Automatic Segmentation of Cross-Sectional Coronary Arterial Images

Ehab Essa[a,b], Xianghua Xie[a,*]

[a]*Department of Computer Science, Swansea University, UK*
[b]*Faculty of Computers and Information Sciences, Mansoura University, Egypt*

Abstract

We present a novel approach to segment coronary cross-sectional images acquired using catheterization imaging techniques, i.e. intra-vascular ultrasound (IVUS) and optical coherence tomography (OCT). The proposed approach combines cross-sectional segmentation with longitudinal tracking in order to tackle various forms of imaging artifacts and to achieve consistent segmentation. A node-weighted directed graph is constructed on two consecutive cross-sectional frames with embedded shape constraints within individual cross-sections or frames and between consecutive frames. The intra-frame constraints are derived from a set of training samples and are embedded in both graph construction and its cost function. The inter-frame constraints are imposed by tracking the borders of interest across multiple frames. The coronary images are transformed from Cartesian coordinates to polar coordinates. Graph partition can then be formulated as searching an optimal interface in the node-weighted directed graph without user initialization. It also allows efficient parametrization of the border using radial basis function (RBF) and thus reduces the tracking of a large number of border points to a very few RBF centers. Moreover, we carry out supervised column-wise tissue classification in order to automatically optimize the feature selection. Instead of empirically assigning weights to different feature detectors, we dynamically and automatically adapt those weighting depending on the tissue compositions in each individual column of pixels. The proposed approach is applied to IVUS and OCT images. Both qualitative and quantitative results show superior performance of the proposed method compared to a number of alternative segmentation techniques.

Keywords: IVUS, OCT, graph cut, image segmentation, shape prior.

1. Introduction

Coronary atherosclerosis is an inflammatory disorder that involves deposition of cholesterol and other fatty substances within the arterial wall. It can lead to progressive narrowing

*corresponding author
Email addresses: ehab_essa@mans.edu.eg (Ehab Essa), x.xie@swansea.ac.uk (Xianghua Xie)
URL: http://csvision.swan.ac.uk (Xianghua Xie)

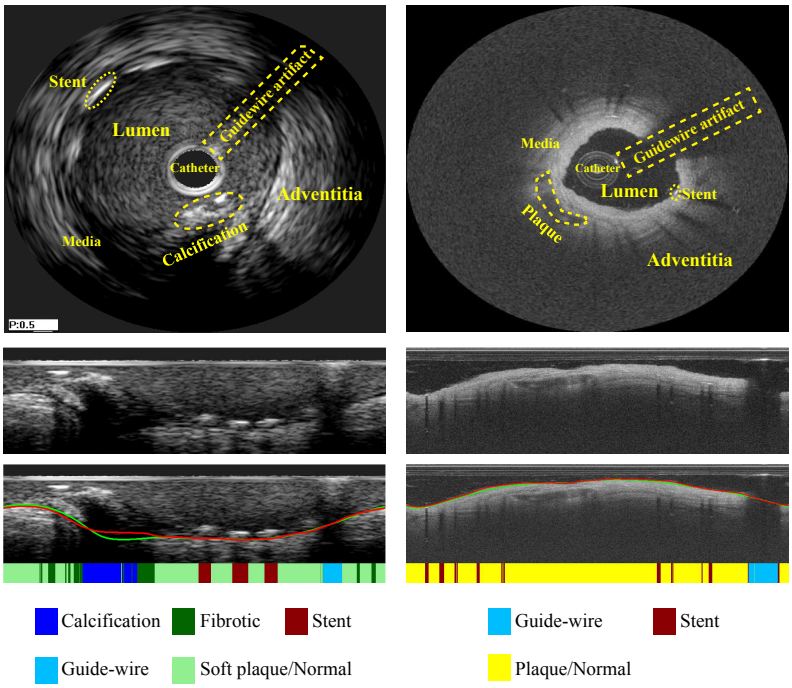

Figure 1: Example IVUS and OCT images. First row: original images. Second row: polar transformed images. Last row: segmented media-adventitia border in IVUS and lumen border in OCT using the proposed method (red); the groundtruth is shown in green. The bottom of each segmented image visualizes our tissue classification results.

of coronary arteries, which can cause angina, or sudden blockage of the coronary arteries leading to acute myocardial infarction. Intra-vascular Ultrasound (IVUS) and optical coherence tomography (OCT) are catheter-based technologies, which capture 2D cross-sectional images of the coronary arteries. Both modalities measure the back-scattered signal from the surrounding vessel structure after sending sound wave in IVUS or light in OCT. These provides a much detailed visualization of lumen, stent strut location, and plaque morphology. A coronary cross-section is generally seen as a lumen and a coronary vessel wall, the latter consisting of three layers: intima, media and adventitia. There are two types of borders of clinical interest: the lumen-intima border that corresponds to the inner coronary arterial wall and the media-adventitia border that represents the outer coronary arterial wall located between the media and adventitia (see Fig. 1). This work is concerned with segmenting the media-adventitia boder in IVUS and lumen boder in OCT. The segmentation of the lumen border in OCT allows, for instance, quantitative analysis of vessel narrowing and its impact on blood supply to myocardium, and the localization of the media-advetitia border in IVUS provides both the exterior geometry of the coronary vessel and the region of interest for virtual histology.

Various techniques have been developed to segment IVUS images, e.g. [1, 2, 3, 4, 5, 6], [7, 8, 9, 10]. Methods that fit contours to local image gradients, e.g. [1], are susceptible to common artifacts, such as speckle noise and acoustic shadow. Moreover, user initialization

is often necessary in order to converge to meaningful local minima. There are also methods that use global regional information, instead of intensity discontinuities. For instance in [2], Cardinal *et al.* assumed that each region consists of uniform scattering in tissue and has its own statistically distinctive Rayleigh distribution. However, it is expected that arteries captured in IVUS contains various forms of diseased tissue which greatly compromises their assumption. Textural information has also been used. Papadogiorgaki *et al.* [3] proposed to use discrete wavelet frames to construct decomposition trees to identify vessel wall borders. RBF is then used to smooth the initial contour obtained by applying a threshold on those texture features. However, the method suffers from the presence of stent or severe calcium plaque. Recently in [9, 10], pixel-wise classification was used to obtain initial wall borders. Su *et al.* [9] used double neural networks to segment wall borders, while Zakeri *et al.* [10] used sparse representation based classification and then refined the result using an active contour model. To deal with heterogeneous texture in high-frequency IVUS images remains a challenge.

Similar approaches have been applied to OCT imaging, e.g. [11, 12, 13, 14, 15, 16, 17], which has become increasingly popular due to its high resolution. Parametric active contours can cope with small gaps introduced by stent shadow as long as the landmark points are localized on the lumen border. For example, the methods proposed in [11, 13, 12] rely on active contour parametric interpolation to cope with the acoustic shadow caused by stent. However, guide-wire and blood residue can cause much larger and more irregular acoustic shadow that interpolation alone may not be sufficient. The work in [11] requires occasional user inspection and intervention based on assumptions of the regularity of the lumen border. Tung *et al.* [14] used a convex hull based approach to identify guide-wire shadow casting on the lumen border after an initial segmentation. It is assumed that most part of the lumen border in OCT is clearly visible and continuous. Thus, a large discrepancy between derived convex hull and initial segmentation indicates optical shadow. However, the lumen border may not always be convex, particularly at bifurcations. Although more sophisticated texture analysis techniques may improve the segmentation performance, e.g. [16], these data driven approaches generally suffer from the imperfection in imaging and natural variations in anatomical structure and tissue composition. Others also resorted to user interaction to eliminate the ambiguity in imaging, e.g. [15].

In an attempt to overcome the shortcomings of imaging features, anatomical and imaging priors have been used to constrain the segmentation. Sonka *et al.* [4] requires the user to draw an elliptic shape to identify the region-of-interest (ROI) and uses parametrized prior knowledge on arterial wall thickness and double echo pattern in objective function to carry out segmentation. However, these hard constraints may not be valid in some cases, e.g. media thickness. Takagi *et al.* [5] extended the work by incorporating spatio-temporal filters to reduce blood speckles in order to enhance contrast. This however does not address the issue of acoustic shadow or scattering due to stent or calcification. More recently in [6], an auxiliary border is used to tackle the distractions caused by stent and calcification and assumed the real media-adventitia border is beneath the auxiliary border in a simultaneous segmentation. The behavior of the auxiliary border can be hard to predict, particularly when there is no such distractions.

Learning and using appropriate priors are hence important. One approach is to adopt user interaction and directly impose prior knowledge through initialization and/or user adjustment, e.g. [4, 1, 2, 18, 19, 20]. Sun *et al.* [20] proposed a semi-automatic graph-based method, in which a pre-segmentation of the lumen is necessary to construct the graph. A combination of edge and region based costs are assigned to each node. The method requires the user to interactively correct the segmentation result on the longitudinal view. Incremental user input are allowed until satisfactory segmentation is achieved. The maximum inter-frame difference is set as a global constant in order to impose the smooth constraint. An alternative is to generalize priors and impose them as constraints in order to achieve automated segmentation. Unal *et al.* [21] used signed distance transform to implicitly represent prior shapes and applied principal component analysis (PCA) to generalize the shape variation. Its automated initialization of the media-adventitia border, however, is based on the maximum gradient information which is susceptible to imaging artifacts. In [22], short vertical image segments are collected along media-adventitia borders to score image segments in unseen images based on the frequency of occurrence. Those score values are then used as the basis for the data term in a cost function, minimized using graph cut. The distribution of image segments however is not generalized, and its discrete form inevitably produces undesirable oscillations in score values.

Moreover, incorporating shape prior into graph based segmentation attracted much attention, e.g. [23, 24, 25, 18]. A set of user specified landmarks are used to define a shape template as a distance function and integrated with neighboring edges in the graph to impose shape prior [23]. Veksler [18] introduced a star shape prior to graph cut, also through user interaction. The user is required to specify the center 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 necessary to discourage the inherent bias in graph construction towards smaller size segmentation. Decomposition techniques, e.g. kernel PCA, have been used to generalize shape priors. In [24], graph cut is iteratively computed with the terminal edges updated according to PCA subspace projection of segmented shape from previous iteration. However, user initialization is generally necessary, as shown in [24], in order to impose shape constraint when computing graph cut. Vu and Manjunath [25] used non-parametric density estimation to generalize the shape prior from multiple templates. The terminal edges are iteratively updated according to the similarity between current shape obtained through graph cut and training shapes. Several graph cuts have to be computed before reaching the final result.

In this paper, we introduce a fully automatic segmentation approach to delineate the outer vessel border in IVUS and lumen border in OCT without any user initialization or interaction. The segmentation of the lumen border in OCT allows, for instance, quantitative analysis of lumen narrowing, and the localization of outer vessel border (i.e. media-adventitia border) in IVUS provides both the exterior geometry of the coronary vessels and the region of interest for virtual histology. We propose not only to learn image artifacts in order to produce optimal cost function combination for image features but also to learn shape prior to impose global and local shape constraints. We adopt the optimal surface graph construction method [26] in order to avoid user initialization. However, by using optimized cost function

based on tissue classification, this initial segmentation is far more meaningful. The graph construction is then adjusted to impose local shape constraint, and a new shape prior term is incorporated into the cost function to impose global constraint. Moreover, to improve the consistency among cross sections, we integrate Kalman filtering into our segmentation process to produce coherent and consistent results. Through qualitative and quantitative comparisons, we show the proposed approach achieves superior performance compared to the state of the art.

The main contributions of this work can be summarized as:

- We propose a fully automated segmentation method for two different intravascular imaging modalities (IVUS and OCT) to segment the coronary arterial wall by transforming the segmentation problem into energy minimization that can be solved without any user interaction.
- A novel cost function formulation is proposed based on a column-wise supervised classification to dynamically formulate the cost function according to the artery tissue type or artifact.
- Shape prior model is incorporated in both the graph construction and the cost function for segmentation. A non-parametric estimation method is used based on the similarity between an initial segmentation and a set of shape templates to incorporate global shape prior. Graph construction is also data dependent by defining smoothness arcs between every two neighboring columns to allow learning-based transition of the border.
- A novel spatio-temporal segmentation using the Kalman filter is proposed. The Kalman filter is used to provide the temporal information about the border across frames. The predicated border is integrated into the system by adding a new set of arcs between every two consecutive frames.

2. Proposed method

The proposed method combines cross-sectional segmentation and longitudinal tracking. Shape priors are learned in both cross-sectional and longitudinal domains. The segmentation and tracking is formulated as a graph cut problem. The coronary images are transformed from Cartesian coordinates to polar coordinates to facilitate appropriate graph construction and border tracking. The node-weighted directed graph is constructed on two consecutive cross-sectional frames with embedded shape constraints both within individual frame and between consecutive frames. Column-wise supervised tissue classification is carried out so that weightings for different types of features in the cost functions adapt to the tissue compositions in each individual column of pixels. Combined with shape prior based cost functions, the final segmentation is obtained by finding the minimum closed set on the node-weighted graph.

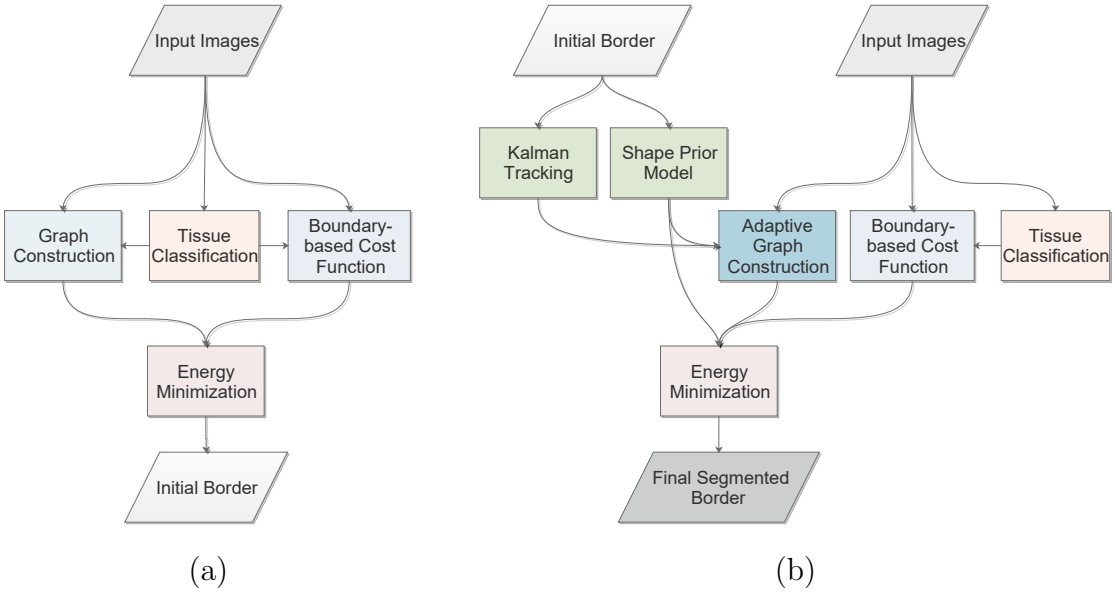

Figure 2: System overview. (a) An initial border segmentation is first computed, then refined by incorporating prior and temporal information (b).

Fig. 2 provides a schematic overview the proposed method, which consists of two stages: data driven initial segmentation as shown in (a) and prior driven refined segmentation as shown in (b). For initial segmentation, the input images are two consecutive frames, and for each frame a node-weighted directed graph is constructed with intrinsic geometric constraints as defined in Section 2.1. The tissue and artifact classification (described in Section 2.2) is a column-wise classification based on random forest and Haar-like features that are used to dynamically formulate the boundary-based cost function. The node cost is inversely proportional to the likelihood of belonging to the border of interest and is defined based on a combination of a set of edge and bar features, which is described in Sections 2.3 and 2.4. The border segmentation is achieved by minimizing the energy function that is equivalent to finding the minimum enclosure set on the graph as described in Section 2.6. Next, the generalized shape prior and the Kalman filter based tracking (Sections 2.5, and2.7) are used to regularize the segmentation, both within individual frames and between consecutive frames. The final border segmentation is obtained by minimizing an energy functional that incorporates both shape prior and temporal constraints.

2.1. Graph Construction

A typical IVUS volume contains a large number of frames which makes it computational expensive and memory demanding for batch-mode processing. Multi-resolution downsampling and processing can be adopted to alleviate the computational burden e.g. [22]. However, our graph construction only requires immediate consecutive frames, which makes it much more scalable compared to existing multi-resolution approaches. The first step is transforming the images to polar coordinates from Cartesian coordinates to facilitate our

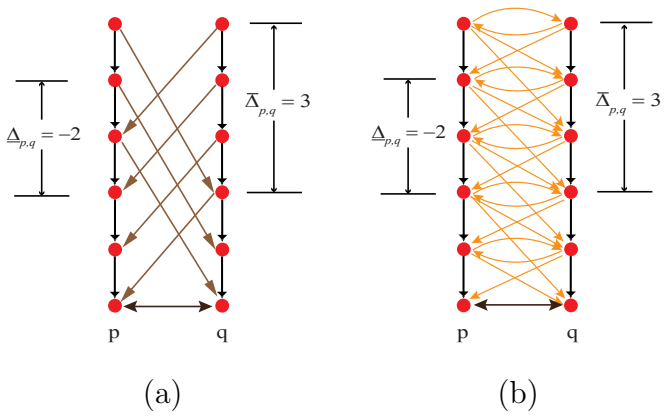

Figure 3: Example graph construction. (a) inter-columns arcs connecting two neighboring columns p and q, where $\underline{\Delta}_{p,q} = -2$ and $\bar{\Delta}_{p,q} = 3$ in this example, shown as brown arcs. (b) smoothness arcs, in orange, for the same columns when $f'_{p,q}(0) = 0$.

feature extraction and classification, as shown in Fig. 1, and the catheter region can be easily removed using thresholding as it is fixed in size and location (center of the image in Cartesian coordinates). The lumen and media-adventitia borders in the polar coordinates are shown as interfaces that intersect with each column once and once only. This allows us to define the segmentation of these borders as an optimal path in the the horizontal direction. We thus follow [26] to construct a node-weighted directed graph, whose minimum closed set provides an optimal interface that does not overlap itself at any column of pixels. We also take a similar approach to [27] to incorporate shape prior into graph construction by using a convex function defined on neighboring columns. However, different to previous approaches, smoothness constraint is imposed at the initial stage of estimating the border of interest and nearest shapes after alignment are used to generalize the shape prior. Moreover, the shape evolution from frame to frame is captured through tracking and the derived prior is embedded in the graph construction by adding an additional layer. This graph construction also eliminates the need for user initialization to specify foreground and background.

Let $G = \langle V, E \rangle$ denote the graph, where each node $V(x, y)$ corresponds to a pixel in the transformed IVUS image $I(x, y)$ in polar coordinates. The graph G consists of three arc types: intra-column, inter-column and smoothness penalty arcs, see Fig. 3. Since the images are unwrapped into the polar coordinates, the first and the last columns are connected by inter-column and smoothness arcs to enforce connectivity.

For intra-column, along each column every node $V(x, y)$, where $y > 0$, has a directed arc to the node $V(x, y-1)$ with $+\infty$ weight assigned to the arc to ensure that the desired interface intersects with each column exactly once.

In the case of inter-column, for each node $V(x, y)$ a directed arc with $+\infty$ weight is established to link with node $V(x+1, y-\bar{\Delta}_{p,q})$. Similarly, node $V(x+1, y)$ is connected to $V(x, y+\underline{\Delta}_{p,q})$, where the difference between two neighboring columns p and q of the desired border is in between the maximum distance $\bar{\Delta}_{p,q}$ and the minimum distance $\underline{\Delta}_{p,q}$ and it acts as a hard constraint. The inter-column arcs are also defined between two consecutive

frames.

The values for $\bar{\Delta}_{p,q}$ and $\underline{\Delta}_{p,q}$ in the same frame are computed from shape prior, which is defined later in Section 2.5.2, and the equivalent for between the frames are computed based on the Kalman filter, which is defined later in Section 2.7. The nodes in the last row of the graph are connected to each other with $+\infty$ weight to maintain a closed graph. Inter-column and intra-column arcs are illustrated in Fig. 3(a). Note, In order to compute the initial border without incorporating the shape prior model, the $\bar{\Delta}_{p,q}$ and $\underline{\Delta}_{p,q}$ are globally fixed for all neighboring columns within and between two consecutive frames. The delta values control the maximum and the minimum distance shift between two neighboring columns. However, this does not guarantee smooth transition as the border can be anywhere within the range. X. Wu and D. Chen [28] showed that incorporating a convex non-decreasing function in a multi-column graph can be solved in polynomial time, instead of NP hard. Ishikawa [29] has previously used a convex function to carry out MRF minimization for multi-label problems. Here, we use a convex function, defined on neighboring columns, to enforce smoothness constraints and embed shape prior in the graph construction.

2.2. Tissue and Artifact Classification

IVUS and OCT contains various forms of imaging artifacts that disrupt the continuity of lumen and media-adventitia borders. Hence it is desirable to detect those regions and treat them differently. In IVUS, calcification causes strong acoustic shadow and stents exhibit strong bright echo. In OCT, stent casts a shadow behind its metallic strut, and calcified plaque has heterogeneous bright tissue with strong borders. There are some reported attempts to detect those artifacts and structures, such as calcification, metallic stent, and plaque, in both IVUS and OCT. In [30], the authors used intensity thresholding and contrast comparison to identify calcification in IVUS. Image gradient magnitude has also been used to detect calcification, e.g. [21]. In OCT, several works were proposed to detect stent by using intensity profile [12, 15], gradient information [11], and ridge detector [31]. An active contour method was also used in [32] to find the location of calcification after automatic segmentation of lumen and guide wire based on the edge map. However, these heuristic driven approaches often rely on manually tuned parameters to make decisions. Machine learning methods have been recently applied to detect stents [33] and plaques [34]. Textural features, such as co-occurrence and local binary patterns, and geometrical features are used for supervised classification to localize those regions. To ensure satisfactory accuracy, these methods require a predetermined segmentation of the lumen and arterial borders, which itself is not a trivial task.

Instead of a usual attempt of localizing those image artifacts, which is problematic, we transform the images into polar coordinates and classify the entire columns of pixels that may contain those imaging artifacts. The detection result will then have an influence on the formulation of the cost function, i.e. dynamically selecting appropriate features for each column.

We classify individual columns of pixels in the polar coordinates into a number of categories. For IVUS, we use five categories, i.e. calcification, fibrotic plaque, stent, guide-wire artifact, and normal tissue or soft plaque, because of both their different feature appearance

and their influence on media-adventitia border. When there are multiple artifacts in the same column, our ground-truth labeling and automated classification for each column are based on the dominant artifact. For example, when there is calcified plaque along with some soft and fibrotic tissues, the columns will be labeled as calcified plaque.

For OCT, since some features such as calcification are below the lumen border, we only need to separate guide-wire artifact, stent, and plaque or normal tissue. Calcification behind the lumen wall creates weaker features in OCT, partly due to its much stronger signal attenuation. Hence, in the case for OCT lumen segmentation, we group normal tissue with soft, fibrotic and calcified plaque to form a single class.

Among others, Support Vector Machines (SVM) and Adaptive Boosting, and Random Forests (RF) are popular choices in supervised classification and have shown effective in dealing with noisy images, e.g. ultrasound [35, 36, 33]. Some are more computationally expensive than others, e.g. pruning may be required to improve SVM efficiency in classification but it increases training time [37, 38, 39]. In this work, we train a Random Forests classifier [40] to perform the task. RF is an ensemble classifier consisting of a set of unpruned decision trees, which significantly improves the generalization ability of the classifier compared to a single decision tree. RF has been shown to be more efficient than boosting with a comparable accuracy [41]. RF provide a robust performance when there is a significant overlap between classes and noise [40] and it is an effective ensemble that can identify nonlinear patterns without overfitting. Moreover, RF can be conveniently extended to solve multi-class classification without the need to train multiple one-vs-all binary classifiers as in SVM.

It is possible to directly use intensity values from each column for classification, since the goal is not to localize imaging artifacts but to determine which columns may contain them. However, particularly for IVUS, the intensity profiles from individual columns may appear similar to each other for certain classes, e.g. stent and calcification. We thus extract both horizontal and vertical Haar-like features at multiple scales. The horizontal template is defined as $[1, -2, 1]$ to take into account neighboring columns, and two vertical templates are defined as $[-1, 1]^T$ and $[1, -1, 1]^T$ to highlight edges and bar-like features. These Haar-like features are similar to those that have been used in, for instance, face recognition [42]. Since we are not dealing with localization, one dimensional templates are sufficient, which ensures efficiency.

2.3. Feature Extraction

In IVUS images, the media layer is usually thin and generally dark, and the adventitia layer tends to be brighter, see Fig. 1 as an example. Hence, edge based features are appropriate to extract the media-adventitia border. In OCT imaging, the lumen appears to be much darker because blood is flushed out before imaging. The intima and other tissues, including plaque, surrounding the lumen have a bright appearance. Hence, the lumen-intima border has a good contrast, i.e. image gradient features may be adopted to highlight the border. However, guide-wire artifact and other interfering image features commonly exist inside the artery and they cast shadows over the border of interest, disrupting its continuity. Extracting image gradient features will be affected by those imaging artifacts. In this work,

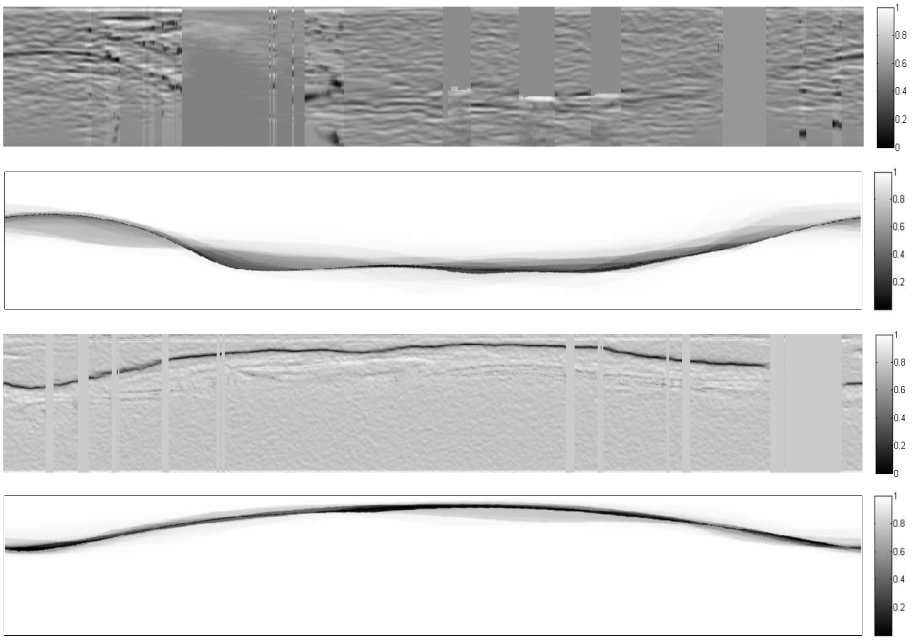

Figure 4: Visualization of the cost functions. The first row and the third row show the boundary term $\mathcal{C}_B$ for IVUS and OCT images, respectively. The second row and the fourth row show the corresponding shape prior terms $\mathcal{C}_P(x, y)$.

we propose to detect and classify those artifacts in both IVUS and OCT and treat them differently when incorporating them into the cost function. The image features are derived from a combination of steerable features and local phase features.

Steerable filters are more effective in highlighting edge features than isotropic band-pass filters, particularly when there is noise interference [43]. A steerable filter is a linear combination of different oriented instances of the base filter. Here, we use a set of first and second order derivatives of Gaussian (GD) filters $G_n(x, y)$ in six different orientations to highlight the edge features along the border, i.e.

$$G_n^{\theta}(x, y) = \sum_{j=1}^{M} k_j(\theta) G_n^{\theta_j}(x, y) \tag{1}$$

where $G_n^{\theta}(x, y)$ is the rotated version of $G_n(x, y)$ at θ orientation and $k_j(\theta), 1 \leq j \leq M$, are interpolation functions.

Local phase features have been shown to be an effective alternative to intensity derived features to deal with inhomogeneity, low image quality, and imaging shadow which are common in IVUS and OCT images. For example, in [44] local phase features were used to highlight acoustic boundaries in echocardiographic images. Local phase features are considered as extrema in Fourier phase components, which can be located as peaks in the local energy function obtained by convolving odd and even symmetric Log Gabor filters, ($o_m(x, y)$, $e_m(x, y)$), to remove DC component and preserve phase in localized frequency. Two types of features can be extracted from phase congruency: feature asymmetry and feature symmetry.

Feature symmetry favors bar-like image patterns and exists in the frequency components at either the minimum or maximum symmetric points in their cycles, which is useful in extracting the thin media layer. Its dark polarity symmetry is used here [45]:

$$\phi_s(x,y) = \max_{\theta} \frac{\sum_m \lfloor [-e_m(x,y) - |o_m(x,y)|] - T_m \rfloor}{\sum_m A_m(x,y) + \varepsilon}, \tag{2}$$

where θ and m denote filter orientation and scale, ε is a small constant, T_m is an orientation-dependent noise threshold, $A_m(x,y) = \sqrt{e_m^2(x,y) + o_m^2(x,y)}$ and $\lfloor . \rfloor$ denotes zeroing negative values. In contrast, feature asymmetry highlights step-like image patterns and corresponds to the point where all the frequency components are at the most asymmetric points in their cycles. It can be defined as:

$$\phi_a(x,y) = \max_{\theta} \frac{\sum_m \lfloor [|o_m(x,y)| - |e_m(x,y)|] - T_m \rfloor}{\sum_m A_m(x,y) + \varepsilon}. \tag{3}$$

Phase asymmetry is useful to highlight lumen border in OCT, where edge features can be seen from darker lumen layer to brighter intima layer.

2.4. Cost Function: Boundary Term

A dynamic cost function based on the classification result is used, instead of a uniform cost function. Each column of pixels is classified into a set of labels, $\mathcal{L}$, each of which has its characteristic features and thus own cost function definition. The cost function can be defined as a combination of costs $cf_l(x,y)$ that is defined based on label $l \in \mathcal{L}$:

$$\mathcal{C}_B = \bigcup_{l \in \mathcal{L}} \bigcup_{(x,y) \in \mathcal{P}_l} cf_l(x,y), \tag{4}$$

where $\mathcal{P}_l$ denotes a set of columns in polar coordinates (x,y) that are classified as label l. We adopt a boundary based energy functional to form the basis of the classification-based cost function, and can be generically formulated as:

$$E_B = \sum_{V \in S} \hat{\mathcal{C}}_B(x,y), \tag{5}$$

where S is a path in the directed graph and $\hat{}$ denotes unit normalization.

2.4.1. IVUS cost function

For normal tissue and soft plaque, the media layer has a good contrast to adventitia. Hence, their cost is defined as

$$cf(x,y) = G_1(x,y), \tag{6}$$

where G_1 is summation of filtering response of the first order GD from six different orientations, i.e. G_1 measures total edge strength. The first order GD filters are designed so that the stronger the border the lower the filtering response.

Calcified plaque exhibits strong edge feature and casts varying degrees of acoustic shadow on media-adventitia border. We use normalized intensity directly as feature, which means that all the shadowing area underneath the calcified plaque have the same cost values. The calcified regions are then penalized, and meanwhile the smoothness constraint in graph construction and shape prior have much greater influence in determining the optimal location of the border.

Fibrous tissue exhibits similarity to calcification, except in majority cases the media layer is still discernible. Hence, combining the phase symmetry feature with the edge feature detection is appropriate to enhance the border, i.e.

$$cf(x, y) = G_1(x, y) - \phi_s(x, y), \tag{7}$$

where ϕ_s is the local phase symmetry feature. Note that G_1 and ϕ_s are generally not collocated when detecting media-adventitia border, i.e. ϕ_s focuses on middle of the media layer which is slightly above the media-adventitia border. To address this discrepancy, we accumulate the ϕ_s response in a short line segment immediately above current pixel so that ϕ_s response is shifted downwards by a few pixels. The length of this line segment is empirically set to be 8 pixels in consideration of the normal expectation of media layer thickness. Similar parametric assumption has been used in the literature, e.g. Garvin *et al.* [46] in segmenting retina layers in OCT. This empirical parameter setting does not have a deterministic effect on the result. When dealing with different resolution for different imaging devices, this can be accommodated by scaling this parameter according to spatial resolution.

The presence of stent causes scattering of ultrasound signals, leading to very bright pixels. Hence, once stent is detected by RF classifier, it is straight forward to localize the stent region which should not be part of media or adventitia, by searching for brightest pixels. The cost for the area above the stent is assigned a positive constant as the desired border is below the stent. Second order GD (G_2) are used to assign cost value for the regions underneath the stent, and G_2 is rotational invariant measurement of bar-like feature. The same transformation for ϕ_s is applied to G_2 due to the similar nature.

As for guide-wire artifact, there are also very bright pixels very close to the catheter, and it casts acoustic shadow or reverberation over its entire column. Hence, we do not extract any feature and a constant is used as their cost value. An example of the boundary term is visualized in the first row of Fig. 4, where the original image and the classification result are illustrated in Fig. 1.

2.4.2. OCT cost function

For OCT lumen segmentation, plaques in the artery pose little impact on the lumen border since they are not obstructing the lumen border. When there is no imaging artifacts, a good contrast along the lumen is achieved because the blood is flushed out before imaging. Hence, we can highlight the border based on intensity discontinuity. We use a combination of asymmetric local phase and the first order GD, both of which have good response to the lumen border. The cost function is thus defined as

$$cf(x, y) = G_1(x, y) - \phi_a(x, y), \tag{8}$$

where ϕ_a is the local phase asymmetry feature. Metal stent is implanted between the lumen and intima. Optical shadow caused by stent in OCT is narrow and generally behind the lumen border, see Fig. 1. The guide-wire artifact in OCT is similar to stent and has a bright echo usually near the catheter, casting large shadows. Hence, it is treated in the same way, i.e. we assign a constant cost value to the whole columns that are detected as containing guide-wire artifact. We rely on other constraints to estimate lumen border at regions where no or very little information is available for the border of interest. A visualization of such a cost function is given in Fig. 4.

2.5. Shape Prior Model

The shape prior model consisted of two terms; cost function and adaptive graph construction. The cost function is defined based on the similarity between the initial segmented shape (obtained through minimizing the boundary cost $\mathcal{C}_B(x, y)$ only) and prior shapes. The graph construction is then adapted so that inter-column arcs define dynamically according to the prior. The energy term for shape prior can be expressed as:

$$E_S = \sum_{(x,y)\in S} \mathcal{C}_P(x, y) + \sum_{(p,q)\in\mathcal{N}} f_{p,q}, \tag{9}$$

where $\mathcal{C}_P$ represents the prior cost function and $f_{p,q}$ is a convex function penalizing abrupt changes in the cut S between neighboring columns p and q in the set $\mathcal{N}$ of neighboring columns in the graph. Notably, in [47], Cremers *et al.* introduced a nonparametric shape prior model based on multiple shape templates to be incorporated in a level set framework. The authors in [25], adapted the method into discrete domain and used graph cut as the segmentation method. However, these methods depend on the user interaction to determine the object and background and need multiple iterations until convergence. In this work, we propose a nonparametric shape prior model that can be solved by computing the minimum s-t cut in a node-weighted directed graph, without any user intervention.

2.5.1. Cost Function

Each shape is represented by a binary template where the interior is set to one and the exterior is set to zero. The distance between two templates ψ^a and ψ^b is defined as sum of squared differences:

$$d^2(\psi^a, \psi^b) = \sum_{\Omega} (\psi^a - \psi^b)^2. \tag{10}$$

where Ω denotes the image domain. This distance measure is a true metric and is not influenced by image size. Let $\Psi = \psi^1, ..., \psi^N$ denote the N number of shapes from the training set. Given a possible cut in the graph which produces an aligned binary shape ψ, its similarity to a shape template ψ^n in the training set is computed as:

$$\alpha(\psi, \psi^n) = exp(-\frac{1}{2\sigma^2} d^2(\psi, \psi^n)). \tag{11}$$

Thus, the likelihood of this particular cut can be evaluated by taking into account of all training shapes:

$$c_{R_0} = \frac{\sum_{n=1}^{N} \alpha(\psi, \psi^n)\psi^n}{\sum_{n=1}^{N} \alpha(\psi, \psi^n)}. \tag{12}$$

Here, an initial cut can be conveniently obtained by minimizing the boundary based cost function alone. The labeling of the shape likelihood and initial cut needs to be compared in order to assign appropriate terminal arcs. The shape prior cost is defined as:

$$\mathcal{C}_P(x, y) = \lambda_2 |c_{R_0}(x, y) - c_{R_1}(x, y)|, \tag{13}$$

where c_{R_0} and c_{R_1} denote the cost associated to prior for the inferior region (the region under the border) and superior region (the region above the border) respectively, and λ_2 is the weight for the shape prior cost. The summation of the normalized weighted templates c_{R_0} is in effect the inferior-region cost and is inversely proportional to the likelihood of a pixel belong to the region underneath the border of interest. To define the superior-region prior cost c_{R_1}, we simply compute the complement of c_{R_0}, i.e.

$$c_{R_1} = \max_{x,y} c_{R_0}(x, y) - c_{R_0}(x, y). \tag{14}$$

As shown in Section 2.6, the shape prior cost $\mathcal{C}_P(x, y)$ is used to assign weights for each pixel according to its position from the border. By assigning the shape prior cost in this way, we eliminate the need to identify the terminal connection type. The images in second row of Fig. 4 provide an example of IVUS and OCT shape prior cost $\mathcal{C}_P(x, y)$.

2.5.2. Adaptive Graph Construction

A convex function $f_{p,q}(h)$ is defined between neighboring columns p and q to penalize inter-column changes that deviate from expectation (obtained from prior) and impose smooth inter-column transitions (see orange arcs in Fig. 3 (b)). To that end, the smoothness penalty arcs are constructed with intermediate links, i.e. $h \in (\underline{\Delta}_{p,q}, \bar{\Delta}_{p,q})$ for inter-column arcs. The direction of these arcs is based on the first order derivative of the convex function $f_{p,q}(h)$:

$$f_{p,q}(h) = \lambda_1 (h - \mu_{p,q})^2, \tag{15}$$

where λ_1 is a weighting factor for smoothness constraint and $\mu_{p,q}$ is the expected difference in border position between neighboring columns p and q. If the inter-column difference matches expectation, i.e. $f'_{p,q}(h) = 0$, two directional arcs of opposite direction are established between $V(x, y)$ and $V(x + 1, y - h)$. The weights on the arcs are determined by the second order derivative of the convex function $((f''_{p,q}(h))$ and are equally distributed. If $f'_{p,q}(h) > 0$, an arc from $V(x, y)$ to $V(x+1, y-h)$ is established and its weight is assigned as $f''_{p,q}(h)$; and if $f'_{p,q}(h) < 0$, the arc is connected from $V(x + 1, y)$ to $V(x, y + h)$ with its weight assigned as $f''_{p,q}(h)$.

In this work, we use this mean difference, $\mu_{p,q}$, to impose shape prior in graph construction, in addition to the shape prior term in the cost function. The aligned training shapes are also transformed into polar coordinates. The nearest neighbors are used to impose the shape

constraint. The inter-column differences are then generalized using mean $\mu_{p,q}$ and standard deviation $\sigma_{p,q}$ at individual columns. These statistics are then used in determining maximum and minimum distances when connecting neighboring columns, i.e. $\bar{\Delta}_{p,q} = \mu_{p,q} + c \cdot \sigma_{p,q}$, $\underline{\Delta}_{p,q} = \mu_{p,q} - c \cdot \sigma_{p,q}$, and c is a real constant. The smoothness arcs are also defined between $\underline{\Delta}_{p,q}, \bar{\Delta}_{p,q}$. These inter-column arcs impose a hard constraint, enforcing a general variation in shape with its details determined by minimizing the cost function that is based on both image feature and shape prior.

2.6. Energy minimization

Each graph node is weighted by a value $w(x, y)$ representing its rank to be selected in the minimum closed set graph. The cost function C is defined as $C(x, y) = \mathcal{C}_B(x, y) + \mathcal{C}_P(x, y)$. The weight assignment is carried out according to:

$$w(x,y) = \begin{cases} C(x,y) & \text{if } y = 0, \\ C(x,y) - C(x,y-1) & \text{otherwise.} \end{cases} \tag{16}$$

For a feasible path $\mathcal{P}$ in the graph, the subset of nodes on or below $\mathcal{P}$ form a closed set and it can be shown that the cost of the s-t cut in the graph is equivalent to the cost of nodes in the corresponding subset differ by a constant [26]. Hence, segmenting the border of interest is equivalent to finding the minimum closed set in the directed graph. The s-t cut algorithm [48] can then be used to find the minimum closed set, based on the fact that the weight can be used as the base for dividing the nodes into nonnegative and negative sets. The source s is connected to each negative node and every nonnegative node is connected to the sink t, both through a directed arc that carries the absolute value of the cost node itself. Since the graph has $O(n)$ nodes and $O(m)$ edges, the minimum closed set can be computed in $T(n, m)$ time that depends on the choice of the s-t cut algorithm. An efficient implementation has been proposed by Boykov and Kolmogorov [48], which showed to be more efficient than all the previous max-flow/min-cut algorithms.

2.7. Spatio-Temporal Tracking and Segmentation

Thus far, segmentation is carried out on a frame by frame basis, which is commonly used in object localization, e.g. [3], [4], [21]. However, it ignores the temporal information which could help to refine the object estimation, particularly when there is large uncertainty due to image artifact or noise. In the optimal surface segmentation [26], the authors added the shape constraints between the columns in neighboring frames as a constant value, whose optimal value is difficult to choose. The optimal surface method would also require substantial memory to construct 3D volume for thousands of IVUS or OCT frames. In [27], the authors used mean and standard deviation to model the shape variation between neighboring frames.

Here, we propose a spatio-temporal segmentation that is adaptive and only requires additional graph construction at adjacent frames. The spatial information is obtained by the proposed graph-based segmentation using both boundary and shape prior cost as described previously and the temporal information is extracted by using a Kalman filter to predict the border location in polar coordinates across frames and use this prior information to find

the final location of the border by adjusting the graph construction constraints between two consecutive frames.

However, instead of tracking all the border pixels in their polar coordinates, which is computationally expensive, we interpolate the border shape using RBFs. The RBF centers are uniformly placed along the angular coordinate. Thus, tracking the border of interest becomes localizing those RBF centers at a set of angular positions in the polar coordinates, i.e. the Kalman filter only needs to track and predict those few RBFs, instead of over a thousand individual pixels on the interface in polar coordinates. The thin plate local compact support RBF is used for interpolation [49]. The movement of the border across frames is in effect determined by the locations of those RBF centers.

The Kalman filter is an optimal estimator based on a linear state space system that recursively perform two processes: prediction and measurement. Generally the state estimation is defined as [50]:

$$x_t = Ax_{t-1} + w_t \approx N(0, Q), \tag{17}$$

and the measurement:

$$z_t = Hx_t + v_t \approx N(0, R), \tag{18}$$

where A is the state transition matrix, H is the measurement matrix, w_t denotes zero mean white Gaussian process with covariance Q, and v_t is the measurement noise with zero mean and covariance R.

A continuous Wiener process acceleration (CWPA) model [50] is used to estimate the RBF motion model and define the A and Q matrices. The desired border is sampling across n points of the border location in polar coordinates with fixed angle location so that only the radial coordinate needs to be predicated. The state vector has three components, i.e. location y_n, velocity $\dot{y_n}$ and acceleration $\ddot{y_n}$, which can be defined as:

$$x_k = [y_1 \dots y_n, \dot{y_1} \dots \dot{y_n}, \ddot{y_1} \dots \ddot{y_n}]^T. \tag{19}$$

The prediction phase is to estimate the prior state $\hat{x}_k^-$ and its covariance matrix P_k^- according to the motion model:

$$\begin{aligned} \hat{x}_k^- &= A\hat{x}_{k-1}, \\ P_k^- &= AP_{k-1}A^T + Q. \end{aligned} \tag{20}$$

At the measurement phase, the posterior state $\hat{x}_k$ is computed by correcting the prior state $\hat{x}_k^-$ based on the current observation z_k and controlled by the Kalman gain K:

$$\begin{aligned} \hat{x}_k &= \hat{x}_k^- + K(z_k - H\hat{x}_k^-), \\ K &= P_k^- H^T (HP_k^- H^T + R)^{-1}, \\ P_k &= (1 - KH)P_k^-. \end{aligned} \tag{21}$$

The observation z_k of the Kalman measurement is acquired by temporal graph cut segmentation, in which new inter-column arcs are defined between the two adjacent frames. These arcs impose hard constraints to control the maximum distance, Δ_{p_{k-1},p_k} at column

p, between the observed border S^* of frame k and the final border S of frame $k-1$, i.e. $|S_k^*(p) - S_{k-1}(p)| \leq \Delta_{p_{k-1},p_k}$. The value of Δ_{p_{k-1},p_k} is calculated based on the classification result in frame k so that smaller inter-frame movement is allowed when image artifacts are detected. An example of the temporal graph construction is shown in Fig. 5.

To incorporate the Kalman predicated border $\hat{S}$ into the final segmentation, we define the convex function between the two adjacent frames as following:

$$E_T = \sum_{(p_{k-1},p_k)\in\mathcal{N}_t} f_{p_{k-1},p_k}, \tag{22}$$

where $f_{p_{k-1},p_k}(h) = \lambda_3(h - \mu_{p_{k-1,k}})^2$ is penalizing abrupt changes between the same columns p in two adjacent frames $k-1, k$ in the set $\mathcal{N}_t$ of all columns in two adjacent graphs. The arcs between the two frames are defined as described in Section 2.1, where $\mu_{p_{k-1,k}}$ is the difference between the posterior estimation of border $\hat{S}$ in the current frame k and the final border segmentation in the previous frame $k-1$ and the inter-frames constraints is defined as

$$\begin{aligned} \bar{\Delta}_{p_{k-1,k}} &= \mu_{p_{k-1,k}} + c \cdot \sigma_{p_{k-1,k}}, \\ \underline{\Delta}_{p_{k-1,k}} &= \mu_{p_{k-1,k}} - c \cdot \sigma_{p_{k-1,k}}, \end{aligned} \tag{23}$$

where c is a real constant and $\sigma_{p_{k-1,k}}$ is defined based on the Kalman covariance P_k.

The final segmentation is then obtained by minimizing the three energy functions (5), (9), and (22). To make the Kalman filter more consistent, the final segmented border is used to update the measurement process. This process is known in literature [51] as a sequential Kalman filter.

Note, the final border segmentation for frame $k-1$ is not changed after segmenting the frame k. This can be achieved by redefining the cost function $\hat{\mathcal{C}}_B(x,y)$ for frame $k-1$ so that negative infinite cost is assigned for the nodes $V(x,y)$ corresponding to the border in frame $k-1$ and the rest nodes are assigned value one. In other words, we connect the nodes just above the border to the sink and the border nodes itself to the source, both with infinite weight, so that the result of the minimum closed set graph for frame $k-1$ is fixed.

3. Experimental Results

3.1. Dataset and Evaluation

We have collected a total of 33,220 IVUS images and 2,303 OCT images in order to evaluate the proposed method in segmenting media-adventitia border in IVUS and lumen border in OCT. The IVUS images were acquired from 10 *in vivo* pullbacks of human subjects were acquired by a 40 MHz transducer Boston Scientific ultrasound machine with Atlantis SR Pro Catheter. The media-adventitia border is of interest in these images. Manual labeling of the borders was carried out on every 10th frame, i.e. 3,322 frames to establish ground-truth for quantitative analysis[1]. These images contain various forms of fibrous plaque,

[1]Manual labeling the entire dataset is too time consuming. However, quantitative comparison on every 10th frame should provide reliable assessment.

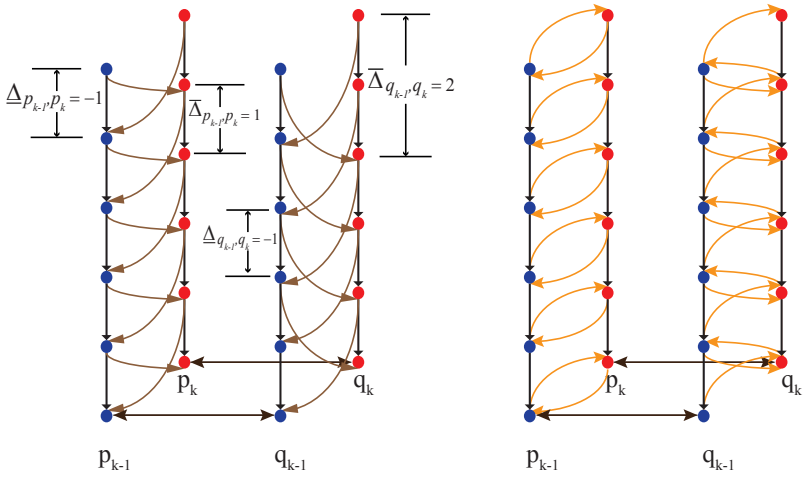

Figure 5: Example temporal graph construction between two frames, $k-1$ and k. (a) inter-columns arcs (brown) connecting the same columns in two neighboring frames, p_{k-1} and p_k, where $\bar{\Delta}_{p_{k-1},p_k} = 1, \underline{\Delta}_{p_{k-1},p_k} = -1$, and q_{k-1} and q_k where $\bar{\Delta}_{q_{k-1},q_k} = 2, \underline{\Delta}_{q_{k-1},q_k} = -1$. (b) smoothness arcs (orange) for the same columns when $f'_{p_{k-1},p_k}(0) = 0$ and $f'_{q_{k-1},q_k}(0) = 1$. Note, Figure 3 shows graph arcs within one frame and here the temporal arcs are shown between two frames to complete the whole set of graph arcs.

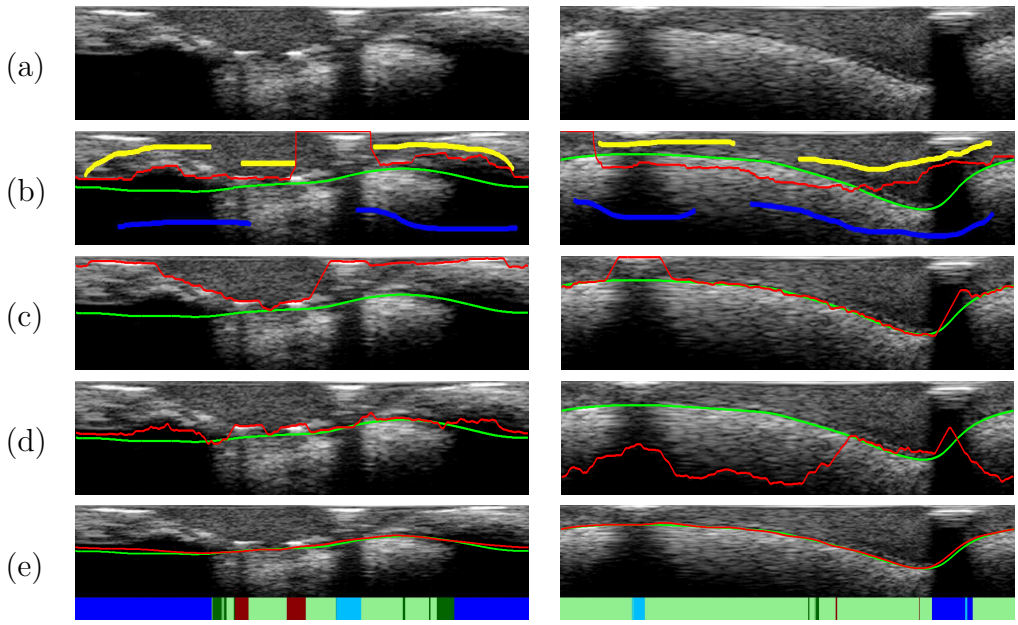

Figure 6: Comparison between ground-truth (green) and segmentation results (red): (a) Original image; (b) $s-t$ cut [52] (foreground: yellow, background: blue); (c) Optimal surface [26] with 1st order GD; (d) Optimal surface [26] with 2nd order GD; (e) Proposed method: bottom of each image also visualizes the RF classification result: calcified plaque (blue), fibrotic plaque (dark green), stent (dark red), guide-wire shadowing (cyan), and soft plaque/normal tissue (light green).

calcification, stent, and acoustic shadow. From the unlabeled IVUS images, 615 images were selected to train the tissue classifier. From the 3,322 labeled images, 667 images were used to generalize the shape and temporal prior and the rest 2,655 images formed the testing set. It is worth noting that the proposed framework can be adapted to multi-surface IVUS segmentation so that both the lumen border and the media-adventitia border can be segmented simultaneously. However, the blood speckles under high frequency of the transducer (40 MHz) become prominent and the unavailability of the raw radiofrequency signals of our dataset make it a difficult to differentiate the lumen border. In practice, a different frequency imaging can be carried out in order to segment the lumen border.

The OCT dataset consists of 13 *in vivo* pullbacks of human subjects acquired from C7-XR LightLab Imaging system. The primary use for OCT in coronary imaging is lumenology as it has a good visualization of the arterial wall. OCT has a much lower penetration depth compared to IVUS, which can limit its ability to assess large plaque burden. For our dataset, the media layer is not visible and our objective is to segment the lumen border. Column-wise labeling was carried out on 457 images to train the OCT tissue classifer, and all the rest 1,846 images were manually labeled for both training, 457 images, and testing, 1,389 images.

Furthermore, we evaluate our method on the 40Mhz IVUS dataset introduced by the MICCAI 2011 IVUS segmentation challenge workshop [53] to segment media-adventitia borders. This data set contains 77 images extracted from 22 pullbacks. Four adjacent frames of each of those images are also provided. A quarter of the images are provided as the training set. A further 291 images from our own images were added to the training set. This is in line with the challenge protocol and used by their participants, since the training set is relatively small. The testing set consists of 58 images. The image quality of this dataset varies significantly from our own datasets. The images in the IVUS challenge dataset have undergone pre-processing of blood filtering which leads to different appearances. Thus, our appearance based tissue classification needs to be re-trained. However, RF is efficient to train and it is possible to build an adaptive RF. Our shape prior however remains the same.

The proposed method was compared against a number of recent and state-of-the-art techniques, which include optimal surface graph cut [26], star graph cut [18] and texture-RBF for IVUS segmentation [3]. Quantitative comparison to methods that were reported in the challenge dataset was also carried out. The evaluation metrics include absolute mean difference (AMD), Hausdorff distance (HD) in pixel, sensitivity (Sens.), specificity (Spec.), Jaccard measure (JM), and percentage of area difference (PAD). The inner region is considered as positive for Sens. and Spec. measurements. The evaluation on the challenge dataset exactly followed their protocol with provided evaluation code.

3.2. Column-wise Tissue Classification

To perform column-wise tissue classification, we randomly selected a small number of images for supervised training, i.e. less than 2% IVUS images. The total number of trees in RF is set to be 500 and the number of features used to make the decision at a node of the tree is $4\sqrt{L}$, where L is the length of feature vector. The IVUS feature length is 3808 computed from vertical window of size of 128 and horizontal window size of 512.

Table 1: IVUS Column-wise tissue classification.

Accuracy(%)	IVUS					
Class	Fibrotic	Calcification	Stent	Guide-wire	Others	**Overall**
Intensity features	62.42	88.16	59.97	72.75	89.87	80.64
Haar-like features	72.52	89.67	64.42	80.84	89.72	**83.78**

Table 2: OCT Column-wise tissue classification.

Accuracy(%)	OCT			
Class	Stent	Guide-wire	Others	**Overall**
Intensity features	44.08	95.69	97.64	89.47
Haar-like features	82.82	95.91	99.25	**96.27**

Using the proposed Haar-like features from 8 scales, we achieved a modest improvement in overall accuracy from 80.64% using intensity based feature in our earlier work [54] to 83.78% in classifying individual columns of pixels into five different categories. However, since most columns of pixels are classified as "others" (mainly contain normal tissues), there are substantial improvements in classifying columns that contain imaging artifacts, e.g. fibrotic tissue classification increased from 62.42% to 72.52%. Table 1 summarizes the results.

For OCT tissue classification, the overall accuracy using the proposed features is 96.27%, compared to an overall accuracy of using pixel intensity profile of 89.47%. The main improvement is in detecting stent, i.e. a substantial improvement from 44.08% to 82.82%. Detailed classification results for each type of tissue can be seen in Table 2.

3.3. IVUS Segmentation

First, we provide a comparison to a conventional graph cut method that does not use shape prior constraints but relies on user intervention to guide segmentation. In Fig. 6, we qualitatively compare our proposed method (row (e)) to the conventional $s-t$ cut algorithm (row (b)). The $s-t$ cut does not incorporate shape prior into the cost function. Instead, prior information is incorporated through user initialization by providing strokes on foreground and background regions. Even with careful initialization, the conventional $s-t$ cut did not produce satisfactory results. Due to its high initialization dependency and relatively intensive manual process, we did not carry out a full quantitative analysis. However, later we show a full quantitative comparison to an alternative graph cut which also adopts user initialization but we believe is more suited for IVUS segmentation. The failure of standard $s-t$ cut suggests that a stronger prior is necessary and also the image features and its cost function may need to be tailored.

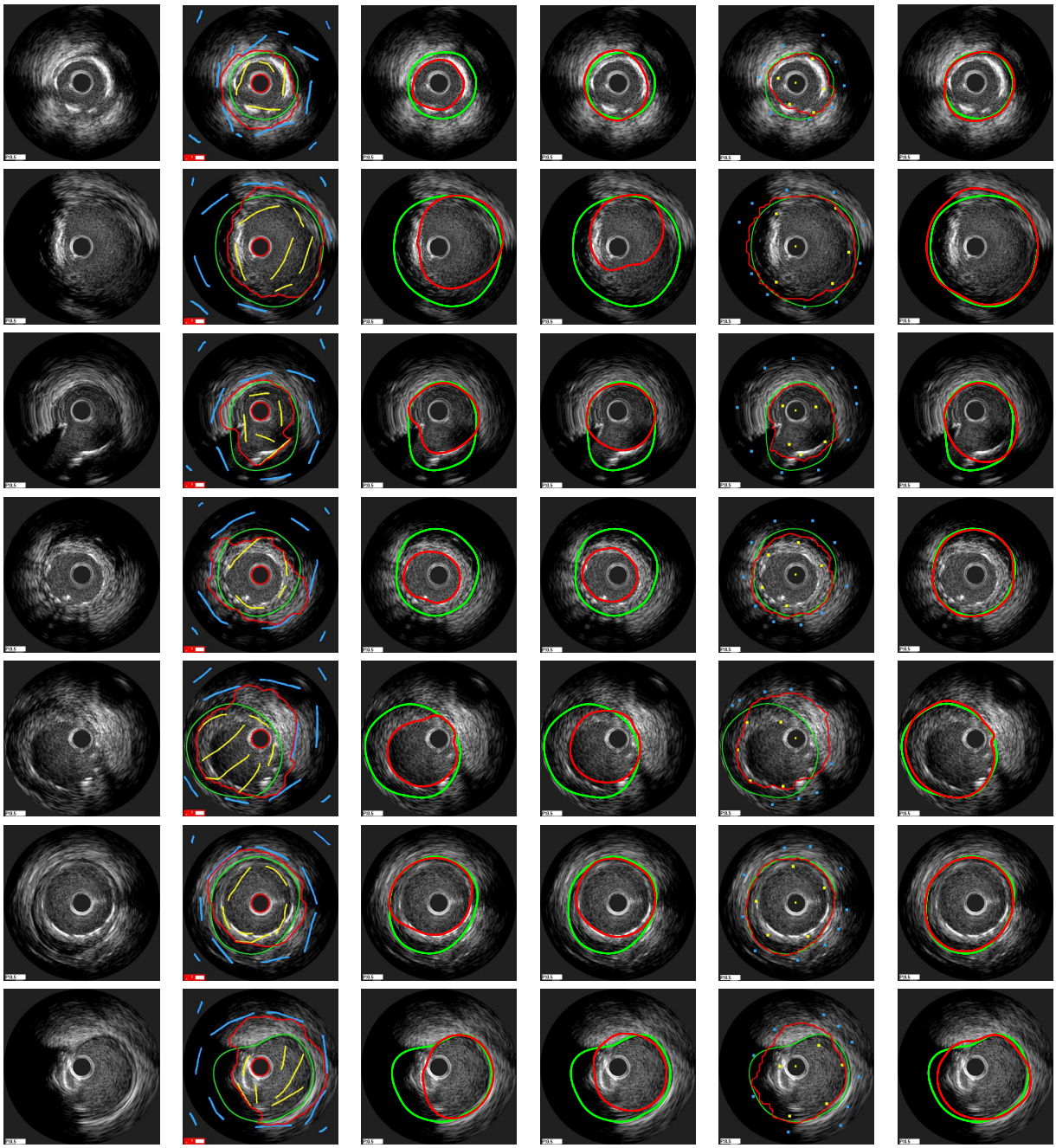

Figure 7: Comparison between ground-truth (green) and segmentation results (red): (a) original image. (b) $s-t$ cut (foreground user strokes: yellow, background user strokes: blue). (c) optimal surface [26] with first order GD. (d) Texture-RBF method. (e) star graph cut (foreground user points: yellow, background user points: blue). (f) proposed method.

Next, we employ the optimal surface graph [26], which is the basis of our graph construction, to examine the role of image features. The first and second order GD features

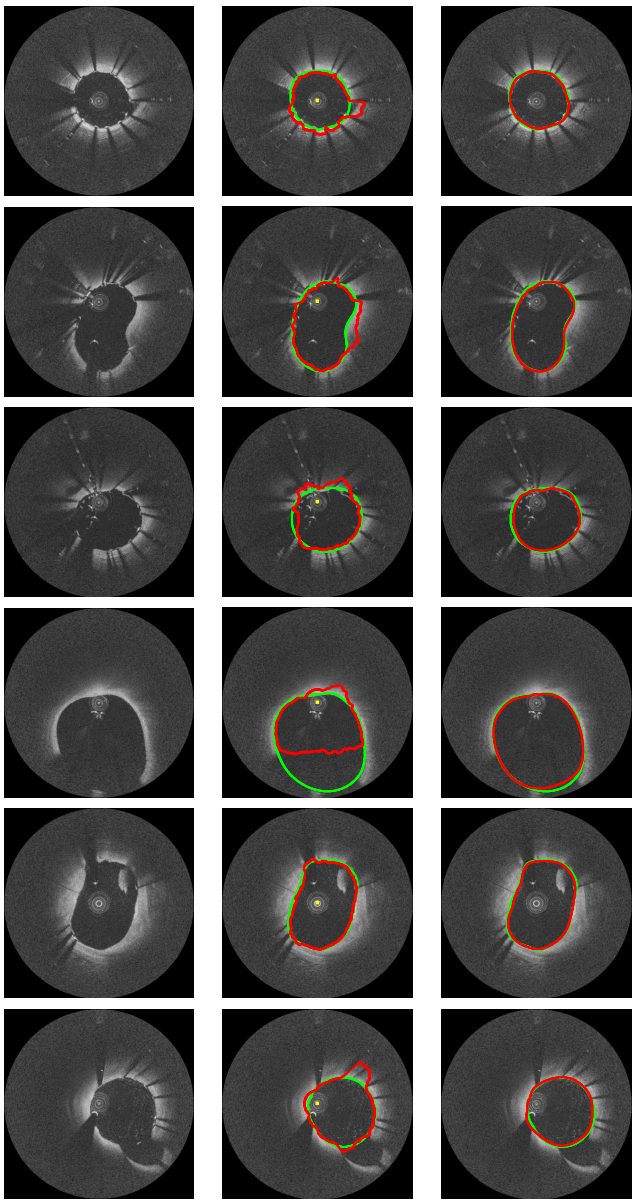

Figure 8: Comparison between ground-truth (green) and segmentation results (red): (a) original image. (b) star graph cut (star point: yellow). (c) proposed method.

are the fundamental features in our method. Hence, we applied these two features separately with optimal surface graph, with the same local shape constraint as in our proposed method. Rows (c) and (d) in Fig. 6 show typical results of applying first order GD and second order GD with optimal surface graph, respectively. With the proposed method, we used RF to classify five different types of tissue and adapted the cost function accordingly. The bottom of each image in row (e) visualizes the classification result. In this example, there are multiple artifacts in some of the regions, e.g. the left side of the image is classified

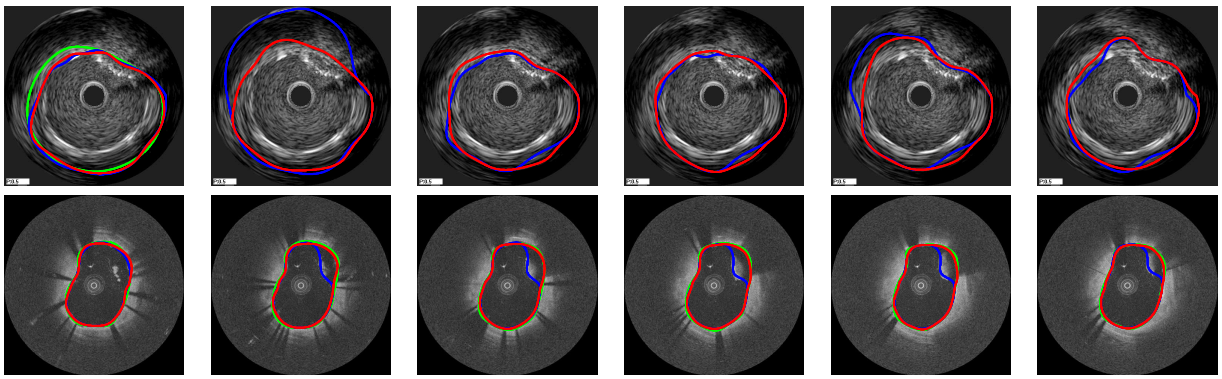

Figure 9: Segmentation with longitudinal tracking using the proposed method (red) performs better than segmenting the images on a frame by frame basis (blue). The groundtruth is shown in green.

Table 3: IVUS quantitative comparison. Mean (standard deviation).

	AMD	HD	JM	Sens.	Spec.
Optimal surface [26] with 1st GD	21.94 (19.36)	53.77 (39.64)	0.84 (0.13)	84.57 (13.95)	**98.87 (1.55)**
Optimal surface [26] with 2nd GD	35.52 (32.56)	68.97 (44.71)	0.79 (0.17)	94.46 (6.11)	77.77 (25.94)
Texture-RBF method [3]	21.68 (13.37)	49.82 (27.99)	0.84 (0.09)	87.92 10.14	95.60 (4.70)
Star Graph cut [18]	14.70 (8.75)	38.98 (17.53)	0.89 (0.06)	94.90 (4.23)	92.84 (7.29)
Proposed method	**6.46 (5.90)**	**19.23 (16.63)**	**0.95 (0.04)**	**96.60 (3.68)**	98.01 (2.87)

as calcified plaque and two stent strut are classified in the center. Our ground-truth labeling and automated classification for each column are based on the dominant artifact. In this case, the dominant effect on the left side of the image is the shadowing effect caused by calcification.

Further examples of optimal surface method can be found in column (c), Fig. 7. The quantitative results are shown in Table 3. It is evidently clear that the proposed method achieved significantly superior performance, e.g. the absolute mean difference for the proposed method is 6.46 compared to 21.94 and 35.52 for optimal surface with two different feature sets.

We also implemented the Texture-RBF method [3], which is designed for segmenting IVUS. Its overall performance is very similar to optimal surface graph with first order GD, see Table 3. Example results are shown in Fig. 7(d). This method consistently underestimated the media-adventitia border.

Finally, we compare our method against the semi-automatic star graph cut [18]. Since

Table 4: OCT quantitative comparison. Mean (standard deviation).

	AMD	HD	JM	Sens.	Spec.
Optimal surface [26] with 1st GD	7.25 (5.34)	37.06 (29.05)	0.89 (0.06)	89.76 (5.99)	99.80 (0.56)
Star Graph cut [18]	5.86 (4.41)	22.23 (17.67)	0.91 (0.06)	**96.28 (4.86)**	98.16 (1.61)
Proposed method	**3.26 (2.74)**	**11.01 (11.93)**	**0.95 (0.03)**	95.55 (3.19)	**99.84 (0.29)**

the media-adventitia border is assumed on the radial direction from the catheter center without obstruction, the intrinsic shape regularization in star graph cut is well suited for IVUS segmentation, i.e. the catheter center is treated as the star point and the media-adventitia border lies on the radial spikes from the star point. Segmentation using a single user point (star point) however resulted failure in most cases. Hence, additional user points were carefully placed on each cross section: 5 foreground points and 10 background points, see column (e) in Fig. 7. These additional constraints improved its performance, but it was still adversely affected by IVUS artifacts. The proposed method, column (f), showed very close segmentation to the ground-truth. Comparative quantitative results are shown in Table 3. It clearly shows statistically significant improvements of the proposed method.

3.4. OCT Segmentation

Table 4 shows the quantitative comparison between the star graph [18] and our proposed method on OCT segmentation. Here, the star graph method requires the catheter region to be removed before segmentation in order to obtain meaningful results. Similar to IVUS segmentation, the proposed method achieved superior performance in segmenting the lumen border in OCT. Fig. 8 shows typical results of star graph cut and ours.

Moreover, Fig. 9 demonstrates the benefit of imposing temporal shape constraint using the proposed method compared to segmentation on a frame-by-frame basis, especially at bifurcations and where blood is not completely flushed out during OCT imaging. A sudden change of segmented lumen border can be seen for the frame-by-frame segmentation, see the first two images in the second row. An IVUS example is also given in the top row. If however the artifacts exit in the starting frame, mis-segmentation can occur at this point due to lack of temporal information. As more frames are made available, temporal segmentation reduces border oscillation in the longitudinal direction.

3.5. Leave-one-subject-out Results and Performance Analysis

So far, the experiments conducted involve all subject samples, which is a common evaluation method used in these studies. Next, we carry out leave-one-subject-out cross validation, that is the labelled samples of the testing sequence are completely excluded from the training. This applies to both shape prior generalization and tissue classification. The results

Table 5: IVUS Leave-one-subject-out quantitative result. Mean (standard deviation)

	AMD	HD	JM	Sens.	Spec.
Classification-based Segmentation	9.24 (6.54)	32.25 (20.35)	0.928 (0.05)	95.32 (0.04)	96.99 (0.04)
Shape prior based Segmentation	8.04 (6.90)	23.79 (18.05)	0.938 (0.05)	96.24 (0.04)	96.88 (0.05)
Temporal based Segmentation	7.90 (6.09)	25.04 (17.51)	0.938 (0.04)	96.02 (0.04)	97.30 (0.04)
Shape prior and Temporal based Segmentation	7.51 (5.85)	22.51 (16.37)	0.942 (0.04)	96.38 (0.03)	97.17 (0.04)

obtained thus will demonstrate the generalization capability of the proposed method. We also carry out a comprehensive performance analysis in order to understand the contributions of individual components of the proposed method.

Tables 5 and 6 show the quantitative results using leave-one-subject-out validation on IVUS and OCT. We also analyze each component of the proposed method, i.e. their contributions to the overall performance.

We first examine the performance of segmentation based on the RF classifier to dynamically define the cost function. The graph construction is defined frame by frame. Its performances on IVUS and OCT in leave-one-subject-out validation are shown in the first row of Tables 5 and 6. Compared to the optimal surface method using fixed cost function shown in Tables 3 and 4, the classification based cost function achieved significant improvements to segment both the media-adventitia border in IVUS and the lumen border in OCT.

Next, we separately examine the role of shape prior and temporal constraint. The shape prior works by adapting the graph construction to add smoothness transition of the border and also by exerting constraints based on the shape similarity metric. The initial segmentation is computed first using the prior adaptive graph and then aligned to training shapes to compute the shape prior cost. The results are shown in the second row in the tables. In the third experiment, the temporal constraints is incorporated without the shape prior by segmenting two adjacent frames and adding additional inter-frame arcs, and its results are presented in the third row. Both shape prior and temporal constraint performed better than the classification based segmentation and hence optimal surface. The temporal constraints in IVUS worked out slightly better than in OCT probably due to the reason that IVUS contains more frames than OCT in each pullback.

Finally, the last row shows the performances of the proposed method with all components integrated. Improvements were observed on almost all evaluation metrics.

Table 6: OCT Leave-one-subject-out quantitative result. Mean (standard deviation)

	AMD	HD	JM	Sens.	Spec.
Classification-based Segmentation	5.99 (3.42)	26.49 (16.50)	0.908 (0.05)	91.34 (0.04)	99.80 (0.01)
Shape prior based Segmentation	3.68 (3.22)	11.71 (13.44)	0.945 (0.04)	95.18 (0.04)	99.73 (0.08)
Temporal based Segmentation	4.01 (2.62)	14.71 (13.80)	0.939 (0.04)	94.45 (0.03)	99.78 (0.01)
Shape prior and Temporal based Segmentation	3.56 (2.46)	11.64 (10.75)	0.946 (0.03)	95.24 (0.03)	99.78 (0.01)

3.6. Parameter settings

The parameters λ_1, λ_2, and λ_3 are used as weightings for inter-column smoothing, shape prior, and Kalman based inter-frame smoothing, respectively. Weightings are necessary as those attributes are different from each other (i.e. at different scales) and their contributions to the cost function should be treated differently as well. Larger λ_1 value will limit inter-column movements, and smaller the value the more flexible the segmentation moving from column to the next. λ_2 is used to balance the contribution between image feature and shape prior; the larger the value the stronger the shape constraint. λ_3 controls the contribution from Kalman prediction. The employment of these weighting parameters is common in energy minimization based approaches, and the selection of these parameter values is largely heuristic driven. The constant c is used in defining upper and lower bound and used in the context of a parameterized normal distribution. Its role is well documented in the literature and it typically takes the value 2 or 3 (value 3 is used in our work). The standard deviation for the Gaussian function in computing shape similarity is empirically chosen, and is directly related to bandwidth selection in Parzen window density estimation. Automatically model order selection has been reported, e.g. MDL, but it is computationally very expensive. The number of RBFs is selected so that they are evenly spread horizontally. The number of centers is chosen so that it is sufficient in describing the overall shape and it does not affect the fundamental performance of the method. The interpolation problem is very simple and in 1D; the selection of the basis function does not make a material difference.

3.7. IVUS challenge dataset

The proposed method was also applied to the 40Mhz IVUS dataset from this IVUS challenge [53]. Table 7 summarizes the quantitative comparison between the proposed method and other methods reported in [53] that were tested on the same IVUS images in segmenting the media-adventitia border. Note, the method proposed by participant 3 (P3) is a semi-automatic method where the user must initialize the contour position for the first image of each sequence. The other two methods (P6 and P8) automatically segment the border. The

Table 7: IVUS challenge quantitative comparison. Mean (standard deviation)

	JM	HD	PAD
P3 [53]	0.86(0.11)	1.18(1.02)	0.10(0.11)
P6 [53]	0.76(0.11)	1.78(0.83)	0.17(0.14)
P8 [53]	0.80(0.13)	1.57(1.03)	0.14(0.16)
Proposed method	0.84(0.10)	1.22(0.72)	0.13(0.15)

statistical result shows that our proposed method outperforms both automatic segmentation methods and closely follows the semi-automated results. Note, this dataset does not allow us to implement our proposed temporal constraints due to the images are not completed pullbacks.

4. Conclusions

We presented an automated segmentation method for two different intravascular imaging modalities in order to delineate media-adventitia borders in IVUS pullbacks, and lumen borders in OCT pullbacks. The proposed method uses local and global shape constraints in both graph construction and adaptive cost function implementation. Tissue classification proved useful to optimize the cost function and we show that column-wise classification is sufficient, rather than pixel-wise classification which is problematic. The segmentation is also regularized by temporal border tracking to ensure consistency among consecutive frames. Quantitative and qualitative comparison on three coronary imaging datasets that contains over 35,000 images showed clear advantages of the proposed method when compared to both conventional graph cut methods and state-of-the-art approaches, including optimal surface and star graph cut.

References

[1] J. D. Klingensmith, R. Shekhar, D. G. Vince, Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial adventitial borders in intravascular ultrasound images, IEEE T-MI 19 (10) (2000) 996–1011.

[2] M.-H. R. Cardinal, et al., Intravascular ultrasound image segmentation: a three-dimensional fast-marching method based on gray level distributions, IEEE T-MI 25 (1) (2006) 590–601.

[3] M. Papadogiorgaki, et al., Image analysis techniques for automated ivus contour detection, Ultrasound in Medicine and Biology 9 (34) (2008) 1482–1498.

[4] M. Sonka, et al., Segmentation of intravascular ultrasound images: A knowledge-based approach, IEEE T-MI 14 (1995) 719–732.

[5] A. Takagi, et al., Automated contour detection for high frequency intravascular ultrasound imaging: A technique with blood noise reduction for edge enhancement, Ultrasound in Medicine and Biology 26 (6) (2000) 1033–1041.

[6] E. Essa, X. Xie, I. Sazonov, P. Nithiarasu, Automatic IVUS media-adventitia border extraction using double interface graph cut segmentation, in: ICIP, 2011, pp. 69–72.

[7] F. Destrempes, et al., Segmentation method of intravascular ultrasound images of human coronary arteries, Computerized Medical Imaging and Graphics 38 (2) (2014) 91 – 103.
[8] Z. Gao, et al., Automated framework for detecting lumen and mediaadventitia borders in intravascular ultrasound images, Ultrasound in Medicine & Biology 41 (7) (2015) 2001 – 2021.
[9] S. Su, Z. Hu, Q. Lin, W. K. Hau, Z. Gao, H. Zhang, An artificial neural network method for lumen and media-adventitia border detection in IVUS, Computerized Medical Imaging and Graphics 57 (2017) 29 – 39.
[10] F. S. Zakeri, S. K. Setarehdan, S. Norouzi, Automatic media-adventitia IVUS image segmentation based on sparse representation framework and dynamic directional active contour model, Computers in Biology and Medicine (2017) –doi:https://doi.org/10.1016/j.compbiomed.2017.03.022.
[11] C. Kauffmann, P. Motreff, L. Sarry, in vivo supervised analysis of stent reendothelialization from optical coherence tomography, IEEE T-MI 29 (3) (2010) 807–818.
[12] S. Gurmeric, G. G. Isguder, S. Carlier, G. Unal, A new 3-d automated computational method to evaluate in-stent neointimal hyperplasia in in-vivo intravascular optical coherence tomography pullbacks, in: Med Image Comput Comput Assist Interv., 2009, pp. 776–785.
[13] G. Unal, S. Gurmeric, S. G. Carlier, Stent implant follow-up in intravascular optical coherence tomography images, Int. J. Cardiovasc Imaging 26 (2010) 809–816.
[14] K.-P. Tung, W.-Z. Shi, R. D. Silva, E. Edwards, D. Rueckert, Automatical vessel wall detection in intravascular coronary oct, in: IEEE ISBI, 2011, pp. 610–613.
[15] G. J. Ughi, et al., Automatic segmentation of in-vivo intra-coronary optical coherence tomography images to assess stent strut apposition and coverage, Int. J. Cardiovasc Imaging. 28 (2) (2012) 229–241.
[16] S. Tsantis, et al., Automatic vessel lumen segmentation and stent strut detection in intravascular optical coherence tomography, Med. Phys. 39 (1) (2012) 503–513.
[17] Y. Cao, et al., Automatic lumen segmentation in intravascular optical coherence tomography images using level set, Computational and Mathematical Methods in Medicine (2017) – doi:https://doi.org/10.1155/2017/4710305.
[18] O. Veksler, Star shape prior for graph-cut image segmentation, in: ECCV, 2008, pp. 454–467.
[19] J.-L. Jones, X. Xie, E. Essa, Combining region-based and imprecise boundary-based cues for interactive medical image segmentation, International Journal for Numerical Methods in Biomedical Engineering 30 (12) (2014) 1649–1666.
[20] S. Sun, M. Sonka, R. Beichel, Graph-based ivus segmentation with efficient computer-aided refinement, IEEE Transactions on Medical Imaging 32 (8) (2013) 1536–1549.
[21] G. Unal, et al., Shape-driven segmentation of the arterial wall in intravascular ultrasound images, IEEE Trans. Info. Tech. in Biomed. 12 (3) (2008) 335–347.
[22] A. Wahle, et al., Plaque development, vessel curvature, and wall shear stress in coronary arteries assessed by x-ray angiography and intravascular ultrasound, MIA 10 (1) (2006) 615–631.
[23] D. Freedman, T. Zhang, Interactive graph cut based segmentation with shape priors, in: CVPR, 2005, pp. 755–762.
[24] J. Malcolm, Y. Rathi, A. Tannenbaum, Graph cut segmentation with nonlinear shape priors, in: ICIP, 2007, pp. 365–368.
[25] N. Vu, B. S. Manjunath, Shape prior segmentation of multiple objects with graph cuts, in: CVPR, 2008, pp. 1–8.
[26] K. Li, X. Wu, D. Z. Chen, M. Sonka, Optimal surface segmentation in volumetric images-a graph-theoretic approach, IEEE T-PAMI 28 (1) (2006) 119–134.
[27] Q. Song, X. Wu, Y. Liu, M. Garvin, M. Sonka, Simultaneous searching of globally optimal interacting surfaces with shape priors, in: CVPR, 2010, pp. 2879–2886.
[28] X. Wu, D. Chen, Optimal net surface problems with applications, in: Automata, Languages and Programming, 2002, pp. 1029–1042.
[29] H. Ishikawa, Exact optimization for markov random fields with convex priors, IEEE T-PAMI 25 (10) (2003) 1333–1336.
[30] E. Filho, Y. Saijo, A. Tanaka, M. Yoshizawa, Detection and quantification of calcifications in intravas-

cular ultrasound images by automatic thresholding, Ultrasound in Medicine and Biology 34 (1) (2008) 160–165.
[31] C. Xu, et al., Automatic detection of stent struts with thick neointimal growth in intravascular optical coherence tomography image sequences, Phys. Med. Biol. 56 (20) (2011) 6665–6675.
[32] T. Kubo, C. Xu, Z. Wang, N. S. van Ditzhuijzen, H. G. Bezerra, Plaque and thrombus evaluation by optical coherence tomography, Int. J. Cardiovasc Imaging. 27 (2) (2011) 289–298.
[33] D. Rotger, P. Radeva, N. Bruining, Automatic detection of bioabsorbable coronary stents in ivus images using a cascade of classifiers, IEEE Transactions on Information Technology in Biomedicine 14 (2) (2010) 535–537.
[34] L. S. Athanasiou, et al., A novel semiautomated atherosclerotic plaque characterization method using grayscale intravascular ultrasound images: Comparison with virtual histology, IEEE Transactions on Information Technology in Biomedicine 16 (3) (2012) 391–400.
[35] V. S. Lempitsky, M. Verhoek, J. A. Noble, A. Blake, Random forest classification for automatic delineation of myocardium in real-time 3d echocardiography, in: FIMH, 2009, pp. 447–456.
[36] E. G. Mendizabal-Ruiz, M. Rivera, I. A. Kakadiaris, Segmentation of the luminal border in intravascular ultrasound b-mode images using a probabilistic approach, Medical Image Analysis 17 (6) (2013) 649 – 670.
[37] A. Bosch, A. Zisserman, X. Muoz, Image classification using random forests and ferns, in: ICCV, 2007, pp. 1–8.
[38] J. Weston, G. BakIr, Fast binary and multi-output reduced set selection, Tech. Rep. TR-132, MPI (2004).
[39] X. Liang, An effective method of pruning support vector machine classifiers, IEEE Transactions on Neural Networks 21 (1) (2010) 26–38.
[40] L. Breiman, Random forests, Machine Learning 45 (1) (2001) 5–32.
[41] R. Caruana, A. Niculescu-Mizil, An empirical comparison of supervised learning algorithms, in: Proceedings of the 23rd International Conference on Machine Learning, ICML '06, 2006, pp. 161–168.
[42] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: cvpr, Vol. 1, 2001, pp. I–511–I–518.
[43] W. T. Freeman, E. H. Adelson, The design and use of steerable filters, IEEE T-PAMI 13 (9) (1991) 891–906.
[44] M. Mulet-Parada, A. Noble, 2D + T acoustic boundary detection in echocardiography, MIA 4 (1) (2000) 21–30.
[45] P. Kovesi, Symmetry and asymmetry from local phase, in: Tenth Australian Joint Conference on Artificial Intelligence, 1997, pp. 185–190.
[46] M. K. Garvin, M. D. Abramoff, R. Kardon, S. R. Russell, X. Wu, M. Sonka, Intraretinal layer segmentation of macular optical coherence tomography images using optimal 3-d graph search, IEEE Transactions on Medical Imaging 27 (10) (2008) 1495–1505.
[47] D. Cremers, S. Osher, S. Soatto, Kernel density estimation and intrinsic alignment for shape priors in level set segmentation, International Journal of Computer Vision 69 (3) (2006) 335–351.
[48] Y. Boykov, V. Kolmogorov, An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, IEEE T-PAMI 26 (9) (2004) 1124–1137.
[49] G. Turk, J. F. O'brien, Modelling with implicit surfaces that interpolate, ACM Trans. Graph. 21 (4) (2002) 855–873.
[50] Y. Bar-Shalom, X. R. Li, T. Kirubarajan, Estimation with Applica-tions to Tracking and Navigation, Wiley-Interscience, 2001.
[51] D. Simon, Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches, Wiley, 2006.
[52] Y. Boykov, G. Funka-Lea, Graph cuts and efficient N-D image segmentation, IJCV 70 (1) (2006) 109–131.
[53] S. Balocco, et al., Standardized evaluation methodology and reference database for evaluating IVUS image segmentation, Computerized Medical Imaging and Graphics 38 (2) (2014) 70 – 90.
[54] E. Essa, X. Xie, I. Sazonov, P. Nithiarasu, D. Smith, Shape prior model for media-adventitia border

segmentation in IVUS using graph cut, in: MICCAI-MCV, 2012, pp. 114–123.