

Multi-scale Statistical Grey Value Modelling for Thrombus Segmentation from CTA

Silvia D. Olabbarriaga¹, Marcel Breeuwer², and Wiro J. Niessen¹

¹ University Medical Center Utrecht, Image Sciences Institute,

Heidelberglaan 100, 3584 CX Utrecht, NL {silvia,wiro}@isi.uu.nl

² Philips Medical Systems, Medical IT - Advanced Development, Best, NL

Abstract. In this paper we present, evaluate, and discuss two multi-scale schemes for modelling grey-level appearance in a deformable model for the segmentation of abdominal aortic aneurysm thrombus from CT angiography scans. The methods are initialised with the lumen boundary, and the image-based deformation force is based on a non-parametric statistical grey level model built from training data obtained at multiple scales. The image force direction and magnitude are based on a fit between the trained model and the data. Two multi-scale schemes are used for deformation. In one scheme, the boundary is progressively refined based on training data obtained at each scale, in a coarse-to-fine approach, and in the other, the training data obtained at all scales are used simultaneously to drive the deformation. The two schemes are evaluated and compared to the single scale scheme based on a leave-one-out study of nine patients.

1 Introduction

Abdominal aortic aneurysm (AAA) is a relatively common disease among the elderly population. Because the rupture of an AAA can be fatal, there is clinical interest in risk analysis, simulation, surgery planning, and treatment follow-up based on minimally invasive imaging, such as CT angiography (CTA). Segmentation of lumen, thrombus and calcifications is necessary for the reconstruction of a patient-specific model based on CTA scans. While lumen and calcifications can be rather easily isolated with thresholding techniques, this is not true for the thrombus. Due to the low contrast between thrombus and surrounding tissue, automated segmentation methods have difficulty to delineate the correct boundary. Only a few thrombus segmentation methods have been reported in the literature [1,2,3,4]. The methods proposed by De Bruijne et al. [1,3] are based on an active shape model (ASM), while those proposed by Subasic et al. [2] and Giachetti & Zanetti [4] are based on deformable models. These methods, however, suffer from shortcomings: either they require significant amount of user intervention [1,3,4], or the reported results are not sufficiently accurate [2].

In [5] we introduced a new AAA segmentation method in which the lumen and thrombus are segmented automatically based on two positions clicked by the user. The lumen surface is used to initialize the thrombus segmentation

method, which is based on a discrete deformable model approach. A statistical grey level model built from training data is adopted to generate the image-based deformation force. This force expands (or shrinks) the boundary to include (or exclude) image locations with intensity patterns that correspond to the inside (or outside) of the object, respectively. A non-parametric technique is used for intensity pattern classification, and the most likely class determines the force direction (inwards or outwards the boundary). That method is now extended by determining the force magnitude also from the class with largest likelihood, creating a natural link between training and deformation in a multi-scale framework. In this paper, we propose two multi-scale schemes and investigate their added value with respect to the single-scale case. The first scheme consists of progressively refining the boundary based on models built from training data, using a coarse-to-fine approach. The second is a new scheme in which a single model, built at multiple scales simultaneously, is used to guide deformation. In both schemes, the image force direction and magnitude are determined based on training data. The schemes are evaluated on nine CTA scans, and the results are compared with manual segmentations.

2 Thrombus Segmentation Method

The proposed method is based on a deformable model approach, in particular the discrete formulation proposed in [6] and extended in [7]. The deformation of an initial boundary, represented with a polygonal mesh, is performed at discrete time steps t by changing the position of its vertices \mathbf{x}_j according to the evolution equation:

$$\mathbf{x}_j^{t+1} = \mathbf{x}_j^t + (1 - \lambda)(\mathbf{x}_j^t - \mathbf{x}_j^{t-1}) + \lambda (\alpha_j F_{int}(\mathbf{x}_j^t) + \beta_j F_{ext}(\mathbf{x}_j^t)), \quad (1)$$

where λ is a damping factor, F_{int} and F_{ext} are respectively the internal (shape-based) and external (image-based) forces acting on the vertex, and α_j and β_j indicate the forces relative importance in the deformation process. The initial boundary is the lumen surface obtained automatically with another deformable model initialised with minimal user interaction (see details in [5]). The adopted internal force (continuity of mean curvature) enforces mesh smoothness and regularization. The external force moves the mesh into image positions with “boundary-like” properties. Typically, this force would be determined from the thrombus edge (e.g. image gradient) because object boundaries are expected at locations of sharp intensity variation. In the case of the AAA thrombus, however, this assumption does not hold. Firstly, several other neighbouring structures (e.g., spine and calcifications) appear with strong contrast in an abdominal CTA image and may attract the boundary to the wrong position. Secondly, the contrast between thrombus and background can be extremely weak in some regions (e.g. bowels), and image derivatives at such locations are not easily distinguishable from noise. Therefore, a more complex image force is required for AAA thrombus segmentation.

2.1 Multi-scale Image Force

The image force adopted in this work corresponds to an inflating (or shrinking) vector along the vertex normal \mathbf{n}_j :

$$F_{ext}(\mathbf{x}_j) = \begin{cases} S\mathbf{n}_j & : \mathbf{x}_j \text{ is inside the object} \\ -S\mathbf{n}_j & : \mathbf{x}_j \text{ is outside the object} \\ 0 & : \mathbf{x}_j \text{ is at the object boundary} \end{cases} \quad (2)$$

where S is the magnitude of the force (deformation step size). A non-parametric pattern classification approach, namely k -nearest neighbours (k NN [8]), is used to determine the most likely class (inside/outside/boundary) corresponding to the intensity pattern at the vertex. The posterior probability of an intensity pattern \mathbf{y} belonging to class ω is determined as follows:

$$P(\omega|\mathbf{y}) = \frac{k_\omega}{k} \quad (3)$$

where k_ω is the number of points with class label ω among the k nearest neighbours (closest training points) of \mathbf{y} . The classification features are image intensity values (intensity profiles of length l and sampling spacing δ) collected along the vertex normal \mathbf{n} – see fig. 1.

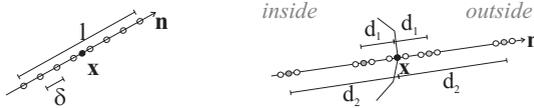


Fig. 1. Profile sampling. Left: normal vector \mathbf{n} , profile length l and sampling spacing δ . Right: Multi-scale training scheme with 5 classes ($n = 2$).

During training, profile samples are collected at the correct boundary (delineated manually) and at shifted positions inside and outside the thrombus. In our previous work [5], a three-class model is adopted, in which two shifted samples are collected respectively at $\mathbf{x} - d\mathbf{n}$ and $\mathbf{x} + d\mathbf{n}$. Here we extend this notion to a $2n + 1$ -class model in which shifted samples are collected at increasing distances $d_q, q \in [1, n]$ from the boundary position \mathbf{x} – see fig. 1. The class labels are assigned as follows: ω_0 for boundary samples, ω_{-q} for inside samples and ω_{+q} for outside ones.

During deformation, the image force at a vertex \mathbf{x}_j is calculated in the following way: the intensity profile \mathbf{y}_j at the vertex is sampled, and the most likely class ω_i is determined by $\max_{i \in [-n, +n]} P(\omega_i|\mathbf{y}_j)$ – see eq. 3. The class label ω_i indicates directly whether the vertex is most likely to be inside ($i < 0$), outside ($i > 0$) or at the boundary ($i = 0$) – see eq. 2. The image force strength is also scaled according to the most likely class ω_i :

$$S = d_q c f \quad (4)$$

where d_q is the shift distance for class ω_i , with $q = |i|$; $c = P(\omega_i|\mathbf{y}_j)$ is the confidence in the classification step; and f is a factor that regulates the maximum force magnitude. Note that larger (or smaller) deformation steps are taken when the intensity pattern best fits a class sampled far from (or near to) the correct boundary. Moreover, the force magnitude is reduced when the intensity pattern does not clearly indicate the situation at hand. Note that the multiple training samples obtained at varying d_q could be used in several ways during deformation. Here we consider two possibilities:

- A refinement scheme in which the surface is progressively deformed by running the method n times using the non-boundary training profiles obtained at $d_p, p \in [n, 1]$ (similarly to [9,3]), and
- A new multi-scale modelling scheme in which the training profiles obtained at all scales d_q are used simultaneously, in a single deformation run.

3 Evaluation Study

The method is tested on nine contrast-enhanced CTA scans of different patients in which the AAA is located between the renal and the iliac arteries and no implants. The scans consist of a random sample of patients from the University Medical Centre Utrecht (UMCU) endovascular aneurysm repair programme [10]. All images are acquired with Philips Medical Systems (PMS) spiral CT scanners with resolution of $0.5 \times 0.5 \times 2$ mm and contain circa 125 slices of 512×512 voxels.

The training samples are collected based on manual segmentations performed by experts of the Dept. of Vascular Surgery and the Image Sciences Institute, UMCU, using the PMS EasyVision workstation contouring facilities. The original manual meshes are smoothed and resampled before profile extraction and comparison with the obtained results. The leave-one-out strategy is used, i.e., the segmented scan is never part of the training set. Each training set contains 80 000 to 200 000 profiles.

The profile configuration is constant in all experiments, with $l = 5$ mm and $\delta = 1.0$ mm (6 features per sample). The number of neighbours is fixed to $k = 21$, since the behaviour with respect to k in the multi-scale schemes is rather robust [5]. The thrombus deformable model parameters are set as follows: $\lambda = 0.8$, $\beta = 0.1$, and $\alpha = 1 - \beta$ (see eq. 1); and $f = 0.5$ mm (see eq. 4). The deformation stops automatically if the volume difference between two iterations is smaller than 0.1%.

In the experiments we investigate the behaviour of each multi-scale scheme with respect to the shift distance d_n and to the number of scales ($n = 1, 2, 3$). The closest sample is obtained at $d_1 = 0.5$ mm (in-slice image resolution) and the farthest at varying distances ($d_n \in \{1, 2, 3, 4, 5\}$ mm). Distances increase exponentially for $n = 3$, with $d_2 \in \{0.7, 1.0, 1.2, 1.4, 1.6\}$ mm. Results are analysed separately for the single scale, refinement and the multi-modelling schemes, and finally the schemes are compared. The error measure adopted in the evaluation corresponds to the distance (in mm) between vertices in a given result and

Table 1. Segmentation error (mm) for the single and multi-scale schemes for varying n and d_n : mean \pm standard deviation of rmse for nine scans. The star marks the configuration with smallest average error.

d_n	Single Scale	Refinement		Multi-model	
		$n = 2$	$n = 3$	$n = 2$	$n = 3$
1	4.3 \pm 2.2	4.3 \pm 2.2	4.3 \pm 2.2	5.2 \pm 2.2	5.0 \pm 2.2
2	2.0 \pm 1.4	2.0 \pm 1.4	2.0 \pm 1.4	2.9 \pm 2.1	2.9 \pm 2.1
3 (*)	1.4 \pm 0.4	1.2 \pm 0.3	1.2 \pm 0.4	1.2 \pm 0.3	1.2 \pm 0.4
4	3.0 \pm 3.2	3.0 \pm 3.3	2.7 \pm 3.4	2.6 \pm 2.8	2.3 \pm 2.5
5	3.1 \pm 3.1	2.8 \pm 3.3	2.5 \pm 3.4	3.2 \pm 3.1	2.8 \pm 3.2

the surface obtained with manual segmentation. The root mean squared error (rmse) of all vertices characterizes the overall segmentation quality. The statistical significance of differences between results obtained with different settings is evaluated with a paired T-test and expressed by the null hypothesis probability (p value). The relative volume of overlap between a segmentation A and the manual segmentation B is calculated with $2 \frac{V(A \cap B)}{V(A) + V(B)}$.

4 Results

Table 1 summarizes the segmentation errors obtained for all scans with varying number of scales n , shift distance d_n , and deformation schemes.

In average, smaller errors are obtained for $d_n = 3$ mm in all configurations. The difference is more significant with respect to $d_n = 1, 2$ ($p < 0.05$) than to $d_n = 4, 5$ ($p > 0.15$). Figure 2 illustrates a typical case: for larger d_n , the boundary erroneously deforms into neighbouring structures, while for smaller d_n , it stays near the lumen boundary. In the remainder of the paper we only compare results obtained with $d_n = 3$ mm.

Figure 3 presents the segmentation error (rmse) obtained for each scan in all deformation schemes. Although a small reduction in error is obtained with the multi-scale schemes in comparison to single scale, the difference is not significant. The same observation holds for the improvement obtained by adopting 3 scales instead of 2 in both multi-scale schemes ($n = 2, 3$). Note, however, that the segmentation error is rather small in all schemes, with $\text{rmse} \leq 2.2$ mm, and that this is significantly smaller than the initial error ($\text{rmse} \geq 2.8$ mm). Table 2 presents the average volume of overlap between these segmentation results and the manual segmentations and the errors in the volume calculated based on these segmentations. Note that the average segmentation error (rmse), volume accuracy and volume errors are comparable to those reported in [3] (respectively 1.9 mm, 95% and 4.1%).

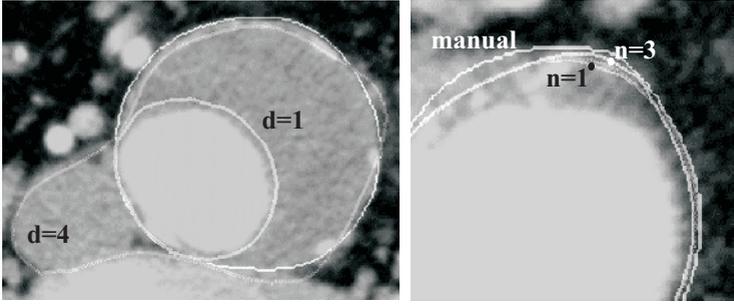


Fig. 2. Illustration of segmentation results and manual segmentation at selected slices of different scans. Left: varying d_n (monoscale). Right: varying n (refinement scheme, $d_n = 3$).

Table 2. Volume of overlap (%) and difference in volume calculation (%) between segmentation results and manual segmentation ($d_n = 3$ mm).

	Single Scale	Refinement		Multi-model	
		$n = 2$	$n = 3$	$n = 2$	$n = 3$
Overlap	95.5 ± 1.3	96.2 ± 1.0	96.7 ± 1.2	96.2 ± 1.3	96.3 ± 1.6
Error	3.9 ± 2.9	2.6 ± 2.9	2.7 ± 3.1	2.9 ± 3.1	3.1 ± 3.6

5 Discussion and Conclusions

The image force adopted here for segmentation is inspired by approaches in which an appearance model built from training data is fit to the image. In ASMs, as introduced by Cootes et al. [11], a parametric statistical model is adopted for modelling grey values along the boundary profile; the model is built based on training samples of correct boundary profiles only. Such a simple model is not applicable for AAA thrombus segmentation, as shown in [3]. Extensions to grey level modelling for ASMs are proposed by Ginneken et al. [9] and De Bruijne et al. [3], in which the training is based on profile samples obtained at boundary positions and at multiple positions inside and outside the object. In both cases, a non-parametric classification technique (k NN) is used to determine the fit between model and data, and the landmark is moved into a position with best fit among a number of considered options. While in [9] the classification features are automatically selected from a large bank of Gaussian filters to discriminate between patterns inside and outside the object, in [3] image intensity is adopted. Finally, in the deformable model method proposed by Pardo et al. [12], the model is trained in one slice based on a 2-D contour, with samples at boundary and non-boundary positions. The most discriminating classification features are selected automatically from a Gaussian bank, and the model fit is measured with a parametric model for deformation in the adjacent slice. The methods above [11,9,3,12] are similar with respect to the following aspects: (1)

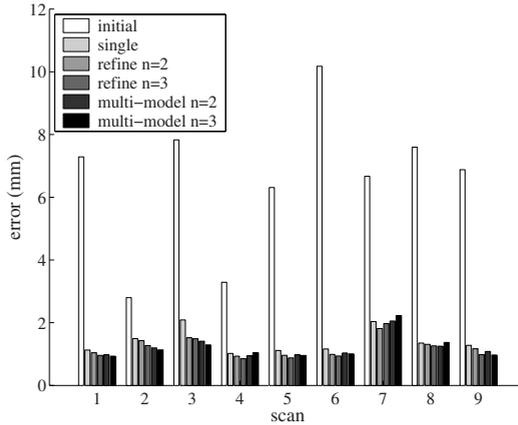


Fig. 3. Root mean squared error for the initial boundary (lumen) and the segmentation results obtained with all schemes ($d_n = 3$ mm).

the deformation is driven by a quality of fit that only takes into account the boundary/non-boundary distinction, and (2) a coarse-to-fine scheme is adopted, in which the boundary is progressively refined with grey level models built at increasing resolution. Compared to these works, two novelties are introduced here. First, the method uses a three-type class model that also distinguishes inside and outside non-boundary profiles, and in this way steers both the direction and magnitude of deformation. Our approach imposes fewer requirements on the initial surface than when an on-off boundary model is adopted. In particular, no successful AAA thrombus segmentation could be obtained when using such an on-off boundary model for deformation starting from the lumen boundary. Secondly, a novel multi-scale modelling approach is introduced that simultaneously considers profiles trained at multiple distances.

The method was evaluated on nine scans using a leave-one-out strategy, and the results were compared to manual segmentations. A limited number of parameters for the coarse-to-fine and multi-scale modelling schemes were considered. With respect to the largest shift distance d_n , better results were obtained in average with $d_n = 3$ mm in both schemes. Accurate results were obtained with all schemes; segmentation errors ([1.2, 1.4]mm) and volume accuracy ([95.5, 96.5]%) are in the same range as reported in [3] (1.9 mm and 95%), and the volume error ([2.6, 3.9]%) is comparable to the intra-observer variability reported in [13] (3.2%).

In conclusion, the results indicate that the three-type class model, that distinguishes inside and outside profiles, is very well suited for thrombus segmentation in CTA. With the lumen as initialization, the thrombus was found as accurate as the previously reported variability in manual outlining. We have shown previously [5] that the lumen can automatically be obtained using only two mouse

clicks as initialization. The three-type class model was essential for obtaining these results: we found that these results could not be obtained with an on-off boundary profile training. Interestingly, the multi-scale approaches only slightly improved on the single scale approach, as long as an appropriate scale is selected. It remains to be investigated what the additional value of multi-scale schemes is in other applications.

Acknowledgments. We thank Prof. Dr. J. Blankensteijn and M. Prinssen, formerly with the Dept. of Vascular Surgery at the UMCU, for kindly providing datasets and manual segmentations. We also thank Dr. de Bruijne for the valuable discussions and suggestions. Finally, we are grateful to the developers of several libraries: 3D Active Objects, by Philips Research Paris; Approximate Nearest Neighbors (ANN), by Univ. of Maryland; and isiImage library, by Image Sciences Institute, UMCU.

References

1. M. de Bruijne et al. Active shape model based segmentation of abdominal aortic aneurysms in CTA images. In *SPIE Medical Imaging*, volume 4684, pages 463–474. SPIE, 2002.
2. M. Subasic, S. Loncaric, and E. Sorantin. 3-D image analysis of abdominal aortic aneurysm. In *SPIE Medical Imaging - Image Processing*, pages 1681–1689, 2002.
3. M. de Bruijne et al. Adapting active shape models for 3D segmentation of tubular structures in medical images. In *Information Processing in Medical Imaging (IPMI)*, volume 2732 of *LNCS*, pages 136–147. Springer, 2003.
4. A. Giachetti and G. Zanetti. AQUATICS reconstruction software: the design of a diagnostic tool based on computer vision algorithms. In *CVAMIA and MMBIA Workshop*, 2004.
5. S.D. Olabbarriaga, M. Breeuwer, and W.J. Niessen. Segmentation of abdominal aortic aneurysms with a non-parametric appearance model. In *CVAMIA and MMBIA Workshop*, 2004.
6. H. Delingette. General object reconstruction based on simplex meshes. *International Journal of Computer Vision*, 32(2):111–146, 1999.
7. O. Gérard et al. Efficient model-based quantifications of left ventricular function in 3-d echocardiography. *IEEE Trans. Medical Imaging*, 21(9):1059–1068, 2002.
8. R. Duda, P. Hart, and D. Stork. *Pattern Classification*. John Wiley & Sons, 2001.
9. B.v.Ginneken et al. Active shape model segmentation with optimal features. *IEEE TMI*, 21(8):924–933, 2002.
10. M. Prinssen et al. Concerns for the durability of the proximal abdominal aortic aneurysm endograft fixation from a 2-year and 3-year longitudinal computed tomography angiography study. *J Vasc Surg*, 33:64–69, 2001.
11. T.F. Cootes et al. The use of active shape models for locating structures in medical images. *Imaging and Vision Computing*, 12(6):355–366, 1994.
12. X.M. Pardo, P. Radeva, and D. Cabello. Discriminant snakes for 3d reconstruction of anatomical organs. *Medical Image Analysis*, 7(3):293–310, 2003.
13. J.J. Wever et al. Inter- and intraobserver variability of CT measurements obtained after endovascular repair of abdominal aortic aneurysms. *American Journal of Roentgenology*, 175:1279–1282, 2000.