tvst

https://doi.org/10.1167/tvst.8.1.10

Article

The Utility of Frame Averaging for Automated Algorithms in Analyzing Retinal Vascular Biomarkers in AngioVue OCTA

Taly Gilat Schmidt[1,*], Rachel E. Linderman[2,*], Margaret R. Strampe[3,4], Toco Y. P. Chui[5,6], Richard B. Rosen[5,6], and Joseph Carroll[2,3]

[1] Department of Biomedical Engineering, Marquette University and Medical College of Wisconsin, Milwaukee, WI, USA
[2] Department of Cell Biology, Neurobiology, & Anatomy, Medical College of Wisconsin, Milwaukee, WI, USA
[3] Department of Ophthalmology & Visual Sciences, Medical College of Wisconsin, Milwaukee, WI, USA
[4] University of Minnesota Medical School, Minneapolis, MN, USA
[5] Department of Ophthalmology, New York Eye and Ear Infirmary of Mount Sinai, New York, NY, USA
[6] Icahn School of Medicine at Mount Sinai, New York, NY, USA

Correspondence: Taly Gilat Schmidt, Department of Biomedical Engineering, Marquette University and Medical College of Wisconsin, PO Box 1881, Milwaukee, WI 53201-1881. e-mail: tal.gilat-schmidt@marquette.edu

Received: 10 May 2018
Accepted: 12 November 2018
Published: 18 January 2019

Keywords: optical coherence tomography angiography; foveal avascular zone; imaging; averaging

Citation: Schmidt TG, Linderman RE, Strampe MR, Chui TYP, Rosen RB, Carroll J. The utility of frame averaging for automated algorithms in analyzing retinal vascular biomarkers in AngioVue OCTA. Trans Vis Sci Tech. 2019;8(1):10, https://doi.org/10.1167/tvst.8.1.10

Purpose: This study proposes an optical coherence tomography angiography (OCTA) frame-averaging method and investigates the effects of the number of frames acquired and averaged on metrics quantifying the foveal avascular zone (FAZ), vessel morphology, and parafoveal intercapillary area (PICA).

Methods: Ten OCTA frames were acquired for each of the 19 subjects without known retinal disease using the AngioVue OCTA system. For each subject, acquired frames were ranked by an image quality metric. A subset of frames was then registered and averaged. The effects of the number of frames acquired and averaged on FAZ segmentation and metrics of FAZ geometry, vessel morphology, and PICA were analyzed.

Results: Frame averaging increased the accuracy of the automatically segmented FAZ region; for example, the absolute error in FAZ area decreased from 0.026 mm^2 (1 frame) to 0.005 mm^2 (5 frames). Averaging multiple frames exponentially decreased the estimated number of vessel endpoints and increased the average vessel length with a 32% decrease in number of endpoints and 14% increase in average vessel length when averaging five frames compared with one. Frame averaging also improved the precision of PICA estimates.

Conclusions: Averaging multiple OCTA frames using the Optovue AngioVue system reduced error in FAZ segmentation and improved the robustness of OCTA vessel morphology and perfusion metrics. The study demonstrated limited benefit in acquiring and averaging more than five frames.

Translational Relevance: Averaging multiple OCTA frames improved the robustness of OCTA foveal biomarkers with limited benefit when averaging more than five frames.

Introduction

The retina has one of the highest oxygen demands in the body.[1] Due to the demands placed on the retinal vasculature, it is susceptible to pathologic changes in a multitude of diseases both ocular[2–4] and systemic[5,6] in nature. Capillary dropout can be found in a variety of diseases, including sickle cell retinopathy[7–9] and diabetic retinopathy,[10–12] whereas pathologic growth of blood vessels is well documented in wet age-related macular degeneration,[3] proliferative diabetic retinopathy,[12–14] and retinopathy of prematurity.[15–17] Fluorescein angiography has been the gold standard for assessing the retinal vasculature since the method was first described in the 1960s.[18] Yet, use of this test is limited due its invasive nature, along with the risks posed by injection of the dye, which include nausea, vomiting, and anaphylactic shock.[19,20] Furthermore, visualization of microvasculature within

the macula is limited in comparison to other devices either due to the resolution of the fundus camera[6,21] or due to the leakage of dye caused by damaged vessels.[22,23]

Because of the limited resolution of the camera when using fluorescein angiography, measurements of the foveal avascular zone (FAZ), an area devoid of capillaries, including the fovea, is limited.[6,21] With the increased resolution available using optical coherence tomography angiography (OCTA), measurements of the FAZ and parafoveal microvasculature have become more accurate, leading to their increased use in clinical practice. With more detailed imaging available, it is of interest to both clinicians and scientists to develop robust biomarkers that provide useful information about retinal health.

Image noise and blur may make robust estimation of OCTA metrics more difficult. Image noise may lead to erroneous detection of a vessel structure or may erroneously fragment an identified vessel. Image blurring, commonly caused by eye motion, weakens vessel edges. Recent studies have shown that averaging multiple OCTA images improves metrics of image quality for the peripapillary area,[24] the choriocapillaris,[25] and the parafoveal region.[26] These studies averaged up to 10 frames after registering all frames to the subjectively selected highest-quality frame. Despite the clear improvements in image quality, the computational overhead to perform image averaging as well as the added time it takes to acquire multiple images makes such an approach challenging for clinical applications. It is critical to assess the "return on investment" for image averaging, as clinical translation of the method will require using the fewest number of images necessary.

This study proposes an OCTA averaging protocol of (1) acquiring multiple OCTA frames, (2) ranking frames using an image quality metric, (3) registering the OCTA frames, and (4) averaging a subset of registered OCTA frames. This study investigated the effects of the number of frames acquired and averaged on metrics quantifying the FAZ, foveal vessel geometry, and parafoveal intercapillary area (PICA).

Methods

Subjects

This study was approved by the institutional review board of the Medical College of Wisconsin (PRO 23999) and was conducted in accordance with the tenets of the Declaration of Helsinki. Informed consent was obtained for all subjects once the nature and risks of the study were explained. Exclusion criteria included any prior history or clinical evidence of retinal or systemic vascular disease. We imaged 19 Caucasian subjects (3 male, 16 female) with ages ranging from 23- to 49-years old.

OCTA Imaging

Subjects were imaged with the AngioVue OCTA system (software version: 2016.2.0.16; Optovue Inc., Fremont, CA). Two scans, one fast horizontal and one fast vertical were acquired, centered on the fovea of the right eye. The nominal size of each scan was 3×3 mm. The two scans were then co-registered to create one volume scan to minimize motion artifacts. Ten such volume OCTA images were acquired for each subject. The en face image of the superficial slab (304×304 pixels), measuring 3 μm above the inner limiting membrane to 16 μm below the inner plexiform layer, was exported for further analysis.

Multiple Frame Acquisition and Automated OCTA Frame Averaging

Figure 1 illustrates the developed OCTA frame processing workflow, for which M acquired OCTA frames are input for processing. For each frame, the gradient magnitude image was calculated using the Sobel operator (v2017a; MATLAB; MathWorks, Natick, MA).[27] The mean gradient magnitude of each frame was calculated as a metric of frame image quality to quantify the sharpness and contrast in the image. The image frames were then sorted from highest to lowest mean gradient magnitude. The N frames with highest mean gradient magnitude were rigidly registered to the frame with highest mean gradient magnitude (StackReg,[28] ImageJ[29]). The N registered frames were then averaged, and the averaged image was further processed to calculate the OCTA biomarkers described in the following sections.

We performed two studies to investigate the effects of multiple frame acquisition and multiple frame averaging. In the first study, the number of frames input to the algorithm, M, was set equal to 10, while the number of frames averaged, N, was varied between one and 10. The second study investigated the combined effects of multiple frame acquisition and averaging. The number of frames input to the workflow, M, was varied between one and 10, and the

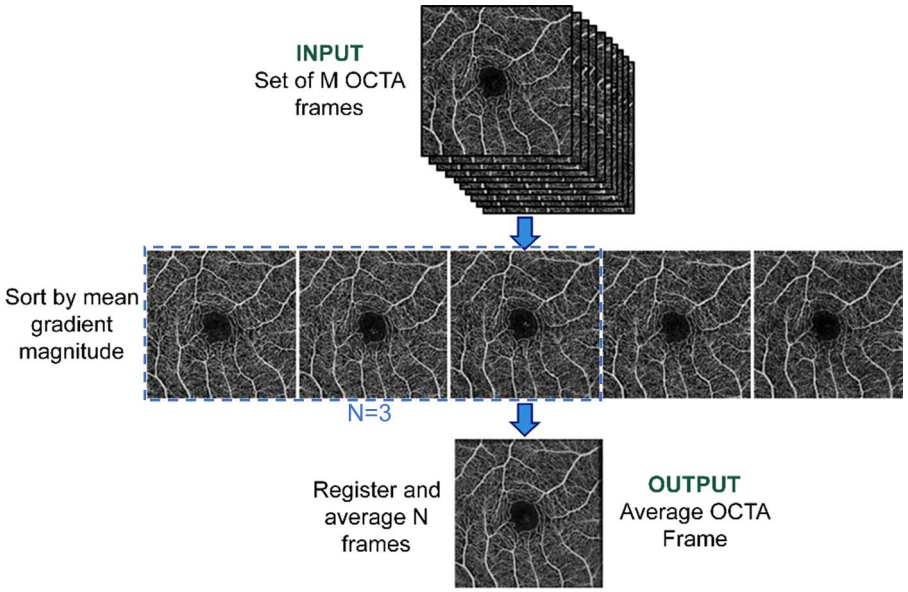

Figure 1. Overview of OCTA frame averaging workflow. The figure shows the example where $M = 5$ acquired frames were input to the algorithm and sorted by mean gradient magnitude. The $N = 3$ frames with highest mean gradient magnitude were registered and averaged.

number of averaged frames, N, was varied between one and M. For this second study, the M frames input to the algorithm for each subject were selected from the 10 acquired frames in the chronologic order in which they were acquired.

Effect of Multiple Frame Acquisition and Averaging on FAZ Segmentation and Quantification

For each of the 19 subjects, the FAZ region was manually segmented by three masked, expert readers, with manual segmentation performed on the image resulting from registering and averaging the five frames with highest mean gradient magnitude. We selected the average of the five highest quality frames as the reference image for ground-truth segmentation because averaging additional lower-quality frames could degrade the resulting image quality when combined with potential registration errors. Two metrics were analyzed to evaluate the agreement between readers. The Dice coefficient[30] of similarity was calculated between the FAZ regions segmented by all pairs of readers for each subject. The coefficient of variation (CV) of the segmented FAZ area was also calculated across readers for each subject. The average Dice coefficient of similarity between readers was 0.986 ± 0.01, with a minimum Dice coefficient of 0.946. The average CV across readers was 1% (95% confidence interval = 0.55%–1.34%). The high agreement between readers suggest that the manual FAZ segmentations provide a reasonable ground truth for assessing the impact of averaging on the image metrics. The three manually segmented regions were processed by the Simultaneous Truth and Performance Level Estimation algorithm[31] to produce the final ground-truth FAZ region as all pixels with greater than 90% probability of being in the FAZ region.

For each case of the number of frames averaged and acquired, the averaged frame was input to a custom FAZ segmentation algorithm, which is described in the Appendix. A custom FAZ segmentation algorithm was developed because commercial algorithms are not available for offline processing of the resulting averaged image. To investigate the effects of frame averaging on FAZ metrics, the agreement between the algorithm-generated FAZ segmentation and ground-truth FAZ segmentation was quantified by Dice coefficient.[30] The area, perimeter, and centroid of the FAZ regions were calculated as FAZ metrics using previously described methods and a custom MATLAB script.[32] All FAZ region metrics were corrected for axial length, as previously described.[32] Errors in the FAZ region metrics were calculated as the difference between the values estimated using the algorithm-segmented FAZ and the values estimated using the ground-truth FAZ.

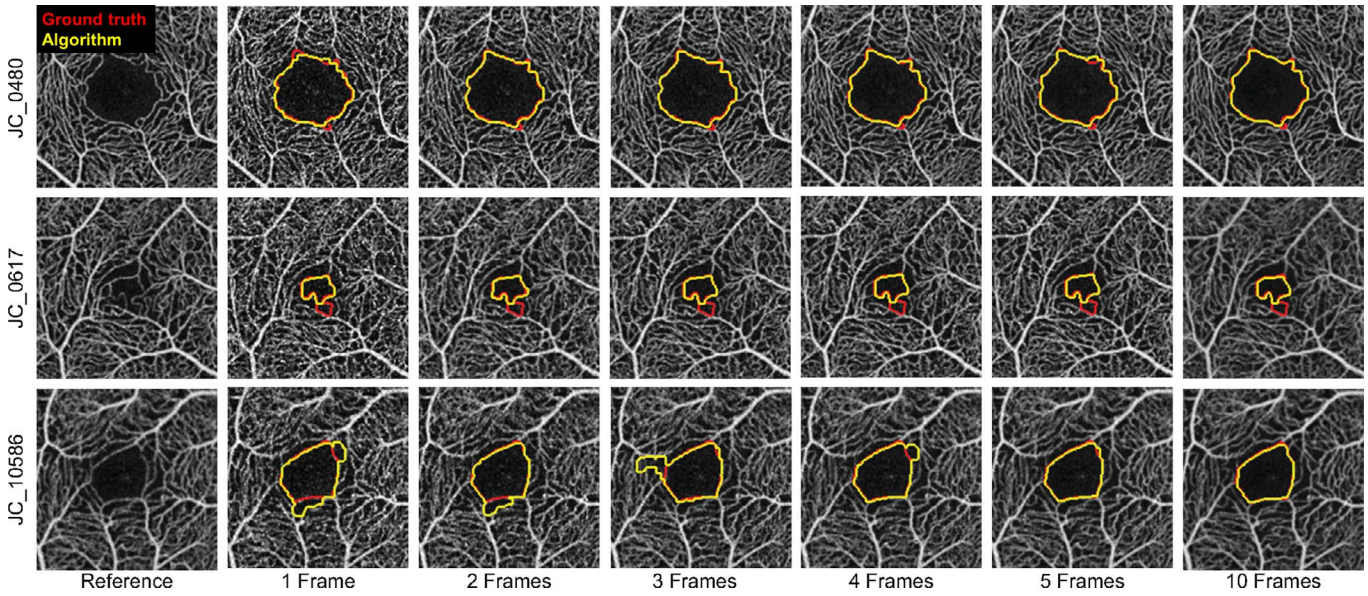

Figure 2. FAZ segmentation results for three subjects for which ten OCTA frames were acquired. The reference image on the left is the average of the five frames identified with highest mean gradient magnitude. The reference image was used to obtain the ground truth reader FAZ segmentations. For each subject, the average of one to five OCTA frames identified with highest image quality is presented, with the ground-truth FAZ segmentation in the *red contour* and algorithm-generated FAZ segmentation in the *yellow contour*. The result of averaging all 10 acquired frames is displayed on the *right*.

Effect of Multiple Frame Acquisition and Averaging on Parafoveal Intercapillary Area (PICA)

Additional analysis was performed to quantify the PICA on the averaged OCTA images. This analysis required the segmented FAZ region identified by the automated algorithm. The input OCTA image and FAZ region were magnified by a factor of six to increase sensitivity. The average PICA within 500 µm of the FAZ was computed using a previously described, custom MATLAB script[33] along with the standard deviation of the average PICA. The effect of frame averaging on the precision of the estimated PICA was evaluated by calculating the CV as the ratio of the standard deviation of the average PICA to the average PICA.

Effect of Multiple Frame Acquisition and Averaging on FIJI AnalyzeSkeleton Metrics

Vessel morphology was assessed by processing each image using the AnalyzeSkeleton 2D/3D[34] Fiji plugin.[35] The total number of endpoints, average vessel length (pixels), and longest vessel length (pixels) were extracted as quantitative metrics. The change in the metrics with respect to the number of averaged frames was analyzed with 10 acquired frames input for processing, and with the number of input frames varied between one and 10.

Results

Results of FAZ Segmentation Algorithm for Ten Acquired Frames and Varying Number of Averaged Frames

Figure 2 displays the results of the automated FAZ segmentation algorithm for varying numbers of averaged frames for three example subjects. Subject JC_0480 represents a case of excellent agreement between the algorithm and expert-segmented FAZ regions, with negligible improvements when averaging more than two frames. For subject JC_0617, the algorithm underestimated the FAZ region by excluding a small protrusion. For this subject, averaging additional frames did not improve the segmentation accuracy. Subject JC_10586 represents a case where averaging multiple frames improved the accuracy of the algorithm-segmented FAZ region, with the average of five frames required for accurate segmentation.

Figure 3 plots the Dice coefficient between the algorithm-segmented FAZ region and the ground-truth FAZ region for the case where 10 frames were input into the algorithm and the number of averaged

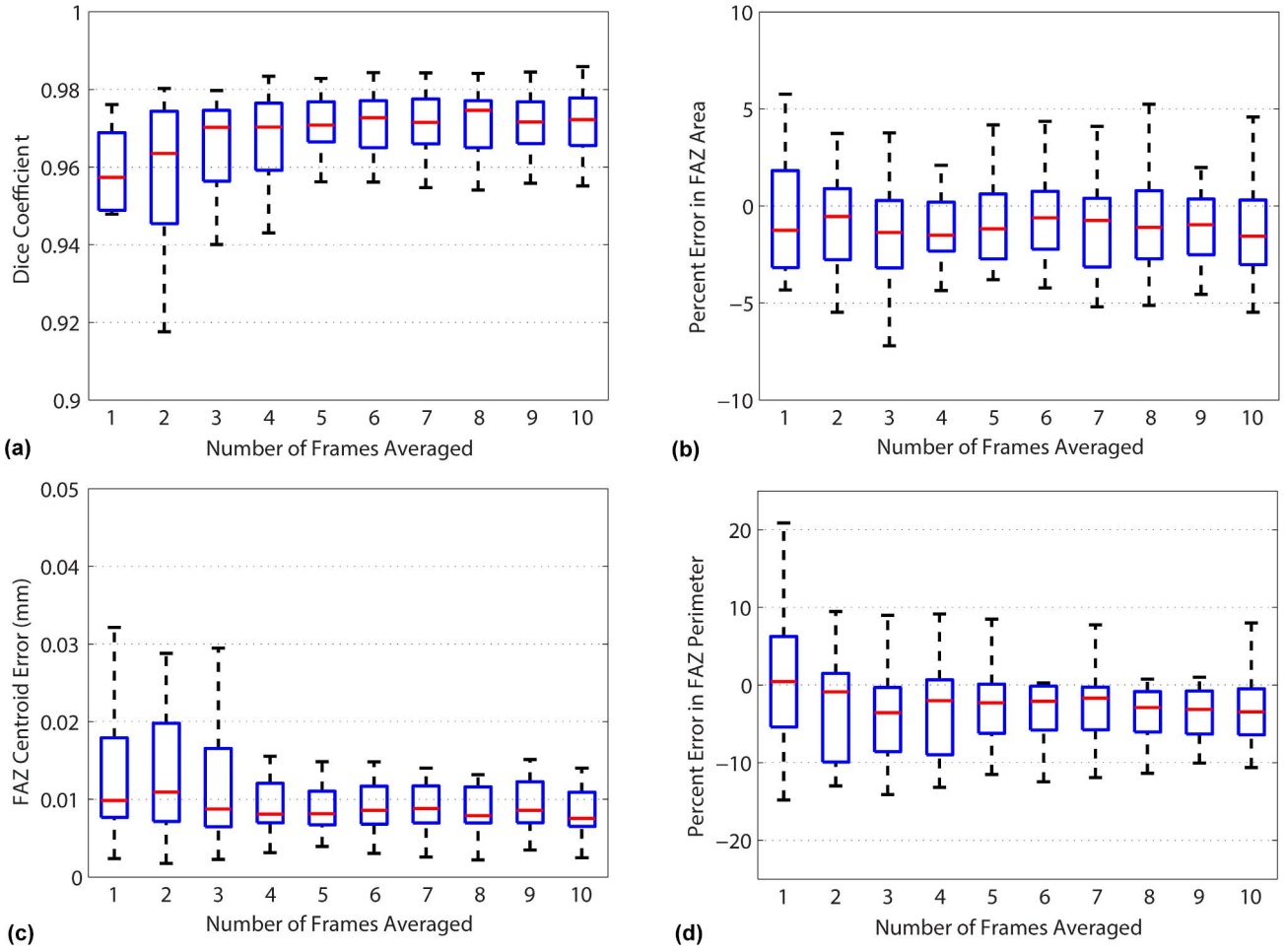

Figure 3. FAZ segmentation accuracy metrics of (a) Dice coefficient between algorithm and ground-truth FAZ segmentations, (b) percent error in FAZ area, (c) error in FAZ centroid location, and (d) percent error in FAZ area for different numbers of averaged frames. Ten OCTA frames were input to the algorithm for all cases. These results include the subject with ambiguous FAZ displayed in Figure 10.

frames varied between one and 10. A Dice coefficient of one signifies perfect agreement between the algorithm and expert FAZ regions, while a Dice coefficient of zero signifies no overlap between the two segmented regions. The median Dice coefficient increased with the number of averaged frames. Figure 3 also plots the distribution across all subjects of the percent error in FAZ area, error in FAZ centroid, and percent error in FAZ perimeter for the algorithm-segmented FAZ region relative to the ground-truth FAZ region. Increasing the number of averaged frames reduced the range of FAZ centroid error but did not systematically affect the error in FAZ area or perimeter, with the interquartile range between −3% to 2% for the area error and −10% to 7% for the perimeter error.

One subject presented with an ambiguous FAZ region, as detailed in the Appendix. The data from this subject are included in Figure 3 but were excluded from further analysis of the effects of frame averaging on FAZ metrics. Table 1 presents the FAZ segmentation accuracy metrics (Dice coefficient, absolute and percent error in FAZ area, absolute and percent error in FAZ perimeter, and FAZ centroid error). Statistical significance of the effect of averaging five frames compared with one frame was tested with the Wilcoxon signed rank test for each metric, as was the significance of the effect of averaging 10 frames compared with one frame. The Dice coefficients when averaging five (0.964) and 10 frames (0.964) were each significantly higher than when using one frame (0.948), although this improvement was small in magnitude. Averaging multiple frames did not significantly affect the other segmentation accuracy metrics in this study, for which 10 frames were input to the algorithm for all subjects.

Table 1. FAZ Segmentation Accuracy Metrics Compared for Segmenting 1 Frame, the Average of 5 Frames, and the Average of 10 Frames, Where 10 Frames Were Input to the Algorithm

FAZ Accuracy Metric	1 Frame With Highest Image Quality	Average of 5 Frames With Highest Image Quality	Average of 10 Frames With Highest Image Quality
Dice coefficient			
Mean ± standard deviation	0.948 ± 0.039	0.964 ± 0.036	0.964 ± 0.036
Range	0.824–0.976	0.823–0.983	0.825–0.986
P value (compared with 1 frame)	n/a	0.005	0.0004
Absolute error in FAZ area, mm^2			
Mean ± standard deviation	0.0091 ± 0.0089	0.0057 ± 0.0036	0.0064 ± 0.0042
Range	0.0008–0.0382	0.0007–0.0134	0.0004–0.0147
P value (compared with 1 frame)	n/a	0.170	0.286
Absolute percent error in FAZ area			
Mean ± standard deviation	4.42 ± 5.41	2.92 ± 4.12	3.54 ± 5.58
Range	0.30–21.15	0.54–18.74	0.18–25.05
P value (compared with 1 frame)	n/a	0.145	0.446
Absolute error in FAZ perimeter, mm			
Mean ± standard deviation	0.1452 ± 0.1105	0.1198 ± 0.1134	0.1226 ± 0.1148
Range	0.0022–0.3850	0.0101–0.3631	0.0032–0.3745
P value (compared with one frame)	n/a	0.372	0.249
Absolute percent error in FAZ perimeter			
Mean ± standard deviation	7.44 ± 6.97	5.98 ± 6.70	6.24 ± 7.22
Range	0.09–25.44	0.44–26.28	0.16–29.40
P value (compared with one frame)	n/a	0.679	0.372
Error in FAZ centroid, mm			
Mean ± standard deviation	0.0132 ± 0.0109	0.0095 ± 0.0054	0.0092 ± 0.0053
Range	0.0024–0.0447	0.0039–0.0286	0.0025–0.0271
P value (compared with one frame)	n/a	0.349	0.170

The outlier case (JC_11068) of ambiguous FAZ shown in Figure 10 was not included in this analysis. Significance in difference in error between acquiring one frame and acquiring 5 or 10 frames was tested with the Wilcoxon Signed-Rank Test.

Effect of the Number of Acquired Frames on FAZ Segmentation

This section presents the results of the algorithm when the number of frames input to the algorithm was varied between one and 10, with the frames input in the same sequence as they were acquired. This study demonstrates the effects of acquiring multiple frames and selecting the best frame, compared with acquiring multiple frames and averaging the *N* highest quality multiple frames. Figure 4a plots the mean absolute error in FAZ area obtained with the automated FAZ segmentation algorithm as the number of acquired frames was varied from one to 10. Figure 4b similarly plots the mean absolute percent error in FAZ perimeter. In both plots, the blue data result from segmenting the FAZ in the one frame identified with highest image quality from the acquired frames. The yellow data result from segmenting the FAZ in the average of up to five frames with highest image quality (i.e., average all frames if the number of frames is $\leq$5, otherwise average 5 frames). The number of averaged frames was limited to five, as the results in the previous section suggest minimal benefit when averaging more than five frames. Analysis of Dice coefficient and centroid location was not possible for this portion of the study, as the frame to which the ground-truth FAZ segmentation was registered was not necessarily input to the algorithm.

Figure 4 demonstrates that the error in FAZ area and perimeter was reduced when multiple frames were acquired and the FAZ was segmented using the frame with highest image quality (without frame averaging). For example, the mean absolute error in FAZ area

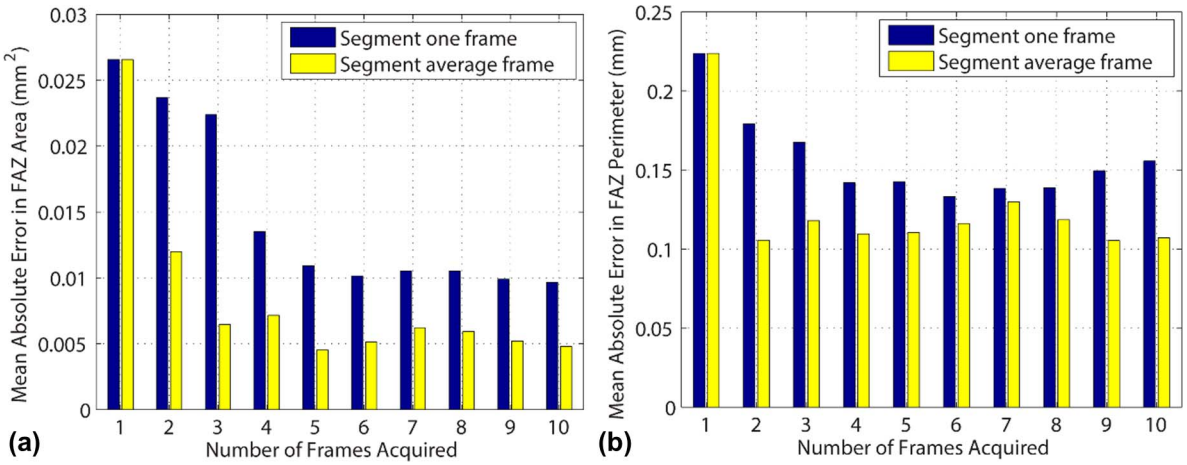

Figure 4. The mean absolute percent error in FAZ (a) area and (b) perimeter plotted with the number of acquired frames varied between one and 10. The *blue* data result from segmenting the FAZ in the one frame identified with highest-image quality from the acquired frames. The *yellow* data result from segmenting the FAZ in the average of up to five frames with highest image quality. The subject with ambiguous FAZ, shown in Figure 10, was not included in this analysis.

decreased from 0.026 mm^2 (1 acquired frame) to 0.024 (2 acquired frames) to 0.010 (5 acquired frames), without frame averaging. However, the reduction in error in FAZ area due to acquiring multiple frames (without averaging) was statistically significant only when nine or more frames were acquired ($P < 0.05$, Wilcoxon signed rank test), while the reduction in error in FAZ perimeter was not significant for any number of acquired frames ($P > 0.1$).

Frame averaging further reduced the error in FAZ area and perimeter, as seen in Figure 4. When five frames were acquired, the mean absolute error in FAZ area decreased from 0.009 to 0.006 mm^2 when the five frames were averaged prior to segmentation compared to using the one frame with highest image quality. The results in Figure 4 suggest minimal benefit in acquiring more than five frames. The reduction in FAZ error when acquiring and averaging multiple frames compared with acquiring one frame was statistically significant when the number of acquired frames was greater than two ($P < 0.047$). Table 2 presents additional evaluation of the FAZ segmentation accuracy metrics, comparing the acquisition of 1, 5, and 10 frames. In Table 1, the results for one frame correspond to segmenting the FAZ on the frame with highest image quality of the 10 acquired frames. In Table 2, the results for one frame correspond to segmenting the FAZ on the first acquired frame, which may not be of high quality. For all error metrics, acquiring five and 10 frames and averaging five frames resulted in a significantly more accurate FAZ region than acquiring one frame ($P <$ 0.025, Wilcoxon signed-rank test). There was no significant difference in FAZ segmentation error obtained by acquiring five and 10 frames.

Figure 5 displays an example comparing the FAZ region segmented by the algorithm from one frame (the first acquired frame) and the FAZ region segmented by the algorithm from the average of the first five acquired frames. The FAZ region is over-segmented when using a single frame. Averaging five frames strengthened the weak vessel edges, resulting in more accurate FAZ segmentation.

Effect of Multiple Frame Acquisition and Averaging on Parafoveal Intercapillary Area

Figure 6a displays the normalized average PICA for each subject for the study of 10 acquired frames with the number of averaged frames varied between one and 10. The area was normalized relative to the value obtained from one frame in order to better visualize the change in this metric with the number of averaged frames. Figure 6b displays the normalized average PICA when the number of acquired frames, M, was varied between one and 10. For this second study, the number of averaged frames was equal to the lesser of number of acquired frames or five (i.e., a maximum of 5 frames were averaged). The results suggest that the average PICA was unaffected by frame averaging.

Figure 7 plots the CV of the average PICA for different numbers of averaged frames. The range of CV values across the subjects decreased when five or more frames were averaged, regardless of the number

Table 2. FAZ Segmentation Accuracy Metrics Compared for Acquiring 1 Frame, Acquiring and Averaging 5 Frames, and Acquiring 10 Frames While Averaging 5 Frames

FAZ Accuracy Metric	1 Frame Acquired	5 Frames Acquired (5 Averaged)	10 Frames Acquired (5 Averaged)
Absolute error in FAZ area, mm^2			
Mean ± standard deviation	0.0257 ± 0.0339	0.0045 ± 0.0051	0.0048 ± 0.0044
Range	0.0024–0.1436	0.0002–0.0205	0.0003–0.0170
P value (compared with 1 frame)	n/a	0.004	0.003
Percent error in FAZ area			
Mean ± standard deviation	9.90 ± 12.04	1.96 ± 2.01	2.36 ± 3.10
Range	0.74–50.97	0.12–6.39	0.07–13.17
P value (compared with 1 frame)	n/a	0.018	0.020
Absolute error in FAZ perimeter, mm			
Mean ± standard deviation	0.2237 ± 0.2010	0.1106 ± 0.0978	0.1071 ± 0.0950
Range	0.1143–0.7082	0.0040–0.3287	0.0045–0.3006
P value (compared with 1 frame)	n/a	0.014	0.025
Percent error in FAZ perimeter			
Mean ± standard deviation	10.73 ± 8.92	5.67 ± 6.64	5.14 ± 4.77
Range	0.09–28.20	0.26–27.91	0.20–14.92
P value (compared with 1 frame)	n/a	0.018	0.010

The outlier case of ambiguous FAZ shown in Figure 4 (JC_11068) was not included in this analysis. The FAZ area and perimeter were significantly reduced when acquiring 5 and 10 frames. Significance in difference in error between acquiring 1 frame and acquiring 5 or 10 frames was tested with the Wilcoxon signed-rank test.

of acquired frames. This reduction in CV suggests that frame averaging improves the precision of the average PICA estimates.

Effect of Multiple Frame Acquisition and Averaging on FIJI AnalyzeSkeleton Metrics

Figure 8 plots the skeletonized vessel metrics for varying numbers of acquired and averaged frames. To better visualize the effect of frame averaging, all metrics are displayed normalized to value calculated from one frame. Increasing the number of averaged frames exponentially decreased the number of endpoints. For the study with 10 acquired frames, averaging five frames reduced the estimated number of endpoints by 32%, averaged across all subjects, relative to the value estimated from one frame. Averaging 10 frames reduced the number of endpoints by an average of 37%, relative to the value estimated at one frame. Similar results were obtained when the number of acquired frames was varied, with 32% reduction in the number of endpoints when acquiring and averaging five frames relative to one frame, and 36% reduction when acquiring 10 frames and averaging five frames. Averaging five frames estimated the number of endpoints to within 8% and 5% of the asymptote of the fitted exponential curve for the case of acquiring five and 10 frames, respectively.

As seen in Figures 8c and 8d, frame averaging increased the estimated average vessel length. The

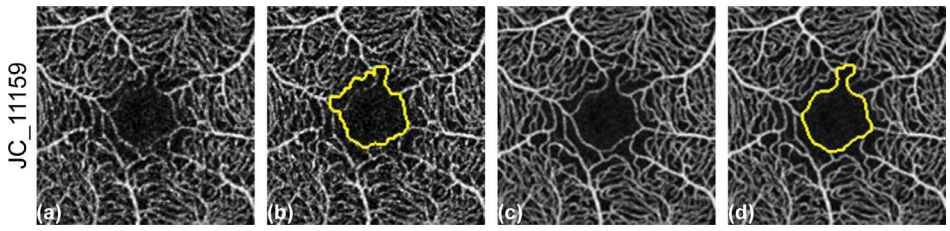

Figure 5. (a) OCTA image resulting from acquiring one frame, with algorithm-segmented FAZ shown in (b). (c) OCTA image resulting from acquiring and averaging five frames. (d) The algorithm-segmented FAZ for the average of five frames.

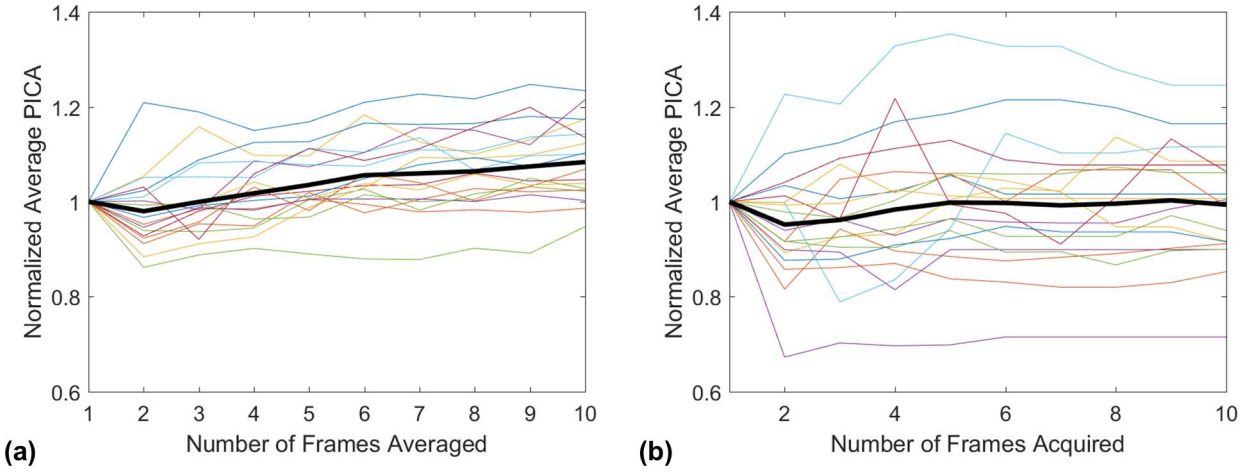

Figure 6. (a) The normalized average PICA calculated for different numbers of averaged frames of 10 total acquired frames, where the values are normalized relative to the value obtained from one frame. (b) The normalized average PICA calculated for different numbers of acquired frames, with the number of averaged frames equal to the lesser of the number of acquired frames or five. In plots (a) and (b) each *colored line* represents one subject, with the average across subjects represented by the *thick black line*.

average vessel length increased by an average of 14% when averaging five frames compared with one frame and increased by an average of 17% when averaging 10 frames. Similar trends were seen when the number of acquired frames was varied between one and 10. Averaging five frames estimated the number of endpoints to within 4% and 3% of the asymptote of the fitted exponential curve for the case of acquiring five and 10 frames, respectively. As seen in Figures 8e and 8f, the longest vessel length varied as the number of averaged frames was varied but did not increase or decrease as the number of averaged frames increased.

Discussion

The proposed workflow of acquiring multiple OCTA images, ranking the images by a gradient magnitude-based image quality metric, and averaging multiple frames improved the accuracy of the automated FAZ segmentation algorithm and affected the vessel morphology and perfusion metrics. This study builds upon an existing body of work that supports the benefits of OCTA frame averaging. For example, frame averaging was shown in previous studies to improve subjective image quality in the parafovea,[26] increase signal-to-noise ratio in the peripapillary

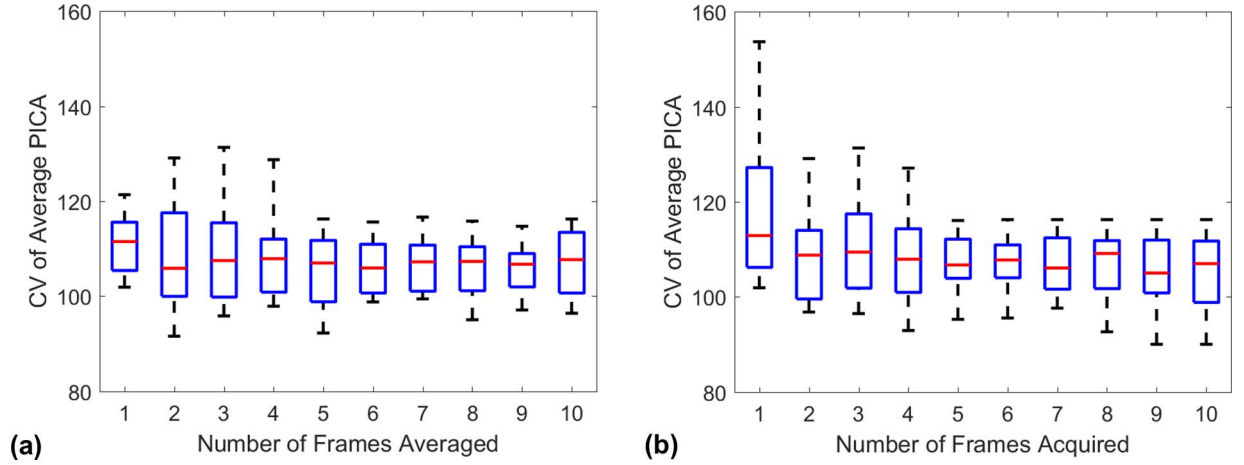

Figure 7. (a) The CV of the average PICA for different numbers of averaged frames of 10 total acquired frames. (b) The CV of the average PICA for different numbers of acquired frames with the number of averaged frames equal to the lesser of the number of acquired frames or five. The boxplots in (a) and (b) represent the distribution across the 19 subjects.

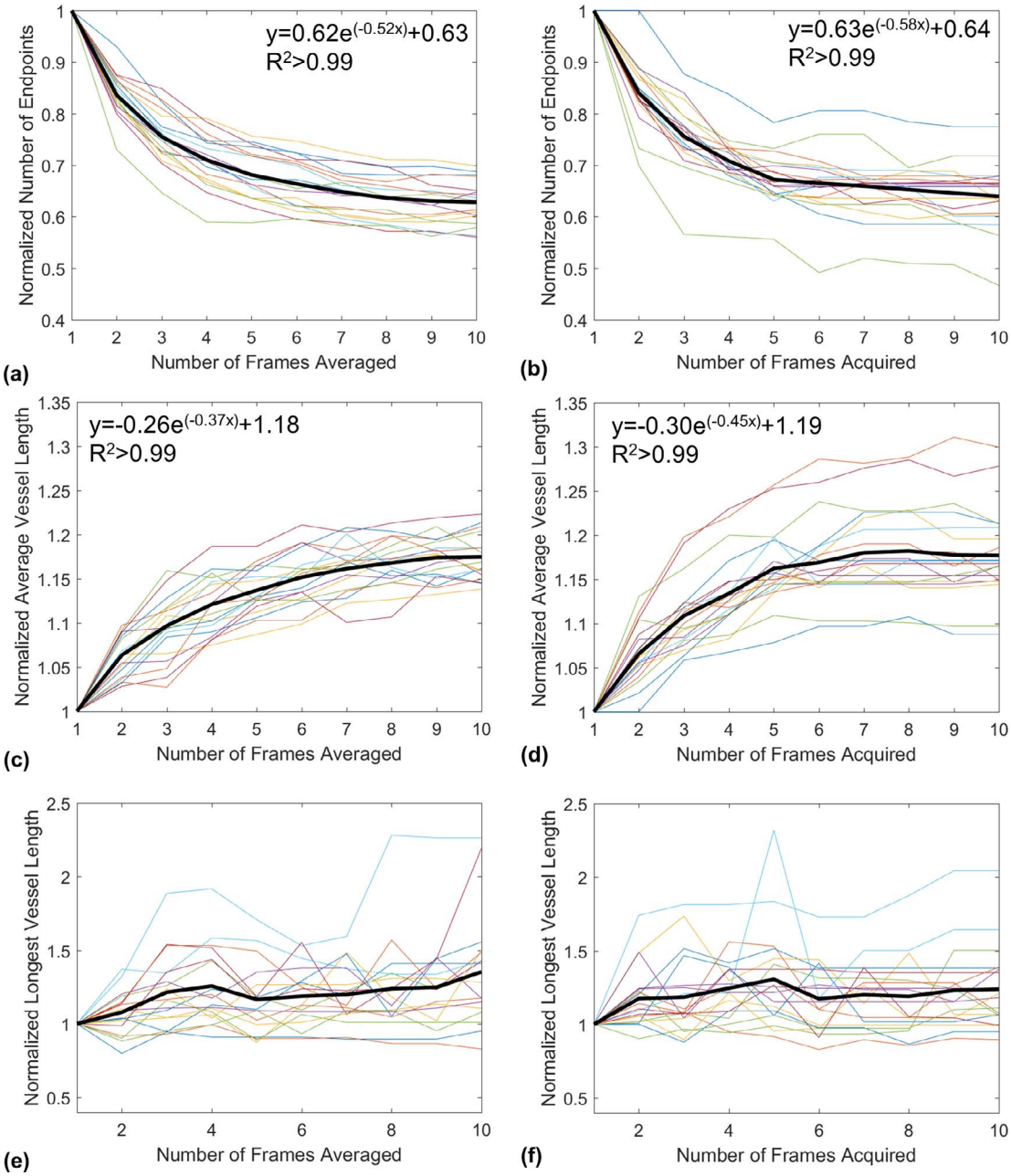

Figure 8. The *left column* (a, c, e), displays the effects of frame averaging on the calculated skeleton metrics for the study that acquired 10 frames and averaged a varying number of frames. The right column (b, d, f) displays results of the study in which the number of acquired frames was varied between one and 10 and the number of averaged frames was equal to the lesser of the number of acquired frames or five. In all plots, each *colored line* represents the results from one subject. The *thick black line* represents the mean across all subjects. Graphs (a, b) plot the number of endpoints, normalized by the number calculated from one frame. Graphs (c, d) plot the normalized average vessel length, and (e, f) plot the longest vessel length, normalized by the value calculated from one frame. Data in plots (a–d) were fit to exponential functions using cftool in MATLAB.

region,[24] and improve the visualization of the choriocapillaries.[25] Frame averaging was found to affect metrics of vessel density, vessel length density, vessel diameter index, and fractal dimension in the parafoveal region[26] as well as metrics of vessel length density, intercapillary distance, mean segment length, and number of endpoints in the peripapillary region.[24] This current study adds to the body of previous work by quantifying the effects of frame averaging on FAZ segmentation, FAZ metrics as well as estimation of the

number of endpoints, average vessel length, longest vessel length, and average PICA.

It is important to note that acquiring multiple OCTA volumes and averaging multiple en face images, as was performed in this study, provides additional benefit for the overall en face image compared with acquiring and averaging multiple B-scans at each retinal location. Acquiring additional B-scans improves the estimation of the flow signal by OCTA algorithms at one retinal location. By acquiring multiple OCTA volumes, each volume provides information about more retinal locations than simply the number of B-scans in one volume, due to eye motion resulting in slightly different retinal locations being sampled across the different volumes. Thus, each en face image contains slightly different flow information at slightly different locations along the vessel. The vessel definition and continuity are therefore improved when averaging multiple en face images.

Whereas previous studies registered the acquired OCTA frames to a subjectively selected reference frame,[24–26] this study proposes the novel use of an image quality metric to rank the acquired frames prior to registration and averaging. A method to automatically select the reference frame will be important for translation into clinical practice. In this study, the mean gradient magnitude of the OCTA frame was used as the metric of image quality for ranking the acquired frames. This metric is designed to provide high scores to frames with high-contrast and sharp edges. However, the mean gradient magnitude may be sensitive to frames with high noise. Also, while the mean gradient magnitude metric is appropriate for ranking image quality of frames from the same subject, the metric is limited in quantifying absolute image quality across subjects, due to changes in vessel density. An absolute metric of OCTA image quality could be used in the future to only select high-quality frames for averaging and FAZ segmentation. Another limitation of the gradient magnitude metric is that it decreases with the number of averaged frames because averaging introduces smoothing and decreases the gradient. Therefore, the gradient magnitude metric cannot be used to determine when to stop including additional frames for averaging. Overall, development of improved OCTA image quality metrics to guide frame averaging would be an area of interesting future study.

In general, the results suggest limited benefit in acquiring and averaging more than five frames, which was also demonstrated in a previous study of frame averaging in the superficial layer of the parafovea.[26] A previous study of OCTA frame averaging in the peripapillary region determined that the optimal number of averaged frames varied with quantitative metric and region of interest location relative to the optic disk.[24] For example, metrics of skeletonized vessel morphology required five to seven frames depending on location, which is similar to the results determined in this current study. Both the current study and the previous study demonstrated similar trends of an exponential decrease in the estimated number of endpoints and exponential increase in the average vessel length with an increasing number of averaged frames.[24] These results demonstrate that frame averaging increases the integrity of the vessel segments, which thus reduces endpoints and increases mean vessel length.

Acquiring and averaging multiple OCTA frames did not affect the average PICA estimated for each patient but did improve the precision of the estimate when more than five frames were averaged. The results of the skeleton analysis demonstrate that frame averaging increases the number of intact vessels. While the intact vessels may alter the area of parafoveal intercapillary subregions, the average PICA may be unaffected. The reduction in noise due to frame averaging may also cause a slight increase in the PICA and may cause the improved precision in the PICA estimates.

Acquiring multiple frames may pose challenges in practice due to the added acquisition time. The presented results enable evaluating the difference between OCTA metrics obtained with varying numbers of frames to determine the correct operating point for specific applications. Understanding which metrics are most affected by frame averaging, such as FAZ area (Fig. 4), number of endpoints (Fig. 8), and average vessel length (Fig. 8), may help point clinicians toward metrics that may be more robust and of greater utility when averaging is not possible. If the goal is to use these metrics to detect small changes over time, the results suggest using averaged images where possible to increase the sensitivity of the longitudinal comparisons. Sometimes multiple OCTA volumes are acquired in practice while attempting to obtain a single volume with minimal motion artifacts. The results of this and previous OCTA averaging studies[24–26] suggest that it may be beneficial to average the portions of the volumes devoid of visible motion artifacts rather than discarding those that are not "perfect" in the eyes

of the operator. Even if multiple frames are acquired, new workflow and software capabilities to rank, select, and average frames are needed within commercial OCTA systems to achieve the benefits of frame averaging.

One limitation of this study is the relatively small number of subjects, though the observed effects were fairly robust. The study was further limited to subjects without known retinal or vascular disease, and the effect of acquiring and averaging multiple OCTA frames in subjects with pathology requires future study. A further limitation of this study is the relatively narrow age range of the subjects (23–49), as vascular biomarkers may change with age.[36,37] The FAZ results in this study are also limited to the specific implemented FAZ segmentation algorithm developed in this work and a single OCTA device (Optovue). Frame averaging is expected to improve FAZ segmentation accuracy for other segmentation approaches; however, future studies are required to investigate frame averaging combined with commercial FAZ segmentation algorithms and deep learning segmentation algorithms that have recently been developed.[38,39] Likewise, there are differences in image quality across commercial OCTA devices,[40–42] due in part to different algorithms (e.g., OMAG, SSADA) and different scan parameters (e.g., scan density, motion correction). The benefit of frame averaging may be greater or lesser when using images obtained from other devices.

Conclusions

Acquiring and averaging multiple OCTA frames reduced the error in automated FAZ segmentation and improved the robustness of OCTA skeleton and perfusion metrics. The results demonstrate significant reduction in the error of the segmented FAZ region when acquiring and averaging more than two frames. The study determined limited benefit in acquiring and averaging more than five frames.

Acknowledgments

The authors thanks Optovue for providing the AngioVue OCTA system. They also thank Erin Curran and Phyllis Summerfelt for their assistance with this study.

Supported in part by the National Eye Institute of the National Institutes of Health (NIH) under award numbers R01EY024969, R01EY027301, and P30EY001931 and by the National Institute of Aging of the NIH under award number T35AG029793. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Rachel E. Linderman is the recipient of a Fight for Sight–Nick Cacciola Summer Student Fellowship Award. Additional support from a New York Eye and Ear Infirmary Foundation Grant and the Marrus Family Foundation. The effort of Taly Gilat Schmidt was supported by the Marquette University Way Klingler Sabbatical fellowship.

Disclosure: **T.G. Schmidt**, None; **R.E. Linderman**, Optovue (C); **M.R. Strampe**, None; **T.Y.P Chui**, None; **R.B. Rosen**, Optovue (C), Astellas (C), Boehringer-Ingerheim (C), NanoRetina (C), OD-OS (C), Opticology (I), Regeneron (C), Diopsys (C), Guardion Health (C), Bayer (C), Genentech-Roche (C); **J. Carroll**, Optovue (F)

*Taly Gilat Schmidt and Rachel E. Linderman contributed equally to this work.

Appendix: Automated FAZ Segmentation Algorithm

Methods

While automated and semiautomated algorithms for segmenting the FAZ are available on commercial OCTA systems, these algorithms are not generally accessible for offline research purposes. Therefore, we developed an automated FAZ segmentation algorithm for use in this study (MATLAB v2017a). The developed automated FAZ segmentation algorithm, illustrated in Figure 9, is based on finding the largest nonvessel region in the OCTA image. The algorithm first identifies the vessel edge regions. The Canny edge detection method[43] was investigated for this purpose but was found to be insufficiently sensitive to weak edges. Therefore, a modified edge detection algorithm was developed that is based on the Canny edge detector but excludes the initial smoothing operation and nonmaxima suppression steps, which were found to remove weaker edges. The averaged OCTA image was first filtered with the Sobel filter to enhance the contrast of vessel edges. The Sobel-filtered image was then thresholded using a hysteresis thresholding method, with two thresholds values selected for creating a mask of strong edges (pixels $> T_{strong}$) and a mask of weak edges (pixels $> T_{weak}$), where

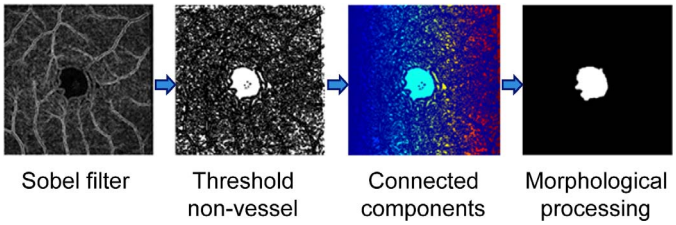

Figure 9. Overview of automated FAZ segmentation algorithm. In the first step, the OCTA image is processed with a Sobel filter to enhance edges. Thresholding is performed to identify nonvessel regions, which are the labeled with a connected components algorithm, where each component is labeled with an integer value that is displayed as a unique color in the figure. The largest central component is selected as the FAZ. In the final step of morphologic processing, the FAZ region is closed to fill small gaps and dilated to extend to the centerlines of the bordering vessels.

$T_{strong} > T_{weak}$. The thresholds were calculated adaptively as proportional to the most common intensity in the Sobel-filtered image (i.e., the intensity corresponding to the peak of the histogram of the Sobel-filtered image). A sensitivity study found that a low threshold setting between 0.5 and 0.6 of the histogram peak and a high thresholding setting between 0.8 and 1.3 of the histogram peak returned similar segmentation results (Dice coefficient > 0.95). Thresholds settings of 0.5 (T_{weak}) and 1.0 (T_{strong}) of the histogram peak were selected for all further analyses. The final vessel edge mask included all identified strong edges and all identified weak edges that were connected to a strong edge. The complement of the edge mask is a mask of nonvessel-edge regions. The connected components of this nonvessel mask were identified assuming four-neighbor connectivity. The FAZ region was selected as the component with the largest area and whose centroid was within 25 pixels of the image center, which assumes that the FAZ is generally near the center of the acquired image frame. The following morphologic processing was performed to refine the identified FAZ region. The image was closed with a 3×3-pixel cross-structuring element to fill small gaps and then dilated by a 5×5-pixel square structuring element so that the region extended to the centerlines of the bordering vessels, followed by a hole filling step.

Results and Discussion

The FAZ segmentation algorithm correctly identified the FAZ region in 18 of 19 subjects. Figure 10 displays the results of the subject for which the algorithm failed to identify the ground-truth FAZ region. The expert readers selected the most central avascular region for this subject, while the proposed algorithm, as part of the morphologic processing criteria, selected the largest avascular region near the center of the image. Because the FAZ for this subject was ambiguous, an additional segmentation was performed with a commercial algorithm (Optovue AngioAnalytics). The commercial FAZ segmentation selected the union of the FAZ regions selected by the readers and the proposed algorithm.

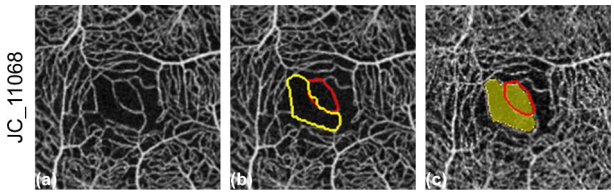

Figure 10. Subject for which the readers, proposed algorithm, and commercial FAZ segmentation algorithm identified different FAZ regions. (a) Reference image that is the average of five OCTA frames. (b) Reference image with ground-truth FAZ as the *red contour* and the algorithm-generated FAZ as the *yellow contour*. (c) Single OCTA frame with *shaded region* representing the FAZ as segmented by the Optovue AngioAnalytics software and *red contour* representing the reader-generated ground-truth FAZ.

A previous study of a different FAZ segmentation algorithm operating on a single frame estimated a FAZ area error of 0.0240 ± 0.0259 mm^2, percent error in FAZ area ranging from 2.2% to 54.4%, and an absolute maximum error of 0.145 mm^2.[32] The performance of the algorithm proposed in this paper for the case of one acquired frame is similar to the previously studied algorithm (mean absolute error: 0.0257 ± 0.0339 mm^2, error range: 0.74%–50.97%, maximum error: 0.1436 mm^2), suggesting that the proposed algorithm has comparable segmentation performance.

As seen in Figure 10, the error in FAZ identification for one subject was caused by the criteria used to select the FAZ from the candidate avascular regions, rather than an error in segmenting avascular regions. In the future, the proposed algorithm could use alternate criteria and additional OCT features to identify the FAZ region among the possible central avascular regions.

References

1. Wangsa-Wirawan ND, Linsenmeier RA. Retinal oxygen: fundamental and clinical aspects. *Arch Ophthalmol*. 2003;121:547–557.
2. Gupta P, Ting DSW, Thakku SG, et al. Detailed characterization of choroidal morphologic and

vascular features in age-related macular degeneration and polypoidal choroidal vasculopathy. *Retina*. 2017;37:2269–2280.
3. Lindner M, Fang PP, Steinberg JS, et al. OCT angiography-based detection and quantification of the neovascular network in exudative AMD. *Invest Ophthalmol Vis Sci*. 2016;57:6342–6348.
4. Sridhar J, Shahlaee A, Rahimy E, et al. Optical coherence tomography angiography and en face optical coherence tomography features of paracentral acute middle maculopathy. *Am J Ophthalmol*. 2015;160:1259–1268.e2.
5. Popovic N, Radunovic M, Badnjar J, Popovic T. Fractal dimension and lacunarity analysis of retinal microvascular morphology in hypertension and diabetes. *Microvasc Res*. 2018;118:36–43.
6. Lee J, Rosen R. Optical coherence tomography angiography in diabetes. *Curr Diab Rep*. 2016;16:123.
7. Jung JJ, Chen MH, Frambach CR, Rofagha S, Lee SS. Spectral domain versus swept source optical coherence tomography angiography of the retinal capillary plexuses in sickle cell maculopathy. *Retin Cases Brief Rep*. 2018;12:87–92.
8. Brasileiro F, Martins TT, Campos SB, et al. Macular and peripapillary spectral domain optical coherence tomography changes in sickle cell retinopathy. *Retina*. 2015;35:257–263.
9. Scott AW. Ophthalmic Manifestations of Sickle Cell Disease. *South Med J*. 2016;109:542–548.
10. Calderon GD, Juarez OH, Hernandez GE, Punzo SM, De la Cruz ZD. Oxidative stress and diabetic retinopathy: development and treatment. *Eye (Lond)*. 2017;31:1122–1130.
11. Heng LZ, Comyn O, Peto T, et al. Diabetic retinopathy: pathogenesis, clinical grading, management and future developments. *Diabet Med*. 2013;30:640–650.
12. Burton DSM, Carrim ZI. Angiographic evidence of peripheral ischemia in diabetic retinopathy and the risk of impending neovascularisation. *Can J Diabetes*. 2015;39:14–15.
13. Yokota R, Inoue M, Itoh Y, Rii T, Hirota K, Hirakata A. Comparison of microinsicion vitrectomy and conventional 20-gauge vitrectomy for severe proliferative diabetic retinopathy. *Jpn J Ophthalmol*. 2015;59:288–294.
14. Abu El-Asrar AM, Mohammad G, Nawaz MI, et al. Relationship between vitreous levels of matrix metalloproteinases and vascular endothelial growth factor in proliferative diabetic retinopathy. *PLoS One*. 2013;8:e85857.
15. Vinekar A, Jayadev C, Mangalesh S, et al. Comparing the outcome of single versus multiple session laswer photoablation of flat neovascularization in zone 1 aggressive posterior retinopathy of prematurity: a prospective randomized study. *Retina*. 2015;35:2130–2136.
16. Vavvas DG. Anti-VEGF in retinopathy of prematurity, need to titrate. *Invest Ophthalmol Vis Sci*. 2013;54:2027.
17. Joyal J-S, Omri S, Sitaras N, Rivera J-C, Sapieha P, Chemtob S. Neovascularization in retinopathy of prematurity: opposing actions of neuronal factors GPR91 and semaphorins 3A. *Acta Paediatr*. 2012;101:819–826.
18. Novotny HR, Alvis DL. A method of photographing fluorescence in circulating blood in the human retina. *Circulation*. 1961;24:82–86.
19. Lopez-Saez MP, Ordoqui E, Tornero P, et al. Fluorescein-induced allergic reaction. *Ann Allergy Asthma Immunol*. 1998;81:428–430.
20. Ha SO, Kim DY, Sohn CH, Lim KS. Anaphylaxis caused by intravenous fluorescein: clinical characteristics and review of literature. *Intern Emerg Med*. 2014;9(3):325–330.
21. Mendis KR, Balaratnasingam C, Yu P, et al. Correlation of histologic and clinical images to determine the diagnostic value of fluorescein angiography for studying retinal capillary detail. *Invest Ophthalmol Vis Sci*. 2010;51:5864–5869.
22. Tsai M-T, Lee C-K, Lin K-M, et al. Quantitative observation of focused-ultrasound-induced vascular leakage and deformation via fluorescein angiography and optical coherence tomography. *J Biomed Opt*. 2013;18:101307.
23. Lee CS, Lee AY, Sim DA, et al. Reevaluating the definition of intraretinal microvascular abnormalities and neovascularization elsewhere in diabetic retinopathy using optical coherence tomography and fluorescein angiography. *Am J Ophthalmol*. 2015;159:101–10.e1.
24. Mo S, Phillips E, Krawitz BD, et al. Visualization of radial peripapillary capillaries using optical coherence tomography angiography: the effect of image averaging. *PLoS One*. 2017;12:e0169385.
25. Uji A, Balasubramanian S, Lei J, Baghdasaryan E, Al-Sheikh M, Sadda SR. Choriocapillaris imaging using multiple en face optical coherence tomography angiography image averaging. *JAMA Ophthalmol*. 2017;135:1197.
26. Uji A, Balasubramanian S, Lei J, Baghdasaryan E, Al-Sheikh M, Sadda SR. Impact of multiple en face image averaging on quantitative assessment from optical coherence tomography angiography images. *Ophthalmology*. 2017;124:944–952, https://doi.org/10.1016/j.ophtha.2017.02.006.

27. Gonzales RC, Woods RE. *Digital Image Processing*. Prentice Hall; 2002.
28. Thevenaz P, Ruttimann UE, Unser M. A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process*. 1998;7: 27–41.
29. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9:671–675.
30. Zou KH, Warfield SK, Bharatha A, et al. Statistical validation of image segmentation quality based on a spatial overlap index1. *Acad Radiol*. 2004;11:178–189.
31. Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans Med Imaging*. 2004;23: 903–921.
32. Linderman R, Salmon AE, Strampe M, Russillo M, Khan J, Carroll J. Assessing the accuracy of foveal avascular zone measurements using optical coherence tomography angiography: segmentation and scaling. *Transl Vis Sci Technol*. 2017; 6(3):16, https://doi.org/10.1167/tvst.6.3.16.
33. Krawitz BD, Phillips E, Bavier RD, et al. Parafoveal nonperfusion analysis in diabetic retinopathy using optical coherence tomography angiography. *Transl Vis Sci Technol*. 2018;7(4):4.
34. Arganda-Carreras I, Fernández-González R, Muñoz-Barrutia A, Ortiz-De-Solorzano C. 3D reconstruction of histological sections: application to mammary gland tissue. *Microsc Res Tech*. 2010;73:1019–1029.
35. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676–682.
36. Iafe NA, Phasukkijwatana N, Chen X, Sarraf D. Retinal capillary density and foveal avascular zone area are age-dependent: quantitative analysis using optical coherence tomography angiography. *Invest Ophthalmol Vis Sci*. 2016;57:5780–5787.
37. Pinhas A, Linderman R, Mo S, et al. A method for age-matched OCT angiography deviation mapping in the assessment of disease- related changes to the radial peripapillary capillaries. *PLoS One*. 2018;13:e0197062.
38. Prentašic P, Heisler M, Mammo Z, et al. Segmentation of the foveal microvasculature using deep learning networks. *J Biomed Opt*. 2016;21:075008.
39. Guo Y, Camino A, Wang J, Huang D, Hwang TS, Jia Y. MEDnet, a neural network for automated detection of avascular area in OCT angiography. *Biomed Opt Express*. 2018;9:5147.
40. Munk MR, Giannakaki-Zimmermann H, Berger L, et al. OCT-angiography: a qualitative and quantitative comparison of 4 OCT-A devices. *PLoS One*. 2017;12:e0177059.
41. Pilotto E, Frizziero L, Crepaldi A, et al. Repeatability and reproducibility of foveal avascular zone area measurement on normal eyes by different optical coherence tomography angiography instruments. *Ophthalmic Res*. 2018;59:206–211.
42. Shiihara H, Sakamoto T, Yamashita T, et al. Reproducibility and differences in area of foveal avascular zone measured by three different optical coherence tomographic angiography instruments. *Sci Rep*. 2017;7:9853.
43. Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1986;6:679–698.