Reconstructing Cerebrovascular Networks under Local Physiological Constraints by Integer Programming

Markus Rempfler[a,b,*], Matthias Schneider[b], Giovanna D. Ielacqua[c], Xianghui Xiao[d], Stuart R. Stock[e], Jan Klohs[c], Gábor Székely[b], Bjoern Andres[f], Bjoern H. Menze[a,g]

[a] *Department of Computer Science, TU München, Germany*
[b] *Computer Vision Laboratory, ETH Zürich, Switzerland*
[c] *Institute for Biomedical Engineering, University and ETH Zürich, Switzerland*
[d] *Advanced Photon Source, Argonne National Laboratory, USA*
[e] *Feinberg School of Medicine, Northwestern University, Chicago IL, USA*
[f] *Max Planck Institute for Informatics, Saarbrücken, Germany*
[g] *Institute for Advanced Studies, TU München, Germany*

Abstract

We introduce a probabilistic approach to vessel network extraction that enforces physiological constraints on the vessel structure. The method accounts for both image evidence and geometric relationships between vessels by solving an integer program, which is shown to yield the maximum a posteriori (MAP) estimate to the probabilistic model. Starting from an over-connected network, it is pruning vessel stumps and spurious connections by evaluating the local geometry and the global connectivity of the graph. We utilize a high-resolution micro computed tomography (µCT) dataset of a cerebrovascular corrosion cast to obtain a reference network and learn the prior distributions of our probabilistic model. We perform experiments on micro magnetic resonance angiography (µMRA) images of mouse brains and discuss properties of the networks obtained under different tracking and pruning approaches.

Keywords: vascular network extraction, vessel tracking, cerebrovascular networks, integer programming

*Corresponding author

1. Introduction

Many diseases affect general properties of the cerebrovascular network, examples are arteriosclerosis and dilative vascular malformations changing vessel shape and diameter, but also Alzheimer's and related neuro-degenerative diseases are suspected to affect the general vascularity and global network properties [1, 2]. Studies investigating such diseases freqently use mouse models for experiments and commonly acquire in-vivo cerebrovascular imagery by means of micro magnetic resonance angiography (µMRA). While segmenting and tracing tubular structures is a longstanding field of interest in medical image computing [3, 4, 5, 6], we approach here the wider – and somewhat neglected [7] – problem of extracting the full vascular network from image volumes under consideration of local geometric properties and global constraints of the vascular structure.

Most vessel segmentation techniques rely on tubularity measures or other vessel enhancement filters [3], and then apply rule-based or learned decision algorithms to segment the vessels [5, 6, 8]. Networks – in terms of a graph representing vessels by their centerline, complemented with additional information such as local radii – can be extracted from binary segmentations using morphological operators [9, 10], or by tracking vessels directly by minimal path techniques [11], e.g. by applying a fast marching algorithm [12] or a Dijkstra-like scheme [13]. We point the interested reader to extensive reviews in [5, 6]. In most applications, however, the extracted graphs need further post-processing: Lu et al. [14], for example, incorporated discriminative classifiers that examine local geometrical features of segments into a hierarchical approach for vessel-structure parsing. In order to deal with imperfections in vascular connectivity of extracted networks, Kaufhold et al. [15] discussed a supervised learning approach to gap filling and network pruning, whereas Schneider et al. [16] recently proposed a generative approach to gap in-fill that is guided by a simplified angiogenesis model. While segmentation algorithms are likely to enforce expected local vessel shape and geometry, only few approaches consider both local

properties and global network connectivity when extracting the full network: Jiang et al. [17] incorporated assumptions about vessel diameters (Murray's hypothesis [18]) in a global optimization problem restricted to arterial trees. In a different application, Türetken et al. [19] introduced recently an integer programming approach that evaluates path coherence and connectivity of general curvilinear structures, such as streets in remote-sensing images or vessels in confocal image stacks. Starting from an over-connected graph, they are pruning edges that do not fulfill desired structural relationships of neighbouring segment pairs using a path classifier that is trained from annotated 3-D networks.

All of these approaches enforce local coherence within the extracted network – a general property of the vascular network. More complex local properties of a structural network, however, can be described by network motifs – frequently recurring subgraphs that are characteristic for a type of network [20, 21].

In this paper, we enforce local geometrical properties similar to Jiang et al. [17], exploring the relevance of two basic motifs of vascular networks, i.e., the geometrical properties of continuing segment pairs and of vessel bifurcations.We present a probabilistic model which combines this geometric prior with local vessel evidence obtained from a segmentation algorithm [8], and show that the maximum a posteriori (MAP) estimate can be computed by an integer linear program (ILP). We simply learn the global statistic of desired geometrical properties of the network motifs from a high resolution dataset. Identifying a more efficient scheme for the integer programming problem, we can now scale it to large datasets and illustrate its application for delineating vessels in µMRA images of the murine brain.

2. Methods

In this section, we detail on the proposed vessel network extraction method that estimates the most probable network under consideration of image evidence and physiological prior knowledge. As depicted in the workflow (Fig. 1), this method starts from an overconnected graph G^{over}, hence we briefly review the

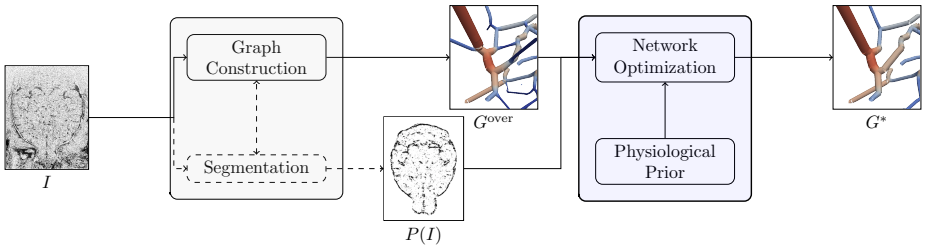


Figure 1: Workflow: In a first stage (gray box), the image volume I is processed so as to obtain an overconnected graph G^{over} as well as a confidence measure for vessels such as the confidence map $P(I)$. In the following step, the network G^* is extracted from G^{over} in a optimization that considers both image evidence (according to $P(I)$) and geometric-physiological prior knowledge. In this paper, we focus on the network optimization step (blue), where both image evidence and geometrical relationships of certain network motifs, namely continuing pairs and bifurcations are considered.

applied segmentation framework and skeletonization method.

2.1. Vessel Segmentation Method and Construction of the Overconnected Graph

As a first stage, we transform image intensities into confidence maps by using the framework of Scheider et al. [22, 8]: In this approach, multiscale steerable filter templates (SFT) are used as efficient directional filters, offering features that are invariant with respect to the local vessel direction. An oblique random forest (RF) [23], which determines splits by solving a linear regression with elastic net penalty in each node, is used for a subsequent classification indicating the local presence of a vessel-like structure. The RF assigns each voxel v in an image volume I to a probability $p_v \in [0, 1]$.

We apply a threshold θ to the probability volume $P(I)$ and skeletonize the resulting binary volume using distance-ordered homotopic thinning (DOHT) [9], a method that iteratively removes voxels without altering the objects topology, to derive a network graph $G(\theta)$. We obtain an over-connected network by generating multiple binary segmentations from $P(I)$ with different thresholds $\{\theta_i\}$, skeletonizing each of them by DOHT to $G(\theta_i)$ and superposing them into one network $G^{\text{over}}(\{\theta_i\})$. The resulting network contains both segments with low

confidence (contributed by graphs from low thresholds θ close to 0), but maintains the high spatial accuracy of a graph that is generated from conservative thresholds (i.e., with θ close to 1).

Note that any method which generates an overconnected graph G^{over} by proposing local vessel connections could be used instead.

2.2. Vessel Network Extraction

The goal of our method is to find the most plausible network G^* out of an overconnected network graph $G^{\text{over}} = (V, E)$ with edges $E = \{e_i\}$ and given image evidence $P(I)$. We encode subgraphs of G^{over} with a set of binary variables $X = \{x_i\}$ where each x_i indicates whether or not the corresponding segment $e_i \in E$ is active (i.e. $x_i = 1$). Therefore, we arrive at the equivalent problem of determining the MAP estimate of $\mathbf{x} \in \{0, 1\}^{|E|}$, for which we describe a probabilistic model (Sect. 2.2.1) that considers image evidence, local properties of specific network motifs as well as global connectivity, and derive an ILP that allows estimating the MAP network (Sect. 2.2.2).

2.2.1. Probabilistic Model

We formulate a probabilistic model according to Fig. 2, where I is the image evidence, G is the given (overconnected) Graph and X is the set of binary variables denoting subgraphs of G. Ω is the set of all feasible solutions of $\mathbf{x}$:

$$\Omega = \{\mathbf{x} \in \{0, 1\}^{|E|} : \mathbf{A}\mathbf{x} \geq \mathbf{b}\} \ , \tag{1}$$

with $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ being the short notation for all hard constraints that will be considered such as, for example, those that enforce connectivity. This introduces a probabilistic interpretation, as in [24], of the hard constraints that we impose on the extracted networks.

According to the given probabilistic model, we arrive at the posterior distribution for $\mathbf{x}$:

$$P\left(\mathbf{X} = \mathbf{x}|I, G, \Omega\right) \propto P\left(\Omega|\mathbf{X} = \mathbf{x}\right) P\left(\mathbf{X} = \mathbf{x}|I, G\right) \ . \tag{2}$$

$$I, G \longrightarrow X \longrightarrow \Omega$$

Figure 2: Probabilistic model. I: Image, G: Graph (overconnected); X: Set of binary variables denoting subgraphs of G; Ω: Set of feasible configurations of $\mathbf{x}$.

Next, we model $P(\mathbf{X} = \mathbf{x}|I, G)$ as a Markov random field (MRF):

$$P(\mathbf{X} = \mathbf{x}|I, G) = \frac{1}{Z} \prod_{x_i \in X} \phi_i\left(x_i; I, G\right) \prod_{\substack{x_i, x_j \in X: \\ e_i, e_j \text{ adjacent} \\ \nexists e_k \text{ adjacent to } e_i \wedge e_j}} \phi_{i,j}\left(x_i, x_j; I, G\right) \prod_{\substack{x_i, x_j, x_k \in X: \\ e_i, e_j, e_k \text{ adjacent}}} \phi_{i,j,k}\left(x_i, x_j, x_k; I, G\right) \quad , \tag{3}$$

where Z is the partition function and $\phi\left(.\right)$ are the potentials, which are defined in the following. For each segment represented by x_i, we set the unary potential:

$$\phi_i\left(x_i; I, G\right) = \begin{cases} P\left(x_i = 1|I\right) & \text{if } \; x_i = 1 \; , \\ P\left(x_i = 0|I\right) & \text{else} \; , \end{cases} \tag{4}$$

where $P\left(x_i = 1|I\right)$ can be understood as image evidence for the segment x_i being part of the vasculature. The higher-order potentials are chosen as:

$$\phi_{i,j}\left(x_i, x_j; I, G\right) = \begin{cases} p_{C,ij} & \text{if } \; x_i x_j = 1 \; , \\ p_T & \text{else} \; , \end{cases} \tag{5}$$

and

$$\phi_{i,j,k}\left(x_i, x_j, x_k; I, G\right) = \begin{cases} p_{B,ijk} & \text{if } \; x_i x_j x_k = 1 \; , \\ p_{C,ij} & \text{if } \; x_i x_j - x_i x_j x_k = 1 \; , \\ p_T & \text{else} \; . \end{cases} \tag{6}$$

where $p_{C,ij}$ is the likelihood of e_i continuing in e_j, $p_{B,ijk}$ the likelihood of a bifurcation involving e_i, e_j and e_k, whereas p_T represents the possibility that neither of them occur. Note that both binary and ternary potentials in (3) account for the relationships of multiple segments, hence we can split $P\left(\mathbf{X} = \mathbf{x}|I, G\right)$ into a pure image evidence term – containing only the unary potentials $\phi_i\left(.\right)$ – and a prior term consisting of both $\phi_{ij}\left(.\right)$ and $\phi_{ijk}\left(.\right)$.

2.2.2. Maximum A Posteriori Estimation by Integer Programming.

So far, we translated the problem of finding the most plausible subnetwork in G^{over} into determining the MAP estimate $\mathbf{x}^*$ of (2):

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} P\left(\mathbf{X}=\mathbf{x}|I,G,\Omega\right) = \arg\max_{\mathbf{x}\in\mathcal{X}} P\left(\Omega|\mathbf{X}=\mathbf{x}\right) P\left(\mathbf{X}=\mathbf{x}|I,G\right), \quad (7)$$

with $\mathcal{X}=\{0,1\}^n$. In the following, we are going to derive an ILP to determine $\mathbf{x}^*$. We start by specifying the likelihood $P(\Omega|\mathbf{X}=\mathbf{x})$ to be equal for all feasible $\mathbf{x}$ and 0 else, i.e.

$$P(\Omega|\mathbf{X}=\mathbf{x}) \propto \begin{cases} 1 & \text{if } \mathbf{x}\in\Omega \ , \\ 0 & \text{otherwise} \ . \end{cases} \quad (8)$$

Applying this definition leads to:

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} P\left(\Omega|\mathbf{X}=\mathbf{x}\right) P\left(\mathbf{X}=\mathbf{x}|I,G\right) \quad (9)$$

$$= \arg\max_{\mathbf{x}\in\Omega} P(\mathbf{X}=\mathbf{x}|I,G) \ . \quad (10)$$

From (10), the definition of the MRF (3) and the fact that each pseudo-boolean function has a unique multilinear polynomial form, it follows that the MAP estimate $\mathbf{x}^*$ takes the form of the integer program:

$$\min_{\mathbf{x}} \quad J(\mathbf{x}) = \sum_{x_i\in X} w_i x_i + \sum_{\substack{x_i,x_j\in X: \\ e_i,e_j \text{ adjacent}}} w_{ij} x_i x_j + \sum_{\substack{x_i,x_j,x_k\in X: \\ e_i,e_j,e_k \text{ adjacent}}} w_{ijk} x_i x_j x_k \ , \quad (11)$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \geq \mathbf{b} \ , \quad (12)$$

$$x_i \in \{0,1\} \qquad \forall x_i \in X \ , \quad (13)$$

with the second and third order weights w_{ij} and w_{ijk} derived as

$$w_{ij} = -\log\frac{p_{C,ij}}{p_T} \ , \quad (14)$$

$$w_{ijk} = -\log\frac{p_{B,ijk}p_T^2}{p_{C,ij}p_{C,ik}p_{C,jk}} \ . \quad (15)$$

The derivation is given in the appendix.

105 *2.2.3. Computing the Weights*

Image Evidence. The weights derived from the unary potentials of the MRF allow to account for image evidence observed for each segment represented by

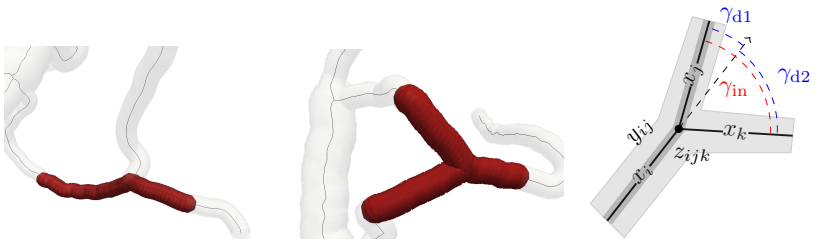


Figure 3: Illustration of network motifs that are considered by the physiological model and the variables at a potential bifurcation. a) Pairs of continuing segments and b) triplets that form a bifurcation (illustrated in red). c) Variables at a potential bifurcation: Segments, e.g. x_i, are drawn as solid black lines, while a pairwise variable y_{ij} that represents x_i continuing in x_j is depicted by the dark grey overlay (y_{ik}, y_{jk} are omitted for clarity). z_{ijk} corresponds to all three adjacent edges in the bifurcation (light grey). Furthermore, both deviation angles γ_{d} (blue) and the inner angle γ_{in} (red) are shown. Radii are estimated perpendicular to the edge direction. In our experiments, we consider geometric features $\gamma_{ijk} = (\gamma_{\mathrm{in}}, \gamma_{\mathrm{d1}}, \gamma_{\mathrm{d2}})_{ijk}$ for bifurcations and $\gamma_{ij} = (\gamma_{\mathrm{d}})_{ij}$ for continuing segments, while radius estimates are used to determine the main trunk.

x_i separately. To infer $P(x_i = 1|I_i, E_i)$ from the image evidence, we average the probabilistic output of the voxelwise classification along the segment e_i. We define $P(x_i = 1|I_i, E_i) = p_i$ and compute:

$$w_i = -\log \frac{p_i}{1 - p_i} \quad . \tag{16}$$

Geometric Prior. As depicted in Fig. 3, we consider two network motifs and weight them accordingly with the derived w_{ij} and w_{ijk}. In order to compute these, we evaluate angles between the involved segments under consideration of their estimated radii (cf. Fig. 3), resulting in:

$$w_{ij} = -\log \frac{p_{C,ij}}{p_T} = -\log \frac{P(\gamma_{ij}|\text{continue}, \boldsymbol{\Theta})\, P(\text{continue}|\boldsymbol{\Theta})}{P(\gamma_{ij}|\text{terminate}, \boldsymbol{\Theta})\, P(\text{terminate}|\boldsymbol{\Theta})} \quad , \tag{17}$$

$$w_{ijk} = -\log \frac{p_{B,ijk} p_T^2}{p_{C,ij} p_{C,ik} p_{C,jk}} \tag{18}$$

$$= -\log \frac{P(\gamma_{ijk}|\text{branch}, \boldsymbol{\Theta})\, P(\text{branch}|\boldsymbol{\Theta})\, P(\text{terminate}|\boldsymbol{\Theta})^2}{\prod_{(i',j') \in 2^{\{i,j,k\}}} P(\gamma_{i'j'}|\text{continue}, \boldsymbol{\Theta})\, P(\text{continue}|\boldsymbol{\Theta})} \quad , \tag{19}$$

where $\boldsymbol{\Theta}$ is the parametric model that encodes physiologically realistic geometric properties of the network motifs in terms of distributions over the considered

geometric features γ_{ij} and γ_{ijk}. In our experiments (Sect. 3.3), we will discuss the choice of such a model $\boldsymbol{\Theta}$ and fit it to evidence from high-resoltion network data.

2.2.4. Global Connectivity

An essential aspect when extracting vascular networks from noisy or incomplete data is to enforce connectivity between the observed components in the network. In our approach, we enforce this property by hard constraints:

$$\sum_{x_i \in M} x_i < |M| + \sum_{x_j \in N} x_j \qquad \forall M \subset X \setminus x_{\text{seed}} \ , \tag{20}$$

where M is a set of connected segments and N its neighbourhood. In other words, a subset M of segments that form one connected structure can only be active, if there is an incoming segment (or it is adjacent to the seed). As there are exponentially many constraints, we follow a lazy constraint generation approach and iteratively add those which are required (cf. Sect. 2.2.6).

2.2.5. Linear Formulation

We note that the integer program in (11) contains second and third order relations between variables. To deal with these, we exploit the binary nature of the variables x_i and introduce additional auxiliary variables $Y = \{y_{ij}\}$ and $Z = \{z_{ijk}\}$ to substitute these products ($x_i x_j$ and $x_i x_j x_k$) in the objective. A set of linear constraints ties the auxiliary variables to the corresponding indicator variables $\{x_i\}$ such that $y_{ij} = x_i x_j$ and $z_{ijk} = x_i x_j x_k$ holds for all feasible solutions:

$$y_{ij} \leq x \qquad \forall x \in \{x_i, x_j\}, \forall y_{ij} \in Y \ , \tag{21}$$

$$y_{ij} \geq x_i + x_j - 1 \qquad \forall y_{ij} \in Y \ , \tag{22}$$

$$z_{ijk} \leq x \qquad \forall x \in \{x_i, x_j, x_k\}, \forall z_{ijk} \in Z \ , \tag{23}$$

$$z_{ijk} \geq x_i + x_j + x_k - 2 \qquad \forall z_{ijk} \in Z \ , \tag{24}$$

which leads to the integer linear program

$$J(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{x_i \in X} w_i x_i + \sum_{y_{ij} \in Y} w_{ij} y_{ij} + \sum_{z_{ijk} \in Z} w_{ijk} z_{ijk} \ , \tag{25}$$

$$\text{s.t.} \quad \mathbf{A}'(\mathbf{x},\mathbf{y},\mathbf{z}) \geq \mathbf{b}' \qquad , \tag{26}$$

$$x_i \in \{0,1\} \qquad \forall x_i \in X \ , \tag{27}$$

$$y_{ij} \in \{0,1\} \qquad \forall y_{ij} \in Y \ , \tag{28}$$

$$z_{ijk} \in \{0,1\} \qquad \forall z_{ijk} \in Z \ , \tag{29}$$

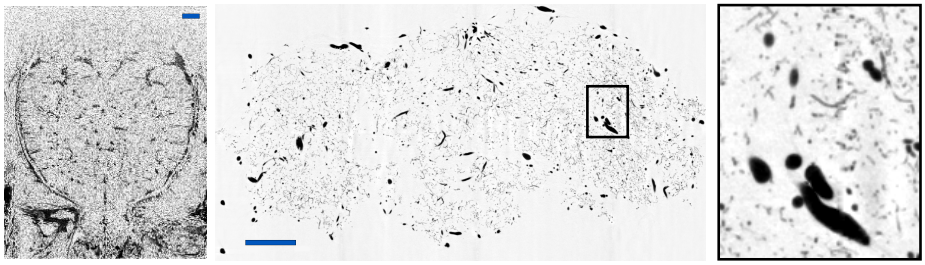
where we summarized the constraints (21) to (24) together with (12) in (26). $(\mathbf{x},\mathbf{y},\mathbf{z})$ is the concatenation of all binary variables to a column vector of all variables. We can solve the linear problem of (25) by a branch and cut algorithm implemented in libraries such as [25].

2.2.6. Solving the Integer Programming Problem for Large Datasets.

The described integer programming problem of (25) with the associated constraints might grow quickly in number for larger graphs that result from whole brain scans. To tackle this problem, we propose to employ a lazy constraint generation scheme together with the following approach:

1. Given the variable set X of the ILP, define a graph $A = (V_A, E_A)$ with a vertex $v_i \in V_A$ for every $x_i \in X$.
2. Add an edge e_{ij} to E_A if and only if there exists a constraint that contains both variables x_i and x_j.
3. Determine the connected components in A. Vertices of each connected component represent a sub-problem that can be solved independently – of course with their according constraints.

Whenever a constraint or variable is added, we adjust this structure dynamically. Then only sub-problems – i.e. connected components of A that are affected by the change – need to be solved, speeding up the performance over approaches that solve the complete problem in every run.



(a) (b)

Figure 4: (a) Example slice of a whole-brain µMRA dataset, (b) Central slice of the corrosion cast µCT with a magnified subregion (right). All images are gray-scale inverted, the blue bar represents 1 mm.

3. Experiments

3.1. Image Data

We use four 3-D µMRA images of the murine brain of size 248 px × 248 px × 109 px with an isotropic voxel spacing of 60 µm, and a corrosion cast micro computed tomography (µCT) of the cerebral vessel network of a mouse with a volume of 2048 px × 2048 px × 3714 px voxels, with a spacing of 2.9 µm. Both image data types are depicted in Fig. 4. The µCT is downsampled by a factor of 2 for the subsequent steps.

3.2. Preprocessing: Segmentation and Graph Construction

In our experiments, we use SFT and RF [8] to obtain an initial segmentation. Parameters, such as filter template order and scales as well as random forest parameters, are adjusted in a leave-one-out cross-validation using manually annotated ground truth labels. Probability maps $P(I)$ are binarized for different thresholds θ and transformed into network graphs using DOHT as discussed above. We segment and track vessels in both µMRA and the µCT volume. The non-overconnected results, obtained from single thresholds $\{0.2, 0.5, 0.9\}$, – denoted as *alternative* below – serve as comparison in the experiments.

3.3. Training: Learning the Geometric Prior from the High Resolution Network

We use the geometrical prior to support bifurcations in our over-connected graph that are valid with respect to their diameter and relative angle, and remove those that are not. To this end, we learn the relative frequencies of radii and deviation angles of vessel segments from the high resolution µCT (see Fig. 3c for angles). We find $p(\gamma_{ij}|\text{continue}, \boldsymbol{\Theta})$ to be well represented by an exponential distribution (where γ_{ij} is the deviation angle between two continuing vessels), and $p(\gamma_{ijk}|\text{branch}, \boldsymbol{\Theta})$ to be well represented by a Gaussian (where γ_{ijk} are the three angles of a bifurcation) while $\boldsymbol{\Theta}$ is the joint set of parameters of the two distribution models. Radius estimates are utilized to determine which vessel is the main trunk in a bifurcation and the reference for the angles considered here. An illustration of the considered angles is given in Fig. 3. Furthermore, we estimate the relative frequencies of the discussed network motifs, $P\,(\text{continue}|\boldsymbol{\Theta})$, $P\,(\text{branch}|\boldsymbol{\Theta})$ and $P\,(\text{terminate}|\boldsymbol{\Theta})$. For this, we take into account that paths in the high-resolution data (n-times lower voxel spacing) are more densely sampled due to the lower spacing, thereby leading to more continuing pairs over a certain physical distance. Hence, we consider only every n-th pair for estimation of $P\,(\text{continue}|\boldsymbol{\Theta})$. Otherwise, $P\,(\text{continue}|\boldsymbol{\Theta})$ would be strongly overestimated for the low-resolution data. Parameters $\boldsymbol{\Theta}$ are fitted to the distributions observed in the µCT using the maximum likelihood estimate. The fitted model is then used to determine weights w_{ij} and w_{ijk} as in (17) and (19). Samples of favourable network motifs according to the learned distributions are depicted in red in Fig. 5.

3.4. Application: Extracting Networks from Low Resolution µMRA Data

We generate an over-connected graph for each of the µMRA test sets using the approach described in Sect. 2.1 with multiple thresholds $\theta_i \in \{0.2, 0.5, 0.9\}$. After optimization, we compare the extracted network to networks obtained from individual thresholds $\theta = 0.2, 0.5$ and 0.9 (alternative).

Figure 7a shows a optimized network, while Fig. 7b-c provides close-up views before and after optimization, respectively. A number of spurious sprouts and

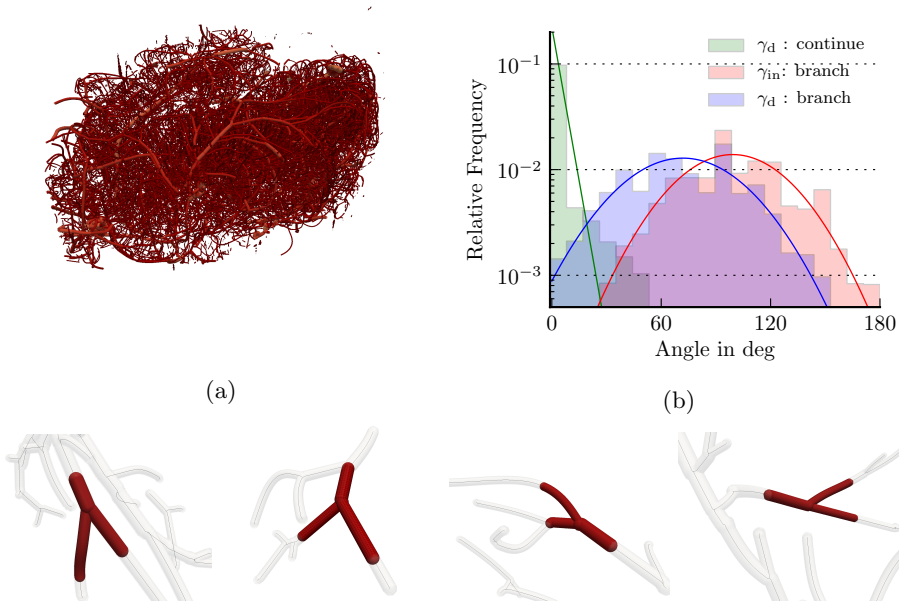


Figure 5: (a) Processed network of the corrosion cast µCT. (b) Angle histograms computed on the extracted network of the corrosion cast dataset with inferred distributions (solid). Shown are deviation angles of continuing segments (green), deviation (blue) and inner (red) bifurcation angles (see Fig. 3c). (c) Favourable configurations of bifurcations (red) according to the fitted model of the corrosion cast µCT.

loops are visible that are removed during the process. Details in Fig. 7d-g show differences between the two thresholded and an optimized network. We find that both thresholded networks (Fig. 7e-f) lose the connection of the large branch (center to top-right), while it is retained in the optimized network (Fig. 7g).

Figure 8 reports quantitative measures of global network properties, such as total vessel volume, average distance to the next vessel (i.e. extravascular distance) and Dice score when rasterized (using a tubular model) and compared with voxel-based segmentation labels. We find the optimized network (opt, $\alpha = 1$) to always group with the more favourable of the alternative, which is a low vessel volume, a moderate, i.e. neither too small nor too large, extravascular distance and a high Dice score, while each of the alternative approaches provides poor results in at least one of these scores. Comparing the result of our optimization with $\alpha = 1$ and a maximal $\alpha_{\max}$, we observe that both relative vessel volume vary only slightly, while the average distance increases when maximum emphasis is put on the image evidence, hence resulting in networks with a decreased vascular density. Note that connectivity is enforced by hard constraints in both cases.

In Fig. 6, the behaviour of our regularization method with varying weight parameter α is shown. For large values of α, the image evidence is emphasized (the network is still connected), while choosing a small value of α puts more emphasis on geometric weights. Although varying α has limited impact on the scores we calculate for our test data, it is recommendable for other problems to determine the most suitable α depending on the previously applied processing. Comparing the processing time of the proposed solver (Sect. 2.2.6) and a standard method (IBM ILOG CPLEX V12.51 [25]) as given in Table 1, we find that our approach of handling subproblems during the optimization leads to an improvement: run times (on a quadcore CPU, 16 GB RAM) were shorter in three out of four cases, whereas the standard method took between 44 % and 107 % longer. We find that absolute run times may vary widely depending on data and parametrization due to the dynamic processing scheme. The results indicate that the approach of partitioning the ILP into independent subproblems

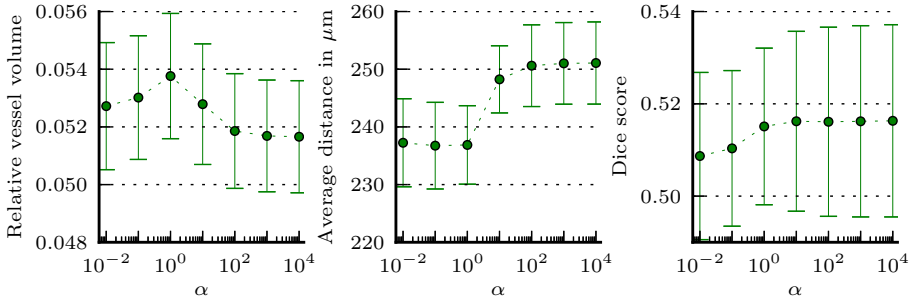


Figure 6: Sensitivity study on the weight parameter α performed on two µMRA datasets. The parameter α has a limited impact with respect to the depicted scores. Averages are within 5.16 % and 5.37 % for relative vessel volume, 236 µm and 251 µm average distance and 50.8 % and 51.6 % Dice (cf. Fig. 8).

Dataset	# 1	# 2	# 3	# 4
t_{ilpcc}	2060.6	186.0	18.6	8.8
t_{cplex}	2958.6	385.0	31.6	8.7
$\frac{t_{\text{cplex}}-t_{\text{ilpcc}}}{t_{\text{ilpcc}}}$	0.44	1.07	0.70	-0.01

Table 1: Runtimes in min on different µMRA datesets for our proposed solving scheme (t_{ilpcc}) and an off-the-shelf solver (t_{cplex}). We observe that our method is in the majority of cases faster as the off-the-shelf solver.

yields larger gains on datasets which required more iterations in the constraint generation, i.e. had longer absolute runtimes.

4. Conclusions

We have introduced a probabilistic approach for extracting vessel networks and have shown that we can compute the MAP estimate by solving an ILP. We presented experiments on µMRA images, where we learned physiological-geometric properties from a high-resolution corrosion cast µCT of a murine cerebrovascular network. Our method can be applied as post-processing step to existing vessel segmentation pipelines in order to incorporate physiological knowledge for improved network extraction. Regarding the construction of the

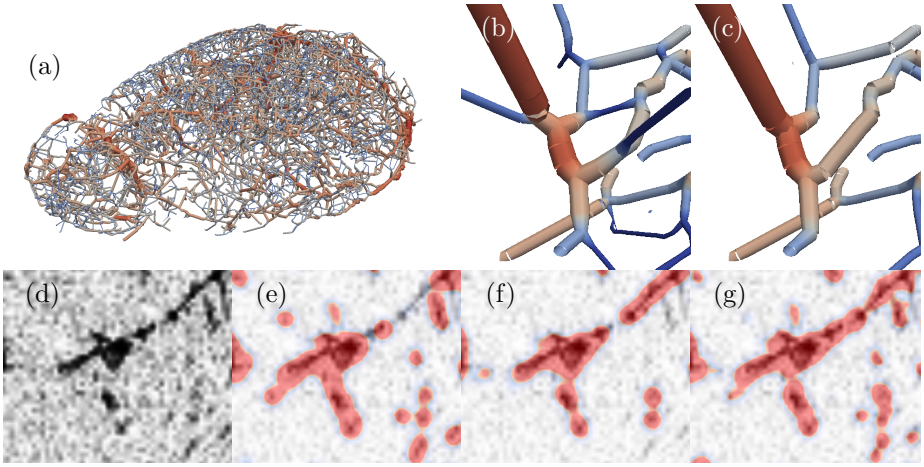


Figure 7: Visualisation of the results. (a) Rendered vascular network extracted with our method. (b) Detail view before and (c) after optimization. Colours change with vessel diameter. (d) Raw image. (e) Rasterized DOHT network $\theta = 0.5$ and (f) with $\theta = 0.9$. (g) Postprocessed network with our method. Note that for the rasterization, a simple tube model is used and therefore, not a perfect voxel-grained segmentation is to be expected but rather a qualitative visualization that indicates whether or not a structure is present in the network model (shown as red overlay).

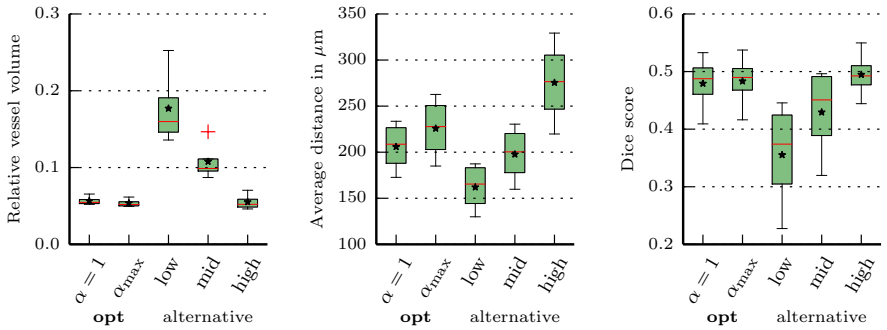


Figure 8: Comparison between the optimized networks (opt) with $\alpha = 1$, $\alpha_{\max} = \alpha \to \infty$ and simple networks obtained at thresholds $\theta = 0.2, 0.5$ and 0.9 (alternative: low= 0.2, mid= 0.5, high= 0.9). Boxplots (median in red, mean as $\star$) depict the statistics on all four µMRA datasets. The Dice score has been computed by rasterizing the network using a simple tube model (negatively affecting absolute Dice values for all approaches to the same degree) and comparing it to voxel-based segmentation labels. Our regularization always groups with the more favourable of the alternatives, i.e. it has a small vessel volume, a moderate average distance and yet an acceptable high Dice score.

overconnected graph, the pipeline could be also be complemented with geometrical or physiological in-fill schemes such as discussed in [15, 16]. The local geometrical properties are easily extendable, for example considering vessel shape, length, curvature and flow direction, or a non-parametric learning as in [14]. Furthermore, higher-order network motifs that occur frequently in vasculature could be incorporated into the prior model in the same way as bifurcations.

Acknowledgements.

This research was supported by the Technische Universität München - Institute for Advanced Study (funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement n 291763, the Marie Curie COFUND program of the the European Union), by grants from the EMDO foundation, Swiss National Science Foundation grant 136822, and the Swiss National Center of Competence in Research on Computer Aided and Image Guided Medical Interventions (NCCR Co-Me) supported by

the Swiss National Science Foundation. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

Appendix

In Sect. 2.2.2, we found that the MAP estimate $\mathbf{x}^*$ is given by

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\in\Omega} P(\mathbf{X}=\mathbf{x}|I,G) \ . \tag{30}$$

Here, we are going to derive the cost function from the definition of the MRF. We start by using the definition of the potential functions in the logarithm of $P(\mathbf{X}=\mathbf{x}|I,G)$:

$$\begin{aligned}
&\log P(\mathbf{X}=\mathbf{x}|I,G) \propto \\
&\quad \sum_{x_i\in X} \log\left(P(x_i=1|I)^{x_i}\, P(x_i=0|I)^{1-x_i}\right) + \sum_{\substack{x_i,x_j\in X:\\ e_i,e_j \text{ adjacent}\\ \nexists e_k \text{ adjacent to } e_i\wedge e_j}} \log\left(p_{C,ij}^{x_i x_j} p_T^{1-x_i x_j}\right) \\
&\quad + \sum_{\substack{x_i,x_j,x_k\in X:\\ e_i,e_j,e_k \text{ adjacent}}} \log\left(p_{B,ijk}^{x_i x_j x_k} p_T^{(1-x_i)(1-x_j)(1-x_k)} \prod_{(i',j')\in S} p_{C,i'j'}^{x_{i'}x_{j'}-x_i x_j x_k}\right) \qquad (31) \\
&= \sum_{x_i\in X} x_i \log\frac{P(x_i=1|I)}{P(x_i=0|I)} + \log P(x_i=0|I) + \sum_{\substack{x_i,x_j\in X:\\ e_i,e_j \text{ adjacent}\\ \nexists e_k \text{ adjacent to } e_i\wedge e_j}} x_i x_j \log\frac{p_{C,ij}}{p_T} + \log p_T \\
&\quad + \sum_{\substack{x_i,x_j,x_k\in X:\\ e_i,e_j,e_k \text{ adjacent}}} \left(x_i x_j x_k \log\frac{p_{B,ijk} p_T^2}{\prod_{(i',j')\in S} p_{C,i'j'}} + \sum_{(i',j')\in S}\left(x_{i'}x_{j'}\log\frac{p_{C,i'j'}}{p_T}\right) + \log p_T \right) \ , \qquad (32)
\end{aligned}$$

where we applied multilinear representations for pseudo-boolean functions and S is the set of pairs out of $\{i,j,k\}$, i.e. $S = 2^{\{i,j,k\}}$. Discarding the constant terms, rearranging the pairwise terms originating from the ternary potentials and merging them with the uniquely pairwise terms – which generates per def-

inition no doubles – then leads to

$$\log P\left(\mathbf{X}=\mathbf{x}|I,G\right) \propto \sum_{x_i\in X} x_i \log\frac{P\left(x_i=1|I\right)}{P\left(x_i=0|I\right)} + \sum_{\substack{x_i,x_j\in X:\\ e_i,e_j \text{ adjacent}}} x_i x_j \log\frac{p_{C,ij}}{p_T} + \sum_{\substack{x_i,x_j,x_k\in X:\\ e_i,e_j,e_k \text{ adjacent}}} x_i x_j x_k \log\frac{p_{B,ijk}p_T^2}{p_{C,ij}p_{C,ik}p_{C,jk}} \ . \tag{33}$$

Finally, taking the negative logarithm turns the maximization into a minimization:

$$\mathbf{x}^* = \arg\min_{\mathbf{x}\in\Omega} -\log P(\mathbf{X}=\mathbf{x}|I,G) \tag{34}$$

$$= \arg\min_{\mathbf{x}\in\Omega} \sum_{x_i\in X} -\log\frac{P\left(x_i=1|I\right)}{P\left(x_i=0|I\right)} x_i + \sum_{\substack{x_i,x_j\in X:\\ e_i,e_j \text{ adjacent}}} -\log\frac{p_{C,ij}}{p_T} x_i x_j \tag{35}$$

$$+ \sum_{\substack{x_i,x_j,x_k\in X:\\ e_i,e_j,e_k \text{ adjacent}}} -\log\frac{p_{B,ijk}p_T^2}{p_{C,ij}p_{C,ik}p_{C,jk}} x_i x_j x_k \ , \tag{36}$$

which is equivalent to the integer program:

$$\min_{\mathbf{x}} \quad J(\mathbf{x}) = \sum_{x_i\in X} w_i x_i + \sum_{\substack{x_i,x_j\in X:\\ e_i,e_j \text{ adjacent}}} w_{ij} x_i x_j + \sum_{\substack{x_i,x_j,x_k\in X:\\ e_i,e_j,e_k \text{ adjacent}}} w_{ijk} x_i x_j x_k \ , \tag{37}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \geq \mathbf{b} \ , \tag{38}$$

$$x_i \in \{0,1\} \qquad \forall x_i \in X \ , \tag{39}$$

with the weights given as:

$$w_i = -\log\frac{P\left(x_i=1|I\right)}{P\left(x_i=0|I\right)} \ , \tag{40}$$

$$w_{ij} = -\log\frac{p_{C,ij}}{p_T} \ , \tag{41}$$

$$w_{ijk} = -\log\frac{p_{B,ijk}p_T^2}{p_{C,ij}p_{C,ik}p_{C,jk}} \ . \tag{42}$$

References

[1] J. Klohs, C. Baltes, F. Princz-Kranz, D. Ratering, R. M. Nitsch, I. Knuesel, M. Rudin, Contrast-enhanced magnetic resonance microangiography

reveals remodeling of the cerebral microvasculature in transgenic ArcAβ mice, The Journal of Neuroscience 32 (5) (2012) 1705–1713.

[2] J. M. Hunter, J. Kwan, M. Malek-Ahmadi, C. L. Maarouf, T. A. Kokjohn, C. Belden, M. N. Sabbagh, T. G. Beach, A. E. Roher, Morphological and pathological evolution of the brain microcirculation in aging and Alzheimer's disease, PLoS ONE 7.

[3] A. Frangi, W. Niessen, K. Vincken, M. Viergever, Multiscale vessel enhancement filtering, in: W. Wells, A. Colchester, S. Delp (Eds.), Medical Image Computing and Computer-Assisted Interventation MICCAI98 SE - 14, Vol. 1496 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 1998, pp. 130–137.

[4] S. R. Aylward, E. Bullitt, Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction., IEEE transactions on medical imaging 21 (2002) 61–75.

[5] C. Kirbas, F. Quek, A review of vessel extraction techniques and algorithms, ACM Computing Surveys (CSUR) 36 (2) (2004) 81–121.

[6] D. Lesage, E. Angelini, I. Bloch, G. Funka-Lea, A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes, Medical image analysis 13 (6) (2009) 819–845.

[7] Y. Jiang, Z. Zhuang, A. Sinusas, X. Papademetris, Vascular tree reconstruction by minimizing a physiological functional cost, Computer Vision and Pattern Recognition Workshops (CVPRW), 2010

[8] M. Schneider, S. Hirsch, B. Weber, G. Székely, B. H. Menze, Joint 3-D vessel segmentation and centerline extraction using oblique hough forests with steerable filters, Medical Image Analysis, in press

[9] C. Pudney, Distance-ordered homotopic thinning: a skeletonization algorithm for 3D digital images, Computer Vision and Image Understanding 72 (3) (1998) 404–413.

[10] T. Lee, R. Kashyap, C. Chu, Building skeleton models via 3-D medial surface axis thinning algorithms, CVGIP: Graphical Models and Image Processing.

[11] L. D. Cohen, R. Kimmel, Global Minimum for Active Contour Models : A Minimal Path Approach, International Journal of Computer Vision 24 (1) (1997) 57–78.

[12] F. Benmansour, L. D. Cohen, Tubular structure segmentation based on minimal path method and anisotropic enhancement, International Journal of Computer Vision 92 (2) (2011) 192–210.

[13] M. A. Gülsün, H. Tek, Robust vessel tree modeling, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2008, Springer, 2008, pp. 602–611.

[14] L. Lu, J. Bi, S. Yu, Z. Peng, A. Krishnan, X. S. Zhou, Hierarchical learning for tubular structure parsing in medical imaging: A study on coronary arteries using 3D CT angiography, in: Computer Vision, 2009 IEEE 12th International Conference on, 2009, pp. 2021–2028.

[15] J. P. Kaufhold, P. S. Tsai, P. Blinder, D. Kleinfeld, Vectorization of optically sectioned brain microvasculature: Learning aids completion of vascular graphs by connecting gaps and deleting open-ended segments, Medical Image Analysis 16 (6) (2012) 1241–1258.

[16] M. Schneider, S. Hirsch, B. Weber, G. Székely, B. H. Menze, TGIF: Topological Gap In-Fill for Correction of Vascular Connectivity – A Generative Physiological Modeling Approach, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014, 2014.

[17] Y. Jiang, Z. W. Zhuang, A. J. Sinusas, L. H. Staib, X. Papademetris, Vessel connectivity using Murray's hypothesis., in: Medical image computing and computer-assisted intervention : MICCAI 2011, Vol. 14, 2011, pp. 528–36.

[18] C. D. Murray, The physiological principle of minimum work: I. The vascular system and the cost of blood volume, Proceedings of the National Academy of Sciences of the United States of America 12 (3) (1926) 207.

[19] E. Turetken, F. Benmansour, B. Andres, H. Pfister, P. Fua, Reconstructing loopy curvilinear structures using integer programming, in: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, IEEE, 2013, pp. 1822–1829.

[20] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network motifs: simple building blocks of complex networks., Science (New York, N.Y.) 298 (5594) (2002) 824–7.

[21] U. Alon, Network motifs: theory and experimental approaches., Nature reviews. Genetics 8 (6) (2007) 450–61.

[22] M. Schneider, S. Hirsch, G. Székely, B. Weber, B. H. Menze, Oblique random forests for 3-D vessel detection using steerable filters and orthogonal subspace filtering, in: B. H. Menze, G. Langs, A. Montillo, Z. Tu, A. Criminisi (Eds.), MICCAI 2012 - Workshop on Medical Computer Vision, Springer Heidelberg / Berlin, 2013, pp. 142–154.

[23] B. Menze, B. Kelm, D. Splitthoff, U. Koethe, F. Hamprecht, On oblique random forests, in: D. Gunopulos, T. Hofmann, D. Malerba, M. Vazirgiannis (Eds.), Machine Learning and Knowledge Discovery in Databases, Vol. 6912 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2011, pp. 453–469.

[24] L. Qu, B. Andres, Estimating Maximally Probable Constrained Relations by Mathematical Programming (2014) 16 http://arxiv.org/abs/1408.0838

[25] IBM ILOG CPLEX Optimizer. http://www.ibm.com/software/integration/optimization/cplex-optimizer/ (2014)