Globally-Optimal Anatomical Tree Extraction from 3D Medical Images Using Pictorial Structures and Minimal Paths

Zahra Mirikharaji$^{(\boxtimes)}$, Mengliu Zhao, and Ghassan Hamarneh

Medical Image Analysis Lab, Simon Fraser University, Burnaby, Canada
{zmirikha,hamarneh}@sfu.ca

Abstract. Extracting centerlines of anatomical trees (e.g., vasculature and airways) from 3D medical images is a crucial preliminary step for various medical applications. We propose an automatic tree extraction method that leverages prior knowledge of tree topology and geometry and ensures globally-optimal solutions. We define a pictorial structure with a corresponding cost function to detect tree bifurcations in anatomical trees with predefined topology. The tree bifurcations are encoded as nodes in the pictorial structure and are associated with an artificial neural network (ANN) based unary term. The geometrical (direction and length) statistics of tree branches are learned from a training set and encoded as geometrical priors for regularizing the pictorial structure edges. Finally, detected bifurcations as well as the ANN tubularity scores, are leveraged to trace globally optimal minimal paths along 3D tree centrelines. Our method outperforms competing state-of-the-art when evaluated on 3D synthesized vasculature and lung airways in CT and our results demonstrate the advantages of incorporating tree statistics and global optimization for this task.

Keywords: Tree-like · Pictorial structure · Geometrical prior · Bifurcation detection · Centerline extraction · Global optimization

1 Introduction

Branching tree-like anatomical structures are abundant in the human body (i.e., vascular and airway trees of circulatory and respiratory systems) and analyzing their properties is important for various clinical applications, e.g., diagnosis and surgical planning. A necessary precursor to morphological tree analysis is segmenting the trees from 3D medical images. However, 3D segmentation of tree structures is challenging due to, e.g., insufficient contrast between vessels or airways and background, neighbouring/touching tissue, and geometrical variability. Extracting the trees amounts primarily to identifying the bifurcations and the curvilinear paths between them.

Z. Mirikharaji and M. Zhao—Joint first author.

Several previous works on segmentation of tree-like structures relied on local, voxel/patch-level information. For example, Frangi et al. proposed a filter based on local Hessian matrix [7]; Law et al. estimated branch direction based on optimal local inward flux orientation [10]; Schneider et al. used steerable filters and random forest for pixel-wise classification [22]; and Wu et al. proposed a deep learning framework to classify local patches for tracking [26].

Tracking based methods, on the other hand, provide better structural information, but they generally fail to build a global tree structure. Lesage et al. proposed a particle filtering method to track coronary vessels, which incorporate vessel geometry using flux based features [11]. Macedo et al. [13] proposed a centerline-tracking method, on top of a 2D feature based bifurcation detector.

Incorporating prior knowledge, like geometry and topology, into optimization based image segmentation algorithms has been proven useful for obtaining more accurate and plausible results [17]. However, these priors typically introduce non-convexities in the objective functions.

Although tree-like structures were extracted in [19,20,23–25], in contrast to our work, their trees need to be seeded in tubularity measurement maps and the initial tree topologies are seeding-dependent. In [24,25], edge pair-wise geometrical, instead of topological, prior dominated the optimization process, which makes it impossible to maintain a desired, fixed anatomical structure. While in [19,20,23], the topological priors were interpreted as 2-tuple or 3-tuple of neighboring edges, instead of constructing the whole anatomical tree structure. To segment coronary vessels on 2D xray sequences, M'hiri et al. used temporal and spatial prior inherited from an earlier xray image, but the method is difficult to extend to 3D [15]. Beriault et al. proposed a CRF framework that used brain structures (e.g., basal ganglia) and sinuses locations as anatomical priors for segmenting the cerebral vasculature [2]. But none of these methods addressed the global branching aspect of anatomical tree structures.

Our goal in this work is to perform 3D tree extraction while satisfying these two important objectives: (i) encode the geometrical and topological priors of trees and (ii) ensure a globally optimal tree extraction solution. In this paper, we achieve both objectives by adopting, for the first time, pictorial structures for tree extraction.

Pictorial structure were introduced into the computer vision community in 2005 (Felzenszwalb et al. [5]) and an extensive literature (e.g., Belagiannis et al. [1], Burenius et al. [3]) has been established based on this concept since then. To ensure global optimality, pictorial structures require a model with a tree-like topology. This property makes them a natural and ideal fit to the problem of anatomical tree extraction.

2   Methodology

At a high level, our automatic approach comprises two key steps (Fig. 1): (i) detecting bifurcations and (ii) extracting centerlines of branches connecting bifurcations. Step (i) is achieved via fitting the pictorial structure to the 3D
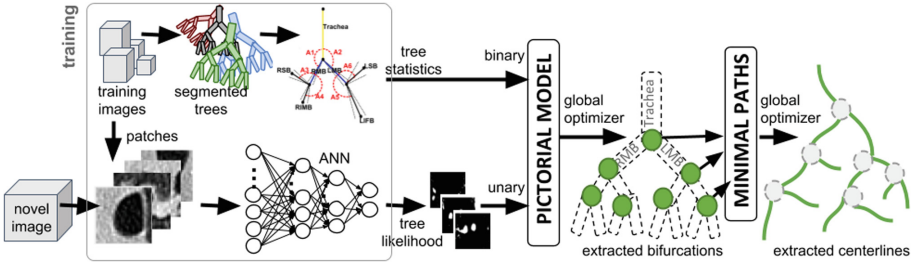
Fig. 1. An overview of the proposed method.

image data by globally optimizing an energy function with an artificial neural network (ANN) derived unary term and a geometrical statistics based binary term. Step (ii) is achieved via a globally optimal minimal path extraction.

2.1   Bifurcation Detection

We formulate the problem of bifurcation detection in 3D as a pictorial structure optimization. A pictorial structure models a deformable object by a set of connected parts. This technique finds the instances of an object in an image by measuring the matching cost for each part and a deformation cost between each two connected components. Felzenszwalb et al. [5] restricted the connection of components to form an acyclic graph $G = (V, E)$, where each vertex v_i corresponds to a component and each edge $e_{ij} = (v_i, v_j)$ models a connection between vertices v_i and v_j. We encode the 3D anatomical tree bifurcations as the nodes of the pictorial structure whereas branch directions and lengths statistics are learned as geometrical priors for regularizing pictorial edges. Let $\mathcal{I}(\mathbf{x})$ be an N dimensional image and $\mathbf{x} \in \mathbb{R}^N$, we optimize an energy function over the location of n nodes in N dimensional space, as follows:

$$\mathcal{L}^* = \underset{\mathcal{L}=\{\mathcal{L}_1,\ldots,\mathcal{L}_n\}}{\arg\min} (\underbrace{\sum_{i=1}^{n}\mathcal{U}(\mathcal{L}_i|\mathcal{I})}_{unary\ term} + \underbrace{\sum_{e_{ij}\in E}\mathcal{B}(\mathcal{L}_i,\mathcal{L}_j))}_{binary\ term} \tag{1}$$

where $\mathcal{U}(\mathcal{I}, \mathcal{L}_i)$ is the unary term penalizing locating v_i at location $\mathcal{L}_i$ and $\mathcal{B}(\mathcal{L}_i, \mathcal{L}_j)$ is the binary term penalizing the deformation of the vector $\mathcal{L}_{ij} = \mathcal{L}_j - \mathcal{L}_i$ away from geometrical priors learned from training data. By leveraging the generalized distance transform [6], the pictorial energy function in (1) is efficiently and globally minimized.

Unary Term via an Artificial Neural Network: We train a three layer neural network stacked with Restricted Boltzmann Machines (RBM) to build a distance map and use it as the unary term in (1). RBM is a two-layer network

of visible and hidden units with no intra-layer connections and symmetrically weighted inter-layer connections. Instead of initializing the network by small random weights, we pre-train the network unsupervisely using RBMs [8]. RBMs compute the joint probability of visible and hidden units and provide a high level representation of data in an unsupervised manner. We construct the network by stacking three RBMs, considering hidden units of preceding ones as visible units of following RBMs. We then fine-tune an ANN, end-to-end, to predict tree voxels by minimizing the total cross entropy between predicted and ground truth segmentations of a training dataset. To encourage a detected bifurcation to be close to the center of its neighboring branches, we instead compute and predict a distance map from segmented edges rather than the segmentation maps themselves.

Binary Term from Geometrical Statistical Priors: We learn distribution of branch angles and lengths of anatomical trees from the skeletons of the ground truth segmentations of a training dataset. Anatomical branch angles and lengths are encoded as three dimensional displacement vectors pointing from bifurcations at lower generations of the tree to upper generations. We model the joint prior distribution of locations of two pictorial connected components as a multivariate Gaussian. The mean vector μ_{ij} and covariance matrix Σ_{ij} of a displacement vector between nodes v_i and v_j are estimated from the training data. By applying singular value decomposition, i.e., $\Sigma_{ij} = U'_{ij} M_{ij}^{-1} U_{ij}$, we write the following joint likelihood estimation in the form of the Mahalanobis distance:

$$-\log p(\mathcal{L}_i, \mathcal{L}_j) \; \propto \; d_{ij}(\mathcal{L}_i, \mathcal{L}_j) = [T_{ij}(\mathcal{L}_i) - T_{ji}(\mathcal{L}_j)]' M_{ij}^{-1} [T_{ij}(\mathcal{L}_i) - T_{ji}(\mathcal{L}_j)] \tag{2}$$

where $\mathcal{L}_i$ and $\mathcal{L}_j$ are the locations of nodes v_i and v_j, respectively; $T_{ij} = U'_{ij}(\mathcal{L}_i - \mu_{ij})$ and $T_{ji} = U'_{ij}(\mathcal{L}_j)$ are rigid (i.e. 6 DOF) spatial transformations [5]; and M_{ij} is a diagonal matrix weighting the deformation cost of connection e_{ij}.

To efficiently find a global solution to (1), we must pick a set of connections between pictorial components that form an acyclic graph (tree). One natural option is to adopt the anatomical tree connectivity as the pictorial structure tree connectivity (option 1). However, it is more informative to connect pairs of nodes with a consistent behaviour across the training data [6]. So, alternatively (option 2), we construct a complete weighted graph over all vertices, assign to edge e_{ij} a weight w_{ij} equal to 2-norm of covariance matrix of three dimensional displacement vectors $\|\Sigma_{ij}\|_2$, and finally find the Minimum spanning tree (MSP) of this graph by Prim's algorithm [18]. We found that the detected tree structure is the same as anatomical tree.

2.2   Branch Centerline Extraction

We use a globally optimal minimal path extraction, based on the fast marching method, to extract the centerlines of all tree branches. While most minimal path

extraction methods are semi-automatic and require users to provide a path's start and end points [9, 12, 14, 16], we used the bifurcation locations detected by the pictorial model optimization to initialize the minimal path extraction.

In tubular structures, if the speed function of the minimal path algorithm is homogeneous or has a small variation near the actual centerline, the shortest path is detected as the *Euclidean* path instead of the medial path or centerline. To ensure that the detected paths pass along the centers of branches, we adapted Deschamps' path centering algorithm [4]. Deschamps first extracted a rough centerline for tubular structures and then used the detected centerline to achieve a rough binary segmentation of vasculatures. A distance transform of the detected edges in the segmented vasculatures is computed and is fed to the minimal path algorithm as a new speed function. In this paper, instead of segmenting using estimated centerline, we use the output mask from our ANN as an approximate segmentation.

3   Experiments

3.1   Data Description

Synthetic Data: We generated 50 volumes, each of size 150^3 and containing a binary tree structure with 4 levels. The tree statistics are set as: mean branch lengths of $\{50, 40, 30, 20\}$ (voxels) and standard deviation (std) of 2; the mean angles between the neighboring levels are $\{\frac{\pi}{4}, \frac{\pi}{6}, \frac{\pi}{12}\}$ with std $= \frac{\pi}{36}$.

Clinical CT Data: We used 19 chest computed tomography (CT) scans[1]. We performed 3-fold cross validation to train the model. The ANN is trained with 240,000 2D patches of size 29×29 along axial planes around each voxel. We chose our ANNs based on empirical evidence and previous works. Other ANN designs, or even other non-ANN methods for calculating the unary term, may yield even better results. A grid search is performed to set and fix the ANN hyperparameters. Our pictorial model consists of the seven components of the first four levels of an airway tree (i.e., trachea, left and right main bronchi, etc.)

3.2   Evaluation Measures

Bifurcation Detection evaluation: The performance of the proposed approach is assessed by (i) N_D: number of detected bifurcations with distance less than D from the ground truth locations; (ii) M: mean distance between the ground truth bifurcations and the corresponding closest detected bifurcations.

Branch Extraction Evaluation: We measure how well the detected curvilinear centerlines match the ground truth centerlines by computing μ_D: the average distance between centerlines [21].

---

[1] http://image.diku.dk/exact/.

Table 1. Performance of different methods on clinical data with measure M and μ_D. Distance unit in mm and values shown in format $mean \pm std$.

	Method	M (mm)	μ_D (mm)	bif. det./path init
A	Tracker [13]	9.41 ± 6.84	9.21 ± 9	Manual root seed
B	OOF [10]	NR†	3.24 ± 1.68	Manual bif. det.
C	Minimal path $+$ ANN $+$ DT	NR†	3.09 ± 1.5	Manual bif. det.
D	Ours (pict w/o stats $+$ minimal path)	14.54 ± 16.54	4.87 ± 4.84	Automatic
E	Ours (pict with stats $+$ minimal path)	8.39 ± 7.41	3.51 ± 2.4	Automatic

†NR: Not reported since bifurcations are manually selected.

Table 2. Effect of SNR on measure M for synthetic data ($mean \pm std$).

SNR	M (voxel)
∞ (noise-free)	5.19 ± 3.30
10	5.29 ± 3.23
5	6.69 ± 11.29
3.3	7.83 ± 9.3
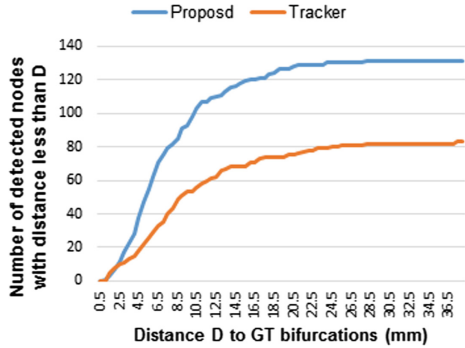
Fig. 2. Variation of N_D on real data for proposed method and Tracker [13].

3.3   Experimental Results

Advantage of Statistics: To confirm the advantage of incorporating tree statistics, we removed the tree statistics of real data and globally optimized the objective function. In practice, we mimicked a uniform geometrical prior by scaling up the covariance matrix elements by a factor of 20. Rows D and E in Table 1 show that incorporating statistics improves M and μ_d by 42% and 28%, respectively.

Robustness to Noise: To evaluate the robustness of our bifurcation detector to noise, we added three levels of Gaussian noise with SNR= $[10, 5, 3.3]$ to the noise-free synthetic data. A distance map from the edges of the tree mask is used as the unary term. Table 2 shows that our method is stable even to high level of noise. For example, in columns 3 and 4 of Table 2, SNR is doubled while M is increased about one voxel.

Advantage of Globally Optimal Model Fitting: We examined the drop in performance when a gradient descent local optimizer approach is used. As expected, the bifurcation localization result is highly sensitive to initialization even if the initialization is close to the ground truth locations. This sensitivity can be attributed to not having a reliable and clean (noise-free) data term, causing

the local optimizer to get trapped in local optima. It is also worth noting that our algorithm is linear in both the number of branching points and the number of possible locations for each node. We also compare the proposed method to two competing methods, a tracker based on a bifurcation estimator (Tracker) [13] and the model-based optimally oriented flux method (OOF) [10].

Comparing to Tracker: The root seed point of Tracker is manually set in the trachea trunk. Since the tracker doesn't have a built-in anatomical tree topology, we match each ground truth bifurcation to its closest one among all the detected bifurcations. Figure 2 reports N_D as a function of D for the proposed method (blue curve) and Tracker (brown curve) on real data. The two different plateau levels of the two curves illustrate that not all the bifurcations are detected by Tracker. Also, rows A and E in Table 1 illustrate that the proposed method outperforms Tracker, reducing the error by 10% in M and 62% in μ_D.

Comparing to OOF: To trace the centerlines using the tubularity score of OOF, we had to manually select bifurcations before using the fast marching algorithm to generate the path between those bifurcations. A Naïve comparison of rows B and E in Table 1 shows that OOF outperforms our method by about 3 mm in μ_D on real data. However, our proposed method doesn't need initialization and is fully automatic. So, for a fair comparison, we used the same manually selected bifurcations and the centerline extraction approach in Sect. 2.2. The result is reported in row C. Now, comparing rows B and C shows that, using the same set of bifurcations, the minimal path on distance transform of ANN output (i.e., our approach) outperforms OOF. Moreover, since the variation of μ_D for rows B, C and E is less than the average voxel size of our clinical data $(0.67*0.67*0.95 \, \text{mm}^3)$, these experiments confirm that by detecting bifurcations using the pictorial structure, the tracing algorithm becomes fully automatic while the accuracy of the centerline detection remains practically unchanged.

4   Conclusions

We presented the first global method for extracting tree-like structures from 3D medical images while encoding geometrical tree priors. The global model-to-data fitting made centerline tracing free from any initialization and the incorporation of priors made the method more robust to noise. Incorporating fixed topological priors for consistent branches is advantageous of this paper. In the existence of topological variability, e.g. pathology cases or generations deep down the tree; our method is not designed to handle these cases. Nevertheless, we note that the pictorial algorithm [5] is stable to occlusions, so even when the tree model has a fixed topology, it should still be able to locate actual trees with slight topological variations. Future work will involve integrating the minimal path optimization within the pictorial algorithm and encoding more elaborate branch statistics (e.g. medial curvature and radii). It is also interesting to explore automatic ways to detect pathological deviations from priors that are supported by image evidence, as these may indicate pathology.

References

1. Belagiannis, et al.: 3D pictorial structures for multiple human pose estimation. In: CVPR, pp. 1669–1676 (2014)
2. Bériault, S., et al.: Automatic SWI venography segmentation using conditional random fields. IEEE TMI **34**(12), 2478–2491 (2015)
3. Burenius, M., et al.: 3D pictorial structures for multiple view articulated pose estimation. In: CVPR, pp. 3618–3625 (2013)
4. Deschamps, T.: Curve and shape extraction with minimal path an level-sets techniques: applications to 3D Medical Imaging. Ph.D. thesis, University of Paris Dauphine (2001)
5. Felzenszwalb, P.F., et al.: Pictorial structures for object recognition. IJCV **61**(1), 55–79 (2005)
6. Felzenszwalb, P., Huttenlocher, D.: Distance transforms of sampled functions. Cornell University, Technical report (2004)
7. Frangi, A.F., Niessen, W.J., Vincken, K.L., Viergever, M.A.: Multiscale vessel enhancement filtering. In: Wells, W.M., Colchester, A., Delp, S. (eds.) MICCAI 1998. LNCS, vol. 1496, pp. 130–137. Springer, Heidelberg (1998). doi:10.1007/BFb0056195
8. Hinton, G.E., et al.: Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006)
9. Krissian, K., et al.: A minimal cost path and level set evolution approach for carotid bifurcation segmentation. Midas J. (2009)
10. Law, M.W.K., Chung, A.C.S.: Three dimensional curvilinear structure detection using optimally oriented flux. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008. LNCS, vol. 5305, pp. 368–382. Springer, Heidelberg (2008). doi:10.1007/978-3-540-88693-8_27
11. Lesage, D., et al.: Adaptive particle filtering for coronary artery segmentation from 3D CT angiograms. CVIU **151**, 29–46 (2016)
12. Li, H., Yezzi, A.: Vessels as 4D curves: Global minimal 4D paths to extract 3D tubular surfaces and centerlines. IEEE TMI **26**(9), 1213–1223 (2007)
13. Macedo, et al.: A centerline-based estimator of vessel bifurcations in angiography images. In: SPIE Medical Imaging, p. 86703K (2013)
14. Metz, C., et al.: Two point minimum cost path approach for CTA coronary centerline extraction. Insight J., 123 (2008)
15. M'hiri, F., et al.: A graph-based approach for spatio-temporal segmentation of coronary arteries in X-ray angiographic sequences. Comput. Biol. Med. **79**, 45–58 (2016)
16. Mueller, D.: Fast marching minimal path extraction in ITK. Insight J., 1–8 (2008)
17. Nosrati, et al.: Incorporating prior knowledge in medical image segmentation: a survey. CoRR abs/1607.01092 (2016). http://arxiv.org/abs/1607.01092
18. Prim, R.C., et al.: Shortest connection networks and some generalizations. Bell Labs Tech. J. **36**(6), 1389–1401 (1957)
19. Rempfler, M., et al.: Extracting vascular networks under physiological constraints via integer programming. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014. LNCS, vol. 8674, pp. 505–512. Springer, Cham (2014). doi:10.1007/978-3-319-10470-6_63

20. Robben, D., Türetken, E., Sunaert, S., Thijs, V., Wilms, G., Fua, P., Maes, F., Suetens, P.: Simultaneous segmentation and anatomical labeling of the cerebral vasculature. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014. LNCS, vol. 8673, pp. 307–314. Springer, Cham (2014). doi:10.1007/978-3-319-10404-1_39

21. Schaap, M., et al.: Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. MIA **13**(5), 701–714 (2009)

22. Schneider, M., et al.: Joint 3D vessel segmentation and centerline extraction using oblique Hough forests with steerable filters. MIA **19**(1), 220–249 (2015)

23. Skibbe, H., et al.: Efficient Monte Carlo image analysis for the location of vascular entity. IEEE TMI **34**(2), 628–643 (2015)

24. Türetken, E., et al.: Automated reconstruction of dendritic and axonal trees by global optimization with geometric priors. Neuroinformatics **9**(2–3), 279–302 (2011)

25. Türetken, E., et al.: Automated reconstruction of tree structures using path classifiers and mixed integer programming. In: CVPR, pp. 566–573. IEEE (2012)

26. Wu, A., et al.: Deep vessel tracking: a generalized probabilistic approach via deep learning. In: IEEE ISBI, pp. 1363–1367 (2016)