

ORIGINAL ARTICLES

Image Metrics for Predicting Subjective Image Quality

LI CHEN, PhD, BEN SINGER, PhD, ANTONIO GUIRAO, PhD, JASON PORTER, PhD, and
DAVID R. WILLIAMS, PhD

Center for Visual Science, University of Rochester, Rochester, New York (LC, BS, JP, DRW) and Departamento de Física, Universidad de Murcia, Murcia, Spain (AG)

ABSTRACT: *Purpose.* Despite the proliferation of wavefront sensors to characterize the optical quality of individual eyes, there is not yet an accurate way to determine from a wave aberration how severely it will impact the patient's vision. Some of the most commonly used metrics, such as RMS wavefront error and the Strehl ratio, predict subjective image quality poorly. Our goal is to establish a better metric to predict subjective image quality from the wave aberration. *Methods.* We describe three kinds of experiments designed to compare the effectiveness of different metrics in determining the subjective impact of the wave aberration. Subjects viewed a visual stimulus through a deformable mirror in an adaptive optics system that compensated for the subject's wave aberration. In the first experiment, we show that some Zernike modes such as spherical aberration and defocus interact strongly in determining subjective image quality. In the second experiment, the subject's wave aberration was replaced by the wave aberration corresponding to an individual Zernike mode. The subject then adjusted the coefficient of the Zernike mode to match the blur of a standard stimulus. In the third experiment, the subject viewed the same stimulus through the wave aberration of one of 59 different postoperative patients who had undergone LASIK and matched the blur by adjusting defocus. We then determined which among many image quality metrics best predicted these matching data. *Results.* RMS wavefront error was a poor predictor of the data, as was the Strehl ratio. *Conclusions.* The neural sharpness metric best described the subjective sharpness of images viewed through the wave aberrations of real eyes. This metric can provide a single number that describes the subjective impact of each patient's wave aberration and will also increase the accuracy of refraction estimates from wavefront-based autorefractors and phoropters. (*Optom Vis Sci* 2005;82:358–369)

Key Words: image quality, adaptive optics, visual performance, wave aberration, subjective refraction

Conventional phoropters are typically used to subjectively refract the eye. Subjective refractions, however, can be lengthy procedures. Objective autorefractors can automatically refract the eye with useful estimates of lower-order aberrations,^{1, 2} but they generally are not accurate and repeatable enough for a final refraction. Their errors can exceed 0.50 D when measuring large refractive errors,³ and the typical test–retest variability obtained with autorefraction and subjective refractions can be as high as ± 0.75 D.⁴

Compared with other techniques,^{5–9} the complete description of the eye's optics provided by objective wavefront sensors¹⁰ offers the possibility that such devices will completely replace tedious subjective refractions. A critical component required to achieve this goal is the development of an algorithm that can transform the wave aberration into optimum values for sphere, cylinder, and axis.

Accurate, automated refraction depends on replacing the patient with an algorithm that can predict the best refraction as well or better than the patient him- or herself.

Mouroulis et al.¹¹ examined subjective contrast and edge sharpness discrimination in the presence of different amounts of only two aberrations, coma and astigmatism. They simulated aberrations on a screen with targets of different contrast or edge sharpness. They measured the correlation between subjective and objective image quality and found that the Strehl ratio and wavefront variance show low correlation with subjective performance, whereas the modulation transfer function across the frequency range of interest had a good correlation in the presence of asymmetric aberrations.

Guirao and Williams¹² showed that higher-order aberrations can influence the refraction of the eye. They proposed that the

most insightful and useful metrics will incorporate known stages in visual processing, favoring metrics that include a computational step corresponding to retinal image formation rather than an arbitrary manipulation of the wave aberration in the pupil plane. Their experimental data showed that, in five eyes, image plane metrics produced a more accurate estimate of the subjective refraction than pupil plane metrics such as RMS wavefront error.

Applegate and colleagues^{13, 14} developed a psychophysical methodology for studying the visual influence of higher-order aberrations based on simulation. They aberrated acuity charts by convolving the chart with the point spread functions corresponding to individual Zernike modes. These simulations were then viewed through a small pupil to approximate the retinal images that would have been formed had these aberrations existed in the eye's pupil plane. Their experiments have shown that individual Zernike modes differ greatly in their influence on visual acuity. They also showed that Zernike modes can interact strongly in their effects on visual acuity.¹⁵ Williams and colleagues^{16–18} obtained similar results as those of Applegate using a deformable mirror to produce aberrations instead of convolution. Investigators from Indiana University^{19–21} studied subjective optical quality metrics for visual acuity. They found that wavefront variance was the best and the Strehl ratio was the worst metric for quantifying visual performance. They also predicted refractions from patients' wave aberrations using different image metrics.²²

We have adopted a somewhat different approach that complements the methods discussed here. First, we used adaptive optics to manipulate the aberrations in the eye. Second, we used a method based on the subject's subjective image quality when viewing stimuli through different wave aberrations rather than visual acuity or refraction. In this article, we introduce a psychophysical procedure that takes advantage of adaptive optics and allowed us to compare different metrics with the subjective image quality of individual

subjects. This procedure led to a neural sharpness metric, which can be implemented in a wavefront sensor to tell the clinician what kind of subjective image quality the patient has from the wave aberration alone.

METHODS

Subjects

Measurements were performed in the right eyes of six subjects. Subject consent was obtained according to the Declaration of Helsinki. During the measurement, the subject's head was stabilized with a bite bar, and the subject's pupil was dilated with tropicamide (1%).

Matching Experiments with Adaptive Optics

Usually adaptive optics is used to correct the eye's wave aberration for retinal imaging or for delivering a sharp visual stimulus to the retina. However, adaptive optics can also be used not only to correct the eye's wave aberration, but also to generate specific wave aberrations in the eye. Figure 1 shows the setup of the adaptive optics system for the matching experiment. This adaptive optics system²³ uses a Shack-Hartmann wavefront sensor, placed conjugate with the subject's pupil, to make measurements of the eye's wave aberration at 30 Hz. This Shack-Hartmann wavefront sensor has 177 lenslets arranged in a square array, which can measure aberrations for a 6-mm pupil up to the tenth order, corresponding to 63 Zernike modes. The wave aberration measurements were made at a wavelength of 810 nm. A deformable mirror with 97 PMN actuators, also conjugate with the subject's pupil plane, is used to correct the subject's wave aberration based on the measurements from the Shack-Hartmann wavefront sensor. In our experiments, in addition to removing the higher-order aberrations in the

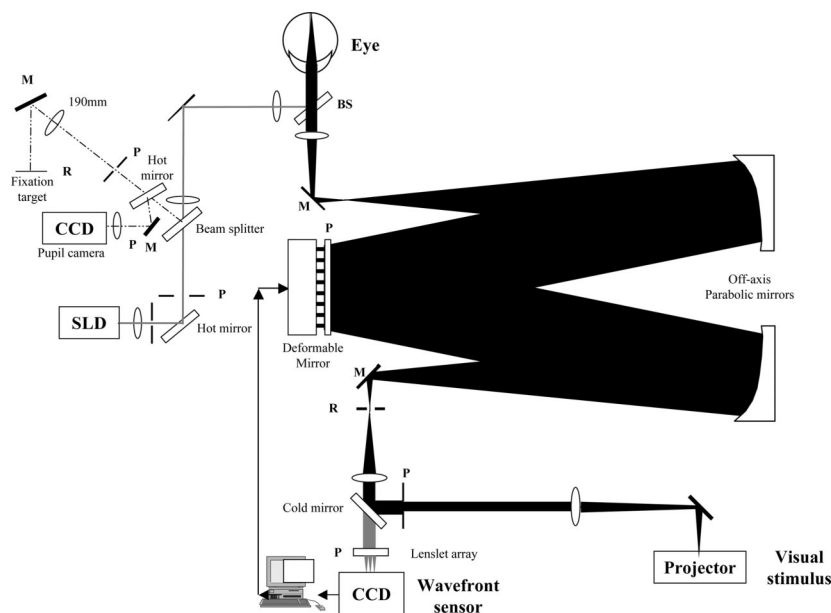


FIGURE 1.

Schematic diagram of the adaptive optics system for the blur-matching experiments. The black path is used to deliver the stimulus image from the projector to the eye for the visual experiments. The gray path, overlapped by the black path from the beam splitter (BS) before the eye to the cold mirror, represents the infrared light used for the wave aberration measurement and correction.

eye on each trial, the deformable mirror also acted as an aberration generator to blur the subject's vision. This process is done in a closed-loop fashion so that the AO system is working at 30 Hz to fix the aberrations we want to their desired values.

Subjects viewed a test field subtending 1° in 550-nm monochromatic light at an eccentricity of 1° through a 6-mm artificial pupil in the adaptive optics system. A trial lens was placed in the pupil plane (denoted with "P") between the cold mirror and the projector to compensate for the chromatic aberration difference between the wavefront sensing wavelength of 810 nm and the stimulus presentation at 550 nm. The stimulus, shown in Figure 2, contained a binary noise pattern with sharp edges at random orientations. The binary noise pattern was produced from a uniform noise distribution spatially filtered with an annular filter in the frequency domain. The inner radius of the annular filter was 6.4 cycles/deg and the outer radius was 9.6 cycles/deg. This gray-level image was transformed to a binary image with only black and white values by setting all gray levels above 50% to white and those below to black. A Gaussian function smoothed the edge of the field. A different noise pattern was used to generate a new stimulus on each trial so that edges at all orientations were presented over the course of the experiment. The retinal illuminance of the stimulus display is 390 Trolands measured by IL17000 Radiometer.

Experiment 1. We used adaptive optics to demonstrate the interactions between pairs of aberrations in real eyes. On each trial, a deformable mirror was used to remove all of the eye's monochromatic aberrations while simultaneously presenting one aberration of fixed value to the eye. The subject then maximized the subjective image quality of the stimulus by either increasing or decreasing the value of a different aberration, also generated by the deformable mirror.

Experiment 2. This experiment measured the variation in subjective blur produced by different Zernike modes by matching the blur produced by a standard aberration. Each trial consisted of two 500-ms intervals during which the binary noise pattern was

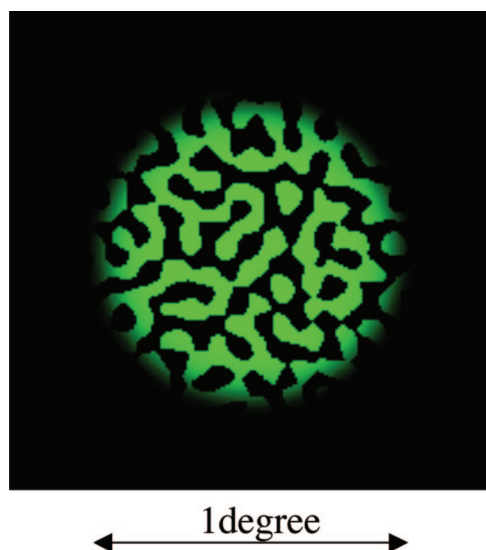


FIGURE 2.

Stimulus in monochromatic light of 1° viewed through an adaptive optics system for psychophysical experiments. The color version of this figure is available online only at www.optvissci.com.

displayed. During the first interval, the binary noise stimulus was viewed with a wave aberration corresponding to a single Zernike mode. During the second interval, the stimulus was viewed with a wave aberration corresponding to a standard aberration. The standard aberration was created by combining all 18 Zernike modes from the second through fifth order, each having an amplitude of $0.1 \mu\text{m}$. The standard aberration was chosen to include a large number of aberrations of modest amplitude but was otherwise arbitrarily selected. At the time, this experiment was performed, our AO system used Noll Zernike polynomials²⁴ to describe the wave aberration, in which case all the amplitudes of the standard aberration were positive. After converting the standard aberration into the OSA standard for Zernike polynomials,²⁵ some of the Zernike terms in the standard aberrations were positive and some of them were negative as shown in Figure 3. The test aberration consisted of a single Zernike mode whose coefficient was adjusted by the subject to produce the same subjective blur in the stimulus as the standard aberration. In between stimulus presentations, the subject viewed a uniform field for 300 ms when the adaptive optics system was generating the desired wave aberration in closed loop. The wavefront sensor laser beacon was positioned at the edge of the field and turned off during the stimulus presentation to avoid disrupting sensitivity to blur. Each mode had two matching values, one positive and one negative. The matching measurement for each mode was performed eight times, four times to match the positive value and four times to match the negative value. The final matching value of one mode to the standard aberration is the average of the absolute values of these eight matches. This experiment was then repeated for 18 Zernike modes corresponding to those of the second through fifth orders.

Experiment 3. A similar matching procedure was used to measure the subjective blur produced by wave aberrations measured from real patients. The key differences were that the standard aberration was replaced by one of 59 wave aberrations from post-operative patients who had undergone LASIK, and defocus was always used as the test aberration for blur matching. Each LASIK patient's wave aberration, containing 18 Zernike modes, was measured with a wavefront sensor. The defocus of each patient's wave aberration was set to zero in the standard aberration. For each match, the adaptive optics system replaced the subject's wave aberration with that of one of the postLASIK patients. Figure 4 shows a sample of some postLASIK aberrations in the top row along with the same aberrations generated in the eye of one of the subjects with adaptive optics. The similarity between the corresponding top and bottom images shows that adaptive optics can do a good job of generating many different wave aberrations. The residual aberrations between the patient's wave aberration and the wave aberration generated by adaptive optics are $<0.1 \mu\text{m}$.

Figure 5 shows the matching procedure in which the subject changed the value of defocus to match the blur caused by the postLASIK patient's aberrations. Defocus was chosen as the test aberration to quantify the blur caused by the patient's wave aberration because its magnitude can be expressed in diopters as well as microns of RMS wavefront error. The matching value of defocus to each patient's aberration was also measured four times at the positive value and four times at the negative value. The final matching value was the average of the absolute values of these eight

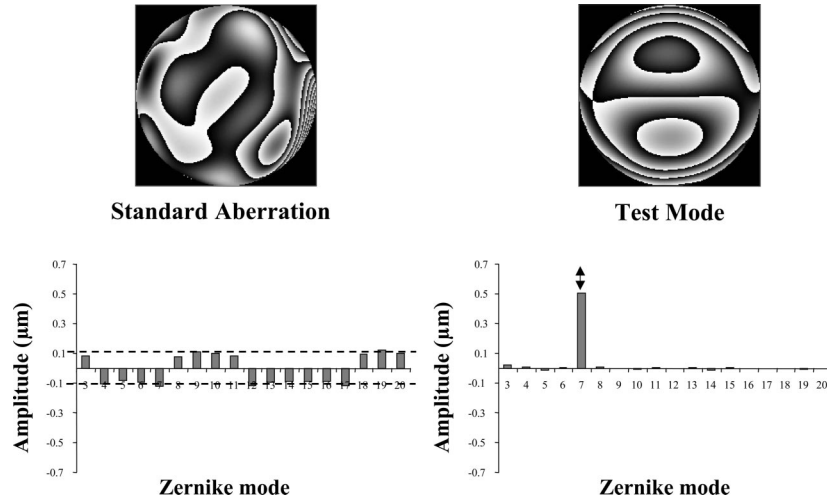


FIGURE 3.

Experimental procedure of blur matching with an individual Zernike mode. The subject alternately viewed the stimulus through a standard wave aberration (top left) and a single testing Zernike mode (top right). The amplitudes of the aberrations are shown below each pattern. In this example, the test mode corresponds to vertical coma ($j = 7$).

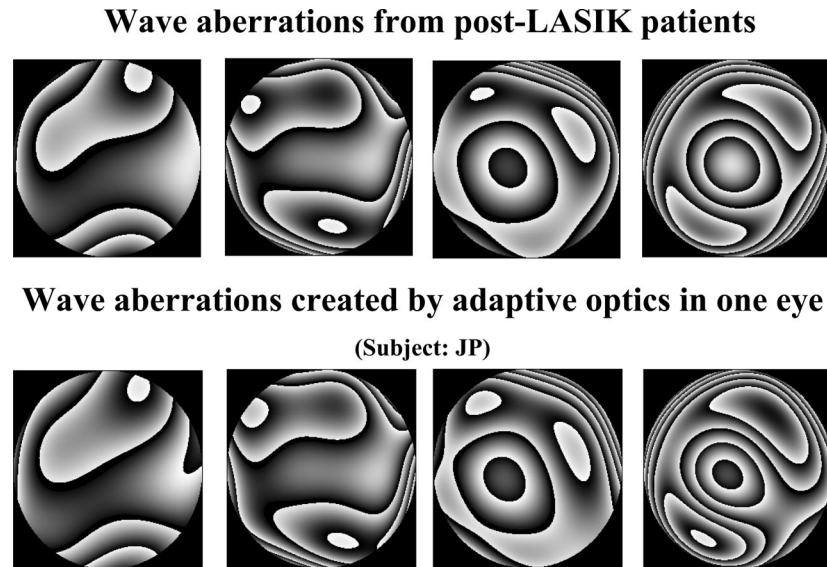


FIGURE 4.

Samples of wave aberrations from postoperative patients who had undergone LASIK generated in one real eye with adaptive optics. At the top of this figure are examples of wave aberrations measured from four postLASIK patients. The bottom row shows these same wave aberrations, created in the eye of a single subject with adaptive optics.

measurements. We tested several different metrics, listed subsequently, to determine which best described the matching data.

Definitions of Metrics

Wavefront sensors normally express the eye’s wave aberration, WA, using a set of normalized Zernike polynomials, Z_n^m , as recommended by the Optical Society of America,²⁵

$$WA = \sum_{n, m} c_n^m Z_n^m \tag{1}$$

where c_n^m are the Zernike coefficients.

RMS Wavefront Error. Because the Zernike polynomials are orthonormal over a unit pupil, the RMS wavefront error, defined

as the standard deviation of the wave aberration surface, can be calculated as the square root of the sum of squared Zernike coefficients.

$$RMS = \sqrt{\sum_{n, m} (c_n^m)^2} \tag{2}$$

Strehl Ratio. The Strehl ratio can also be used to quantify the severity of a given wave aberration. To compute the Strehl ratio, the generalized pupil function is first calculated from the wave aberration as

$$PupilFunction = pe^{i(2\pi/\lambda)(WA)} \tag{3}$$

where p is a circular pupil aperture with a unit amplitude transmittance. The point spread function (PSF) is calculated as the

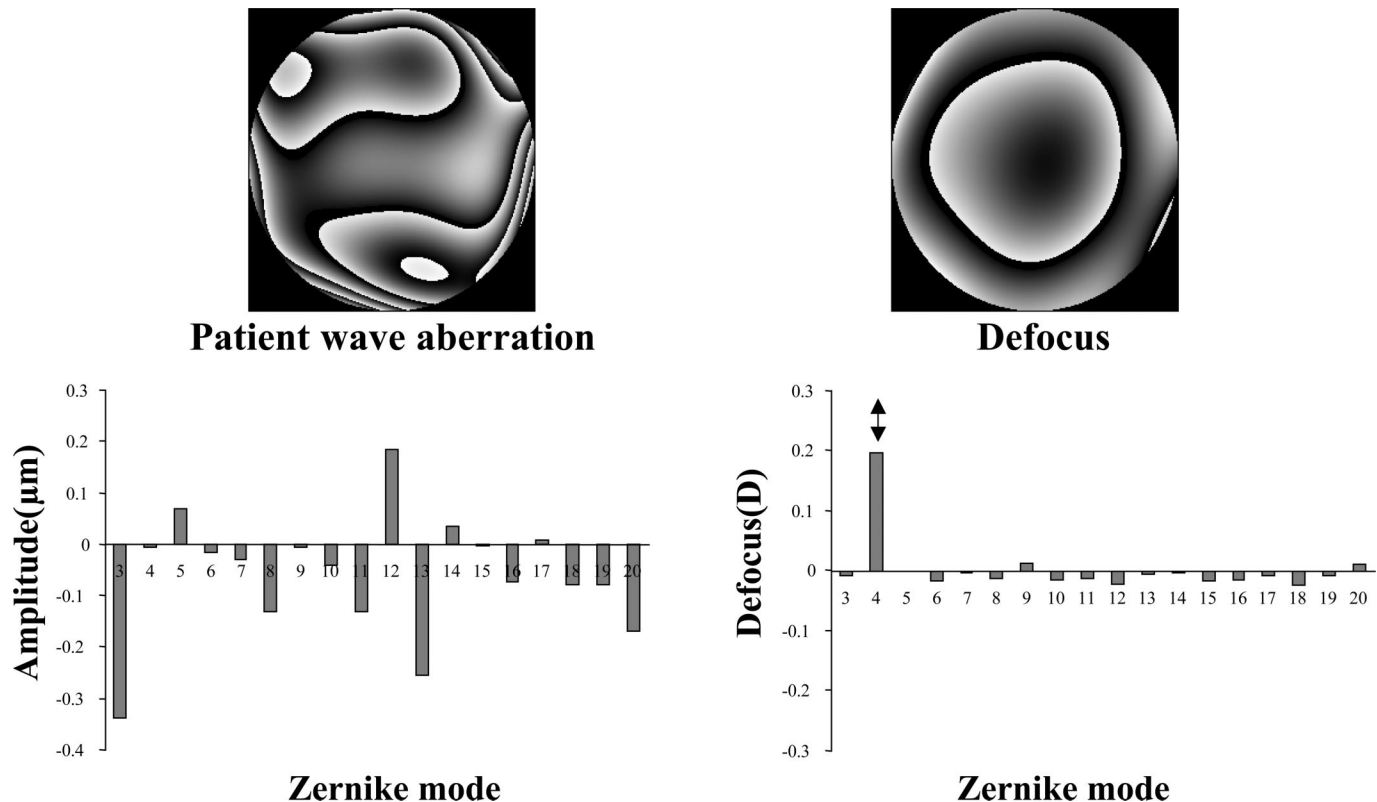


FIGURE 5. Experimental procedure of blur matching with wave aberrations from postoperative patients who had undergone LASIK. The subject alternatively viewed the stimulus through the wave aberration from a postLASIK eye generated by the adaptive optics system (top left) and through Zernike defocus (top right). The values of the constituent aberrations for one postLASIK eye are shown in the bottom left. The amplitude of defocus could be adjusted by the subject to match the blur created by the postLASIK wave aberration.

squared modulus of the Fourier transform (FT) of the generalized pupil function:

$$PSF = |FT(PupilFunction)|^2 \tag{4}$$

The ratio of the peak intensity of the eye’s PSF in the presence of aberrations to the peak intensity of an aberration-free PSF for the same pupil size is the Strehl ratio,

$$\text{Strehl Ratio} = \frac{Max(PSF_{intensity})_{Aberrated\ eye}}{Max(PSF_{intensity})_{Aberration-free}} \tag{5}$$

The Intensity Variance of the Point Spread Function. The intensity of the PSF is calculated as the average value of the squared PSF minus the squared value of the average PSF:

$$\text{var}(PSF) = \overline{PSF^2} - (\overline{PSF})^2 \tag{6}$$

The entropy of the PSF. The entropy is mathematically calculated as follows:

$$\text{Entropy} = - \sum_{x,y} PSF(x,y) \times \log[PSF(x,y)] \tag{7}$$

where (x,y) indicates position (in pixels) and log is the decimal logarithm. This metric is a measure of the spatial variance of the PSF, i.e., a measurement of how the energy is distributed in the retinal image. An aberration-free PSF has the minimum amount of

entropy possible. Aberrations increase the entropy from this minimal value.

AreaMTF. AreaMTF is the area under the MTF between 0 and 60 cycles/deg. The MTF is the modulus of the Fourier transform of the PSF (or, alternatively, is the autocorrelation of the generalized pupil function):

$$MTF = |FT(PSF)| \tag{8}$$

Neural Sharpness. Neural sharpness (NS) captures the effectiveness of a PSF for simulating the neural portion of the visual system. It is the maximum value of the convolution of the eye’s PSF and a spatial sensitivity function with a Gaussian profile that represents the neural visual system.

$$NS = Max(PSF(x,y) * g(x,y)) \tag{9}$$

where,

$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{[(x-\mu_x)^2 + (y-\mu_y)]^2}{2\sigma^2}} \tag{10}$$

is a bivariate-Gaussian neural weighting function with a standard deviation (σ) of 1 min of arc, selected to optimize the performance of this metric for the dataset.

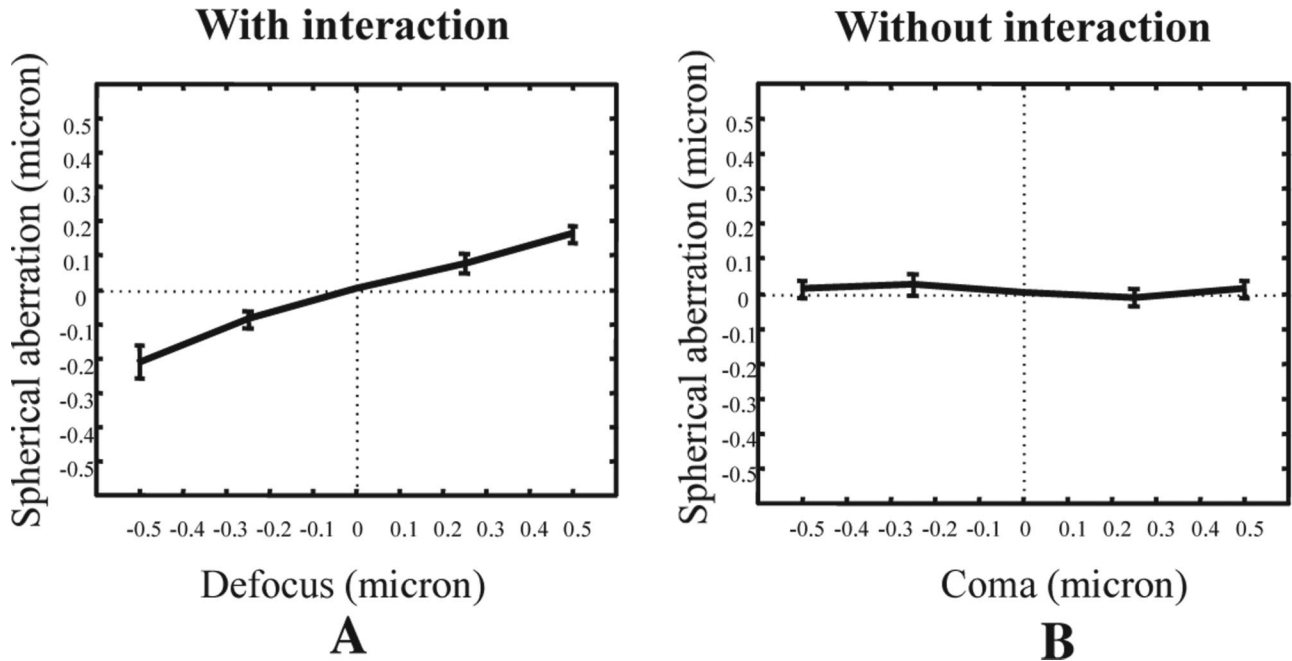


FIGURE 6.

Mode interactions between spherical aberration and defocus from two subjects (A) and between spherical aberration and coma (B) measured in a real eye with adaptive optics. There was no interaction between spherical aberration and coma, but spherical aberration and defocus did interact.

RESULTS

In the first experiment, we used adaptive optics to demonstrate that defocus and spherical aberration can sometimes improve or degrade image quality more than would be expected from the individual components. Figure 6A shows that subjective image quality was maximized in two subjects initially presented with 0.5 μm of defocus by adding an average of 0.18 μm of spherical aberration (6-mm pupil). When initially presented with -0.5 μm of defocus, subjects added -0.19 μm of spherical aberration, on average, to optimize subjective image quality. A similar relationship, which is not shown here, was found for coma and secondary coma. On the other hand, we also found pairs of aberrations that do not constructively interact to enhance image quality such as coma and spherical aberration, shown in Figure 6B, and astigmatism and coma. As illustrated in Figure 6, subjective image quality became worse when subjects added any amount of spherical aberration to a fixed value of coma. Retinal blur was minimized when no spherical aberration was present, indicating a lack of interaction between these two particular modes.

Besides interactions between defocus and spherical aberration, Figure 7 shows that modes two radial orders apart with the same sign and angular frequency can be combined to enhance subjective image quality compared with loading the same magnitude of RMS error into either component individually.

Figure 8 gives the psychophysical mode matching value for individual Zernike modes. One can see that aberrations at the edges have larger matching values than those at the center. This indicates that Zernike modes at the center of each order are more blurred than those along the flanks. These results qualitatively agree with the simulation in Figure 9 showing that equal amplitudes of RMS wavefront error in different modes produce large differences in

subjective blur. These data are in agreement with those of Applegate et al.,¹³ who performed a related experiment in which a logarithm of the minimum angle of resolution acuity chart was convolved with the point spread functions of individual Zernike modes. Figure 7 also shows that aberrations at the edges of the Zernike pyramid contain central regions that are relatively flat compared with those in the center of the pyramid.

Figure 10 shows the fitting results of the mode matching data when using the RMS wavefront error, the Strehl ratio, the entropy of the PSF, the intensity variance of the PSF, the AreaMTF, and the neural sharpness metrics. The predicted value of each Zernike mode is adjusted iteratively until its calculated metric value matches the metric value of the standard aberration. The correlation coefficient between the psychophysical mode matching value and the predicted value from different metrics and different subjects is shown in Table 1. This is also shown in Figure 10. For the fitting of each individual subject's data, the neural sharpness metric did the best for three subjects, whereas Entropy did the best for another three subjects. Entropy showed large individual differences for reasons that are not clear. Averaged across the six subjects, the correlation coefficient of the neural sharpness metric was highest out of all the metrics, although PSF variance also had a high average correlation coefficient.

We also used the RMS wavefront error, the Strehl ratio, PSF variance, entropy, AreaMTF, and the neural sharpness metric to fit the matching results from six subjects for the subjective blur created by real aberrations from 59 postoperative patients who had undergone LASIK. Figure 11 shows the correlation between the blur matching value and the predicted value from six subjects for each metric. The correlation coefficient between the psychophysical blur matching value and the predicted value from different

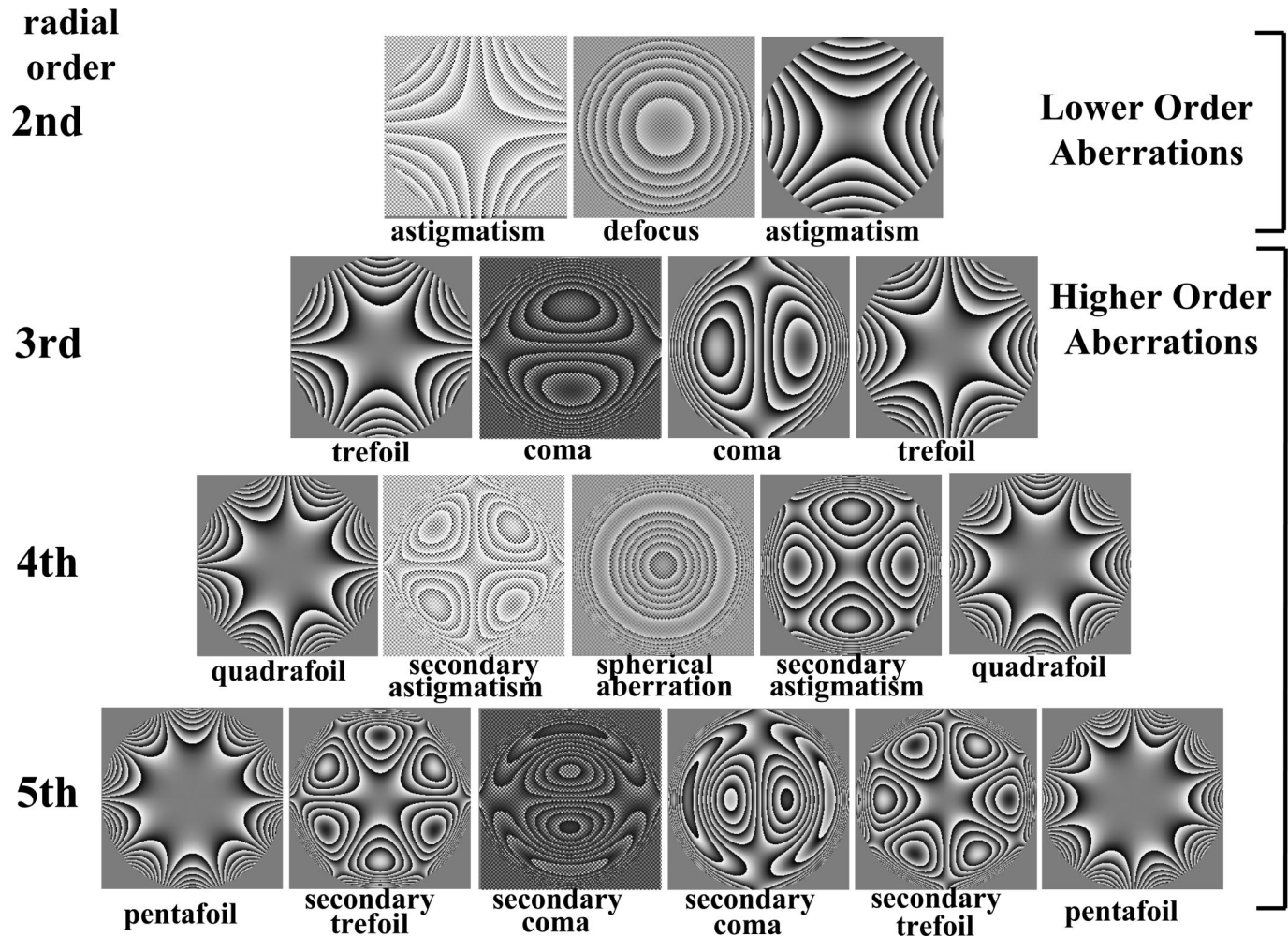


FIGURE 7. Examples of Zernike mode interaction pairs are illustrated between astigmatism and secondary astigmatism, defocus and spherical aberration, and coma and secondary coma. Mode interactions can occur between modes two radial orders apart with the same sign and angular frequency.

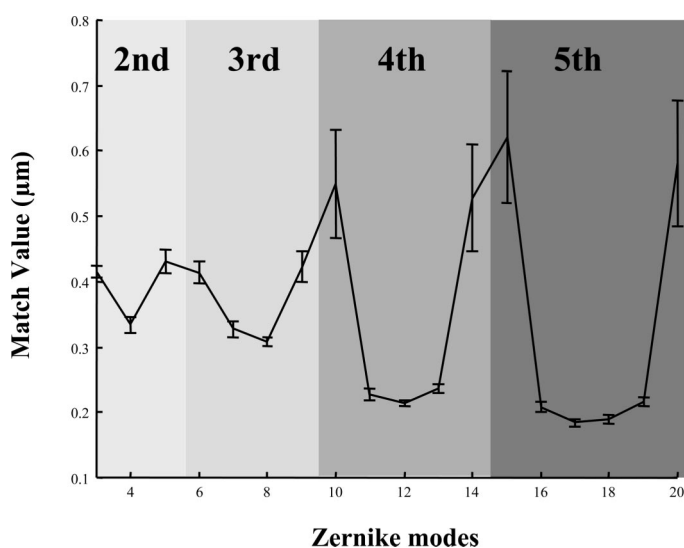


FIGURE 8. Experimental results obtained by measuring the impact of individual Zernike modes on subjective image quality using the mode-matching technique. Aberrations at the center of each order have smaller matching values, indicating that they are more potent at generating blur.

metrics and different subjects is shown in Table 2. All the correlation coefficients are significant having p values < 0.02. The correlation coefficient of the neural sharpness metric was also highest of all the metrics when examining the data from the average of six subjects. For the fitting of individual subject's data, the neural sharpness metric did the best for three subjects, whereas entropy did the best for another three subjects.

DISCUSSION

Applegate et al.^{13–15} explored different image-quality metrics by identifying which metrics best predicted logarithm of the minimum angle of resolution acuity or letters convolved by the PSF of individual Zernike modes. Our study complements this study by focusing on subjective image quality rather than acuity. Although subjective image quality and acuity are presumably correlated, the effect of an aberration on the subjective impression of image quality need not correspond to its effect on acuity. We used adaptive optics not only to correct the eye's wave aberration as it is usually used, but also to generate specific wave aberrations in the eye. This psychophysical procedure allows us to compare different metrics with the subjective image quality of individual subjects to identify

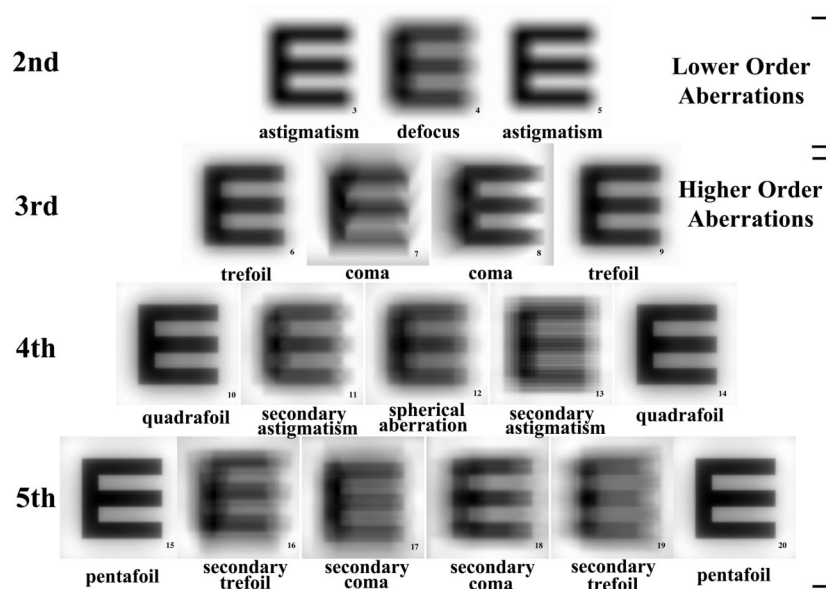


FIGURE 9. Images simulated using different aberrations consisting of individual Zernike modes with the same RMS. All Zernike modes are not equally effective at blurring image quality.

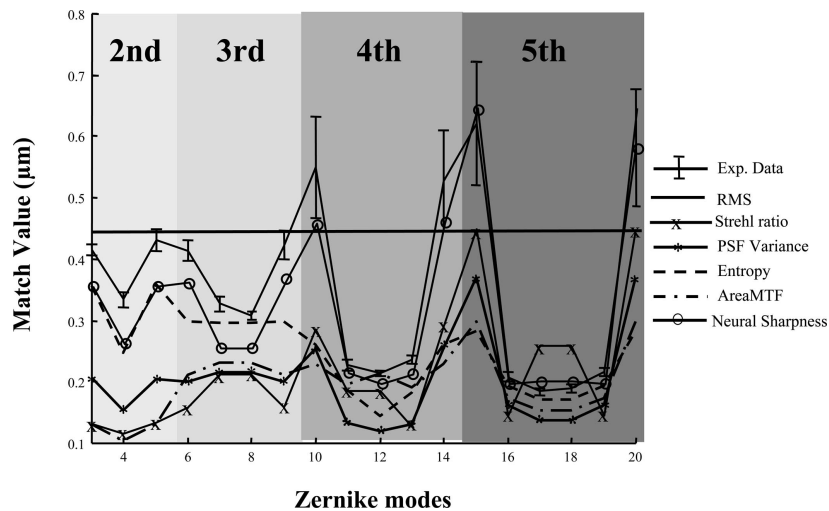


FIGURE 10. Subjective image quality measured and predicted by different metrics. (The error bar represents one standard error of the mean measurement from six subjects.)

a robust metric to predict the subjective impact of the eye’s wave aberration. The adaptive optics method has the advantage over the simulation method in that, in principle, it is possible to exactly recreate the light distribution on the retina that would result from a particular wave aberration. The use of a small pupil in the simulation method means that the light distribution caused by a particular wave aberration can only be approximated. Both methods are subject to inaccuracies in practice such as the ability to create the appropriate light distribution with display technology in the simulation and inaccuracies in deformable mirror compensation with the adaptive optics approach.

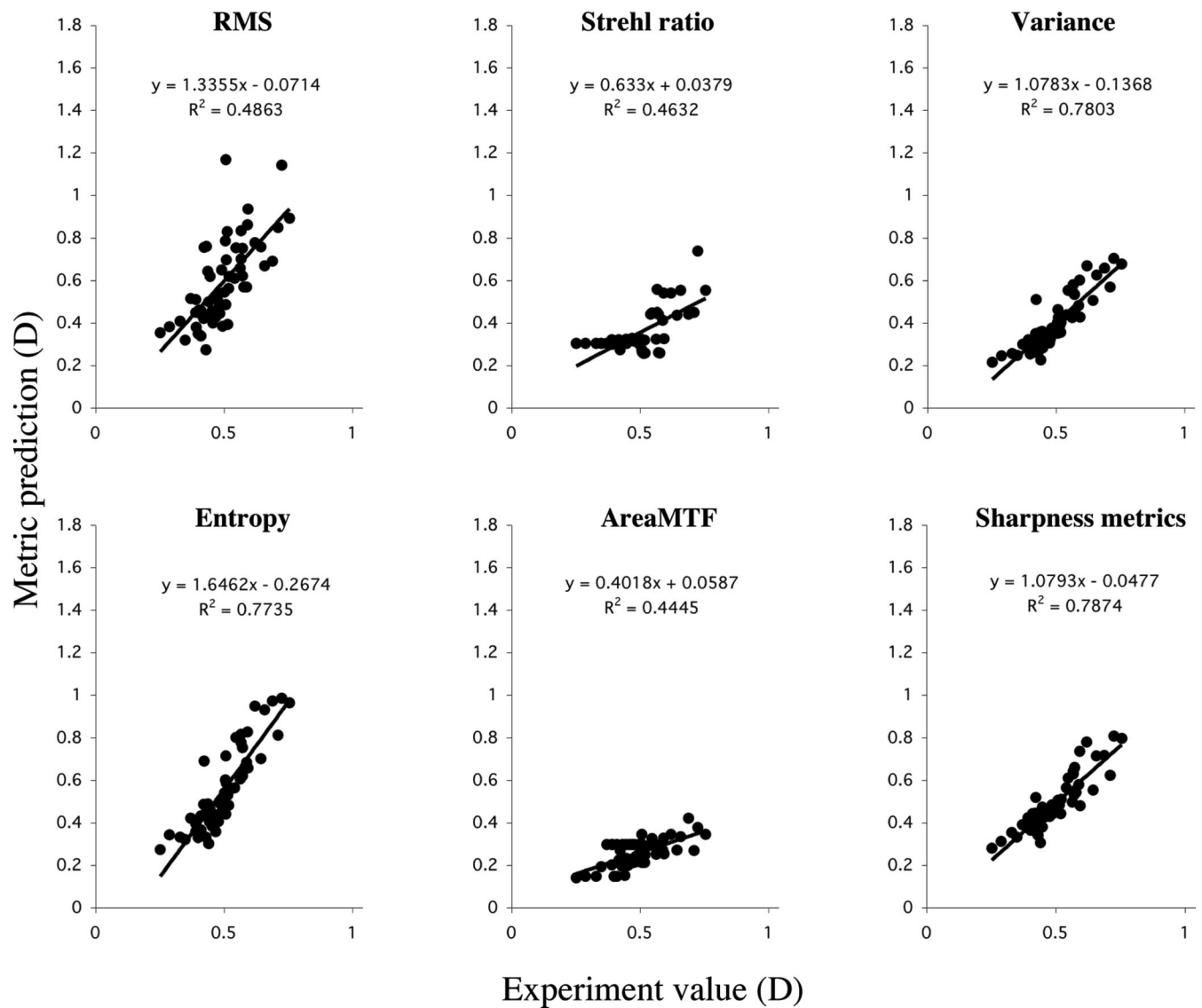
The RMS wavefront error provides an estimate of the flatness of the wave aberration. Most wavefront sensors today rely on the RMS wavefront error to predict the severity of the wave aberration.

For a perfect wavefront, RMS wavefront error is zero and increases as the wavefront becomes increasingly aberrated. From the RMS wavefront error, clinicians can roughly assess the quality of the patient’s vision. Unfortunately, RMS wavefront error is not an especially useful metric for describing the subjective impact of the eye’s wave aberration. Figure 12 illustrates two retinal images of the letter E for one hypothetical eye with only spherical aberration (Fig. 12A), and a second eye with the same amount of spherical aberration plus an additional amount of defocus (Fig. 12B). Although the RMS wavefront error is higher in the second eye (Fig. 12B), image quality is better than that in Figure 12A. This shows that higher-order aberrations such as spherical aberration in this case can influence the optimum refraction. The reason for this is that Zernike modes, as measured by the wavefront sensor at the

TABLE 1.

The correlation coefficient between the psychophysical blur matching value and the predicted value from different metrics and different subjects for predicting aberration of 18 individual Zernike modes

Subject	RMS	Strehl ratio	PSF variance	Entropy	AreaMTF	Neural sharpness
LC	0	0.0845	0.5225	0.6466	0.0767	0.6439
HH	0	0.6098	0.8381	0.1815	0.4823	0.9205
ND	0	0.0421	0.4401	0.8195	0.0497	0.4889
YY	0	0.476	0.8625	0.3595	0.4498	0.9451
LN	0	0.0173	0.3704	0.8442	0.0177	0.4658
YJ	0	0.2094	0.6693	0.5864	0.208	0.7927
Average	0	0.362	0.8109	0.4669	0.208	0.9171

**FIGURE 11.**

Correlation between the measured experimental blur matching value of defocus and the value predicted by six different metrics. The neural sharpness metric yielded the highest correlation coefficient.

patient's pupil, can interact strongly with each other to affect image quality at the patient's retina.¹⁵ Their effects at the retina do not add together in a simple way. Figure 6 shows the similar qualitative effects in normal viewing. Experiment 1 demonstrated that the

complexities of the interactions between Zernike modes for predicting subjective blur means that RMS wavefront error is not a productive avenue for deriving a metric of subjective image quality.

Typically, eyes with better optical quality will have larger Strehl

TABLE 2.

The correlation coefficient between the psychophysical blur matching value and the predicted value from different metrics and different subjects for predicting aberration of 59 postoperative LASIK patients

Subject	RMS	Strehl ratio	PSF variance	Entropy	AreaMTF	Neural sharpness
ND	0.1488	0.1059	0.3769	0.4012	0.3624	0.3662
JC	0.3885	0.3579	0.6529	0.6348	0.3968	0.6583
JP	0.4508	0.6042	0.7118	0.6835	0.4146	0.7664
LC	0.3541	0.3482	0.5622	0.5513	0.2686	0.5703
LN	0.4472	0.4711	0.7435	0.7501	0.4057	0.7392
AA	0.3501	0.1473	0.3706	0.3714	0.1455	0.3383
Average	0.4863	0.4632	0.7803	0.7735	0.4445	0.7874

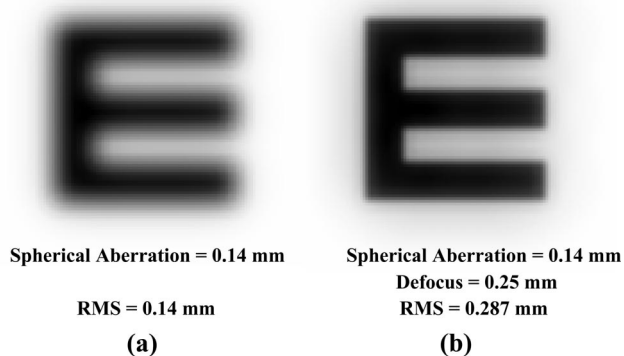


FIGURE 12.

Simulated retinal images of the letter E for two hypothetical eyes: (A) an eye with spherical aberration and (B) an eye with the same amount of spherical aberration as in (A), but also with an additional amount of defocus. RMS wavefront error does not necessarily directly reflect subjective image quality, because image quality is better in (B) when the RMS is largest.

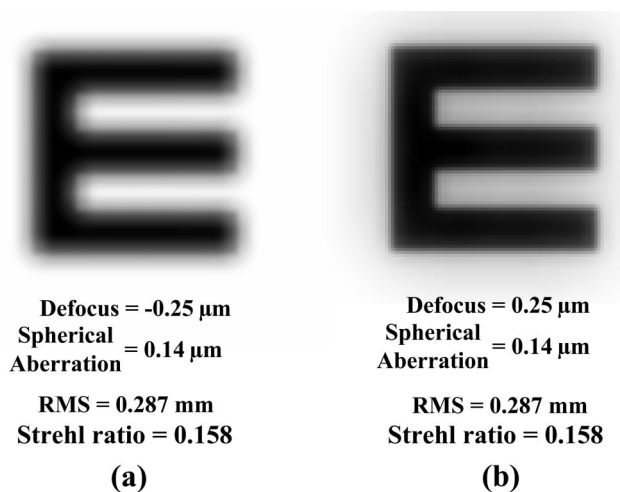


FIGURE 13.

Strehl ratio does not necessarily and adequately reflect subjective image quality. Both the RMS and Strehl ratio are identical for the two hypothetically aberrated eyes in (A) and (B). However, subjective image quality is better for (B).

ratios and more compact PSFs than those with poor optics. However, it is easy to find examples in which the Strehl ratio is not a good metric for estimating the subjective impact of the wave aberration.

Figure 13 simulates two Es that are blurred by different wave aberrations and yield identical Strehl ratios. However, image quality is better in Figure 13B than in Figure 13A.

Neither the RMS wavefront error nor the Strehl ratio predicts consistently and accurately subjective image quality for individual Zernike modes and real aberrations from 59 postoperative patients who had undergone LASIK. The predictions from the entropy of the PSF, the intensity variance of the PSF, and the AreaMTF metrics are not good either. Therefore, we needed to develop a new metric that could better predict the effect of the wave aberration on image quality and transform it into an optimum refraction for the patient. The neural sharpness metric combines the eye’s optics and neural blur, better predicting subjective image quality. Because our goal is to replace the patient’s visual system in refracting the eye, it makes sense that our algorithm should include the basic stages that we know exist in the visual system, that is, an image formation step followed by neural processing. Of course, there are many other metrics that one might investigate,¹⁸ but this one captures the fundamental features that we believe a good metric should have. There is great practical value in the simplicity of a single-value metric for subjective image quality. For example, refraction can be optimized through a search that maximizes a single number. However, blur is not a unity perceptual experience, a point that the subjects in our experiments repeatedly made to us because of the difficulty they had in performing the blur matching task as a result of the sometimes quite different appearance of the stimuli at the match point. A multivariate scheme would more accurately describe the subjective effect of a given wave aberration. For example, our experience with different wave aberrations suggests that some of them reduce the overall contrast of the image while keeping edges crisp. Others keep contrast high but sharp edges become fuzzy. Still other aberrations, especially odd-order aberrations like coma, produce asymmetry in images such as flaring away from the object in one direction. This study has focused on the development of a single metric because a single metric is required in applications such as the use of a wavefront sensor to predict the best refraction for all viewing conditions and visual tasks. However, multivalued metrics have the advantage of providing a more detailed description of the subjective experience of the observer.

It would be possible for the subject to undertake matching experiments with adaptive optics in which sharpness, contrast, and symmetry are each assessed, so that the importance of each of these subjective dimensions for subjective image quality could be determined. Retinal image quality also depends on the spectral nature of

light because of chromatic aberration in the eye. The image metrics can be generalized to include chromatic aberration, and this is important because chromatic aberration interacts with the eye's monochromatic aberrations to determine overall image quality.

Schwiegerling²⁸ developed a visual performance metric by weighting the pupil function of the eye with a Gaussian filter and demonstrated that the variance of the Gaussian-weighted wavefront error is well correlated with visual performance. Guirao and Williams¹² argued that image plane metrics are superior to pupil plane metrics. This was true in the limited sample of metrics they used. Our approach is an extension of that philosophy, incorporating a neural component because metrics for subjective image quality might also need to incorporate the fact that neural processing is plastic, changing its performance depending on the wave aberration it currently sees the world through. Webster et al.²⁶ found that the neural visual system adapts to prismatic distortions, contrast, or blur. A recent study by Artal et al.²⁷ reveals that this plasticity extends to the monochromatic aberrations of the eye as well. Artal used the Rochester Adaptive Optics System to replace the wave aberration from a subject, either in its original orientation or rotated version. Subjects viewing the world through their own wave aberration reported that it was much sharper than when the wave aberration was rotated. Despite the fact that the rotation only changes the orientation of the aberrations and not the objective amount of retinal blur, the subjective blur changed dramatically. These observations support clinical wisdom that patients will often reject astigmatic corrections that improve objective image quality but cause too large a departure from their normal experience of the world. We do not yet know to what extent neural factors have influenced the subjective judgments of our subjects in the blur-matching paradigm. For example, one might expect that subjects would perceive less blur in LASIK patient wave aberrations that were more similar to their own, a potential source of bias in our experiments. Further analysis will be required to sort out the effect of possible neural factors on the blur-matching paradigm.

What is especially lacking is a clear understanding of the limitations of a single value metric and whether it is generalizable across tasks. It also remains to be seen to what extent such a metric is generalizable across different visual stimuli and viewing conditions. The binary noise stimulus we used is very similar to letter targets more commonly used in acuity tasks because it contains sharp edges at high contrast. We anticipate, although we have not shown empirically, that our results would not have been greatly influenced by the use of letters instead. We chose the binary noise stimulus because, unlike letters, it does not contain any bias toward edges at particular orientations.

CONCLUSIONS

The neural sharpness metric is effective at describing the subjective sharpness of images viewed through the wave aberrations of both individual Zernike modes and real eyes. The metric is biologically plausible and fast to compute. It does not require the interpretation of individual Zernike modes and can be incorporated in the software of current wavefront sensors. This metric shows an example that such a metric, which combines the eye's optics and neural blur, will provide clinicians with a description of the subjective impact of each patient's wave aberration and will also in-

crease the accuracy of refraction estimates from wavefront-based autorefractors and phoropters. Although the neural sharpness metric performed the best on average in this study, there is no doubt that improvements in metrics can be achieved.

ACKNOWLEDGMENTS

This work has been supported in part by the National Science Foundation Science and Technology Center for Adaptive Optics, managed by the University of California at Santa Cruz under cooperative agreement no. AST-9876783 and through the National Eye Institute, grant no. 8R1EY04367D-18 and grant no. 2R01EY05901-12. Also supported in part by Ministerio de Educación y Ciencia grant FIS04-01018 (to AG). We thank Ian Cox, Scott MacRae, Yasuki Yamauchi, Lourdes Llorence, and Lana Nagy for helpful discussion and assistance.

Received July 12, 2004; accepted January 6, 2005.

REFERENCES

1. Wesemann W, Rasso B. Automatic infrared refractors—a comparative study. *Am J Optom Physiol Opt* 1987;64:627–38.
2. Goss DA, Grosvenor T. Reliability of refraction—a literature review. *J Am Optom Assoc* 1996;67:619–30.
3. Bullimore MA, Fusaro RE, Adams CW. The repeatability of automated and clinician refraction. *Optom Vis Sci* 1998;75:617–22.
4. Cheng X, Himebaugh NL, Kollbaum PS, et al. Test–retest reliability of clinical Shack-Hartmann measurements. *Invest Ophthalmol Vis Sci* 2004;45:351–60.
5. Smirnov MS. Measurement of the wave aberration of the human eye. *Biofizika* 1961;6:776–95.
6. Campbell FW, Gubisch RW. Optical quality of the human eye. *J Physiol (Lond)* 1966;186:558–78.
7. Howland B, Howland HC. Subjective measurement of high-order aberrations of the eye. *Science* 1976;193:580–2.
8. Walsh G, Charman WN, Howland HC. Objective technique for the determination of monochromatic aberrations of the human eye. *J Opt Soc Am (A)* 1984;1:987–92.
9. He JC, Marcos S, Webb RH, et al. Measurement of the wave-front aberration of the eye by a fast psychophysical procedure. *J Opt Soc Am (A)* 1998;15:2449–56.
10. Liang J, Grimm B, Goelz S, et al. Objective measurement of wave aberrations of the human eye with the use of a Hartmann-Shack wave-front sensor. *J Opt Soc Am (A)* 1994;11:1949–57.
11. Mouroulis P, Zhang HP. Visual instrument image quality metrics and the effects of coma and astigmatism. *J Opt Soc Am (A)* 1992;9:34–42.
12. Guirao A, Williams DR. A method to predict refractive errors from wave aberration data. *Optom Vis Sci* 2003;80:36–42.
13. Applegate RA, Sarver EJ, Khemsara V. Are all aberrations equal? *J Refract Surg* 2002;18:S556–62.
14. Applegate RA, Ballentine C, Gross H, et al. Visual acuity as a function of Zernike mode and level of root mean square error. *Optom Vis Sci* 2003;80:97–105.
15. Applegate RA, Marsack JD, Ramos R, et al. Interaction between aberrations to improve or reduce visual performance. *J Cataract Refract Surg* 2003;29:1487–95.
16. Williams DR. What adaptive optics can do for the eye. *Rev Refract Surg* 2002;3:14–20.
17. Williams DR. Wavefront basics. In: Koury JB, ed. *Wavefront and Emerging Refractive Technologies*. The Hague: Kugler Publications; 2003:3–16.
18. Williams DR, Applegate RA, Thibos LN. Metrics to predict the

- subjective impact of the eye's wave aberration. In: Krueger RR, Applegate RA, MacRae SM, eds. *Wavefront Customized Visual Correction: The Quest for Supervision II*. Thorofare, NJ: Slack Inc; 2003:77–84.
19. Thibos LN, Hong X, Bradley A, et al. Statistical variation of aberration structure and image quality in a normal population of healthy eyes. *J Opt Soc Am A Opt Image Sci Vis* 2002;19:2329–48.
 20. Hong X, Himebaugh N, Thibos LN. On-eye evaluation of optical performance of rigid and soft contact lenses. *Optom Vis Sci* 2001;78:872–80.
 21. Thibos LN, Applegate RA. Assessment of optical quality. In: Krueger RR, Applegate RA, MacRae SM, eds. *Customized Corneal Ablation: The Quest for Super Vision*. Thorofare, NJ: Slack Inc; 2001:55–63.
 22. Cheng X, Bradley A, Thibos LN. Predicting subjective judgment of best focus with objective image quality metrics. *J Vis* 2004;4:310–21.
 23. Hofer H, Chen L, Yoon GY, et al. Improvement in retinal image quality with dynamic correction of the eye's aberrations. *Opt Express* 2001;8:631–3.
 24. Noll RJ. Zernike polynomials and atmospheric turbulence. *J Opt Soc Am* 1976;66:207–11.
 25. Thibos LN, Applegate RA, Schwiegerling JT, et al. Standards for reporting the optical aberrations of eyes. *J Refract Surg* 2002;18:S652–60.
 26. Webster MA, Georgeson MA, Webster SM. Neural adjustments to image blur. *Nat Neurosci* 2002;5:839–40.
 27. Artal P, Chen L, Fernandez EJ, et al. Neural compensation for the eye's optical aberrations. *J Vis* 2004;4:281–7.
 28. Schwiegerling J. Gaussian weighting of ocular wave-front measurements. *J Opt Soc Am* 2004;21:2065–2072.

Li Chen

Center for Visual Science

University of Rochester

RC Box 270270

Rochester, NY 14627-0270

e-mail: lchen@cvs.rochester.edu