tvst

Article

Automated Assessment of Photoreceptor Visibility in Adaptive Optics Split-Detection Images Using Edge Detection

Min Chen[1], Yu You Jiang[2,3], James C. Gee[1], David H. Brainard[4], and Jessica I. W. Morgan[2,3]

[1] Department of Radiology, University of Pennsylvania, Philadelphia, PA, USA
[2] Scheie Eye Institute, Department of Ophthalmology, University of Pennsylvania, Philadelphia, PA, USA
[3] Center for Advanced Retinal and Ocular Therapeutics, University of Pennsylvania, Philadelphia, PA, USA
[4] Department of Psychology, University of Pennsylvania, Philadelphia, PA, USA

Correspondence: Min Chen, Richards Medical Research Laboratories, 3700 Hamilton Walk, 6th Floor, Philadelphia, PA 19104, USA. e-mail: minchen1@upenn.edu
Jessica I. W. Morgan, 3400 Civic Center Blvd. Ophthalmology, 3rd Floor West, 3-112W, Philadelphia, PA 19104, USA. e-mail: jwmorgan@pennmedicine.upenn.edu

Received: September 28, 2021
Accepted: April 18, 2022
Published: May 24, 2022

Keywords: adaptive optics; photoreceptor visibility; image assessment

Citation: Chen M, Jiang YY, Gee JC, Brainard DH, Morgan JIW. Automated assessment of photoreceptor visibility in adaptive optics split-detection images using edge detection. Transl Vis Sci Technol. 2022;11(5):25, https://doi.org/10.1167/tvst.11.5.25

Purpose: Adaptive optics scanning laser ophthalmoscopy (AOSLO) is a high-resolution imaging modality that allows measurements of cellular-level retinal changes in living patients. In retinal diseases, the visibility of photoreceptors in AOSLO images is affected by pathology, patient motion, and optics, which can lead to variability in analyses of the photoreceptor mosaic. Current best practice for AOSLO mosaic quantification requires manual assessment of photoreceptor visibility across overlapping images, a laborious and time-consuming task.

Methods: We propose an automated measure for quantification of photoreceptor visibility in AOSLO. Our method detects salient edge features, which can represent visible photoreceptor boundaries in each image. We evaluate our measure against two human graders and two standard automated image quality assessment algorithms.

Results: We evaluate the accuracy of pairwise ordering (PO) and the correlation of ordinal rankings (ORs) of photoreceptor visibility in 29 retinal regions, taken from five subjects with choroideremia. The proposed measure had high association with manual assessments (Grader 1: PO = 0.71, OR = 0.61; Grader 2: PO = 0.67, OR = 0.62), which is comparable with intergrader reliability (PO = 0.76, OR = 0.75) and outperforms the top standard approach (PO = 0.57; OR = 0.46).

Conclusions: Our edge-based measure can automatically assess photoreceptor visibility and order overlapping images within AOSLO montages. This can significantly reduce the manual labor required to generate high-quality AOSLO montages and enables higher throughput for quantitative studies of photoreceptors.

Translational Relevance: Automated assessment of photoreceptor visibility allows us to more rapidly quantify photoreceptor morphology in the living eye. This has applications to ophthalmic medicine by allowing detailed characterization of retinal degenerations, thus yielding potential biomarkers of treatment safety and efficacy.

Introduction

Adaptive optics scanning laser ophthalmoscopy (AOSLO) is a cutting-edge, high-resolution in vivo retinal imaging technique.[1,2] By providing subcellular scale resolution, AOSLO imaging can resolve individual photoreceptor cells and provide the ability to monitor and detect cellular-level retinal changes in living patients. Multimodal AOSLO imaging, such as simultaneous non-confocal split-detection and confocal imaging, has enabled imaging of the photoreceptor inner segment[3] mosaic in addition to photoreceptor waveguided reflectance. It has proven useful for understanding retinal disease phenotypes, especially in cases where photoreceptor waveguiding is compromised. For

Copyright 2022 The Authors

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Downloaded from intl.iovs.org on 04/23/2024

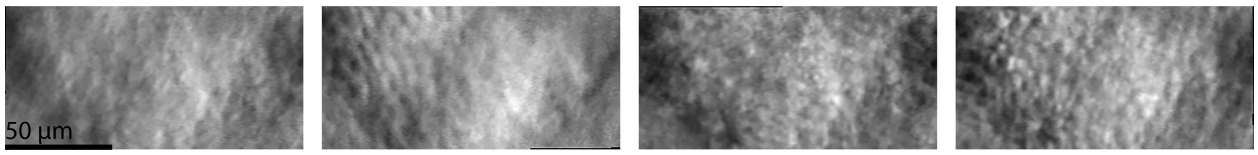

Figure 1. Examples of four split-detection AOSLO images acquired in a subject with CHM at the same retinal location and imaging session, with increasing photoreceptor visibility from *left* to *right*.

example, the photoreceptor inner segment mosaic has been shown to remain intact throughout the central island of retained retina in choroideremia (CHM), an X-linked inherited retinal degeneration, even when the outer segment waveguided reflectance in this disease is abnormal.[4–6] Similarly, AOSLO studies of achromatopsia have revealed remnant cone inner segments in split-detection images despite almost complete absence of the waveguided reflectance that mediates confocal images.[3,7] The capability to observe photoreceptor morphology in the living eye enables significant applications to ophthalmic medicine by allowing detailed characterization of retinal degenerations, which can serve as potential biomarkers of treatment safety and efficacy for experimental therapeutic interventions.[8,9]

A feature of AOSLO imaging is that its high-resolution acquisitions result in small fields of view (less than 400 × 400 μm of retina) for each individual AOSLO image. Thus, to cover a sufficiently large region of the retina for analysis, a series of AOSLO images must be acquired across the retina, with adjacent overlapping images, then montaged and analyzed together. Although several methods can largely automate the montaging process,[10–12] some steps in the post-montage analyses, such as the ordering of overlapping images for display purposes and the selection of high-quality regions of interest for quantitative analysis of the photoreceptor mosaic at a given location, still rely on time-consuming human intervention.

More specifically, one important processing step that has yet to be automated for AOSLO image analysis is to quantify differences in *photoreceptor visibility* for overlapping images within a montage. This currently requires a subjective evaluation of how well one can observe the separation between distinct photoreceptor cells in an image. Although this is not often a problem for images from healthy subjects, photoreceptor visibility may vary significantly across images of the same retinal location when analyzing AOSLO images from a patient population. The variation arises particularly because eye motion is often larger in subjects with disease and also because the adaptive optics correction may fluctuate both with eye movements and with local phenotypes of disease, such as loss of pigmentation in the retinal pigment epithelium/choroid, thinning/thickening of the inner retina, or thinning of the outer nuclear layer.[13–15] As an example, Figure 1 shows four images taken of the same location within a single retina from a patient with CHM, with photoreceptor visibility increasing from left to right. We can see from this example that, despite being the same location and imaging session, there is considerable variability in photoreceptor visibility among the images. Importantly, these images were each collected with the intention of visualizing and quantifying the photoreceptor mosaic, which demonstrates the importance of assessing photoreceptor visibility in patient images. The reliability of downstream analysis depends on determining the image with the greatest photoreceptor visibility, as images with high photoreceptor visibility will be easier to grade and quantify, but poor photoreceptor visibility can lead to less accurate and more highly variable mosaic quantifications. Current standard practice is to assign a human grader to manually assess photoreceptor visibility across each region where there are overlapping images within a montage and to bring the image with the most visible photoreceptors to the top layer of the montage for further analysis and display. This manual ranking is a time-intensive operation that, for large datasets, can take upward of 6 hours for a typical montage. The significant time required for manual assessment of photoreceptor visibility directly limits analysis throughput and the size of AOSLO studies.

In this work, we aim to automate the process of assessing the visibility of photoreceptor cells in overlapping AOSLO images. To our knowledge, no other method has yet been proposed specifically for this purpose, although there have been general methods developed for automatically assessing image quality, which can describe the general level of noise, blurriness, or artifacts in an image. These existing approaches typically rely on finding salient features or estimating the signal-to-noise ratio in the image as a surrogate for the quality of the image.[16–19] In our experience, however, these methods do not work well for assessing photoreceptor visibility in AOSLO images, because the majority of existing methods for estimating image quality were developed for the purpose of assessing images of natural scenes. Their underlying statisti-

Downloaded from intl.iovs.org on 04/23/2024

cal models assume that the objects being imaged are on significantly different scales than potential noise or blurring in the image. In contrast, the photoreceptor mosaics in split-detection AOSLO images are a repetitive pattern of small circular objects, which are generally not congruent with such models.

Here, we propose an edge-based measure designed specifically for automatically assessing photoreceptor visibility in AOSLO split-detection images. Our method is motivated by the observation that highly visible photoreceptors in AOSLO are defined by clear boundaries between the cells. These boundaries then present themselves as strong edges in the image that can be found by an edge detection algorithm. We validate our method using AOSLO images from CHM patients and compare the automated results to manual rankings of photoreceptor visibility for the same images.

Methods

Edge-Based Assessment of Photoreceptor Visibility

This work presents an automated approach for assessing the visibility of photoreceptor cells in AOSLO split-detection images. Our method uses automated edge detection as an initial step for finding salient edge features in the image that provide a proxy for visible separation between photoreceptors. For each image, we then convert the number of detected edge features in the entire image into an assessment score that can be used to quantify photoreceptor visibility in each AOSLO split-detection image and thereby rank overlapping images within a montage.

For the detection of edges, we use a Canny edge detector,[20] which finds locations in the image where there are significant intensity gradients after applying a Gaussian smoothing filter. For our method, we used the MATLAB 2020b (MathWorks, Natick, MA) implementation of the Canny detector, which uses a Gaussian kernel with standard deviation of sqrt(2) pixels. This value roughly translates to a 3 × 3-pixel patch, which we found to work well for removing noise while preserving the edges of the cone photoreceptors in our images (which have an average radius of around 6 pixels). The sensitivity of the Canny detector is controlled using an upper and a lower threshold on the gradients to determine which pixels are considered a strong edge in the image (above the upper threshold) and which are a weak edge in the image (between the two thresholds), as well as what gradient values are suppressed (below the lower threshold). The final edges detected by the algorithm consist of all strong edges and the subset of weak edges directly connected to one or more strong edge. Because the role of the edge detection in our method is to identify all salient edge features in the image, we set the lower threshold to zero to prevent any features from being suppressed. The upper threshold is then used, as described in the next paragraphs, to derive our assessment score. To account for images of different sizes, we calculate the total number of edge pixels detected in the region divided by the number of possible pixels total in the image to get a percentage value, which effectively normalizes our measure by image size.

We propose two approaches for converting the edges detected in each AOLSO image into an assessment score. The first is a static threshold (ST) approach. In this approach, the upper Canny threshold (T) is set to a constant value, and the percent of edge pixels (P%) detected in the image is used as the assessment score for the image. This idea is demonstrated in the red outlined images in Figure 2, where for images with higher photoreceptor visibility (low to high, viewing left to right) we can detect a greater percentage of edge pixels for the same threshold level.

One limitation of using a static threshold T is that the expected percentage of edge pixels may vary with retinal eccentricity. For example, images in the periphery of the retina or in regions of pathology will have more sparsely distributed photoreceptors, causing fewer edges to be detected. Other locations, such as images in the parafovea containing densely packed photoreceptors, may yield a high percentage of edge pixels to be detected. Thus, it may be difficult to set a single static threshold for the whole dataset, because setting the threshold too high will lead to a floor effect, where no edges can be detected in the sparse retinal regions. Likewise, setting the threshold too low will lead to a ceiling effect, where the number of pixels detected will be saturated in the dense retinal regions. As a result, the assessment score may not provide adequate differentiation for images at the low and high ends of photoreceptor visibility. For this reason, we propose a second approach for converting the edge detection into an assessment score. In this alternate approach, we use an adaptive threshold (AT), where the Canny threshold is chosen such that the P% of the image is filled with edge pixels. In this case, the assessment score is the maximum threshold that maintains P% edge pixels in the image. This idea is demonstrated by the blue outlined images in Figure 2, where the same percent of edge pixels can be found for different degrees of photoreceptor visibility with different values of the AT.

The two approaches each have a single static parameter that must be set prior to running the algorithm

Downloaded from intl.iovs.org on 04/23/2024

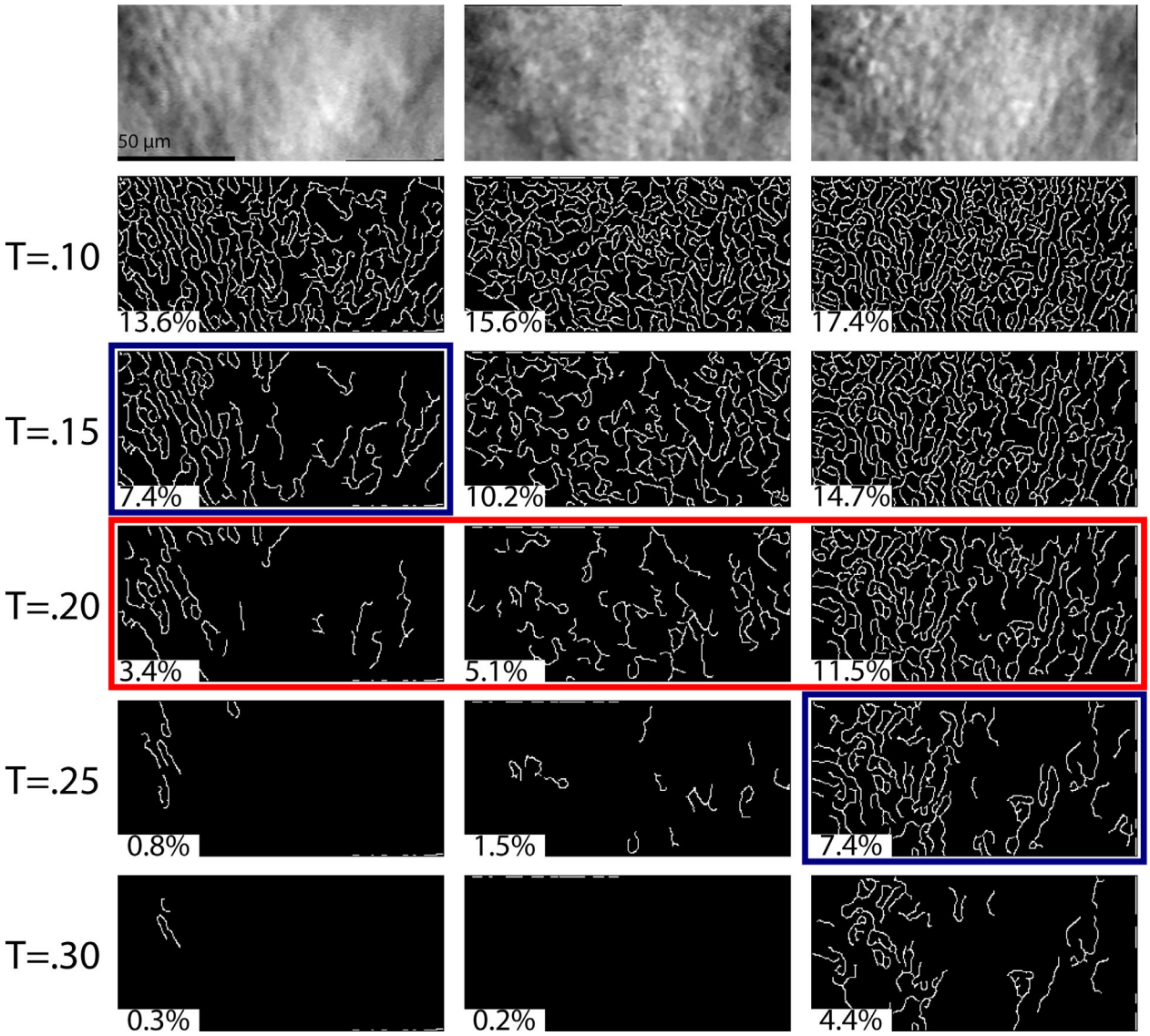

Figure 2. Different values for the Canny threshold parameter (*T*) affect the detection of edges in split-detection AOSLO images. (*Top row*) Three split-detection AOSLO images of the same retinal location from Figure 1 with different levels of photoreceptor visibility. The percent value at the bottom left of each image indicates the percent of edge pixels (*P*%) found in the image. The *red outlined box* demonstrates the ST approach, where for the same threshold *T* images with high photoreceptor visibility will have an increased percent of edge pixels detected. The two *blue outlined boxes* demonstrate the AT approach, where images with high photoreceptor visibility can use a more stringent (higher) threshold to find the same percent of edge pixels.

(*T* for ST and *P*% for AT). For ST, the goal is to set a *T* value such that numerous edge pixels are detected when there is high photoreceptor visibility and few edge pixels are detected when there is low visibility. In our analysis, we found that if *T* was set too low, edges (likely noise) were detected even when there was poor visibility. And, conversely, if the parameter was set too high, edges were not found even if the photoreceptor visibility was great. We found that an effective compromise for both high and low photoreceptor visibility images was to tune the parameter such that the algorithm detects about half of the edge pixels for an image with average photoreceptor visibility. This gives the measure room to swing higher or lower depending on how much better or worse the photoreceptor visibility is in the image. The same concept was applied to the AT and setting the static *P*% parameter, where we tuned the parameter such that it would produce a result at the center of the output range when run on an image with average visibility.

Prior to conducting our validation experiments, the values for these parameters were determined

translational vision science & technology

Downloaded from intl.iovs.org on 04/23/2024

empirically using a separate development imaging dataset that has no subject overlap with the images used in the validation. The empirically determined parameters were $T = 0.17$ when using a static threshold, and $P\% = 5\%$ when using an adaptive threshold. These values were set in our pre-registration prior to running our validation experiments and remained unchanged post hoc.

Data

AOSLO datasets from five CHM subjects ages 14 to 38 years were included for validation of our method for assessing photoreceptor visibility in AOSLO images. All subjects gave informed consent prior to being enrolled, the institutional review board at the University of Pennsylvania approved the study, and the study followed the tenets of the Declaration of Helsinki. Subjects were imaged using a custom-built multimodal AOSLO equipped with both confocal and split-detection imaging modalities. This system has been described previously.[21] Briefly, wavefront sensing was done using an 848 Δ26 nm superluminescent diode (Superlum, Cork, Ireland), and aberration compensation was completed using a 97-actuator deformable mirror. Confocal and non-confocal split-detection reflectance imaging was obtained using a 795 Δ15.3 nm superluminescent diode imaging source. Three photomultiplier tubes were arranged with one detector centered to collect the confocal reflectance, and two detectors were arranged to collect the non-confocal reflectance split between right and left halves, as first reported by Scoles et al.[3] Subjects were cyclopledged using one drop each of tropicimide and phenylephrine, and their viewing position in the apparatus was stabilized using a bite bar. Subjects were instructed to fixate on a target while AOSLO video sequences of the fovea, parafovea, and meridians in the macula were acquired at a rate of 18 Hz. A reference frame was automatically selected from the image sequence using a custom algorithm.[22] Image sequences were desinuoided, aligned to the reference frame, and averaged using a strip-based registration algorithm to account for both intra- and inter-frame eye motion as previously described.[23] Images then were de-distorted to compensate for the estimated eye motion that occurred within the reference frame.[12,24] Finally, registered images from adjacent retinal locations were automatically montaged using a feature-matching algorithm.[10]

To compile a test dataset for validation of our algorithm, we identified five to seven locations that contained a minimum of four overlapping images from the montages for each of the five CHM subjects. These locations were selected such that, for each montage, at minimum two locations were from the foveal region, two locations were from the arms away from the fovea, and the remaining locations were from anywhere in the montage that met the minimum image overlap condition. Each location served as a different test, resulting in 29 test locations across the five montages. The test locations each had between four and 10 overlapping images, with a mean of 5.76 images for the 29 locations.

Manual Ground-Truth Comparisons

For each test location in the dataset, two human graders performed comprehensive pairwise comparisons between all overlapping pairs of images at the location, resulting in 469 comparisons in total across all locations in the dataset. The pairwise comparisons were generated using a custom software program that presented each image pair side-by-side to the grader and then recorded the grader's evaluation for which image in each pair had the better photoreceptor visibility. These pairwise comparisons served to provide ground truth in our evaluation.

To determine a full human grader ranking of photoreceptor visibility across all the images at each test location, each grader's pairwise comparisons were converted into an ordinal ranking using the Bradley–Terry model.[25] The model takes a probabilistic approach that uses the pairwise results as prior information to perform a log-likelihood estimation of a set of parameters representing relative scores for each image. These estimated scores can then be ordered to provide ground-truth ordinal rankings over the images at each test location.

In addition to the proposed methods (ST and AT), we evaluated for comparison two standard automated image quality assessment measures. The first method, known as the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE),[18] was developed using support vector regression to learn the quality score from a database of images with known image distortions, including compression artifacts, blurring, and noise. The second method, known as the Perception-Based Image Quality Evaluator (PIQE),[19] is an approach based on detecting human visual system–inspired features in the image and does not rely on training images. For BRISQUE and PIQE, we used MATLAB 2020b implementations with default parameters.

Preregistration

Our evaluation experiments were preregistered at https://osf.io/vmh6s/ prior to initiating the validation, and completion of the experiments did not require

Downloaded from intl.iovs.org on 04/23/2024

any post hoc modifications to the pre-registered experimental plan. (The nomenclature in the pre-registration document referred to photoreceptor visibility as "image quality," which was a broad term that we have since refined.)

Results

Pairwise Evaluation

We first evaluated the accuracy of each algorithm for simple pairwise comparisons of photoreceptor visibility between images with overlapping regions. For each pair of images, the automated algorithms were tasked with deciding which image in the pair had higher photoreceptor visibility. This was repeated for each of the 469 pairs of images in the validation dataset. Then, for each algorithm, we calculated the number of pairs where the algorithm ordering agreed with each grader. Table 1 shows the accuracy rate for each of the automated algorithms and between the two graders. We observed that the proposed Canny edge detection methods (both ST and AT) had higher agreement accuracy than BRISQUE and PIQUE, with performance that approached inter-grader accuracy.

Table 1. Accuracy Rates Across 469 Pairwise Comparisons of Photoreceptor Visibility

	Accuracy
Grader 1 vs. Grader 2	0.755
Grader 1 vs. proposed (ST)	0.665
Grader 1 vs. proposed (AT)	**0.667**
Grader 1 vs. BRISQUE	0.478
Grader 1 vs. PIQE	0.412
Grader 2 vs. proposed (ST)	0.680
Grader 2 vs. proposed (AT)	**0.708**
Grader 2 vs. BRISQUE	0.565
Grader 2 vs. PIQE	0.495

Two manual graders (Graders 1 and 2) served as the ground-truth ordering, which was used to evaluate the results from the proposed algorithm using the ST and AT, as well as two existing standard methods, BRISQUE[18] and PIQE.[19] The best performing algorithm result relative to each grader is highlighted in bold.

Ranking Evaluation

Our second analysis aimed to represent a more realistic scenario for application of the algorithm as it is applied to AOSLO montages. Instead of observing the pairwise accuracy, we evaluated how each algorithm performed a ranking of the overlapping images. For each of the 29 test locations with overlapping images, we used the algorithms to score every image at the location and then ordered the images according to their photoreceptor visibility score. For example, Figure 3 demonstrates the photoreceptor visibility order determined by the Bradley–Terry model[25] from the pairwise

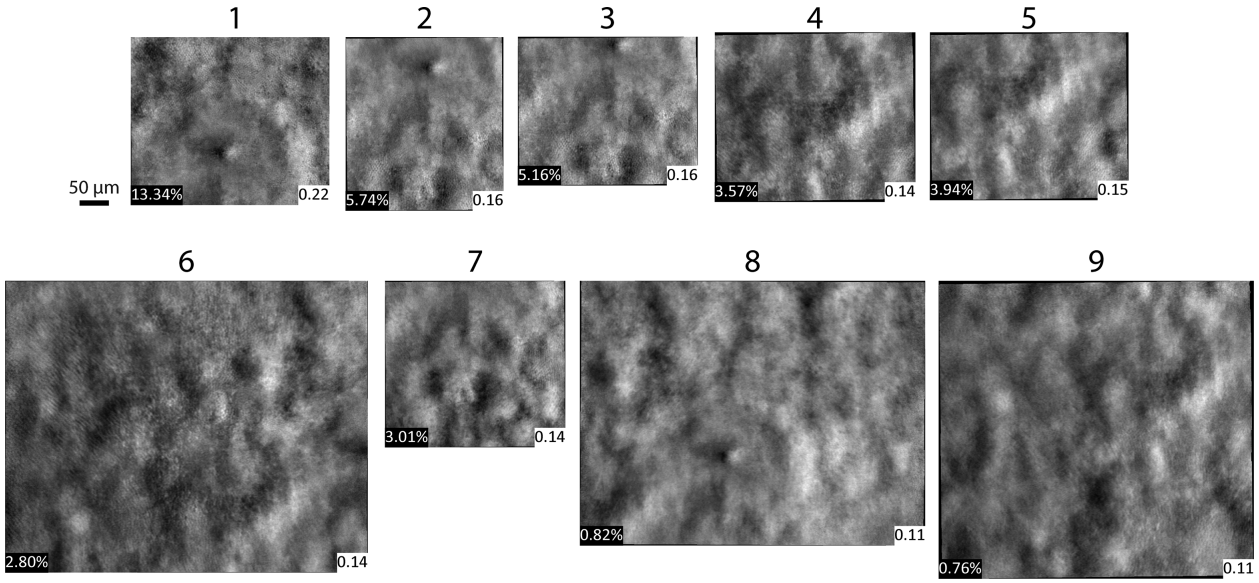

Figure 3. Example ranking of a retinal location with overlapping images with different photoreceptor visibility. The numbers above each image show the Bradley–Terry model ranking from the pairwise comparisons made by Grader 1, with 1 representing the best photoreceptor visibility and 9 representing the worst. The *left corner* of each image shows the percent edge pixel (*P*%) detected in the image (range, 0.76%–13.3%), and the *right corner* of each image shows the adaptive threshold (*T*) estimated to find 5% of the edge pixels in the image (range, 0.11–0.22).

translational vision science & technology

Downloaded from intl.iovs.org on 04/23/2024

Table 2. Example Rankings At a Retinal Location With Nine Overlapping Images of Different Photoreceptor Visibility

Image No.	Grader 1	Grader 2	Proposed (ST)	Proposed (AT)	BRISQUE	PIQE
1	1st	2nd	1st	1st	3rd	2nd
2	2nd	1st	2nd	2nd	2nd	1st
3	3rd	3rd	3rd	3rd	7th	8th
4	4th	4th	5th	5th	1st	3rd
5	5th	5th	4th	4th	6th	7th
6	6th	6th	7th	7th	4th	5th
7	7th	7th	6th	6th	5th	4th
8	8th	8th	8th	9th	8th	6th
9	9th	9th	9th	8th	9th	9th

Shown are all rankings for each image in Figure 3. Rankings were determined by each manual grader (Graders 1 and 2), the proposed algorithm using the ST and AT, and two existing standard methods, BRISQUE[18] and PIQE.[19] The numbers in each row indicate the rankings assigned for each image by the graders or algorithms, where 1st indicates the image with the best photoreceptor visibility and 9th indicates the worst photoreceptor visibility.

comparisons made by Grader 1. Table 2 shows the ordering of the images for this same location for each of the automated algorithms along with the ordering from Grader 2's pairwise comparisons. The rankings at this location for Grader 2 and the Canny edge detection algorithms (both ST and AT) correlated strongly with Grader 1's rankings, whereas the rankings from the BRISQUE and PIQE algorithms did not. We then evaluated the overall and mean Spearman correlations at each test location for the derived ordinal rankings of each automated algorithm and each grader's derived ordinal rankings. Table 3 shows the overall and mean correlation between the automated and manual rankings. Both the ST and AT variations of the proposed approach outperformed the BRISQUE and PIQE algorithms and approached the correlation observed between the manual graders' orderings.

AOSLO Montage Ordering

One goal of this work is to present an assessment measure that can be successfully applied to AOSLO montages, such that the images with the best photoreceptor visibility can be raised automatically to the top of the montage for downstream photoreceptor analysis. In Figure 4, we show a qualitative demonstration of applying the proposed measure for this purpose. We first constructed a full montage of an AOSLO dataset acquired from the retina of a patient with CHM (age 29 years). The montaging was performed using a previously presented, open-source software algorithm (https://github.com/BrainardLab/AOAutomontaging)[10] we developed, which detects, compares, and matches structural intensity patterns (known as SIFT features) between images. We then evaluated the proposed (AT) assessment measure on

Table 3. Spearman Correlations Between the Automated and Manual Ordinal Rankings at 29 Test Locations With Overlapping Images

		Grader 1 vs.				Grader 2 vs.			
	G1 vs. G2	Proposed (ST)	Proposed (AT)	BRISQUE	PIQE	Proposed (ST)	Proposed (AT)	BRISQUE	PIQE
Overall	0.748	0.603	**0.609**	0.335	0.191	0.588	**0.615**	0.458	0.325
Mean	0.566	0.408	**0.434**	−0.217	−0.370	0.406	**0.500**	−0.009	−0.133

"Overall" is the total correlation of all rankings across all locations, and "mean" is the average across the separate correlations at each of the 29 locations. Two manual graders (G1 vs. G2) provided ground-truth orderings generated by analysis of their pairwise comparisons, which were used to evaluate the orderings from the proposed algorithm using the ST and AT, and two existing standard methods, BRISQUE[18] and PIQE.[19] The best performing algorithm result relative to each grader is highlighted in bold.

translational vision science & technology

Downloaded from intl.iovs.org on 04/23/2024

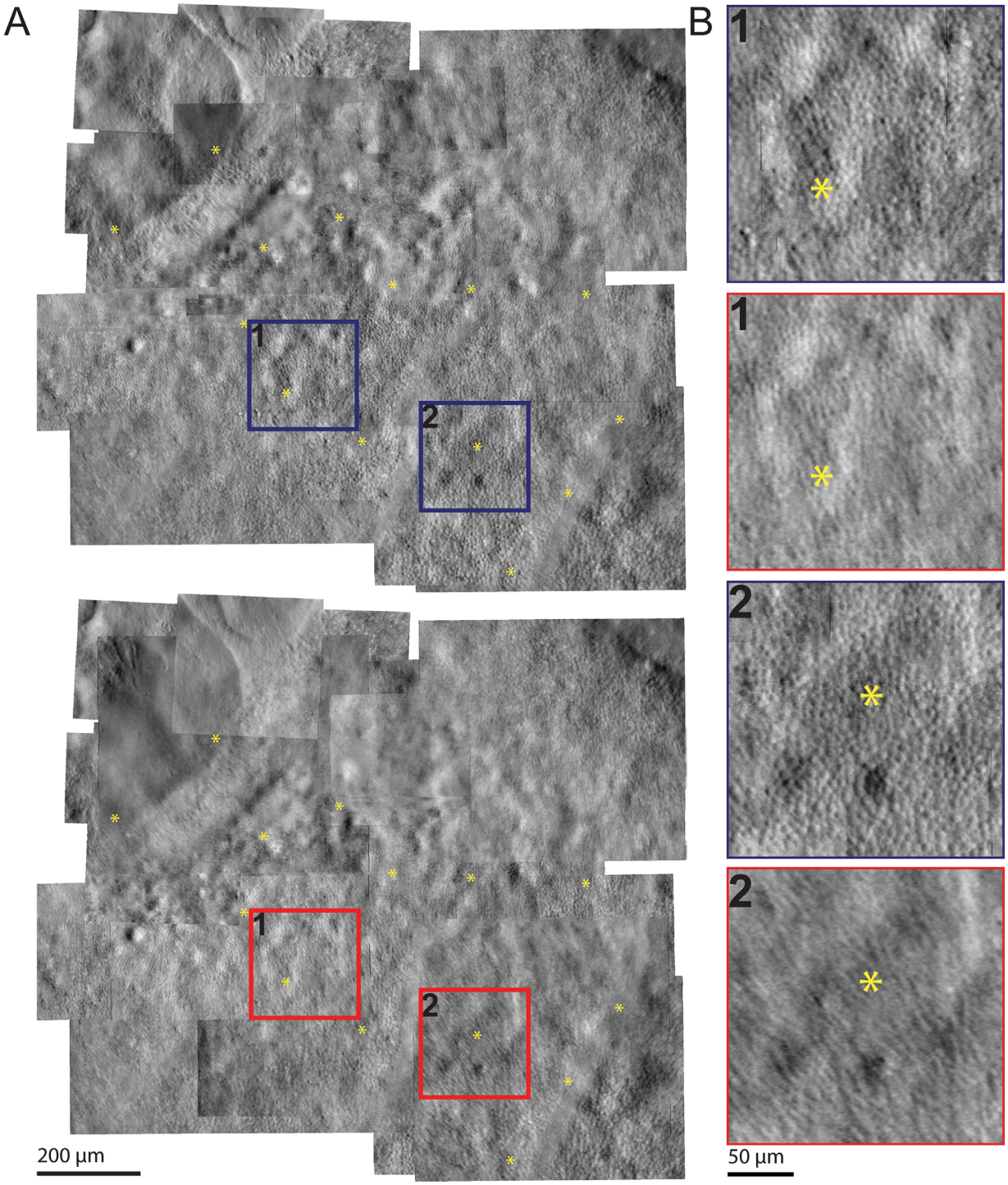

Figure 4. (A) Example AOSLO split-detection montage from CHM subject 13048 (age 29 years), with images ordered automatically using our proposed edge-based assessment measure (AT) and displayed with highest to lowest (*top montage*) or lowest to highest (*bottom montage*) photoreceptor visibility. (B) Zoomed in images at the corresponding locations indicated by the *red* and *blue boxes* in the two montages. *Yellow asterisks* mark example locations where the photoreceptor mosaic is more readily visible in the top montage.

each image in the montage and ordered the images according to their score. In Figure 4A, the top montage displays the AOSLO images ordered from highest to lowest photoreceptor visibility according to the AT score, and the bottom montage displays the images order from lowest to highest visibility. Yellow asterisks denote example locations in the montage where photoreceptor visibility is altered based on which

Downloaded from intl.iovs.org on 04/23/2024

image is displayed in the montage. We can observe from the zoomed in regions in Figure 4B that ordering the images with the highest score on top yields a montage with significantly better photoreceptor visibility.

Discussion

This work aims to develop an automated method for quantitatively assessing the visibility of photoreceptors in AOSLO split-detection images. This automatic assessment allows us to reduce the amount of manual intervention required to produce a high-quality AOSLO montage before further downstream analysis of the cone mosaic. Our experiments evaluated the performance of the proposed method using both the grader pairwise comparison data directly and rankings derived from those comparisons. From these results, we observed that the proposed method performed photoreceptor visibility assessment approaching that of human graders and significantly outperformed two standard baseline image quality methods (BRISQUE and PIQE) that were not designed specifically to assess photoreceptor visibility. Good performance was found using both the ST and AT variants of the proposed method. We found, however, that the AT approach provided slightly better performance than the ST approach. This was likely due to the AT allowing for a more robust window of analysis such that images with very high or very low cone density did not approach the floor or ceiling for possible edges detected in the images.

Applications for AOSLO Montaging

The target application of our method is to automatically order images within AOSLO montages such that the AOSLO images with high cone visibility are displayed on the top of the montage. Figure 4 shows a qualitative example of how the proposed algorithm can be used for automatically improving photoreceptor visibility within a montage. By algorithmically ordering the images within the montage, more regions with clear images of cones are observed. This suggests that the features found by the proposed method are congruent with our interpretation of high photoreceptor visibility for AOSLO split-detection images. A default ordering without a photoreceptor visibility assessment, such as a random ordering, or an ordering based on the timing of image acquisition, would likely result in a montage with photoreceptor visibility in between the best and worst case examples shown in Figure 4. The advantage of using automated assessments for image ordering comes in the form of reducing the time required for manual input. From our experience, running on a standard desktop computer (four-core Intel i7 with 16 GB ram), the proposed algorithm can operate on a full AOSLO montage dataset in 30 to 60 seconds depending on the number of images in the dataset. This represents a considerable speedup in throughput relative to manual processing which is on the order of hours. One consideration, however, is that, although the algorithm can in principle be operated automatically without manual intervention, we can observe from our evaluation results that the ordering is not perfect relative to a human grader. In practice, the proposed algorithm can achieve higher accuracy if used in a semi-automated fashion, where the data are processed automatically in an initial run through, and then manual adjustments are made to optimize the final ordering. However, even when used semi-automatically in this manner, based on informal reports, our graders observed a reduction of several hours of processing per montage dataset.

We also note that, in our montage demonstration, we only provide a basic application of the proposed measure for ordering the montage, in which entire images are ordered and the highest ranked images are brought to the top. More advanced techniques can potentially be implemented where image overlap information from the montage can be used to restrict the regions where the measure is evaluated. This can potentially allow more nuanced comparisons between images and may provide better results for cases where a single image has both high- and low-visibility regions. However, implementation of such a technique is not straightforward. For regions with more than two overlapping images, there can be varying combinations of intersections between the images that can potentially contradict each other and create loops when used for ordering (e.g., image A has better visibility than parts of image B, and image B has better visibility than parts of image C, but image C also has better visibility than parts of image A). Alternatively, we can attempt to divide the images into pieces according to their overlapping regions and bring the best piece from each image to the top of the montage. However, due to the inevitable small differences in alignment, inherent image distortions, and ancillary image differences (e.g., intensity) in each image, such an approach might produce an inordinately patchy-appearing montage that may not be suitable for downstream photoreceptor quantification. Further investigation is necessary to develop such applications for the proposed measure.

Limitations

One challenge with this work is that manual assessment of photoreceptor visibility is a highly subjective analysis. We see from Table 1 that, even between two

Downloaded from intl.iovs.org on 04/23/2024

trained graders, we only observed a 75% agreement on which image in a pair had the best photoreceptor visibility. This variance is a result of several factors. First, in the AOSLO montages, two images are rarely directly on top of each other. This means each image will have non-overlapping regions that may have different photoreceptor visibility characteristics than the overlapping regions. Different graders might focus on different regions of an image, resulting in differences during the evaluation. Second, some overlapping images may have similar or equal photoreceptor visibility, so that differences between them are below the threshold at which a grader can make reliable judgments. Agreement between graders is likely higher for image pairs that differ considerably in photoreceptor visibility and approaches chance for those that do not differ much. Figure 3 illustrates a case where the overlapping images span a wide range of quality and where grader agreement is high (97%). Conversely, in the Supplementary Materials (Supplementary Fig. S1 and Supplementary Table S1) we show an example location with low variability between images, where the grader agreement fell to 57%. The images in our evaluation set were chosen to represent a typical use case and include both images that differ considerably and those that do not differ much.

We have not explored the application of the proposed measure for AOSLO confocal images. Although the confocal modality can offer higher resolution of photoreceptors in healthy subjects (particularly for rod mosaics and foveal cones), there are significant challenges associated with assessing photoreceptor visibility in confocal images of patients with disease. Foremost, the photoreceptor waveguiding observed in confocal AOSLO images is variably compromised in disease, with some disease states yielding normally waveguiding photoreceptors, others yielding dim or mottled cone reflectance profiles, and still others resulting in non-waveguiding gaps in the photoreceptor mosaic. We reasoned that an automated algorithm for confocal AOSLO images may be more dependent on disease state and thus require more disease-specific tuning and evaluation or have more limited applicability. Thus, we limited our evaluation to split-detection images because in many disease states photoreceptor inner segments appear with similarly distinct edges in the split-detection images.

Although we only formally evaluated our algorithm on eyes with choroideremia, our intention was to develop a broadly applicable algorithm for assessing photoreceptor visibility in disease, irrespective of the specific diagnosis. For example, we have successfully used the algorithm to order images obtained from subjects with achromatopsia and (with smaller *N*) other inherited retinal degenerations. For these cases, we qualitatively found photoreceptor visibility in the ordered montage surpasses that of the default montage with no ordering (data not shown). Regardless of the fact that different pathologies yield different levels of photoreceptor visibility, we predict our algorithm will identify which images within a dataset contain the highest photoreceptor visibility. Thus, we expect that our proposed Canny edge detection–based photoreceptor visibility assessment measure can be broadly applied to improve automated visualization of photoreceptors within an AOSLO montage regardless of disease state. The primary reason healthy control data was not included in the analysis presented here is that the photoreceptor mosaic in AOSLO images of healthy eyes is almost always visible. This fact not only negates the need for photoreceptor visibility assessment when forming AOSLO montages of healthy eyes but also means it would be difficult to establish a reliable ground-truth ranking of images based upon manual agreement.

Finally, we observe anecdotally that, because the confocal and split-detection modalities are acquired simultaneously, high-quality confocal images typically go hand in hand with high-quality split-detection images. Thus, we predict that a manual ordering of confocal images based on photoreceptor visibility would correlate with an ordering of their corresponding split-detection images. This would allow an ordering of split-detection images by our algorithm to be also used indirectly to order confocal images, as well. Testing this hypothesis, however, is outside of the scope of the present study and should be pursued in future work. Other future applications of our assessment measure could include using the score to filter AOLSO images prior to being montaged, thus saving on the computation time required to form the montage itself. Alternatively, the scores could potentially help determine when downstream analysis of an image may be unreliable for further analysis.

Software

We provide our edge-based photoreceptor visibility assessment software as an open-source repository. It is available to download at https://github.com/PennVRC/AOQualitySorter.

Conclusions

We present an approach for automatically assessing the photoreceptor visibility of AOSLO split-detection images using edge-based features. Our automated assessment is comparable to manual rankings and

Downloaded from intl.iovs.org on 04/23/2024

provides significant improvement over two standard general image quality assessment measures. This finding suggests that edge-based assessment of AOSLO images can be used to automatically order images within AOSLO montages for improved photoreceptor visibility. This will result in a significant reduction of manual labor required to finalize high-quality AOSLO montages and will enable higher throughput for studies requiring downstream analysis and quantification of the photoreceptor mosaic.

Acknowledgments

The authors thank Alfredo Dubra for sharing the adaptive optics scanning laser ophthalmoscopy optical design and the adaptive optics control, image acquisition, and image registration software.

Supported by grants from the National Institutes of Health (NIH R01EY028601, R01EY030227, and P30EY001583), Foundation Fighting Blindness, F. M. Kirby Foundation, Research to Prevent Blindness, and the Paul and Evanina Mackall Foundation Trust.

Disclosure: **M. Chen**, None; **Y.Y. Jiang**, None; **J.C. Gee**, None; **D.H. Brainard**, U.S. Patent Application 16/389,942 (P); **J.I.W. Morgan**, U.S. Patent 8226236 (P), U.S. Patent Application 16/389,942 (P), and Applied Genetic Technologies Corporation (F)

References

1. Morgan JI. The fundus photo has met its match: optical coherence tomography and adaptive optics ophthalmoscopy are here to stay. *Ophthalmic Physiol Opt*. 2016;36:218–239.
2. Roorda A, Duncan JL. Adaptive optics ophthalmoscopy. *Annu Rev Vis Sci*. 2015;1:19–50.
3. Scoles D, Sulai YN, Langlo CS, et al. In vivo imaging of human cone photoreceptor inner segments. *Invest Ophthalmol Vis Sci*. 2014;55:4244–4425.
4. Sun LW, Johnson RD, Williams V, et al. Multimodal imaging of photoreceptor structure in choroideremia. *PLoS One*. 2016;11:e0167526.
5. Tuten WS, Vergilio GK, Young GJ, et al. Visual function at the atrophic border in choroideremia assessed with adaptive optics microperimetry. *Ophthalmol Retina*. 2019;3:888–899.
6. Morgan JI, Han G, Klinman E, et al. High-resolution adaptive optics retinal imaging of cellular structure in choroideremia. *Invest Ophthalmol Vis Sci*. 2014;55:6381–6397.
7. Georgiou M, Litts KM, Kalitzeos A, et al. Adaptive optics retinal imaging in CNGA3-associated achromatopsia: retinal characterization, interocular symmetry, and intrafamilial variability. *Invest Ophthalmol Vis Sci*. 2019;60(1):383–396.
8. Talcott KE, Ratnam K, Sundquist SM, et al. Longitudinal study of cone photoreceptors during retinal degeneration and in response to ciliary neurotrophic factor treatment. *Invest Ophthalmol Vis Sci*. 2011;52(5):2219–2226.
9. Morgan JI, Jiang YY, Serrano LW, et al. Cone mosaic integrity in choroideremia following gene augmentation via subretinal injection of AAV2-hCHM. *Invest Ophthalmol Vis Sci*. 2021;62(8):47–47.
10. Chen M, Cooper RF, Han GK, Gee J, Brainard DH, Morgan JI. Multi-modal automatic montaging of adaptive optics retinal images. *Biomed Opt Express*. 2016;7(12):4899–4918.
11. Davidson B, Kalitzeos A, Carroll J, et al. Fast adaptive optics scanning light ophthalmoscope retinal montaging. *Biomed Opt Express*. 2018;9(9):4317–4328.
12. Chen M, Cooper RF, Gee JC, Brainard DH, Morgan JI. Automatic longitudinal montaging of adaptive optics retinal images using constellation matching. *Biomed Opt Express*. 2019;10(12):6476–6496.
13. Bidaut Garnier M, Flores M, Debellemanière G, et al. Reliability of cone counts using an adaptive optics retinal camera. *Clin Exp Ophthalmol*. 2014;42(9):833–840.
14. Morgan JI, Chen M, Huang AM, Jiang YY, Cooper RF. Cone identification in choroideremia: repeatability, reliability, and automation through use of a convolutional neural network. *Transl Vis Sci Technol*. 2020;9(2):40.
15. Tanna P, Kasilian M, Strauss R, et al. Reliability and repeatability of cone density measurements in patients with Stargardt disease and RPGR-associated retinopathy. *Invest Ophthalmol Vis Sci*. 2017;58(9):3608–3615.
16. Wang Z, Bovik AC. *Modern Image Quality Assessment: Synthesis Lectures on Image, Video, and Multimedia Processing*. Williston, VT: Morgan & Claypool Publishers; 2006:1–56.
17. Zhai G, Min X. Perceptual image quality assessment: a survey. *Sci China Inform Sci*. 2020;63(11):211301.
18. Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain.

Downloaded from intl.iovs.org on 04/23/2024

IEEE Trans Image Process. 2012;21(12):4695–4708.
19. Venkatanath N, Praneeth D, Bh MC, Channappayya SS, Medasani SS. Blind image quality evaluation using perception based features. In: *2015 Twenty First National Conference on Communications (NCC)*. Piscataway, NJ: Institute of Electrical and Electronics Engineers; 2015:1–6.
20. Canny J. A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell*. 1986(6):679–698.
21. Dubra A, Sulai Y. Reflective afocal broadband adaptive optics scanning ophthalmoscope. *Biomed Opt Express*. 2011;2:1757–1768.
22. Salmon AE, Cooper RF, Langlo CS, Baghaie A, Dubra A, Carroll J. An automated reference frame selection (ARFS) algorithm for cone imaging with adaptive optics scanning light ophthalmoscopy. *Transl Vis Sci Technol*. 2017;6:9.
23. Dubra A, Harvey Z. Registration of 2D images from fast scanning ophthalmic instruments. *Lect Notes Comput Sci*. 2010;6204:60–71.
24. Bedggood P, Metha A. De-warping of images and improved eye tracking for the scanning laser ophthalmoscope. *PLoS One*. 2017;12:e0174617.
25. Bradley RA, Terry ME. Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*. 1952;39(3/4):324–345.

Downloaded from intl.iovs.org on 04/23/2024