HHS Public Access

Author manuscript

IEEE Trans Biomed Eng. Author manuscript; available in PMC 2020 May 12.

Discontinuity Preserving Liver MR Registration with 3D Active Contour Motion Segmentation

Dongxiao Li [Member, IEEE],
College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China.

Wenxiong Zhong,
College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China.

Kofi M. Deh,
Department of Radiology, Weill Cornell Medical College, New York, NY 10021, USA.

Thanh D. Nguyen,
Department of Radiology, Weill Cornell Medical College, New York, NY 10021, USA.

Martin R. Prince,
Department of Radiology, Weill Cornell Medical College, New York, NY 10021, USA.

Yi Wang [Fellow, IEEE], and
Department of Radiology, Weill Cornell Medical College, New York, NY 10021, USA., Department of Biomedical Engineering, Cornell University, Ithaca, NY 14853, USA.

Pascal Spincemaille* **[Member, IEEE]**
Department of Radiology, Weill Cornell Medical College, New York, NY 10021, USA

Abstract

Objective: The sliding motion of the liver during respiration violates the homogeneous motion smoothness assumption in conventional non-rigid image registration and commonly results in compromised registration accuracy. This paper presents a novel approach, registration with 3D active contour motion segmentation (RAMS), to improve registration accuracy with discontinuity-aware motion regularization.

Methods: A Markov random field-based discrete optimization with dense displacement sampling and self-similarity context metric is used for registration, while a graph cuts-based 3D active contour approach is applied to segment the sliding interface. In the first registration pass, a mask-free L1 regularization on an image-derived minimum spanning tree is performed to allow motion discontinuity. Based on the motion field estimates, a coarse segmentation finds the motion boundaries. Next, based on MR signal intensity, a fine segmentation aligns the motion boundaries with anatomical boundaries. In the second registration pass, smoothness constraints across the

*correspondence pas2018@med.cornell.edu.

segmented sliding interface are removed by masked regularization on a minimum spanning forest and masked interpolation of the motion field.

Results: For in vivo breath-hold abdominal MRI data, the motion masks calculated by RAMS are highly consistent with manual segmentations in terms of Dice similarity and bidirectional local distance measure. These automatically obtained masks are shown to substantially improve registration accuracy for both the proposed discrete registration as well as conventional continuous non-rigid algorithms.

Conclusion/Significance: The presented results demonstrated the feasibility of automated segmentation of the respiratory sliding motion interface in liver MR images and the effectiveness of using the derived motion masks to preserve motion discontinuity.

Keywords

Active contour; Liver; magnetic resonance imaging (MRI); motion segmentation; nonrigid registration

I. Introduction

RSPIRATORY motion registration of liver magnetic resonance imaging (MRI) data is needed for many clinical applications, such as hepatic perfusion analysis using dynamic contrast enhanced MRI (DCE-MRI) [1], longitudinal hepatic disease and treatment monitoring [2], image guided liver therapy and surgery [3], and atlas-based liver segmentation [4].

Non-rigid motion registration of medical images has been of great interest for the past three decades [5–7] but it still remains challenging because it is an ill-posed, nonconvex optimization problem with several million degrees-of-freedom. This is typically overcome by regularizing the motion field. In general, various spatial smoothness constraints are enforced to avoid physically implausible motion fields and undesirable local minima [7]. For abdominal image registration, respiration results in a superior-inferior sliding motion of the organs – such as the liver – against the inward-outward moving abdominal wall and the stationary spine. The resulting discontinuities in the motion field prevent conventional algorithms that use homogeneous smoothness regularization from finding an accurate motion field [8].

In recent years, regularization for discontinuous motion fields has been proposed for lung computed tomography (CT) registration using publicly available reference datasets such as EMPIRE10 Challenge [8] or DIR-Lab [9–11] for validation. For liver MR images, discontinuity preserving registration has gained attention recently, but no reference dataset for validating registration accuracy is publicly available [12–14].

The algorithms that have been proposed for discontinuity preserving respiratory motion registration can be generally classified into two main categories: mask-free and mask-based. In the first category, conventional homogeneous smoothness constraints in continuous optimization-based registration were replaced by various locally adaptive regularization approaches, such as direction-dependent regularization [15], piecewise-diffeomorphic

registration [16], anisotropic diffusion regularization [17], bilateral filtering [18], SLIC Demons [19], total variation (TV) regularization [20], and eXtended Free-Form Deformation (XFFD) [21].

Recently, Markov random field (MRF)-based discrete optimization [22] has attracted increasing interest. Traditional continuous optimization is more susceptible to local minima, particularly in respiratory motion registration where the motion magnitude can be large compared to the typical scale of image features. In contrast, global minima can be achieved using discrete optimization with dense displacement sampling (DEEDS) [23]. In DEEDS, a relaxed graph structure – an image-derived minimum spanning tree (MST) – is used to represent the underlying anatomical connectivity and only the MST edges are included in the regularization. The adoption of an MST graph has two important advantages: first, it greatly reduces the time for solving the MRF problem; second, it allows sliding motion at lung surfaces by removing most motion smoothness constraints between nodes with large intensity differences, such as those between the chest wall and the lung. As evaluated in [24], among six registration tools, DEEDS yielded the best registration performance for atlas-based segmentation of 13 abdominal organs in CT scans.

Nevertheless, registration accuracy remains compromised without using explicit motion masks in the regularization. With prior knowledge of where the motion discontinuity occurs, smoothness constraints across the boundary can be avoided. As reported on the DIR-Lab website (http://www.dir-lab.com), most of the top-ranking algorithms use prior lung segmentations. To cope with large motion, prior lung segmentation has been combined with DEEDS to restrict MST-based MRF regularization on key points only inside the lung volume [25].

Because manual segmentation is time consuming and requires anatomical expertise, considerable effort has been spent on developing automatic motion mask extraction, either by image-based segmentation before registration or by joint registration and segmentation. In lung CT images, the high contrast between the air-filled lungs and the chest wall leads to the successful application of image-based segmentation [26], [27]. However, in abdominal CT and MR images, due to the low contrast between adjacent organs and the abdominal wall, automated image-based segmentation remains challenging. Recently, model-based methods incorporating liver shape prior knowledge have been investigated to obtain accurate and robust segmentation, but mostly focused on abdominal CT images. Wang *et al.* proposed a shape-intensity prior level set combining probabilistic atlas and probability map constraints in [28], Shi *et al.* proposed a hierarchical local region-based sparse shape composition model in [29] and a low-rank plus sparse decomposition-based shape model in [30], all for automated liver segmentation in CT scans. Manual intervention was usually needed to achieve the desired precision for liver segmentation on CT and MR [31].

Joint registration and segmentation techniques aim to solve these two problems together. For lung CT registration, Preston *et al.* presented a piece-wise diffeomorphic deformation model with automatic sliding boundary computation [32]. Jud *et al.* introduced a regularity criterion which incorporated local directional statistics-based motion segmentation [33]. Swierczynski *et al.* developed a level-set approach, which merged a classic Chan-Vese

segmentation with the active dense displacement field estimation [34]. However, image-based prior segmentation is satisfactory for lung CT, and joint segmentation and registration did not achieve improved registration accuracy compared to the state-of-the-art, as reported in [34].

For liver MR registration, Kiriyanthan *et al.* proposed a combination model [12] of optical flow based registration [35] and convex segmentation [36]. A Primal-Dual approach [37] was introduced to solve both motion segmentation and registration using pure TV-L1 regularization [13]. Nevertheless, for optimal performance, incorporation of manually annotated corresponding landmarks was necessary [14]. This avoided unreliable motion fields, especially for large motion.

For MRF-based registration framework, Mahapatra *et al.* proposed to integrate segmentation by combining registration and segmentation labels into both the regularization and data terms [38]. Gass *et al.* formulated simultaneous segmentation and multiresolution nonrigid atlas registration as an MRF-based maximum a-posteriori problem [39]. However, due to computational complexity, the total label space was limited, and is thus not applicable to large nonrigid motion registration.

In this work, we present a novel approach for liver MR images: registration with 3D active contour motion segmentation (RAMS). In contrast to previous studies [12–14], motion registration and segmentation are optimized alternatingly in a pure discrete optimization framework with just two registration passes and one embedded segmentation process (Fig. 1). The DEEDS [23] algorithm using the self-similarity context (SSC) similarity metric [40] and a graph cuts based 3D active contour (GCBAC) [41] approach are adapted for liver MR registration and sliding motion interface segmentation, respectively. First, mask-free L1 regularization on an image-derived MST is performed to obtained motion fields that allow for motion discontinuity. Next, segmentation is performed in two stages: 1) based on the motion fields, a coarse segmentation finds motion boundaries, and 2) based on the MR signal intensity, this segmentation is refined by aligning motion boundaries with anatomical boundaries. In a second registration pass, smoothness constraints across boundaries are removed by masked regularization on a minimum spanning forest (MSF) and masked motion field interpolation. The proposed RAMS approach was validated on in vivo liver MRI data and additionally evaluated using a public liver CT dataset.

II. Methods

A. MRF-based Registration Using DEEDS-SSC-L1

As shown in Fig. 1, $\boldsymbol{I}_t$ denotes the fixed target image and I_s denotes the moving source image for registration. Following DEEDS, first order 3D (trilinear) B-spline meshes are used to model the motion fields, and MRF-based discrete optimization is used to search the motion vector (MV) of each control node (abbreviated as node below). $\boldsymbol{I}_t$ is divided into non-overlapping identical-sized 3D patches centered on the nodes. Voxel-wise MVs are generated from node MVs using B-spline interpolation. A graph G is defined using the nodes as vertices with edges connecting each node to its 26 immediate 3D neighbors. This

graph model is chosen to improve the performance of the graph cuts-based segmentation in RAMS. A set of all possible discretized 3D MV candidates are assigned to each vertex.

The optimization of node MVs is formulated as:

$$\boldsymbol{u}^* = \underset{\boldsymbol{u}_p,\ p \in \mathscr{P}}{\operatorname{argmin}}\ E(u) = \underset{u_p,\ p \in \mathscr{P}}{\operatorname{argmin}} \left(\underbrace{\sum_{p \in \mathscr{P}} D\left(\boldsymbol{I}_{\mathrm{t}}, \boldsymbol{I}_{\mathrm{s}}, \boldsymbol{r}_p, \boldsymbol{u}_p\right)}_{\text{data term}} + \alpha \underbrace{\sum\nolimits_{(p,q) \in \mathscr{N}} R\left(\boldsymbol{u}_p, \boldsymbol{u}_q, \boldsymbol{r}_p, \boldsymbol{r}_q\right)}_{\text{regularization term}} \right) \quad (1)$$

where $\mathscr{P}$ is the set of nodes, $\boldsymbol{r}_{\mathrm{p}}$ the position of node p, $\boldsymbol{u}_{\mathrm{p}}$ the node's MV, $\mathscr{N}$ the set of neighboring node pairs, and α is a weighting parameter. The data and regularization terms are constructed as follows.

1) Data Similarity Using SSC—To handle possible intensity changes between target and source, the SSC data similarity metric [40] is used, which has robust performance in object shape matching with respect to changes in color, texture, edges, and other photometric properties [42], [43]. SSC builds on the modality independent neighborhood descriptor (MIND) multi-modal similarity metric [44] by refining the neighborhood layout to improve matching robustness, and is efficiently quantized to a 60-bits coded 12-element vector descriptor for each voxel [40]. To enable an MV search at subvoxel precision, quantized SSC descriptor images $\boldsymbol{Q}_{\mathrm{t}}$ and $\boldsymbol{Q}_{\mathrm{s}}$ are precomputed at a higher spatial resolution after upsampling of $\boldsymbol{I}_{\mathbf{t}}$ and I_{s}, respectively.

For each node p, the data term in (1) is defined as

$$D\left(\boldsymbol{I}_t, \boldsymbol{I}_s, \boldsymbol{r}_p, \boldsymbol{u}_p\right) = \frac{1}{\left|\mathscr{R}\left(\boldsymbol{r}_p\right)\right|} \sum_{\boldsymbol{r} \in \mathscr{R}\left(\boldsymbol{r}_p\right)} bitcount\left(\boldsymbol{Q}_{\mathrm{t}}(\boldsymbol{r}) \oplus \boldsymbol{Q}_{\mathrm{s}}\left(\boldsymbol{r} + \boldsymbol{u}_p\right)\right) \quad (2)$$

where $\oplus$ denotes bit-wise exclusive OR operation, *bitcount* counts the number of nonzero bits, and $\mathscr{R}\left(r_p\right)$ is a random subset of voxel coordinates within the patch at node p.

2) Fast L1 Regularization on an Image-derived MST—To allow motion discontinuity, L1 regularization [20], [45], is used:

$$R\left(\boldsymbol{u}_p, \boldsymbol{u}_q, \boldsymbol{r}_p, \boldsymbol{r}_q\right) = \left\|\boldsymbol{u}_p - \boldsymbol{u}_q\right\|_1 / \left\|\boldsymbol{r}_p - \boldsymbol{r}_q\right\|_2. \quad (3)$$

Edge weights $W(p, q)$ are defined as the gradient between patches:

$$W(p,q)=\sqrt{\sum_{r\in\mathcal{S}(r_p)}\left(I_t(r)-I_t\left(r-r_p+r_q\right)\right)^2}/\left\|r_p-r_q\right\|_2 \quad (4)$$

where $\mathcal{S}(\boldsymbol{r}_p)$ is the set of voxel coordinates within the patch at node p. From the graph G, a subgraph is extracted connecting all vertices of G but with minimum total edge weight $\sum_{(p,q)}W(p,q)$. This is the minimum spanning tree (MST) of G and is extracted using Prim's algorithm [46]. By simplifying the graph to an MST, the global optimum of (1) can be found efficiently using message passing [23]. For each node p, given the MV $\boldsymbol{u}_q$ of its parent node q, the energy E' of the optimal MV $\boldsymbol{u}_p$ is found using min-sum message passing [45]:

$$E'(p,q)=\min_{\boldsymbol{u}_p}\left(D\left(\boldsymbol{I}_t,\boldsymbol{I}_s,\boldsymbol{r}_p,\boldsymbol{u}_p\right)+\alpha R\left(\boldsymbol{u}_p,\boldsymbol{u}_q,\boldsymbol{r}_p,\boldsymbol{r}_q\right)+\sum_c E'(c,p)\right) \quad (5)$$

where c represents the child nodes of p. In a first pass, $E'(p,q)$ is calculated for all possible $\boldsymbol{u}_q$, from leaf nodes to its tree root. Once the tree root is reached, the optimal $\boldsymbol{u}_p$ is selected for each node p in a second pass, traversing from the tree root to the leaf nodes.

For each tree edge, the message passing of (5) naively requires $O(|\mathscr{L}|^2)$ calculations, where $\mathscr{L}$ is the label search space. A lower envelope computation-based min-convolution method [45] with $O(|\mathscr{L}|)$ cost is employed for fast L1 regularization in RAMS.

3) Inverse Consistent Symmetric Mapping—Enforcing inverse consistency as an additional regularization has been shown to improve registration accuracy [23], [25], [40], [47]. By swapping the role of target and source, a backward motion field $\boldsymbol{v}(\boldsymbol{r})$, which spatially aligns $\boldsymbol{I}_t(r+\boldsymbol{v}(\boldsymbol{r}))$ with $\boldsymbol{I}_s(\boldsymbol{r})$, is solved using the same MRF-based approach described above. Then the consistency between $\boldsymbol{u}(\boldsymbol{r})$ and $\boldsymbol{v}(\boldsymbol{r})$ is enforced by the following fixed point scheme [40] for several iterations:

$$\boldsymbol{u}^{n+1}(\boldsymbol{r})=0.5\left(\boldsymbol{u}^n(\boldsymbol{r})-\boldsymbol{v}^n\left(\boldsymbol{r}+\boldsymbol{u}^n(\boldsymbol{r})\right)\right), \quad (6a)$$

$$\boldsymbol{v}^{n+1}(\boldsymbol{r})=0.5\left(\boldsymbol{v}^n(\boldsymbol{r})-\boldsymbol{u}^n\left(\boldsymbol{r}+\boldsymbol{v}^n(\boldsymbol{r})\right)\right). \quad (6b)$$

4) Multi-resolution Optimization—To reduce computational complexity, a multi-resolution scheme [23], [45] is used in RAMS. From the second level on, each node MV is decomposed into two additive components: a prior value and an incremental value. The prior component is obtained by B-spline interpolation of the node MVs of the previous level. Only the incremental component is searched. From lower to higher levels, node spacing, MV quantization, and label search range are successively decreased.

Isotropic level parameters were used in previous studies [23], [45]. However, in MRI, voxel sizes are commonly anisotropic, and respiratory motion magnitudes are usually anisotropic as well. To optimize computer resource utilization, especially for large motion registration, anisotropic node spacing, MV quantization, and label search ranges are used in RAMS.

B. Coarse Segmentation based on Node MV Estimates

For motion pattern analysis in computer vision, motion segmentation is usually formulated as a clustering problem of the motion flow fields [48–50]. In our previous work [51], an MST clustering with maximum standard deviation reduction (MSDR) [52] was used. To boost the robustness in motion clustering using inaccurate motion estimates, we make use of prior knowledge about the sliding motion interface by introducing active contour models [53]. Active contours have been widely used in continuous optimization based image segmentation, e.g., automatic left ventricle segmentation [54] and the user-guided segmentation software ITK-SNAP [55]. To avoid the inherent sensitivity to local minima in continuous optimization, and inspired by the GCBAC approach to object segmentation [41], we propose a discrete optimization based method (Fig. 2). Based on the node MV estimates at each resolution level in the first registration pass, a coarse segmentation partitions the nodes of G into an inner and an outer region. The optimal 3D contour corresponds to the segmentation boundary, ideally the sliding motion interface. It is computed by a series of s-t min-cut based graph partitioning [41].

1) Edge Capacity Generation—The capacity $C_{\mathrm{m}}(p,q)$ for each edge (p,q) of G is set to

$$C_{\mathrm{m}}(p,q) = \left\lfloor c_{\mathrm{m}} \exp\left(-\left(\gamma_{\mathrm{m}} G_{\mathrm{m}}(p,q)/\max_{(p,q)\in\mathcal{N}} G_{\mathrm{m}}(p,q)\right)^2\right)\right\rfloor \quad (7)$$

$$G_{\mathrm{m}}(p,q) = \left\|\boldsymbol{u}_p - \boldsymbol{u}_q\right\|_2 / \left\|\boldsymbol{r}_p - \boldsymbol{r}_q\right\|_2 \quad (8)$$

where $G_{\mathrm{m}}(p,q)$ is the node MV gradient along the edge (p,q), c_{m} and γ_{m} constant integers, and $\lfloor \cdot \rfloor$ denotes the floor operation such that the edge capacity is rounded to the nearest integer less than or equal to that value. Here the subscript m is used to identify the parameters for motion field-based segmentation. The maximum value of $G_{\mathrm{m}}(p,q)$ in the entire graph is used for normalization of the gradient magnitude. γ_{m} determines the sensitivity in boundary detection, and c_{m} scales the edge capacity before discretization.

2) Contour Initialization—At the first level, the initial contour CR_0 is initialized with the surfaces of a shrunk image cuboid, which is d_{m} away from the image bounds in the right, left, anterior, and posterior directions, shown as the black dotted lines in Fig. 2. From the second level on, the contour CR_k is initialized by upsampling the contour CR_{k-1} obtained at the previous level.

3) Contour Dilation—The contour CR_{k-1} is morphologically dilated to a contour neighborhood CN_k using a $(2d_{\mathrm{m}}+1)^3$ sized 3D patch of ones as a structuring element, where d_{m} is the isotropic dilation half-width.

4) Subgraph Extraction—A subgraph, which consists of the nodes only in CN_k and their associated edges, is extracted from the original graph G. As shown in Fig. 2, the nodes in the inner and outer contour of CN_k are set as the source and sink vertices, respectively, and the remainder nodes in CN_k are set as the intermediate vertices. This multi-source multi-sink subgraph is then converted to a single-source single-sink subgraph by collapsing all source and sink vertices to a single source and a sink vertex, respectively [41].

5) Subgraph Partitioning—A new optimal contour CR_k is computed by finding the s-t min-cut of the subgraph:

$$cut^* = \mathrm{argmin}_{cut} F(cut) = \mathrm{argmin}_{cut} \sum_{(p,q) \in cut} C_{\mathrm{m}}(p,q) \quad (9)$$

where *cut* is a set of edges such that, by removing them, the subgraph is split into a source partition and a sink partition, and *F(cut)* is the sum of edge capacities of *cut*. The preflow-push algorithm with excess scaling [56] is used to compute the s-t min-cut. The total capacity of the min-cut F_k is exactly the maximum flow between the source and the sink. CR_k is set to the set of outer nodes of the inner source partition.

6) Stopping Criteria—Iterations are stopped when $F_k = F_{k-1}$. From the first to the final level, the spatial resolution of coarse segmentation is refined successively.

C. Fine Segmentation based on Image Intensity

Based on the final level coarse segmentation output, a subsequent fine segmentation is performed at voxel resolution using the MR signal intensity. The algorithm for fine segmentation is similar to the coarse segmentation, except for the following three differences:

1. Graph vertices correspond to voxels.
2. To improve noise robustness, the MR image is low pass filtered before computing the edge capacity $C_{\mathrm{i}}(p, q)$:

$$C_{\mathrm{i}}(p,q) = \left\lfloor c_{\mathrm{i}} \exp\left(-\left(\gamma_{\mathrm{i}} G_{\mathrm{i}}(p,q) / \max_{(p,q) \in \mathcal{N}} G_{\mathrm{i}}(p,q)\right)^2\right)\right\rfloor \quad (10)$$

$$G_{\mathrm{i}}(p,q) = \left|\boldsymbol{I}'\left(\boldsymbol{r}_p\right) - \boldsymbol{I}'\left(\boldsymbol{r}_q\right)\right| / \left\|\boldsymbol{r}_p - \boldsymbol{r}_q\right\|_2 \quad (11)$$

where $G_{\mathrm{i}}(p,q)$ is the voxel intensity gradient along the edge (p,q), c_{i} and γ_{i} constant integers, and $\boldsymbol{I}'$ is the low pass filtered image of $\boldsymbol{I}$. Here the subscript i indicates image-based segmentation. A windowed isotropic 3D Gaussian filter, consisting of a series of 1D Gaussian filtering steps, is used for low pass filtering. In each dimension, the 1D Gaussian filter $g(x) = \exp\left(-x^2/\left(2\sigma^2\right)\right)/(\sqrt{2\pi}\sigma)$ is sampled and normalized with $\sigma = 1$ and $x = [-2, \quad -1, \quad 0, \quad 1, \quad 2]$, resulted in a 5-tap filter with coefficients [0.0545, 0.2442, 0.4026, 0.2442, 0.0545].

3. Because $F(cut)$ tends to decrease when fewer edges are included in cut [41], the optimal contour tends to shrink inwards and the obtained motion boundaries are generally smaller than the actual sliding interface (see Fig. 3). Therefore, the dilation of the 3D contour CR_{k-1} is constrained outwards at each *iteration k*. This is implemented by excluding the vertices in the inner source partition after an isotropic dilation using a $\left(2d_i + 1\right)^3$ sized 3D patch of ones as a structuring element, where d_i is the dilation half-width.

After the fine segmentation, a binary image mask $\boldsymbol{M}_{\mathrm{i}}(\boldsymbol{r})$ is obtained for each image, with $\boldsymbol{M}\mathrm{i}(r) = 1$ for voxels in the inner region. When the motion discontinuity is small, segmentation is not necessary for registration. Therefore, in RAMS, if $\max_{(p,q)\in\mathcal{N}} G_{\mathrm{m}}(p,q)$ of the final level is less than a preset threshold τ, then the motion field segmentation and the second registration (described below) pass are not carried out.

D. Masked Regularization on an MSF

To incorporate motion masks into the regularization, the MST graph is generalized to an MSF graph, which is a union of MSTs [57]. At each level in the second registration pass, a mask for nodes, $\boldsymbol{M}_{\mathrm{n}}(p)$, is derived by down-sampling the corresponding image mask $\boldsymbol{M}_{\mathrm{i}}(\boldsymbol{r})$:

$$\boldsymbol{M}_{\mathrm{n}}(p) = \left\lfloor 0.5 + \frac{1}{\left|\mathcal{S}\left(r_p\right)\right|} \sum\nolimits_{r\in\mathcal{S}\left(r_p\right)} \boldsymbol{M}_{\mathrm{i}}(\boldsymbol{r}) \right\rfloor \quad (12)$$

Those edges connecting nodes of different mask values are then removed from the initial 26-neighbor graph, resulting in two disconnected subgraphs. Using the edge weight defined in (4), an MST is extracted for each subgraph using Prim's algorithm [46]. To simplify the reuse of data structure and MRF solver, a virtual tree is used to represent MSF by introducing an additional virtual root vertex as the parent of the MST root vertices. The data and regularization terms associated with the virtual root are both set to zero. Therefore, fast L1 regularization on an MSF is solved similarly by message passing over a virtual tree instead of manipulating separate trees.

E. Masked Interpolation of Motion Vector

MV interpolation is performed in two subtasks: the inverse consistent iterations of $\boldsymbol{u}(\boldsymbol{r})$ and $\boldsymbol{v}(\boldsymbol{r})$ at each level using (6a) and (6b); and the upsampling of node MVs for generating the

prior component of the next level node MVs and the final voxel-wise MVs. In the first registration pass, trilinear B-spline interpolation is performed using all eight surrounding nodes as references. In the second pass, when a motion mask is present, a neighboring node is only included if it belongs to the same mask as the interpolated point. Candidate reference nodes are searched on the two nearest neighboring axial node planes simultaneously. The search radius is increased in step-wise fashion until valid reference nodes become available. If the nearest eight surrounding reference nodes are all valid, trilinear interpolation is performed. Otherwise, inverse distance weighted (IDW) interpolation [58] is performed using the valid reference nodes.

III. Experiments

A. Liver MRI Data

In vivo abdominal MRI data were acquired for the validation of the proposed algorithm. This study was institutional review board approved and HIPAA compliant. Ten healthy volunteers (5 males and 5 females, 20–30 years old) were imaged after obtaining informed consent. From each volunteer, two breath-hold 3D abdominal images were acquired using a 1.5T scanner, 8-channel cardiac coil, LAVA-FLEX sequence, slice thickness 2 mm, matrix size 512 × 512 × 88–96, in plane resolution 0.66 to 0.78 mm. The volunteers were instructed to hold their breath at the end of maximum exhalation and then at maximum inhalation. The field of view was tuned for each individual so that it covered the whole liver. From the first volunteer, three additional breath-hold images were obtained at different inspiration levels, approximately evenly distributed between maximum exhalation and maximum inhalation.

Registration experiments were performed on intra-subject image pairs. For each subject, the maximum expiratory breath-hold image was set as the target for registration. 13 cases were constructed: cases 1–10 corresponded to 10 maximum expiratory and maximum inspiratory image pairs from the 10 subjects, and cases 1.1–1.3 corresponded to the three additional image pairs from the first subject by setting the other three breath-hold images as the source images in order of increasing inspiration level.

B. Parameter Setting in Liver MR Registration

A fixed parameter setting was used for all of the liver MR registration experiments. Three resolution levels were used for each registration pass, and the level parameters were the same for the first and the second pass. Using the right-left (RL), anterior-posterior (AP), inferior-superior (IS) coordinate system, the upsampling rate for the input images was $\boldsymbol{s} = (1,1,2)$. From the first level to the third level, the node spacing was $\boldsymbol{n} = (8,8,8),(6,6,6),(4,4,4)$ voxels, respectively, the MV quantization was $\boldsymbol{q} = (4,4,4),(2,2,2),(1,1,1)$, respectively, and the half-width of MV label search space was $\boldsymbol{l} = (8,15,18),(6,6,6),(4,4,4)$ respectively. For the i-th dimension, the label search subspace was defined as $\mathscr{L}(i) = \{-\boldsymbol{l}(i), \cdots, -1, 0, +1, \cdots, \boldsymbol{l}(i)\}$, and the corresponding space in voxel units was $\mathscr{L}'(i) = \mathscr{L}(i)\boldsymbol{q}(i)/\boldsymbol{s}(i)$. Note that $\boldsymbol{l}$ was anisotropic at the first level in order to cover maximum motion magnitudes in different directions while minimizing computational burden. The weighting parameter α in (1) was set to 16, which was empirically determined

by checking the standard deviation of the determinant of the Jacobian, *std(**J**)*, in the regions of smooth motion field.

The parameters for coarse segmentation were empirically determined as $c_m = 128$, $\gamma_m = 15$, $d_m = 4$, and fixed for all three levels. The fine segmentation parameters were $c_i = 128$, $\gamma_i = 15$,$d_i = 4$. The Gaussian filter for image smoothing in (11) was 3D isotropic with $\sigma = 1$ and a window length of 5. The threshold for activating RAMS segmentation and the second registration pass was $\tau = 4.0$.

C. Performance Evaluation on Liver MRI Data

For quantitative evaluation of automated segmentation, manual segmentation of the anatomical sliding interface between the internal organs and the abdominal/thoracic wall was performed for the target (maximum expiratory) image in each subject using ITK-SNAP [55]. The anatomical boundaries of the sliding interface were traced slice by slice in the high-resolution axial cross-section under the supervision of body radiologist with 24 years of experience (Dr. Martin R. Prince).

The Dice similarity coefficient (DSC) [59] was used to compare the resulted masks between automatic segmentation and manual segmentation:

$$DSC(\boldsymbol{X}, \boldsymbol{Y}) = \frac{2|\boldsymbol{X} \cap \boldsymbol{Y}|}{|\boldsymbol{X}| + |\boldsymbol{Y}|} \quad (13)$$

where $\boldsymbol{X}$ and $\boldsymbol{Y}$ represent the binary image masks, and $|\cdot|$ denotes the number of non-zero voxels. A higher DSC indicates a better matching.

In addition, the bidirectional local distance (BLD) [60] was calculated to measure the local surface differences by setting the manually segmented boundary as the reference surface:

$$BLD\left(p_{ref}, \boldsymbol{T}\right) = \max\left(FMinD\left(p_{ref}, \boldsymbol{T}\right), BMaxD\left(\boldsymbol{T}, p_{ref}\right)\right) \quad (14)$$

where the forward minimum distance $FMinD\left(p_{ref}, \boldsymbol{T}\right)$ is defined as the minimum of Euclidean distances from a point p_{ref} on a reference surface $\boldsymbol{R}$ to all points on a test surface $\boldsymbol{T}$, and the backward maximum distance $BMaxD\left(\boldsymbol{T}, p_{ref}\right)$ is defined as the maximum of those $FMinD\left(p_t, R\right)$, $p_t \in \boldsymbol{T}$ which equals $\left\|p_t - p_{ref}\right\|_2$.

To evaluate registration accuracy, the results of the first and second registration passes were compared under the same parameter settings. For additional validation of the obtained motion masks, the open source registration software Elastix v4.8 [61] was tested in two modes using the same recommended parameter file: affine transform followed by cubic B-spline transform, Normalized Correlation Coefficient (NCC) similarity metric, and adaptive stochastic gradient descent optimizer. In the first mode, a single cubic B-spline transform

was used for the entire image. In the second mode, the motion fields of the inner and the outer region were estimated in two separate registrations using the target image mask calculated by RAMS. Mutual information between image pairs was used to evaluate image alignment quantitatively, as in [12–14].

D. Additional Evaluation Using Liver CT Data

As presented in [62], a more reliable quantitative evaluation of registration accuracy is comparing the estimated MVs with the reference MVs derived from a set of densely distributed voxel-wise manually annotated corresponding landmark pairs. Unfortunately, no such dataset is currently publicly available for liver MR registration. Therefore, we evaluated RAMS using an open liver CT dataset hosted on ITK's medical development database (http://midas.kitware.com/community/view/47) as a supplement. Following previous work [17, 19, 21], RAMS was performed on the same four preprocessed image pairs with isotropic voxel spacing of 2×2×2 mm^3 and image dimension of (230~250)x(154~171)x(150~200) using the following parameters (see above for notation): $\boldsymbol{s}$ = (2,2,2), $\boldsymbol{n}$ = (5,5,5), (4,4,4), (3,3,3), $\boldsymbol{q}$ = (3,3,3), (2,2,2), (1,1,1), $\boldsymbol{l}$ = (4,4,7), (3,3,4), (3,3,3), $c_{\mathrm{m}} = 128$, $\gamma_{\mathrm{m}} = 18$, $d_{\mathrm{m}} = 2$, $c_{\mathrm{i}} = 128$, $\gamma_{\mathrm{i}} = 18$, and $d_{\mathrm{i}} = 2$. A series of $\alpha = 2^k, k = 0, 1, 2, 3, 4, 5, 6, 7$ were tested for each patient. To quantify the registration accuracy, the total registration error (TRE) was calculated for each test using the well-distributed lung landmarks and abdominal landmarks provided with this dataset [17], and compared with state-of-the-art results reported in the literature [17, 19, 21].

IV. Results

A. Segmentation Performance in Liver MR Registration

Fig. 3 shows the RAMS segmentation result of the target and source images for case 1. The sliding motion interface in each image was globally detected by coarse segmentation using the forward or backward motion field estimates. Only the final level coarse segmentation result is shown. The motion boundaries were aligned with anatomical boundaries of the sliding interface by the subsequent fine segmentation.

A 3D visualization of the segmented boundaries for the target image in case 1 is shown in Fig. 4. Without the motion based coarse segmentation, MR signal intensity based GCBAC found the outer body surface of the subject (Fig. 4a). In contrast, the boundaries segmented by RAMS (Fig. 4c) coincided well with the manually segmented sliding motion interface (Fig. 4b).

Table I summarizes the DSC results for the target image of each case. In case 1.1, RAMS failed to obtain a satisfactory segmentation due to small motion discontinuity with the final level $\max_{(p,q) \in \mathcal{N}} G_{\mathrm{m}}(p,q) = 3.7$. As presented in Section II.C.3), RAMS segmentation and the second registration pass are unnecessary and should be disabled for such cases, therefore case 1.1 is not included in the statistics. In all other cases, the coarse segmentation (with overall mean DSC of 92.8%±0.4%) greatly outperformed MR signal intensity-based segmentation (with overall mean DSC of 67.1%±1.1%), and the fine segmentation (with overall mean DSC of 96.5%±0.2%) further improved the result. From cases 1.2 and 1.3 to 1,

DSC of the coarse segmentation decreased slightly. In contrast, after image based fine segmentation, no such trend in DSC was observed.

Table II summarizes the averaged BLD for the segmentation of target images. In all cases except case 1.1, the coarse segmentation (with overall mean BLD of 4.3±0.3 mm) greatly outperformed MR signal intensity-based segmentation (with overall mean BLD of 36.5±1.2 mm), and the fine segmentation (with overall mean BLD of 2.3±0.1 mm) further improved the result. Both DSC and BLD metrics show a high consistency between RAMS segmentation and manual segmentation.

B. Registration Performance on Liver MRI Data

1) Comparison between DEEDS-SSC-L1 and RAMS—Fig. 5 shows a comparison of images before and after registration in case 2. Fig. 6 compares the resulting motion fields in the same axial, coronal, and sagittal slices of Fig. 5. After the first registration pass (DEEDS-SSC-L1), small features inside the liver in the source image were well aligned with those in the target image, whereas the tissues around the sliding interface were still misaligned (Fig. 5, third column). After the second registration pass (RAMS), the spatial alignment in this neighborhood was well preserved.

The improvement in image alignment resulted from the improvement in motion estimation accuracy, as shown in Fig. 6. In DEEDS-SSC-L1, the sliding motion interface was over-smoothed to an erroneous wide transition zone. This led to an incorrect transformation of the voxels in the source image, resulting in the misalignment of image features. By using explicit masked regularization and masked interpolation in the second registration pass, the registration accuracy around the sliding interface has been substantially improved in RAMS (Fig. 5, fourth column).

2) Elastix with and without the RAMS Derived Mask—Fig. 7 shows the transformed source images after Elastix registration in the same cross sections of Figs. 5–6. A large neighborhood around the sliding interface (arrows in Fig. 7, left column) was not aligned with the target image (Fig. 5, first column), due to incorrect homogeneous smoothness constraints when using a single cubic B-spline transform. By introducing the target image mask calculated by RAMS into Elastix registration, two separate cubic B-spline transforms were used to model the motion field in two regions, so that motion smoothness constraints across the sliding interface were removed, leading to a substantial improvement in motion estimation accuracy and better image alignment after registration (Fig. 7, right column).

Fig. 8 presents volumetric 3D rendering of the inner regions of the images in case 2. Image alignment around the sliding interface is significantly improved by using explicit motion masks derived from RAMS segmentation in both registration methods.

Table III summarizes the mutual information (MI) between each image pair. For each method and for every case (except case 1.1), under the same parameter settings, the MI after registration using a mask was greater than without a mask, indicating improved registration accuracy. The incremental MI was larger in Elastix than in RAMS, showing more benefits

brought by using a motion mask in Elastix. By incorporating the motion mask information, Elastix achieved a similar level of spatial alignment after registration. We also observed that in case 1.1, the MI decreased on the contrary after the second registration pass in RAMS. This might be attributed to inaccurate segmentation of the sliding interface. However, in RAMS, the segmentation and the second registration pass are not performed in this case since $\max_{(p,q)\in\mathcal{N}} G_{\mathrm{m}}(p,q) < \tau$, and the accuracy of the first registration pass (DEES-SCC-L1) is sufficient.

3) Quantification of Deformation Smoothness—To examine spatial smoothness of the result motion field of RAMS, we computed voxel-wise Jacobians of the forward motion field in each case. Using the fine segmentation derived by RAMS, voxels in a target image were partitioned into three regions: the inner region, the sliding interface, and the outer region. To include boundary voxels on both sides, the initial sliding interface was morphologically dilated using a spherical structuring element whose radius was 1 voxel. The standard deviation of Jacobians (Std(J)) was calculated separately for individual regions and summarized in Table IV.

According to Table IV, large transformation complexity is observed around the sliding interface (with overall mean Std(J) of 5.21±0.53), whereas smooth deformation is observed both in the inner region (with overall mean Std(J) of 0.39±0.02) and the outer region (with overall mean Std(J) of 0.19±0.01). By incorporating motion masks into registration, RAMS is able to preserve motion discontinuity along the sliding interface while obtaining plausible transformations without folding inside each of the two separate domains.

4) Quantification of Sliding Motion—To quantify the locations and level of sliding motion detected by RAMS, we used a sliding motion measure proposed in [63] which calculates voxel-wise maximum shear stretch $\gamma_{\max}$ of the estimated motion field. It has been shown that respiratory sliding motion between lung boundaries and the chest wall could be well characterized by voxel-wise $\gamma_{\max}$ based only on the motion field obtained from lung CT registration [18, 63]. the maximum shear stretch $\gamma_{\max}$ is defined at each voxel as follows:

$$\lambda_i = \sqrt{\mathrm{eigenvalues}\left(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F}\right)},\ i = 1,2,3 \quad (15)$$

$$\gamma_{\max} = \left(\lambda_1 - \lambda_3\right)/2, \quad (16)$$

where **F** is 3×3 matrix representing the gradient tensor of 3D motion field at this voxel, λ_1 and λ_3 are the maximal and minimal principal stretch components, respectively [63].

Fig. 9 visualizes voxel-wise $\gamma_{\max}$ of the forward motion field $\boldsymbol{u}(\boldsymbol{r})$ estimated by DEEDS-SSC-L1 and RAMS for case 10 in an example coronal section and an example axial section. The manual segmentation and RAMS fine segmentation of the same cross-sections are also

shown in Fig. 9(a) for reference. The value of $\gamma_{\max}$ is linearly mapped to pseudo color using "jet" colormap. As presented in [18, 63], $\gamma_{\max} = 5$ can be set as a noticeable level of sliding motion. Distinct contours, which correspond to very high values of $\gamma_{\max}$ ($\gg 5$), can be easily observed in Fig. 9(c). And the locations of these contours are highly consistent with RAMS segmentation of the sliding interface shown in Fig. 9(a). In contrast, $\gamma_{\max}$ are smaller and the data boundaries are smooth in DEEDS-SSC-L1 as shown in Fig. 9(b).

5) Evaluation of Image-adaptive MST/MSF—In MST/MSF extraction, edge weights $W(p, q)$ in (4) are defined as the gradient between image patches. Therefore, locally adaptive regularization is achieved by removing most edges (thus motion smoothness constraints) between nodes with large intensity difference. We investigated the influence of using image gradient as edge cost in MST/MSF construction on RAMS segmentation and registration performance by comparing with the results using randomized edge cost. Fig. 10 visually compares DSC, average BLD, and MI results on our liver MRI dataset, where the data in red, blue and cyan color are copied from Table I-III.

As shown in Fig. 10(a) and (b), the accuracy of coarse segmentation decreases when using randomized edge cost, probably due to the drop in accuracy of motion field estimates obtained in the first registration pass. Nevertheless, after the subsequent fine segmentation, similar accuracy is achieved as usual. This demonstrates the robustness of image-based fine segmentation with respect to small changes in the initial boundaries output from coarse segmentation.

As seen in Fig. 10(c), the registration accuracy in terms of image matching decreases in both the first and second registration pass when using random edge cost in MST/MSF construction. More interestingly, RAMS result using random edge cost (green curve) is better than DEEDS-SSC-L1 result using image-adaptive edge cost (blue curve). This demonstrates the effectiveness of using explicit motion masks in respiratory motion registration.

C. Registration Performance on Liver CT Data—For each α, the TRE was calculated using all of the lung and abdominal landmarks from the four patients. The overall TRE was observed to reach a minimum when $\alpha = 8$, and increase gradually when deviating from 8. Table V compares the TRE statistics of RAMS ($\alpha = 8$) based on various set of landmarks to the state-of-the-art registration results reported in the literature (Anisotropic Diffusion [17], SLIC Demons [19], and XFFD [21]). Using Wilcoxon rank sum test, the TREs obtained by RAMS are statistically significant lower ($p < 0.05$) for both the lung landmarks and the abdominal landmarks.

D. Running Time

RAMS was implemented in C++, compiled and run on a desktop computer using an Intel i7–5820K CPU at 3.30GHz with 64GB RAM running Windows 7 Enterprise. For cases 2–10 in liver MR registration, where each image had the same dimension of 512×512×96, the average running time was 39 seconds for SSC precomputation, 712 seconds for the first

registration pass, 6 seconds for the coarse segmentation, 274 seconds for the fine segmentation, and 347 seconds for the second registration pass. The average total registration time for each case was about 23 minutes. The coarse segmentation across three levels converged in maximum 22 iterations, and the fine one converged in maximum 37 iterations. The time increase in fine segmentation was due to the vast increase in the number of vertices in the s-t min-cut computation. The time for the second registration pass was less than half of that for the first pass. The saved time mainly came from the reuse of data costs at the first level, which consumed nearly half of the first pass running time in order to cover the large MV search range.

V. Discussion

The main contribution of RAMS approach is the novel adaptation and combination of two discrete optimization-based methods – DEEDS registration and GCBAC segmentation – for discontinuity preserving respiratory motion registration of liver MR images. And the key innovation is the proposed two-stage segmentation process. Using the motion field estimates obtained in the first mask-free registration pass, a coarse segmentation globally locates motion boundaries. Using the MR signal intensity, a subsequent fine segmentation aligns the motion boundaries with anatomical boundaries of the sliding motion interface.

As validated on in vivo breath-hold liver MRI data acquired in 10 healthy subjects across a large range of respiratory displacements, the motion masks calculated by RAMS agreed well with manual segmentations. By incorporating the motion segmentation information in the second registration pass, motion discontinuity around the sliding interface was preserved substantially better, leading to a significant improvement in image alignment after registration.

RAMS requires just two registration passes and one embedded segmentation process. This is in contrast to recently reported joint registration and segmentation methods for liver MR [12–14], where an iterative continuous optimization scheme was used at each of a series of coarse-to-fine pyramid levels to avoid local minima. At each level, many alternating optimizations of motion field estimation and segmentation were needed for convergence. The efficiency of RAMS can be primarily attributed to the proposed two-stage segmentation approach, which exploits both the information from the motion field estimates and from the MR signal intensity. The motion-based segmentation coarsely finds the global optimum, and the subsequent image-based segmentation refines the result to anatomical boundaries. This novel combination achieves fast convergence and avoids time-consuming iterations between motion registration and motion segmentation. Because the motion field-based coarse segmentation is crucial, it requires a high-performance mask-free registration algorithm in the first pass. In RAMS, this is accomplished by DEEDS-SSC-L1.

The improvement of registration accuracy by explicitly using motion masks in the regularization has been extensively validated in the literature, particularly for the respiratory motion registration of thoracic and abdominal images [7], [8], [38], [26]. Compared to thoracic CT images, automated segmentation of the sliding interface in abdominal MR images is more challenging due to the low contrast between internal organs and the

abdominal wall. As validated using Elastix, the motion masks calculated by RAMS can also benefit other registration algorithms, including continuous algorithms. As a byproduct of registration, the obtained segmentation may be used for other medical image analysis and clinical applications.

There are some limitations in the current work. The quantitative assessment of registration accuracy using well-distributed manually annotated landmark pairs was only done on the liver CT data which is publicly available. Nevertheless, under the same parameter settings, the comparison of image alignment between MR registrations with and without the motion mask is also a good metric. Additionally, a high agreement between RAMS segmentation and manual segmentation was observed.

The proposed two-stage GCBAC segmentation is designed to work for images with strong motion discontinuities. As observed in case 1.1, RAMS coarse segmentation fails to detect the sliding interface when motion discontinuity is small. A subsequent intensity-based segmentation would therefore have decreased accuracy. For this reason, in RAMS, such situation is detected automatically and the segmentation and the second registration pass are not carried out. In RAMS, the data similarity metric in registration uses the self-similarity context, which has robust performance in geometric matching with respect to intensity changes. However, like other intensity-based segmentation algorithms, the fine segmentation is sensitive to MRI artifacts such as image noise, partial volume effect (PVE), and bias field effect. We suppose that these artifacts could be effectively reduced by using modern MRI scanners with more advanced acquisition technology. As validated on the real MRI data acquired in our work, these artefacts are tolerable to the presented method. Additional preprocessing steps, e.g. bias field correction, will be further investigated to improve the robustness to MRI artefacts.

Recently, model-based methods incorporating liver shape prior knowledge have demonstrated superior performance on automated liver segmentation in CT scans, compared to conventional intensity-based segmentation methods [28–30]. We will investigate the potential to improve the accuracy and robustness of motion mask extraction in abdominal MR images by incorporating the model-based idea in our future work.

Running time is often comprised to improve registration accuracy in respiratory motion registration by introducing more sophisticated locally adaptive regularization. To enable realtime intra-procedure 3D motion estimations in MR-guided interventions, C. Baumgartner *et al.* [64] proposed a novel auto-adaptive 3D motion modeling using 2D motion fields estimated from sagittal MR slices. Nevertheless, this model-based motion estimation still requires a long calibration time of 12 min and a large prediction latency of 800 ms. By using complete discrete optimization, the registration time of RAMS is predictable. Compared to the running time of continuous optimization-based approaches with locally adaptive regularization, e.g. more than 60 minutes for a typical lung CT registration in [18], the running time of RAMS is competitive. And the current implementation of RAMS is an initial non-optimized version for algorithm validation. For clinical application, RAMS could be accelerated by parallel processing, e.g. using General Purpose Graphics Processing Units (GPGPUs).

VI. Conclusion

In summary, RAMS is a discontinuity preserving liver MR registration approach that, by combining motion field estimates and MR signal intensity, is shown to efficiently find the anatomical boundaries of the sliding motion interface and thereby improve overall motion estimation accuracy.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

We would like to thank Dr. Mattias P. Heinrich for sharing the DEEDS code (http://www.mpheinrich.de/software.html) and thank Dr. Danielle F. Pace for sharing manually annotated landmarks for the MIDAS liver CT dataset. We would also like to thank the anonymous reviewers for their valuable comments and suggestions that helped improve the quality of this paper.

References

[1]. Hamy V et al., "Respiratory motion correction in dynamic MRI using robust data decomposition registration - application to DCE-MRI," Med Image Anal, vol. 18, no. 2, pp. 301–313, Feb. 2014. [PubMed: 24322575]

[2]. Coenegrachts K et al., "Prediction and monitoring of treatment effect using T1-weighted dynamic contrast-enhanced magnetic resonance imaging in colorectal liver metastases: potential of whole tumour ROI and selective ROI analysis," Eur J Radiol, vol. 81, no. 12, pp. 3870–3876, 12 2012. [PubMed: 22944331]

[3]. Carrillo A et al., "Semiautomatic 3-D image registration as applied to interventional MRI liver cancer treatment," IEEE Trans Med Imaging, vol. 19, no. 3, pp. 175–185, 3 2000. [PubMed: 10875702]

[4]. Yan Z et al., "Atlas-based liver segmentation and hepatic fat-fraction assessment for clinical trials," Computerized Medical Imaging and Graphics, vol. 41, pp. 80–92, 2015. [PubMed: 24962337]

[5]. Brown LG, "A survey of image registration techniques," ACM Comput. Surv, vol. 24, no. 4, pp. 325–376, 1992.

[6]. Maintz JB, and Viergever MA, "A survey of medical image registration," Med Image Anal, vol. 2, no. 1, pp. 1–36, 3 1998. [PubMed: 10638851]

[7]. Sotiras A, Davatzikos C, and Paragios N, "Deformable medical image registration: a survey," IEEE Trans Med Imaging, vol. 32, no. 7, pp. 1153–1190, 7 2013. [PubMed: 23739795]

[8]. Murphy K et al., "Evaluation of registration methods on thoracic CT: the EMPIRE10 challenge," IEEE Trans Med Imaging, vol. 30, no. 11, pp. 1901–1920, 11 2011. [PubMed: 21632295]

[9]. Castillo R et al., "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets," Phys Med Biol, vol. 54, no. 7, pp. 1849–1870, 4 2009. [PubMed: 19265208]

[10]. Castillo E et al., "Four-dimensional deformable image registration using trajectory modeling," Phys Med Biol, vol. 55, no. 1, pp. 305–327, 1 2010. [PubMed: 20009196]

[11]. Castillo R et al., "A reference dataset for deformable image registration spatial accuracy evaluation using the COPDgene study archive," Phys Med Biol, vol. 58, no. 9, pp. 2861–2877, 5 2013. [PubMed: 23571679]

[12]. Kiriyanthan S, Fundana K, and Cattin P, "Discontinuity Preserving Registration of Abdominal MR Images with Apparent Sliding Organ Motion," in Abdominal Imaging. Computational and Clinical Applications New York, NY, USA: Springer, 2012, pp. 231–239.

[13]. Kiriyanthan S. A primal-dual approach for discontinuity preserving registration. in Proc. IEEE 9th Int. Symp. Biomed. Imaging; May 2012; 350–353.

[14]. Kiriyanthan S. A landmark-based primal-dual approach for discontinuity preserving registration. in Proc. 4th Int. Conf. Abdominal Imag.: Comput. Clin. Appl.; 2012; 137–146.

[15]. Schmidt-Richberg A et al., "Estimation of slipping organ motion by registration with direction-dependent regularization," Med Image Anal, vol. 16, no. 1, pp. 150–159, 1 2012. [PubMed: 21764627]

[16]. Risser L et al., "Piecewise-diffeomorphic image registration: application to the motion estimation between 3D CT lung images with sliding conditions," Med Image Anal, vol. 17, no. 2, pp. 182–193, 2 2013. [PubMed: 23177000]

[17]. Pace DF, Aylward SR, and Niethammer M, "A Locally Adaptive Regularization Based on Anisotropic Diffusion for Deformable Image Registration of Sliding Organs," IEEE Trans Med Imaging, vol. 32, no. 11, pp. 2114–2126, 11 2013. [PubMed: 23899632]

[18]. Papiez BW et al., "An implicit sliding-motion preserving regularisation via bilateral filtering for deformable image registration," Med Image Anal, vol. 18, no. 8, pp. 1299–3311, 12 2014. [PubMed: 24968741]

[19]. Papież BW. Liver Motion Estimation via Locally Adaptive Over-Segmentation Regularization. in Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assisted Intervention; 2015; 427–434.

[20]. Vishnevskiy V et al., "Isotropic Total Variation Regularization of Displacements in Parametric Image Registration," IEEE Trans Med Imaging, vol. 36, no. 2, pp. 385–395, 2 2017. [PubMed: 27654322]

[21]. Hua R et al., "Multiresolution eXtended Free-Form Deformations (XFFD) for non-rigid registration with discontinuous transforms," Med Image Anal, vol. 36, pp. 113–122, 2 2017. [PubMed: 27894001]

[22]. Glocker B et al., "Dense image registration through MRFs and efficient linear programming," Med Image Anal, vol. 12, no. 6, pp. 731–41, 12 2008. [PubMed: 18482858]

[23]. Heinrich MP et al., "MRF-based deformable registration and ventilation estimation of lung CT," IEEE Trans Med Imaging, vol. 32, no. 7, pp. 1239–1248, 7 2013. [PubMed: 23475350]

[24]. Xu Z et al., "Evaluation of Six Registration Methods for the Human Abdomen on Clinically Acquired CT," IEEE Trans Biomed Eng, vol. 63, no. 8, pp. 1563–1572, 8 2016. [PubMed: 27254856]

[25]. Heinrich M, Handels H, and Simpson I, "Estimating Large Lung Motion in COPD Patients by Symmetric Regularised Correspondence Fields," in Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assisted Intervention, 2015, pp. 338–345.

[26]. Vandemeulebroucke J et al., "Automated segmentation of a motion mask to preserve sliding motion in deformable registration of thoracic CT," Med Phys, vol. 39, no. 2, pp. 1006–1015, 2 2012. [PubMed: 22320810]

[27]. Hosseini-Asl E et al., "3-D Lung Segmentation by Incremental Constrained Nonnegative Matrix Factorization," IEEE Trans Biomed Eng, vol. 63, no. 5, pp. 952–963, 5 2016. [PubMed: 26415200]

[28]. Wang J et al., "Shape–intensity prior level set combining probabilistic atlas and probability map constrains for automatic liver segmentation from abdominal CT images," INT J Comput Ass Rad, vol. 11, no. 5, pp. 817–826, 5 2016.

[29]. Shi C et al., "A hierarchical local region-based sparse shape composition for liver segmentation in CT scans," Pattern Recogn, vol. 50, pp. 88–106, 2 2016.

[30]. Shi C et al., "Low-rank and sparse decomposition based shape model and probabilistic atlas for automatic pathological organ segmentation," Med Image Anal, vol. 38, pp. 30–49, May 2017 [PubMed: 28279915]

[31]. Chartrand G et al., "Liver Segmentation on CT and MR Using Laplacian Mesh Optimization," IEEE Trans Biomed Eng, vol. 64, no. 9, pp. 2110–2121, 9 2017. [PubMed: 27893375]

[32]. Preston JS, Joshi S, and Whitaker R, "Deformation Estimation with Automatic Sliding Boundary Computation," in Proc. 19th Int. Conf. Med. Image Comput. Comput.-Assisted Intervention, 2016, pp. 72–80.

[33]. Jud C. Directional Averages for Motion Segmentation in Discontinuity Preserving Image Registration. in Proc. 20th Int. Conf. Med. Image Comput. Comput.-Assisted Intervention; 2017; 249–256.

[34]. Swierczynski P, Papież BW, Schnabel JA, and Macdonald C, "A level-set approach to joint image segmentation and registration with application to CT lung imaging," Computerized Medical Imaging and Graphics, vol. 65, pp. 58–68, 2018/04/01/, 2018. [PubMed: 28705410]

[35]. Brox T. High Accuracy Optical Flow Estimation Based on a Theory for Warping. in Proc. 8th European Conf. Comput. Vis.; 2004. 25–36.

[36]. Chan TF, Esedoglu S, and Nikolova M, "Algorithms for finding global minimizers of image segmentation and denoising models," Siam J Appl Math, vol. 66, no. 5, pp. 1632–1648, 2006.

[37]. Chambolle A, and Pock T, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," J Math Imaging Vis, vol. 40, no. 1, pp. 120–145, 2011.

[38]. Mahapatra D, and Sun Y, "Integrating segmentation information for improved MRF-based elastic image registration," IEEE Trans Image Process, vol. 21, no. 1, pp. 170–183, 1 2012. [PubMed: 21791411]

[39]. Gass T, Székely G, and Goksel O, "Simultaneous Segmentation and Multiresolution Nonrigid Atlas Registration," IEEE Trans Image Process, vol. 23, no. 7, pp. 2931–2943, 7 2014. [PubMed: 24816586]

[40]. Heinrich M. Towards Realtime Multimodal Fusion for Image-Guided Interventions Using Self-similarities. in Proc. 16th Int. Conf. Med. Image Comput. Comput.-Assisted Intervention; 2013. 187–194.

[41]. Xu N, Ahuja N, and Bansal R, "Object segmentation using graph cuts based active contours," Comput Vis Image Underst, vol. 107, no. 3, pp. 210–224, 2007.

[42]. Chatfield K, Philbin J, and Zisserman A, "Efficient retrieval of deformable shape classes using local self-similarities," in Proc. IEEE 12th Int. Conf. Comput. Vis. Workshops, 2009, pp. 264–271.

[43]. Shechtman E, and Irani M, "Matching Local Self-Similarities across Images and Videos," in Proc. IEEE Conf. Comput. Vis. and Pattern Recognit., 2007, pp. 1–8.

[44]. Heinrich MP et al., "MIND: modality independent neighbourhood descriptor for multi-modal deformable registration," Med Image Anal, vol. 16, no. 7, pp. 1423–1435, Oct. 2012. [PubMed: 22722056]

[45]. Felzenszwalb PF, and Huttenlocher DP, "Efficient belief propagation for early vision," Int J Comput Vision, vol. 70, no. 1, pp. 41–54, 10 2006.

[46]. Prim RC, "Shortest Connection Networks and Some Generalizations," Bell System Technical Journal, vol. 36, no. 6, pp. 1389–1401, 1957.

[47]. Christensen GE, and Johnson HJ, "Consistent image registration," IEEE Trans Med Imaging, vol. 20, no. 7, pp. 568–582, 7 2001. [PubMed: 11465464]

[48]. Hu W et al., "A system for learning statistical motion patterns," IEEE Trans Pattern Anal Mach Intell, vol. 28, no. 9, pp. 1450–1464, 9 2006. [PubMed: 16929731]

[49]. Wu S, and Wong HS, "Crowd motion partitioning in a scattered motion field," IEEE Trans Syst Man Cybern B Cybern, vol. 42, no. 5, pp. 1443–1454, 10 2012. [PubMed: 22562767]

[50]. Chen K, Zhou Z, and Wu W, "Progressive Motion Vector Clustering for Motion Estimation and Auxiliary Tracking," ACM Trans. Multimedia Comput. Commun. Appl, vol. 11, no. 3, pp. 1–23, 2015.

[51]. Li D. Discontinuity preserving registration using truncated L1 regularization and minimum spanning tree based motion clustering. in Proc. ISMRM 24th Ann. Meeting; 2016. 0781

[52]. Grygorash O, Zhou Y, and Jorgensen Z, "Minimum spanning tree based clustering algorithms," in Proc. 8th Int. Conf. Tools Artif. Intell., 2006, pp. 73–81.

[53]. Kass M, Witkin A, and Terzopoulos D, "Snakes - Active Contour Models," Int J Comput Vision, vol. 1, no. 4, pp. 321–331, 1987.

[54]. Lee HY et al., "Automatic left ventricle segmentation using iterative thresholding and an active contour model with adaptation on short-axis cardiac MRI," IEEE Trans Biomed Eng, vol. 57, no. 4, pp. 905–913, 4 2010. [PubMed: 19203875]

[55]. Yushkevich PA et al., "User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability," Neuroimage, vol. 31, no. 3, pp. 1116–1128, 7 2006. [PubMed: 16545965]

[56]. Ahuja RK, and Orlin JB, "A Fast and Simple Algorithm for the Maximum Flow Problem," Operations Research, vol. 37, no. 5, pp. 748–759, 9 1989.

[57]. Cousty J et al., "Watershed cuts: minimum spanning forests and the drop of water principle," IEEE Trans Pattern Anal Mach Intell, vol. 31, no. 8, pp. 1362–1374, 8 2009. [PubMed: 19542572]

[58]. Shepard D, "A two-dimensional interpolation function for irregularly-spaced data," in Proc. 23rd ACM National Conf., 1968, pp. 517–524.

[59]. Pluempitiwiriyawej C et al., "STACS: new active contour scheme for cardiac MR image segmentation," IEEE Trans Med Imaging, vol. 24, no. 5, pp. 593–603, May 2005. [PubMed: 15889547]

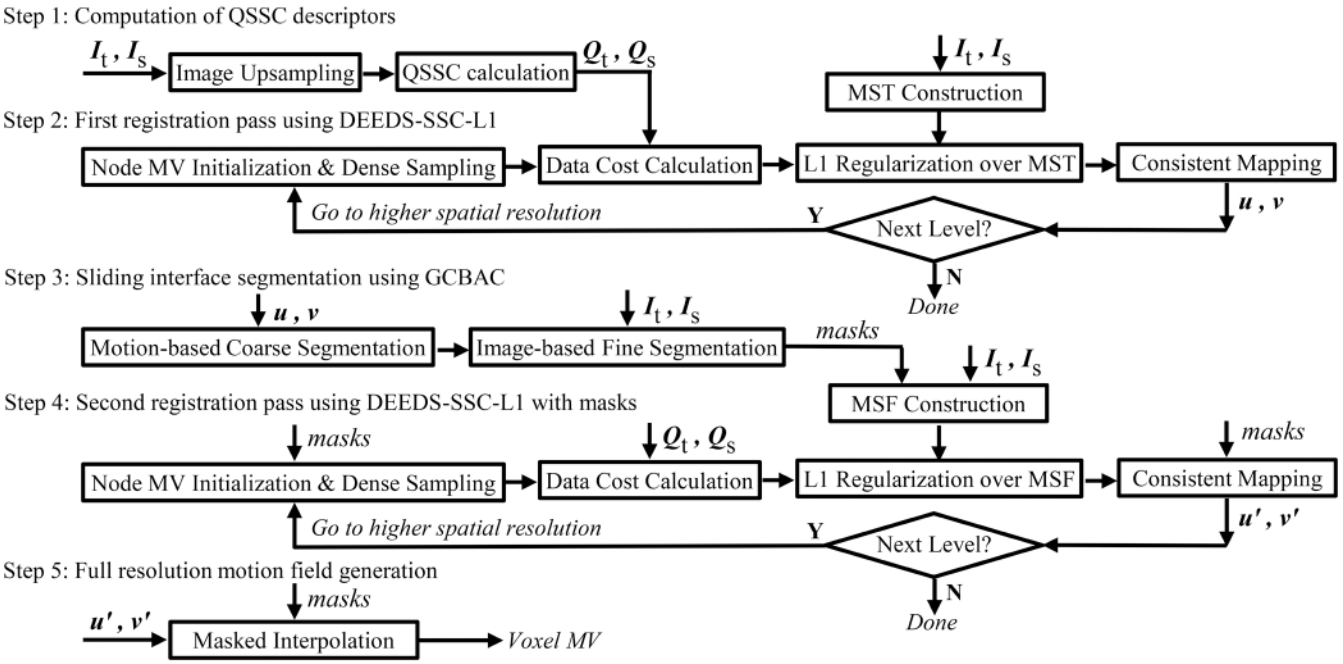
[60]. Kim HS et al., "Bidirectional local distance measure for comparing segmentations," Med Phys, vol. 39, no. 11, pp. 6779–6790, 11 2012. [PubMed: 23127072]

[61]. Klein S et al., "elastix: a toolbox for intensity-based medical image registration," IEEE Trans Med Imaging, vol. 29, no. 1, pp. 196–205, 1 2010. [PubMed: 19923044]

[62]. Rohlfing T, "Image similarity and tissue overlaps as surrogates for image registration accuracy: widely used but unreliable," IEEE Trans Med Imaging, vol. 31, no. 2, pp.153–163, 2 2012. [PubMed: 21827972]

[63]. Amelon RE et al., "A Measure for Characterizing Sliding on Lung Boundaries," Ann Biomed Eng, vol. 42, no. 3, pp. 642–650, 3 2014. [PubMed: 24114112]

[64]. Baumgartner CF et al., "Autoadaptive motion modelling for MR-based respiratory motion estimation," Med Image Anal, vol. 35, pp. 83–100, Jan, 2017. [PubMed: 27343436]



Abbreviations: DEEDS – dense displacement sampling, GCBAC – graph cuts based active contour, MSF – minimum spanning forest, MST – minimum spanning tree, MV – motion vector, QSSC – quantized self-similarity context.

Fig. 1.

Flowchart of the proposed registration with 3D active contour motion segmentation (RAMS) approach for liver MR registration.

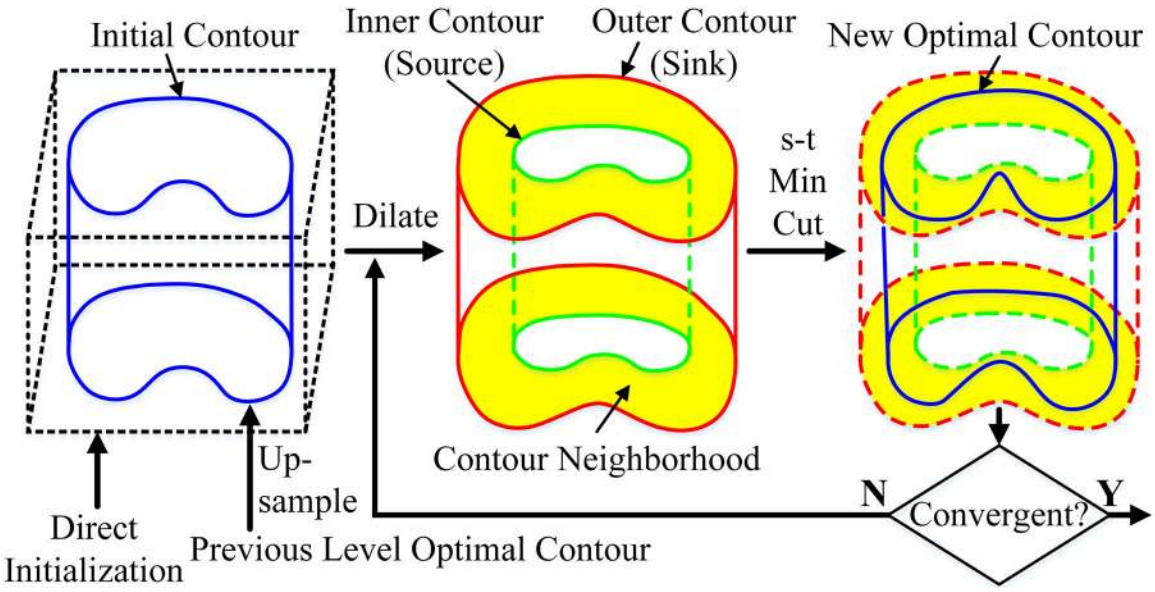


Fig. 2.

Illustration of RAMS segmentation process at each resolution level using graph cuts-based active contour (GCBAC) approach.

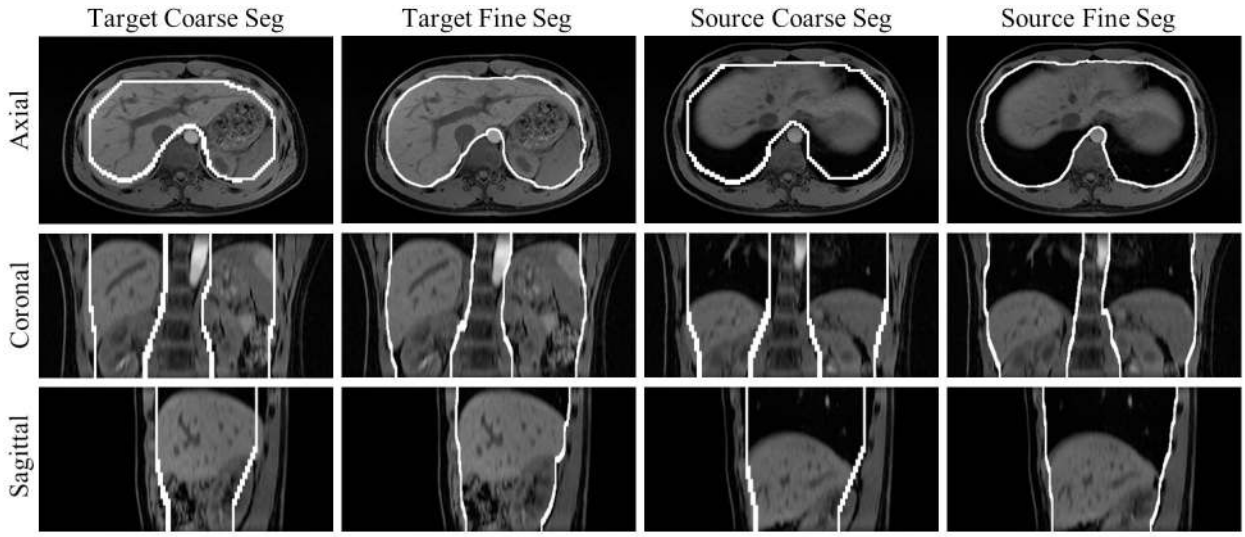


Fig. 3.

RAMS segmentation result for case 1. The left two columns show the final level coarse segmentation and the fine segmentation of the target image, and the right two columns show those of the source image. The rows show an example axial slice (top), coronal slice (middle), and sagittal slice (bottom), with the same slice coordinates for all columns. The white contours represent the outer boundaries of the inner segments.

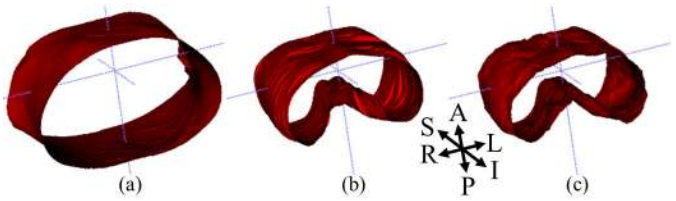


Fig. 4.

3D visualization of the target image mask boundaries for case 1, generated by: (a) GCBAC segmentation using MR image only, (b) manual segmentation, and (c) RAMS segmentation.

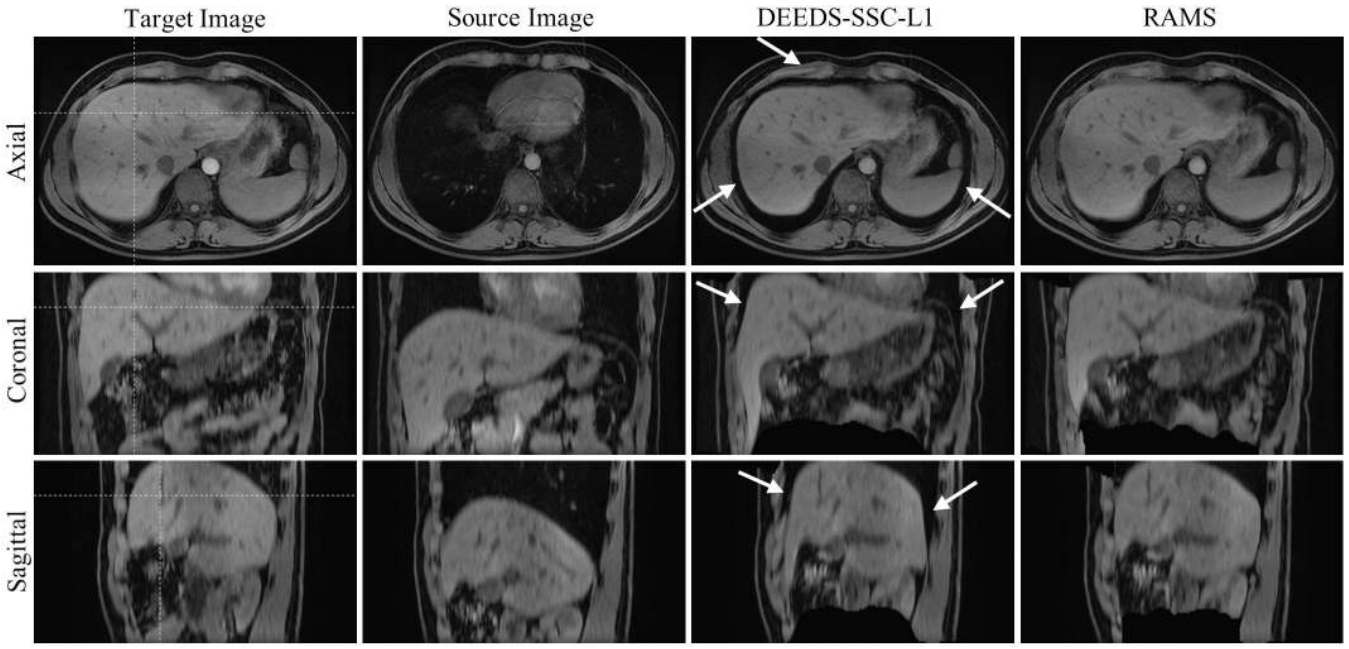


Fig. 5.

Comparison of images before and after registration using DEEDS-SSC-L1 and RAMS in case 2. The left two columns show the target image and the source image, respectively. The right two columns show the transformed source images of DEEDS-SSC-L1 and RAMS. The rows show an example axial slice (top), coronal slice (middle), and sagittal slice (bottom). The white dotted lines in the leftmost column indicate the location of cross sections. The white arrows (third column) indicate decreased image alignment with DEES-SSC-L1 at the sliding motion interface. These are resolved when using RAMS (fourth column).

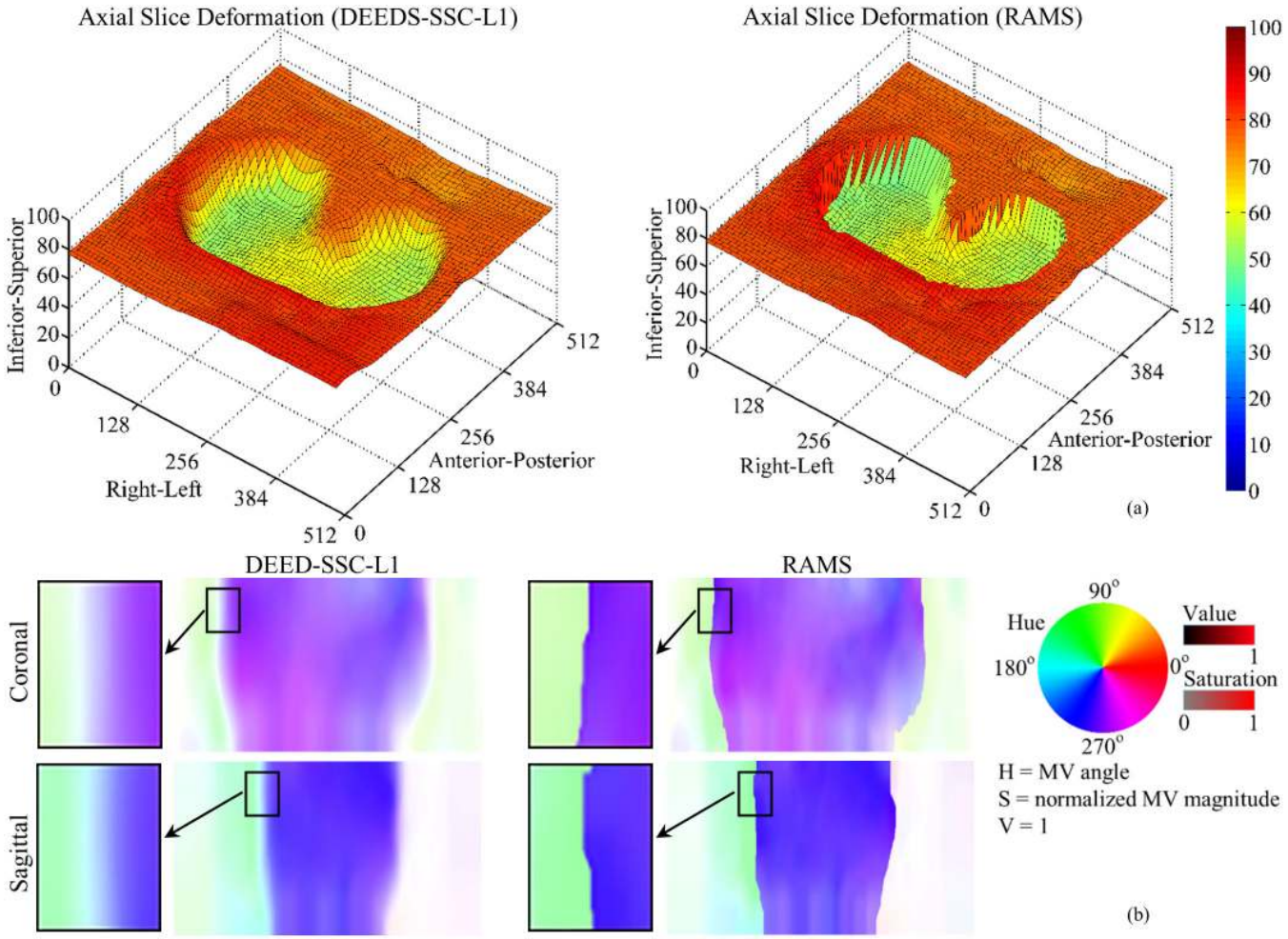


Fig. 6.

Comparison of the result motion fields of DEEDS-SSC-L1 and RAMS in case 2 in the same cross sections of Fig. 5: (a) deformed target axial slices in the source image, and (b) motion field projections on the coronal and the sagittal slices using HSV color mapping, with magnified regions of interest. The sliding motion interface is over-smoothed in DEEDS-SSC-L1 whereas it is well preserved in RAMS, leading to substantially improved image alignment, as seen in Fig. 5.

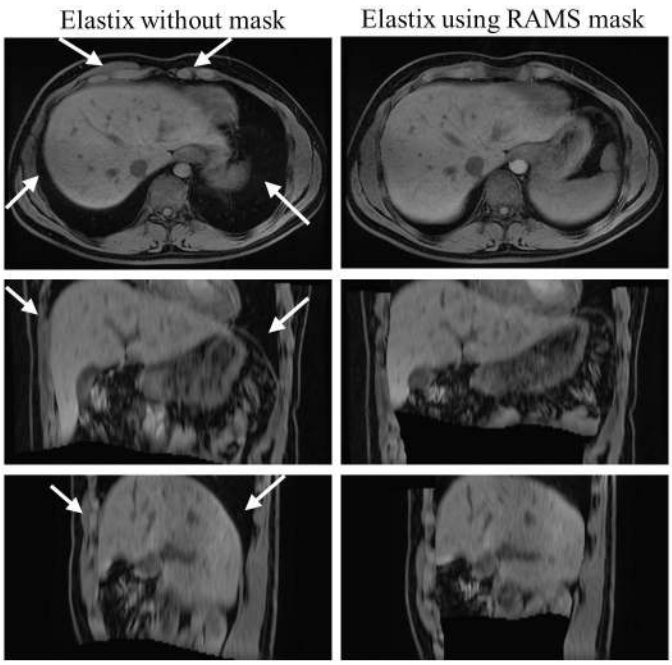


Fig. 7.

Comparison of the transformed source images after Elastix registration for case 2 in the same cross sections of Fig. 5. Compared to the target image in Fig. 5, regions around the sliding interface are misaligned when not using a mask (white arrows, left column). Using the RAMS derived mask, the result of Elastix is significantly improved (right column).

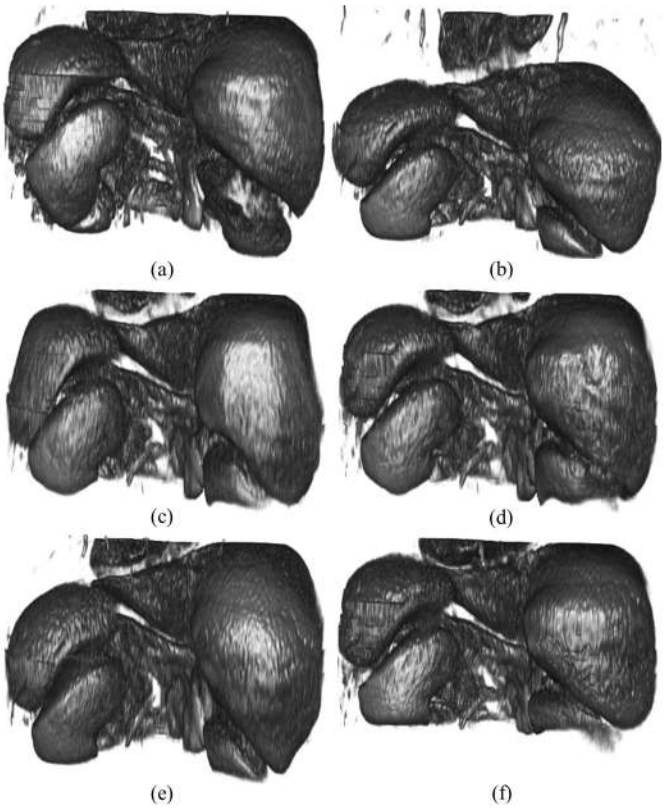


Fig. 8.

Volumetric rendering of the inner regions of the following images in case 2: (a) the target image, (b) the source image, the transformed source image of (c) DEEDS-SSC-L1, (d) RAMS, (e) Elastix without mask, and (f) Elastix using the RAMS derived mask. Image alignment around the sliding interface is significantly improved comparing (d) to (c), and (f) to (e), respectively.

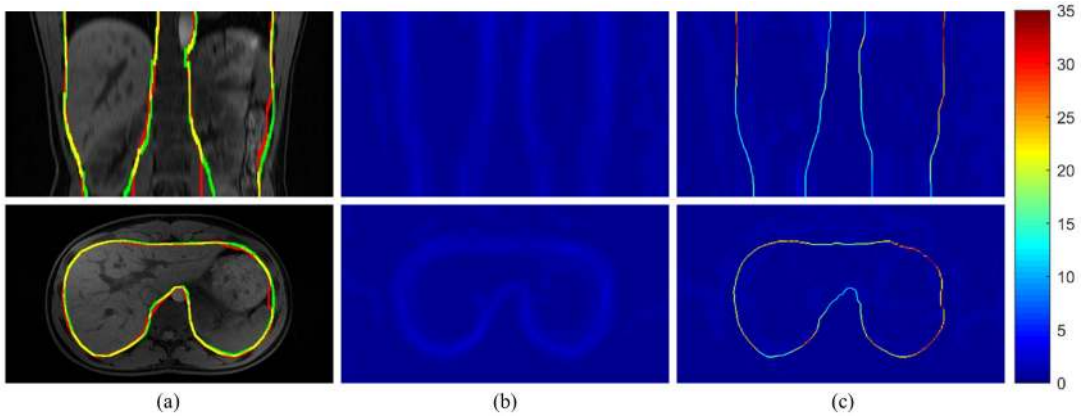


Fig. 9.

Quantification of the sliding motion for case 10 of the liver MRI dataset. The top row shows an example coronal slice and the bottom shows an example axial slice of the target image: (a) MR slice images overlaid with contours of RAMS fine segmentation in red, manual segmentation in green, and their overlaps in yellow; the maximal shear stretch of the forward motion field output from (b) DEEDS-SSC-L1 and (c) RAMS registration. Through the same linear pseudo color mapping, it can be easily observed that RAMS preserves sliding motion with sharp values of the maximal shear stretch at the sliding interface, while motion discontinuity is smoothed in DEEDS-SSC-L1.

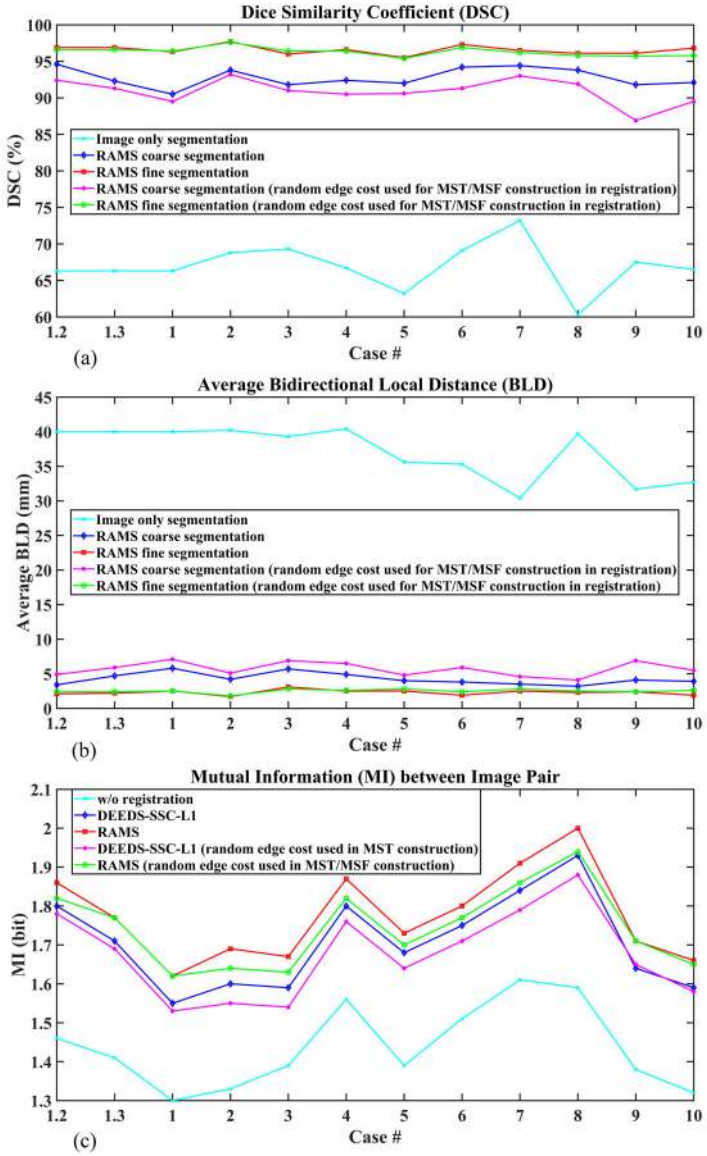


Fig. 10.

Evaluation of the impact of using image gradient as edge cost in MST/ MSF construction on RAMS segmentation and registration performance by comparing with the results using randomized edge cost on the liver MRI dataset: (a) DSC, (b) average BLD and (c) MI. According to (a) and (b), the segmentation accuracy decreases in motion-based coarse segmentation when using randomized edge cost, but recovers after image-based fine segmentation. According to (c), the registration accuracy decreases in both the first and second registration pass when using random edge cost.

TABLE I

DSC (%) of the Masks Generated by GCBAC Segmentation Using MR Image Only, the Coarse Segmentation of RAMS, and the Subsequent Fine Segmentation of RAMS, Compared to the Corresponding Manually Segmented Masks. For each Case, Only the Target Image Was Manually Segmented and Used in Evaluation. The Last Line Lists the Maximal Node MV Gradient of the Final Level. Case 1.1 Is Excluded from the Statistics of Mean and Standard Error of DSC Because RAMS Segmentation Was Not Applicable Due to Small Motion Discontinuity.

Case #	1.1[a]	1.2	1.3	1	2	3	4	5	6	7	8	9	10	Mean ± SE
Image only segmentation	66.3	66.3	66.3	66.3	68.8	69.3	66.7	63.2	69.1	73.2	60.3	67.5	66.5	67.1 ± 1.1
RAMS coarse segmentation	83.7	94.6	92.3	90.5	93.8	91.8	92.4	92.0	94.2	94.4	93.8	91.8	92.1	92.8 ±0.4
RAMS fine segmentation	83.3	96.9	96.9	96.3	97.7	96.0	96.6	95.5	97.3	96.5	96.1	96.1	96.8	96.5 ± 0.2
$\max_{(p,q)\in\mathcal{N}} G_{\mathrm{m}}(p,q)$	3.7	4.0	4.4	5.4	5.6	6.2	4.6	5.2	6.0	4.7	4.6	5.5	6.1	—

[a] Case 1.1 is not included in the statistics of mean and standard error.

TABLE II

Average Bidirectional Local Distance (BLD) between the Boundary Contours of GCBAC Segmentation and the Corresponding Reference Contours of Manual Segmentation, Given in Millimeters. For each Case, Only the Target Image Was Manually Segmented and Used in Evaluation. Case 1.1 Is Excluded from the Statistics Because RAMS Segmentation Was Not Applicable Due to Small Motion Discontinuity.

Case #	1.1[a]	1.2	1.3	1	2	3	4	5	6	7	8	9	10	Mean ± SE
Image only segmentation	40.0	40.0	40.0	40.0	40.2	39.3	40.4	35.6	35.3	30.4	39.7	31.7	32.7	36.5 ± 1.2
RAMS coarse segmentation	17.0	3.4	4.7	5.8	4.2	5.7	4.9	4.0	3.8	3.5	3.2	4.1	3.9	4.3 ±0.3
RAMS fine segmentation	19.7	2.1	2.2	2.5	1.7	3.1	2.5	2.5	1.9	2.5	2.3	2.4	1.9	2.3 ±0.1

[a]Case 1.1 is not included in the statistics of mean and standard error.

TABLE III

Comparison of Mutual Information (MI) between the Transformed Source Images and the Corresponding Target Images in 13 Cases, Given in Bits. For Each Registration Approach, the Same Parameter Setting Were Used in the Two Modes, i.e. without Mask and Using Mask. The Incremental MI between the Two Modes Is Calculated for Each Approach and Each Case Separately.

Case #	1.1	1.2	1.3	1	2	3	4	5	6	7	8	9	10
I_0 (w/o registration)	1.56	1.46	1.41	1.30	1.33	1.39	1.56	1.39	1.51	1.61	1.59	1.38	1.32
I_{r1} (DEEDS-SSC-L1)	1.94	1.80	1.71	1.55	1.60	1.59	1.80	1.68	1.75	1.84	1.93	1.64	1.59
I_{f2} (RAMS)	E51	1.86	1.77	1.62	1.69	1.67	1.87	1.73	1.80	1.91	2.00	1.71	1.66
$\Delta I_r = I_{f2} - I_{r1}$	–0.43	0.06	0.06	0.07	0.09	0.08	0.07	0.05	0.05	0.07	0.07	0.07	0.07
I_{e1} (Elastix w/o mask)	1.81	1.61	1.42	1.20	1.12	1.18	1.53	1.56	1.45	1.52	1.77	1.58	1.47
I_{e2} (Elastix using RAMS mask)	1.94	1.86	1.60	1.47	1.57	1.56	1.67	1.67	1.80	1.84	1.92	1.71	1.59
$\Delta I_e = I_{e2} - I_{e1}$	0.23	0.25	0.18	0.27	0.45	0.38	0.14	0.11	0.35	0.32	0.15	0.13	0.12

TABLE IV

Standard Deviation of Jacobians (Std(J)) of each Forward Motion Field Output from RAMS Registration. Using the Fine Segmentation Derived by RAMS, the Target Image of each Case Is Partitioned into Three Regions: the Inner Region, the Sliding Interface, and The Outer Region. Std(J) Is Calculated Separately for Three Regions in each Case.

Case#	1.2	1.3	1	2	3	4	5	6	7	8	9	10	Mean ± SE
Inner region	0.26	0.31	0.38	0.37	0.49	0.40	0.37	0.44	0.37	0.35	0.44	0.54	0.39 ±0.02
Sliding interface	2.52	3.51	5.96	5.37	9.61	4.38	5.64	5.89	4.90	3.03	5.74	5.94	5.21 ±0.53
Outer region	0.13	0.15	0.18	0.17	0.22	0.23	0.22	0.20	0.18	0.19	0.18	0.18	0.19 ± 0.01

TABLE V

Comparison of Total Registration Error (TRE) Evaluated on the MIDAS Liver CT Dataset. All Values Are in Millimeters. A Statistically Significant Lower TRE ($p < 0.05$ Using Wilcoxon Rank Sum Test) for Both the Lung Landmarks and the Abdominal Landmarks Was Obtained by RAMS when Compared to State-of-the-art Results Reported in the Literature (Anisotropic Diffusion [17], SLIC Demons [19] and XFFD [21]).

Landmark Location		**Lung Landmarks**					**Abdominal Landmarks**				
Patient ID		P0	P1	P2	P4	All	P0	P1	P2	P4	All
Landmark Count		52	55	56	56	219	21	22	21	20	84
w/o Registration	Mean	10.89	6.83	5.11	5.61	7.04	9.08	5.89	6.30	4.42	6.44
	Stdev	3.81	2.57	3.01	4.82	4.27	2.89	3.15	2.76	3.30	3.42
Anisotropic Diffusion	Mean	2.89	1.64	2.12	1.99	2.15	2.27	2.38	2.79	2.80	2.56
	Stdev	2.03	0.66	1.01	1.39	1.42	1.23	1.55	1.75	1.94	1.62
SLIC Demons	Mean	—[a]	—	—	—	2.08	—	—	—	—	2.19
	Stdev	—	—	—	—	—	—	—	—	—	—
XFFD	Mean	—	—	—	—	—	**1.54**	1.56	2.25	2.41	1.94
	Stdev	—	—	—	—	—	0.87	0.93	1.28	0.98	1.01
RAMS (α=8)	Mean	**1.46**	**1.41**	**1.45**	**1.40**	**1.43**	**1.54**	**1.16**	**1.73**	**1.86**	**1.56**
	Stdev	0.63	0.54	0.66	0.69	0.63	0.88	0.53	0.88	1.20	0.92

[a] Not reported