Errors in Zernike Transformations and Non-modal Reconstruction Methods

Daniel R. Neal, PhD; C. David Baer, Dipl-Ing; Daniel M. Topa, BA

ABSTRACT

PURPOSE: Wavefront aberrometers often represent the wavefront in terms of Zernike polynomials. The fit coefficients for these polynomials are pupil-size dependent; hence it is desirable to be able to calculate polynomial coefficients at a different pupil size from the initial measurement.

METHODS: We developed a method for scaling the Zernike fit coefficients for different sized pupils and compared them to the coefficients calculated by reanalyzing the original data. We evaluated cases where the initial set of Zernike coefficients was an accurate representation of the data as well as cases where the Zernike polynomials fit the data poorly.

RESULTS: The rescaling methods work well in all cases with a good original fit (as measured by the gradient root-mean-square [RMS] fit error). Rescaling and reanalyzing also agree well over a small scaling range. However, if the original fit was poor, large errors result if the rescaling is more than ~25%.

CONCLUSIONS: Rescaling works well and is useful as long as the RMS fit error can be evaluated. The RMS difference between the zonal reconstruction and the Zernike fit is useful for determining this fit error. [*J Refract Surg*. 2005;21:S558-S562.]

A number of instruments are now in common use for measuring wavefront aberrations of optical systems. The wavefront often is described in terms of Zernike polynomials, which provides for a convenient, compact notation and for measurement spaces where the polynomials are orthogonal, separation of individual phenomena.

A particular example is the measurement of the human eye. A number of wavefront aberrometers have been developed for diagnostic and treatment applications. These aberrometers are used for customized refractive surgery and other procedures, as well as for diagnostics on the surgical and treatment process.

Although the lower order aberrations, focus and astigmatism, can be written in such a way that they are pupil independent, the higher order aberrations can not, ie, the Zernike fit coefficient has a strong dependence on the measurement area. Even for the same patient, a constant variation occurs in the pupil size over a series of successive measurements. To allow comparison of measurements taken at different pupil sizes, a number of transformation methods have been developed. These allow for a calculation of the full set of Zernike coefficients at some new pupil diameter D_2 from a set measured at another diameter D_1. However, it is important to know under what conditions these transformations are valid and yield accurate results and where other methods should be used.

ZERNIKE TRANSFORMATIONS

Several authors have derived methods for calculating Zernike coefficients from values determined from measurement over a different pupil.[1,2] Another simple method for resizing the pupil is presented here. This method takes advantage of the fact that Zernike polynomials are constructed from simple monomial or Taylor polynomials. The most common calculation is resizing the pupil to a (usually) smaller diam-

From WaveFront Sciences Inc, Albuquerque, NM.

The authors have financial and proprietary interests in WaveFront Sciences Inc, Albuquerque, NM.

Correspondence: Daniel R. Neal, PhD, WaveFront Sciences Inc, 14810 Central SE, Albuquerque, NM 87123. Tel: 505.275.4747; Fax: 505.275.4749; E-mail: drneal@wavefrontsciences.com

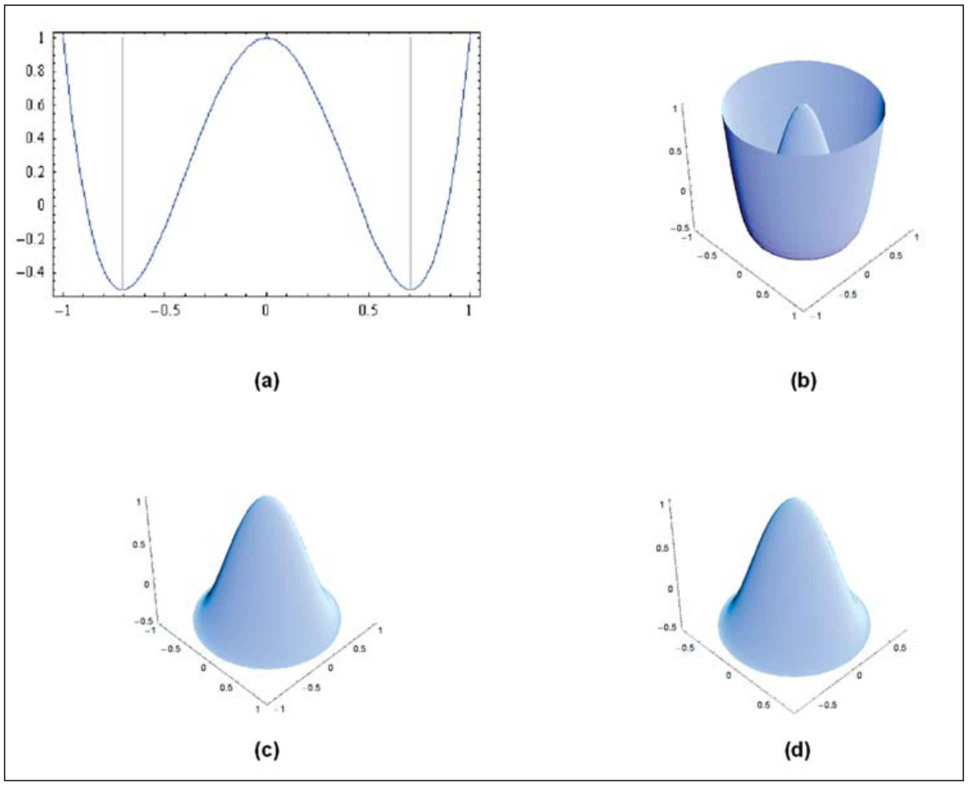

Figure 1. Resizing from one size pupil to a smaller size is actually a rescaling of the coordinate system. **A)** A one-dimensional slice of a wavefront measured over physical region Σ. **B)** A measurement of a Zernike eigenstate where the unit disk D_2 is mapped over some physical region Σ. **C)** The unit disk was remapped to the reduced region σ. Notice the change in plot domains between the figure and the one preceding it. **D)** Note the concentric region bounded by the minimum in this same measurement. The plot domain was reduced to show this region.

eter; however, pupil shift and rotation have also been considered.

Pupil Resize Transformation

Consider an ideal measurement of a wavefront as shown below. For the purpose of the measurement, the unit disk D_2 has been mapped over physical region Σ. The measurement then tells us about our observation, which is confined to Σ. Figure 1A shows an ideal measurement described by the Zernike polynomial Z_4^0.

Now suppose we wish to look at some concentric sub space of the region Σ. For example, consider the region between the minima in Figure 1A. Because the surface is ideally described by the Zernike polynomial, the necessary data for the reduced region σ are already present. Figure 1A shows the boundaries of σ. We wish to describe the wavefront in the reduced region σ delineated by the gray lines. We have all of the information about the wavefront in this region and do not need a new measurement.

The unit disk must be mapped to the region between the minima. This remapping is simply a rescaling of the coordinate system. To reduce the coverage of the unit disk over Σ by some fraction f where $0 < f < 1$, the rescaling is

$$x' \leftarrow fx. \qquad (1)$$

Imagine that we have measured a set of Zernike coefficients c_k over the region Σ. Zernike polynomials are constructed from a corresponding set of Taylor monomials.

$$\begin{pmatrix} 1 \\ x \\ y \\ y^2 - x^2 \\ 2x^2 + 2y^2 - 1 \\ 2xy \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 \\ -1 & 0 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{pmatrix}, \text{ or } Z_j = \mathbf{A}_{jk} T_k \quad (2)$$

Note that the coefficients are transformed using the same matrix. The inverse transformation can also be readily constructed.

We note that the rescaling is a natural operation in the Taylor basis as each term contains only a single polynomial order. Thus the Zernike polynomial coefficients are first mapped to a set of Taylor polynomial coefficients a_k (unless the original data were directly fit to Taylor polynomials). The coefficients b_k over the reduced region σ are generated by multiplying the coefficient vector by a diagonal rescale matrix. Observe that the power of f matches the order of the amplitude it multiplies. For example, the second order terms are multiplied by f_2.

$$\begin{pmatrix} b_{00} \\ b_{10} \\ b_{01} \\ b_{20} \\ b_{11} \\ b_{02} \end{pmatrix} = \begin{pmatrix} f^0 & & & & & \\ & f^1 & & & 0 & \\ & & f^1 & & & \\ & & & f^2 & & \\ & 0 & & & f^2 & \\ & & & & & f^2 \end{pmatrix} \begin{pmatrix} a_{00} \\ a_{10} \\ a_{01} \\ a_{20} \\ a_{11} \\ a_{02} \end{pmatrix} \quad (3)$$

Because the rescale matrix is diagonal and none of the diagonal entries are zero, this matrix has an inverse. The inverse matrix is also a diagonal matrix with negative powers of *f*. This means that the coordinate system rescaling can be reversed. The rescaled Zernike coefficients are obtained by an inverse mapping from the scaled Taylor coefficients b_k. Scaling the unit disk beyond the region Σ is an extrapolation to regions that were not included in the original Zernike computation w_k and could thus lead to erroneous results.

The process is now shown in two dimensions. Figure 1B shows a Z_4^0 mode plotted where the unit disk maps the full region Σ. The next plot, Figure 1C, shows the same data but plotted over the region σ. Notice that the plot domain is now $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. Again, in these figures, the unit disk maps to Σ. In Figure 1D, we have rescaled the coordinate system so the unit disk now maps to σ and the domain is [-1,1]. Observe that Figures 1C and 1D look exactly alike, and only the *x* and *y* axes are different.

ZERNIKE FIT ERROR

The results of the previous section are exact as long as the wavefront surface is precisely described by Zernike polynomials. However, in any real measurement, measurement noise, sampling error, and higher-order effects may not be well represented by lower order Zernike surfaces. To characterize these effects, the Zernike Fit Error can be defined as

$$w^2_{fit} = \frac{1}{N}\sum_k \left(wk - \sum_{m=2}^{M} C_m P_m(x_k, y_k)\right)^2 \quad (4)$$

where the values w_k are the actual wavefront surface values at each point (x_k, y_k). However, for a slope measurement system, it is unlikely that w_k is known directly. Rather, the wavefront gradients have been measured, and the fit forms the basis of the reconstruction. An alternative is to define a fit error in terms of the measured values, ie,

$$\beta^2_{fit} = \frac{1}{N}\sum_k \left(\beta_k^x - \sum_{m=2}^{M} C_m \frac{\partial P_m(x_k, y_k)}{\partial x}\right)^2 + \frac{1}{N}\sum_k \left(\beta_k^y - \sum_{m=2}^{M} C_m \frac{\partial P_m(x_k, y_k)}{\partial y}\right)^2 \quad (5)$$

TABLE 1

Patients Chosen for Wavefront Measurements

Patient	Sex	Age	Comments
EB	F	29	Normal
SB	F	46	Post LASIK with 6-mm ablation
SS	M	51	Corneal scars from Herpes virus

This has units of slope (dimensionless). For a set of wavefront gradient measurements that are well represented by the Zernike set under consideration, this will have a low value. If the gradients contain a significant amount of higher order information that is not well represented by the Zernike polynomial set, this will have a higher value. In fact, for least squares methods, this often is the parameter that is minimized with respect to the fit coefficients. It may be possible to approximate the wavefront fit error from the gradient fit error using the grid or measurement area size *d* (the lenslet size for Shack-Hartmann systems):

$$w_{fit} \approx \beta_{fit} d. \quad (6)$$

However, it is unlikely that this can be exactly evaluated as it would require measurement of both the wavefront and its gradient.

SCALING RESULTS

A number of cases for wavefront measurements of real human eyes were considered. Three different patients were chosen on the basis of age, sex, and eye condition (Table 1).

In Figure 2, we compared the results of the resizing algorithm against direct recalculation of the wavefront fit form data. In both cases, the Zernike fit coefficients were calculated over the appropriate smaller region. The results for several key coefficients are shown in Figure 2. For patient EB, the overall fit error, which was calculated according to Eq.(5), is low, totaling only 0.4 mrad RMS. This is a normal eye that is well represented by lower order Zernike polynomials. Even for a 4th order fit, the fit error is low. Clearly, the various coefficients are reproduced between the two methods, and both methods agree throughout the entire range of 3 to 6.25 mm.

Consider patient SB, who is 5 years post-LASIK (Fig 3). The patient was dilated using cycloplegic drops to make measurements over a larger pupil; however, we were unable to achieve such a low fit error as in a normal eye. The RMS fit error was >1 mrad for the 4th

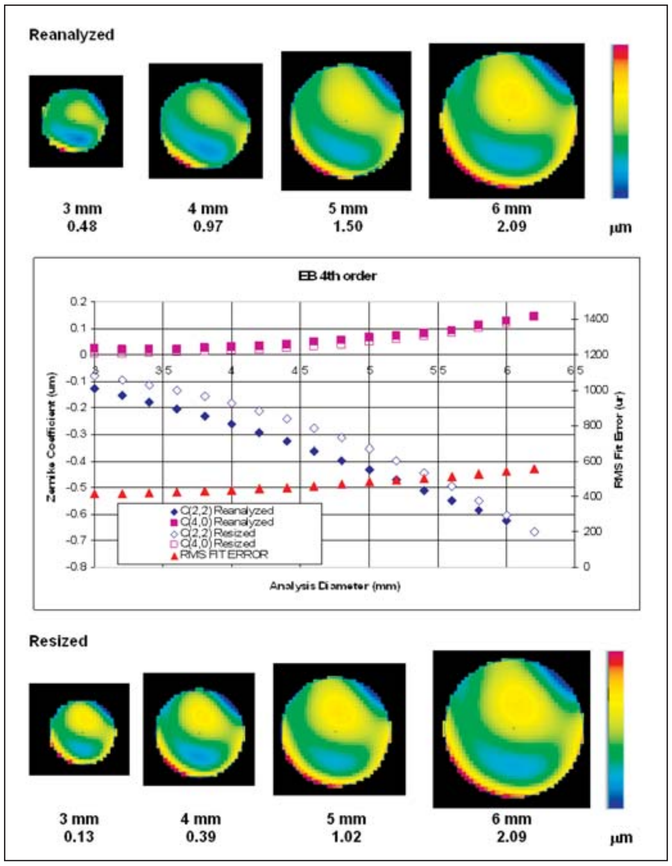

Figure 2. Reanalyzed vs resized comparison for patient EB, a 29-year-old woman.

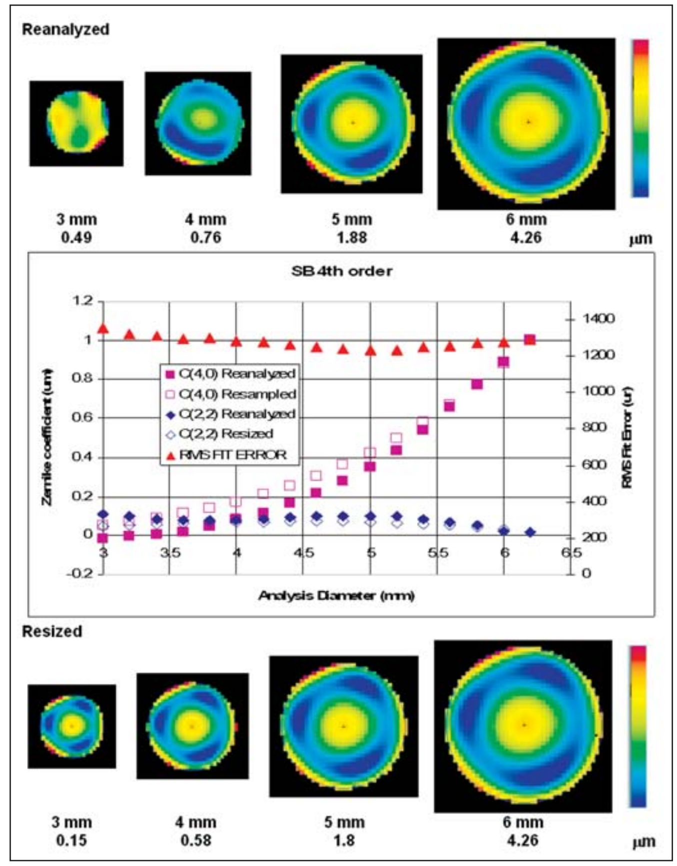

Figure 3. Reanalyzed vs resized comparison for patient SB, a 46-year-old woman who is 5 years post-LASIK.

order fit. In this case, a direct comparison between the two methods does not produce the same result. The high spherical aberration term (typical of post-LASIK patients) is attenuated by the resizing method but does not drop fully to zero as does the recalculation method. This is because the actual wavefront structure is very flat in the central region but has sharp, steep edges. This sudden change is not well-captured by the Zernike fit; hence the values are attenuated but do not approach the correct values at small pupil size. Even so, it should be noted that the resizing worked well over a small range and would thus serve the purpose of adjusting the pupil size to allow comparison of population data.

The third case is an eye that has significant corneal scarring resulting from infection with the Herpes virus. The central portion of the pupil has good optical quality; however, a region with strong wavefront variations is noted at 2.5-3 mm radius. The fit error (Fig 4) increases rapidly for pupil sizes between 5 and 6 mm for this patient. At the largest pupil diameter, it is >1.7 mrad RMS. As can be seen by examining the various predicted Zernike terms, the resizing algorithm does not work well in this case.

The resizing method cannot predict the correct wavefront at smaller pupil sizes for two reasons. The first is that the residual fit error leads to poor representation by Zernike polynomials. The other reason is the increased spatial frequency information that is contained in Zernike polynomials at smaller pupil sizes. This can be seen in the wavefront plots in Figures 2 and 3. In each case, we have varied the pupil size by a factor of more than two. Consider the Z_4^0 Zernike term. It has essentially two cycles of spatial frequency across the aperture. The number of cycles is fixed, but, as the aperture size decreases, the spatial scale changes. So although the Z_4^0 polynomial contains 0.3 c/mm at 6-mm diameter, it contains 0.6 c/mm at 3-mm diameter. If the original data are reanalyzed, the fit will pick up local variations that are at high spatial frequency, and hence different fit coefficients will result. The resizing algorithm contains only the original spatial frequency information, and it cannot account for the increase in spatial frequency with reduction in pupil size.

ZONAL RECONSTRUCTION

Some wavefront reconstruction techniques provide wavefront information at the resolution of the sensor.[3] These are commonly referred to as zonal methods (af-

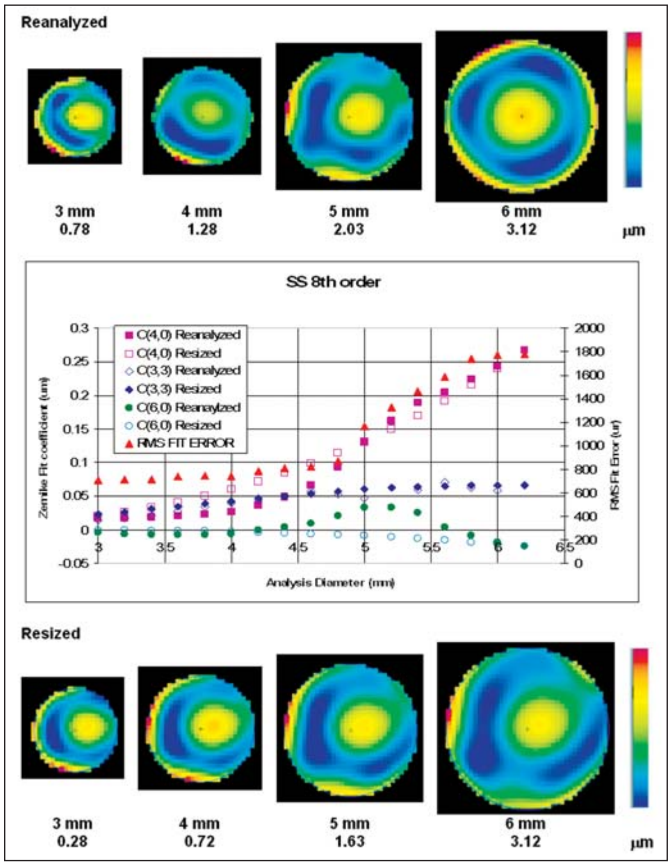

Figure 4. Reanalyzed vs resized comparison for patient SS, a 46-year-old man who has corneal scars.

ter Southwell[4]). The problem with these methods is that they provide only the wavefront distribution and provide no information that can readily be used for refraction or quantitative comparison of effects. Given an accurate sensor, the zonal reconstruction represents the wavefront that would presumably also be represented by the Zernike surface. We propose the use of the difference between the two methods as a metric to describe how well the wavefront information can be represented by the Zernike surface. The difference metric thus becomes:

$$\sigma^2_{Z-M} = \frac{1}{N}\sum\left(w^z_k - \sum_n C_n Z_n(x_k,y_k)\right)^2. \qquad (7)$$

This has the advantage in that the units are in micrometers RMS, which is easier to interpret, and it does not suffer from the approximations inherent in Eq.(6). This difference metric is presented for each patient in

TABLE 2

Comparison of Fit Errors Using Different Equations

Patient	Fit Order	w_{fit} Eq.(6)	σ_{z-n} Eq.(7)
EB	4	0.088	0.14
SB	4	0.20	0.31
SS	8	0.28	0.16

Table 2. Although the numbers have the same approximate magnitude, it is evident that Eq.(6) does not adequately predict how the spatial structure adds to form the residual fit error. The zonal-modal difference can be higher or lower depending on the spatial structure. Note that the assumption that the sensor is accurate implies that the zonal reconstruction does not amplify the noise in some unpredictable fashion. However, the noise would be present in the modal reconstructor as well and would affect the Zernike fit in the same manner as the zonal. Thus the difference would have a similar meaning.

We have demonstrated a simple method for determining a new set of Zernike fit coefficients from an initial fit. In all cases (even with some extreme examples), we found that the method worked well over a small range of diameters (~80% of the original pupil). However, the resize method (as compared with recalculating the fit from the data) did not work as well when the fit error was large. Furthermore, if regions with large wavefront variation are present near the edge of the pupil, then even a small scaling of pupil size would result in significant error.

We have also proposed the calculation of the zonal-modal difference as a metric for determining whether the wavefront information can be properly represented by Zernike polynomials.

REFERENCES

1. Campbell C. Matrix method to find a new set of Zernike coefficients from an original set when the aperture radius is changed. *J Opt Soc Am A Opt Image Sci Vis*. 2003;20:209-217.
2. Schwiegerling J. Scaling Zernike expansion coefficients to different pupil sizes. *J Opt Soc A Am Opt Image Sci Vis*. 2002;19:1937-1945.
3. Panagopoulou SI, Neal DR. Zonal matrix iterative method for wavefront reconstruction from gradient measurements. *J Refract Surg*. 2005;21:S563-S569.
4. Southwell WH. Wavefront estimation from wave-front slope measurements. *J Opt Soc Am A Opt Image Sci Vis*. 1980;70:998-1006.