Improveing the Repeatability of Topographic Height Measurements in Confocal Scanning Laser Imaging Using Maximum-Likelihood Deconvolution

Andrew J. Patterson,1 David F. Garway-Heath,2 and David P. Crabb1

Purpose. To evaluate maximum likelihood (ML) blind deconvolution as a technique for improving the repeatability of topographic height measurements obtained from scanning laser tomography (Heidelberg Retinal Tomograph [HRT]; Heidelberg Engineering, Heidelberg, Germany).

Methods. ML blind deconvolution is an image-processing technique that estimates the original scene from a degraded image. This technique has been used in confocal scanning laser microscopy to remove “out-of-focus” haze in three-dimensional confocal image stacks. ML blind deconvolution requires no prior estimation of the point-spread function (PSF), as opposed to classic linear deconvolution methods. Instead, the algorithm estimates an initial PSF based on the optical setup of the confocal scanning device and optics of the eye and iteratively proceeds to a solution. The improvement in repeatability of height measurements from mean topography images within scan (intrascan) and between scans (interscan) afforded by ML deconvolution was evaluated in a test–retest series of HRT images from 40 ocular hypertensive and glaucomatous patients with varying degrees of media opacity.

Results. There was an improvement in intrascan repeatability in 38 out of the 40 mean topography images (median improvement 2.5 μm, interquartile range 2.19, P < 0.001), and an improvement in interscan repeatability in 33 of the 40 mean topographies (median improvement, 1.0 μm, interquartile range 3.49, P < 0.001). There was a positive association between the magnitude of the improvement in repeatability and the level of mean pixel height standard deviation (MPHSD), intrascan (P = 0.004) and interscan (P = 0.002).

Conclusions. ML blind deconvolution algorithm improves the repeatability of topographic height measurements from the HRT. This improvement was greater in patients with poorer quality images. (Invest Ophthalmol Vis Sci. 2006;47:4415–4421) DOI:10.1167/iovs.06-0191

From the 1Department of Optometry and Visual Science, City University, London, United Kingdom; and 2Glaucoma Research Unit, Moorfields Eye Hospital, London, United Kingdom.


Supported in part by funding from the Moorfields Eye Hospital Special Trustees, London.

Submitted for publication February 22, 2006; revised May 8 and 31, 2006; accepted August 3, 2006.

Disclosure: A.J. Patterson, AutoQuant, Inc. (F); D.F. Garway-Heath, Heidelberg Engineering (F); D.P. Crabb, None

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked “advertisement” in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: David P. Crabb, Department of Optometry and Visual Science, City University, Northampton Square, London E1C 0HB, UK; d.crbab@city.ac.uk.

Confocal scanning laser tomography (CSLT) is an established technology for the examination of the posterior segment of the eye.1-2 The commercially available Heidelberg Retina Tomograph (HRT; Heidelberg Engineering, Heidelberg, Germany) is an exemplar of a clinically used version of this instrumentation. It is widely used in the assessment of eye diseases and is of particular importance in monitoring the glaucomatous optic nerve head (ONH).3-5 Analysis of these images has been used to distinguish between normal and glaucomatous eyes.3-6 Longitudinal series of images from a single patient can be used to monitor glaucomatous damage over time.7-9 However, a recent large population study using scanning laser tomography showed that satisfactory images (defined as average repeatability of topographic height <68 μm) could not be obtained in 10% of a normal elderly population.10 A test–retest study of the HRT in an ocular hypertensive (OHT) and glaucomatous population indicated similar results (11% >68 μm).11 Computer simulations show that improving the repeatability of image series increases the sensitivity of techniques in detecting glaucomatous damage.7

Confocal scanning laser tomography is a special application of confocal microscopy that regards the ocular fundus as the object. The resolution of the raw 3-D images obtained by the technology is roughly pencil-shaped, with a lateral resolution of ∼10 μm in diameter and a depth resolution of ∼300 μm. A topography image is generated by determining the position of peak reflectance in the confocal stack at each pixel. Reflecting layers within the retina (e.g., pigment epithelium, nerve fiber layer, inner limiting membrane) vary in thickness from 200 to 500 μm. Because of the limited depth resolution (300 μm) of the technology, the topography image can be thought of as the position of the mean depth of the interfaces between these layers.12

Campbell and Gubisch13 demonstrated that the reflective properties of the fundus best approximate a diffuse surface, they also showed with Fourier analysis that the optimal pupil diameter is between 2 and 3 mm for imaging applications. If the resolution of optical sectioning of the human retina achieved by CSLT were only limited by diffraction then, according to estimates from theoretical models, the lateral resolution would be between 1.6 and 2.3 μm, and the depth resolution would be between 45 and 48 μm.14,15 These resolutions would lead to topographic height measurements with an axial position precision of ∼2.3 μm (whereas the current optimal precision is ∼15 μm).16 However, the resolution achieved in practice is limited by the optics of the human eye, specifically the lens, cornea, and tear film, whereas laser safety standards for the human eye and the confocal pinhole size also inhibit resolution. The lens, cornea, and tear film produce wave aberrations that are unique for each eye. Low-order aberrations can be compensated for by spectacles or lenses placed in front of the instrument optics; however, high-order aberrations remain. Pupil dilation improves resolution up to a critical point. The pupil dilation for optimal lateral resolution has been shown to be different from that which achieves optimal depth resolution.17 Venkateswaran et al.15 demon-
strated how the achievable depth resolution increases with increasing pinhole size.

If we now consider an ideal imaging system, the image acquired would be a scaled version identical to the true object being imaged. Image-restoration techniques are a collection of image-processing techniques used to estimate the true object of interest given the image obtained and information of the nature of the blurring and noise in the imaging system. Image-restoration algorithms can be classified as linear and nonlinear. Nonlinear techniques, based on computationally intensive methods have increasingly gained acceptance, with Jansson et al. finding they deliver superior results to linear methods in a broad spectrum of applications. As computational resources are now faster and cheaper, the historic objections against the beamlinear techniques have become inappropriate. Maximum-likelihood (ML) deconvolution is an example of these techniques: this estimates both the true object and the blur from the degraded image using partial information about the imaging system. Holmes et al. were the first to apply the techniques to images obtained in confocal microscopy. In this study, we applied ML blind deconvolution to series of images obtained from a test-retest study of HRT topography reproducibility. Our motivation is to investigate whether this technique can improve the repeatability of topographic height measurements and hence increase the utility of the HRT as a technique to assess the ONH.

**Methods**

Images of the ONH were acquired by scanning laser tomography with the HRT. The confocal arrangement of the HRT includes a diode laser, beam-splitter, and two optically conjugate pinholes. The beam passes through the center of the pupil and is focused at the retina to a specific point size, its limit set by the properties of the human eye. Reflected light from the beam exits the eye through the pupil, passes a beam-splitter and then goes through a confocal pinhole before reaching a solid state detector. This setup is designed to block scattered light and light from other sources from entering the detector. A deflector mirror then moves the laser beam horizontally so an adjacent point can be imaged. After one line has been acquired, a second deflector mirror moves the beam vertically before acquiring another horizontal line. A two-dimensional (2-D) scan is acquired in this rasterlike fashion in approximately 32 ms (a total of 256 × 256 pixels). The plane perpendicular to the optic axis in which the beam is focused, called the focal plane, can be changed and is moved from anterior to posterior, acquiring a total of 32 equally spaced scans. The resultant 3-D (256 × 256 × 32-pixel) image is referred to as a confocal image. The total acquisition time is 1.6 seconds. More details are provided by Zinser.

The 3-D confocal image acquired is postprocessed in a step to compensate for random eye movements, such as microsaccades and slow drift, known to occur when a subject is asked to fixate on a target. A topography image is then formed by calculating the position of maximum reflectivity at each z-profile (a one-dimensional signal of 32 intensity values parallel to the optical axis). The intensity at each pixel in the resultant topography image represents the surface height of the ONH or surrounding papillary retina. Typically, three topography images are acquired at each visit. The topographies are normally merged to calculate a mean topography. Image registration algorithms within the proprietary HRT software align the topography images for interscan differences in examination positions.

The repeatability of the images obtained was quantified by the mean pixel height standard deviation (MPHSD). This metric is effectively a gauge of the variability of each pixel height measurement across the three topographies used to make up the mean topography and is used in this study to report intrascan repeatability. It is calculated from the standard deviations at each pixel across the mean topographic image (i.e., the MPHSD is the mean of 256 × 256 pixel height standard deviations). Previous studies have used MPHSD to evaluate the repeatability of the technology in normal subjects and patients with glaucoma. It has been shown that MPHSD is influenced by lens opacity, age, and degree of astigmatism. In this study, MPHSD was also used to report the repeatability of mean topographies, referred throughout as interscan repeatability.

Confocal scanning laser tomography has known limitations. For example, although the optical setup is designed to reject most light from outside the focal plane, it by no means rejects all it and an out-of-focus haze remains. The resolution in the confocal images is higher in the x and y directions compared with poorer resolution along the optical axis (z-axis). The resolution obtained is also limited by the optics of the eye, aberrations are generated by the cornea and the lens. This results in axial smearing. For example, if a spherical point object is being imaged with constant reflectivity properties, the resultant image obtained will appear elongated in the z-axis. Another limitation of the technology is that the detector in the optical setup, prone to Poisson noise, primarily due to quantum variations in the number of photons recorded. This noise obscures real data and randomly creates impossible features such as high-intensity data only one pixel in size. A further discussion on the principals and limitation of imaging systems is given by Goodman.

Deconvolution is an example of an image restoration algorithm that models the imaging system with:

\[
g(r) = f(b(r)r)/r^2dr + n(r)
\]

where \(g(r)\) is the image obtained in direction \(r\), \(f(r)\) is the true image, \(n(r)\) is noise, \(b(r)r\) is the point-spread function (PSF): the image brightness at location \(r\) of a point source located at position \(r'\). The PSF describes how much a single-point source of light is spread through the focal planes. The image formed by a system \(g(r)\) is a convolution of the PSF (across the whole geometrical image area) with the brightness at each point source. The wider the PSF the more blur the image will contain. In confocal scanning laser tomography, the PSF is assumed to have a three dimensional hour-glass shape, orientated along the optical axis and of highest intensity in the central “narrow” area. The objective of image restoration algorithms is to obtain an estimate of the true image \(f(r)\), given the image obtained \(g(r)\). In classic linear deconvolution, the PSF \(b(r)r\) is assumed to be known explicitly before the procedure. A long list of these techniques is available, such as the inverse filter and Wiener filter. Unfortunately, in our situation, the blur \(b(r)r\) is unknown, along with much information about the true image \(f(r)\). Blind deconvolution refers to the task of separating two convolved signals \(f(r)\) and \(b(r)r\), when both signals are either unknown or partially known.

**ML Deconvolution**

The ML estimation is a mathematical optimization strategy designed to produce the best guess of true data that have been corrupted by random noise. It is an adaptation of the Richardson-Lucy optimization strategy. The ML deconvolution approach is known to fail, however, unless strong constraints can be applied to the properties of the PSF. Holmes et al. published suitable constraints for data obtained by confocal microscopy, using assumptions about the shape of the hour-glass PSF and optimized for the mathematical nature of the noise present. Previous studies evaluated ML deconvolution in confocal microscopy on simulated and real data. In this study we applied ML deconvolution, using software developed by Holmes et al., with a commercially available system (AutoDeblur; ver. 9.3.6; AutoQuant Imaging, Inc., now Media Cybernetics, Silver Spring, MD).

The deconvolution software requires input of the optical setup and image medium to allow it to approximate and iteratively constrain the solution of the PSF. The numerical aperture of the lens (0.08), wavelength of the laser beam (680 nm), refractive index of the medium being imaged (assumed to be close to water, 1.33), and spacing of the image obtained is input into the software. For a 10° scan, the x and y spacing are 11.4 μm With the HRT, the depth of scan can be varied.
from 2 to 4 mm, while always acquiring 32 scans. The z-spacing is set appropriately, depending on the depth of the scan. The z-spacing is set as scan depth divided by 32, and this is directly entered into the deconvolution software.

Clinical Data

The techniques were applied to a test–retest data set. The study protocol adhered to the Declaration of Helsinki and had local ethics committee approval; the informed consent of subjects was obtained. All subjects attended the OHT clinic at Moorfields Eye Hospital, London. Strouthidis et al.11,38 originally collected the data to evaluate the test–retest variability of the HRT and HRT II. The study design was originally developed to test interoperator and intervisit repeatability. Patients were not excluded on ONH appearance; patients were excluded on the basis of myopia greater than 12 D of spherical power or any history of intraocular surgery. The eye with greater media opacity was selected preferentially. A total of five mean topographies was acquired from each subject, with three single topography images on the basis of myopia greater than 12 D of spherical power or any history of intraocular surgery. The eye with greater media opacity was selected preferentially. A total of five mean topographies was acquired from each subject, with three single topography images combined to form a mean topography. The images were obtained over a 6-week period in two visits to the clinic. On the first visit, three mean topographies were obtained, and in the follow-up visit a further two were obtained. All subjects had had experience with scanning laser tomography, having been imaged at least three times previously.

Comparison

In this study, we investigated whether ML deconvolution would result in an improvement of intrascan (within-scan) and interscan (between-scan) topographic height measurement repeatability. For this objective, 40 patients’ image series (HRT Classic) were randomly selected from the test–retest dataset.

Measuring Intrascan Repeatability

The mean topography on the first visit was obtained, and the MPHSD was recorded from the HRT software (version 2.01b). Three confocal stacks used to generate the mean topography images were deconvolved. The images were then re-entered into the HRT software and three single topographies were calculated. It was observed that deconvolution had induced an artifact at the border of the single topographies. This is a known phenomenon and Gonzalez et al.59 suggest using an edge-taper function to blur the edge of the image to minimize this effect. As a local spatial filter may affect the results, the artifact at the edges was removed using an image-processing erosion algorithm.52 In this application the erosion algorithm proceeds by labeling pixels with an intensity of 0 as background (or not available), this is the “windowed” area typically seen at the edges of a topography image. The rest of the pixel intensities (1 to 255) are set as foreground. The erosion algorithm proceeds by removing a layer of foreground pixels at the boundary edge between the foreground and background pixels. This makes the windowed area at the edges of the topography image larger. In this application, 5 pixels were eroded from the boundary edge. The mean topography was then calculated, and the MPHSD was recorded for comparison with the preprocessed data.

Measuring Interscan Repeatability

Three mean topographies were randomly selected for each patient. For the purpose of quantifying the repeatability across scans, the mean topography images were input into the HRT database as single topographies. The proprietary image registration algorithms therefore spatially aligned the mean topographies and the computed MPHSD was recorded. Nine confocal images associated with the three mean topography images were deconvolved. Nine single topographies were then generated and the edge artifact was removed as before. Three mean topographies were generated and input into the HRT software as single topographies, and the MPHSD quantifying the interscan repeatability was recorded.

Deconvolution in Scanning Laser Tomography 4417

The association between lens opacity and the effect deconvolution had on repeatability was investigated. The lens opacity was measured subjectively with the Lens Opacity Classification System (LOCS) III grading system. Nuclear opalescence (NO), nuclear color (NC), and posterior subcapsular (PS) and cortical (C) scores were recorded.

A paired nonparametric (Wilcoxon) test was performed to determine whether the improvement in repeatability was statistically significant, because the distributions of differences in MPHSD had a positive skew. Spearman’s rank correlation test was used to detect whether there was an association between the average MPHSD and the improvement in repeatability before and after deconvolution. The association between improvement and lens opacity was also quantified in this way.

RESULTS

An example of a preprocessed and postprocessed confocal image from different viewing perspectives is shown in Figure 1. The raw confocal image is in the left-hand column, and the deconvolved image is the right-hand column. Figures 1c and 1d compare the maximum reflectance images from the x-z plane; this viewing perspective can be thought of as a side elevation of the 3-D confocal stack. The reduction in axial smearing is apparent after deconvolution. An individual slice in Figures 1e and 1f suggests a reduction in high-frequency noise. Two z-profiles, before and after deconvolution, are plotted in Figures 1g and 1h. The two z-profiles represent a pixel located in the neuroretinal rim (marked by an arrow in Fig. 1a), an area of the image that typically has low light reflectance. In this patient with lens opacity of NO = 2.7 and NC = 2.2, a high amount of noise was seen in the raw z-profile, whereas the deconvolved z-profile showed low noise. Deconvolution improved the resolution—most apparent along the optical (z) axis, and reduced high frequency noise.

The effect deconvolution had on intrascan repeatability is summarized in Figure 2, which shows the average MPHSD plotted against the difference between MPHSD before and after deconvolution. An improvement in repeatability after deconvolution results in a point being above the zero line parallel to the x-axis. An improvement in MPHSD occurred in 38 of the 40 images. The median improvement of 2.5 μm (interquartile range, 2.19) is statistically significant (P < 0.001). The figure also demonstrates that the improvement was greater in subjects with higher MPHSD. There is a statistically significant association between average intrascan MPHSD and the improvement in repeatability r = 0.45 (P = 0.004). The association between LOCS III scores and improvement in repeatability was NO: r = 0.37, P = 0.019; NC: r = 0.35, P = 0.040; C: r = 0.22, P = 0.162; and PS: r = 0.10, P = 0.526.

In Figure 3 the same figure format shows the effect deconvolution had on interscan repeatability. An improvement in interscan repeatability of topographic height measurements occurred in 35 of the 40 images, with a median improvement of 1.80 μm (P < 0.001). There was an association between average interscan MPHSD and the difference in MPHSD, before and after deconvolution (r = 0.49; P = 0.002). The association between LOCS III scores and improvement in repeatability was NO: r = 0.25, P = 0.125; NC: r = 0.18, P = 0.254; C: r = 0.15, P = 0.358; and PS: r = 0.12, 0.459.

It takes approximately 3 minutes of computer processing to deconvolve a single HRT image (256 × 256 × 32 pixels) using a standard desk-top computer (Pentium IV 1.6 GHz, Intel, Mountain View, CA).

DISCUSSION

Previously ML deconvolution has been used in confocal microscopy, wide-field epifluorescence microscopy and transmit-
FIGURE 1. Left: raw stack of optic nerve head confocal images acquired by HRT; right: confocal image stack after 30 iterations of ML deconvolution. The maximum projections are in the x-y plane of the raw data and are otherwise known as reflectance images for the original image (a) and deconvolved images (b). The maximum projection in the x-z plane: original image (c) and deconvolved image (d) show axial smearing associated with confocal scanning laser tomography in the original image. There was better discrimination between slices in the deconvolved image. Slice number 15 in the original (e) and deconvolved (f) images shows a reduction in high-frequency noise. Two z-profiles (g) pre- and (h) postdeconvolution are shown at a position in the rim area (marked by the arrow in a).
Deconvolution in Scanning Laser Tomography

Figure 2. Effect of deconvolution on intrascan repeatability of topographic height measures. The plot shows the difference in average MPHSD against the difference in MPHSD, before and after deconvolution. An improvement in repeatability is represented by a point above the x-axis. An improvement in repeatability occurred in 38 of the 40 images (P < 0.001).

Figure 3. Effect of deconvolution on the interscan repeatability of topographic height measures. An improvement in repeatability occurred in 33 of the 40 images (P < 0.001).
highlighted in a recent study that showed that 10% had a MPHSD above 68 μm and a mean MPHSD of 26.8 ± 13.3 (SD) μm with the worst 10% removed. Strouthidis et al. reported in a population with OHT and primary open angle glaucoma, over a mean of five visits, that MPHSD had a mean of 33.5 ± 23.6 (SD) μm and range of 12 to 130 μm. The instrument guidelines categorize MPHSD < 10 as excellent; 10 to 20, very good; 20 to 30, good; 30 to 40, acceptable; 40 to 50, look for ways to improve; >50, low-quality image and suggest not to use as a baseline image. The results highlighted in Figures 2 and 3 suggest that the improvement obtained by deconvolution was greater in these least-repeatable images. For example, the mean improvement in the five patients with an original MPHSD greater than 60 μm was 10.4 μm. Therefore, these techniques may have important real clinical impact in the use of images that are of poor quality, and it is hoped that the results from this study combined with further work to optimize and test the techniques on further data sets will result in scanning laser tomography's having wider utility in the assessment of the ONH.

Acknowledgments

The authors thank Nick Strouthidis for collecting HRT imaging data and lens opacity scores, AutoQuant, Inc. (Media Cybernetics) for providing software and technical support, and the reviewers for comments and advice on the manuscript.

References