A multi-atlas segmentation using graph-cuts with applications to liver segmentation in CT scans

C. Platero, M.C. Tobar

Department of Computer Sciences, Technical University of Madrid, Spain

Abstract

An atlas-based segmentation approach is presented. A combination of low level operations, an affine probabilistic atlas and a multi-atlas based segmentation is proposed. The proposed combination provides highly accurate segmentation due to registrations and atlas selections based on the regions of interest (ROIs). Furthermore, our approach shares common elements between the probabilistic atlas and the multi-atlas segmentation: a) the spatial normalization and b) the segmentation method which is based on minimizing a discrete energy function using graph cuts.

The method is evaluated in the segmentation of the liver in Computed Tomography (CT) images. Low level operations define a ROI around the liver from an abdominal CT. Next, a coarse segmentation of the liver is obtained from the probabilistic atlas with a low computational effort. We generate a probabilistic atlas using an affine registration based on geometry moments from manually labeled data. An iterative scheme of registration-segmentation is applied for obtaining a coarse segmentation of the target image using this probabilistic atlas. Then, a multi-atlas segmentation approach improves the accuracy of segmentation. Both the atlas selections and the non-rigid registrations of the multi-atlas approach use a binary mask defined by the coarse segmentation. Experimentally we show that this approach performs better than atlas selections and non-rigid registrations in the whole ROI. The segmentation results are comparable to expert human performance and other recently published results.

Keywords:
Atlas-based segmentation, Graph cuts, Liver, CT

1. Introduction

Segmentation of 3D CT liver images is usually the first step in computer-assisted diagnosis and surgery systems for liver diseases. Segmentation of the liver in such scans is a challenging issue due to the large anatomical variability among patients. Some overviews of segmentation of CT liver scans are given in [1, 2]. These methods have been classified into five main categories: gray level based, graph cuts, model fitting, level set approaches and atlas registration. Although, it can be grouped into two categories: data-driven approaches and model-based approaches. The first group is based on gray level intensities such as thresholding, clustering or region growing. Their major drawback is the adjacent organ separations (e.g., stomach, kidney and heart) which may have intensities similar to the liver. For these images, automatic segmentation of the liver based on gray-value alone is almost infeasible. The relation between pixel intensities and their assigned labels are usually weak. For example, region-growing approaches leak into surrounding tissue and require subsequent manual corrections [3, 4]. For such images, the intensity alone is insufficient for obtaining a robust segmentation. These cases demand the incorporation of prior knowledge of the structures to be segmented. The second group, i.e. model-based segmentation, is a global approach which would match a prior model into given data. It is usually based on a geometrical or statistical model. Normally, building a good prior shape model is difficult because of large inter-patient variability. Statistical shape model is used for liver segmentation frequently because of its ability to constrain the segmentation to approximately match previously seen shapes of a training database [5, 6]. In many case, these approaches do not make full use of the appearance. To overcome this drawback, level set based variational approaches allow to incorporates shape priors into edge-based and region-based models [7, 8]. A popular method to incorporate prior information in the segmentation process is the atlas-based segmentation. In this paper, following Aljabar et al [9], we consider that an atlas is an image in one modality with its respective labeling (which are often generated by manual segmentation). Atlas-based methods start by registering an atlas image into the target image (usually with an intensity-based similarity measure). The resulting transformation is then used to deform the atlas labeled image into the domain of the target image. This process is often called label propagation. These approaches have a simple process compared to other generic segmentation techniques: only a registration method and a number

of pre-segmented data sets are required. However, the main source of error in these approaches depends of the registration techniques. Indeed, insufficient similarity between the registered atlas and the target image often produces unreliable segmentations. An improvement could be done by registering the atlases only near the object of interest and not in the whole image [10]. Taking this into account, we propose a process of refinement of the segmentation. An initial solution is obtained using low level operations which defines the regions of interest (ROIs). Next, a combination of strategies based on several atlases can be applied for each ROI, making that the registration and segmentation operations can be more successful. Liver segmentation from CT is a good example of these difficulties in registration based approaches. The success of a registration in an abdominal CT is compromised by the complexity of this scenario. Before performing any registration, it needs to define the ROI around the liver.

On the other hand, segmentations with a single atlas are intrinsically biased towards the shape and the appearance of a subject [11]. Several studies have shown that approaches which incorporate properties of a group of atlases outperform the use of a single atlas [12, 13]. There are two different atlas-based segmentation strategies using a group of atlases: a) Probabilistic atlas and b) Multi-atlas segmentation. We discuss the advantages and disadvantages of these two atlas-based segmentation strategies using multiple atlases and how they can be combined.

Probabilistic atlas

A probabilistic atlas, in the widest sense, is a spatial statistical model of the appearance and shape of some structures to be studied. The construction of a probabilistic atlas consists of a spatial-normalization step and usually of an intensity-normalization step as well. The spatial normalization is needed to capture the inter-subject variability of the structures. Spatial normalization of the training atlases can be achieved with different registration algorithms. Due to the fact that registration methods are a tradeoff between warp regularization and fidelity term, this leads to a probabilistic atlas of arbitrary sharpness: weak regularization leads to a sharp atlas while strong regularization yields a blurry atlas [14]. Given a smoothness parameter that controlled the registration, an iterative atlas generation scheme is usually employed [12, 15, 16].

Subsequently, the statistical parameter spatial maps, belonging to the probabilistic atlas, are computed for each label. The simplest probabilistic

atlas provides only the prior probability of labels at a spatial position and no information about the expected appearance of the image [15, 16]. More complex probabilistic atlases provide statistics on the relationship between the labels and the intensities [17, 18]. We focus on a probabilistic atlas with modeling both the appearance and shape of the objects to be segmented.

Once constructed the probabilistic atlas and given an image to be segmented, the probabilistic atlas is registered into the target image and then it is used in the segmentation task as prior information in the Bayesian formulation [15, 16, 17]. The advantage of this approach is that once the probabilistic atlas has been generated, only a single registration from the atlas to the target image is required for obtaining a segmentation. However, this method depends on the success of a single registration. To overcome this drawback, the new approaches join the registration and segmentation of a unseen image as an iterative process of estimating the labeling and calculating the registration parameters [14, 19, 20, 21]. In this work, we build a probabilistic atlas and a coarse segmentation of the target image is obtained by applying the probabilistic atlas, using an iterative process of affine registration and segmentation.

Multi-atlas segmentation

Atlases within a database can be registered to a target image and their segmentations can be transformed and subsequently fused to provide a consensus segmentation. The main benefit of the multi-atlas segmentation approach is that the effect of errors associated with any single atlas propagation can be reduced in the process of combination. The main drawback is the computational complexity. Indeed, the computational time of segmentation increases linearly with the number of the atlases that have to be registered. Some authors [13, 9] have demonstrated that the precision in segmentation is improved as more atlases are combined. The overlap accuracy of a multi-atlas segmentation starts increasing rapidly and finally, becomes very slowly when more segmentations are fused. Therefore, an atlas selection is required so that the number of atlas is as low as possible, in a manner that no further improvement is expected when adding more atlases. Finally, a label fusion method is also required to obtain a consensus segmentation. The fusion of the propagated segmentations can be achieved in different ways: majority voting rule [12], STAPLE [22] or minimization of an energy function with intensity and prior terms [23]. Therefore, the non-rigid registrations, the atlas selection and the label fusion method have to be investigated in order to

improve the performance of multi-atlas segmentation approach.

The paper is organized as follows. In Section 2, the proposed method of combining low level operations with a probabilistic atlas and a multi-atlas segmentation approach is presented. The experiments that are performed for the liver segmentation are described in Section 3. The results and conclusions are presented in Section 4.

2. Method

A flow chart of the proposed framework is shown in figure 1. Given an initial solution of the object of interest by using low-level operations, a ROI is determined. Next, a fast probabilistic atlas is applied in the ROI and a coarse binary segmentation (S_C) is calculated by minimizing a discrete energy function using a graph-cut technique. S_C is a binary mask image, which is used to define the domain of non-rigid registrations and similarity measures of the atlas selection. Finally, the atlases are ranked and the selected atlases are propagated to the target image and a label fusion method, which is also based on minimization of a discrete energy function, improves the segmentation with higher accuracy. Below we present how the ROI of the liver is obtained, the segmentation method and the combination of a probabilistic atlas with a multi-atlas segmentation.

2.1. Initial solution

An initial solution is required to define the region of interest around the anatomical structure to be segmented. The initial solution allows us to introduce prior knowledge about the segmentation problem using low-level operations. For example, in the case of the liver segmentation from these CT scans, liver and heart have nearly the same intensity and a liver-heart separation algorithm has to be applied for preventing over-segmentation [3]. In particular, a combination of conventional and specific techniques are applied for obtaining the initial solution: liver-heart separation, non-linear diffusion, 3D edge detector and morphological post-processing. Firstly, a liver-heart separation surface is computed following [3]. The segmentations of the lungs are calculated. For each coronal slice, a minimal length curve is found, which connects the bottom of the right lung lobe with the bottom of the left lung lobe. The set of these curves define the liver-heart separation surface. Then, the target image is filtered by means of a nonlinear diffusion filter with selection of the stopping time [24]. The filtered image is close to a piecewise smooth model.

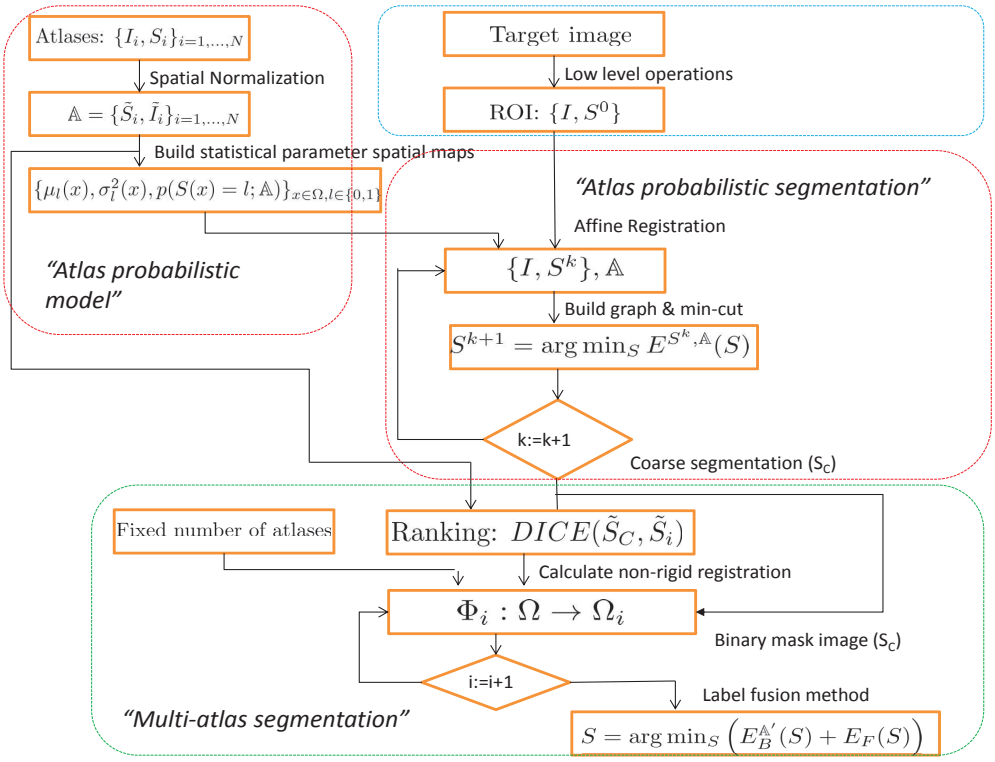

Figure 1: Flow chart summarizing the three main steps of the proposed method.

Histogram analysis, 3D edge detection and the liver-heart separation surface are applied to the filtered image which produces a partition in isolated regions. The segmentation is followed by various post-processing steps so that the size (largest organ) and the location (mostly on the right side) are used to determinate the initial solution. We denote this initial segmentation as S_0.

2.2. Segmentation method

Both the probabilistic atlas and the multi-atlas segmentation run a segmentation method based on minimizing a pseudo-boolean function by using a graph-cut technique. A Conditional Random Field (CRF) model [25] is used to incorporate terms of appearance and shape, which are estimated from the training atlases. Previously, other authors have used this framework [26, 23, 27]. Our method incorporates the following differences: a) A generative appearance model based on intensity from each pixel and its

neighborhood, b) A label prior probability is estimated using a majority voting rule and c) A spatial regularizer that minimizes the surface of separation between two different labels using a Finsler metric [28].

Consider a set of N training atlases for a region of interest with binary labeling $\{A_i\}_{i=1,\dots,N} = \{I_i, S_i\}_{i=1,\dots,N}$ and a target image I with initial solution S^0, where $I_i : \Omega_i \subset \mathbb{R}^n \to \mathbb{R}$, $n = 3$ and $S_i : \Omega_i \subset \mathbb{R}^n \to \{0, 1\}$ are the labeled maps. In the labeled images, object pixels are labeled $S(x) = 1$, background pixels by $S(x) = 0$. Let $\Phi_i : \Omega \to \Omega_i$ be the spatial mapping from the target image coordinates to the coordinates of $i - th$ training subject. For simplicity, we assume that $\{\Phi_i\}_{i=1,\dots,N}$ have been pre-computed by using a pairwise registration procedure. This assumption allow us to shorthand $\mathbb{A} = \{\tilde{S}_i = S_i \circ \Phi_i, \tilde{I}_i = I_i \circ \Phi_i\}_{i=1,\dots,N}$ as the training set into the common coordinates.

The segmentation of an image I, based on image intensities and prior knowledge, is computed by the minimization of a discrete energy function:

$$S^* = \arg\min_S E^{\mathbb{A}}(S), \qquad E^{\mathbb{A}}(S) = E_B^{\mathbb{A}}(S) + E_F(S), \tag{1}$$

where the term $E_B^{\mathbb{A}}(S)$ is derived from $\mathbb{A}$ using the framework of Bayesian estimation theory and $E_F(S)$ is associated with an image-based Finsler metric. We consider S as a discrete random field with a neighbourhood system $\mathcal{N}$, which is the set of edges connecting variables in the random field. The CRF model is defined by unary and pairwise potentials:

$$E(S) = \sum_{x \in \Omega} \psi_x(S(x); I, \mathbb{A}) + \sum_{x,y \in \mathcal{E}} \psi_{xy}(S(x), S(y); I) \tag{2}$$

The unary potentials of the CRF model are defined as the negative log of the likelihood of a label being assigned to a pixel. It is computed from an appearance model and a label prior. The pairwise edge potentials of the CRF take the form of a contrast sensitive Potts model.

The unary potentials

The Bayesian formulation allows incorporate prior information about the shape and appearance of structures to be segmented. To find the maximum a posteriori probability (MAP) estimation is equivalent to minimize the following energy function where the Bayes theorem is applied

$$E_B^{\mathbb{A}}(S) = -\log\left(p(S|I;\mathbb{A})\right) = -\log\left(\frac{p(I|S;\mathbb{A})p(S;\mathbb{A})}{p(I;\mathbb{A})}\right).$$

We assume that the observed intensities of I are independent random variables. The image likelihood $p(I|S;\mathbb{A},S^k)$ can then be written as a product of the likelihoods of the individual pixels

$$p(I|S;\mathbb{A}) = \prod_{x\in\Omega} p(I(x)|S(x);\mathbb{A}).$$

Usually, the intensity distribution is modeled by a mixture of Gaussians [20, 19]. Alternatively, we use a Gaussian distribution for each pixel and for each label [18, 21]:

$$p(I(x)|S(x)=l;\mathbb{A}) = \frac{\exp\left(-\frac{(I(x)-\mu_l(x))^2}{2\sigma_l^2(x)}\right)/\sigma_l(x)}{\sum_{j\in\{0,1\}}\exp\left(-\frac{(I(x)-\mu_j(x))^2}{2\sigma_j^2(x)}\right)/\sigma_j(x)}$$

where $l\in\{0,1\}$. We estimate the statistical parameters $\mu_l(x)$ and $\sigma_l^2(x)$ from the registered atlases as follows:

$$\mu_l(x) = \frac{\sum_{i\in Q_l(x)}\tilde{I}_i(x)}{\#Q_l(x)}, \qquad \sigma_l^2(x) = \frac{\sum_{i\in Q_l(x)}(\tilde{I}_i(x)-\mu_l(x))^2}{\#Q_l(x)-1}.$$

where $Q_l(x)=\{i|\tilde{S}_i(x)=l\}$. The means and variances are estimated using a variable number of samples $\#Q_l(x)$. To overcome bias in the statistical maps, they are smoothed by means of linear diffusion with Neumann boundary conditions:

$$\mu_l(x) = h_{\sigma_s}(x) * \frac{\sum_{i\in Q_l(x)}\tilde{I}_i(x)}{\#Q_l(x)}, \qquad \sigma_l^2(x) = h_{\sigma_s}(x) * \frac{\sum_{i\in Q_l(x)}(\tilde{I}_i(x)-\mu_l(x))^2}{\#Q_l(x)-1} \tag{3}$$

where $*$ denotes the convolution operator and $h_{\sigma_s}(x)$ is a Gaussian mask, being σ_s the scale parameter [29], which is a tunable parameter.

The label prior probability $p(S;\mathbb{A})$ models the joint probability of all pixels in a particular label configuration. Instead, we assume that the prior probability that a pixel x has label l only depends on its position:

$$p(S;\mathbb{A}) = \prod_{x\in\Omega} p(S(x);\mathbb{A}).$$

For each pixel x and each label $l\in\{0,1\}$, we define

$$p(S(x)=l;\mathbb{A}) = \frac{\#Q_l(x)}{N}. \tag{4}$$

Image likelihood and label prior terms are combined to define the unary potentials $\psi_x(S(x); I, \mathbb{A})$:

$$\psi_x(S(x); I, \mathbb{A}) = -\log\left(\frac{p(I(x)|S(x); \mathbb{A})p(S(x); I, \mathbb{A})}{p(I(x))}\right)$$

Spatial regularization

Following the work of Boykov and Kolmogorov [28, 30], the smoothness term E_F of the energy function is defined from a Finsler metric. These authors decomposed the energy into E_R and E_f with weights $\lambda_1, \lambda_2 \in \mathbb{R}$, that is,

$$E_F(S) = \lambda_1 E_R(S) + \lambda_2 E_f(S).$$

The first part minimizes the segmentation surface by a Riemannian metric and the second one takes into account the orientation of the segmentation surface in the metric. We consider that the isotropic Riemannian metric from the image is defined by $D(x) = g(\|\nabla I(x)\|)\mathbb{I}$, where $\mathbb{I}$ is an identity matrix, $g(x) = (\exp(-x/\gamma))^{1/3}$ and γ is estimated as the average of $\|\nabla I(x)\|$. The pairwise potentials are defined by

$$\begin{aligned}\psi_{xy}(S(x), S(y); I) = \lambda_1 \omega_x^R(y)(1{-}S(x))S(y){+}\\ \lambda_2 \omega_x^f(y)(S(x)(1{-}S(y)){-}S(y)(1{-}S(x))),\end{aligned}$$

where $\omega_x^R(y) = \dfrac{g(\|\nabla I(x)\|)}{\|x - y\|}$ and $\omega_x^f(y)$ is assigned by the dot product between $\nabla I(x)$ and the vector defined by x and y.

Therefore, the proposed model is characterized by a pseudo-boolean function defined on unary and pairwise potentials and the optimal labeling is found by applying the min-cut/max-flow algorithm of [31].

2.3. Getting the coarse segmentation

The coarse segmentation S_C is used to define the domain of non-rigid registrations and similarity measures of the atlas selection. S_C is obtained by means of a probabilistic atlas. We have observed that S_C improves the performances of the multi-atlas segmentation with respect to S_0 or the conventional approach (i.e. without defining the domain using binary masks for registering and atlas selections) [32].

In a probabilistic atlas, the information from atlases is usually combined into a generative model in a common coordinate system. In this paper, an atlas is selected as a reference to which all atlases are then co-registered using a particular registration method [17]. An affine registration is applied for spatial normalization. The principle that we have adopted for spatially normalizing transformations is that these transformations should align the anatomical structures with low computational effort. We only expect a coarse segmentation of the target image that improves the results of the multi-atlas segmentation. For this propose, the affine transformations are sufficient. Usually, the affine registration is done with the modality images. Since we have an initial solution (S_0) and iterative segmentation-registration approaches work better than a single propagation of the probabilistic atlas [20], we use an affine registration based on the alignment of the labeled images, where its parameters are calculated by geometric moments of the labels [33]. The atlases are co-registered to a manual segmentation and the statistical parameter maps are calculated.

Due to the complex dependencies between the unknown segmentation, the target image and the registration parameters, this problem is simplified using an EM framework [34]. The E-Step captures the posterior probability of the structure which depends the registration and the M-Step updates the registration parameters [20]. A coarse segmentation is obtained by an iterative method of registration-segmentation. Given an initial solution of the ROI (S^0), belonging to the target image, $\{I, S^0\}$ are aligned into the reference system by using the same affine transformation method, which was used to built the probabilistic atlas [1]. A segmentation of I is computed by minimizing our energy function. The resulting segmentation is again used for registering $\{I, S^1\}$ into the all co-registered atlases.

This process converges with a few iterations. In each iteration, the centroid and the axes of the structure to be segmented are better estimated and thus the registration between $\{I, S^k\}$ and the co-registered atlases is less biased, which also implies an improvement in the segmentation. Certainly, there is a feedback effect between the affine registration and the minimization of the energy function. Finally, an inverse affine transformation is applied to

[1] Although in the segmentation method, the atlases are co-registered into the target image (see subsection 2.2), there is no loss of generality in assuming that $\{I, S^k\}$ are aligned into the normalized atlases using affine transformations

return the segmentation into the native space of the target image.

2.4. Multi-atlas segmentation

Given the coarse segmentation, a multi-atlas approach improves the segmentation with higher accuracy. All the atlases are not needed to be registered into the ROI of the target image [13, 9]. An atlas selection framework is required for selecting the atlases that best propagate their labels. Several methods to rank atlases have been tested [9, 35]. They are usually based on a similarity measure between each atlas with respect to the target image. Other criteria, which are not based on the similarity between images, are not considered in this study (e.g. the age of the patients in the medical image analysis). The best strategy as tradeoff between reliability and computational cost consists in the registration of the target image into a reference where the atlases were previously registered. This approach requires only one registration in runtime. Most often, the ranking is done with an intensity-based similarity measure computed between the target image and each atlas. We propose to use DICE coefficient [36] as a similarity measure between an aligned coarse segmentation of the target image and each aligned atlas labeled image. The above affine registration, which is based on geometric moments of the labeled images, is used for the alignments of images due to it was already used for calculating the coarse segmentation and now it is also employed for the atlas selection.

Once the atlases are ranked, the number of the selected atlases is required. Aljabar et al [9] showed that given a ordered list of atlases, the accuracy of the final segmentation increases rapidly up to a maximum level following by a gradual decline according to the number of fused atlases. This indicates that a fixed number of atlas can be determined for each application.

Non rigid registration

We have chosen a technique based on the maximization of an intensity-based similarity measure, in combination with a deformation field parameterized by cubic B-splines. Klein et al [37] have demonstrated for some intensity-based similarity measures that the optimization converges to the solution when a very small number of random samples of intensity pairs are used. The flexibility of the control point grid also allows to introduce a binary mask image, where is only applied the nonrigid transformation. Here, this binary mask is defined using a dilated version of the coarse segmentation,

S_C. The dilation is used for including borders of the anatomical structure and some surrounding tissues.

Given a target image, the selected atlases are firstly aligned to the target image using the above affine registration. These affine parameters are calculated with the coarse segmentation of I, i.e. S_C, and the corresponding atlas labeled image. This step is faster than the rest of the nonrigid registration. Next, a multi-resolution scheme is used in the nonrigid registration step to avoid local minima.

Label fusion method

Several methods have been proposed to combine the propagation segmentations of the selected atlases into a single segmentation. Among them, majority voting rule [12], STAPLE [22] or the minimization of an energy function with intensity and prior terms [23]. We consider our segmentation method that uses graph cuts to optimize a discrete energy function based on registered atlases. Given the registered atlases to the target image, a statistical model of appearance and shape is computed. Subsequently, a graph-cut technique is used to minimize the energy function defined in (2) and the consensual segmentation is obtained from the graph cut [23]. The difference with the segmentation based on the proposed probabilistic atlas is that the fusion method does not require iterations of registration-segmentation. Now the statistical parameter spatial maps are not biased because the selected atlases are considered better registered into the target image. Therefore, given the set of selected registered atlases into the target image, the statistical parameter maps are calculated and a consensual segmentation is obtained by $S^* = \arg\min_S \left(E_B^{\mathbb{A}'}(S) + E_F(S)\right)$ by using graph-cuts.

3. Experiments with liver CT data

The atlases and the test data were taken from a public database for liver segmentation [38](*http://www.sliver07.org/*). We have used this public database, which allows us to compare our approach to a large number of other segmentation methods, including other atlas-based schemes [32]. A total of 30 images were divided randomly into 20 images used for the training set and the other 10 ones are used for the test set. All CT images are enhanced with a contrast agent and scanned in the central venous phase using diverse scanners (machines with 4, 16, and 64 detector rows). The pixel spacing varies from 0.55 to 0.88 mm, while the inter-slice distance varies from 1 to

3 mm. Most images in this study exhibited pathologies, including tumors, metastasis and cysts of different sizes.

To evaluate the quality of a given segmentation, we have used the following five error measures [38]: overlap error (m_1), relative absolute surface difference (m_2), average symmetric distance (m_3), root mean square symmetric distance (m_4) and maximum symmetric distance (m_5). The main advantage of using multiple measures instead of a single measure is that different measures detect various aspects of the segmentation quality. The problem of how to combine the different measures to produce a ranking of segmentation results is solved by transforming the result of each error measure to a common scale and averaging the resulting values to obtain the final score [38]. Each measure is converted to a scale ranging from 0 to 100 by:

$$\alpha_j(X,Y) = \max\left(100 - 25\frac{m_j(X,Y)}{\tilde{m}_j}, 0\right), \quad j = 1,\ldots,5$$

where X and Y represent the manual and automatic segmentation binary images, respectively, m_j is an error measure and $\tilde{m}_j$ is the corresponding reference value which was obtained by the average of manual segmentations. An $\alpha_j = 100$ points is a perfect match with the reference segmentation and a score around 75 is equivalent to human performance. The final score is the average of the individual measure scores: $\alpha(X,Y) = \sum_{j=1}^{j=5} 0.2 \cdot \alpha_j(X,Y)$. In order to compare our approach with other methods and applications, the resulting segmentations have also been measured using the DICE coefficient.

3.1. Experiments with ground truth

The training atlases are used for experimental validation in a leave-one-out fashion: one atlas is used as the target image and the other 19 as training atlases. The manual segmentation of the target image is used as the ground truth. This procedure is repeated in all atlases that belong to the training set.

For the coarse segmentation and the atlas selection, the atlases and the target image are firstly sub-sampled by a factor of two in each dimension in order to reduce computation time. Pilot experiments showed that using the full resolution data increased computational times and hardly improved the results. However, the non-rigid registrations of the atlases to the target image are done within the original resolution.

3.1.1. Setting the probabilistic atlas parameters through training

Three parameters of the probabilistic atlas are tuned: a) the scale parameter σ_s of (3), b) the multipliers λ_1 and λ_2 of the energy function and c) the number k of iterations in (1). Twenty leave-one-out segmentations on the training atlases are performed to determinate the tunable parameters. These parameters are varied in certain ranges and their effects are measured by the overlap between the resulting segmentation and the ground truth. DICE is chosen as measure of the segmentation overlaps. In our experiments, k varies from 1 to 10. Since σ_s and the multipliers are coupled together, an iterative adjustment is used. We have observed that the Riemannian metric is more influential than the surface orientation term in the optimization process. Therefore, λ_1 is firstly tuned and then λ_2 is adjusted. Considering 3D grid-graphs with 6 neighborhood system in the CRF model, $\sigma_s = 1$, $k = 1$, $\lambda_2 = 0$ and $\nabla I(x)$ is calculated by Gaussian derivatives at scale 1, λ_1 is varied and DICE is used for detecting the optimal value (see Fig. 2 (a)), which is tuned with $\lambda_1 = 10$. With a 6 neighborhood system, $\nabla I(x)$ is easily decomposed into $\omega_x^f(y)$ and λ_2 can take positive or negative values due to the orientation of $\nabla I(x)$. Indeed, the gray level of the liver may be brighter than in other adjacent tissues (the majority of the times) but it may also getting to be more darker in other areas (e.g. contact liver-kidney). Fig. 2 (b) shows the trend of λ_2 ($\sigma_s = 1$, $k = 1$, $\lambda_1 = 10$). It is clear the least impact on the success of the segmentation. With fixed values $\lambda_1 = 10$ and $\lambda_2 = 4$, σ_s and k are tuned to 2 and 4 respectively (Fig. 2 (c)). The tuning process is repeated with $\sigma_s = 2$. The parameters did not vary substantially ($\lambda_1 = 8$, $\lambda_2 = 4$) and performances are virtually the same.

The experiment shows the utility of performing an iterative scheme of segmentation-registration for obtaining S_C (see Fig. 2 (c)). In comparison to the classical registration-segmentation with a probabilistic atlas ($k = 1$), the iterative method given in (1) performs better for $k > 1$, due to more reliable estimation of the posterior probabilities which is produced by the better registration, between the probabilistic atlas and the target image. We also observe the effect of spatial regularization $\{\mu(x,l), \sigma^2(x,l)\}_{l\in\{0,1\},x\in\Omega_{ref}}$ of $\mathbb{A}$ by linear diffusion. With the linear diffusion, the parameter estimations of the appearance model are less biased and the iterative process is convergent and stable, otherwise ($\sigma_s = 0$) it is not.

Furthermore, we have experimentally observed that if our probabilistic atlas does not have the appearance model as in [15], the numerical scheme

does not work. This is because there is a weak feedback effect between segmentation and registration. Indeed, the intensities of the target image are only used in the regularization term of the optimization process, i.e. pairwise potentials are only updated in each iteration of the EM framework. Therefore, there are no substantial changes in the new segmentation and consequently do not vary the parameters of the affine registration.

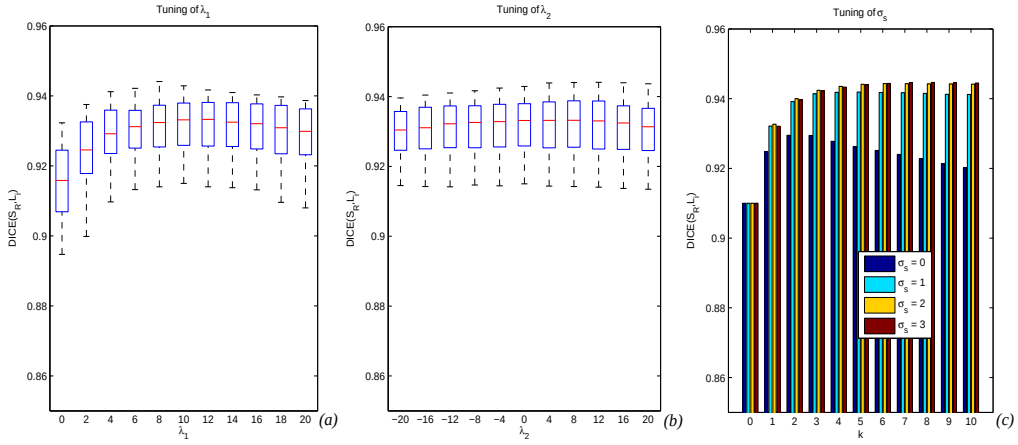

Figure 2: Tuning of parameters λ_1, λ_2 and σ_s: (a) The distributions of $DICE(S_R, L_i)$ with the values of λ_1, where S_R is the reference segmentation and L_i is the automatic segmentation obtained with λ_1 ($\sigma_s = 1$, $k = 1$, $\lambda_2 = 0$). (b) The distributions of $DICE(S_R, L_i)$ with the values of λ_2 ($\sigma_s = 1$, $k = 1$, $\lambda_1 = 10$). (c) The average DICE in segmenting all target images for σ_s and for each iteration ($\lambda_1 = 10$, $\lambda_2 = 4$).

3.1.2. The ranking of atlases for the multi-atlas segmentation

Any ranking of atlases based on similarity measures requires a spatial normalization step. In this paper, the atlases and the target image are registered into a common reference using affine transformations. An atlas of the training set is chosen as reference for the spatial normalization [39].

Three criteria for ranking the atlases are tested: a) a random order which does not require any registration task, b) using Mutual Information (MI) as in [9] and c) our method based on a similarity measure between each atlas labeled image and S_C with DICE coefficient. Once ranked the atlases, they are registered into the target image. The same non-rigid registration method is applied for the three ranking criteria. All nonrigid registrations are computed using *Elastix* [40], an available public package for medical image

registration. This framework is based on techniques described in [41, 37, 42]. We employ a four level multi-resolution scheme. In our experiments, the B-Spline grid spacing are 64, 32, 16 and 8 mm in all directions, for the four respective resolutions. The negative MI is used as the cost function, which is implemented according to [42]. For the optimization of the cost function, an iterative stochastic gradient descent optimizer is used [37]. In each iteration, 2000 random samples are used to calculate the derivative of the cost function. Random samples are acquired from a domain defined by a binary mask, which is set by a dilated version of S_C.

We denote by Φ_i the combined transformation of affine and non-rigid registrations from the domain of I to A_i. Fig. 3 shows the relationship between the individual atlases and their performance in segmenting the target images. The results are shown in two graphs: a) the distributions of $\alpha(S_R, \tilde{S}_i)$ for a given rank (random, MI, proposed or optimal), where S_R is the ground-truth segmentation of the target image, i is the order of atlas in the database from the similarity to the target image and $\tilde{S}_i = S_i \circ \Phi_i$ is the corresponding deformed labeled image and b) the average of $\alpha(S_R, \tilde{S}_i)$ or $DICE(S_R, \tilde{S}_i)$ against the rank where each plotted point shows the average score or DICE for all target images at a given rank. The optimal ranking is obtained by sorting in descending order according to the score of each label using S_R. Our atlas selection is the most similar to the optimal ranking. Although the segmentations are not fused in this experiment, there are relationships between the atlas rankings and the segmentation accuracies derived from fusing these selected atlases (it will be shown in the following experiment).

3.1.3. Stopping criterion

The following experiment describes a test of the effect of varying the number of selected atlases for generating the consensual segmentation. Given a rank of atlases and a number of selected atlases, the single consensual segmentation is computed. The same method of non-rigid registration and fusion are applied for all samples and for all rankings. Registrations is done under the same conditions as in the previous experiment. The label fusion method is based on minimizing our energy function as it was explained in subsection 2.4. Fig. 4 shows how the segmentation accuracy varies with the number of the fused atlases. In the upper part of the figure, each box and whiskers in the upper part of the graph visualizes the distributions of $\alpha(S_R, L_i)$ for each criterion depend on the ranking of atlases. L_i is a binary image of the target image that is obtained by the fusion method with the

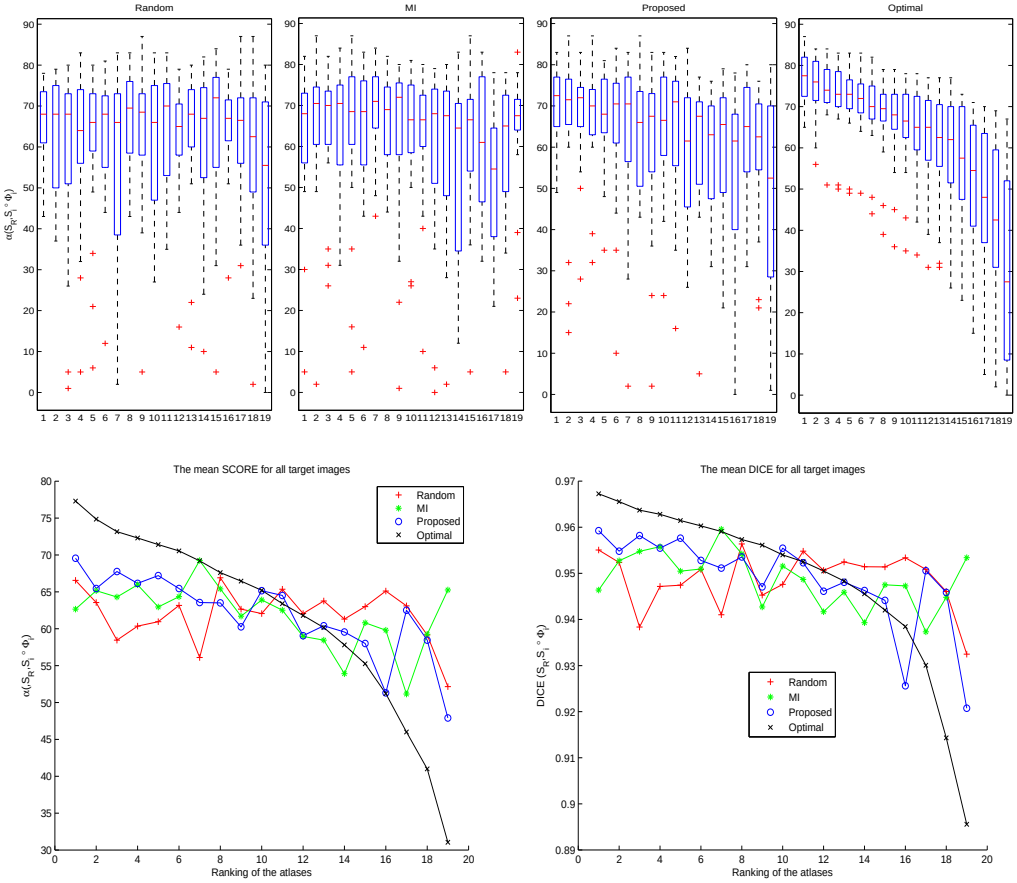

Figure 3: Relationship between the individual atlases and their performance in segmenting the target image. The upper part of the graph visualizes the distribution of score in segmenting by label propagation for a given rank (random, MI, proposed or optimal) and for an individual atlas, $\alpha(S_R, \tilde{S}_i)$. The lower part of the figure shows the average scoring and the average DICE on all target images against the rank.

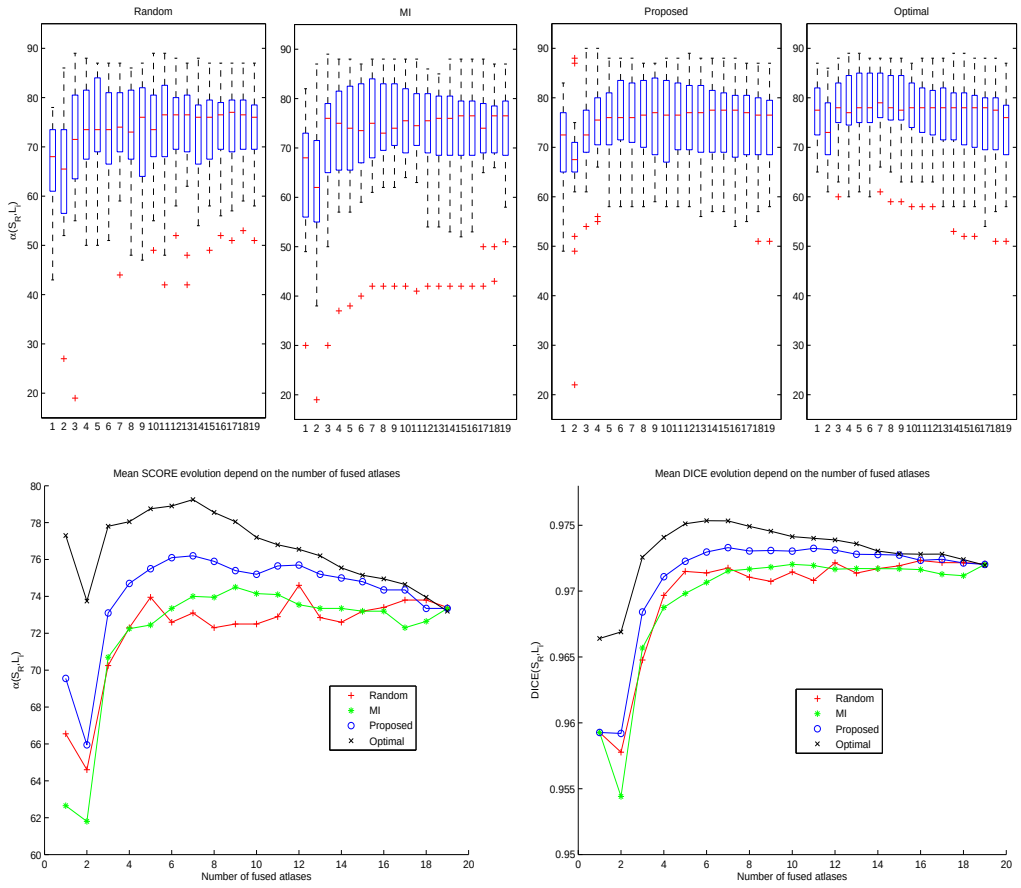

Figure 4: Relationship between the segmentation accuracy and the number of the fused atlases. The upper part of the graph visualizes the scoring distribution in segmenting by label propagation for a number of fused atlases depending on the type of rankings, $\alpha(S_R, L_i)$. The lower part of the figure shows the mean scoring and DICE in all target images against the different number of fused atlases and for different atlas selection criterion.

first i selected atlases. In the lower part of the figure, each plotted point shows the average score or DICE coefficient in segmenting all target images for the number of fused atlases. The general pattern shows a sharp initial increase up to a maximum score followed by a gradual decline, as in [9]. The decrease in the score when two atlases are fused is due to increased uncertainty. Discrimination with only two registered atlas can not rely on other sources to doubtful cases. We also observe that the score system is more sensitive than the DICE coefficient. The DICE coefficient does not show a fall when two atlas are fused.

Our approach requires 6 or 7 atlases for getting the maximum score and DICE ($\alpha = 76.3 \pm 8.6$ and $DICE = 0.973 \pm 0.007$). The other atlas selections (MI and random) require a greater number of atlases for getting the maximum score (MI: 9, random: 12) and their scores are worse (MI: $\alpha = 74.50 \pm 10.9$ and $DICE = 0.972 \pm 0.01$, random: $\alpha = 74.80 \pm 9.5$ and $DICE = 0.972 \pm 0.008$). However, no significant improvements are obtained between our proposal and the other atlas selections. Only the effect of atlas selection has been evaluated. Recall that the method of registration and fusion are applied in the same manner in all atlas selection frameworks. Statistical significance is evaluated using a paired two-sided Wilcoxon test where $p-$value < 0.05 shows significant improvement. Given a fix number of atlases equal to 6 and compared to our proposal, the $p-$values are 0.456 and 0.297 corresponding to the MI-based and random-based atlas selections respectively.

3.1.4. Comparison with a conventional multi-atlas segmentation

In Heimann et al [1], this atlas database is used for showing a comparison among several methods for liver segmentation from CT images. One of the presented methods is based on multi-atlas segmentation [32]. A fix number of twelve selected atlases are registered to the target image using an affine transformation followed by a B-spline approach in multiple resolutions. The cost function used for the registrations is the negative MI. Using these ideas, we implement a conventional multi-atlas segmentation. The atlases are registered to the whole target image as in [32], and the ranking of the atlases is done according to [9]. We compare this approach with our proposal. Our proposed label fusion is used in both approaches. The upper part of Fig. 5 illustrates the distributions of $\alpha(S_R, L_i)$ with the number of the fused atlases. The lower part of the figure shows the average scores and DICE in segmenting all target images for the number of fused atlases and for each approach. Set-

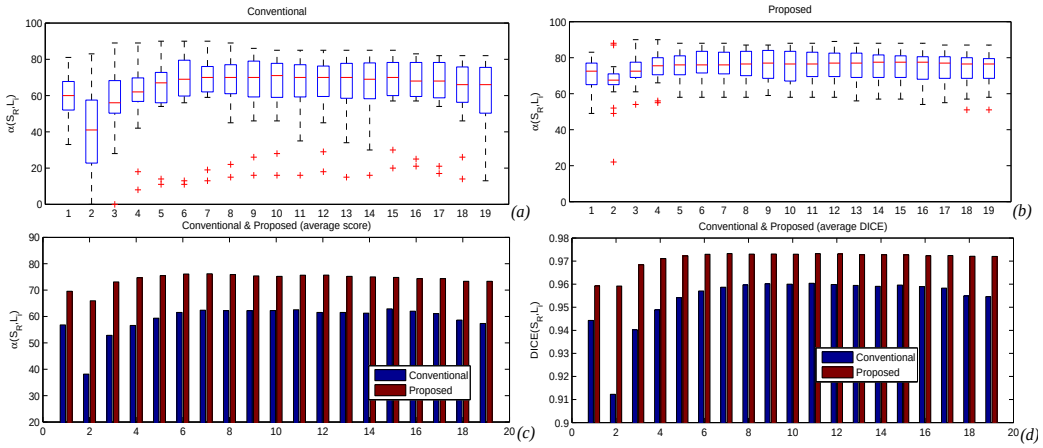

Figure 5: Comparison between a conventional multi-atlas segmentation and the proposed approach depending on the number of fused atlases: (a) The distributions of $\alpha(S_R, L_i)$ with S_R is the ground-truth segmentation and L_i is the consensual segmentation with the first i atlases. (b) The distributions of $\alpha(S_R, L_i)$ for the proposed approach. (c) The average scores in segmenting all target images for the number of fused atlases and for each approach. (d) The average DICE in segmenting all target images for the number of fused atlases and for each approach.

ting the number of fused atlases to 6, the conventional approach gives worse results (α : 61.5 ± 25.1, DICE: 0.957 ± 0.037) than the proposed approach ($\alpha : 76.3 \pm 8.6$, DICE: 0.973 ± 0.007). For all cases, $p-$values are always less than 0.05, indicating significant improvement between our approach and the conventional one for any number of fused atlases.

3.2. Results

The performances of the three stages (initial solution, coarse segmentation and label propagations) are evaluated by computing $\{m_j, \alpha_j\}_{j=1,\ldots,5}$ for the case of liver segmentation. Table 1 gives the mean values and standard deviations for each step of the measures and scores of the segmentations of the target images. The DICE coefficients are 0.910 ± 0.024, 0.943 ± 0.016 and 0.973 ± 0.007 for initial solution, coarse segmentation and multi-atlas segmentation method respectively and the scores are $34.4 \pm 14.3, 56.3 \pm 15.1$ and 76.3 ± 8.6 (see Fig. 6). The evaluation of the test images is performed by an external team by submitting the results to the web site as in [1]. The results of the test images are lower than that of the training set, $\alpha : 70.5 \pm 16.4$ (test), 76.3 ± 8.6 (training).

Table 1: Average values of the measures and scores for all 20 training images and 10 test images: volumetric overlap error (m_1), relative absolute volume difference (m_2), average symmetric surface distance (m_3), root mean square symmetric surface distance (m_4), and maximum symmetric surface distance (m_5).

Type		$m_1[\%]$	$m_2[\%]$	$m_3[mm]$	$m_4[mm]$	$m_5[mm]$	$\alpha()$
Initial	measures	16.5 ± 3.8	8.6 ± 5.8	3.6 ± 1.6	7.1 ± 3.9	50.2 ± 24.8	
Solution	scores	37	56	21	19	39	34.4
Affine Prob.	measures	11.2 ± 2.8	2.3 ± 1.8	2.2 ± 0.9	4.4 ± 1.9	35.8 ± 15.5	
Atlas	scores	56	88	44	40	53	56.3
the proposed approach	measures	5.2 ± 1.2	0.9 ± 1.2	1.0 ± 0.4	2.2 ± 1.1	26.9 ± 10.7	
20 training images	scores	79	91	76	70	65	76.3
the proposed approach	measures	7.6 ± 3.2	-0.5 ± 3.9	1.3 ± 0.7	2.9 ± 1.8	24.7 ± 10.7	
10 test images	scores	70	87	68	60	68	70.5

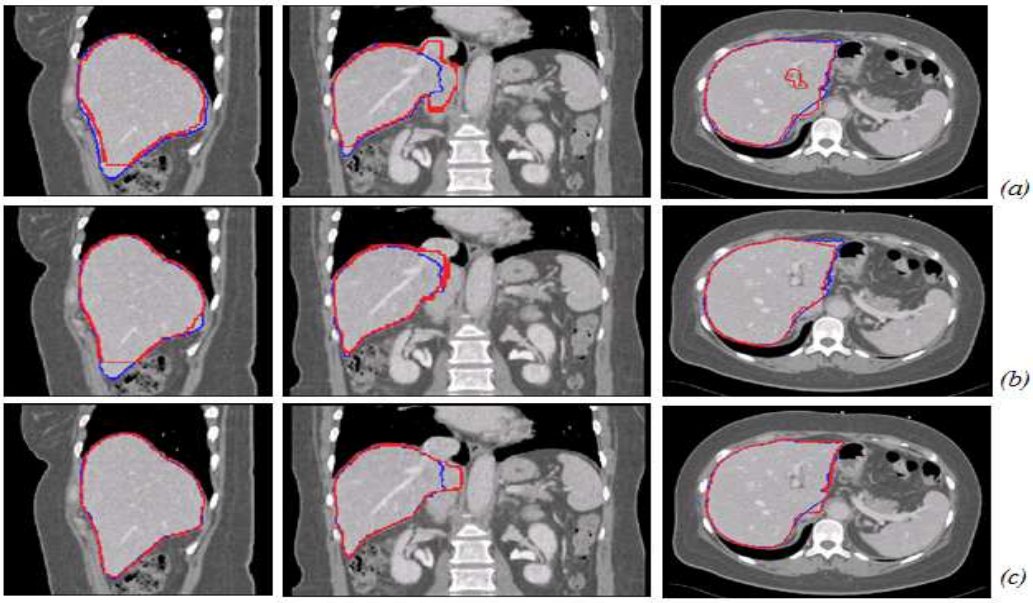

Figure 6: Resulting segmentations for a representative patient from the training database. From left to right, a sagittal, coronal and axial slice for (a) Initial solution (b) Coarse segmentation and (c) multi-atlas segmentation. The outline of the ground-truth segmentation is in blue, the outline of the segmentation of the method described in this paper is in red.

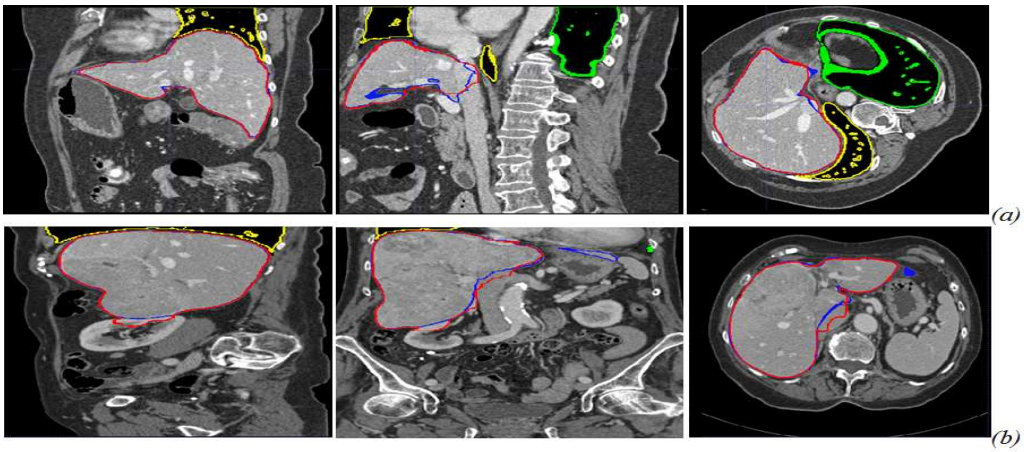

Figure 7: From left to right, a sagittal, coronal and axial slice for an easy case (a) and a difficult one (b). The outline of the ground-truth segmentation is in blue, the outline of the segmentation of the method described in this paper is in red. It is also illustrated the outline of the lung segmentations in green (left lung) and yellow (right lung).

Fig. 7 shows slices from two cases, drawing the result of the method (in red) and the manual segmentation (in blue) from the training database. Fig. 8 shows two cases with tumors from the test database. The proposed approach is shown robust to the first case, but not for the second case when tumors appear in the boundary of the liver. Fig. 8(b) shows the worst score of all images. The results listed at *http://www.sliver07.org/* show that our approach is comparable to expert human performance and other recently published results. The method is implemented with the ITK library and some procedures are parallelized with OpenMP. The average computational times are $17.16 \pm 6.18s$ and $33.29 \pm 9.27s$ for initial solution and course segmentation (including initial solution) respectively ([Dual CPU] Intel Xeon E5520 @ 2.27GHz). Obviously, the major computational cost is the multi-atlas segmentation method, the average computational time is $208.33 \pm 17.60s$ per atlas registered. Fusing six atlas into the target image, the total runtime to segment one sample is 22.13 ± 1.86 minutes.

4. Discussion and conclusions

An atlas-based segmentation framework is proposed that combines low level operations, a fast probabilistic atlas with a multi-atlas segmentation. The proposed combination provides high accuracy in segmentation due to

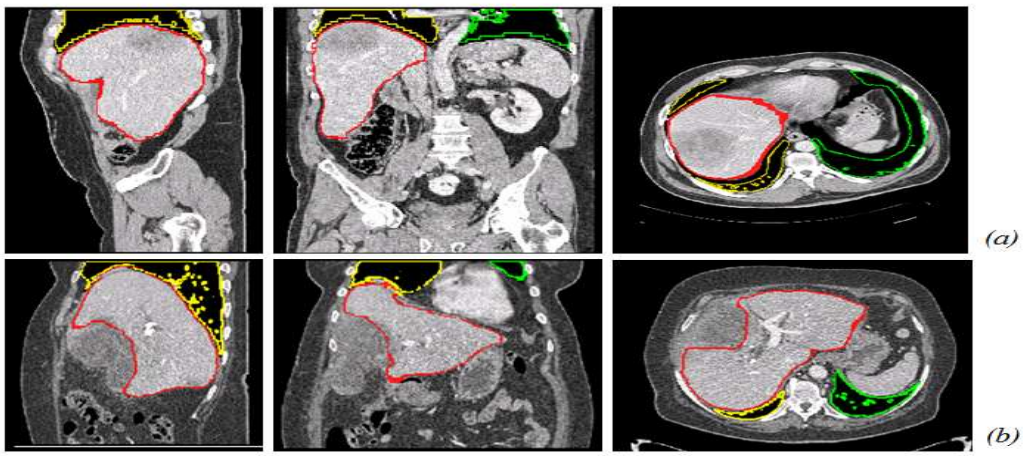

Figure 8: Large tumors in the boundary of the livers from the test database. From left to right, a sagittal, coronal and axial slice for the two worst cases, which give the worst scores. The outline of the ground-truth segmentation is in blue, the outline of the segmentation of the method described in this paper is in red. It is also illustrated the outline of the lung segmentations in green (left lung) and yellow (right lung).

registrations and atlas selection based on ROIs. Furthermore, our approach shares common elements between the proposed probabilistic atlas and the multi-atlas segmentation: a) the spatial normalization and b) the segmentation method. The spatial normalization is used both in building the probabilistic atlas and for obtaining the atlas selection for each ROI. The segmentation method is based on standard CRF models allowing incorporate appearance and shape in a single unified manner.

Specifically, ROIs of the target image are obtained by using low level operations. A probabilistic atlas is built for each ROI by means of affine registrations based on geometric moments from labeled images. The label prior probability has been estimated by voting and the image likelihood is modeled by independent Gaussian distributions for each pixel and for each label. To overcome bias in the statistical parameters of the probabilistic atlas, they have been smoothed by means of spatial linear diffusion. Other statistical models can be used as in [23], where the image likelihood is approximated using a Parzen window estimator. An advantage of our approach is that there is only one probabilistic atlas and it is not necessary to recalculate the statistical parameters for each target image.

Given an initial solution in a ROI by using low level operations, an EM

framework is used for obtaining a coarse segmentation. In each iteration, a new segmentation is computed by minimizing an discrete energy function and this labeling is again aligned with the the probabilistic atlas by geometric moments belonging to labeled images. The energy function, which is applied for segmenting, combines the maximization of the posterior probability and the minimal area of the separation surfaces between object and background under an image-based Finsler metric. The resulting energy function has been globally minimized using graph cuts. Few iterations have been sufficient to converge this process. We have experimentally observed that the iterative method outperforms the classical approach of a single registration and obtaining the segmentation by means of the probabilistic atlas. We have also observed that the effect of spatial regularization the statistical parameters of the probabilistic atlas by linear diffusion causes that the process converges with a few iterations.

The computing time devoted to obtain the coarse segmentation is low due to affine registrations based on geometric moments of the labeled images and the min-cut/max flow algorithm.

The coarse segmentation is employed in the next step for defining the registering mask and for obtaining the rankings of atlases more similar to the target image for each region. Three aspects of multi-atlas approach have been analyzed: image registration, atlas selection and fusion methods. In image registration, we have proposed to align the selected atlases to the target image by geometrical moments using S_C followed by a local deformation in the regions of interest. For atlas selection, DICE coefficient has been used to rank the atlases. Different atlas selection methods have been compared. Although our atlas selection framework is the most similar to the optimum one and gives better segmentation results, the improvements are not statistically significant to the other atlas selections. However, our approach shows significant improvements in relation to the conventional framework for any number of fused atlases in liver segmentation. In label fusion, a standard CRF model is used for MAP inference.

Comparing segmentation results between different published methods is always difficult. The quality of the databases used for validation, the anatomical definition of the structure, the quality of expert segmentations, the populations studied and the different measures reported all make it difficult to compare results. With these caveats in mind, we compare our segmentation results with other approaches that used the same database, i.e. *http://www.sliver07.org/* [38]. Maklad et al [43] have used the blood vessel

information to segment the liver through the portal phase of an abdominal CT dataset. This semiautomatic method requires a small group of manual seeds. Their results are evaluated with an overall score of 85.7, which is ranked as the best on this public database. Peng et al [8] show another semi-automatic method, which is a level set based variational approach. Their model is not restricted by training data and can be applied to livers of any shape. They have reported a score of 80 ± 4. Rusk et al [3] have proposed a fully automatic method for liver segmentation. Their method is essentially an advanced region growing and performs an average 61 ± 21 of total score. Linguraru et al [44] present an automated segmentation of livers. An affine invariant shape parameterization is combined with a geodesic active contour and graph cuts. They have reported a score of 76 ± 6. Therefore, our approach is comparable to expert human performance and other recently published results for liver segmentation in CT image. The proposed method has an intermediate position among the best methods for automatic and semi-automatic segmentation. Our results provides high accuracy in automatic segmentation and the computational time depends on the level of accuracy requested by user.

On the other hand, the issue of liver segmentation has only a single ROI and does not show the full potential of the proposed method. New problems that require two or more regions of interest should be analyzed. We have applied the label fusion method to the hippocampal segmentation from magnetic resonance imaging [45]. It is partially shown that the proposed methods in this work are generic and could be incorporated to other applications.

References

[1] Heimann T, van Ginneken B, Styner M, Arzhaeva Y, Aurich V, Bauer C, et al. Comparison and evaluation of methods for liver segmentation from ct datasets. Medical Imaging, IEEE Transactions on 2009;28(8):1251–65.

[2] Campadelli P, Casiraghi E, Esposito A. Liver segmentation from computed tomography scans: A survey and a new algorithm. Artificial Intelligence in Medicine 2009;45(2-3):185–96.

[3] Ruskó L, Bekes G, Fidrich M. Automatic segmentation of the liver from multi-and single-phase contrast-enhanced ct images. Medical Image Analysis 2009;13(6):871–82.

[4] Kumar S, Moni R, Rajeesh J. Automatic liver and lesion segmentation: a primary step in diagnosis of liver diseases. Signal, Image and Video Processing 2013;7(1):163–72.

[5] Zhang X, Tian J, Deng K, Wu Y, Li X. Automatic liver segmentation using a statistical shape model with optimal surface detection. Biomedical Engineering, IEEE Transactions on 2010;57(10):2622–6.

[6] Lu J, Shi L, Deng M, Yu SCH, Heng PA. An interactive approach to liver segmentation in ct based on deformable model integrated with attractor force. In: Machine Learning and Cybernetics (ICMLC), 2011 International Conference on; vol. 4. IEEE; 2011, p. 1660–5.

[7] Platero C, Tobar MC, Sanguino J, Poncela JM, Velasco O. Level set segmentation with shape and appearance models using affine moment descriptors. In: Pattern Recognition and Image Analysis. Springer; 2011, p. 109–16.

[8] Peng J, Dong F, Chen Y, Kong D. A region-appearance-based adaptive variational model for 3d liver segmentation. Medical physics 2014;41(4):435–02.

[9] Aljabar P, Heckemann R, Hammers A, Hajnal J, Rueckert D. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. Neuroimage 2009;46(3):726–38.

[10] Shi F, Yap PT, Fan Y, Gilmore JH, Lin W, Shen D. Construction of multi-region-multi-reference atlases for neonatal brain mri segmentation. Neuroimage 2010;51(2):684–93.

[11] Rohlfing T, Brandt R, Menzel R, Maurer Jr CR. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. Neuroimage 2004;21(4):1428–42.

[12] Rohlfing T, Brandt R, Menzel R, Russakoff D, Maurer C. Quo vadis, atlas-based segmentation? Handbook of Biomedical Image Analysis 2005;:435–86.

[13] Heckemann R, Hajnal J, Aljabar P, Rueckert D, Hammers A. Automatic anatomical brain mri segmentation combining label propagation and decision fusion. Neuroimage 2006;33(1):115–26.

[14] Thomas Yeo B, Sabuncu M, Desikan R, Fischl B, Golland P. Effects of registration regularization and atlas sharpness on segmentation accuracy. Medical Image Analysis 2008;12(5):603–15.

[15] Joshi S, Davis B, Jomier M, Gerig G. Unbiased diffeomorphic atlas construction for computational anatomy. Neuroimage 2004;23:S151–60.

[16] Wang Q, Seghers D, DAgostino E, Maes F, Vandermeulen D, Suetens P, et al. Construction and validation of mean shape atlas templates for atlas-based brain image segmentation. In: Information Processing in Medical Imaging. Springer; 2005, p. 689–700.

[17] Park H, Bland P, Meyer C. Construction of an abdominal probabilistic atlas and its application in segmentation. Medical Imaging, IEEE Transactions on 2003;22(4):483–92.

[18] Fischl B, Salat D, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 2002;33(3):341–55.

[19] Ashburner J, Friston K. Unified segmentation. Neuroimage 2005;26(3):839–51.

[20] Pohl K, Fisher J, Grimson W, Kikinis R, Wells W. A bayesian model for joint segmentation and registration. Neuroimage 2006;31(1):228–39.

[21] Han X, Fischl B. Atlas renormalization for improved brain MR image segmentation across scanner platforms. Medical Imaging, IEEE Transactions on 2007;26(4):479–86.

[22] Warfield S, Zou K, Wells W. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. Medical Imaging, IEEE Transactions on 2004;23(7):903–21.

[23] van der Lijn F, den Heijer T, Breteler M, Niessen W. Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. NeuroImage 2008;43(4):708–20.

[24] Platero C, Sanguino J, Velasco O. Nonlinear diffusion filters without parameters for image segmentation. In: Computer Analysis of Images and Patterns, CAIP 2009; vol. 5702 of *LNCS*. Springer; 2009, p. 517–24.

[25] Lafferty J, McCallum A, Pereira FC. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. ICML 2001;:282–9.

[26] Song Z, Tustison N, Avants B, Gee J. Integrated graph cuts for brain mri segmentation. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2006 2006;:831–8.

[27] Wolz R, Heckemann RA, Aljabar P, Hajnal JV, Hammers A, Lötjönen J, et al. Measurement of hippocampal atrophy using 4d graph-cut segmentation: Application to adni. Neuroimage 2010;52(1):109–18.

[28] Kolmogorov V, Boykov Y. What metrics can be approximated by geo-cuts, or global optimization of length/area and flux. In: Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on; vol. 1. 2005, p. 564–71.

[29] Koenderink J. The structure of images. Biological cybernetics 1984;50(5):363–70.

[30] Boykov Y, Kolmogorov V. Computing geodesics and minimal surfaces via graph cuts. In: Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on. 2003, p. 26–33.

[31] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Transactions on Pattern Analysis and Machine Intelligence 2004;26(9):1124–37.

[32] van Rikxoort E, Arzhaeva Y, van Ginneken B. Automatic segmentation of the liver in computed tomography scans with voxel classification and atlas matching. In: MICCAI 2007 workshop proceedings: 3D segmentation in the clinic. Citeseer; 2007, p. 101–8.

[33] Pei S, Lin C. Image normalization for pattern recognition. Image and Vision computing 1995;13(10):711–23.

[34] Dempster AP, Laird NM, Rubin DB, et al. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal statistical Society 1977;39(1):1–38.

[35] Lotjonen J, Wolz R, Koikkalainen J, Thurfjell L, Waldemar G, Soininen H, et al. Fast and robust multi-atlas segmentation of brain magnetic resonance images. Neuroimage 2010;49(3):2352–65.

[36] Dice L. Measures of the amount of ecologic association between species. Ecology 1945;26(3):297–302.

[37] Klein S, Staring M, Pluim J. Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines. Image Processing, IEEE Transactions on 2007;16(12):2879–90.

[38] van Ginneken B, Heimann T, Styner M. 3D segmentation in the clinic: A grand challenge. In: 3D Segmentation in the Clinic: A Grand Challenge, MICCAI Workshop, Proceedings. 2007, p. 7–15.

[39] Park H, Bland P, Hero A, Meyer C. Least biased target selection in probabilistic atlas construction. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2005 2005;:419–26.

[40] Klein S, Staring M, Murphy K, Viergever M, Pluim J. Elastix: a toolbox for intensity-based medical image registration. Medical imaging, IEEE transactions on 2010;29(1).

[41] Rueckert D, Sonoda L, Hayes C, Hill D, Leach M, Hawkes D. Nonrigid registration using free-form deformations: application to breast MR images. Medical Imaging, IEEE Transactions on 1999;18(8):712–21.

[42] Thévenaz P, Unser M. Optimization of mutual information for multiresolution image registration. Image Processing, IEEE Transactions on 2000;9(12):2083–99.

[43] Maklad AS, Matsuhiro M, Suzuki H, Kawata Y, Niki N, Satake M, et al. Blood vessel-based liver segmentation using the portal phase of an abdominal ct dataset. Medical physics 2013;40(11):113–501.

[44] Linguraru MG, Richbourg WJ, Watt JM, Pamulapati V, Summers RM. Liver and tumor segmentation and analysis from ct of diseased patients via a generic affine invariant shape parameterization and graph cuts. In: Abdominal Imaging. Computational and Clinical Applications. Springer; 2012, p. 198–206.

[45] Platero C, Tobar M, Sanguino J, Velasco O. A new label fusion method using graph cuts: Application to hippocampus segmentation. In: XIII Mediterranean Conference on Medical and Biological Engineering and Computing 2013. Springer; 2014, p. 174–7.