A Statistical Deformable Model for the Segmentation of Liver CT Volumes

Tobias Heimann, Hans-Peter Meinzer, and Ivo Wolf

Div. Medical and Biological Informatics, German Cancer Research Center,
69120 Heidelberg, Germany
t.heimann@dkfz.de

Abstract. We present a fully automated method based on an evolutionary algorithm, a statistical shape model (SSM), and a deformable mesh to tackle the liver segmentation task of the MICCAI Grand Challenge workshop. To model the expected shape and appearance, the SSM is trained on the 20 provided training datasets. Segmentation is started by a global search with the evolutionary algorithm, which provides the initial parameters for the SSM. Subsequently, a local search similar to the Active Shape method is used to refine the detected parameters. The resulting model is used to initialize the main component of our approach: a deformable mesh that strives for an equilibrium between internal and external forces. The internal forces describe the deviation of the mesh from the underlying SSM, while the external forces model the fit to the image data. To constrain the allowed deformation, we employ a graph-based optimal surface detection during calculation of the external forces. Applied to the ten test datasets of the workshop, our method delivers comparable results to the human second rater in six cases and scores an average of 59 points.

1 Introduction

The automated segmentation of the liver from clinical CT images is a challenging undertaking: Contrast to neighboring structures as stomach or kidney is often low to non-existent, large variations in shape make an accurate modeling of the population difficult and, last but not least, existing lesions (e.g. tumors) considerably change the appearance of the organ. In this study, we tackle the task using a shape-guided deformable model that we recently published in [1]. The backbone of our method is a statistical model of shape and appearance, similar to the *Active Shape Model* (ASM) introduced by Cootes et al. [2]. In contrast to ASMs, we allow additional deformations of the model – i.e. which have not been captured during the training phase – by employing a deformable surface framework. For details of the method, we refer the reader to the comprehensive publication [1]; below, we will only briefly outline the employed techniques.

T. Heimann, M. Styner, B. van Ginneken (Eds.): 3D Segmentation in The Clinic:
A Grand Challenge, pp. 161-166, 2007.

2 Method

2.1 Construction of the Statistical Model

To construct an SSM from the 20 provided training volumes, the binary segmentations first had to be converted to surfaces. Since all surfaces have to be of the same topology (genus zero in our case), the small tunnels present in some training datasets were closed using a morphological closing operation with an ellipsoidal kernel approximating a sphere of 4mm radius. After running the Marching Cubes algorithm [3] for surface extraction, 2500 landmarks were spread equally on each training surface. The required point correspondences between all shapes were determined automatically using an optimization approach based on the description length of the model [4]. As in most shape modeling approaches, a principal component analysis on the covariance matrix of landmark positions yielded the main modes of variation for the training set, which are used to constrain the deformation of the model during segmentation.

To model the gray-value appearance of the liver, we employed a kNN-classifier trained on intensity profiles as suggested in [5]. At each landmark, profiles perpendicular to the surface were sampled from all training volumes and stored as boundary samples. Shifting the profiles towards the inside and outside of the liver yielded additional non-boundary samples. Using a moderated kNN-classifier, the probability of a profile g lying on the boundary can then be estimated by the number of nearby boundary samples $b_k(g)$ relative to k:

$$p(b|g) = \frac{b_k(g) + 1}{k + 2} \tag{1}$$

In order to improve the performance of the classifier, similar profile models were clustered. While we used the k-Means algorithm in our previous publication [1], we opted for the Mean Shift algorithm [6] this time, mainly due to its ability to determine the optimal number of clusters automatically.

To support a multi-resolution approach during search, appearance models were built for the native image resolution R_0 and for four down-sampled versions (R_1 to R_4). All profiles consisted of seven intensity samples, spaced at 1mm in R_0 (distance doubled with each down-sampling step).

2.2 Global Search Using Evolutionary Algorithm

When analyzing an unseen image, the first step is to determine suitable initial parameters for the SSM. We use an evolutionary algorithm for that purpose, which optimizes a population of 1000 solutions over 40 iterations. A solution consists of three translation values, a scale factor, a rotation angle around the z-axis (to cope with the rotated images) and ten shape parameters. In each iteration, all solutions are evaluated by combining probabilities from relevant appearance models in R_4:

$$w_s = \exp\left(\frac{v}{n}\sum_i \log p_i(b|g_i)\right) \tag{2}$$

Here, n holds the number of employed appearance models and $v = 5$ is a constant that determines the speed of convergence of the search process (see below). After all solutions have been evaluated, random sampling with the weights w_s is used to assemble the population for the next iteration. Note that chosing larger values for v amplifies the probability differences between different solutions and thus influences the amount of well-fitting solutions which is selected in this step. Each chosen solution is mutated by adding Gaussian noise with specific standard deviation to all parameters. The standard deviation is reduced in each optimization step, which makes the search exploit a large space at the beginning and more local regions towards the end.

In order to improve the runtime of the global search, we do not evaluate all appearance models in Eq. 2, but only a subset containing the most reliable ones (approximately 200). Appearance model reliability is estimated by an evaluation on the training data. Mesh decimation techniques ensure that the liver is still adequately sampled by the reduced landmark set.

2.3 Local Search Using Deformable Mesh

The local search is initialized with the best-rated solution from the global search. To improve the initialization, this solution is optimized locally using an ASM search [2] in R_4 until the maximum vertex movement D_{max} is less than 4mm and subsequently in R_3 until $D_{max} < 2$mm. The ASM search determines the optimal position for each landmark (according to the corresponding appearance model) and projects the resulting coordinates into the shape space, so that the result always is a valid model instance.

Once this local initialization has converged, the deformable mesh is initialized with the geometry of the current shape model and a forces-equilibrium evolution is started. In each iteration, all vertices are updated according to the Lagrangian equation of motion, which in our case means the application of regularizing internal forces and data-fitting external forces.

The internal forces strive to keep the deformable mesh close to the best-fitting SSM. They are based on the concepts of tension and rigidity, which have been introduced by Kass et al. [7] for the popular *Snakes* algorithm. In our case of a two-manifold surface, tension forces push or pull the edges in the mesh towards the length of the corresponding edge in the SSM. Since an edge is defined as the line connecting two neighboring vertices, its length is simply the distance between these two points. Rigidity forces act on the angles between neighboring faces, and rotate adjacent faces in the direction to reach the same value as the corresponding angle in the SSM.

The external forces try to move all vertices to the locations where their local appearance model predicts the highest boundary probability. Similar as in the ASM search, only a number of locations perpendicular to the surface are evaluated in each step (alltogether 13 locations spaced at 0.5mm in R_0). In order to improve robustness against outliers, an optimal surface detection [8] is conducted to find the solution with the minimum costs under the hard constraint that neighboring points are not allowed to shift more than a a given distance Δ

Dataset	Overlap Error		Volume Diff.		Avg. Dist.		RMS Dist.		Max. Dist.		Total
	[%]	Score	[%]	Score	[mm]	Score	[mm]	Score	[mm]	Score	Score
1	7.0	73	1.0	95	1.2	70	2.3	69	20.8	73	76
2	10.8	58	-0.7	97	1.8	54	4.1	43	33.6	56	61
3	28.4	0	-24.9	0	8.1	0	16.9	0	75.7	0	0
4	7.4	71	0.5	97	1.3	67	2.8	61	28.6	62	72
5	5.2	80	-1.2	93	0.9	77	2.0	73	24.2	68	78
6	16.3	36	-1.0	95	4.6	0	10.7	0	72.9	4	27
7	11.7	54	5.2	72	2.0	49	4.7	35	34.4	55	53
8	5.8	77	3.1	83	0.9	76	1.7	76	17.0	78	78
9	8.7	66	1.3	93	1.3	69	2.6	64	22.5	70	72
10	8.7	66	-0.5	97	1.4	65	2.9	60	22.7	70	72
Average	11.0	58	-1.7	82	2.4	53	5.1	48	35.2	54	59

Table 1. Results of the comparison metrics and scores for all ten test cases.

against each other ($\Delta = 1$ sample point for R_3 and 2 points for R_2 to R_0). The optimal surface detection efficiently solves this task by converting it to a graph problem, in which the maximum flow has to be determined.

Starting in R_3, the deformable surface is evolved until the average landmark movement falls below a given threshold, subsequently the search continues in R_2 etc. Over the different resolutions, the weight between internal and external forces is shifted, so that the latter ones continuously gain more influence. For a listing of all involved parameters we refer the reader to [1].

3 Results

The construction of the shape and appearance model using the provided 20 training datasets ran fully automated and took approximately twelve hours on a standard PC, mainly for the optimization of point correspondences. The presented segmentation approach was evaluated on the ten test images of the Grand Challenge workshop. On a standard PC, the algorithm ran for approximately ten minutes per image, with the evolutionary algorithm taking the major part of around six minutes. The resulting error metrics and scores are listed in Table 1. Figure 1 shows the obtained segmentations for three cases.

4 Discussion and Conclusion

The average performance of the algorithm and final score is seriously declined by the two poor results of images 3 and 6. We attribute these failed segmentations to the low number of training samples used to build the statistical model. In former experiments [1], we used 32 training samples and reached significantly better results (1.6mm average surface distance instead of 2.4mm). Although the

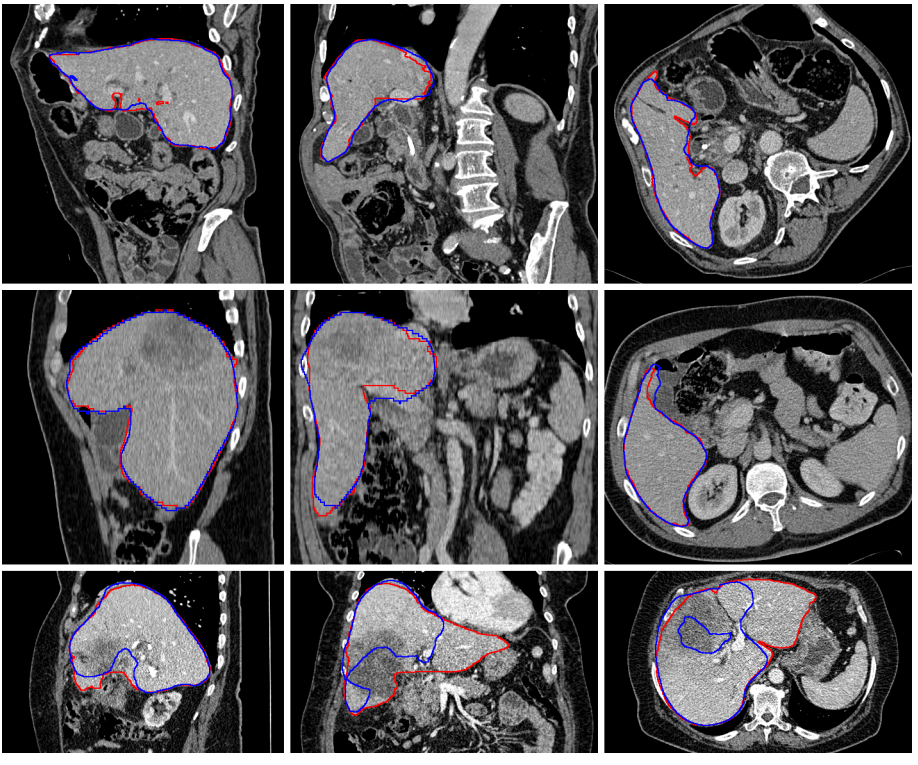

Fig. 1. From left to right, a sagittal, coronal and transversal slice from a relatively easy case (1, top), an average case (4, middle), and a relatively difficult case (3, bottom). The outline of the reference standard segmentation is in red, the outline of the segmentation of the method described in this paper is in blue. Slices are displayed with a window of 400 and a level of 70.

deformable mesh relaxes the strict contraints of the SSM and theoretically allows the adoption to arbitrary shapes, the experiments in this paper demonstrate that in practice, this works only if the SSM is able to at least roughly represent the shape in question. If this is not the case, the deformable mesh may even self-intersect (as it happened for image 6), which makes the resulting segmentation practically unusable. An additional problem for image 3 was the dark intensity of the large tumor: This type of boundary appearance was not present in the training data, thus the appearance model could not locate the boundary correctly (see Fig. 1, lower row).

On the positive side, our method reached results comparable to a human second rater (i.e. around 75 points) on six out of ten images. By enlarging the size of the training set, we hope to be able to increase the number of successful segmentations. The time of ten minutes required to segment one dataset is completely acceptable for the application of surgery planning, which is our current

focus. In case of other applications with less time available, the method could be sped up by e.g. using a different approach for initialization. In the future, we plan to implement a method for intersection detection and prevention, which should improve the performance for cases where even an expanded SSM (based on more samples) cannot approximate the true shape. All in all, we are optimistic that given enough training samples and some technical tweaks, our method will be able to segment 90 percent of clinical cases with an accuracy comparable to manual raters. Computer-based diagnosis and operation planning would highly benefit from these advances.

Acknowledgements

This research has been supported by the German Research Foundation DFG under grant WO 1218/2-1.

References

1. Heimann, T., Münzing, S., Meinzer, H.P., Wolf, I.: A shape-guided deformable model with evolutionary algorithm initialization for 3D soft tissue segmentation. In: Proc IPMI. Volume 4584 of LNCS., Springer (2007) 1–12
2. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models – their training and application. Comput Vis Image Underst **61**(1) (1995) 38–59
3. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3D surface construction algorithm. In: SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques, New York, NY, USA, ACM Press (1987) 163–169
4. Heimann, T., Wolf, I., Meinzer, H.P.: Automatic generation of 3D statistical shape models with optimal landmark distributions. Methods Inf Med **46**(3) (2007) 275–281
5. de Bruijne, M., van Ginneken, B., Viergever, M.A., Niessen, W.J.: Adapting active shape models for 3D segmentation of tubular structures in medical images. In: Proc IPMI. Volume 2732 of LNCS., Springer (2003) 136–147
6. Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell **24**(5) (2002) 603–619
7. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. Int J Comp Vis **1**(4) (1988) 321–331
8. Li, K., Millington, S., Wu, X., Chen, D.Z., Sonka, M.: Simultaneous segmentation of multiple closed surfaces using optimal graph searching. In: Proc IPMI. Volume 3565 of LNCS., Springer (2005) 406–417