Comparison of sorting algorithms to increase the range of Hartmann-Shack aberrometry

Phillip Bedggood
Andrew Metha
University of Melbourne
Department of Optometry and Vision Sciences
Corner Keppel and Cardigan Street
Carlton 3053, Victoria, Australia

Abstract. Recently many software-based approaches have been suggested for improving the range and accuracy of Hartmann-Shack aberrometry. We compare the performance of four representative algorithms, with a focus on aberrometry for the human eye. Algorithms vary in complexity from the simplistic traditional approach to iterative spline extrapolation based on prior spot measurements. Range is assessed for a variety of aberration types in isolation using computer modeling, and also for complex wavefront shapes using a real adaptive optics system. The effects of common sources of error for ocular wavefront sensing are explored. The results show that the simplest possible iterative algorithm produces comparable range and robustness compared to the more complicated algorithms, while keeping processing time minimal to afford real-time analysis. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3516706]

Keywords: Hartmann-Shack; wavefront sensing; range; accuracy.

Paper 10418R received Jul. 23, 2010; revised manuscript received Oct. 5, 2010; accepted for publication Oct. 6, 2010; published online Nov. 29, 2010.

1 Introduction

Hartmann-Shack aberrometry is one of the most common methods used to characterize the wavefront aberrations of an optical system, most especially the aberrations of the human eye. In its most common implementation, an array of lenslets is used to form an array of images of a point source that has passed through some optical media. Each subimage, or “spot,” is deviated away from the axis of its lenslet by an amount proportional to the average wavefront slope across the lenslet. Measuring these deviations across the pupil enables the wavefront shape to be reconstructed and/or an appropriate adaptive optics correction to be determined.1

Locating the position of each spot is relatively trivial using standard image processing techniques. The more challenging task is to ensure that each spot is assigned to the correct lenslet, a process referred to as “sorting” the Hartmann-Shack pattern. The easiest way to accomplish this is to assume that each spot image is formed on the detector plane within the bounds of its lenslet. The problem is then reduced to a series of small search windows of fixed location and size. We refer to this as the “conventional” algorithm. With this algorithm, when the local wavefront slope is too steep, a spot will appear outside the bounds of its lenslet and therefore be ascribed to the wrong lenslet. This causes errors in the wavefront fit and in any adaptive optics correction.

The range of this conventional algorithm is sufficient for the majority of normal human eyes when measuring aberrations close to the optical axis. However, the algorithm can fail when measurements are sufficiently off-axis, when refractive error is moderate to high, or when pathology is present that distorts the refractive components of the eye.

Several hardware changes to Hartmann-Shack systems have been proposed to circumvent the limitations of the conventional approach. Larger magnification of the pupil can be used to reduce wavefront slope and hence spot deviation, but this requires a wider detector.2 Lenslet focal length can be shortened to again provide less deviation at the detector plane, but this decreases accuracy due to pixelation effects.2 Lenslet size can be made larger, but at the expense of reduced lateral spatial resolution of the wavefront.2 Yoon et al. suggested a rapidly translating plate that could block certain lenslets in alternating frames, although this sacrifices temporal resolution.3 Lindlein and Pfund suggested an array of astigmatic lenslets with different orientations to uniquely identify each spot in a local area.4 Ares et al. described the use of an array of astigmatic lenslets together with a software algorithm that was able to trace the resultant line foci outside the bounds of the conventional search area.5 For aberrometry on the human eye, appropriate trial lenses are often put into place. This can inject unwanted reflections and is time consuming in that a refraction must be performed first, and it does not help with imaging the peripheral fundus where refraction changes markedly, or for higher order effects such as in patients with corneal disease.

Software-based approaches can also increase the range, while avoiding alterations to existing systems and minimizing the cost of system construction. For example, Pfund et al. described a sorting algorithm that was inspired by a common approach used in unwrapping phase images in interferometry.7 Groening et al.8 and Lundström and Unsbro9 described more complex algorithms where a spline surface was fit to central spots, and then iteratively extrapolated to predict the locations of subsequent spots. Leroux and Dainty used a similar approach but fit Zernike polynomials to the spot displacement data.
instead of a spline, producing a smoothing effect that suppresses noise. Lee et al. and Smith and Greivenkamp each described algorithms in which spots are located in a predetermined order in such a way as to reduce ambiguity as to lenslet-spot correspondence.

These software methods generally reported dramatic improvements in the range of aberrometry, but it is difficult to compare algorithms directly due to differences in system parameters, aberrations considered, and criteria for failure. In particular, no range data at all appeared with the Smith and Greivenkamp algorithm. The major aim of this review is to facilitate the direct comparison of several of these algorithms by providing further range data obtained under consistent conditions. We also explore the robustness of the considered algorithms in the context of ocular aberrometry by considering the effects of poorly sensed spots, decentration of the pupil (in the absence of a pupil monitor), and decentration of the input beam.

2 Methods—The Algorithms

Three of the preceding algorithms were selected to reflect the most modern tactics that provide large increases in range over the conventional approach, and to cover a range of complexity in terms of processing required. For each algorithm, it is assumed that the spot positions are already known via some image processing method, and that the only remaining problem is the assignment of the spots to their corresponding lenslets. Figure 1 depicts the order in which spots are located for each algorithm. For a more thorough treatment of a given algorithm, the reader should consult the original paper.

It is noteworthy that each of the algorithms explored here approach the spot allocation process in a generally similar way. An initial assumption is made in each case that a small number of spots nearest to the pupil center correspond to their directly overlying lenslets. This is a valid assumption since aberrations are smallest at the pupil center. From this anchor point, each algorithm expands iteratively outward from the pupil center. On a given iteration only spots that are adjacent to previously allocated spots are sorted, which enables the algorithms to proceed with minimum uncertainty since the wavefront changes minimally from one lenslet to the next.

2.1 Lundström and Unsbo—“B-Spline”

This algorithm assumes that the central nine spots of the Hartmann pattern correspond to their overlying lenslets. A B-spline function is fit to the spot displacements and extrapolated to predict the locations of surrounding spots that have not yet been ascribed a lenslet. This is followed by a new B-spline fit and further gradual extrapolation, until all spots have been sorted. We refer to this as the “B-spline” algorithm. The B-spline algorithm is the most general in its processing—it
is therefore predicted to offer the greatest range, but the slowest processing time. Upon request, the authors generously supplied us with a MATLAB implementation of their algorithm, which was used to generate data here.

2.2 Leroux and Dainty\textsuperscript{10} — “Zernike”
This algorithm follows an approach similar to that of Lundström and Unso.\textsuperscript{9} In place of a B-spline fit, Zernike polynomials are fit to the actual wavefront data. A second-order fit is employed for the central nine spots, and a fourth-order fit beyond that. The smoothing effect of the Zernike fit was reported to improve performance in the presence of noisy measurements and/or missing spots. We refer to this as the “Zernike” algorithm. The code for this algorithm was kindly made available on the author’s website, and was used to generate results here after making some minor changes to match the parameters of our chosen lenslet array.

2.3 Smith and Greivenkamp\textsuperscript{12} — “Spiral”
This algorithm was originally published using a hexagonal lenslet array, but is equally valid for the square lenslet array considered here. The algorithm initially assigns the three spots closest to the pupil center to their directly overlying lenslets. Following this, the algorithm selects each new lenslet one at a time in a predetermined order that spirals outward from the pupil center and matches the geometry of the lenslet array. The location of the spot for each new lenslet is predicted using the spot locations of the closest three lenslets that have already been sorted. The closest actual spot that matches this prediction is then assigned to the lenslet, as long as it is within some maximum distance from the prediction (e.g., one lenslet diameter). The geometry of an outward spiral from the pupil center ensures that changes in wavefront slope between subsequent lenslets are kept as gradual as possible, enabling the algorithm to correctly sort spots even in the presence of high levels of aberration. We refer to this as the “spiral” algorithm.

We adopted an approach in our aberrometry software virtually identical to that suggested by Smith and Greivenkamp; MATLAB code was downloaded from our website at http://www.optometry.unimelb.edu.au/research/biophotonics.html. Our implementation of this algorithm is simplified further in that only the closest two spots are used to predict the location of each new spot in the series. The use of two spots represents the maximum computational simplicity for any solution in which previously located spots are used to predict subsequent spot locations. The other difference in our implementation is that when no spot is found for a given lenslet, instead of recording no data we record the predicted spot location as a “dummy” entry that can be used to make future spot predictions. Such entries enable the algorithm to robustly skip over poorly sensed spots within the pupil (e.g., due to media opacities or low signal strength), and also aid in navigating the circular edge of the pupil with a square-shaped spiral.

2.4 A Note on Fourier-Based Algorithms
The preceding approaches operate in the spatial domain, directly on the Hartmann-Shack images. The strong periodicity of Hartmann-Shack images make them amenable to Fourier analysis also. For example, the first harmonic in the x and y directions can be isolated and shifted to the origin, at which point the inverse Fourier transform is calculated. The phase of the resulting complex array, when scaled appropriately, gives the wavefront slope across the image.\textsuperscript{13,14} However, this slope data is “wrapped” modulo 2\pi and must be “unwrapped.” In a highly analogous way to the spot sorting problem already described, this can be done by stepping in a fast, simple manner through each data point or by adopting more advanced approaches that begin in the pupil center and work outward.\textsuperscript{15} The core problem therefore remains the same as for the other approaches, and so we do not specifically investigate Fourier-based methods here. Note that Fourier-based methods become more desirable when the array size and number of spots is large (due to savings in processing time), and when there is significant noise or a large number of poorly sensed spots.\textsuperscript{13

3 Methods—Analysis
To initially compare each of the preceding algorithms, custom MATLAB code was used to simulate double-pass Hartmann-Shack images obtained from a human eye. Monochromatic aberrations were simulated at 830 nm using a phase plate located in the pupil plane, with the retina located 20.2 mm behind this. The retinal space refractive index was 1.33. Fourier optics was used to propagate a 1.0-mm-diam entry beam to form a realistically blurred point source on the retina. Light from a point on this retinal object was then propagated back out of the “eye” with an exit pupil diameter of 6.0 mm. The pupil plane wavefront was sectioned into 0.4-mm-diam squares to simulate a 0.4-mm-pitch square lenslet array. For each section, the point spread function was computed for a lenslet with a 24.0-mm focal length, and convolved with the retinal spot to simulate the double-pass image of the spot on the detector.

Spot centroids were located by standard image processing functions in MATLAB in two steps. First, a binary image was generated by thresholding at intensity levels of 20\% and the MATLAB function “bwlabel” was used to label the centers of the resulting binary spots. Second, these approximate labels were supplied as arguments to the MATLAB “regionprops” function, along with the original gray-scale image, to determine the center of mass for each spot.

To quantify the range for a given aberration, the aberration was varied in magnitude with a trial-and-error approach until the cut-off point was reached (to within 0.05 \(\mu\)m) at which spots began to be incorrectly assigned to a lenslet. Only negative Zernike coefficients were considered—e.g., the results for defocus represent a myopic eye. Negative units were chosen because the range with the improved algorithms is much larger for hyperopic defocus; since spots do not overlap here it is actually the size of the detector array that limits the range.

As well as determining range for a well-centered eye, we modeled the effects on the achievable range of some common sources of errors in ocular aberrometry. First, without a separate pupil monitor, there may be as much as a 0.2-mm error in lateral pupil alignment for our 0.4-mm-pitch lenslet array; we considered the effects of a 0.2-mm displacement on the range. Second, the incident beam is usually displaced from the optical axis in ocular aberrometry to avoid the corneal reflection. This causes the least-displaced Hartmann-Shack spot to lie away...
from the pupil center, which can introduce errors if not taken into account. To simulate this effect we modeled a 1.2-mm error in the deviation of the beam from the optical axis. Finally, when the SNR is low, some spots may not be adequately detected. We randomly excluded 10 of the 177 spots in the pupil from the analysis and measured the range achievable (the same 10 spots were excluded for each analysis).

Following the computational modeling just described, we employed a real adaptive optics system to offer a more “real-world” test of each algorithm. The system consisted of a Hartmann-Shack wavefront sensor with identical parameters to that modeled here, in combination with a Mirao 52-d deformable mirror from Imagine Eyes. A model eye was used, consisting of a +20 D lens positioned 3 mm in front of an adjustable aperture (the pupil). A piece of white paper acted as the retina.

To sample the large number of wavefront shapes achievable with the deformable mirror, we systematically applied random voltage signals to the 52 actuators. The random commands were generated on the range –1 to +1 (the maximum signal possible to send to the mirror), scaled by some scaling factor. Scaling factors from 0.05 to 0.50 were used, in increments of 0.05. This means that in the highest amplitude condition it was possible for neighboring actuators to have a difference in voltage of 1.0. For each scaling factor, 100 random sets were generated and applied to the mirror actuators, and the resulting Hartmann-Shack image was recorded. We therefore recorded 1000 separate images resulting from the mirror undergoing various shapes and amplitudes of deformation. Each image was processed in the same way as for the preceding modeling (a higher threshold was used due to the presence of noise in the real images) and analyzed with each of the algorithms. In the case of the Zernike algorithm, it is critical to know the location of the input beam; to measure this we placed a white card in the pupil plane and estimated the centre of the ~1.2-mm-diam spot that was produced on the wavefront sensor CCD.

Since in this real system case we did not have definitive knowledge of spot-lenslet correspondence, we relied on the internal error checking of each algorithm to reject spots that were probably not correctly assigned. This is a possible source of error, although each of the algorithms used rejects spots in a similar and generally sensible way (primarily by comparing predicted positions to found positions and rejecting those that are greater than a certain number of pixels away). Relying on this internal error checking, we arbitrarily defined failure as 5 or more spots being missed of the ~280 spots across the 7.5-mm pupil. This criterion is slightly more lax than the single-spot criterion used in the preceding computational modeling, but it seemed necessary to account for occasional poorly sensed spots.

4 Results and Discussion

The results for the computational modeling are graphed in Fig. 2 under the various conditions already described for defocus, horizontal-vertical astigmatism, vertical coma, and spherical aberration. Data for the various aberration types were stacked to aid visualization. All units are normalized Zernike coefficients in micrometers, in accordance with reporting standards for vision science, and were fit over a 6.0-mm pupil. Coefficients are all negative in sign (e.g., the defocus term represents a myopic eye).

---

**Fig. 2** Range for each Hartmann-Shack sorting algorithm for various types of aberration, expressed in units of (negative) normalized Zernike coefficients for a 6.0-mm pupil. Astigmatism is horizontal-vertical and coma is vertical. In the “decentered pupil” case, there was a 0.2-mm error in assessment of the pupil center; in the “decentered beam” case there was an uncompensated 1.2-mm decentration of the input beam to the eye; in the “spots missing” case there were 10 randomly chosen spots missing from across the pupil.
For the well-centered “normal” configuration, the ranges of each iterative algorithm were greatly improved over the conventional algorithm and were generally comparable to each other. As predicted, the B-spline algorithm showed the greatest range. Up to 16.6 D of defocus and 32.9 D of astigmatism could be accurately measured with this algorithm and our chosen lenslet array. Compared to the B-spline algorithm, the spiral algorithm showed identical range for defocus and coma, but 15 to 17% less range for astigmatism and spherical aberration. The Zernike algorithm showed equivalent range for coma and spherical aberration, but 16 to 18% less range for defocus and astigmatism, again compared to the B-spline algorithm.

The robustness of each algorithm to common sources of error is also explored in Fig. 2. We can see that none of the algorithms suffered a significant decrease in range when random spots were missing in the pupil; despite their serial nature, there were no downstream errors for poorly sensed spots. This is by virtue of the large number of spots considered when making each prediction for the B-spline and Zernike algorithms, and by the use of the “dummy” grid for the spiral algorithm. In fact, the B-spline and spiral algorithms were also highly robust in the presence of the other error types considered; range decrease was no more than 6 to 10% across all of the error conditions for these two algorithms. The Zernike algorithm was not robust against decenteration in the estimated position of either the pupil or the input beam, making it advisable when using this method to use a calibrated scale when decentering the input beam, and possibly a separate pupil monitor if maximal range is desired.

It is of interest to know whether the range limits shown here could be improved on with some other as yet undiscovered sorting algorithm, or whether the obtained limits were a result of physical spot overlap (i.e., limited by image processing difficulties rather than the sorting algorithm). To test this, we repeated the range measurements by using the predicted spot locations from geometric optics rather than relying on the MATLAB centroiding functions. The range for the B-spline algorithm increased by 22 to 42%, depending on the aberration type, confirming that for the isolated aberrations considered here this algorithm surpassed the practical limit set by the overlapping of spots. We can therefore confidently state that any future algorithm would not provide a greater range. For reference, the Zernike algorithm showed minimal improvement in range under the same conditions, and the spiral algorithm showed an improvement only for defocus. This means that as far as a pure sorting task is concerned the B-spline algorithm is superior to the other two algorithms, but that this advantage is diminished under practical conditions due to spot overlap.

It is also important to consider the range results in the context of the maximum levels of each type of aberration that the algorithms can correctly sort the spots; the other algorithms show varying amounts of error. Even allowing for a ~2-D variation over the central 40 deg of the visual field, this is still sufficient to account for >99.9% of the normal human eye population. Since we have already shown that the improved algorithms operate at or close to the hard limit set by the physical overlap of spots, a larger range would require a lenslet array with a shorter focal length (or some other hardware alteration as discussed in Sec. 1). The range for defocus for each algorithm can be adjusted as a linear function of the focal length of the lenslet array used. For example, if a lenslet array with half the focal length were used instead, a roughly equivalent spot pattern would result only when twice as much defocus is present. This is because the wavefront slope, and therefore the spot displacement, for pure defocus varies linearly in x and y across the pupil.

Spherical aberration is very rarely greater than ~0.4 μm over a 6.0-mm pupil in the normal population, and changes little off-axis, although it may reach ~1.0 μm in eyes suffering from corneal pathology. This is well within range of the 2.4 to 2.9 μm shown for the improved algorithms here. Coma is very rarely greater than ~0.5 μm in the normal population and can change by as much as ~0.6 μm over the central 40 deg of the visual field, which is similarly well within the improved ranges seen here of ~4.6 to 4.8 μm. However, keratoconic eyes have been reported to feature up to ~3.6 μm of vertical coma, which may give any of the algorithms considered here difficulty when combined with other aberration types, due to potential spot overlap. If it were desirable to accommodate a keratoconic

![Fig. 3 Example wavefront from the set of 1000 randomly generated wavefronts. This wavefront was generated with the maximum amplitude considered here. Large circles indicate spots that were correctly located according to the internal error checking of each algorithm. We can see that in this specific example, the spiral algorithm was able to correctly sort the spots; the other algorithms show varying amounts of error.](https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics/15/6/067004-5/067004-067004.jpg)
The preceding results quantify the performance of each algorithm for particular aberration types in isolation. We next used a real adaptive optics system with a model eye to simulate a worse-case scenario—randomized, irregular wavefronts composed of many different types of aberration. The amplitude of the random wavefronts was varied systematically by altering the allowed amplitude of the commands applied to the deformable mirror, with 100 different wavefronts generated for each of the 10 amplitude conditions. Figure 3 shows an example Hartmann-Shack image that was generated using the maximum amplitude, and illustrates the attempts of each algorithm to analyze it. Closed symbols denote spot positions located by the image processing routine. Circles indicate spots that a given algorithm judged were unambiguously assigned to a lenslet. In this particular example, the spiral algorithm did not make any errors, the B-spline made a small number of errors, and the conventional and Zernike algorithms made a large number of errors.

Figure 4 plots the average rate of failure for each algorithm as a function of the maximum allowed actuator amplitude. The failure criterion was met if 5 or more spots (of the ~280 in the pupil) were missed. The error bars indicate standard error of the mean.

The spiral and B-spline algorithms achieved similar results over much of the amplitude range tested; the spiral algorithm was superior at the highest end of the scale. This is probably because the spiral algorithm is far more local in nature and so is somewhat better suited to handle large wavefront deviations that are not well correlated with deviations in other parts of the pupil (the wavefront measured in Fig. 3 is a good example of this). There is a large difference evident between the Zernike algorithm and the spiral and B-spline algorithms. This is likely because the randomly generated mirror shapes would not often be well represented by the Zernike expansion. Note that this algorithm still performed significantly better than the conventional algorithm, and that the high-amplitude, irregular aberrations produced here are unlikely to be seen in eyes without significant disease affecting the ocular media.

Finally, we report on the processing time for each algorithm. Absolute processing times, of course, depend highly on the computer hardware used. However, even the relative processing times given here should be taken as a rough approximation, since the MATLAB code for each algorithm has not necessarily been optimized for speed. With that in mind, with 10 trials for the well-centered case on an Intel Core i7 2.66 GHz processor with 2 GB RAM and using a 17 x 17 lenslet array (i.e., detector size of 6.8 mm), we obtained processing times of 2.3, 4.2, 28, and 41 ms for the conventional, spiral, Zernike, and B-spline algorithms, respectively. This excluded components of each algorithm that involved the initialization of variables, which would optimally be performed only once in a session. The spiral algorithm was faster than the other improved algorithms, as predicted due to the maximally simplistic calculations involved in predicting each spot location.

However, the magnitude of temporal variations in ocular aberrations for a given retinal point are generally small and well within the range of the conventional algorithm. A fast frame rate could therefore be achieved by employing a single iteration of any improved algorithm and feeding these results into the conventional algorithm to guide its search. This approach offers a clear benefit to frame rate because of the unique ability of the conventional algorithm to take advantage of parallel processing and so achieve even greater speed than reported here.

5 Conclusion

The range of aberrometry was similar for the improved algorithms considered here, despite the differences in processing...
complexity. These algorithms approached the physical limits to the achievable range imposed by the overlap of Hartmann-Shack spots. The results presented here show that each algorithm is more than sufficient to accurately characterize the vast majority of human eye aberration, both on- and off-axis, with our chosen lenslet array parameters. The B-spline algorithm produced a high range and was highly robust to error, but was an order of magnitude slower than simpler algorithms. The Zernike algorithm was somewhat faster, but care must be taken to accurately estimate decentration of the pupil and of the input beam, and performance is noticeably reduced for wavefronts that are not well approximated by the Zernike expansion. The spiral algorithm was highly robust to error and to irregular aberration types. It was also 7 to 10 times faster than the other improved algorithms, affording the potential for real-time analysis. Real-time analysis for ocular aberrometry could alternatively be achieved by employing only a single initial iteration of any improved algorithm to inform subsequent iterations of the conventional algorithm, since the vast majority of human eye aberration is static.

Acknowledgments

This work was supported by the Australian Research Council (Discovery Project Grant DP0984649), and the Rowden White Benevolent Bequest.

References