A Generic Probabilistic Active Shape Model for Organ Segmentation

Andreas Wimmer[1,2], Grzegorz Soza[2], and Joachim Hornegger[1]

[1] Chair of Pattern Recognition, Department of Computer Science,
Friedrich-Alexander University Erlangen-Nuremberg
`andreas.wimmer@informatik.uni-erlangen.de`
[2] Siemens Healthcare Sector, Computed Tomography, Forchheim, Germany

Abstract. Probabilistic models are extensively used in medical image segmentation. Most of them employ parametric representations of densities and make idealizing assumptions, e.g. normal distribution of data. Often, such assumptions are inadequate and limit a broader application. We propose here a novel probabilistic active shape model for organ segmentation, which is entirely built upon non-parametric density estimates. In particular, a nearest neighbor boundary appearance model is complemented by a cascade of boosted classifiers for region information and combined with a shape model based on Parzen density estimation. Image and shape terms are integrated into a single level set equation. Our approach has been evaluated for 3-D liver segmentation using a public data base originating from a competition (`http://sliver07.org`). With an average surface distance of 1.0 mm and an average volume overlap error of 6.5 %, it outperforms other automatic methods and provides accuracy close to interactive ones. Since no adaptions specific to liver segmentation have been made, our probabilistic active shape model can be applied to other segmentation tasks easily.

1 Introduction

Probabilistic modeling plays an important role in medical image segmentation. When dealing with images of complex anatomical structures, often impaired with noise or low contrast, it is convincing to search for the most likely boundary, region, or shape. In a related fashion, active shape models [1,2] use statistical knowledge about an object's shape and local gray level appearance for segmentation, traditionally assuming normal distribution of both. Deciding for a certain parametric distribution may be necessary because the model is not manageable without, or because little knowledge is available about the true nature of data. Often, a model works under the given assumptions for a specific problem, but a generalization to other fields of application is not possible.

In order to overcome such limitations, we propose a novel active shape model, which is entirely built upon non-parametric estimates of probabilities. In particular, we use a nearest neighbor appearance model of the organ boundary, a cascade of boosted classifiers to detect the object region, and Parzen density estimation

to deduce a shape model. No assumptions about the underlying distribution of data are made, and no specific training, such as establishing correspondences between landmark points, is required. All image and shape terms are combined in a single equation within the level set framework [3]. Employing an implicit shape representation moreover avoids a parameterization of the target shape. While the proposed model is highly generic, it can nevertheless compete with or even outperform methods tailored to a specific problem, which is demonstrated for liver segmentation using a public data base [4].

The remainder of this paper is organized as follows. In the next section, related work is reviewed. The proposed boundary and region models are described in Sect. 3, along with the shape model. In Sect. 4, the application of the active shape model to liver segmentation is detailed, and evaluation results are presented. The paper concludes with a discussion of the obtained results and directions for further research.

2 Related Work

At the 2007 MICCAI conference, a competition for liver segmentation was held at a conference workshop [4,5]. The majority of automatic methods ranking in the highest quantile incorporated shape knowledge. They were mostly based on active shape models (ASM) [1], such as the approach proposed in [6], where a heuristic intensity model is used for incorporating image information.

In order to increase flexibility, a recursive subdivision of the liver shape into smaller patches was proposed in [7]. Individual models are trained, and during segmentation, an adhesiveness constraint ensures that patches overlap smoothly.

The ASM principle has been widely extended in [8] with marginal space learning and steerable features for rapidly detecting the liver pose and initializing the first modes of the shape model. Segmentation is then performed in a hierarchical manner from coarse to fine scales, with image information being inferred from a probabilistic boosting tree.

3 Probabilistic Active Shape Model

We segment an organ by evolving an active surface, which is implicitly described through the zero level set of a level set function $\phi(\boldsymbol{x})$, with $\boldsymbol{x} \in \mathbb{R}^3$ [3]. In contrast to explicit shape representations, e.g. through meshes, the implicit representation is parameter free and consequently does not depend on the target shape. Furthermore, when deforming the surface through the level set framework, self intersections do not pose a problem.

A popular edge based level set segmentation method is geodesic active contours [9], which is associated with the following partial differential equation to track the evolving boundary over time.

$$\frac{\partial}{\partial t}\phi = g\left|\nabla\phi\right|\operatorname{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \nabla g \cdot \nabla\phi \tag{1}$$

The first term accounts for the smoothness of the boundary, while the second part constitutes an advection term. The stopping function g is the reciprocal of an edge detector, which in turn depends on the image I. A common choice for g is $1/\left(1 + |\nabla I|^2\right)$.

3.1   Boundary Model

In most cases, a simple edge detector based on the image gradient is not sufficient to deal with poorly discriminable tissues or noise. In contrast, the information contained in an intensity profile sampled at the organ boundary is much richer and more specific. An appearance model may be obtained by sampling profile vectors perpendicular to the boundary in a training stage. Traditionally [2], profiles were assumed to be normally distributed. In that case, the appearance model is given by the mean vector and covariance matrix. However, a normal distribution of profiles often is not present. In Fig. 1(a), profiles sampled from the liver boundary in computed tomography (CT) images are shown. The mean vector is not characteristic for the sample set.

As a non-parametric alternative, we instead build a nearest neighbor appearance model of the organ boundary. This approach, originally proposed in [10], allows capturing arbitrary distributions of profile vectors. During the training stage, additional false profiles are sampled off the true organ boundary. In order to evaluate the boundary probability, the k nearest neighbors found in the training set are determined using the L_2 norm. Amongst these, the number of true boundary profiles divided by k yields the desired probability. How to incorporate this procedure into a level set evolution has been described in [11]. For each point which is part of the narrow band around the active surface, an intensity profile is sampled. The profile orientation is given by the gradient of the level set function $\nabla \phi$, which is collinear to the normal. The boundary probability is then determined by finding the nearest neighbors in the training set and calculating the ratio mentioned before. The result is a probability image, from which a stopping function is calculated, as described for (1).

3.2   Region Model

Since the capture range of the boundary model is limited, an inaccurate initialization can result in slow convergence to the targeted boundary. In the worst case, the algorithm will stop at a local extremum. We therefore complement the boundary model with a region model, which leads to a global optimization.

Organs may have inhomogeneous intensities and contain tumors and vessels filled with contrast agent. Hence, a region model should be able to represent complex intensity distributions. Since distributions of different tissues may overlap, it is also advantageous to incorporate contextual information. As an example, high intensity values may be observed for voxels within vessels and for voxels within bones. While in the former case, the voxel shall be included in the segmentation if surrounded by organ tissue, in the latter case it shall be excluded.
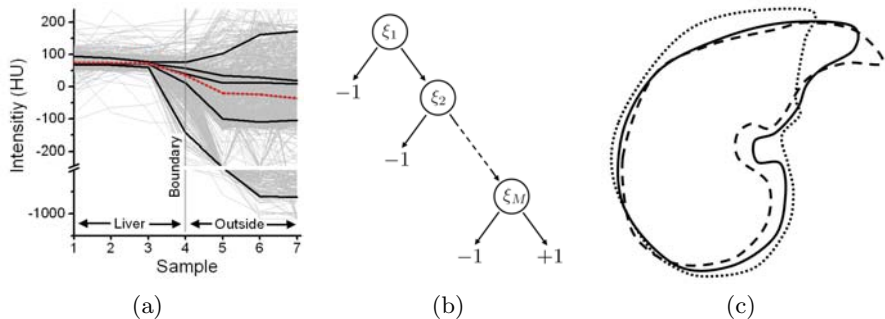
Fig. 1. (a) Profiles sampled from the liver boundary in CT images. The intensity scale is given in Hounsfield units (HU). Characteristic modes are highlighted (solid lines). The mean profile (dotted line) is not representative. (b) Cascade of boosted classifiers. At each stage ξ_i, a sample is either rejected (-1) or passed on to the next stage. Samples accepted at the final stage $(+1)$ are considered to be part of the organ. (c) Principle of the kernel density shape model depicted for the 2-D case. Evolving active contour (solid line) and reference shapes (dashed lines) are rigidly aligned. The active contour is moved towards each reference shape, weighted according to their similarity.

How to define meaningful features and train a classifier to meet these requirements is not evident in advance. We therefore build a cascade χ of M boosted classifiers, as illustrated in Fig. 1(b). At each stage $i = 1, 2, \ldots, M$, a strong classifier ξ_i either rejects a sample or passes it on to the next stage. Only when the sample is accepted at the final stage, it is considered to be part of the organ. A strong classifier ξ_i is a weighted combination of weak classifiers and is determined by the AdaBoost [12] algorithm. Classification and regression trees (CART) [13] are used here as weak classifiers in order to further increase flexibility and handle complex region statistics. Within a CART, the decision which branch to take is based on Haar-like features [14]. They resolve to simple intensity differences, which are defined in large combinatorial numbers for a neighborhood around the sample voxel. Thereby, contextual information is integrated in the region model. While training a cascade of boosted classifiers may take several days, the application is extremely fast through the use of integral images. Further details can be found in [14]. Our implementation is based on OpenCV[1].

The result $\chi(\boldsymbol{x})$ of the cascade is $+1$, if the sample voxel $\boldsymbol{x}$ is accepted and -1, if it is rejected. This region information is incorporated as a dilatation or erosion of the surface into the level set equation (1), which leads to

$$\frac{\partial}{\partial t}\phi = g\,|\nabla\phi|\,\mathrm{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \nabla g \cdot \nabla\phi + \chi g\,|\nabla\phi|\,. \tag{2}$$

---

[1] http://opencvlibrary.sourceforge.net

3.3   Shape Model

We further increase the robustness of the proposed active surface by incorporating a shape model. Shape knowledge is highly valuable in regions with ambiguous or missing image information. For example, the boundaries between liver and stomach or liver and muscle tissue may be virtually non-existent, depending on the level of contrast enhancement. When the surface is about to leak into other tissue, shape knowledge allows constraining such an unlikely deformation.

Most ASMs assume a normal distribution of shapes. The according shape space captured by the model is spanned by the mean shape and a linear combination of Eigen modes, which are obtained through principal component analysis performed on a training set [1]. We pursue the more generic non-parametric approach of [15], which is to determine the probability of observing a shape through Parzen density estimation [16]. The employed distance metric for shapes represented through level set functions is

$$\Delta\left(\phi_1, \phi_2\right) = \int_{\boldsymbol{x}} \left(H\left(\phi_1\right) - H\left(\phi_2\right)\right)^2 d\boldsymbol{x}. \tag{3}$$

$H\left(.\right)$ denotes the Heaviside function, which is 1 if the argument is non-negative and 0 otherwise. When evaluating the distance, both shapes have to be aligned. This is achieved intrinsically by normalizing the center of gravity, scaling, and principal axes of both shapes, i.e. of $H\left(\phi_1\right)$ and $H\left(\phi_2\right)$.

Given a set of N reference shapes $\{\phi_i\}_{i=1}^N$, a Parzen estimate of the probability of a shape ϕ is given by

$$P\left(\phi\right) \propto \frac{1}{N} \sum_{i=1}^N K_\sigma\left(\Delta\left(\phi, \phi_i\right)\right), \tag{4}$$

where K_σ is a Gaussian kernel, whose standard deviation σ is chosen according to the nearest neighbor distances: $\sigma^2 = \frac{1}{N} \sum_{i=1}^N \min_{j \neq i} \Delta^2\left(\phi_i, \phi_j\right)$. By optimizing this shape probability with respect to ϕ, a constraint is derived for the level set equation which evolves the surface towards similar shapes. The principle is illustrated in Fig. 1(c). Instead of maximizing (4), a minimization problem is obtained by taking the negative logarithm. Calculating the functional derivative of $-\log P\left(\phi\right)$ with respect to ϕ leads to a shape constraint term, which extends the active surface of (2) to a probabilistic active shape model.

$$\frac{\partial}{\partial t}\phi = g\left|\nabla\phi\right| \operatorname{div}\left(\frac{\nabla\phi}{\left|\nabla\phi\right|}\right) + \nabla g \cdot \nabla\phi + \chi g\left|\nabla\phi\right| + \frac{\sum_{i=1}^N K_\sigma\left(\Delta\left(\phi, \phi_i\right)\right)\frac{\partial}{\partial\phi}\Delta^2\left(\phi, \phi_i\right)}{2\sigma^2 \sum_{i=1}^N K_\sigma\left(\Delta\left(\phi, \phi_i\right)\right)} \tag{5}$$

We refer to [15] for further details on the derivation of the shape constraint.

4   Evaluation

The proposed algorithm has been evaluated for liver segmentation from CT scans. Both training and test images were taken from a public data base, which

had been assembled for a segmentation competition held at a workshop of the 2007 MICCAI conference [4,5]. Since ground truth segmentations are not provided for test images and the evaluation is conducted by the organizers, results are meaningful and comparable. The 20 training and 10 test images have a slice resolution of 512×512 voxel and 64 to 502 slices. Intra slice spacing varies between 0.56 mm $\times$ 0.56 mm and 0.86 mm $\times$ 0.86 mm, inter slice spacing varies between 0.7 mm and 5 mm. The data base comprises both normal and abnormal liver shapes with tumors and varying levels of contrast enhancement.

The boundary appearance model was built by sampling intensity profiles of length 7 with 1 mm spacing. In addition to true boundary profiles, negative examples were obtained with offsets of ± 1 mm and ± 3 mm. During segmentation, the $k = 10$ nearest neighbors were used to estimate the boundary probabilities. A cascade with $M = 5$ stages was trained for the region model. With the inter slice spacing being large compared to the intra slice spacing, we decided for $2-$D Haar-like features. In order to train with a large context while at the same time preventing the classifier from over fitting, we conducted training and application for slices down sampled to a resolution of 128×128 voxel. Using a window of 12×12 voxel, the Haar feature pool consisted of about 16000 features. In contrast to PCA based approaches, no training of the shape model is required here.

For segmentation, the initial pose of the shape model was determined by extracting the largest object detected by the region model. In order to accelerate convergence, segmentation was performed in a multi-scale manner, starting with an image down sampled by a factor of 4. The whole process ran without user interaction and took about 3 min. per image on a 3 GHz CPU.

The obtained segmentation results were submitted to the organizers of the competition and evaluated with respect to volume overlap and difference as

Table 1. Results of the evaluation metrics and scores for all test cases. The maximum score, corresponding to a segmentation identical with the reference, would be 100. While interactive systems have reached scores in the range of 73 to 82, the majority of automatic systems falls within the range of 52 to 69. See http://sliver07.org for all results and [4,5] for details on the metrics and the score scale.

| Case	Ovrl. Error [%]	Score	Vol. Diff. [%]	Score	Avg. Dist. [mm]	Score	RMS Dist. [mm]	Score	Max. Dist. [mm]	Score	Total Score
1	6.41	74.98	1.16	93.82	1.02	74.51	1.99	72.32	18.61	75.51	78.23
2	6.41	74.95	-0.64	96.59	0.98	75.62	2.31	67.93	22.29	70.68	77.15
3	5.10	80.08	0.54	97.11	1.00	75.10	1.83	74.65	15.82	79.19	81.23
4	7.27	71.62	4.31	77.09	1.20	70.05	2.49	65.47	28.99	61.86	69.22
5	5.61	78.10	1.86	90.13	0.96	76.04	2.12	70.55	21.91	71.17	77.20
6	7.99	68.77	-3.02	83.94	1.27	68.25	2.37	67.14	18.28	75.95	72.81
7	5.04	80.33	2.79	85.17	0.73	81.71	1.42	80.25	13.00	82.89	82.07
8	7.06	72.42	3.45	81.65	1.18	70.59	2.17	69.90	14.16	81.37	75.19
9	6.79	73.49	3.85	79.55	0.84	78.97	1.47	79.64	16.01	78.94	78.12
10	7.04	72.52	-3.92	79.17	1.01	74.76	1.80	75.01	14.15	81.38	76.57
Avg.	6.47	74.72	2.55	86.42	1.02	74.56	2.00	72.29	18.32	75.89	76.78
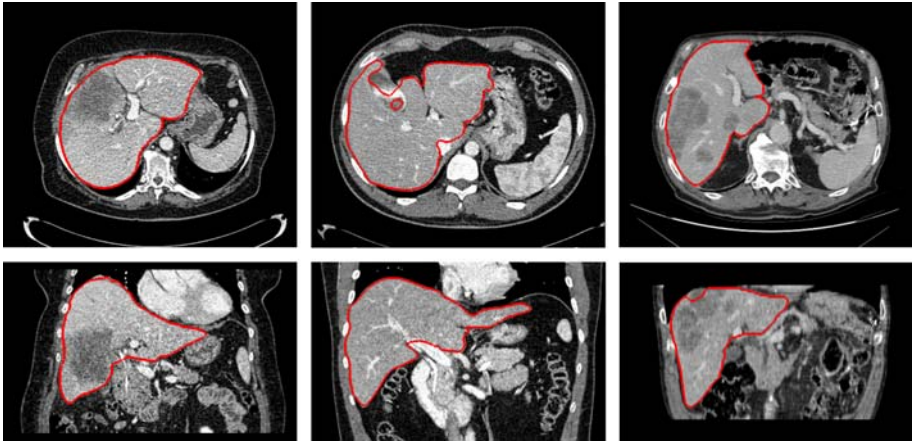
Fig. 2. From left to right, segmentation results for cases 3, 6, and 10 in transversal (top) and coronal (bottom) views. The large tumor in case 3 and the majority of tumors in case 10 were correctly included in the segmentation. For case 6, the shape model was too restrictive in the region of the gall bladder, which therefore was not completely excluded. What appears as a circle in the transversal view is actually a bulge in 3-D.

well as symmetric average, root mean squared, and maximum surface distance. Results are listed in Table 1 and displayed in Fig. 2 for several data sets.

5   Discussion and Conclusion

We have presented a novel probabilistic active shape model for organ segmentation, which combines boundary, region, and shape information in a single level set equation. Non-parametric estimates are used for all involved densities, which leads to a generic approach that can be applied to various segmentation tasks.

The proposed algorithm has been evaluated for liver segmentation from CT images using a public data base (cf. Sect. 4). As shown in Table 1, our method achieved constantly high scores for all test cases. On average, the surface distance is 1.0 mm and the overlap error is 6.5 %, which is close to interactive methods [4,5] and meets clinical requirements.

Failure to exclude the portal vein resulted in a higher overlap error and surface distance in case 4. For some cases where heart and liver share similar levels of contrast enhancement, the region term was not able to discriminate between both of them, leading to slight over segmentation indicated by positive volume differences in Table 1.

The advantages of our non-parametric approach are most evident when dealing with images that differ from the ordinary. For example, case 6 was accurately segmented except for a small part of the gall bladder, although the liver is unlike most training shapes. In contrast, ASM approaches perform considerably worse for organs significantly different from the mean, since the underlying PCA inherently

assumes a normal distribution of shapes. This can be observed also for the top ranking method of the competition [6], even though it was trained with an extensive data base of 112 images [5], much more than the 20 images we used.

For the future, we are confident that results can be further improved by increasing the training set, from which all three terms will benefit. In addition, we plan to apply the proposed approach to the segmentation of other structures. Preliminary results show a high accuracy also for the segmentation of the nucleus caudate. No changes to the core system were made, which emphasizes the broad applicability of our method.

References

1. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models – their training and application. CVIU 61(1), 38–59 (1995)
2. Cootes, T.F., Taylor, C.J.: Statistical models of appearance for medical image analysis and computer vision. In: SPIE Medical Imaging, vol. 4322, pp. 236–248 (2001)
3. Sethian, J.A.: Level Set Methods and Fast Marching Methods, 2nd edn. Cambridge University Press, New York (1999)
4. Heimann, T., Styner, M., van Ginneken, B.: 3D Segmentation in the Clinic – A Grand Challenge. In: MICCAI Workshop Proceedings (2007)
5. Heimann, T., van Ginneken, B., Styner, M., et al.: Comparison and evaluation of methods for liver segmentation from CT datasets. In: IEEE TMI (2009)
6. Kainmüller, D., Lange, T., Lamecker, H.: Shape constrained automatic segmentation of the liver based on a heuristic intensity model. In: 3D Segmentation in the Clinic – A Grand Challenge, pp. 109–116 (2007)
7. Okada, T., Shimada, R., Sato, Y., Hori, M., Yokota, K., Nakamoto, M., et al.: Automated segmentation of the liver from 3D CT images using probabilistic atlas and multi-level statistical shape model. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007, Part I. LNCS, vol. 4791, pp. 86–93. Springer, Heidelberg (2007)
8. Ling, H., Zhou, S.K., Zheng, Y., Georgescu, B., Suehling, M., Comaniciu, D.: Hierarchical, learning-based automatic liver segmentation. In: CVPR (2008)
9. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. International Journal of Computer Vision 22(1), 61–79 (1997)
10. van Ginneken, B., Frangi, A.F., Staal, J.J., ter Haar Romeny, B.M., Viergever, M.A.: Active shape model segmentation with optimal features. IEEE Transactions on Medical Imaging 21(8), 924–933 (2002)
11. Wimmer, A., Soza, G., Hornegger, J.: Implicit active shape model employing boundary classifier. In: ICPR (2008)
12. Freund, Y., Schapire, R.E.: A decision-theoretic generalization of on-line learning and an application to boosting. JCSS 55(1), 119–139 (1997)
13. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Trees. Chapman & Hall, New York (1984)
14. Viola, P., Jones, M.: Robust real-time face detection. International Journal of Computer Vision 57(2), 137–154 (2004)
15. Cremers, D., Osher, S.J., Soatto, S.: Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. IJCV 69(3), 335–351 (2006)
16. Parzen, E.: On the estimation of a probability density function and the mode. Annals of Mathematical Statistics 33, 1065–1076 (1962)