Digital Stereo Image Analyzer for Generating Automated 3-D Measures of Optic Disc Deformation in Glaucoma

Enrique Corona, Sunanda Mitra*, *Senior Member, IEEE*, Mark Wilson, Thomas Krile, *Senior Member, IEEE*, Young H. Kwon, and Peter Soliz

Abstract—The major limitations of precise evaluation of retinal structures in present clinical situations are the lack of standardization, the inherent subjectivity involved in the interpretation of retinal images, and intra- as well as interobserver variability. While evaluating optic disc deformation in glaucoma, these limitations could be overcome by using advanced digital image analysis techniques to generate precise metrics from stereo optic disc image pairs. A digital stereovision system for visualizing the topography of the optic nerve head from stereo optic disc images is presented. We have developed an algorithm, combining power cepstrum and zero-mean-normalized cross correlation techniques, which extracts depth information using coarse-to-fine disparity between corresponding windows in a stereo pair. The gray level encoded sparse disparity matrix is subjected to a cubic B-spline operation to generate smooth representations of the optic cup/disc surfaces and new three-dimensional (3-D) metrics from isodisparity contours. Despite the challenges involved in 3-D surface recovery, the robustness of our algorithm in finding disparities within the constraints used has been validated using stereo pairs with known disparities. In a preliminary longitudinal study of glaucoma patients, a strong correlation is found between the computer-generated quantitative cup/disc volume metrics and manual metrics commonly used in a clinic. The computer generated new metrics, however, eliminate the subjective variability and greatly reduce the time and cost involved in manual metric generation in follow-up studies of glaucoma.

Index Terms—Glaucoma, isodisparity contour, new three-dimensional (3-D) metrics of optic nerve head, stereo optic disc photography.

I. INTRODUCTION

DEVELOPMENT of a medical diagnostic system usually requires long term evaluation and willingness from the physicians to accept a new technology. Retinal diseases such as diabetic retinopathy, age-related macular degeneration (ARMD) and glaucoma are common causes of early visual loss and blindness. Early detection of diabetic retinopathy and glaucoma is particularly significant since it allows timely treatment to prevent major visual field loss and prolongs the effective years of usable vision. Therefore, development of precise as well as automated methods of clinical measures for evaluating changes in retinal features, such as the optic nerve head (ONH) topography in glaucoma, is essential. Because the ophthalmoscope is still the most widely used instrument employed by clinicians to diagnose retinal diseases, the diagnostic information from its use and the interpretation of retinal images from funduscopes are dependent on the expertise of the clinician. The major limitation to precise evaluation of early retinopathy in the present clinical situations still remains the inability of the human vision system to detect subtle changes and make precise estimates of size, shape, and color of pathological features. This limitation leads to intra- as well as interobserver variability in sensitivity and specificity. The subjective evaluation and interpretation of the optic nerve head have been reported with documentation of inter- and intraobserver variations [1], [2]. Zangwill *et al.* [3] developed a quantitative grading procedure for measuring cupping, defined by cup contour, for use in population based studies. To overcome the human vision limitations, a number of advanced retinal imaging systems have been developed over the years for early detection of glaucoma. Ophthalmic systems using laser scanning and optical interferometric techniques have been designed to detect nerve fiber layer (NFL) loss in glaucoma [4]–[6]. However, the clinical utility of the aforementioned quantitative approaches have not been shown to be superior to expert interpretation of stereo disc photography, and the latter is still the technique commonly employed in clinical settings for documenting optic disc changes [22]–[26]. Digital stereo image interpretation in clinics by physicians, to derive two-dimensional (2-D) measures such as the optic cup/disc ratios, is not yet fully automated. The need for human intervention continues to be a principal source of variability [7], [8], [21]–[25], [27]–[30].

The requirement for a low cost, easy-to-use, and widely accepted diagnostic measure for glaucoma is met in most clinics principally by standard photography, either with a stereo camera that collects the image pair simultaneously or through the use of standard retinal imaging cameras that collect the image pair sequentially. Accurate measures of the optic nerve head are cru-

E. Corona and T. Krile are with the Computer Vision and Image Analysis Laboratory, Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX 79409-3102 USA.

*S. Mitra is with the Computer Vision and Image Analysis Laboratory, Department of Electrical and Computer Engineering, M.S. 3102, Texas Tech University, Lubbock, TX 79409-3102 USA (Sunanda.Mitra@coe.ttu.edu).

M. Wilson and P. Soliz are with the Kestrel Corporation, Albuquerque, NM 87109 USA.

Y. H. Kwon is with the Department of Ophthalmology and Visual Sciences, University of Iowa Hospitals and Clinics, Iowa City, IA 52242-1091 USA.

cial in the early diagnosis and follow-up management of chronic glaucoma. Glaucoma can cause damage in the optic nerve and ultimately loss of vision. Currently, the condition of the ONH is qualitatively evaluated by observation of a pair of stereoscopic fundus photographs by an ophthalmologist in addition to measurements of cup to disc ratios from the same photographs. Early diagnosis of glaucoma is based on observations of experienced clinical observers and manual drawing of the contours of the ONH and, thus, the procedure required is tedious and time-consuming and, most of all, prone to variations in their interpretation that may mask subtle changes due to the disease progression. We have previously addressed fundus image enhancement and stereo image analysis based on advanced signal/image processing methodologies such as cepstrum analysis for registration of stereo images, and disparity to depth mappings and three-dimensional (3-D) surface recovery of the optic disc using cubic B-spline interpolation [9]–[16]. The present work focuses on developing an automated computerized technique for precise 3-D measures of the cup/disc ratios of the ONH from stereoscopic pairs of fundus images based on advanced image analysis techniques involving 3-D surface recovery from stereo disparity and registration using Fourier methods [17], [18]. The computer generated measures are compared with manual measures used by the ophthalmologists for a test data set of longitudinal stereo fundus images of glaucoma patients spanning over 20 years to determine the validity of computer generated 2-D and 3-D cup/disc measures in monitoring progression of glaucoma. The methodology is described in Section II and Section III presents the new 3-D measures of the ONH and the segmented 2-D measures obtained from the 3-D surface for comparison with the existing clinical measures. The preliminary results show the potential of such automated 3-D ONH measures in consistent and improved evaluation and follow-up of glaucomatous ONH's in clinical trials for drug therapy. Section IV discusses future developments and conclusions. Our methodology yields quantitative evaluation of deformation in the ONH in terms of additional measures such as the change in the volume of the cup/disc in a longitudinal follow-up study of a patient.

II. METHODOLOGY

A. Overall System Design Background

Early detection of glaucoma is particularly significant since it allows timely treatment to prevent major visual field loss and prolongs the effective years of usable vision. Although traditionally an elevated intraocular pressure is often associated with glaucoma, current literature [19] indicates that the measurement of intraocular pressure cannot be a reliable predictor of visual function loss from glaucoma. The major limitation of precise evaluation of early glaucomatous changes in present clinical situations still remains the inability of the human vision system to detect subtle changes and make precise estimates of size, shape, and color of pathological features. Change in cupping of the optic disc represents a valuable indicator for the ophthalmologists to diagnose and monitor the disease. This work is devoted to deriving a 3-D surface from 2-D stereo data. Several stages are involved in this, from preprocessing and initial registration of the stereo pair, to disparity finding and interpolation
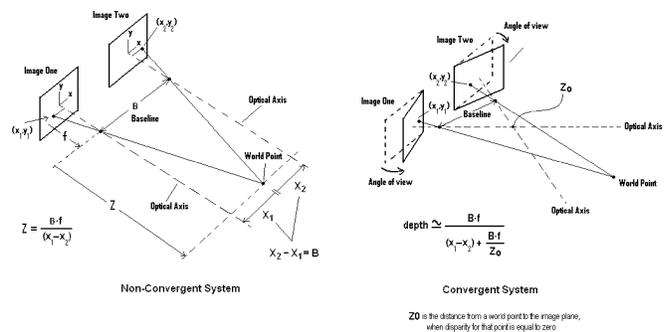

Fig. 1. Convergent and nonconvergent visual systems.

of the sparse disparity maps. Two cameras capture the same 3-D real world image from different perspectives, providing a pair of stereo images. The coordinate associated with depth of this scene can be extracted by triangulation of corresponding points in the stereoscopic images. In the process of finding disparities between conjugate pairs of points, image-matching strategies are used. According to the matching strategy used (to find these disparities) the processes of searching can be either area based or feature based. Area based strategies intend to match image areas, while feature based processes try to match whatever feature seems to be in the stereo pair. Sun in [20] suggests an area based matching technique using zero mean normalized cross correlation (ZNCC) as the disparity measure. The disparity measure used in this paper involves ZNCC that is expressed as follows:

$$\text{ZNCC}(i,j) = \frac{\text{cov}_{i,j}(f,g)}{\sigma_{i,j}(f) \times \sigma_{i,j}(g)} \quad (1)$$

$$\text{cov}_{i,j}(f,g) = \frac{1}{((2K+1)(2L+1)-1)}$$
$$\times \sum_{m=i-K}^{i+K} \sum_{n=j-L}^{j+L} \left(f_{m,n} - \overline{f_{i,j}}\right)(g_{m,n} - \overline{g_{i,j}}). \quad (2)$$

Here, f and g are the arrays of pixels to be tested. K and L define the size of these arrays (windows), and the indexes for the pixels within the windows are i and j. Note that the array is rectangular-shaped. $\sigma(f)$ and $\sigma(g)$ equal the square roots of $\text{cov}(f,f)$ and $\text{cov}(g,g)$, respectively. There are primarily two vision systems for implementing stereo vision, referred to as convergent and nonconvergent systems. In the nonconvergent system, which approximates the clinical imaging situation better, the disparity is inversely proportional to depth. In the convergent system the disparity between corresponding points can also be shown to be inversely proportional to the depth. However, some other parameters are involved as well. Fig. 1 shows the convergent and nonconvergent systems and their relationships with the calculated disparities in the stereo pair.

This paper assumes a nonconvergent system as a good approximation to the clinical imaging situation and uses both feature and area based matching techniques to compute disparity between conjugate pairs of the same optic disc. The stereo disc photographs used in this paper were taken with a Zeiss fundus

camera (C. Zeiss, Thornwood, NY) on Kodachrome 25 or Ektrachrome 100 film. Stereopsis was achieved by decentration of the camera angle. The color photographs were digitized using a Nikon LS-2000 slide scanner. The original slides were cropped to 15-degree fields of view during the scanning process to produce 512×512 pixel images saved as 8-bit TIFF files. The mapping from 12-bit to 8-bit was done by the following function $\text{Remap} = \log(\varepsilon + \alpha \cdot \text{Input})$ where epsilon is a small tolerance since input values include zero and alpha takes values of 0.1, 0.01, 0.001, and 0.0001. The Remap is then stretched to values between 0 and 255 with a linear transform. The conversion modes included a linear stretch and four others stretched with separate quasigamma values [30].

Although the initial inputs to the digital stereo image analyzer are color images, extraction of binary features (blood vessels) is necessary for better cepstral matching. The power cepstrum works best as a matching technique if the frequency of the signal is greater than or equal to the frequency of the noise present [31]. Thus, the signal requires high-frequency edge features (such as binary features) for satisfactory matching. However, when pyramidal-structured correlation coefficients between the stereo pair pixels are computed, windows without large featureless regions are used for finer disparity search. A very important constraint applied by the vision model used is that no disparity is expected on the vertical axis (only horizontal shifts are allowed). To assure that this statement is true, the pair of images for which disparity is to be calculated is, first, vertically registered. This registration combines power cepstrum and frequency spectrum analysis to compensate for unwanted rotations and shifts that exist within the stereo pair.

At a given level of coarseness, the pair of images is divided into square windows of a given size according to the current level of coarseness. For corresponding windows in the stereo pair, the power cepstrum of the sum of both windows is calculated, as well as the cross correlation (using ZNCC) along a range of pixels varying from minus one half of the window size to plus one half of the window size. From the set of possible horizontally shifted pixel positions obtained by cepstral analysis, the one with the highest correlation coefficient is considered to be the disparity associated with each element in the whole window. After every disparity has been calculated for each pixel in the pair, a new stereo pair is generated consisting of the same size windows as the old stereo pair but shifted by the number of pixels in the previous window. Then, the disparity map is stored, the search window size is halved, and the search is performed again on the new stereo pair. When the window size is reduced to the size of 8×8, the algorithm applies a low-pass filter to each disparity map stored for each window size, and adds them up to get a resulting sparse disparity matrix. Finally, these data are fed into a cubic B-spline interpolation routine that smoothes the computed disparity matrix. Then features such as blood vessels are superimposed to help as landmarks in the final 3-D representation of the ONH. Computer generated measures such as cup to disc ratios in volume and area can be calculated from these data by segmenting the cup and disc contours from iso-disparity contours generated at each depth. Fig. 2 shows the overall process for the 3-D visualization of the stereo image pair. Although the registration and the disparity search processes are fully auto-
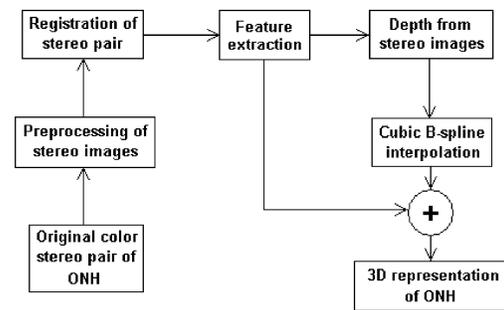

Fig. 2. A schematic block diagram of the analysis and modeling involved for 3-D visualization of a stereo image pair.
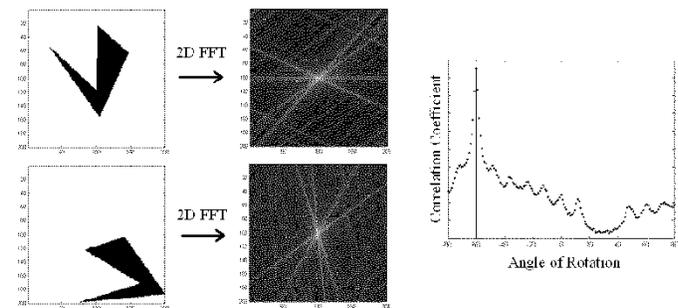

Fig. 3. Plot of ZNCC coefficients versus angle of rotation. Highest coefficient shows the amount of degrees that must be used for rotational compensation.

mated, the cup and disc contours can be interactively adjusted from the iso-disparity contours for better matching with the contours manually generated by the ophthalmologists.

B. Preprocessing and Stereo Pair Registration

Preprocessing steps are explained in more detail here. Three channel (RGB) decomposition is performed on the original color pair. Only the green channel is processed since it is the one that carries the most information. Red and blue channels have low entropy in relation to the green channel and therefore are not taken into account. The registration process removes all vertical displacements leaving only the horizontal shifts arising from the different positions of the camera while taking the stereo fundus images. A good registration is crucial to obtain accurate disparity maps. We employ a power cepstrum based registration that uses Fourier spectrum properties to correct rotational errors that may be present in the stereo pair. This process begins by extracting the most relevant features such as the blood vessels in both images. These features are extracted by subtracting a low-pass filtered version of the original stereo pair from the original (unsharp masking). After binarizing this new stereo pair, multiple passes of a median filter are used to eliminate some of the resulting noise in the images. Compensation for rotational differences is also performed here via ZNCC correlation of the Fourier spectrum of the images. According to the inherent Fourier spectrum properties, a rotation in the spatial image results in the same amount of rotation of its spectrum. Thus, it is possible to find the angle of rotation of one image in the stereo pair with respect to the other by performing step-by-step rotations and cross correlating their Fourier transforms. The actual angle of rotation will be
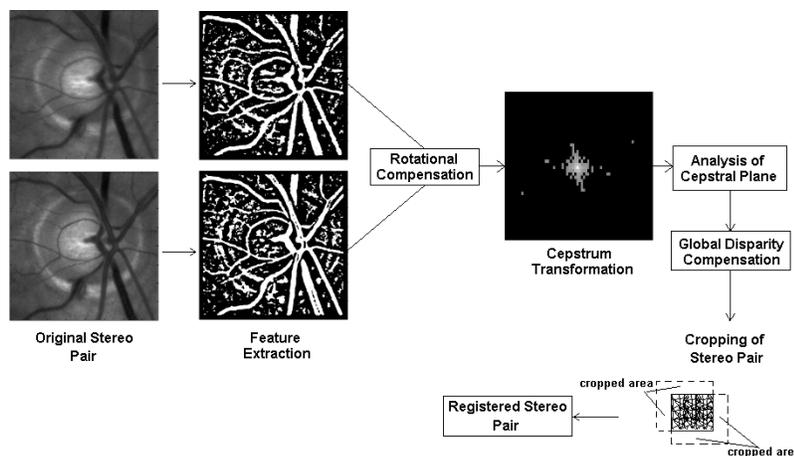
Fig. 4. Compensation for translation and rotation within the stereo pair of images.

the one with the highest cross correlation obtained. Rotational compensation is applied once the angle of rotation has been found. Fig. 3 shows an example plot of ZNCC coefficients at each angle step for which the search was performed. In the cases chosen, the search is performed from a range of $-2°$ to $+2°$ in $0.1°$ steps.

After the rotational correction, a cepstrum transformation is applied to the sum of the binary-featured stereo pair images. The power cepstrum P is defined as in [14]

$$P\left[i(x,y)\right] = \left| \mathbf{F}\left(\ln\left\{\left|\mathbf{F}\left[i(x,y)\right]\right|^2\right\}\right)\right|^2 \tag{3}$$

where $\mathbf{F}$ represents the Fourier transform operation. Let $w(x,y)$ be the reference image, $w(x+x_0, y+y_0)$ be the shifted image, and $i(x,y) = w(x,y) + w(x+x_0, y+y_0)$. Then, the power cepstrum of the sum of both images is given as

$$P\left[i(x,y)\right] = P\left[w(x,y)\right] + A\delta(x,y) + B\delta(x \pm x_0, y \pm y_0) \\ + C\delta(x \pm 2x_0, y \pm 2y_0) + \cdots \tag{4}$$

where $\delta(x,y)$ is the Kronecker delta and A, B, and C are the first three coefficients for this power cepstrum expansion series [16]. Equation (4) shows that the displacement between images results in the sum of the power cepstrum of the original image $w(x,y)$ plus a multitude of delta functions. Each delta is separated from the others by an integer multiple of the actual displacement we are looking for. As (4) shows, the cepstrum of the reference must be subtracted from the cepstrum of the stereo pair in order to leave only the deltas related to the displacement. With this in mind, a fixed number of deltas are chosen from the resulting cepstrum. Each delta represents a translational shift, or an integer multiple of the shift, of a pixel in the shifted image from the corresponding pixel in the reference image. All points are tested by cross correlating the reference image with the other image shifted by the number of pixels (in the vertical and horizontal directions) indicated by the current point being tested. The highest correlation will correspond to the most probable relative translation between both images. Once the translation of one image with respect to the other is known, compensation is performed. Some cropping is necessary after compensation
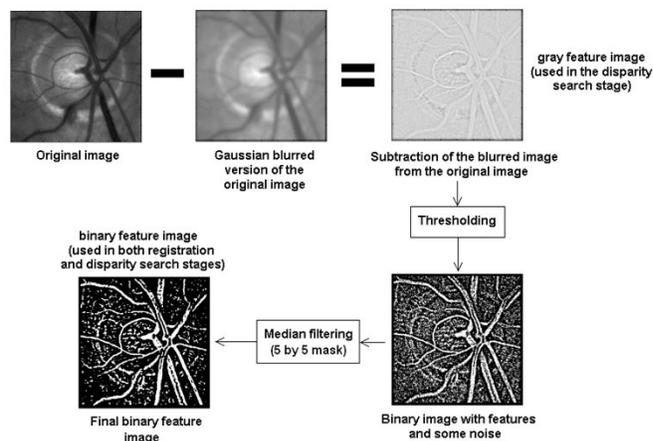

Fig. 5. Feature extraction process.

to get rid of regions with no common information. After compensation of these displacements and cropping, the stereo pair is ready for further processing. The registration process is shown in Fig. 4.

C. Feature Extraction

Preprocessing of the green channel of the multispectral stereo fundus images to extract salient features in the images is performed next. The same unsharp masking procedure as in the previous registration stage is used again. This enhances the salient features in the images (blood vessel edges) that can be segmented later by thresholding. Although the green channel is chosen here (because of better contrast in this channel rather than in the blue or red), a gray-scale conversion taken directly from the color stereo pair can also be used without noticeable loss of detail in the binary (feature) images. After thresholding the stereo pair, blood vessels are segmented along with some noise intrinsic to the unsharp masking method used. To filter out some of the remaining noise, a median filter is applied to get a clearer binary representation of the features. Fig. 5 shows the steps followed by the binary feature extraction technique. It is important to mention that a combination of binary and gray=scale features is used in the disparity search stage. The gray=scale features are obtained using the unsharp masking as

done with the binary features, but no thresholding is performed. Since both the power cepstrum and ZNCC are used to find disparities, the feature stereo pair must have characteristics that favor both methods. Cepstrum performs well as long as the bandwidth of the noise remains below the bandwidth of the signal, showing the necessity for sharp edges that can be provided by a binary image. The correlation coefficient method works well as long as there are no large flat regions. This is the reason for including the unsharp masked gray image also. The superposition of the binary image over the gray=scale image will be the final feature image from which disparities are calculated.

D. Depth From Stereo and Cubic B-Spline Interpolation

This step receives the feature stereo pair from the aforementioned stage and outputs a disparity map showing right and left displacements between corresponding points. In this step, only horizontal shifts are expected between the two images in the stereo pair. The algorithm developed for the search of disparities first divides both images into square windows of a given size (multiple of two), say $N \times N$. Cross correlation (ZNCC) will be performed between the windows in one image with the windows in the other image. If cross correlation is larger than a certain threshold, it is assumed that the windows at that position in the image are similar, so the cepstrum is applied to those windows to check for possible shifts. Otherwise, if the cross correlation is smaller than or equal to the threshold mentioned, zero disparity is assigned to every pixel in the window. Only a specified number of horizontal points shown in the cepstrum are taken into account for analysis. Let us say that, for an N by N window, only $N/4$ horizontal points are chosen for analysis in the cepstral plane. This is because for an $N \times N$ window the maximum horizontal displacement that can be detected is $N/2$ (either to the right or to the left, making a full range of N pixels), so checking all $N/2$ points for right and left shifts will be very time consuming. Instead, only the most probable $N/4$ horizontal shifts found by the cepstrum will be tested using the cross correlation technique. One of the images of the stereo pair is considered as the reference image and the other is the test image. Then, for every point chosen (from the cepstrum), cross correlation is applied between the reference window (in the reference image) and the other window (in the test image) shifted by the number of pixels determined by the cepstral shift. Since the cepstrum can only detect the amount of the shifts but not their direction (the cepstrum is symmetric about the origin), each point should be tested for left and right shifts. So, when checking $N/4$ cepstral points, actually $N/2$ positions are analyzed. The highest value in the cross correlation will be the most probable shift that will be assigned to all elements in the window currently being tested for disparity. The number of cepstral points is not a fixed parameter and can be modified. This modification will affect the processing time and the accuracy of the disparity map. Once all disparities have been calculated with a window size of $N \times N$ pixels, the size of the window is reduced by a factor of two and the whole process is repeated until the windows reach a predetermined size. Each disparity map (calculated at a given resolution) is accumulated by adding

it to the previous disparity map. At the end of the process, the final disparity map is the total accumulated disparity map. Usually the starting window size is 64 by 64 and the stopping size is 8 by 8. Smaller sizes of a window may not be worth computing because of the much longer time required and the small impact of it on the final disparity map. Also, since the window is so small, chances are that noise becomes a serious issue. The cepstrum is, in fact, a very noise tolerant technique that is suitable for finding disparities in chosen regions [31] while cross correlation is noise sensitive and finds disparities using a procedure in a pixel-by-pixel fashion. A combination of both techniques results in an accurate and noise tolerant algorithm. It is quite important to mention that in the very first coarseness level, disparities found to be equal to one half of the size of the window (either positive or negative) are checked in an alternative way. Such a procedure is followed since the disparity reported on the boundary of the search may be only an indication that the window is too small to catch a larger disparity. In this specific case, the disparity (for that specific window) will be calculated as an average of the neighboring windows. In order to get an accurate 3-D representation from a stereo pair of images, disparities must be known for each point (pixel) of one image with respect to the other. Since the disparity search algorithm only finds disparities for the features or regions, disparities of all individual pixels are not known. The interpolation used here gives an estimate of the other missing disparities. Cubic B-spline is the interpolation technique applied to the sparse matrix resulting from the disparity search. It can be shown that the cubic B-spline can be modeled by three successive convolutions with a constant mask [16], [32]. In this case, a mask consisting of all ones of size 32×32 or 64is used. After filtering the original sparse disparity matrix three times with the mask described above, a smooth representation results. This is the final 3-D surface of the ONH. With this surface, measures such as the disc and cup volume can be made. The disparity search process and the interpolation are shown in Fig. 6. In this figure, only for illustrative purposes, the three convolutions are shown. The actual implementation is performed in the frequency domain by multiplying the Fourier transform of the flat kernel mask by itself three times. This kernel mask is then padded with zeros to fit the size of the disparity map. The Fourier transform of the disparity map is taken and multiplied by the padded kernel. Finally, an inverse Fourier transform is taken for mapping the data back to the spatial domain.

E. Superposition of Features

A contrast reversed, linearly stretched feature image is superimposed on the depth image in order to visualize the blood vessels as landmarks on the 3-D representation of the ONH. This step is very important for ophthalmologists because longitudinal data can be more easily analyzed when reference points can be distinguished on ONH 3-D representations. A mask of the same size as the image is constructed from one of the images in the binary stereo feature pair. This mask is multiplied by the original gray=scale image, leaving only blood vessels along with some noise that can be filtered out by thresholding. Once this is done, the resulting picture is added to the final 3-D surface
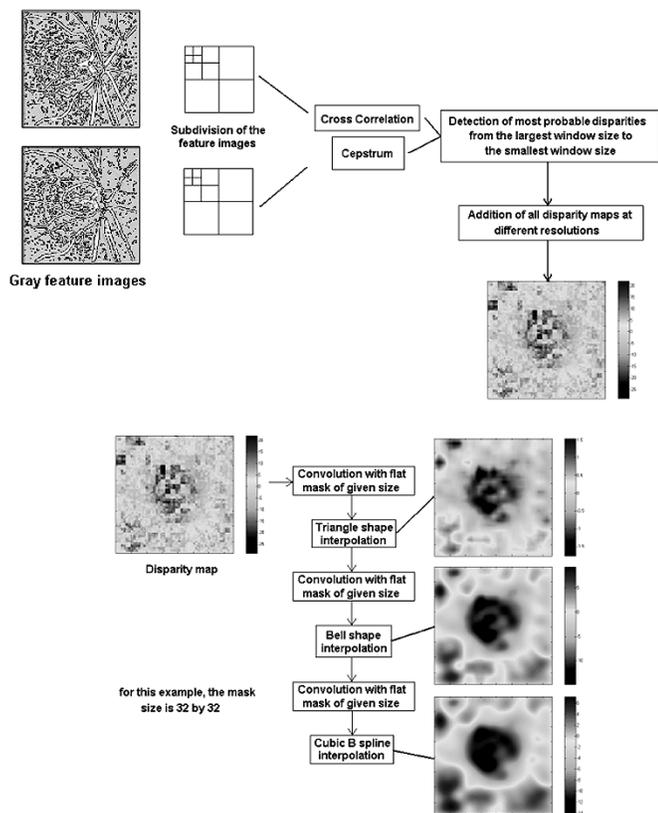
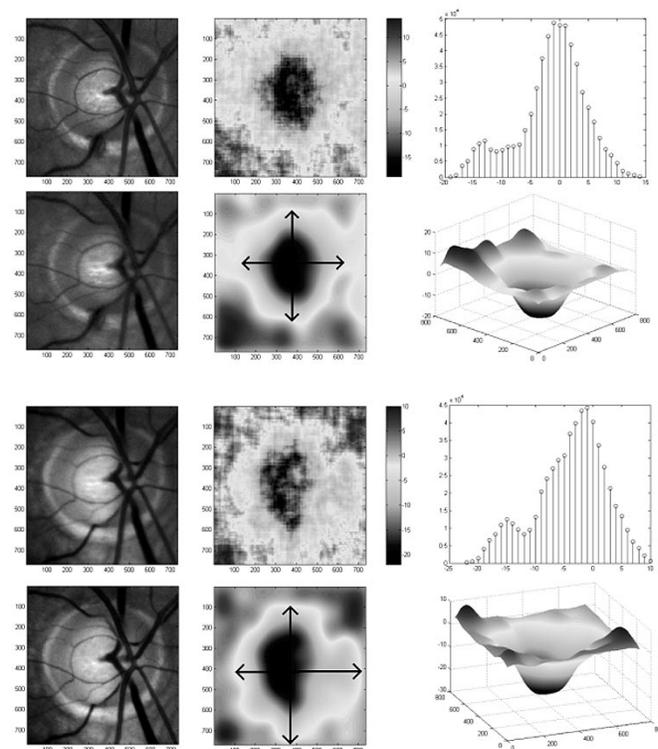Fig. 6.   Disparity search process and interpolation.

Fig. 8.   The first column shows a stereo pair of fundus images of a glaucoma patient taken in 1994 on top, and the stereo pair of the same eye taken in 1999 on the bottom. The second column shows the disparity maps and the corresponding isodisparity contours from the two stereo pairs. The third column shows the disparity distributions and depth representations of the ONH.
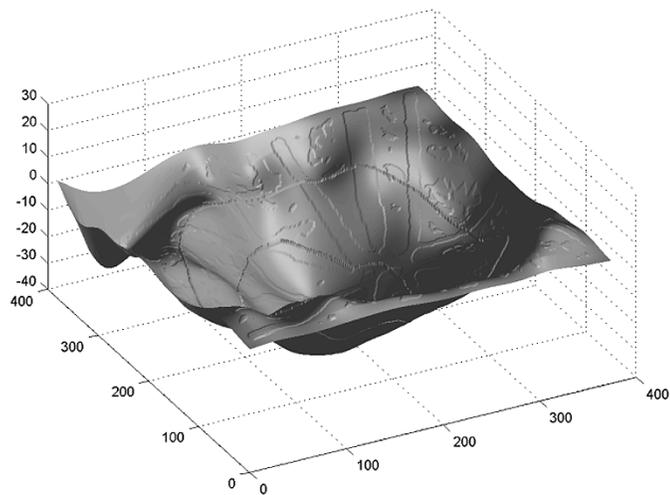

Fig. 7.   Final 3-D representation of the ONH.

of the ONH. This is the final 3-D representation of the ONH. Fig. 7 shows the final representation obtained with the methodology previously described.

III. RESULTS

Fig. 8 shows the disparity distributions and the corresponding 2-D disparity maps for the same eye of a glaucoma patient from stereo disc photographs taken in 1994 on the top and the same taken in 1999 on the bottom. Further spreading of vertical and
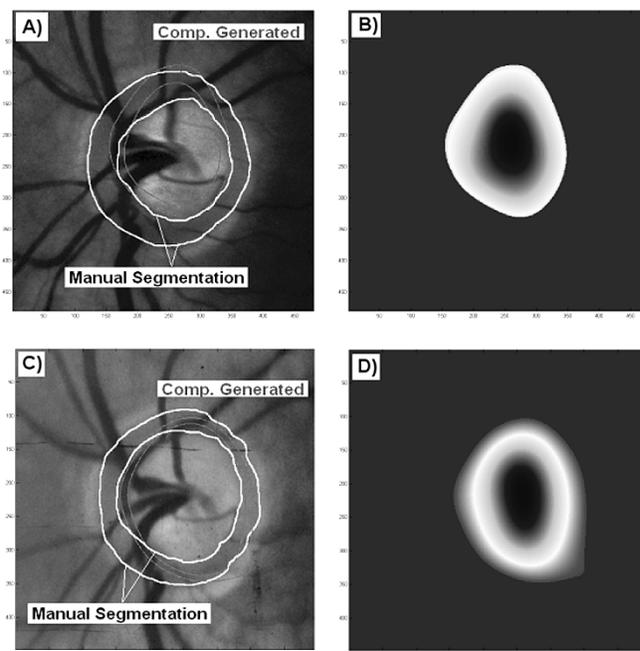

Fig. 9.   The left column shows the computer generated and manual disc and cup contours marked by an ophthalmologist superimposed on the fundus images taken in 1989 and 1999 of patient number 1 (Tables I and II). The right column shows the corresponding isodisparity contours.

horizontal elongations of the disparity maps through the years suggests the progression of the disease.

TABLE I

MEASURES IN PIXELS BASED ON OPHTHALMOLOGIST'S MANUAL SEGMENTATION (MO)

Data File (patient number-year of study)	Vertical cup length	Vertical disc length	Horiz. cup length	Horiz. disc length	Cup area	Disc area	Cup volume	Disc volume
Patient 1 1989	196	280	185	261	27009	56250	234305	312619
Patient 1 1993	198	270	176	248	25959	52981	263573	351062
Patient 1 1996	211	281	197	255	32866	55911	291897	350259
Patient 1 1999	198	264	188	240	29002	49757	192314	206171

TABLE II

MEASURES IN PIXELS BASED ON OUR COMPUTER GENERATED SEGMENTATION (CG)

Data File (patient number-year of study)	Vertical cup length	Vertical disc length	Horiz. cup length	Horiz. disc length	Cup area	Disc area	Cup volume	Disc volume
Patient 1 1989	198	243	153	194	22792	34779	218818	283568
Patient 1 1993	209	251	185	247	30032	46515	2904001	345188
Patient 1 1996	223	242	180	200	32110	39087	298944	331957
Patient 1 1999	226	244	176	197	31001	38174	203727	215083

Our semiautomatic segmentation method for the cup and disc consists of manually looking for the iso-disparity contours from the automatically generated smoothed disparity matrices that better enclose the cup and disc. In this context, the best contour is the one that has more points in common with the manual segmentation from the ophthalmologist. Prior to the automation of our algorithm, initial results were evaluated by our semiautomatic segmentation of the cup and disc using the interpolated version of the disparity maps found in the previous stages. There is also a manual segmentation performed by a group of trained ophthalmologists. In Fig. 9, both segmentations (clinical and semiautomatic) are shown for the same patient with glaucoma. The stereo pairs were taken in 1989 and 1999. Notice that the contours are not exactly the same but they have similar shapes. Prior to developing a fully automated algorithm for cup/disc contours, statistical validity tests such as correlation coefficients between clinical and automated measures generated from our algorithm with a large sample of stereo pairs evaluated by several ophthalmologists need to be performed.

Segmented cup and disc (both computer generated and ophthalmologist segmentation) are used to obtain volume measures from the 3-D ONH final representation. The cup volume is that contained in the portion of the cup within the depth representation. The disc volume is obtained in the same manner. The cup (or disc) volume is found simply by accumulating the intensity values on the depth map enclosed in the segmented cup (or disc). If more precision is needed, subpixel interpolation can be performed, thus increasing the number of intensity values to be accumulated within the disc or cup segmentation. The cup (or disc) area is calculated as the number of pixels contained within the cup (or disc) segmentation. Maximum length measures are calculated as the maximum length of the row or column con-

tained in the cup or disc. The length and area ratio measures of the optic disc, based on manually drawn contours of the cup and disc (by the ophthalmologist) show an increment in deformation of the ONH in the follow-up measures as listed in Tables I and II. The new computer generated volume measures (using our semiautomatic segmentation) also follow the same trend, namely an increase in the deformation of the optic disc, in the same eye of patient one during ten years of evaluation between 1989–1999.

We have extended this study initially to six additional patients with data available over a period spanning 18 years as test data sets to verify this trend and determine the correlation as well as statistical significance levels of computer generated measures with manual clinical measures. Other measures from longitudinal data sets from a patient (patient one) are shown in Table III. Both manual (clinical) and computer generated 2-D and 3-D measures show a consistent increase in deformation of the ONH in the patient. However, the actual values of the measures from ophthalmologists and computer generated measures of the cup and disc differ somewhat. Investigation of the modeling involved in the automated segmentation process needs to be carried out to interpret these differences. Results from clinical and computer generated measures for one patient over ten years are shown in plots of Fig. 10, suggesting good correlation for our new volume measures from clinical and computer generated cup/disc contours. Table IV demonstrates examples of such good correlations for two additional patients over a period of 20 years for each patient. Recently, we have also completed analysis of two additional sets of stereo fundus photographs with a total of 159 stereo pairs. For the first data set consisting of 86 stereo image pairs, the average correlation coefficient between the ophthalmologist and the computer generated cup to disc ratios was found to be 91.4%. For the second data set consisting of 73 stereo image pairs, the
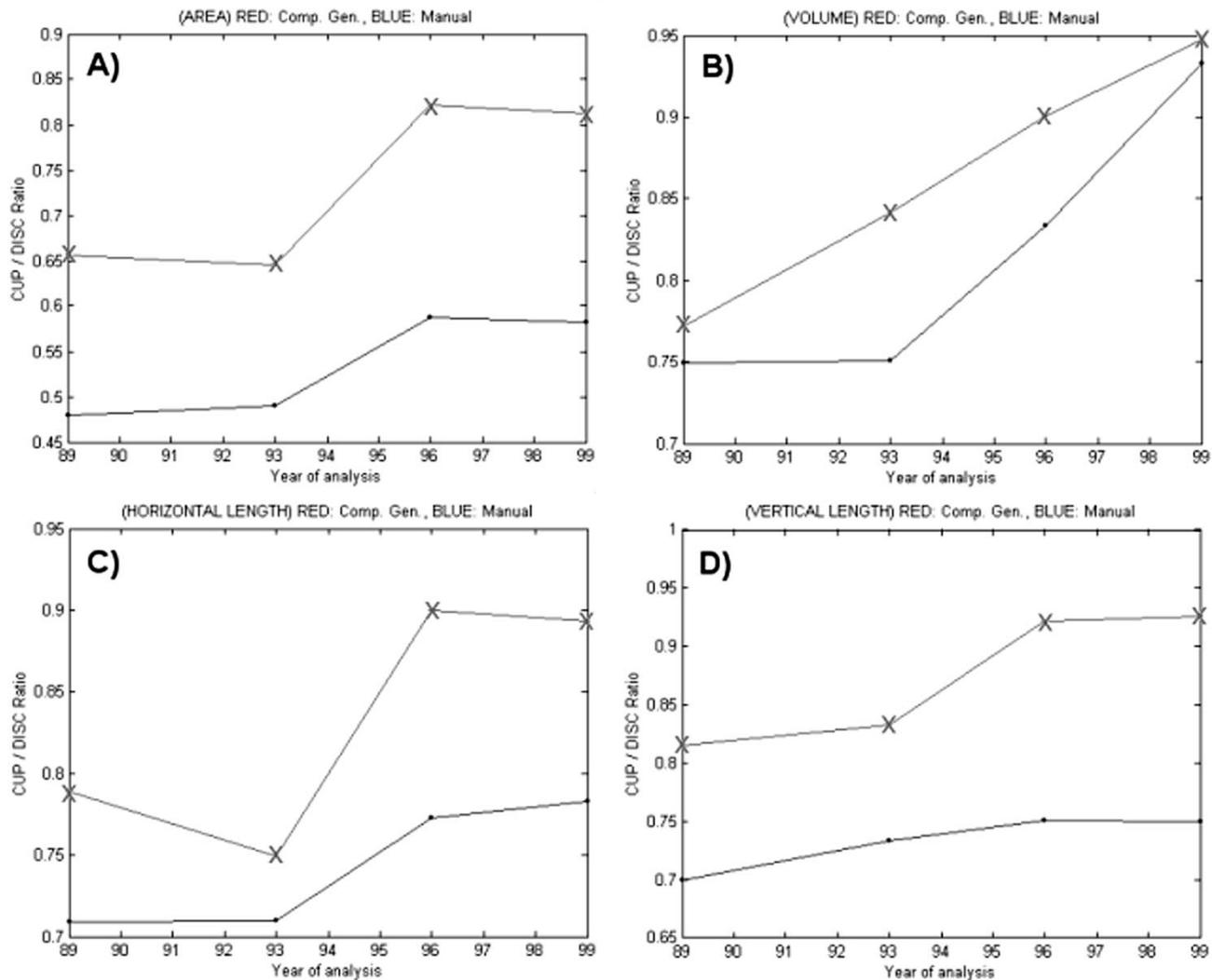
Fig. 10. Cup to disc ratios versus years of study for patient number 1. A) Area cup/disc ratio. B) Volume cup/disc ratio. C) Horizontal length cup/disc ratio. D) Vertical cup/disc ratio. xxx: computer generated, ...: manual.

TABLE III
2-D AND 3-D CUP/DISC RATIOS BY COMPUTER GENERATED AND MANUAL SEGMENTATIONS

Data File (patient number-year of study)	Cup to disc vertical length ratio **(CG)**	Cup to disc vertical length ratio **(MO)**	Cup to disc horizontal length ratio **(CG)**	Cup to disc horizontal length ratio **(MO)**	Cup to disc area ratio **(CG)**	Cup to disc area ratio **(MO)**	Cup to disc volume ratio **(CG)**	Cup to disc volume ratio **(MO)**
Patient 1 1989	0.81	0.7	0.79	0.71	0.66	0.48	0.77	0.75
Patient 1 1993	0.83	0.73	0.75	0.71	0.65	0.49	0.84	0.75
Patient 1 1996	0.92	0.75	0.90	0.77	0.82	0.59	0.9	0.83
Patient 1 1999	0.93	0.75	0.89	0.78	0.81	0.58	0.95	0.93
Correlation coefficient between **MO** and **CG**	0.91		0.96		0.99		0.91	
(**MO**)=Manually Segmented by the Ophthalmologist				(**CG**)=Computer Generated				

average correlation coefficient for the same measures was 95.7%. However, more studies need to be performed for comparing the variability of the computer generated automated measures with the manually generated measures by the ophthalmologists.

TABLE IV

COMPARISON OF CORRELATION COEFFICIENTS BETWEEN COMPUTER-GENERATED VERSUS MANUAL SEGMENTATION OF VARIOUS MEASURES AMONG EACH PATIENT FROM LONGITUDINAL STUDIES USING FOUR STEREO PAIRS SPANNING 10 TO 20 YEARS

Longitudinal study of patient #	Spanning years	Cup to disc **vertical length** correlation	Cup to disc **horizontal length** correlation	Cup to disc **area** ratio correlation	Cup to disc **volume** ratio correlation
Patient 1	89-93-96-99	0.91	0.96	0.99	0.91
Patient 2	78-87-94-98	0.99	0.99	0.98	0.96
Patient 3	74-79-84-94	0.96	0.93	0.99	0.86

IV. CONCLUSION

The 2-D measures of cup-to-disc ratios from stereo fundus images currently may be obtained in clinical practice by visual estimates and time consuming manual drawing of the cup/disc contours by skilled ophthalmologists based on their subjective perception of the 3-D shape of the ONH. Such measures are prone to inter- and intrasubject variability and are quite tedious to acquire when large data sets are involved for monitoring the patients over a long period of time, as in evaluating the effect of drug therapy. We have developed a semi-automated method of obtaining such 2-D measures, as well as additional new 3-D measures from a number of robust signal/image analysis approaches for feature extraction, registration and 3-D visualization of the ONH from digital stereo fundus images. Fully automatic segmentation of the cup and disc is an extremely challenging task and currently in progress. Our preliminary results show good correlation between clinical measures and computer generated 2-D measures, even though the method still needs validation and testing over large samples. Table IV shows quite good correlation for three glaucoma patients. Although 3-D surface recovery from stereo images is an ill-posed problem, the use of proper constraints and optimization results in repeatable disparity contours and hence depth maps assuming a nonconvergent stereo vision system. This algorithm demonstrates the potential of such an approach in long-term follow-up of glaucoma patients and in evaluation of existing as well as new drug therapies. Although this methodology has been proposed previously [15], [16] the current approach has been further developed to yield more precise and repeatable measures to qualify for clinical trials. Digital image analysis has also gained more popularity and confidence from the ophthalmologists at the present time which allows us access to large medical data sets for carrying on long-term follow-up studies for more efficient evaluation of drug therapy in glaucoma. Simultaneous research is also being conducted in investigating several issues in optimization of the entire process.

ACKNOWLEDGMENT

REFERENCES

[1] J. Tielsch, J. Katz, and H. Quigley, "Intraobserver and interobserver agreement in measurement of optic disc characteristics," *Opthalmol.*, vol. 95, p. 350, 1988.

[2] P. Lichter, "Variability of expert observers in evaluating the optic disc," *Trans. Amer. Opthalmol. Soc.*, vol. 74, p. 532, 1976.

[3] L. Zangwill *et al.*, "Agreement between clinicians and a confocal scanning laser ophthalmoscope in estimating cup/disk ratios," *Amer. J. Opthalmol.*, vol. 199, pp. 415–421, 1995.

[4] F. Midelberg *et al.*, "Ability of the Heidelberg retina tomograph to detect early glaucomatous visual field loss," *J. Glaucoma*, vol. 4, pp. 242–247, 1995.

[5] D. Huang *et al.*, "Optical coherence tomography," *Science*, vol. 254, p. 1178, 1991.

[6] M. Hee *et al.*, "Optical coherence tomography of the human retina," *Arch. Ophthalmol.*, vol. 113, pp. 325–332, 1995.

[7] L. Dandona, H. Quigley, and H. Jampel, "Variablility of depth measurements of the optic nerve head and peripapillary retina with computerized image analysis," *Arch. Ophthalmol.*, vol. 107, pp. 1786–1792, 1989.

[8] R. Varma and G. Spaeth, "The PAR is 2000: A new system for retinal digital image analysis," *Opthalmic Surg.*, vol. 19, no. 3, pp. 183–192, 1988.

[9] S. Whiteside *et al.*, "Fundus image processing system for early detection of glaucoma," in *Proc. SPIE*, vol. 697, 1986, pp. 123–128.

[10] S. Whiteside, S. Mitra, and T. F. Krile, *Differential Retinal Structural Damage Exhibited by Image Enhancement of Fundus Photographs, Low Vision: Principles and Applications*, G. C. Woo, Ed. New York: Springer-Verlag, 1986, p. 125.

[11] D. J. Lee, S. Mitra, and T. F. Krile, "Digital registration techniques for sequential fundus images," in *Proc. SPIE*, vol. 829, 1987, pp. 293–300.

[12] V. Chandran, D. J. Lee, and S. Mitra, "Fundus image registration for a PC-based image processing system," in *Proc. Topical Meeting Non-Invasive Assessment Visual System*, Lake Tahoe, NV, Jan. 1987, pp. 26–28.

[13] S. Mitra, B. S. Nutter, and T. F. Krile, "Automated method for fundus image registration and analysis," *Appl. Optics*, vol. 27, pp. 1107–1112, 1988.

[14] D. J. Lee, T. F. Krile, and S. Mitra, "Power cepstrum and spectrum techniques applied to image registration," *Appl. Optics*, vol. 27, pp. 1099–1106, 1988.

[15] M. Ramirez *et al.*, "3-D digital surface recovery of the optic nerve head from stereo fundus images," in *Proc. 5th Ann. IEEE Symp. Computer-Based Medical Systems*, June 1992, pp. 284–291.

[16] J. Ramirez, S. Mitra, and J. Morales, "Visualization of the three dimensional topography of the optic nerve head through a passive stereo vision model," *J. Electron. Imag.*, vol. 8, no. 1, pp. 92–97, 1999.

[17] L. G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 4, pp. 325–376, Dec. 1992.

[18] E. Corona *et al.*, "Digital stereo optic disc image analyzer for monitoring progression of glaucoma," presented at the Int. Soc. Optical Engineering (SPIE), Feb. 23–25, 2002.

[19] J. Caprioli, "Recognizing structural damage to the optic nerve head and nerve fiber layer in glaucoma," *Amer. J. Ophthalmol.*, vol. 124, no. 4, pp. 516–520, Oct. 1997.

[20] C. Sun, "A fast stereo matching method," in *Digital Image Computing: Techniques and Applications*. Auckland, New Zealand: Massey Univ., 1997, pp. 95–100.

[21] W. Sponsel *et al.*, "A clinically practical reproducible and rapid means for obtaining and comparing glaucomatous stereo disc images," *Invest. Opthamol. Vis. Sci.*, vol. 39, p. 117, 1998.

[22] G. Shuttleworth, C. Khong, and J. Diamond, "A new digital optic disc stereo camera: Intra-observer and inter-observer repeatability of optic disc measurements," *Br. J. Ophthalmol.*, vol. 84, pp. 403–407, 2000.

[23] M. Greaney, D. Garway-Heath, and D. Hoffman, "Comparison of imaging methods to distinguish normal eyes from those with early glaucoma," presented at the Eur. Glaucoma Society Meeting Association for Research Vision Ophthalmology Ann. Meeting, Fort Lauderdale, FL., 2000.

[24] L. Pecora *et al.*, "Intra and inter-observer variability of optic nerve head by using the discam acquisition system," *Invest. Opthamol. Vis. Sci.*, vol. 41, p. 2985, 2001.

[25] M. Lim *et al.*, "Digital stereoscopic photography with Chevron drawings versus standard film photography: Estimate of cup to disc measurements," *Invest. Opthamol. Vis. Sci.*, vol. 41, p. 697, 2001.

[26] J. Nicholl *et al.*, "Comparison of Heidelbeng retina tomograph and Marcher Discam analysis results," *Invest. Opthamol. Vis. Sci.*, vol. 41, p. 101, 2001.

[27] T. Nguyen *et al.*, "Chevron ratio affords accurate measures of progressive optic disc cupping," *Invest. Opthamol. Vis. Sci.*, vol. 39, p. 3870, 1998.

[28] J. Mensah *et al.*, "The Chevron ratio: A new method of analyzing the optic disc from stereo color photographs using vascular landmarks," *Invest. Opthamol. Vis. Sci.*, vol. 38, p. 3870, 1997.

[29] P. Soliz *et al.*, *Considerations of Stereo Cues, Resolution and Contrast in a Computer-Based Grading System for Age-Related Macular Degeneration*. Ft. Lauderdale, FL: Association for Research in Vision and Ophthalmology (ARVO), 2000.

[30] M. Wilson *et al.*, "Optimizing retinal image digitization for improved digital processing and visualization," in *Proc. 14th IEEE Symp. Computer-Based Medical Systems (CBMS)*, Bethesda, MD, July 2001, pp. 26–27.

[31] P. W. Smith and N. Nandhakumar, "An improved power cepstrum based stereo correspondence method for textured scenes," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, pp. 338–348, Mar. 1996.

[32] W. K. Pratt, *Digital Image Processing*, 2nd ed. New York: Wiley, 1991, pp. 112–117.