Segmentation of Thrombus in Abdominal Aortic Aneurysms From CTA With Nonparametric Statistical Grey Level Appearance Modeling

Sílvia D. Olabarriaga*, Jean-Michel Rouet, Maxim Fradkin, Marcel Breeuwer, and Wiro J. Niessen, *Member, IEEE*

Abstract—This paper presents a new method for deformable model-based segmentation of lumen and thrombus in abdominal aortic aneurysms from computed tomography (CT) angiography (CTA) scans. First the lumen is segmented based on two positions indicated by the user, and subsequently the resulting surface is used to initialize the automated thrombus segmentation method. For the lumen, the image-derived deformation term is based on a simple grey level model (two thresholds). For the more complex problem of thrombus segmentation, a grey level modeling approach with a nonparametric pattern classification technique is used, namely k-nearest neighbors. The intensity profile sampled along the surface normal is used as classification feature. Manual segmentations are used for training the classifier: samples are collected inside, outside, and at the given boundary positions. The deformation is steered by the most likely class corresponding to the intensity profile at each vertex on the surface. A parameter optimization study is conducted, followed by experiments to assess the overall segmentation quality and the robustness of results against variation in user input. Results obtained in a study of 17 patients show that the agreement with respect to manual segmentations is comparable to previous values reported in the literature, with considerable less user interaction.

Index Terms—Abdominal aortic aneurysm, deformable models, image segmentation, statistical grey level modeling, thrombus segmentation.

I. INTRODUCTION

AN ABDOMINAL aortic aneurysm (AAA) is an abnormal ballooning of the abdominal portion of the aorta, which can break open and cause death. Contrast CT angiography (CTA) has been the preferred imaging modality in AAA treatment because it provides detailed information about the aortic anatomy, making it possible to visualise lumen, calcifications, and thrombus in a minimally invasive manner (see Fig. 1). Segmenting these structures enables the reconstruction of three-dimensional (3-D) patient-specific models that can support measurements needed at various stages of AAA
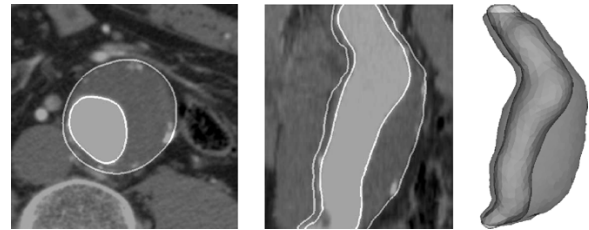

Fig. 1. Example of an abdominal CTA scan: axial (left) and coronal slices (center) showing AAA lumen (inner contour) and thrombus (outer coutour) delineated by experts. On the right, the corresponding 3-D AAA model is shown.

treatment. In particular, these models facilitate the assessment of the rupture risk, which is based on measurements of the native AAA, such as diameter and volume [1] and, more recently, based on indicators derived from hemodynamics simulation [2]. Moreover, the dimensions of the aneurysm neck are important to assess the suitability for endovascular repair, as well as to choose the appropriate stent type and size [1]. Finally, the change of AAA size, preferably volume ([3] and [4]), is an important indicator of rupture risk in postoperative follow-up. In spite of all developments in the field of medical image segmentation, the reconstruction of 3-D AAA models still demands a significant amount of human intervention. In practice, most of the preoperative AAA measurements are obtained by inspecting the image in a slice-by-slice fashion and manually pointing out the structures of interest. Typically, the complete contour is drawn slice-by-slice for volume measurement. Besides being time-consuming and prone to large intraobserver and interobserver variability, this approach is becoming extremely impractical with the increasing data size (e.g., multislice CT). Automated tools for AAA model reconstruction, such as [5] are, therefore, desired for achieving higher efficiency and precision in the assessment of AAA parameters in all treatment phases.

Many highly automated methods for lumen segmentation and tracking have been reported in the literature (e.g., [5]–[7]). Also calcification detection and segmentation have been achieved in an automated fashion [5], [8]. Thrombus segmentation, however, remains as a difficult segmentation problem due to the low contrast between thrombus and surrounding tissue in CTA images (see Fig. 2). Approaches based on image gradient fail because strong responses from neighboring objects, such as the spine, lumen, and calcifications, distract the method from finding the correct boundary. Threshold-based approaches also fail, since similar image intensity can be found inside the

*S. D. Olabarriaga is with the University Medical Center Utrecht, Image Sciences Institute, Q.S. 459, Heidelberglaan 100, 3584CX Utrecht, The Netherlands (e-mail: silvia@isi.uu.nl).

J.-M. Rouet and M. Fradkin are with PMS Research Paris, 92156 Suresnes, France

M. Breeuwer is with PMS, Medical IT – Advanced Development, 5680 DA Best, The Netherlands.

W. J. Niessen is with the University Medical Center Utrecht, Image Sciences Institute, 3584 Utrecht, The Netherlands
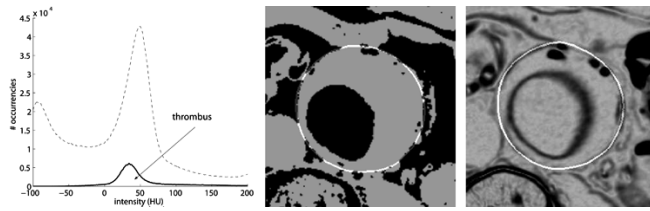
Fig. 2. Difficulties for thrombus segmentation. (Left) Histogram of values inside the thrombus and in the background, showing that this structure is not separable based only on image intensity. (Center) Result of bi-level thresholding in [0,100] HU. (Right) Gradient magnitude of image intensity (inverted image). Note the strong responses (dark regions) at the lumen, the spine and calcifications, and the weak responses at the thrombus boundary.

Fig. 3. Overview of the AAA lumen and thrombus segmentation method.

thrombus and in the neighboring structures. For these reasons, the level of automation and/or accuracy of existing thrombus segmentation methods remains low, motivating our research.

In this paper, a new approach for thrombus segmentation is presented. It is based on a 3-D discrete deformable model (DM) and adopts a nonparametric statistical grey level appearance model built from training data with a supervised pattern classification technique. This segmentation method is initialized with minimal user intervention and generates plausible results in interactive response time. By appropriate training of the statistical model, the method can be adapted in a straightforward manner to varied imaging circumstances, such as when different scanners or acquisition protocols are employed.

This paper initially provides an overview of existing methods in Section II, motivating for the adopted approach described in Section III. The method is evaluated based on 17 patient scans and corresponding manual segmentations using the experimental setup described in Section IV. A parameter optimization study is presented in Section V. An evaluation of the method's overall performance (segmentation quality and robustness to user variation) is described in Section VI and the results presented in Section VII. A discussion and conclusions close the paper in Section VIII.

II. RELATED WORK AND MOTIVATION

Only a few methods for thrombus segmentation have appeared in the literature. The method by De Bruijne *et al.* in [9], [10] is based on active shape models (ASM): a two-dimensional (2-D) contour drawn by the user in one slice is propagated to the adjacent one based on grey values similarity. The reported results are accurate, but the amount of user intervention is large, since slice-by-slice control is required. The method by Subasic *et al.* in [11] is based on 3-D level-sets. A sphere positioned by the user is deformed into the lumen, which is used to initialize the thrombus segmentation method. Simple image features combined in a preprocessing step are used for the thrombus level-set. Although user intervention is minimal, the reported results are not accurate. Another method proposed by De Bruijne *et al.* in [12] uses a 3-D ASM with a grey level appearance model based on a nonparametric pattern classification technique. The user has to draw the top and bottom contours, and indicate the approximate aneurysm center. Reported results are accurate, but the amount of interaction for initialization is still significant. The method recently proposed by Giachetti
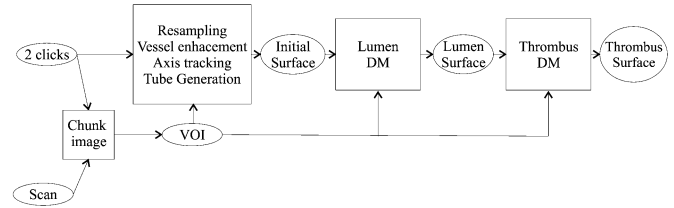
and Zanetti in [5] segments the thrombus slice-by-slice with a 2-D DM initialized with dilated cross-sections of the lumen boundary, which is segmented from a sphere, similarly to [11]. The thrombus is deformed based on a strong shape constraint (truncated Fourier series) and an image feature combining intensity gradient and threshold. The amount of user intervention is significant, due to slice-by-slice control, and accuracy is not reported.

The above methods suffer either from low accuracy or low automation level. In this work, we aim at simultaneously obtaining high automation and segmentation accuracy by combining three concepts present in these methods. First, the thrombus segmentation method consists of a 3-D DM incorporating local shape constraints, avoiding slice-by-slice processing and extra effort for shape training and for alignment with the data at initialization. Second, the method is initialized with the lumen boundary, which can be generated with minimal user intervention. Finally, the DM employs a novel image-based constraint, adopting a grey level appearance modeling strategy based on a supervised pattern classification technique. The adoption of such a sophisticated model is motivated by the observation that simpler ones are unlikely to generate accurate results for the reasons illustrated in Fig. 2. A new grey level appearance model built from training data is, therefore, adopted to generate external constraints that steer the thrombus deformation process efficiently.

III. DEFORMABLE MODEL FOR THROMBUS SEGMENTATION

Fig. 3 shows an overview of the method for lumen and thrombus segmentation, which consists of three steps. First, the user indicates the vessel segment by clicking two points inside the lumen. Second, the user interactively segments the lumen boundary with a DM based on a simple image feature. Finally, the resulting lumen boundary is used to initialize another DM for thrombus segmentation.

A. Thrombus Deformable Model

In segmentation methods based on DMs, also called active objects, an initial boundary is deformed under internal (shape-based) and external (image-based) constraints until an equilibrium is reached (see [13], [14] for reviews). In this work, we adopt a discrete formulation of DMs in which the object boundary is represented by a polygonal mesh with particular topology (2-simplex). The method adopted here, also called 3-D active object (3DAO), was introduced by Delingette in [15] and extended by the MedISys Group, Philips Medical Systems Research Paris [16], [17].

In short, the deformation process consists of considering all vertices in the mesh as physical masses submitted to the

Newtonian law of motion. In a formulation with discrete time steps t, using central finite differences and an explicit scheme, the position of a vertex $\mathbf{x}$ is changed according to the evolution equation

$$\mathbf{x}^{t+1} = \mathbf{x}^t + (1 - \lambda)(\mathbf{x}^t - \mathbf{x}^{t-1}) \\ + \lambda \left(\alpha F_{\text{int}}(\mathbf{x}^t) + \beta F_{\text{ext}}(\mathbf{x^t}) \right), \quad (1)$$

where λ is a damping factor, F_{int} and F_{ext} are, respectively, the internal and external forces acting on the vertex, and α and β indicate their relative importance in the deformation process. Typically, $\alpha + \beta = 1$, and their values are optimized for the application at hand. As in any method based on DMs, the following main aspects determine the success of segmentation: the initial boundary, the proper choice of internal and external forces, their relative balance, and the stopping criterion of the deformation process. Moreover, the mesh resolution is also important in this case, since a discrete implementation is adopted. These aspects are discussed in more detail below.

B. Initial Boundary

The initial boundary corresponds to the lumen surface, which is generated as follows (see, also, [18]). Two positions are clicked by the user at the proximal and distal slices of the vessel segment. From these positions, the volume-of-interest (VOI), in particular the slice interval for thrombus and lumen segmentation, is determined. Also, the two clicks are used to determine the lumen axis with vessel enhancement [19] followed by minimum cost path tracking [20]. A tube is built around this axis and adopted as initial boundary for lumen segmentation.

The lumen DM adopts a simple external force based on bi-level thresholds of image intensity, which are set by default to the expected range of Hounsfield unit (HU) values for contrast-enhanced blood and adapted interactively, if necessary. The deformation process converges automatically and successfully to the lumen boundary in a few seconds. The resolution of the lumen boundary is adapted and used directly to initialize the thrombus DM.

C. Mesh Resolution

The mesh resolution is controlled by a parameter indicating the target mean face area, and the mesh is adapted automatically with a scheme based on split and merge operations applied sequentially. The operations are performed similarly as proposed in [15], while the whole procedure tries to maintain the given mesh resolution. Particularly, split and merge operations are repeated until all mesh faces respect the desired resolution (face area) with a given tolerance. Since 2-simplex faces are not planar, the face area is approximated as a sum of triangles connecting consecutive face vertices and face gravity center.

For efficiency reasons, it is important to find an appropriate trade-off between the number of vertices evaluated during deformation and the loss of accuracy caused by smoothing when a low mesh resolution is adopted. We currently adopt a coarse-to-fine approach: first a low-resolution mesh is deformed until convergence, and then the mesh is refined and deformed again.

D. Internal Force

The internal force imposes regularization and smoothness constraints on the polygonal mesh, with a tangential and a normal component. The tangential component enforces the uniform distribution of vertices on the mesh surface, while the normal component minimizes the mesh mean curvature in a neighborhood of a given size. In the adopted representation, mean curvature is approximated from the elevation angle of the vertex with respect to the tangent plane determined by its three neighbors. Neighborhood size here refers to the topological distance among the vertices – see details in [15] and [21]. As a consequence, the magnitude of the internal forces varies with the mesh resolution: for rougher representations, the internal forces are stronger, and α, β must be adapted accordingly. Typically, α is reduced in a similar ratio as the mesh resolution increases.

E. External Force

The external force attracts the mesh to image positions where certain boundary properties are found. In order to favor a stable evolution process, the external force F_{ext} is constrained to be collinear with the surface normal at the location of the vertex $\mathbf{x}$, corresponding to an inflating (or shrinking) force along the vertex normal $\mathbf{n}$

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

where S is the maximum force strength. Furthermore, a shrinking force is always generated when the vertex is outside the VOI.

To determine whether the vertex is inside, outside, or at the boundary, a grey level appearance modeling strategy inspired by [12] is adopted. A pattern classification approach is used to determine the most likely class corresponding to the intensity pattern at the vertex. Patterns are characterized by the image intensity profiles sampled perpendicularly to the boundary.

For the classification, a nonparametric technique is adopted, namely the k-nearest neighbors (k-NN, see [22]). In this supervised learning technique, the arbitrary probability density functions for each class are estimated by dividing the feature space into variable cells. A cell corresponds to the k "neighbors" (or closest training points) of a given feature point $\mathbf{y}$. This classifier was chosen due to its suitability for dealing with complex decision boundaries in the feature space, which is the case of AAA thrombus appearance.

The probability of $\mathbf{y}$ belonging to a class ω is determined from the density of training points in the cell

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

where k_ω is the number of points belonging to class ω among the k nearest neighbors of $\mathbf{y}$.

Three classes are considered: *inside* (ω_I), *outside* (ω_O), and *boundary* (ω_B). The classification features are image intensity profiles of length l and spacing δ sampled with tri-linear interpolation along the vertex normal $\mathbf{n}$ [Fig. 4 (left)]. Note that l and δ determine the number of classification features.
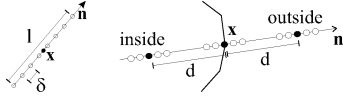
Fig. 4. Intensity profiles used as classification features: profile parameters (left) and distance d for nonboundary samples (right).

During training, profile samples are collected for three classes: one at the correct boundary position $\mathbf{x}$, delineated manually, and two at shifted positions inside and outside the thrombus, respectively, at $\mathbf{x} - d\mathbf{n}$ and $\mathbf{x} + d\mathbf{n}$ [Fig. 4 (right)]. Note that the distance parameter d affects only the classifier training stage.

During deformation, the intensity profile $\mathbf{y}$ at the vertex is sampled, and the most likely class ω_j is determined by $\arg\max_{j \in [I,O,B]} P(\omega_j | \mathbf{y})$. The force strength S is determined as follows:

$$S = sc \tag{4}$$

where s is a constant maximum strength and $c = P(\omega_j | \mathbf{y})$ is the confidence in the classification step. The force magnitude is reduced by the scale factor c when the intensity profile pattern does not clearly indicate the situation at hand. Therefore, the external force magnitude can be small (and cause the deformation to stop) because the vertices have reached the boundary position or due to large classification uncertainty.

Note that the computation of F_{ext} is an expensive step, since the k-nearest neighbors of $\mathbf{y}$ must be determined. In this work, we adopt the approximate nearest neighbors (ANN) implementation by Arya and Mount [23], which allows very fast retrieval of closest neighbors by introducing an adjustable approximation parameter ϵ. Moreover, we limit the training set to a random subset of the available profiles. In this way, the extra computational burden introduced by the classification step is reduced and interactive response time can be achieved. We currently adopt $\epsilon = 1$ HU and 5% of the profiles are included in the training, but we observed that results do not improve when larger training sets or smaller ϵ are used.

F. Stopping Criterion

A simple convergence test is used to stop the deformation automatically when the mesh does not change significantly. Two convergence detection criteria were investigated: the difference in volume and the total mesh displacement between two iterations. When the difference or displacement falls below a given threshold T, the deformation is interrupted. Results are robust with both criteria and a large range of values of T, provided that the T is smaller than the amount of typical change at each iteration; otherwise, the deformation will be interrupted too early. On the other hand, T should not be too small to prevent useless iterations.

We observed experimentally that the proper choice of threshold depends on the damping factor λ and the magnitude of F_{ext}, since these influence the amount of deformation in each step. When λ is large, the deformation develops faster, with larger displacements and fewer iterations. As shown in [15], however, there is a risk of converging to a wrong solution. In all experiments $\lambda = 0.8$ is adopted, although we observed that,
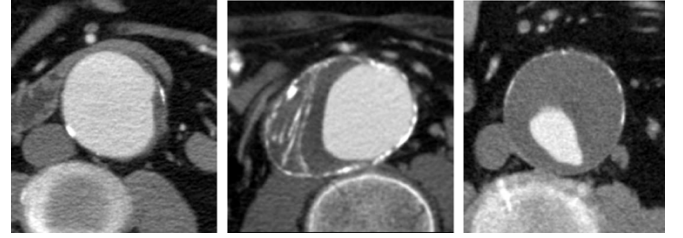

Fig. 5. Types of AAA thrombus found in the patient scans: small thrombus (left), large thrombus with a large amount of calcifications (center) and thrombus containing mainly soft tissue (right).

for weak F_{ext}, the results are quite insensitive to the choice of λ in a large range. The magnitude of F_{ext} is mostly regulated by its strength s. When $T > s$, the deformation may stop too early, while when $T \ll s$, the deformation continues by the action of the internal forces only, leading to under-segmentation. As a compromise, we currently adopt $s = 1.0$ mm and $T = 0.5$ mm; however, no differences are visible for a large range of $T < 1.0$ mm.

IV. EXPERIMENTAL SETUP

In Sections V and VI a large number of experiments are described that investigate the behavior of the method with varying parameter settings. These experiments are performed based on the data and evaluation criteria described below.

A. Patient Data

The images are CTA scans of patients obtained from the endovascular aneurysm repair programme of the University Medical Centre Utrecht (UMCU) [24]. They are acquired with Philips spiral CT scanners [Tomoscan series, Philips Medical Systems, Best, The Netherlands (PMS)], with a resolution of $0.5 \times 0.5 \times 2$ mm. The scans are composed of circa 125 slices of 512×512 voxels. Seventeen preoperative scans containing a native AAA were chosen from this study. In all cases, no stents or other implants are present, the AAA is located in the aortic segment between the branching to the renal and to the iliac arteries, and both lumen and thrombus are visible. While the AAA lumen is typically dilated, the thrombus size and the amount of calcifications vary largely among these patients – see examples in Fig. 5. In some cases (3 scans), the thrombus is small, and its outer contour is located near the lumen boundary. In other scans, the thrombus is large, containing a large amount of calcifications (2 cases) or mainly soft tissue (12 cases).

Manual segmentations of the thrombus boundary are obtained with the PMS EasyVision workstation contouring facilities by one trained observer (SDO). The contours drawn manually on each slice are used to reconstruct a 3-D polygonal mesh, which is then smoothed and resampled, and to generate a binary mask. The 3-D mesh and mask are adopted as reference segmentation in the evaluation and training.

A subset of five randomly selected scans is used for parameter optimization (Section V), and the complete set is used for reporting the quality of segmentation results (Section VI). In all cases, the leave-one-out strategy is adopted, in which all available data is used for training, with the exception of the scan for which the results are reported.

B. Evaluation Measures

The results obtained in the experiments concerning thrombus segmentation are assessed with three objective measures that compare a given segmentation result A and the reference (manual) segmentation R.

- Segmentation Overlap (Ψ): relative volume of overlap between A and R

$$\Psi = 2\frac{V(A \cap R)}{V(A) + V(R)} \tag{5}$$

where A, R are binary masks and $V(\cdot)$ is the number of "on" voxels.

- Volume Error Υ: difference in volume between A and R

$$\Upsilon = \frac{|V(A) - V(R)|}{V(R)}. \tag{6}$$

- Segmentation Error (D_i): distance (in mm) between the vertices $\mathbf{x}_i$ in a given result and the closest (triangular) face in the reconstructed mesh corresponding to R.

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).

V. Parameter Optimization

In this section, experiments are performed with five scans to determine optimal parameters with respect to mesh resolution, profile configuration and balance of external/internal forces.

A. Mesh Resolution

This experiment investigates the loss of segmentation accuracy as a function of mesh resolution. For the scans in the test set, the manual mesh is resampled with increasing mean face area a and used to initialize a simple DM based on the binary (manual) segmentation. The image-based force in this case moves the vertex outwards when it is located inside the thrombus and inwards otherwise. The remaining parameters for this simple DM are $\beta = \alpha = 0.5$ and $s = 1$ mm.

Fig. 6 (left) shows the results: as a increased, the segmentation overlap Ψ dropped due to boundary smoothing. Fig. 6 (right) illustrates the smoothing effect at high and low resolutions. Some loss of accuracy was observed even for the highest mesh resolution ($a = 1$ mm^2), but these differences were negligible in a slice-by-slice visual inspection ($\mu_\Psi = 99.1 \pm 0.2\%$). Note that accuracy was acceptable ($\mu_\Psi > 98\%$) for $a \leq 20$ mm^2.

In subsequent experiments, different mesh resolutions are chosen for training and for deformation. The highest resolution ($a = 1$ mm^2) is adopted for training the classifier and as reference segmentation, where maximum accuracy is desired. During deformation, the coarse and fine mesh resolutions are, respectively, set to $a_C = 50$ and $a_F = 20$ mm^2.

B. Profile Configuration

Here, we investigate the optimal profile parameters (length l and sampling interval δ) and the optimal distance for obtaining nonboundary samples based on the test scans (d).
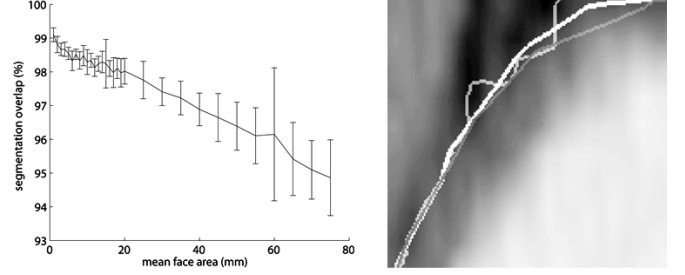

Fig. 6. Segmentation overlap as a function of mean face area (left) and effect of mesh smoothing with $a = 1, 20, 50$ mm^2 in a cross-sectional view where the differences are most noticeable (right).
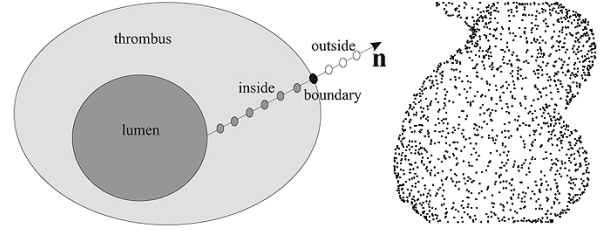

Fig. 7. Positions of test profiles samples, simulating configurations possibly encountered during the deformation process: cross-sectional (left) and 3-D (right) illustrations.

Initially, the classification accuracy Ω (percent of correctly classified samples) is measured for all possible combinations (72) of the following parameter settings: $l \in \{2, 3, 4, 5, 6, 7\}$ mm, $\delta \in \{.5, 1\}$ mm, and $d = \{.5, 1, 2, 3, 4, 5\}$ mm. The training set is composed of profiles collected as described in Section III-E. The classified profiles are collected as illustrated in Fig. 7 at 10% of randomly chosen vertices in the reference segmentation. One sample of ω_B is collected at each position, and the other samples are collected with 1.5 mm spacing. A margin of 4.5 mm around the thrombus boundary is covered for collecting samples of ω_O. For ω_I, samples are collected until the lumen is reached. The classifier is applied with $k = 23$, but the overall classification quality is quite stable for a large range of values $k \in [10, 50]$. The "true" classification is given by a labeled image indicating the location of the reference thrombus, lumen and background. The results obtained with this experiment are discussed below.

With respect to the differences in Ω for $\delta = 0.5$ and 1 mm, we observed that they were very small (in average $0.05 \pm 0.66\%$) and not statistically significant in a paired T-test ($p = 0.18$). Results obtained with $\delta = 0.5$ mm (the in-slice image resolution) were not more accurate than those obtained with $\delta = 1.0$ mm, but the first configuration required twice as many features, and, consequently, a more costly classification process. In the subsequent analysis we, therefore, concentrate on the results obtained with $\delta = 1$ mm.

Table I and Fig. 8 show the average Ω obtained with different l and d for the test scans. The overall classification performance was typically higher for $d = 3$ mm, but no preferred value of l became evident from these results.

In the second part of this experiment, the segmentation method is applied to the test scans for profiles with varying l and fixed $\delta = 1$ mm, $d = 3$ mm. The DM parameters are set

TABLE I
CLASSIFICATION ACCURACY OF TEST PROFILES FOR VARYING l AND d,
WITH $\delta = 1.0$. MEAN AND STANDARD DEVIATION FOR 5 SCANS

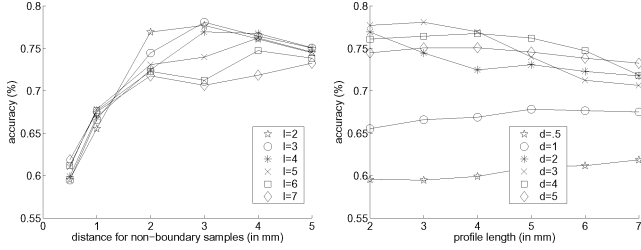| 1	$d = .5$	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$
2	59±4.9	65±4.5	76±3.9	77±5.0	76±4.7	74± 5.0
3	59±4.7	66±4.4	74±2.5	78±4.5	76±4.6	75±4.8
4	59±4.9	66±4.3	72±2.9	77±3.8	76±4.6	75±5.0
5	61±4.1	67±4.1	73±3.0	74±2.7	76±4.7	74±5.0
6	61±4.3	67±4.0	72±2.8	71±2.5	74±4.5	73±5.1
7	61±3.9	67±4.0	71±3.0	70±2.5	71±3.8	73±5.0

Fig. 8. Classification accuracy (Ω) for varying d (left) and l (right), with $\delta = 1$ mm. Mean results for 5 scans.
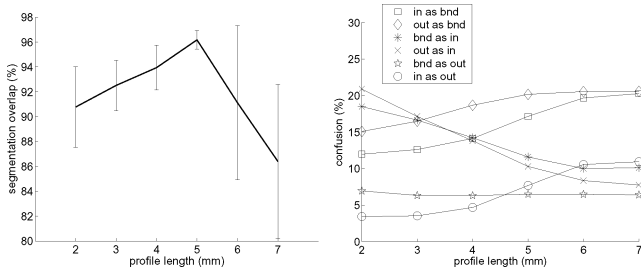

Fig. 9. (Left) Mean Ψ obtained with $\delta = 1$ mm, $d = 3$ mm and varying l. (Right) Mean classification confusion for each class obtained with the same configuration.

as described in Section III-E, with $a_C = 50$, $\beta_C = 0.4$ for the coarse resolution, $a_F = 20$, $\beta_F = 0.2$ for the fine resolution, and $\alpha = 1 - \beta$. The training set and classifier parameters are the same as above.

Fig. 9 (left) shows the average segmentation overlap obtained with varying l for the test scans, indicating that the average performance was superior for $l = 5$ mm. Fig. 10 illustrates the behavior observed for varying l: the boundary expanded too much for smaller l ("explosion") and, inversely, it did not expand enough for larger l ("shrinking"). This behavior can be understood by analysing the classification confusion determined in the previous experiment (see Fig. 9 (right), which corresponds to the number of samples (percent) of class ω_X that are classified as ω_Y by the k-NN classifier. During the deformation, when a boundary or outside profile is misclassified as inside, an external force pointing outwards is generated and the boundary expands erroneously (explosion). The inverse effect (shrinking) occurs when a boundary or inside profile is misclassified as outside, and an external force pointing inwards is generated. Note that the misclassification of inside as outside class and vice-versa was roughly balanced for $l = 5$ mm, possibly explaining why optimal segmentation results are obtained with this configuration.
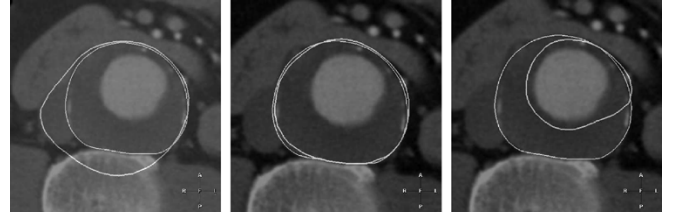

Fig. 10. Illustration of the method's behavior for varying profile length: cross-section showing the manual and the automatic segmentation obtained for $l = 2$ mm (left), $l = 5$ mm (center) and $l = 7$ mm (right) with $\delta = 0.5$ mm and $d = 3$ mm.
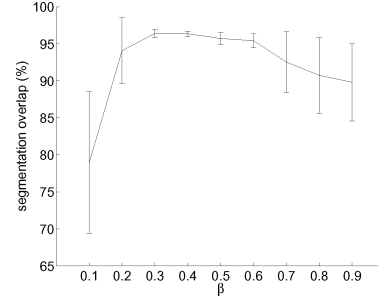

Fig. 11. Segmentation overlap (Ψ) for varying β_C: mean and standard deviation for five test scans.

In conclusion, the following profile configuration provided optimal results on the test scans: $\delta = 1$ mm, $l = 5$ mm, and $d = 3$ mm.

C. Balance of Internal and External Forces

This experiment investigates the behavior with respect to the parameters β and α, which regulate the balance of internal and external forces. With all the other parameters set as described in Section V-B, the value of β is varied for the coarse resolution in the range $\beta_C \in [0.1, 0.9]$ with spacing of 0.1. For the fine resolution, $\beta_F = \beta_C/2$.

Fig. 11 shows the average segmentation overlap for 5 scans. In average, results were better for $\beta_C = 0.4$, although similar segmentations were obtained for $\beta_C \in [0.3, 0.5]$. For smaller β, the external force was sometimes too weak and the boundary remained close to the lumen. For larger β_C, F_{ext} may be too strong and lead the boundary into the vena cava and bowels.

VI. PERFORMANCE EVALUATION

In these experiments, the goal is to assess the quality of segmentation results obtained with the proposed method, as well as their robustness with respect to small perturbations in the user input. The method is applied to all 17 scans with the optimal DM parameters determined in Section V: $\delta = 1$ mm, $l = 5$ mm, and $d = 3$ mm; $a_C = 50$ and $a_F = 20$ mm²; and $\beta_C = 0.4$ and $\beta_F = 0.2$. Convergence is detected automatically as explained in Section III-F.

A. Quality of Thrombus Segmentation Results

The purpose of this experiment is to assess segmentation quality via the measures in Section IV-B. Additionally, the number of iterations and the elapsed time for deformation (total for coarse and fine resolutions) are presented to provide an

TABLE II
SUMMARY OF RESULTS: MEAN, STANDARD DEVIATION,
MINIMUM AND MAXIMUM VALUES FOR 17 SCANS

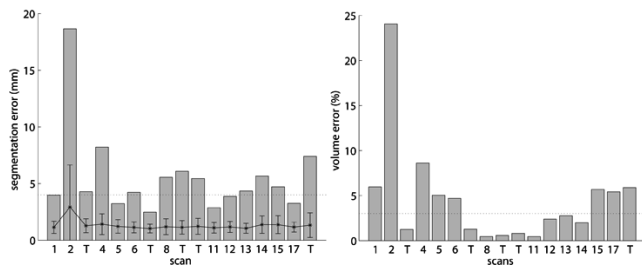|	$\mu \pm$ std. dev.	[min , max]
Segmentation Overlap (Ψ,%)	95.0 ± 3.3	[82.8, 96.8]
Volume Error (Υ, %)	4.5 ± 5.6	[0.4, 24.1]
Mean Segmentation Error (μ_{D_i}, mm)	1.3 ± 0.4	[1.0, 2.9]
Maximum Distance ($\max_i D_i$, mm)	5.5 ± 3.7	[2.5, 18.7]
$D_i \leq 1$ mm (% vertices)	42.5 ± 5.7	[31.1, 50.6]
$D_i \leq 2$ mm (% vertices)	89.8 ± 7.2	[69.3, 98.5]
Number of iterations	301 ± 274	[83, 1242]
Elapsed deformation time (s)	41 ± 69.3	[6.7, 305.5]

Fig. 12. Segmentation and volume error for each scan, respectively, D_i (left) and Υ (right). The five test scans are marked with T. For D_i: maximum (bars), mean (dots), and standard deviation (error bars). The horizontal lines indicate, respectively, $\max_i D_i = 4$ mm and $\Upsilon = 2\%$.
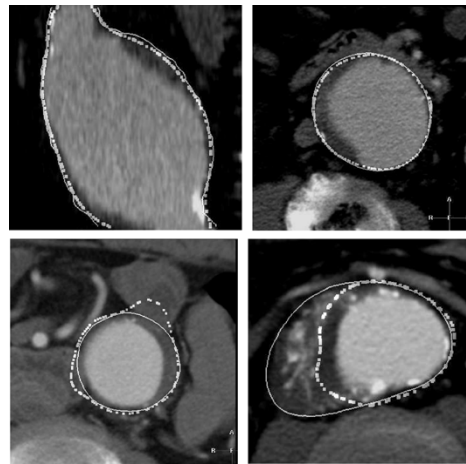

Fig. 13. Selected segmentation results (dots) and the manual segmentation (line). *Upper*: coronal and axial slices of a successful segmentation (scan 6). *Lower left*: bump into bowels and vena cava (scan 4). *Lower right*: worst result (scan 2).

impression about the computational effort required to generate the reported results.

B. Robustness Against User Variation

Here, we investigate how unintentional variations in the user input influence the thrombus segmentation quality, namely small perturbations in the two clicks provided for lumen initialization. The two positions used in the previous experiments are randomly displaced in a circle with radius of 6 mm in the axial plane and used to generate the initial lumen surface (smaller perturbations did not affect the lumen result). Points outside the lumen were rejected.

In total, five combinations of perturbed clicks are used to initialize the method as described in Section III-B for all 17 scans. The resulting variation is measured by the largest difference in segmentation overlap $V = \max_i \Psi_i - \min_i \Psi_i$ observed for the five segmentation results i.

VII. RESULTS

Table II summarizes the average results obtained for 17 scans with the experiment described in Section VI-A. In average, 95% volume overlap and 4.5% volume error were obtained, closely to the limit imposed by the chosen boundary representation (see Section V-A). Moreover, the average segmentation error was 1.3 mm, and the large majority of vertices (89.8%) were within a distance of 2 mm from the correct boundary.

Fig. 12 presents the segmentation error and the error in volume calculation for each scan. In six cases, the maximum D_i was below 4 mm, which was considered an acceptable error in the experiment simulating interactive segmentation described

in [10]. Fig. 13 (upper) illustrates one of these results, which possibly displays acceptable accuracy for most clinical applications. In the remaining scans, errors typically corresponded to bumps where the boundary expanded into the vena cava and bowels – Fig. 13 (lower left) shows an extreme example. In no scan the thrombus boundary completely leaked into neighboring structures. Depending on the accuracy requirements of the application, these localized errors could be manually corrected with efficient 3-D dragging tools. Note, however, that in nine cases $\Upsilon \leq 3\%$ and possibly would require no correction at all, since the error is below the average intraobserver volume difference of 3.2% reported in [25]. Only for one scan, with a large amount of calcifications inside the thrombus, the result was not satisfactory at all because the deformation stopped too early [see Fig. 13 (lower right)].

With respect to computational effort, the number of iterations and the deformation time varied largely among scans. In 10 cases, the deformation converged in less than 30 s, and only in one case (scan 14) it took more than 1 min. If this outlier is ignored, the average deformation time drops to 24.5 ± 12.8 s, which we consider admissible in an interactive setting.

Finally, with respect to robustness against user variation (Section VI-B), results were very encouraging. Although the resulting boundaries displayed minor differences, the variation V in the segmentation overlap was small, in average $1 \pm 0.8\%$. In only one case, out of 85, the generated result was significantly different from the remaining four produced for that scan.

VIII. DISCUSSION AND CONCLUSIONS

The proposed segmentation method consists of a discrete DM that adopts an image-based force inspired by approaches where an appearance model built from grey value information contained in a training set is fitted to the data. This approach is an adaptation of ASM introduced by Cootes *et al.* [26], in which a parametric statistical 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

[12]. Extensions to grey level modeling for ASMs were proposed in [12], [27], in which the model is trained based not only on samples at boundary positions, but also inside and outside the object. In both cases, a nonparametric classification technique (k-NN) is used to determine the new boundary position by taking the best fit between the model and the data among a number of candidates. Finally, in the DM method proposed by Pardo *et al.* [28], the model is trained in one slice based on a 2-D contour, with samples at boundary and nonboundary positions. The most discriminating classification features are selected automatically from a Gaussian bank, and the obtained parametric model is propagated to the adjacent slice.

The methods above are similar with respect to the following aspect: the deformation is driven by a quality of fit that only takes into account the boundary/nonboundary distinction. Compared to these works, one novelty is introduced here: the proposed method uses a three-type class model that also distinguishes inside and outside nonboundary profiles, and in this way steers the deformation direction directly and efficiently. We found that our approach imposes fewer requirements on the initial surface than when an on-off boundary model is adopted. In particular, in experiments we observed that the AAA thrombus segmentation was not always successful based the on-off boundary model introduced in [9] when the DM is initialized at the lumen boundary. When the thrombus is large, all the candidate positions may fall inside the thrombus, and no deformation occurs. A larger search range would reduce this problem, however, at a high computational cost, since the expensive classification step would have to be performed for a large number of candidate positions. In the proposed deformation scheme, on the other hand, the distance between the initial and the target boundaries is less relevant. When properly trained, the method is in principle capable of identifying intensity patterns located far from the target and move the boundary outwards or inwards accordingly.

The novel grey level modeling and deformation scheme has been successfully applied to the difficult problem of AAA thrombus segmentation in 17 patient scans (Section IV). A large number of experiments were performed to determine the optimal parameter configuration for five test scans (Section V). With this configuration, the thrombus DM produced plausible segmentation results in interactive response time (< 1 min) for 16 out of 17 scans (Section VII). Depending on the accuracy requirements imposed by the application, six (or possibly nine) results obtained with minimal user interaction would be considered acceptable. In ten (or seven) cases, small and localized bumps into the bowels and vena cava were observed, and would possibly require manual correction (for example, with an efficient 3-D dragging tool). Only in one scan the obtained result was not satisfactory, and would demand more manual editing. Finally, these results were shown to be robust with respect to the initial clicked positions (Section VI-B), favoring segmentation precision.

With respect to the average quality of results obtained in this initial study (see Table II), it seems compatible with values found in the literature. The average segmentation overlap (95%) and volume error (4.5%) are similar to results reported in [12] (respectively, 95% and 5.1%) and comparable to the

intraobserver variability for AAA thrombus volume assessment from CTA reported in [25] (3.2%). The segmentation error also seems to be similar to [12], although the values are not directly comparable. Note, however, that the results reported here were generated with significantly less user intervention, favoring segmentation efficiency. The three-type class model was essential for obtaining these results: we found that such performance could not be obtained with a DM using on-off boundary profile training, for which the demands on the initial boundary are higher. The same is true for simpler, threshold-based or gradient-based image forces, which often generate implausible results. In particular, we observed that, with bi-level threshold, the boundary often embraces other structures with similar appearance. For the best results obtained with this image force (when the balance between internal and external constraints is manually tuned for each scan), an average segmentation accuracy of $89.6 \pm 3.7\%$ and volume error of $10.2 \pm 7.3\%$ were obtained. When gradient magnitude was adopted, the surface was easily attracted back to the lumen boundary. Results obtained with the proposed method are, therefore, encouraging and indicate the feasibility of achieving efficient and precise segmentation of AAA thrombus and lumen. We believe that this method is an efficient alternative for manual segmentation that could be valuable for the measurement of AAA volume, facilitating the assessment of AAA rupture risk in a more sensitive manner [3], [4].

The proposed method must now be validated for a larger set of scans for assessing its real utility in a more general context. With the same training, the method is currently applied to images of the Hemodyn project [2], which are acquired at a different hospital with a different scanner and protocol. In particular, the slice thickness varies between 2 and 5 mm. Initial results are promising, indicating similar performance as found in our limited study: segmentation accuracy of $95.8 \pm 1.0\%$ and volume error of $3.8 \pm 2.2\%$ (average for three scans).

Finally, the grey level model adopted here is built from training data using a statistical pattern classification technique with supervised learning. As such, the segmentation method's performance depends on the quality of training, which should include sufficient and representative examples of the target patterns. In other words, the proposed method must (and can) be customized by training to particular imaging conditions. The evaluation presented here, for example, is based on preoperative scans obtained in the context of a particular study that adopted a standardized imaging protocol. Although initial results obtained for the Hemodyn project indicate that this training is also valid to some degree for other scanning protocols, applying the proposed method to postoperative scans, or to scans acquired with multi-slice CT scanners, would possibly benefit from new training, since the AAA thrombus appearance is likely to be different. Note that the need of training is at the same time a "curse" and a "blessing" of supervised learning approaches. While training typically requires significant effort for manual annotation, on the other hand, it provides a structured and straightforward manner of adapting the method to new circumstances. We feel that adopting a structured approach for appearance modeling is preferable to *ad hoc* procedures, which also require revision when the imaging circumstances change.

References

[1] S. C. Whitaker, "Imaging abdominal aortic aneurysm before and after endoluminal stent-graft repair," *Eur. J. Radiol.*, vol. 39, pp. 3–15, 2001.

[2] M. Breeuwer, U. Gotte, R. Hoogeveen, B. J. B. M. Wolters, S. de Putter, H. v.d. Bosch, J. Buth, J.-M. Rouet, and F. Laffargue, "Assessment of the rupture risk of abdominal aortic aneurysms by patient-specific hemodynamic modeling? initial results," in *Computer Assisted Radiology and Surgery*, ser. Int. Congr. Amsterdam, The Netherlands: Elsevier, 2004, vol. 1268, pp. 1090–1095.

[3] J. J. Wever, J. D. Blankensteijn, W. P. T. M. Mali, and B. C. Eikelboom, "Maximal aneurysm diameter follow-up is inadequate after endovascular abdominal aortic aneurysm repair," *Eur. J. Vasc. Endovasc. Surg.*, vol. 20, pp. 177–182, 2000.

[4] B. Kritpracha, H. G. Beebe, and A. J. Comerota, "Aortic diameter is an insensitive measurement of early aneurysm expansion after endografting," *J. Endovasc. Ther.*, vol. 11, no. 2, pp. 184–190, 2004.

[5] A. Giachetti and G. Zanetti, "AQUATICS reconstruction software: The design of a diagnostic tool based on computer vision algorithms," in *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*. Berlin, Germany: Springer-Verlag, 2004, vol. 3117, LNCS, pp. 48–63.

[6] O. Wink, W. J. Niessen, and M. A. Viergever, "Fast delineation and visualization of vessels in 3-D angiographic images," *IEEE Trans. Med. Imag.*, vol. 19, no. 4, pp. 337–346, Apr. 2000.

[7] A. Giachetti, M. Tuveri, and G. Zanetti, "Reconstruction and web distribution of measurable arterial models," *Med. Image Anal.*, vol. 7, no. 1, pp. 79–93, 2003.

[8] I. Isgum, B. van Ginneken, and M. Olree, "Automatic detection of calcifications in the aorta from CT scans of the abdomen," *Academic Radiol.*, vol. 11, no. 3, pp. 247–257, 2004.

[9] M. de Bruijne, B. van Ginneken, W. J. Niessen, J. B. A. Maintz, and M. A. Viergever, "Active shape model based segmentation of abdominal aortic aneurysms in CTA images," *Proc. SPIE Med. Imag.*, vol. 4684, pp. 463–474, 2002.

[10] M. de Bruijne, B. van Ginneken, M. A. Viergever, and W. J. Niessen, "Interactive segmentation of abdominal aortic aneurysms in CTA data," *Med. Image Anal.*, vol. 8, no. 2, pp. 127–138, 2004.

[11] M. Subasic, S. Loncaric, and E. Sorantin, "3-D image analysis of abdominal aortic aneurism," *Proc. SPIE Med. Imag.*, vol. 4684, pp. 1681–1689, 2002.

[12] M. de Bruijne, B. van Ginneken, W. J. Niessen, and M. A. Viergever, "Adapting active shape models for 3D segmentation of tubular structures in medical images," in *Information Processing in Medical Imaging*. Berlin, Germany: Springer-Verlag, 2003, vol. 2732, LNCS, pp. 136–147.

[13] C. Xu, D. L. Pham, and J. L. Prince, "Image segmentation using deformable models," in *SPIE Handbook on Med. Imag.*. Bellingham, WA: SPIE, 2000, vol. 2, ch. 3, pp. 129–174.

[14] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Med. Image Anal.*, vol. 1, no. 2, pp. 91–108, 1996.

[15] H. Delingette, "General object reconstruction based on simplex meshes," *Int. J. Comput. Vis.*, vol. 32, no. 2, pp. 111–146, 1999.

[16] O. Gérard, M. Fradkin, A. C. Billon, M. Jacob, J.-M. Rouet, and S. Makram-Ebeid, "Automatic analysis of the left ventricle in the time sequence of 3D echo-cardiographic images," in *Medical Image Computing and Computer-Assisted Interventions*. Berlin, Germany: Springer-Verlag, 2001, vol. 2208, LNCS, pp. 1224–1225.

[17] O. Gérard, A. C. Billon, J.-M. Rouet, M. Jacob, M. Fradkin, and C. Allouche, "Efficient model-based quantifications of left ventricular function in 3–D echocardiography," *IEEE Trans. Med. Imag.*, vol. 21, no. 9, pp. 1059–1068, Sep. 2002.

[18] S. D. Olabarriaga, M. Breeuwer, and W. J. Niessen, "Segmentation of abdominal aortic aneurysms with a nonparametric appearance model," in *Computer Vision and Mathematical Methods in Medical and Biomedical Image Analysis*. Berlin, Germany: Springer-Verlag, 2004, vol. 3117, LNCS, pp. 257–268.

[19] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention*. Berlin, Germany: Springer-Verlag, 1998, vol. 1496, LNCS, pp. 130–137.

[20] O. Wink, W. J. A. Niessen, and M. A. Viergever, "Minimum cost path determination using a simple heuristic function," in *ICPR: Image, Speech and Signal Processing*. Los Alamitos, NM: IEEE Comput. Soc., 2000, vol. 3, pp. 1010–1013.

[21] J. Montagnat, H. Delingette, N. Scapel, and N. Ayache, "Representation, Shape, Topology and Evolution of Deformable Surfaces. Application to 3D Medical Image Segmentation," INRIA, Sophia Antipolis, France, Tech. Rep. 3954, 2000.

[22] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley, 2001.

[23] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *J. ACM*, vol. 45, no. 6, pp. 891–923, 1998.

[24] M. Prinssen, J. J. Wever, W. P. T. M. Mali, B. C. Eikelboom, and J. D. Blankensteijn, "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.*, vol. 33, pp. 64–69, 2001.

[25] J. J. Wever, J. D. Blankensteijn, J. C. van Rijn, I. A. M. J. Broeders, B. C. Eikelboom, and W. P. T. M. Mali, "Inter- and intraobserver variability of CT measurements obtained after endovascular repair of abdominal aortic aneurysms," *Am. J. Roentgenol.*, vol. 175, pp. 1279–1282, 2000.

[26] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," *Imag. Vis. Comput.*, vol. 12, no. 6, pp. 355–366, 1994.

[27] 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. Med. Imag.*, vol. 21, no. 8, pp. 924–933, Aug. 2002.

[28] X. M. Pardo, P. Radeva, and D. Cabello, "Discriminant snakes for 3D reconstruction of anatomical organs," *Med. Image Anal.*, vol. 7, no. 3, pp. 293–310, 2003.