Cerebral Vascular Tree Matching of 3D-RA Data Based on Tree Edit Distance

W.H. Tang and Albert C.S. Chung

Lo Kwee-Seong Medical Image Analysis Laboratory,
Department of Computer Science and Engineering,
The Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong
{cstommy, achung}@cse.ust.hk

Abstract. In this paper, we present a novel approach to matching cerebral vascular trees obtained from 3D-RA data-sets based on minimization of tree edit distance. Our approach is fully automatic which requires zero human intervention. Tree edit distance is a term used in the field of theoretical computer science to describe the similarity between two labeled trees. In our approach, we abstract the geometry and morphology of vessel branches into the labels of tree nodes and then use combinatorial optimization strategies to compute the approximated edit distance between the trees. Once the optimal tree edit distance is computed, the spatial correspondences between the vessels can be established. By visual inspection to the experimental results, we find that our approach is accurate.

1   Introduction

1.1   Why 3D-RA data-sets are used?

Three-Dimensional Rotational Angiography (3D-RA) images are playing an increasingly important role in the study of intracranial vasculatures. Due to its relatively higher voxel resolution and contrasting property than other image modalities, a more accurate surface can be segmented out for further geometric and morphological studies.

1.2   Clinical Motivations

Properties of intracranial aneurysms are under active research in the medical field. Beck et al. [1] and Weir et al. [2] studied the relations between the probabilities of rupture and the size and site of aneurysms. In specifying the locations of aneurysms, they simply used the names of the arteries where the aneurysms were located. We believe that a more accurate and specific coordinates for specifying the locations of aneurysms is needed. In specifying the structure of aneurysms, they usually used the diameters of the aneurysms. However, information about the size and orientation of the vessels where the aneurysms were located was not

considered. We believe that an accurate matching between normal vessel and problematic vessels can help locate and quantify the position of defection and help give more information to assess the chance of aneurysm rupture.

1.3   Previous Work

The goal of vascular matching is to find spatial correspondences between two sets of vessels. If they are acquired from the same patient, the problem can be solved easily by using rigid registration. We are therefore focusing on vascular mapping across different patients. This problem is ill-posed since the underlying transformation is unknown and the vascular structures can vary significantly across patients. Up to now, there is no standardized procedure for quantifying the goodness or error of a matching. There is not much research on vascular mapping in the literature. We present some previous work below.

Antiga & Steinman [3] suggested a methodology for mapping bifurcating vessels. Vessel junction was decomposed into a geometric reference system. However, the emphasis is on the bifurcating junction, but not on the whole vascular tree. Cool et al. [4] presented a tissue-based registration approach to forming vascular density atlas. However, the registration was based on brain tissues but not the real geometry of the vessels. It only maps the position of the vessels with respect to the embedded brain tissue. The geometry and morphology of the vessels, such as radius and orientation information, were not used in the registration process. Chillet et al. [5] suggested a method to form a vascular atlas using vessel-to-image affine registration. However, the affine registration used may not be sufficient to represent the underlying deformation to match the vessels, especially for those which are highly curved and rolled. Jomier & Aylward [6] presented a model-to-image registration approach to mapping vessels. In their work, a global rigid transformation was performed first followed by local and piece-wise deformation of branches via propagation. However, if the structural difference between the vessels is significant, this method will not work since a mismatch of branch points will propagate to the children branches.

Therefore, a global matching method which makes use of the geometry and morphology of the vessels is desirable.

1.4   Outline

In our method, we abstract the geometry and morphology of vessel branches into the labels of tree nodes. By doing so, we are able to use a tree edit distance approach to find the best matching in a global perspective but without ignoring the geometric and morphological properties of each branch. We will then use combinatorial optimization strategies to compute the approximated edit distance between the trees. Once the optimal tree edit distance is computed, the spatial correspondences between the vessels can be found. Detailed procedures will be explained in Section 2. In Section 3, experimental results will be shown and explained. Finally, in Section 4, a conclusion will be made and possible future work will be suggested.

2 Methodology

2.1 Centreline Extraction

The first step is to obtain the centreline from a pre-segmented 3D-RA data-set. Since we are interested in a tree representation of centreline. We use a Voronoi diagram-based approach. The Voronoi diagram of the surface points on the segmented vessel is computed. It is a graph $G = (V, E, W)$, where each node $v \in V$ is representing a point in $\mathbb{R}^3$, each edge $e \in E$ is representing a line segment in $\mathbb{R}^3$. $W : E \to \mathbb{R}$ is a weight assignment function to each edge. Once W is defined, we can find the weighted shortest path tree $G' = (V, E')$ of G where $E' \subset E$. We set $W(e)$ to be the reciprocal of the average Euclidean distance from the edge e to the segmented surface. It is not difficult to observe that by using this weight assignment function, for every pair of nodes $v_1, v_2 \in V$, the unique simple path between v_1 and v_2 defined in G' is one of the many paths between v_1 and v_2 defined in G; and this simple path tends to keep the distance from the segmented surface as large as possible throughout it. The next step is to detect endpoints of the vessel. This can be done automatically as the major vessels are usually intersecting with the bounding box of the 3D-RA data-set, since 3D-RA considers only the region of interest in the brain. Once the endpoints are detected, the centreline is just the union of all paths from the endpoints to the root in the shortest path tree G'. Note that the root is just an endpoint where the internal carotid artery intersects with the bounding box of the 3D-RA data-set, which can also be detected easily, because of its large cross-sectional radius. Besides the centreline itself, the cross-sectional radius information of each centreline point, which can be obtained from the Voronoi diagram, is also stored.

2.2 Branch Representation, Similarity and Importance Measure

A branch in a vascular tree is a vessel whose both ends are either a vessel junction or an endpoint on the bounding box of the 3D-RA data-set, with no other junctions in between. The whole vascular tree can be represented by a union of all branches, which intersect at junctions. Since we will abstract the geometry of branches into nodes in a theoretical tree, before outlining the tree edit distance algorithm, we first present the *similarity function* between two different branches and the *importance function* of a branch, which are equivalent to node relabeling cost and node removal cost respectively in the theoretical tree.

A branch b can be represented by two parametric functions in range [0,1], $\mathbf{x}_b(s)$ and $r_b(s)$. $\mathbf{x}_b(s)$ is a vector function which represents the normalized arc length parameterized centreline curve of b, with $\mathbf{x}_b(0)$ and $\mathbf{x}_b(1)$ representing the end closer to the root and the end further away from the root respectively. $r_b(s)$ is a scalar function which represents the radius of b along the centreline curve $\mathbf{x}_b(s)$, with the same parameterizations, i.e., $r_b(t)$ is the radius associated with the centreline point $\mathbf{x}_b(t)$, where $t \in [0, 1]$.

The similarity function (matching cost) $S(b_1, b_2)$ of two branches b_1 and b_2 with length l_1 and l_2 respectively is defined as,

$$S(b_1, b_2) = \alpha S_1(b_1, b_2) + \beta S_2(b_1, b_2) + \gamma S_3(b_1, b_2) \tag{1}$$

where

$$S_1(b_1, b_2) = \frac{l_1 + l_2}{2} \int_0^1 |r_{b_1} - r_{b_2}| \, ds \qquad (2)$$

is the cost of matching the radii of the two branches b_1 and b_2; and,

$$S_2(b_1, b_2) = \frac{l_1 + l_2}{2} \int_0^1 \left(\frac{r_{b_1} + r_{b_2}}{2} \cdot \frac{1}{\pi} \cos^{-1} \frac{\mathbf{x}'_{b_1} \bullet \mathbf{x}'_{b_2}}{\|\mathbf{x}'_{b_1}\| \cdot \|\mathbf{x}'_{b_2}\|} \right) ds \qquad (3)$$

is the cost of matching the orientations of the two branches b_1 and b_2, weighted by their average radius; and,

$$S_3(b_1, b_2) = |l_1 - l_2| \qquad (4)$$

is the cost of matching the lengths of the two branches.

α, β and γ are weighting parameters and we set $\alpha = \beta = \gamma = 1$ in our experiment. Note that $S(b_1, b_2)$ is zero if and only if the two branches b_1 and b_1 are identical, otherwise $S(b_1, b_2) > 0$.

The importance function (removal cost) $D(b)$ of a branch b with length l is defined as,

$$D(b) = l \int_0^1 r_b \, ds \qquad (5)$$

which can be viewed as the surface area of the branch.

2.3 Tree Edit Distance

With the definition of matching cost and removal cost of branches in the previous sub-section, now we can abstract the geometry of the branches into nodes in a theoretical tree, and two nodes are connected together if and only if the two branches they represent form a junction. Therefore the theoretical tree formed can be viewed as a dual graph of the geometric tree of the centreline. For an example, please see Fig. 1. Note that since it is difficult to assign an order to the children of a node in a 3D tree, the theoretical tree is unordered, i.e, the order of the children for any node is not important, but only the parent-child relationship. Once we have it, everything about tree edit distance is manipulated on the theoretical tree. The returned result will be the set of nodes to be removed and the correspondences between the unremoved nodes of the two theoretical trees, which are respectively equivalent to the set of branches to be removed in order to yield the best matching and the correspondences between the unremoved branches in the two vascular structures.

We will apply the algorithm in [7] to perform unordered tree matching. Let T_1 and T_2 be two trees with sets of nodes N_1 and N_2 respectively. The paper suggested a concept called *marking*. A marking $K = (S_1, S_2)$ is two sets of nodes $S_1 \in N_1$ and $S_2 \in N_2$ to be removed from T_1 and T_2 respectively. Removing a node n means setting the parent of all the children of n to be the parent of n and then ignoring n. Note that for a geometric vascular tree, removing a branch may
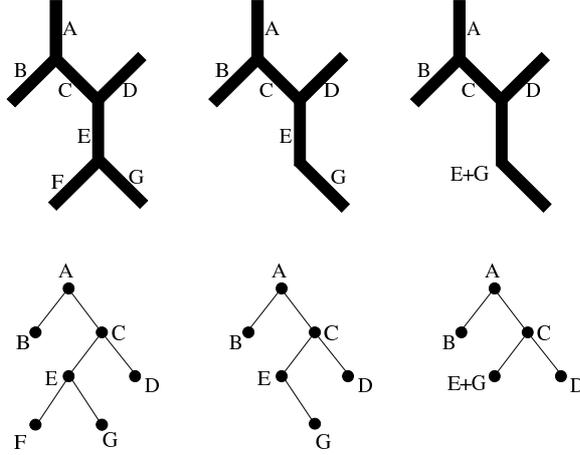
Fig. 1. The top row shows three geometric trees with the branches labeled and the bottom row shows the corresponding theoretical trees with the nodes labeled. The left column shows the trees before removing F, the middle column shows the trees after removing F but before joining E and G, the right column shows the trees after joining E and G.

cause two existing unremoved branches join together to form a longer branch; in the theoretical tree context, this is handled by joining the node which has only one child into its parent. For a visual example, please see Fig. 1. A marking $K = (S_1, S_2)$ is legal if and only if the resultant trees after the removal of S_1 and S_2, denoted by $K(T_1)$ and $K(T_2)$ respectively, are isomorphic, i.e., there exists a one-to-one correspondence of all nodes between $K(T_1)$ and $K(T_2)$ such that parent-child relationship is preserved in the mapping. Note that since we are considering unordered trees, the number of such isomorphic mappings can be more than one in a marking. The removal cost of a marking $K = (S_1, S_2)$, $D(K)$, is defined as,

$$D(K) = \sum_{s \in S_1} D(s) + \sum_{s \in S_2} D(s) \tag{6}$$

which is the total removal costs for all removed nodes. The matching cost of a marking K, $S(K)$, is defined as,

$$S(K) = \min_{im \in IM} \sum_{(n_1, n_2) \in im} S(n_1, n_2) \tag{7}$$

where IM is the set of all isomorphic mappings for $K(T_1)$ and $K(T_2)$, (n_1, n_2) is a corresponding pair of nodes in the isomorphic mapping im, and $S(n_1, n_2)$ is the matching cost between the nodes n_1 and n_2. Also, $S(K)$ is ∞ if K is not a legal marking. So $S(K)$ is the total matching cost of all corresponding pairs in the best isomorphic mapping between $K(T_1)$ and $K(T_2)$. To find $S(K)$, we use the bottom-up matching algorithm suggested in [7], which uses bipartite
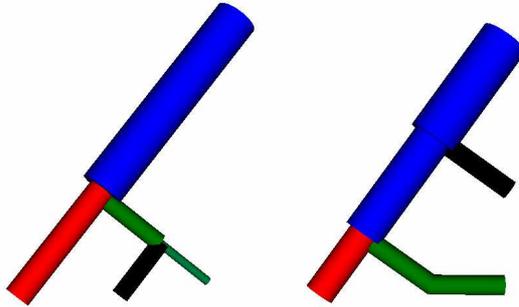
Fig. 2. *(Color image)* The two figures are two synthetic simple structures to be matched. The branches in black color are those to be removed in order to obtain the best matching. Two branches of same color across the figures are matched branch pairs.

matchings to find the best mapping between the children sets of the two nodes. Finally, the total tree edit distance between two trees T_1 and T_2, $Dist(T_1, T_2)$ is then defined as,

$$Dist(T_1, T_2) = \min_{K \in 2^{N_1} \times 2^{N_2}} [D(K) + S(K)] \qquad (8)$$

which is the minimum of the sum of the delete cost and matching cost over all possible markings. Note that $2^{N_1} \times 2^{N_2}$ in the equation is the cartesian product of the two power sets, which is the set of all possible markings. Since it is exponential to the total number of nodes in the two trees, it is computationally infeasible to find the exact tree edit distance which requires an exhaustive search of all possible markings. Instead, we use an iterative improvement approach suggested in [7], which randomly adds or removes a node from either or both trees to find an approximated solution by a downhill optimization. Since this approach is stochastic and easily gets trapped in local minima, it is executed several times and the best solution is taken.

3 Experimental Results

We first apply our method on a pair of simple synthetic tubular structures. Since the structures just consist of a few branches, we use the exhaustive search method to enumerate through all possible markings. Fig. 2 shows the two structures and the matching result. The branches in black color are those to be removed in order to obtain the best matching. Branches of same color across the two structures are matched branch pairs. It shows that the matching result agrees with the intuitive mapping from visual inspection.

Then we apply our method to two sets of real 3D-RA data from two different patients. There are more than 30 total branches in the data-sets, making ex-
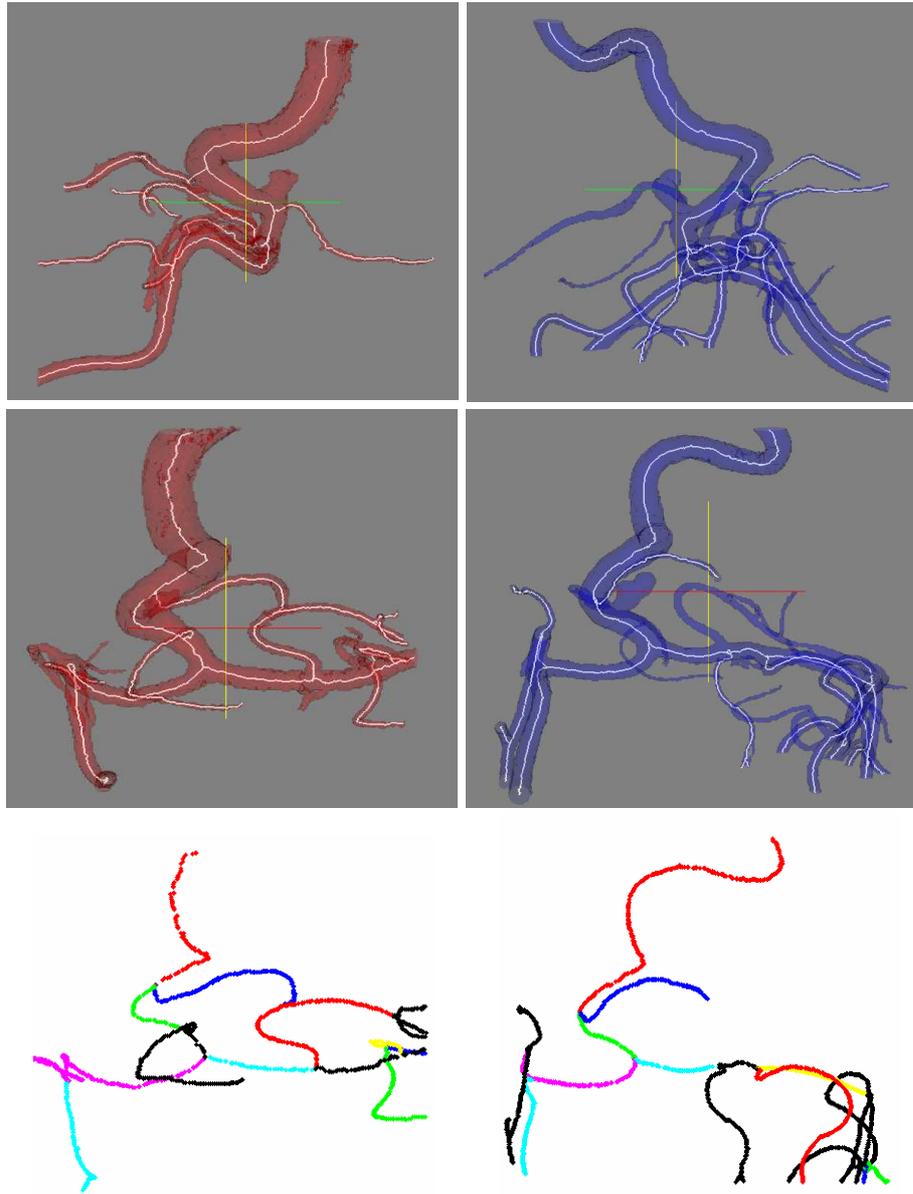
Fig. 3. *(Color image)* The left and right columns respectively show the vascular data of two different patients. The top and middle rows are two different views of the segmented surfaces superimposed with the centerlines. The bottom row shows the matching result of our method. The branches in black color are those to be removed in order to obtain the best matching. Two branches of same color across the figures are matched branch pairs (Some colors are used twice in the same figure).

haustive search impossible. So we execute the stochastic iterative improvement method several times and take the best result. Fig. 3 shows the data and the result. We can see that the major branches are matched correctly. For other branches, the matching is consistent with their relative positions, although we do not apply any positional mismatch penalty in our cost function. Therefore by visual inspection, our method is accurate. This is our current research work to further test our method on a larger set of 3D-RA data.

4   Conclusion

We have presented a novel method for matching vascular trees. Our method is based on the concepts of theoretical tree edit distance. The geometry and morphology of the vascular branches are abstracted into the nodes of the theoretical trees. The *branches similarity* and *importance measures* are incorporated into the nodes relabeling and removal costs of the theoretical trees. These make our method able to perform the matching in a global perspective but still able to consider the geometry of every branches. Experimental results show that our method can work on both synthetic data and two sets of segmented 3D-RA vessels from different patients. In the future, we will use different optimization strategies to improve the suboptimal solution and perform experiments on more real data-sets.

References

1. Beck, J., Rohde, S., Berkefeld, J., Seifert, V., Raabe, A.: Size and location of ruptured and unruptured intracranial aneurysms measured by 3-dimensional rotational angiography. Surg. Neurol. **65** (2006) pp. 18–25
2. Weir, B., Disney, L., Karrison, T.: Sizes of ruptured and unruptured aneurysms in relation to their sites and the ages of patients. J. Neurosurg. **96** (2002) pp. 64–70
3. Antiga, L., Steinman, D.A.: Robust and objective decomposition and mapping of bifurcating vessels. IEEE Trans. on Med. Img. **23** (2004) pp. 704–713
4. Cool, D., Chillet, D., Kim, J., Guyon, J.P., Foskey, M., Aylward, S.: Tissue-based affine registration of brain images to form a vascular density atlas. MICCAI (2003) pp. 9–15
5. Chillet, D., Jomier, J., Cool, D., Aylward, S.: Vascular atlas formation using a vessel-to-image affine registration method. MICCAI (2003) pp. 335–342
6. Jomier, J., Aylward, S.R.: Rigid and deformable vasculature-to-image registration: A hierarchical approach. MICCAI (2004) pp. 829–836
7. Shasha, D., Wang, T.L., Zhang, K., Shih, F.Y.: Exact and approximate algorithms for unordered tree matching. IEEE Trans on Systems, Man, and Cybernetics **24** (1994) pp. 668–678