

LIVER SEGMENTATION USING STRUCTURED SPARSE REPRESENTATIONS

Vimal Singh[†] Dan Wang[†] Ahmed H. Tewfik[†] Bradley J. Erickson^{*}

[†] University of Texas, Austin, Texas, USA

^{*} Mayo Clinic, Rochester, Minnesota, USA

ABSTRACT

Segmentation of liver from volumetric images forms the basis for surgical planning required for living donor transplantations and tumor resections surgeries. This paper introduces a novel idea of using sparse representations of liver shapes in a learned structured dictionary to produce an accurate preliminary segmentation, which is further evolved using a joint image and shape based level-set framework to obtain the final segmented volume. Structured dictionary for liver shapes can be learned from an available training dataset. The proposed approach requires only 3 orthogonal segmented masks as user-input, which is less than half the number required by current state-of-the-art interaction-based methods. The increased accuracy of the preliminary segmentation translates into faster convergence of the evolution step and highly accurate final segmentations with mean average symmetric surface distances (ASSD) [1] of (1.03 ± 0.3) mm when tested on a challenging dataset containing 62 volumes. Our approach segments a volume on an average of 5 mins and, is 25% (approx.) faster than comparably performing techniques.

Index Terms— Structured Sparsity, Sparse Representations, Level-set Evolution, Semi-Automatic Segmentation, Subspace Clustering

1. INTRODUCTION

An accurate liver segmentation method is critical for avoiding donor-recipient volume mismatches in case of living donor liver transplantation (LDLT) surgeries and in defining the precise course of action for tumor resection procedures [2]. Liver segmentation is a challenging task. In abdominal computed tomography (CT) or magnetic resonance (MR) images, there is little difference in the gray-value intensities of adjacent tissues. This leads to loss of boundary for liver in regions close to organs such as: diaphragm, kidney, pancreas, stomach, and heart. At such boundary-less regions, segmentation using simple intensity based heuristic approaches, seed-initiated region growing techniques and edge-detection based evolution algorithms leak into the surrounding organs and need to be corrected manually. Due to the limitations of established techniques for liver segmentation, commonly used systems in clinical practice rely either on manual segmentations or on tools requiring an expert to segment the organ correctly. This process is time consuming, tedious and generally not-reproducible [3]. This paper presents a radically novel approach to the liver segmentation problem by first producing an initial accurate segmentation from learned structured sparse representations of liver surfaces, which is further evolved using a regularized 3D level-set formulation to achieve final segmentation.

Automatic segmentation methods rely on heuristics based assumptions for their initialization, which generally fail in the presence of pathological structures and result in poor generalization and performance of these methods over varying datasets [4]. On the other

hand, semi-automatic methods require user-input for initialization which can be used to regularize the underlying evolution techniques to avoid complete failures resulting in comparatively more robust and accurate methods than the automatic approaches [4]. In general, for semi-automatic approaches there is a trade-off between the amount of user-interaction employed and the segmentation performance (speed and accuracy) that can be achieved. The challenge is to improve their speed and performance, while keeping the user-interaction minimal. In this project, a novel approach for recovery of 3D deformable organs using structured sparse representations is seamlessly coupled with a simple user-input scheme to achieve volumetric segmentation of organs. The inherent coupling between the requirements for stable reconstruction of deformable organs using the proposed sparse representation based approach and, the possibility of acquiring segmentation masks in orthogonal planes allows for reducing the amount of user-input required to accurately segment complexly shaped organs like liver. An important point to be noted is that the domain in which the input is acquired (i.e., the sampling domain) and the domain in which the shape is approximated (i.e., the representation domain) are different in our method.

To the best of our knowledge, the approach presented in [5] is the current state-of-the-art semi-automatic method and it has been classified as a medium interaction approach in [4]. This method requires 6-8 orthogonal segmentation masks as user-input and uses radial basis functions (RBFs) to interpolate a liver estimate from user-provided masks. As RBFs are smooth functions (low spatial frequencies), the interpolated shape in regions far from input masks is a poor approximation of the true liver shape. As a result the segmentation performance of [5] even after an evolution step is sub-optimal and cannot be used for clinical purposes. The approach presented in [6] is based on a similar idea of reducing the amount of user-interaction needed through identification of a Sparse Information Model (SIM) for liver segmentation. The SIM developed requires the user to provide binary masks in 5 longitudinal slices and interpolates the remaining longitudinal masks from these masks using a linear function learned through statistical analysis of the training dataset. The performance achieved using the SIM method is better than that of the methods based on simple active shape models (Principal Component Analysis) [7], but is poorer than the approach of [5] and far below the clinical standards.

The main innovation of this project consists of leveraging the structured sparse representations of liver surfaces to produce an accurate preliminary segmentation of any given liver from CT data. The approach requires less than half the number of manually selected contours (only 3 binary masks) as user-input compared to traditional methods [5, 6] and leads to substantially more accurate initial estimate for the liver shape. The increased accuracy of the initial segmentation translates into highly accurate final segmentations computed via level-set methods at $\sim 25\%$ faster speeds. The rest of this paper is organized as follows. In section 2, the proposed ap-

proach is presented in detail. Preliminary results obtained using our approach over a limited dataset are presented in section 3. Finally, Section 4 concludes the paper.

2. LIVER SEGMENTATION USING STRUCTURED SPARSE REPRESENTATIONS

The proposed approach relies on structured sparse representations of 3D deformable organs [8]. Sparse representation of liver shape in a learned structured dictionary allows for estimating a 3D shape close to the true liver using limited number of samples. This forms the premise for our semi-automatic segmentation method requiring low user-input. Due to the complexity and variability of liver shapes [4], the initial estimate of the liver is further evolved using a regularized 3D level-set approach to produce clinically actionable liver segmentation. The initial step of our approach provides the following advantages to the underlying evolution framework: (1) it provides a close initialization, thus guarantees a faster convergence; (2) it allows for regularization based on user-input and the estimated liver shape to prevent leakage. Our complete approach as outlined in figure 1, can be divided into two stages: 1) training stage and, 2) segmentation stage.

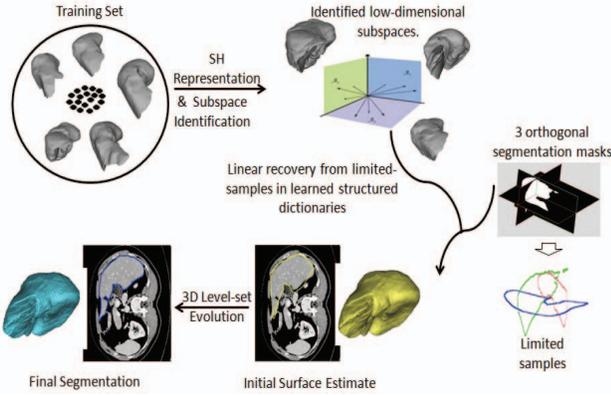


Fig. 1: Segmenting liver using Structured Sparse Representations

2.1. Training Stage

The training stage involves identification of a structured dictionary for representing possible liver shapes from an available training data set. The dictionary is learned using spherical harmonic (SH) transform and iterative subspace identification (ISI) method [9]. During the training stage, all livers in the training data-set are first represented with SH coefficient vectors in the harmonic domain to remove high frequency noise and lower the training vector size. Then, the ISI method is used to identify all low-dimensional subspaces of these SH representations. Each identified subspace represents a family of liver shapes and each family member can be expressed as a linear combination of the basis vectors spanning that subspace. For more details the reader is referred to [8]. Empirically, larger training dataset leads to better generalization of the identified dictionary and thus results in closer initial estimation of the liver shape. Let $\{\mathbf{G}_i\}_{i=1}^J$ represent the J identified shape subspaces and \mathbf{G} denote the learned structured dictionary obtained by concatenating all \mathbf{G}_i . Then, a 3D liver \mathbf{V} residing in the i^{th} subspace \mathbf{G}_i can be represented using a

unit-block sparse vector p with low-approximation error as in (1), where c_i are its non-zero description coefficients in \mathbf{G}_i .

$$\begin{aligned} \mathbf{V} &\cong \mathbf{G}_i * c_i \\ &= \mathbf{G} * p \\ p &= [\mathbf{0}^T \dots c_i^T \dots \mathbf{0}^T]^T \end{aligned} \quad (1)$$

Note, the dimension of the descriptor p is far less than the original surface data dimension due to the low-dimensionality of identified subspaces $\{\mathbf{G}_i\}_{i=1}^J$: $\dim(p) \ll \dim(\mathbf{V})$.

2.2. Segmentation Stage

The segmentation stage consists of two main steps. The first step involves identification of the optimal subspace and reconstructing an estimate of liver shape from limited samples in the identified subspace. Estimating the initial shape corresponds to recovering the structured sparse representation p in (1) by solving the following over-determined system:

$$\begin{aligned} \min_p \quad & \|\mathbf{V} - \mathbf{G} * p\|_2 \\ \text{s.t.} \quad & \|p\|_{0,B} \leq 1 \end{aligned} \quad (2)$$

where, $\|p\|_{0,B} \leq 1$ represents the unit-block sparsity constraint. Problem (2) assumes that all liver samples are available. But, since $\dim(p) \ll \dim(\mathbf{V})$, one can do with using small number of 3D samples only, as long as they contain information for all the needed coefficients and the over-completeness of the problem is maintained. Let, \mathbf{V}_{input} denote such limited sample points and $\tilde{\mathbf{G}}_i$ represent the sub-matrix obtained from selecting rows in \mathbf{G}_i corresponding to them. Then, estimation of the initial liver using available limited samples can be reformulated as solving the following minimization problems:

$$\min_{1 \leq i \leq J} \|\mathbf{V}_{input} - \tilde{\mathbf{G}}_i * \hat{c}_i\|_2 \quad (3)$$

In the proposed approach, the samples used for recovery correspond to the boundary contour points of input masks provided by the user, as shown in Fig. 1. In order to ensure closeness of \hat{c}_i to true c_i , it is required to choose sampling locations such that $\tilde{\mathbf{G}}_i$ is well-conditioned. Empirically, such recovery problems provide a better fit with more data-knowledge and it has been observed that limited-samples corresponding to 3 input orthogonal planes result in well-behaved linear systems, i.e., the conditioning of $\tilde{\mathbf{G}}_i$ is close to that of \mathbf{G}_i . For this reason, 3 segmented masks at the center of liver are used as the minimum user-input in order to capture as much shape topology as possible to ensure more accurate estimation of the structured sparse representations. The initial estimate of liver shape can be obtained as:

$$\mathbf{V}_{estimate} = \mathbf{G}_{i^*} * \hat{c}_{i^*} \quad (4)$$

where, i^* is the active shape-subspace yielding the minimum 2-norm error at the sampled locations \mathbf{V}_{input} . Our approach is different from simple active shape model (ASM) based approaches [7], as the active shape space is represented as a union of multiple dominant low-dimensional subspaces instead of a single high-dimensional principal component space. Since, there is a direct trade off between the conditioning of the linear inverse problem (3) and, the dimensionality of the shape space (\mathbf{G}_i), the proposed algorithm yields more accurate and smooth approximate solutions than ASM based methods.

The second step of the segmentation stage is to evolve the estimated shape ($\mathbf{V}_{estimate}$) from the first step into a final segmentation using a level-set formulation based on the 3D image data. The

level-set formulation of equation (5) is used, which weighs the user-input and the estimated shape during evolution to ensure prevention of leakage into surrounding organs. In (5), $\phi(\mathbf{x}, t)$ is the embedding function whose zero level-set corresponds to the evolving surface, $\mathbf{V} = \{\mathbf{x} | \phi(\mathbf{x}, t) = 0\}$, \mathbf{I} is the non-linearly diffused 3D gray-scale CT image, g is a stopping function as in [5], “div” is the divergence operator, $\delta(\mathbf{x})$ is a distance function and $\mu(\mathbf{x})$ weights the influence between shape and image components during the evolution. The spatial position \mathbf{x} is omitted in (5) for a simpler notation.

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \operatorname{div} \left((\mu(\delta) g(\nabla \mathbf{I}) + (1 - \mu(\delta)) \delta) \frac{\nabla \phi}{|\nabla \phi|} \right) \quad (5)$$

The first term in (5) is the image term which controls smoothness of the evolving surface and stops it at strong edges. The second term corresponds to shape evolution, where $\delta(\mathbf{x})$ is a distance function of the evolving surface from the initial liver estimate. $\mu(\mathbf{x})$ is a sigmoid function whose shift and slope is calculated based on the distance of estimated masks in the previous step corresponding to the user provided input masks. For more details on the joint image and shape based evolution, the reader is referred to [5].

3. RESULTS

The proposed method is applied to a combined dataset comprising of 62 livers acquired from various sources [1, 10]. A variety of different CT scanners were used for acquisition and most images in these studies were pathologic and included tumors and cysts of different sizes. For all results reported, we randomly pick 5 livers as test data and the remaining volumes are used for training purpose. This experiment is repeated multiple times and results are averaged. Segmentation performance is evaluated using reference segmentations available in the datasets.

	Florin <i>et al.</i>	Wimmer <i>et al.</i>	Proposed approach
Median(%)	11.5	14.96	10.41
Maximum(%)	17.1	26.09	13.84
Minimum(%)	9.5	9.06	6.96

Table 1: Comparison of the initial liver estimation step using the symmetric distance (SD) measure.

To evaluate the true innovation of the proposed approach, we quantify the closeness of the initial liver estimate obtained in the first step to the actual liver shape using the symmetric distance (SD) measure (6), as defined in [6] using dice’s coefficient.

$$SD = \left(1 - \frac{2 * |\mathbf{V}_1 \cap \mathbf{V}_2|}{|\mathbf{V}_1| + |\mathbf{V}_2|} \right) \quad (6)$$

Dice Coefficient is defined as twice the shared information over the combined cardinalities. Table 1 compares the SD measure obtained after the initial step for our approach with those of [5, 6]. Three orthogonal binary segmentation masks at liver center mass are used to recover the sparse representation in a learned structured dictionary to obtain an initial liver estimate for the **proposed approach** and, a shape is interpolated using linear RBFs based on these masks for the Wimmer *et al.* approach as explained in [5]. Florin *et al.* use 5 key slices in the longitudinal direction to estimate a complete liver from a learned Sparse Information Model (SIM) [6]. Table 1 illustrates that our approach identifies a SIM yielding sparser representations

for liver shapes than [6] and, when coupled with optimal samples it produces estimates closer to true liver than existing state-of-the-art methods.

Fig. 2 shows the segmentations achieved for two test livers using the proposed method where the user-input is restricted to 3 masks as described previously. The first row of Fig. 2 corresponds to a normal difficult case which is segmented with good accuracy. The second row shows an example result exhibiting the robustness of our approach to lesions. Such robustness is attributed to the innovative method of estimating the initial liver using structured sparse representations. The green contours show the initial shape obtained at each plane in the first step of the segmentation stage. It can be observed that the estimates obtained are very close to the reference segmentation contours, which are shown in red color. The blue contours show the final shape to which our liver estimate evolves using the level-set formulation presented in Sec. 2.2. To quantify the segmentation performance, following metrics are used: volumetric overlap error (VOE), relative volume difference (RVD), average symmetric surface distance (ASSD), root mean square symmetric surface distance (RMS-SSD) and maximum symmetric surface distance (MSSD). ASSD, RMS-SSD and MSSD are statistical derivatives of symmetric surface distances (SSD) which are calculated as follows: border voxels for the final segmentation output and the reference segmentations are determined. For each voxel along one border, distance to the closest voxel along the other border is determined. All these distances for border voxels from both output and reference segmentations are termed as symmetric surface distances [1]. Table 2 summarizes the metrics obtained for the proposed approach and their comparison with state-of-the-art semi-automatic method of [5] using the results reported in [4] and the average human segmentation performance as described in [1].

Metric	Avg. human	Wimmer <i>et al.</i>	Proposed approach
VOE(%)	6.4	8.1 ± 1.1	6.04 ± 1.3
RVD(%)	4.7	6.1 ± 2.6	1.85 ± 1.7
ASSD(mm)	1.0	1.3 ± 0.2	1.03 ± 0.3
RMS-SSD(mm)	1.8	2.2 ± 0.4	1.79 ± 0.4
MSSD(mm)	19.0	18.1 ± 4.6	18.6 ± 4.1
Input (Masks)		6 - 8	3
Run Time		(~ 7 mins)	(~ 5 mins)

Table 2: Comparison of performance metrics for segmentation using the proposed approach.

Table 2 and Fig. 2 show that the preliminary results of the proposed method over a limited dataset are significantly better than the performance of Wimmer *et al.* method [5], while requiring less than half the amount of user-interaction. Our complete approach runs ~25% times faster due to the reduced user-interaction time and, the faster convergence of the evolution step as it is initialized closer to the true liver. The proposed approach also performs better than many state-of-the-art automatic approaches in terms of the mean ASSD metric (best: 1.37mm) based on a comparison of the performances as reported in [11].

4. CONCLUSION AND FUTURE WORK

In this paper, a low user-input based two-step semi-automatic liver segmentation method is presented. Novelty of the proposed method lies in its first step, where sparse representations of 3D liver shapes in

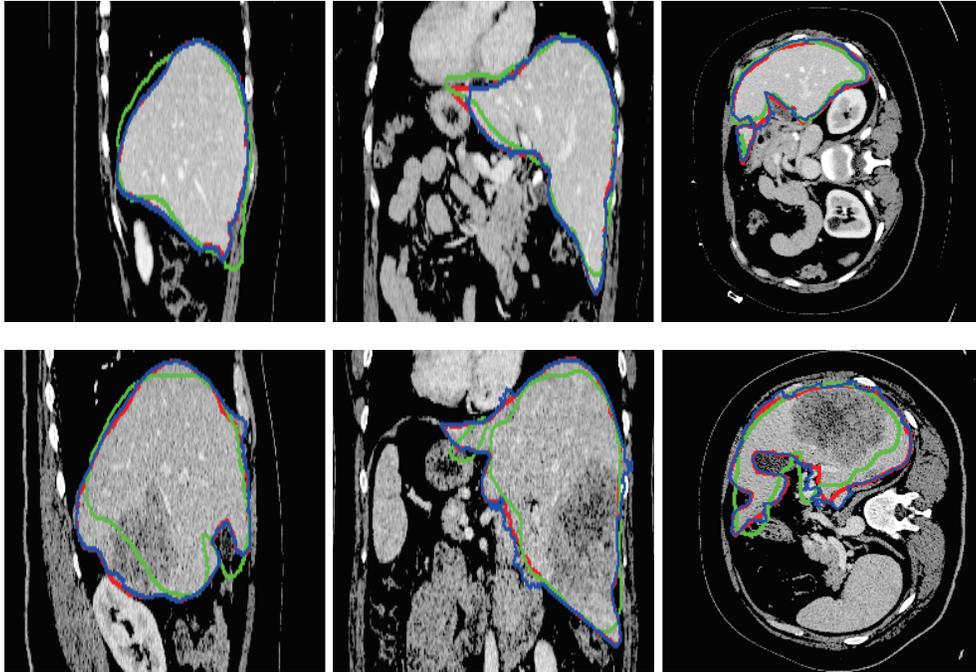


Fig. 2: From left to right, a sagittal, a coronal and a transversal slice for two different test livers. Green contours show initial estimate of the liver shape obtained in the first step of the segmentation stage. Red contours correspond to the reference segmentations and, Blue contours are the final segmentations to which green contours evolve to.

a learned structured dictionary are used to produce a preliminary segmentation close to the true liver. The inherent coupling between the requirements for stable recovery using sparse representations and, the possibility of acquiring segmentation masks in orthogonal directions allows for obtaining an accurate preliminary segmentation with as low as 3 masks, half the number of masks required by state-of-the-art methods [5, 6]. In the second step, the liver estimate is further evolved using a joint image and shape based level-set formulation to achieve final segmentation. Preliminary results obtained using the proposed approach over a limited dataset show that it achieves better performance than state-of-the-art semi-automatic and automatic methods. For future work, more data from different datasets (such as: MRI) should be included to further improve the performance and generalization of the proposed approach.

5. REFERENCES

- [1] B. V. Ginneken, T. Heimann, and M. Styner, “3D Segmentation in the Clinic: A Grand Challenge,” in *MICCAI Wshp. 3D Segmentation in the Clinic: A Grand Challenge*, 2007, pp. 7–15.
- [2] B. Reitingner, A. Bornik, R. Beichel, and D. Schmalstieg, “Liver Surgery Planning Using Virtual Reality,” *Computer Graphics and Applications, IEEE*, vol. 26, no. 6, pp. 36–47, nov.-dec. 2006.
- [3] M. A. Styner, H. C. Charles, J. Park, and G. Gerig, “Multi-site validation of image analysis methods: assessing intra- and inter-site variability,” in *Medical Imaging 2002: Image Processing*, Milan Sonka and J. Michael Fitzpatrick, Eds. 2002, vol. 4684, pp. 278–286, SPIE.
- [4] T. Heimann *et. al.*, “Comparison and Evaluation of Methods for Liver Segmentation From CT Datasets,” *Medical Imaging, IEEE Transactions on*, vol. 28, no. 8, pp. 1251–1265, Aug 2009.
- [5] A. Wimmer, G. Soza, and J. Hornegger, “Two-stage semi-automatic organ segmentation framework using radial basis functions and level sets,” in *MICCAI Wshp. 3D Segmentation in the Clinic: A Grand Challenge*, 2007, pp. 179–188.
- [6] C. Florin, N. Paragios, G. Funka-lea, and J. Williams, “Liver Segmentation Using Sparse 3D Prior Models with Optimal Data Support,” in *Information Processing in Medical Imaging (IPMI07)*, in. 2007, Press.
- [7] H. A. N. S. Lamecker, T. Lange, and M. Seebass, “Segmentation of the Liver using a 3D Statistical Shape Model,” Tech. Rep., April 2004.
- [8] D. Wang and A. H. Tewfik, “Sparse representation of deformable 3D organs,” in *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on*, july 2009, pp. 887–890.
- [9] B. Gowreesunker and A. H. Tewfik, “A Novel Subspace Clustering Method for Dictionary Design,” in *Independent Component Analysis and Signal Separation*, vol. 5441 of *Lecture Notes in Computer Science*, pp. 34–41. Springer Berlin / Heidelberg, 2009.
- [10] “<http://www.ircad.fr>,” .
- [11] H. Ling, S.K. Zhou, Y. Zheng, B. Georgescu, M. Suehling, and D. Comaniciu, “Hierarchical, learning-based automatic liver segmentation,” in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, June 2008, pp. 1–8.