

# Integrating stereo and photometric stereo to monitor the development of glaucoma

Simon Lee and Michael Brady

Robotics Research Group,  
Department of Engineering Science,  
University of Oxford

---

*We report progress on a system to monitor the development of glaucoma by measuring the topography of the optic disk. The need for an accurate method for doing this using passive vision is explained. Sparse depth measurements from stereo matching of blood vessels provide insufficient constraint for reconstructing the surface of the optic disk. Shape-from-shading has to contend with a spatially-varying albedo. We show how a particular arrangement of fundus cameras allows us to apply a technique akin to homomorphic filtering to recover estimates of  $\partial z/\partial x$  that can be smoothed by appropriate regularisation. Stereo is integrated with this photometric stereo depth estimates. Examples of an artificial object and optic disk surface reconstructions are presented.*

---

This paper reports progress on a system to aid physicians monitor the development of glaucoma, a complex of eye diseases that annually cause blindness in an estimated 50,000 people in the United States, with similar figures for Europe. In most cases, blindness could be prevented if patients were treated at an early stage; in practice this would require automated analysis and monitoring.

The next section summarises some basic facts about glaucoma. The bottom line is that measuring and monitoring changes in the topography of a structure called the optic disk is key to assisting in the diagnosis and treatment of glaucoma. The challenge for computer vision is that the optic disk is small, at a relatively large distance from any imager external to the eye, the albedo of the disk varies spatially, there is a sparse network of blood vessels, and there is considerable variation in the shape and size of healthy disks. As the next section explains, ophthalmologists often take stereo pairs of photographs to enable them to view the cup and 3D shape of the optic disk. Stereo matching of the blood vessels shown in Figure 1 is relatively straightforward, and is discussed in Section 4. The need for accurate depth measurements mean that excessive smoothing cannot be tolerated; but current edge-based photogrammetric methods suffice. Sparse depth measurements from stereo matching of blood vessels provide insufficient constraint for reliable accurate reconstruction of the surface of the optic disk, which is mostly smoothly shaded. Shape-from-

shading offers hope of dense depth measurements; however, it has to contend with the spatially varying albedo of the surface of the optic disk. Fortunately, we can control the lighting and positions of the stereo pair of cameras. Section 2 describes how we can exploit this degree of freedom to apply a technique akin to homomorphic filtering to separate the contribution of albedo and reflectance map. This provides an estimate of a component of the gradient  $\nabla z$  of the disk surface, and after integration, an estimate of the topography of the disk. The gradient components are noisy, so Section 3 describes how we regularise them using a weak prior model of the structure of the disk.

Section 5 reports work on integrating the sparse stereo and dense photometric stereo estimates. Finally, Section 6 presents results using the system we have developed.

## 1 MEDICAL BACKGROUND

Glaucoma refers to a complex of diseases that have in common an increase in intraocular pressure. This elevation of pressure causes excavation of the optic disk, which leads to loss of the visual field and eventually blindness may result.

The optic disk composed of the anterior termination of the optic nerve. It is located an average of  $3.42 \pm 0.34$  mm from the fovea; the disk itself has an average horizontal diameter of  $1.75 \pm 0.19$  mm and an average vertical diameter of  $1.86 \pm 0.21$  mm. (The average diameter of an eye ball is approximately 23–25mm). The internal disk surface is not in general flat. Almost all eyes have some form of *cup* or physiological excavation of the disk.

Observation of the optic disk is fundamental to the management of any patient suspected of glaucoma. In a glaucomatous eye, the deformation of the optic nerve head causes the death of the nerve fibres or the ganglion cells. It is thought that cupping occurs before nerve damage, so that it is vital to monitor the topography of the optic nerve head in the ocular hypertensive patient if glaucoma is to be detected early. Since the size of the optic nerve head is governed by many physiological factors, *absolute*

measurements are not much use for diagnosis. To differentiate the pathological from the physiological features, the patient's eyes are examined periodically, usually every six months. If on successive examinations, there are changes in the size of the cup or the color, the pallor, the disease is diagnosed. In order to record and compare the changes, ophthalmologists take photographs of the fundus and make measurements on parameters such as the cup/disk ratio, the rim area, etc. In addition, stereoscopic photographs allow them to assess and compare the depth of the cup as well as the 3D shape of the optic disk.

The fundus images we used were taken by the Zeiss fundus camera with the Allen stereo separator. The separator is a rotating glass piece and stereo photographs were taken with the separator at two different positions. The illumination source comes from the camera, and the Allen separator also changes the effective direction of the source. The photographs were digitized by a CCD camera.

As noted in the Introduction, the bottom line is that measuring and monitoring changes in the topography of the optic disk is key to assisting in the diagnosis and treatment of glaucoma. The challenge for computer vision is that the optic disk is small (1.8 mm diameter), at a distance of at least 25 mm from any imager, the albedo of the disk varies spatially, there is a sparse latticework of blood vessels, and there is considerable variation in healthy disks.

## 2 PHOTOMETRIC STEREO

Shape from shading relates the shape  $z(x, y)$  of an imaged surface (or its gradient) to image irradiance (grey level shading). Typically, image irradiance is proportional to scene radiance, which depends upon the surface albedo  $\alpha(x, y)$ , the energy  $E_s$  of the scene irradiance, and the reflectance function  $R(p(x, y), q(x, y))$ , where  $p(x, y) = \partial z / \partial x$  and  $q(x, y) = \partial z / \partial y$ . Shape from shading from a single image is underconstrained [8].

The majority of shape from shading algorithms require all the factors listed above to be known or well-controlled. For example, the surface reflectance function  $R$  is usually assumed to be Lambertian; the albedo is assumed constant over the surface; the light source position and intensity must be known; and shadowing and mutual illumination are not allowed. In the case of optic disk images, the albedo most definitely is not constant, though the assumption of Lambertian reflectance appears to be quite reasonable. No direct evidence is available to support this hypothesis, particularly as living tissue has different reflectance characteristics than dead tissue (available pathologically). However, in our

previous work on cataract [7], a light source was focussed on the optic disk and used to reflect and diffuse the light so that the lens is illuminated from behind. The polarisation characteristics of the lens and fundus images exploited in that work strongly suggest that the reflectance of the cup is approximately Lambertian.

If the brightness incident on the optic disk (scene irradiance)  $I_s$  and the direction of illumination  $\mathbf{n}_s$  are known, there remain three unknowns: the surface normal  $\mathbf{n}(x, y)$  and the albedo. Assuming Lambertian reflectance, we have

$$I(x, y) = \alpha(x, y) I_s (\mathbf{n}_s \cdot \mathbf{n}(x, y))$$

Inspired by photometric stereo [8, 15], we consider taking two images, from different vantage points  $\mathbf{n}_{s_i} = [p_{s_i}, q_{s_i}]^T$ . Then:

$$I_1(x, y) = \alpha(x, y) I_{s_1} \frac{1 + p_{s_1} p + q_{s_1} q}{[(1 + p^2 + q^2)(1 + p_{s_1}^2 + q_{s_1}^2)]^{\frac{1}{2}}}$$

$$I_2(x, y) = \alpha(x, y) I_{s_2} \frac{1 + p_{s_2} p + q_{s_2} q}{[(1 + p^2 + q^2)(1 + p_{s_2}^2 + q_{s_2}^2)]^{\frac{1}{2}}}$$

We now *choose* to place the cameras symmetrically (about a cyclopean centre), oriented in the  $x - z$  plane and to illuminate the scene with equal energy light sources: that is,

$$\begin{aligned} q_{s_1} &= q_{s_2} = 0, & \text{and} \\ p_{s_1} &= -p_{s_2} = p_s, & \text{and} \\ I_{s_1} &= I_{s_2} = I \end{aligned}$$

Then the equations simplify to:

$$I_1(x, y) = \alpha(x, y) I \frac{1 + p_s p}{[(1 + p^2 + q^2)(1 + p_s^2)]^{\frac{1}{2}}}$$

$$I_2(x, y) = \alpha(x, y) I \frac{1 - p_s p}{[(1 + p^2 + q^2)(1 + p_s^2)]^{\frac{1}{2}}}$$

so that

$$\begin{aligned} \frac{I_1}{I_2} &= \frac{1 + p_s p}{1 - p_s p} = M \quad (\text{Homomorphic filter}) \\ p &= \frac{M - 1}{p_s (M + 1)} \end{aligned}$$

Only  $p(x, y)$  is recovered by this method, but it is sufficient to estimate depth by integration:

$$z(x, y_i) = z_o(x_o, y_i) + \int_{x_o}^x p(x, y_i) dx$$

(for  $i = 0, \dots, N$ )

Our photometric stereo technique provides no information about  $q$  of the surface gradient, therefore surface reconstruction is not possible without the knowledge of  $z_o$  in each profile. This is supplied by binocular stereo results which is described in Section 4. We find that

this technique gives good, though noisy, estimates for the horizontal component  $p$  of the surface gradient; but we should be aware of the errors associated with the technique:

1. The photometric stereo technique requires exact correspondence of each point in the images. To define a point in a region where the brightness is varying slowly and to match it with its conjugate pair in the other image is extremely difficult. In the special camera system we are designing, a single camera with two illumination sources are used, the correspondence problem is then trivial. For the moment, however, the fundus images are binocular pairs instead of photometric stereo pairs. The matching problem is solved by matching edges (unwanted edges are removed by hand), measuring the disparity, and using linear approximation to estimate the disparity of the image points between edges. The error may not be as large as it seems since the brightness changes slowly in general, and in places where it changes rapidly, matching is usually assisted by edge points.
2. Even when the illumination directions of the cameras are known accurately, the effective illumination direction will be changed by the optics of the eye. This produces incorrect scaling of the depth profiles.
3. To calculate depth, the gradient of each point has to be integrated; but this also integrates up the error. Next section describes a method which will smooth out most of the random noise.
4. The surface of the retina is modelled as a Lambertian reflector, and though we argued above that this appears to be a good approximation, it is not known for certain.
5. We assume that the scene irradiance  $I_s$  is equal for both images. It is difficult to ensure that two light source intensities are equal. However, standard techniques for normalising image brightness appear to cope well with this problem.
6. The surface of the cup is concave, so mutual illumination occurs. Forsyth [5] have shown that ignoring mutual illumination leads to underestimation of depth. It has been found that the extent of mutual illumination is related to the albedo. If the surface reflectivity is large, the mutual illumination effect is large; but if the albedo is low, the mutual illumination effect is low. The albedo cannot be changed, but is a function of the wavelength of illumination. Since the retinal surface is red and yellow in colour, green colour light is a good candidate for illumination since the reflectivity of green light is low and green light can increase the contrast between the optic disk and blood vessels.

### 3 REGULARISATION

The technique described in the previous section gives good, though noisy, estimates of the surface gradient component  $p(x, y)$ . To smooth the estimates we apply regularisation, that is, generalised least squares. The simplest approach is to minimise:

$$\iint (\mathcal{P} - p)^2 + \lambda (p_x^2 + p_y^2) dx dy$$

where  $\mathcal{P}(x, y)$  is the initial estimate computed by the technique outlined above, and where  $\lambda$  is a non-negative weighting constant. We find that if  $\lambda$  is assumed constant, the surface is smoothed too severely, especially near the edge of the optic cup where the surface curvature is large. Instead, we suppose that  $\lambda(x, y)$  is a function which is allowed to vary spatially and we minimise  $p(x, y)$  and  $\lambda(x, y)$  simultaneously. That is, we minimise:

$$\iint (\mathcal{P} - p)^2 + \lambda (p_x^2 + p_y^2) + \mu (\lambda_x^2 + \lambda_y^2) + \gamma (\Lambda - \lambda)^2 dx dy$$

where,  $\Lambda(x, y)$  is a prior estimate of  $\lambda(x, y)$ . One possible prior estimate is an idealised model such as that shown in Figure 2f(top). In practice, we compute the approximation to the ideal as a function of the gradient  $\nabla \mathcal{P}$  (Figure 2f bottom) which we find gives good results. The Euler-Lagrange equations are then:

$$\begin{aligned} F_p - \frac{\partial}{\partial x} F_{p_x} - \frac{\partial}{\partial y} F_{p_y} &= 0 \\ F_\lambda - \frac{\partial}{\partial x} F_{\lambda_x} - \frac{\partial}{\partial y} F_{\lambda_y} &= 0 \end{aligned}$$

that is:

$$\begin{aligned} (\mathcal{P} - p) + \lambda \nabla^2 p + \lambda_x p_x &= 0 \\ (p_x^2 + p_y^2) + 2\gamma(\lambda - \Lambda) - 2\mu \nabla^2 \lambda &= 0 \end{aligned}$$

The corresponding finite difference iterative scheme is:

$$\begin{aligned} p_{i,j}^{k+1} &= \frac{1}{1 + \lambda_{i,j-1}^k + \lambda_{i-1,j}^k + 2\lambda_{i,j}^k} (\mathcal{P}_{i,j} + \lambda_{i-1,j}^k p_{i-1,j}^{k+1} \\ &\quad + \lambda_{i,j-1}^k p_{i,j-1}^{k+1} + \lambda_{i,j}^k p_{i+1,j}^k + \lambda_{i,j}^k p_{i,j+1}^k) \\ \lambda_{i,j}^{k+1} &= \frac{1}{2\gamma + 8\mu} [2\gamma \Lambda_{i,j} + 2\mu (\lambda_{i-1,j}^{k+1} + \lambda_{i,j-1}^{k+1} + \lambda_{i+1,j}^k \\ &\quad + \lambda_{i,j+1}^k) - (p_x^2 + p_y^2)] \end{aligned}$$

$\mu$  controls the smoothness of the function  $\lambda$ , and  $\gamma$  prevents  $\lambda$  deviates from its prior estimate. There is no strict rule to choose  $\mu$  and  $\gamma$ ; their choices depend on the type of surface and how noisy the images are. For fundus images, surface discontinuities are not common, thus  $\lambda$  should not vary too rapidly and  $\mu$  is usually chosen to be large. The images we obtained at present are noisy, the prior estimate of  $\Lambda$  based on these noisy images are therefore not very accurate and  $\gamma$  is usually chosen to be small. Experiments have been done and it is found that the result is not very sensitive to a large range of  $\mu$  and  $\gamma$ .

## 4 BINOCULAR STEREO

Many approaches to stereo vision have been developed, including grey level correlation matching and edge-based stereo. Each technique has advantages and disadvantages, its efficacy depending upon the application. The optic nerve head is essentially featureless apart from blood vessels and the edge of the optic disk. Since the shading usually varies slowly over large areas of the cup, much of which has a slowly varying gradient, grey level correlation is inappropriate. We choose to use the edge-based stereo algorithm PMF [12], so that matches are confined to blood vessels and the edge of the optic disk.

Initially we used the Canny edge detector to determine the set of edges for stereo matching. There are a number of problems with this. First, the scale of the Gaussian smoothing filter trades localisation for signal-to-noise. The poor localisation that results in a “clean” edge map is inappropriate for an (essentially photogrammetric) application such as ours in which depth accuracy is crucial. Second, the Canny edge detector fails to detect junctions reliably, usually leaving them unconnected. Instead, we use the morphological edge detector developed by Noble [11]. It does not require the smoothing step, thus avoiding the problems with the Canny edge detector. The output is clean, the position of the edges more accurate, and junctions are preserved. This is important for edge-based stereo since it reduces the chances of mismatch and also helps to provide accurate depth measurement.

To get the best view of the fundus, the optical axes of the fundus cameras are pointed at, and focussed on, the optic nerve head, so are not parallel. (Although the Allen stereo separator gives effectively parallel camera axes, the beam is bended by the optics of the eye.) This arrangement maximises stereo separation to give maximum depth sensitivity. However, the stereo images need to be rectified to facilitate matching. Rectification is done crudely at present since the camera system was designed for monocular viewing and due to the movement of the eye during the image acquisition process, the camera is therefore impossible to calibrate. We use the registration process developed by Hanna [7] in his work on monitoring the development of cataract. Matching vectors are computed between the edges, and the stereo pair are then rotated and translated accordingly. This yields epipolars that are approximately parallel and horizontal.

Result of using PMF are shown in Figure 1d. As expected, depth measurements are sparse and are only available along the edges of blood vessels and on the edge of the optic disk. Since the camera is not calibrated at present, disparity is relative rather than absolute, so the depth computed is not accurate. This will be achieved by a CCD stereo imaging system we have designed.

## 5 INTEGRATION OF DEPTH ESTIMATES

Two independent depth measurements are available from photometric stereo and edge-based stereo. The depth map from photometric stereo is dense, but is subject to the errors enumerated in Section 2. On the other hand, the depth estimates from edge-based stereo are sparse but (relatively) accurate (and can be expected to be sufficiently accurate once the imaging system is available). A more accurate reconstruction of the fundus surface can be achieved if the two measurements are combined. The question to be addressed is: how?

A popular approach Grimson, Terzopoulos, Blake and Zisserman is to formulate a suitable physical analogy, minimising the corresponding energy equation. In this vein, suppose that we are given a thin elastic plate, whose unstressed (initial) shape is that determined by the dense depth estimates given by photometric stereo. This surface is to deform such that the corresponding surface points conform to the depths given by binocular stereo. More precisely,

$$(x, y, z) \rightarrow (x, y, \hat{z}) = (x, y, z + f(x, y, z))$$

subject to  $(x, y, \hat{z}) = (x, y, z_{binocular})$ ,  $\forall (x, y) \in \zeta$  where

$$\zeta = \{(x, y) \mid \text{there is a binocular stereo depth measurement at the point } (x, y)\}$$

We next suppose that the deformation energy is proportional to the integral of the quadratic variation [6]:

$$\iiint \left( \left( \frac{\partial^2 U}{\partial x^2} \right)^2 + \left( \frac{\partial^2 U}{\partial y^2} \right)^2 + \left( \frac{\partial^2 U}{\partial z^2} \right)^2 + 2 \left( \frac{\partial^2 U}{\partial x \partial y} \right)^2 + 2 \left( \frac{\partial^2 U}{\partial x \partial z} \right)^2 + 2 \left( \frac{\partial^2 U}{\partial y \partial z} \right)^2 \right) dx dy dz$$

The aim is to find the functional  $U(x, y, z)$  that minimises the quadratic variation subject to the constraint of conforming to the stereo data. The Euler-Lagrange equation for the quadratic variation is the biharmonic equation:

$$\nabla^4 U = 0$$

$U = |r|$  is a solution to the biharmonic equation. Suppose that  $U$  is any such solution, Bookstein [3] has recently pointed out that:

$$g(x, y, z) = a_0 + a_x x + a_y y + a_z z + \sum_{i=1}^n w_i U(|P_i - (x, y, z)|)$$

also satisfies the biharmonic equation, where  $P_i$  denotes the set of depth measurements from binocular stereo and  $w_i$  are scalars. The initial (linear) part of the expression

for  $g$  defines an affine transform whose quadratic variation is zero; that is, it contributes nothing to the energy expression. The affine transform allows the initial surface defined by the photometric stereo depth estimates to be translated and rotated before deforming it in depth.

## 6 EXAMPLES

Two examples are shown. The first experiment uses a pair of fundus stereo photographs, captured by a Zeiss fundus camera. Photometric stereo and binocular stereo results are shown in Figure 1c and 2d respectively. The photometric stereo result is not as good as that of the second experiment due to the fact that the fundus images used are binocular stereo pair and the matching method described in Section 2 is used to solve the correspondence problem. Cross-sections of the two set of results near the center of the optic cup are shown in Figure 1e, with the result of integration superimposed.

The second experiment was set up using a half tennis ball sit on a piece of cardboard, under relatively well-controlled photometric stereo illumination condition. The raw images, the output from photometric stereo, and the surface reconstructed are shown in Figure 2a, b, c, d respectively. A cross-section of the gradient map is also shown (Figure 2e) and the result of the regularisation smoothing is superimposed. The idealised  $\Lambda$  and its approximation are shown in Figure 2f. Finally, the surface reconstructed is displayed in Figure 2g.

## 7 CONCLUSIONS AND FURTHER WORK

We are currently performing a small clinical assessment of our technique. We hope to compare results using our technique with those obtained for the same eyes (ours) using the Rodenstock Analyzer. As was noted in Section 2 and 4, the accuracy of our technique will be increased substantially once we have constructed the stereo-pair of fundus cameras that we have designed. A number of technical improvements to our current work are foreshadowed. First, we expect to replace the finite difference scheme developed in Section 3 with a corresponding finite element scheme. Second, we propose to investigate the integrability constraint of Horn Brooks and Chellappa that estimates  $z(x, y)$  directly instead of the two-step scheme presented here in which  $p(x, y)$  is first estimated, then regularised, and finally integrated. Finally, we expect to develop further our work on integrating binocular stereo and photometric stereo estimates to allow for errors in the binocular stereo depth measurements.

## ACKNOWLEDGEMENTS

We thank Professor Tony Bron and Mr Nick Phelps-Brown of the Radcliffe Eye Infirmary for their continuing encouragement of our work and allowing us access to the Fundus camera. Lionel Tarassenko, Andrew Blake, and Andrew Zisserman have contributed useful comments and advice. Alan Simpson, Lynda Lindsell, and Paul Parker provided fundus images. Klaus Buchenrieder effected an introduction to representatives of the Rodenstock Instrument Corporation. SL thanks the Croucher Foundation for supporting his graduate study at Oxford. JMB thanks Sarah Harrington for heroic defence.

## References

- [1] Baker H. H. and Binford T. O. "Depth from edge and intensity based stereo" *Proc. 7th Int. Joint Conf. on A. I.* (1981) pp 631-616.
- [2] Bishop K. I. "Variability and reproducibility of the optic disc topographic measurements with the Rodenstock Optic Nerve Head Analyzer" *American Journal of Ophthalmol.* (1988) pp 696-702.
- [3] Bookstein F. R. "Principle warps: thin-plate splines and the decomposition of deformations" *IEEE transactions on pattern analysis and machine intelligence* Vol. 11 (1989) pp 567-584.
- [4] Courant R. and Hilbert D. *Methods of mathematical physics* Vol. 1, Interscience Publishers, New York (1953).
- [5] Forsyth D. and Zisserman A. "Mutual illumination" *CVPR* (1989).
- [6] Grimson W. E. L. *From images to surface - A computational study of the human early visual system* The MIT Press Series in Artificial Intelligence, The MIT Press, Massachusetts (1981).
- [7] Hanna K. J. *Monitoring Cataract Changes* D.Phil thesis, University of Oxford (1990).
- [8] Horn B. K. P. *Robot vision* The MIT Electrical and Computer Science Series, The MIT Press and McGraw-Hill Book Company (1986).
- [9] Littmann H. "Determination of the real size of an object on the fundus of the living eye" *Klin. Mbl. Augenheilk* (1982) pp 286-289.
- [10] Mikelberg F. S. and Douglas G. R. and Schulzer M. "The correlation between cup-disc ratio, neuroretinal rim area, and optic disc area measured by the video-phthalmograph (Rodenstock Analyzer) and clinical measurement" *American Journal of Ophthalmol* (1986) pp 7-12.

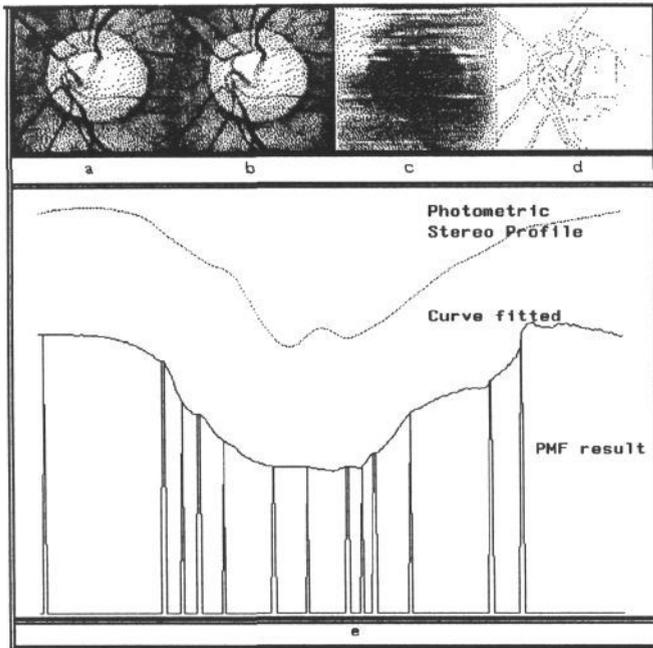


Figure 1: A stereo pair of the optic disk are shown in a) and b). c) is the surface reconstruction of the optic disk by the photometric stereo method. d) is the sparse depth measurement by PMF. e) shows cross-sections of c) and d) at approximately the center of the optic disk, and the integration of the two results is shown.

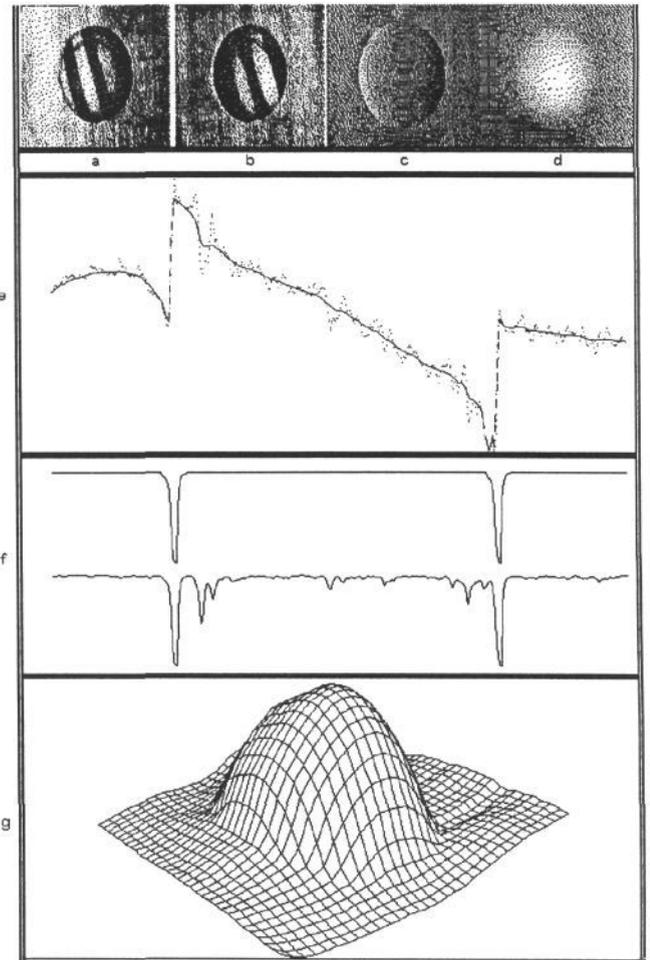


Figure 2: a) and b) show the photometric stereo pair of a half tennis ball sit on a piece of cardboard under illumination from the left and right respectively. The gradient map c) is computed from the gray level images by photometric stereo. After filtering by the regularisation technique, the surface of the scene d) can be reconstructed from the smoothed gradient map. e) shows cross-section of the gradient map before and after smoothing. Note the rapid changes in gradient at the edges of the tennis ball are preserved. f) shows the idealised model for  $\lambda$  (top) and its approximation (bottom) of the regularisation smoothing method. g) shows the surface of the scene reconstructed.

- [11] Noble J. A. *Descriptions of image surfaces* D.Phil thesis, University of Oxford (1989).
- [12] Pollard S. B. and Mayhew J. E. W. and Frisby J. P. "Disparity gradient, Lipschitz continuity, and computing binocular correspondences" *Perception* Vol. 14 (1985) pp 449-470.
- [13] Smith G. D. *Numerical solution of partial differential equations: Finite Difference Methods* Oxford applied mathematics and computing science series, Oxford University Press, Oxford, third edition (1985).
- [14] Terzopoulos D. "Regularisation of inverse visual problems involving discontinuities" *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 8 (1986) pp 413-424.
- [15] Woodham R. J. "Photometric method for determining surface orientation from multiple images" *Optical Engineering* (1980) pp 139-144.