BIFURCATION DETECTION IN 3D VASCULAR IMAGES USING NOVEL FEATURES AND RANDOM FOREST

Mengliu Zhao and Ghassan Hamarneh

Medical Image Analysis Lab, School of Computing Science, Simon Fraser University, Canada
{mengliuz, hamarneh}@sfu.ca

ABSTRACT
Bifurcation detection is important in medical image analysis for mainly two reasons: 1) plaques are easy to accumulate at artery bifurcations, which leads to atherosclerosis and strokes; 2) for quantification (e.g., branch length, thickness, tortuosity), visualization, and blood flow simulation, it’s necessary to extract all the branches and their connectivity in a vessel tree, which makes bifurcation localization crucial. In this paper, several novel features are designed for classifying bifurcations in 3D vascular images using random forest. Encouraging results with both synthetic and real datasets are obtained.

Index Terms— bifurcation detection; 3D vascular images; random forest; classification

1. INTRODUCTION

According to World Health Statistics 2012 [16], cardiovascular diseases have lead to the largest proportion (48%) of non-communicable deaths in the world. One of the common cardiovascular diseases is atherosclerosis. When atherosclerosis plaques accumulate inside the artery, the walls become fragile and might cause strokes and cardiac-related deaths. The plaques are more apt to happen at bifurcation locations, thus making it important to detect bifurcations in vascular images. It also matters in problems like blood vessel quantification (e.g., measuring branch length, thickness, tortuosity), visualization, and blood flow simulation, to detect all the bifurcation locations and extract the whole set of vessels [13].

Many bifurcation extraction methods have been proposed in recent years. For 2D retinal fundus images, a classification method was proposed by Saha et al. [17], counting local connectivity in the binarized image. For 3D images, Wette et al. [18] improved the Corkscrew algorithm for bifurcation extraction in coronary arteries, which is a semi-automatic region growing algorithm requiring users input such as start/end points and intensity value bounds. In [11], a bifurcation is treated as a junction: first the vessels are tracked from a single seed given by the user, a skeletonization step is applied, then junctions are detected by examining each pixel’s 26-neighborhood. Brozio et al. [5] evaluated bifurcations by the lengths and shapes of the minimal paths among vessel candidates. Baboiu et al. [3] designed a bifurcation filter based on modeling bifurcations in scale space as the superposition of three straight vessels. CETIN et al. [6] defined a new vessel tensor and detected bifurcations by clustering the direction vectors generated by the tensors. In [15], the authors proposed a vessel crawler organism model, in which image features and prior anatomical information is fed into an artificial cognitive layer that controls the crawling deformations, detection of bifurcations, and spawning of child vessel crawlers. In [4], carotid bifurcations are detected by mapping surrounding vessel branch information to a known anatomical database.

As pointed out by Lee et al. [12], the asymmetric characteristic of bifurcations made it hard to fix parameters for most algorithms. This has given rise to some machine learning based approaches. In [1], the authors proposed to classify bifurcations in 2D sequential IVUS images using features based on, e.g., Gabor filters and cross-correlation; comparisons were made across different classifiers, e.g., AdaBoost, random forest and Support Vector Machines (SVM). In [20], Gaussian and its 1st and 2nd derivatives on cross-sectional planes were used for detecting bifurcations with AdaBoost.

However, in the aforementioned machine learning based methods, features in [1] were specially designed for IVUS images and hard to be adapted to other modalities. The novelty of this paper lies in the three features designed for bifurcation detection in 3D CT images. In Section 2, our new features, as well as four others, are combined as input of a random forest classifier for building a bifurcation detector, which is further incorporated in a general tracking pipeline (Algorithm 1). In Section 3, two datasets, VascuSynth1 and VESSEL12 2 are examined. It is shown that our proposed features outperform those in [20] by a large margin in cross-validation. Both numerical and visual results show that the proposed bifurcation detector gives superior result to those obtained by the method of Macedo et al. [14].

1http://vascusynth.cs.sfu.ca
2http://vessel12.grand-challenge.org/
2. METHODOLOGY

In this section, we detail three novel features as well as summarize existing ones that we adapted and used for bifurcation detection via a random forest classifier, which is eventually implemented within a general vessel tracking algorithm.

Von Mises-Fisher (VMF) Distribution. VMF distribution is a directional distribution defined on a p-dimensional unit hyper-sphere \( S^{p-1} \) (p \( \geq 2 \)) [8]:

\[
P(X) = c_p(\kappa)\kappa^p e^{|\theta^T X|}, \quad X \in S^{p-1},
\]

where \( \theta \) stands for the mean direction, \( \kappa \) is the concentration coefficient and \( c_p(\kappa) \) is a normalizing constant. Similar to the definition of Gaussian Mixture Model (GMM), a mixture of \( C \) VMF distributions could be written as follows:

\[
p(X|\Theta) = \sum_{i=1}^{C} \alpha_i p(X|\theta_i),
\]

where \( \alpha_i \) and \( \theta_i \) represent mixture weights and the corresponding mean directions, respectively for \( i = 1, ..., C \).

For each pixel \((x, y, z)\) in a 3D image, we define \( N_k(x, y, z) \) as the \( k \times k \times k \) neighborhood and \( \vec{v}(x', y', z') \) as the vessel direction at each pixel \((x', y', z')\) in \( N_k(x, y, z) \). Then the observation data for this mixtures distribution (2) is \( \{\vec{v}(x', y', z') : (x', y', z') \in N_k(x, y, z)\} \): all parameters could be calculated using Expectation-Maximization (EM) [8]. Once mean directions \( \vec{v}_1, ..., \vec{v}_C \) are obtained, we measure the differences between each pair of directions by calculating a \( C \times C \) matrix \( \Theta \), where \( \Theta_{ij} = (\theta_i \cdot \theta_j) / (|\theta_i||\theta_j|) \). Standard deviation of the elements of matrix \( \Theta \) is denoted as \( \theta_{std} \), for measuring the coherence of the vessel directions. We’re using both \{\alpha_i : i = 1, ..., C\} and \( \theta_{std} \) as our VMF variable. For a neighborhood containing a single vessel segment, the weights would ideally contain two peak values (positive and negative vessel directions) and \( \theta_{std} \) is relatively large (since the direction cosine matrix is sparse); while for a bifurcation, the six weight values would be more homogeneous and \( \theta_{std} \) would be small (direction cosine matrix is denser).

Histogram of Eigenvectors (HOE). Inspired by the Histogram of Gradient (HOG) feature [7], we propose a similar feature but using estimated vessel directions instead. For the pixel \((x, y, z)\) and its neighborhood \( N_k(x, y, z) \), the HOE feature is obtained through the following steps:

1. Calculate vesselness and vessel direction (eigenvector of Hessian with the smallest eigenvalue), for every pixel in \( N_k(x, y, z) \), as in [9];

2. Construct a spherical histogram by casting a vote for each eigenvector on the sphere surface and weight it by the corresponding vesselness [9];

3. Normalize each bin count on the sphere surface so the total count adds up to 1;

4. Define HOE \((x, y, z)\) as the vector whose elements are only the largest \( C \) values of the sorted bin counts.

Similar to VMF, the HOE of a bifurcation neighborhood would have six (antipodally symmetric) peaks versus only two for a vessel segment.

Concurrency of Eigenvectors (COE). To ensure that the bifurcation is at the center of the ROI, the COE feature (with value in [0,1]) is designed as follows:

\[
COE(X) = K \left( \int_{X' \in N_k(X)} |\vec{v}(X) \cdot \vec{v}(X')|/|\vec{v}(X')| dX' \right)
\]

\[X = (x, y, z), \quad X' = (x', y', z') \quad \text{and} \quad K = \left( \int_{X' \in N_k(X)} dX' \right)^{-1}
\]

Equation (3) will reach its maximum when the bifurcation is right in the center of the ROI and decrease elsewhere.

Other Existing Features for 3D Bifurcation Detection. The following features are also included when training the random forest for bifurcation detection: (1) Spherical Shell Intersection (SSI): This feature is originally described in [19] as the number of connected components between a spherical shell model (composed by two concentric spheres) and its neighborhood \( N_k(x, y, z) \); (2) Convex Hull Ratio (CHR): This feature is defined similar to the “junction-ness degree” in [2], only here we are using vesselness above a certain threshold, \( T_{th} \), to determine foreground pixels; (3) Principal Component Analysis (PCA): The PCA feature vector includes the three eigenvalues of the covariance matrix of all vessel directions in the neighborhood; (4) Scale: the Gaussian scale with total count adds up to 1; (5) Hull area of all connected components.

In summary, our feature vector that is used to train the random forest is of length \( 2C + 8 \).

For fairness in comparing with other bifurcation detection methods, we propose a simplified tracking framework (Algorithm 1) similar to [6]. Note that the bifurcation detection step (Line 8) can be implemented via our or a competing method. The daughter branch detection (Line 10) is implemented using method in [14]. Centralization step (Line 16) is performed using the method in [6]. The vessel end detection (Line 18) is obtained by thresholding the image intensity by \( Th_{1} \).

3. EVALUATION

Our method is evaluated on two datasets, VascuSynth [10] (synthetic, 3D) and VESSEL12 (clinical CT chest images, 3D), against two competing algorithms, [20] and [14]. In all the following experiments, we fix these parameters by empirical tests: \( \alpha = 0.5, \beta = 0.5, \sigma = 5 (\sigma = 50 \text{ for VESSEL12} \).
Algorithm 1: Vessel Tracking Algorithm

**Input**: volume image, seed point, initial tracking direction and trained classifier

**Output**: bifurcation set, centerline set

1. **Initialization**: active set (AC) = \{seedpoint\}, bifurcation set (BF) = ∅, centerline set (CL) = ∅, vesselness, scale and vessel direction

2. while AC ≠ ∅ do
   3. Active point \leftarrow AC;
   4. New tracking starts;
   6. while Current tracking in progress do
      7. Trace along vessel direction;
      8. Classify newly detected point;
      9. if Bifurcation detected then
         10. Detect start points (SP) of two daughter branches;
         11. if Two daughter branches found then
             12. AC \leftarrow SP;
             13. BF \leftarrow active point;
             14. End current tracking
         15. else Continue with current tracking;
      16. Centralization;
      17. CL \leftarrow active point;
      18. if End of vessel reached then Stop;
     19. else Continue with current tracking;

In our first experiment (Figure 1), we evaluate classification performance via leave-one-out cross-validation. Although AdaBoost was used in [20], only random forest is used in this experiment since we want to focus on the contribution of the proposed features. To this end we collected training samples, where each sample is a subvolume \((11 \times 11 \times 11\), automatically located in VascuSynth data or manually picked from VESSEL12 data). For VascuSynth, 100 non-bifurcation samples and 100 bifurcation samples are used. For VESSEL12 dataset, 86 non-bifurcation samples and 105 bifurcation samples are used. 200 decision trees are trained each time, with bagging ratio as 0.8 and minimal leaf number as 4.

In Figure 2, the two sub-figures in the top row show the cumulative histogram of distances from detected bifurcation (DB) points to the closest ground truth (GT) bifurcation location, while the two sub-figures in the bottom row show the accumulated histogram of distances from GT to DB. The first column shows results of VascuSynth – the fact that curves are steeper for our proposed algorithm implies a higher accuracy over the competing algorithm. The second column shows results of VESSEL12 dataset, although the two curves almost overlap at the top, the red is much steeper than the blue at the bottom, which shows a lower tendency of detecting false positives than the competing algorithm. Figure 3 shows a visual comparison between tracking results using two bifurcation detectors, green rings indicate errors in bifurcation detection of the competing method of Macedo et al. [14].
Fig. 3: Visual Comparison Between Two Methods: Left – Proposed; Right – Competing Method [14].

4. CONCLUSION

In this paper, we proposed several novel features for detecting bifurcation in 3D vascular images and showed how they improve bifurcation detection and localization accuracy.

Acknowledgement. Partial funding for this research was provided by the National Science and Engineering Research Council (NSERC) of Canada.

5. REFERENCES