1040-5488/04/8105-0383/0 VOL. 81, NO. 5, PP. 383–388
OPTOMETRY AND VISION SCIENCE


ORIGINAL ARTICLE

# Unwrapping Hartmann-Shack Images from Highly Aberrated Eyes Using an Iterative B-spline Based Extrapolation Method

**LINDA LUNDSTRÖM, MSc and PETER UNSBO, PhD**

*Biomedical and X-ray Physics, Royal Institute of Technology, Stockholm, Sweden*

**ABSTRACT: *Purpose.* When the wavefront aberrations of the eye are measured with a Hartmann-Shack (HS) sensor, the resulting spot pattern must be unwrapped, that is, for each lenslet the corresponding spot must be identified. This puts a limitation on the measurable amount of aberrations. To extend the range of an HS sensor, a powerful unwrapping algorithm has been developed. *Methods.* The unwrapping algorithm starts by connecting the central HS spots to the central lenslets. It then fits a B-spline function through a least squares estimate to the deviations of the central HS spots. This function is then extrapolated to find the expected locations of HS spots for the unconnected lenslets. The extrapolation is performed gradually in an iterative manner; the closest unconnected lenslets are extrapolated and connected, and then the B-spline function is least squares fitted to all connected HS spots and extrapolated again. *Results.* Wavefront aberrations from eyes with high aberrations can be successfully unwrapped with the developed algorithm. The dynamic range of a typical HS sensor increases 3.5 to 13 times compared with a simple unwrapping algorithm. *Conclusions.* The implemented algorithm is an efficient unwrapping tool and allows the use of lenslets with a low numerical aperture and thus gives a relatively higher accuracy of measurements of the ocular aberrations. (Optom Vis Sci 2004;81:383–388)**

Key Words: B-spline function, eye, Hartmann-Shack sensor, ocular wavefront aberrations, unwrapping

It has become increasingly popular in vision research to measure the ocular aberrations with a Hartmann-Shack (HS) wavefront sensor.[1] The HS sensor consists of a lenslet array and a charge-coupled device (CCD) detector. Each lenslet focuses a part of the wavefront emerging from an eye onto the CCD detector in the focal plane of the lenslets. If the wavefront is flat and thus free from aberrations, each lenslet will focus its section of the wavefront to a spot right behind the lenslet, to the projected lenslet center. If the wavefront has aberrations, however, its shape will be distorted and the focused light spots will move away from the projected lenslet centers (Fig. 1). The displacement of a focused spot is proportional to the average tilt of the wavefront over the area of that lenslet. The shape of the wavefront can then be reconstructed from these local tilts. This is often done by least squares fitting of derivatives of Zernike polynomials.[1, 2]

A prerequisite of wavefront reconstruction is that each HS spot can be assigned to the correct lenslet. This puts a limit on the measurable shape of a wavefront. If a spot has moved outside the region of its lenslet and closer to another lenslet, it will be harder to assign the spot to the right lenslet: the so-called unwrapping problem. The unwrapping problem occurs when the aberrations are large, for example in subjects with a high degree of uncorrected ametropia, keratoconus, penetrating keratoplasty, or central scotomas and eccentric fixation.[3]

Different approaches to solve the problem with unwrapping involve changes in the optics of the HS system or software-based algorithms. The disadvantage with optical solutions is that they introduce a higher complexity into the system or lower the accuracy of the measurements by decreasing the f-number of the lenslets. Thus, an algorithm that handles the unwrapping is an attractive solution.

This article describes in detail the implementation of an unwrapping algorithm based on extrapolation of B-spline polynomials[4, 5] adapted and extended for measurements of large aberrations in the human eye. This algorithm handles the pure unwrapping problem and, thus, still requires a subsequence wavefront reconstruction with Zernike polynomials. The specific aim in this work was to be able to measure the large off-axis aberrations in eyes with central visual field loss.[3]

## METHODS

The developed unwrapping algorithm needs two inputs: the HS spot positions in the HS image, in arbitrary order, and the positions of the projected lenslet centers. An iterative B-spline extrap-

olation method then arranges the HS spots according to their corresponding lenslets.

The method starts by connecting the central HS spots with the closest projected lenslet centers. From these connected pairs the locations of the HS spots, corresponding to the other lenslets, are predicted. The prediction is performed by fitting a polynomial function to the displacements of the connected HS spots. This function is then extrapolated outward to the closest not-yet-connected lenslets, to give expected locations of their HS spots. A new HS spot will be connected if it lies close enough to one of these expected locations. The algorithm continues in an iterative manner and ends when no more pairs can be connected. The main structure of the method is described in the flowchart in Fig. 2. The following subsections give a more detailed description of the four main parts: creating a grid of parameter points corresponding to the projected lenslets, connecting the central starting points, creating the extrapolation function (the B-spline function), and fitting the B-spline function and extrapolating iteratively out to the closest not-yet-connected lenslets.

The fifth and last subsection gives numerical values for the various parameters used and discusses the implementation of the algorithm.

## Parameter Grid

The algorithm starts by creating a grid of parameter points, which is used instead of the projected lenslet centers in the extrapolation. The parameter points correspond to the lenslets, with the same spacing and rotation, but with a displaced location. The central point of the parameter grid coincides with the central HS spot, that is, the spot closest to the center of mass of the set of HS spots and thus close to the middle of the pupil. This central parameter point and its HS spot are defined to correspond to the closest lenslet, which means that the wavefront is assumed to be almost flat in the center of the pupil. This is assumed because it is not, by software-based means, possible to unwrap a large overall tilt of the wavefront.[6]

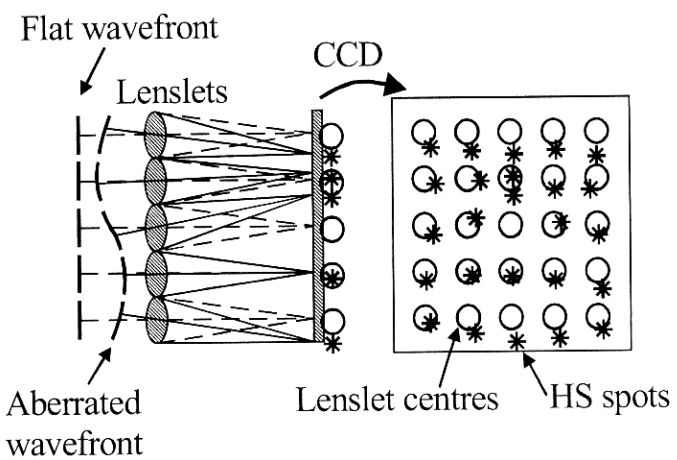


**FIGURE 1.**
The function of an HS sensor. The unwrapping problem arises when the HS spots have moved too far away from the projected lenslets.

## Starting Pairs

The extrapolation procedure needs some initial connections between parameter points and HS spots. Because the wavefront is assumed to have smaller aberrations in the middle of the pupil, it is convenient to start here. The central parameter point and the central HS spot are thus the first connected point pair. The algorithm then connects the parameter points around the first with the closest HS spots, as shown in the enlarged square in Fig. 3. The displacements of these HS spots from their corresponding parameter points serve as the starting values for the iteration. If the wavefront is tilted, these displacements will only be correct by a modulus of a lenslet diameter. This will affect the first-order polynomials, but not the higher-order information.

## B-spline Function

When an extrapolation to connect the parameter points and the HS spots is performed, the quality of the prediction depends on

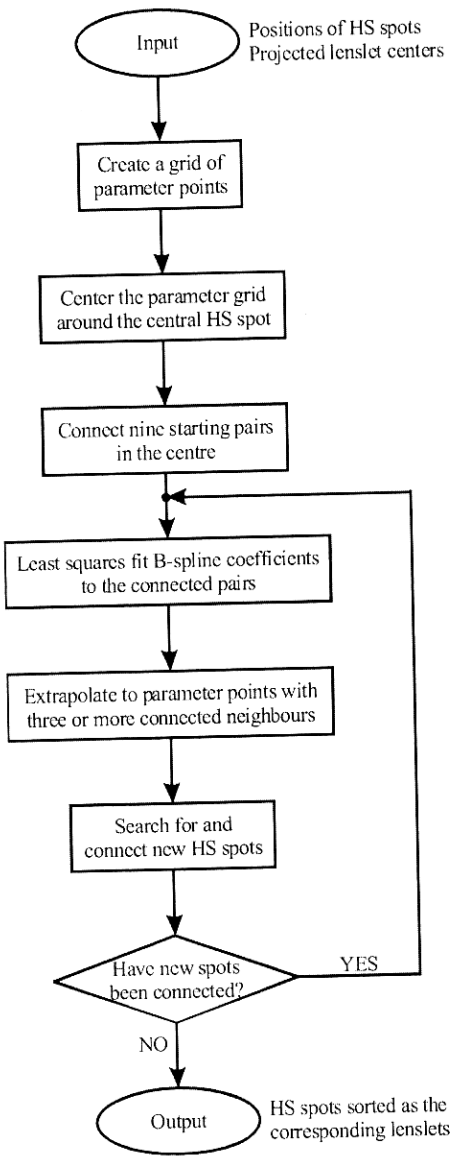


**FIGURE 2.**
The structure of the unwrapping algorithm.

how well suited the extrapolation function is for the wavefront aberrations it should describe. Spline functions are especially appropriate for describing aberrations with different characteristics in different regions of the pupil. A spline function consists of a number of subfunctions, each describing the displacements of the HS spots in different areas of the pupil. The current algorithm uses a special kind of spline functions, B-spline functions, which are well suited for extrapolation because they are well behaved and do not diverge at the edges of the pupil.[4] B-spline functions are preferable in comparison with polynomials that cover the whole pupil (e.g., Zernike or Taylor polynomials), because the latter need high orders to follow the spot pattern with an increased risk of divergence during extrapolation.

In a B-spline function, $\Phi(x,y)$, the pupil is divided into quadratic intervals (Fig. 3), with one basis function, $B_i(x,y)$, associated with each interval, $i$. The linear combination of these basis functions forms the total B-spline function:

$$\Phi^n(x,y) = \sum_{i=1}^{m} c_i B_i^n(x,y) \tag{1}$$

where $m$ is the number of intervals, $c_i$ are constants, and $n$ is the chosen B-spline order. Each basis function is a complete function, defined over the whole pupil. A basis thus exists everywhere, but is only nonzero in a part of the pupil, where it is a polynomial of order $n$. B-spline functions of the zeroth, first, and second order have basis functions, which in one dimension are defined as[4]:

$$B_i^0(x) = 1 \quad [-0.5 < x < 0.5] \tag{2}$$

$$B_i^1(x) = \begin{cases} (x + 1.5) & [-1.5 < x < -0.5] \\ (0.5 - x) & [-0.5 < x < 0.5] \end{cases} \tag{3}$$

$$B_i^2(x) = \begin{cases} \dfrac{(x+1.5)^2}{2} & [-1.5 < x < -0.5] \\ \dfrac{(x+1.5)(0.5-x)}{2} + \dfrac{(1.5-x)(x+0.5)}{2} & [-0.5 < x < 0.5] \\ \dfrac{(1.5-x)^2}{2} & [0.5 < x < 1.5] \end{cases} \tag{4}$$

Two-dimensional basis functions are then constructed by multiplying the basis in the x-dimension with the same basis in the y-dimension:

$$B_i^n(x,y) = B_i^n(x) B_i^n(y) \tag{5}$$

These basis functions are plotted in one and two dimensions in Fig. 4. In equations 2 to 4, the width of the intervals is normalized to unity and the coordinate x is set to zero in the middle of the associated interval of the basis. The interval borders are thus at x equal to $\pm$ 0.5, $\pm$ 1.5. . .. As can be seen from the equations and Fig. 4, a basis function can be nonzero over more than one interval depending on the order of the basis; the zeroth-order basis, $B^0(x)$, is only nonzero in one interval ($-0.5 < x < 0.5$). The first-order basis, $B^1(x)$, covers two intervals ($-1.5 < x < -0.5$ and $-0.5 <$

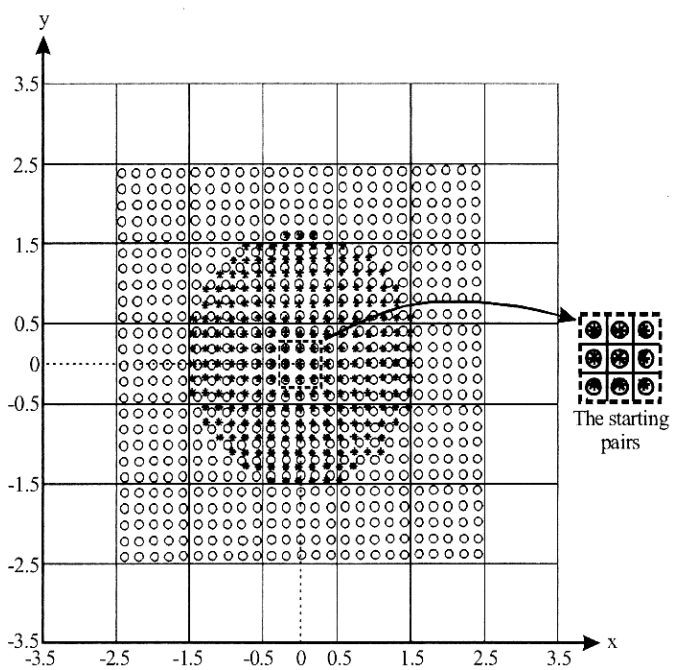


**FIGURE 3.**

The division of the pupil into intervals. The rings represent the parameter points. The stars indicate the HS spots, and the lines show the interval boarders. The enlarged square shows the nine starting pairs in the middle of the pupil.

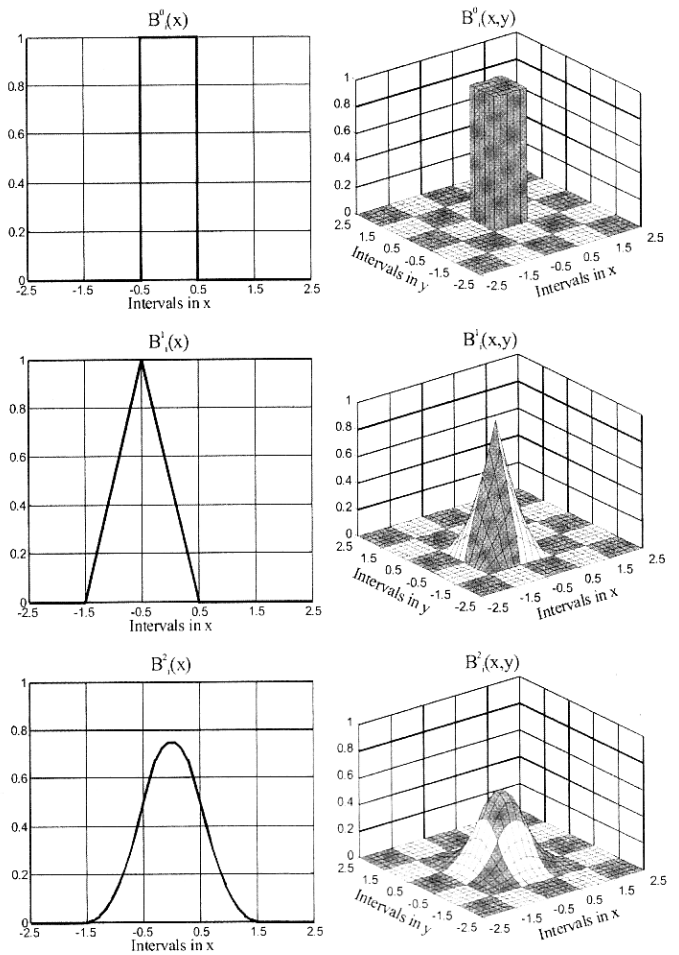


**FIGURE 4.**

B-spline basis functions of the zeroth, first, and second order plotted in one and two dimensions according to equations 2 to 5. The chessboard pattern shows the 25 quadratic intervals.

$x < 0.5$). The second-order basis, $B^2(x)$, covers three intervals and so forth. When the basis functions cover more than one interval, they will partly overlap. This makes the total B-spline function smooth and more flexible.

## Iteration: Least Squares Fit and Extrapolation

When the central HS spots have been connected to the right lenslets and thus to the corresponding parameter points, the rest of the HS spots are unwrapped and connected to the corresponding lenslets in an iterative manner with B-spline extrapolation.

In each iteration, the coefficients, $c_i$, multiplying the B-spline basis functions, are least squares fitted over the region with connected point pairs. This fit is performed with the help of a matrix, $A$, which has one row for each parameter point and one column for each basis function. A matrix element $A_{pi}$ thus contains the value of the $i$th basis in parameter point $p$, with coordinates $x_p$ and $y_p$:

$$A_{pi} = B_i^n(x_p, y_p) \tag{6}$$

$A$ is a sparse matrix and must be calculated only once. The fit is performed for two sets of coefficients: One set, $\mathbf{c_x}$, is fitted to the displacements between the connected parameter points and their corresponding HS spots in the x-direction, $\mathbf{\Delta x}$, and one set, $\mathbf{c_y}$, is fitted to the displacements in the y-direction, $\mathbf{\Delta y}$. These least squares fits are described in matrix formalism as:

$$A^T A\mathbf{c_x} = A^T\mathbf{\Delta x}$$
$$A^T A\mathbf{c_y} = A^T\Delta_y \tag{7}$$

where $\underline{A}$ contains the rows of matrix $A$ that correspond to already connected parameter points. The coefficients are then used in the extrapolation to find the expected locations of the HS spots for neighboring unconnected parameter points. The extrapolation is also performed twice, once with the x-coefficients to give the expected displacements in the x-direction and once with the y-coefficients:

$$\mathbf{\Delta x} = A\mathbf{c_x}$$
$$\mathbf{\Delta y} = A\mathbf{c_y} \tag{8}$$

For each unconnected parameter point, $p$, the displacements $\Delta x_p$ and $\Delta y_p$ define the expected location of a HS spot. If the closest unconnected HS spot lies within a certain radius of the expected location, it is connected to the parameter point. Otherwise the algorithm will leave the parameter point unconnected.

When all neighboring expected locations have been checked, the algorithm starts on the next iteration, with fit and extrapolation to new neighboring parameter points (Fig. 5). The iteration stops when the algorithm cannot find any new HS spots to connect. The output of the algorithm is the HS spots ordered as the parameter points and thus the lenslets.

The algorithm was tested on a number of eyes and on computer-generated spot patterns. The measurements where performed with an HS sensor with 325 × 325 μm lenslets of focal length 18 mm. The study followed the tenets of the Declaration of Helsinki and was approved by the local Research Ethics Committee. All subjects gave informed consent before participation.

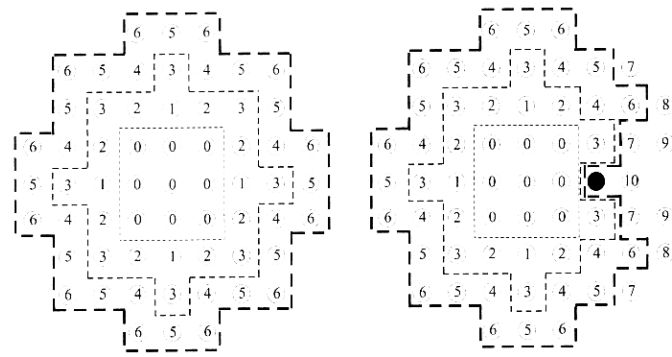

**FIGURE 5.**
The progression of the iteration. 0 = starting points; 1 = parameter points connected in the first iteration; 2 = points connected in the second iteration. The left figure shows the progression when all parameter points are connected, and the right shows how the algorithm goes around a parameter point that has no HS spot (filled circle).

## Implementation

Our implementation uses a two-dimensional B-spline function of the second order, n = 2, as an extrapolation function. This means that each two-dimensional B-spline basis function is nonzero over 3 × 3 intervals (Fig. 4) and that each interval contains nine nonzero basis. On the borders of the 3 × 3 intervals of a basis, the basis is zero with its first derivative also equal to zero. The pupil area that a basis covers depends on the chosen size of the intervals. In this implementation, intervals of the size of 5 × 5 parameter points (lenslets) are used, and thus, each second-order basis is nonzero over 15 × 15 parameter points. The intervals are placed with one interval centered over the nine starting pairs (the starting parameter points are simply connected to the closest HS spot) (Fig. 3).

The second order B-spline function and the interval size of 5 × 5 parameter points were chosen for three reasons. First of all, the second order is the lowest order, which has a derivative continuous over the interval borders. Secondly, higher-order basis would lead to an underdetermined least squares fit of the nine starting pairs. The third-order basis, for example, would require at least 16 starting pairs (16 nonzero basis in each interval) to be well defined. However, it is desirable to have as few starting pairs as possible since the correct connection of the starting pairs is a delicate operation. The third reason is to have a reasonable number of degrees of freedom. The number of degrees of freedom of a B-spline function equals the number of basis functions and, thus, depends on the order of the basis and on the size of the intervals. During the unwrapping, the number of degrees of freedom is increased adaptively as the number of intervals increases when more spots are connected. A pupil of 6 mm covers about 4 × 4 intervals of 5 × 5 parameter points (with a lenslet diameter of 325 μm). With a second-order B-spline function, this gives in total 72 (in x- and y-direction: 6 × 6 × 2) degrees of freedom. This number would correspond to using Zernike polynomials up to the order of 10. If the intervals are reduced to 3 × 3 parameter points, this would give

128 (8 × 8 × 2) degrees of freedom, corresponding to Zernike polynomials up to the order of 14.

During the extrapolation two things must be investigated to ensure that the predicted locations of HS spots are reliable. First, the extrapolation should use only B-spline basis functions, which have a reasonable support in the region of the already connected points. A two-dimensional B-spline basis is defined over a region of 15 × 15 parameter points. If only a few of these 225 points are connected, the least squares fit of the coefficient will be of low quality. The quality will be especially low if the basis gets support only from a corner, where the numerical value of the basis function is very low. The implemented algorithm only includes a basis $i$, if some of its elements in matrix $\underline{A}$ (column $i$) has a value of 0.01 or larger. This means that the basis has support from connected parameter points in a reasonable area of its nonzero region. Second, the extrapolation should not be extended far away from the already connected points. The accuracy of the predicted location of an HS spot depends on the number of neighboring connected point pairs. If, for example, an unconnected parameter point is surrounded by connected pairs, the least squares fit and extrapolation will give a very accurate prediction. Conversely, an extrapolation for a parameter point, without connected neighbors, will produce a predicted location with lower accuracy. It is therefore important to make sure that the parameter points, to be extrapolated, have some connected neighbors. The method proposed here labels each parameter point with the number of connected neighbors. The extrapolation is then performed only for parameter points, which have at least three already connected neighbors. This means that the algorithm will iterate outward in circles (left image of Fig. 5). If an HS spot is missing or too far away from the predicted location, the corresponding parameter point will be left unconnected in that iteration, but will be rechecked in each of the following iterations until it is connected or all HS spots are identified. Thus, this technique enables the algorithm to work around difficult parts in the HS spot pattern and to handle missing HS spots (right image of Fig. 5).

After the extrapolation, the algorithm searches for an HS spot in a circle around the predicted location and the closest HS spot is connected. The radius of the circle sets the maximal distance between the predicted location and the HS spot. In this implementation, the radius is half a lenslet diameter.

## RESULTS

The algorithm has been tested successfully on a number of eyes. It can, without any modifications, handle measured HS spot patterns from myopic, hyperopic, and astigmatic eyes, both on the visual axis and in eccentric angles. Fig. 6 shows some examples of successfully unwrapped spot patterns representing particularly large wavefront aberrations. In Fig. 6A, the HS spot pattern from an eye with a large central scotoma has been measured and unwrapped in an eccentric fixation angle of 22°. The main aberration here is coma. Fig. 6B is a measurement on axis in an eye after penetrating keratoplasty. Although this spot pattern shows a large amount of refractive errors and higher-order aberrations, the unwrapping algorithm is successful. Fig. 6C shows the limit of what the algorithm can handle. It is an eye with keratoconus, and the extreme aberrations in the lower left corner have caused missing spots. This example shows the strength of the developed algorithm; it can work around a missing spot and connect the other spots correctly. The last example, Fig. 6D, corresponds to a computer-generated wavefront and shows that the algorithm can handle a large amount of the most common higher-order aberrations, coma and spherical aberration.

To compare the developed unwrapping algorithm with a simple algorithm, which just connects the parameter points with the closest HS spots, computer-simulated spot patterns from single Zernike mode wavefronts were generated. The maximal Zernike coefficients for defocus, coma, and spherical aberration were found empirically by increasing each coefficient separately, with all others set to zero, until the algorithm could no longer unwrap the spot pattern. For a 6 mm pupil, lenslets of size 325 × 325 μm and a focal length of 18 mm, the B-spline algorithm could handle 9 times more defocus (3.9 μm vs. 35.8 μm), 13 times more coma (1.1 μm vs. 14.3 μm), and 3.5 times more spherical aberration (1.0 μm vs. 3.5 μm).

## DISCUSSION

The unwrapping problem has been discussed much in the field of interferometry, and several different software-based methods have been developed to perform phase unwrapping.[6] In interferometry, the unwrapping problem arises because of the modulo $2\pi$ ambiguity of the fringes. In the field of wavefront measure-

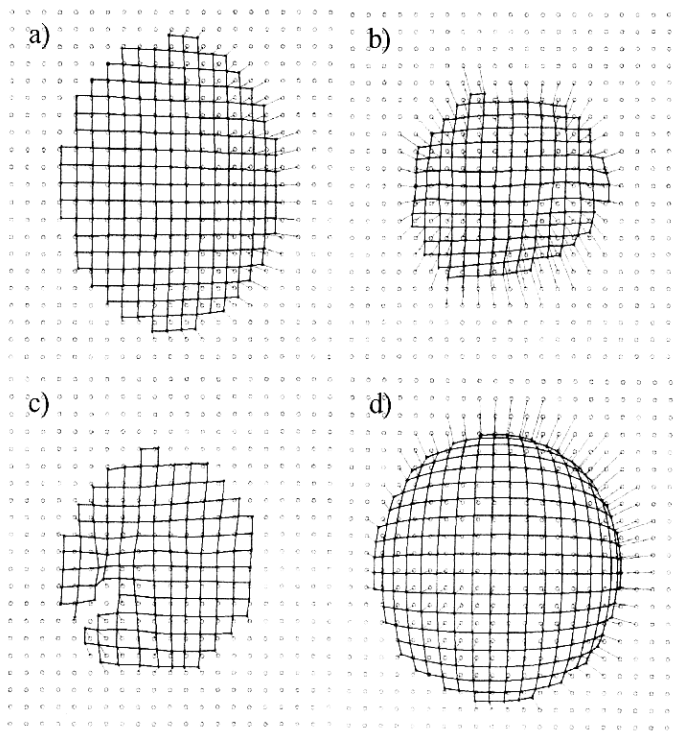


**FIGURE 6.**
Example of HS spot patterns unwrapped by the algorithm. The grid shows how the HS spots have been unwrapped. The HS spots (stars) are connected to their corresponding parameter points (rings) with narrow lines. A: Eccentric fixation in 22°, right eye −2.75 −2.00 ×80, together with considerable coma. B: Eye after penetrating keratoplasty, right eye −6.00 −5.75 ×165. C: Eye with keratoconus and thus very large and irregular aberrations. D: Computer-generated spot pattern with coma and spherical aberration.

ments on eyes with HS sensors, it is more common to avoid the unwrapping problem by changing the optics in the HS system. One approach is to choose lenslets with a lower f-number, that is, a shorter focal length and larger diameter. This will enable the system to measure larger wavefront aberrations without the need of unwrapping. At the same time, however, it will decrease the accuracy of the measurements because of the lower number of sampling points and the relatively smaller movements of the HS spots for the same amount of aberrations. Therefore, some HS sensor systems include extra optical hardware to avoid or solve the unwrapping problem. Some examples are optics to compensate for defocus and astigmatism, a moveable aperture transmitting the spot from one lenslet at a time,[7] astigmatic lenslets for easier identification of the spots,[8] and an additional measurement of the spot positions in a plane between the lenslets and the detector. All these methods are time-consuming, entail a larger complexity, or lower the accuracy. Therefore, there is a lot to gain if the unwrapping can be made by software-based means instead. One simple solution is to rescale the grid of the projected lenslets to cope with more defocus. However, this rescaling cannot handle astigmatism and higher-order aberrations.

Two more advanced methods for software-based unwrapping have previously been suggested: a method with B-spline extrapolation[5] and a sorting method.[9] The method proposed herein has implemented and developed the key ideas with B-spline extrapolation by Groening et al.[5] To our knowledge, it represents the first advanced algorithm for unwrapping HS images from human eyes.

One of the basic assumptions is that the measured wavefront is continuous over the sampling scale; the B-spline function is not suited to handle true discontinuities, and neither are Zernike polynomials. However, as long as the sampling is dense enough, the method should be able to handle most naturally occurring functions.

The main limitation of the implemented algorithm is that if it makes one wrong connection, there is a high risk of more false connections. It is better that the algorithm leaves a spot unconnected instead of connecting it incorrectly. A further development of the algorithm could be to make it possible to go back and recheck the assignments already performed.

## CONCLUSIONS

This article presents an effective software-based method to unwrap HS spot patterns from human eyes without any loss in accuracy. The proposed algorithm performs a least squares fit of a B-spline function to the starting pairs of connected parameter points and HS spots in the center of the pupil. It then extrapolates outward in a stepwise manner for those parameter points that have at least three connected neighboring parameter points. The algorithm can handle large amounts of regular and irregular wavefront aberrations and is thus a useful tool for investigating high aberrations in, for example, peripheral vision.

## ACKNOWLEDGMENTS

*Supported by the Göran Gustafsson Foundation, the Carl Trygger Foundation, and the Carl-Johan and Berit Wettergren Foundation.*

*Presented, in part, as a poster at the annual meeting of the Association for Research in Vision and Ophthalmology on May 7, 2002, in Fort Lauderdale, FL.*

*Received April 17, 2003; revision received January 22, 2004.*

## REFERENCES

1. Liang J, Grimm B, Goelz S, Bille JF. 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.
2. Malacara D. Optical Shop Testing, 2nd ed. New York: Wiley, 1992.
3. Gustafsson J, Unsbo P. Eccentric correction for off-axis vision in central visual field loss. Optom Vis Sci 2003;80:535–41.
4. De Boor C. A Practical Guide to Splines. New York: Springer-Verlag, 1978.
5. Groening S, Sick B, Donner K, Pfund J, Lindlein N, Schwider J. Wave-front reconstruction with a Shack-Hartmann sensor with an iterative spline fitting method. Appl Optics 2000;39:561–7.
6. Ghiglia DC, Pritt MD. Two-Dimensional Phase Unwrapping: Theory, Algorithms and Software. New York: Wiley, 1998.
7. Olivier S, Laude V, Huignard JP. Liquid-crystal Hartmann wave-front scanner. Appl Optics 2000;39:3838–46.
8. Lindlein N, Pfund J. Experimental results for expanding the dynamic range of a Shack-Hartmann sensor using astigmatic microlenses. Opt Eng 2002;41:529–33.
9. Pfund J, Lindlein N, Schwider J. Dynamic range expansion of a Shack-Hartmann sensor by use of a modified unwrapping algorithm. Opt Lett 1998;23:995–7.

**Linda Lundström**
*BIOX*
*Biomedical and X-ray Physics*
*Royal Institute of Technology*
*AlbaNova*
*Roslagstullsbacken 21*
*106 91 Stockholm, Sweden*
*e-mail: linda@biox.kth.se*