2014 22nd International Conference on Pattern Recognition

Using Object Probabilities in Deformable Model Fitting

Christoph Jud, Thomas Vetter
Department of Mathematics and Computer Science
University of Basel
Basel, Switzerland
Email: christoph.jud@unibas.ch, thomas.vetter@unibas.ch

***Abstract*—We present a novel image segmentation method based on statistical shape model fitting. Instead of fitting the model to raw intensity values we consider object probabilities. The abstraction from the plain intensity images to probability maps makes the segmentation more robust against misleading texture inside the object or surrounding background. The target object probability is predicted based on random forest regression trained with neighborhood dependent features of sample images. In contrast to similar approaches, both, the object boundary as well as the whole object and background region are considered for segmentation. We apply our approach to a 3D cone beam computed tomography image dataset of the jaw region where we segment the wisdom tooth shape. Compared to a boundary- and a region-based method we obtain superior segmentation performance.**

***Keywords*—medical image segmentation, statistical shape model, nonparametric appearance model, random forest regression**

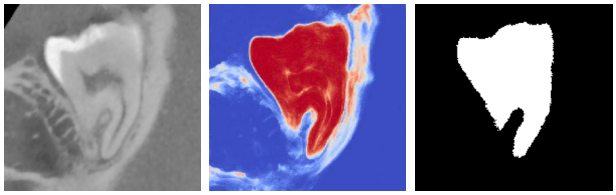
I. Introduction

Object segmentation is a frequent task in medical image analysis. To cope with background clutter, missing edges or noise in the image, statistical shape priors have become indispensable. However, if objects have heterogeneous intensity values or share intensity values with other objects, their assignment based on plain intensities is ambiguous.

In this paper, we present a deformable model-based image segmentation method where the shape model is fitted to object probabilities instead of raw intensity values. This makes the segmentation more robust to misleading texture inside the object or background clutter. Formally, the segmentation problem can be formulated as follows: A reference label map $L_R : \mathbb{R}^d \to \{0, 1\}$ is deformed by a sought deformation field $u : \mathbb{R}^d \to \mathbb{R}^d$, such that it matches the unknown target label map L_T, which assigns each point of a corresponding target image $I_T : \mathbb{R}^d \to \mathbb{R}$ as belonging to object or background. Since L_T is not known, we predict the object probability from the target image I_T and represent the resulting predicted target shape by a probabilistic shape function $P_T : \mathbb{R}^d \to [0, 1]$, where $P_T(x)$ denotes the probability of a point $x \in \mathbb{R}^d$ being part of the object of interest. We formulate the joint minimization functional for segmentation

$$\hat{u} := \arg\min_{u \in \mathcal{U}} \mathcal{D}[P_T, L_R, u] + \gamma \mathcal{R}[u], \tag{1}$$

where $\mathcal{D}$ is a similarity term, which quantifies the matching between the deformed reference label map L_R and the predicted target probability map P_T. $\mathcal{R}$ is a regularizer, which measures how well the current solution u fits to our prior assumption about the space of admissible deformations $\mathcal{U}$.



CT Image Probability Map Label Map

Fig. 1. Slice through a test sample where the whole tooth should be segmented from the jaw bone. (a) In the CT image, the dark pulp structure inside the tooth as well as the bright tooth enamel are visible. The tooth's dentin shares intensity values with the jaw bone. (b) Probability map of the object prediction (red=1, blue=0). (c) Ground truth label map.

The important part of the data term is how to obtain P_T out of the target image I_T. We propose to learn such a transformation from training images in a discriminative manner using random forest regression. This has the advantage, that not only intensity distributions but also higher level structures using advanced image features can be learned from sample data. These are, for example, parts inside the object or surrounding background which have no well-defined correspondence over all examples i.e. parts whose variability cannot be explained by the deformable model.

The shape variation is modeled by the reference label map L_R which is deformed by a parametrized deformation field of the form $u_\alpha(x) = \mu(x) + \sum_i \alpha_i \phi_i(x)$ which models the space of admissible deformations. The minimization problem becomes a parametric problem

$$\hat{\alpha} := \arg\min_{\alpha \in \mathbb{R}^n} \mathcal{D}[P_T, L_R, u_\alpha] + \gamma \mathcal{R}[\alpha]. \tag{2}$$

While this parametrization is very general, we focus on basis functions ϕ_i, which are learned from nonrigidly registered training label maps using principal component analysis. As such, the space of admissible deformations is restricted to the span of the training samples, which makes the optimization robust to noise and boundary insufficiencies, as for instance missing edges.

Our approach is mainly inspired by Cremers et al. [1], where the probabilistic definition of shape is introduced. They use a linear model of probabilistic training shapes as a prior and derive a convex energy term for segmentation. In our work

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.570

however, we consider a deformation model on a reference shape, which leads to a nonconvex optimization, but allows to generate shapes of a broader class of objects. Furthermore, in [1] color histograms are learned to predict object probabilities, whereas in [2] conditional random fields are used for capturing local pixel dependencies for prediction.

In parametric approaches, based on point distributions [3]–[5], appearance profiles orthogonal to the object boundary drive the model fitting. One problem common to these approaches is, that they do not model object and background explicitly since they only consider information around the object boundary. Region based appearance methods [3], [6], [7] model the whole inside of the object using the learned intensity or appearance distribution of the object from training samples. However, no background is modeled. In [8], the background is considered, but only based on few feature landmarks. In [9], the shape evolution is driven by point-wise random forest regression. However, in our method, no explicit shape model is needed. Atlas-based segmentation methods [10]–[12] try to construct a reference shape having a representative appearance, which is fitted to unseen data. An advantage of atlas-based approaches is that they are straight forward to optimize. However, an atlas has to be selected or constructed and it is difficult to cover the background in an atlas. Geometric approaches [13] are topology independent and have been very successful. However, because of a large degree of freedom, the optimization is not straight forward and computationally intensive [14]. For a comprehensive overview about deformable model-based image segmentation we refer to the review of Gavriil Tsechpenakis [15].

We tested our method on a cone beam computed tomography (CBCT) image dataset of the jaw region for segmenting the wisdom tooth. A challenging background is the jaw bone, which is difficult to distinguish from the tooth's roots, because of lack of intensity or texture contrast. Furthermore, inside the tooth there is a misleading pulp structure which is difficult to jointly model with the tooth shape variation. In Fig. 1, a typical example is shown. Among others, for the object prediction, we use Harmonic Filters [16] as feature extraction method. We show a significant performance gain using our method compared to the Active Shape Model of Cootes et al. [3] and the intensity model of Cremers et al. [1].

II. Method

A. Similarity Measure

For the Functional (2), we define a (dis)-similarity term $\mathcal{D}$, which consists out of the following parts. The object and background prediction for each point in the image is estimated separately and treated as independent. Hence, for the full object resp. background probability, we integrate over the logarithms of all object resp. background points. Additionally, as object boundary term, a weighted total variation norm [17] is added to the metric. An edge indicator serves as the weighting.

Let the deformed reference label map $L_R(x+u(x))$ be denoted as L_u and P_{ob}, P_{bg} as the object and background probability maps of the target image I_T. We define the similarity measure

$$\mathcal{D}[u] = -\lambda \int_{\mathbb{R}^d} \log\left(P_{ob}(x)\right) L_u(x) dx \tag{3}$$

$$-\lambda \int_{\mathbb{R}^d} \log\left(P_{bg}(x)\right) \left(1 - L_u(x)\right) dx \tag{4}$$

$$+ (1-\lambda) \int_{\mathbb{R}^d} \frac{1}{1+|\nabla P_{ob}(x)|} |\nabla L_u(x)| dx, \tag{5}$$

where $|\nabla \cdot|$ denotes the gradient magnitude. Since L_u is a binary label map, the first term (3) measures the segmentation quality on the region inside the object, (4) on the background and the last term (5) along the boundary of the predicted shape. The trade-off between the region terms and the boundary is controlled by λ.

B. Random Forest Regression

The important part to define is how P_{ob} and P_{bg} are predicted, given a target image I_T. Let a list of n training images $\{I_i\}_{i=1}^n$ and corresponding ground truth label maps $\{L_i\}_{i=1}^n$ be given. Let further $F_i : \mathbb{R}^d \to \mathbb{R}^m$ be a feature image of I_i obtained by a general feature extraction method. Hence, for each point in the image, there is n-times an m-dimensional feature vector with a corresponding binary label indicating the point being object or background. With this training data the object and background prediction can be learned by any regression method. In this paper, we use random forest regression. The obtained prediction accuracy hinges on the choice of the extracted features. If only the image intensities are considered, the intensity distribution (histogram) is estimated. However, more advanced features as e.g. Laplacian of Gaussians, Sobel edges or Harmonic Filters [16] incorporate point neighborhoods and are able to capture long range structural dependencies of different parts of the object which improves the prediction performance (see Result Section III).

C. Statistical Deformation Model

Given a training set $\mathcal{X} = \{u_1, \ldots, u_n\}$ of deformation fields, the variation of shape $L_R(x+u_\alpha(x))$ is modeled by a linear model

$$u_\alpha(x) = \mu(x) + \sum_{i=1}^{n} \alpha_i \phi_i(x), \tag{6}$$

where μ is the mean deformation $\mu(x) = \frac{1}{n}\sum_{i=1}^n u_i(x)$ and $\phi_i : \mathbb{R}^d \to \mathbb{R}^d$ are the eigenmodes of the training set $\mathcal{X}$. With this parametrization over α, a simple Gaussian prior is used as regularization $\mathcal{R}[\alpha] = \|\alpha\|^2$. The joint functional (2) is minimized with respect to the parameters α using the following energy gradient: Let rewrite $\mathcal{D}$ as

$$\mathcal{D}[u] = \lambda \int_{\mathbb{R}^d} P(x) L_R(x+u(x)) dx \tag{7}$$

$$+ (1-\lambda) \int_{\mathbb{R}^d} Q(x) |L_R(x+u(x))| dx, \tag{8}$$

where $P(x) = \log \frac{P_{bg}(x)}{P_{ob}(x)}$ and $Q(x) = \frac{1}{1+|\nabla P_{ob}(x)|}$. Let further $(\nabla L_R)(x+u(x))$ be denoted as $(\nabla L)_u(x)$ and $(\nabla^2 L_R)(x+u(x))$ as $(\nabla^2 L)_u(x)$. The energy gradient becomes

$$d\mathcal{D}[u] = \lambda \cdot d\mathcal{D}_A[u] + (1-\lambda) \cdot d\mathcal{D}_B[u], \tag{9}$$

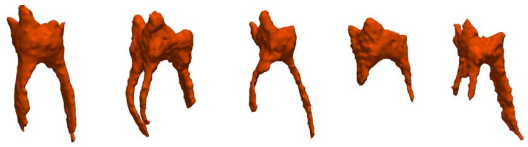


Fig. 2. Sample pulp shapes.

where the areal part

$$d\mathcal{D}_A[u] = \int_{\mathbb{R}^d} P(x)(\nabla L)_u(x)\mathbf{J}_u(x)dx \tag{10}$$

and the boundary part

$$d\mathcal{D}_B[u] = \int_{\mathbb{R}^d} Q(x)\frac{(\nabla L)_u(x)\mathbf{J}_u(x)}{|(\nabla L)_u(x)\mathbf{J}_u(x)|} \cdot \Big(((\nabla^2 L)_u(x)\mathbf{J}_u(x))^T\mathbf{J}_u(x) + (\nabla L)_u(x)\mathbf{J}_u(x)\mathbf{H}_u(x)\Big)dx. \tag{11}$$

$\mathbf{J}$ denotes the Jacobian. Since u is linear in the parameters α, the Hessian $\mathbf{H}$ vanishes and the derivative can be efficiently computed.

Having given n training label maps $\{L_i\}_{i=1}^n$, we switch their representation to signed distance maps $\{S_i\}_{i=1}^n$ where the zero level set is the object boundary. Using nonrigid Bspline registration (see Rueckert et al. [18]) the shapes are matched to a reference S_R such that $S_i(x + u_i(x)) \approx S_R(x), \forall x \in \mathbb{R}$. A reference sample is selected from the training distance maps to which all other training samples are registered. The obtained average deformation is applied to the reference label, resulting in an average label map which again is used in the registration as final reference for all training samples, similar to the atlas construction in [12].

III. Results

In this section, we evaluate the segmentation performance of our method. We show the performance obtained by our object prediction approach and the impact of choosing an arbitrary sample as reference in comparison to the above mentioned average reference. Further, we compare our approach with the classical Active Shape Model [3] which serves as baseline. Moreover, we show comparisons with the method of Cremers et al. [1] and an extension of their approach using the object prediction of our method. We evaluate the results using two quantitative measures. We sample the boundary contour of the ground truth label maps as well as the ones of our segmentation results to get meshes, which can be compared. As a quantitative measure we apply the bidirectional local distance (BLD) which was proposed by Kim et al. [19] in order to compare segmentations. Additionally, we evaluate our segmentation performance using the dice coefficient between the ground truth and the segmentation label maps. The statistical deformation model is implemented using *Statismo* [20].

A. Description of the Data

We performed all experiments on a dataset of 26 training and 20 test CBCT images of the jaw region with the wisdom tooth as the object of interest. For all images a ground truth label map is given by experts for following labels: whole tooth, pulp (nerve structure inside the tooth) and the joint structure dentin/enamel (whole tooth except pulp) including the ones of neighboring teeth. The dataset is restricted to double-rooted teeth. The images have been similarity aligned and their histograms aligned to a training sample. The pulp structure is jointly modeled as tooth shape together with dentin and enamel, because the different number of tracts and hunches makes the registration of the pulp difficult. The point-wise correspondence assumption for this kind of structure is incongruous. In Fig. 2, sample pulp shapes are shown to illustrate this issue.

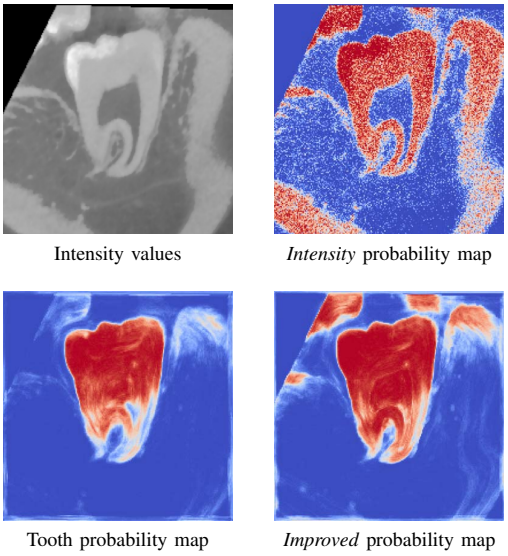


Fig. 3. Example probability maps.

B. Object Prediction

We tested our method with two object prediction setups. *Intensity*: A random forest regressor on the whole tooth labels using the intensity values only. *Improved*: Two random forest regressors; one on the pulp and one on the dentin/enamel labels. The following features were used: intensity values, Laplacian of Gaussian on three scales ($\sigma = \{0.5, 1, 2\}$), Gaussian Gradient Magnitudes, Sobel edges and Multi-channel Harmonic Filters [16] of the order 3 on four scales ($\sigma = \{2, 5, 8, 11\}, \gamma = \{1, 1, 1, 1\}$). For prediction, the maximum of the two regression outputs is used. All random forests were trained with 50000 points per label uniformly sampled as training data (a border of 0.75mm has been excluded from sampling to reduce image boundary effects). 256 trees were learned per forest, where 30% of the sampled points were used per tree. As splitting criterion, the Gini impurity is used.

In Fig. 3, example probability maps of the two different prediction methods *intensity* and *improved* applied to a test sample, are shown. Additionally, we compare the probability map to a prediction where the same features were used as in *improved*, but with only one single random forest trained on the whole tooth label. In the *intensity* probability map the pulp

TABLE I. FORMAT: AVERAGE PERFORMANCE ± 1 STDEV.

	Method	Test	Training
Υ_{ob}	*intensity*	$0.658 \pm 8.19e-3$	$0.692 \pm 9.41e-3$
	tooth	$0.810 \pm 8.77e-3$	$0.879 \pm 2.78e-3$
	improved	$0.834 \pm 5.77e-3$	$0.886 \pm 4.71e-3$
$\Upsilon_{\text{ob}}, \Upsilon_{\text{bg}}$	*intensity*	$0.601 \pm 1.44e-3$	$0.605 \pm 1.81e-3$
	tooth	$0.776 \pm 6.75e-3$	$0.842 \pm 1.49e-3$
	improved	$0.794 \pm 2.31e-3$	$0.825 \pm 1.11e-3$

TABLE II. FORMAT: AVERAGE PERFORMANCE ±1 STDEV.

Method	Test (BLD)	Training (BLD)
ASM	$0.418 \pm 18.1e-3$	$0.376 \pm 15.7e-3$
Cremers et al. (intensity)	$0.366 \pm 4.00e-3$	$0.345 \pm 5.53e-3$
Cremers et al. (improved)	$0.341 \pm 3.72e-3$	$0.328 \pm 5.19e-3$
Our (intensity, arb. ref.)	$0.387 \pm 65.5e-3$	$0.189 \pm 3.18e-3$
Our (intensity, no bound.)	$0.365 \pm 44.0e-3$	$0.170 \pm 2.47e-3$
Our (intensity)	$\mathbf{0.350} \pm \mathbf{38.6}e-3$	$0.174 \pm 2.13e-3$
Our (improved, arb. ref.)	$0.362 \pm 20.3e-3$	$0.235 \pm 7.49e-3$
Our (improved, no bound.)	$0.338 \pm 22.5e-3$	$0.193 \pm 1.17e-3$
Our (improved)	$\mathbf{0.334} \pm \mathbf{20.8}e-3$	$0.199 \pm 1.45e-3$
Cremers et al. (ground truth)	$0.310 \pm 3.99e-3$	$0.233 \pm 5.72e-3$
Our (ground truth)	$0.211 \pm 3.28e-3$	$0.091 \pm 0.13e-3$

Method	Test (dice)	Training (dice)
Cremers et al. (intensity)	$0.881 \pm 0.40e-3$	$0.886 \pm 0.68e-3$
Cremers et al. (improved)	$0.892 \pm 0.26e-3$	$0.896 \pm 0.44e-3$
Our (intensity, arb. ref.)	$0.905 \pm 1.54e-3$	$0.942 \pm 0.26e-3$
Our (intensity, no bound.)	$0.907 \pm 1.07e-3$	$0.946 \pm 0.19e-3$
Our (intensity)	$\mathbf{0.910} \pm \mathbf{0.86}e-3$	$0.945 \pm 0.17e-3$
Our (improved, arb. ref.)	$0.904 \pm 0.51e-3$	$0.932 \pm 0.30e-3$
Our (improved, no bound.)	$0.909 \pm 0.64e-3$	$0.941 \pm 0.06e-3$
Our (improved)	$\mathbf{0.910} \pm \mathbf{0.54}e-3$	$0.939 \pm 0.07e-3$
Cremers et al. (ground truth)	$0.901 \pm 0.28e-3$	$0.927 \pm 0.54e-3$
Our (ground truth)	$0.936 \pm 0.16e-3$	$0.968 \pm 0.02e-3$

could not be predicted as belonging to the object, because it shares intensity values with the background. Furthermore, the bone structure of the jaw is not fully predicted as background. In the tooth probability map, the background and the neighboring teeth (upper left of the tooth crown) could be greatly reduced. But, there is some uncertainty left in the region of the roots. In the *improved* probability map, the background is predicted accurately, while the neighboring teeth are predicted as object as well. Nevertheless, the overall tooth structure including the roots is better predicted than in the other two probability maps. In Tab. I, the prediction performance of the different methods are compared using the performance measures

$$\Upsilon_{\text{ob}} = \frac{\sum_x L_T(x) P_{ob}(x)}{\sum_x L_T(x)}, \quad \Upsilon_{\text{bg}} = \frac{\sum_x (1 - L_T(x)) P_{bg}(x)}{\sum_x (1 - L_T(x))}, \tag{12}$$

where L_T is the ground truth target label map. The *improved* method yields the best object prediction.

C. Evaluation of Segmentation

In the following, we discuss the segmentation performance of our method on a few qualitative test cases. We quantitatively evaluate our approach and compare the results with the ASM method [21]. Further, we compare them to the method of Cremers et al. [1] where only the model fitting without the global optimization was used, since the images are pre-aligned. For our method we have chosen the following parameters: weighting of the regularization $\gamma = 0.0001$ and the trade-off between region and boundary terms $\lambda = 0.75$. Image gradients have been calculated by convolving the first order derivative of a Gaussian, where we used $\sigma = 0.15$. In Fig. 4, we show

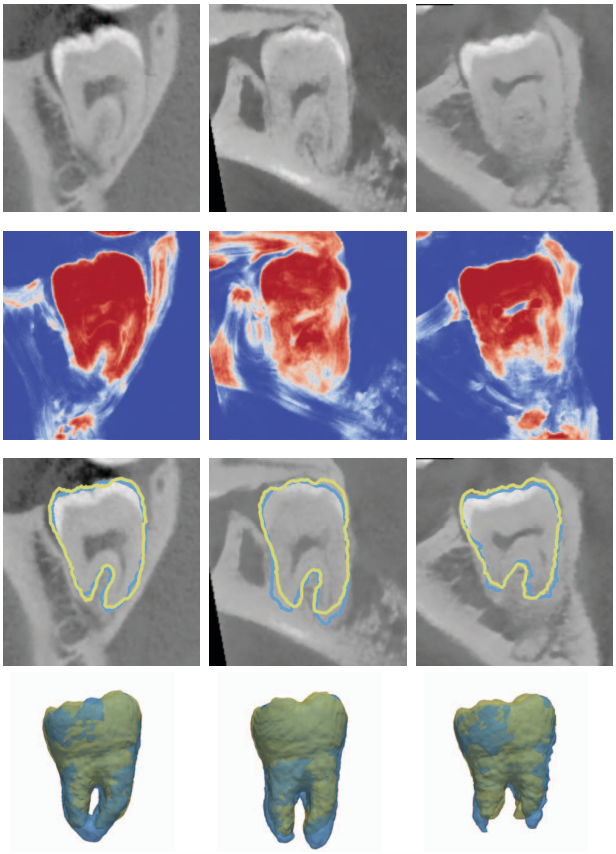


Fig. 4. Three difficult test images. In the upper row, the original CBCT images are shown. The second row shows the *improved* probability maps. The third row shows the outline of the ground truth label in blue and the segmentation results in yellow. In the last row, the teeth are visualized as 3D meshes in the same colors.

qualitative results of three difficult test examples which have been well segmented using our approach. The first is difficult, because of touching root tips. In the second one, there is almost no texture contrast between jaw and root structure. And in the last example there is a lot of noise present in the region of the roots. The bottom row shows the ground truth shape (blue) overlayed with our segmented shape (yellow).

In Tab. II, numerical results of the following experiments are provided. In the upper part, the BLD between the ground truth and resulting shapes was used as performance measure, while in the bottom part, the dice coefficient on the whole image domain serves as segmentation accuracy measure. The ASM method was performed on the histogram matched images and the method of Cremers et al. [1] using the *intensity* and the *improved* prediction. We evaluate two additional variants for our approach: one using a training label map as reference (*arb. ref.*) and one without the boundary term ($\lambda = 1$) (*no bound.*). In the two last rows the ground truth label maps are used as probability maps (*ground truth*). Compared to the *intensity* experiments, the *improved* object prediction could always improve the test segmentation performance even in the method of Cremers et al. [1]. The two additional experiments *arb. ref.* and *no bound.* show that the constructed reference

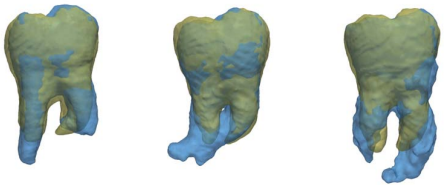


Fig. 5. Three examples of outliers.

shape as well as the incorporation of the boundary term improve the test segmentation. Since the ASM method only considers object boundary its segmentation performance is minor. Compared to the intensity model of Cremers et al. [1] the deformation model and the incorporation of the object boundary in our method leads to better segmentation results.

In Fig. 5, three bad examples are depicted. The first one is caused by the expressiveness of the statistical shape model. In the training data, there is no sample with such difference in the two root lengths. The second one originates from a fusion of almost the full roots, which is difficult to handle with our method. In the last example, the roots could not be predicted as object well enough.

IV. Conclusion

We presented a new segmentation method based on a probabilistic shape representation and an implicit statistical shape model. A main part of the method is the object and background prediction, which is learned from training samples to get robust against object boundary insufficiencies, noise and background clutter. We compared our approach to the ASM method and the method of Cremers et al. [1]. In our experiments on 3D CBCT images, we show superior segmentation performance using our method. We plan to integrate feature selection into the object prediction to leaner its computational costs. The tuning of the tree count and depth in the random forest training is planned as future work. We plan also to apply our method to multi-modality datasets where the same structure is learned from different image modalities for the object prediction. To reduce the statistical deformation model bias for small sample sets, we plan to integrate the Gaussian process framework of Lüthi et al. [22]. Furthermore, an automatic initialization of the model, similar to [1] is currently ongoing work.

References

[1] D. Cremers, F. R. Schmidt, and F. Barthel, "Shape priors in variational image segmentation: Convexity, lipschitz continuity and globally optimal solutions," in *CVPR*, 2008.

[2] M. Ruggeri, G. Tsechpenakis, S. Jiao, M. E. Jockovich, C. Cebulla, E. Hernandez, T. G. Murray, and C. A. Puliafito, "Retinal tumor imaging and volume quantification in mouse model using spectral-domain optical coherence tomography," *Optics express*, vol. 17, no. 5, p. 4074, 2009.

[3] T. F. Cootes, G. Edwards, and C. Taylor, "Comparing active shape models with active appearance models," in *Proc. of British Machine Vision Conference*. BMVA Press, 1999, pp. 173–182.

[4] M. Brejl and M. Sonka, "Object localization and border detection criteria design in edge-based image segmentation: automated learning from examples," *IEEE Trans. on Medical Imaging*, vol. 19, no. 10, pp. 973–985, 2000.

[5] H. Li and O. Chutatape, "Automated feature extraction in color retinal images by a model based approach," *IEEE Trans. on Biomedical Engineering*, vol. 51, no. 2, pp. 246–254, 2004.

[6] J. G. Bosch, S. C. Mitchell, B. P. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. Reiber, "Automatic segmentation of echocardiographic sequences by active appearance motion models," *IEEE Trans. on Medical Imaging*, vol. 21, no. 11, pp. 1374–1383, 2002.

[7] R. Larsen, M. B. Stegmann, S. Darkner, S. Forchhammer, T. F. Cootes, and B. Kjær Ersbøll, "Texture enhanced appearance models," *Computer Vision and Image Understanding*, vol. 106, 2007.

[8] B. Van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny, and M. A. Viergever, "Active shape model segmentation with optimal features," *IEEE Trans. on Medical Imaging*, vol. 21, no. 8, pp. 924–933, 2002.

[9] T. F. Cootes, M. C. Ionita, C. Lindner, and P. Sauer, "Robust and accurate shape model fitting using random forest regression voting," in *Computer Vision–ECCV 2012*. Springer, 2012, pp. 278–291.

[10] J. C. Gee and R. K. Bajcsy, "Elastic matching: Continuum mechanical and probabilistic analysis," *Brain warping*, vol. 2, 1998.

[11] H. Park, P. H. Bland, and C. R. Meyer, "Construction of an abdominal probabilistic atlas and its application in segmentation," *IEEE Trans. on Medical Imaging*, vol. 22, no. 4, pp. 483–492, 2003.

[12] A. F. Frangi, D. Rueckert, J. A. Schnabel, and W. J. Niessen, "Automatic construction of multiple-object three-dimensional statistical shape models: Application to cardiac modeling," *IEEE Trans. on Medical Imaging*, vol. 21, no. 9, pp. 1151–1166, 2002.

[13] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on pure and applied mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[14] M. Rousson and N. Paragios, "Prior knowledge, level set representations & visual grouping," *International Journal of Computer Vision*, vol. 76, no. 3, pp. 231–243, 2008.

[15] G. Tsechpenakis, "Deformable model-based medical image segmentation," in *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*, 2011, pp. 33–67.

[16] M. Schlachter, M. Reisert, C. Herz, F. Schlurmann, S. Lassmann, M. Werner, H. Burkhardt, and O. Ronneberger, "Harmonic filters for 3d multichannel data: Rotation invariant detection of mitoses in colorectal cancer," *IEEE Trans. on Medical Imaging*, vol. 29, no. 8, pp. 1485–1495, 2010.

[17] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[18] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images," *IEEE Trans. on Medical Imaging*, vol. 18, no. 8, pp. 712–721, 1999.

[19] H. S. Kim, S. B. Park, S. S. Lo, J. I. Monroe, and J. W. Sohn, "Bidirectional local distance measure for comparing segmentations," *Medical physics*, vol. 39, p. 6779, 2012.

[20] M. Lüthi, R. Blanc, T. Albrecht, T. Gass, O. Goksel, P. Buchler, M. Kistler, H. Bousleiman, M. Reyes, and P. C. Cattin, "Statismo-a framework for pca based statistical models." *The Insight Journal*, pp. 1–18, 2012.

[21] D.-J. Kroon, "Active shape model (asm) and active appearance model (aam)," *MATLAB implementation: http://www.mathworks.com*, vol. 8, p. 22, 2010.

[22] M. Lüthi, C. Jud, and T. Vetter, "A unified approach to shape model fitting and non-rigid registration," in *Machine Learning in Medical Imaging*. Springer, 2013, pp. 66–73.