Volume-Fused Super-Resolution OCT Reveals Cellular-Level Structures in the Outer Segment Across Wider Fields Without Adaptive Optics

Stefan B. Ploner[1,2,✉], Weijia Fan[3], Hao F. Zhang[3], James G. Fujimoto[2], and Andreas Maier[1]

[1] Pattern Recognition Lab, FAU Erlangen-Nürnberg, Germany
stefan.ploner@fau.de
[2] Department of Electrical Engineering and Computer Science, and Research Laboratory of Electronics, Massachussets Institute of Technology, Cambridge, USA
[3] Department of Biomedical Engineering, Northwestern University, Evanston, USA

Abstract. Due to its historic limitations in imaging speed, for a long time, ophthalmic Optical Coherence Tomography (OCT) image analysis focused on features such as retinal layer thickness or fluids, which are visible in 2D B-scans. Although OCT angiography raised interest in en face (C-scan) images to visualize vascular layers, analysis of structural en face features like nerve fibers or photoreceptors was limited by transverse resolution, shot noise, and eye motion creating non-uniform displacements and distortion in raster-scanned volume images. Imaging of these structures is limited to much more complex adaptive optics imaging devices, which place a major hurdle for clinical translation, and complicate imaging by requiring pupil dilation. In this contribution, we integrate computational methods for motion, distortion and illumination compensation of repeated OCT volumes with orthogonally alternating fast scan directions, resulting in a consistent 3D point cloud, and apply the first fully-automatic inverse model for reconstructing 3D (multi-frame) super-resolved ophthalmic OCT images with higher sample density and reduced blur. In particular, we visualize reflections from individual cones in the outer segment with visible-light OCT across a $3 \times 3\,\mathrm{mm}^2$ field. Our method is unique in its comprehensive description of confounders of the imaging process, unlocks a novel paradigm of widefield OCT acquisition, and significantly broadens access to features of high clinical interest.

Keywords: Image Reconstruction · Super resolution · Ophthalmology.

1 Introduction

Optical coherence tomography (OCT) is a non-invasive 3D optical imaging modality that is standard of care in ophthalmology [6]. Modern OCT acquires individual 1D depth profiles (A-scans). Scanning along a line forms 2D cross-sectional images (B-scans), and raster scanning provides volumetric images. In ophthalmology, scanned imaging is prone to various image artifacts [25]. Most

importantly, various types of eye motion cause non-uniform and discontinuous displacements [14]. Slight changes of the optical path from eye motion, as well as temporal opacity variations of the vitreous, which is traversed by the beam before and after being reflected at the retina, modulate signal strength by a mostly low frequency bias between repeats. Vignetting, a partial or complete blockage of the beam by the iris, as well as eye blinks, can lead to substantial signal reduction or complete signal loss. Image artifacts increase for longer periods of fixation and limit the practically possible acquisition time [12]. Finally, images suffer from various sources of noise, especially shot noise and speckle.

Motion is typically compensated by eye tracking based on an additional high frame-rate modality like scanning laser ophthalmoscopy. This allows repositioning the A-scan locations during scanning and interruption of scanning during eye blinks. Tracking is typically limited by latency and only compensates for motion along the two transverse axes. An alternative approach that bypasses latency is retrospective registration of repeated volume scans [26]. Use of orthogonal raster scans allows 3D motion correction based on a small number of volumes [11,20]. While registering and merging additional scans increases signal-to-noise ratio (SNR), residual misregistrations and signal variation can cause new artifacts. This hindered the clinical translation of the first such method [25].

Although 3D super resolution was demonstrated in ophthalmic OCT by a pioneering work [22], evaluation is limited to a single healthy subject, manual selection of the 4 out of 12 best scans was required, the low resolution of the presented features had fewer requirements on motion compensation accuracy, and the aforementioned image artifacts were not considered. Especially lack of modeling the signal variation between samples from different volumes can cause high frequency artifacts in supersampled data, which may be further worsened by deconvolution. Super resolution increases motion correction requirements by the same factor as its increase in pixel density. This may explain its earlier adoption in OCT imaging domains with less motion [24] and why ophthalmic applications are limited to repeated B-scans (XZ plane) [5]. B-scans are much easier to align because of their minimal acquisition time. Due to the high noise levels in OCT, single frame methods usually only perform super sampling, but not deblurring [30,29]. All the latter methods do not apply to 3D images and the thereof extracted en face images (XY plane), which are the focus of this paper.

Substantial increase in transverse resolution is possible with adaptive optics (AO) [15], which compensates for aberrations in the optics of the eye. This allows wider beam diameters at the pupil, which also improve the SNR. However, AO-OCT requires additional expensive hardware, is limited to small imaging fields [1], and wide beams require prior pupil dilation. Therefore, current application is limited to research institutes. Shorter wavelengths, as used in visible-light OCT (visOCT), provide an alternative but limited resolution enhancement, which is less explored for transverse (lateral) imaging.

This work describes a computational method for reconstructing higher resolution 3D images from repeated (non-AO) retinal OCT volumes by inverting a unique, comprehensive imaging model. We combine recent developments in ret-

rospective motion correction [18], distortion compensation [16,21], illumination compensation [19] and artifact removal [17] with a novel method for 3D super resolution. Based on the consistent, oversampled point cloud resulting from the initial steps, we perform regularized inverse model optimization to reconstruct a deconvolved volume at a multiple of the pixel density of the inputs. We achieve substantial improvements in image quality and improve the visibility of small retinal features such as capillary vasculature. In particular, we demonstrate imaging of (mostly S-)cones across large fovea-centered fields in the outer segment (OS) in visOCT, $\gtrsim 3\,\mu\text{m}$ posterior to the (M/L-cone) inner segment (IS)/OS junction. This was previously limited to smaller fields imaged by AO devices.

2 Methods

We model the image formation of the v-th vectorized volume $\boldsymbol{y}^v \in (N_L \times 1)$ as

$$\boldsymbol{y}^v = \boldsymbol{B}(\tilde{\boldsymbol{a}}^v, \boldsymbol{\rho}^v)\boldsymbol{F}(\boldsymbol{d}^v)\boldsymbol{H}\boldsymbol{x} + \boldsymbol{e}^v \tag{1}$$

where $\boldsymbol{x} \in (N_H \times 1)$ represents the unknown vectorized high-resolution image of dimensions $N_H = \tilde{w} \cdot \tilde{h} \cdot \tilde{d}$. $\boldsymbol{H} \in (N_H \times N_H)$ models the point spread function, which we assume to be a location-independent anisotropic Gaussian with device-specific axial and transverse standard deviations σ_{ax} and σ_{tv}. The downsampling and backward warping operator $\boldsymbol{F} \in (N_L \times N_H)$ resamples the blurred high-resolution image at the $N_L = w \cdot h \cdot d$ low resolution voxel locations defined by the 3D A-scan displacement coefficients $\boldsymbol{d}^v \in (w \times h \times 3)$ [18]. The illumination bias operator $\boldsymbol{B}(\tilde{\boldsymbol{a}}, \boldsymbol{\rho}) = \text{diag}(\tilde{\boldsymbol{a}}.^{\boldsymbol{\rho}}) \in (N_L \times N_L)$ handles variations in signal strength (illumination), where $.^{\square}$ is the element-wise power, $\tilde{\boldsymbol{a}}^v \in (w \times h)$ are A-scan illumination coefficients, and $\boldsymbol{\rho}^v \in (w \times h \times d)$ is a foreground mask that cancels out compensation in the background [19]. $\boldsymbol{e}^v \in (N_L \times 1)$ is a noise term. In the following, we describe our method for restoring $\boldsymbol{x}$. It is outlined in Fig. 1.

2.1 Mapping of the input data to a consistent point cloud

The A-scan coefficients are computed following the methods described in [18] and [19], which we briefly summarize in this section. The methods require volumes to be raster-scanned with orthogonally alternating B-scan directions. The A-scan displacement coefficients $\boldsymbol{d}^v$ are embedded in a lower-dimensional temporal parameterization to model time-dependent motion during scanning with minimal dimensionality [18]. Although the retina is assumed to be rigid, due to optical effects, observed displacements are not. We use the extended distortion model of [16,21], which includes additional (time-dependent) parameters to model spatial sheering, curvature, and scaling. In contrast, illumination varies non-uniformly both in terms of time and space. Therefore, the illumination coefficients $\tilde{\boldsymbol{a}}^v$ are embedded in a spatiotemporal parameterization [19]. Spatial illumination variation is dominated by low frequencies, which enables substantial reduction of the parameters' dimensionality by utilizing their continuity. Spline

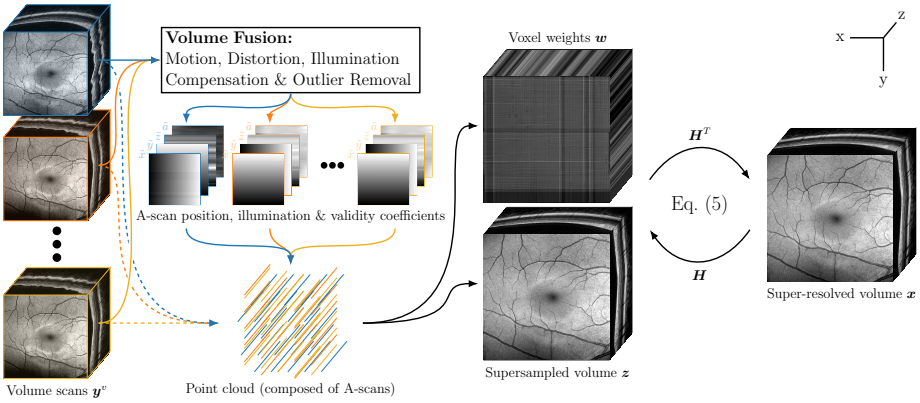

Fig. 1. High-level sketch of volume fusion & iterative super-resolution. Gridding the point cloud in $\boldsymbol{w}$ and $\boldsymbol{z}$ benefits memory requirements ($\sim 3{\cdot}N_H$) when the oversampling rate is greater than the upsampling factor, and runtime.

interpolation is used to map both types of sparse parameters back to dense A-scan coefficients. Because all volumes are affected by non-uniform displacements and illumination variation, there is no fixed registration target.

All parameters are then iteratively optimized to minimize sum-of-squared differences between log-scale intensities following an expectation-maximization scheme. For this, the current parameters are mapped to A-scan coefficients and then used to invert the illumination bias and warping operations. The backward warping of the image model is inverted by computing a corresponding forward warping. (Note that this corresponds to per-volume estimates of a blurry low-resolution image $\boldsymbol{\ell}^v \hat{=} \boldsymbol{DHx}$, where $\boldsymbol{D}$ downsamples to the input resolution.) Updates of the parameters of (moving) volume m are then computed by backward warping each orthogonal target volume $\boldsymbol{\ell}^t$ (target index t) using the A-scan coefficients of volume m. This maps the target volumes to the voxels of m, which allows computation of the similarity metric. The metric is derived for the A-scan coefficients of m, and updates are propagated through the motion/illumination models and (spatio-)temporal parameterizations to the parameters via the chain rule. In this process, compensation of aliasing effects in the objective function is crucial to achieve sufficient accuracy for the subsequent super-resolution task. This is achieved with marginal overhead by sampling $\boldsymbol{\ell}^t$ at pseudo-random voxel locations to break the accumulation of aliasing patterns that could occur between uniform grids. At this point, we refer the interested reader to [18].

In contrast to the sequential motion and illumination optimization in [19], we perform both tasks jointly and in a multi-resolution pyramid. Joint optimization is important for reliable motion estimation in elderly subjects, which exhibit stronger illumination variation. This reduces the number of required iterations to those of the more challenging task (motion estimation). At this point, preliminary evaluation suggests that most A-scan position errors are reduced

to sub-micron levels or ˜1/10th of the original pixel spacing [21, Fig. 3], and moderate variations in signal strength are compensated. In a final step following [17], we remove inconsistent data such as blinks, severe signal variation, or misregistration, by thresholding the B-scan-aggregated similarity metric. This part of the pipeline proved its robustness in several clinical studies [27,7,28].

2.2 Super-resolution image reconstruction

In this section, we treat all non-excluded data equally as an oversampled point cloud and therefore drop the volume superscripts. Since outliers were removed, we formulate the objective function as a classic least squares minimization

$$\operatorname{argmin}_{\boldsymbol{x}} ||\boldsymbol{y} - \boldsymbol{B}(\tilde{\boldsymbol{a}}, \boldsymbol{\rho})\boldsymbol{F}(\boldsymbol{d})\boldsymbol{H}\boldsymbol{x}||_2^2 . \tag{2}$$

Following the general concept of the fast super resolution formulation in [3], we split the problem into non-iterative fusion, followed by iterative deblurring and gap interpolation. The fused (high-resolution) image is computed as

$$\boldsymbol{z} = \boldsymbol{F}^{-1}(\boldsymbol{d})\boldsymbol{B}^{-1}(\tilde{\boldsymbol{a}}, \boldsymbol{\rho})\boldsymbol{y}. \tag{3}$$

where the diagonal matrix $\boldsymbol{B}$ is trivially invertible. The inverse of the backward warping and downsampling operator $\boldsymbol{F}$ is computed as hermite spline interpolation along each (undistorted) A-scan in axial direction, followed by forward warping with a nearest-neighbor splatting kernel in the transverse directions. Note that $\boldsymbol{F}$ is typically not invertible, especially in super-resolution settings, which manifests as gaps in $\boldsymbol{z}$. We then optimize the simplified substitute problem

$$\operatorname{argmin}_{\boldsymbol{x}} ||\boldsymbol{W}\,(\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x})||_2^2 + \lambda||\boldsymbol{\Gamma}\boldsymbol{x}||_2^2, \tag{4}$$

where $\boldsymbol{W} = \operatorname{diag}(\boldsymbol{w})$ with $\boldsymbol{w} = \operatorname{sqrt}\left(\boldsymbol{F}^{-1}(\boldsymbol{d}) \cdot \boldsymbol{1}\right)$, $\dim(\boldsymbol{1}) = \dim(\boldsymbol{y})$ is a weight matrix whose diagonal entries correspond to the summed weights of the splatting kernel of $\boldsymbol{F}^{-1}$ that correspond to the voxels in $\boldsymbol{z}$. The square root is intended to make the weights more related to the SNR of $\boldsymbol{z}$. To compensate the non-invertability of $\boldsymbol{F}$, we add an anisotropic Tikhonov regularization term. $\boldsymbol{\Gamma}$ defines 3D 7-neighborhood Laplace kernels, where the axial neighbors have a weight factor γ. This fills gaps and evens out the noise levels in combination with $\boldsymbol{W}$.

We initialize the super-resolved image $\boldsymbol{x}$ with $\boldsymbol{z}$, fill missing values in $\boldsymbol{x}$ via interpolation to speed up convergence, and replace NaN values in $\boldsymbol{z}$ with zeros (to silently ignore zero-weighted entries). The objective function is minimized by stepping along the negative gradient according to the following update scheme:

$$\boldsymbol{x}^{(x+1)} = \boldsymbol{x}^{(x)} + \delta \cdot \left(2\boldsymbol{H}^T\boldsymbol{W}^2(\boldsymbol{z} - \boldsymbol{H}\boldsymbol{x}) + \lambda\boldsymbol{\Gamma}\boldsymbol{x}\right) \tag{5}$$

Here, $\boldsymbol{H}^T = \boldsymbol{H}$, the diagonal matrix $\boldsymbol{W}$ is trivially squarable, and δ is the step size. We used 3^2 (similar to volume repeats) transverse upsampling ($(\tilde{w}, \tilde{h}, \tilde{d}) = (3w, 3h, d)$), deconvolution $\sigma_{\mathrm{tv}} = 2.9\,\mu\mathrm{m}$ ($\lesssim$ upsampled spacing) and $\sigma_{\mathrm{ax}} = \mathrm{FWHM}/(2\sqrt{2\ln 2}) = 0.6\,\mu\mathrm{m}$ with full width at half maximum $\mathrm{FWHM} = 1.4\,\mu\mathrm{m}$,

regularizer weights $\lambda = 0.2$ and $\gamma = 0.1$ (small to avoid blur), and a step size of $\delta = 0.12$ (safe margin below observed divergence, scale with inverse value range).

The entire method was implemented end-to-end in CUDA 12.8. If a specific en face slice is of interest, a layer depth map can be incorporated into the fusion step by adding it to the axial A-scan displacement coefficients $\boldsymbol{d}_z^v$. We compute layer segmentations in fused volumes, interpret the resulting depth map as an en face image, and backward warp to the input A-scans via $\boldsymbol{F}$. The impact of performing fusion and, consequently, image reconstruction in such a flattened volume is usually low, but can reduce blur at tilted layer boundaries as introduced by the forward warping transverse splatting kernel and regularization in Eqs. 3 and 4.

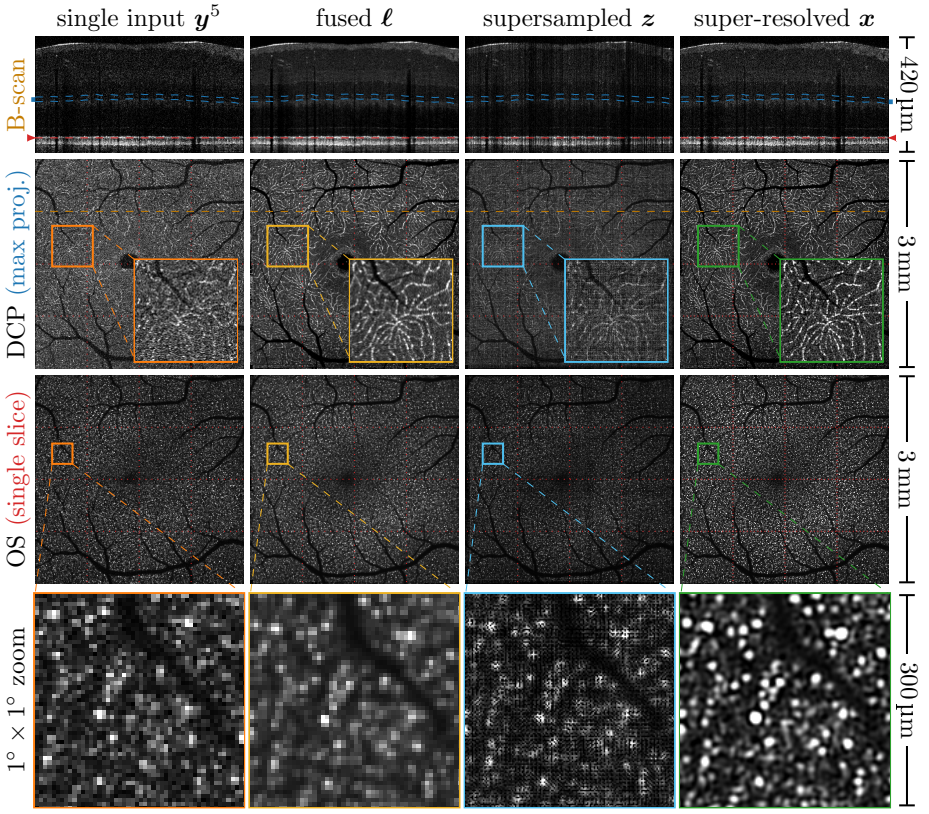

Fig. 2. Comparison of intermediate steps. $\boldsymbol{y}^5$: Representative input volume. $\boldsymbol{\ell}$: 10 volumes fused at original sampling density. $\boldsymbol{z}$: Fused densly sampled volume, gaps in black. $\boldsymbol{x}$: Super-resolved volume. Top: Logarithmic intensity-scale B-scans. Dashed lines specify the projected depth range (DCP, blue) / the depth position (OS, red) of the below en face images. Middle: Max projection over the 17 µm anterior to the inner nuclear layer / outer plexiform layer boundary showing the deep capillary plexus (DCP). Horizontal dashed lines mark the position of the B-scans. Bottom: Single slice, ˜3 µm posterior to the IS/OS junction. Dotted gridlines were added for orientation.

3 Results

Eight visible-light (510 – 610 nm) OCT datasets were acquired from healthy volunteers with a prototype device at Northwestern University [23]. 10 repeated volumes of 512×512 A-scans were acquired over a $3\,\text{mm} \times 3\,\text{mm}$ field ($10° \times 10°$) at an A-scan rate of 100 kHz and a 1.1 mm beam diameter ($\frac{1}{e^2}$) at the pupil (no dilation). Imaging was performed in agreement with the declaration of Helsinki, and in accordance with the institutional review board's approved protocols.

Figure 2 shows a comparison of intermediate steps. Volumes were fused and reconstructed in a flattened volume relative to the posterior inner nuclear layer boundary and the IS/OS junction (ellipsoid zone), respectively. The input image A-scan depth positions were interpolated at the depth of the fused volume to prevent layer segmentation-related differences, while their transverse positions were kept original and therefore contain displacement artifacts. The super-resolved image shows substantial improvements in contrast, sharpness, and roundness of the reflections in the OS. Note that reflections originate from different depths. Therefore, some reflections appear with reduced brightness in this specific slice. While the reflections cannot be fully resolved in the fovea exactly at the (M/L-

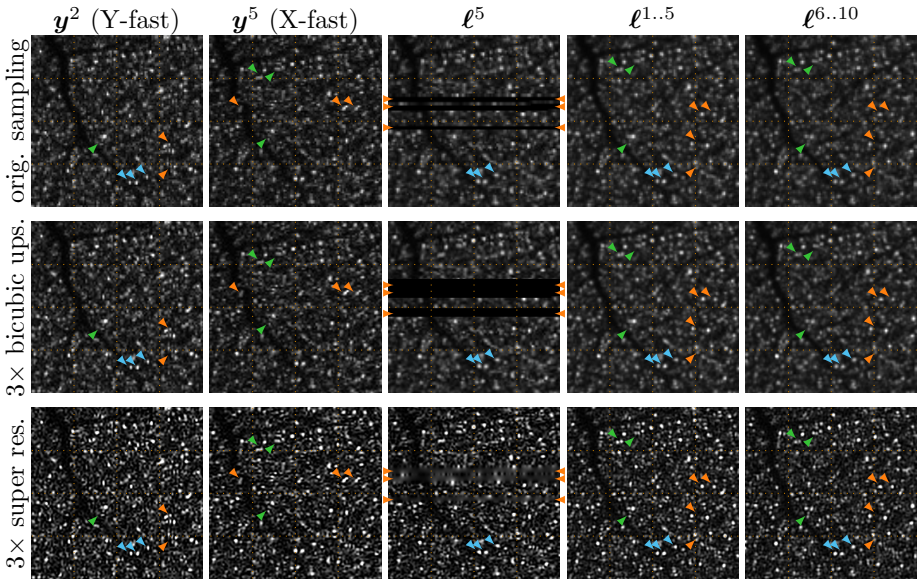

Fig. 3. Reproducibility and ablation in the outer segment, cutouts (586 µm^2) similar to Fig. 2. $\boldsymbol{y}^2$, $\boldsymbol{y}^5$: Second/fifth scan without motion compensation show displacements. $\boldsymbol{\ell}^5$, $\boldsymbol{\ell}^{1..5}$, $\boldsymbol{\ell}^{6..10}$: Fifth / first five / second five scan(s) after compensation. Motion correction removes doubling (orange markers) and improves cone shape (blue markers). Fusion (right) fills gaps (orange markers in $\boldsymbol{\ell}^5$) and improves consistency of reflections. Rows compare original sampling, naive bicubic upsampling, and super resolution (SR). SR improves contrast and sharpness globally, which improves separability (green markers).

cone) IS/OS junction and cone outer segment tips bands, the density thins out to imagable levels within 2 – 3 microns and imaging is possible throughout the outer segment. At ˜4.5 – 8 µm posterior to the IS/OS, our images look very similar to the AO OCT slice labeled "IS/OS: S cones" in [15, Fig. 5].

Figure 3 demonstrates reproducibility of the super resolution image enhancement (bottom) in the OS by performing it separately on a distinct subset of the fused volumes (first five vs. second five in the right two columns). Compared to the original images, motion correction removes doubling and improves shape, while volume fusion fills gaps and improves consistency of reflections. SR improves contrast and sharpness, which also improves separability of reflections.

4 Discussion

Our framework for 3D motion correction, volume fusion, and multi-frame super resolution consistently removes scanning-related displacements and further artifacts from in-vivo human retinal images, and can substantially improve visibility of small image features. Regular volume fusion primarily improves SNR. Since OCT is dominated by intensity-dependent shot noise, this particularly benefits low signal regions and, e. g., aided in the discovery of a very weakly reflecting OCT band [7]. In contrast, super resolution is particularly beneficial in strongly-reflecting, higher-SNR regions like photoreceptor cone reflections. To the best of our knowledge, the enhancement of such reflections in the outer segment is unique in non-adaptive optics OCT data with standard scanning speed and beam diameter, i. e., without pupil dilation. The reflections in the outer segment may be dominated by S-cones [15], which are sparser than M- and L-cones, and are known to differ in inner and outer segment lengths by, on average, 5 and -5 µm [2,8]. This shifts their junction away from the denser reflections in the typical (M/L-cone) IS/OS band. However, a definitive per-cell distinction cannot be made due to variations in depth position. When based on visible-light OCT, our approach benefits from ˜1.5× increased resolution in the transverse directions over regular NIR OCT, and is expected to outperform current AO OCTs in axial resolution (1.4 µm vs., e. g., 2.4 µm in [8]). Initial high-(axial-)resolution near infrared (NIR) OCT data also suggests potential for imaging the sparse reflections in the OS. In addition, being able to enhance such minute features proves the high effectiveness in motion and distortion compensation of our overall pipeline.

Although our method achieves reproducible enhancement under equivalent imaging conditions, photoreceptor imaging is challenged by their waveguide property, particularly at the IS/OS junction [13]. Additional variation is caused by scintillation, a short-term reflectance change after functional stimulation that is triggered by the scanning laser beam [9], and cell renewal (phagocytosis) [10].

A further limitation of our method is that the denser M- and L-cones are not resolved in the fovea, and particularly weakly reflecting structures such as ganglion cells cannot be imaged. This is possible with AO imaging, which compensates aberrations of the eye at the cost of additional system complexity and, therefore, can use larger beam sizes at the pupil to enhance SNR. While our im-

provement is on a different level, our method is purely computational and works in comparatively large imaging fields (limited by imaging time). Note that a limited degree of aberrations may be averaged out due to motion-induced differing ray paths between volumes. By not requiring pupil dilation, wait times of 15 to 30 minutes, as well as restrictions after imaging due to the associated glare, e. g., in driving, can be avoided. It would be interesting to investigate whether a combination of both approaches can further push the boundaries of high-resolution retinal imaging. Although being able to image only one cone type may seem as a strong limitation, since S-cones are known to be more vulnerable than other cone types in many diseases [4], they are most relevant as an early biomarker, and may predict the degeneration of the other cone types.

Author contributions Conceptualization, methodology, implementation, visualization, and original manuscript: SP. Imaging, raw signal reconstruction, data curation and manuscript review: WF. Funding: SP, AM, JF, HZ.

Acknowledgments. This study was funded by the German Research Foundation (508075009), the National Institutes of Health (5-R01-EY011289-39, 5-U01-EY033001) and Northwestern University (Christina Enroth-Cugell Graduate Research Award).

Disclosure of Interests. Author HZ is co-founder of Opticent, Inc. Author JF has IP on orthogonal motion correction, receives research funding from Topcon Healthcare, Inc., and has personal financial interest in Pixcel, Inc.

References

1. Bedggood, P., Daaboul, M., Ashman, R.A., et al.: Characteristics of the human isoplanatic patch and implications for adaptive optics retinal imaging. J. Biomed. Opt. **13**(2), 024008 (2008). https://doi.org/10.1117/1.2907211
2. Curcio, C.A., Allen, K.A., Sloan, K.R., et al: Distribution and morphology of human cone photoreceptors stained with anti-blue opsin. J. Comp. Neurol. **312**(4), 610–624 (1991). https://doi.org/https://doi.org/10.1002/cne.903120411
3. Farsiu, S., Robinson, M., Elad, M., Milanfar, P.: Fast and robust multiframe super resolution. IEEE Trans. Image Process. **13**(10), 1327–1344 (2004). https://doi.org/10.1109/TIP.2004.834669
4. Greenstein, V.C., Hood, D.C., Ritch, R., Steinberger, D., Carr, R.E.: S (blue) cone pathway vulnerability in retinitis pigmentosa, diabetes and glaucoma. Invest. Ophthalmol. Vis. Sci. **30**(8), 1732–1737 (08 1989)
5. He, Z., Xiao, Z., Xu, Z., et al: MFGAN: OCT image super-resolution and enhancement with blind degradation and multi-frame fusion. In: International Workshop on Advanced Imaging Technology (IWAIT) 2025. vol. 13510, p. 1351005. SPIE (2025). https://doi.org/10.1117/12.3057230
6. Huang, D., Swanson, E., Lin, C., et al.: Optical Coherence Tomography. Science **254**(5035), 1178–1181 (1991). https://doi.org/10.1126/science.1957169
7. Jamil, M., Won, J., Ploner, S., et al.: High-Resolution OCT Reveals Age-Associated Variation in the Region Posterior to the External Limiting Membrane. Transl. Vis. Sci. Technol. **14**(1), 16–16 (01 2025). https://doi.org/10.1167/tvst.14.1.16

8. Ji, Q., Bernucci, M.T., Liu, Y., et al: Structural analysis of cone photoreceptors in AO-OCT enables S-cone identification by a support vector machine. Biomed. Opt. Express **17**(1), 346–364 (Jan 2026). https://doi.org/10.1364/BOE.581923
9. Jonnal, R.S., Rha, J., Zhang, Y., Cense, B., Gao, W., Miller, D.T.: In vivo functional imaging of human cone photoreceptors. Opt. Express **15**(24), 16141–16160 (Nov 2007). https://doi.org/10.1364/OE.15.016141
10. Kocaoglu, O.P., Lee, S., Jonnal, R.S., Wang, Q., Herde, A.E., Besecker, J., Gao, W., Miller, D.T.: 3D imaging of cone photoreceptors over extended time periods using optical coherence tomography with adaptive optics. In: Manns, F., Söderberg, P.G., Ho, A. (eds.) Ophthalmic Technologies XXI. vol. 7885, p. 78850C. International Society for Optics and Photonics, SPIE (2011). https://doi.org/10.1117/12.878229
11. Kraus, M.F., Liu, J.J., et al: Quantitative 3D-OCT motion correction with tilt and illumination correction, robust similarity measure and regularization. Biomed. Opt. Express **5**(8), 2591–2613 (Aug 2014). https://doi.org/10.1364/BOE.5.002591
12. Lin, W.C., Coyner, A.S., Amankwa, C.E., et al: High prevalence of artifacts in optical coherence tomography with adequate signal strength. Transl. Vis. Sci. Technol. **13**(8), 43–43 (08 2024). https://doi.org/10.1167/tvst.13.8.43
13. Marsh-Armstrong, B., Murrell, K., Valente, D., Jonnal, R.S.: Using directional OCT to analyze photoreceptor visibility over AMD-related drusen. Sci. Rep. **12**(1), 9763 (2022). https://doi.org/10.1038/s41598-022-13106-3
14. Martinez-Conde, S., Macknik, S.L., Hubel, D.H.: The role of fixational eye movements in visual perception. Nat. Rev. Neurosci. **5**(3), 229–240 (2004). https://doi.org/10.1038/nrn1348
15. Miller, D.T., Kurokawa, K.: Cellular-scale imaging of transparent retinal structures and processes using adaptive optics optical coherence tomography. Annu. Rev. Vis. Sci. **6**(Volume 6, 2020), 115–148 (2020). https://doi.org/https://doi.org/10.1146/annurev-vision-030320-041255
16. Ploner, S., Babiker, F., Hwang, Y., et al: Quantification and compensation of 2nd order image distortions in high resolution OCT volume fusion and longitudinal registration. Invest. Ophthalmol. Vis. Sci. **66**, PB0088 (7 2025)
17. Ploner, S., Won, J., Takahashi, H., et al: A reliable, fully-automatic pipeline for 3D motion correction and volume fusion enables investigation of smaller and lower-contrast OCT features. Invest. Ophthalmol. Vis. Sci. **65**, 5909 (6 2024)
18. Ploner, S., Chen, S., Won, J., et al: A Spatiotemporal Model for Precise and Efficient Fully-Automatic 3D Motion Correction in OCT. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2022. pp. 517–527. Springer Nature Switzerland, Cham (2022). https://doi.org/10.1007/978-3-031-16434-7_50
19. Ploner, S., Won, J., Schottenhamml, J., et al: A Spatiotemporal Illumination Model for 3D Image Fusion in Optical Coherence Tomography. In: 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI). pp. 1–5 (2023). https://doi.org/10.1109/ISBI53787.2023.10230526
20. Ploner, S.B., Kraus, M.F., Moult, E.M., et al: Efficient and high accuracy 3-D OCT angiography motion correction in pathology. Biomed. Opt. Express **12**(1), 125–146 (Jan 2021). https://doi.org/10.1364/BOE.411117
21. Ploner, S.B., Schneider, L.S., Hwang, Y., et al.: Validation of simulation-based evaluation by inclusion of unlabeled data and meta-ablation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2026. Springer Nature Switzerland, Cham (2026)
22. Robinson, M., Chiu, S., Lo, J., et al: Novel applications of super-resolution in medical imaging. In: Super-Resolution Imaging, pp. 383–412. CRC Press (2010)

23. Rubinoff, I., Miller, D.A., Kuranov, R., et al.: High-speed balanced-detection visible-light optical coherence tomography in the human retina using subpixel spectrometer calibration. IEEE Trans. Med. Imag. **41**(7), 1724–1734 (2022). https://doi.org/10.1109/TMI.2022.3147497
24. Shen, K., Lu, H., Baig, S., Wang, M.R.: Improving lateral resolution and image quality of optical coherence tomography by the multi-frame superresolution technique for 3D tissue imaging. Biomed. Opt. Express **8**(11), 4887–4918 (Nov 2017). https://doi.org/10.1364/BOE.8.004887
25. Spaide, R.F., Fujimoto, J.G., Waheed, N.K.: Image artifacts in optical coherence tomography angiography. Retina **35**(11), 2163–2180 (2015). https://doi.org/10.1097/IAE.0000000000000765
26. Sánchez Brea, L., Andrade De Jesus, D., Shirazi, M.F., et al.: Review on Retrospective Procedures to Correct Retinal Motion Artefacts in OCT Imaging. Appl. Sci. **9**(13) (2019). https://doi.org/10.3390/app9132700
27. Won, J., Takahashi, H., Ploner, S.B., et al.: Topographic measurement of the subretinal pigment epithelium space in normal aging and age-related macular degeneration using high-resolution oct. Invest. Ophthalmol. Vis. Sci. **65**(10), 18–18 (08 2024). https://doi.org/10.1167/iovs.65.10.18
28. Won, J., Yaghy, A., Ploner, S.B., et al.: High-resolution, motion-corrected, volume-fused OCT for investigating longitudinal changes in subretinal drusenoid deposits in intermediate amd. Transl. Vis. Sci. Technol. **14**(6), 15–15 (06 2025). https://doi.org/10.1167/tvst.14.6.15
29. Xia, W., Han, T., Tao, K., et al.: Super-resolution reconstruction of OCT images based on frequency and spatial information in adversarial neural networks. Phys. Med. Biol. **70**(23), 235006 (11 2025). https://doi.org/10.1088/1361-6560/ae1c17
30. Yao, B., Jin, L., Hu, J., Liu, Y., Yan, Y., Li, Q., Lu, Y.: PSCAT: a lightweight transformer for simultaneous denoising and super-resolution of OCT images. Biomed. Opt. Express **15**(5), 2958–2976 (May 2024). https://doi.org/10.1364/BOE.521453