Accurate and Robust Fully-Automatic QCA: Method and Numerical Validation

Antonio Hernández-Vela[1,2], Carlo Gatta[1,2], Sergio Escalera[1,2], Laura Igual[1,2], Victoria Martin-Yuste[3], and Petia Radeva[1,2]

[1] Dept. MAIA, Universitat de Barcelona, Gran Via 585, 08007 Barcelona, Spain
[2] Centre de Visió per Computador, Edifici O, Campus UAB, 08193 Bellaterra, Spain
{ahernandez,cgatta,sescalera,ligual,petia}@cvc.uab.cat
[3] Institut Clinic del Torax, Hospital Clinic Barcelona,
Villarroel 170, 08036 Barcelona, Spain
27700vmy@comb.cat

Abstract. The Quantitative Coronary Angiography (QCA) is a methodology used to evaluate the arterial diseases and, in particular, the degree of stenosis. In this paper we propose AQCA, a fully automatic method for vessel segmentation based on graph cut theory. Vesselness, geodesic paths and a new multi-scale edgeness map are used to compute a globally optimal artery segmentation. We evaluate the method performance in a rigorous numerical way on two datasets. The method can detect an artery with precision $92.9 \pm 5\%$ and sensitivity $94.2 \pm 6\%$. The average absolute distance error between detected and ground truth centerline is 1.13 ± 0.11 pixels (about 0.27 ± 0.025mm) and the absolute relative error in the vessel caliber estimation is 2.93% with almost no bias. Moreover, the method can discriminate between arteries and catheter with an accuracy of 96.4%.

Keywords: Vessel segmentation, centerline extraction, QCA, GraphCut.

1 Introduction

The enhancement and segmentation of tubular structures and/or vessel-like structures is a prolific topic in the medical imaging research; many methods, exploiting photometric and structural properties of tubular structures have been proposed (see *e.g.* a complete review of recent methodologies for CTA segmentation in [11,1]). Nonetheless, in the case of vessel segmentation in angiography sequences, the problem is still very hard; highly reliable, fully automatic methods are not established yet [5]. Finally, accurate segmentation is still a hot topic and far from being solved, as demonstrated by the excellent scale selection method proposed in [10]. An extensive overview of different methods for vessel extraction is provided in [7]. Recently, an interesting approach to vessel segmentation has been proposed in [12], which fuses local features with local directional information; unfortunately, authors do not provide a quantitative evaluation of their method. Nonetheless, most works are based on local image analysis to extract vessels or employ an a-priori model to help vessel extraction. In contrast, graph cut (GC) technique is an optimal segmentation tool that combines local and contextual image information analysis by modeling relations between neighboring pixels. The GC algorithm [2,8] has been used in many computer vision problems and, in particular, it can

G. Fichtinger, A. Martel, and T. Peters (Eds.): MICCAI 2011, Part III, LNCS 6893, pp. 496–503, 2011.

be applied to binary-segmentation of images, obtaining a solution which corresponds to the global minimum of an energy function. The goodness of the solution depends on the suitability of the energy terms and their reliable computation.

In this paper, we use GC theory to obtain a globally optimal segmentation of the coronary tree in angiography images. Differently than other methods, we are interested in accurate detection of both the centerline and the vessel borders. To this aim, we propose a novel energy function tailored to the artery segmentation problem. The energy takes into account (1) the vessel local appearance, using any vesselness measure, (2) the local connectivity to other vessel regions, using geodesic paths and, (3) a measure of edgeness based on a novel multi-scale version of the adaptive Canny detector [4]. Moreover, we propose a machine learning based approach to automatically detect the catheter; providing a methodology that is far more general than the one in [12]. Finally, we propose two datasets, which allow the quantitative evaluation of the method in terms of ability to detect the artery (Precision and Sensitivity), errors in the centerline detection and caliber estimation, and ability to discriminate between arteries and catheters[1]. To the best of our knowledge, there is no such dataset available in the literature. In next section, we provide a brief introduction to GC, then detail our contributions.

2 Method

2.1 Graph Cut

Let us define $\mathcal{X} = (\mathbf{x}_1, ..., \mathbf{x}_i, ..., \mathbf{x}_{|\mathcal{P}|})$ the set of pixels for a given angiography gray-scale image I; $\mathcal{P} = (1, ..., i, ..., |\mathcal{P}|)$ the set of indexes of I; $\mathcal{N}$ the set of unordered pairs $\{i, j\}$ of neighboring pixels of $\mathcal{P}$ under a 4-(8-) neighborhood system, and $L = (L_1, ..., L_i, ..., L_{|\mathcal{P}|})$ a binary vector whose components L_i specify assignments to pixels $i \in \mathcal{P}$. Each L_i can be either "vess" or "back" indicating if it belongs to vessel or background, respectively. GC formulation [2] defines the cost function $E(L)$, which describes soft constraints imposed on boundary and region properties of L as $E(L) = U(L) + \lambda B(L)$. The unary term is denoted with $U(L) = \sum_{i \in \mathcal{P}} U_i(L_i)$, the boundary term with $B(L) = \sum_{\{i,j\} \in \mathcal{N}} B_{\{i,j\}}\ \Omega(L_i = L_j)$, where the characteristic function $\Omega(L_i = L_j)$ is 0 if $L_i \neq L_j$ and 1, otherwise. The unary term $U(L)$ is defined assuming that individual penalties for assigning pixel i to "vess" and "back", (i.e. $U_i(\text{"vess"})$ and $U_i(\text{"back"})$) are given by vessel and background models. The term $B(L)$ comprises the boundary properties of segmentation L. Any $B_{\{i,j\}} \geq 0$ should be interpreted as a penalty for a discontinuity between i and j. Finally, the coefficient $\lambda \in \mathbb{R}^+, \lambda \geq 0$ specifies the relative importance of the boundary term against the unary term. GC algorithm imposes hard constrains on the segmentation result by means of the definition of seed points where labels are predefined and can not be modified. The subsets $\mathcal{V} \subset \mathcal{P}, \mathcal{B} \subset \mathcal{P}, \mathcal{V} \cap \mathcal{B} = \emptyset$ denote the subsets of vessel and background seeds, respectively.

Boykov et al. [2] show how to efficiently compute the global minimum of $E(L)$ among all segmentations L satisfying the hard constraints $\forall i \in \mathcal{V}, L_i = \text{"}vess\text{"}, \forall i \in \mathcal{B}, L_i = \text{"}back\text{"}$, using a minimum cut algorithm on a certain graph defined by nodes

[1] The datasets and the evaluation methodology will be provided under request.

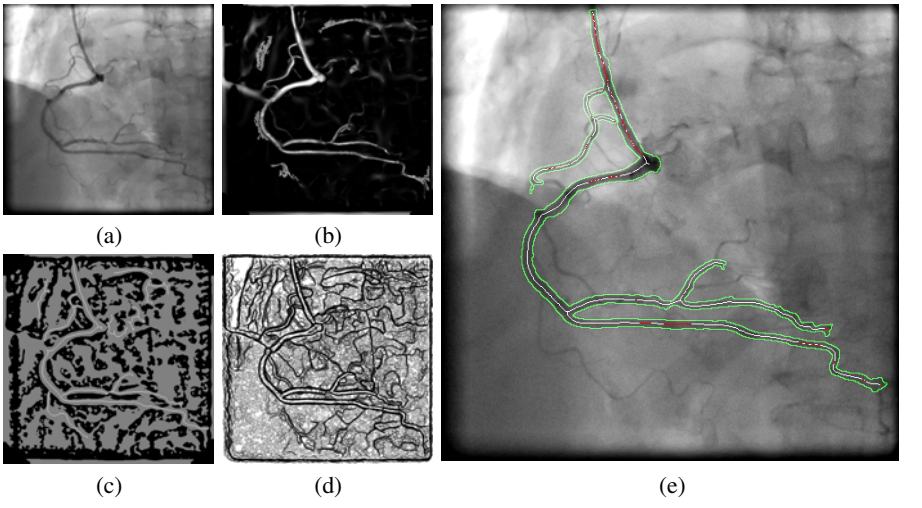


Fig. 1. AQCA approach: (a) Input image, (b) Unary potential: Vesselness, (c) Vessel-Background seeds (in white and black), (d) Boundary potential: Multi-scale edgeness map, (e) Final segmentation, centerline estimation, and catheter detection (in green, white, and red).

and edges being image pixels and pixel relations, respectively. We use the algorithm presented in [3] for computing the minimum cut.

2.2 Vessel Segmentation Algorithm

In this section, we describe in detail the seed initialization and the definition of unary and boundary potentials.

Seed initialization. In order to achieve a fully automatic methodology, we exploit the inherent structure of vessels to define vessel seeds based on valleys, and background seeds based on low probabilities of the vesselness image [6]. In particular, vessel seeds $\mathcal{V}$ correspond to those pixels corresponding to the highest responses on a multilocal valley detector with structure tensor, namely $\mathcal{S}_t$, as described in [9], $\mathcal{V} = \{i|\mathcal{S}_{t,i} > \Theta_v\}$, where $\mathcal{S}_{t,i}$ is the valley response at pixel i, and Θ_v is a sensitivity valley threshold. The background seeds $\mathcal{B}$ are the pixels corresponding to low probabilities in the vesselness image V, $\mathcal{B} = \{i|V_i \leq \Theta_b\}$, where Θ_b is a sensitivity vesselness threshold and the vesselness measure at pixel i, V_i, is computed as in [6]. Fig. 1(c) shows the selection of $\mathcal{V}$ and $\mathcal{B}$ seeds for the input image in Fig. 1(a).

Unary term. We define the vessel and background models using the vesselness map V. However, some vessel regions (especially those corresponding to bifurcations), can contain low vessel probability. To avoid this problem, we introduce the computation of geodesic paths among vessels seeds. We initialize the unary potentials at each pixel i as $U_i(\text{“vess”}) = -\ln(p(L_i = \text{“}vess\text{”}))$, $U_i(\text{“back”}) = -\ln(p(L_i = \text{“}back\text{”}))$. The probability of a pixel to be marked as “vess” is computed using the vesselness-geodesic measure VG, $p(L_i = \text{“}vess\text{”}) = VG(\mathbf{x}_i)$ and the opposite probability as

$p(L_i = \text{“}back\text{”}) = 1 - p(L_i = \text{“}vess\text{”})$. In particular, the map VG is computed for pixel i as the maximum between the vesselness value and the inverse of the geodesic distance,

$$VG_i = \max\left(V_{o,i}, \frac{\frac{1}{D_i+\mu(D)}}{\max(\frac{1}{D_i+\mu(D)})}\right),$$

where D and $\mu(D)$ correspond to the geodesic distance map and its mean, respectively. D is computed as explained in the following subsection.

Geodesic map. Given an arbitrary parameterized discrete path $\Gamma = \{i, ..., j\}$ defined by $|\Gamma| = R$ pixels, we define the geodesic distance D of Γ as,

$$D(\Gamma) = m(||\nabla I(\Gamma)||^2)\left(\sum_{i=1}^{R-1}\frac{||\nabla I_i||^2}{R}\right), \tag{1}$$

where the quantity $||\nabla I_i||$ is a finite difference approximation of the image gradient between points $(\mathbf{x}_i, \mathbf{x}_{i+1})$, and the function $m(\mathbf{z})$ represents the maximum variance of the R-dimensional vector $\mathbf{z}$, $m(\mathbf{z}) = \max_{i,j}|\mathbf{z}_i - \mathbf{z}_j|, i, j \in \Gamma$. The measure defined in Eq.(1) is normalized by the length of the path, allowing any path length to be considered. However, it is penalized by the maximum variance of image gradients within the path. The selected geodesic path is given by, $\Gamma^*_{\{i,j\}} = \text{argmin}_{\Gamma\in\mathcal{C}_{\{i,j\}}} D(\Gamma)$, and its distance measure is $d(i,j) = \min_{\Gamma\in\mathcal{C}_{\{i,j\}}} D(\Gamma)$, being $\mathcal{C}_{\{i,j\}}$ the set of all possible paths between points i and j using short-path algorithm. We proceed as follows: after computing the partial path $\Gamma^*_{\{i,j\}}$, in order to select the next path point $j^* \in \mathcal{G}_j$, where $\mathcal{G}_j$ is the set of 8-neighbor of j, we use the following criterion, $j^* = \text{argmin}_{\ell\in\mathcal{G}_j}(d(i,j) + \sigma_{\{j,\ell\}}D(\Gamma_{\{j,\ell\}}))$, where $\sigma_{\{j,\ell\}}$ is the variance between j and ℓ. Once the next point has been selected, we continue the path only if $d(i,j^*) < \Theta_d$, where $\Theta_d = 0.05$ is an XRay-dependent empirically set threshold. Since different geodesic maps can be found for different initialization pixels j, the geodesic map, for each pixel i, is computed as $D_i = \min_{j\in\Gamma} D(\Gamma_{\{i,j\}})$. These pixels are the centroids of a k-means clustering over the vessel seeds. An example of the VG map is shown in Fig. 1(b).

Boundary term. We propose an image-dependent multi-scale edgeness measure. First, we run the canny edge detector algorithm on the observed image at different threshold levels. Then, we compute the edge probability at each pixel by the linear average of the edge thresholds and for different scales as follows, $J^*_i = \min_j \frac{1}{n}\sum_{k=1}^{n} J_{p,\gamma_k,\sigma_j}$, where J_{p,γ_k,σ_j} is the binary edge map using the threshold $\gamma_k \in [0.02, 0.03, ..., 0.3]$ and scale $\sigma_j \in [0.5, 1, ..., 5]$ for pixel i. If pixel i is labeled as an edge pixel for most of the threshold levels at a significant scale, it has a high probability of being an edge pixel. The final boundary potential over the multi-scale edgeness map is computed as $B_{\{i,j\}} = J^*_i$. An example is shown in Fig. 1(d).

2.3 Postprocessing

We perform a post-filtering step consisting in keeping only the biggest connected component in the final segmentation. The main aim of this step is to get rid of possible false

positive (FP) regions that could be introduced by the seed initialization. An example of the final segmentation is shown in green in Fig. 1(e).

The centerline (CR) is extracted as follows: given the binary segmentation $L(\mathbf{x}_i)$, we compute its distance map $M(\mathbf{x}_i)$. Then, a non-maxima suppression is applied to find local maxima and a classic ridge transversal method is applied to connect the local maxima. The ridge transversal stops when it finds another centerline or it exits the segmented area. Fig. 1(e) shows an example of an extracted centerline in white.

Vessel caliber is estimated by applying a local Laplacian of Gaussian (LoG) filtering at CRs locations at different scales. The scale space computed using $\sigma^2 \text{LoG}(x, y; \sigma)$ has a minimum at $\sigma = w/2$, where w is the width of the ridge.

Catheter detection. By merely its appearance, the catheter is not easily distinguishable from arteries. This causes that the proposed segmentation method tends to segment the catheter as an artery. In order to detect the catheter, from each point of the centerline path we extract: (1) its position $\mathbf{x}$, (2) its curvature $K(\mathbf{x})$, (3) its angular direction $\alpha(\mathbf{x})$, and (4) its caliber $C(\mathbf{x})$. A classification is performed in a point wise way using a Bayesian classifier. Being $c = \{\text{“cat”}, \text{“vess”}\}$ the catheter or artery class, we modeled (1) $p(\mathbf{x}|c)$ using a Kernel Density Estimator with $\sigma_{\text{KDE}} = 15$pixels, (2) $p(\log(\epsilon + K(\mathbf{x}))|c)$ using a Gaussian Mixture Model[2] (10 Gaussians), (3-4) $p(\alpha(\mathbf{x})|c)$ and $p(C(\mathbf{x}|c))$ as two discrete histograms. In Fig. 1(e) an example of catheter detection is shown in red.

The time complexity of the algorithm most expensive part is $O(|\mathcal{P}| + \frac{|\mathcal{N}|}{|\mathcal{P}|} \log |\mathcal{P}|)$.

3 Validation

Material. We defined two datasets: DS1 and DS2. DS1 is formed by 20 images acquired with a single plane Philips INTEGRIS Allura Flat Detector, of RCAs, where three experts have blindly annotated the centerlines. The experts had to annotate the centerlines with different labels: “vess”: the arteries that potentially can present a clinical interest (with a caliber of, at least, 1mm); “don’t care”: all other arteries in the image, and “cat”: the catheter guide. DS2 is formed by 31 images from 27 patients, acquired with a SIEMENS Artis zee, of 10 RCAs, 10 LADs, and 11 Cxs. Two experts blindly segmented a total of 41 lesions (12 LADs, 13 Cxs and 16 RCAs) assisted by a semi-automatic method (QCA-CMS Version 6.0, MEVIS). The experts were asked to manually correct unsatisfactory segmentations.

Methods. We compare our proposed method (AQCA) against a classic ridge transversal centerline extraction method (RT), and the state of the art GC method. Furthermore, we also compare the obtained results with the inter-observer (IO) variability of the experts ground truth (GT). In the case of AQCA and GC, parameters Θ_v, Θ_b, and λ are tuned via cross-validation over DS1. Given N patients, the tuning is performed using a Leave One Patient Out (LOPO) methodology by maximizing $0.5P + 0.5S$ in order to provide a balance between precision (P) and sensitivity (S).

[2] The curvature distribution was far from being Gaussian, so we applied the logarithm to “Gaussianize” it.

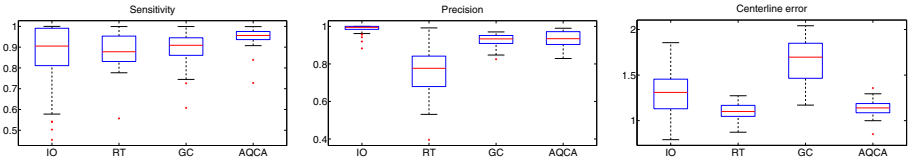


Fig. 2. Centerline evaluation results for DS1

Validation protocol. For the evaluation of centerline detection, we compute P, S, localization error ($\mathcal{E}_L$) on DS1. Additionally we evaluate the caliber estimation on DS2, measuring the mean signed relative error and mean absolute relative error. Finally, the catheter detection is evaluated in terms of P, S, and accuracy.

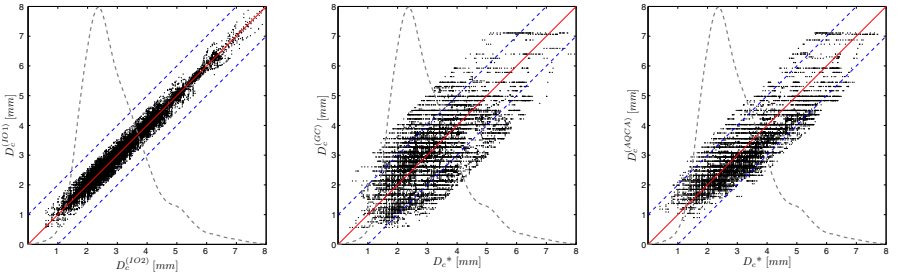
Centerline evaluation: To evaluate the CR detection, we computed P, S, and $\mathcal{E}_L$ considering Θ_c, which defines the maximal distance between the GT and the detected CR. To compute the S, we check for every CR point in the GT if there exists a detected CR point in a distance smaller than Θ_c; if this happens, this point is considered as a True Positive (TP). Similarly, P is computed by checking the detected points instead than the ones in the GT. Θ_c has been set to 5 pixels to allow large $\mathcal{E}_L$ errors.

Caliber evaluation: Vessel caliber evaluation was performed over DS2. We approximated two cubic splines to the borders annotated by the experts. Using these splines, we determined the CR and extracted the caliber for each point [10]. For each point in the GT CR, we localized the nearest point in the detected CR and evaluated the caliber estimation error using the Euclidean distance. We computed the signed error $\Delta D_c = D_c - D_c^*$, where D_c^* is the ground truth caliber, in millimeters. Finally, we defined the average absolute and signed relative errors $\frac{|\Delta D_c|}{D_c^*}$ and $\frac{\Delta D_c}{D_c^*}$, respectively.

4 Results

Figure 2 shows S, P, and $\mathcal{E}_L$ for the IO variability, the RT, GC, and AQCA for the dataset DS1[3]. The RT method has the lowest S and a very low P, while the $\mathcal{E}_L$ is very low; this confirms that the vesselness measure is well suited to accurately detect the CR, but it has the disadvantage to produce many FPs as confirmed by the low P. A basic GC approach increases both S and P while $\mathcal{E}_L$ is increased due to inaccurate border detection using a gray-level based boundary term. Our proposed method shows the highest S and P, and a $\mathcal{E}_L$ that is very close to RT, while actually detecting more vessel pixels than both RT and GC (higher S, less false negatives). It is also interesting to note that the proposed method has a lower P than the IO variability but higher S: this means that the proposed method still produces some FPs but tends to detect clinically relevant arteries in a way that is the "average" of the observers. Figure 3 shows scatter plots of the caliber estimation on dataset DS2 for the IO variability, the basic GC, and AQCA, respectively. The gray dashed curve shows the density of points w.r.t. the caliber. It can

[3] The results for both datasets are available at `www.cvc.uab.es/~ahernandez/AQCA.zip`.



(a) (b) (c)

Fig. 3. GT and estimated calibers in scatter plots for (a) IO variability, (b) GC, and (c) AQCA

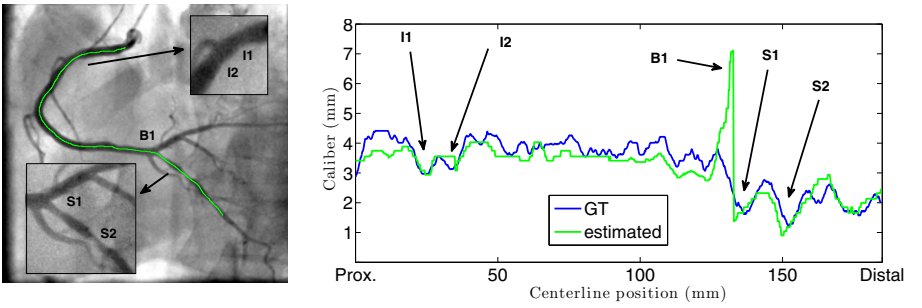


Fig. 4. Extracted centerline and GT and estimated caliber along the extracted centerline

be noticed that the two experts show a very high correlation. This is due to the fact that the segmentation has been performed thanks to a computer assisted method: if there is a bias in the measurement, it is equal for both experts. The basic GC method performs badly, as confirmed by the large average relative error of 6.04%, while the proposed method performs much better having an average relative error of 2.93%. It has to be noted that, while GC present a bias of -1.6%, our method presents almost no bias (-0.035%). Moreover, for the calibers which are more frequent in X-Ray angiography, our proposed method has a very high correlation with the ground truth data. Figure 4 shows an example of AQCA result: the estimated caliber from proximal to distal positions accurately follows the GT. At proximal position the method correctly detects two clinically irrelevant caliber variations (I1 and I2). Then, at the bifurcation B1, our method estimates the bifurcation "caliber" that is obviously not relevant, and can be easily detected as an outlier. Just after the bifurcation, the proposed method accurately measures the vessel caliber at the stenosis S1 and S2. Finally, we collected ground truth data (56406 samples) from DS1 and evaluated the catheter detection methodology on a leave-one-patient-out methodology, obtaining an average S of 70.9%, P of 90.1% and accuracy of 96.4%.

5 Conclusion

We presented AQCA, a novel segmentation method for X-Ray angiography images that takes into account vessel appearance, artery tree continuity, and borders appearance within the graph-cut theory. The algorithm has been tested on two new data sets.

Despite it has been tuned on DS1, it provided excellent results on DS2, showing the inherent robustness of the approach. While being applied on accurate QCA, the method could be profitably used as preprocessing for non-rigid multimodal registration algorithms. Moreover, it can be easily adapted to detect tubular structures in other kind of images. Future lines of research encompass the use of an high order potential to deal with irregularity at bifurcations and crossings; a supervised method to optimize the seed selection; an intelligent procedure to semantically tag the artery in order to obtain an automated QCA report for all relevant branches.

Acknowledgments. This work has been supported in part by the projects: La Marató de TV3 082131, TIN2009-14404-C02, and CONSOLIDER-INGENIO CSD 2007-00018. The work of C. Gatta is supported by a Beatriu de Pinos Fellowship.

References

1. Benmansour, F., Cohen, L.: Tubular structure segmentation based on minimal path method and anisotropic enhancement, vol. 92, pp. 192–210. Springer, Netherlands (2011)
2. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient n-d image segmentation, vol. 70, pp. 109–131. Kluwer Academic Publishers, Hingham (2006)
3. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, vol. 26, pp. 359–374. IEEE Computer Society, Los Alamitos (2001)
4. Candemir, S., Akgul, Y.: Adaptive regularization parameter for graph cut segmentation. In: Campilho, A., Kamel, M. (eds.) ICIAR 2010. LNCS, vol. 6111, pp. I:117–I:126. Springer, Heidelberg (2010)
5. Fallavollita, P., Cheriet, F.: Towards an automatic coronary artery segmentation algorithm. In: EMBS 2006, pp. 3037–3040. IEEE Computer Society, Los Alamitos (August 2006)
6. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A.C.F., Delp, S.L. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
7. Kirbas, C., Quek, F.: A review of vessel extraction techniques and algorithms, vol. 36, pp. 81–121. ACM, New York (2004)
8. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts, vol. 26, pp. 65–81. IEEE Computer Society, Los Alamitos (2004)
9. Lopez, A.M., Lumbreras, F., Serrat, J., Villanueva, J.J.: Evaluation of methods for ridge and valley detection, vol. 21, pp. 327–335. IEEE Computer Society, Los Alamitos (1999)
10. Mirzaalian, H., Hamarneh, G.: Vessel scale-selection using mrf optimization. In: CVPR, pp. 3273–3279. IEEE Computer Society, Los Alamitos (June 2010)
11. Schaap, M., van Walsum, T., Niessen, W.: Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms, vol. 13, pp. 701–714. Elsevier, Amsterdam (2009)
12. Schneider, M., Sundar, H.: Automatic global vessel segmentation and catheter removal using local geometry information and vector field integration. In: IEEE ISBI, pp. 45–48. IEEE Press, Los Alamitos (2010)