Automatic Multi-organ Segmentation Using Learning-Based Segmentation and Level Set Optimization

Timo Kohlberger[1], Michal Sofka[1], Jingdan Zhang[1], Neil Birkbeck[1],
Jens Wetzl[1], Jens Kaftan[2], Jérôme Declerck[2], and S. Kevin Zhou[1]

[1] Image Analytics and Informatics, Siemens Corporate Research, Princeton, NJ, USA
[2] Molecular Imaging, Siemens Healthcare, Oxford, UK

Abstract. We present a novel generic segmentation system for the fully automatic multi-organ segmentation from CT medical images. Thereby we combine the advantages of learning-based approaches on point cloud-based shape representation, such a speed, robustness, point correspondences, with those of PDE-optimization-based level set approaches, such as high accuracy and the straightforward prevention of segment overlaps. In a benchmark on 10–100 annotated datasets for the liver, the lungs, and the kidneys we show that the proposed system yields segmentation accuracies of 1.17–2.89mm average surface errors. Thereby the level set segmentation (which is initialized by the learning-based segmentations) contributes with an 20%-40% increase in accuracy.

1 Introduction

Discriminative segmentation approaches have proven to give reliable, fully-automatic, and fast detections of anatomical landmarks within volumetric images, as well as the accurate determination of organ boundaries, such as of the inner and outer walls of the heart, [10], or of the liver, [4]. Usually, the segmenting surface is represented by relatively low number of explicit control points, such as they are used in Active Shape Models.

Besides restrictions in topology, other well-known disadvantages of point cloud-based shape representations are the dependence of the local detailedness on the local density of control points. The latter often are non-homogeneously distributed across the shape boundary, and thus yield varying levels of segmentation accuracy. Level set-based shape representations, see [1] and the references therein, on the other hand, provide a well-known mean to encode segment boundaries at a homogeneous resolution, with simple up- and down-sample schemes. Moreover, in the case of multiple objects, the detection and formulation of constraints to prevent overlaps between adjacent segment boundaries can be achieved much simpler by a level set representation where signed distance functions are employed.

In the following we will present a fully automatic segmentation system for multiple organs on CT data, that combines the advantages of both segmentation

approaches and their employed shape representations in an optimal manner. In particular the point-to-point correspondences, which are estimated during the learning-based segmentation, will be preserved in the level set segmentation. With regards to the latter, we will present novel terms which allow to impose region-specific geometric constraints between adjacent boundaries.

In Section 2 we first give an outline of the learning-based detection and segmentation stages, all of which have been published in the mentioned citations. After a brief description on the conversion of their output meshes to implicit level set maps, we start to build an energy-based minimization approach for multiple organs as a segmentation refinement stage. In the experimental Section 3, we first show the impact of the new constraints at qualitative examples, and finally evaluate the overall improvement of the level set refinement stage over the detection-based results at 10–100 annotated cases for the liver, lungs and kidneys.

2   Approach

2.1   Anatomical Landmark Detection and Learning-Based Segmentation of Organ Boundaries

We initialize the multi-region level set segmentation from explicitly represented boundary surfaces stemming from an existing learning-based detection framework, which itself consists of several stages. In the first stage, a landmark detection system estimates key organ landmarks ranging from the abdominal to upper body region, see [5] for more details. These landmarks then serve a initializations for a hierarchical bounding box detection system based on Marginal Space Learning [10] in connection with Probabilistic Boosting Trees [9]. The latter yields bounding box estimates for the liver, the left and right lung, the heart, and the kidneys. In the third stage, organ-specific boundary detectors are employed to evolve the correct organ boundaries on a coarse scale and in subsequently on a fine scale, see [4]. In addition, PCA-based statistical shape model are used to regularize the boundary shape on the coarse resolution. Thereby, segment boundaries of each organ are represented by a triangulated mesh, i.e. a connected point-cloud as being used in Active Shape Models.

2.2   From Meshes to Zero-Crossings of Signed Distance Maps

Although the learning-based segmentation sub-system already provides good individual organ segmentations, see Fig 2(a) for example, they usually exhibit small overlaps between adjacent organ boundaries, or gaps where the true organ boundaries coincide. Given representations of only the adjacent segments' boundaries, those deficiencies are difficult to detect and remove. Instead we initialize signed distance functions $\phi_i : \mathbb{R}^3 \Rightarrow \mathbb{R}$ from each of the result meshes C_i, for $i, \dots, N$ organs by employing a fast mesh voxelization algorithm. The boundary information then is encoded implicitly in the zero-crossings of the ϕ_i, i.e.

$C_i := \{\mathbf{x} \,|\, \phi_i(\mathbf{x}) = 0, \, |\nabla\phi| = 1\}$, with $|\nabla\phi| = 1$ denoting the so-called distance-property, and $\phi > 0$ inside the object and < 0 outside, see [1] and the references therein. The distance functions are discretized on a regular grid, which, in the following, is assumed to be the same for all organs. Furthermore, we employ a narrow-banded level set scheme, which maintains the distance-property in a small narrow-band of ± 2 voxels from the zero crossing. In addition to the distance functions, we still keep the mesh points, denoted by $\{\boldsymbol{p}_i\}$, and are tracking them along with the evolving zero crossing as described in [3], since they provide point-wise correspondences to the mean shape of the PCA model employed in the preceding learned-based boundary detection step.

2.3   Data and Smoothness Term of the Level Set Approach

Having the triangulated boundary meshes $\{C_i\}$ of the detection stages transfered to distance functions functions $\{\phi_i\}$, detecting and removing local overlaps and gaps between them can be realized much easier. The ultimate target of course is to find the correct separating boundary between two neighboring organs. To that end, in the following, we propose a level set segmentation approach which not only refines the segmentation boundary accuracy, removes local overlaps and gaps, but also finds the true separating boundary given that enough image information is available.

For each organ O_i, this refining level set segmentation is realized by employing gradient descent iteration to converge to a minimum of an associated energy functional $E_i(\phi_i)$, given the initial distance maps as starting points, see [1] and the references therein. As data-dependent energy term, we here employ

$$E_p(\phi) = -\alpha \int_\Omega H(\phi) \log p_{in}(I(\mathbf{x})|\phi) + \big(1 - H(\phi)\big) \log p_{out}(I(\mathbf{x})|\phi) \, d\mathbf{x}\,,$$

with H denoting the Heaviside step function, and $p_{in/out}$ referring to non-parametric probability estimates of the intensities inside and outside, respectively, of the current segment ϕ using a Parzen density estimator with a Gaussian kernel, see [2] for further details, and α being a constant weight. In order to add robustness against noisy data, we furthermore incorporate the boundary smoothness regularization term

$$E_c(\phi) = \int_\Omega \gamma_{l(\mathbf{x})} \, |\nabla H(\phi)| \, d\mathbf{x}, \quad \text{with} \quad l(\mathbf{x}) = \arg \min_{i=1,\dots,N} \|\mathbf{x} - \mathbf{p}_i\|_{\mathcal{L}_2}, \qquad (1)$$

which employs a weight $\gamma_{l(\mathbf{x})}$ that varies with the location on the boundary. The latter is realized by assigning fixed weights $\{\gamma_i\}$ to each of the correspondence points $\{\boldsymbol{p}_i\}$ on the mean shape of the PCA shape model, which then are tracked along during the zero-crossing evolution, see [3] for more details.

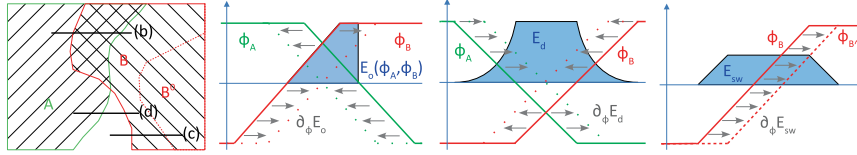### 2.4   Disjoint Constraint to Remove Overlaps

Let C_A^0 and C_B^0 be detected boundaries of the learning-based stage in the case of two adjacent organs A and B, and assume that C_A^0 and C_B^0 imperfectly overlap

each other to a certain degree, see, e.g., Fig. 1(a). By representing these two surfaces using signed distance functions ϕ_A^0 and ϕ_B^0, locations $\tilde{\mathbf{x}}$ inside the overlapping region are exclusively characterized by $\phi_A(\tilde{\mathbf{x}}) > 0$ and $\phi_B(\tilde{\mathbf{x}}) > 0$, and thus provide a much simpler overlap indicator compared to any other based on an explicit shape representation. Subsequently, additional energy terms which explicitly penalize overlaps usually are of the form

$$E_o(\phi_A, \phi_B) := \int_\Omega H\big(\phi_A(\mathbf{x})\big)\, H\big(\phi_B(\tilde{\mathbf{x}})\big)\, \phi_B(\tilde{\mathbf{x}})\, d\mathbf{x} \tag{2}$$

where the first product in the integrand is unequal zero only inside the overlap regions, see Fig. 1(b). In addition to similar terms such as proposed in [6], we propose to also multiply with the second distance function ϕ_B which makes E_O smoother at the presence of small overlaps, and thereby decreases oscillations during gradient descent. The corresponding energy gradient reads (cf. [7]):

$$\partial\phi_A/\partial t = -\partial E_o/\partial\phi_A = -\delta_\epsilon(\phi_A)\, H_\epsilon\big(\phi_B(\tilde{\mathbf{x}})\big)\phi_B(\tilde{\mathbf{x}}) \,. \tag{3}$$

Fig. 1. Imposing geometric constraints to remove overlaps (b) and gaps (c) from an existing segmentation, while controlling the deviation from a given shape locally (d). (Note that (b)–(d) are 1D-cuts along the black lines in (a)).

### 2.5   Local Proximity Constraint to Fill Gaps

With regards to removing erroneous gaps between adjacent segmentation boundaries, we add the following energy to the total energy functional:

$$E_d(\phi_A, \phi_B) := \frac{1}{2} \int_\Omega \beta_{l(\mathbf{x})}\big(\phi_A(\mathbf{x}) + \phi_B(\tilde{\mathbf{x}}) + D\big)^2 d\mathbf{x} \tag{4}$$

with $D = 0$ for the time being, and $\{\beta_i\}$ being correspondence points-bound weights with $\beta_i = 0$ at points where no boundary coincidence ought to be enforced, and $\beta_i > 0$ at locations where boundaries of A and B ought to coincide. As illustrated in Fig. 1(c), ϕ_A and ϕ_B cancel each other out if their zero crossings coincide and thus the integrand becomes zero. As an extension, one can enforce the two boundaries to not touch but stay in a predefined distance $D > 0$ from each other. The gradient descend PDE of E_d w.r.t. ϕ_A reads:

$$\partial\phi/\partial t = -\partial E_d/\partial\phi = -\beta_{l(\mathbf{x})}\big(\phi_A(\mathbf{x}) + \phi_B(\tilde{\mathbf{x}}) + D\big) \tag{5}$$

which shows that of ϕ_A increases at locations where $\phi_B < D$, and thus expands its representing boundary, and descreases at locations where $\phi_B > D$, i.e. shrinks the boundary.

(a)   (b)   (c)
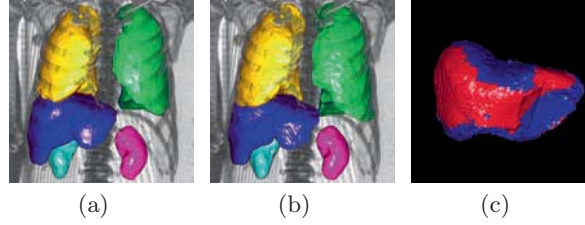
Fig. 2. (a) Segmentation result of the machine learning system. (b) Refined segmentations after applying level set segmentation to (a). (c) Visualization of the manually set local weights $\{\omega_i^{out}\}$ of the outward template constraint (red: 50, blue: 0.2). These weights are bound to point-based shape correspondences in relation to a fixed model shape, and thereby allow for a region-specific control over the different geometrical constraints during the level set segmentation.

2.6   Template Constraint to Control Deviation of the Refinement

Finally, we added a third geometric term, which ensures that the level set result is sufficiently similar to learning-based contour, that is, the refined boundary is sought only in the vicinity of its initialization. To that end, we use the term

$$E_{sw}(\phi, \phi_P) := \frac{1}{2} \int_{\Omega} \omega_{l(\mathbf{x})}^{in} \, H\big(\phi_P(\mathbf{x}) - \phi(\mathbf{x})\big) \; + \omega_{l(\mathbf{x})}^{out} \, H\big(\phi(\mathbf{x}) - \phi_P(\mathbf{x})\big) \; d\mathbf{x} \; . \quad (6)$$

which is an extension of the approach shown in [8] in the sense that it applies region-specific weights $\{\omega_i^{in}\}$ to the shape dissimilarity measure between the current ϕ and the prior shape ϕ_P (which is the initial one here), as well as applying different weights for deviations outside or inside of C_P (we refer to C_P as "template" shape in the following). See Fig. 2(c) for a local weight map. Technically, note that the first term of the integrand is non-zero only if the zero-crossing of ϕ resides inside the zero-crossing of ϕ_P, that is the current boundary C is smaller than the prior boundary C_P, see Fig. 1(d). Vice-versa, the second term measures local expansions relative to C_P, by becoming non-zero only where $\phi(\mathbf{x}) > \phi_P(\mathbf{x})$.

The corresponding energy gradient clearly shows that the proposed energy term has the desired effect:

$$\partial\phi/\partial t = -\partial E_{sw}/\partial \phi = \omega_{l(\mathbf{x})}^{in} \, \delta_{\epsilon}\big(\phi_P(\mathbf{x}) - \phi(\mathbf{x})\big) \; - \omega_{l(\mathbf{x})}^{out} \, \delta_{\epsilon}\big(\phi(\mathbf{x}) - \phi_P(\mathbf{x})\big) \; , \quad (7)$$

i.e. increasing ϕ at locations where $\phi < \phi_P$, and decreasing it in the opposite case.

2.7   Interleaved Multi-energy Minimization

Finally, all of the proposed energy terms are combined into energy minimizations for each organ $O_{i=1,...,N}$:

$$\min_{\phi_i} E_p(\phi_i) + E_c(\phi_i) + \sum_{j \in \mathcal{N}_i(j)} E_o(\phi, \phi_j) + \sum_{j \in \mathcal{P}_i(j)} E_d(\phi, \phi_j) + E_{sw}(\phi_i, \phi_i^0) \; , \quad (8)$$

which are mutually coupled by the disjoint and proximity terms ($\mathcal{N}_i$: indices of organs adjacent to O_i, $\mathcal{P}_i$: indices of organs with which O_i shares a mutual proximity constraint). Consequently, minimizers $\{\tilde{\phi}_i\}$ of these individual energies depend on each other. To that end, we found interleaved gradient descent iterations to yield the desired segmentation improvements in practice. Specifically, we carry out a descent along the negative gradients of the N per-organ energies in lockstep, while using the segmentation results $\{\phi_i^{t-1}\}$ of the previous joint iteration to compute the coupled energy gradients $\partial E_i(\phi_i; \{\phi_i^{t-1}\})/\partial\phi_i$. The descent for a particular energy is terminated if a given maximum number of iterations has been reached, or if the maximum norm of its gradient falls below a given threshold, i.e. the segmentation boundary ϕ_i changes less than a chosen tolerance.

3   Experimental Evaluation

3.1   Parameter Selection

In a first series of experiments, we studied the effect of the proposed new energy terms qualitatively on a few data sets. Thereby we also manually selected all the involved weights in order achieve optimal results on a small set of test cases. Specifically, we weighted the data term E_p with a factor of 2 for all organs and set the lowest weight γ_i in the smoothness term E_c to 0.7 at locations where a high curvature is desired (such as at the lung tips), and to a value of 1.5 where low curvatures are to enforced. For the disjoint energy term E_o we found a weighting of 1000 to remove any existing overlaps while not producing any oscillations. The proximity term E_d turned out to yield the desired results when setting β_i to 10 at correspondence points where it ought to be active and to zero elsewhere. For the template energy E_{sw} the inward and outward deviation-penalizing weights ω^{in} and ω^{out} were set according to the overall quality and robustness of the learning-based boundary result. Specifically, at locations where the latter tends to under-segment, such as at the lower tips of the lung wings, ω^{out} was set to the low value of 0.2 in order to allow the level set result to deviate outwards. Vice-versa, at locations where the learning-based stage tends to over-segment, such as at the lower side of the liver ω^{in} is given a low value. Finally, at locations where the learning-based stage already yields highly accurate results, such as adjacent to the ribs for the liver and lung boundaries, both ω^{in} and ω^{out} where set to high values of 50 in order to bind the level set-based boundary close to it. Fig. 3(c) visualizes the weights $\{\omega_i^{out}\}$ for a final segmentation of the liver.

3.2   Accuracy Benchmark

In a next step, we benchmarked the overall system at manually annotated data sets in order to study the overall accuracy improvement yielded by level set refinement. Test sets for the different organs listed in Table 1 were drawn randomly from a set of 434 annotated CT cases consisting of low to high contrast scans, and an average voxel spacing of 1.02/1.02/2.16mm (x/y/z). Cases not drawn for
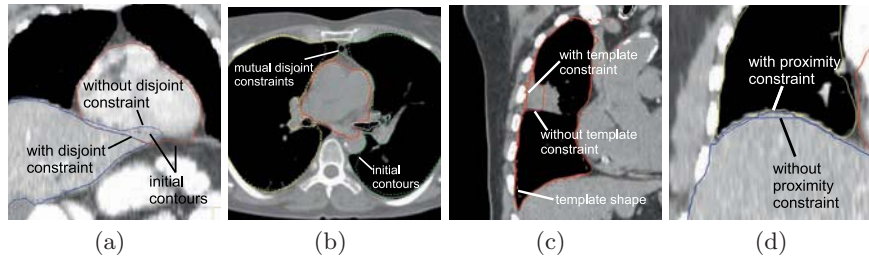
Fig. 3. Effect of the different proposed geometric constraints. Whereas the disjoint constraint (a)+(b) can be used to remove overlaps between initial segmentations, the template constraint (c) can bind the level set zero-crossing to the initial one in a location-specific manner. With the proximity constraint (d), coincidence of shared boundaries can be imposed locally.

testing were used to train the learning-based pipeline, which were 313 boundary annotations for the liver and about 130 annotations for the other organs.

After the the joint gradient descent had converged w.r.t. each energy of the level set segmentation, the final meshes were extracted from the level set maps via Marching Cubes. As error measure we first computed the shortest Euclidean distances between each result mesh and its corresponding annotated mesh at every vertex of the former as well as every vertex of the latter, and then averaged all such distances. The results in Table 1 show that the presented segmentation system yields state-of-the-art accuracies ranging from 1.17mm average surface error for left kidney to 2.89mm for the liver. Thereby the proposed level set segmentation contributes with an improvement of 20% for the liver, 30% for the left lung and left kidney, and 40% for the right lung and right kidney. Average run-times for the full detection pipeline is 2–3 min, which includes about 1 min for the only coarsely multi-threaded level set stage on a 2.0 GHz eight-core Xeon machine.

Table 1. Symmetric surface errors using machine learning-based segmentation and after applying level set-based refinement

Sym. surface error (mm)	Mean	Std.dev.	Median	Worst 80%	# cases
Liver, learning-based only	3.5	1.7	3	4.0	100
Liver, level set-refined	2.9	1.7	2.6	3.6	100
Left lung, learning-based only	2.1	0.5	1.9	2.5	60
Left lung, level set-refined	1.5	0.3	1.4	1.7	60
Right lung, learning-based only	2.7	0.9	2.4	3.0	60
Right lung, level set-refined	1.6	0.6	1.5	1.8	60
Right kidney, learning-based only	1.9	0.9	1.8	2.0	10
Right kidney, level set-refined	1.1	0.9	0.8	3.9	10
Left kidney, learning-based only	1.9	1.0	1.7	2.3	10
Left kidney, level set-refined	1.3	1.0	1.0	1.9	10

4 Conclusion

In our experiments we found the proposed algorithm to combine the robustness and speed of a machine learning approach with the high accuracy and advantageous distance map representation of a level set approach. Furthermore, the novel level set constraints allow to impose region-specific geometrical priors in the refinement stage. Yet the involved localized weights, as well as the global ones, need to be set manually in one or more parameter tuning sessions. One approach to automatize this step could be to minimize a sum energies in Equ. (8) over a set of fixed shapes gained from the learning-based stage with respect to these parameters. Finally, our experimental results show state-of-the-art accuracy and robustness of the proposed algorithm for five different organs on various unseen data sets.

References

1. Chan, T.F., Vese, L.A.: Active Contours without Edges. IEEE Trans. on Image Processing 10(2), 266–277 (2001)
2. Cremers, D., Rousson, M., Deriche, R.: A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape. IJCV 72(2), 195–215 (2007)
3. Kohlberger, T., Uzunbas, G., Alvino, C., Kadir, T., Slosman, D., Funka-Lea, G.: Organ Segmentation with Level Sets Using Local Shape and Appearance Priors. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 34–42. Springer, Heidelberg (2009)
4. Ling, H., Zhou, S.K., Zheng, Y., Georgescu, B., Suehling, M., Comaniciu, D.: Hierarchical, Learning-based Automatic Liver Segmentation. In: CVPR 2008, pp. 1–8. IEEE Press, New York (2008)
5. Liu, D., Zhou, S.K., Bernhardt, D., Comaniciu, D.: Search Strategies for Multiple Landmark Detection by Submodular Maximization. In: CVPR 2010, pp. 2831–2838. IEEE Press, New York (2010)
6. Mansouri, A.-R., Mitiche, A., Vázquez, C.: Multiregion Competition: A Level Set Extension of Region Competition to Multiple Region Image Partitioning. Computer Vision and Image Understanding 101(3), 137–150 (2005)
7. Paragios, N.: A Variational Approach for the Segmentation of the Left Ventricle in Cardiac Image Analysis. IJCV 50(3), 345–362 (2002)
8. Rousson, M., Paragios, N.: Shape Priors for Level Set Representations. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2351, pp. 78–92. Springer, Heidelberg (2002)
9. Tu, Z.: Probabilistic Boosting-Tree: Learning Discriminative Models for Classification, Recognition, and Clustering. In: ICCV 2005, vol. 2, pp. 1589–1596. IEEE Press, New York (2005)
10. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Four-Chamber Heart Modeling and Automatic Segmentation for 3-D Cardiac CT Volumes Using Marginal Space Learning and Steerable Features. IEEE Trans. on Medical Imaging 27(11), 1668–1681 (2008)