Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Vessel tree tracking in angiographic sequences

Dong Zhang
Shanhui Sun
Ziyan Wu
Bor-Jeng Chen
Terrence Chen

Vessel tree tracking in angiographic sequences

Dong Zhang,^a Shanhui Sun,^{a,*} Ziyang Wu,^b Bor-Jeng Chen,^a and Terrence Chen^a

^aSiemens Healthcare, Medical Imaging Technologies, Princeton, New Jersey, United States

^bSiemens Corporation, Corporate Technology, Princeton, New Jersey, United States

Abstract. We present a method to track vessels in angiography [contrast filled vessels in two-dimensional (2-D) x-ray fluoroscopy]. Finding correspondence of a vessel tree from consecutive angiogram frames provides significant value in computer-aided clinical applications such as fast vessel tree segmentation, three-dimensional (3-D) vessel topology reconstruction from corresponding centerlines, cardiac motion understanding, etc. However, establishing an accurate vessel tree correspondence (vessel tree tracking) is a nontrivial problem due to nonlinear periodic cardiac and breathing motion in 2-D views, foreshortening, false bifurcations due to 3-D to 2-D projection, occlusion from other anatomies, etc. The vessel tree is represented by BSpline curves. The control points of the BSpline curves are landmarks that are the tracking targets. Our method maximizes the appearance similarity while preserving the vessel structure. A directed acyclic graph (DAG) is employed to represent the appearance and shape structure of the vessel tree: nodes from the DAG encode the appearance of the vessel tree landmarks, and the edges encode the relative locations between landmarks. The vessel tree tracking problem turns into finding the most similar tree from the DAG in the next frame, and it is solved using an efficient dynamic programming algorithm. We performed evaluations on 62 x-ray angiography sequences (above 1000 frames). Experiment results show our algorithm is robust to these challenges and delivers better performance, compared to four existing methods. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.4.2.025001]

Keywords: vessel tree tracking; coronary tree tracking; angiography.

Paper 16193RR received Aug. 25, 2016; accepted for publication Mar. 21, 2017; published online Apr. 10, 2017.

1 Introduction

Accurate and automated tracking of vessel tree structures in angiographic sequences provides significant values in many computer-aided clinical applications in cardiac surgery planning and interventional procedures, for instance, quantitative analysis of vascular tree structure,^{1–3} percutaneous coronary intervention (PCI) guidance,^{4–6} heart dynamics learning from vessel motion,^{7,8} three-dimensional (3-D) vessel tree reconstruction^{9–11} from tree structure of motion, etc. As shown in Fig. 1, the vessel tree tracking problem for angiographic sequences is defined as follows: given the target vessel tree in the first frame of an angiography sequence either manually or automatically, the vessel tree structures in the remaining frames are automatically identified. To the best of our knowledge, there are very few prior literatures dealing with the vessel tree tracking problem. The vessel tree tracking problem is challenging due to large nonlinear periodic cardiac motion, vessel crossing, foreshortening, occlusion, and similar appearance along a vessel branch. Thus, traditional tracking methods developed in the computer vision domain cannot be directly applied. This work is among the first attempts to solve the problem robustly, and we validated it in a quantitative manner with 62 real clinical scenes.

Vessel tree tracking in angiographic sequences has many applications, and there are great clinical values if the vessel tree structures can be tracked robustly in two-dimensional (2-D) angiography. (a) We can learn dynamic vessel tree cardiac motion model from a tracked vessel tree. Once we correlate an electrocardiography (ECG) signal with these cardiac motions, we can predict vessel motion in certain cardiac phases

using ECG signal. Predicting vessel position from an ECG signal is clinically important in the PCI procedure. Usually, the PCI procedure is performed as the cardiologist injects radiopaque dyes into the coronary artery to guide the catheter. Since the radiopaque dyes are harmful, by tracking and predicting the vessels in angiographic sequences, it will reduce the injection of radiopaque dyes dramatically and have significant benefits for the patients. Note that breathing motion is included in the model learning process; however, breathing motion is a rigid transformation and can be easily compensated by detecting reference objects such as the guide-wire, which is visible in both the Cine stage and the fluoroscopic stage. (b) For 3-D vessel center-line reconstruction from 2-D views, a practical approach is to obtain the 2-D point correspondences, and 2-D vessel tracking plays an important role here. It is very difficult to match vessel tree directly since the structures look very similar, and tracking techniques can be applied to get the correspondences while the vessel structure constraints can be utilized. If the vessel tree can be robustly tracked in a rotational scan, one can do a motion compensated reconstruction of the entire vessel tree in 3-D. This can lead to a 3-D or four-dimensional angiography that has higher resolution than regular computed tomography (CT) angiography. (c) Via tracking, we can get vessel centerline segmentation as a by-product.

To overcome the above challenges, we proposed a method to model the vessel tracking problem for angiographic sequences as an optimization problem based on directed acyclic graph (DAG). The contributions of our method are fourfold. (1) To the best of our knowledge, this is the first method that can robustly track the vessels in 2-D angiographic sequences for both a single

*Address all correspondence to: Shanhui Sun, E-mail: sunshanhui@gmail.com

Fig. 1 Overview of the vessel tree tracking for angiographic sequences. The input (a) shows an angiographic sequence and the initialization (blue curves) in the first frame. Initialization could be performed by computer-aided method or manual annotation. Illustrations of tracking results (b): landmarks (red markers) and vessel tree structure (green curves).

branch and a tree structure. (2) We propose a formulation of the vessel tree tracking problem for angiographic sequences using DAG optimization, which can be solved efficiently by dynamic programming. (3) The tracking algorithm only requires historical and current frames as the input, and it can process the data at a speed of two frames per second, which is sufficient for most of the clinical applications mentioned above. (4) We have collected and annotated 62 angiographic sequences, consisting of about 1000 frames. We compared our method with four competing tracking methods [i.e., SNAKE,¹² Struck,¹³ tracking-learning-detection (TLD),¹⁴ and structure preserving object tracking (SPOT)¹⁵] qualitatively and quantitatively and showed significant improvements.

Following the introduction, we introduce some related works of vessel tree tracking for angiographic sequences in Sec. 2. We present details of our method in Sec. 3; qualitative and quantitative results on 62 clinical sequences in Sec. 4; and the conclusion in Sec. 5.

2 Related Works

Only a few papers that deal with tracking of 2-D vessels in the angiographic sequences have been published. None of the existing methods directly targets 2-D vessel tree tracking. In this section, we review relevant state-of-the-art tracking techniques in the fields of 2-D vessel branch tracking, 3-D vessel tracking, and general object tracking.

For 2-D vessel tracking, there are some related methods that have the potential to be applied to single vessel tracking (e.g., SNAKE)¹⁶ or vessel tree tracking (e.g., Junction landmark).¹⁷ Active contour models such as SNAKE^{12,16,18,19} and enhanced SNAKEs^{20,21} are very insightful methods for capturing the deformation of objects, and they minimize the internal energy (shape smoothness) and external energy (image appearance) to track objects. However, SNAKE-based methods can only track single curvatures and must have an initial proximity to the target. Heibel et al.²² proposed modeling BSpline-based curvatures as a Markov random field for interventional-wire tracking, which has a very elegant formulation. In Refs. 23 and 24, the authors proposed new methods to track a single curvilinear structure (guide-wire). These methods can only track single curvatures

that do not have as much deformations as vessel trees have. A combined vessel extraction and tracking method to track single vessels was proposed in Ref. 18, which can be applied for 3-D vessel reconstruction; however, this method only handles a single vessel and highly relies on the vessel extraction. Junction landmark-based tracking methods such as Refs. 17 and 25 can track bifurcations of vessel trees effectively and are also very useful in some clinical practices. However, these methods are not suitable for single vessel branch tracking, and they are unable to track curved structures. They can only track bifurcations, so a long vessel is very difficult to be tracked if it does not have enough bifurcations. Additionally, they do not have a mechanism to effectively interpolate the structures between the bifurcations, and therefore, cannot track the vessel body well. Optical flow-based methods^{26,27} were also employed to track the principal coronary arteries; however, in angiographic sequences, the frames do not have enough contrast, and the thin vessel structures (not just principal coronary arteries) are difficult to capture by optical flow, especially when they are undergoing large motion.

In contrast, there are many 3-D vessel tracking methods.^{28,29} They cannot be applied to 2-D vessel tracking since 3-D voxel features are much richer and more discriminative, but these methods are very insightful. In Refs. 28 and 29, a registration framework in which the vessel centerline model deforms in an external potential field of a multiscale analysis response map was formulated. A two-step vessel extraction and nonrigid template matching method was proposed in Ref. 30 with 3-D cardiac CT data. In Ref. 31, the coronary trees were first extracted and related by a point-based matching algorithm with feature landmarks constraint and then applied to match the 3-D extracted centerlines.

For general object tracking across image sequences, optical flow³² plays an important role. Optical flow calculation can be very fast and integrated into the real-time tracking system.³³ It can be integrated³⁴ into the active contour tracking framework. Furthermore, Gaussian mixture models can be used to improve the optical flow-based tracking.³⁵ Another important direction for object tracking across image sequences is feature design and fusion. A collaborative feature fusion model using both holistic templates and local representations is proposed.³⁶ In addition, sparse prototypes³⁷ are employed to exploit both classic principal

component analysis algorithms with recent sparse representation schemes for learning effective appearance models. Intuitively, structural cues¹⁵ of the objects are also important for object tracking. More recently, many multiple object tracking techniques^{14,15} have been introduced; they are potentially very useful for vessel tracking since vessel trees can be models as deformable multiple object structures that consist of landmarks. The state-of-the-art trackers have never been evaluated for vessel tree tracking problems; such trackers include the SPOT,¹⁵ which uses an online structural support vector machine (SVM) incorporating structural constraints from multiple objects (landmarks), the Struck tracker¹³ using a kernelized structural SVM, and the TLD tracker.¹⁴ The reconstruction of tree-like³⁸ and curvilinear^{39,40} structures are also related to our work.

3 Method

3.1 Problem Formulation

Figure 2 shows components and problem formulation of our proposed method. In our work, “vessel tree” is a collection of vessel branches connected without simple cycles. Note that a single vessel branch is considered a degenerated case of the vessel tree. We model the vessel tree using BSpline curves controlled by the landmarks [Fig. 2(a), left image]. The vessel tree in frame f is represented mathematically by a directed tree $\mathcal{T}_f = (\mathcal{N}_f, \mathcal{E}_f)$, where $\mathcal{N}_f = \{n_f^i | i = 1, \dots, L\}$ is the set of landmarks, L is the number of landmarks, and $\mathcal{E}_f = \{(i, j) | n_f^j = \text{Parent}(n_f^i)\}$ is the set of directed edges depicting child–parent relationship for the landmarks.

Given the vessel tree $\mathcal{T}_{f-1}$ in frame $f - 1$, the tracking problem becomes finding the vessel tree $\mathcal{T}_f$ in frame f . A directed acyclic weighted graph $G_f = (V_f, E_f, W_f)$ is built to solve this problem, where V_f , E_f , and W_f denote the “hypothesis nodes,”

“hypothesis edges,” and “hypothesis weights,” respectively. N hypothesis nodes $V_f^i = \{v_f^{k,i} | k = 1, \dots, N\}$ [Fig. 2(b)] are generated for each landmark n_{f-1}^i . Hypothesis edges connect hypothesis nodes of a landmark to all the hypothesis nodes of its parent landmark [Figs. 2(c) and 2(d)]. Hypothesis weights consist of unary scores and binary scores. Unary scores are assigned to the hypothesis nodes, indicating the appearance similarity between the hypothesis nodes and their corresponding landmark in frame $f - 1$ (Sec. 3.3). Binary scores are assigned to the hypothesis edges (Sec. 3.3), enforcing structural constraints on child–parent landmarks with spatial relationship between adjacent landmarks. Let $U_f^i(v_f^{k,i})$ be the unary score assigned to the $v_f^{k,i}$ and $B_f(v_f^{k,i}, v_f^{l,j})$ be the binary score assigned to the edge between hypothesis node $v_f^{k,i}$ and hypothesis node $v_f^{l,j}$. The objective is to select one hypothesis node for each landmark forming a tree such that the total combined unary–binary scores of this tree are maximized. Given an arbitrary tree $\mathcal{T}_f^S = (V_f^S, E_f^S)$ in which $S = \{v_f^{s,i} | s_i \in [1, N], i = 1, \dots, L\}$ is the set of selected hypothesis nodes, the energy function is

$$\mathcal{D}_f(S) = \left[\sum_{i=1, \dots, L} U_f^i(v_f^{s_i,i}) + \lambda_B \sum_{(i,j) \in \mathcal{E}_{f-1}} B_f(v_f^{s_i,i}, v_f^{s_j,j}) \right], \quad (1)$$

where $v_f^{s_i,i}$ and $v_f^{s_j,j}$ are the hypotheses for landmarks n_f^i and n_f^j , respectively, and λ_B is a weighting parameter. Equation (1) is a combination of the unary weights (for the nodes) and binary weights (for the edges). For an optimal tracking of the vessel tree to the next frame, we want to maximize the combined unary and binary weights. The tracked vessel tree (optimal solution) is

Fig. 2 Components of the proposed method. (a) Two consecutive frames containing the vessel of interest. The vessel landmarks and center-lines are shown on the left image. (b) Landmark hypotheses. Hypotheses for different landmarks are marked with different colors. (c) Hypothesis edges connecting landmark hypotheses. (d) Representation of the hypothesis graph. (e) Tracked vessel tree in frame f .

$$T_f^* = \arg \max_{T_f^s} \mathcal{D}_f. \quad (2)$$

3.2 Hypotheses Generation

In this section, we describe how we automatically generate the hypothesis nodes for the landmarks in the subsequent image. We combine a local intensity similarity constraint, a Frangi vesselness constraint, and a Gaussian location constraint to generate the hypothesis nodes. The ideas behind these constraints are: in the new frame, vessel landmarks can only occur at vessel region (Frangi vesselness constraint) with similar intensity appearance (intensity similarity: template matching) as on the previous frame; in addition, we assume a prior Gaussian motion model, and we thus force a Gaussian locality constraint. Hypotheses for landmark n_f^i in frame f are generated by sampling from a likelihood map U_f^i . We model U_f^i as a fusion of local intensity similarity C_f^i , Frangi vesselness response R_f ,⁴¹ and location constraint Q_f^i

$$U_f^i = \lambda^C \cdot C_f^i + \lambda^R \cdot R_f + \lambda^Q \cdot Q_f^i, \quad (3)$$

where λ^C , λ^R , and λ^Q are the fusion weights, respectively. The intensity-based similarity map C_f^i in frame f is generated using normalized cross correlation with a patch around n_{f-1}^i . To constrain location for temporal smoothness, we model Q_f^i as a zero mean Gaussian at n_{f-1}^i . We use both global and local (eight-neighborhood) maxima in U_f^i to generate the hypotheses. Figure 3 shows the score map fusion and hypotheses generation for a landmark. In the score map, since all the pixels near the global maximum have large values, the correct hypothesis can be in these neighboring pixels rather than the single global maximum, with the consideration of the binary weights. Therefore, we randomly sample some hypotheses near the global maximum from the score map.

3.3 Unary and Binary Scores

The unary score assigned $v_f^{k,i}$ is the value sampled from score map U_f^i [Eq. (3)]. The binary score is computed as follows. For two adjacent landmarks n_{f-1}^i and n_{f-1}^j in frame $f-1$, let vector $\mathbf{n}_{f-1}^{i \rightarrow j}$ be their relative location, and vector $\mathbf{v}_f^{k,i \rightarrow l,j}$ be the relative location of their corresponding hypotheses $v_f^{k,i}$ and $v_f^{l,j}$ in frame f . The binary score for edge between $v_f^{k,i}$ and $v_f^{l,j}$ is defined as

$$B_f(v_f^{k,i}, v_f^{l,j}) = - \frac{\frac{\|\mathbf{v}_f^{k,i \rightarrow l,j}\|_2 - \|\mathbf{n}_{f-1}^{i \rightarrow j}\|_2}{\|\mathbf{n}_{f-1}^{i \rightarrow j}\|_2}}{\frac{|\angle(\mathbf{v}_f^{k,i \rightarrow l,j}, \mathbf{n}_{f-1}^{i \rightarrow j})|}{\pi}}, \quad (4)$$

where $\|\cdot\|$ is L2 norm and $\angle(\cdot, \cdot)$ is the angle ($\in [0, \pi]$) between two vectors. For two adjacent hypotheses, Eq. (4) assigns larger value to the edge if their relative distance and direction are similar to the corresponding landmarks.

3.4 Optimal Solution

We solve Eq. (2) with dynamic programming similar to Ref. 42. Let $d_f(k, i)$ be the maximized value for hypothesis node $v_f^{k,i}$ in frame f . According to Eq. (1), the recursive function of the dynamic programming is

$$d_f(k, i) = U_f^i(v_f^{k,i}) + \sum_{(l,j) \in \mathcal{E}_{f-1}} \max_{l=1, \dots, N} [\lambda_B B_f(v_f^{l,j}, v_f^{k,i}) + d_f(l, j)]. \quad (5)$$

$d_f(k, i)$ can be obtained recursively. The computational complexity of the algorithm is $O(LN)$. The optimal solution is then obtained by backtracking with a complexity $O(L)$. Therefore, the overall time complexity is linear with respect to the number of nodes in the graph. To optimize Eq. (1), we make sure that the selected hypotheses look very similar to the landmarks in the previous frame and that they satisfy the spatial constraints between the landmarks.

4 Experiments

4.1 Datasets

We have collected two groups of datasets to evaluate the proposed method. All the angiographic sequences were annotated frame by frame by an expert.

1. The vessel tree dataset is collected for evaluation of the complex vessel tree tracking task, which includes 15 2-D angiographic sequences. Each angiographic sequence contains a vessel tree with at least two branches, which are visible in 23 frames, on average. The frame size is 512×512 pixels, and the average pixel size is 0.29 mm (min: 0.216 mm and max:

Fig. 3 Score map fusion U_f^i and hypotheses generation for a landmark. (a) A patch around n_{f-1}^i . (b) Local intensity similarity C_f^i , Frangi vesselness response R_f ,⁴¹ and location constraint Q_f^i for n_f^i . (c) Score map fusion with examples of landmark hypotheses. Global and local maxima are marked with a green or red “x” respectively.

0.368 mm). The frame rates for these angiographic sequences are all 8 frames/s.

2. The single branch dataset is collected for evaluation of robust vessel branch tracking task in which 47 2-D angiographic sequences obtained during PCI procedures are included. The target vessel branches on which PCI is going to be performed are visible in 24 frames on average. The frame size is either 512×512 or 600×600 pixels with average pixel size of 0.25 mm (min: 0.184 mm and max: 0.368 mm). The frame rates for these angiographic sequences are all 8 frames/s.

Annotations: We asked an expert to annotate landmarks for vessel tree structures in each frame, and at least five landmarks per vessel branch must be given. We interpolate the landmark locations by BSpline and use these points as the ground truth.

4.2 Experimental Setup

The vessel tree in the first frame can be manually annotated or automatically detected.^{43–45} In our experiments, we use the landmark annotations in the first frame of each sequence as an initialization and track them throughout the rest of the frames. Typically, we annotate around five landmarks for each vessel branch. Note that the first frame of each sequence is excluded from evaluations.

Evaluation metric: Similar to Ref. 24, a quantitative error index is utilized

$$d_e = \frac{d(G, R) + d(R, G)}{2.0}, \quad (6)$$

where $d(X, Y)$ is an average shortest distance of the points in set X to points in set Y , R is the tracking result, and G is the ground truth for the points on the vessel tree after interpolation between landmarks.

4.3 Results

We compared our method to four state-of-the-art tracking methods: SNAKE-based curvature tracking method,¹² Struck,¹³ TLD,¹⁴ and SPOT.¹⁵ The reasons we selected these four methods as comparisons are as follows. SNAKE¹² is good at tracking deformable structures (e.g., curves, organs) and it is intuitive to employ it for 2-D vessel tracking. Since the vessel structure can be models by landmarks, it is natural to consider the problem as a multiple object tracking problem, and we employ Struck¹³ as one of the best methods⁴⁶ for this task. We also want to see the performance if we track each landmark separately, so we employ TLD¹⁴ as one of the best single object tracking methods as a comparison. As we mentioned, the structural cues of the vessel trees are important for accurate tracking; we compare with SPOT¹⁵ since it takes advantage of the structural cues explicitly to improve the tracking performance. Tracking performances averaged over all sequences in each dataset are summarized in Tables 1 and 2 for all approaches. A student's t -test at a significance level of 0.05 indicates that our tracking algorithm significantly outperformed compared methods (p -values $\ll 0.001$).

Figure 4 shows qualitative results for all evaluated methods, and we show detailed results for the proposed method in Fig. 5.

Table 1 Vessel tree tracking results on the vessel tree dataset. “Ours,” the proposed method; “w/o shape,” the proposed method without enforcing structural constraints; Struck;¹³ TLD;¹⁴ SNAKE;¹² and SPOT.¹⁵ Mean and standard deviation (Std.) are given for d_e .

Method	d_e			
	Vessel tree dataset			
	Mean (mm)	Std. (mm)	Mean (pixels)	Std. (pixels)
Ours	1.57	1.04	7.82	5.33
w/o shape	4.77	1.66	24.15	8.93
Struck ¹³	2.24	0.81	11.26	4.14
TLD ¹⁴	2.76	1.06	14.05	5.74
SNAKE ¹²	16.83	7.58	83.41	34.79
SPOT ¹⁵	22.35	27.61	106.02	128.69

Table 2 Vessel tree tracking results on the single branch dataset. “Ours,” the proposed method; “w/o shape,” proposed method without enforcing structural constraints; Struck;¹³ TLD;¹⁴ SNAKE;¹² and SPOT.¹⁵ Mean and standard deviation (Std.) are given for d_e .

Method	d_e			
	Single branch dataset			
	Mean (mm)	Std. (mm)	Mean (pixels)	Std. (pixels)
Ours	0.85	0.40	4.77	2.11
w/o shape	4.79	2.20	28.25	16.25
Struck ¹³	2.28	1.33	13.22	8.10
TLD ¹⁴	3.05	2.31	18.17	15.55
SNAKE ¹²	16.71	9.10	97.78	56.87
SPOT ¹⁵	39.82	17.33	224.94	93.09

4.4 Discussions

From the results in Tables 1 and 2, we can see that the proposed method outperforms the state-of-the-art tracking methods with a large margin. In Fig. 4, our method gives visually very similar results compared to the ground truth. Note that, by incorporating structural constraints, the tracking performance is significantly improved, which can be seen from the comparisons between “ours” and “w/o shape” in Fig. 4. For single branch tracking (last two columns in Fig. 4), the performances with structural constraints are slightly better, and, for the vessel tree tracking (first three columns in Fig. 4), it performs significantly better than without the constraints, which is very intuitive since vessel trees have more structural information than a single branch. Struck¹³ performs the best in the comparisons; it can roughly track the vessel trees but cannot adjust to the deformation accurately; however, the spatial relationships between the landmarks

Fig. 4 Examples of vessel tree tracking results on five different angiographic sequences. Row 1 is the original angiographic sequence image. Row 2 is the ground truth. Row 3 is our results. Row 4 is the results of the proposed method without enforcing the structural constraint. Rows 5 to 8 are the results using Struck,¹³ TLD,¹⁴ SNAKE,¹² and SPOT,¹⁵ respectively. Angiographic sequences 1 to 3 are from the vessel tree dataset while the other two are from the single branch dataset.

Fig. 5 Detailed results for sampled frames of the proposed method on four angiographic sequences in the datasets. Angiographic sequences 1 and 2 are from the vessel tree dataset. Angiographic sequences 3 and 4 are from the single branch dataset. The red curves represent the initialization of the vessel (tree), and green curves are the tracking results.

are not tracked very well. This is intuitive since Struct¹³ is not good at modeling the structural constraints but is rather good at multiple object tracking. For TLD,¹⁴ the performance differs a lot from landmark to landmark. Since it is a single object tracking method, for some of the landmarks that are easy for tracking,

TLD can perform well; however, for some harder landmarks, it totally failed. This phenomenon is straightforward since the tracking for each landmark is independent. SNAKE¹² cannot track the vessels correctly due to the low contrast and rapid motion of the vessel trees. SPOT¹⁵ can track the vessel trees with

relatively small error since it also takes advantage of the structural information; however, it failed in the single branch tracking since there are less structural constraints for such branches.

4.5 Limitations

There are two limitations of the proposed method. (1) The speed is 2 frames/s, which prohibits it from some real-time tracking applications (e.g., PCI). It can be sped up by a graphics processing unit (GPU) implementation. (2) Currently, it is using the manual annotation of the first frame as an initialization for tracking, and fully automatic initialization^{43–45} would be highly preferable in some of the clinical applications. We will improve the system by incorporating these techniques.

4.6 Implementation Details

The experiments were performed on a Windows PC with a four-core 3.7 GHz Xeon central processing unit (CPU). The number of hypotheses is a trade-off between accuracy and efficiency, and we generated $N = 100$ hypotheses for each landmark. For the parameters in Eqs. (1) and (3), we set $\lambda^C = \lambda^R = \lambda^Q = \lambda_B = 1$. We used $w = 151$ (pixels) as the window size, and $0.5w$ as the standard deviation of the Gaussian functions. We performed some experiments to choose the parameters mentioned above. We found that the results are not very sensitive to the hypothesis node sampling process, and the parameters in Eqs. (1) and (3) do not have a big impact on the results if enough hypotheses are generated.

For the comparisons, the binary scores in the DAG are set to 0 for w/o shape experiments. For fair comparisons, we adapted the four methods (TLD, SNAKE, SPOT, Struck, all of them are open source) to our vessel tracking problem.

The proposed method runs at a speed of 2 frames/s. Currently the speed is enough for our applications, and if necessary, the frame rate can be significantly improved by a GPU implementation. The computational bottleneck is the hypotheses generation, which takes <95% of the CPU time. The hypothesis generation is a highly parallelable process since the hypotheses are independent from each other, which leads a potential 20 to 100 times faster GPU implementation, which is 40 to 200 frames/s.

5 Conclusion

In this paper, we proposed a method to track vessel tree and single vessel branch in 2-D x-ray angiographic sequences. By fully utilizing appearance and structural constraints and finding a globally optimal solution for the proposed energy function, our method greatly outperformed four state-of-the-art tracking algorithms. Not only are we among the first to attempt to solve this problem, our result is robust and validated using clinical datasets. Many other clinical relevant applications can potentially benefit from our result. The underlying technologies from our methods are not limited to vessel tracking and can easily be generalized to track other wire-like structures (e.g., guide-wire) with different energy functions.

Disclosures

Dr. Shanhui Sun and Terrence Chen are full time employees in Siemens Healthcare. Siemens Healthcare is a medical technology company and manufactures angiography and interventional x-ray systems. The image data used in this manuscript are generated from some of such systems. All other authors declare no conflicts of interest.

Disclaimer

The outlined concepts are not commercially available. Due to regulatory reasons, their future availability cannot be guaranteed.

References

1. G. Agam, S. G. Armato, and C. Wu, "Vessel tree reconstruction in thoracic ct scans with application to nodule detection," *IEEE Trans. Med. Imaging* **24**(4), 486–499 (2005).
2. X.-C. Pan et al., "3D liver vessel reconstruction from CT images," in *Proc. of the 2014 Second Int. Conf. on 3D Vision*, Vol. 2, pp. 75–81, IEEE Computer Society (2014).
3. K.-S. Lin et al., "Retinal vascular tree reconstruction with anatomical realism," *IEEE Trans. Biomed. Eng.* **59**(12), 3337–3347 (2012).
4. F. W. Mohr et al., "Coronary artery bypass graft surgery versus percutaneous coronary intervention in patients with three-vessel disease and left main coronary disease: 5-year follow-up of the randomised, clinical syntax trial," *Lancet* **381**(9867), 629–638 (2013).
5. P. B. Shah, "Management of coronary chronic total occlusion," *Circulation* **123**, 1780–1784 (2011).
6. W. J. Dewilde et al., "Use of clopidogrel with or without aspirin in patients taking oral anticoagulant therapy and undergoing percutaneous coronary intervention: an open-label, randomised, controlled trial," *Lancet* **381**(9872), 1107–1115 (2013).
7. G. Pennati et al., "Computational fluid dynamics models and congenital heart diseases," *Front. Pediatr.* **1**, 4 (2013).
8. V. Mihalef, "Patient-specific modelling of whole heart anatomy, dynamics and haemodynamics from four-dimensional cardiac ct images," *Interface Focus* **1**(3), 286–296 (2011).
9. F. Li et al., "Segmentation and reconstruction of cerebral vessels from 3D rotational angiography for AVM embolization planning," in *36th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC '14)*, pp. 5522–5525, IEEE (2014).
10. N. Baka et al., "Statistical coronary motion models for 2D+ t/3D registration of x-ray coronary angiography and CTA," *Med. Image Anal.* **17**(6), 698–709 (2013).
11. S. B. Kutty et al., "A review of 3D reconstruction of coronary arteries based on the co-registration of IVUS and coronary angiogram," in *Int. Conf. on Computer Assisted System in Health (CASH '14)*, pp. 1–5, IEEE (2014).
12. S. Zheng and D. Weirong, "Analysis of coronary arterial dynamics from x-ray angiographic sequences," in *Second Int. Symp. on Computational Intelligence and Design*, Vol. 1, pp. 201–204 (2009).
13. S. Hare, A. Saffari, and P. H. Torr, "Struck: structured output tracking with kernels," in *IEEE Int. Conf. on Computer Vision (ICCV '11)*, pp. 263–270, IEEE (2011).
14. Z. Kalal, K. Mikolajczyk, and J. Matas, "Tracking-learning-detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(7), 1409–1422 (2012).
15. L. Zhang and L. van der Maaten, "Structure preserving object tracking," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '13)*, pp. 1838–1845, IEEE (2013).
16. M.-P. Dubuisson-Jolly, C.-C. Liang, and A. Gupta, "Optimal polyline tracking for artery motion compensation in coronary angiography," in *Sixth Int. Conf. on Computer Vision (ICCV '11)*, pp. 414–419, IEEE (1998).
17. W. Yumei, "Tracking of coronary arteries in angiogram sequence by structural matching of junctions" Master's Thesis, National University of Singapore (2011).
18. I. Nwogu and L. Lorigo, "Fast temporal tracking and 3D reconstruction of a single coronary vessel," in *IEEE Int. Conf. on Image Processing (ICIP '07)*, Vol. 5, p. V–537 (2007).
19. B. C. Tom, S. N. Efstratiadis, and A. Katsaggelos, "Motion estimation of skeletonized angiographic images using elastic registration," *IEEE Trans. Med. Imaging* **13**(3), 450–460 (1994).
20. S. Zheng and Y. Daoyin, "Tracking vessels in x-ray angiogram sequences based on deformable model," *Iran. J. Electr. Comput. Eng.* **4**(2), 137 (2005).
21. H. Zhu and M. H. Friedman, "Vessel tracking by template string in angiography," *Proc. SPIE* **4549**, 29 (2001).

22. H. Heibel et al., "Interventional tool tracking using discrete optimization," *IEEE Trans. Med. Imaging* **32**(3), 544–555 (2013).
 23. E. Cheng et al., "Curvilinear structure tracking by low rank tensor approximation with model propagation," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 3057–3064 (2013).
 24. P. Wang et al., "Robust guidewire tracking in fluoroscopy," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 691–698, IEEE (2009).
 25. C. Bellemare and J. Meunier, "Structural method for tracking coronary arteries in coronary cineangiograms," in *First Canadian Conf. on Computer and Robot Vision*, pp. 378–384, IEEE (2004).
 26. D. Decarlo and D. Metaxas, "Optical flow constraints on deformable models with applications to face tracking," *Int. J. Comput. Vision* **38**(2), 99–127 (2000).
 27. P. Fallavollita and F. Cheriet, "Robust coronary artery tracking from fluoroscopic image sequences," in *Int. Conf. Image Analysis and Recognition*, pp. 889–898, Springer (2007).
 28. G. Shechter et al., "Three-dimensional motion tracking of coronary arteries in biplane cineangiograms," *IEEE Trans. Med. Imaging* **22**(4), 493–503 (2003).
 29. G. Shechter et al., "Temporal tracking of 3D coronary arteries in projection angiograms," *Proc. SPIE* **4684**, 612 (2002).
 30. D. P. Zhang et al., "Coronary artery motion modeling from 3d cardiac ct sequences using template matching and graph search," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pp. 1053–1056, IEEE (2010).
 31. G. Yang et al., "Characterization of 3-D coronary tree motion from msct angiography," *IEEE Trans. Inf. Technol. Biomed.* **14**(1), 101–106 (2010).
 32. Y. Mae et al., "Object tracking in cluttered background based on optical flow and edges," in *Proc. of the 13th Int. Conf. on Pattern Recognition*, Vol. 1, pp. 196–200, IEEE (1996).
 33. J. Shin et al., "Optical flow-based real-time object tracking using non-prior training active feature model," *Real-Time Imaging* **11**(3), 204–218 (2005).
 34. I. Mikic, S. Krucinski, and J. D. Thomas, "Segmentation and tracking in echocardiographic sequences: active contours guided by optical flow estimates," *IEEE Trans. Med. Imaging* **17**(2), 274–284 (1998).
 35. A. K. Chauhan and P. Krishan, "Moving object tracking using gaussian mixture model and optical flow," *Int. J. Adv. Res. Comput. Sci. Software Eng.* **3**(4), 243–246 (2013).
 36. W. Zhong, H. Lu, and M.-H. Yang, "Robust object tracking via sparsity-based collaborative model," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 1838–1845, IEEE (2012).
 37. D. Wang, H. Lu, and M.-H. Yang, "Online object tracking with sparse prototypes," *IEEE Trans. Image Process.* **22**(1), 314–325 (2013).
 38. E. Türetken, F. Benmansour, and P. Fua, "Automated reconstruction of tree structures using path classifiers and mixed integer programming," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 566–573, IEEE (2012).
 39. M. El Saban and B. Manjunath, "Tracking curvilinear structures using active contours and application to microtubule videos," Technical Report, http://vision.ece.ucsb.edu/publications/saban_icip2005.pdf (2005).
 40. N. Honnorat, "Curvilinear structures segmentation and tracking in interventional imaging," PhD thesis, Ecole Centrale Paris (2013).
 41. A. F. Frangi et al., "Multiscale vessel enhancement filtering," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '98)*, pp. 130–137 (1998).
 42. Y. Yang and D. Ramanan, "Articulated pose estimation with flexible mixtures-of-parts," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 1385–1392 (2011).
 43. S.-Y. Sun et al., "Model-guided extraction of coronary vessel structures in 2D x-ray angiograms," in *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI '14)* (2014).
 44. D. Lesage et al., "A review of 3D vessel lumen segmentation techniques: models, features and extraction schemes," *Med. Image Anal.* **13**(6), 819–845 (2009).
 45. T. Behrens, K. Rohr, and H. S. Stiehl, "Segmentation of tubular structures in 3D images using a combination of the Hough transform and a Kalman filter," in *Joint Pattern Recognition Symp.*, pp. 406–413 (2001).
 46. A. W. Smeulders et al., "Visual tracking: an experimental survey," *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(7), 1442–1468 (2014).
- Dong Zhang** received his BS degree in automation and his PhD in control science and engineering from Zhejiang University, China, in 2007 and 2013, respectively, and his PhD in computer science from the University of Central Florida in 2016. He is a postdoc associate at the Center for Research in Computer Vision, UCF. His research interests are in computer vision and medical imaging.
- Shanhui Sun** received his PhD in electrical and computer engineering (ECE) from the University of Iowa. He is currently a staff scientist at Siemens Healthcare, Medical Imaging Technologies, Princeton, New Jersey, USA. Previous affiliation name was Siemens Corporation, Corporate Technology, and Siemens Corporate Research.
- Ziyan Wu** received his BS degree in electrical engineering and automation and his MS degree in measurement technology and instruments from Beihang University, China, in 2006 and 2009, respectively, and his PhD in computer and systems engineering from Rensselaer Polytechnic Institute, Troy, New York, USA, in 2014. He is a research scientist in the Vision Technologies and Solutions Group, Siemens Corporate Technology, Princeton, New Jersey, USA.
- Bor-Jeng Chen** received his BS and MS degrees in electrical engineering from the National Taiwan University in 2009 and 2012, respectively. He is currently pursuing his PhD in electrical engineering at Viterbi School of Engineering, University of Southern California (USC). His research interests include visual tracking and object recognition in imagery.
- Terrence Chen** received his PhD in computer science from the University of Illinois at Urbana Champaign in 2006. He is the head of Vision Technologies and Solutions Research Group, Siemens Corporate Research, Princeton, New Jersey, USA. He is a senior member of IEEE. His expertise is in computer vision, medical imaging, pattern recognition, and machine learning, and he has published more than 30 papers in leading journals and conferences.