SELECTED PAPER FROM THE 3RD INTERNATIONAL
CONFERENCE ON COMPUTATIONAL & MATHEMATICAL BIOMEDICAL
ENGINEERING 2013

Combining region-based and imprecise boundary-based cues for interactive medical image segmentation

Jonathan-Lee Jones, Xianghua Xie*,† and Ehab Essa

Department of Computer Science, Swansea University, Swansea, UK

SUMMARY

In this paper, we present an approach combining both region selection and user point selection for user-assisted segmentation as either an enclosed object or an open curve, investigate the method of image segmentation in specific medical applications (user-assisted segmentation of the media–adventitia border in intravascular ultrasound images, and lumen border in optical coherence tomography images), and then demonstrate the method with generic images to show how it could be utilized in other types of medical image and is not limited to the applications described. The proposed method combines point-based soft constraint on object boundary and stroke-based regional constraint. The user points act as attraction points and are treated as soft constraints rather than hard constraints that the segmented boundary has to pass through. The user can also use strokes to specify region of interest. The probabilities of region of interest for each pixel are then calculated, and their discontinuity is used to indicate object boundary. The combinations of different types of user constraints and image features allow flexible and robust segmentation, which is formulated as an energy minimization problem on a multilayered graph and is solved using a shortest path search algorithm. We show that this combinatorial approach allows efficient and effective interactive segmentation, which can be used with both open and closed curves to segment a variety of images in different ways. The proposed method is demonstrated in the two medical applications, that is, intravascular ultrasound and optical coherence tomography images, where image artefacts such as acoustic shadow and calcification are commonplace and thus user guidance is desirable. We carried out both qualitative and quantitative analysis of the results for the medical data; comparing the proposed method against a number of interactive segmentation techniques. Copyright © 2014 John Wiley & Sons, Ltd.

1. INTRODUCTION

Effectively separating objects of interest from background in images is of vital importance to many medical applications. Automated techniques are appealing in terms of efficiency. More often than not, prior knowledge about object appearance and/or shape is necessary to achieve meaningful results. However, it is not always practical, or even possible, to obtain comprehensive prior information and a sufficiently robust learning algorithm to deal with large and sometimes unpredictable variations in real world images. An alternative approach to automated segmentation is to allow and encourage user input and provide interactive segmentation results to suffice user demand. Often, one dilemma is to balance user involvement and interaction flexibility, particularly given the ubiquitous imaging device and ever-increasing amount of images in modern age. To effectively and efficiently capture user intent is vitally important.

---

*Correspondence to: X. Xie, Department of Computer Science, Swansea University, Singleton Park, Swansea, UK,
  SA2 8PP.
†E-mail: x.xie@swansea.ac.uk

The user interaction is conventionally made either by simple mouse click or drag operations on the region of interest (ROI) or on the object boundary. An example is Intelligent Paint [1], which is a simple interactive method that allows the user to identify all the regions inside an object. The object region is interactively expanded by a simple click and drag operation. Then a homogeneous area that has the same intensity profile is selected, which is very similar to those techniques adopted in commercial software. Similarly, with Magic Wand, for example in Adobe Photoshop, the user by simple clicks highlights all the homogeneous area that has the same intensity profile. The Intelligent Scissors [2] and Live Wire [3] methods are among early methods to perform on the fly segmentation by allowing the user to follow the object boundary, instead of region, through a few mouse clicks. These methods are based on well-known shortest path algorithms, such as Dijkstra's method, to find the optimal shortest path between two user points. Shortest path methods have an advantage of segmenting both open and close end objects. However, often only edge-based features are used to find the shortest path, and more importantly usually those user points are treated as anchor points that the segmented path has to go through.

With the help of powerful optimization techniques, the method of user interaction has been expanded, for example, adding object/background strokes, at the same time simplifying user involvement compared with painstakingly tracing the object boundary [3–9]. For example, the user can simply draw multiple strokes inside and outside the object, then the segmentation model can learn the distribution of pixel intensities for both object and background. These techniques usually are more suited for segmenting closed objects and not for open curve segmentation, but in the proposed method we adapt the method so that it can be used with both open and closed objects.

Graph-cut algorithms are widely used to find optimal solution in interactive segmentation at polynomial time complexity and usually for segmenting closed objects. Boykov and Jolly [4] introduced a graph-cut-based interactive segmentation method by defining unary and pairwise costs of each pixel. The unary cost is inversely proportional to the probability of each pixel to be in the object or in the background, while the pairwise cost is based on the intensity difference between two neighbouring pixels. Many methods have been introduced to extend this method, such as GrabCut [6] and Lazy Snapping [10]. In the GrabCut, the authors proposed to use a Gaussian mixture model to build a local colour model to enhance the unary cost. It reduces the user intervention by allowing the user to define a rectangular window surrounding the object. Lazy Snapping is also based on graph cut over a pre-segmented image using a watershed algorithm. K-means is used to cluster the foreground/background colours and assign each pixel to the nearest cluster. The method also has a boundary editing tool to refine the result. However, this method usually needs a multiple user intervention to correctly cut out the object because of the simplicity in cost function.

Shortest path is another optimization technique that has been used in interactive segmentation, for example [2, 3, 9]. These methods place the emphasis on boundary-based features; edge-based features are used to define the cost between pixels. The user interactively identifies a starting point of the path and iteratively adds more seeds around the outline of the object. On the other hand, the intelligent paint method [1] allows the user to identify regions inside the object instead of the boundary. The region is interactively expanded by simple click and dragging operations. Shortest path has an advantage of segmenting both open and close end objects; however, many of them just use edge-based features to find the shortest path, and the user points are generally treated as hard constraints.

Incorporating shape prior to graph-based segmentation has also shown to improve the segmentation result, for example [5, 11–13]. Veksler [5] introduced a star shape prior to graph cut, also through user interaction. The user is required to specify the centre of ROI as the star point, and hence all boundary points of ROI lie on the radial spikes from the star point. Additional points, specifying foreground and background, are often necessary. A ballooning term is also used to discourage bias toward small segment. However, the method can only segment the convex shapes. Gulshan *et al.* [14] have extended the method to multiple stars by using geodesic paths instead of Euclidean rays. Other interactive segmentation methods such as a transductive framework of Laplacian graph regularizer [15] have been also introduced.

In this work, we propose an approach to combine two different types of user interactions, that is, boundary-based interaction (utilizing the user input control points) and region-based stroke
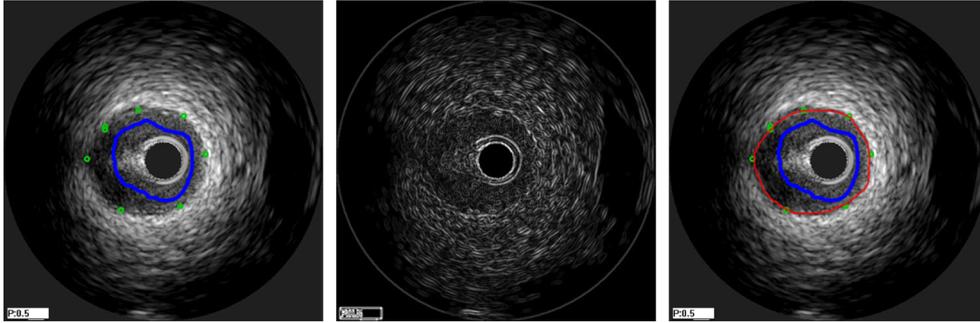
Figure 1. Edge-based detection in IVUS images. From left to right: initial image with user selections added, resultant edge map obtained, and the segmentation produced (shown in red).
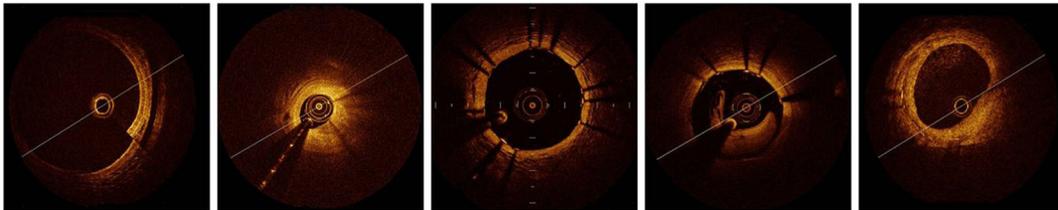

Figure 2. Examples of artefacts in OCT images. These are (from left to right) sew-up (shearing) error, guide-wire reflection/shadow, stents, swirls, and speckles caused by blood in lumen occluding the light source.

interaction, to segment the image. Unlike previous edge-based methods, we utilize a series of soft constraints to guide the image segmentation. By switching to soft constraint, imprecise user input is allowed, without adversely affecting the segmentation. To further augment this edge-based approach, we allow the user to select regions for foreground interest with strokes. We show that this allows effective combination of boundary and region-based features. The user points give the user control over the segmentation process, allowing errors in segmentation to be easily prevented and a more desirable result to be obtained; see Figure 1. We investigate two medical applications for our proposed method: intravascular ultrasound (IVUS) imaging and optical coherence tomography (OCT). Both of these modalities are catheter based and used in cardiology diagnosis, with IVUS being more commonplace. These catheter-based approaches can be used to assess the severity of any stenoses present, to categorize their morphology, and also allow for the measurement of vessel diameter allowing the severity of any occlusions to be assessed by a clinician and the location of any lesions, as well as many other clinical and therapeutic studies [16]. In most IVUS images, a cross-section of the arterial wall is proceeded, with three regions: the lumen, the vessel (made up of the intima and media layers), and the adventitia surrounding the vessel wall. The media–adventitia border is the dividing layer representing the outer arterial wall. In IVUS images, the media can be seen as a dark band, with no other distinct features. It is encapsulated by the adventitia, which is a wide layer of fibrous connective tissue. OCT images are segmented to reveal the lumen border. This is clearly visible for the most part, but in a similar fashion to IVUS images there are a number of artefacts present. These are reflections/shadows caused by the guide wire, reflections/shadows caused by stents, and other anomalies such as shearing (sew-up errors), presence of blood in lumen etc.; see Figure 2 for some examples.

Because of the nature of these artefacts, automatic methods employed in these applications require a significant amount of pre-processing [17, 18]: a large data set for training set [19] or a method to remove the artefacts [20]. Our proposed method, by giving the user control to help the process, removes the need for extensive pre-processing of our large data sets for training. The nature of the IVUS and OCT images, with very pronounced artefacts, lend themselves well to this semi-automatic approach, with most parts of the image being such that the automatic process will be suitable, but by allowing user input, the difficult regions (such as shadows and various artefacts)

can be accommodated easily. Interactive segmentation that can efficiently and effectively transfer user knowledge to segmentation is thus highly desirable in this application. It also allows us to quantitatively evaluate our method, including both efficiency and accuracy.

The rest of the paper is organized as follows. Section 2 presents the proposed method, including user input, superpixel segmentation, and multilayered graph segmentation. Experimental results from segmenting medical data sets with ground truth are presented in Section 3. These show the results for IVUS and OCT images and quantitive data. Also included are results for generic images. Section 4 concludes the paper.

2. PROPOSED METHOD

The proposed method involves the user selecting a series of user control points on the image. These represent the start and the end point for the segmentation, and the user-selected points act as the attraction points in the shortest path search that results in the segmentation. These user-selected points act in a fashion similar to an elastic band, pulling the segmentation toward them. In this way, it is possible for the user to influence the segmentation process allowing them to preferentially select features that they want. In order to enhance the image segmentation, the user can also select areas for foreground using strokes.

An energy functional is then formulated on the basis of the combination of the attraction force that is computed using distance transform and the discontinuity in foreground probability. By assuming that the user points are in a sequential order (as without this the problem will be NP-hard (Non-deterministic Polynomial-time hard) and computationally intensive), we construct a multilayer graph with each layer encapsulating a single individual user point. Effectively, we create multiple identical layers, all made up of duplicates of the image, for each user point added. The segmentation problem is then transformed into searching the shortest path in this layered graph; that is, the resulting segmentation is obtained through searching a minimum path in this stack of layers in a manner similar to a 3D object. This layered approach allows the segmentation to be carried out in polynomial time, instead of an NP-hard optimization problem, at the same time achieving global minima.

2.1. User input

The proposed method allows two different types of user input: attraction points to indicate the edge of the desired object and strokes to indicate the ROI. Figure 1 provides an example of segmentation using the proposed method. Conventionally, user input to segmentation is focused on foreground and background specification [4–7]. For example, in [6], the user interaction consists of dragging a rectangle around the object of interest and in doing so the user specifies a region of background that is modelled in separating the foreground object. Several other methods require the user to specify points on the object boundaries instead [3, 8, 9] examples of other common methods are shown in Figure 3. However, more often than not, these boundary-based user points are treated as anchor points and the segmentation path has to go through them. This kind of hard constraint is not always desirable. It does not allow imprecise user input, and it can lead to difficulties in combining region-based and boundary-based approaches as discrepancy between different object descriptions is generally expected. Notably, in [9] the authors introduced soft constraint user point by embedding the user constraint in distance functions. The segmentation result is considered to be the shortest path to loosely connect the user points. However, it is known to be a NP-hard problem. Hence, it is assumed that the user points are placed in a sequential order and such a constraint reduces the computational complexity to polynomial time. This user input constraint can be seen to be generally acceptable as it is intuitive to follow the outline of an object rather than skipping around. In this work, we follow this approach to treat boundary-based user points. However, we also allow the user to place region-based strokes. These strokes are used to model foreground probability, and the discontinuity in foreground probability indicates the presence of object boundary. We combine these two types of user input with image features in an energy functional that is then optimized using graph partitioning through finding the shortest path from the first to the last user points. Moreover,
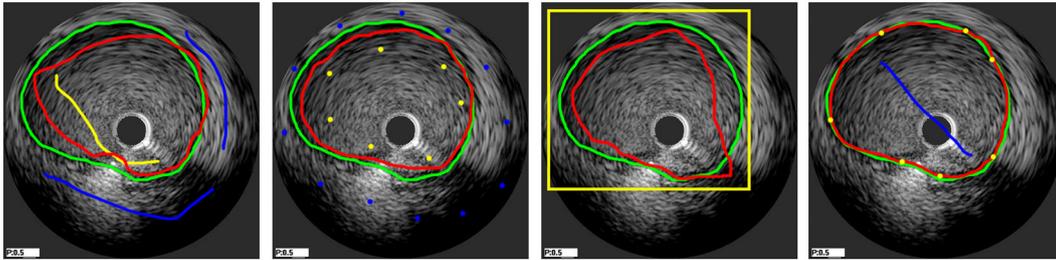
Figure 3. Examples of different segmentation methods on an IVUS image. From left: graph cut [4], seeded star graph cut [5], GrabCut [6], and proposed method. Red curve shows the segmentation result; blue or yellow for user points, background/foreground selections, and the initial window of the GrabCut. Green shows the ground truth from manually labelled image.

we apply a superpixel segmentation in order to generate a much coarser, but irregular, multilayer graph so that the computational cost is drastically reduced. It also provides a regional support at a low level for the shortest path search in the graph.

2.2. Superpixel segmentation

Efficient search for the shortest path, for instance, using Dijkstra's algorithm on a multidimensional graph is not a trivial task. Many researchers attempted to speed up the Dijkstra's algorithm by, for example, using multilevel scaling [21] or restricting the search space [22] by deciding whether or not the edge will be considered during the searching process. One mechanism that can be used to minimize the amount of nodes on the graph is superpixel segmentation. Superpixel segmentation is a process that groups a set of homogeneous neighbouring pixels together to reduce the complexity of solving further image processing such as segmentation [1, 10] and object localization [23]. Superpixel segmentation algorithms vary from graph-based [24, 25] to gradient descent methods [26, 27]. In this paper, we speed up the Dijkstra's algorithm by using the mean shift method to over-segment the image and thus create the superpixels. This over-segmented image is then used to create the graph, by only considering the boundary of the superpixel regions as a potential path that can be used to find the shortest path between two points; the whole process is thus far more efficient. Additionally, this superpixel segmentation provides low-level regional information to the graph search which relies significantly on edge information.

Mean shift algorithm [26] is a non-parametric gradient descent method that iteratively shifts the mean of the region toward the local maxima of the density for a given set of samples. Mean shifting method is suitable for clustering data without any assumption of the cluster shape. It has been widely used in many applications, such as segmentation [26] and tracking [28]. Given n data points (pixels) of x_i in the d-dimensional space R^d, the non-parametric probability function is defined by kernel density estimator as the following:

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \tag{1}$$

where h is the bandwidth parameter and K is the radially symmetric kernel such as Gaussian kernel $K(x) = (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\|x\|^2\right)$. The local maxima of density is located among the zeros of the gradient $\|\nabla f(x)\| \cong 0$. So the mean shift can be derived as the following:

$$m_{h,G(x)}(x) = \frac{\sum_{i=1}^{n} x_i G\left(\left\|\frac{x - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} G\left(\left\|\frac{x - x_i}{h}\right\|^2\right)} - x, \tag{2}$$

where $G(x) = -K'(x)$ and $m_{h,G(x)}(x)$ is the difference between the weighted mean, using kernel G, and x, the centre of the kernel. The mean shift vector points toward the maximum increase of the density, and it converges at a nearby point where the density estimate has zero gradient.
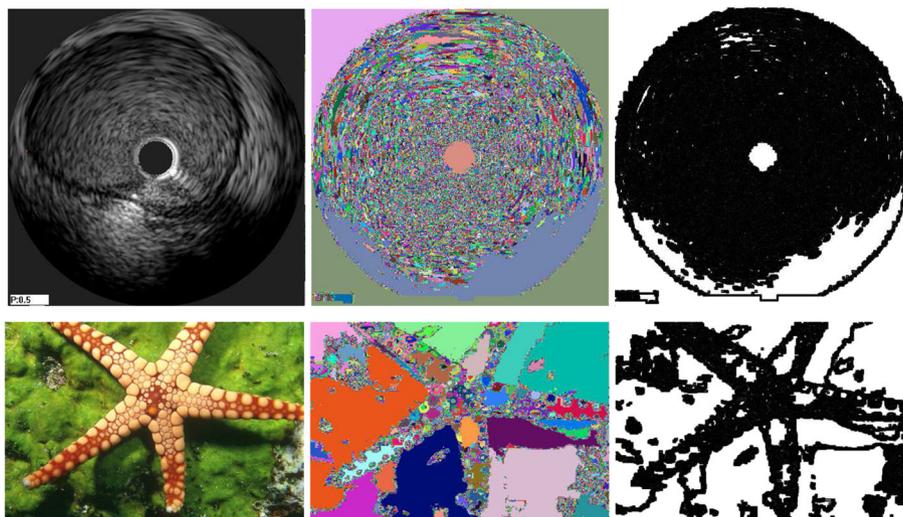
Figure 4. The use of superpixel segmentation to identify graph nodes in a single layer. Top row shows IVUS example, bottom row natural image. From left to right: initial image, superpixel segmentation, and binary representation of graph nodes.

Figure 4 provides an example of the mean shift segmentation. Mean shift is preserving the edge features in the image. The black region, shown in the rightmost column of the figure, represents areas on or close to edges in the superpixel segmentation and are used to construct the graph as it is discussed in the next section. In order to prevent the segmentation being too jagged in IVUS images (which are inherently more noisy than the real world examples used), it was necessary to increase the amount of segmentation obtained in this stage (as can be seen when comparing the middle column for the IVUS image and generic image).

2.3. Layered graph construction

In order to impose soft constraint for user point, we follow the approach proposed in [9] to construct a layered graph so that given a set of attraction points we fit a curve to follow the edges in the image and pass through the vicinity of the given points. The user points are assumed to be placed in a sequential order, which is acceptable in most applications. The computational complexity, however, is reduced from being NP-hard to polynomial time.

For each user point, $X_i, i \in \{1, 2, \ldots, k\}$, we create a new layer of directed graph. This is a copy of the image layer, with the same edge-based weighting. In that way we have a series of layers equal to the number of user points, n, plus an additional layer in order for the weighting of the last user point to be used, as shown in Figure 5. This results in a multilayer directed graph, $G = (V, E)$, where V is the set of vertices and E the set of weighted edges. For each pixel p, there exits an edge e to each of its neighbouring pixels on the same layer, providing that they are on the boundaries of the superpixels. Therefore, a pair of neighbouring pixels $(p, q) \in V$ with a corresponding edge $e = (v_p, v_q)$ also have an edge to the corresponding point on the superseding layer $e = (v_{p_i}, v_{p_{i+1}})$, where i represents the current layer of the image. For each edge, we assign a weight w to build a weighted graph (V, E). These weights are calculated on the basis of whether the edge is internal to a layer (w_i) or trans-layer (w_x). By creating the graph in this way, an order is established with the user points yet allowing for a global minimum to be found rather than a series of pairwise local minima. Edges of zero weight are added from the start node s to each pixel in the first layer and from the last layer $k + 1$ to the terminal node t. This has the effect of making the first and the last user points elastic and not hard constraints. For example, if the first user point X_1 is not located on an edge, then in the overall minimization it would be of lower cost to enter the first layer at the nearest point to X_1 on a strong edge. In this way, all user points act as soft constraints.

If P is the set of pixels in the image, P_s is therefore the subset of pixels that also fall on the boundaries of our super pixels, and p_i and q_i are pixels in layer i giving v_{p_i} as the vertex p in layer
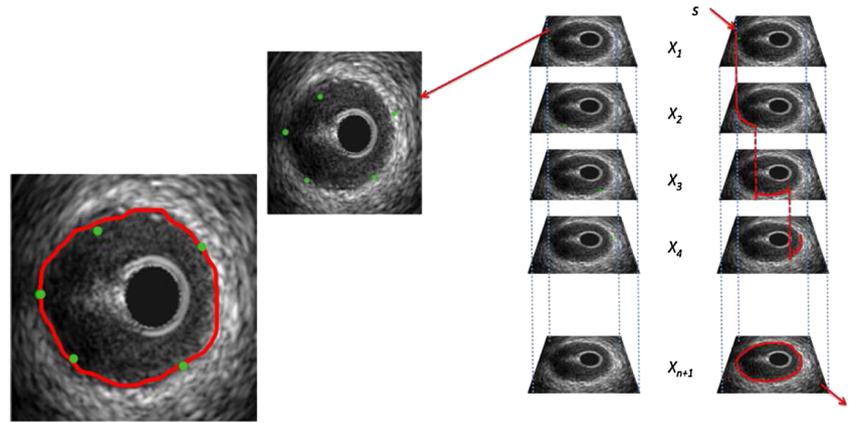
Figure 5. Example of 3D graph traversal. The stack of images on the right show how the graph is constructed out of a number of layers corresponding to the number of user points $n + 1$. The algorithm finds the shortest path through the layers, minimizing the costs of edges, boundaries, and distance from user points. The final result of the segmentation is shown on the left.

i; we can define the set of nodes V as

$$V = \{s, t\} \cup \{p_i \in P_s \wedge 1 \leq i \leq k + 1\} \tag{3}$$

and thus the set of edges as

$$E = \begin{cases} (s, v_{p_1}) | p \in P_s & \cup \\ (v_{p_{k+1}}, t) | p \in P_s & \cup \\ (v_{p_i}, v_{q_i}) | (p, q) \in N \wedge 1 \leq i \leq k + 1 & \cup \\ (v_{p_i}, v_{p_{i+1}}) | p \in P_s \wedge 1 \leq i \leq k + 1. \end{cases} \tag{4}$$

The segmentation is thus to find the shortest path from the start point s to the end point t; see Figure 5 across the 3D graph.

The edges on the directed layered graph are categorized as internal edges w_i within individual layers and interlayer edges w_x. The weighting for these two types of edges is assigned differently. The internal edges are assigned with two types of weights, that is, boundary-based edge weights and region-based edge weights. The boundary-based edge weights are calculated on the basis of the magnitude of image gradients, that is, using an edge detection function $g_e = 1/(1 + \nabla I)$ where I denotes the image or its smoothed version using, for instance, Gaussian filtering. Hence, for any given edge between neighbouring pixels (v_p, v_q), we assign a weight (w_e) according to

$$w_e((v_p, v_q)) := \frac{1}{2} ||p - q|| (g_e(p) + g_e(q)). \tag{5}$$

The region-based edge weights are computed from foreground probabilities. The user strokes placed in the foreground provide an estimation for foreground intensity distribution, which is then used to evaluate each pixel in the image. The discontinuity in this generated probability map is then used to compute the region-based edge weight in a similar fashion to image intensity, that is,

$$w_f((v_p, v_q)) := \frac{1}{2} ||p - q|| (g_f(p) + g_f(q)), \tag{6}$$

where g_f is the edge detection function based on probability values.

By combining in a weight derived from this discontinuity map, we add an extra level of robustness to the method. As regions not selected as foreground will have an increased cost, if there exist strong edges that may cause errors in our segmentation but are not selected as foreground, then the likelihood of the segmentation following these edges is greatly reduced.

The internal edge weight is thus the linear combination of the boundary-based weight and region-based weight:

$$w_i = w_e + w_f. \tag{7}$$

The attraction force imposed by user points is materialized through the interlayer edge weights w_x. We apply distance transform to the user points in each layer of the graph, and the interlayer edge weight is assigned as $w_x = d(v_{p_i}, v_{p_j})$ where d denotes the distance transform function. This distance weighting produces isolinear bands of weight around the user point, with increasing weight to go through to the next layer as the distance from the user point increases.

2.4. Energy minimization

The energy function for any curve C in our method is a combination of three terms, that is, for any arc C between two points p_i and q_j where the points are

$$C(p, q) = w_x + w_i, \tag{8}$$

as w_x can be written as

$$\alpha \sum_{i=1}^{k} ||C(s_i) - X_i||, \tag{9}$$

and likewise w_i can be written as

$$\beta \int_0^{L(C)} g(C(s))ds + \int_0^{L(C)} g_f(C(s))ds. \tag{10}$$

This is all providing that the points are treated as being in a sequential order and that the interconnections between layers are unidirectional. The overall energy function can then be expressed as

$$\begin{aligned}
\mathcal{E}(C, s_1, \ldots, s_k) = &\ \alpha \sum_{i=1}^{k} ||C(s_i) - X_i|| \\
&+ \beta \int_0^{L(C)} g(C(s))ds \\
&+ \int_0^{L(C)} g_f(C(s))ds, \\
&\ s.t. s_i < s_j, \forall i < j,
\end{aligned} \tag{11}$$

where α and β are real constants used to weigh the effects of the edge-based and distance-based terms. Figure 6 shows how changing the ratio of these constants can be used to put more or less
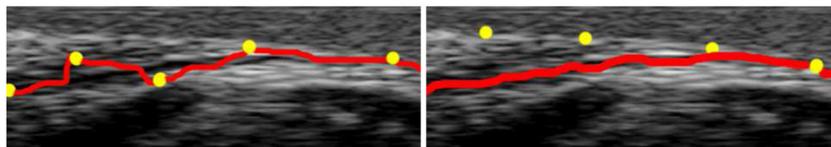

Figure 6. Effect of changing α and β. The Image on the left shows the ratio of these two constants skewed toward adding more emphasis on the user points, thus reducing the 'elastic' properties. The image on the right shows the other extreme, with the user points being bypassed in favour of the stronger edge.

emphasis on the control points (even going so far as to remove the elastic property when the ratio is skewed largely in favour of α .

The first term is used to enforce the soft constraint by the user points, and it penalizes the paths further away from the user points. The second term is the boundary-based data term that prefers the path passing through strong edges, while the last term is the region-based data term that prefers path travelling through abrupt changes in foreground probability. By using the layered graph construction, the minimization of the energy functional is achieved by finding the shortest path from the start point s to the end point t. The Dijkstra's algorithm is used to calculate the shortest path in the layered directed graph. Note that the interlayer edges are unidirectional so that the path cannot travel back to previously visited layers.

The Dijkstra's algorithm works on a directed graph $G(V, E)$ to find the shortest path between two defined nodes; the algorithm divides the nodes of the graph into two sets: visited and unvisited nodes. Once the node is marked as visited node, it will not be checked again. The algorithm starts searching from the starting node s, assigns an initial tentative distance of zero to the starting node and infinity to all other nodes, and then calculates the tentative distances for all neighbouring nodes; these tentative distances are defined as the summation of the edge weight w_i and the current distance of the beginning node of that edge. The edge weight must be a non-negative value. The algorithm will mark the node that has the minimum distance as a visited node. The algorithm will repeat the process by calculating the tentative distance for all neighbouring nodes for all visited nodes and only mark the node having the minimum distance as a visited node until the terminal node t is reached. The running time of Dijkstra's algorithm is $O(|E| + |V|log|V|)$, where E is the number of edges and V is the number of nodes.

3. EXPERIMENTAL RESULTS

To show the effectiveness of the proposed method, we test our method for two different medical applications. The two medical applications use a medical data set of IVUS images that has ground truth available for quantitative comparisons and a set of OCT images. The segmentation of images in the medical application is very challenging as it generally requires anatomical prior knowledge, as well as other expert knowledge in some cases, in order to sufficiently perform the segmentation task. This makes interactive segmentation the preferred approach to this application. Wherever appropriate, we present comparisons with other interactive segmentation techniques. We also show a more generic set of images to illustrate the versatility of our method and to highlight how it can be used for other (medical) applications.

3.1. Intravascular ultrasound image segmentation

To study the efficiency and efficacy of the proposed method, we apply our method to a medical image segmentation problem where expert prior knowledge in anatomy is necessary but also often subjective. Here we need to interactively identify the media–adventitia border in IVUS images where imaging artefacts are commonplace. Figure 12 provides several examples of IVUS images (the ground truth of media–adventitia layer can be seen in green).

There have been many different approaches to the problem of segmenting IVUS images, for example [16, 19, 29–32]. These can be broadly categorized into fully automatic methods or methods that allow user interactions. In [29], the authors used contour detection and tracing with smoothness constraint and circular dynamic programming optimization to segment the lumen border. It assumes homogeneity of the lumen region and high contrast between lumen and artery wall. Katouzian *et al.* [30] applied complex brushlet transform and constructed magnitudes-phase histograms of coefficients that contain distinct peaks corresponding to lumen and non-lumen regions. The lumen region is then segmented on the basis of K-means classification and a parametric deformable model. Homogeneity of the lumen region is critical to the success of the method. Methods based on region growing, for example [32], also suffer from such limitations, as artefacts and irregularities are very common in IVUS images. Particularly for media–adventitia border, the region inside the border is non-uniform as seen in Figure 7. Calcification in arterial wall leads to acoustic shadowing and high

reflectance, as well as catheter and guild wire occlusion and artefacts. A stent placed against the internal wall also produces strong features and acoustic shadows that break homogeneity. Incorporating user prior knowledge into segmentation hence is often necessary and has been shown to be an effective approach. For instance, Essa *et al.* [19] incorporated a shape prior to graph-cut construction to regularize segmentation of media–adventitia border. However, these approaches generally require significant amount of training data, and model re-training is often necessary in order to adapt to new data sets. User initialization is an alternative approach to transfer expert knowledge into segmentation, for example [3, 5–9, 33]. However, most user interactions are limited to either boundary-based landmark placement or strokes indicating foreground and background regions. We adopt the approach of combining two different types of user interactions, that is, boundary-based and region-based, to segment media–adventitia border in IVUS. The user points are treated as soft constraint instead of hard constraint in most interactive segmentation methods. We show that this soft user constraint allows effective combination of boundary and region-based features. The method is evaluated on an IVUS data set with manually labelled ground truth and compared against state-of-the-art techniques. Figure 8 illustrates the benefit of using user interaction to effectively influence segmentation result. Through simple user input, the expert knowledge of the user can be put into place while being augmented by the automatic process.

3.2. Optical coherence tomography image segmentation

We also compared our proposed method with other methods in OCT images (shown in Figure 9). OCT is another catheter-based modality used in cardiology. In these images, we are segmenting the luminal border rather than the media–adventitia border that was targeted in IVUS. We chose this as
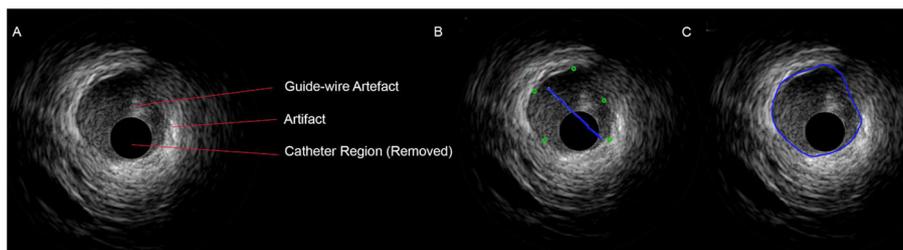

Figure 7. Overview of an IVUS image and segmentation by the proposed method. (a) Original IVUS image. (b) User input. (c) Segmentation result.
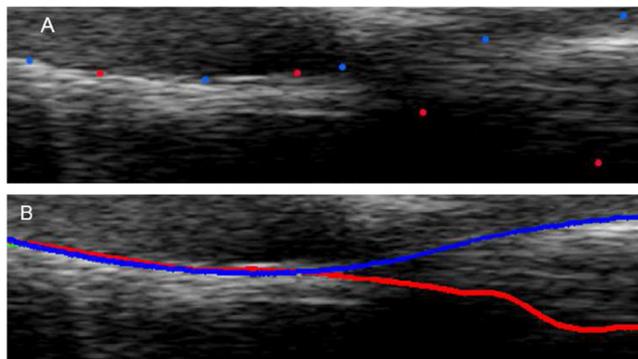

Figure 8. Effectiveness in imposing user prior knowledge. As in the natural image segmentation, the user can select different edges by the use of user points. In this case, it can be seen that there are two possibilities for the media–adventitia border in the image. By placing a couple of points, the user can steer the segmentation along the path they desire.
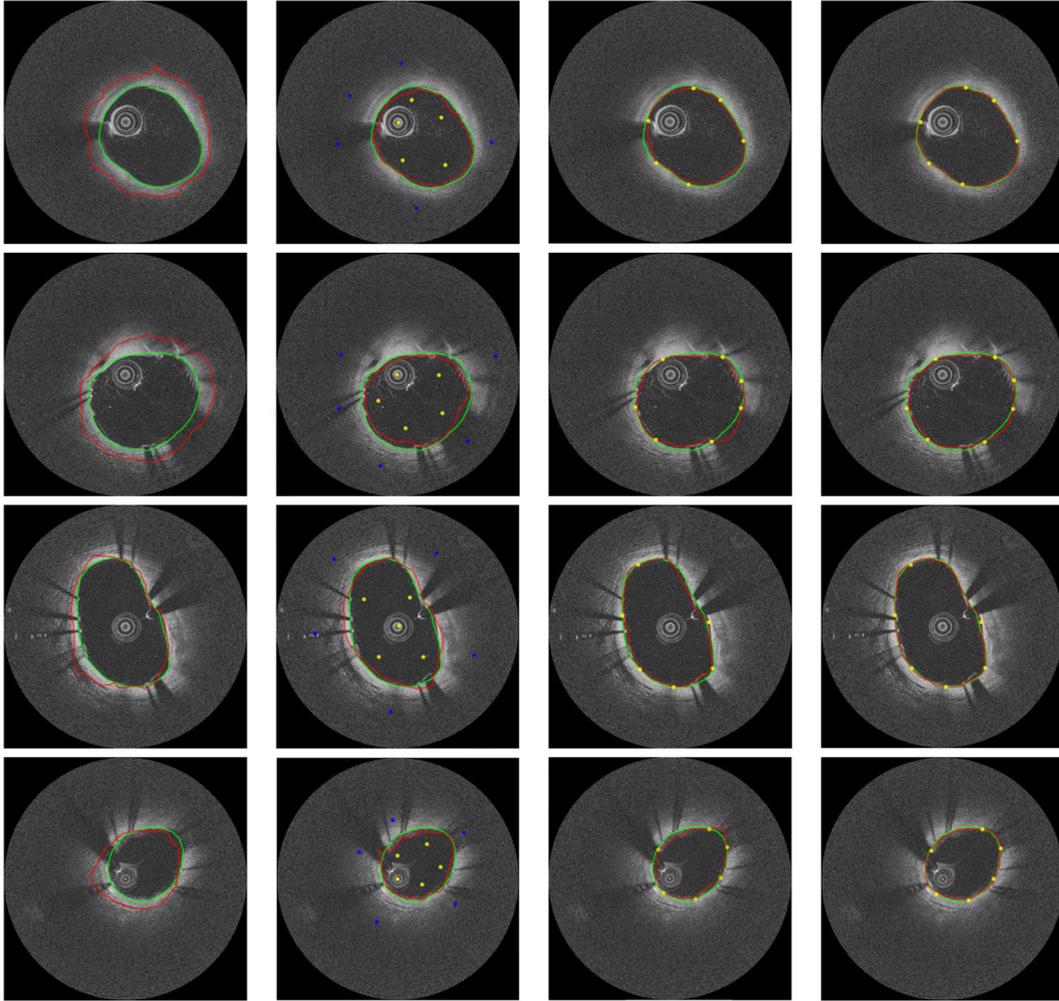
Figure 9. Comparison between ground truth (green) and (from left to right) star graph cut, seeded star graph cut, single method (no regional constraints), and proposed method (red) for OCT images showing segmentation of the lumen border.
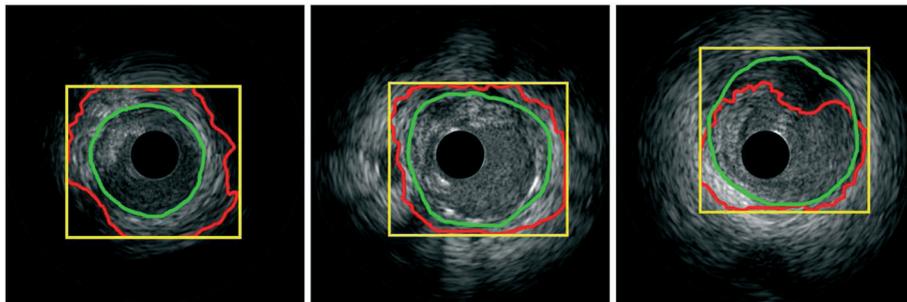

Figure 10. Typical GrabCut segmentation results (red) on IVUS images. Ground truth is shown in green. It can be clearly seen that the obtained results are considerably out from the ground truth.

OCT has far lower penetrance than IVUS images but yields a much higher resolution view of the luminal border. This makes it favourable for assessing stent placement and other surface lesions.

Table I. Quantitative comparison of the IVUS data set.

Method		HD	AOR	Spec	Sens	Acc
Star graph cut	Mean	60.57	81.22	89.22	89.86	89.29
	STD	15.64	1.00	12.00	1.00	6.52
Star graph cut with F/B labelling	Mean	43.81	86.05	90.39	**93.99**	92.17
	STD	23.89	9.00	9.00	5.00	5.65
Single method without F/B labelling	Mean	46.28	69.34	84.92	89.43	88.12
	STD	9.73	9.24	5.82	10.53	8.76
Proposed method with F/B labelling	Mean	**33.57**	**89.93**	**94.21**	93.14	**94.41**
	STD	5.35	9.16	3.88	5.37	7.67

Bold font indicates best performance.
HD, Hausdorff distance (pixels); AOR, area overlap ratio (%); Spec, specificity (%); Sens, sensitivity (%); Acc, accuracy (%); F/B, foreground and background.

Table II. Quantitative comparison of the OCT data set.

Method		HD	AOR	Spec	Sens	Acc
Star graph cut	Mean	21.55	91.75	98.09	96.73	97.41
	STD	16.07	5.36	1.61	3.99	2.80
Single method without F/B labelling	mean	21.47	91.66	98.01	97.11	97.56
	STD	14.94	8.21	2.32	4.25	3.28
Proposed method with F/B labelling	Mean	**20.97**	**92.31**	**98.66**	**97.92**	**98.29**
	STD	15.84	4.99	1.76	3.60	2.68

Bold font indicates best performance.
HD, Hausdorff distance (pixels); AOR, area overlap ratio (%); Spec, specificity (%); Sens, sensitivity (%); Acc, accuracy (%); F/B, foreground and background.
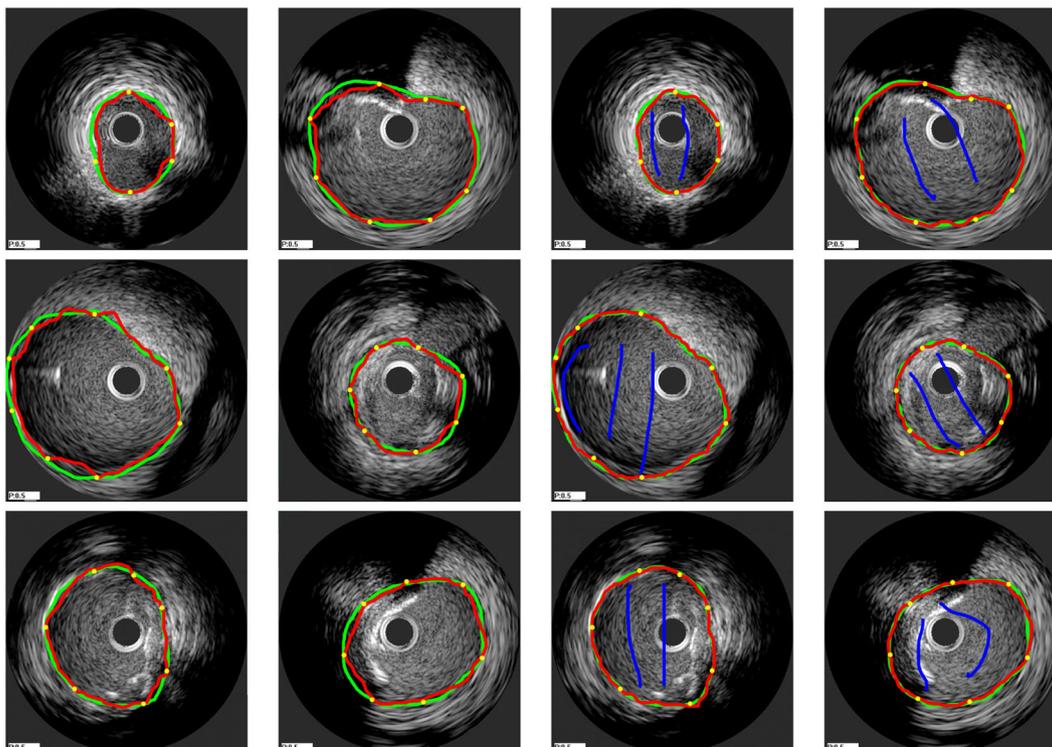

Figure 11. Comparison between ground truth (green) and single method [9] versus our proposed method (red). Images on the left two columns show the single method, and the right two columns our proposed method. Note that the user control points used remain the same for both the methods.
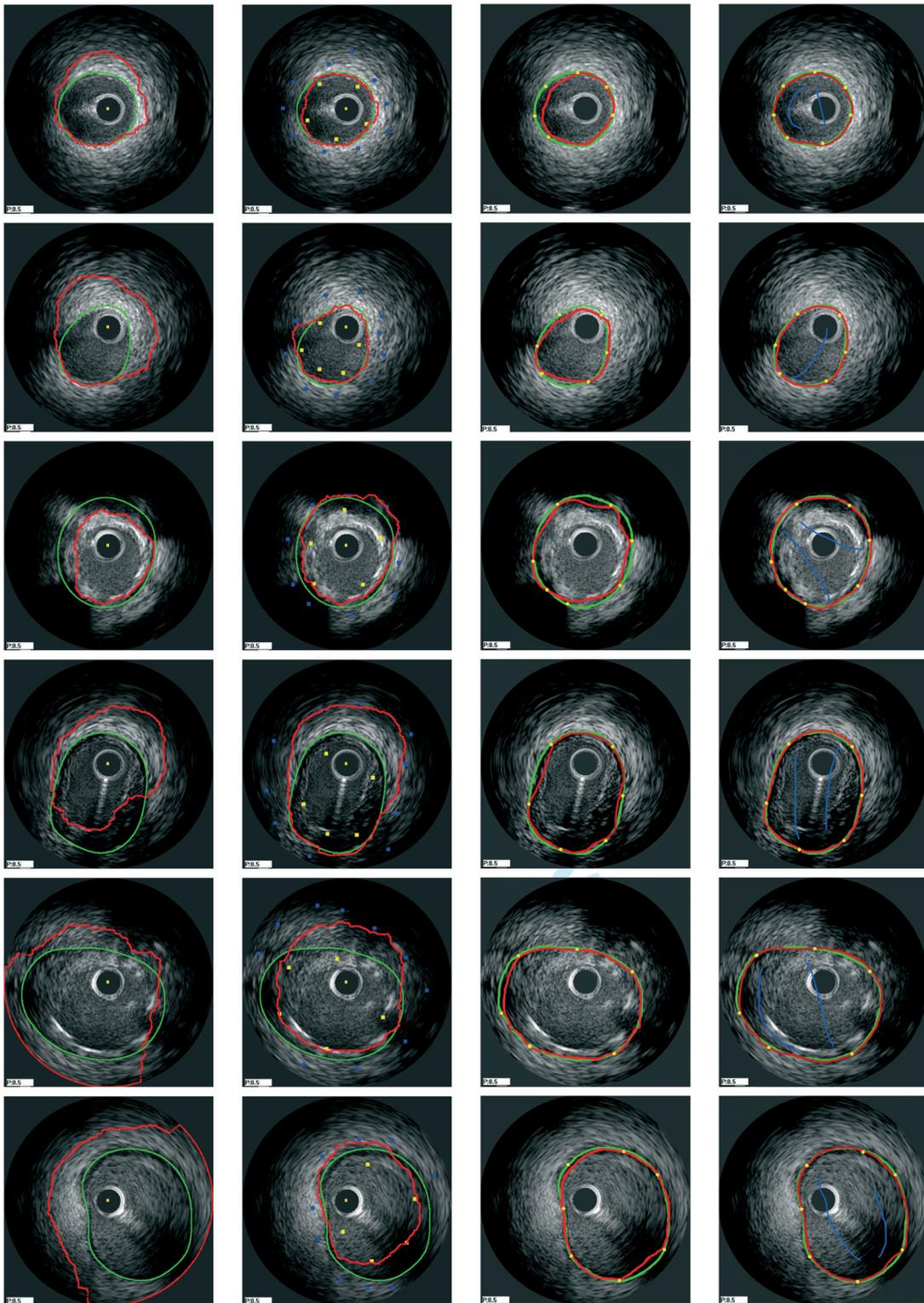
Figure 12. Comparison between ground truth (green) and (from left to right) star graph cut [5], seeded star graph cut [5], single method [9], and proposed method (red). Note that the user control points used remain the same for both the single and proposed methods.
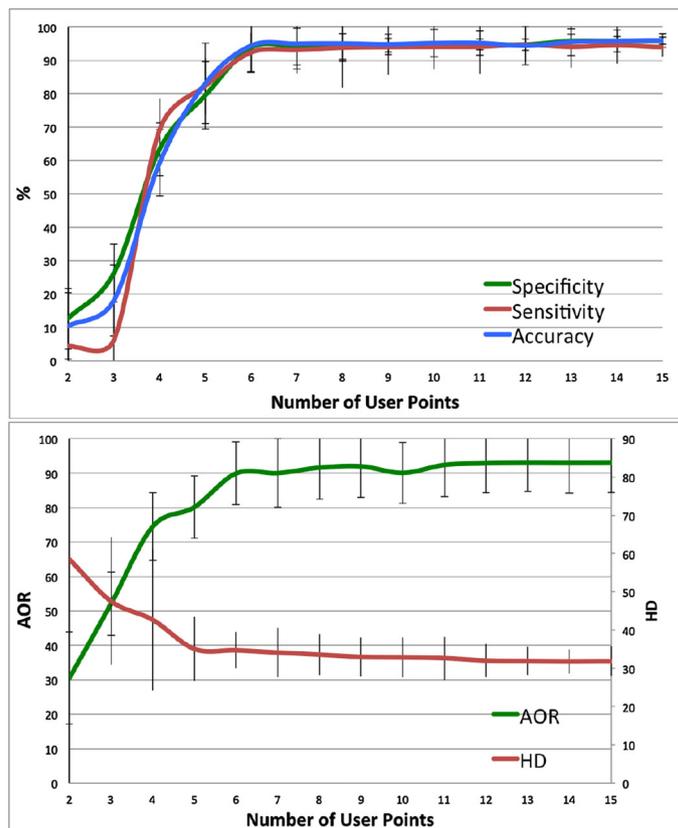
Figure 13. Initialization dependency test. This test was carried out using the IVUS data. The number of user points placed on the image was increased, and the effect of this on the observed metrics was recorded. It can be seen that only after a very few number of user points have been added, the accuracy reaches a plateau, which is an indication of good automation.

3.3. Quantitive results

In order to evaluate the results, we used a set of 248 IVUS images with ground truth. They were taken from pull-backs on seven different patients. The ground-truth labelling was obtained through manual labelling of the border of interest. These were then segmented using the proposed method, which was compared with the method using only the edge detection and not the background/foreground weighting (single method) [9] and star graph cut (both with single and multiple seed points) [5]. GrabCut segmentation was not used for the quantitative analysis, because as can be seen from Figure 10 it performs very poorly in this application. The quantitive analysis was also carried out on the results obtained from the OCT data. In this case, we were segmenting the luminal edge, which is quite well defined and regular, so all methods saw an improvement in accuracy. The results were obtained from a similar number of images as before (280) from seven different pull-backs.

The quantitative comparison is based on a number of metrics, including Hausdorff distance, area overlap ratio, specificity, sensitivity, and accuracy. Table I shows the quantitative results obtained from the IVUS data, and Table II shows the OCT data. The star graph-cut method performed reasonably well with both foreground and background labelling. The implicit shape prior in star graph construction proved to be a useful constraint in segmenting media–adventitia border that conforms well to this shape constraint. Comparable performance was achieved for the proposed method without regional support example results shown in Figure 11. However, the full proposed method outperformed the rest. Several typical segmentation results are shown in Figure 12.

To study the robustness of the proposed method, we carried out an initialization dependency test. We tested our method with 15 user points as initialization. We then randomly remove one user point each time for testing until we only have two points left for initialization. The overall results using

five different metrics are shown in Figure 13. The proposed method achieved good performance with just six user points. Considering in actual application where user input is far more experienced than this random process, even less points may be needed.

3.4. Generic image segmentation

In order to demonstrate the versatility of the proposed method, we carried out segmentation of generic real world images. We believe that our method could equally be used in other types of medical image analysis, other than the two examples we provided. By showing the results of segmentation of a series of real world images, we can further show the versatility and robustness of this method.

The proposed method was evaluated using the Berkely Image Database [34]. This data set contains images of various types. The methods were used to perform a selection/segmentation based on features in the image that would be a realistic segmentation to be carried out (for example, object selection, horizon selection etc.) The results from the proposed method were then compared with other available methods, namely $s - t$ graph cut [4], seeded star graph cut [5], GrabCut [6], and layered graph search with only point-based interaction [9]. A selection of open and closed curve segmentations were used to demonstrate and compare the results.

The proposed method showed a very favourable segmentation performance compared with the methods we tested it against. The combination of a background/foreground separation, combined
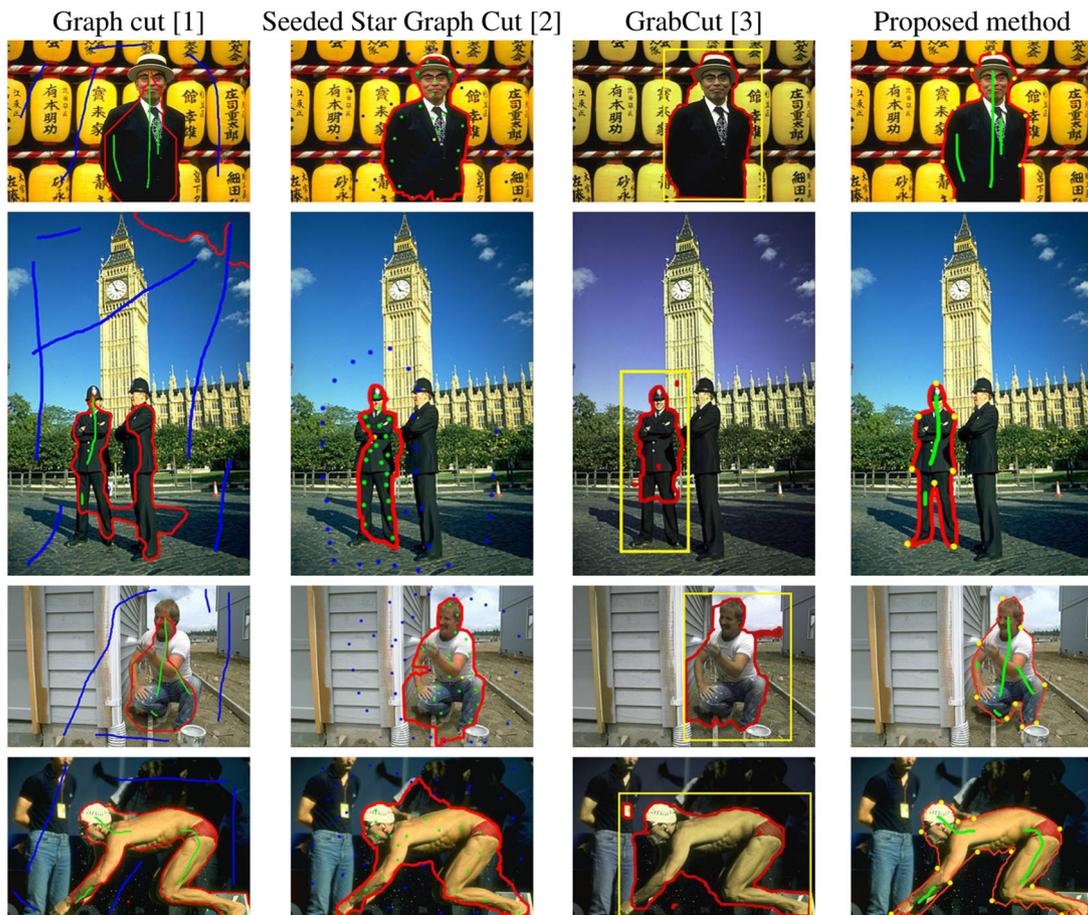

Figure 14. Segmenting humans from complex scenes. From left: original image, graph cut [4], seeded star graph cut [5], GrabCut [6], and proposed method. Red curve shows the segmentation result, blue for the background strokes, green for foreground strokes, and yellow for star point and the initial window of the GrabCut.

Figure 15. Comparison between single method [9] and the proposed combined approach.

with the edge-based approach, gave the method a very robust segmentation, being able to segment an object of interest from an image that other (single methodology based) techniques found difficult to handle, for example if colours were closely related to the background or if there were many conflicting edges. The advantage in being able to perform with open or closed curves is shown in some of the images, for example dividing the image on the horizon or segmenting figures that extend to the edge of the image. Figure 14 provides several comparative results in segmenting images of humans from complex backgrounds.

Figure 15 demonstrates the advantages of the inclusion of regional data into the algorithm, in comparison with [9]. In some cases, where there are edges other than those required, without the regional selection, the segmentation can lose accuracy. By selecting the region specificity, in most cases these edges can be ignored without the requirement for more user points, which would increase the complexity of the graph.

4. CONCLUSION

We presented an interactive segmentation technique that combines boundary-based and region-based object representations. We used this method to segment two sets of medical images (IVUS and OCT). We adopted a layered graph representation to simplify computation and a superpixel method to improve segmentation speed and efficiency. The proposed method was compared against a number of recent methods and the standard graph-cut techniques, showing improved versatility and better results. Where other methods had difficulty with certain image types, the combined approach was able to segment the desired information. We also show our method being used to segment generic real world images. This we believe shows how versatile the method is and how easily it could be used for other (medical) applications.

REFERENCES

1. Reese LJ. Intelligent paint: region-based interactive image segmentation. *Masters Thesis*, Department of Computer Science, Brigham Young University, 1999.
2. Mortensen EN, Barrett WA. Interactive segmentation with Intelligent Scissors. *Graphical Models and Image Processing* 1998; **60**(5):349–384.
3. Falco AX, Udupa JK, Samarasekera S, Sharma S, Hirsch BE, Lotufo, R de A. User-steered image segmentation paradigms: live wire and live lane. *Graphical Models and Image Processing* 1998; **60**(4):233 –260.
4. Boykov YY, Jolly M-P. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. *ICCV*, Vol. 1, Vancouver, Canada, 2001; 105–112.
5. Veksler O. Star shape prior for graph-cut image segmentation. *ECCV*, Marseille, France, 2008; 454–467.
6. Blake A, Rother C, Kolmogorov V. GrabCut: interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics* 2004; **23**(3):309–314.
7. Sinop AK, Grady L. A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. *ICCV*, Rio de Janeiro, Brazil, 2007; 1–8.
8. Cohen LD, Kimmel R. Global minimum for active contour models: a minimal path approach. *IJCV* 1997; **24**(1): 57–78.
9. Windheuser T, Schoenemann T, Cremers D. Beyond connecting the dots: a polynomial-time algorithm for segmentation and boundary estimation with imprecise user input. *Iccv*, 2009; 717–722.
10. Li Y, Sun J, Tang C-K, Shum H-Y. Lazy snapping. *ACM Transactions on Graphics* 2004; **23**(3):303–308.
11. Freedman D, Zhang T. Interactive graph cut based segmentation with shape priors. *CVPR*, San Diego, CA, USA, 2005; 755–762.
12. Malcolm J, Rathi Y, Tannenbaum A. Graph cut segmentation with nonlinear shape priors. *ICIP*, San Antonio, TX, USA, 2007; 365–368.
13. Vu N, Manjunath BS. Shape prior segmentation of multiple objects with graph cuts. *CVPR*, Anchorage, Al, USA, 2008; 1–8.
14. Gulshan V, Rother C, Criminisi A, Blake A, Zisserman A. Geodesic star convexity for interactive image segmentation. *CVPR*, San Francisco, CA, USA, 2010; 3129–3136.
15. Duchenne O, Audibert J-Y, Keriven R, Ponce J, Segonne F. Segmentation by transduction. *CVPR*, Anchorage, AL, USA, 2008; 1–8.
16. Katouzian A, Angelini ED, Carlier SG, Suri JS, Navab N, Laine AF. A state-of-the-art review on segmentation algorithms in intravascular ultrasound (IVUS) images. *IEEE Transactions on Information Technology in Biomedicine* 2012; **16**(5):823–834.
17. Moraes MC, Furuie SS. An automatic media–adventitia border segmentation approach for IVUS images. *Computing in Cardiology*, Belfast, United Kingdom, 2010; 389–392.
18. Atahnasiou LS, Bourantas CV, Rigas GA, Exarchos TP. Fully automated calcium detection using optical coherence tomography. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Osaka, Japan, 2013; 1430–1433.
19. Essa E, Xie X, Sazonov I, Nithiarasu P, Smith D. Shape prior model for media–adventitia border segmentation in IVUS using graph cut. *LNCS, Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging* 2013; **7766**:114–123.
20. Tung K-P, Shi W-Z, De Silva R, Edwards E. Automatical vessel wall detection in intravascular coronary OCT. *2011 IEEE International Symposium on Biomedical Imaging: From Nano To Macro*, Chicago, Il, USA, 2011; 610–613.
21. Holzer M, Schulz F, Wagner D. Engineering multilevel overlay graphs for shortest-path queries. *Journal of Experimental Algorithmics* 2009; **13**:5:2.5–5:2.26.
22. Wagner D, Willhalm T. Geometric speed-up techniques for finding shortest paths in large sparse graphs. *European Symposium on Algorithms (ESA)* 2003; **2832**:776–787.
23. Fulkerson B, Vedaldi A, Soatto S. Class segmentation and object localization with superpixel neighborhoods. *ICCV*, Kyoto, Japan, 2009; 670–677.
24. Shi J, Malik J. Normalized cuts and image segmentation. *T-PAMI* 2000; **22**(8):888–905.
25. Moore AP, Prince SJD, Warrell J. Lattice cut—constructing superpixels using layer constraints. *CVPR*, San Francisco, CA, USA, 2010; 2117–2124.
26. Comaniciu D, Meer P. Mean shift: a robust approach toward feature space analysis. *T-PAMI* 2002; **24**(5):603–619.
27. Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *T-PAMI* 1991; **13**(6):583–598.
28. Comaniciu D, Ramesh V, Meer P. Real-time tracking of non-rigid objects using mean shift. *CVPR*, Hilton Head, SC, USA, 2000; 142–149.

29. Luo Z, Wang Y, Wang W. Estimating coronary artery lumen area with optimization-based contour detection. *T-MI* 2003; **22**(4):564–566.
30. Katouzian A, Angelini ED, Sturm B, Laine AF. Automatic detection of luminal borders in IVUS images by magnitude-phase histograms of complex brushlet coefficients. *EMBC*, Buenos Aires, Argentina, 2010; 3072–3076.
31. Sonka M, Zhang X, Siebes M, Bissing MS, DeJong SC, Collins SM, McKay CR. Segmentation of intravascular ultrasound images: a knowledge-based approach. *T-MI* 1995; **14**(4):719–732.
32. Brathwaite PA, Chandran KB, McPherson DD, Dove EL. Lumen detection in human IVUS images using region-growing. *Computers in Cardiology*, Indianapolis, USA, 1996; 37–40.
33. Boykov Y, Jolly M-P. Interactive organ segmentation using graph cuts. *MICCAI*, Pittsburgh, Pennsylvania, USA, 2000; 276–286.
34. Martin D, Fowlkes C, Tal D, Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *Proceedings of the 8th International Conference of Computer Vision*, Vol. 2, Vancouver, Canada, 2001; 416–423.