Comparison of Hartmann analysis methods

Carmen Canovas and Erez N. Ribak

Analysis of Hartmann–Shack wavefront sensors for the eye is traditionally performed by locating and centroiding the sensor spots. These centroids provide the gradient, which is integrated to yield the ocular aberration. Fourier methods can replace the centroid stage, and Fourier integration can replace the direct integration. The two—demodulation and integration—can be combined to directly retrieve the wavefront, all in the Fourier domain. Now we applied this full Fourier analysis to circular apertures and real images. We performed a comparison between it and previous methods of convolution, interpolation, and Fourier demodulation. We also compared it with a centroid method, which yields the Zernike coefficients of the wavefront. The best performance was achieved for ocular pupils with a small boundary slope or far from the boundary and acceptable results for images missing part of the pupil. The other Fourier analysis methods had much higher tolerance to noncentrosymmetric apertures.

OCIS codes: 010.7350, 120.2650, 120.3890, 120.5050, 170.4460.

1. Background

Measurements of optical wavefronts are now becoming popular, spreading from optical manufacturing to astronomy, and more and more for ophthalmology. A common wavefront sensor is the Hartmann–Shack device, where the wavefront is sampled by a lenslet array, providing an array of focal spots on a camera. From such an image, it is possible to obtain the wavefront's slopes, by calculating the spot movement. The next stage is integration of the final wavefront from these slopes. Both steps are calculation processes that have been widely described either in the image[1,2] or in the Fourier domain.[3–6] For Fourier analysis, it is better if the lenslets are along a Cartesian grid, and the number of lenslets and the number of detector pixels are both maximized. Fourier methods also allow very long focal lengths of the lenslets, improving the sensitivity, because adherence of the spots to the lenslet locations is no longer required. The only condition, similar to centroiding, is that the wavefront and its derivatives are well behaved, and beams do not cross each other before the detector (the adiabatic condition).

To recap and understand the mathematical basis of wavefront reconstruction methods, we write the intensity of the spots in the Hartmann–Shack detector, in a flat wavefront case, as a regular grid,[3–6]

$$I_0(x, y) = \sum_{m,n} a_{m,n}(\cos 2\pi mx/P + \cos 2\pi ny/P), \quad (1)$$

where $P = 2\pi/k$ is the lenslet pitch (in pixel units). To develop the retrieval methods, the spots are assumed circularly symmetric and equal, so not all the harmonics are necessary. We also borrow the name sidelobes for these harmonics, as in the similar superheterodyne demodulation method. The first harmonics in x and y provide us with enough information to retrieve the wavefront, because they represent the spots layout but not their internal structure. Hence, in Eq. (1), it is possible to work with $m = n = 1$:

$$I_0(x, y) \approx \cos kx + \cos ky. \quad (2)$$

If the wavefront contains some aberrations, the spots are shifted by a quantity directly related to the aberrations. These aberrations modify the periodicity of the array:

$$I(x, y) \approx \cos[kx + FW_x(x, y)] + \cos[ky + FW_y(x, y)], \quad (3)$$

where W_x and W_y are the phase x and y derivatives that we want to determine in order to retrieve the final wavefront $W(x, y)$ after the integration. F is the

C. Canovas is with the Departamento de Física, Laboratorio de Optica, Universidad de Murcia, 30071 Murcia, Spain. E. N. Ribak (eribak@physics.technion.ac.il) is with the Department of Physics, Technion-Israel Institute of Technology, Haifa 32000, Israel.

Received 23 June 2006; revised 9 October 2006; accepted 23 October 2006; posted 26 October 2006 (Doc. ID 72244); published 1 April 2007.

0003-6935/07/100001-06$15.00/0

focal length of the lenslets. In centroiding, Eq. (3) is sampled at the lenslet locations only, whereas here we assume that the slopes are continuous even between the measured spots.

It is rather easy to isolate these terms of Eq. (3) in the Fourier transform of the spot pattern, as they are contained in only four symmetric sidelobes [Fig. 1(b)]. The centers of the sidelobes are at distance k from the Fourier origin (in the discrete representation), as there are k lenslets across the pupil:

$$\mathcal{F}\{I(x,\,y)\} \approx \mathcal{F}\{\exp(-iFW_x)\}\delta(u-k,\,v) + \mathcal{F}\{\exp(iFW_x)\} \times \delta(u+k,\,v) + \mathcal{F}\{\exp(-iFW_y)\}\delta(u,\,v-k) + \mathcal{F}\{\exp(iFW_y)\}\delta(u,\,v+k), \quad (4)$$

where u, v signifies the coordinates in the Fourier domain, and $\mathcal{F}\{a\}$ is the Fourier transform of a. The different harmonics ($m, n = 0, \pm1, \pm2, \ldots$) are well separated in the Fourier domain when the number of lenslets is large, and the first harmonics contain the slopes. This is exactly the property that the Fourier demodulation technique exploits in order to get the phase derivatives. In this technique, the calculations of the slope components are performed by a rigid translation of the whole Fourier transform of the image in order to place the first sidelobe, corresponding to the lenslet frequency k, in the center. To remove all other sidelobes, a low-pass filter multiplies the centered sidelobe, obtaining the filtered and centered Fourier transform of the image

$$\mathcal{F}\{I_x'(x,\,y)\} \approx \mathcal{F}\{\exp(-iFW_x)\}\delta(u,\,v) \quad (5)$$

and similarly for the y direction. After isolation, the x and y sidelobes are inverse transformed separately, and their arguments provide the horizontal and vertical slopes.[3] As opposed to the centroiding method where the slope is given only at the sites of the lenslets, here the slope components are found (interpolated) for each pixel in the original frame.

We chose to calculate the final wavefront in the Fourier domain as well, by employing the fact that a Fourier transform of a function is known if its derivatives in x and y are known,[7]

$$\mathcal{F}\{W(x,\,y)\} = -[iu\mathcal{F}\{W_x(x,\,y)\} + iv\mathcal{F}\{W_y(x,\,y)\}]/(u^2+v^2). \quad (6)$$

Then, to get the final wavefront, an inverse Fourier transform is used. We actually employed a similar least-squares solution,[6,8] where u and v in Eq. (6) are replaced everywhere by $\sin u$ and $\sin v$. This accurate method can be slow, because to process one image, it is necessary to perform six Fourier transforms.

In some application, such as in adaptive optics, speed is important. In an effort to provide faster solutions, it was realized that the demodulation (but not the integration) can be performed directly on the Hartmann–Shack pattern in a manner similar to Fourier analysis[5] [Eqs. (1)–(5)]: The Fourier shift of the x lobe to the center can be obtained by multiplying the original pattern by an exponential phase $I^{cx}(x, y) = I(x, y)\exp(-2i\pi x/P)$, and similarly for y. The low-pass filtering can be performed next by a convolution of I^{cx} with a kernel of the size of pitch P. Alternatively, a low-pass filter can be performed by smoothing the same array[5] to interpolate the (complex) values between the spots. In both methods, convolution and smoothing (or interpolation below), the slope components are the *arguments* of the resulting arrays. To retrieve the final wavefront, some type of integration is needed, so Eq. (6) is also applicable, requiring three Fourier transforms. Another type of integration could be useful, such as modal approximation, say by Zernike polynomials. However, Eq. (6) ensures Laplace's property, whereas in modal calculation, only low-order terms are extracted from the wavefront slopes, and from them the wavefront is retrieved.[9]

Another method was developed in order to minimize the calculation time, the fast Fourier demodu-

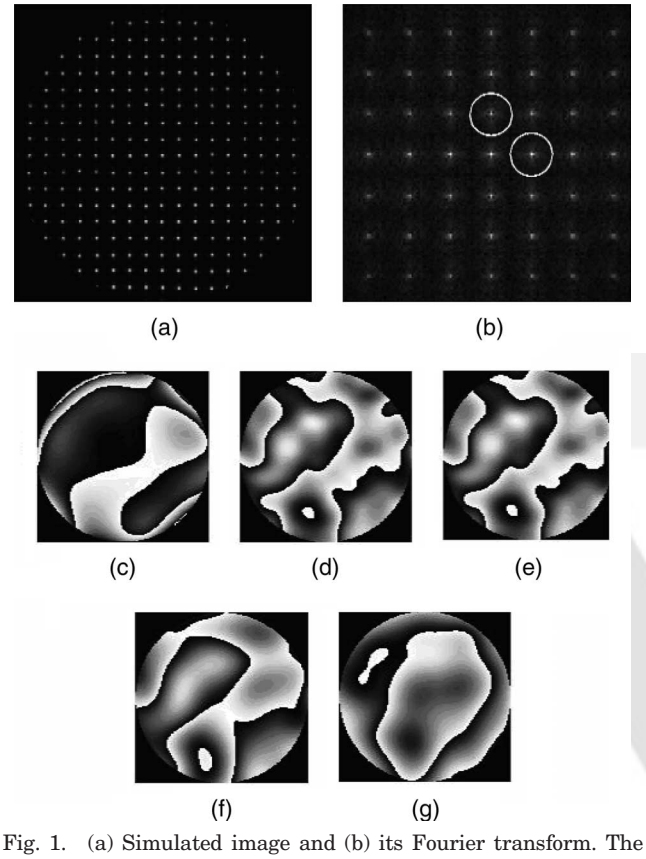

Fig. 1. (a) Simulated image and (b) its Fourier transform. The distance of the first sidelobes from the Fourier origin is equal to the number of spots across the pattern. Fourier modulation consists in isolating each circled sidelobe, centering it, and transforming it back to get the corresponding slope component. The resultant wavefronts are from five processing methods (c) traditional centroiding, (d) convolution, (e) smoothing, (f) Fourier demodulation, and (g) fast Fourier demodulation. All the methods consisted of first obtaining the wavefront slopes. In (c), Zernike polynomials were fit to the slopes, in (d)–(f) integrating the slopes in the Fourier domain, and in (g) all processing was achieved in the Fourier domain. Note the different residual tilts.

lation.[4] Here both demodulation and integration are calculated in the Fourier domain, returning to the image domain only to express the final results. In this case, the adiabatic condition[3,5] is not enough. The amplitude is expected to be nearly constant and the slopes to vary rather slowly,

$$\exp(iFW_x) \approx 1 + iFW_x, \tag{7}$$

so the Fourier transform of the wavefront after the demodulation described before could be written as

$$\mathcal{F}\{I^{cx}(x, y)\} = \mathcal{F}\{\exp(iFW_x)\},$$

$$\delta(u, v) \approx \mathcal{F}\{1 + iFW_x\}\delta(u, v). \tag{8}$$

This allows extraction of the anti-Hermitian part of the filtered and centered Fourier sidelobes, where all the information about the transform of the slope components is contained:

$$\mathcal{F}\{iFW_x(x, y)\} \approx [\mathcal{F}\{\exp(iFW_x)\} - \mathcal{F}^*\{\exp(iFW_x)\}]/2. \tag{9}$$

This information is exactly what is needed in the integration of the slopes [Eq. (6)], so the whole calculation is achieved without leaving the Fourier domain. Thus, for every image, only two Fourier transforms are needed, making the calculation much faster. This method might also be called modal, as it calculates the Fourier modes of the slope (albeit directly from the Hartmann–Shack pattern), then integrates these Fourier modes to yield the wavefront. Compare this with the calculation of the Zernike modes of the slope (as obtained by the centroids), and from them, the Zernike description of the wavefront.

2. Methods

The main objective of the study is to perform a comparison between the methods in terms of speed and accuracy, and conditions thereof. To make the comparison full, we also included a calculation based on the centroid method, performed by a well-tested program.[9] As this program yields its results as Zernike coefficients, we also expressed the retrieved wavefronts through the other methods as Zernike terms.

A MATLAB program was written with the theoretical basis of all methods. Each image was processed by the four methods to yield four final wavefronts as well as the centroid result. This program was run for images with circular apertures (ocular pupils), which could not be accomplished before with the method of fast Fourier demodulation. The main reason for writing the program was to see how this method (and previous ones) handles real images and what the conditions for good results are.

There are some important issues to stress that are related to the processing. Theoretically, the method to obtain the slopes in the Fourier domain is by extraction of the anti-Hermitian part. To put this into practice, it is important to take into account the fact that an anti-Hermitian matrix has an antisymmetrical real part and a symmetrical imaginary part. Then it is possible to obtain a matrix with these properties from the matrix, which holds the Fourier transform of the image. Toward this end, a center of symmetry needs to be well defined after shifting the sidelobe to the Fourier origin. Using that center of symmetry, another matrix is calculated where every point relates, in its symmetric form, to the matrix center of symmetry. The anti-Hermitian matrix retrieved from the original one is the result of the subtraction of this one and the new one generated by the center of symmetry:

$$\mathcal{F}\{I'(\mathbf{r})\} = [\mathcal{F}\{I'_R(\mathbf{r})\} - \mathcal{F}\{I'_R(-\mathbf{r})\} + i(\mathcal{F}\{I'_i(\mathbf{r})\} + \mathcal{F}\{I'_i(-\mathbf{r})\})]/2, \tag{10}$$

where $\mathcal{F}\{I'(\mathbf{r})\}$ is the filtered and centered Fourier transform of the image with relation to the center of symmetry, and subscripts I and R stand for real and imaginary parts.

Two conditions have to be fulfilled in order to get the anti-Hermitian part, both related to the requirement for a center of symmetry. The first one is that the aperture is symmetrical within the frame to be processed. Otherwise, it is not possible to define a center of symmetry, and the slopes cannot be retrieved. The second condition is that the ocular pupil must not be cut by the optics or by low eyelids, again losing symmetry.

If these conditions are not fulfilled, the results from the fast Fourier demodulation will only be qualitatively acceptable. Thus we had to find the pupil inside the image, defining the diameter and location of a box holding the centered image. First, the image is low-pass filtered in the Fourier domain, leaving only the central lobe but excluding the spot-related lobes (a smoothing method in the image plane is just as good, provided it obliterates the spot pattern). After that, the edge of the resulting image is calculated. Finally, a circle is fitted to this edge. This circle is defined as the pupil, and from it, the containing box is readily calculated (Fig. 2). With this subroutine, two objectives are achieved: the unimportant information outside the pupil is discarded, and the remaining image is centrosymmetric. Of course, all this is based on the assumption that the pupil is indeed round.

A second issue is the subtraction of the reference. As every system has some aberrations and systematic errors in it, a reference image is first taken through the system from a model of a perfect eye, processed and stored. To get the real ocular wavefront, it is necessary to subtract from the processed wavefront this reference wavefront. For fast Fourier demodulation, the direct subtraction of the reference wavefront can be performed in the Fourier domain because of the linearity of Fourier transforms. However, it is also important to take into account the average derivative phase generated by the shift of the real from the reference image. This shift can arise

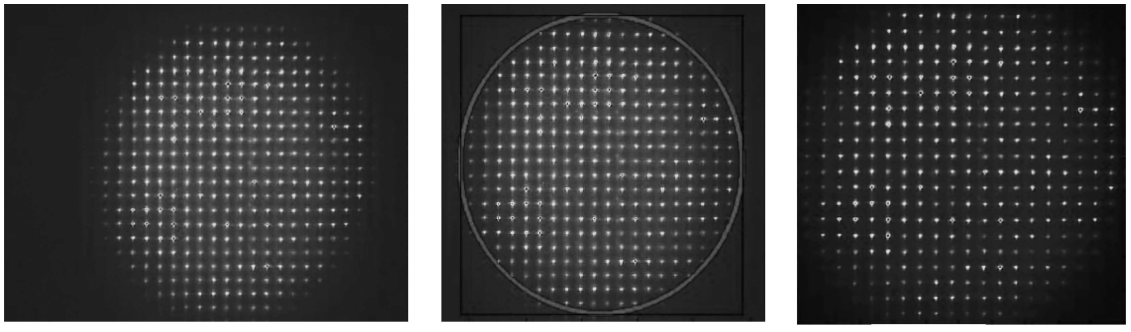

Fig. 2. Defining a centrosymmetric pupil: (a) original image, (b) calculated pupil edge, (c) calculated box containing all the image points.

from the fact that the two patterns, the reference and the final one, had different centers: the grid of spots is the same, but the pupil can appear in various locations as the eye moves, and cut out different spots. In such a case, the spot pattern can look as if shifted by the difference in the centers. If this is not removed, then an average derivative is added, and after integration, the final result for fast Fourier demodulation is totally different from the other methods.

Another point is related to the edges. It has been shown that extending the edges, by copying the peripheral spots outside the pupil, is a technique that can lower the boundary-related errors.[5] This has been used for all the methods that we tested except for centroiding. As we show in the results, this technique is less beneficial for fast Fourier demodulation, which seems to be more sensitive to small extension errors. We believe that extrapolating the edge by a true lenslet pitch and not by integer pixels, as we do now, could remove this problem. Furthermore, the true lenslet pitch is calculated from the perfect reference image. For eyes with large edge slopes, the common result of defocus, astigmatism or spherical aberration, the actual local pitch is different from that of reference pitch, and spot extrapolation can harm the results.

Finally, aliasing effects must be considered. Except for centroiding, there is at least one Fourier stage in the calculation, where aliasing can reduce the quality of the final results. Zero padding is used to remove these effects: a matrix of zeros of twice the pupil size is defined. The image, resulting from the centering subroutine, is planted in the center of the matrix of zeros. The size of the matrix, even if necessary, slows all Fourier calculations.

3. Comparison

We now present results for two types of image, all processed with circular apertures: computer-created images and real images. For this last case, we processed both round and cut pupils and checked if those theoretical assumptions, described before, were correct and sufficient. Finally, we expressed the final wavefronts retrieved in each method in Zernike terms and calculated their first 30 coefficients to perform a quantitative comparison of all the methods and to study the reliability of fast Fourier demodulation.

Hartmann–Shack images were simulated by creating wavefronts with Gaussian distribution and a power spectrum of $-\frac{5}{6}$ to $-\frac{5}{3}$. These wavefronts were then properly propagated through a lenslet array, and their power spectrum was taken with Poisson noise added to it (no speckle was assumed). In other simulations, the Airy patterns at the foci of the lenslets were replaced by equal-magnitude delta functions [Fig. 1(a)]. Processing results are shown in Figs. 1(c)–1(g) for one of the ten realization runs. Also, the

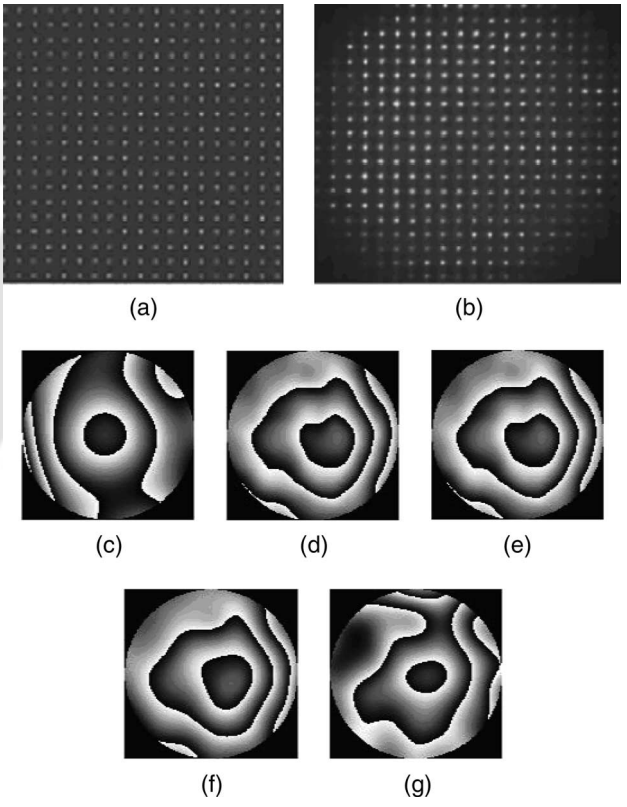

Fig. 3. Comparison of wavefronts: (a) reference image, (b) part of the image used in analysis from the eye's measured spot pattern, (c)–(g) results for the five methods corresponding to Figs. 1(c)–1(g).

wavefronts were retrieved with the centroiding program, and the results that we show were obtained after the subtraction of a reference image, generated by the same simulation program with much weaker aberrations. The results are similar in all the methods. There is a very small tilt difference in the full Fourier demodulation and centroiding, and the two Fourier demodulations have a slightly narrower low-pass filter. We calculated the rms differences between the original wavefront and the results, compared with the rms of the wavefront for the low noise case. The ratio was less than 7% for the convolution and smoothing methods, 9% for the Fourier method, and 10% for the centroid and full Fourier method. These numbers are influenced by the slope and bandwidth differences, and the slightly smaller centroid diameter and are accurate to ~3%. Overall, we can conclude that, for the first time to the best of our knowledge, fast Fourier demodulation was applied successfully for circular apertures.

We now turn to real images. We show that fast Fourier demodulation works properly for whole, uncut pupils. Figure 3(a) displays the reference and the experimental data. The result of the centering subroutine and the final image to be processed (with later zero padding) are also given in Fig. 3(b). The final panels compare the final wavefronts for all the methods. The images have been processed for a 3.5 mm pupil. The results are similar qualitatively, despite edge errors that have to do with the extension of the edges and with pupil centering.

We also processed images where the 4.2 mm pupil was cut by the camera (Fig. 4). In this case, as expected from the asymmetric amplitude, the results of the full Fourier demodulation are worse. On the other hand, the shape of the wavefront is similar for all the methods, if the sign is not taken into account. We show this as an example: it happens because the program locates automatically the brightest sidelobes, and sometimes it finds an equal sidelobe in the opposite quadrant. This can easily be remedied by forcing it always to choose the sidelobes in the same quadrants.

To perform a quantitative comparison, a set of experimental images was processed in all the methods, and the final wavefronts were calculated as coeffi-

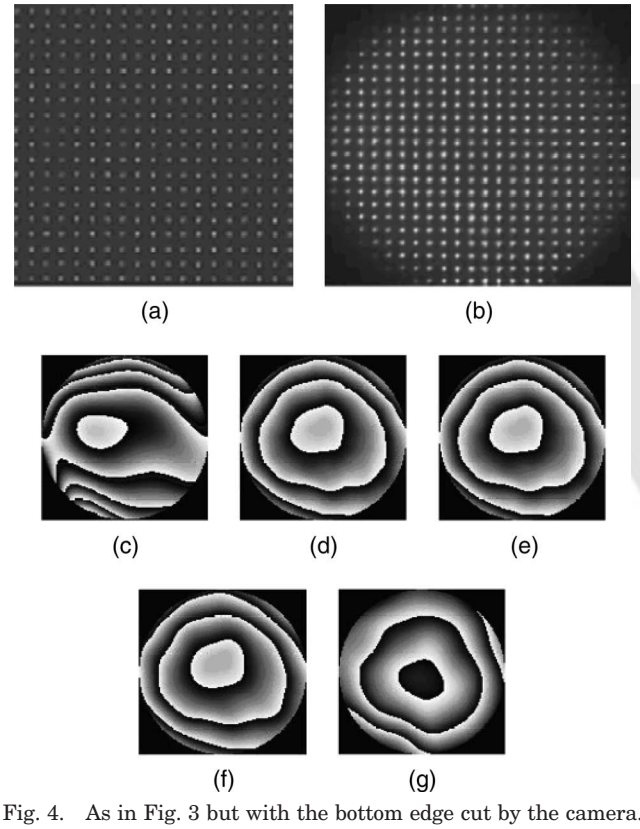

Fig. 4. As in Fig. 3 but with the bottom edge cut by the camera. The fast Fourier demodulation picked up the opposite sidelobes and inverted the sign.

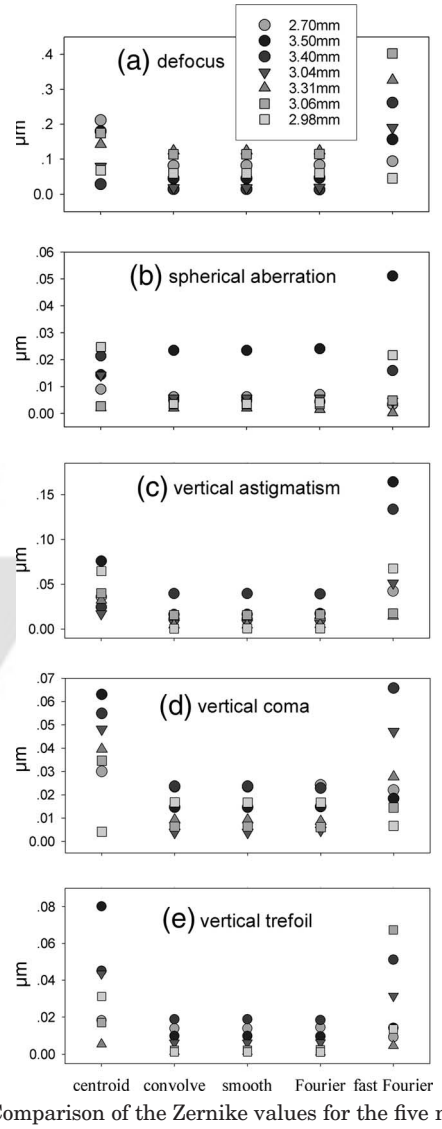

Fig. 5. Comparison of the Zernike values for the five methods as in Figs. 2–4. Shown from top are the absolute values of (a) defocus, (b) spherical aberration, (c) vertical astigmatism, (d) vertical coma, (e) vertical trefoil. The apertures were limited to 2.70, 3.50, 3.40, 3.04, 3.30, 3.06, and 2.98 mm. The largest pupils show the largest errors, as more boundary values were included.

cients for 30 Zernike polynomials. The main coefficients were plotted together. The results for defocus, spherical, vertical astigmatism, and vertical coma for different sizes of the pupil are shown in Fig. 5 for pupils that vary from 2.7 to 3.5 mm, and the results are similar for all the methods. For bigger pupils, the results deviated, as is visible for the spherical aberrations in the 3.5 mm pupil. In the wavefronts shown in Fig. 3, some edge effects were also present. We suspect that they arise from the fact that for large pupils fast Fourier demodulation becomes even more sensitive to precise extension of the boundary and to accurate centroiding. The bigger the pupil, the more lenslets are visible, and the higher the frequency k of the sidelobes. The same errors in the spot locations lead to larger relative errors for denser spots. In addition, aperture centroiding becomes more sensitive to pixelation errors, and comparison with the reference wavefront becomes less accurate. We plan to probe into these error sources and remove them, for example, by shifting the images by noninteger pixel steps when extending their edges.

We also ran comparisons on the calculation time for all the methods. For Hartmann–Shack experimental images, where the image size is 768×576 pixels, the processing time of fast Fourier demodulation, similar to an interpolation method and slightly lower than a convolution method, is half of the Fourier demodulation. This method employs two Fourier transforms, as compared with the six by a simple demodulation method. The balance is due to the fact that in using MATLAB, the fast Fourier transform routines are written efficiently in C as compared with the rest of the processing.[5]

As a conclusion, the fast Fourier demodulation performs well with actual ocular circular pupils. Its results are currently accurate for round pupils having well-behaved boundary slopes. Otherwise, all the methods yield similar results, with differences of less than 10% among them. These differences occur mostly at the boundary, where the extrapolation of spots by the reference pitch, rounded to integer pixels, might not match the actual ocular pitch at the edges. We now develop the next analysis version to deal with this problem as well as the occasional shifting of the pupil across the reference image.

The final choice of algorithm depends now on the application: for adaptive optics, either the smoothing or the fast Fourier method will be acceptable; for accurate wavefront sensing, the slower full Fourier method or the convolution with a large kernel.

This work was performed at the Technion with cooperation within the Sharp-Eye Research and Training Network (European Framework V program), with a contribution from the Israeli Ministry of Science. We thank Pablo Artal for the results from the ocular wavefront analysis program.

References

1. R. K. Tyson, *Principles of Adaptive Optics,* 2nd ed. (Academic, 1997).
2. J. W. Hardy, *Adaptive Optics for Astronomical Telescopes* (Oxford U. Press, 1998).
3. Y. Carmon and E. N. Ribak, "Phase retrieval by demodulation of a Hartmann–Shack sensor," Opt. Commun. **215,** 285–288 (2003).
4. Y. Carmon and E. N. Ribak, "Fast Fourier demodulation," Appl. Phys. Lett. **84,** 4656–4567 (2004).
5. A. Talmi and E. N. Ribak, "Direct demodulation of Hartmann–Shack patterns," J. Opt. Soc. Am. A **21,** 632–639 (2004).
6. A. Talmi and E. N. Ribak, "Wavefront reconstruction from its gradients," J. Opt. Soc. Am. A **23,** 288–297 (2006).
7. F. Roddier and C. Roddier, "Wavefront reconstruction using iterative Fourier transforms," Appl. Opt. **30,** 1325–1327 (1991).
8. K. R. Freischlad and C. L. Koliopoulos, "Modal estimation of a wavefront from difference measurements using the discrete Fourier transform," J. Opt. Soc. Am. A **3,** 1852–1861 (1986).
9. P. M. Prieto, F. Vargas-Marin, S. Goelz, and P. Artal, "Analysis of the performance of the Hartmann–Shack sensor in the human eye," J. Opt. Soc. Am. A **17,** 1388–1398 (2000).