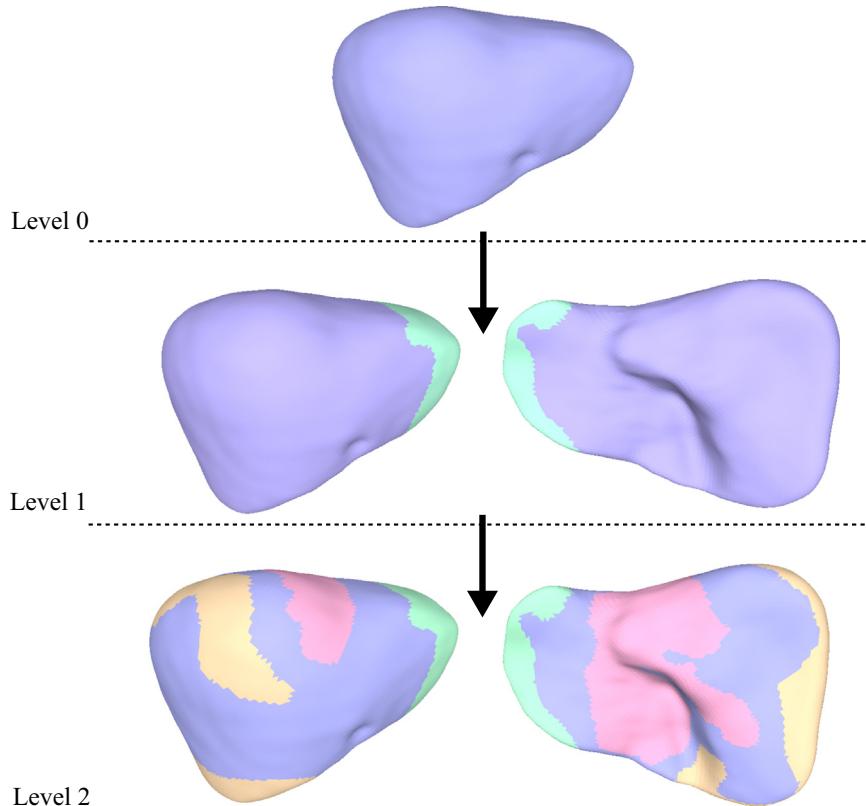


Fig. 4. An example of the multilevel shape division of the mean liver shape with the total number of levels $L=3$, and each level l has 2^l regions indicated by different colors. Left and right columns are the anterior and posterior view, respectively. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

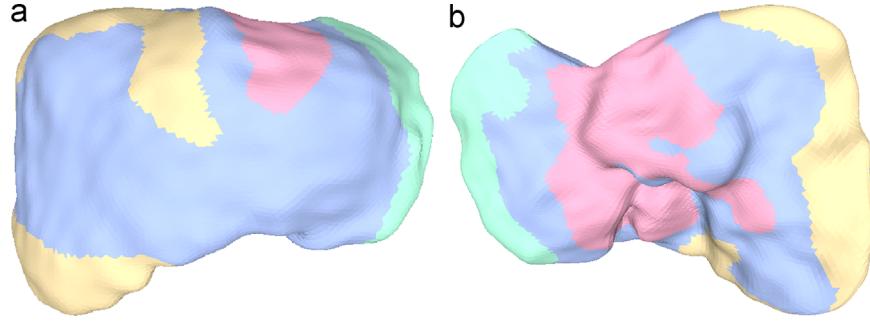


Fig. 5. A shape division example of an input shape with large shape deviation from the mean shape at level $l=2$ with $n_l=4$ regions indicated by different colors. (a) The anterior view. (b) The posterior view. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

region with homogeneous shape variation. Firstly, the same j th regions at level l of all training shapes $\{\mathcal{M}_{lj}^i\mid i=1,2,\dots,n_s\}$ are spatially aligned into a common coordinate frame using the generalized Procrustes analysis (GPA) [40]. A shape vector $\mathbf{d}_{lj}^i \in \mathbb{R}^{3N_l}$ (N_l is the number of vertices of the region at level l) is used to denote the pre-aligned \mathcal{M}_{lj}^i by concatenating the coordinates of all its vertices. Then shape vectors representing the same j th regions at level l of all training shapes are stacked column-wisely to build a local shape repository for j th region at level l : $\mathbf{D}_{lj} = [\mathbf{d}_{lj}^1, \mathbf{d}_{lj}^2, \dots, \mathbf{d}_{lj}^K] \in \mathbb{R}^{3N_l \times K}$, where K is the number of training shapes. The overall shape repository at level l \mathbf{D}_l is thus defined as the concatenation of all the region-specific shape repositories: $\mathbf{D}_l = [\mathbf{D}_{l1}, \mathbf{D}_{l2}, \dots, \mathbf{D}_{ln_l}] \in \mathbb{R}^{(3N_l \times K) \times n_l}$.

Given an input shape to be refined, we firstly establish its shape correspondence with the mean liver shape and align it to the mean shape using Procrustes analysis, then we obtain its multi-level shape division by propagating shape division information of the mean liver surface. Fig. 5 shows a shape division example of an input shape with large shape deviation from the mean shape at level $l=2$ with $n_l=4$ regions. We can see that even though this liver case shows large shape variation, its shape division correlates well with that of the mean shape at the same level as shown in Fig. 4. The input shape at level l can then be represented by a shape vector $\mathbf{y}_l = [\mathbf{y}_{l1}, \mathbf{y}_{l2}, \dots, \mathbf{y}_{ln_l}]$, where $\mathbf{y}_{lj} \in \mathbb{R}^{3N_l}$ denotes j th region at level l of the input shape. Based on the same two sparsity properties of the input shape as in SSC method (i.e., the input shape can be approximately represented by a sparse linear combination of existing training shapes in a shape repository; the input shape may include sparse gross errors), we can refine the input shape at level l \mathbf{y}_l by solving the following extended ℓ_1 -minimization problem:

$$(\hat{\mathbf{x}}, \hat{\mathbf{e}}) = \underset{\mathbf{x}, \mathbf{e}}{\operatorname{argmin}} \|\mathbf{x}\|_1 + \lambda \|\mathbf{e}\|_1 \quad \text{s.t.} \quad \|\mathbf{y}_l - \mathbf{D}_l \mathbf{x} - \mathbf{e}\|_2^2 \leq \epsilon, \quad (4)$$

where $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n_l}]$ and $\mathbf{x}_j \in \mathbb{R}^{3N_l \times K \times n_l}$ denotes the coefficients of linear combination for \mathbf{y}_{lj} , $\mathbf{e} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{n_l}]$ and $\mathbf{e}_j \in \mathbb{R}^{3N_l}$ represents the sparse gross errors for \mathbf{y}_{lj} , λ is a positive regularization parameter that controls the trade-off between sparsity of \mathbf{x} and \mathbf{e} , and ϵ is the reconstruction error tolerance. Then the input shape at level l \mathbf{y}_l is refined as $\mathbf{D}_l \hat{\mathbf{x}}$ and transformed to its original coordinate frame. Eq. (4) is an extension of the minimization problem presented by Nguyen and Tran [41]. Essentially we use the following data model:

$$\mathbf{y}_l = \mathbf{D}_l \mathbf{x} + \mathbf{e} + \epsilon. \quad (5)$$

Instead of only assuming small dense Gaussian noise ϵ as in PCA-based SSMs and the SSC model employed by Shao et al. [36] (i.e., Eq. (2)), we also explicitly model the outliers with the sparse error term \mathbf{e} , which are often caused by misleading appearance cues. When Eq. (4) is optimized, $\hat{\mathbf{e}}$ will capture the sparse outliers if present, otherwise, the ℓ_2 -norm will account for the dense Gaussian noise with bounded energy ϵ and all entries in $\hat{\mathbf{e}}$ will be zero.

However, a drawback of Eq. (4) is that to refine a specific region of the input shape at level l \mathbf{y}_{lj} , the overall shape repository at level l \mathbf{D}_l is employed. To sparsely represent \mathbf{y}_{lj} , thus, training shapes that belong to region-specific shape repositories other than \mathbf{D}_{lj} could be selected. Also, the sparse optimization problem has low computational efficiency, especially when the number of regions at level l n_l is large. We can assume that each region of the input shape has the same two sparsity properties as the global input shape. Therefore, we can divide the above sparse optimization problem (Eq. (4)) defined on all regions at level l into n_l subproblems and refine the input shape in a region-by-region manner. This is the key idea behind our proposed MLR-SSC. Since local region-based shape statistics often lie in a more compact subspace than that of global shape, it can describe the detailed local shape information more accurately. Moreover, because the number of vertices of each region becomes smaller, the computational efficiency of the sparse optimization will be improved. To refine j th region of the input shape at level l \mathbf{y}_{lj} , we solve the following extended ℓ_1 -minimization problem by using the region-specific shape repository \mathbf{D}_{lj} instead of \mathbf{D}_l :

$$\begin{aligned} (\hat{\mathbf{x}}_j, \hat{\mathbf{e}}_j) &= \underset{\mathbf{x}_j, \mathbf{e}_j}{\operatorname{argmin}} \|\mathbf{x}_j\|_1 + \lambda \|\mathbf{e}_j\|_1 \\ \text{s.t. } &\|\mathbf{y}_{lj} - \mathbf{D}_{lj} \mathbf{x}_j - \mathbf{e}_j\|_2^2 \leq \epsilon. \end{aligned} \quad (6)$$

Then \mathbf{y}_{lj} is refined as $\mathbf{D}_{lj} \hat{\mathbf{x}}_j$. Eq. (6) is the objective function of our proposed MLR-SSC. Because the ℓ_1 -norm is a sparsity-inducing norm, this optimization problem will find the sparsest solution (i.e., most entries in $\hat{\mathbf{x}}_j$ and $\hat{\mathbf{e}}_j$ will be zero). The ℓ_2 -norm is used to model the dense Gaussian noise with bounded energy ϵ . Therefore, MLR-SSC is robust to handle both large sparse error and small dense Gaussian noise, complex shape variation, and detailed local shape information recovery. An example of refining the first region at level 1 of an input shape \mathbf{y}_{11} using MLR-SSC is illustrated in Fig. 6.

By defining $\mathbf{B}_j = [\mathbf{D}_{lj}, \frac{1}{\lambda} \mathbf{I}] \in \mathbb{R}^{3N_l \times (K+3N_l)}$ (\mathbf{I} is the $3N_l \times 3N_l$ identity matrix) and $\mathbf{z}_j = [\mathbf{x}_j^T, \lambda \mathbf{e}_j^T]^T \in \mathbb{R}^{K+3N_l}$, we can reformulate Eq. (6) as the Basis Pursuit Denoising (BPDN) problem [42] which only deals with one variable \mathbf{z}_j

$$\hat{\mathbf{z}}_j = \underset{\mathbf{z}_j}{\operatorname{argmin}} \|\mathbf{z}_j\|_1 \quad \text{s.t.} \quad \|\mathbf{y}_{lj} - \mathbf{B}_j \mathbf{z}_j\|_2^2 \leq \epsilon \quad (7)$$

Many optimization algorithms have been proposed to solve this sparse representation problem and stable recoverability of both \mathbf{x}_j and \mathbf{e}_j can be achieved [43,41]. In consideration of both efficiency and robustness, we use the LARS-Homotopy method [34] to solve the BPDN problem in this paper, whose computational complexity is $O(N_l K)$ [43]. Once each region of the input shape has been independently refined, the results are combined to form a whole refined input shape, which is then transformed to its original coordinate frame. Because the input shape is refined in a region-by-region manner, to ensure a smooth transition between

neighboring regions in the refined input surface, we extend each region to overlap neighboring regions along the boundaries, and average the refined regions in these overlapping areas.

3.2. Hierarchical liver segmentation framework

So far we have made appropriate preparations: we build a multilevel shape division and a multilevel local shape repository. The only thing remaining is to design an optimization algorithm to solve the objective function of shape model deformation. To make our ASM-based segmentation framework more efficient and more robust to local minima, we present a hierarchical deformable shape optimization strategy. Our segmentation framework consists of two main components: liver shape initialization and a hierarchical optimization algorithm.

3.2.1. Liver shape initialization

In order to derive a more patient-specific initial shape, we propose a blood vessel-based liver shape initialization method. The key idea behind this method is that hepatic vessels (i.e., portal vein and hepatic vein) of each liver tissue provide unique prior information in terms of its shape and size. Since in the portal

phase of abdominal CT images, hepatic vessels have higher contrast than liver parenchyma, they are relatively easy to segment. Specifically, the main branches of hepatic vessels are firstly extracted from the portal phase of abdominal CT images, then the core region of the liver shape is derived by computing the convex hull of these vessels, finally global SSC is employed to infer an initial liver shape.

The main steps of liver shape initialization are as follows:

- Vessel segmentation:** A median filter is applied to the input CT images to reduce the noise, and Frangi's vesseness measure [44] is employed to enhance the hepatic vessels. Because we only need rough hepatic vessels to initialize the liver shape, a simple confidence connected region growing (CCRG) method is used to segment them. Two seeds are manually selected at the main branches of portal vein and hepatic vein, respectively. Fig. 7a shows an initial segmentation result for an input CT volume, from which we can see that the result leaks into neighboring organs. To remove these organs, we manually cut off neighboring organs at the entrances to portal vein and hepatic vein, respectively. Fig. 7b shows the final segmented hepatic vessels.

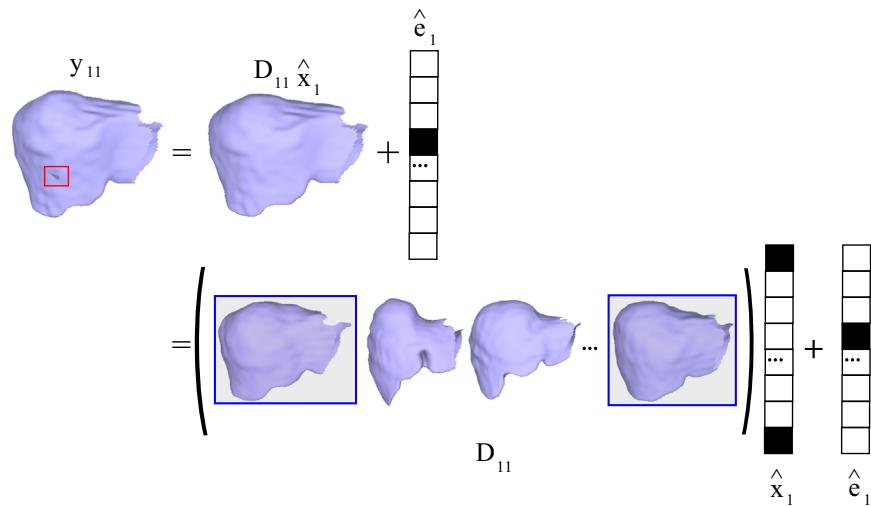


Fig. 6. An example of refining the first region of an input shape at level 1 \mathbf{y}_{11} as $\mathbf{D}_{11}\hat{\mathbf{x}}_1$ using our multilevel local region-based sparse shape composition (MLR-SSC). Filled-in blocks in $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{e}}_1$ indicate nonzero entries. Therefore, both $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{e}}_1$ are sparse in this case, which means \mathbf{y}_{11} can be approximately represented by a sparse linear combination of existing training shapes in shape repository \mathbf{D}_{11} (indicated by blue rectangles in \mathbf{D}_{11}), and \mathbf{y}_{11} contains sparse gross errors (indicated by a red rectangle in \mathbf{y}_{11}). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

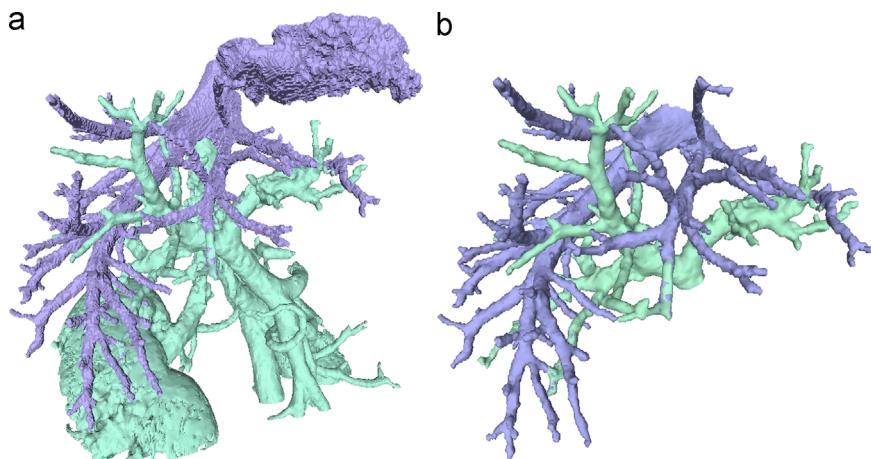


Fig. 7. An example of segmenting the hepatic vessels using confidence connected region growing (CCRG) method. (a) The initial result. (b) The final result after cutting off the neighboring organs.

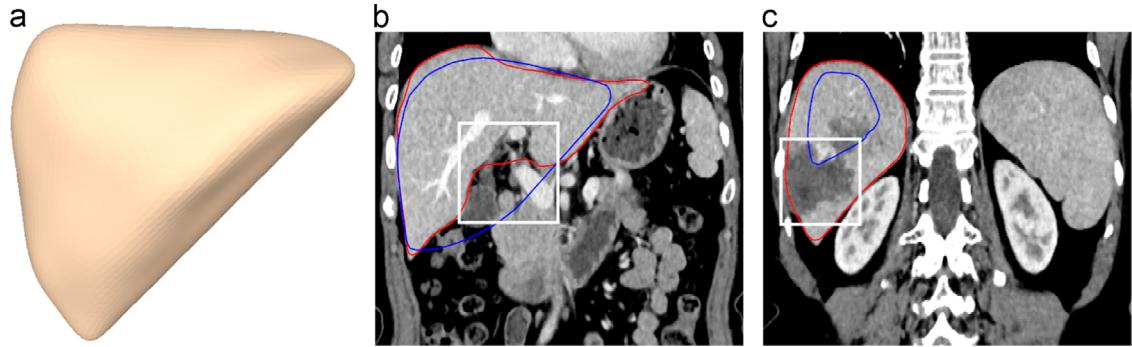


Fig. 8. Examples of the core regions for three different liver cases. (a) 3-D view of the core region derived from extracted hepatic vessels of Fig. 7b. (b) and (c) 2-D images of two typical inaccurate core regions with large gross errors that locate at the entrance to portal vein and regions with large tumors (indicated by the white rectangles), respectively. The red and blue contours show the ground truth and the core region, respectively. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

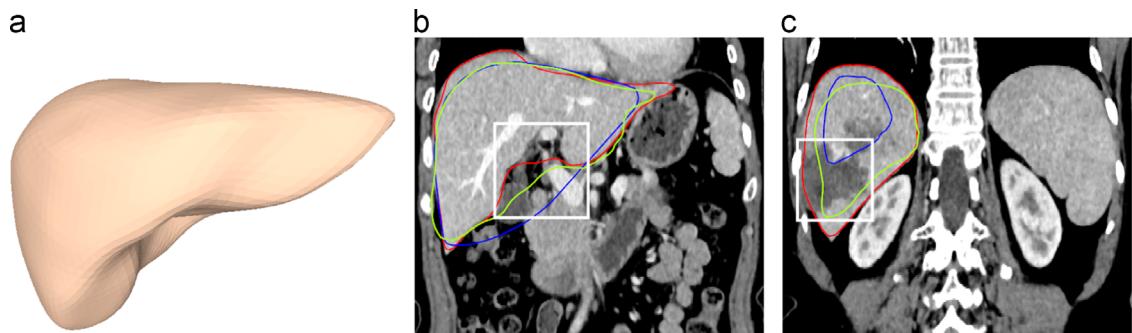


Fig. 9. Graph indicating the effect of the shape inference by using global SSC on the final liver shape initialization. (a) 3-D view of the initial shape inferred from the core region of Fig. 8a by using global SSC. (b) and (c) 2-D images of the initial shapes inferred from the core regions of Figs. 8b and c, respectively. The red and yellow contours show the ground truth and the inferred initial shape, respectively. For illustration purpose, the original core regions are also shown in blue. All the inferred initial shapes are more accurate and more consistent with the ground truth, especially in regions indicated by the white rectangles. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

- (ii) *Deriving the core region of the liver shape:* The extracted hepatic vessels are converted into a triangular mesh using marching cubes method [45], then the core region of the liver shape is derived by computing the convex hull of these vessels using the Quickhull algorithm [46]. The convex hull of a set of points is the smallest convex set that contains the points. Fig. 8 shows the examples of the derived core regions for three different liver cases.
- (iii) *Shape inference using global SSC:* As shown in Fig. 8, the derived core regions of the liver shape are not good enough for shape initialization. In particular, the core regions (shown in Figs. 8 b and c) have large gross errors that locate at the entrance to portal vein and regions with large tumors. These gross errors can result in over- and under-segmentation in the final results, due to the ASM's local search strategy. Therefore, based on the same two sparsity properties of the input shape as in SSC method, we use our proposed MLR-SSC to infer a more accurate liver shape. In order to avoid over-fit problems and make the results more robust, we employ the global shape-based SSC. Specifically, we set the level of MLR-SSC $l=0$ with $n_l=1$ region, and the core region (pre-corresponded and aligned with the mean liver shape using Procrustes analysis) is refined by the global SSC. Fig. 9 shows the effect of shape inference using global SSC on the final liver shape initialization. It can be observed that, all the inferred initial shapes are more accurate and more consistent with the ground truth than the original core regions, especially in regions indicated by the white rectangles.

3.2.2. Hierarchical optimization strategy

We describe a hierarchical optimization strategy where the shape model is deformed in a coarse-to-fine manner. And our proposed MLR-SSC inherently supports this multi-resolution strategy.

Specifically, multi-resolution Gaussian pyramids of all training volumes are firstly constructed, such that MLR-SSC at different levels correspond to volumes at different resolutions. Then separate appearance models based on normalized gradient profiles are learned for volumes at different resolutions as in [39]. Given an input CT volume to be segmented, we also build a multi-resolution Gaussian pyramid for it. At the beginning of ASM search, we set the level of MLR-SSC $l=0$ with $n_l=1$ region (i.e., the global shape-based SSC), then we gradually increase the level l with more regions, it becomes the local SSC, and more detailed local shape information can be reconstructed. The ASM search starts at the lowest resolution input volume using MLR-SSC and appearance model corresponding to that level, and moves to a higher resolution input volume after convergence or a predefined number of iterations [39]. And the segmentation result of each level is directly employed as initialization for the next level. The process is repeated until convergence is achieved on the original input volume in the pyramid. The details of our hierarchical liver segmentation framework are showed in Algorithm 2.

Algorithm 2. Hierarchical liver segmentation framework.

Input: \mathcal{V} : an input portal phase of abdominal CT volume,
 L_{max} : number of resolution levels,
 N_{max} : number of iterations per resolution level.

```

Output: Segmentation result  $\mathcal{M}_{(L_{max}-1)}$ .
 $\mathcal{M}_0 \leftarrow$  Blood vessel-based liver shape initialization.
 $\{\mathcal{V}_0, \mathcal{V}_1, \dots, \mathcal{V}_{(L_{max}-1)} = \mathcal{V}\} \leftarrow$  Construction of a multi-resolution Gaussian pyramid for  $\mathcal{V}$ .
for  $l = 0$  to  $(L_{max}-1)$  do
  for  $i = 0$  to  $(N_{max}-1)$  do
    for all landmark points  $\mathbf{v}_j \in \mathcal{M}_l$  do
       $\mathbf{q}_j \leftarrow$  Find the target point for  $\mathbf{v}_j$  in  $\mathcal{V}_l$  using local ASM search based on appearance model corresponding to level  $l$ .
       $\mathbf{v}_j \leftarrow \mathbf{q}_j$ .
    end for
     $\mathcal{M}_l \leftarrow$  Refine the intermediate deformed shape using MLR-SSC at level  $l$  with  $n_l$  regions in a region-by-region manner.
     $i \leftarrow i+1$ .
  end for
   $l \leftarrow l+1$ .
end for
return Segmentation result  $\mathcal{M}_{(L_{max}-1)}$ .

```

4. Experiments

4.1. Datasets

To evaluate the accuracy and performance of our proposed method, we have applied it to a set of 60 portal phase contrast-enhanced CT scans of the abdomen, coming from two publicly available databases and one non-public database. The first public database, SLIVER07 (<http://www.sliver07.org>), consists of 30 CT scans provided by the organizers of the Segmentation of the Liver Competition 2007 (SLIVER07) [2]. SLIVER07 includes 20 training data (SLIVER07-Train) with corresponding ground truth, and 10 test images (SLIVER07-Test) without ground truth. Most datasets in this database were pathological, including tumors, metastases and cysts of different sizes [2]. The second public database, 3Dircadb1 (<http://www.ircad.fr/softwares/3Dircadb/3Dircadb1/index.php?lng=en>), consists of 20 CT scans with corresponding ground truth provided by IRCCAD, the French Research Institute against Digestive Cancer. 75% of the cases in this database contained pathologies, mainly including hepatocellular carcinoma (HCC), metastases and hemangiomas. The CT scans in these two databases were acquired using a variety of different CT scanners from different vendors. The Non-Public database consists of 10 CT scans from different subjects with corresponding expert segmentations manually annotated by our clinical partner. 10 hepatic CT cases were acquired using a 64-detector row and dual-source CT scanner (SOMATOM Definition Flash; Siemens Healthcare, Forchheim, Germany), and 3 cases were pathological, including HCC, cyst and hemangioma. The CT scans from this database featured a relatively higher image resolution. The detailed information of all the liver datasets used in this study is given in Table 1.

Table 1

The detailed information of all the liver datasets used in this study.

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
3Dircadb1	20	512×512	0.56–0.86	74–260	1.0–4.0
Non-Public	10	512×512	0.58–0.67	184–221	1.0

In this study, the SLIVER07-Test database was only used for the liver segmentation competition and comparing our method with state-of-the-art methods. Considering that SLIVER07-Train, 3Dircadb1, and Non-Public databases were acquired using different protocols and from different centers, to minimize training bias towards a certain center and make the method comparisons more meaningful, we employed the stratified holdout method [47] to obtain both training and test data. Specifically, we stratified the selection of 20 training and 30 test datasets from these three sources according to their relative proportions (i.e., 2:2:1). To make the selection more representative of the population, we also stratified the selection of training and test data from each database according to the ratio of normal to pathological cases in that database. We denote the final derived databases for training and testing as Hybrid-Train and Hybrid-Test, respectively. Note that when comparing our method with state-of-the-art methods based on the 3Dircadb1 (Section 5.4.1) and SLIVER07-Test (Section 5.4.2) databases, we only used SLIVER07-Train database as training data.

4.2. Evaluation metrics and statistical analysis

As in the Segmentation of the Liver Competition 2007 (SLIVER07) [2], to quantitatively evaluate the performance of our proposed method, we compared the segmentation result with the expert-generated ground truth according to the following five volume and surface based metrics: 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 [2]. The volume and surface based metrics are given in percent and millimeters, respectively. For all these evaluation metrics, the smaller the value is, the better the segmentation result.

Differences in method segmentation accuracy (our method differing from other methods) are assessed using paired *t*-tests. $p < 0.05$ is considered as the significant level for all statistical tests. The null hypothesis is that the difference in the mean values of the same evaluation metric between our method and the compared methods is zero. Note that all statistical hypothesis tests are based only on VOE and ASD, because among all the employed metrics, these two are the most commonly used measures to evaluate segmentation accuracy [2].

4.3. Implementation details

We used the SPHARM-PDM package [48] and the Minimum Description Length (MDL) method implemented by Heimann et al. [49] to establish shape correspondence between different training shapes. In this procedure, all the training samples were resampled to an isotropic resolution of 1.5 mm, and any holes fully enclosed in the training samples were filled to ensure spherical topology. One of the training samples judged as most similar to the mean shape was chosen as the template for the Procrustes analysis. The same procedure was also employed to establish shape correspondence between the mean liver shape and the input shapes to be refined. We modified the VSA method implemented in MEPP package [50] to decompose the liver shapes into multiple homogeneous regions in a multilevel fashion.

The parameter settings for the hierarchical liver segmentation framework were determined offline using the SLIVER07-Train database to obtain the optimal performance. These settings were the same for all the test data in our implementation. All the parameter settings deployed in our method are listed in Table 2. The procedure for choosing the value of sparsity parameter λ in Eq. (6), the number of shape division levels L and resolution levels L_{max} will be discussed in detail in Sections 5.1.3 and 5.1.4, respectively.

Table 2

Parameter settings for the proposed hierarchical liver segmentation framework.

Parameter	Value	Detailed information
n_p	2562	Number of landmark points
L	5	Number of shape division levels
λ	100	Sparcity parameter in Eq. (6)
L_{max}	5	Number of resolution levels in the hierarchical optimization strategy
N_{max}	10	Number of iterations per resolution level
k	9	Number of sample points in a normalized gradient profile on either side of the landmark point
n_e	6	Number of new positions to evaluate on either side of the landmark point during each iteration

In this study, we also compared our proposed MLR-SSC model with two closely related methods: (1) SSM, statistical shape model [16] uses principal component analysis (PCA) [18] to learn the shape prior models (see Appendix A for more details). In our implementation, the used modes of variation captured 98% of the total shape variance in the dataset; (2) SSC, sparse shape composition [29] uses global training shapes in the shape repository \mathbf{D} to refine an input shape (see Eq. (1)). It is just a special case of our proposed method when the number of shape division levels $L = 1$.

All methods were implemented in C++¹ on Ubuntu platform and tested on a desktop PC with a 2.33 GHz Intel quad-core processor and 6 GB of RAM. To solve the sparse representation problems in our proposed method, we deployed the C++ version of the SPAMS library [43], which is an open source optimization toolbox for solving various sparse representation problems.

5. Results

5.1. Evaluation of the MLR-SSC Model

Before applying our proposed MLR-SSC model to segment liver from CT images, we evaluate its performance based on the SLIVER07-Train database using generalization ability measure under different circumstances. Note that in all experiments conducted in this section, we only used the MLR-SSC at level $l = 2$ with $n_l = 4$ regions.

5.1.1. Generalization ability

Generalization ability, specificity and compactness are the three most used measures for quantifying the quality of a shape modeling method [51]. In this study, we only employed the generalization ability to evaluate the quality of our proposed method, because the other two are not applicable to our method. The generalization ability quantifies the capability of the model to represent unknown shapes of the same object class. It can be measured by performing a series of leave-one-out tests on the training set, where the difference between the omitted training shape S_i and the reconstructed shape S_i^* of S_i is measured. The final result is derived by averaging over the complete set of tests.

To quantify the difference between two corresponding shapes, we employed the binary volume based metrics instead of the commonly used landmark based metrics, since volume based metrics are independent of the underlying landmark distributions and allow for more precise measurement of global and local shape differences [13]. Specifically, the reconstructed shape S_i^* is firstly converted to a binary volume V_i^* , and the difference between it and the omitted training shape S_i is computed as the average

Table 3

The mean and standard deviation of the generalization ability and computation time of the three different shape prior modeling methods for the SLIVER07-Train database.

Method	Generalization ability (mm)	Time (s)
SSM	$2.87 \pm 0.46^*$	0.629 ± 0.035
SSC	$2.69 \pm 0.41^*$	0.210 ± 0.005
MLR-SSC	1.19 ± 0.20	0.208 ± 0.004

Bold values are the best result in that column.

* This indicates a statistically significant difference between the marked result and the corresponding one of MLR-SSC at a significance level of 0.05.

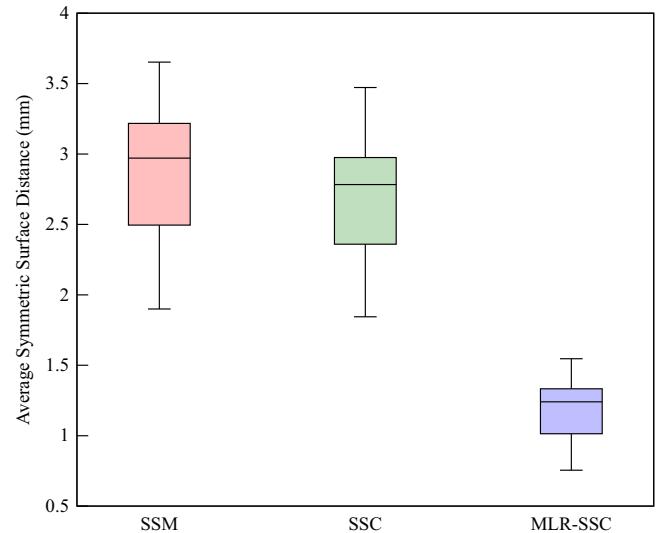


Fig. 10. Box plots of the generalization ability for the three different shape prior modeling methods.

symmetric surface distance (ASD) between V_i^* and the binary training volume V_i corresponding to S_i .

Thus, the generalization ability can be defined as

$$\mathcal{G} = \frac{1}{n_s} \sum_{i=1}^{n_s} ASD(V_i^*, V_i). \quad (8)$$

The smaller the value of the generalization ability, the better the shape modeling method. The mean and standard deviation of the generalization ability for the SLIVER07-Train database were calculated. We also tested our model against two other shape modeling methods: SSM and SSC.

Quantitative comparisons of the generalization ability are shown in Table 3 and Fig. 10. These indicate that our method achieves the best generalization ability, and the performance differences between our method and the other two are statistically significant ($p < 0.05$). Table 3 also shows the time needed to reconstruct each test shape. Our method is highly efficient and only takes 0.208 s to reconstruct a test shape, while SSM needs two times more for the same task. Compared with the time needed to refine each test shape in the original SSC [29] using the interior point methods, which is about 8 s (using the author's Python implementation), our method shows much better time efficiency.

5.1.2. Effect of the size of training data on the generalization ability

The main motivation of our proposed MLR-SSC model is to overcome the over-restrictive issues of traditional shape modeling methods when only a limited number of training samples is available, especially for volumetric images where the manual segmentation is difficult to derive. To evaluate the sensitivity of

¹ The source code is available at <https://github.com/ivanshih/MLR-SSC>.

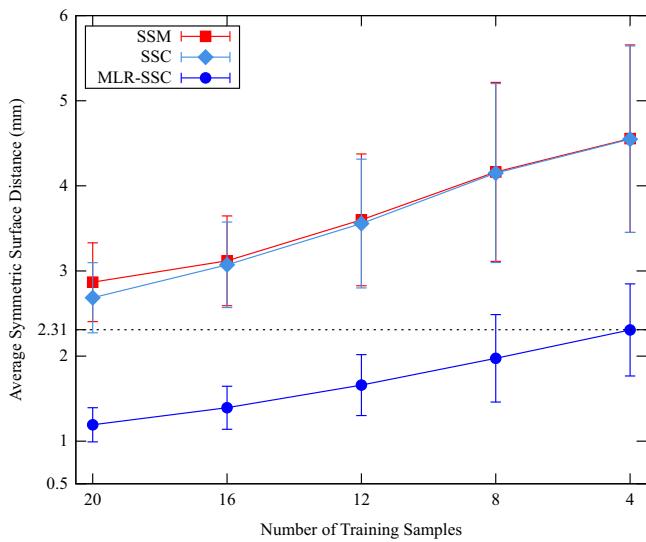


Fig. 11. Generalization ability of the three different shape prior modeling methods with different number of training samples.

our method with respect to the size of training data, we have tested its generalization ability in a series of experiments where the number of training samples was decreased gradually (i.e., from 20 to 4). We also tested our method against SSM and SSC.

Fig. 11 shows the results of the generalization ability of all the compared methods with different number of training samples. The value of the generalization ability rises when the number of training samples decreases for all the compared methods, but our method's rate of change is smaller than the other two methods. When the number of training samples is four, our method's generalization ability is 2.31 mm, which is even smaller than SSM and SSC's generalization ability (2.87 mm and 2.69 mm, respectively) when all the 20 training samples are employed. Therefore, our method can achieve reasonable performance even with as few as four training samples.

5.1.3. Influence of the sparsity parameter λ on the generalization ability

An important implementation detail in our proposed MLR-SSC model is the choice of the sparsity parameter λ , which balances the sparsity of \mathbf{x} and \mathbf{e} in Eq. (6). It is recommended to set $\lambda = \sqrt{\frac{n}{m(\log n)^{1/2}}}$

as a “good-for-all” parameter in [41], where m is the dimension of the shape vectors in the shape repository, and n is the number of training shapes. In this study, we choose to determine its value empirically. Specifically, we have tested the effect of the sparsity parameter λ on the generalization ability of our method on test shapes with and without sparse gross errors.

Fig. 12 shows the results of the generalization ability of our method with different choices for the sparsity parameter λ (i.e., from 1 to 400). From this we can see that the choice of the sparsity parameter λ is critical for the success of our method. If it is too small, the error term \mathbf{e} will have more priority than the coefficients term \mathbf{x} to be regularized. Because the larger column vectors in \mathbf{B} (i.e., $\frac{1}{\lambda}\mathbf{I}$) will result in smaller coefficients, that is less ℓ_1 costs. If it is too large, the model cannot retrieve any gross errors. Fortunately, our method achieves stable results in a wide range of values (i.e., from 10 to 170), which means our method is not sensitive to the values of the sparsity parameter λ in a certain large range. Therefore, we choose a fixed value of $\lambda=100$ for all the test data in our implementation.

5.1.4. Influence of the number of levels and regions in MLR-SSC on liver shape reconstruction

The main feature of the proposed MLR-SSC model is to preserve the detailed local shape information even if it is not statistically significant in the training data, thanks to the use of multilevel shape division. The choice of the number of regions in each level of MLR-SSC thus plays an important role in achieving this goal. If it is too small, MLR-SSC will only have little ability to reconstruct the finer local details. If it is too large, lots of inconsistency will occur at areas where different subdivisions intersect, resulting in highly rough liver surfaces. To test the sensitivity of our method with regard to the choice of number of regions, we have evaluated its generalization ability with a varying number of regions (i.e., even numbers between 2 and 32). Fig. 13 shows the results of the generalization ability with different choices of number of regions. We can see that at first as the number of regions increases, the value of the generalization ability drops, but after 20 regions, it begins to rise, and our method achieves the best results in the range of 12–20 regions.

Considering that all datasets employed in this study have an in-plane size of 512×512 , the maximum possible number of

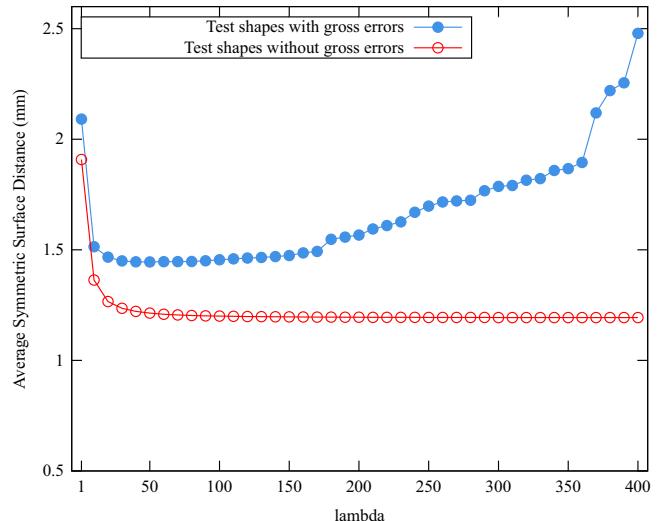


Fig. 12. Generalization ability of our proposed shape prior modeling method with different choices for the sparsity parameter λ (from 1 to 400) on test shapes with and without sparse gross errors.

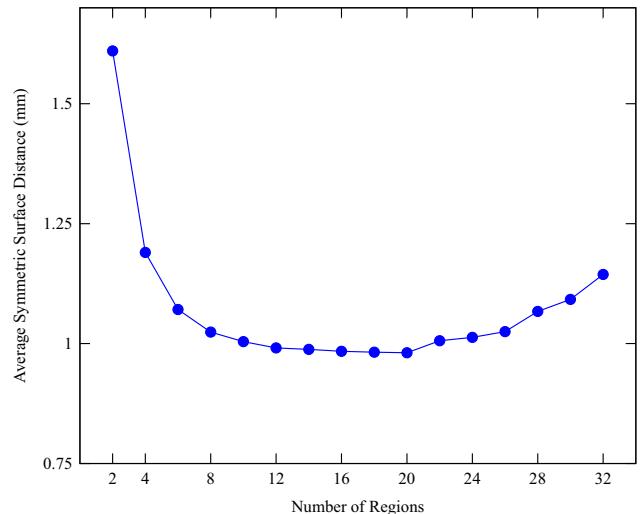


Fig. 13. Influence of the number of regions in MLR-SSC on liver shape reconstruction.

resolution levels that can be chosen in our hierarchical segmentation framework is 6 (i.e., level 0–5). Because input volumes in the next level (i.e., level 6) of the Gaussian pyramid have an in-plane size of 8×8 , which is too low to have any details about the liver tissue to be segmented. Also in our hierarchical segmentation framework, MLR-SSC at different levels correspond to input volumes at different resolutions (i.e., input volumes at level l of the Gaussian pyramid are segmented using MLR-SSC at level l with $n_l = 2^l$ regions). Therefore, the maximum possible number of levels in MLR-SSC is 6 with a maximum of $2^5 = 32$ regions.

Regarding the above-mentioned aspects, we choose the number of shape division levels $L = 5$ with a maximum of $2^4 = 16$ regions. The chosen number of regions is within the range of 12–20 that achieve the best shape reconstruction results. Accordingly, we select the number of resolution levels in our hierarchical framework $L_{max} = 5$ in our implementation.

5.2. Liver segmentation from CT images

Here, we evaluate our proposed MLR-SSC model in a specific application of segmenting the liver tissue from CT Images based on the Hybrid-Test database to illustrate the effectiveness of our

method. Specifically, MLR-SSC was used to refine the intermediate deformed shape during the hierarchical optimization procedure. We also compared our model with two closely related methods: SSM and SSC. To make a fair comparison, the same hierarchical liver segmentation framework is used in all methods, i.e., the same shape correspondence establishment, liver shape initialization, and hierarchical optimization strategy. The only difference is the shape prior modeling method used to refine the intermediate deformed shapes. We also carefully tuned the parameters for SSM and SSC to achieve optimal performance.

5.2.1. Qualitative results

Fig. 14 compares the segmentation results between shape prior based on SSM and our method in challenging cases where the livers contain severe pathological abnormalities, which include tumors in the right lobe and in the quadrate lobe, respectively. Each column shows one case. The 3-D visualization of errors is based on the average symmetric surface distance (ASD) between the segmentation result and ground truth, in which the red and blue regions indicate over- and under-segmentation, respectively. Both methods can restore parts of the liver tissue with tumors, because liver shapes without these regions do not appear in the

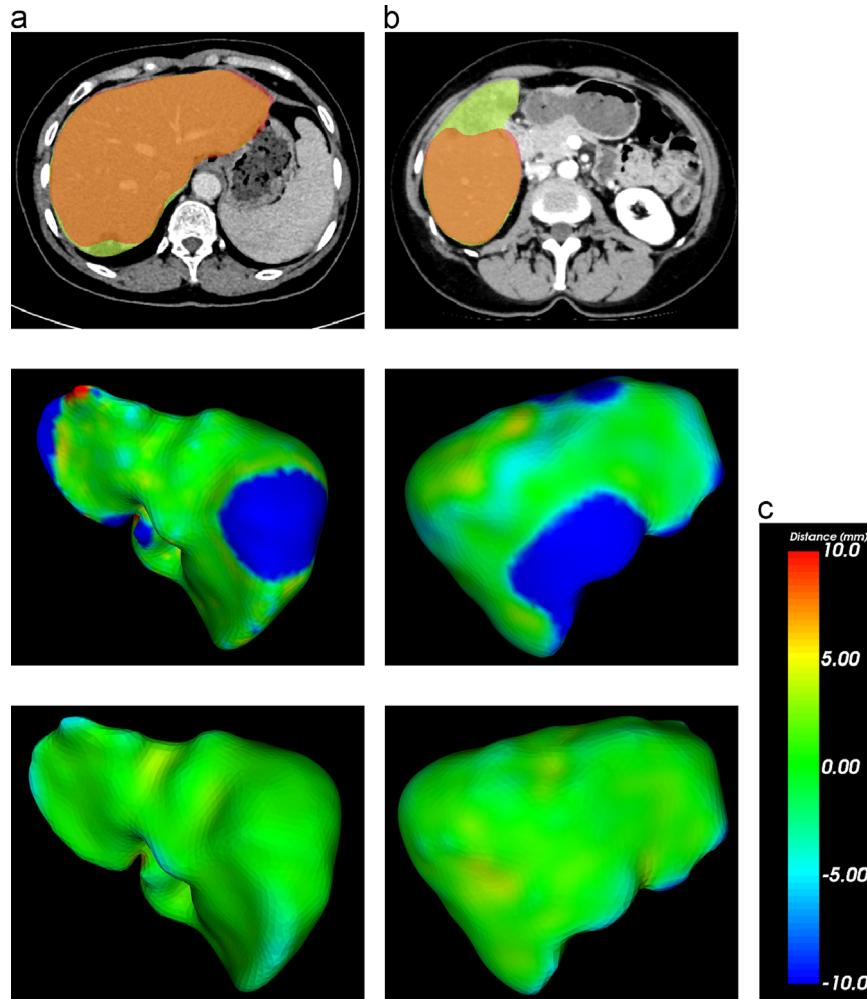


Fig. 14. Visual comparisons of liver segmentation results between shape prior based on SSM and our proposed MLR-SSC method in two challenging cases where the livers contain severe pathological abnormalities. A liver with (a) a large tumor in the right lobe, and (b) a large tumor in the quadrate lobe. Each column shows one case. The first row is the superimposition of results of SSM (drawn as red) and our method (drawn as green) in 2-D slice images, and the overlapping results are drawn as brown. The second row and the third row are the 3-D visualization of average symmetric surface distance (ASD) errors of SSM and our method, respectively. The red and blue regions indicate over- and under-segmentation, respectively. (c) The used distance to color bar. SSM based method cannot restore some local shape details, while our method successfully preserves and reconstructs these finer local details. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

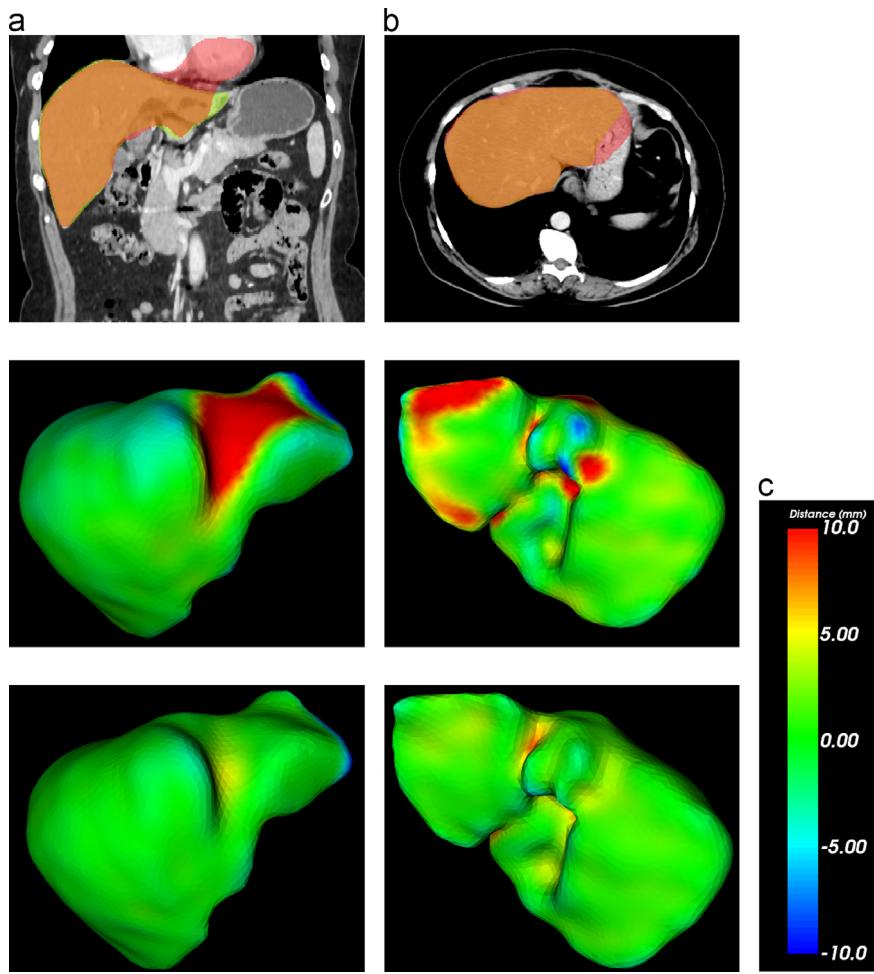


Fig. 15. Visual comparisons of liver segmentation results between shape prior based on SSM and our proposed MLR-SSC method in two challenging cases where the liver tissue has to be separated from neighboring organs with similar intensity values, which include (a) the heart and (b) the stomach. Each column shows one case. The first row is the superimposition of results of SSM (drawn as red) and our method (drawn as green) in 2-D slice images, and the overlapping results are drawn as brown. The second row and the third row are the 3-D visualization of average symmetric surface distance (ASD) errors of SSM and our method, respectively. The red and blue regions indicate over- and under-segmentation, respectively. (c) The used distance to color bar. SSM based method easily leaks into these neighboring organs with similar intensity values, while our method successfully excludes them. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

training samples. However, SSM based method fails to reconstruct some local shape details, because 20 training samples are not sufficient for it to capture the full range of liver shape variability. Consequently, it prevents the shape model from moving to these local details that are beyond the principal modes of shape variation. While our method successfully preserves and reconstructs these finer local details, since it refines the deformed shapes in a region-by-region manner.

Fig. 15 compares the segmentation results between shape prior based on SSM and our method in challenging cases where the liver tissue has to be separated from neighboring organs with similar intensity values, which include the heart and the stomach, respectively. Each column is for one case. SSM based method easily leaks into these neighboring organs with similar intensity values and over-segments the live tissue, while our method successfully excludes them and achieves the least ASD errors. The reason is that in SSM based method, these over-segmented regions are roughly within the principal modes of shape variation. Consequently, these regions can be well reconstructed by the SSM based method during the deformation procedure. By contrast, in our method these over-segmented regions are explicitly modeled and reconstructed as the sparse gross errors. Therefore, our method can exclude these regions from the liver tissue effectively. These

experimental results demonstrate the robustness of our method against both over- and under-segmentation.

Fig. 16 compares the refinement results of intermediate deformed liver shapes between SSC and our method. The results of both methods have a high overlap with the ground truth, but our method delineates the local shape details more accurately, such as the concave region at the entrance to portal vein (indicated by the white rectangles). The reason is that in SSC, the limited number of global training shapes are directly incorporated in the shape repository **D** to refine the deformed shapes. As a result, the detailed local shape information cannot be accurately reconstructed. By contrast, in our method the deformed shapes are decomposed into multiple regions with homogeneous shape variation and refined in a region-by-region manner. Therefore, our method can restore these finer local details more effectively.

5.2.2. Quantitative results

Table 4 shows the quantitative comparative results of the liver segmentation with three different shape prior modeling methods based on the Hybrid-Test database. Results for each measure are represented as mean and standard deviation of the overall data-sets. As shown in Table 4, our method achieves the best performance with the smallest variances according to the deployed

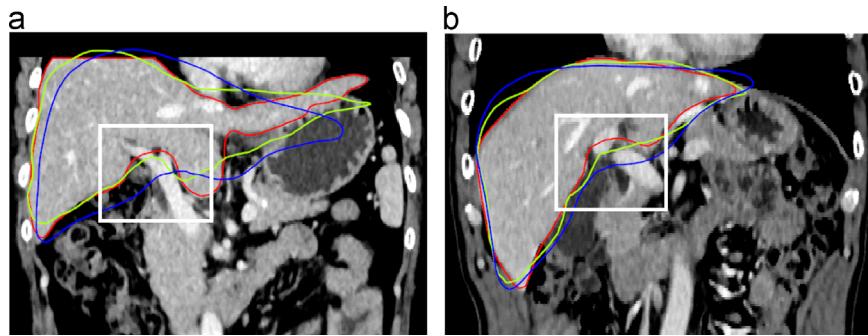


Fig. 16. Visual comparisons of refinement results of intermediate deformed liver shapes between shape prior based on SSC and our proposed MLR-SSC. All images are of 2-D coronal slices. The red contour shows the ground truth, the blue and yellow contours show the refinement results of SSC and our method, respectively. The results of both methods have a high overlap with the ground truth, but our method delineates the local shape details more accurately, such as the concave region at the entrance to portal vein (indicated by the white rectangles). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Table 4

Quantitative comparative results of the liver segmentation with three different shape prior modeling methods based on the Hybrid-Test database.

Method	VOE (%)	SRVD (%)	ASD (mm)	RMSD (mm)	MSD (mm)
SSM	$16.24 \pm 6.94^*$	9.20 ± 9.52	$2.94 \pm 1.38^*$	5.18 ± 2.67	33.64 ± 13.71
SSC	$9.20 \pm 4.31^*$	2.76 ± 4.29	$1.78 \pm 1.11^*$	3.83 ± 2.38	26.58 ± 11.01
Our MLR-SSC	7.05 ± 2.22	1.73 ± 4.33	1.16 ± 0.39	2.33 ± 0.92	22.68 ± 9.27

Results for each measure are represented as mean and standard deviation of the overall datasets. Bold values are the best result in that column for the corresponding database.

* This indicates a statistically significant difference between the marked result and the corresponding one of MLR-SSC at a significance level of 0.05.

Table 5

Quantitative liver segmentation results of our proposed method under different types of pathologies included in the Hybrid-Test database.

Metric	HCC	Metastasis	Cyst	Hemangioma
VOE (%)	7.40 ± 1.98	8.19 ± 2.46	6.95 ± 1.45	6.57 ± 0.53
ASD (mm)	1.23 ± 0.34	1.26 ± 0.38	1.21 ± 0.24	1.18 ± 0.20

Results for each measure are represented as mean and standard deviation of the overall datasets belong to that particular pathology. HCC stands for hepatocellular carcinoma.

measures, indicating its robustness on diverse test data. Also our method outperform the other two compared methods with statistically significant differences ($p < 0.05$). Specifically, the mean ASD and VOE of our method are 1.16 mm and 7.05%, respectively. The mean ASD and VOE of SSM based method are more than two times those of our method, and the mean ASD of SSC based method is nearly 50% higher than that of our method. These indicate that our proposed method is more accurate and robust than the other two methods in the presence of small training sets.

In the previous study using SSM [8,12,13], they used more advanced appearance models such as classifiers [8,13] and clustering techniques [12], so that satisfactory segmentation results were achieved. In the present study, we focus on the shape prior model, so the SSM based method employed the simple appearance model based on normalized gradient profiles. This introduced a low performance of SSM result listed in Table 4.

Table 5 shows our method's quantitative liver segmentation results under different types of pathologies included in the Hybrid-Test database. Results for each measure are represented as mean and standard deviation of the overall datasets belong to that particular pathology. For all the types of pathologies, our method achieves relatively high accuracy according to VOE and ASD, indicating our method's robustness in the presence of pathology.

5.3. Evaluation of the liver shape initialization method

To assess the effectiveness of our proposed blood vessel-based liver shape initialization method, we compared it with mean shape based initialization method based on the Hybrid-Test database. To make a fair comparison, the same hierarchical segmentation framework is used, only the shape initialization module is different. Specifically, Gaussian mixture models (GMM) based non-rigid registration between the mean shape and ground truth of test data is employed to initialize the mean shape, which is implemented in the gmmreg package [52].

Fig. 17 demonstrates the comparisons of shape initialization and segmentation results between mean shape-based method and our method. As shown in Fig. 17, the initialization of mean shape based-method has a low overlap with the ground truth. This method over-segments neighboring organs with similar intensity values, and under-segments long and thin regions (e.g., the left tip of the liver). By contrast, the initial shape of our method is much closer to the ground truth and segmentation result is more accurate, especially in the regions that deviate largely from the mean shape. Fig. 18 also shows the segmentation results of SSM and SSC for the same liver case when starting from our blood vessel-based shape initialization. Compared with the result of our proposed MLR-SSC (i.e., the yellow contour in Fig. 17b), these two modeling methods show lower accuracy, especially for SSM.

Table 6 shows the quantitative comparative results of liver shape initialization and segmentation between mean shape based method and our method based on the Hybrid-Test database. Results for each measure are represented as mean and standard deviation of the overall datasets. For shape initialization results, the accuracy of our method is statistically significantly higher ($p < 0.05$) than that of mean shape based method. However, our method obtains a larger mean SRVD. Specifically, the mean ASD and VOE of our method are 4.83 mm and 23.64%, respectively. For segmentation results, our method performed statistically significantly better ($p < 0.05$) than mean shape based method. The mean ASD and VOE of mean shape based method are nearly two times those of our method. Specifically, the mean ASD and VOE of

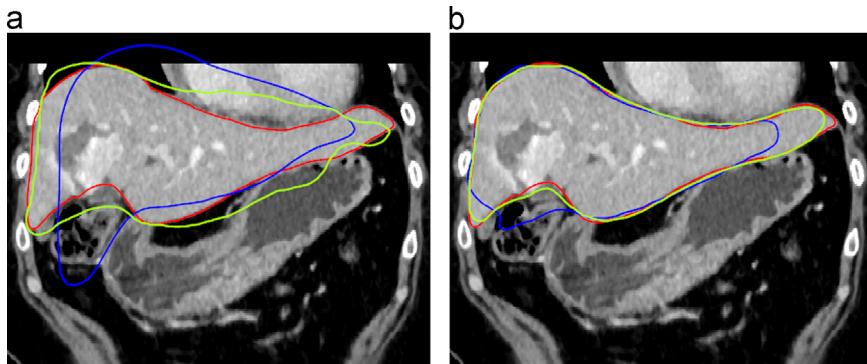


Fig. 17. Visual comparisons of liver shape initialization and segmentation results between mean shape based and our proposed blood vessel-based shape initialization method. All images are of 2-D coronal slices. The red contour shows the ground truth, the blue and yellow contours show the results of mean shape based method and our method, respectively. (a) Initialization results. (b) Segmentation results using the initializations from (a). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

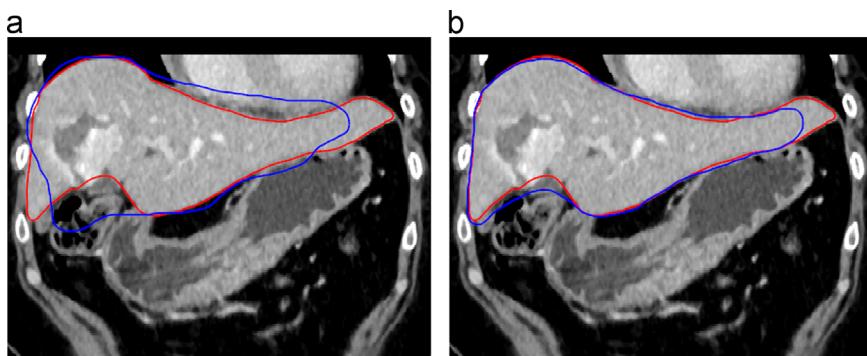


Fig. 18. Segmentation results of SSM and SSC for the same liver case as in Fig. 17 when starting from our proposed blood vessel-based shape initialization (i.e., the yellow contour in Fig. 17a). All images are of 2-D coronal slices. The red and blue contours show the ground truth and the segmentation result, respectively. (a) SSM result. (b) SSC result. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Table 6
Quantitative comparative results of liver shape initialization and segmentation between two different shape initialization methods based on the Hybrid-Test database.

Method	VOE (%)	SRVD (%)	ASD (mm)	RMSD (mm)	MSD (mm)
<i>Mean shape</i>					
Initialization	$29.07 \pm 8.11^*$	9.48 ± 10.87	$6.49 \pm 2.26^*$	9.13 ± 4.08	38.87 ± 19.99
Final result	$13.08 \pm 5.50^*$	12.14 ± 10.00	$2.30 \pm 0.95^*$	4.48 ± 1.59	28.40 ± 16.06
<i>Our method</i>					
Initialization	23.64 ± 5.66	20.74 ± 10.58	4.83 ± 1.15	6.64 ± 1.60	31.68 ± 8.51
Final result	7.05 ± 2.22	1.73 ± 4.33	1.16 ± 0.39	2.33 ± 0.92	22.68 ± 9.27

Results for each measure are represented as mean and standard deviation of the overall datasets. Bold values are the best result in that column.

* This indicates a statistically significant difference between the marked result and the corresponding one of our method at a significance level of 0.05.

mean shape based method are 2.30 mm and 13.08%, respectively. Also the mean SRVD and MSD of mean shape based method are much higher than ours. This is mainly because its segmentation results are easily affected by neighboring organs with similar intensity values and leak into those regions.

Therefore, our proposed method provides a more patient-specific initial shape for the hierarchical segmentation framework and achieves much better segmentation accuracy than the mean shape based method.

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

To assess the performance of our proposed hierarchical liver segmentation framework within the larger context of existing literature, we compare it with state-of-the-art methods in this section.

5.4.1. Evaluation on the 3Dircadb1 database

Here, we compare our hierarchical liver segmentation framework with three recently published model-based methods based on the 3Dircadb1 database: Chung's method [12], Kirschner's method [13] and Erdt's method [14], whose detailed modular information is listed in Table 7.

Table 8 shows the quantitative comparative results of the liver segmentation with all compared methods based on the 3Dircadb1 database. Results for each measure are represented as mean and standard deviation of the overall datasets. As shown in Table 8, paired *t*-test on the segmentation accuracy demonstrates that our results are significantly different from Chung's method ($p < 0.01$), and from Kirschner's method ($p < 0.05$). Our method shows marginal performance improvement ($p > 0.05$) over Erdt's method, even though they used considerably more training samples (220) than ours (20). Also our method obtains smaller variances,

Table 7

The detailed modular information of three compared state-of-the-art model-based liver segmentation methods and our method.

Method	Number of training samples	Shape correspondences	Shape initialization	Shape model	Appearance model	Search algorithm
Chung et al. [12]	23	Currents-based mesh-to-mesh registration	Mean shape-based	Single reference SSM	MPAM (intensity profiles)	Deformable simplex meshes
Kirschner et al. [13]	47	Parameter space propagation + shape similarity tree	Machine learning-based	SSM	Classification based (intensity profiles)	Probabilistic ASM
Erdt et al. [14]	220	Manual mesh-to-volume registration	Machine learning-based	SSM	Local gradient threshold	Multi-tiered ASM
Our method	20	SPHARM + MDL	Blood vessel-based	MLR-SSC	SAM (normalized gradient profiles)	Multi-resolution ASM

SPHARM, MPAM and SAM stand for spherical harmonic, multimodal prior appearance model and statistical appearance model, respectively.

Table 8

Quantitative comparative results between our proposed method and three state-of-the-art model-based liver segmentation methods based on the 3Dircadb1 database.

Method	VOE (%)	SRVD (%)	ASD (mm)	RMSD (mm)	MSD (mm)	Time
Chung et al. [12]	$12.99 \pm 5.04^{**}$	-5.66 ± 5.59	$2.24 \pm 1.08^{**}$	N/A	25.74 ± 8.85	N/A
Kirschner et al. [13]	N/A	-3.62 ± 5.50	$1.94 \pm 1.10^*$	4.47 ± 3.30	34.60 ± 17.70	41 ± 3.8 s
Erdt et al. [14]	10.34 ± 3.11	1.55 ± 6.49	1.74 ± 0.59	3.51 ± 1.16	26.83 ± 8.87	45 s
Our method	8.74 ± 2.37	2.41 ± 1.71	1.45 ± 0.37	2.55 ± 0.59	26.91 ± 7.72	8.5 ± 1.3 min

Results for each measure are represented as mean and standard deviation of the overall datasets. Bold values are the best result in that column. N/A stands for Not Available information.

* This indicates a statistically significant difference between the marked result and the corresponding one of our method at a significance level of 0.05.

** This indicates a statistically significant difference between the marked result and the corresponding one of our method at a significance level of 0.01.

Table 9

Quantitative comparative results between our proposed method and the published semi-automatic methods based on the SLIVER07-Test database.

Method	VOE (%)	SRVD (%)	ASD (mm)	RMSD (mm)	MSD (mm)	Total score
Dawant et al. [53] (Med)	7.2 ± 1.2	2.5 ± 2.3	1.1 ± 0.2	1.9 ± 0.5	17.1 ± 5.4	76 ± 5
Lee et al. [54] (Low)	6.9 ± 1.4	1.3 ± 2.9	1.1 ± 0.3	2.1 ± 0.5	21.3 ± 4.0	75 ± 5
Our method (Med)	7.6 ± 1.0	0.2 ± 2.9	1.3 ± 0.2	2.1 ± 0.4	20.9 ± 4.5	74 ± 4
Beichel et al. [55] (Med)	6.5 ± 1.1	1.1 ± 1.9	1.1 ± 0.4	2.5 ± 1.2	23.4 ± 10.5	74 ± 9
Wimmer et al. [56] (Med)	8.1 ± 1.1	6.1 ± 2.6	1.3 ± 0.2	2.2 ± 0.4	18.7 ± 4.6	69 ± 5
Slagmolen et al. [57] (Med)	10.4 ± 3.1	3.7 ± 6.2	2.0 ± 0.7	5.0 ± 2.4	40.5 ± 18.2	52 ± 19
Beichel et al. [55] (Low)	14.3 ± 9.4	3.1 ± 10.7	3.6 ± 3.1	7.9 ± 5.9	49.2 ± 20.4	41 ± 27

Results for each measure are represented as mean and standard deviation of the overall datasets. Bold values are the best result in that column. Med and Low indicate methods with medium and low user interaction, respectively.

indicating that it is more robust on diverse datasets. Specifically, the mean ASD and VOE of our method are 1.45 mm and 8.74%, respectively. However, the SRVD of our method (2.41%) is slightly higher than Erdt's method (1.55%), which means that our method tends to over-segment the liver tissue. And the MSD of our method (26.91 mm) is slightly higher compared to Chung's method (25.74 mm). These results indicate that our proposed method is more accurate and robust than the state-of-the-art model-based methods even though we used a much smaller training set.

Table 8 also shows the computation time of runtime testing stage (i.e., shape initialization and shape deformation). Our method takes more time than Kirschner's and Erdt's method. Specifically, the average computation time of our method was 8.5 min, most of which was taken by the shape initialization module (5.4 min per dataset, including 3.5 min for the interactive hepatic vessel segmentation). In the offline training stage, the total time for establishing shape correspondence was 400 h. On average, the time for the multilevel shape division was 5 min. The total time required for training the appearance model was 30 min.

5.4.2. Evaluation on the SLIVER07-Test database

We compared our hierarchical liver segmentation framework with the published semi-automatic methods from the "MICCAI 2007 Grand Challenge" workshop [2]. Considering that the

interactive hepatic vessel segmentation in our method takes an average of 3.5 min, we only compared our method to semi-automatic methods with medium or low user interaction (i.e., the interaction time is less than 5 min). Because semi-automatic methods with high user interaction employed extensive manual refinement of the binary segmentation mask [2]. Specifically, we also evaluated our method on the SLIVER07-Test database which is designed specifically for the purpose of liver segmentation competition. Since the ground truth for the database is not available, our results were submitted to the competition organizers for evaluating the segmentation accuracy.

Table 9 lists the statistics for the 'rank' among the 7 semi-automatic methods. Our algorithm ranks third in the 7 methods. Our method achieves a total score of 74, which is only two points lower than the best score of 76 reached by Dawant's method [53]. Specifically, the mean ASD and VOE of our method are 1.3 mm and 7.6%, respectively (Team HIT-liver in SLIVER07 competition). Also, the results of our proposed framework are comparable to the manual segmentation of a second independent human rater (Team Lara in SLIVER07 competition) whose total score is 75. Therefore, our method's performance is comparable with that of human rater, and it can be deployed for effective liver tissue segmentation.

6. Discussion and conclusion

In this paper, we have presented a novel framework for liver segmentation in portal phase of abdominal CT images that extends ASMs by introducing: (1) a new MLR-SSC shape prior model; (2) a blood vessel-based liver shape initialization method; and (3) a hierarchical deformable shape optimization strategy. The proposed segmentation framework was extensively evaluated on a dataset of 60 clinical CT scans. The experimental results demonstrate that (1) our proposed MLR-SSC model achieves much better accuracy and efficiency than both the conventional shape model and the original SSC model in the presence of small training sets; (2) our blood vessel-based liver shape initialization method provides a more patient-specific initial shape for the hierarchical segmentation framework and achieves much better accuracy than the mean shape based method; and (3) our method shows slightly superior performance to state-of-the-art model-based methods even though our models are learned from a much smaller training set. Therefore, our proposed framework can be deployed for effective liver tissue segmentation in the presence of small training sets.

Despite the promising results, the segmentation framework presented in this paper could be further extended in several ways: (1) While we have empirically found that using shape variation information works well for shape division, it is not necessarily the optimal criterion. It would be worthwhile to evaluate other criteria which may potentially further improve the results; (2) In runtime testing stage, most of the computation time is taken by the shape initialization module (almost 64%). The computational efficiency of our framework would be further improved if more efficient or even automatic hepatic vessel segmentation methods are employed.

Experimental results show that our method does not require an accurate shape initialization; it works well when initial shapes have a relatively low overlap with the ground truth. Nevertheless, our approach has some limitations. The proposed blood vessel-based shape initialization method is sensitive to the presence of very large liver tumors. In such situations, large parts of the hepatic vessels will be occluded by the tumors. This results in an unsuccessful extraction of vessels in these regions. Consequently, the derived initial shapes cannot capture the full shape details, and the deformable models tend to under-segment these regions. In our experiments, the derived initial shapes correlate well with the ground truth despite the presence of large tumors as shown in Fig. 9c. Even in this case, however, the tumor volume is around 27% of the whole liver volume. Our segmentation method is successful. In addition, an attempt was made that when the tumor volume is 50% or more, our initialization method gave a very poor initial shape, which caused a large segmentation error in the final results. Our future work will focus on extending the method to the above situations.

Conflict of interest

None declared.

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 61401246; the National Basic Research Program of

China (973) under Grant No. 2011CB707701; the Nature Science Foundation of Heilongjiang Province of China (No. F2015005), and the Scientific Research Personnel Project of Heilongjiang Provincial Education Department (No. 12541164), and the Fundamental Research Funds for the Central Universities under Grant No. HIT. IBRSEM. 201328.

Appendix A. PCA-based statistical shape model building

Assume that we have a set of n_s corresponding liver training shapes $\{S_i | i = 1, 2, \dots, n_s\}$, and each shape S_i is represented by a shape vector $\mathbf{x}_i = (x_1, y_1, z_1, \dots, x_{n_p}, y_{n_p}, z_{n_p})^T$ with n_p landmark points. Firstly, all training shapes are spatially aligned into a common coordinate frame using the generalized Procrustes analysis (GPA) [40]. We define the aligned landmark configuration matrix as $\mathbf{L} = ((\mathbf{x}_1 - \bar{\mathbf{x}}), \dots, (\mathbf{x}_{n_s} - \bar{\mathbf{x}}))$, where $\bar{\mathbf{x}}$ is the mean shape vector: $\bar{\mathbf{x}} = \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{x}_i$. Because of its high numerical stability, singular value decomposition (SVD) of the landmark configuration matrix \mathbf{L} is employed to build the shape model, instead of using the usual eigen-decomposition of the covariance matrix of \mathbf{L}

$$\frac{1}{\sqrt{n_s - 1}} \mathbf{L} = \mathbf{UDV}^T, \quad (A.1)$$

where columns of matrix of \mathbf{U} form the principal modes of variation ϕ_m (eigenvectors), and diagonal entries of \mathbf{D}^2 are their respective variances λ_m (eigenvalues). All valid shapes of liver structure can be approximated by a linear combination of the first c modes of variation

$$\mathbf{x} \approx \bar{\mathbf{x}} + \Phi \mathbf{b} = \bar{\mathbf{x}} + \sum_{m=1}^c b_m \phi_m, \quad (A.2)$$

where $c = \min \{t | \sum_{i=1}^t \lambda_i / \sum_{i=1}^{n_s-1} \lambda_i > 0.98\}$, \mathbf{b} contains the shape parameters given by $\mathbf{b} = \Phi^T(\mathbf{x} - \bar{\mathbf{x}})$, and each b_m is limited to the interval $b_m \in [-3\sqrt{\lambda_m}, 3\sqrt{\lambda_m}]$.

References

- [1] P. Campadelli, E. Casiraghi, A. Esposito, Liver segmentation from computed tomography scans: A survey and a new algorithm, *Artif. Intell. Med.* 45 (2) (2009) 185–196.
- [2] T. Heimann, B. van Ginneken, M. Styner, Y. Arzhueva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes, F. Bello, G.K. Binnig, H. Bischof, A. Bornik, P. Cashman, Y. Chi, A. Cordova, B.M. Dawant, M. Fidrich, J.D. Furst, D. Furukawa, L. Grenacher, J. Hornegger, D. Kainmüller, R. Kitney, H. Kobatake, H. Lamecker, T. Lange, J. Lee, B. Lennon, R. Li, S. Li, H.-P. Meinzer, G. Németh, D. S. Raicu, A. Rau, E.M. van Rikxoort, M. Rousson, L. Ruskó, K.A. Saddi, G. Schmidt, D. Seghers, A. Shimizu, P. Slagmolen, E. Sorantin, G. Soza, R. Susomboon, J.M. Waite, A. Wimmer, I. Wolf, Comparison and evaluation of methods for liver segmentation from CT datasets, *IEEE Trans. Med. Imaging* 28 (8) (2009) 1251–1265.
- [3] M. Kobashi, L.G. Shapiro, Knowledge-based organ identification from CT images, *Pattern Recognit.* 28 (4) (1995) 475–491.
- [4] L. Ruskó, G. Bekes, M. Fidrich, Automatic segmentation of the liver from multi- and single-phase contrast-enhanced CT images, *Med. Image Anal.* 13 (6) (2009) 871–882.
- [5] L. Soler, H. Delingette, G. Malandain, J. Montagnat, N. Ayache, C. Koehl, O. Dourthe, B. Malassagne, M. Smith, D. Mutter, J. Marescaux, Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery, *Comput. Aided Surg.* 6 (3) (2001) 131–142.
- [6] A. Afifi, T. Nakaguchi, Liver segmentation approach using graph cuts and iteratively estimated shape and intensity constrains, in: Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI'12), Nice, France, 2012, pp. 395–403.
- [7] H. Lamecker, T. Lange, M. Seebass, Segmentation of the liver using a 3D statistical shape model, ZIB-Report 04-09, Konrad-Zuse-Zentrum für Informationstechnik Berlin, April 2004.
- [8] T. Heimann, S. Müning, H.-P. Meinzer, I. Wolf, A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation, in: Proceedings of Information Processing in Medical Imaging (IPMI), Kerkrade, The Netherlands, 2007, pp. 1–12.

- [9] T. Okada, R. Shimada, M. Hori, M. Nakamoto, Y.-W. Chen, H. Nakamura, Y. Sato, Automated segmentation of the liver from 3D CT images using probabilistic atlas and multilevel statistical shape model, *Acad. Radiol.* 15 (11) (2008) 1390–1403.
- [10] J. Feng, H.H. Ip, A multi-resolution statistical deformable model (MISTO) for soft-tissue organ reconstruction, *Pattern Recognit.* 42 (7) (2009) 1543–1558.
- [11] X. Zhang, J. Tian, K. Deng, Y. Wu, X. Li, Automatic liver segmentation using a statistical shape model with optimal surface detection, *IEEE Trans. Biomed. Eng.* 57 (10) (2010) 2622–2626.
- [12] F. Chung, H. Delingette, Regional appearance modeling based on the clustering of intensity profiles, *Comput. Vis. Image Underst.* 117 (6) (2013) 705–717.
- [13] M. Kirschner, The probabilistic active shape model: from model construction to flexible medical image segmentation (Ph.D. thesis), Technische Universität Darmstadt, Darmstadt, Germany, 2013.
- [14] M. Erdt, M. Kirschner, Fast automatic liver segmentation combining learned shape priors with observed shape deviation, in: Proceedings of IEEE 23rd International Symposium on Computer-Based Medical Systems (CBMS 2010), Perth, Australia, 2010, pp. 249–254.
- [15] H. Park, P.H. Bland, C.R. Meyer, Construction of an abdominal probabilistic atlas and its application in segmentation, *IEEE Trans. Med. Imaging* 22 (4) (2003) 483–492.
- [16] T.F. Cootes, C.J. Taylor, D.H. Cooper, J. Graham, Active shape models—their training and application, *Comput. Vis. Image Underst.* 61 (1) (1995) 38–59.
- [17] T. Heimann, H.-P. Meinzer, Statistical shape models for 3D medical image segmentation: a review, *Med. Image Anal.* 13 (4) (2009) 543–563.
- [18] I. Jolliffe, Principal component analysis, Springer Series in Statistics, 2nd edition, Springer, Berlin, Germany, 2002.
- [19] J. Koikkalainen, T. Tölli, K. Lauerma, K. Antila, E.M. Mattila, M. Lilja, J. Lötjönen, Methods of artificial enlargement of the training set for statistical shape models, *IEEE Trans. Med. Imaging* 27 (11) (2008) 1643–1654.
- [20] C. Davatzikos, X. Tao, D. Shen, Hierarchical active shape models, using the wavelet transform, *IEEE Trans. Med. Imaging* 22 (3) (2003) 414–423.
- [21] Y. Wang, L.H. Staib, Boundary finding with prior shape and smoothness models, *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (7) (2000) 738–743.
- [22] T.F. Cootes, C.J. Taylor, A mixture model for representing shape variation, *Image Vis. Comput.* 17 (8) (1999) 567–573.
- [23] P.D. Sozou, T.F. Cootes, C.J. Taylor, E.C. Di Mauro, Non-linear generalization of point distribution models using polynomial regression, *Image Vis. Comput.* 13 (5) (1995) 451–457.
- [24] C.J. Twining, C.J. Taylor, The use of kernel principal component analysis to model data distributions, *Pattern Recognit.* 36 (1) (2003) 217–227.
- [25] K. Lekadir, R.D. Merrifield, G. Yang, Outlier detection and handling for robust 3-D active shape models search, *IEEE Trans. Med. Imag.* 26 (2) (2007) 212–222.
- [26] M. Rogers, J. Graham, Robust active shape model search, in: Proceedings of European Conference on Computer Vision (ECCV'02), Copenhagen, Denmark, 2002, pp. 517–530.
- [27] J. Abi-Nahed, M. Jolly, G. Yang, Robust active shape models: a robust, generic and simple automatic segmentation tool, in: Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI'06), Copenhagen, Denmark, 2006, pp. 1–8.
- [28] J. Wright, Y. Ma, J. Mairal, G. Sapiro, T.S. Huang, S. Yan, Sparse representation for computer vision and pattern recognition, *Proc. IEEE* 98 (6) (2010) 1031–1044.
- [29] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D.N. Metaxas, X.S. Zhou, Towards robust and effective shape modeling: sparse shape composition, *Med. Image Anal.* 16 (1) (2012) 265–277.
- [30] S.K. Zhou, Discriminative anatomy detection: classification vs regression, *Pattern Recognit. Lett.* 43 (2014) 25–38.
- [31] H. Ling, S.K. Zhou, Y. Zheng, B. Georgescu, M. Suehling, D. Comaniciu, Hierarchical, learning-based automatic liver segmentation, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'08), Anchorage, Alaska, USA, 2008, pp. 1–8.
- [32] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, D. Comaniciu, Four-chamber heart modeling and automatic segmentation for 3-D cardiac CT volumes using marginal space learning and steerable features, *IEEE Trans. Med. Imaging* 27 (11) (2008) 1668–1681.
- [33] A.S. Maklad, M. Matsuhiro, H. Suzuki, Y. Kawata, N. Niki, M. Satake, N. Moriyama, T. Utsunomiya, M. Shimada, Blood vessel-based liver segmentation using the portal phase of an abdominal CT dataset, *Med. Phys.* 40 (11) (2013) 113501–1–113501–17.
- [34] B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, Least angle regression, *Ann. Stat.* 32 (2) (2004) 407–499.
- [35] S. Zhang, Y. Zhan, D.N. Metaxas, Deformable segmentation via sparse representation and dictionary learning, *Med. Image Anal.* 16 (7) (2012) 1385–1396.
- [36] Y. Shao, Y. Gao, Y. Guo, Y. Shi, X. Yang, D. Shen, Hierarchical lung field segmentation with joint shape and appearance sparse learning, *IEEE Trans. Med. Imaging* 33 (9) (2014) 1761–1780.
- [37] S. Lloyd, Least squares quantization in PCM, *IEEE Trans. Inf. Theory* 28 (2) (1982) 129–137.
- [38] D. Cohen-Steiner, P. Alliez, M. Desbrun, Variational shape approximation, *ACM Trans. Graph. (TOG)* 23 (3) (2004) 905–914.
- [39] T.F. Cootes, C.J. Taylor, Statistical Models of Appearance for Computer Vision, Tech. Rep., University of Manchester, October 2001.
- [40] C. Goodall, Procrustes methods in the statistical analysis of shape, *J. R. Stat. Soc. Ser. B (Methodol.)* 53 (2) (1991) 285–339.
- [41] N. Nguyen, T. Tran, Exact recoverability from dense corrupted observations via ℓ_1 -minimization, *IEEE Trans. Inf. Theory* 59 (4) (2013) 2017–2035.
- [42] S.S. Chen, D.L. Donoho, M.A. Saunders, Atomic decomposition by basis pursuit, *SIAM J. Sci. Comput.* 20 (1) (1998) 33–61.
- [43] J. Mairal, F. Bach, J. Ponce, Sparse modeling for image and vision processing, *Found. Trends Comput. Graph. Vis.* 8 (2–3) (2014) 85–283.
- [44] A.F. Frangi, W.J. Niessen, K.L. Vincken, M.A. Viergever, Multiscale vessel enhancement filtering, in: Proceedings of Medical Image Computing and Computer-Assisted Intervention (MICCAI'98), Cambridge, MA, USA, 1998, pp. 130–137.
- [45] W.E. Lorensen, H.E. Cline, Marching cubes: a high resolution 3D surface construction algorithm, in: Proceedings of SIGGRAPH '87, New York, NY, USA, 1987, pp. 163–169.
- [46] C.B. Barber, D.P. Dobkin, H. Huhdanpaa, The quickhull algorithm for convex hulls, *ACM Trans. Math. Soft. (TOMS)* 22 (4) (1996) 469–483.
- [47] I.H. Witten, E. Frank, Data Mining: Practical Machine Learning Tools and Techniques, second edition, The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann, San Francisco, USA, 2005.
- [48] M. Styner, I. Oguz, S. Xu, C. Brechbühler, D. Pantazis, J.J. Levitt, M.E. Shenton, G. Gerig, 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, 2006, <http://dx.doi.org/1926/215>.
- [49] T. Heimann, I. Oguz, I. Wolf, M. Styner, H.-P. Meinzer, 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, 2006, <http://dx.doi.org/1926/224>.
- [50] G. Lavoué, M. Tola, F. Dupont, MEPP - 3D mesh processing platform, in: Proceedings of International Conference on Computer Graphics Theory and Applications, Rome, Italy, 2012, pp. 206–210.
- [51] R.H. Davies, Learning shape: optimal models for analyzing natural variability (Ph.D. thesis), University of Manchester, Manchester, UK, 2002.
- [52] B. Jian, B.C. Vemuri, Robust point set registration using Gaussian mixture models, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (8) (2011) 1633–1645.
- [53] B.M. Dawant, R. Li, B. Lennon, S. Li, Semi-automatic segmentation of the liver and its evaluation on the MICCAI 2007 grand challenge data set, in: Proceedings of MICCAI Workshop 3-D Segmentation in the Clinic: A Grand Challenge, 2007, pp. 215–221.
- [54] J. Lee, N. Kim, H. Lee, J.B. Seo, H.J. Won, Y.M. Shin, Y.G. Shin, Efficient liver segmentation exploiting level-set speed images with 2.5D shape propagation, in: Proceedings of MICCAI Workshop 3-D Segmentation in the Clinic: A Grand Challenge, 2007, pp. 189–196.
- [55] R. Beichel, C. Bauer, A. Bornik, E. Sorantin, H. Bischof, Liver segmentation in CT data: a segmentation refinement approach, in: Proceedings of MICCAI Workshop 3-D Segmentation in the Clinic: A Grand Challenge, 2007, pp. 235–245.
- [56] A. Wimmer, G. Soza, J. Hornegger, Two-stage semi-automatic organ segmentation framework using radial basis functions and level sets, in: Proceedings of MICCAI Workshop 3-D Segmentation in the Clinic: A Grand Challenge, 2007, pp. 179–188.
- [57] P. Slagmolen, A. Elen, D. Seghers, D. Loeckx, F. Maes, K. Haustermans, Atlas based liver segmentation using nonrigid registration with a B-spline transformation model, in: Proceedings of MICCAI Workshop 3-D Segmentation in the Clinic: A Grand Challenge, 2007, pp. 197–206.

Changfa Shi received the M.S. degree from Harbin Institute of Technology, Harbin, China, in 2010. He is now working towards the Ph.D. degree in School of Computer Science and Technology, Harbin Institute of Technology. His current research interests include medical image processing, analysis and sparse representation.

Yuanzhi Cheng is currently a Professor with the School of Computer Science and Technology, Harbin Institute of Technology. He has authored over 30 papers in scientific journals. His research interests are concentrated on pattern recognition, image processing, and computer-assisted surgical system.

Fei Liu received the B.S. degree from Zhejiang University, Zhejiang, China, in 2008. She is currently a Ph.D. candidate in the Biomedical Engineering Department, Tsinghua University. Her research interest is medical image analysis.

Yadong Wang is currently a Professor in School of Computer Science and Technology, Harbin Institute of Technology. His research interests include pattern recognition, machine learning and Bioinformatics. He has published more than 60 papers in scientific journals.

Jing Bai received the Ph.D. from Drexel University, Philadelphia, in 1985. From 1985 to 1987 she was a Research Associate and Assistant Professor with the Biomedical Engineering and Science Institute of Drexel University. She is currently a Professor at Biomedical Engineering Department of Tsinghua University, Beijing, China. Her research activities have included mathematical modeling and simulation of cardiovascular system, optimization of cardiac assist devices, medical ultrasound, telemedicine, home health care network and home monitoring devices, and Infrared imaging. She has published four books and over 200 journal papers. She is a Fellow of IEEE and a Fellow of AIMBE. From 1997, she became an Associate Editor for IEEE Transactions on Information Technology in Biomedicine.

Shinichi Tamura received the Ph.D. degree in electrical engineering from Osaka University, Osaka, Japan, in 1971. He is currently a Professor in the Center for Advanced Medical Engineering and Informatics, Osaka University. He has published more than 260 papers in scientific journals and received several awards from journals including Pattern Recognition and Investigative Radiology. His current research activities include works in the field of medical image analysis and its applications. Dr. Tamura is a Fellow of IEEE.