Topology Reconstruction of Tree-like Structure in Images via Structural Similarity Measure and Dominant Set Clustering

Jianyang Xie[1], Yitian Zhao[1*], Yonghuai Liu[2], Pan Su[1], Yifan Zhao[3*],
Jun Cheng[1], Yalin Zheng[4], Jiang Liu[1]

[1]Ningbo Institute of Industrial Technology
Chinese Academy of Sciences
`yitian.zhao@nimte.ac.cn`

[2]Department of Computer Science
Edge Hill University
`liuyo@edgehill.ac.uk`

[3]School of Aerospace, Transport and Manufacturing
Cranfield University
`yifan.zhao@cranfield.ac.uk`

[4]Department of Eye and Vision Science
University of Liverpool
`yzheng@liv.ac.uk`

Abstract

The reconstruction and analysis of tree-like topological structures in the biomedical images is crucial for biologists and surgeons to understand biomedical conditions and plan surgical procedures. The underlying tree-structure topology reveals how different curvilinear components are anatomically connected to each other. Existing automated topology reconstruction methods have great difficulty in identifying the connectivity when two or more curvilinear components cross or bifurcate, due to their projection ambiguity, imaging noise and low contrast. In this paper, we propose a novel curvilinear structural similarity measure to guide a dominant-set clustering approach to address this indispensable issue. The novel similarity measure takes into account both intensity and geometric properties in representing the curvilinear structure locally and globally, and group curvilinear objects at crossover points into different connected branches by dominant-set clustering. The proposed method is applicable to different imaging modalities, and quantitative and qualitative results on retinal vessel, plant root, and neuronal network datasets show that our methodology is capable of advancing the current state-of-the-art techniques.

1. Introduction

Accurate reconstruction of the pervasive *tree-like* structures commonly encountered in biomedical imagery, such as blood vessels in medical images, the roots of plants in color photography, or neuronal arbors in optical microscopy images, is a fundamental step in providing valuable clinical information in diagnosis of pathological con-

ditions [25, 30], revealing important information about the delivery of nutrients to different plant tissues [14, 1], or screening the neuronal connectivity to study the relation between neuronal structure and function [3, 26], respectively. In consequence, extracting and representing the geometrical (radii, lengths or tortuosity) and topological (branching connectivity) properties of these tree-like structures is critical for diagnosis, or other analytic purposes.

Rapid development has taken place in the automatic geometric extraction of the properties of tree-like structures, as evidenced by extensive reviews [16, 32]: but automated topology reconstruction from a single, two-dimensional (2-D) image of a three-dimensional (3-D) tree-structure is relatively unexplored, and more challenging. The main reason is that while a 3-D tree-structure projects readily to a 2-D image plane, the problem of then reconstructing the topology of the tree-like structure from it is ill-posed and error-prone [14], due to their depth loss in projection, imaging noise and low contrast in appearance and geometric structures.

The underlying tree-structure topology must be able to identify and differentiate individual curvilinear objects from the entire tree network. The top row of Figure 1 shows examples of tree-like structures in biomedical imagery: the bottom row gives their derived topological information. Manual reconstruction of these tree-structure topologies is time consuming, and commercial software, e.g., Vaa3D[1], still largely relies on manual annotation and so is subject to human error [11]. Thus, the catalyst for the development of algorithms for accurate automatic topology reconstruction has been primarily the need to overcome time constraints and avoid human error.

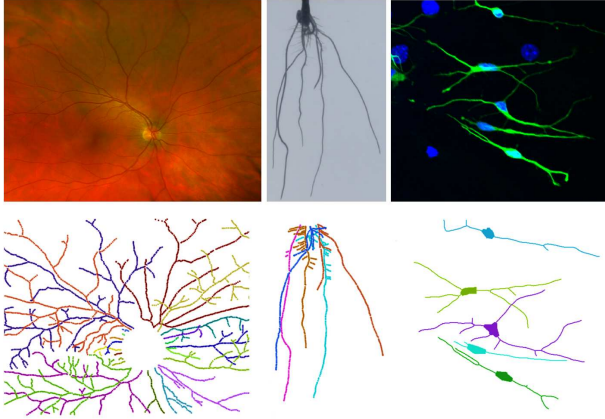[1]`http://home.penglab.com/proj/vaa3d/home`

Figure 1. Examples of biomedical images exhibiting tree-like structures, and their derived topological information. From left to right: retinal vessels, a plant root system, and neuronal arbors. Each individual tree branch is marked with a distinct color.

1.1. Related works

Automated tree-structure topology reconstruction from three-dimensional images is a well-posed problem [18], and most existing efforts in tree-structure topology reconstruction have focused on three-dimensional data, such as the lung airways from computed tomography (CT) [6, 19], cerebral vascular network from magnetic resonance angiogram (MRA) [4, 27], and nerves or neurons from microscopic image stacks [5]. However, topology estimation from a single 2-D image of a 3-D tree-structure is more challenging, and has received relatively little attention. To the best of our knowledge, only a small number of studies have addressed the topic directly.

The topological reconstruction of tree structures in neuronal and retinal images was formalized as label propagation over a two step graph-theoretical approach by De et al. [11, 12], and each tree branch could be separated from the tree-like network. Cheng et al. [8] formed a directed graph to represent the retinal vessel network. The entire vessel graph was separated into several subgraphs with individual labels, by propagating tree labels from known root nodes to the entire graph. Estrada et al. [14] regularized the topology reconstruction task with a parametric tree-growth model by using a heuristic search algorithm, and applied it to the retinal vessels, plant roots and synthetic leafy trees. The authors further utilized the topological information to classify all the vessels into arteries or veins [13]. Dashtbozorg et al. [10] proposed a graph-based approach for retinal vessel topology estimation, and the vessel orientation and angle features were used to group the nodes at crossover into two pairs.

The automated methods so far developed often fail to identify the connectivity correctly in instances of curvilinear *crossover*: i.e., two curvilinear objects overlap or touch-

ing due to the projection [10]. Crossovers lead to difficulty in predicting whether a given curvilinear object contacting a crossover belongs to the same tree, or a different tree, due both to their projective ambiguity and subtle changes in appearance, contrast and geometry in the imaging process. Moreover, the inhomogeneous intensity and poor contrast due to imaging noise or uneven illumination, often leads to a tree with apparent discontinuities, or with fuzzy boundaries, and which aggravates the difficulty in the estimation of topology.

1.2. Proposed approach

Recently, it has been shown in [33] that the topology of the retinal vessel tree can be estimated by exploiting the concept of dominant set clustering, which formalizes the task of topology estimation as a pairwise clustering problem. However, this method still makes topological mistakes, because it suffers from the fact that the dominant set clustering approach uses the Euclidean distance at the pixel level without attending to global properties, such as tree shape. Such a similarity measure cannot fully capture the structural information conveyed in terms of intensity or shape.

To overcome these limitations, we propose a novel tree-like structure topology reconstruction framework that works automatically from a single 2-D image. In particular, we focus on addressing the bottleneck crossover problem. The proposed method includes three main contributions:

- The formalization of the topology reconstruction of a tree-like structure from a single projective image.

- A novel curvilinear structural similarity measure is introduced by aggregating three different similarities, which together can take into account both intensity and geometric properties to represent curvilinear structures locally and globally. It ensures that the extracted similarity is diverse in features.

- The re-conceptualization of the topology estimation task as a pairwise clustering problem: a curvilinear structural similarity-enabled dominant set clustering, which proves to be an effective way of addressing the problem of tracing branches at crossovers.

The proposed method has been validated quantitatively using four publicly accessible datasets with three different imaging modalities - microscopic images, retinal color fundus images, and RGB source images.

2. Method

2.1. Graph Generation

The proposed topology reconstruction approach for tree structure requires a pre-processing step, which encompasses the following four phases.
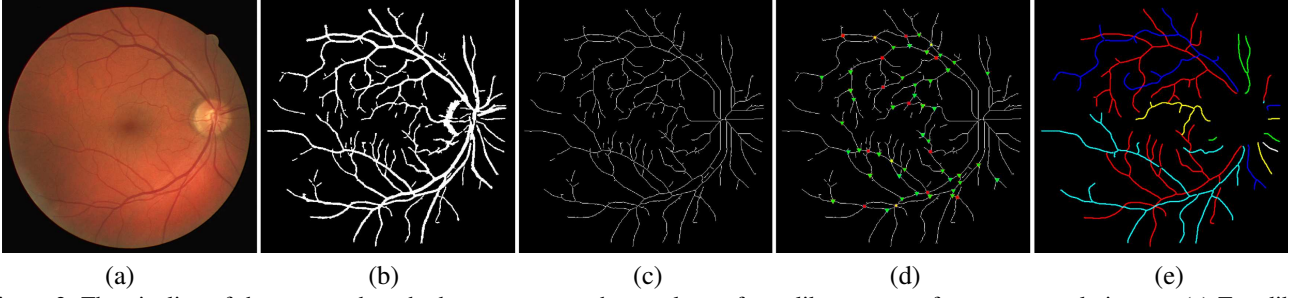
Figure 2. The pipeline of the proposed method to reconstruct the topology of tree-like structure from an example image. (a) Tree-like structure (retinal vessel); (b) Extracted tree-like structure; (c) Skeletonized tree-like structure; (d) Localization of junctions, in which green triangles indicate bifurcations, red squares and yellow stars crossovers. (e) Reconstructed topology of tree network.

Segmentation: The infinite perimeter active contour model using hybrid region information [34] was employed to automatically segment the tree-like structures. (Note: this step may be skipped for some datasets where manual annotation of the tree-like structures is provided with the data source.) *Skeletonization*: The centerlines of tree-structures provide an accurate representation of their topologies. Therefore, an iterative morphology thinning operation [2] is performed on the extracted tree-structures to obtain a single-pixel-wide skeleton map. *Junction localization*: The bifurcations, crossovers, and filament ends (terminal points) may then be extracted from the skeleton map by locating junction points (pixels with more than two neighbors) and terminal points (pixels with one neighbor). All the junction points and their neighbors are then removed from the skeleton map, producing an image with clearly separated segments. Junctions are categorized into four groups, according to the number of points involved in each junction - terminal points (1), connecting points (2), bifurcation points (3), and crossovers/meeting points (4 and above). The bracketed number indicates the number of filament segments connected to each intersection. *Graph generation*: A graph can be generated from the skeleton map by linking first and last nodes in the same segment. The typical misrepresentations from the generated graph, such as node splitting, missing links and false links, are corrected by employing a graph modification approach proposed in [10].

Fig. 2 (a)-(d) shows an example of the outline of the graph preparation for retinal vessels.

2.2. Dominant-set Clustering for Topology Reconstruction

The concept of dominant sets arises from the study of graph theory, by which a continuous formulation of the maximum clique problem is defined [23]. An undirected graph G with weighted edges is represented by $G = (V, E, \omega)$, where V is a set of nodes/pixels, edge set $E \subseteq V \times V$ indicates all the possible connections, and $\omega : E \to \mathbb{R}$ is the positive weight function and represents the similarity among the pixels in V. A $|V| \times |V|$ symmetric matrix $A = (a_{ij})$ is used to represent the weighted graph G as an adjacency matrix.

A dominant set can be formally defined based on the values of similarity among nodes in V. Let $S \subseteq V$ be a nonempty subset of nodes, $p_i \in V$ and $p_j \in S$. The relative similarity between p_i and p_j with respect to the average similarity between p_j and its neighbours in S is defined as:

$$\phi_S(i,j) = a_{ij} - \frac{1}{|S|} \sum_{p_k \in S} a_{jk}. \quad (1)$$

The edge weight definition plays a crucial role in the proposed method. In [33], the inverse Euclidean distance between the two end points of an edge in the feature space is used, where each node is represented as a 23-dimensional feature vector. However, this approach lacks of robustness due to failure to capture global characteristics. In our work, a_{ij} is derived by a curvilinear structural similarity metric which is aggregated from pixel level, region level, and geometrical similarities. The derivation of this novel similarity will be proposed below in Sec. 3.

It is known from definition that $\phi_S(i,j)$ in Eqn. (1) can be either positive or negative. The weight of p_i with regard to S is assigned recursively as:

$$W_S(i) = \begin{cases} 1 & \text{if } |S| = 1 \\ \sum_{p_j \in S \setminus \{p_i\}} \phi_{S \setminus \{p_i\}}(i,j) W_{S \setminus \{p_i\}}(j) & \text{otherwise.} \end{cases}$$
$$(2)$$

where $S \setminus \{p_i\}$ indicates the nodes set S excluding the node p_i, and $W_S(i)$ expresses the similarity between node p_i and the nodes of $S \setminus \{p_i\}$ with respect to the mutual similarity amongst the nodes in $S \setminus \{p_i\}$. Finally, the total weight of S is calculated by $W(S) = \sum_{p_i \in S} W_S(i)$.

Formally, a non-empty subset of nodes S, $S \subseteq V$ such that $W(S') > 0$ for any non-empty $S' \subseteq S$ is said to be a dominant set if:

$$W_S(i) > 0, \text{ for all } p_i \in S \quad (3)$$

Algorithm 1 Topology Estimation(I)

Inputs: I: a set of points p_i that is associated with intersections;
Outputs: S^*: a partition of I, each element of which includes points of the same branch;

1: initialize the $|I| \times |I|$ symmetric matrix $A = (a_{ij})$ by calculating $a_{ij} = \omega(i, j)$ with respect to given features;
2: $S^* = \emptyset$
3: **while** $I \neq \emptyset$ **do**
4:     initialize a $|I|$ dimensional vector $\mathbf{x}^{(0)} \in \Delta$
5:     **for** $t = 0 : MaxIteration - 1$ **do**
6:         **for** $i = 1 : |I|$ **do**
7:             $x_i^{(t+1)} = x_i^{(t)} \frac{(A\mathbf{x}^{(t)})_i}{\mathbf{x}^{(t)\top} A\mathbf{x}^{(t)}}$
8:         **end for**
9:     **end for**
10:    $S = \emptyset$
11:    **for** each $x_i^{(MaxIteration)} > 0$ **do**
12:        remove the column and row with respect to p_i from A
13:        $S = S \bigcup p_i$
14:    **end for**
15:    $I = I \setminus S$
16:    $S^* = S^* \bigcup \{S\}$
17: **end while**

and

$$W_{S \cup \{p_i\}}(i) < 0, \text{for all } p_i \notin S. \tag{4}$$

In general, the weights of edges within the dominant set of a edge-weighted graph should be large, representing high internal homogeneity or similarity [23]. By contrast, the weights of edges which link to the dominant sets externally will be small. Therefore, the dominant set is a proper solution to identify branches of a tree, because the similarity of two points from the same branch should be large, while that of two points belonging to different branches should be small.

Dominant sets can be identified by local solutions of the program:

$$\begin{aligned} \text{maximize} \quad & f(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} \\ \text{subject to} \quad & \mathbf{x} \in \Delta, \end{aligned} \tag{5}$$

where

$$\Delta = \left\{ \mathbf{x} \in \mathbb{R}^d : \sum_{i=1}^d x_i = 1 \text{ and } x_i \geq 0 \text{ for all } i = 1, \cdots, d \right\}$$

and $d = |I \setminus S|$. A strict local solution $\mathbf{x}^*$ of Eqn. (5) indicates a dominant set $\mathbb{S}$ of G, where $x_i > 0$ means that the node in question $p_i \in \mathbb{S}$. An effective optimization approach for solving Eqn. (5) is given by the so-called *replicator dynamics* [24]:

$$x_i^{(t+1)} = x_i^{(t)} \frac{(A\mathbf{x}^{(t)})_i}{\mathbf{x}^{(t)\top} A\mathbf{x}^{(t)}}, \tag{6}$$

where $i = 1, 2, \cdots, |\mathbb{S}|$. It has been proven that for any initialization of $\mathbf{x} \in \Delta$, its trajectory will remain in Δ with the increase of iteration t. As t in Eqn. (6) increases, the objective function $f(\mathbf{x})$ in Eqn. (5) is either strictly increasing or constant. In practice, the stopping criterion of the dynamic system can be set as a maximal number of iteration

t, or as a minimal increment of $f(\mathbf{x})$. Algorithm 1 shows the complete procedure for partitioning a tree structure into branches S.

It is worth noticing that a peeling-off strategy has been adopted: i.e., this method iteratively extracts a subset of points belonging to the same branch (a dominant set S) each time by using Eqn. (6) and repeating the process with the remaining points within the intersection $I = I \setminus S$. The identification of different branches at an intersection is carried out by identifying the most obvious branch first, and then the second most obvious, and so on. Therefore, the peeling-off strategy is a direct, intuitive implementation of this procedure [24].

3. Curvilinear structural similarity measure

The similarity metric $a_{ij} = \omega(i, j)$ in Algorithm 1 is as yet to be defined. In this section, we will detail this similarity measure.

Traditional similarity measures, such as the Euclidean distance, cannot fully capture the local and global, or intensity-based and geometry-based structural information of the given objects, and it is hard to provide reliable information for data clustering. Therefore, a fine similarity measure with informative local and global structural details is essential in the reconstruction of a complicated tree-like structure. In this work, we adopt the concept of structural similarity (SSIM) [29], and re-conceptualize it as a new tree-structural similarity measurement, which further exploits the regional and geometric characteristics of curvilinear segments.

Two different curvilinear segments will be deemed to be from the same tree branch if they satisfy the following anatomical characteristics, which are supported by clinical and biological evidences [21, 26, 14]:

- High local similarity: their local contrast is small;

- High global similarity: entire objects consist of similar curvilinear compoenents;

- High geometric similarity: the global orientations and tortuosity of the curvilinear object are similar.

In consequence, we propose a novel curvilinear structural similarity measure to satisfy the above properties. This similarity measurement is aggregated from the similarity measured at pixel level, region level, and geometrical level.

3.1. Pixel-aware structural similarity measure

SSIM is a similarity measure that is able to compare the local pattern of pixel intensities from the perspective of image formation. Let $\mathbf{u}$ and $\mathbf{v}$ be the pixel values in the segments where the nodes p_i and p_j are located, respectively.

Then the $\text{SSIM}(i, j)$ of two segments containing nodes p_i and p_j can be defined as:

$$\text{SSIM}(i, j) = \frac{2\mathbf{u}'_i \mathbf{v}'_j}{(\mathbf{u}'_i)^2 + (\mathbf{v}'_j)^2} \cdot \frac{2\sigma_{\mathbf{u}_i}\sigma_{\mathbf{v}_j}}{(\sigma_{\mathbf{u}_i})^2 + (\sigma_{\mathbf{v}_j})^2} \cdot \frac{\sigma_{\mathbf{u}_i \mathbf{u}_j}}{\sigma_{\mathbf{u}_i}\sigma_{\mathbf{v}_j}}. \tag{7}$$

The three terms of SSIM are luminance, contrast, and structural components, respectively [29]. The $\mathbf{u}'$, $\sigma_{\mathbf{u}}$, $\mathbf{v}'$, and $\sigma_{\mathbf{v}}$ are the mean and standard deviation of the pixel values of the two segments containing nodes p_i and p_j, respectively, and $\sigma_{\mathbf{uv}}$ is the covariance between them. (Note, in the real application, the size of the two curvilinear segments are usually in different, as thus they are straightened and resized into a certain shape.) It can be observed that the three components are relatively independent. For example, the smaller the difference between $\mathbf{u}'$ and $\mathbf{v}'$, or $\sigma_{\mathbf{u}}$ and $\sigma_{\mathbf{v}}$, the greater the similarity of luminance or contrast. Meanwhile, the change of luminance or contrast will not affect the structures measurement. In [29], the structural similarity of two different objects is equivalent to the correlation coefficient between $\mathbf{u}$ and $\mathbf{v}$, and it is associated with the two normalized vectors $(\mathbf{u} - \mathbf{u}')/\sigma_{\mathbf{u}}$ and $(\mathbf{v} - \mathbf{v}')/\sigma_{\mathbf{v}}$.

3.2. Region-aware structural similarity measure

The pixel level structural similarity measure discussed above is mainly defined by the intensity values of each pixel, and so can capture local details of curvilinear objects. However, a purely pixel level structural similarity measure cannot well account for regional similarities [9, 15]. Therefore, in this section, we propose a novel region level structural similarity measure.

Assume that the given segment i is partitioned into M superpixels $\{\mathbf{p}_m^i\}_{m=1}^M$, and segment j is superpixelized into N patches $\{\mathbf{p}_n^j\}_{M=1}^N$ using the Simple Linear Iterative Clustering (SLIC) method. The region-aware structure similarity (RSIM) of these two segments can be formulated as:

$$\text{RSIM}(i, j) = \sum_{m=1}^M \sum_{n=1}^N \text{SSIM}(\mathbf{p}_m^i, \mathbf{p}_n^j) \cdot w(\mathbf{p}_m^i, \mathbf{p}_n^j), \tag{8}$$

where $w(\mathbf{p}_m^i, \mathbf{p}_n^j)$ is a standard Gaussian weighting function to model the local contrast in terms of geometric distances between superpixels m and n, so that local contrast can be effectively combined with control over the influence radius:

$$w(\mathbf{p}_m^i, \mathbf{p}_n^j) = \frac{1}{\mathbf{Z}}\exp\left\{-\frac{\|\mathbf{p}_m^i - \mathbf{p}_n^j\|^2}{2\sigma^2}\right\}, \tag{9}$$

where $\mathbf{Z}$ is the normalization term to ensure that $\sum_{n=1, n\neq m}^N w(\mathbf{p}_m^i, \mathbf{p}_n^j) = 1$. The range of the regional similarity may be controlled by the standard deviation σ^2 from 0 to 1. Larger values of σ increase the spatial weighting of the distant segments, and vice versa. In our implementation, we set $\sigma^2 = 0.3$ empirically. The operator $\|\cdot\|$ indicates the spatial distance between $\mathbf{p}_m^i$ and $\mathbf{p}_n^j$.

The RSIM considers not only the local intensity values, but also how the pixel values are distributed in each patch, and how similar such distributions are between different patches. This method treats patches (superpixels) instead as units for similarity measurement by considering how similar different patches are and comparing them directly to each other [15] in order to determine their relative contrast and similarity.

3.3. Geometry-aware structural similarity measure

Dividing the intensity-based structural similarity into pixel level and region level, helps in obtaining the object structural similarity. However, in the field of biomedical image analysis, the object's shape or geometry level similarity is crucial to revealing abnormalities [35]. To this end, we propose a novel geometrical structural similarity measure (GSIM) in this work. This GSIM can capture the overall geometry of a shape, or the tortuosity of a curvilinear component, and is more robust to variation in imaging quality and scale.

In our work, the *shape index* is proposed in order to capture the intuitive notion of 'shape' locally and globally [20]. The shape index $\mathscr{E}$ of each centerline pixel in a given segment may be defined as

$$\mathscr{E} = \frac{2}{\pi}\arctan\frac{\mathscr{C}_2 + \mathscr{C}_1}{\mathscr{C}_2 - \mathscr{C}_1}, \tag{10}$$

where $\mathscr{C}_1$ and $\mathscr{C}_2$ are the curvatures of each side of a curvilinear segment. The curvature obtained from eigensystem analysis of the Hessian matrix is commonly used in the computation of vectors tangent to oriented structures [31]. The measurement of shape index in a 3D surface is better able to represent the concavity or convexity of a given region, and in 2D or 1D applications, it may be employed to reveal the global tortuosity of a given shape or curve.

The GSIM can be obtained by calculating the relative distance between their arguments:

$$\text{GSIM}(i, j) = \exp\left\{-\frac{rd(\mathscr{E}_i', \mathscr{E}_j')^2}{2g^2}\right\}, \tag{11}$$

where $\mathscr{E}'$ indicates the mean shape index value of a segment, and g is a control parameter that is able to reveal how distant pixels and dissimilar features will impact the shape index of the current pixel: here, $g^2 = 3$ is empirically chosen. The relative distance $rd(\cdot)$ of two shape indices between two objects is defined as:

$$rd(\mathscr{E}_i, \mathscr{E}_j) = \frac{\|\mathscr{E}_i, \mathscr{E}_j\|}{\text{ave}_{k \subset S}(\|\mathscr{E}_i, \mathscr{E}_j\|)}, \tag{12}$$

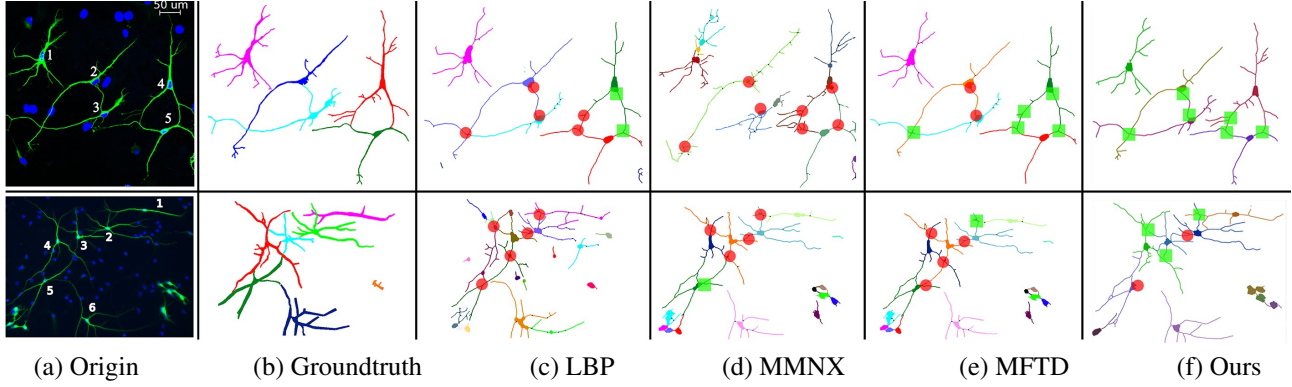|					
(a) Origin	(b) Groundtruth	(c) LBP	(d) MMNX	(e) MFTD	(f) Ours

Figure 3. Two examples of topology reconstruction by different methods on neuron arbors from the **NeurB1** and **NeurB2** dataset.

Table 1. Performances of different methods in identifying the connectivity at junctions, and their DIADEM scores for the **NeurB1** and **NeurB2** datasets.

Dataset	# Junction	Method	# Correctly Identified	Acc_I	Acc_{II}	DIADEM
NeurB1	557	Ours	501	90.0%	93.6%	0.67
		LBP	-	-	-	0.42
		MFTD	389	69.8%	-	0.52
		MMNX	207	37.2%	-	0.43
NeurB2	254	Ours	207	81.5%	92.0%	0.63
		LBP	-	-	-	0.19
		MFTD	199	78.4%	-	0.26
		MMNX	67	26.4%	-	0.22

where $\mathrm{ave}_{k \subset S}(\|\mathscr{E}_i - \mathscr{E}_j\|)$ is the average Euclidean distance of shape index values between segment i and other candidate segment k in intersection S.

3.4. Similarity Fusion via OWA aggregation

Each of the three aforementioned curvilinear structural similarity measures has disadvantages if used alone in real applications. However, they can be aggregated to provide a more robust solution. Weighted average operators (linear summation, Hadamard product, average operator, etc.) are commonly used to implement the aggregation process. In this work, we formulate a combined curvilinear structural similarity by using an Ordered Weighted Averaging (OWA) operator to combat the scale and illumination change from one segment to another.

The fundamental aspect of an OWA operator is the ranking step, in which the extraneous variables are ranked in descending order, with their values subsequently integrated into a single aggregated value [28]. An OWA mapping $A_{\mathrm{OWA}} \colon \mathbb{R}^q \to \mathbb{R}$ is defined as

$$A_{\mathrm{OWA}}(a_1, \cdots, a_Q) = \sum_{q=1}^{Q} w_q a_{\pi(q)}, \qquad (13)$$

where $a_{\pi(q)}$ is a permutation of a_q, in which $a_{\pi(q)}$ is the q-th largest value of the a_q. The ordering of input arguments gives the OWA a nonlinear character. w_q is a collection of weights that satisfies $\sum_q w_q = 1, q = 1, \cdots, Q, Q > 1$.

Different choices of the weight w lead to different aggregation results: in our work, $w_1 = 0.25$, $w_2 = 0.35$, and $w_3 = 0.4$ were chosen empirically. By applying the OWA aggregation to a real problem of this kind, users will have the degree of freedom to control both the t-transitivity of the operator [28], by tuning the stress function, or by tuning the weights in a weighting vector.

4. Datasets and Experimental Results

4.1. Datasets

Four tree-like image datasets were used in this work to validate the effectiveness of our approach: two neuronal image datasets (NeuB1 and NeuB2 [11]), one retinal image dataset (WIDE [14]), and one plant root dataset (RICE [14]).

The neuronal datasets NeuB1 and NeuB2 contain 112 and 98 neuronal images, respectively. (Note: images from the NeuB2 dataset are more challenging, due to noise and blurring.) The manual annotations of structure topology were made by using the annotation tool Neuromantic[2].

The WIDE dataset [14] comprises 15 high resolution color fundus images, each of 3900×3072 pixels. The RICE dataset [14] comprises of 18 rice-root RGB source images of 1300×900 pixels. Both datasets are available from Duke University, and the manual annotations of the topological

---

[2]https://www.reading.ac.uk/neuromantic/body_index.php

trees were generated by a sketching software [7].

4.2. Results on NeurB1 and NeurB2 datasets

Figure 3 presents for visual comparison the topology reconstruction results of the competing methods on two example images from the NeurB1 and NeurB2 neuronal dataset: a commercially available neurite tracer module of Metamorph NX (MMNX)[3]; and two label propagation methods for tree-structure topology estimation: Loopy Belief Propagation (LBP) [22], and Matrix Forest Theorem of Directed (MFTD) [11]. Pink discs indicate incorrectly identified junctions, and gray square demonstrate correctly identified junctions of neuronal images. Compared with their gold standards, it is clear from visual inspection that the proposed method made consistently far fewer mistakes.

To facilitate better observation and objective evaluation of the performance of the proposed method in the topology reconstruction of tree-like structures, we calculate the accuracy of the identified junctions and their associated curvilinear centerline pixels of the tree-like structure as follows. Let A be the total number of the junctions or centerline pixels of a tree-like structure, and B be the number of the junctions or centerline pixels that have been correctly identified (the vertices of the estimated topology tree have been assigned labels identical with groundtruth). The accuracy is then computed by $Acc = \frac{B}{A} \times 100\%$, and here we refer to Acc_I and Acc_{II} as the accuracy at junction level, and centerline pixel level, respectively. In addition, the DIADEM score [17] is utilized to measure the similarity between a groundtruth tree-like structure and the corresponding topology estimation result. The DIADEM score is a widely-used metric to measure the similarity of two neuronal structures, and the obtained score falls in the range [0, 1], where 0 indicates a completely mismatch and 1 indicates perfect agreement. The x-y threshold of the DIADEM score is set as 30 pixels in this experiment.

Table 1 presents the performance measurements of the proposed method and competing methods in identifying connectivity at junctions. It may be observed that the proposed method is able to detect the most of the junctions correctly, with $Acc_I = 90.0\%$, $Acc_{II} = 93.6\%$, and DIADEM score = 0.67 in the NeurB1 dataset. Overall, our approach consistently outperforms the commercial software and the other topology estimation methods by a rather large margin with respect to crossover issues. (Note: we quote the visual and numerical results reported in [11] in the belief that their results were the best achievable.)

Table 2. Performances of the proposed method in identifying connectivity at junctions from the **WIDE** and **RICE** datasets.

Dataset	# Junction	# Correctly Identified	Acc_I	Acc_{II}
WIDE	4908	4743	96.6%	93.6%
RICE	962	949	98.7%	98.9%

Table 3. Topology reconstruction performances of different methods from the **WIDE** and **RICE** datasets.

Dataset	Method	s_p	s_f	s_a	s_r
WIDE	Human	0.988	0.991	0.826	0.940
	Ours	0.971	0.980	0.788	0.901
	HSA	0.966	0.972	0.732	0.860
	GLA	0.906	0.920	0.571	0.748
RICE	Ours	0.988	0.992	0.921	0.984
	HSA	0.983	0.991	0.898	0.972
	GLA	0.941	0.958	0.795	0.895

4.3. Results on the WIDE and RICE datasets

Figures 4 and 5 illustrate the topology reconstruction results on the WIDE and RICE datasets. Again, we first measure the performance of the proposed method in identifying connectivity in the WIDE and RICE datasets, as shown in Table 2, by counting the number of correctly identified junctions. As expected, the accuracy of the topology reconstruction at vessel centerline level (i.e. $Acc_{II} = 93.6\%$ in WIDE dataset) is lower than the result obtained at junction level ($Acc_I = 96.6\%$). This is because the number of curvilinear segments is much larger than the number of junctions.

In order to provide an objective evaluation, we have compared the proposed method with a Heuristic Search Algorithm (HSA) [14] that efficiently explores the space of possible trees starting from the Greedy Linear-time Algorithm (GLA) [14], and these are the only two methods in the literature that have reported topology reconstruction performances on these two datasets. For the sake of fair comparison, four weighted scores were obtained to measure the similarity between each reconstructed tree and the ground-truth tree: parent similarity (s_p), flow similarity (s_f), absolute similarity (s_a), and relative similarity (s_r). The first two demonstrate local performance in connectivity, while the latter two capture more global differences. For more detail, we refer the reader to [14].

Table 3 reports the comparison of our method with the state-of-the-art topology reconstruction methods on the WIDE and RICE datasets. For the WIDE dataset, the inter-observer scores were reported in [14]. It can be seen that the similarity scores of the proposed method are very close to those of the human observer. All three methods - the proposed method and its two competitors - achieve superior performances on the RICE dataset. That is because the roots of rice plants have fewer branches than the other datasets, and these root structures have a tendency to grow downwards through the ground (right direction of Figure 5), giv-

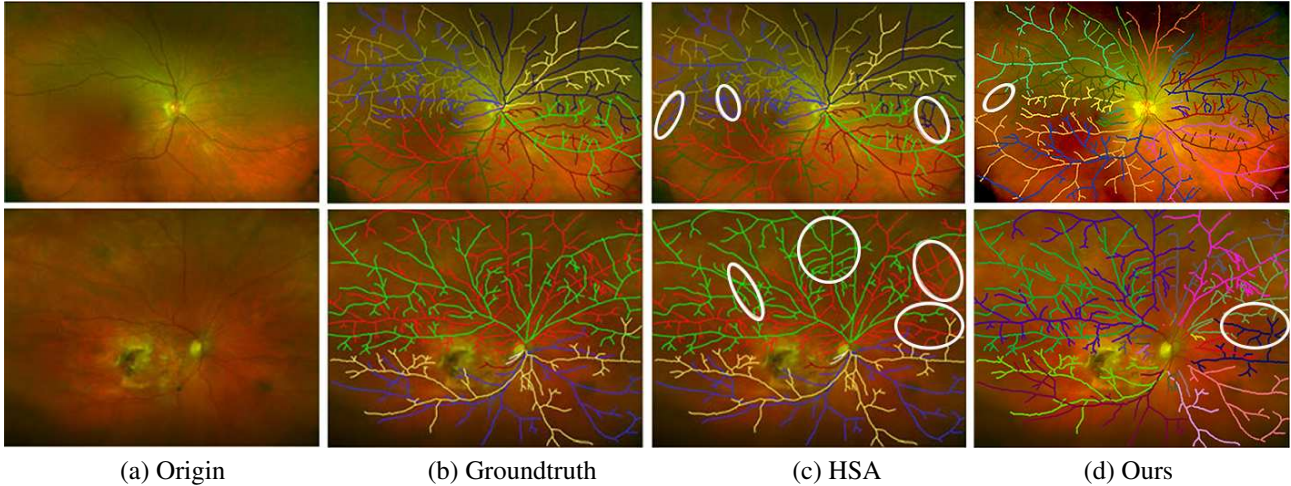|			
(a) Origin	(b) Groundtruth	(c) HSA	(d) Ours

Figure 4. Two examples of topology reconstruction by different methods on retinal vessels from the **WIDE** dataset. The white circles indicate the incorrectly identified junctions.

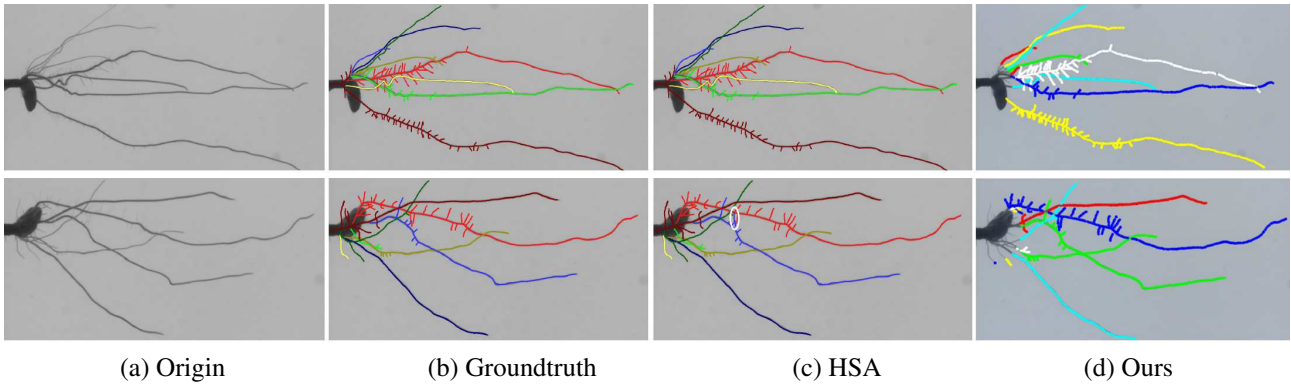|			
(a) Origin	(b) Groundtruth	(c) HSA	(d) Ours

Figure 5. Two examples of topology reconstruction by different methods on plant roots from the **RICE** dataset.

ing a radially symmetrical appearance. In consequence, the less variable images tend to lead to more accurate results in the topology estimation task.

5. Conclusions

The topology reconstruction from a single 2-D image of a 3-D tree-structure is challenging, since the full spatial location of each tree branch is lost after projection. In this paper, we have proposed a novel topology estimation method for tree-like structures, formulated as a pairwise clustering problem. First, a novel curvilinear structure similarity measurement is proposed, then it is utilized to guide the dominant-set clustering approach to identify the connectivity at the junctions, before finally achieving an accurate topology estimation of the tree network.

It is worth noting that by using this method the number of points processed in the course of identifying a dominant set is greatly reduced, from the number of pixels in an image to the number of junctions in a tree network. This leads to an increase of time-efficiency in topology estimation. The

competitiveness of our approach is demonstrated by the fact that it accurately reconstructed tree topologies from four publicly accessible biomedical image datasets with different imaging modalities. The results demonstrated that our method achieves superior performances when compared directly with the existing state-of-the-art ones. It is believed that the proposed method could be a powerful tool for analyzing tree-like structures in the biomedical images.

6. Acknowledgments

References

[1] L. Band et al. Multiscale systems analysis of root growth and development: Modeling beyond the network and cellular scales. *The Plant Cell Online*, 24(10):3892–3906, 2012. 1

[2] P. Bankhead, J. McGeown, and T. Curtis. Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PLoS ONE*, 7:e32435, 2009. 3

[3] S. Basu et al. Neurite tracing with object process. *IEEE Trans. Med. Imaging*, 35(6):1443–1451, 2016. 1

[4] H. Bogunovic et al. Anatomical labeling of the circle of willis using maximum a posteriori probability estimation. *IEEE Trans. Med. Imag.*, 32(9):1587–1599, 2013. 2

[5] K. Brown et al. The DIADEM data sets: Representative light microscopy images of neuronal morphology to advance automation of digital reconstructions. *Neuroinformatics*, 9(2-3):143–157, 2011. 2

[6] J. P. Carson, D. R. Einstein, and R. A. Corley. High resolution lung airway cast segmentation with proper topology suitable for computational fluid dynamic simulations. *Comp. Med. Imag. and Graph.*, 34(7):572–578, 2010. 2

[7] X. Chen et al. Sketch-based tree modeling using markov random field. *ACM Trans. Graph.*, 27(5):109:1–109:9, 2008. 7

[8] L. Cheng, J. De, X. Zhang, F. Lin, and H. Li. Tracing retinal blood vessels by matrix-forest theorem of directed graphs. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2014*, pages 626–633, 2014. 2

[9] M. Cheng et al. Global contrast based salient region detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(3):569–582, 2015. 5

[10] B. Dashtbozorg et al. An automatic graph-based approach for artery/vein classification in retinal images. *IEEE Trans. Image Processing*, 23(3):1073–1083, 2014. 2, 3

[11] J. De et al. A graph-theoretical approach for tracing filamentary structures in neuronal and retinal images. *IEEE Trans. Med. Imaging*, 35(1):257–72, 2016. 1, 2, 6, 7

[12] J. De et al. Transduction on directed graphs via absorbing random walks. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40(7):1770–1784, 2018. 2

[13] R. Estrada and et al. Retinal artery-vein classification via topology estimation. *IEEE Trans. Med. Imaging*, 34(12):2518–2534, 2015. 2

[14] R. Estrada et al. Tree topology estimation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 37(8):1688–1701, 2015. 1, 2, 4, 6, 7

[15] D. Fan et al. Structure-measure: A new way to evaluate foreground maps. In *IEEE International Conference on Computer Vision, ICCV 2017*, pages 4558–4567, 2017. 5

[16] M. Fraz and et al. Blood vessel segmentation methodologies in retinal images - a survey. *Comput. Meth. Prog. Bio.*, 108:407–433, 2012. 1

[17] T. A. Gillette, K. M. Brown, and G. A. Ascoli. The DIADEM metric: Comparing multiple reconstructions of the same neuron. *Neuroinformatics*, 9(2-3):233–245, 2011. 7

[18] P. Glowacki et al. Reconstructing evolving tree structures in time lapse sequences by enforcing time-consistency. *IEEE Trans. Pattern Anal. Mach. Intell.*, 40(3):755–761, 2018. 2

[19] R. Grothausmann et al. Method for 3d airway topology extraction. *Comp. Math. Methods in Medicine*, 2015:127010:1–127010:7, 2015. 2

[20] J. J. Koenderink. *Solid shape*. MIT Press, 1990. 5

[21] A. Mosinska-Domanska et al. Active learning for delineation of curvilinear structures. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016*, pages 5231–5239, 2016. 4

[22] K. P. Murphy, Y. Weiss, and M. I. Jordan. Loopy belief propagation for approximate inference: An empirical study. In *Proc. ICUAL*, pages 467–475, 1999. 7

[23] M. Pavan and M. Pelillo. Dominant sets and hierarchical clustering. In *IEEE International Conference on Computer Vision (ICCV 2003)*,, pages 362–369, 2003. 3, 4

[24] M. Pavan and M. Pelillo. Dominant sets and pairwise clustering. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(1):167–172, 2007. 4

[25] T. Qureshi and B. Al-Diri. A bayesian framework for the local configuration of retinal junctions. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2014*, pages 3105–3110, 2014. 1

[26] M. Radojevic and E. Meijering. Automated neuron tracing using probability hypothesis density filtering. *Bioinformatics*, 33(7):1073–1080, 2017. 1, 4

[27] D. Robben and et al. Simultaneous segmentation and anatomical labeling of the cerebral vasculature. *Medical Image Analysis*, 32:201–215, 2016. 2

[28] P. Su et al. Exploiting data reliability and fuzzy clustering for journal ranking. *IEEE Trans. Fuzzy Systems*, 25(5):1306–1319, 2017. 6

[29] Z. Wang et al. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Processing*, 13(4):600–612, 2004. 4, 5

[30] J. Xie, Y. Zhao, and Y. Wang. Retinal vascular topology estimation via dominant sets clustering. In *IEEE International Symposium on Biomedical Imaging, ISBI 2018*, pages 1458–1462, 2018. 1

[31] J. Zhang et al. Reconnection of interrupted curvilinear structures via cortically inspired completion for ophthalmologic images. *IEEE Trans. Biomed. Engineering*, 65(5):1151–1165, 2018. 5

[32] Y. Zhao et al. Automatic 2D/3D vessel enhancement in multiple modality images using a weighted symmetry filter. *IEEE Trans. Med. Imaging*, pages 1–1, 2017. 1

[33] Y. Zhao et al. Retinal artery and vein classification via dominant sets clustering-based vascular topology estimation. In *International Conference on Medical Image Computing and Computer Assisted Intervention - MICCAI 2018*, pages 56–64, 2018. 2, 3

[34] Y. Zhao, S. P. Harding, and Y. Zheng. Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images. *IEEE Trans. Med. Imaging*, 34(9):1797–1807, 2015. 3

[35] Y. Zhao, Y. Zheng, Y. Liu, J. Yang, Y. Zhao, and Y. Wang. Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy. *IEEE Trans. Med. Imaging*, 36(1):51–63, 2017. 5