

Automated Segmentation of Abdominal Aortic Aneurysms in Multi-spectral MR Images

Marleen de Bruijne^{1,2,3}, Bram van Ginneken², Lambertus W. Bartels², Maarten J. van der Laan³, Jan D. Blankensteijn³, Wiro J. Niessen², and M.A. Viergever²

¹ IT University of Copenhagen, Denmark

² Image Sciences Institute, University Medical Center Utrecht, The Netherlands

³ Department of Surgery, University Medical Center Utrecht, The Netherlands

Abstract. An automated method for segmenting the outer boundary of abdominal aortic aneurysms in MR images is presented. The method is based on the well known Active Shape Models (ASM), which fit a global landmark-based shape model on the basis of local boundary appearance models. The original three-dimensional ASM scheme is modified to deal with multi-spectral image information and inconsistent boundary appearance in a principled way, with only a limited amount of training data. In addition, a framework for user interaction is proposed. If required, the obtained segmentation can be corrected in an interactive manner by indicating points on the desired boundary.

The methods are evaluated in leave-one-out experiments on 21 datasets. A segmentation scheme combining gray level information from two or three MR sequences produces significantly better results than a single-scan model. Average volume errors with respect to the manual segmentation are 4.0%, in 19 out of 21 datasets. In the cases in which the obtained error is large, results can easily be improved using the interactive scheme.

1 Introduction

After endovascular repair of abdominal aortic aneurysms (AAA), regular post-operative imaging is needed for the rest of the patient's life, in order to timely detect complications. One of the most important factors to monitor is aneurysm sac size. Volume measurements are thought to be more accurate than diameter measurements, but require a very time-consuming slice-by-slice manual segmentation. Although most centers use CTA follow-up for AAA patients, MRI is attracting attention as a modality that may replace CTA in this application. Advantages of MRI over CTA are the lack of ionizing radiation and nephrotoxic contrast agent, increased sensitivity to endoleaks, and improved soft-tissue contrast which enables assessment of thrombus consistency [1]. Most publications on computerized AAA segmentation describe segmentation of the contrast-filled lumen in CTA images. Aneurysm volume measurement requires segmentation of the aneurysm outer boundary, which is a more complex task. A few authors have addressed this problem in CTA images, using three-dimensional active surfaces or level-sets [2,3,4]. To our knowledge, no research has concentrated on automated aneurysm volume measurements in MRI scans. This paper presents a method for automated delineation of the outer

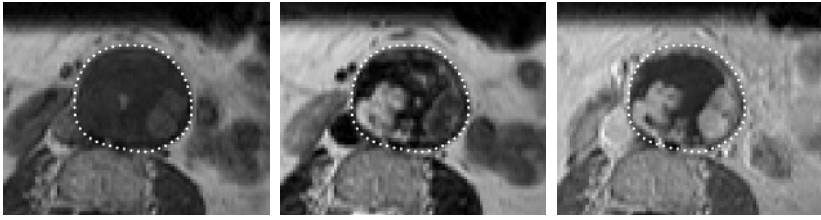


Fig. 1. T1 (left), T2 (middle) and T1C slice (right), taken at the same level in the same patient. The white dots denote the expert segmentation obtained in the T1C image. In the T1 scan, the signal of both thrombus and lumen is lower than most of the background. The T2 scan shows differences in thrombus consistency; fresh clot is bright, old thrombus is dark. The aortic lumen shows up bright only in the T1C scan. In the latter scan, the thrombus is dark just as in the T1 image, except for the regions where gadolinium was deposited, notably in endoleaks and at the vessel wall.

aneurysm boundary in multiple MR sequences. The method is inspired by the Active Shape Model (ASM) framework as proposed by Cootes and Taylor [5]. ASMs consist of a landmark based shape model, linear models of gray value appearance around the landmarks, and an iterative optimization scheme. Both the shape model and the boundary appearance model are derived from segmented example images. The components of the original ASM scheme are modified to enable AAA segmentation. The landmark based shape model, called Point Distribution Model (PDM), is adapted to better describe tubular objects if the training set is small. A non-linear gray value model is proposed which can deal with the highly variable boundary appearance of AAA and exploits information of different MR images. The shape parameters are more robustly estimated using dynamic programming regularization [6] and a weighted fit. To increase segmentation speed and robustness, a multi-resolution approach is used. A similar method, applied to segmenting aneurysms in CTA scans, was presented in [7]. In addition, a framework for user interaction is proposed. If the operator does not approve of a segmentation obtained automatically, he or she can guide the segmentation process by indicating points on the desired boundary. The automated segmentation as well as the added value of user interaction and the use of multiple MR sequences are evaluated in leave-one-out experiments on 21 MR studies of different patients. Each study includes a T1 and T2-weighted spin echo scan and a post contrast T1 (T1C) scan. The T1C images are segmented manually by an expert. Examples of the three scans are given in Figure 1.

2 Methods

2.1 Preprocessing

To allow for multi-spectral segmentation, so as to exploit the information present in all three scans, the scans must be registered so that the same voxel corresponds to the same position in the patient. We perform a rigid registration by maximizing the normalized mutual information [8]. To account for intensity differences between scans and intensity

inhomogeneities within scans, the images are locally normalized to zero mean and unit variance:

$$L_{\text{norm}}(\mathbf{x}) = \frac{L(\mathbf{x}) - G(\mathbf{x}, \sigma) * L(\mathbf{x})}{\sqrt{G(\mathbf{x}, \sigma) * (L(\mathbf{x}) - G(\mathbf{x}, \sigma) * L(\mathbf{x}))^2}}, \quad (1)$$

with $L(\mathbf{x})$ the original gray valued image and $G(\mathbf{x}, \sigma)$ a three-dimensional Gaussian kernel of width σ . In the experiments, we use $\sigma = 20$ mm.

2.2 Segmentation Scheme

This section briefly reviews the three components of the segmentation scheme. The last subsection discusses how user interaction can be incorporated into the segmentation process.

Shape model. Shapes are described by a set of landmark points which, in the ideal case, denote the same anatomical points in different instances. To construct the AAA model, we interpolate a fixed number of slices and place a fixed number of landmarks in each slice, equidistantly along contours that are drawn manually by an expert. Each shape can be approximated by a linear combination of the mean shape and several *modes of shape variation* which describe a joint displacement of all landmarks. In PDMs, the modes of variation are given by the principal components of a collection of example shapes. A common problem in statistical shape modeling is that the model can be too specific to fit to new shapes properly, owing to a limited amount of training data. We have applied two generalizations for PDMs of tubular structures that were discussed in more detail in [7]. First, the cross-sectional and axis shape variations of the training set are modeled independently. This results in a model containing twice the number of modes of a normal PDM. Second, synthetic deformation modes, describing low frequency deformations of the object surface, are added.

Appearance model. Fitting the shape model to a new image requires a notion of object boundary appearance. To derive the boundary appearance from the training set, gray value profiles are sampled around each landmark, perpendicular to the object contour. The original ASM formulation uses the Mahalanobis distance to the mean training profile as a boundary similarity measure [5]. The effect of global intensity changes is reduced by sampling the first derivative and normalizing the profile. This model produces accurate results in many applications, but encounters problems in images where boundary appearance is highly variable and neighboring objects induce similar edges. We propose a non-parametric appearance model in which the goodness of fit of a given profile is determined not only by the similarity to boundary profiles of the training set, but also by the dissimilarity to non-boundary profiles. The model is constructed from profiles around the contours in the training set as well as profiles that are shifted outwards or inwards with respect to the correct contour. Raw intensity profiles are used instead of the normalized derivative profiles of the linear model. The probability that a given profile belongs to the boundary is obtained using k nearest neighbor (k NN) probability density estimation. Multi-spectral appearance models are constructed from the concatenation of

gray value profiles sampled at the same position in the three separate, registered and normalized images. To enable coarse-to-fine fitting, separate models are built for different levels of resolution.

Optimization. The shape model is fitted to new images using a fast iterative optimization scheme, in which boundary gray value optimization and shape approximation are alternated. The appearance model determines for each landmark its optimal new position, among candidate positions along the line perpendicular to the contour, on both sides. To reduce the effect of outliers, consistent displacement of neighboring landmarks is enforced by computing the global optimal path through the evaluated positions [6]. The shape approximation itself is also an iterative procedure. It alternates the computation of optimal pose parameters (translation, rotation and scaling) with computation of the optimal shape parameters, thus minimizing the (weighted) sum of squared distances between the landmarks of the approximated shape and the optimal positions according to the appearance model. The process of alternating gray value and shape optimization is repeated a fixed number of times, whereupon it is repeated at the next level of resolution. The user initializes and constrains the segmentation process by drawing the top and bottom contours of the aneurysm manually. To aid the model in establishing the correct object axis an additional point is placed in the approximate aneurysm center of the central slice. The shape model is iteratively fitted to these points using a constrained ASM scheme, in which after each iteration the landmarks of the manually drawn slices are replaced to their original position and the landmarks of the central slice are translated such that their average position coincides with the manually identified center point.

User interaction. In some cases, the algorithm may fail to find a correct segmentation. An interactive segmentation scheme could be valuable in such a case. The user can guide the segmentation process by manually correcting part of the boundary and approving other parts, without having to perform the full segmentation manually. We propose an interaction framework in which the user identifies one or several points at the boundary of the aneurysm where the obtained automatic segmentation is far off. The nearest landmarks on the contour are determined, and are fixed at the indicated positions. Subsequently, a constrained ASM scheme is run that optimizes the shape and appearance, given the fixed points. After each iteration of ASM optimization, the point of the new shape that is nearest to the indicated boundary point is fixed while the point that was fixed previously is allowed to move freely. The optimization process is iterated until the mean squared distance between all user indicated points and the nearest landmarks is smaller than a voxel. We assume that the first segmentation is always reasonably close to the correct boundary, so that fixating the nearest point gives a reliable indication in which direction the shape should deform. In the experiments, that assumption was always valid. This interactive segmentation is an extension to the Interactive Shape Models (iASM) as proposed by van Ginneken et al. [9], in which the operator drags incorrectly placed landmarks to their correct positions. In the absence of anatomical landmarks, the correct position of a landmark in an AAA model is not clearly defined. Forcing a landmark at a position that is shifted along the contour with respect to its expected position may impose problems in other parts of the shape.

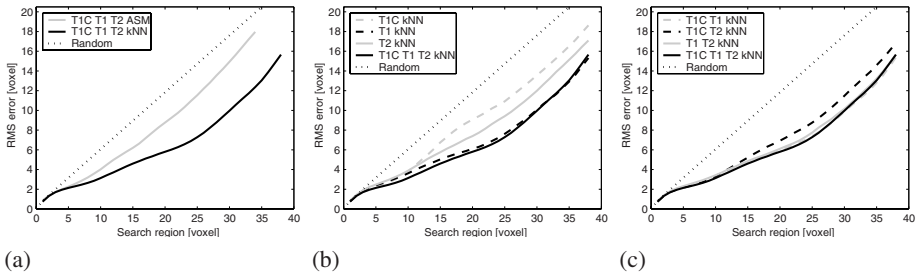


Fig. 2. Performance of different appearance models. The lines denote the root mean squared error of landmark positioning as a function of the length of the search region on either side of the contour. The dotted line corresponds to the error expectation for random position selection. (a) The original ASM gray value model (gray) and the multi-class k NN model (black), both constructed from all three scans. (b) Single scan k NN models compared with a model of all three scans. (c) Two-scan k NN models compared with a model of all three scans.

3 Experiments

Leave-one-out experiments are performed on 21 post-operative MR studies of different patients. Follow-up ranges from 2 days to 54 months. All scans were acquired according to a fixed protocol on a clinical 1.5-T scanner (Gyrosan NT, Philips). The scan resolution is $1.5 \times 1.5 \times 6.6$ mm.

Appearance model. First, performance of the appearance models is evaluated independently of the rest of the segmentation process. Starting from the landmarks on the contour drawn manually, the optimal landmark positions according to the appearance model are determined. The distances from these points to the landmark positions provided by the manual tracings are measured. The resulting root mean squared distances, as a function of the size of the search region, are shown in Figure 2. Figure 2a compares the performance of the linear gray value model that is originally used in ASM with the new k NN model, for profiles that combine the gray values of T1, T2 and T1C scans. The k NN model performs significantly better in all cases. Figure 2b and 2c compare appearance models of different combinations of scans. All single-scan models produce larger errors than the model combining information from the three scans, but the accuracy of the T1 model comes close to that of the combined model. The two-scan models are more accurate than the single-scan models; both combinations with the T1 perform similarly to the three-scan model.

Automated segmentation. Segmentation experiments were performed using the k NN appearance model for the best single-scan model (T1), the best two-scan model (T1C-T1), and the combination of all three scans. Figure 4 shows the resulting segmentation errors with respect to the manual segmentation. The root mean squared error averaged over all datasets is 3.7 mm (2.5 voxel) for the T1 model and 2.9 mm (1.9 voxel) and

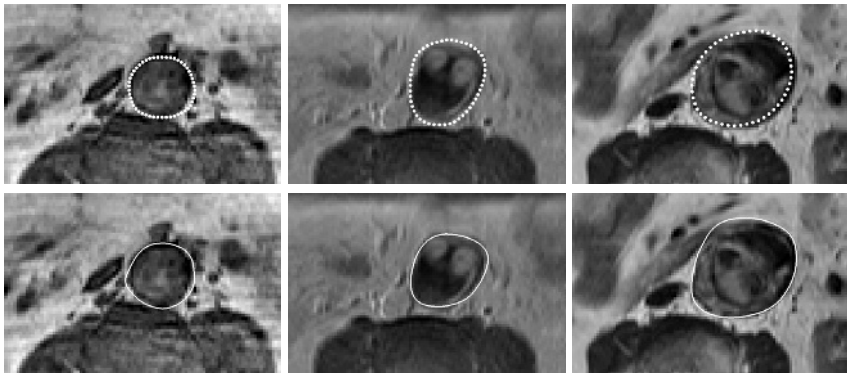


Fig. 3. Examples of obtained segmentations. The top row shows the expert segmentations, the second row the automatic segmentations. From left to right examples of a T1 (error 1.1 mm), a T1C (error 2.0 mm) and a T2 scan (error 2.1 mm).

2.6 mm (1.7 voxel) for the two-scan and three-scan model respectively. Both the two-scan model and the three-scan model produce significantly smaller errors than the one-scan model ($p = 0.008$ and $p = 0.0002$ in a paired t -test). The difference in performance between the three-scan and the two-scan model is not significant ($p = 0.13$). The errors obtained in datasets 10 and 20 are considerably larger than all others. In one of these, some of the surrounding structures, with intensity similar to the aneurysm, are included in the contour (see Figure 6). In the other dataset the thrombus gives a high signal on T1, which has not been observed in any of the training images. If these two cases of failure are left out of consideration, the average error of the three-scan model for the remaining datasets is 2.0 mm (1.3 voxel). The corresponding average overlap volume is 94%, the mean signed volume error is -1.3% (-2.0 ml), and the mean unsigned volume error is 4.0% (5.3 ml). The magnitude of inter-observer variation in these datasets is yet unknown. A difference of means analysis on primary results yielded an inter-modality repeatability coefficient (RC) of 15.9 ml for MR compared to CTA [10], while inter-observer repeatability of CTA alone was reported to be 10.3 ml [11]. Our automatic segmentations have RC=13.4 ml compared to the manual segmentations, which suggests that the accuracy of the automated method is comparable to manual segmentation. Examples of segmentations obtained using the gray value model of all three scans are given in Figure 3.

Interactive segmentation. The interactive segmentation procedure is evaluated through simulated observer interaction using the expert tracings. After automated segmentation, the point on the manually drawn contour that is furthest away from the automated segmentation is determined. This point is the fixed position in the next constrained segmentation step. The segmentation results obtained with an increasing number of fixed points are shown in Figure 5. In the majority of datasets, the difference between the manual and automatic segmentation is already small, and fixing one or several points

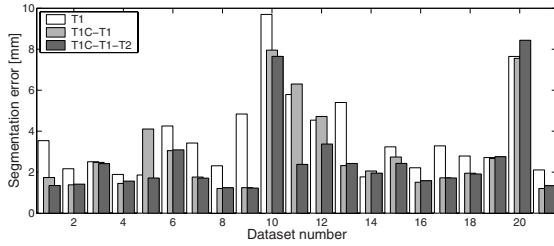


Fig. 4. Root mean squared landmark-to-contour segmentation error for each dataset, using either a single-scan, a two-scan or three-scan k NN appearance model.

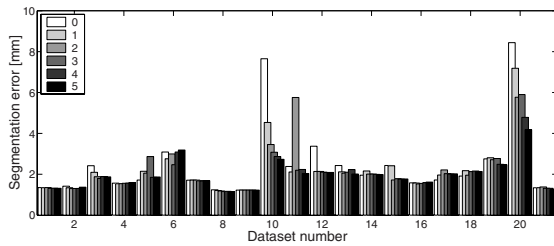


Fig. 5. Segmentation error after fixing 0 to 5 points on the correct boundary. Results obtained using the three-scan k NN appearance model.

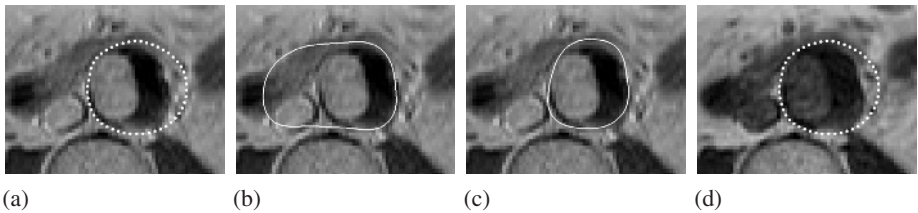


Fig. 6. Illustration of the interactive segmentation scheme. (a) TIC image with the expert segmentation. (b) First automatic segmentation result. The vena cava and part of the bowels have been included in the aneurysm volume. One point on the left side of the aneurysm (in a different slice) is fixed. (c) Second automatic segmentation result. The contour is much closer to the expert segmentation, but does not include the bright border on the right. (d) T1 scan, registered to the TIC, with the expert segmentation superimposed. The bright border turns out to be outside the aneurysm sac, while the operator—judging on basis of the T1C only—had interpreted it as gadolinium deposit at the vessel wall. The interactive segmentation is in fact more accurate than the expert segmentation performed on the T1C alone.

hardly influences the results. However, if we look at the two datasets in which the first segmentation failed, errors are greatly reduced with this user interaction. An example is given in Figure 6. This figure also illustrates the flaws in the expert segmentation performed on TIC.

4 Conclusions

An automated method for segmenting AAA in MR images is presented. Information of different MR sequences is exploited, which results in improved accuracy. Obtained volume errors with respect to a manual segmentation are comparable to manual inter-observer errors in roughly 90% of the cases. In the remaining ten percent of difficult images, only a small amount of user-interaction is required to correct the initial results.

References

1. S. Haulon, C. Lions, E. McFadden, M. Koussa, V. Gaxotte, P. Halna, and J. Beregi, "Prospective evaluation of magnetic resonance imaging after endovascular treatment of infrarenal aortic aneurysms," *European Journal of Vascular and Endovascular Surgery* **22**(1), pp. 62–69, 2001.
2. L. Wilson, S. Brown, J. Young, R. Li, and L. Brandt, "Three-dimensional computer models of abdominal aortic aneurysms by knowledge-based segmentation," in *Computer Assisted Radiology and Surgery*, H. Lemke, M. Vannier, K. Inamura, and A. Farman, eds., *Excerpta Medica International Congress Series* **1191**, pp. 213–217, Elsevier Publishers, 1999.
3. M. Subasic, S. Loncaric, and E. Sorantin, "3D image analysis of abdominal aortic aneurysm," in *Medical Imaging: Image Processing*, M. Sonka and M. Fitzpatrick, eds., *Proceedings of SPIE* **4684**, pp. 1681–1689, SPIE Press, 2002.
4. K. Subramanyan, M. Steinmiller, D. Sifri, and D. Boll, "Automatic aortic vessel tree extraction and thrombus detection in multi-slice CT," in *Medical Imaging: Image Processing*, M. Sonka and M. Fitzpatrick, eds., *Proceedings of SPIE* **5032**, pp. 1629–1638, SPIE Press, 2003.
5. T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models – their training and application," *Computer Vision and Image Understanding* **61**(1), pp. 38–59, 1995.
6. G. Behiels, F. Maes, D. Vandermeulen, and P. Suetens, "Evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models," *Medical Image Analysis* **6**(1), pp. 47–62, 2002.
7. M. de Bruijne, B. van Ginneken, M. Viergever, and W. Niessen, "Adapting active shape models for 3D segmentation of tubular structures in medical images," in *Information Processing in Medical Imaging*, C. Taylor and A. Noble, eds., *Lecture Notes in Computer Science* **2732**, pp. 136–147, Springer, 2003.
8. C. Studholme, D. Hill, and D. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition* **32**(1), pp. 71–86, 1999.
9. B. van Ginneken, M. de Bruijne, M. Loog, and M. Viergever, "Interactive shape models," in *Medical Imaging: Image Processing*, M. Sonka and M. Fitzpatrick, eds., *Proceedings of SPIE* **5032**, pp. 1206–1216, SPIE Press, 2003.
10. M. van der Laan, L. Bartels, M. Olree, C. Bakker, M. Viergever, and J. Blankensteijn, "MR-techniques for follow-up after endovascular abdominal aortic aneurysm repair," submitted.
11. J. Wever, J. Blankensteijn, J. van Rijn, I. Broeders, B. Eikelboom, and W. Mali, "Inter- and intra-observer variability of CTA measurements obtained after endovascular repair of abdominal aortic aneurysms," *American Journal of Roentgenology* **175**(5), pp. 1297–1282, 2000.