Chapter 11
3D Reconstruction of the Retinal Arterial Tree Using Subject-Specific Fundus Images

D. Liu, N.B. Wood, X.Y. Xu, N. Witt, A.D. Hughes, and Thom SAMcG

Abstract Systemic diseases, such as hypertension and diabetes, are associated with changes in the retinal microvasculature. Although a number of studies have been performed on the quantitative assessment of the geometrical patterns of the retinal vasculature, previous work has been confined to 2 dimensional (2D) analyses. In this paper, we present an approach to obtain a 3D reconstruction of the retinal arteries from a pair of 2D retinal images acquired in vivo. A simple essential matrix based self-calibration approach was employed for the "fundus camera-eye" system. Vessel segmentation was performed using a semi-automatic approach and correspondence between points from different images was calculated. The results of 3D reconstruction show the centreline of retinal vessels and their 3D curvature clearly. Three-dimensional reconstruction of the retinal vessels is feasible and may be useful in future studies of the retinal vasculature in disease.

11.1 Introduction

The retina, which lies at the posterior fundal surface of the eye, has the highest oxygen requirement per unit weight of any tissue in the body [10] and this makes it particularly vulnerable to vascular insults impairing oxygen and nutrient supply. Retinal vascular anatomy and net-work structure are adversely affected by high blood pressure, diabetes mellitus, ageing and atherosclerosis. Diabetic eye disease is one of the commonest causes of blindness in UK. A number of studies have shown that generalized arteriolar narrowing and retinopathy are associated with in-creased risk of stroke, ischaemic heart disease, heart failure, renal dysfunction and cardio-vascular mortality [13]. Therefore, quantitative assessment of the retinal vascular network is very important.

D. Liu, N.B. Wood, and X.Y. Xu
Department of Chemical Engineering, Imperial College London, UK

N. Witt, A.D. Hughes, and Thom SAMcG
International Centre for Circulatory Health, Imperial College London, UK

The geometric patterns of the retinal microvascular network are readily observed in vivo using fundal photography [12]. However, quantitative analysis of the geometrical patterns requires vessel segmentation and reconstruction. The reconstruction of the retina, especially the area of optic disc, has been performed by several researchers [6, 15], but most pathological changes in the microvasculature occur away from this region.

3D reconstruction of the retinal vascular tree from fundal images is a considerable challenge and only a few such attempts have been made so far [9]. When subject-specific retinal images are obtained with a fundus camera, the intrinsic parameters of the fundus camera-eye system will be altered by the relative displacement between the camera and the eye of the subject. These changes can be reduced by acquiring retinal images with relatively small displacements of the camera in a plane which may be assumed parallel with the surface of the lens. Consequently the intrinsic parameters may be assumed to be fixed. The retinal vessels of interest can be segmented using a semi-automatic approach [8] and the point-by-point correspondence between different images can also be calculated. In order to acquire a metric reconstruction result, an essential matrix based self-calibration approach was performed to estimate the intrinsic parameters of the "fundus camera-eye" optical system. For this the pixels of the camera-eye system are assumed to be nearly perfectly rectangular (which means that the aspect ratio is considered to be one and there is no skew) and the principal point of the camera-eye system is assumed to lie at the centre of the final retinal image. With these assumptions the self-calibration approach can be reduced to a simple system and a final metric 3D reconstruction can be recovered after retrieving the projection matrix from the essential matrix.

11.2 Methodology

11.2.1 Image Acquisition

The retinal images for reconstruction were obtained in a normal subject following mydriasis with tropicamide (1% eye drops) using a commercial retinal fundus camera (Zeiss FF 450+ with a 30° field of view (Fig. 11.1). The fundus camera is a specialized low power microscope with an attached camera designed to photograph the interior surface of the eye, including the retina, optic disc, macula, and other structures. Digitized images were captured using a CCD camera and transferred to a PC for analysis. The principle of the paired image acquisition is illustrated in Fig. 11.2 and a stereo (approximate) pair of retinal images is shown in Fig. 11.3.

Although the eye and fundus camera are very complex, they are combined and simplified as one single lens in the analysis described here. Because this special system combines the eye and the fundus camera, the displacement between them, such as the change in relative distance and rotation, will alter the intrinsic parameters of the combined eye-camera system (Fig. 11.2). In order to minimize these changes, the distance between the camera and eye was held approximately constant and only a
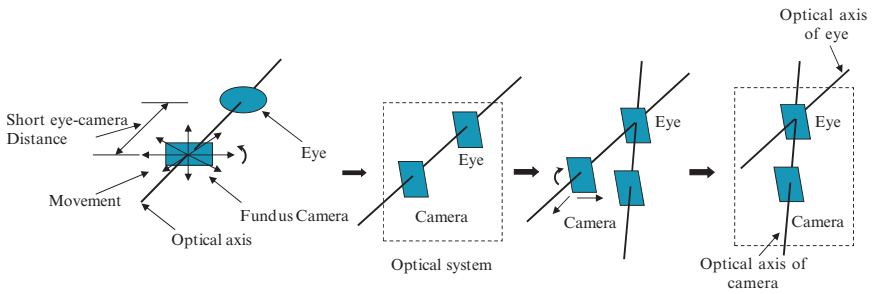
Fig. 11.1 Retinal imaging

Fig. 11.2 Schematic diagram illustrating the approach to retinal imaging
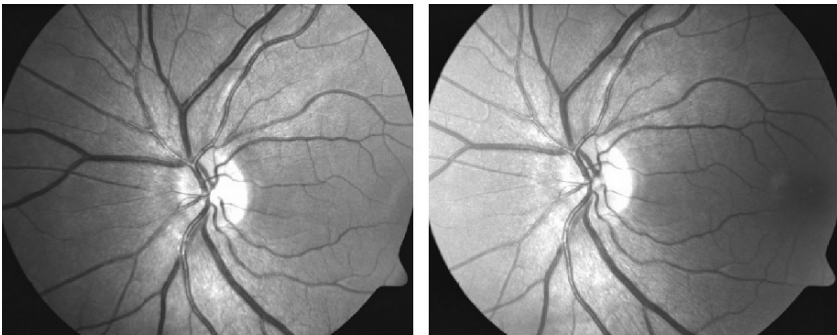

Fig. 11.3 A stereo pair of retinal images for reconstruction

small relative displacement of the fundus camera was made when the retinal images were acquired. Retinal images were acquired with a resolution of $1{,}280 \times 1{,}030$ pixels but were reduced to 499×402 pixels prior to analysis in order to reduce the computational time.
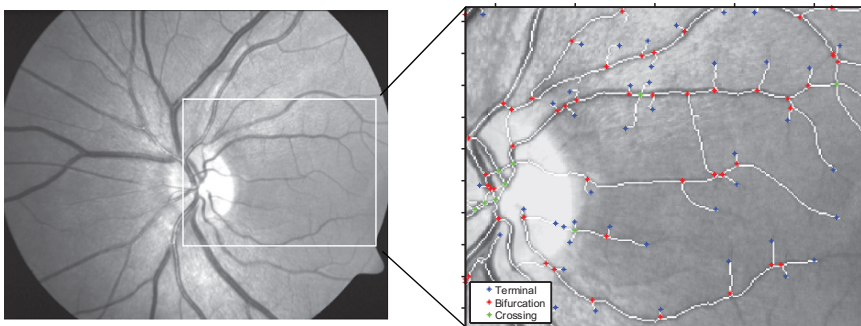
11.2.2 Feature Points Extraction and Vessel Segmentation

Depending on the theory of 3D reconstruction from 2D images, two images of a single scene are related by the epipolar geometry, which may be represented by a 3×3 singular matrix called the fundamental matrix F. It captures all geometric information contained in the two images, and must first be estimated for the reconstruction. A standard linear camera calibration matrix K has the following entries:

$$K = \begin{bmatrix} f & s & u_0 \\ 0 & \delta f & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where f is the focal length in pixels and δ is the aspect ratio. (u_0, v_0) are the coordinates of the principal points, and s is the skew factor which is zero for rectangular pixels.

Generally in order to obtain the fundamental matrix F, at least 7 corresponding fiducial points (i.e. at bifurcations) should be obtained. In fact, more than seven matches are required for an accurate estimation. Therefore a semi-automatic method [8] was used to perform the feature points extraction and vessel segmentation by scale-space analysis of the 1st and 2nd derivatives of the image intensity profiles (Fig. 11.4). The coordinates of the individual pixels corresponding to the centrelines of the segmented vessels are recorded. Because of the limited resolution of the fundus camera, only trunk branches of retinal vessels such as the 1st to 4th generations were clear enough to be analyzed (Fig. 11.5). Matching points were selected by an operator and are marked out in Fig. 11.6. If the same vessel were segmented from different images, the correspondence between them could be obtained.

Fig. 11.4 Feature points extraction. Left: original retinal image; right: analyzed vessels
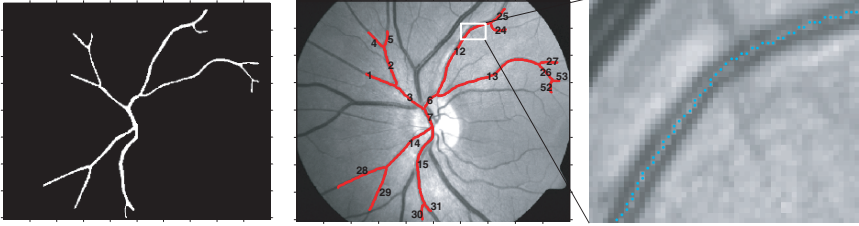
Fig. 11.5 Vessel segmentation. Middle: original retinal images; left: the segmentation result of the red marked vessels; right: the recorded vascular centreline points showing the vessel sect in the white rectangle area
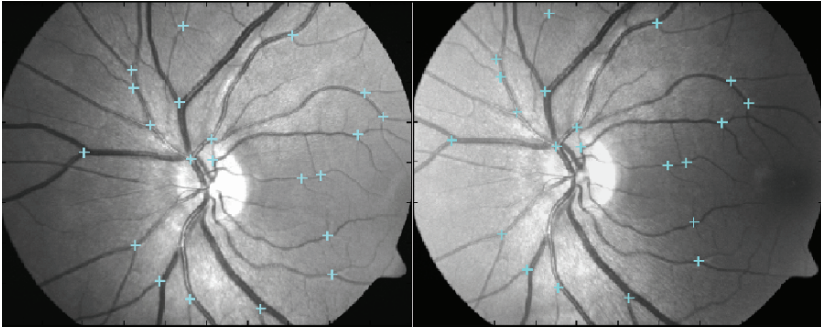
Fig. 11.6 Retinal images with marked corresponding points

11.2.3 Estimation of Epipolar Geometry

Based on the marked corresponding points (crossing and bifurcation points), the fundamental matrix F was recovered by applying the gold standard algorithm developed by Hartley and Zisserman [5] by minimizing the re-projection geometric error:

$$\sum_i d(x_i, \hat{x}_i)^2 + d(x'_i, \hat{x}'_i)^2 \tag{11.1}$$

where $x_i(u_i, v_i) \leftrightarrow x'_i(u'_i, v'_i)$ are the marked correspondences, and $\hat{x}_i \leftrightarrow \hat{x}'_i$ are the estimated correspondences that satisfy $\hat{x}_i F \hat{x}'_i = 0$ exactly for rank-2 estimated fundamental matrix F.

The gold standard algorithm was implemented by taking the following steps:

1. Obtain the initial estimated rank 2 fundamental matrix $\hat{F}$ using the normalized 8-points algorithm [3]:

    – Normalization: transform the corresponding points according to $\hat{x}_i = T x_i$ and $\hat{x}'_i = T' x'_i$, where T and T' are normalized transforms consisting of a translation and scaling.

- Find the fundamental matrix $\hat{F}$ corresponding to the matches $\hat{x}_i \leftrightarrow \hat{x}_i'$ with a linear solution and enforce the rank 2 constraint to it by Singular Value Decomposition (SVD).
- Set final fundamental matrix $F = T'^{\mathrm{T}}\hat{F}T$, where superscript T represents the transpose of the vector.

2. Compute an initial estimated subsidiary variables $\{\hat{x}_i \leftrightarrow \hat{x}_i'\}$ as follows:

   - Define two projection matrices as $P = [I\,|\,0]$ and $P' = [[e']_\times F\,|\,e']$, where e', the epipole of the second image, could be obtained from F.
   - From the correspondence $x_i \leftrightarrow x_i'$, the 3D points $\hat{X}_i$ are obtained by an iterative linear-eigen triangulation method [4] in order to save computational cost.
   - The projective correspondence consistent with F is obtained by $\hat{x}_i = P\hat{X}_i$, $\hat{x}_i' = P'\hat{X}_i$.

3. Minimize the cost function in Eq. (11.1) by varying $P' = [M|t]$ and $\hat{X}_i$ with the Levenberg-Marquardt algorithm.
4. Compute the fundamental matrix F as $F = [t]_\times M$.

This process of estimating the fundamental matrix is effectively equivalent to a projective reconstruction. Epipolar lines of the two retinal images calculated by the above algorithm are shown in Fig. 11.7.

11.2.4 Self-Calibration

In principle, a projective reconstruction can be obtained without the calibration matrix based on the fundamental matrix, F, but, in practice, due to the ambiguity of projective reconstruction, results may not be sufficiently accurate. Therefore the 3D reconstruction was made on the basis of a metric projection. It is known that a metric reconstruction of a scene may be computed by using the essential matrix E which could be derived from the calibration matrix K [5]:
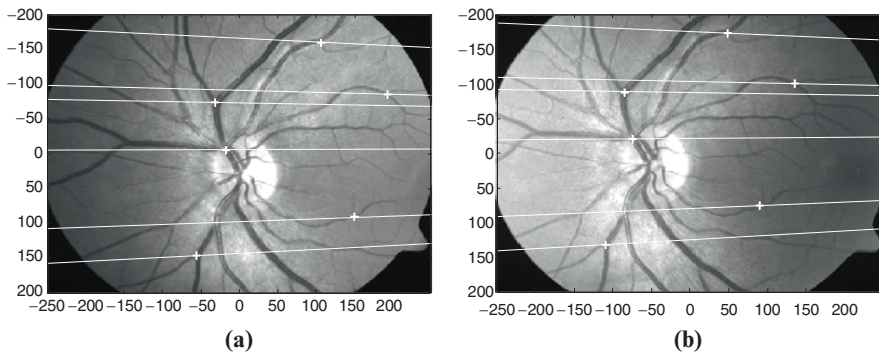

Fig. 11.7 The epipolar lines of the two retinal images

$$E = K^T F K \tag{11.2}$$

For the special fundus camera-eye system, a general photogrammetric calibration method, which depends on a calibration object with a known 3-D geometry [14, 16], is not available. Instead a self-calibration method was employed. Since the image acquisition process was specially designed to minimize alteration of the camera-eye system, we can assume that the intrinsic parameters of the camera-eye system remained constant, and the pixels could be considered as nearly perfectly rectangular with an aspect ratio of 1 and no skew. The principal point of the camera-eye system is assumed to be at the centre of the final retinal image. Therefore the only unknown parameter of the calibration matrix is the focal length.

It has been recognized that if n_k is the number of intrinsic parameters known in all views and n_f is the unknown but constant intrinsic parameter, the number of views, m, required for self-calibration will be:

$$mn_k + (m-1)n_f \geq 8 \tag{11.3}$$

Therefore, a minimum of two views is required in this case [11].

A self-calibration method, based on the characteristics of the essential matrix E, was used to recover the unknown focal length: two of the three singular values of E should be identical and the other should be zero [11]. The cost function:

$$C = \omega_{12} \frac{\sigma_1 - \sigma_2}{\sigma_2} \tag{11.4}$$

was minimized by a direct search algorithm. $\sigma_1 > \sigma_2$ are the non-zero singular values of $E = K^T F K$, and ω_{12} is a weighting factor which represents the degree of confidence in the estimation of the fundamental matrix F. There are several possible choices for ω_{12}: (i) the residual of the estimation of F – the inverse of the mean geometric distance between the image points and their corresponding epipolar lines, (ii) the number of points used in the computation of F, and (iii) simply set it to one [1]. After obtaining the focal length by this self-calibration method, the essential matrix E could be calculated from Eq. (11.2).

Apart from sensitivity to the noise of images, the application of self-calibration is always dependent on the issue of initialization. Since the nonlinear minimization is used for self-calibration, convergence to the global minimum can be guaranteed only if the algorithm is initialized properly. Although this algorithm has a good global convergence according to Fusiello [2], we still employed a planar chessboard based photogrammetric calibration approach [7] to generate initial values for self-calibration. A series of images of a 2-D chessboard plane were acquired by moving and rotating the fundus camera (Fig. 11.8). The chessboard should not be exactly parallel to the image plane, because the calibration method applied here depends on the vanishing points. Otherwise, the vanishing points for the horizontal and vertical lines of the planar chessboard would both be at infinity, and no solution would exist (Fig. 11.9). The guessed values for K here are based only on the optics of the fundus camera. Therefore, it does not include the optics of eye, but should be close enough to the values for the whole optical system.
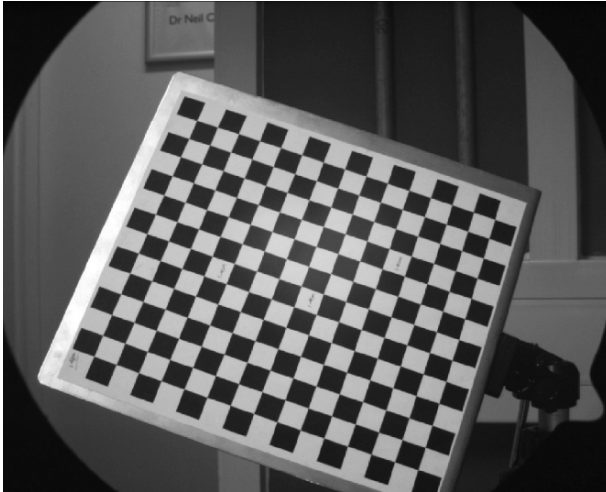
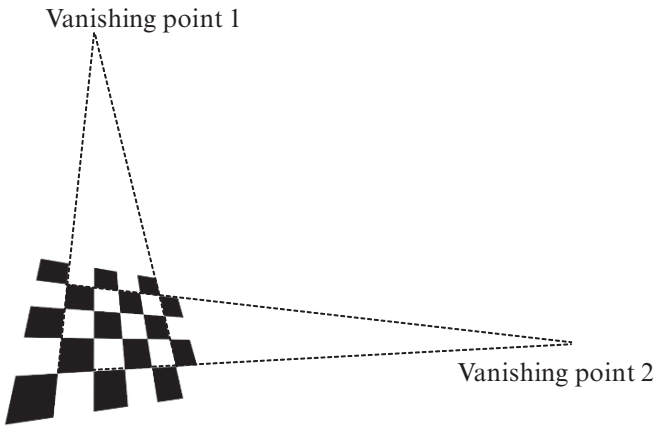**Fig. 11.8** Planar chessboard used for calibration

Fig. 11.9 Vanishing points of a perspective chessboard

11.2.5 Recovery of the Projection Matrix

The corresponding points $x_i \leftrightarrow x_i'$ in the 2D images and the unknown 3D points, X_i, on the object have the relationship:

$$x_i = PX_i, \quad x_i' = P'X_i \tag{11.5}$$

here P and P' are the two projective matrices. If the essential matrix E of the camera-eye system was obtained, the matrices P and P' could be retrieved from E as follows [5]:

1. Suppose the origin of the coordinate system is at the first camera centre, then the two projection matrices could be factorized as $P = K[I\,|0]$ and $P' = K[R\,|t]$, where R and t are the 3D displacements (rotation and translation) from the global coordinate system to the camera coordinate system. If the calibration matrix K was known, its inverse can be applied to the points x_i and x'_i to obtain their normalized coordinates:

$$\bar{x}_i = K^{-1}x_i = [I\,|\,0]X_i, \quad \bar{x}'_i = K^{-1}x'_i = [R\,|\,t]X_i$$

Thus the corresponding normalized projection matrices are

$$\bar{P} = [I\,|\,0], \quad \bar{P}' = [R\,|\,t]$$

2. The essential matrix can be calculated from Eq. (11.2), or

$$\text{expressed as } E = [t]_\times R = SR$$

where S is the skew-symmetric matrix of t.
Let the SVD of E be UDV^T where $D = \text{diag}(k,\ k,\ 0)$, then the possible factorization of $E = SR$ is one of the following:

$$S = UZU^T; \ R = UWV^T \text{ or } UW^TV^T$$

where

$$Z = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Hence the normalized projection matrix $\bar{P}'$ has four possible choices based on SVD of E as follows:

$$\bar{P}' = [UWV^T\,|+u_3] \text{ or } [UWV^T\,|-u_3] \text{ or}$$
$$[UW^TV^T\,|+u_3] \text{ or } [UW^TV^T\,|-u_3]$$

here $u_3 = U(0,0,1)^T$, the last column of U.
3. Finally the P and P' can be calculated by

$$P = K\bar{P} \text{ and } P' = K\bar{P}'.$$

Based on the fact that reconstructed points should be in front of both cameras, the correct solution may be determined by testing a single point if it is in front of both cameras.

At the same time the extrinsic parameters, the rotation axis l and the angle of rotation λ may be obtained as:

$$l = (R_{32} - R_{23}, R_{13} - R_{31}, R_{21} - R_{12})^T \tag{11.6}$$

$$\lambda = \arccos\left(\frac{trace(R) - 1}{2}\right) \tag{11.7}$$

The projection matrix P and P' are computed according to the essential matrix E obtained above.

11.2.6 Parameter Refinement

Based on the corresponding points obtained from vessel segmentation, the extrinsic parameters R and t, and the intrinsic parameter f were refined by minimizing the first-order geometric error cost function using the Levenberg-Marquardt algorithm.

$$C_r = \sum_i \frac{(x_i' \tilde{F} x_i)^2}{(\tilde{F} x_i)_1^2 + (\tilde{F} x_i)_2^2 + (\tilde{F}^T x_i')_1^2 + (\tilde{F}^T x_i')_2^2} \tag{11.8}$$
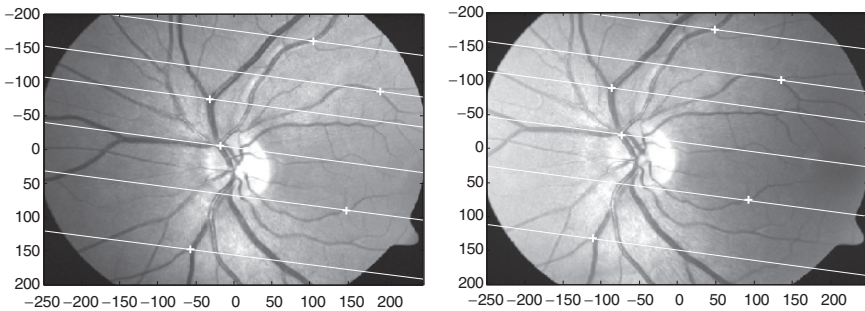
Here the fundamental matrix $\tilde{F}$ was calculated as

$$\tilde{F} = K^{-T}[t]_\times R' K^{-1} \tag{11.9}$$

where R' is the rotation matrix calculated from Rodrigues' formula [5]

$$R' = I + \sin\lambda \left[\frac{l}{\|l\|}\right]_\times + (1 - \cos\lambda)\left[\frac{l}{\|l\|}\right]_\times^2. \tag{11.10}$$

After this refinement, the final projection matrix could be obtained from R, t and K. The corresponding epipolar lines are displayed in Fig. 11.10.

Fig. 11.10 The epipolar lines after parameter refinement

11.2.7 Reconstruction of 3D Points

Knowing the projection matrices for two images separately, the 3D coordinates of each point, X_i, can be calculated. In order to obtain a smooth reconstruction, the corresponding points from vessel segmentation were smoothed and interpolated using cubic splines.

An iterative linear method (Iterative-Eigen) [4] was used to perform the triangulation of 3D points. Equation (11.5) can be written as:

$$\begin{cases} \frac{1}{w_{i,j}}(u_i p^{3T} X_{i,j} - p^{1T} X_{i,j}) = 0 \\ \frac{1}{w_{i,j}}(v_i p^{3T} X_{i,j} - p^{2T} X_{i,j}) = 0 \\ \frac{1}{w'_{i,j}}(u'_i p'^{3T} X_{i,j} - p'^{1T} X_{i,j}) = 0 \\ \frac{1}{w'_{i,j}}(v'_i p'^{3T} X_{i,j} - p'^{2T} X_{i,j}) = 0 \end{cases} \tag{11.11}$$

where p^{iT} and p'^{iT} are the ith rows of P and P' respectively. $w_{i,j}$ and $w'_{i,j}$ are the weight factors at the jth step of iteration which have the form:

$$w_{i,j} = p^{3T} X_{i,j-1}, \quad w'_{i,j} = p'^{3T} X_{i,j-1} \tag{11.12}$$

At the beginning of the iteration $w_{i,0}$ and $w'_{i,0}$ were set to be 1 in order to find the initial solution of $X_{i,0}$.

All calculations were performed using Matlab (The Mathworks). The parameter refinement was based on the codes from the Oxford Brookes toolbox (http://cms.brookes.ac.uk/staff/PhilipTorr/). The flowchart of a complete reconstruction process is shown in Fig. 11.11.

11.3 Experiment and Results

A cylinder with a chessboard attached to its external surface was adopted to understand and obtain a preliminary validation of the metric projective reconstruction procedure. The focal length and aperture of the camera (Nikon D50) were fixed when pictures of the chessboard were taken, which means that the intrinsic parameters of the camera optical system were kept constant. Four images were acquired and used (Fig. 11.12 left) for self-calibration. The corners of the chessboard were extracted automatically for reconstruction (Fig. 11.12 right). The epipolar geometry between the image pairs was calculated by the normalized 8-points algorithm. After self-calibration, the 3D coordinates of the corners of the chessboard were calculated by the Iterative-Eigen triangulation (Fig. 11.13). Although the final triangulation of the 3D object coordinates was only based on two images, the self-calibration could utilize more images to improve the accuracy of the estimated camera intrinsic parameters. Quantitative comparison between the measurements and reconstruction results is summarized in Table 11.1. It is clear that with good estimation of the
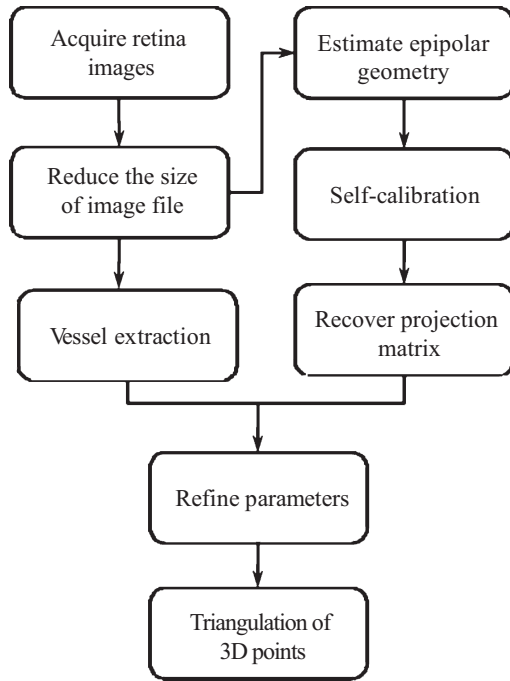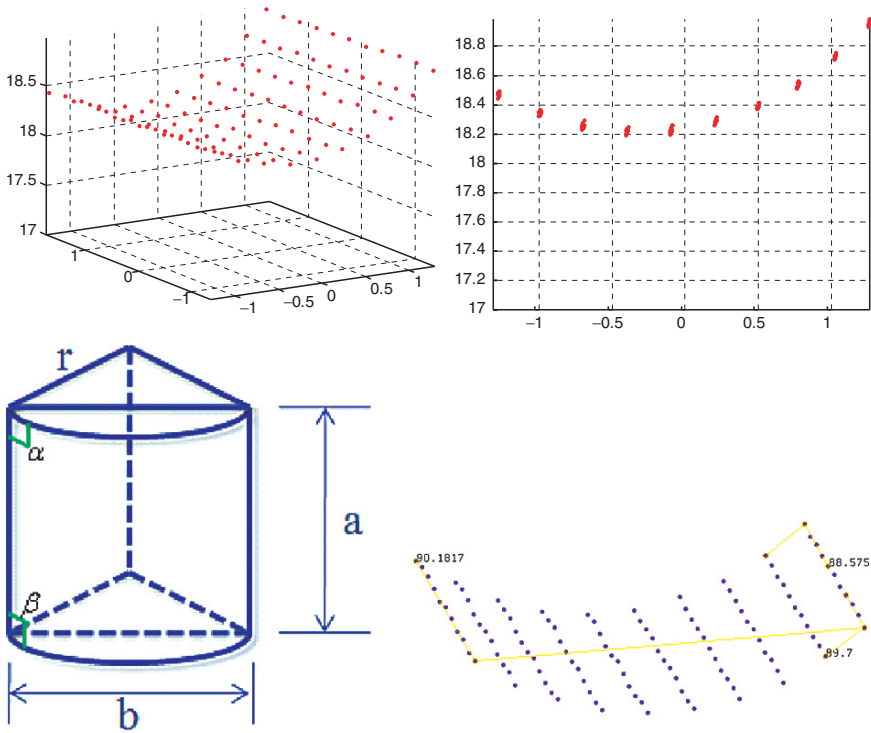
Fig. 11.11 The reconstruction process

Fig. 11.12 Object image (left) and image with extracted corners (right)

epipolar geometry and self-calibration, the reconstruction procedure implemented here is capable of recovering very well most of the geometric features such as height to length ratio (a/b), angles (α and β) and curvature (r/b). The percentage error in curvature (r/b) seems to be larger than that in the other parameters, possibly due to the uncorrected distortion in the images and the smaller value for r making it more sensitive to errors.

Figure 11.14 shows the reconstruction results of the segmented vessels in the central region of the image presented in Fig. 11.5. Figure 11.14a is the 2-D view of
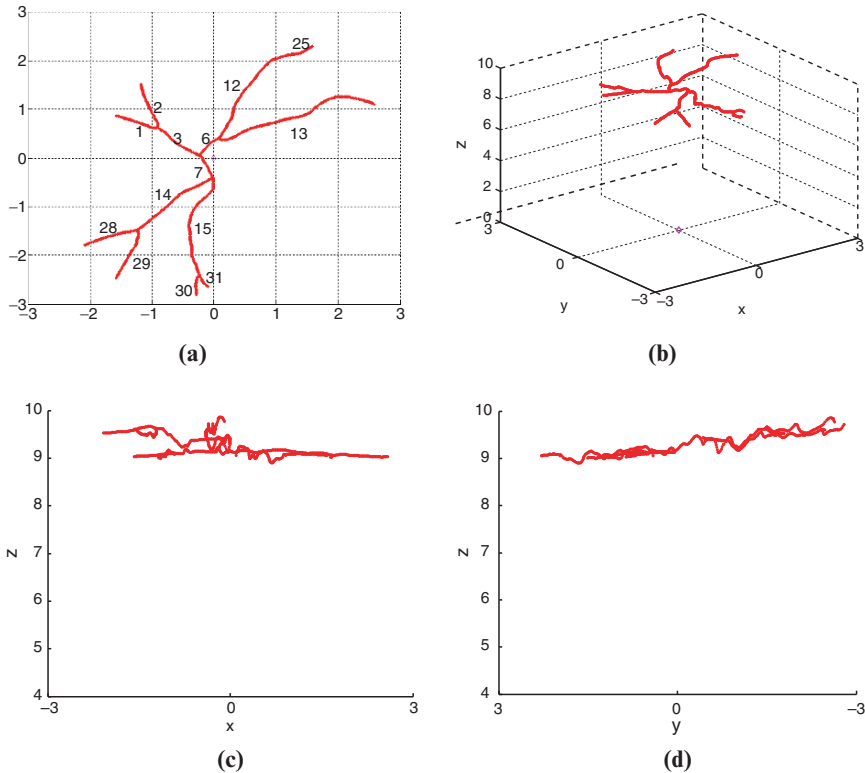
Fig. 11.13 Three-dimensional view of the reconstruction result (top left) and its side view (top right). Geometry model of the cylindrical chessboard (bottom left) and reconstruction result with angle measurements in ICEM (bottom right)

Table 11.1 Quantitative comparison of the reconstructed results and measurements. See Fig. 11.13 for definition of geometric parameters

	Measured (M)	Reconstructed (R)	Error $= |(M-R)/M|*100\%$
a/b	1.1951	1.1778	1.45%
r/b	0.8415	0.7771	7.66%
α	90	89.7	0.3%
β	90	88.575	1.58%

the recovered centreline points of the marked 3-D vessels. The numbers (according to Fig. 11.5) indicate which segments of the marked vessels are retrieved. Figure 11.14b is the 3-D view of the reconstructed vessel centrelines. Figure 11.14c and d show the side projections of the centrelines of the reconstructed vessels. These preliminary results demonstrate that 3D reconstruction of the retinal vessels is feasible and may be useful in future studies of the retinal microvasculature in health and disease. Future studies will attempt to validate the 3-D reconstruction ex vivo using a model eye with known geometry and in vivo using ultrasound to measure geometric features.

Fig. 11.14 Three-dimensional reconstruction results of the centre-lines of the marked retinal vessels: (**a**) top view of the reconstructed vessels; (**b**) 3D view; (**c**), (**d**) side projections

References

1. Dick A, Torr P, Cipolla R (2000) Automatic 3d modelling of architecture. In: Proceedings of 11th British Machine Vision Conference (BMVC'00), Bristol, pp 372–381.
2. Fusiello A (2001) A new autocalibration algorithm: experimental evaluation. Computer analysis of Images and Patterns 2001, Lecture Notes in Computer Science 2124: 717–724.
3. Hartley RI (1997) In defense of the eight-point algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence 19: 580–595.
4. Hartley RI, Sturm P (1997) Triangulation. Computer Vision and Image Understanding 68(2): 146–157.
5. Hartley R, Zisserman A (2003) Multiple View Geometry in Computer Vision. Cambridge: Cambridge University Press.
6. Kai Z, Xu X, Zhang L, Wang GP (2005) Stereo matching and 3-D reconstruction for optic disk images. CVBIA, LNCS 3765: 517–525.
7. Liu Y, Wu YX, Wu MP, Hu XP (2004) Planar vanishing points based camera calibration. In: Proceedings of the Third International Conference on Image and Graphics (ICIG'04), Hong Kong, China, pp 460–463.

8. Martinez-Perez ME, Hughes AD, Stanton AV, Thom SA, Chapman N, Bharath AA, Parker KH (2002) Retinal vascular tree morphology: a semi-automatic quantification. IEEE Transactions on Biomedical Engineering 49(8): 912–917.
9. Espinosa-Romero A, Martinez-Perez ME (2005) Optical 3D reconstruction of retinal blood vessels from a sequence of views. In: Proceedings of SPIE 5776, Merida, Mexico, pp 605–612.
10. Masters B (2004) Fractal analysis of the vascular tree in the human retina. Annual Review of Biomedical Engineering 6: 427–452.
11. Mendonca PRS, Cipolla R (1999) A simple technique for self-calibration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, Colorado, pp 500–505.
12. Stanton AV, Wasan B, Cerutti A, Ford S, Marsh R, Sever PP, Thom SA, Hughes AD (1995) Vascular network changes in the retina with age and hypertension. Journal of Hypertension 13:1724–1728.
13. Wong TY, McIntosh R (2005) Systemic associations of retinal microvascular signs: a review of recent population-based studies. Ophthalmic and Physiological Optics 25: 195–204.
14. Xu J, Chutatape O (2003) Comparative study of two calibration methods on fundus camera. In: Proceedings of the 25 Annual International Conference of the IEEE EMBS, Cancun, Mexico, pp 17–21.
15. Xu J, Chutatape O (2006) Auto-adjusted 3D optic disk viewing from low-resolution stereo fundus image. Computers in Biology and Medicine 36: 921–940.
16. Zhang Z (1999) Flexible camera calibration by viewing a plane from unknown orientations. In: Proceedings of the International Conference on Computer Vision, Corfu, Greece, pp 666–673.