Automated identification of cone photoreceptors in adaptive optics retinal images

Kaccie Y. Li and Austin Roorda

School of Optometry, University of California, Berkeley, California 94720, USA

In making noninvasive measurements of the human cone mosaic, the task of labeling each individual cone is unavoidable. Manual labeling is a time-consuming process, setting the motivation for the development of an automated method. An automated algorithm for labeling cones in adaptive optics (AO) retinal images is implemented and tested on real data. The optical fiber properties of cones aided the design of the algorithm. Out of 2153 manually labeled cones from six different images, the automated method correctly identified 94.1% of them. The agreement between the automated and the manual labeling methods varied from 92.7% to 96.2% across the six images. Results between the two methods disagreed for 1.2% to 9.1% of the cones. Voronoi analysis of large montages of AO retinal images confirmed the general hexagonal-packing structure of retinal cones as well as the general cone density variability across portions of the retina. The consistency of our measurements demonstrates the reliability and practicality of having an automated solution to this problem. © 2007 Optical Society of America

OCIS codes: 100.5010, 010.1080, 330.5310.

1. INTRODUCTION

Physical limitations and consequences due to the variability of the packing arrangement of retinal cones were first reported by Yellott[1] in 1983. Since then, the packing structure of retinal cones has been studied both anatomically and psychophysically.[2–6] The advent of retinal imaging systems with adaptive optics (AO) has made it possible to image the living human retina at the microscopic scale.[7] Now that noninvasive studies of the anatomy and physiology of human cones are possible,[8,9] it is reasonable to expect future studies to be done on a much greater scale. Automated methods for data analysis have found application in many medical and scientific fields and have the potential to become a useful tool in the field of retinal imaging. The quantity of data included in any study is of great importance. Reliable automated routines are naturally preferred when large quantities of data need to be analyzed. We want to facilitate the process of determining cone density and cone packing arrangement in AO images to encourage the inclusion of larger data sets in future studies. Manually labeling cones in an image is usually reliable, but it becomes an impractical solution when multiple images are involved. In this paper, we automate the cone labeling process with an algorithm implemented in MATLAB (Mathworks, Inc., Natick, Massachusetts) and functions from the MATLAB Image Processing Toolbox (IPT). Code can be accessed from our research group's Web page.[10] We demonstrate the algorithm's effectiveness in analyzing actual AO retinal images. Algorithm performance is assessed by comparisons with manually labeled results for six different AO retinal images.

Studies that have addressed the sampling limitations of retinal cones usually model the cone mosaic as a hexagonal array of sampling points.[2,6,11] Maximum cone density is achieved with this packing arrangement. It is likely that this is the main reason why retinal cones tend to develop into a hexagonal array in areas where few or no rods are present. A hexagonal sampling grid is the optimal solution for most signal processing applications.[12] Nevertheless, the sampling performance of cones is more commonly associated with cone density rather than arrangement geometry. Knowing that the packing arrangement of cones affects the sampling properties of the retina, why are cones hexagonally arranged only in localized regions and what advantages do such variations in packing arrangement offer?[3,5,6,11] Furthermore, analyses of cones across a variety of eccentricities show that their packing arrangement is also highly variable. Using our algorithm, we analyzed the cones from seven montages constructed from images acquired at various eccentricities using AO. The computed cone locations are then used to generate density contour maps and to make quantified measurements of their packing structure. The consistency of our results indicates the reliability of an automated method for analyzing AO images. Our analysis demonstrates how the process of extracting quantitative information about the density and packing arrangement of cones from images of living retinas can be accomplished efficiently.

2. METHOD

Image formation for an optical system can be described by the convolution of the object with the system's impulse response or point-spread function (PSF):

$$I'(x_1, x_2) = I(x_1, x_2) * \beta(x_1, x_2), \tag{1}$$

where $I'(x_1, x_2)$ is the observed image, $I(x_1, x_2)$ is the actual cone mosaic under observation $*$ is the convolution operator, and $\beta(x_1, x_2)$ is the PSF. AO corrects the aberra-

tions of the eye to minimize the distribution of $\beta(x_1, x_2)$, but residual wavefront error after correction is still often too high for us to see cones near or at the fovea. Image restoration techniques such as deconvolution can enhance the appearance of the image even further,[13] but if retinal cones are not optically resolved, no method can reliably identify them. Under the condition that the cones under observation are optically resolvable, we show that it is only necessary to understand the role of noise in the cone identification process. The observed image $I'(x_1, x_2)$ is converted by the detector into a finite two-dimensional sequence:

$$I'(n_1, n_2) = I'(x_1 T_1, x_2 T_2) + \eta(n_1, n_2), \qquad (2)$$

where T_1 and T_2 are the horizontal and vertical sampling periods determined by the finite pixel size of the detector and $\eta(n_1, n_2)$ is a generalized noise term. The power spectrum of an AO retinal image shown in Fig. 1 contains an approximate hexagonal band region produced by the regular arrangement of cones. Signals that do not correspond to cones, which may originate either from noise in the detection channel or from other features in the retina, are effectively noise. Filtering and morphological image processing are used to isolate signals corresponding only to cone photoreceptors.[14,15]

The intensity of light incident on each pixel of the detector is encoded as integral values that make up the two-dimensional sequence $I'(n_1, n_2)$. This sequence can be viewed as a surface topography of intensity counts over a discrete plane. Cones are single-mode optical fibers oriented toward the direction of the pupil center with very small disarray.[9,16] As a result, reflected light from these fibers can be considered as an array of point sources at the retinal plane. Therefore, the light distributions observed in an AO image are PSFs whose peaks correspond to the actual cone locations. Surface topography features
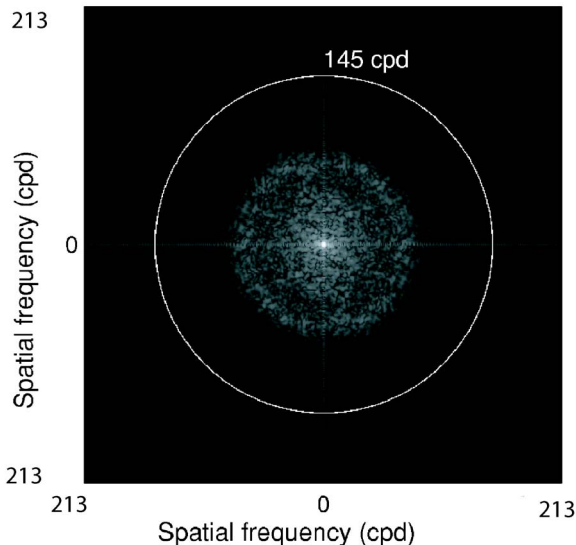

Fig. 1. (Color online) Power spectrum of an AO retinal image (enhanced by log scale) generated using the fast Fourier transform. The shape and size of the band region indicate the hexagonal arrangement and sampling limits of the cones in the image. The periodicity of the tightest packed cones is about 145 cycles per degree (cpd).

that correspond to these PSF peaks are points of local maxima. When cones are optically resolved, we can determine the locations of all the cones by identifying all local maxima in the original image. In an image sampled by an array of finite detector elements, a local maximum is defined to be any pixel whose value exceeds each of its neighboring pixels. IPT function *imregionalmax* effectively accomplishes this task. However, direct application to the observed sequence $I'(n_1, n_2)$ produces a result that exceeds the actual number of cones in the image. These overidentifications are primarily due to the presence of high-frequency noise and can be greatly reduced by applying the appropriate low-pass filter. A low-pass finite-impulse-response (FIR) filter is designed for this purpose and applied to $I'(n_1, n_2)$ prior to further image processing:

$$f(n_1, n_2) = h(n_1, n_2) * I'(n_1, n_2), \qquad (3)$$

where $h(n_1, n_2)$ is the impulse response (15 by 15 pixels). Filter parameters are selected based on the field size of the system and the estimated minimum cone spacing. The minimum cone spacing in the central fovea is approximately $2~\mu$m.[4] Using a conversion factor of $300~\mu$m/deg eccentricity, the frequencies of retinal cones should range from nearly zero to approximately 145 cycles per degree (cpd). The FIR filter $h(n_1, n_2)$ is designed to implement the desired ideal low-pass filter that passes only spatial frequencies within 145 cpd. The densest cones that can be imaged using AO reside within 145 cpd, so filter parameters can be adjusted accordingly. Many FIR filter types exist and can be rather extensive. The most popular filter designs can be found in signal processing texts,[17,18] so we will not state further details for this step.

There is more than one method to implement the convolution operation in Eq. (3), but an appropriate method must provide a justifiable solution for computing the pixel values toward the boundary of $f(n_1, n_2)$. Suppose the sequences $I'(n_1, n_2)$ and $h(n_1, n_2)$ have $N \times N$ and $M \times M$ support, respectively. The convolution of $I'(n_1, n_2)$ with $h(n_1, n_2)$ initially requires extending the bounds of $I'(n_1, n_2)$ to achieve $M+N-1 \times M+N-1$ support. This is typically done by zero padding $I'(n_1, n_2)$, which introduces discontinuities along its boundaries. The resulting $f(n_1, n_2)$ would contain unnecessary blurring and noise near the boundaries. A second method is to take the discrete Fourier transform (DFT) of both $I'(n_1, n_2)$ and $h(n_1, n_2)$ and multiply them in frequency space. The resultant $f(n_1, n_2)$ is then restored by taking the inverse DFT. Due to the nonperiodicity of $I'(n_1, n_2)$, taking the DFT in this manner introduces aliasing along the boundaries of $f(n_1, n_2)$. A practical solution to this problem could be to remove the corrupted pixels by applying a window:

$$f(n_1, n_2) = W(n_1, n_2)[h(n_1, n_2) * I'(n_1, n_2)], \qquad (4)$$

where $W(n_1, n_2)$ is the windowing function. The resultant sequence $f(n_1, n_2)$ is reduced to have only $N-M \times N-M$ support. A reliable alternative exists to avoid this loss in analyzable data. We can extend the bounds of $I'(n_1, n_2)$ with values equal to its nearest border values rather than with zeros. As a result, $N \times N$ support of the input sequence $I'(n_1, n_2)$ is preserved along with continuity at its
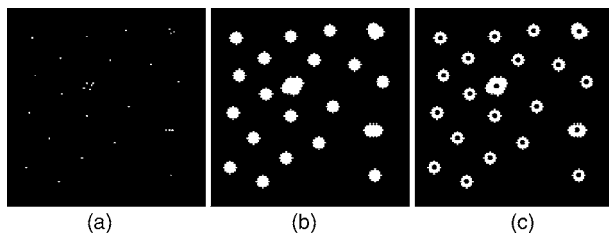
Fig. 2.   Output of IPT function *imregionalmax* is a (a) binary sequence whose values are nonzero at all identified local maxima in the input sequence. (b) This sequence is dilated with a disk-shaped structuring element, and (c) cone locations are computed by computing the center of mass of each object in the sequence after dilation.
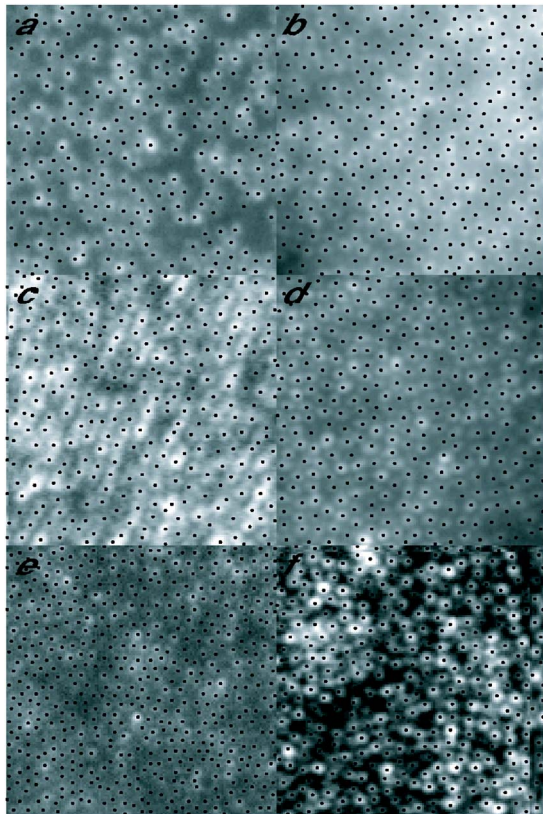

Fig. 3.   (Color online) Every marker in these six grayscale images is manually placed by the authors. Each image is 8 bits and 128×128 pixels, and each marker is accurate to the nearest pixel.

boundaries. IPT functions *fwind2* and *imfilter* are useful for the implementation of these operations.

IPT function *imregionalmax* generates a binary sequence, as shown in Fig. 2(a), where each nonzero pixel corresponds to a local maximum. Missidentifications appear as multiple nonzero pixels that lie within a vicinity that is not physically realizable by the same number of cones. These pixels are grouped into a single object using morphological dilation (IPT function *imdilate*). As shown in Fig. 2(b), this operation translates a binary disk across the domain of the binary sequence, replacing each nonzero element with the disk similar to the convolution. The disk is set at 2 μm in diameter (i.e., the minimum possible cone spacing). The final cone locations, shown in Fig. 2(c), are determined by computing the center of mass of each object after dilation.

We took six retinal images, all with 128×128 pixel support, and labeled them manually as shown in Fig. 3. This diverse set of images varies considerably in cone density, contrast, and, in some cases, biological structure. Comparisons are made with automated results on the same data set to assess the agreement between the two methods. Comparisons are also made between outcomes from five experienced human observers for two of the six images. An agreement is made when a pair of corresponding cones is located within 2 μm of each other. This value was chosen because cone diameters at the observed eccentricities are at least 4 μm. Identified cones that do not satisfy this criterion are defined as disagreements. The level of agreement between the labeling methods is quantified by computing the mean displacement along both the horizontal and the vertical directions and is reported in Table 2. The cone packing arrangement is analyzed graphically using Voronoi diagrams.[19] Voronoi analyses for the two selected images are done on the results from the five human observers and the algorithm.

The image montages analyzed are acquired from one monkey and six human retinas. The images were acquired by both the flood-illuminated AO system at the University of Rochester and an AO scanning laser ophthalmoscope (SLO) (see Table 1).[7,20] Cone density is computed at each cone by counting the number of cones that lie within a defined circular sampling window (approximate radius of 0.07 deg or 20 μm). Cones whose corresponding sampling window extends beyond the bounds of the image are excluded. We generated contour maps of cone density values to observe the variability of cone density across each montage. Linear interpolation is used to fill the spaces between cones with the estimated density values. Since packing structure tends to vary greatly across the retina, each montage is divided into 0.125 deg sections, and Voronoi analysis is done on each individual section. Voronoi regions that extend beyond the bounds of each image section are excluded from all analysis.

3. RESULTS
The number of cones identified in every image of Fig. 3 is reported in Table 2. For each image, the algorithm found approximately equal numbers of cones as the authors did. When analyzing the packing arrangement of retinal cones, we are more concerned with the agreement between the automated and the manual labeling methods. As outlined in Table 2, a 93% to 96% agreement between the two methods is achieved, and the physical locations of

Table 1. Sources for Cone Photoreceptor Images[a]

	Image	System	Eccentricity (deg)
1	Monkey	AO flood illuminated	1.10 to 1.86
2	Subject 1	AO flood illuminated	0.30 to 1.68
3	Subject 2	AO SLO	0.60 to 2.60
4	Subject 3	AO SLO	1.15 to 4.27
5	Subject 4	AO SLO	1.35 to 4.97
6	Subject 5	AO SLO	0.74 to 2.61
7	Subject 6	AO SLO	1.55 to 5.06

[a]Refs. 7, 20, 21.

Table 2. Performance Comparison between Manual and Automated Methods

Image	Number of Cones	% Agreement (%)	% Disagreement (%)	Δx (μm)	Δy (μm)
(a)	274	95.4	9.1	0.06 ± 0.48	-0.07 ± 0.43
(b)	354	94.1	1.2	0.09 ± 0.42	-0.06 ± 0.37
(c)	301	92.7	8.2	-0.02 ± 0.49	-0.04 ± 0.41
(d)	315	96.2	5.6	0.04 ± 0.47	-0.01 ± 0.52
(e)	557	93.5	6.3	0.08 ± 0.44	0.03 ± 0.41
(f)	352	93.8	5.4	0.06 ± 0.38	-0.05 ± 0.42

each corresponding cone pair deviated within 0.52 μm from each other. Disagreements between the two methods, ranging from 1% to 9%, are due to human observer errors and/or unrelated signals that are not removed during the preprocessing step of the algorithm. The highest percentages of disagreement are seen for images (a) and (c). The cone mosaic in (a) clearly has some unusual structures,[22] and the cones present in (c) are borderline resolvable in many areas. In contrast to (a) and (c), other images resulted in better agreement between the two methods. Agreement between outcomes from five experienced human observers spanned from 74.1% to 90.5% for Fig. 3(c) and 92.1% to 96.8% for Fig. 3(d). Analysis of the shape of each Voronoi region allows one to predict the spatial sampling capabilities offered by the observed cone mosaic. Voronoi regions are divided into hexagonal (light gray) and nonhexagonal (dark gray) categories as seen in Figs. 4 and 5. Voronoi diagrams appear to be rather sensitive to modest differences in cone labeling results. Voronoi diagrams of Fig. 4 are of the same image labeled by five human observers and the automated algorithm. Feature variability in these diagrams indicates that uncertainty due to questionable image quality and/or inadequate experience of the observer is not a negligible factor. Among the observers, the percentage of hexagonal Voronoi regions spanned from 39.4% to 52.5% for Fig. 4(c) and 55.36% to 62.71% for Fig. 4(d). The corresponding results for the automated method were at 40.8% and 57.7% for Figs. 4(c) and 4(d), respectively. These observations indicate that decisions made by human observers can significantly influence the outcome of a packing analysis and that the consistency of an automated method may actually be more appropriate for certain images.
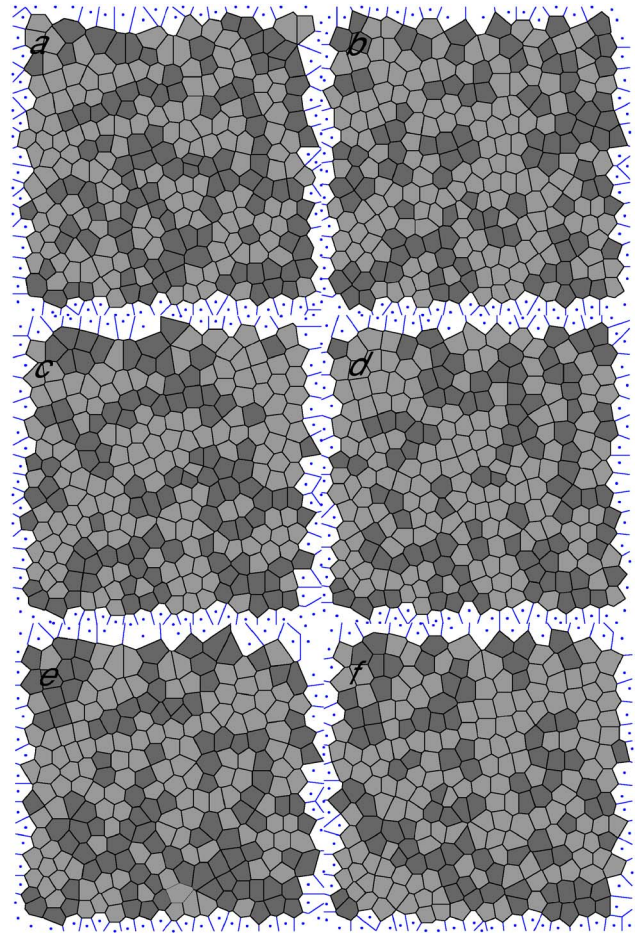

Fig. 4.   (Color online) Voronoi diagrams corresponding to the mosaic of Fig. 3(d) computed from cone locations acquired by (a) the automated method and (b)–(f) the five experienced human observers.
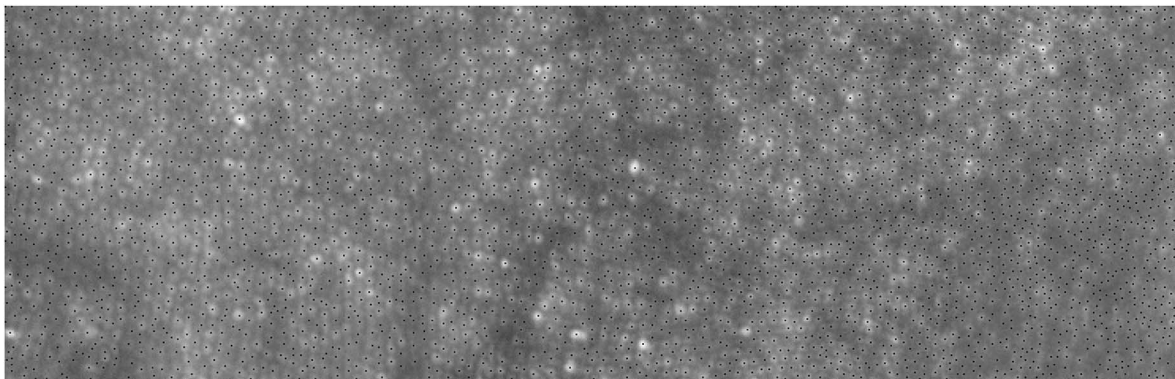

Fig. 5.   Montage image of the monkey retina acquired by the AO flood-illuminated system at the University of Rochester and labeled with the proposed algorithm.

0.17   0.30   0.42   0.55   0.67   0.80   0.92   1.05   1.17   1.30   1.42   1.55   1.67   1.80
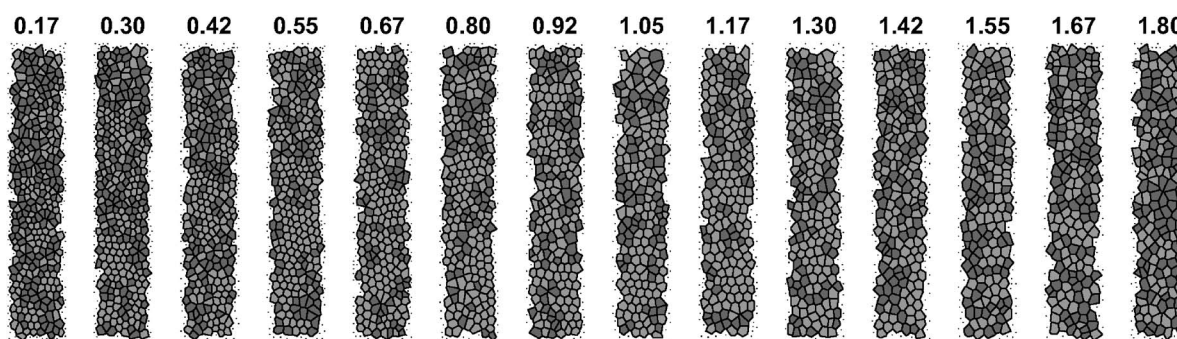

Fig. 6.   Voronoi diagrams generated for the montage given in Fig. 4 at the eccentricities specified above each diagram. Hexagonal regions are shaded in light gray.
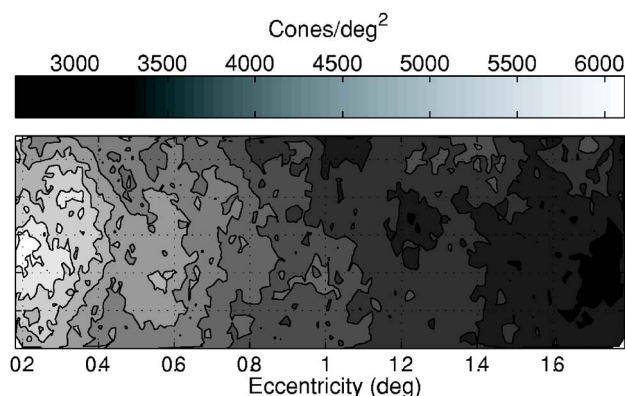

Fig. 7.   (Color online) Topography map of retinal cone density for the montage given in Fig. 4.
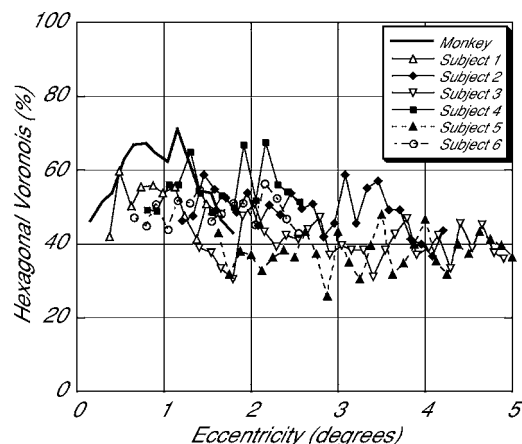

Fig. 8.   Percentages of hexagonal Voronoi regions plotted over eccentricity for all seven montages analyzed in this study.

We computed cone densities and Voronois using the methods described above for seven large montages with well-resolved cones. The montage acquired from the monkey is shown in Fig. 5. The Voronoi diagram for this cone mosaic and its corresponding cone density contour map are given in Figs. 6 and 7. Voronoi diagrams and cone density plots illustrate the variability in cone packing arrangement and density across the retina. Voronoi regions gradually increase in size at higher eccentricities, which signifies how cone density decreases monotonically as eccentricity increases. This can also be seen from the cone

density contour map. The retinal montages cover eccentricities from 0.17 to 5.00 deg. Cone density values within this range varied from approximately 1040 to 6500 cones/deg^2. This converts to approximately 12,300 to 72,200 cones/mm^2, which agrees with the trends reported by Curcio *et al.*[4] In a large montage, it is possible to identify localized patches in the mosaic where there is prominent hexagonal-packing structure. The percentages of hexagonal regions across 0.125 deg sections are plotted in Fig. 8. In general, at least 40% to 50% of retinal cones are hexagonally arranged, but more than 70% hexagonal packing may appear in certain localized regions. In a study done by Curcio and Sloan,[3] an analysis done on one prepared human retina suggested that cones may be more hexagonally arranged near the edge of the fovea (between 0.20 and 0.25 deg eccentricity) rather than at the fovea center. This suggests that, even though cones are more densely packed toward the foveal center, they do not necessarily become more hexagonally arranged. This tendency may also be present in our data for the monkey (71% at 1.17 deg and 46% at 0.17 deg) and subject 1 (60% at 0.50 deg and 42% at 0.37 deg) as shown in Fig. 8. However, better-quality images of the foveal center and its immediate periphery are needed to confirm this hypothesis.

4. DISCUSSION

The abundance of hexagonal Voronoi regions reveals that many localized patches of hexagonally arranged cones appear throughout each mosaic. Hexagonal arrays have interesting sampling properties. Natural scenes generally have circular symmetric power spectrums, and hexagonal sampling arrays provide the most efficient solution for sampling such signals.[12] The Nyquist limit for a normal hexagonal array is $(\sqrt{3}s)^{-1}$ along one axis and $(2s)^{-1}$ along the other, where s is the center-to-center spacing of cones. This means that a computation savings of approximately 13% over standard rectangular sampling is achieved. Our analysis indicated that nonhexagonal Voronois can make up less than 30% of a mosaic region near the fovea, but this percentage increases sharply to 50% to 60% at greater eccentricities. Cones at higher eccentricities are more randomly arranged and offer the visual system some protection from perceiving an aliased signal.[1,5,23] Earlier work by Yellott[1] showed from data taken from the peripheral retina of a monkey that the power spectrum

resembled that of a Poisson (random) distribution. Aliasing does not occur at the fovea because the Nyquist limit, due to higher cone densities, extends beyond frequencies that can pass through the optics of the eye. The cone packing arrangement decreases at higher eccentricities, so, even though cone densities are lower, aliasing is generally replaced by noise. We hope that the methods presented in this paper will encourage further quantitative analysis on the packing arrangement of rods and cones. Such analyses are essential for understanding how photoreceptors migrate into their permanent arrangement structure during development.

Cones are optical fibers oriented in the direction of the pupil center.[9] Reflected light emitted from the cone apertures are effectively point sources that generate an array of PSFs in AO images. For this reason, it is important to understand that not all PSFs correspond to cones in an AO image with questionable quality. When AO fails to resolve individual cones or when interference is present in the system, a single PSF-like intensity distribution may correspond to multiple cones or noise. When individual cones are resolvable, the cone labeling process can be done reliably using the proposed algorithm. This is demonstrated by evaluating the algorithm performance on a diverse set of AO images and comparing the outcomes to manually labeled results. Comparisons between the automated and the manual labeling methods resulted in similar outcomes. The extent to which the two methods as well as different human observers agreed is influenced by the appearance or quality of the analyzed image. This is especially true when performing Voronoi analyses, so packing arrangement studies should be done only with quality images where cones are well resolved. A low-quality image or an image with unique features resulted in lower levels of agreement between manual and automated methods. Equivalent comparisons between several experienced human observers often resulted in greater outcome variability, suggesting that a consistent automated solution may be more reliable in many cases.

Corresponding author K. Y. Li can be reached by e-mail at kaccie@berkeley.edu.

REFERENCES

1. J. I. Yellott, "Spectral consequences of photoreceptor sampling in the rhesus retina," Science **221**, 382–385 (1983).
2. D. R. Williams, "Topography of the foveal cone mosaic in the living human eye," Vision Res. **28**, 433–454 (1988).
3. C. A. Curcio and K. R. Sloan, "Packing geometry of human cone photoreceptors—variation with eccentricity and evidence for local anisotropy," Visual Neurosci. **9**, 169–180 (1992).
4. C. A. Curcio, K. R. Sloan, R. E. Kalina, and A. E. Hendrickson, "Human photoreceptor topography," J. Comp. Neurol. **292**, 497–523 (1990).
5. D. R. Williams and R. Collier, "Consequences of spatial sampling by a human photoreceptor mosaic," Science **221**, 385–387 (1983).
6. D. R. Williams and N. J. Coletta, "Cone spacing and the visual resolution limit," J. Opt. Soc. Am. A **4**, 1514–1523 (1987).
7. J. Z. Liang, D. R. Williams, and D. T. Miller, "Supernormal vision and high-resolution retinal imaging through adaptive optics," J. Opt. Soc. Am. A **14**, 2884–2892 (1997).
8. A. Roorda, A. B. Metha, P. Lennie, and D. R. Williams, "Packing arrangement of the three cone classes in primate retina," Vision Res. **41**, 1291–1306 (2001).
9. A. Roorda and D. R. Williams, "Optical fiber properties of individual human cones," J. Vision **2**, 404–412 (2002).
10. A. Roorda and K. Y. Li, "AO image processing," (2006), retrieved 2006, vision.berkeley.edu/roordalab/Kaccie/KaccieResearch.htm.
11. D. R. Williams, "Aliasing in human foveal vision," Vision Res. **25**, 195–205 (1985).
12. R. M. Mersereau, "Processing of hexagonally sampled two-dimensional signals," Proc. IEEE **67**, 930–949 (1979).
13. J. C. Christou, A. Roorda, and D. R. Williams, "Deconvolution of adaptive optics retinal images," J. Opt. Soc. Am. A **21**, 1393–1401 (2004).
14. R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using Matlab* (Pearson Education, 2004), pp. 65–193.
15. *Image Processing Toolbox, Users Guide, Version 4* (The MathWorks, Inc., 2003).
16. J. M. Enoch, "Optical properties of the retinal receptors," J. Opt. Soc. Am. **53**, 71–85 (1963).
17. J. S. Lim, "Finite impulse response filters," in *Two-Dimensional Signal and Image Processing*, A. V. Oppenheim, ed. (Prentice Hall, 1990), pp. 195–263.
18. R. C. Gonzalez and R. E. Woods, "Image enhancement in the frequency domain," in *Digital Image Processing*, 2nd ed. (Addison-Wesley, 2001), pp. 147–215.
19. M. B. Shapiro, S. J. Schein, and F. M. Demonasterio, "Regularity and structure of the spatial pattern of blue cones of macaque retina," J. Am. Stat. Assoc. **80**, 803–812 (1985).
20. A. Roorda, F. Romero-Borja, W. J. Donnelly, H. Queener, T. J. Hebert, and M. C. W. Campbell, "Adaptive optics scanning laser ophthalmoscopy," Opt. Express **10**, 405–412 (2002).
21. Y. H. Zhang, S. Poonja, and A. Roorda, "MEMS-based adaptive optics scanning laser ophthalmoscopy," Opt. Lett. **31**, 1268–1270 (2006).
22. J. Carroll, M. Neitz, H. Hofer, J. Neitz, and D. R. Williams, "Functional photoreceptor loss revealed with adaptive optics: an alternate cause of color blindness," Proc. Natl. Acad. Sci. U.S.A. **101**, 8461–8466 (2004).
23. R. L. Cook, "Stochastic sampling in computer graphics," ACM Trans. Graphics **5**, 51–72 (1986).