

A spatial structure function with metric embedding kernel for retinal OCT

author order TBD¹, Loan Huynh¹, Ronald Zambrano^{2,3}, Layton Aho¹, Gadi Wollstein², Joel S. Schuman^{2,3}, and Andrew R. Cohen¹

¹Electrical and Computer Engineering, Drexel University, USA

²Wills Eye Hospital, Philadelphia, USA

³Biomedical Engineering, Drexel University, USA

I. ???

possible submission : [IEEE Data Compression Conference?](#)

???visualization aspect ratio??? any strong feelings on stretching images for visualization?

II. INTRODUCTION

A. *Why metric embedding?*

tl;dr RKHS make subsequent optimization learning easier to implement and more likely to generalize better [primer].

B. *Optical Coherence Tomography of the Retina*

The last few decades have ushered in revolutionary advancements in medical imaging which have helped elucidate relationships between biological structures and physiological function in health and disease. Combining computer-aided diagnosis with imaging data further augments physicians' ability to make higher precision clinical decisions. Perhaps no better example of this could be made than with the impact that optical coherence tomography (OCT) has had on eye care. OCT is a non-invasive diagnostic imaging tool which employs principles of optical interferometry using a low-coherence light source to obtain cross-sectional digitally reconstructed images of biological tissue [1]. Often described as an optical analog to ultrasound, OCT images

are generated from detection of electrical field produced from the echoed tissue and reference signals [2]. OCT datasets are comprised of multiple A-scans acquired in rapid succession to form cross-sectional images called B-Scans. These B-scans can be rapidly acquired in a raster pattern to obtain a volumetric image of the subject tissue. Current generations of spectral domain (SD) OCT have an axial resolution of 5-8 microns with a lateral resolution of 6-20 microns and a scanning rate of 100,000 A-scans per second, making them ideal for Capture of 3-D volumetric data in vivo. [3].

Although predominantly used in clinical eye practice, OCT spans beyond ophthalmology. OCT has clinical applications in cardiology, otology, dermatology, and dentistry [4]. OCT's ability to penetrate up to 3mm of tissue and produce high resolution images have provided use cases in diagnosing myocardial infarction with non-obtrusive coronary arteries [5], facilitating cochlear implants surgery [6], diagnosing basal cell carcinoma [7], and even dental cavity detection [8]. More recently, the sub-field of ophthalmic imaging known as Oculomics has shown the capability of using OCT to make prognostications of non-ocular systemic disorders like cardiovascular [9] and neurological disease [10, 11], as well as improved estimation of phenotypic age for predicting mortality [12].

Mainly, this ability to safely and quickly acquire ocular images in a longitudinal and reproducible manner has revolutionized the way clinicians diagnose and manage blindness causing diseases such as glaucoma. Glaucoma is the global leader of irreversible blindness, with some projections indicating that the number of affected individuals could rise to 112 million by 2040 due to population aging [13]. Glaucoma, characterized by persistent microstructural deterioration of the optic nerve and retina, is often asymptomatic until the moderate to severe stages of disease. Therefore, OCT plays a critical role in mitigating disease through early detection of these microscopic structural changes. However, OCT alone is insufficient in determining disease severity. Glaucoma standard of care involves ophthalmic imaging with OCT in conjunction with functional assessment of the patient's visual field (VF) through standardized automated perimetry [14]. One of the prominent ways to categorize disease severity and visual field defect is through visual field mean deviation (VFMD), which is a numerical estimation of light in decibels (dB) that an individual eye can perceive [15]. It is important to note, however, that structural changes in the retina are not linearly correlated with VFMD throughout the disease spectrum. Cross-sectional studies have demonstrated that there is little to no correlation between mean RNFL and VFMD until significant RNFL damage has occurred, at which point

statistically significant associations with VFMD can be observed [16]. Therefore, improving early detection of structural changes through ocular imaging are mission critical in the pursuit of vision preservation.

Artificial intelligence (AI) applications using deep learners (DL) to detect, diagnose and predict disease progression have a potentially enormous impact on public health [17]. The most common approaches look to classify glaucoma based on clinically accepted structural features such as the retinal nerve fiber layer (RNFL), ganglion cell inner plexiform layer (GCIPL) and optic nerve head (ONH) [18–22]. Furthermore, incorporating features such as ganglion cell complex (GCC) thickness, ONH macrostructure, and RNFL reflectance maps into DL model training have been shown to significantly improve diagnostic accuracy [23]. Although there is no consensus on the exact clinical classification of glaucoma, the structure function relationship between OCT and VF are critical in understanding disease severity and progression. This reliability on VF for diagnosis is challenging because the VF test itself is challenging for patients. The test, which can last between 4-8 minutes per eye, requires patients to sit still and fixate on singular point without moving and click a button anytime a point of light is perceived in their field of vision. Test difficulty, combined with an older patient demographic, often makes VF test results more variable than OCT, which takes mere seconds to perform. Therefore, considerable efforts have been made towards predicting VF outcomes based on OCT. Previous studies have shown the ability to derive the spatial relationship between structure and function using OCT and VF point-by-point estimation [14] based on stochastic optimization DL models [24].

Unsupervised feature agnostic approaches using OCT have shown promise in predicting disease progression measured by VF. Chen et al. [25], describe a 3D based ResNet18 Convolutional neural network (CNN) capable of inferring pointwise VF sensitivities directly from segmentation-free OCT 3D volumes. Similarly, attention-guided network approaches have demonstrated improved glaucoma detection by taking OCT volumes and computing dual 3-D gradient class activation heatmaps to predict the Visual Field Index (VFI) estimation [26]. DL approaches using OCT volumes from glaucoma subjects have identified 14 non-clinically defined surface shape patterns near the ONH which are capable of predicting specific VFMD loss rate with a best coefficient of determination of $r^2 = 0.37$ [27]. These clinically agnostic DL approaches using OCT volumes have demonstrated the ability to identify previously undiscovered biomarkers for improving the prediction of glaucomatous functional defects, but, perhaps due to the insufficient confidence indices or inability to explicitly and metrically define the relationships,

remain limited to research applications.

Key observation – inherent anisotropy of imaging implements Frangi-like plate filter along the high resolution axis.

C. Kolmogorov complexity and the normalized information distance (NID)

FLIF [28] ...

D. Semi-supervised spectral learning

E. Kolmogorov structure function vs. multi-scale vessel enhancing image filters

$H_{Kolmogorov}$ [Vereschegin and Vitanyi] vs. H_{Frangi}

III. THE RETINAL STRUCTURE FUNCTION

IV. VALIDATING THE NCD VS. PREDICTION AGAINST FUNCTIONAL CHANGES OF THE VISUAL FIELD

Idea : $NCD(OCT_1, OCT_2)$ to predict $|VFMD_1 - VFMD_2|$

open source code in progress (?). open source image data? what can we release? maybe a small sample, e.g. ~ 100 stacks?

Deep learning state of the art in estimating VFMD from retinal OCT achieves RMSE $\lesssim 2.5$. We pose the question differently, looking instead to predict $|\Delta VFMD|$ given the normalized compression distance between image pairs. Training a single-layer single-node regression network with sigmoid activation yields cross-validation RMSE ~ 0.8 . A second single-layer 10-node regression network with input including one of the two VFMD values (from either image date) achieves cross-validation RMSE ~ 0.7 .

V. FUTURE WORK

Anisotropic denoising on the GPU with HIP.

Clustering by compression: e.g. M vs. F, OD vs. OS, zip code?

RKHS combinations: macula + ONH simultaneously, multi-resolution by channel vs. dimensionality reduction, metadata w/ bzip,...?

VI. ACKNOWLEDGEMENTS

VII. SOFTWARE AND DATA AVAILABILITY

All of the software tools used are available free and open source, see <https://git-bioimage.coe.drexel.edu/opensource/rsf>. The image data together with segmentation and tracking results can be viewed interactively at <https://leverjs.net/ssfCluster>. The LEVERSC 4-D WEBGL viewer [29] renders 3-D kymographs and images, and the web API also supports downloading metadata and results directly.

REFERENCES

1. Huang, D. *et al.* Optical coherence tomography. *Science* **254**, 1178–1181 (1991).
2. Fujimoto, J. & Swanson, E. The development, commercialization, and impact of optical coherence tomography. *Investigative ophthalmology & visual science* **57**, OCT1–OCT13 (2016).
3. Alexopoulos, P., Madu, C., Wollstein, G. & Schuman, J. S. The development and clinical application of innovative optical ophthalmic imaging techniques. *Frontiers in medicine* **9**, 891369 (2022).
4. Ali, S. *et al.* Optical coherence tomography’s current clinical medical and dental applications: a review. *F1000Research* **10** (2021).
5. Reynolds, H. R. *et al.* Coronary optical coherence tomography and cardiac magnetic resonance imaging to determine underlying causes of myocardial infarction with nonobstructive coronary arteries in women. *Circulation* **143**, 624–640 (2021).
6. Starovoyt, A., Putzeys, T., Wouters, J. & Verhaert, N. High-resolution imaging of the human cochlea through the round window by means of optical coherence tomography. *Scientific reports* **9**, 14271 (2019).
7. Chen, S. *et al.* Evaluation of ultrahigh-resolution optical coherence tomography for basal cell carcinoma, seborrheic keratosis, and nevus. *Skin Research and Technology* **27**, 479–485 (2021).
8. Hariri, I. *et al.* Estimation of the enamel and dentin mineral content from the refractive index. *Caries research* **47**, 18–26 (2013).
9. Chan, Y. K., Cheng, C.-Y. & Sabanayagam, C. Eyes as the windows into cardiovascular disease in the era of big data. *Taiwan Journal of Ophthalmology* **13**, 151–167 (2023).

10. Suh, A. *et al.* Retina oculomics in neurodegenerative disease. *Annals of Biomedical Engineering* **51**, 2708–2721 (2023).
11. Lin, T.-Y. *et al.* Individual prognostication of disease activity and disability worsening in multiple sclerosis with retinal layer thickness z scores. *Neurology: Neuroimmunology & Neuroinflammation* **11**, e200269 (2024).
12. Nusinovici, S. *et al.* Retinal photograph-based deep learning predicts biological age, and stratifies morbidity and mortality risk. *Age and ageing* **51**, afac065 (2022).
13. Tham, Y.-C. *et al.* Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology* **121**, 2081–2090 (2014).
14. Chen, Z., Ishikawa, H., Wang, Y., Wollstein, G. & Schuman, J. S. Deep-Learning-Based Group Pointwise Spatial Mapping of Structure to Function in Glaucoma. *Ophthalmology Science* **4**, 100523 (2024).
15. Thirunavukarasu, A. J. *et al.* A validated web-application (GFDC) for automatic classification of glaucomatous visual field defects using Hodapp-Parrish-Anderson criteria. *npj Digital Medicine* **7**, 131 (2024).
16. Wollstein, G. *et al.* Retinal nerve fibre layer and visual function loss in glaucoma: the tipping point. *British Journal of Ophthalmology* **96**, 47–52 (2012).
17. Schuman, J. S. *et al.* A Case for the Use of Artificial Intelligence in Glaucoma Assessment. *Ophthalmology Glaucoma* **5**, e3–e13. ISSN: 2589-4196. <https://www.sciencedirect.com/science/article/pii/S2589419621002805> (2022).
18. Prahs, P. *et al.* OCT-based deep learning algorithm for the evaluation of treatment indication with anti-vascular endothelial growth factor medications. *Graefe's Archive for Clinical and Experimental Ophthalmology* **256**, 91–98 (2018).
19. Grewal, D., Jain, R., Grewal, S. & Rihani, V. Artificial neural network-based glaucoma diagnosis using retinal nerve fiber layer analysis. *European journal of ophthalmology* **18**, 915–921 (2008).
20. Shi, M., Tian, Y., Luo, Y., Elze, T. & Wang, M. RNFLT2Vec: Artifact-corrected representation learning for retinal nerve fiber layer thickness maps. *Medical Image Analysis* **94**, 103110 (2024).

21. Berenguer-Vidal, R. *et al.* Automatic segmentation of the retinal nerve fiber layer by means of mathematical morphology and deformable models in 2d optical coherence tomography imaging. *Sensors* **21**, 8027 (2021).
22. Christopher, M. *et al.* Retinal nerve fiber layer features identified by unsupervised machine learning on optical coherence tomography scans predict glaucoma progression. *Investigative ophthalmology & visual science* **59**, 2748–2756 (2018).
23. Tan, O. *et al.* A Hybrid Deep Learning Classification of Perimetric Glaucoma Using Peripapillary Nerve Fiber Layer Reflectance and Other OCT Parameters from Three Anatomy Regions. *ArXiv* (2024).
24. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
25. Chen, Z. *et al.* Segmentation-free OCT-volume-based deep learning model improves pointwise visual field sensitivity estimation. *Translational Vision Science & Technology* **12**, 28–28 (2023).
26. George, Y. *et al.* Attention-guided 3D-CNN framework for glaucoma detection and structural-functional association using volumetric images. *IEEE journal of biomedical and health informatics* **24**, 3421–3430 (2020).
27. Saini, C. *et al.* Assessing surface shapes of the optic nerve head and peripapillary retinal nerve fiber layer in glaucoma with artificial intelligence. *Ophthalmology Science* **2**, 100161 (2022).
28. Wuille, J. S. & Pieter. FLIF: Free lossless image format based on MANIAC compression. *IEEE International Conference on Image Processing (ICIP)* (2016).
29. Winter, Mark & Cohen, A. R. LEVERSC: Cross-Platform Scriptable Multichannel 3-D Visualization for Fluorescence Microscopy Images. *Frontiers in Bioinformatics* **2**. ISSN: 2673-7647. <https://www.frontiersin.org/article/10.3389/fbinf.2022.740078> (2022).

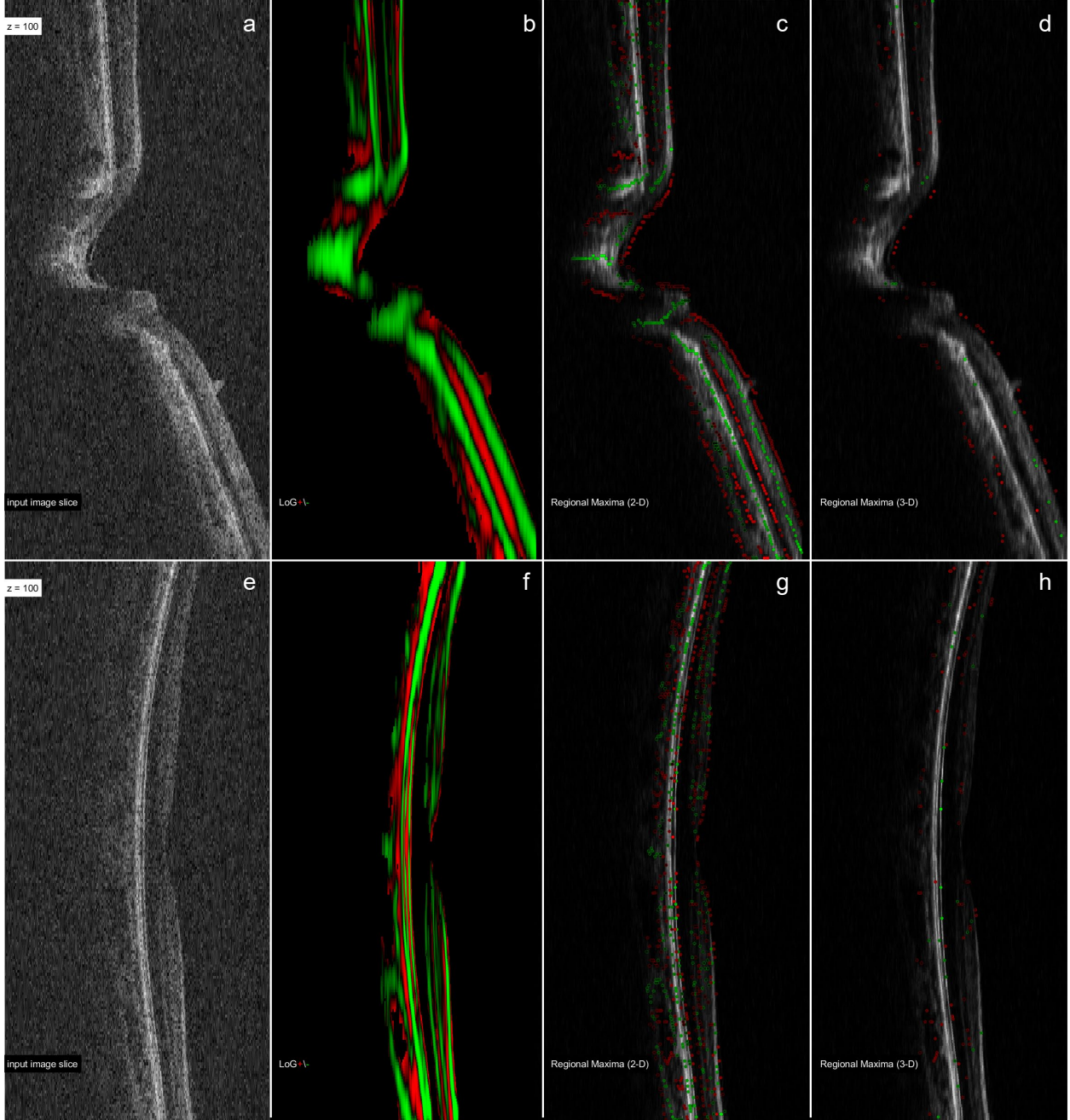


Fig. 1: The retinal structure function (RSF) computes at each image voxel how ‘plate-like’ is the 3-D intensity configuration of the surrounding voxels. The input image (a, single slice shown) is denoised using non-local means (c,d), and then processed with a Laplacian of Gaussian (LoG) convolution kernel (b, single slice shown). The LoG positive response (red) represents dark structure against a bright background, the negative response (green) represents bright structure against a dark background. The 2-D regional maxima weighted by the LoG, with the interior of each dot scaled by filter response, are shown for illustration (e). The 3-D regional maxima weighted by the LoG (f), are the representation used as input to the normalized compression distance (NCD). Top row (a-d) shows optic nerve OCT using LoG of size $[0.5, 10, 10]$ voxels, the bottom row (e-h, same captions as a-d) is macula OCT with LoG of size $[0.35, 7, 7]$ voxels. The goal of the RSF is to extract a local measure of the voxel structure so the NCD best captures patterns of similarity between image pairs. See also [ONH movie](#) and [macula movie](#). TODO brighten macula movie...

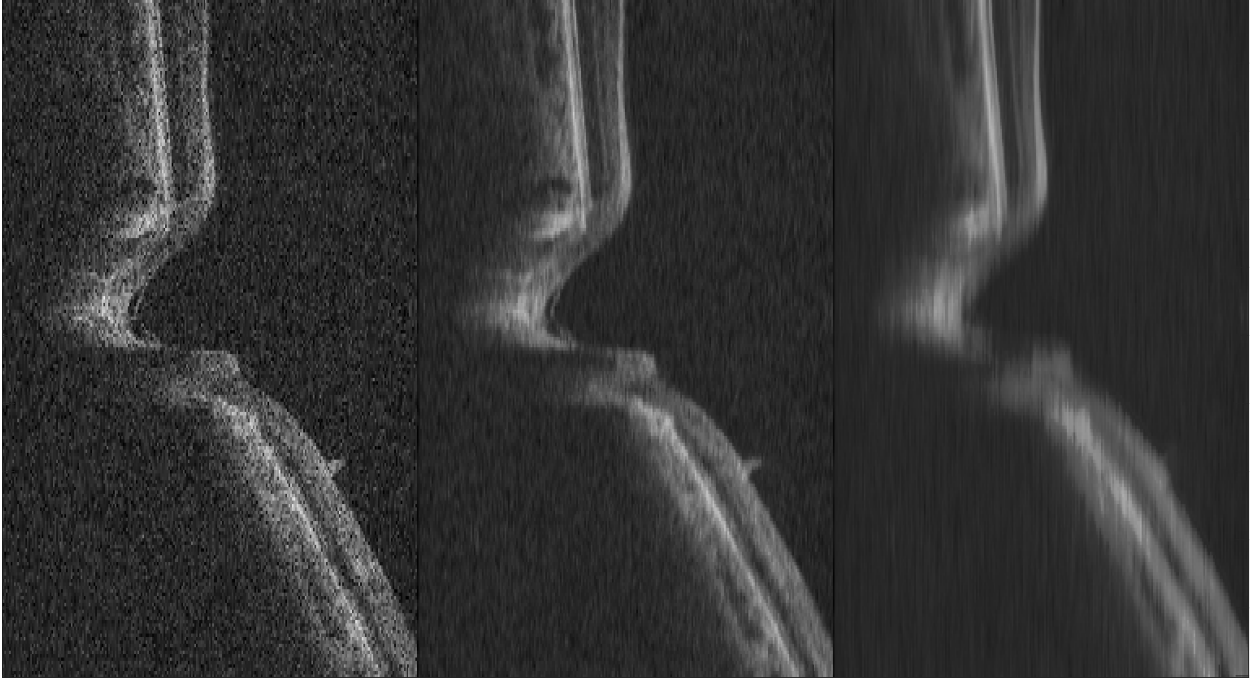


Fig. 2: Inherent resolution anisotropy in retinal OCT enables novel denoising approaches. The input image (a, single slice shown) is denoised using a 3-D non-local means (b) with an isotropic pattern ($[3, 3, 3]$ voxels) and search radius ($[25, 25, 25]$ voxels). With a 'plate-enhancing' anisotropy in both pattern ($[1, 5, 5]$) and search ($[5, 50, 50]$), the orientation of the smoothing direction can be clearly seen along the vertical axis, while along the horizontal or low-res direction (c). The optimal image processing pipeline will be chosen to maximize the retinal structure function, as measured from characteristics of both the embedded and the kernel space. TODO – add arrows indicating axes...Good figure for a grant, not ready for manuscript?

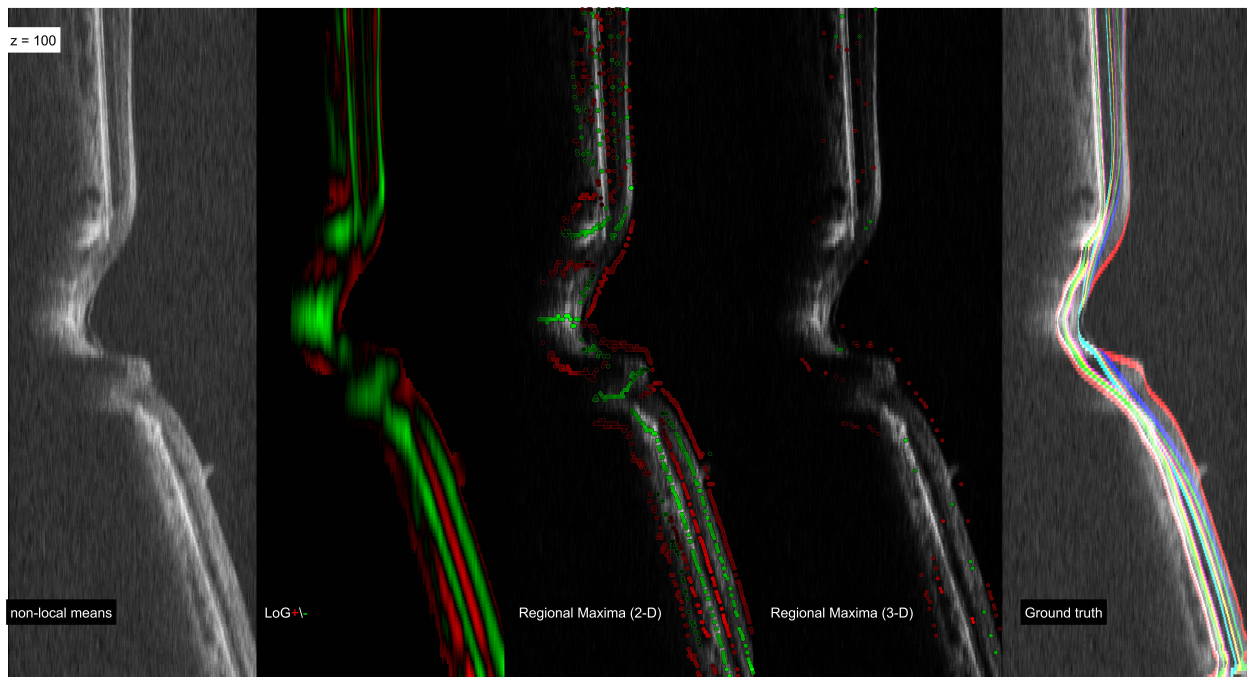


Fig. 3: **Comparison to Iowa segmentation.** The Iowa segmentation is generally gorgeous and does a great job at capturing the 2-D structure. [OCT Explorer](#) fabulous UI! but the algorithms do not appear to be open-source? Need to confirm. Loan – re-render the ground truth panel (let's call this the 'Iowa reference algorithm' [Ronald – refs here]) over the denoised background maybe?