Uncertainty Estimation in Vascular Networks

Markus Rempfler[1,2], Bjoern Andres[3] and Bjoern H. Menze[1,2]

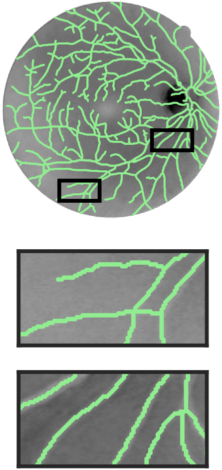
[1]Institute for Advanced Study and [2]Department of Informatics, TU München

[3]Bosch Center for Artificial Intelligence (BCAI), Renningen

14.9.2017

Motivation

- Describe vessel network in terms of a graph.
  - Quantitative analysis of topological properties.
- Why quantify uncertainty in vessel networks?
  - Visualisation for manual inspection and correction.
  - Propagation of uncertainty into analysis of network properties
  - Model parameter inference in a maximum likelihood fashion.



Background

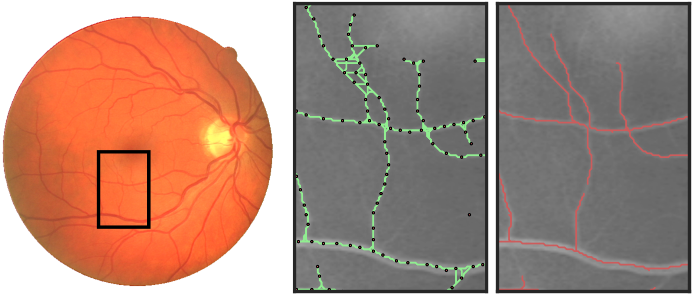
Vessel Network Inference as Subgraph Optimization

Several recent approaches pose the task of inferring a vascular network from image data as a **minimum cost subgraph problem**:

- Türetken et al. TPAMI (2016), Rempfler et al. MedIA (2015), Payer et al. MedIA (2016), Robben et al. MedIA (2016)

Typical workflow:

1. Detect centerlines
2. Construct a hypothesis graph
3. Find the "best" subgraph



Probabilistic Model over Subgraphs

- Subgraphs within $G = (V, E)$ are encoded with binary indicator variables $\mathbf{x} = \{0, 1\}^E$

Feasibility

$$P(\mathbf{x}|\Omega, I, \mathbf{\Theta}) \propto \boxed{P(\Omega|\mathbf{x})} \prod_{ij \in E} P(x_{ij}|I, \mathbf{\Theta}) \prod_{C \in \mathcal{C}(G)} P(x_C|\mathbf{\Theta}),$$

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

- Constraints: *cycle-free*, at most *bifurcations*.

Probabilistic Model over Subgraphs

- Subgraphs within $G = (V, E)$ are encoded with binary indicator variables $\mathbf{x} = \{0, 1\}^E$

Edge terms

$$P(\mathbf{x}|\Omega, I, \mathbf{\Theta}) \propto P(\Omega|\mathbf{x}) \prod_{ij \in E} P(x_{ij}|I, \mathbf{\Theta}) \prod_{C \in \mathcal{C}(G)} P(x_C|\mathbf{\Theta}),$$

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

- Edge probability: discriminatively trained, *local classifier*.

Probabilistic Model over Subgraphs

- Subgraphs within $G = (V, E)$ are encoded with binary indicator variables $\mathbf{x} = \{0, 1\}^E$

Events

$$P(\mathbf{x}|\Omega, I, \boldsymbol{\Theta}) \propto P(\Omega|\mathbf{x}) \prod_{ij \in E} P(x_{ij}|I, \boldsymbol{\Theta}) \boxed{\prod_{C \in \mathcal{C}(G)} P(x_C|\boldsymbol{\Theta})},$$

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

  - Events: penalize *roots*, *terminals* and *bifurcations*.

Probabilistic Model over Subgraphs

- Subgraphs within $G = (V, E)$ are encoded with binary indicator variables $\mathbf{x} = \{0, 1\}^E$

$$P(\mathbf{x}|\Omega, I, \mathbf{\Theta}) \propto P(\Omega|\mathbf{x}) \prod_{ij \in E} P(x_{ij}|I, \mathbf{\Theta}) \prod_{C \in \mathcal{C}(G)} P(x_C|\mathbf{\Theta}),$$

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

- MAP estimator as integer linear programm, optimized with a branch-and-cut algorithm.
- Constraints may be generated on the fly.

Uncertainty Estimation

Sampling

Perturbation Sampler, based on Papandreou and Yuille ICCV (2011)

- Samples are drawn by perturbing the original energy function and optimizing for its perturbed MAP state.
- Original MAP estimator:

$$\begin{aligned} \text{minimize} \quad & \sum_{(i,j)\in E} x_{ij} w_{ij} + \sum_{C\in\mathcal{C}(G)} x_C w_C \\ \text{s.t.} \quad & \mathbf{x} \in \Omega, \;\; [\mathbf{x}, \mathbf{x}_C] \in \Omega_A, \;\; x \in \{0,1\} \end{aligned}$$

Sampling

Perturbation Sampler, based on Papandreou and Yuille ICCV (2011)

- Samples are drawn by perturbing the original energy function and optimizing for its perturbed MAP state.
- (First-order) Perturbed MAP estimator:

$$\begin{aligned} \text{minimize} \quad & \sum_{(i,j)\in E} x_{ij}(w_{ij} + \Delta\gamma_{ij}) + \sum_{C\in\mathcal{C}(G)} x_C w_C \\ \text{s.t.} \quad & \mathbf{x} \in \Omega, \;\; [\mathbf{x}, \mathbf{x}_C] \in \Omega_A, \;\; x \in \{0,1\} \end{aligned}$$

- $\Delta\gamma_{ij}$ is derived as the difference of two Gumbel samples.

Sampling

Gibbs Sampler, Geman and Geman TPAMI (1984)

- Metropolized version of Gibbs sampler (based on Liu (2001)):
  - Randomized scan
  - Acceptance probability: $\alpha = \min\left(1, \frac{1-\pi(x_e|x_{\setminus e})}{1-\pi(x'_e|x_{\setminus e})}\right)$
- Flips that violate $\mathbf{x} \in \Omega$ are never accepted: $\pi(x'_e|x_{\setminus e}) \to 0$.
- Checking feasibility is easier as we know which variable has changed.

Experiments

- Construction of hypothesis graph $G = (V, E)$:
  - Centerline detection with trained regressor following Sironi et al. TPAMI (2015).

Experiments

- Construction of hypothesis graph $G = (V, E)$:
  - Centerline detection with trained regressor following Sironi et al. TPAMI (2015).
  - Nodes $v \in V$: any local maxima after non-maxima suppression.

Experiments

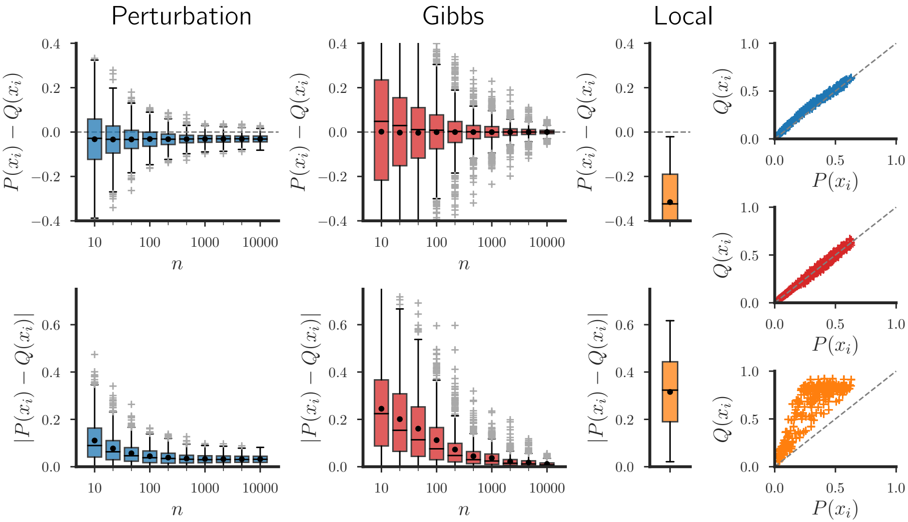
- Construction of hypothesis graph $G = (V, E)$:
  - Centerline detection with trained regressor following Sironi et al. TPAMI (2015).
  - Nodes $v \in V$: any local maxima after non-maxima suppression.
  - Edges $e \in E$: shortest paths between any two nodes $v, w \in V$ s.t. $d(v, w) \leq r_{\text{max}}$.

Experiments

- Construction of hypothesis graph $G = (V, E)$:
  - Centerline detection with trained regressor following Sironi et al. TPAMI (2015).
  - Nodes $v \in V$: any local maxima after non-maxima suppression.
  - Edges $e \in E$: shortest paths between any two nodes $v, w \in V$ s.t. $d(v, w) \leq r_{\max}$.
- Quantitative comparison of approximated marginals:
  - On small subgraphs $\rightarrow$ enables computation of **true marginals** under the given model by **enumeration** of states.

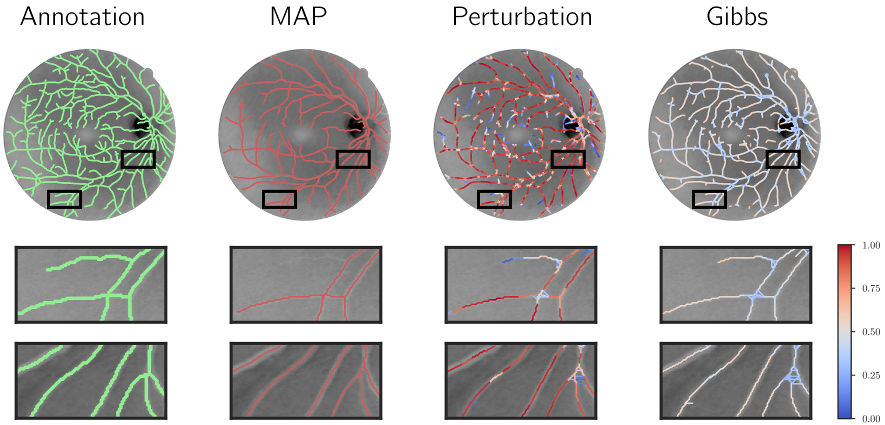
Results

Quantitative Comparison of Approximated Marginals



Results

Qualitative Results



Conclusion & Outlook

Estimating uncertainties under the given model:

- *Gibbs sampler*: unbiased, yet higher variance. Might require adjustments of hyperparameters with varying problem size/density.
- *Perturbation sampler*: biased, but straight forward to apply.

Limitations and **directions** for future work:

- Model choice for most informative marginals.
- Integration of uncertainty estimates in an actual workflow.
- Further work on sampling-based approaches for subgraph problems.