A Model Based Validation Scheme for Organ Segmentation in CT Scan Volumes

Hossein Badakhshannoory, *Student Member, IEEE,* and Parvaneh Saeedi, *Member, IEEE*

***Abstract*—In this work, we propose a novel approach for accurate 3D organ segmentation in the CT scan volumes. Instead of using the organ's prior information directly in the segmentation process, here we utilize the knowledge of the organ to validate a large number of potential segmentation outcomes that are generated by a generic segmentation process. For this, an organ space is generated based on PCA approach using which the fidelity of each segment to the organ is measured. We detail applications of the proposed method for 3D segmentation of human kidney and liver in CT scan volumes. For evaluation, the public database of MICCAI 2007 grand challenge workshop has been incorporated. Implementation results show an average Dice similarity measure of 0.90 for segmentation of the kidney. For the liver segmentation, the proposed algorithm achieves an average volume overlap error of 8.7% and an average surface distance of 1.51 mm.**

***Index Terms*—Model based segmentation, statistical model generation, principal component analysis, model based validation**

I. Introduction

An essential part of any computer-aided surgery is planning prior to the surgery. Planning often involves preparing a patient specific 3D model of the organ under surgery. The 3D model is generated by segmenting the organ from a set of medical images (bundled into a volume) acquired from different modalities such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI). Medical image volumes generated by various sensors at different imaging conditions could be affected by conditions such as non-uniform intensity distribution (both inside a single slice and across the volume) and noise. Therefore, techniques for medical image segmentation that rely only on low-level information are highly dependent on the setting of parameters. Given the intricate anatomical structures, it is more logical to use prior knowledge about the organ of interest than merely relying on low-level image content. In this sense, probabilistic model based segmentation techniques have been proposed to incorporate statistical knowledge of an organ as *a priori* to identify the organ of interest in an image. In this work, we propose a novel approach that incorporates statistical information as a mean of outcome quality measurement. Using a large number of segmentation outcomes and a statistical model hypothesis with the highest fidelity to the organ is identified.

A. Previous Work

Previous work for automatic 2D and 3D segmentation of organs in medical applications of image processing can be divided into non-model and model based. Many of these methods are proposed for extracting kidney and liver in medical images. Non-model based approaches generally rely on local information such as texture, intensity, spatial correlation of the 2D organ image in consecutive slices, and the location of the organ in the abdominal area with respect to neighboring structures such as spine and ribs.

Susomboon *et al.* [1] employed texture features to perform region classification for extracting liver's soft tissue. While this method utilizes several types of features to represent the liver's texture in the CT images, it fails to produce good results on MICCAI's 2007 grand challenge database [2][3]. This is perhaps due to the fact that in this approach there is no organ shape model to compensate for variations of the texture and intensity levels in different slices. Seo *et al.* [4] proposed a multi-modal threshold method based on piecewise linear interpolation that used spine location as a reference point. Forouzan *et al.* [5] reported a multi-layer threshold technique using thresholds that were obtained by statistical analysis of the liver intensity. While both these methods take advantage of the liver's relative position to the spine and ribs (easily identifiable in medical images) they are also threshold dependent. Determining thresholds that perform robustly and consistently for variant imaging conditions and different intensity values (such as those in the MICCAI's data set) is very difficult if not impossible. Pan and Dawant [6] incorporated a geometrical level set method for automatic segmentation of the liver in abdominal CT scans that did not rely on prior knowledge of shape or size. Although this method relies on a model based technique that outperforms threshold based techniques, it did not use prior knowledge of the liver shape. As a result, it performs relatively poorly in the cases where the liver includes tumors. Moreover, this method is verified only on a very limited number of datasets. In [7] a method for liver extraction was proposed that combined non-rigid registration and a multi-layer segmentation technique to identify liver regions based on it's boundary edges. Since this method does not rely on any shape model, it will not be affected by the diversity of the existing liver shapes in the training set. For the same reason though, the presented results are not comparable with those of methods using shape models. Lin *et al.* [8] segmented the kidney based on an adaptive region growing and an elliptical kidney region positioning that used spine as landmark. While they achieve good results based on Dice measure, they fail to compute and/or report comparable measures similar to those of MICCAI's grand challenge workshop. Also since their method is based on region growing technique, it could be sensitive to the lighting conditions (as such conditions exist

H. Badakhshannoory and P. Saeedi are with the School of Engineering Science, Simon Fraser University, Burnaby, BC, V5A1S6 Canada e-mail: hba14@sfu.ca, psaeedi@sfu.ca.

in the MICCAI's dataset volumes). Wu and Sun reported a deformable contour based method using texture classification and shape to identify kidney in ultrasound images [9]. The proposed work was tested and verified only on ultrasound images. Ultrasound images are essentially very different from CT and MRI images which contain other similar looking structures and organs in close vicinity of the kidneys.

Non-model based methods for organ segmentation are subject to inaccuracies due to variation in imaging condition, presence of tumor inside the organ and noise. Relying on texture and image values could cause inaccuracies in the segmentation process as such features could change from one patient to another. Moreover, most of these methods are parameter dependent and for the best performance often these parameters require adjustment from one CT volume to another.

In recent years, model based image segmentation algorithms have been developed for various medical applications. These methods aim to recover an organ based on statistical information. State of the art algorithms on model based segmentation are based on active shape and appearance models [10][11]. Active shape and appearance models incorporate statistical or parametric shape models (from training data) in the actual segmentation process.

Cootes *et al.* [10] described point distribution models to construct statistical models of general shapes from a training set. They exploited a linear formulation of the shape models named Active Shape Models (ASM) to iteratively search for specific shapes in the image. In the work proposed here, a similar method to [10] is utilized by applying a PCA-based (Principal Component Analysis) model on a set of training data to generate a space (called organ space) that represents the object of interest. Here however, we propose a novel approach for incorporating the knowledge of the organ space for segmentation purposes. Kainmuller *et al.* [12] used ASM in combination with a model of the typical intensity distribution around the liver boundary and neighboring structures. The method in [13] incorporated statistical shape model in combination with an evolutionary algorithm to provide initialization for a deformable mesh that searches for human liver boundaries. Both of the two previous methods performed a search strategy similar to ASM with an improved initialization step. Ling *et al.* [14] expanded ASM by two new learning techniques, marginal space learning and steerable features accompanied with a new shape space initialization. They improved the original ASM by handling highly heterogeneous texture patterns. Wimmer proposed a new ASM [15] which was entirely built upon non-parametric estimates of probabilities. Spiegel *et al.* [16] proposed to apply non-rigid registration to remove the problem of establishing point correspondence among training data samples in ASM. Ali *et al.* [17] estimated kidney shape variations using a distance probabilistic model to approximate densities of kidney and its surroundings using Poisson distribution. Algorithms proposed in [12]-[17] are all variations of the original method by Cootes *et al.* [10][11] where the shape model of the organ was utilized as a priori to lead the segmentation algorithms to better results. All of these methods suggest improvements over the original approach either by improving the performance time or simplification of the required steps for constructing the model. Here, we present a different approach for utilizing the ASM that uses the organ space as a selection tool to automatically choose segmentation parameters that lead to segmentation results with highest fidelity to the model.

The main advantage of the model based techniques is in their accuracy and robustness. They are also capable of compensating for the missing image features via interpolation. The performances of these methods however are dependent on the number and the type of the training data. Also if the shape to be segmented lies too far away from the model space, it might not be reachable even by those better methods that incorporate free-form deformation and statistical model based approaches. Therefore for some test volumes and organs, due to diversity in shape and size, their performance could vary. Also, all of the model based organ segmentation techniques proposed up to now have utilized statistical information of the organs as *a priori* to their main segmentation algorithm.

In this paper, we propose a novel alternative approach for utilizing statistical model information in an identification phase that is devised to choose the best segmentation candidates according to its distance from an organ's model space. Here, first a general segmentation algorithm that covers all segmentation results from under to over segmentation is used. After generating segmentation hypotheses, a statistical model, based on PCA, is used to generate an organ space. The distance of each candidate from the organ space is measured to choose the closest candidate as the best segmentation result. The main contributions of this work are as follows:

1) Introducing a novel approach to validate all potential unbiased segmentation outcomes according to the statistical characteristics of the organ, instead of leading the segmentation process towards results that are biased by the prior information.
2) A dynamic scheme that enables variation of segmentation quality control parameters for each slice of a CT volume independently and according to the fidelity of the generated outcome for that slice with the model (organ space).

These two contributions take advantage of both non-model and model based approaches to segment organs in CT scan volumes accurately.

II. Proposed Method

The proposed work in this paper is a model based general algorithm that instead of using the model information to direct the segmentation algorithm towards segmenting an organ in an image, uses the model information to choose a segment (among all possible segmentation outcomes) with the highest fidelity to the organ. The data flow diagram for this work is depicted in Fig. 1.

The algorithm includes two phases of training and testing. In the training phase, using the ground truth provided by the training dataset, an eigen space is constructed that describes the organ (this space is referred to as the organ space). The organ space is a group of eigenvectors that are generated by applying PCA over the ground truth of the training dataset.

In the testing phase, unknown CT volumes are processed to segment the organ of interest.

The testing phase includes four main procedures:

1) Pre-processing: This procedure includes an alignment process in which the orientation of the CT volumes are automatically estimated and corrected to a reference orientation. It is also responsible for establishing a starting point for the volume under process. This starting point relies on unique characteristics of the organ of interest and the general knowledge of the organ that occurs at a specific location within the volume.
2) Candidate Hypotheses Generation: This procedure segments each image slice into a large number of overlapping regions among which one region could potentially be the true region corresponding to the organ of interest on that specific slice.
3) Candidate Hypotheses Refinement: This procedure reduces the search space (candidate hypotheses) by constraints that incorporate relationships between organ's regions in consecutive slices of the volumes.
4) Best Hypothesis Selection: This procedure measures the quality of all generated candidate hypotheses (after step 3) according to the organ space generated in the training phase and identifies the best candidate.

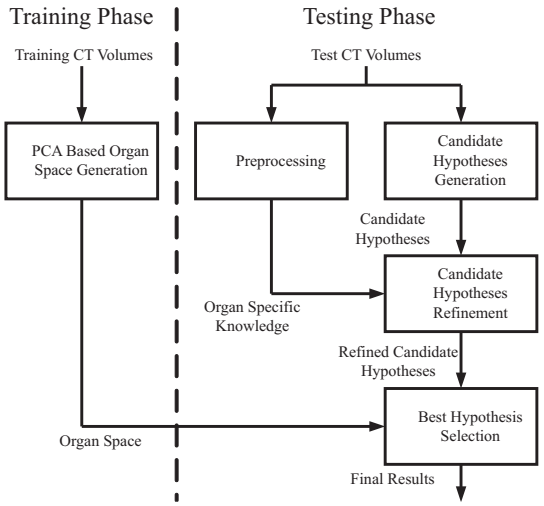


Fig. 1. Flow diagram of the proposed method

In the following subsections, details of the training and testing procedures are described.

A. Training Phase: PCA Based Model Generation

The objective of this section is to create a vector space that represents an organ. Such space is a model that encapsulates the variations of the organ according to the training set volumes. In general an organ space can be created using boundaries, masks or image regions of the organ of interest in CT images. Such information is used to either create masks for the organ or extract the cropped image of the organ from the CT scan data to create the organ space. The space can be created for 2D images/masks of the organ or 3D volumes of images/masks of the organ. Therefore the input data to this module could be 2D or 3D vector data of organ's mask or cropped image.

In this work, we have used different sets of information for the applications of the proposed method. In Section III, 3D kidney organ reconstruction, we have incorporated the cropped images of the kidney region in each training dataset volume. In Section IV we present the application of the proposed method for construction of 3D liver models. For this application, the 2D liver masks from the training data sets are used for creating the organ space. Fig. 2 shows image slice of abdominal area with its 2D mask and the extracted liver's image region. Fig. 3 shows how 2D extracted masks of the liver in the CT slices are piled up together to represent the 3D liver.

Fig. 2. Image of the abdominal area with its corresponding liver.

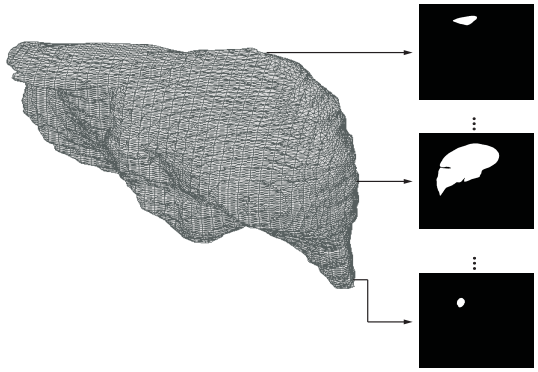


Fig. 3. A 3D liver volume and it's corresponding 2D liver masks.

One approach to create an organ space that encodes variations of an organ in a series of training data is to find the principal components of the distribution. This is equivalent to computing the eigenvectors of the covariance matrix of the set of vector data. Each input vector contributes more or less to each eigenvector. Each eigenvector looks like a ghostly organ mask or image depending on the input type. The eigenvectors form a space that represents the organ where each new organ vector data can be approximated using a linear combination of these eigenvectors. Here, without loss of generality we explain the generation of the model for a training set of 3D input vectors. Note that these vectors can contain either image pixels or binary mask values.

All the volumes in the training set of MICCAI's 2007 grand challenge workshop have the same size in sagittal (X) and coronal (Y) directions (512×512 pixels) but the sampling in the transverse (Z) direction is not the same and therefore each volume has a different number of slices. Also the thickness of each volume is different and it varies from $138mm$ to $213mm$ with a mean of $175mm$. In this work, all the volumes used

in the training set for generating the PCA based model are re-sampled in the transverse direction to the same number of slices (100 for each volume). Because the size of an organ is different in different patients, and also the PCA input vector requires a fixed number of elements, the organ mask volume is re-sampled to 100 slices in the z direction using the nearest-neighbor scheme. Clearly, re-sampling volume masks into a fixed number of samples creates slices with different thicknesses (between different volumes). However, we found that within the range of organ sizes (in both training and test datasets) the PCA sufficiently encapsulates the overall variations of the organ to represent general characteristics of that organ.

We assume that each 3D input vector in the training set is a 3D array of X by Y by Z or equivalently a 1D vector of size $X{\cdot}Y{\cdot}Z$. If L_1, L_2,..., L_n are 1D vectors of the training set and ψ is their mean, the distance of each vector from its mean is defined by ϕ_i = L_i - ψ. Here, we look for the set of n orthonormal vectors u_i that best describe the distribution of the vector data. These vectors are the result of applying PCA over the entire training set. They are eigenvectors of the following covariance matrix:

$$C = \frac{1}{n}\sum_{i=1}^{n}\phi_i\phi_i^T = AA^T \tag{1}$$

where A = [ϕ_1 ϕ_2 ... ϕ_n]. Usually matrix C is large and computing its eigenvectors is exhaustive. [20] introduced a computationally effective way to estimate these vectors. The number of data points in the organ space is n, therefore only $n-1$ meaningful eigenvectors exist and the remaining eigenvectors will have associated eigenvalues of zero. Due to this, the eigenvector equation can be solved for an n by n matrix instead of a $X{\cdot}Y{\cdot}Z$ by $X{\cdot}Y{\cdot}Z$ matrix. Consider the eigenvectors v_i of A^TA such that:

$$A^TAv_i = \mu_i v_i \tag{2}$$

$$AA^TAv_i = \mu_i Av_i \tag{3}$$

Equation 3 shows that Av_i is an eigenvector of AA^T. Using this, instead of calculating the eigenvectors of AA^T, the eigenvectors of A^TA are calculated and are called v_i. Linear combination of v_i and the n training set vectors (ϕ_i) is used to form the eigenvectors that represent the organ space and are called u_i. This is done by:

$$u_i = \frac{1}{n}\sum_{j=1}^{n}v_{ij}\phi_j \tag{4}$$

Once eigenvectors are approximated, they are used to model the organ. This model is used (as explained in the Section II-B4) to measure the similarity of segmented regions (candidate hypotheses) to the organ.

B. Testing Phase

This phase is responsible for processing unknown CT scan volumes to identify regions of the organ of interest. This phase includes four main processes that are explained next.

1) Pre-processing: Most interactive/semi-automatic segmentation algorithms [5][7] require a starting point (on the organ of interest) to be manually identified. This starting point is a crucial element with high impact on the accuracy of the segmentation results. Here the proposed work is fully automatic and therefore does not require identifying such starting point via manual interaction. To initially localize points that belong to the organ of interest, we propose a pre-processing procedure. This procedure utilizes specific and unique physical attributes of the organ and its environment to identify regions of CT volume images (in one slice or more) that belong to that organ. Clearly this procedure must be customized for different organs. We discuss in details examples of such procedure in the applications of the proposed method for kidney and liver 3D modeling.

2) Candidate Hypotheses Generation: For creating potential segment hypotheses, any common segmentation algorithm may be utilized. The basic idea is to create a set of variant segmentation parameters (algorithm dependent) that drives the segmentation results from under segmentation to over segmentation. For instance, let's assume that a segmentation algorithm Seg depends on N parameters p_1,..., p_N. If the lower and upper boundaries of these variables, presented by (v_{1l}, v_{1u}),...,(v_{Nl}, v_{Nu}), drive the output results from under to over segmentation, a combination of values for $p_1 \in [v_{1l}$, v_{1l} + Δl_1,...,v_{1u}],...,$p_N \in [v_{Nl}$ + Δl_N,..., $v_{Nu}]$ will be generated to segment the input image multiple times. Every unique combination of p_1,..., p_N generates one set of segmentation results. The number of overall output segments depends on Δl_i where $i \in 1,...,N$.

3) Candidate Hypotheses Refinement: The purpose of this process is to refine potential candidates by utilizing information that was identified in the pre-processing phase. In general, this phase could rely on organ specific information to remove outliers. It could also involve further analysis (based on local or global features) to verify the fidelity of a hypothesis to the organ.

4) Candidate Selection Based on Similarity to the Organ Space: Similar to the approach proposed by [20], where the face space was used for face detection, the organ space can be used to measure the similarity of a candidate hypothesis to the organ. Each candidate hypothesis is a vector data, same as the training dataset that was used to create the organ space. For measuring the similarity of each candidate hypothesis vector (L_{hypo}) with the organ, first, the mean adjusted hypothesis vector ϕ_{hypo} is projected onto the organ space using:

$$\phi_{hypo} = L_{hypo} - \psi \tag{5}$$

$$\eta_i = u_i\phi_{hypo} \tag{6}$$

The result of this projection is a vector of [η_1,η_2,...,η_n]. Each η_i represents the contribution of an eigenvector u_i in the reconstruction of projected candidate hypothesis. The reconstructed candidate hypothesis is then generated by:

$$\phi_{rec} = \sum_{i=1}^{n}\eta_i u_i \tag{7}$$

For each refined candidate hypothesis, generated in Section II-B3, the Euclidean distance between the mean adjusted volume (ϕ_{hypo}) and the reconstructed version (ϕ_{rec}) is computed by:

$$E = \| \phi_{rec} - \phi_{hypo} \| \tag{8}$$

The candidate hypothesis with the smallest Euclidean distance from its reconstructed version in the organ space is chosen as the best segmentation result.

III. Application of Proposed Method for Kidney Segmentation

In this section, the application of the proposed method for segmentation of the 3D right kidney in CT scan volumes is presented. We follow the same organization as in Section II for presenting processes utilized in this application.

A. Training Phase: PCA based Model Generation

For this application, a kidney space is generated for the 2D cropped images of the kidney from the training set of MICCAI's grand challenge workshop [2][3]. It is worth to mention that different volumes of the data have different thicknesses. They usually range between $0.7mm$ to $5mm$ in the Z direction (transverse). In order to establish correct correspondences between different volumes, all the volumes are re-sampled to the same number of slices (100) in the Z direction. First the CT volume is divided into several sections in the transverse direction with each section containing the same number of slices in them (5 slices in this work). The kidney is then manually extracted for all slices and translated into the center of a blank image with the same size as the image slice. The kidney images of each section, from the training set of MICCAI's CT volumes, are used to train the kidney space for that section of slices. Here, shape, texture, intensity, and size of kidney regions are more similar to each other for closer slices. Therefore, multiple kidney spaces are generated for different transverse sections of CT volumes along the transverse direction. The process of assigning different slices to different sections is shown in Fig 4.

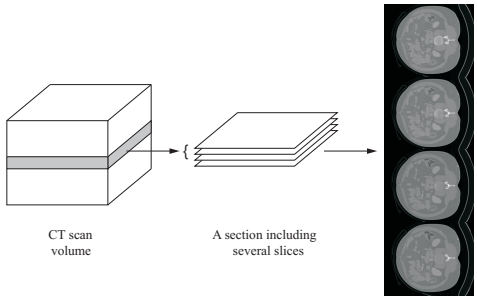


Fig. 4. Different sections of CT scan volumes used for kidney spaces.

B. Testing Phase: Pre-processing

The pre-processing phase includes two main automatic processes: aligning the dataset volumes, and identifying those slices of each volume that indeed include the organ of interest and finding an initial starting point within the organ.

1) Aligning volumes: In order to identify the orientation of the dataset volume and to correct for potentially various orientations the following algorithm is implemented:

i) The area associated with the spine (refer to as spine mask) in the physical middle slice of each test volume in the transverse direction is extracted according to the described procedure in Section III-B2.ii.a. The center of gravity of this mask is also computed.

ii) The lengths of all line segments passing through the center of gravity and limited to the boundaries of the mask ([-90:1:90] degrees) are computed.

iii) The rotation angle corresponding to longest line segment is found. The image is then rotated back by that angle. After the rotation the longest axis of the spine mask becomes parallel with image horizontal axis. To make sure that the spine is always rotated in the correct direction (we want the spine to be on the right side of the image) the distance of the spine mask's center of gravity from image right border is used. Once the rotation angle and the rotation direction are estimated, all slices of the volume are rotated, Fig. 5.

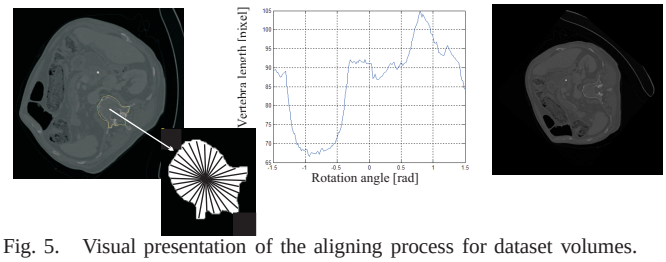


Fig. 5. Visual presentation of the aligning process for dataset volumes.

2) Organ region identification: In this phase, first the upper and lower slices, corresponding to the beginning and end of the kidney in the CT volume, are automatically identified. Once these slices are identified, their middle slice is processed to find the kidney region in it. This section includes two main processes.

i) A fully automated organ reconstruction system requires automatic identification of slices that include that organ within each volume. Such algorithm varies from organ to organ as characteristics of various organs differ. In this section, we introduce a method for identification of the lower and upper boundary slices of the CT volume that contain the kidney (in the transverse direction). This algorithm includes three steps:

a) In the first step, each slice is thresholded in the transverse direction using a dynamic threshold that is determined for each slice by:

$$TR = k \times mean(I) \tag{9}$$

here k is set to 1.35 (a constant value for the entire dataset volumes, found empirically) and I is the image region corresponding the abdominal area of each slice (estimated as described in Section III-B2.ii.b), Fig. 6.

b) Next we identify a slice that includes the kidney (starting slice). Based on the observation from the

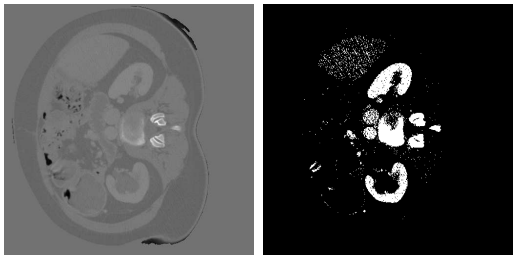


Fig. 6. Left is the slice from transverse direction containing only the abdominal area and right is the thresholded image.

training datasets, usually slices that fall between the bottom slice of the liver and the slice corresponding to the largest 2D liver surface (in the transverse direction) contain a slice that corresponds to a large 2D segment of kidney. Identification of slices containing the largest 2D surface and the bottom slice of the liver are described in section IV-B. Therefore, the slice corresponding to the midpoint between these two slices is chosen as the starting slice of each volume. Using the kidney subregion area, as described in section III-B2.ii.c, a kidney mask guesstimate is extracted for the starting slice. The thresholded image from step III-B2.i.a is then utilized to identify the segment with largest area in kidney subregion.

c) Finally, the algorithm propagates the kidney mask guesstimates in two directions from the starting slice (in the transverse direction). The propagation is performed by choosing the segment with the largest area overlap with the kidney mask guesstimate of the previous slice. This process continues in both directions until the mask guesstimates reach to zero at the two ends.

ii) In this section, the middle slice of the kidney slices is processed to estimate the approximate location of the kidney. The relative location of the kidney with respect to the spine is used for this purpose. This procedure includes the following steps:

a) In CT scan images of the abdominal area, pixels corresponding to bones (spine and ribs) appear substantially brighter than pixels of soft tissue regions. Using intensity based thresholding approach (similar to [8]), the middle slice is thresholded. The threshold value is found empirically based on the training dataset. For CT scan images with maximum and minimum pixel values of 1024 and -1024, this threshold was set to 400. To make the process more robust, 10 slices before and 10 slices after the middle slice are also thresholded. The thresholded results of all 21 images are unionized. The resulting image is a binary mask that holds values of 1 for spine and rib regions and 0 otherwise. In this process, some soft tissues with brighter intensities (Fig. 7.a) could be wrongly thresholded and therefore will be added to the mask. These points are usually sparse and filtered out using a morphological operations that first fills the holes and then removes rib regions (connected to the spine) using a disk structuring element (SE) of radius 4 pixels (Fig. 7.b). The largest connected piece in the resultant image is then identified as spine, Fig. 7.c.

b) Now, we extract the boundary of abdominal area. For this, first the middle slice is enhanced using a Log transform. Canny edge detector is then applied. The NOT of the resulting edge map image is opened with a circular SE of radius 4. The largest piece of the resultant mask is selected. This region represents the entire abdominal area and its boundary is extracted to highlight the region in which the kidney is located (Figs. 7.d, e and f).

c) In the final step, we require to identify a subregion within the abdominal area that most likely contains the kidney or parts of it. Focusing on processing a smaller region of the image, reduces the chance of misidentification of the kidney which is a real problem given it's size and texture in comparison with its neighboring organs such as heart and gall bladder. Usually, the location of the kidney with respect to spine, lies on a line that creates a 70^o angle with the horizontal line passing from the center of the spine (Fig. 8.a). The 70^o is determined empirically. Therefore from the center (center of gravity) of the spine mask, a line is emitted at 70^o and its intersection with the boundary of the abdominal region is found (length of l). Center of the kidney is usually located at a distance of $0.3l$ from the center of spine. A circular region (radius of 50 pixels) that most likely withholds the kidney, is then centered at the location of $0.3l$, Fig. 8.b.

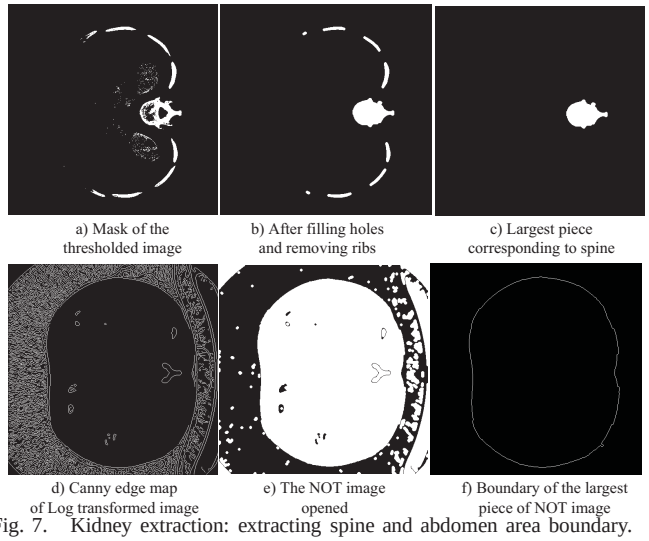
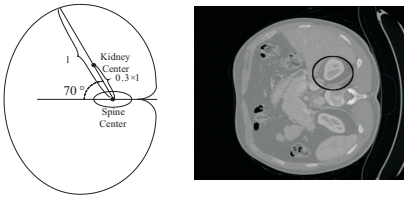


Fig. 7. Kidney extraction: extracting spine and abdomen area boundary.

C. Testing Phase: Candidate Hypotheses Generation

After the pre-processing, segmentation is performed on all images of the CT scan volume and all candidate segments



a) Kidney center calculation

b) Kidney area highlighted

Fig. 8. Kidney extraction pre-processing: extracting kidney location center.

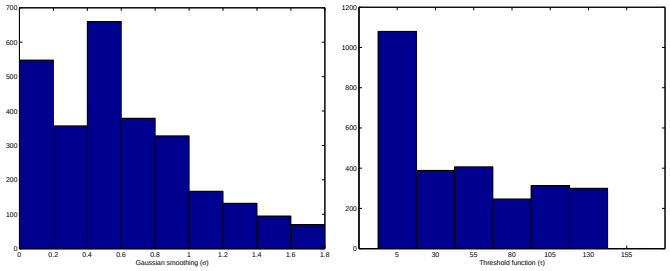


Fig. 9. Distributions of σ and τ values for the kidney.

are identified. These candidates are generated for every 2D slice of each volume of the test set using the graph based segmentation method [18]. Graph-based segmentation is a method capable of preserving details in low-variability image regions while ignoring details in high-variability areas. The algorithm includes following steps:

i) Each image pixel is considered as a region where it corresponds to a node $(v \in V)$ in the overall image graph of $G(V, E)$.

ii) Neighboring pixels are connected by undirected edges $(e \in E)$. For each edge a weight coefficient is computed according to the dissimilarities between pixels.

iii) Similar regions A and B are merged together to produce a larger region if the following condition is held:

$$Dif(A, B) \leq MInt(A, B) \tag{10}$$

Where

$$Dif(A, B) = \min_{v_i \in A, v_j \in B, (v_i, v_j) \in E} w((v_i, v_j)) \tag{11}$$

Here E is the graph edge set and $w((v_i, v_j))$ is the weight between vertex v_i and v_j.

$$MInt(A, B) = min(Int(A) + \tau(A), Int(B) + \tau(B))) \tag{12}$$

$$Int(A) = \max_{e \in MST(A,E)} w(e) \tag{13}$$

MST represents the Minimum Spanning Tree graph $G(V, E)$.

$$\tau(A) = \frac{k}{|A|} \tag{14}$$

The control parameters of this algorithm include: Gaussian smoothing σ, threshold function τ, and scale of observation k.

The range of values selected for σ and τ are found through statistical analysis of the data in the training sets. Initially, values of $\sigma \in [0.1 : 0.2 : 1.7]$ and $\tau \in [5 : 25 : 155]$ were used to segment each slice of every training dataset volume, $9 \times 7 = 63$ segment images for each slice. The overlaps of all segment images of a slice with the mask of that slice (from the ground truth) are computed and the σ and τ of the segment image with the highest overlap are collected. The histograms of the distribution of these two parameters were then estimated, Fig. 9.

Using these histograms, a range of $[0.1 : 0.2 : 1.1]$ for σ and a range of $[5 : 25 : 105]$ for τ are chosen. These ranges cover 90% of the best potential σ and τ according to the training dataset volumes. The observation scale of k proved to have minimal affects on the results and therefore was kept constant at 50.

At the end of this process, for each slice of the CT volume a large number of candidate segments are extracted. These candidates are refined in the next process.

D. Testing Phase: Candidate Hypotheses Refinement

This procedure begins from the middle slice. Using the kidney's potential location (found Section III-B2.i.c), all detected segments of the middle slice, are inspected and only those that fully or partially overlap with that circular region are kept. For the middle slice, all these candidates are passed to Section III-E where only one segment that represents the best kidney region for the middle slice is selected.

The process of hypotheses refinement for other CT slices is slightly different from the middle slice and it is based on the overlap of candidate hypotheses with the best candidate segment from the previous slice. Since the starting point for the processing is the middle slice, this procedure treats CT slices in two batches. Both batches start from the middle slice but move in opposite directions. In each slice, all candidate hypotheses that have overlaps smaller than 50% with the kidney segment (best chosen segment) of the previous slice will be removed.

At the end of this process, for each slice, a number of hypotheses are remained that location- and size-wise are the most probable representative of the kidney region for that slice.

E. Testing Phase: Best Hypothesis Selection

At this point, for each slice there are a number of potential kidney region hypotheses. These candidates are projected into their corresponding kidney space (from III-A) and the hypothesis with the smallest Euclidean distance (as described in Section II-B3) is chosen as the kidney region (kidney mask) of that slice. Once the best candidates for all slices are identified, they are stacked up together to generate a 3D model of the kidney. To refine this 3D model, a post processing based on morphological operation in the sagittal direction is performed. Here the mask slices of the kidney (in the sagittal direction) are first opened by a disk SE of radius 2 pixels and then closed by a disk structuring element of radius 4 pixels. This procedure fills out small holes inside the kidney mask and removes the excess small parts attached to it boundaries.

F. Results

Some segmentation results for different CT slices are shown in Fig. 10. In these results, the ground truth are shown with white solid lines and the contours found by the proposed method with black lines. The quantitative assessment of the algorithm based on the entire test set volumes is presented in Section V.

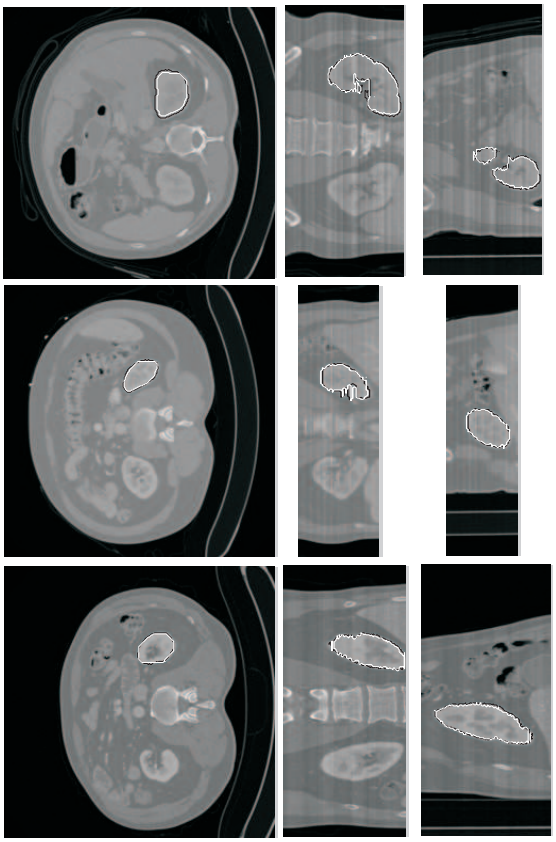


Fig. 10. Results of kidney segmentation in transverse (left), coronal (center) and sagittal (right) directions; black: ground truth, and white: our method).

IV. Application of Proposed Method for Liver Segmentation

In this section, the application of the proposed method for segmentation of 3D liver in CT scan volumes is presented. For this application, the organ space is generated from 3D mask volumes of the liver of MICCAI's 2007 grand challenge workshop training dataset of twenty patients.

Human liver has a soft triangular shape with four lobes that are different in shape and size. Slicing 3D CT volumes of liver's in the transverse direction usually results in images that include more than one liver piece. This adds to the complexity of candidate hypotheses and makes their identification process more complicated. Therefore, CT volumes of liver are processed in the coronal and sagittal directions in which the observed 2D liver regions consist of one single piece.

A. Training Phase: PCA based Model Generation

Similar to Section III-A, in this section a liver space is generated for 3D masks in the training dataset. Liver masks for 2D slices of CT training volumes are extracted and stacked up, Fig. 3. These 3D mask volumes are used to generate the liver space. This liver space is later used for assessment and identification of candidate segment hypotheses.

B. Testing Phase: Pre-processing

The pre-processing phase includes two main automatic processes: aligning the dataset volumes, and identifying those slices of each volume that indeed include the liver. The aligning of dataset volumes is performed by procedure presented in Section III-B1. Since the number of slices and the location of the liver in each volume varies, the following procedure is performed to identify those slices that include the liver. This procedure includes three steps:

i) In the first step, a number of 3D models of the liver are generated using segmentation results of the liver slices in the sagittal and coronal directions. Human's liver is the largest glandular organ of the body within the abdominal area. When looking at the CT volumes, the middle slice (regardless of the number of slices) always includes a segment from the liver. Also in the middle slice, the segment corresponding to the liver usually is the largest segment among all detected segments that correspond to other organs or areas. These two observations are utilized in this step. Therefore, starting from the middle slice of each test volume and moving in two directions, for both sagittal and coronal axes, we generate 21 3D liver models (by changing β in each case). Here each slice of each volume is segmented 21 times. The 3D model for each β is generated by moving from the middle slice in two directions and choosing the segment (for the results of current β) with the maximum overlap with the liver segment from the previous slice. Obviously, the qualities of these models are different and their number depends on the range of segmentation parameters used in the system (in our work 21 3D models).

ii) In this step, the system identifies the slice that most probably corresponds to largest liver region for the test volume (in the transverse direction). For each 3D model generated in step i, the slice (in the transverse direction) corresponding to the largest liver segment is identified. Once all such slices are identified for all the 3D models, a voting scheme is used to identify the most popular slice. We refer to this slice as the center slice although it generally does not correspond to the physical center of the liver. Once this center slice is identified, the next step is incorporated to estimate the upper and lower boundaries of the liver slices.

iii) In this step the liver's upper and the lower boundary slices in the transverse direction are identified as following:

   a) A number of liver mask candidates are generated for the center slice that was found in step ii. As described in Section II-B, using a range of segmentation parameters, a number of liver segments are generated (one segment for each set of parameter).

All these candidate segments are added together and thresholded. The threshold value is dynamically chosen so that the resultant segment has the highest similarity with the 2D mask extracted from the 3D training model (Section II-A). The 2D mask of the 3D training model is chosen according to the liver area (the largest).

b) The remaining slices in the transverse direction are treated in the same way but in two opposite directions. Therefore, each test volume slice image is segmented several times. The segmentation results are added together and thresholded using the threshold value found in iii.a. The segment with the largest overlap with the liver segment from the previous slice is considered as the liver segment for that slice.

c) This process is repeated until the liver mask area becomes zero in both directions. At that point the algorithm has reached to the boundary slices of that data set volume.

C. Testing Phase: Candidate Hypotheses Generation

To highlight the generality of the proposed algorithm, for the liver segmentation application a different segmentation algorithm (the Mean Shift segmentation method [19]) is utilized.

Mean shift segmentation is a non-parametric feature space originally introduced in [22]. In mean shift segmentation usually an attribute of the image (for example color or intensity) is chosen as the feature. Therefore, first a search window is chosen and is centered on an initial location. The mean location of the data in that window represents the new centroid. Therefore, the search window is migrated to a new position which is centered at the location of the centroid found from the previous step. The procedure is repeated until window has reached a local maximum in the density function and the movement of the window's centroid becomes negligible. When segmenting an image, search windows are uniformly positioned over the image data. The converged mean shift window for each initial position is found and windows with the same local maxima are merged together. Mean shift segmentation algorithm controls the quality of segments through three parameters of intensity (h_r) and spatial (h_s) resolutions, and minimum segment size. h_s affects the smoothing, and connectivity of the potential segments and h_r controls the number of segments. The minimum segment size is the area (in pixel) of the smallest individual segment.

In this work the minimum segment size was kept fixed at 20 pixels. The range of values selected for h_r and h_s are found through statistical analysis of the data in the training sets. Initially, values of $h_r \in [1:2:27]$ and $h_s \in [1:2:15]$ were used to segment each slice of every training volume. The overlap of all segmented images of a slice with the mask of that slice (from the ground truth) is estimated and the h_r and h_s of the segmented image with the highest overlap are collected. The histograms of the distribution for these two parameters are then estimated, Fig. 11. Using these histograms, the range of h_r and h_s are then set to $[5:2:23]$ and $[3:2:13]$. Using these two ranges, over 90% of the data in the training volumes will be segmented to their potential segments.

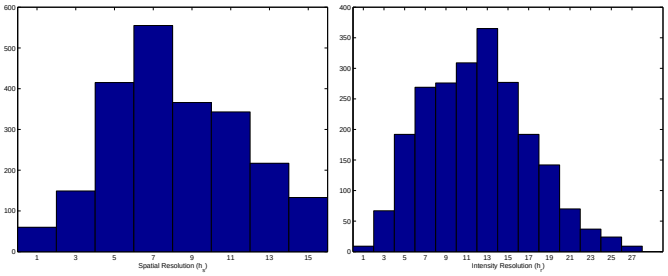


Fig. 11. Distributions of h_r and h_s values for the liver.

D. Testing Phase: Candidate Hypotheses Refinement

The process of candidate hypotheses generation starts from the middle slice of each CT volume in the coronal and sagittal directions where liver has its largest or near largest 2D surface.

Due to the large size and unique texture of the liver region, the candidate hypotheses generation process tends, occasionally, to cut the liver into smaller pieces. To ensure that such condition does not jeopardize the quality of the process, a post processing procedure is proposed in here. This procedure allows reduction of a large number of segments into a smaller number but more consistent set. It includes the following steps:

i) The boundary edges of segment regions (for each set of parameter combination) are extracted to form an edge map (EM).

ii) The EMs of each image (one EM for each set of parameters) are added together to form an accumulative edge map (AEM) for that input image (equation 15).

$$AEM = \sum_{s} EM \tag{15}$$

iii) The contrast of AEM is enhanced using the Log transform. The AEM image includes a few gray levels. The logarithmic correction in here expands the low level range of the AEM image. This allows utilizing the thresholding process at finer levels which leads to a larger number of segment candidates with more precise differences.

iv) Contrast enhanced AEM is then thresholded to form an Enhanced Edge Map (EEM) that includes isolated connected regions. This is described by equation 16:

$$EEM_{\beta}(x,y) = \begin{cases} 1 & Log(AEM(x,y)) > Thresh \\ 0 & otherwise \end{cases} \tag{16}$$

where $Thresh = max(max(Log(AEM)))/\beta$.

The threshold applied here ($Thresh$, the frequency of observation in the segmentation) represents the minimum strength of the boundaries of each segment hypothesis. Naturally, a constant $Thresh$ would not provide same quality results across all input images. Therefore, different values of β are utilized to generate a number of $Thresh$s and EEM images. A range of

[1.5,1.6,...,3.5] for β is used for the liver segmentation in this work. This range was found empirically by inspecting thirty liver CT scan volumes of different patients including a variety of shapes and imaging conditions.

At each slice, different EEMs are generated according to equation 16. If the processed slice is indeed the middle slice of the CT volume, the algorithm selects the segment (in EEM) with the largest area as the best segment representative of the liver. If the current slice is not the middle slice of the volume, the segment (in EEM) with the largest overlap with the liver's segment from the previous slice is considered as the liver segment. This segment is first morphologically opened by a circular SE (radius of 4 pixels) to remove any excess small parts around its boundary. Next a morphological hole filling process is applied to fill any small gaps within this segment. Sample EEM results with their detected liver segments at different βs are shown in Fig. 12.

After extracting all liver segments of a volume for a β value, these masks are stacked up together to form a candidate 3D liver mask volume. This implies that for each $Thresh$ (or β) value, one liver volume hypothesis is generated.

As mentioned earlier, the segmentation candidate generation for liver is performed in the sagittal and coronal directions. If segmentation results in any of those directions include an extra part attached to them, when viewing from transverse direction, such extra part manifest itself as a line of angle zero or 90 degrees. This is clearly under the condition that such an extra piece is not constantly found on consecutive slices. To filter out such those lines, first all detected segments are stacked together to create a 3D model. Then the model is sliced in the transverse direction and filtered using a morphological post processing with a line SE.

Fig. 13 shows three examples of 3D volumes generated for different values of β along with the associated 3D ground truth.

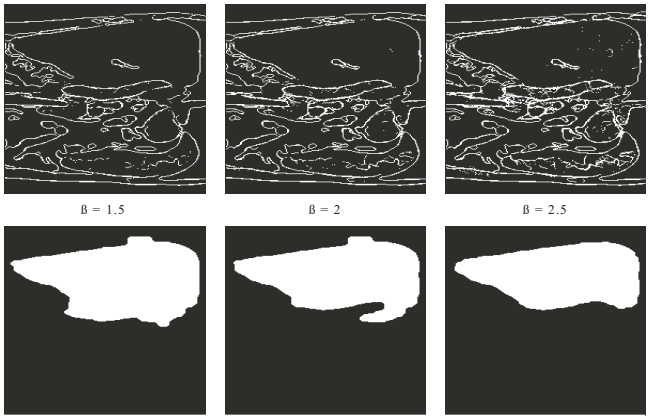


Fig. 12. Sample EEM for different β values (top row) with the corresponding extracted mask candidate (bottom row) for the liver extraction application.

E. Testing Phase: Best Hypotheses Selection

After generating all liver volume candidate hypotheses (including all candidate hypotheses for both sagittal and coronal directions) they are projected into the liver space according to equation 6. They are then reconstructed using equation 7 and the hypothesis with the smallest Euclidean distance from its reconstructed version is chosen as the best 3D reconstructed model of the liver for that CT volume. Note that here we have created two sets of models: one using the data in the sagittal direction and one using the coronal direction. The two sets of models are then measured against the PCA model and the model with the highest similarity will be chosen. Therefore the final selected 3D model could have been originated from either sagittal or coronal directions.

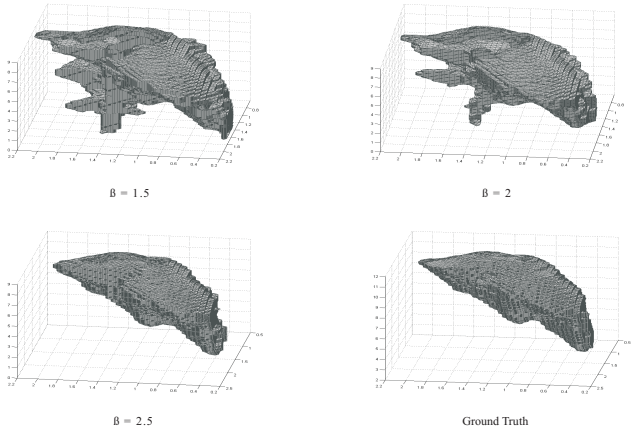


Fig. 13. Sample 3D volume mask representation for different β values in the liver extraction application.

F. Results

Fig. 14 shows some examples of the detected 2D liver mask boundaries along with their corresponding ground truth at transverse, coronal and sagittal directions.

In these results, the ground truth is shown by white lines and the results of the proposed method are shown with black. Quantitative results along with a comparison with the state of art are presented in the next section.

V. Test Results

This section presents complete results and discussions for both applications of the proposed method. In generating segmentation results, the dataset of MICCAI's 2007 grand challenge workshop is used. This dataset includes 20 training and 10 test CT scan volumes.

In the kidney case, we used 23 volumes (15 training and 8 test) provided for the MICCAI's grand challenge. The MICCAI's grand challenge was designed for the liver 3D modeling and therefore some of its volumes did not include the kidney organ completely. Also, the training volumes did not include the ground truth for the kidney. Therefore, the ground truth for both training and test volumes were found manually.

For the liver, in order to generate the organ space for the PCA based model of the liver, first all 20 training volumes were used. The ground truth (liver masks for the training volumes) was provided by the MICCAI's workshop. The ground truth for the test volumes however was not provided but the quality of the results was measured against MICCAI's ground truth by MICCAI's workshop. By exchanging the role

TABLE I
QUANTITATIVE RESULTS FOR THE RIGHT (FIRST ROW) AND LEFT (LAST ROW) KIDNEYS EXTRACTION.

	Training by 7 sets			Training by 10 sets			Training by 15 sets		
	Dice	MSE	SE	Dice	MSE	SE	Dice	MSE	SE
Mean (right kidney)	**0.8964**	**5131**	**0.8873**	**0.9066**	**4686**	**0.8965**	**0.9063**	**4415**	**0.8985**
Mean (left kidney)	**0.9010**	**5096**	**0.8912**	**0.9062**	**4634**	**0.9009**	**0.9042**	**4923**	**0.8939**

TABLE II
QUANTITATIVE RESULTS FOR THE RIGHT KIDNEY EXTRACTION.

Results based on 20 training datasets

	Vol overlap error%	Score	Ave symm diff%	Score	Ave symm surface dist [mm]	Score	RMS symm surface dist [mm]	Score	Max symm surface dist [mm]	Score	Total
Mean[1]	**8.54**	**66.64**	**-2.44**	**82.38**	**1.35**	**66.14**	**2.62**	**63.57**	**22.37**	**70.57**	**69.86**
Mean[2]	**8.70**	**66.00**	**-0.76**	**85.35**	**1.51**	**62.22**	**3.06**	**57.37**	**27.78**	**63.44**	**66.88**

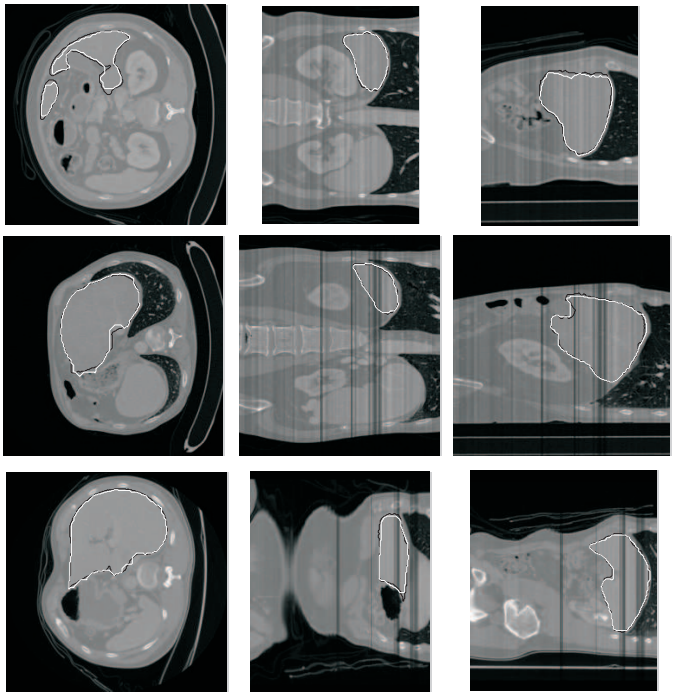


Fig. 14. Results of liver segmentation in transverse (left), coronal (center) and sagittal (right) directions,(black: ground truth, white: our method).

of training and test volumes and utilizing the manually found ground truth for one of the test volumes, we were able to present 30 test cases. Details of this process are presented in Section V-B.

To compare the proposed method with the state of the art, two sets of measures are used. While both these sets can be estimated for each one of these applications, the first set is used for the kidney and the second set for the liver. This was imposed by the state of art, since we could not find any paper that represents both measures for both of these organs.

A. Results for Kidney Segmentation

The proposed solution has been applied on 20 test volumes. Also to make our comparison compatible with the state of art [16], the training phase utilizes 3 models based on 7, 10 and 15 training volumes. The ground truth was prepared manually for both training and test volumes. In order to accommodate the 15 training volumes and 20 test volumes (given that the entire MICCAI's dataset include only 23 (training and test) volumes with both kidneys entirely contained within the sets) the role of training and test volumes are exchanged at several points during the test results generation.

To evaluate the performance of this work on the kidney segmentation, following metrics and sensitivity measure are utilized.

i) Dice Coefficient: this is a similarity measure defined according to the following:

$$Dice = \frac{2|X \bigcap Y|}{|X| + |Y|} \tag{17}$$

here X is the segmentation result by our algorithm and Y is the gold standard (ground truth).

ii) Mean Square Error: this is computed by:

$$MSE = \frac{1}{\Omega} \sum (X - Y)^2 \tag{18}$$

here X and Y are segmentation results by our algorithm and manually found ground truth and Ω is the total number of the pixels in the union of X and Y.

iii) Sensitivity Measure: this is the ratio of correctly identified organ segments to the overall segments in the ground truth.

$$SE = \frac{TP}{TP + FN} \tag{19}$$

here TP is the True Positive and it represents the number of voxels that are segmented consistently (correctly) as kidney tissue by both the proposed method and the ground truth. FN is the False Negative and it represents the number of voxels that exist in the ground truth but were missed by our solution.

These measures are computed and are presented in Table I. The results are separated according to the number of training sets and for left and right kidneys. The mean Dice measure for the left and the right kidneys (across all three models) is 90% which shows slightly better results than the 88.6% presented by [8]. The results for SE (for training with 10 sets) indicate 15% improvement over the results presented by [16] (74%). Interestingly, our proposed method seems to perform rather uniformly regardless of the number of training sets.

The presented MSE results are not compared against values reported by [16] due to the differences in the resolution, size, position and orientation of used datasets.

B. Results for Liver Segmentation

In this section the proposed algorithm has been assessed for liver. In order to generate the organ space for the PCA based model of the liver, first all 20 training volumes are used. The ground truth (liver masks for the training volumes) was provided by the MICCAI's workshop. The ground truth for the test volumes however was not provided but the quality of the results was measured against MICCAI's ground truth by MICCAI's workshop.

To extend the number of test cases, 20 more tests were achieved by using the ground truth of one of the test volumes (manually prepared) and 19 training volumes (leaving one out) to generate the organ space for the PCA model. The left out training volume (acts a test volume) was then modeled using the PCA model. The measurements for these 20 cases were estimated using the assessment program that was provided by the MICCAI's workshop.

To evaluate the performance the evaluation metrics of MICCAI's 2007 workshop [2][3] are adopted. Brief descriptions of these metrics are as follows.

i) Volumetric Overlap Error, in percent. This is the number of voxels in the intersection of segmentation and ground truth divided by the number of voxels in their union, subtracted from 1 and multiplied by 100.
ii) Relative Volume Difference, in percent. This is the total volume difference between segmentation and ground truth divided by total volume of ground truth.
iii) Average Symmetric Surface Distance, in millimeters. The Euclidean distance between every bordering voxel in segmentation and the closest bordering voxel in ground truth is determined.
iv) Symmetric RMS Surface Distance, in millimeters. This measure is similar to the previous measure but here the squared distances are used and the root of the average value is taken.
v) Maximum Symmetric Absolute Surface Distance, in millimeters. This measure is similar to the 2 previous measures but only the maximum of all distances is found.

Table II show the above metrics. In this table, the row Mean[1] represents the mean of the MICCAI's test volumes and the row Mean[2] displays the mean for all the 30 test cases. The proposed algorithm achieves an average of 8.70% for Volume Overlap Error and an average of 1.51 mm for Average Symmetric Surface Distance. The best reported average values for these measures (6.65% and 1.03 mm) by an automatic algorithm on the same dataset are reported by [12]. The average quantitative results for liver segmentation based on all the aforementioned measures show that, at the time of this publication, our algorithm stands among the top four automatic segmentation algorithms for liver extraction and are comparable to those by some of the interactive methods.

C. Execution

All the codes for this project are implemented in MATLAB 7.6.0.324 environment (on a PC with an Intel Core 2 Duo (2 GHz) processor) except for the generic segmentation algorithms (mean shift and graph based), which are done in C++. Since for each slice of each volume, the segmentation algorithm is performed multiple times a large percentage (about 80%) of the running time is spent on the segmentation process. The average runtime for extracting the liver mask of a single slice using the proposed algorithm is about 1 minute. This time for kidney is about 40 seconds. Also the entire system can be programmed in C++ to make the execution time faster.

To ensure the quality of the results the segmentation parameters are chosen to have a large range (covering results from under to over segmentation). This however contributes to substantially higher number of segmentation calls. One way of improving the execution time could be to lower the range of the parameters, concentrating only on the nominal values (as presented in Sections III-C and IV-C). This would cause an overall faster performance, but clearly in some cases the accuracy of the results will be compromised. By assessing the quality of the reconstructed model, we can decide whether further segmentation using an extend range of parameters is required or not. This should allow improvement of the results for such cases while reducing the mean processing time for the entire test volumes.

VI. Conclusion

This paper presented a novel method for identification of organs in CT volumes. The proposed work combined low-level segmentation schemes with a statistical-based modeling approach to accurately identify organ segments. The use of a multi-layer mechanism (through multiple parameter setting combinations) for any generic segmentation algorithm enables the approach to cope with distortions originating from variation in imaging condition and different noise sources. The addition of statistical information (from a training set) provided a unique way to automatically select the most appropriate segmentation parameters (at the slice level) leading to the results that best conformed to the organ's model.

The applications of the proposed method were presented in details for segmentation of kidney and liver in CT scan volumes. Comparison of the performance with the state of art was demonstrated using MICCAI 2007 grand challenge dataset. Currently this work stands among the top four automatic algorithms reported for liver extraction on this dataset.

The proposed solution is a general approach and can be easily customized for general object detection and segmentation applications.

References

[1] R. Susomboon, D. Raicu and J. Furst, *"A Hybrid Approach for Liver Segmentation,"* 3D Segmentation in the Clinic - MICCAI'07 Grand Challenge, pp. 151-160, 2007.
[2] T. Heinmann, M. Styner and B. van Ginneken, *"3D Segmentation in the Clinic - A Grand Challenge,"* International Conference on Medical Image Computing and Computer Assisted Intervention, Workshop Proceedings, pp. 7-15, 2007.
[3] T. Heinmann, B. van Ginneken and M. Styner, *"Comparison and Evaluation of Methods for Liver Segmentation from CT Datasets,"* IEEE Transaction on Medical Imaging, 28(8), pp. 1251-1265, 2009.

[4] K. Seo, L. C. Ludeman, S. Park and J. Park, *"Efficient Liver Segmentation Based on the Spine,"* Advances in Information Systems, vol. 3261, pp. 400-409, 2005.

[5] A. H. Forouzan, R. A. Zoroofi, M. Hori and Y. Sato, *"Liver Segmentation by Intensity Analysis and Anatomical Information in Multi-Slice CT images,"* Proceeding of Liver Segmentation by Intensity Analysis and Anatomical Information in Multi-Slice CT images, volume 4, pp. 287-297, 2009.

[6] S. Pan and B. M. Dawant, *"Automatic 3D segmentation of the liver from abdominal CT images: a level-set approach,"* Society of Photographic Instrumentation Engineers (SPIE) on Medical Imaging, vol. 4322, pp. 128-138, 2001.

[7] H. Badakhshannoory and P. Saeedi, *"Liver Segmentation Based on Deformable Registration and Multi-Layer Segmentation,"* IEEE International Conference on Image Processing, pp. 2549-2552, 2010.

[8] D. T. Lin, C. C. Lei and S. W. Hung, *"Computer-Aided Kidney Segmentation on Abdominal CT Images,"* IEEE Transaction on Information Technology in Biomedicine, pp. 59-65, 2006.

[9] C. H. Wu and Y. N. Sun, *"Segmentation of Kidney from Ultrasound B-mode Images with Texture-Based Classification,"* Journal of Computer Methods and Programs in Biomedicine, vol. 84, pp. 114-123, 2006.

[10] T. F. Cootes, C. J. Taylor, D. H. Cooper and J. Graham, *"Active Shape Models - their Training and Application,"* Journal of Computer Vision and Image Understanding, 61(1), pp. 38-59, 1995.

[11] T. F. Cootes, C. J. Taylor and D. H. Cooper, *"Statistical Models of Appearance for Medical Image Analysis and Computer Vision,"* Society of Photographic Instrumentation Engineers (SPIE) on Medical Imaging, vol. 4322, pp. 236-248, 2001.

[12] D. Kainmuller, T. Lange and H. Lamecker, *"Shape constrained automatic segmentation of the liver based on a heuristic intensity model,"* 3D Segmentation in the Clinic - MICCAI'07 Grand Challenge, pp. 109-116, 2007.

[13] T. Heimann, H. P. Meinzer and I. Wolf, *"A Statistical Deformable Model for the Segmentation of Liver CT Volumes,"* 3D Segmentation in the Clinic - MICCAI'07 Grand Challenge, pp. 161-166, 2007.

[14] H. Ling, S. K. Zhou, Y. Zheng and B. Georgescu, *"Hierarchical, Learning-based Automatic Liver Segmentation"* IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2008.

[15] A. Wimmer, G. Soza and J. Hornegger, *"A Generic Probabilistic Active Shape Model for Organ Segmentation,"* International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 26-33, 2009.

[16] M. Spiegel, D. A. Hahn, V. Daum, J. Wasza and J. Hornegger, *"Segmentation of Kidney Using a New Active Shape Model Generation Technique Based on Non-Rigid Image Registration,"* Journal of Computerized Medical Imaging and Graphics, 33(1), pp. 29-39, 2009.

[17] A. M. Ali, A. A. Farag and A. S. El-Baz, *"Graph Cuts Framework for Kidney Segmentation with Prior Shape Constraints,"* International Conference on Medical Image Computing and Computer Assisted Intervention, pp. 384-392, 2007.

[18] P. F. Felzenszwalb and D. P. Huttenlocher, *"Efficient Graph-Based Image Segmentation,"* International Journal of Computer Vision, 59(2), pp. 167-181, 2004.

[19] D. Comaniciu and P. Meer, *"Mean Shift: A Robust Approach Toward Feature Space Analysis,"* IEEE Transaction on Pattern Analysis and Machine Intelligence, 24(5), pp. 603-619, 2002.

[20] M. Turk and A. Pentland, *"Eigenfaces for Recognition,"* Journal of Cognitive Neuro-science, 3(1), pp. 71-86, 1991.

[21] M. Turk and A. Pentland, *"Face recognition using eigenfaces,"* IEEE Conference on Computer Vision and Pattern Recognition, pp. 586-591, 1991.

[22] F. Keinosuke and L.D. Hostetler, *"The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition,"* IEEE Transactions on Information Theory, 21(1), pp. 3240, 1975.