Performance Comparison of Low-level Vessel Detection Algorithms for Segmentation of X-ray Angiograms

Daniel Paredes Soto[1]
el10dap@sheffield.ac.uk
Philip Chan[2]
p.chan@sheffield.ac.uk
Peter Rockett[1]
p.rockett@sheffield.ac.uk

[1] Department of Electronic and Electrical Engineering
University of Sheffield
Sheffield
S1 3JD, UK

[2] Sheffield Vascular Institute
Northern General Hospital
Sheffield
S5 7AU, UK

Abstract

Using examples generated with a realistic model of the imaging physics, we have assessed and compared the performance of three popular multi-scale vessel detectors due to: Frangi *et al*., Sato *et al*. and Lorenz *et al*. We find that there is only a small difference between the performances of the Frangi and Sato detectors for the conditions considered in this work; for small-diameter vessels the response of these two detectors is complex although both suggest some promise for segmenting the vessels in the collateral bed around an occlusion. The Lorenz detector is shown to be the worst performing across a range of conditions.

1 Introduction

Occlusive vascular disease affecting arterial circulations is the major and fastest growing health problem worldwide, and underlies common conditions such as heart attack, stroke and peripheral vascular disease. The World Heath Organisation estimates that these diseases were responsible for 17.3 million deaths worldwide in 2008, forecast to rise to 23.6 million in 2030. Occlusion of major arteries is naturally compensated, to some extent, by the development of minor channels to carry blood around the occlusion, termed *collateral* circulation.

Traditionally, occlusive disease is treated by surgical methods although recently, pharmacological treatments, especially biological products have been researched. Unlike bypass or angioplasty, pharmacological candidates do not address the occlusion directly but are thought to encourage the formation of additional or better collateral vessels either through *angiogenesis* (the development of new blood vessels) and *arteriogenesis* (the development of flow-carrying arterial circulation from capillary precursors). Currently the success of treatments is judged on clinical endpoints like improved exercise ability or fewer amputations; such endpoints are massively confounded by other variables apart from blood flow changes.

Thus a major roadblock to the clinical evaluation of new drug treatments is the lack of an objective anatomical measurement of any enhancement in the blood conveying capacity of a collateral circulation—the development of such a methodology is our longer term goal. Such a direct measure of anatomical change from before to after-treatment would not be subject to confounding factors, and provide credible evidence of effect.

As a precursor to progress in this area, a means of accurately quantifying the properties the collateral network is needed and a fundamental step in this is the extraction of the, often fine, vessel segments which comprise the collateral bed. We have consequently studied the performance of a number of popular, low-level multi-scale vessel extraction algorithms [2, 4, 5] for digital-subtraction X-ray angiography, a common diagnostic procedure for patients presenting with the symptoms of occlusive vascular disease. This preliminary study seeks to compare performance in the ideal case of an isolated, straight vessel of infinite extent—see, however, Section 4 (Future Work). (We envisage that the low-level detector outputs will ultimately be combined to extract the whole vessel network using tracking, or Markov random field approaches; such higher-level methods are not the immediate concern of this paper.)

Here, we principally address two research questions: i) Which low-level vessel detector has the 'best' properties, and ii) How do these detectors perform on small-diameter vessels since these are a major concern for imaging collateral beds. We find that within the limits of this work, there is little to choose between the Frangi and Sato detectors. As far as we are aware, we are the first to present a *quantitative* comparison of these vessel detectors.

2 Methodology

We compare the performance of three much-used multi-scale detectors due to: Frangi *et al.* [2], Sato *et al.* [5] and Lorenz *et al.* [4]. All three algorithms calculate the eigen-decomposition of the Hessian matrix of the image at multiple scales. They then compute some scalar measure of 'vessel-ness' at each scale and select as the final vessel-ness measure, the maximum response across all scales. Essentially, the detectors only differ in how the hand-crafted scalar measures are calculated from the eigenvalues of the Hessian, λ_1 and λ_2. The Frangi detector [2] computes as a measure of eigenvalue eccentricity $\mathscr{R}_B = \lambda_1/\lambda_2$, where $|\lambda_1| \leq |\lambda_2|$. For some scale σ, the measure of 'vessel-ness', L_F is given by:

$$L_F(\sigma) = \begin{cases} 0 & \lambda_2 > 0 \\ \exp(-\mathscr{R}_B{}^2/2\beta^2)[1 - exp(-\mathscr{S}^2/2c^2)] & \text{otherwise} \end{cases} \quad (1)$$

where $\mathscr{S} = \sqrt{\lambda_1^2 + \lambda_2^2}$. We have used Frangi's suggestions of taking $\beta = 0.5$ and adjusting c for best results.

The vessel-ness measure due to Sato [5] is defined, for $|\lambda_1| < |\lambda_2|$, by:

$$L_S(\sigma) = \begin{cases} -\lambda_2 \exp(-\frac{\lambda_1^2}{2(\alpha_1\lambda_c)^2}) & (\lambda_2 < 0) \wedge (\lambda_1 < 0) \\ -\lambda_2 \exp(-\frac{\lambda_1^2}{2(\alpha_2\lambda_c)^2}) & (\lambda_2 < 0) \wedge (\lambda_1 > 0) \end{cases}$$

where $\alpha_1 < \alpha_2$. We have used Sato's suggested values of $\alpha_1 = 0.5$ and $\alpha_2 = 2$.

For the (parameter-less) detector of Lorenz [4] and $|\lambda_1| \leq |\lambda_2|$, $L_L(\sigma) = |\lambda_1|/|\lambda_2|$.

We compute the Hessian matrix over the ten linearly-spaced scales of $\sigma = 1$ to 10 pixel units and take the overall vessel-ness measure as the maximum detector response across all scales. We then threshold this measure to decide a label (vessel of non-vessel).

Obtaining ground-truth data in medical image processing is an enduring problem due to the uncertainties of inter- and intra-expert hand labelling. Here we adopt the well-established procedure of using synthetic data based on a physically-realistic model of the image formation process. Starting with out assumption of an isolated, straight vessel of infinite extent and circular cross-section irradiated with a uniform X-ray beam normal to the longitudinal vessel direction. We assume the vessel is uniformly filled with contrast agent. As the rays pass through the vessel, the intensity of the X-ray beam will be reduced according to the Beer-Lambert Law, $I = I_0 \exp(-\alpha\ell)$ where I_0 is the incident beam intensity, α is the absorption coefficient, and ℓ is the path length through the (contrast agent in the) vessel. See Figure 1. Cleary $\ell = \ell(x)$, where x is the spatial dimension in Figure 1, which also shows the resulting intensity profile projected onto the camera's focal plane. We selected α to give maximum contrast at the largest vessel diameter considered while allowing $\pm 3\sigma$ 'headroom' for the largest noise perturbations of the pixel values in the output image. The process thus mirrors a radiographer optimising the X-ray imaging conditions.

A distinguishing feature of the present work is the realistic image formation model: previous work, for example Krissian *et al.* [3], has usually assumed the projected vessel intensity profile is Gaussian with a physically-implausible infinite region of support. Drechsler and Laura [1] used the data of Krissian *et al.* in a qualitative comparison of vessel detectors.

We have no *a priori* reason to suppose that a vessel will present at any particular orientation to, or displacement from the image lattice so we select a uniformly-random orientation $\theta \in [0 \ldots 45^o]$ and a displacement $\Delta x \in [-0.5 \text{ pixel} \ldots +0.5 \text{ pixel}]$ from the notional origin in Figure 1. (In practice, these are applied by an affine transformation of the projected intensity image above.) To account for the limitations of the imaging optics, we convolve the affine-transformed image with a Gaussian of $\sigma_{PSF} = 1$ and then add Gaussian-distributed noise of some σ_N which assumes that the dominant noise process is due to thermal noise in the camera/electronics. Finally, we quantise the pixel intensities into the range $[0 \ldots 255]$ to mimic analogue-to-digital conversion in the camera. This final image containing a single vessel is passed to one of the above detectors. In assessing a detector, we consider only the label of the central pixel

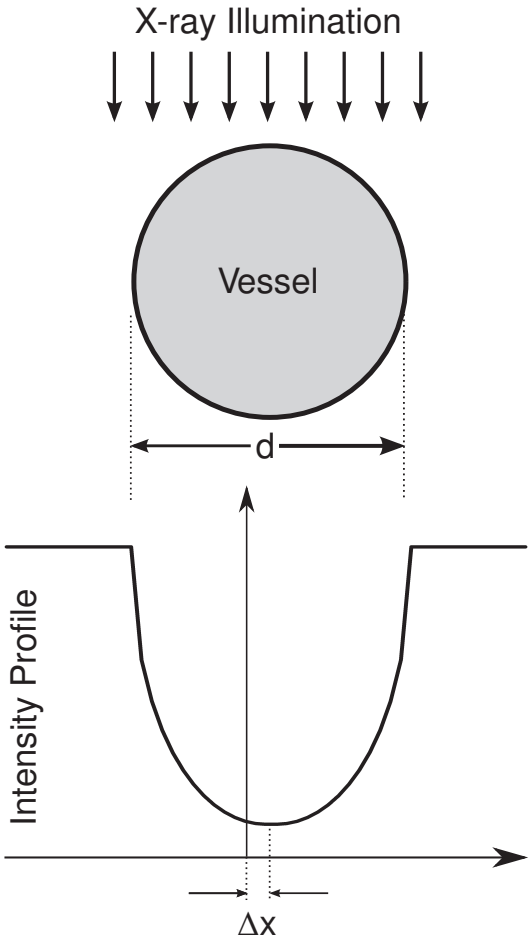

Figure 1: Formation of the vessel intensity model. The lower figure shows the projected intensity profile.

of the image through which the centreline of the vessel passes. Our rationale here is that the signal-to-noise ratio is maximised at the centre of the vessel—if the vessel cannot be detected at its centre, it cannot probably be detected anywhere else across its diameter. We thus implicitly investigate the upper bound of the detectors' performance. Additionally, each image represents an independent Monte Carlo trial.

To form counter-example images (*i.e.* genuine absence of a vessel) we form an image of uniform background intensity but with each pixel corrupted by noise of the same variance as the positive vessel examples. The method of generating positive and negative examples thus closely mirrors the process of digital-subtraction X-ray angiography.

We have repeated the above steps of generating positive and negative vessel examples in a Monte Carlo experiment (1000 trials) and and labelled with each of the detectors using various thresholds to construct receiver operating characteristic (ROC) plots. We also use the area-under-the-curve (AUC) as a summary statistic for the ROC plot.

3 Results

We have considered a range of vessel diameters ($D \in [1,3,5,7\ldots,15]$ pixel units) and noise powers ($\sigma_N \in [5,10,15,20,25]$). For brevity we only show results for the four permutations of small/large vessels, and lowest/highest noise powers since these represent the extremes of detector performance between which there is a more-or-less smooth variation. For the (physically unrealistic) absence of noise ($\sigma_N = 0$) all three detectors perform 'perfectly' (AUC = 1.0) apart from the case of unit vessel width ($D = 1$) where the Frangi and Lorenz detectors are tied (AUC = 0.95) and perform slightly worse than Sato (AUC = 0.97).

Fig. 2(a) and 2(b) show the ROC plots for $D = 1$ and the extremes of (a) low-noise ($\sigma_N = 5$) and (b) high-noise ($\sigma_N = 25$). Interestingly, in the low-noise regime all three detectors perform to some degree although the Sato and Frangi detectors are clearly better than Lorenz. (The straight line segments of the ROC plots are due to the discontinuities introduced by the case of $\lambda_2 = 0$ in Eqn. 1, and so on.) The detectors' behaviour is quite complex with the ROC plots crossing although there is generally little to choose between Frangi and Sato. Note that for the case of high noise, the performance of all three methods reduces to little better than random guessing—the 45^o line—as might be expected since the signal-to-noise (SNR) ratio becomes very low for this case. Since we have use a fixed value of attenuation coefficient (α), the SNR reduces as a function of reducing vessel diameter, hence small vessels are intrinsically harder to label.

For $D = 1$ and 3, and $\sigma_N = 25$ (high noise), there is little to choose between the performances of all three detectors. There is , however, an interesting transition between $D = 3$ and $D = 5$ (see Fig. 2(c) and (d)) where the Frangi and Sato algorithms begin to out-perform Lorenz, with Frangi slightly ahead of Sato, a trend which continues up to large diameters.

Fig. 2(e) and (f) show the corresponding plots for $D = 15$. At this stage, the Sato and Frangi algorithms both perform 'perfectly' (AUC = 1) regardless of noise power. The Lorenz detector, however, is clearly inferior, even for low noise powers.

As a summary, Fig. 3 shows the AUC statistic *vs.* vessel diameter D. The region between these two pairs of curves (low noise and high noise) delineates the useful operating envelope of the two detectors.

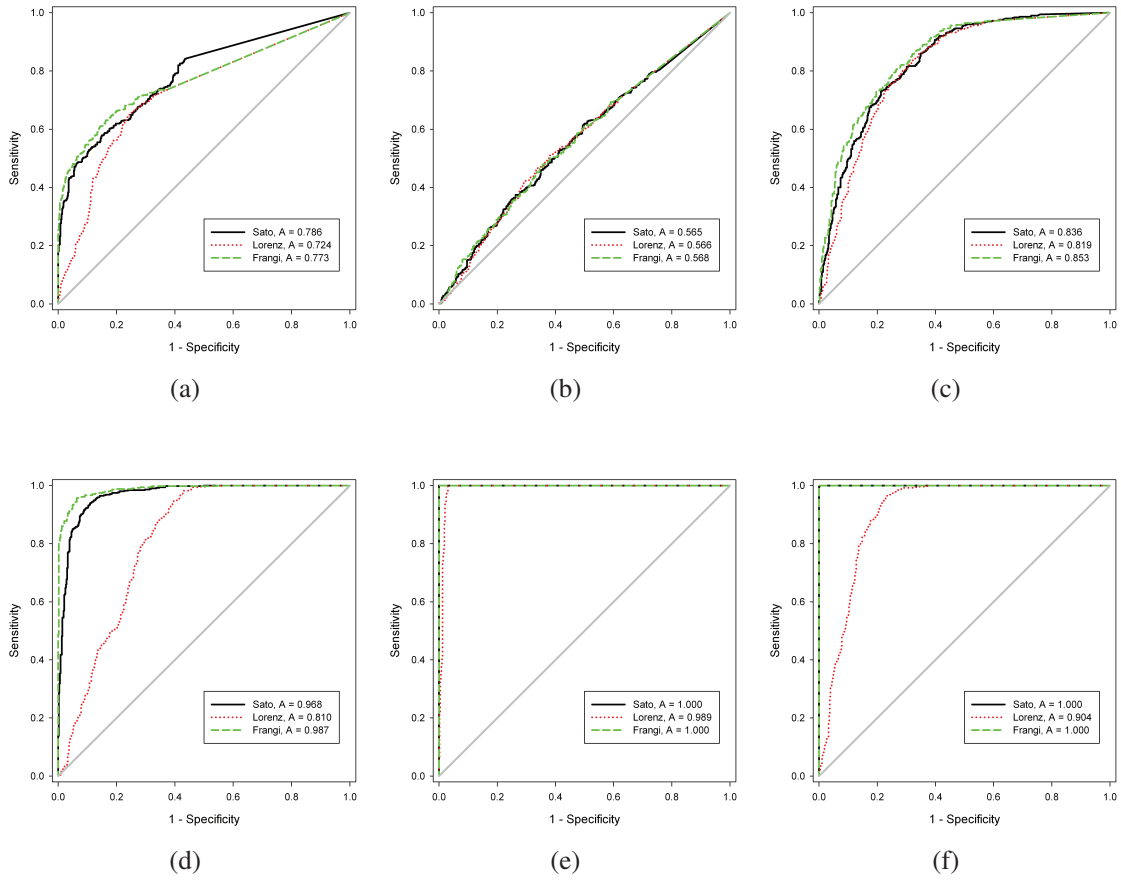

Figure 2: ROC plots for: $D = 1$ and (a) low-noise ($\sigma_N = 5$) and (b) high-noise ($\sigma_N = 25$). (c) $D = 3$ and (d) $D = 5$; high-noise ($\sigma_N = 25$). $D = 15$ and low-noise (e) ($\sigma_N = 5$) and (f) high-noise ($\sigma_N = 25$).

4 Future Work

The present preliminary report considers only the idealised case of a straight, isolated vessel. In future work, we will consider the effects on detector response of curved vessels and bifurcations. In addition, since the detectors are all multi-scale approaches, we anticipate that the responses from two vessels which approach closer than the scale of a Gaussian filter will interact with each other. For extracting vessel *networks* this is an important factor. Moreover, the vessels in a collateral bed frequently present as a 'corkscrew' and it is possible that the response (even) from an isolated corkscrew collateral could be affected by the proximity of other sections of the same vessel. These and other factors will be the subject of future studies.

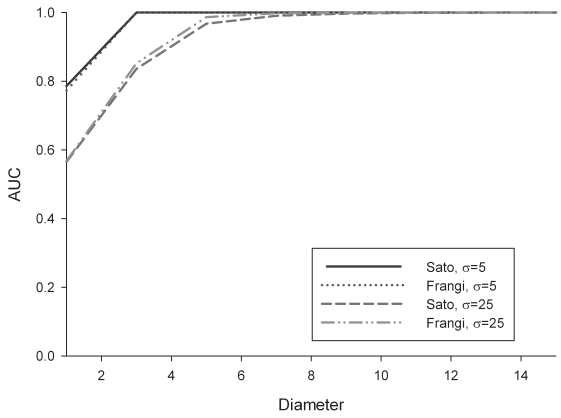

Figure 3: AUC *vs.* vessel diameter D for the Frangi and Sato detectors only.

We are assuming all vessels lie in a plane and can be imaged normal to the longitudinal

vessel direction. In reality, it is highly likely that vessels will have some component passing *into* a 2D image plane so extracting an entire network will be problematic. In this situation, 3D magnetic resonance angiography (MRA) will be highly desirable. We believe, given a suitable MRI simulator, the methodology in this paper can be directly applied to MRA images—this too is an area of future work.

5 Conclusions

From constructing a physics-based model of the process by which the image of a vessel is formed in digital-subtraction X-ray angiography, we have compared the performances of the vessel detectors devised by: Frangi *et al*., Sato *et al*. and Lorenz *et al*. On the basis of the present results, the Lorenz detector is clearly inferior to the Frangi and Sato detectors between which there is little to choose for anything other than the smallest vessel diameters. Whether Frangi or Sato is better for vessels on the 1 pixel scale would seem to depend on exact operating point. Nonetheless, these methods seem to offer some promise for segmenting fine vessels if combined with a contextual approach (*e.g*. Markov random fields).

The present results are limited to straight, isolated vessels. Whether any detector is superior for real, rather than idealised, vessel networks remains to be established in future work. For example, *qualitative* observations suggest the Frangi detector does not perform well at vessel bifurcations [1].

References

[1] K. Drechsler and C. O. Laura. Comparison of vesselness functions for multiscale analysis of the liver vasculature. In 10^{th} *IEEE International Conference on Information Technology and Applications in Biomedicine*, pages 1–5, Corfu, Greece, 2010.

[2] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. In W M Wells, A Colchester, and S L Delp, editors, *Medical Image Computing and Computer-Assisted Intervention*, pages 130–137, Cambridge, MA, 1998.

[3] K. Krissian, G. Malandain, N. Ayache, R. Vaillant, and Y. Trousset. Model-based detection of tubular structures in 3D images. *Computer Vision and Image Understanding*, 80 (2):130–171, November 2000.

[4] C. Lorenz, I.-C. Carlsen, T. M. Buzug, C. Fassnacht, and J. Weese. Multi-scale line segmentation with automatic estimation of width, contrast and tangential direction in 2D and 3D medical images. In 1^{st} *Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medial Robotics and Computer-Assisted Surgery*, pages 233–242, Grenoble, France, 1997.

[5] Y. Sato, S. Nakajima, H. Atsumi, T. Koller, G. Gerig, S. Yoshida, and R. Kikinis. 3D multi-scale line filter for segmentation and visualization of curvilinear structures in medical images. In 1^{st} *Joint Conference on Computer Vision, Virtual Reality and Robotics in Medicine and Medial Robotics and Computer-Assisted Surgery*, pages 213–222, Grenoble, France, 1997.