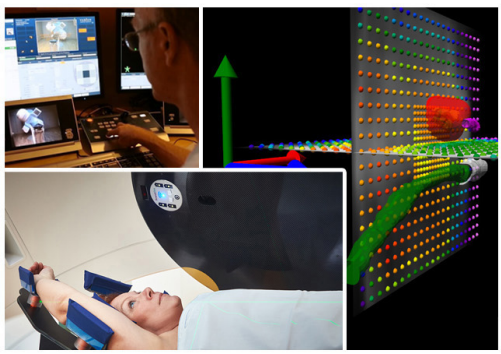
Medical Physics
The International Journal of Medical Physics Research and Practice

3D-SIFT-Flow for atlas-based CT liver image segmentation

Yan Xu, Chenchao Xu, Xiao Kuang, Hongkai Wang, Eric I-Chao Chang, Weimin Huang, and Yubo Fan

Citation: Medical Physics **43**, 2229 (2016); doi: 10.1118/1.4945021
View online: http://dx.doi.org/10.1118/1.4945021
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/43/5?ver=pdfcov
Published by the American Association of Physicists in Medicine



3D-SIFT-Flow for atlas-based CT liver image segmentation

Yan Xu[a)]
State Key Laboratory of Software Development Environment and Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education, Beihang University, Beijing 100191, China and Research Institute of Beihang University in Shenzhen and Microsoft Research, Beijing 100080, China

Chenchao Xu[b)] and Xiao Kuang[c)]
School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China

Hongkai Wang[d)]
Department of Biomedical Engineering, Dalian University of Technology, Dalian 116024, China

Eric I-Chao Chang[e)]
Microsoft Research, Beijing 100080, China

Weimin Huang[f)]
Institute for Infocomm Research (I2R), Singapore 138632

Yubo Fan[g)]
Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education, Beihang University, Beijing 100191, China

(Received 6 April 2015; revised 17 January 2016; accepted for publication 16 March 2016; published 6 April 2016)

Purpose: In this paper, the authors proposed a new 3D registration algorithm, 3D-scale invariant feature transform (SIFT)-Flow, for multiatlas-based liver segmentation in computed tomography (CT) images.
Methods: In the registration work, the authors developed a new registration method that takes advantage of dense correspondence using the informative and robust SIFT feature. The authors computed the dense SIFT features for the source image and the target image and designed an objective function to obtain the correspondence between these two images. Labeling of the source image was then mapped to the target image according to the former correspondence, resulting in accurate segmentation. In the fusion work, the 2D-based nonparametric label transfer method was extended to 3D for fusing the registered 3D atlases.
Results: Compared with existing registration algorithms, 3D-SIFT-Flow has its particular advantage in matching anatomical structures (such as the liver) that observe large variation/deformation. The authors observed consistent improvement over widely adopted state-of-the-art registration methods such as ELASTIX, ANTS, and multiatlas fusion methods such as joint label fusion. Experimental results of liver segmentation on the MICCAI 2007 Grand Challenge are encouraging, e.g., Dice overlap ratio 96.27% ± 0.96% by our method compared with the previous state-of-the-art result of 94.90% ± 2.86%.
Conclusions: Experimental results show that 3D-SIFT-Flow is robust for segmenting the liver from CT images, which has large tissue deformation and blurry boundary, and 3D label transfer is effective and efficient for improving the registration accuracy. [http://dx.doi.org/10.1118/1.4945021]

Key words: SIFT-flow, label transfer, registration, multiatlas, segmentation

1. INTRODUCTION

According to International Agency for Research on Cancer of World Health Organization,[1] liver cancer was estimated to cause nearly 746 000 deaths worldwide (83% in less developed regions) in 2012, making it the second most common cause of death from cancer. It is in general difficulty for doctors to prognose liver cancer (overall ratio of mortality to incidence of 0.95). computed tomography (CT) imaging combined with computer-aided diagnosis (CAD) technologies[2] conveniently provides assistance to prognosis of liver cancer.

The task of liver segmentation from 3D CT images is challenging due to the large deformation of the liver and lack of clear texture patterns, as illustrated in Fig. 1. In addition to performing manual delineations, that is, time-consuming and also results in large variation from expert-to-expert, semiautomatic and fully automatic segmentation algorithms become increasingly popular in the field.[3]

Existing literature in medical image segmentation can be roughly divided into two broad categories: (1) parametric approaches that use model-based methods[4,5] and (2) nonparametric algorithms that use atlas-based registration methods.[6,7]

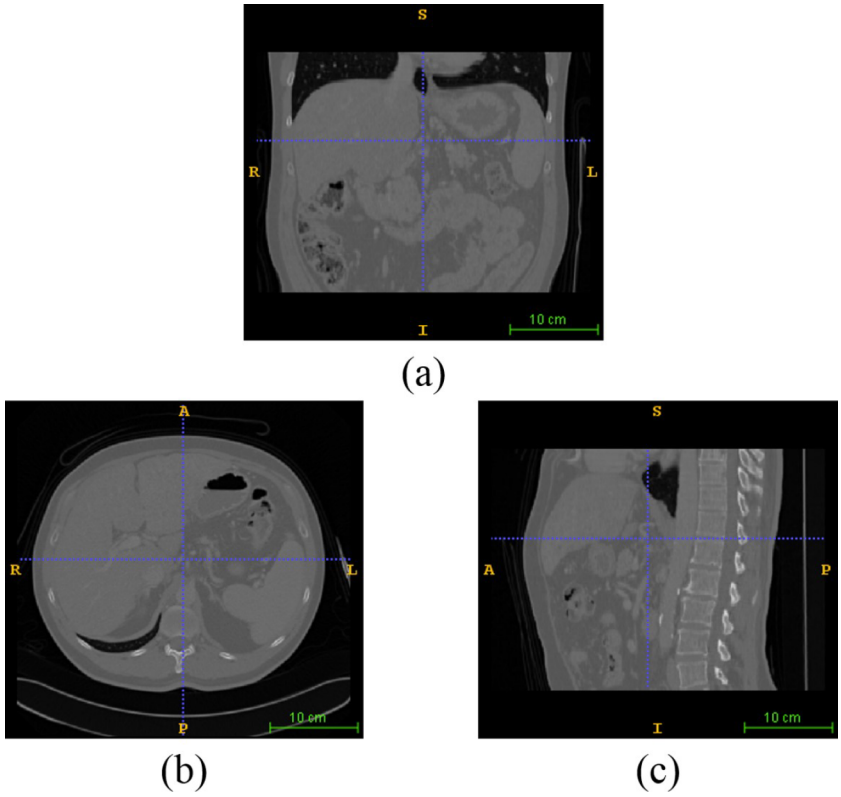


FIG. 1. A clinical CT image. (a)–(c) show livers in different planes. (a) Transverse plane; (b) sagittal plane; (c) coronal plane. The liver in CT image has blurry margin and low gray contrast to the adjacent organs.

In this paper, we study automatic segmentation of liver in 3D CT images by focusing on atlas-based approaches in which an existing atlas is registered to a target (input) image. Ideally, if two images can be perfectly registered, then labeling of the existing atlas can be directly mapped to the input image, resulting in accurate segmentation of the input image. Due to the large variation in individual anatomical structures, it is hard to achieve perfect alignment. Therefore, multiatlas-based segmentation (MAS) methods, which take advantage of having multiple atlases associated with manually delineated labels, have received growing interests.[8] The advantage of multiatlas-based methods over single-atlas-based approaches is evident[9,10] as using multiple atlases more effectively captures the large intrinsic structural variation in CT images. In MAS, each atlas (source) image is registered to a target (input) image; based on the registration (mapping) function, the annotated label map corresponding to the atlas is transformed into the input image spaces; the resulting transformed label maps are then combined to form the segmentation for the input image, typically through a fusion mechanism. Therefore, registration and fusion are the two key steps that critically determine the quality of multiatlas-based segmentation results.[7,11]

Inspired by the success of 2D-scale invariant feature transform (SIFT)-Flow approach in performing large-scale image matching[12–14] and guiding active contours for tracking cell junctions[15] in the biomedical field, we propose a new multiatlas liver segmentation algorithm that extends 2D-SIFT-Flow[12] and the label transfer method[16] to deal with 3D CT images. We call our method, 3D-SIFT-Flow and 3D label transfer, respectively, for registration and label fusion. The SIFT descriptor[17] has been considered as a milestone achievement in computer vision and has been widely adopted in a variety of applications such as object recognition,[18] point tracking,[19,20] and panorama creation.[21] SIFT descriptor is robust, to a certain degree, to local deformation, orientation, scaling, and illumination. SIFT has also been adopted in medical imaging.[22–24] At the same time, optical flow[25] has already successfully been used in medical imaging domain due to its fast computation and precision in mapping structures of interest. Many models based on 2D as well as 3D optical flow methods have been developed and available for medical image registration applications,[26,27] but optical flow is only based on intensity. SIFT Flow[12] combining the SIFT descriptor with optical flow has particular advantage in dealing with images of large difference like the liver in CT images, whereas standard optical flow methods would only be able to cope with images of relatively small variation.[26,27]

2. RELATED WORK

2.A. Liver segmentation

Liver segmentation in CT images is an important task in medical imaging. Accurate segmentation of liver provides liver volume measurements and assists visualization as well as CAD. However, 3D CT liver segmentation remains a challenging task due to existence of large anatomical

deformation and low intensity contrast. In addition, blurry texture of anatomical structures and the background makes the segmentation task difficult. In the past, various techniques for liver segmentation have been proposed.[28] A comprehensive survey on liver CT image segmentation methods is given by Mharib *et al.*,[3] reviewing semiautomatic and automatic liver segmentation, experimenting on several methods (i.e., gray level based techniques, learning techniques, model fitting techniques and probabilistic atlases). As noted by Mharib,[3] the liver segmentation task however still relies greatly on heavy manual intervention and heavy parameter tuning.

A CT liver segmentation challenge was organized for MICCAI 2007 (Ref. 29) in which ten automatic and six interactive methods were proposed, such as region-growing based,[30] shape-constrained segmentation methods[31,32] and gradient vector flow snake based.[33] Most of these methods performed efficiently and robustly in experiments but several problems remained to be improved later (i.e., poor results in real-world clinical images and computational complexity). Methods have been continuously proposed thereafter and results being improved gradually. An automatic method based on a statistical shape mode and an optimal-surface-detection strategy[34] achieved advanced performance,but this method failed in some long and narrow regions of liver. Linguraru *et al.*[35] combined liver with spleen and segmented them using normalized probabilistic atlases and enhancement estimation. Linguraru *et al.*[36] also proposed a novel 3D affine invariant shape parameterization to compare local shape across organs. Danciu *et al.*[37] presented a 3D DCT supervised segmentation method to detect liver volumes. A modified *k*-means segmentation method with a special localized contouring algorithm by Goryawala *et al.*[38] used leveraging of multicore platform to reduce the processing time to one hour, which is still not computationally feasible in clinical applications. Yuan *et al.*[39] proposed geodesic curves to track fuzzy borders between liver and chest wall to deal with the similar HU values on noncontrast-enhanced planning CT images. Cheng *et al.*[40] used random walker algorithm in 3D liver segmentation and model reconstruction, but it largely depends on the labels given by users. Platero and Tobar[41] developed a novel system by combining the spatial normalization with the segmentation method based on standard CRF models for liver segmentation. It used 19 atlases and achieved high Dice coefficient (95%–97%) and a final score of 76 based on the dataset from the MICCAI 2007 Grand Challenge.

2.B. Multiatlas-based segmentation

Good results of MAS depend on excellent registration methods and effective label fusion methods.

Recently, registration techniques in medical imaging start to reach their clinical application and several novel methods have emerged.[42,43] ELASTIX (Ref. 44) is a popular program for registration of medical data. It consists of various transform methods (rigid, affine, nonrigid), similarity measures (mutual information), optimization methods, interpolation methods, and multiresolution schemes. For nonrigid registration, it mainly measures similarity according to image intensity subject to linear or B-spline warping. Despite its popularity, ELASTIX still faces challenge in dealing with large anatomical deformation in clinical CT images. The ANTS system[45] offers open source for large deformation image registration. It was top ranked in a comprehensive brain MRI image registration study[43] and has been widely used in other modalities as well. The ANTS toolkit provides deformable mappings (affine, elastic, diffeomorphic) to find large sharp differences and compute the best correspondence between two medical images. Among these transformation models, symmetric diffeomorphic transformation based on optimizing and integrating a time-varying velocity field is the most flexible and complex model for large deformable registration. However, since the similarity metric in ANTS is still based on intensities of individual pixels/voxels, ANTS also potentially has problems dealing with large structural deformation. HAMMER (Ref. 46) uses several specifically designed features in measuring the similarity, but it has a very specific problem setting that prevents it from being used in the general situations other than brain data.

In multiatlas-based segmentation, effectively fusing the labels after registration is also an important step. In some early attempts, an atlas selection procedure[47] and weight voting[48] were proposed.[8] Extension to weight voting by using global voting and local and semilocal weighting schemes were proposed. Recently, various methods were proposed (i.e., using STAPLE algorithm,[49] based on off-line learning,[50] nonlocal statistical label fusion[51]). Among these methods, a joint label fusion[7] achieved the best performance in the recent MICCAI 2012 Workshop on Multi-Atlas Labeling for the task of hippocampus segmentation, considered as the state-of-the-art method. We call it Multiatlas-based Joint Label Fusion (MAJLF). It combines corrective learning and the nonlocal STAPLE algorithm[49] to handle the limitation of many approaches that different atlases may produce similar label errors when the weights are calculated dependently for each atlas. Due to its good performance, our label fusion will be compared with MAJLF (Ref. 7) in the experiment section.

3. MATERIALS AND METHODS

In this section, we propose the general registration method based on the 3D-SIFT-Flow in detail. In addition, a 3D method based on nonparametric label transfer[16] is implemented for the task.

3.A. 3D SIFT

SIFT is a local descriptor characterizing the local gradient distribution in the image. It was first published by Lowe[18] for feature extraction and object detection. This descriptor is invariant to translation, rotation, scaling (i.e., affine transformation), and local intensity scale of the detected object. Therefore, it is widely used in computer vision for object detection, video tracking, imaging stitching, gesture recognition, and 3D modeling.

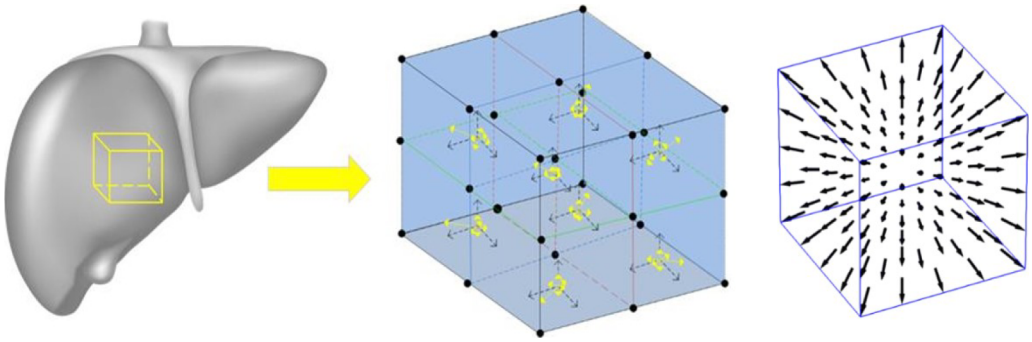


FIG. 2. A 3D SIFT descriptor and a 3D-SIFT-flow illustration. The first figure shows a voxel in liver. The middle one presents how 3D SIFT descriptor is computed and the third one illustrates 3D-SIFT-flow.

In this paper, we use 3D SIFT features to characterize the nonrigid soft tissue deformation in 3D CT images, which can be approximated as a combination of local affine transformations at each pixel. As shown in Fig. 2, to calculate 3D SIFT features, for each pixel $\mathbf{p}$, its local neighborhood is extracted as a cubic region R centered at $\mathbf{p}$. The size of R is determined by the intensity scale of $\mathbf{p}$, and the orientation of R is rotated to the intensity gradient of $\mathbf{p}$. R is further divided into several subblocks ($2\times2\times2$ subblocks in our case), and in each subblock a histogram of gradient magnitudes of six directions is calculated. As a result, the 3D SIFT for $\mathbf{p}$ is a real vector of $2\times2\times2\times6=48$ dimensions.

3.B. General registration method based on 3D-SIFT-flow

We attempt to calculate correspondence between two medical images, i.e., a source image and a target image, and use the correspondence computed from the registration to warp labeling of the source image to the target image. In this paper, we propose a 3D-SIFT-Flow algorithm intended for high-quality 3D medical image registration. The basic idea is to present flow-based registration[25] based on 3D SIFT features instead of intensity similarities. The motivation to present a 3D-SIFT-Flow method (show in Fig. 2) for 3D medical image registration comes from the fact that 2D SIFT Flow method[12] performs robust matching across different scene/object appearances in natural scene images and the discontinuity-preserving spatial model allows matching of objects located at different parts of the scene by describing the image locally using the SIFT descriptor.[18] We use the dense SIFT descriptor to characterize local image structures and encode contextual information. The image with a per-pixel SIFT descriptor is called a SIFT image. We designed an objective function similar to that of 2D-SIFT-Flow[12] to estimate correspondence between two SIFT images. The registration process is driven by minimizing the following energy function:

$$E(\mathbf{w}) = \sum_{\mathbf{p}} \min(\|s_1(\mathbf{p}) - s_2(p+\mathbf{w}(\mathbf{p}))\|_1, t) \quad (1)$$

$$+ \sum_{\mathbf{p}} \eta(|u(\mathbf{p}) + v(\mathbf{p}) + r(\mathbf{p})|) \quad (2)$$

$$+ \sum_{(\mathbf{p},\mathbf{q})\in\epsilon} \min(\alpha|u(\mathbf{p}) - u(\mathbf{q})|, d)$$
$$+ \min(\alpha|v(\mathbf{p}) - v(\mathbf{q})|, d)$$
$$+ \min(\alpha|r(\mathbf{p}) - r(\mathbf{q})|, d), \quad (3)$$

where s_1 and s_2 are the SIFT vectors of two images, $\mathbf{w}(\mathbf{p}) = (u(\mathbf{p}), v(\mathbf{p}), r(\mathbf{p}))$ is the flow vector at pixel location $\mathbf{p} = (x, y, z)$, and ϵ contains all the spatial neighborhoods (a six-neighbor system is used). Parameters η and α are used to maintain the desired balance between the terms. The first term [Eq. (1)] accounts for the dissimilarity of SIFT images, the second term [Eq. (2)] provides a regularization on the first order magnitude of $\mathbf{w}$, and the third term [Eq. (3)] constrains the flow vectors of adjacent pixels to be similar. We use thresholds t and d to account for matching outliers and flow discontinuities, allowing large deformation of soft tissues and discontinuous displacements between adjacent tissues. Due to the large deformation of livers, the weight of Eqs. (2) and (3) should be heavier than Eq. (1). The key to making the energy function more robust for liver segmentation is to decrease the value of α and η.

To optimize the objective function, we use a dual-layer loopy belief propagation (BP-S) method similar to the one used in 2D-SIFT-Flow.[12] Three dimensions are separated by setting up three layers with exactly the same grid. In message passing, we first update the intralayer message in three layers respectively, and then update the interlayer message between them. Different from 2D-SIFT-Flow, when implementing interlayer message updating, we transfer messages in a plane to the third dimension. For example, minimize the messages in the $x-y$ plane and transfer it to z. To speed up the optimization, we use a coarse-to-fine matching scheme similar to the one used in 2D-SIFT-Flow.[12]

3.C. 3D label transfer method

A 3D method based on nonparametric label transfer[16] is proposed to merge the labels of registered atlases based on SIFT features. To facilitate the description, we use following notation: the target image is I and the number of atlases is M. For the ith atlas, we have following information: $\{s_i, c_i, \mathbf{w}_i\}_{i=1:M}$, i.e., the SIFT of original image: s_i, annotation c_i, and wrap mapping function w_i. Our goal is to obtain c, which is the annotation of the target image by the former

information. An energy function will be formulated to integrate the following consideration: position, prior probability, and neighbor relationship. The energy function is similar to 2D Label Transfer.[16] We expand it to three dimensions as following:

$$F = \sum_{\mathbf{p}} \psi(c(\mathbf{p}); s, s_i') + \alpha \sum_{\mathbf{p}} \lambda(c(\mathbf{p})) + \beta \sum_{\substack{\{\mathbf{p}_1,\mathbf{p}_2\}\in\epsilon \\ c(\mathbf{p}_1)\neq c(\mathbf{p}_2)}} \phi(c(\mathbf{p}_1), c(\mathbf{p}_2); I) + \log Z, \tag{4}$$

where Z is the normalization constant.

For the likelihood term ψ, we define

$$\psi = \min_{c(\mathbf{p})=c_i(\mathbf{p})} \{s(\mathbf{p}) - s_i(\mathbf{p}+\mathbf{w}(\mathbf{p}))\}. \tag{5}$$

Note that if we do not have $c(\mathbf{p}) = c_i(\mathbf{p})$ for any i, then ψ will be the maximal value of the difference measure function. For the prior term λ, it will be voted by labels of candidates,

$$\lambda = -\log\left(\frac{|\{i|c(\mathbf{p}) = c_i(\mathbf{p})\}| + 1}{M}\right). \tag{6}$$

We put a negative sign here because the better our solution is, the less F will be. In smooth term ϕ, we use this function

$$\phi = \frac{\xi + e^{-\gamma\|I_{(\mathbf{p}_1)} - I_{(\mathbf{p}_2)}\|^2}}{\xi + 1}, \tag{7}$$

where $\gamma = (2\langle\|I(\mathbf{p}) - I(\mathbf{q})\|^2\rangle)^{-1}$.

This function obtains a maximal value 1 when two pixels are same and converges to $\xi/(\xi+1)$ when two pixels are very different. Note that before this calculation, we use a 3D Gauss filter with a radius of 3 to smooth the target image. To optimize the objective function with respect to the annotations $c(\mathbf{p})$, we still use the BP-S algorithm as mentioned in this section.

3.D. Data

We used CT abdomen images with the liver ground-truth annotation from the MICCAI 2007 Grand Challenge.[52] All images have been acquired in transversal direction and the pixel spacing varies from 0.55 to 0.8 mm and the interslice distance varies from 1 to 3 mm. These images are stored in Metaformat: an ASCII readable header (.mhd) and a separate raw image data file (.raw). ITK-SNAP (Ref. 53) was used to visualize the data in the paper. The dataset contains 40 images (20 training images, 10 testing images for the qualifying, and 10 for the contest). We used 10 from 20 training images (Numbers 1, 2, 3, 8, 11, 16, 17, 18, 19, 20) and then extracted the region of interest ($128\times128\times128$ voxels) containing the liver for registration and label fusion test.

3.E. Implementation details

Ten CT images were used in our experiment. Each original image was used as a target image and the remaining images were used as source images, respectively. We calculated the correspondence between the source images and the target image via 3D-SIFT-Flow. Labeling of source images was then warped according to the correspondence computed by the former step to obtain candidate segmentations for the target image. Finally, these candidates were fused to one final segmentation image via 3D label transfer.

3.F. Comparison

3.F.1. Registration

Three registration methods were compared in our experiments: ELASTIX, ANTS, 3D-SIFT-Flow. We aligned images by affine registration for ELASTIX and ANTS respectively, but did no initial rigid or affine registration for 3D-SIFT-Flow.

ELASTIX: It uses the 3D B-spline registration based on mutual information. A multiresolution registration scheme is used to accelerate the registration. Four levels of resolution are used. The down-sampling ratios for the four resolution levels are 8, 4, 2, and 1. An adaptive stochastic gradient descent algorithm is used for the optimization at each level. We used default parameter file in ELASTIX.[54]

ANTS: We used symmetric normalization model (nonrigid) transformation. The following parameter settings used in the experiment were proven to be effective to handle large deformation.[45]

```
ANTS 3 -m cc[{fix},{mov},1,4] -o output.nii -r Gauss[3,0]
            -t SyN[0.5] -i 100x100x100x20
  WarpImageMultiTransform 3 {mov} {result} -R {fix}
        outputWarp.nii.gz  outputAffine.txt
```

3D-SIFT-Flow: For every pixel in a 3D medical image, its neighborhoods were divided into a $2\times2\times2$ cell array and the gradient orientations were quantized into a 6-bin histogram in each cell. Thus, we obtained a 48D vector as the SIFT representation for a pixel. The down-sampling ratios for the four resolution levels are 8, 4, 2, and 1 in the coarse-to-fine scheme. We used $\alpha = 0.25$, $\eta = 0.0001$ in the energy function and window size = 7 in the BP-S algorithm.

3.F.2. Fusion

3D SIFT Flow was compared with the state-of-the-art label fusion method, MAJLF.[7] We applied both of them to each registration result.

MAJLF:[7] We used the following parameters: 0.1 as the regularization term added to the matrix for inverse, and 2 as the exponent for mapping intensity difference to joint error.

TABLE I. Dice coefficient of three registration methods (mean ± standard deviation %).

ELASTIX	ANTS	3D-SIFT-Flow
86.29 ± 3.86	90.71 ± 3.18	**92.56 ± 2.01**

3D label transfer: We used the following parameters: $\alpha = 0$, $\beta = 0.02$, $M = 9$ in the experiments.

4. RESULTS

4.A. 3D-SIFT-Flow

In our experiments, Dice coefficient (Mean overlap)[55] was used to evaluate the overlapping ratio between the segmentation result and the ground truth. It is the intersection divided by the mean volume of the two regions,

$$\mathrm{DC} = 2\frac{\sum_r |S_r \cap G_r|}{\sum_r (|S_r| + |G_r|)}, \tag{8}$$

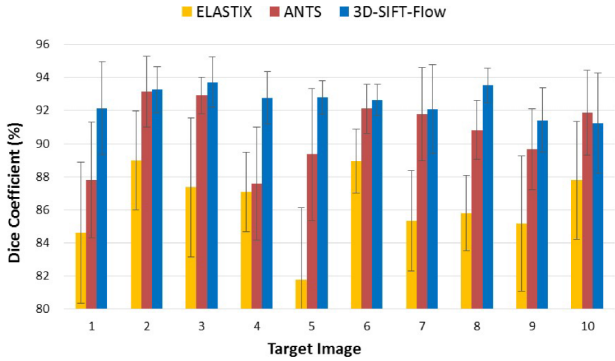


FIG. 3. Comparison of three registration methods. X axis: the number sequence of the target image; Y axis: average Dice coefficient of the results registered by the remaining nine source images (%). The differences between 3D-SIFT-flow and ELASTIX/ANTS are significant (both $p < 0.0001$).

where S_r is the region of the segmentation result, and G_r is the region of the ground truth.

Results of three registration methods are shown in Table I. The Dice coefficient between the manual and automated

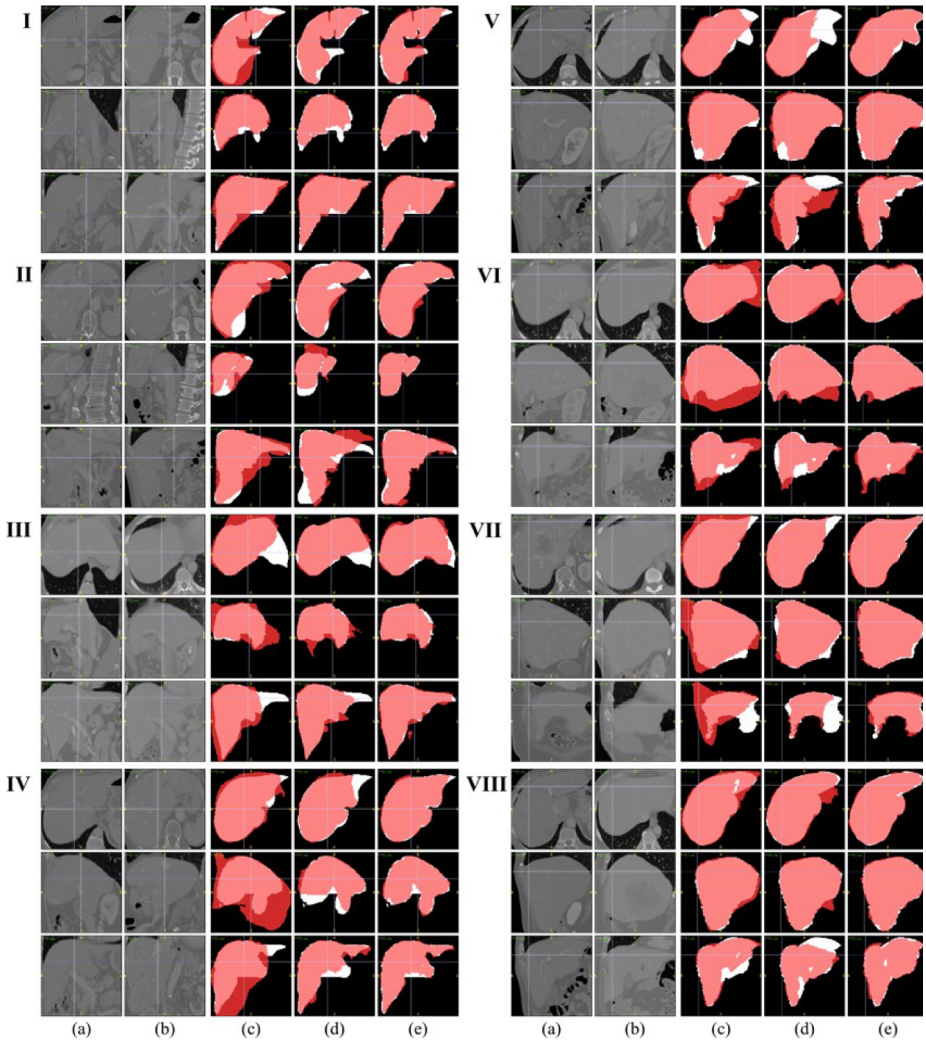


FIG. 4. Comparison of three registration methods with liver annotation: I, II, III, IV, V, VI, VII, VIII are the results of eight CT images. Three rows of the results show liver on different planes (from top to bottom: transverse plane, sagittal plane and coronal plane). (a) The target image; (b) the source image; (c) the result produced by ELASTIX; (d) the result produced by ANTS; (e) the result produced by 3D-SIFT-flow. White: the ground truth; red: the predicted result (color in online version only).

TABLE II. Dice coefficient of the segmentation results produced by two fusion methods (mean ± standard deviation %).

	ELASTIX	ANTS	3D-SIFT-Flow
No fusion[a]	86.29 ± 3.86	90.71 ± 3.18	92.56 ± 2.01
MAJLF (Ref. 7)	94.90 ± 2.86	92.61 ± 1.20	96.25 ± 0.69
3D label transfer	**94.90 ± 1.86**	**95.23 ± 1.49**	**96.27 ± 0.96**

[a]Segmentation performed only by registration.

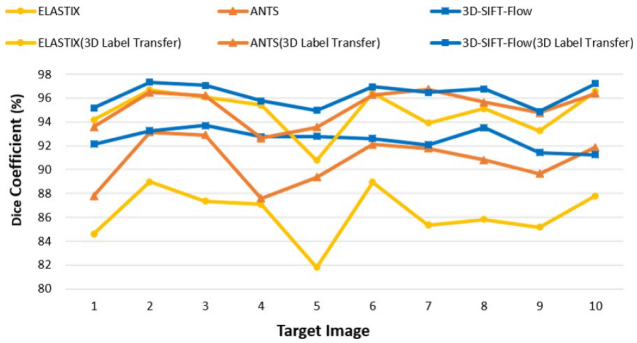


FIG. 5. Dice coefficient of the results produced by three registration methods (ELASTIX, ANTS, and 3D-SIFT-Flow) and further improved by 3D label transfer. Our fusion method increased the Dice coefficient by 4%–8% for three registration methods.

segmentations are 92.56% ± 2.01% by 3D-SIFT-Flow, which verifies the advantage of our registration method, as shown in Fig. 3 and Table I. The highest accuracy and the smallest standard deviation show that 3D-SIFT-Flow performs better and more stable than two state-of-the-art registration methods. Additionally, a *t*-test was conducted between the results of 3D-SIFT-Flow and ANTS/ELASTIX. The *p*-values were both <0.0001, which shows the significant difference between 3D-SIFT-Flow and ELASTIX/ANTS. Additionally, Fig. 3 also indicates that 3D-SIFT-Flow does not always achieves the highest DICE coefficient (e.g., target image 10), ANTS sometimes provide comparable results (e.g., target images 2, 3, 6, 7). We observed two reasons. One is that sharp differences between these target images and template images are weak enough for ANTS to deal with. The other reason is that we reduced the value of parameter α in Eq. (3) which constrains the flow vectors of adjacent pixels to be similar. Such setting can help us detect some blurry margins, at the cost of damaging continuity. Although the discontinuity may cause several errors, this problem can be successfully solved in the label fusion step. For the runtime, 3D-SIFT-Flow was run on a Microsoft Windows HPC cluster (2 quad-core Xeon 2.43-GHz processor, 16 GB RAM), taking about 40 min each). ANTS and ELASTIX took nearly 25 min and 1 min on the some HPC cluster, separately. We will further discuss the runtime of 3D SIFT Flow in Sec. 5.

We show eight typical results in Fig. 4. Compared with the target image, the liver in source image has a different position and large deformation. Take Fig. 4 (VIII) for an example. In transverse plane, the margin of liver in the source image is convex, while it is concave in the target image. The view in coronal plain shows that livers in these two images have large differences in margin and large soft-tissue deformation in the upper right corner. Two state-of-the-art methods failed to find corresponding voxels between two images. However, 3D-SIFT-Flow is capable of compensating for large soft-tissue

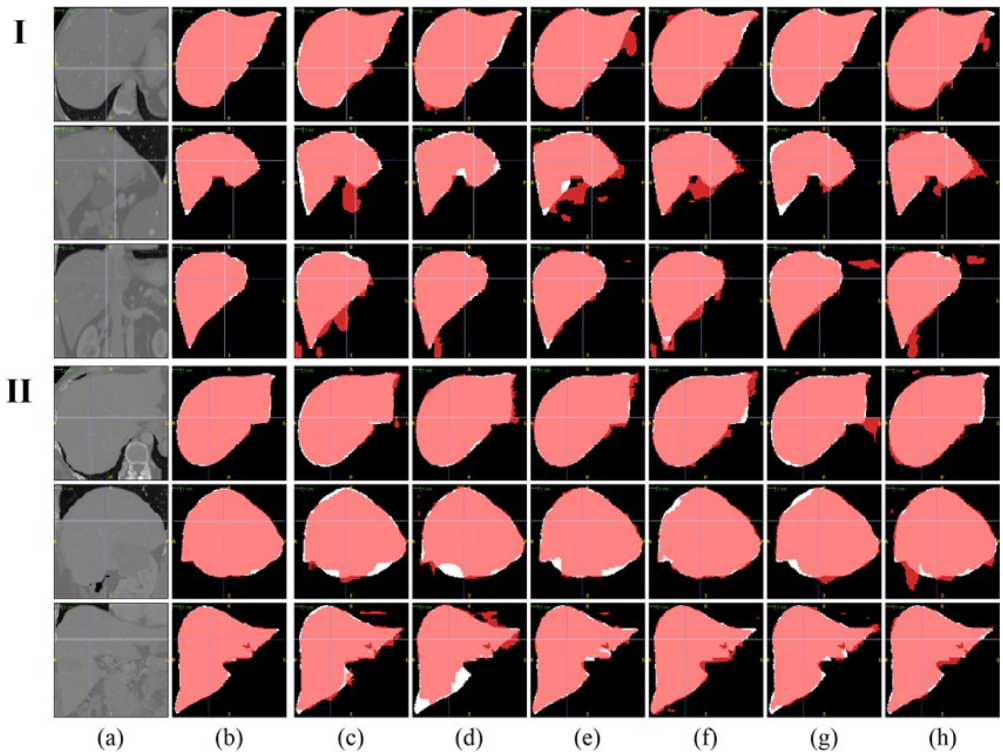


FIG. 6. Some results produced by 3D label transfer and label fusion candidates for the target image. (a) The target image; (b) the label fusion result produced by 3D label transfer; (c)–(h) six candidates registered by 3D-SIFT-flow. White: the ground truth; red: the predicted result (color in online version only).

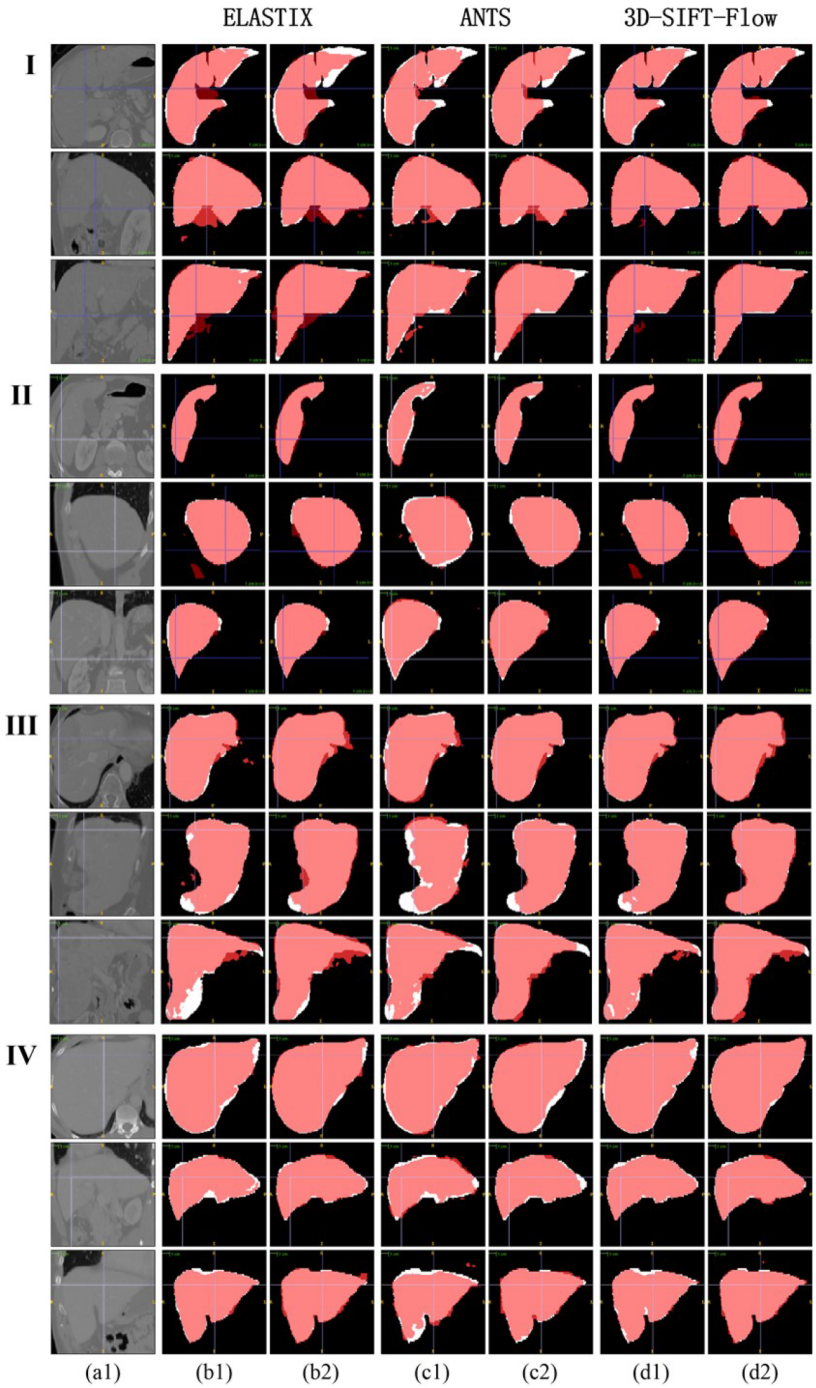


FIG. 7. Comparison of some label fusion results produced by MAJLF (Ref. 7) and 3D label transfer: I, II, III, IV are the results of four CT images. Three rows show liver on different planes (from top to bottom: transverse plane, sagittal plane, and coronal plane). (a) The original image; (b)–(d) the results produced by ELASTIX, ANTS, and 3D-SIFT-flow, respectively [(1) MAJLF (Ref. 7) and (2) 3D label transfer). That is, (d2) are the label fusion results produced by 3D label transfer based on 3D-SIFT-Flow registration method, White: the ground truth; red: the predicted result (color in online version only).

deformation due to the combination of SIFT features and optical flow.

4.B. 3D label transfer

Table II presents the results of 3D label transfer and MAJLF. 3D label transfer further improved Dice coefficient between the manual and automated segmentations to 96.27% ± 0.96% (3D-SIFT-Flow), 95.23% ± 1.49% (ANTS), and 94.90% ± 1.86% (ELASTIX). Our fusion method increased Dice coefficient of the results by 4%–8% for three registration methods. Figure 5 illustrates the Dice coefficient of each group of atlases (before fusion and after fusion). Additionally, Fig. 6 presents the results of 3D label transfer and six candidates based on 3D-SIFT-Flow.

Table II also shows the comparison between 3D label transfer and MAJLF. Two fusion methods both achieved a significant improvement of registration results. The improvements

TABLE III. Average values of 5 measures, scores and the final score for 3D-SIFT-Flow/3D label transfer: volumetric overlap error (m_1), absolute relative volume difference (m_2), average symmetric surface distance (m_3), average symmetric RMS surface distance (m_4), maximum surface distance (m_5).

	m_1 (%)	m_2 (%)	m_3 (mm)	m_4 (mm)	m_5 (mm)	Final score
Measures	7.19 ± 1.78	3.41 ± 2.70	0.83 ± 0.27	1.65 ± 0.57	17.05 ± 5.97	
Scores	71.93	81.89	79.17	77.02	77.57	77.52

of ELASTIX/3D-SIFT-Flow by 3D label transfer and MAJLF were similar. For example, 3D SIFT Flow combined with MAJLF almost yields comparable accuracy with our 3D SIFT Flow/3D label transfer. The reason is that MAJLF sometimes successfully decreases the influence of similar errors between each candidate by combining corrective learning with the nonlocal STAPLE algorithm. However, 3D label transfer increased Dice coefficient by 2.5% in ANTS, which was higher than MAJLF. In Fig. 7, we show four examples. The number sequence of the target image is 1, 2, 5, and 10. It shows the label fusion results by 3D label transfer and MAJLF on three different planes (transverse plane, sagittal plane, and coronal plane). White patterns are the ground truth and red patterns are the predicted result. To evaluate 3D-SIFT-Flow/3D label transfer with other methods in the MICCAI 2007 Grand Challenge, we measure results in standard as the overlap error, relative absolute surface difference, average symmetric distance, root mean square symmetric distance, and maximum symmetric distance, as well as the final score.[56] The result of each measure is transformed to a score (0–100), compared with a standard value which is calculated by averaging the manual segmentations. The final score is the average of scores of 5 different measures. A final score of 75 for liver segmentation can loosely indicate that the proposed method is comparable to human performance. In the experiment, 3D-SIFT-Flow/3D label transfer achieved a score of 77, as shown in Table III. High Dice coefficient and high scores of 5 different measures illustrate the efficiency and effectiveness of our multiatlas-based segmentation system.

For computational complexity, we ran 3D label transfer and MAJLF on the same computer configuration (Intel core i7-4770 CPU 3.40 GHz) and used 9 atlases. The time consumption of two fusion method is shown in Table IV. 3D label transfer cost 5.67 ± 0.10 min, which is more than 5 times less than MAJLF. Moreover, we ran both fusion methods on different numbers of atlas (as shown in Fig. 8). It presents that the speed of MAJLF greatly depends on the number of the atlases. If the number becomes larger, MAJLF requires more time because of the additional step of solving the inverse of the pairwise dependence matrix in its algorithm.[7] Compared with MAJLF, 3D label transfer performs stably and robustly when the number of atlases increases.

TABLE IV. The running time of two fusion methods (min).

	MAJLF (Ref. 7)	3D label transfer
Time	32.36 ± 1.58	**5.67 ± 0.10**

5. DISCUSSION

Liver segmentation using CT data is useful in biomedical applications, but it is still a challenge due to large anatomical deformation, low intensity contrast and blurry margin. In this study, we have successfully proposed 3D-SIFT-Flow to the medical image context and it has proven to effectively handle large anatomical deformation of the liver. The results show that our approach works better for liver CT registration than two state-of-the-art software packages, i.e., ANTS and ELASTIX. The t-test showed significant difference between 3D-SIFT-Flow and ELASTIX/ANTS (both $p < 0.0001$). Among three methods, 3D-SIFT-Flow achieved the highest Dice coefficient and the best visual matching quality (as shown in Fig. 4). Eight examples of the liver from different planes illustrate that our approach successfully finds the voxel-to-voxel correspondence between two CT images although livers have big differences in position, margin, and gray level. The good performance of 3D-SIFT-Flow should be contributed to the adoption of SIFT feature and the usage of optical flow optimization.

As proven by a large number of existing studies, SIFT is good at describing salient image features such as corners and edges. Therefore, our method performs well at matching boundaries and sharp corners. In contrast, ANTS and ELASTIX fail to deal with complex organ boundaries, especially the concave and convex edges. ELASTIX uses mutual information, which is an intensity-based similarity measurement without special consideration of edge features. ANTS uses histogram matching as the similarity metric. It works well for images with several intensity clusters (i.e., the brain MR images), but is not as effective for liver CT images with less contrast. On the other hand, optical flow allows us to capture large intersubject deformation because the mathematical definition of optical flow is more flexible

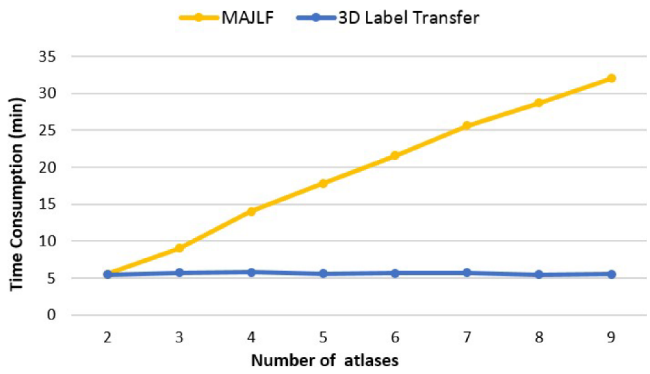


FIG. 8. The figure shows time consumption of two fusion methods. It compares 3D label transfer with MAJLF (Ref. 7) when the number of atlases increases.

than the symmetric normalization model and the B-spline deformation used in ANTS and ELASTIX, respectively.

During registration, several different parameters were used in 3D-SIFT-Flow. We attempted to detect some greatly blurry patterns of liver, so we set several parameters of the Gauss filter in the feature extraction part. However, our registration method achieved similar results when the image preprocessing was changed. Regarding the energy function of SIFT Flow [Eqs. (1)–(3)], α controls the weight of the smoothness term [Eq. (3)] which constrains the similarity of adjacent pixels. We set $\alpha = 1$ at first and then reduced the value of α to 0.25 (used in this paper). Both parameter settings were used on 90 registration experiments. Average of Dice coefficient of registration results increased by nearly 1.0%. Unlike that of natural scene images, the margin of liver is more changeable and deformed [Figs. 4(I-a) and 4(V-c)], so adjacent voxels may not be so similar. However, if we further reduce the value of α, the continuity of 3D-SIFT-Flow will be damaged. In terms of discontinuity, although the segmentation results of $\alpha = 1$ look more continuous than $\alpha = 0.25$, the little destruction to segmentation is acceptable and some blurry margins can be successfully detected. The small displacement term in Eq. (2) constrains the flow vectors to be as small as possible. This term must be heavily weighted due to the large deformation and big offset between the source image and the target image. We also reduced the value of η to lower the weight of Eq. (2). In the coarse-to-fine matching scheme, the window size was enhanced in the experiment to get a bigger searching area in order to match some large deformed parts between two images. In addition, 3D-SIFT-Flow uses the truncated L1-norm regularization to achieve good performance and deformation fields computed by the voxel-to-voxel alignment are likely "blocky" and diffeomorphic.

As a cost of better registration accuracy, 3D-SIFT-Flow needs more time than the other two methods (about 40 min for the image of $128\times128\times128$ voxels on a 2 quad-core Xeon 2.43-GHz processor, 16 GB RAM). Most of the computation is spent on dual-layer loopy belief propagation (BP-S) and the large searching area. Nevertheless, with the progress in computing power, it can be expected that the computation burden will be greatly alleviated in the near future. Moreover, since the BP-S algorithm can be parallelized, we will also focus on parallel computing to speed up 3D-SIFT-Flow. Another drawback of 3D-SIFT-Flow is its comparatively large memory requirement of the dense SIFT feature. To align two typical $512\times512\times256$ CT scans is likely to be $512\times512\times256\times4$ (32 bits float) $\times$ 48 (SIFT features) $\times$ 2 = 24 GB.

For multiatlas fusion, we compared the proposed 3D label transfer method with MAJLF.[7] Compared to the results without atlas fusion, both fusion methods achieved 4%–8% improvements of Dice coefficient for all registration methods (as shown in Fig. 7) and the difference between the results of two fusion methods was not significant. However, considering the running time, 3D label transfer becomes more efficient than MAJLF when the atlas number grows larger (as shown in Fig. 8). The reason is that MAJLF uses an additional step of solving the inverse of the pairwise dependence matrix to its algorithm.

TABLE V. The Dice coefficient of multiatlas segmentation results with 20 training images (mean ± standard deviation %).

Registration method	Label fusion method	DICE coefficient
3D-SIFT-Flow	3D label transfer	**94.99 ± 1.56**
	MAJLF	94.97 ± 2.67
	STAPLE (Ref. 49)	91.56 ± 3.86
ANTS	Major voting	91.64 ± 5.48
	3D label transfer	94.77 ± 1.53
	MAJLF	94.71 ± 2.74
	STAPLE (Ref. 49)	92.78 ± 3.35
ELASTIX	Major voting	92.25 ± 3.49
	3D label transfer	93.01 ± 5.02
	MAJLF	93.88 ± 5.72
	STAPLE (Ref. 49)	86.87 ± 5.23
	Major Voting	89.09 ± 5.58

Furthermore, we ran all 20 training data from MICCAI 2007 Liver Segmentation Grand Challenge (as shown in Table V) and submitted the results of 3D-SIFT-Flow/3D label transfer system on 10 test images.

In the future work, we will investigate extending 3D-SIFT-Flow to other types of medical images, such as magnetic resonance (MR) brain images. For example, our system will be applied to the segmentation of hippocampus (HC) which plays an important role in the diagnosis of Alzheimers disease. Classical registration approaches cannot work well due to the large deformation between abnormal HC and normal HC (Tables VI–VIII). Our method (3D-SIFT-Flow and 3D label transfer), which has been proved to be effective for soft tissue with large deformation in this paper, may perform well in HC segmentation.

6. CONCLUSION

In this paper, we proposed 3D-SIFT-Flow for liver segmentation in CT images and extended the 2D-based nonparametric label transfer method to 3D for fusing the multiple registered atlases. Experimental results showed that 3D-SIFT-Flow is robust for segmenting the liver from CT images, which has large tissue deformation and blurry boundary, and 3D label transfer is effective and efficient for improving the registration accuracy. The binary file of 3D-SIFT-Flow/3D label transfer and samples of CT liver images are given in the supplementary material.[57]

ACKNOWLEDGMENTS

This work was supported by Microsoft Research under eHealth program, Beijing National Science Foundation in China under Grant No. 4152033, Beijing Young Talent Project in China, the Fundamental Research Funds for the Central Universities of China, and Grant No. SKLSDE-2015ZX-27 from the State Key Laboratory of Software Development Environment in Beihang University in China. The authors would like to thank the organizers of MICCAI 2007 Grand Challenge for providing 3D CT liver images.

APPENDIX: RESULTS OF THREE REGISTRATION METHODS

TABLE VI. The Dice coefficient of ELASTIX (mean ± standard deviation %).

Image	1	2	3	4	5	6	7	8	9	10
1	—	89.88	90.84	86.97	72.81	87.40	87.13	86.78	88.15	86.44
2	86.47	—	89.93	90.02	82.24	90.54	86.37	89.70	91.01	90.08
3	89.92	90.69	—	85.60	81.06	89.49	87.26	82.74	86.90	84.18
4	84.51	88.21	85.29	—	79.03	89.56	80.85	85.75	84.08	83.15
5	75.91	82.32	79.03	81.76	—	84.43	80.70	84.37	78.04	83.14
6	83.85	89.32	88.24	88.44	85.31	—	84.10	88.10	85.34	89.67
7	86.33	90.71	92.27	87.17	84.93	90.05	—	85.98	88.69	91.80
8	84.88	89.84	85.15	87.51	87.97	90.40	84.46	—	80.43	91.10
9	88.92	92.93	90.78	89.27	80.88	89.22	88.80	85.95	—	90.55
10	80.73	87.02	84.69	87.05	81.88	89.40	88.41	82.92	83.81	—
Mean	84.61	88.99	87.36	87.09	81.79	88.94	85.34	85.81	85.16	87.79
Standard deviation	4.25	3.00	4.20	2.40	4.33	1.92	3.03	2.27	4.10	3.56

TABLE VII. The Dice coefficient of ANTS (mean ± standard deviation %).

Image	1	2	3	4	5	6	7	8	9	10
1	—	93.03	92.76	89.78	82.19	91.21	89.46	90.34	91.19	89.70
2	90.71	—	93.84	93.25	87.65	92.51	94.20	92.05	89.83	93.44
3	91.69	95.50	—	89.68	93.24	93.47	94.37	90.80	91.48	93.53
4	87.29	94.27	90.42	—	84.38	89.13	85.90	87.66	86.61	87.04
5	81.16	87.99	93.30	80.93	—	92.28	92.78	90.64	86.16	89.90
6	89.65	92.92	93.40	87.76	92.15	—	90.03	93.06	92.89	94.00
7	88.85	94.95	94.24	85.42	93.61	91.10	—	92.72	92.18	95.05
8	86.71	93.56	92.24	86.68	92.09	91.90	93.13	—	87.82	92.58
9	83.70	92.90	93.08	88.16	89.34	93.64	92.10	88.68	—	91.65
10	90.48	93.20	93.00	86.62	89.53	93.70	94.10	91.46	88.79	—
Mean	87.80	93.15	92.92	87.59	89.35	92.10	91.79	90.82	89.66	91.88
Standard deviation	3.49	2.15	1.10	3.40	3.99	1.48	2.83	1.78	2.45	2.56

TABLE VIII. The Dice coefficient of 3D-SIFT-Flow (mean ± standard deviation %).

Image	1	2	3	4	5	6	7	8	9	10
1	—	93.55	94.14	94.13	91.12	93.44	89.39	93.01	91.36	91.63
2	93.07	—	95.29	93.33	93.51	92.99	94.15	94.72	92.93	90.93
3	94.27	95.01	—	94.00	93.41	93.48	94.04	94.42	92.51	91.70
4	94.15	92.53	91.83	—	91.47	93.02	88.84	91.95	89.59	84.72
5	85.01	90.22	90.53	88.87	—	91.82	87.98	91.98	87.46	89.00
6	91.74	93.61	94.58	92.60	93.42	—	92.61	93.12	92.43	91.34
7	92.23	94.69	94.51	92.27	93.30	92.66	—	93.99	93.24	95.32
8	92.44	93.43	93.78	93.01	93.94	93.26	92.70	—	90.25	93.18
9	93.57	93.29	94.09	93.73	91.88	92.41	93.56	94.33	—	93.33
10	92.76	93.13	94.62	92.86	93.11	90.54	95.56	94.19	92.99	—
Mean	92.14	93.27	93.71	92.75	92.80	92.62	92.09	93.52	91.42	91.24
Standard deviation	2.81	1.37	1.53	1.59	1.02	0.94	2.68	1.05	1.96	3.02

[a]Electronic mail: xuyan04@gmail.com
[b]Electronic mail: chenchaoxu33@gmail.com
[c]Electronic mail: kuangxiao.ace@gmail.com
[d]Electronic mail: wang.hongkai@gmail.com
[e]Electronic mail: eric.chang@microsoft.com
[f]Electronic mail: wmhuang@i2r.a-star.edu.sg
[g]Author to whom correspondence should be addressed. Electronic mail: yubofan@buaa.edu.cn

[1]Cancer of World Health Organization, http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx, 2012.
[2]T. Livraghi, L. Solbiati, M. F. Meloni, G. S. Gazelle, E. F. Halpern, and S. N. Goldberg, "Treatment of focal liver tumors with percutaneous radio-frequency ablation: Complications encountered in a multicenter study 1," Radiology **226**, 441–451 (2003).
[3]A. M. Mharib, A. R. Ramli, S. Mashohor, and R. B. Mahmood, "Survey on liver CT image segmentation methods," Artif. Intell. Rev. **37**, 83–95 (2012).
[4]J. Yang, L. H. Staib, and J. S. Duncan, "Neighbor-constrained segmentation with level set based 3D deformable models," IEEE Trans. Med. Imaging **23**, 940–948 (2004).
[5]Z. Tu, K. Narr, P. Dollar, I. Dinov, P. Thompson, and A. Toga, "Brain anatomical structure segmentation by hybrid discriminative/generative models," IEEE Trans. Med. Imaging **27**, 495–508 (2008).
[6]J. H. Morra, Z. Tu, L. G. Apostolova, A. E. Green, A. W. Toga, and P. M. Thompson, "Comparison of adaboost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation," IEEE Trans. Med. Imaging **29**, 30–43 (2010).
[7]H. Wang, J. W. Suh, S. R. Das, J. B. Pluta, C. Craige, and P. A. Yushkevich, "Multi-atlas segmentation with joint label fusion," IEEE Trans. Pattern Anal. **35**, 611–623 (2013).
[8]J. E. Iglesias and M. R. Sabuncu, "Multi-atlas segmentation of biomedical images: A survey," Medical Image Analysis **24**(1), 205–219 (2015).
[9]Y. Hao, T. Jiang, and Y. Fan, "Iterative multi-atlas based segmentation with multi-channel image registration and Jackknife Context Model," in *9th IEEE International Symposium on Biomedical Imaging (ISBI)* (IEEE, Barcelona, 2012), pp. 900–903.
[10]W. Bai, W. Shi, D. P. O'Regan, T. Tong, H. Wang, S. Jamil-Copley, N. S. Peters, and D. Rueckert, "A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images," IEEE Trans. Med. Imaging **32**, 1302–1315 (2013).
[11]X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de Solórzano, "Combination strategies in multi-atlas image segmentation: Application to brain MR data," IEEE Trans. Med. Imaging **28**, 1266–1277 (2009).
[12]C. Liu, J. Yuen, and A. Torralba, "SIFT flow: Dense correspondence across scenes and its applications," IEEE Trans. Pattern Anal. **33**, 978–994 (2011).
[13]M. Drozdzal, L. Igual, J. Vitria, C. Malagelada, F. Azpiroz, and P. Radeva, "Aligning endoluminal scene sequences in wireless capsule endoscopy," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, San Francisco, CA, 2010), pp. 117–124.
[14]C. Gatta, S. Balocco, V. Martin-Yuste, R. Leta, and P. Radeva, "Non-rigid multi-modal registration of coronary arteries using SIFTflow," in *Pattern Recognition and Image Analysis* (Springer, Las Palmas de Gran Canaria, Spain, 2011), pp. 159–166.
[15]S. Kang, C. Lee, M. Goncalves, A. Chisholm, and P. Cosman, "Tracking epithelial cell junctions in c. elegans embryogenesis with active contours guided by SIFT flow," IEEE Trans. Biomed. Eng. **62**, 1020–1033 (2014).
[16]C. Liu, J. Yuen, and A. Torralba, "Nonparametric scene parsing via label transfer," IEEE Trans. Pattern Anal. **33**, 2368–2382 (2011).
[17]D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vision **60**, 91–110 (2004).
[18]D. G. Lowe, "Object recognition from local scale-invariant features," in *IEEE International Conference on Computer Vision* (IEEE, Orlando, FL, 1999), Vol. 2, pp. 1150–1157.
[19]S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, New York, NY, 2006), Vol. 2, pp. 2169–2178.
[20]S. Se, D. Lowe, and J. Little, "Vision-based mobile robot localization and mapping using scale-invariant features," in *IEEE International Conference on Robotics and Automation* (IEEE, Seoul, 2001), Vol. 2, pp. 2051–2058.
[21]M. Brown and D. G. Lowe, "Recognising panoramas," in *IEEE International Conference on Computer Vision* (IEEE, Nice, 2003), Vol. 3, p. 1218.
[22]W. Cheung and G. Hamarneh, "N-SIFT: N-dimensional scale invariant feature transform for matching medical images," in *4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (IEEE, Arlington, 2007), pp. 720–723.
[23]S. Allaire, J. J. Kim, S. L. Breen, D. A. Jaffray, and V. Pekar, "Full orientation invariance and improved feature selectivity of 3D SIFT with application to medical image analysis," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (IEEE, Anchorage, AK, 2008), pp. 1–8.
[24]P. Scovanner, S. Ali, and M. Shah, "A 3-dimensional SIFT descriptor and its application to action recognition," in *Proceedings of the 15th International Conference on Multimedia* (ACM, New York, NY, 2007), pp. 357–360.
[25]B. K. Horn and B. G. Schunck, "Determining optical flow," Artif. Intell. **17**, 185–203 (1981).
[26]J. M. Fitzpatrick, "The existence of geometrical density-image transformations corresponding to object motion," Comput. Vision, Graphics, Image Process. **44**, 155–174 (1988).
[27]S.-H. Huang, S.-T. Wang, and J.-H. Chen, "3D motion analysis of MR imaging using optical flow method," Proc. Annu. Conf. Eng. Med. Biol. **1**, 463–464 (1995).
[28]S. Priyadarsini and D. Selvathi, "Survey on segmentation of liver from CT images," *IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)* (IEEE, Ramanathapuram, 2012), pp. 234–238.
[29]T. Heimann *et al.*, "Comparison and evaluation of methods for liver segmentation from CT datasets," IEEE Trans. Med. Imaging **28**, 1251–1265 (2009).
[30]A. Bornik, R. Beichel, E. Kruijff, B. Reitinger, and D. Schmalstieg, "A hybrid user interface for manipulation of volumetric medical data," in *IEEE Symposium on 3D User Interfaces (3DUI)* (IEEE, Alexandria, 2006), pp. 29–36.
[31]D. Kainmüller, T. Lange, and H. Lamecker, "Shape constrained automatic segmentation of the liver based on a heuristic intensity model," in *Proceedings of MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge* (Springer, Brisbane, 2007), pp. 109–116.
[32]T. Heimann, H.-P. Meinzer, and I. Wolf, "A statistical deformable model for the segmentation of liver CT volumes," Lect. Notes Comput. Sci. 161–166 (2007).
[33]Y. Chi, P. M. Cashman, F. Bello, and R. I. Kitney, "A discussion on the evaluation of a new automatic liver volume segmentation method for specified CT image datasets," in *Proceedings of MICCAI Workshop on 3D Segmentation in the Clinic: A Grand Challenge* (Springer, Brisbane, 2007).
[34]X. Zhang, J. Tian, K. Deng, Y. Wu, and X. Li, "Automatic liver segmentation using a statistical shape model with optimal surface detection," IEEE Trans. Biomed. Eng. **57**, 2622–2626 (2010).
[35]M. G. Linguraru, J. K. Sandberg, Z. Li, F. Shah, and R. M. Summers, "Automated segmentation and quantification of liver and spleen from CT images using normalized probabilistic atlases and enhancement estimation," Med. Phys. **37**, 771–783 (2010).
[36]M. G. Linguraru, W. J. Richbourg, J. Liu, J. M. Watt, V. Pamulapati, S. Wang, and R. M. Summers, "Tumor burden analysis on computed tomography by automated liver and tumor segmentation," IEEE Trans. Med. Imaging **31**, 1965–1976 (2012).
[37]M. Danciu, M. Gordan, C. Florea, and A. Vlaicu, "3D DCT supervised segmentation applied on liver volumes," in *35th International Conference on Telecommunications and Signal Processing (TSP)* (IEEE, Prague, 2012), pp. 779–783.
[38]M. Goryawala, M. R. Guillen, M. Cabrerizo, A. Barreto, S. Gulec, T. C. Barot, R. R. Suthar, R. N. Bhatt, A. Mcgoron, and M. Adjouadi, "A 3-D liver segmentation method with parallel computing for selective internal radiation therapy," IEEE Trans. Inf. Technol. Biomed. **16**, 62–69 (2012).
[39]Y. Yuan, M. Chao, R. Sheu, K. Rosenzweig, and Y. Lo, "Su-c-18a-06: Tracking fuzzy border using geodesic curve and its application to liver segmentation on planning CT," Med. Phys. **41**, 102 (2014).
[40]Y.-Y. Cheng, H.-M. Chang, H.-R. Su, S.-H. Lai, K.-C. Liu, and C.-H. Lin, "3D liver segmentation and model reconstruction from CT images," in *International Conference on Biomedical Engineering and Biotechnology (ICBEB)* (IEEE, Macau, 2012), pp. 654–657.
[41]C. Platero and M. C. Tobar, "A multiatlas segmentation using graph cuts with applications to liver segmentation in CT scans," Comput. Math. Methods Med. **2014**, 1–16.

[42]A. Klein *et al.*, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," NeuroImage **46**, 786–802 (2009).
[43]B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, and J. C. Gee, "A reproducible evaluation of ANTS similarity metric performance in brain image registration," NeuroImage **54**, 2033–2044 (2011).
[44]S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. Pluim, "ELASTIX: A toolbox for intensity-based medical image registration," IEEE Trans. Med. Imaging **29**, 196–205 (2010).
[45]B. B. Avants, N. Tustison, and G. Song, "Advanced normalization tools (ANTS)," Insight J. **2**, 1–35 (2009).
[46]D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," IEEE Trans. Med. Imaging **21**, 1421–1439 (2002).
[47]T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer, "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," NeuroImage **21**, 1428–1442 (2004).
[48]R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," NeuroImage **33**, 115–126 (2006).
[49]S. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation," IEEE Trans. Med. Imaging **23**, 903–921 (2004).
[50]A. R. Khan, N. Cherbuin, W. Wen, K. J. Anstey, P. Sachdev, and M. F. Beg, "Optimal weights for local multi-atlas fusion using supervised learning and dynamic information (superdyn): Validation on hippocampus segmentation," NeuroImage **56**, 126–139 (2011).
[51]A. J. Asman and B. A. Landman, "Non-local statistical label fusion for multi-atlas segmentation," Med. Image Anal. **17**, 194–208 (2013).
[52]"MICCAI (Medical Image Computing and Computer Assisted Intervention) 2007 Grand challenge," http://www.sliver07.org, 2007.
[53]P. A. Yushkevich, J. Piven, H. C. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," NeuroImage **31**, 1116–1128 (2006).
[54]ELASTIX parameter file, http://elastix.isi.uu.nl/download_links.php, 2016.
[55]A. P. Zijdenbos, B. M. Dawant, R. A. Margolin, and A. C. Palmer, "Morphometric analysis of white matter lesions in MR images: Method and validation," IEEE Trans. Med. Imaging **13**, 716–724 (1994).
[56]B. V. Ginneken, T. Heimann, and M. Styner, in *2007 MICCAI Workshop 3D Segmentation in the Clinic: A Grand Challenge* (Springer, Brisbane, 2007), Vol. 39, pp. 93–95.
[57]See supplementary material at http://dx.doi.org/10.1118/1.4945021 for 3D-SIFT-Flow & 3D Label Transfer.