



# Coarse view synthesis using shape-from-shading

Philip L. Worthington, Edwin R. Hancock \*

*Department of Computer Science, University of York, York YO10 5DD, UK*

Received 21 August 2001; accepted 11 January 2002

---

## Abstract

This paper investigates the use of shape-from-shading for coarse view synthesis. The aim of our study is to determine whether needle-maps delivered by a new shape-from-shading (SFS) algorithm can be used as a compact object-representation for the purposes of efficiently generating appearance manifolds. Specifically, we aim to show that the needle-maps can be used to generate novel object views under changing light source and viewer directions. To this end we conduct two sets of experiments. Firstly, we use the recovered needle-maps to re-illuminate objects under varying lighting directions. Here we show that a single input image can be used to construct relatively faithful re-illuminations under radical illumination changes. Secondly, we investigate how the needle-map can be used to generate new object poses. Here we show that needle-maps can be used for both view interpolation and view extrapolation. © 2002 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

*Keywords:* Shape-from-shading; View synthesis; Parametric eigenspaces

---

## 1. Introduction

Appearance-based object recognition has recently attracted considerable interest in the computer vision literature [1,2]. Although there are various realisations of the idea, the unifying principle is to compute a compact representation of the 2D appearance of 3D objects under multiple viewing and illumination conditions. This means that for each object a relatively large number of images must be collected of different object poses or viewing directions and different illumination conditions. For instance, Ullman has shown how different views may be combined for the purposes of recognising intermediate object poses [2]. The parametric eigenspace of Nayar and Murase [1] achieves a degree of data compression by storing the model views implicitly as a manifold in eigenspace. Here the idea is to represent the variability of object appearance under different viewing and lighting directions using the manifold to interpolate the leading components of eigenvectors extracted from the raw images.

One of the criticisms of appearance-based object recognition is the demands it places on data collection. Sufficient image data must be accumulated so that accurate object representations can be constructed. Images must be collected and stored so as to span a sufficient number and range of viewpoints so as to allow recognition from any novel viewpoint that is likely to be encountered. Here, the concepts of characteristic views [3–5] or aspect graphs [6] may prove useful for organising or condensing the amount of information required. In particular, if we treat the idea of a characteristic view (CV) as a natural grouping or clustering of similar views, then it should be possible to represent an object by storing a single representative from within each CV.

In any event, for each viewpoint we must also store, or find some way to model, all appearance changes due to light source variations which are likely to be encountered. It is easy to see that the storage and matching requirements for such a recognition scheme have the potential to grow rapidly beyond the bounds of practicality. Moreover, the view-collection process itself may also prove expensive, and requires that the lighting conditions be carefully controlled. This means that object appearance may only be learned under highly controlled conditions. As a result, autonomously

---

\* Corresponding author. Tel.: +44-1904-43-3374; fax: +44-1904-43-2767.

*E-mail address:* erh@cs.york.ac.uk (E.R. Hancock).

learning object appearance in an uncontrolled environment is difficult.

The observation underpinning this paper is that a more efficient strategy is to collect a small sample of object images. From this sample a set of images is generated so as to span the appearance space for a particular object. In other words, we replace data collection under controlled lighting conditions with view synthesis from a set of representative images. Since they are to be used for the purposes of constructing an appearance-based representation, the synthetic views need only be fairly coarse.

View synthesis has recently attracted considerable interests in the computer vision literature [7,8]. The topic has been the subject of intense activity in the graphics community for some time [9] and has led to the development of a variety of techniques for photo-realistic object rendering. However, the computer vision approach to the problem is somewhat different and revolves around automatically acquiring information concerning object geometry and appearance [9]. For instance, Poggio and Vetter, together with their co-workers have shown how to learn the appearance of faces for the purposes of synthesis [10,11]. This approach uses a linear model. There has recently been interest in the use of least-squares estimation techniques to improve the statistical robustness of the method [12]. Several authors including Sengupta and Ohya [13], and Avidan and Shashua [14] have used affine structure for the purposes of 3D object synthesis from 2D views. There has also been considerable effort aimed at showing how the geometry of different image tokens such as lines [9,15,16] and regions [17] can be exploited for view synthesis. Finally, several authors have used view synthesis as a means of object morphing [18,19].

However, here we require only relatively coarse synthetic views. We take a photometric approach to the problem using the output of shape-from-shading as the starting point. Some steps in this direction have recently been taken by Nayar and Murase [20], and by Georghades and Kriegman [22] who show how an illumination cone can be used for view synthesis after object reconstruction has been performed. The reconstruction process is based on photometric stereo and requires a minimum of three views with known lighting. The second capability is the synthesis of new or intermediate object poses from a representative set of model views. This second point is to some extent addressed by Ullman's view interpolation work [23]. However, this method operates primarily with pictorial descriptions of objects rather than intensity images.

### 1.1. Motivation

The observation underpinning this paper is that shape-from-shading, and by extension other shape-from-X modules such as shape-from-texture, provides an obvious yet hitherto unexplored route to coarse view synthesis. Shape-from-shading has long been a subject of active research within the vision community [24]. Its role has been to

deliver a dense map of local surface orientation information from shading patterns. Shape-from-shading aims to recover the required orientation information by solving the image irradiance equation. However, there are few reported attempts to use shape-from-shading for any practical object recognition or shape analysis tasks [25]. Whilst it appears clear that one of the motivations of early shape-from-shading research was to enable a 3D representation to be derived from a single image, the difficulties encountered in achieving accurate and robust needle-map recovery have proved a serious obstacle to progress in this direction.

Many of the difficulties encountered by existing shape-from-shading schemes can be attributed to the fact that they over-smooth the recovered needle-map. We have recently developed a new framework for shape from shading [26] which addresses this problem. It offers two advantages over existing schemes. Firstly, we can impose compliance with the image irradiance equation as a hard constraint. Secondly, we impose more sophisticated constraints on the local consistency of the recovered needle-map. The needle-maps delivered by the new shape-from-shading framework contain fine surface detail and can be used to identify topographic structure not recoverable with alternative algorithms [27,28]. Moreover, we have recently shown how surface topography information extracted from the needle-maps can be used for the purposes of 3D object recognition from 2D views [29].

Once a needle-map has been obtained, object re-illumination is a straightforward task. All that needs to be done is to modify the light source direction in the image irradiance equation and to compute the resulting image brightness using Lambert's law for matte surface reflectance. It is important to stress that Ullman [2,30] and others have proposed the use of view-interpolation or view-combination techniques for appearance-based object recognition. However, much of the work in this area has focused upon interpolating pictorial descriptions of objects. As Ullman notes [23], interpolation between smooth objects presents greater difficulties, and shape-from-shading represents one possible method of addressing them. However, this has remained only a suggestion with no concrete substantiation in the literature. Here we present concrete experimental results which point to the practical feasibility of this approach. Specifically, we investigate both extrapolation using the needle-map obtained from a single image, and interpolation between two or more views with different object poses. The novel contribution in the current paper is therefore to investigate how our improved needle-maps can be used for object re-illumination and view synthesis.

The outline of the remainder of this paper is as follows. In Section 2 we review the new shape-from-shading algorithm. Section 3 we describe the re-illumination process, and in Section 4 our approach to novel view synthesis. Section 5 shows the experimental results of these approaches, and in Section 6 we draw conclusions and consider the outlook for developing these ideas.

## 2. Data-driven shape-from-shading

Our new shape-from-shading algorithm has been demonstrated to deliver needle-maps which preserve fine surface detail [26,31]. The observation underpinning the method is that for Lambertian reflectance from a matte surface, the image irradiance equation defines a cone of possible surface normal directions. The axis of this cone points in the light-source direction and the opening angle is determined by the measured brightness. If the recovered needle-map is to satisfy the image irradiance equation as a hard constraint, then the surface normals must each fall on their respective reflectance cones. Initially, the surface normals are positioned so that their projections onto the image plane point in the direction of the image gradient. Subsequently there is iterative adjustment of the surface normal directions so as to improve the consistency of the needle-map. In other words, each surface normal is free to rotate about its reflectance cone in such a way as to improve its consistency with its neighbours. This rotation is a two-step process. First, we apply a smoothing process to the current surface normal estimates. This may be done in a number of ways. The simplest is local averaging. More sophisticated alternatives include robust smoothing with outlier reject and, smoothing with curvature or image gradient consistency constraints. This results in an off-cone direction for the surface normal. The hard data-closeness constraint of the image irradiance equation is restored by projecting the smoothed off-cone surface normal back onto the nearest position on the reflectance cone.

To be more formal let  $\mathbf{s}$  be a unit vector in the light source direction and let  $E_{i,j}$  be the brightness at the image location  $(i, j)$ . Further, suppose that  $\mathbf{n}_{i,j}^k$  is the corresponding estimate of the surface normal at iteration  $k$  of the algorithm. The image irradiance equation is

$$E_{i,j} = \mathbf{n}_{i,j}^k \cdot \mathbf{s}. \quad (1)$$

As a result, the reflectance cone has opening angle  $\cos^{-1} E_{i,j}$ . After local smoothing, the off-cone surface normal is  $\bar{\mathbf{n}}_{i,j}^k$ . The updated on-cone surface normal which satisfies the image irradiance equation as a hard constraint is obtained via the rotation

$$\mathbf{n}_{i,j}^{k+1} = \Phi \bar{\mathbf{n}}_{i,j}^k. \quad (2)$$

The matrix  $\Phi$  rotates the smoothed off-cone surface normal estimate by the angle difference between the apex angle of the cone, and the angle subtended between the off-cone normal and the light source direction. This angle is equal to

$$\theta = \cos^{-1} E - \cos^{-1} \frac{\bar{\mathbf{n}}_{i,j}^k \cdot \mathbf{s}}{\|\bar{\mathbf{n}}_{i,j}^k\| \cdot \|\mathbf{s}\|}. \quad (3)$$

This rotation takes place about the axis whose direction is given by the vector

$$(u, v, w)^T = \bar{\mathbf{n}}_{i,j}^k \times \mathbf{s}. \quad (4)$$

This rotation axis is perpendicular to both the light source direction and the off-cone normal. Hence, the rotation matrix is

$$\Phi = \begin{pmatrix} c + u^2 c' & -ws + uv c' & vs + uwc' \\ ws + uv c' & c + v^2 c' & -us + vwc' \\ -vs + uwc' & us + vwc' & c + w^2 c' \end{pmatrix},$$

where  $c = \cos(\theta)$ ,  $c' = 1 - c$ , and  $s = \sin(\theta)$ .

The off-cone surface normal is recovered through a process of robust-smoothing. The smoothness error or consistency of the field of surface normals is measured using the derivatives of the needle-map in the  $x$  and  $y$  directions by the penalty function

$$I = \iint \left\{ \left( \rho_\sigma \left( \left\| \frac{\partial \mathbf{n}}{\partial x} \right\| \right) + \rho_\sigma \left( \left\| \frac{\partial \mathbf{n}}{\partial y} \right\| \right) \right) \right\} dx dy. \quad (5)$$

In the above measure,  $\rho_\sigma(\eta)$  is the robust error kernel used to gauge the local consistency of the needle-map or field of surface normals. The argument of the kernel  $\eta$  is the measured error and the parameter  $\sigma$  controls the width of the kernel. It is important to note the robust-error kernels are applied separately to the magnitudes of the derivatives of the needle-map in the  $x$  and  $y$  directions. Applying variational calculus, the penalty function is minimized by updating the surface normals using the following fixed-point iteration equation

$$\begin{aligned} \bar{\mathbf{n}}_{i,j}^{(k+1)} = & \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial x} \right\|^{-1} \left[ \frac{\partial}{\partial x} \left( \rho'_\sigma \left( \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial x} \right\| \right) \right) \right. \\ & + \rho'_\sigma \left( \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial x} \right\| \right) \left( \mathbf{n}_{i+1,j}^{(k)} + \mathbf{n}_{i-1,j}^{(k)} - \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial x} \right\|^{-2} \right. \\ & \left. \left. \times \left( \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial x} \cdot \frac{\partial^2 \mathbf{n}_{i,j}^{(k)}}{\partial x^2} \right) \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial x} \right) \right] \\ & + \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial y} \right\|^{-1} \left[ \frac{\partial}{\partial y} \left( \rho'_\sigma \left( \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial y} \right\| \right) \right) + \rho'_\sigma \left( \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial y} \right\| \right) \right. \\ & \left. \times \left( \mathbf{n}_{i,j+1}^{(k)} + \mathbf{n}_{i,j-1}^{(k)} - \left\| \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial y} \right\|^{-2} \right. \right. \\ & \left. \left. \times \left( \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial y} \cdot \frac{\partial^2 \mathbf{n}_{i,j}^{(k)}}{\partial y^2} \right) \frac{\partial \mathbf{n}_{i,j}^{(k)}}{\partial y} \right) \right]. \quad (6) \end{aligned}$$

Stated in this way the smoothing process is entirely general. However, in our previous work we found that the most effective error kernel was the log-cosh sigmoidal-derivative M-estimator

$$\rho_\sigma(\eta) = \frac{\sigma}{\pi} \log \cosh \left( \frac{\pi \eta}{\sigma} \right). \quad (7)$$

### 3. Object re-illumination using shape-from-shading

The needle map returned by shape-form-shading may be re-illuminated by a new light source from any chosen direction. This provides an estimate of the appearance of the object under the new lighting conditions. The input information required is a single image and a single application of the shape-from-shading algorithm. Suppose that  $\mathbf{n}_{i,j}^{(final)}$  is the recovered surface normal at the pixel indexed  $(i, j)$  when the shape-from-shading scheme has reached convergence. Further suppose that the new light source direction used for re-illumination points from the direction specified by the unit vector  $\mathbf{s}_{synth}$ . With these ingredients the synthesised image intensity is given by

$$E_{synth}(i, j) = \mathbf{n}_{i,j}^{(final)} \cdot \mathbf{s}_{synth}$$

Of course, re-illuminating in this fashion ignores the possibility of self-occlusion. However, occlusion effects may be taken into account by integrating the needle-map and performing hidden-surface calculations. This adds considerable computational expense, and takes the scheme into the realm of model-based recognition. Here, we consider the simple case only, and assume that self-occlusion effects are relatively small. This is the case for objects that are predominantly convex.

A sequence of images may be generated by re-illumination using different light sources. The idea underpinning this paper is that the set of images generated in this way can be used in an appearance-based object recognition scheme. This process of view synthesis replaces that of gathering many images under varying and controlled lighting conditions. Indeed, it may be possible to calculate sufficiently good estimates of the appearance under all illuminations from a single image. This should be contrasted with the need for three or more images to calculate the illumination cone [22].

Here, we evaluate the accuracy of the re-illumination process within the framework provided by the parametric eigenspace approach of Murase and Nayar [1]. Specifically, we construct an eigenspace using illuminated range images from the USC database of busts of famous composers.

Given that this set of images is relatively small and homogeneous, the resulting eigenspace may be expected to produce excellent discrimination between appearances. Hence, comparing the eigenspace projections of the re-illuminated needle-maps within this eigenspace represents a tough test, since even small differences in appearance may result in large distances in eigenspace.

### 4. Novel view generation using shape-from-shading

The generation of novel views using shape-from-shading information is significantly more complex than re-illumination, and here involves a hybrid approach between appearance-based and model-based recognition. We con-

sider both extrapolation from a single view, and interpolation using two views. However, both operate in much the same manner. To perform these two tasks we require height data. We extract height data using the method of Wu and Li [32]. This is a relatively simple method in which the needle-map is integrated to produce height data by simply summing the needle-map components in the  $x$  and  $y$  directions to generate two surfaces, which are then averaged to reduce noise artifacts. This rough surface approximation may then be rotated by the desired angle about an arbitrary axis. However, the boundary of the surface, where regions may appear or disappear from view, requires special treatment. In this paper, we simply reflect the surface across the image plane to produce a closed object prior to rotation. An alternative approach would be to extrapolate the surface where it meets the image plane, using either the tangent to the surface at the boundary points, or a spline approximation. Such approaches may well yield improved results. This follows intuitively from the assumption that most real-world surfaces are piecewise smooth and that most objects are predominantly convex. Hence, the probability that the current occluding boundary is coincident with a surface discontinuity is assumed small relative to the probability that it continues smoothly, and the surface extrapolated accordingly.

Irrespective of the way in which the surface boundary is treated, after the rotation we wish to obtain a new needle-map. One option is to calculate this from the rotated surface, but this will introduce further numerical errors and results in an essentially model-based approach.

An alternative is to use an approach similar to the curvature method of Ullman [23]. According to this approach, several views of an object are used to establish the local radius of curvature at each point. This information is subsequently used to estimate the change in the position of each point under rotation.

It is certainly possible to estimate local curvature from the needle-map derived from a single image. However, rather than explicitly calculate the local curvature vectors, we take the following computational approach. The rotation of the object,  $\theta$ , is split into three components:  $\delta\theta_1, \theta', \delta\theta_2$ , where  $\theta = \delta\theta_1 + \theta' + \delta\theta_2$  and  $\{\delta\theta_1, \delta\theta_2\} \ll \theta'$ . The small rotation,  $\delta\theta_1$ , is used to estimate the initial direction of motion,  $\{\mathbf{v}_1\}$ , of each initial point,  $\{P_1\}$ . Similarly, the final small rotation,  $\delta\theta_2$ , is used to calculate the final trajectories,  $\{\mathbf{v}_2\}$ , as each point  $\{P_1\}$  maps to a new location,  $\{P_2\}$ . The transformation which maps the unit vector  $\mathbf{v}_1$  to  $\mathbf{v}_2$  is then applied to the surface normal at  $P_1$  to yield the estimated surface normal at the corresponding point  $P_2$ .

Essentially, each initial vector  $\{\mathbf{v}_1\}$  is taken to be an estimate of the tangent to the vector connecting the surface point  $\{P_1\}$  to the centre of rotation of the object, and similarly for  $\{\mathbf{v}_2\}$  at  $\{P_2\}$ . Therefore, we assume that the surface normals remain rigid with respect to these vectors, and apply the transformation  $\{\mathbf{v}_1\} \rightarrow \{\mathbf{v}_2\}$  to obtain a new needle-map for the rotated object.

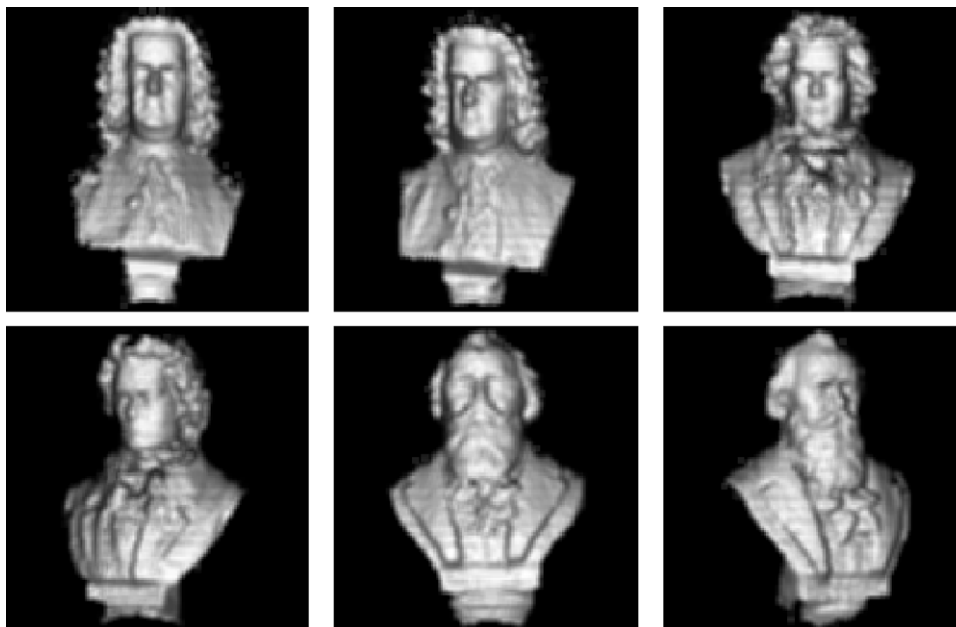


Fig. 1. A subset of the test images used in this paper. Range images of the busts of composers are illuminated using the Lambertian model, in this case by a light source in the direction  $\mathbf{s}=(0, 0, 1)^T$ . They are subsequently blurred using a  $3 \times 3$  Gaussian to smooth the surface noise.

Using this approach, we can extrapolate novel views from a single image using shape-from-shading information. However, this involves making strong assumptions about what happens to surface points that come into view due to the rotation of the object. Interpolation between two views reduces this problem. However, to achieve proper interpolation, some form of correspondence must be established between features in the two images. This is difficult to achieve reliably and automatically between greyscale images of smooth objects. In the early experiments reported here, we have taken a simpler and less accurate approach. Our crude interpolation is achieved by extrapolating the two model images towards the mid-view. The two needle-maps calculated for the interpolated view are then compared on a needle-by-needle basis, and in each case the surface normal which subtends the smallest angle to the light source direction is selected. This simple, non-correspondence-based interpolation method, is implicitly biased towards generating bright objects. Nonetheless, our goal here is coarse view synthesis and as we shall demonstrate Section 5.2 the resulting images provide useful appearance manifolds for the purposes of recognition.

## 5. Experiments

We have investigated using shape-from-shading information for both re-illumination and novel-view generation, using the USC range image database of busts of famous composers. Fig. 1 shows some of the images used in this

study. Each bust has several range images which have been collected in different viewing directions. For each range image we have generated illuminated intensity images using the Lambertian lighting model outlined in Section 3.

Our experiments have focused upon using the images generated from shape-from-shading to approximate the manifolds of objects in the parametric eigenspace [1]. The eigenspace approach is an elegant and accurate approach to object recognition, but requires large numbers of model images to construct a manifold which describes the appearance of an object under a wide range of viewing conditions. Any reduction in this number which does not adversely impact upon recognition accuracy is clearly important in reducing the time and expense of collecting model data.

### 5.1. Re-illumination

In the first instance, we consider the possibility of using shape-from-shading information to significantly reduce the number of model images which must be obtained in order to capture the appearance of an object under variable lighting conditions. The question of modelling illumination variability is under-represented in the literature, and yet is of considerable importance. Specifically, in the real world it is usually easier to control viewpoint than illumination, so in many cases it is not only time-consuming, but impractical to gather the data needed to represent an object under all likely illumination conditions.

Fig. 2 shows the results of re-illuminating the needle-map of two busts obtained by applying the shape-from-shading



Fig. 2. Re-illumination results. The two sequences of images compare the results of re-illuminating the two busts using the surface normals extracted directly from the range image and those extracted using shape-from-shading. In each sequence the single leftmost image is the frontal image of the bust, illuminated by a light source in direction  $\mathbf{s} = (0, 0, 1)^T$ . This image is used as input to our shape-from-shading scheme. The recovered needle-map is then re-illuminated. In each sequence the top row shows the illumination of the range image for light-source directions of (left-to-right)  $0^\circ, 20^\circ, 40^\circ, 60^\circ, 80^\circ$  to vertical in the  $x$  direction. The lower row shows the corresponding re-illumination results obtained using shape-from-shading.

scheme of Section 2. In each of the two sequences the images in the different rows compare the range image illuminations and re-illuminations obtained using shape-from-shading as the light source swings from left to right across the bust. In each case the upper image is the ground-truth illumination obtained from the range data while the lower image is the re-illumination generated from the needle-map. For each of the re-illuminations the needle-map is computed from an image in which the light-source direction is  $\mathbf{s} = (0, 0, 1)^T$ . These two images are shown individually to the left of the figure. The re-illuminations preserve most of the gross image structure even when the light source is moved through a large angle from the image plane normal. The range images contain a lot of fine surface detail. As the re-illumination angle is increased then this is lost. This is largely attributable to our relatively crude image reconstruction method, which does not allow for hidden surface removal when the light source direction grazes the surface at small angles. However, the regions of light and shadow roughly agree throughout, and much of the large-scale image structure is preserved.

In our next set of experiments we study the parametric eigenspace for the Bach bust under varying illumination

direction [1,20,21]. The data shown here is produced using the Columbia University SLAM system. For each image the leading three eigenvalues are computed. The points obtained under successive illumination conditions are interpolated by a trajectory. Fig. 3 shows the trajectories for the range-data illuminations and the shape-from-shading re-illuminations as the lighting direction is varied. The light source direction swings from  $\mathbf{s} = (0, 0, 1)^T$  (i.e. the image plane normal to  $\mathbf{s} = (1, 0, 0)^T$ ) as the points move from the bottom left-hand corner of the eigenspace to the top right-hand corner. The divergence between the trajectories is not significant until the light source direction has moved by  $60^\circ$ . This would suggest that re-illumination can result in a significant reduction in the amount of training data required.

Fig. 4 offers a quantitative analysis of the eigenspace plots of Fig. 3. The top plot is produced by considering the first 3 eigenspace dimensions only, whilst the bottom plot uses 20 dimensions. In each case, the two solid lines show the distances between the true image and the corresponding re-illuminated image. As can be seen from Fig. 3, this distance increases as the illumination direction changes from  $0^\circ$  to  $90^\circ$ .

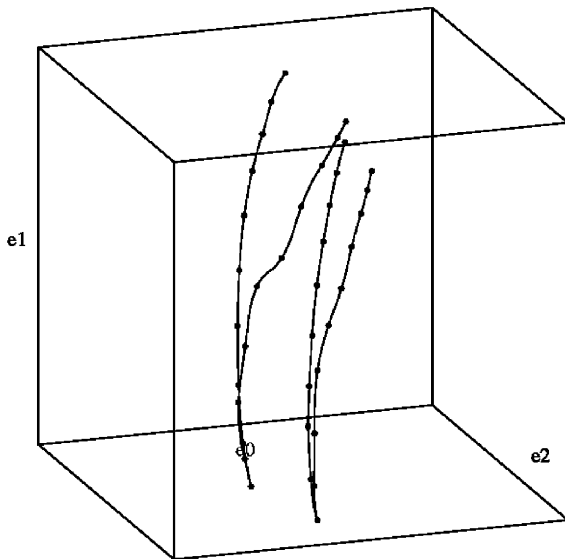


Fig. 3. The two views of the Beethoven bust are treated as in Fig. 2. The trajectories of the true images and the re-illuminated images are plotted for light source directions  $0-90^\circ$  to vertical in the  $x$  direction. The smoother curve of each pair corresponds to the true images, whilst the re-illumination curve diverges from this as the light source is rotated away from the vertical.

The dot-dashed lines are the distances between successive images for each of the true illumination trajectories. Hence, we see in the top plot that the error introduced by the re-illumination process is less than the distance between two images with  $10^\circ$  change in illumination for re-illumination angles upto around  $50-60^\circ$ . For comparison, we also include the distance between the two viewpoints for  $ss = (0, 0, 1)^T$  illumination (the horizontal dashed line). The views are approximately  $20^\circ$  apart, resulting in a distance greater than that due to small changes in illumination. In the bottom plot, using 20 dimensional eigenvectors, the re-illumination error is greater than the distance between two adjacent illumination images. This reflects the fact that the re-illuminated images, whilst retaining the gross detail of the object, differ considerably in small scale detail due to the smoothing effect of the shape-from-shading process. However, the re-illumination error remains within an order of magnitude of the difference between successive illumination images, and of a similar value to the distance between the two views.

In Fig. 5 we show the more complex, 3D manifolds which can be generated when the viewing direction in addition to the light source direction. These manifolds are obtained using three different views of the Bach bust when the light source direction is varied as described above. The inner manifold corresponds to the re-illumination of the 3 initial views for the light source incidence angles in the range  $0-90^\circ$ . For re-illumination angles of around  $0-60^\circ$ , the correlation between the manifolds is fairly good, although the true manifold has greater extent.

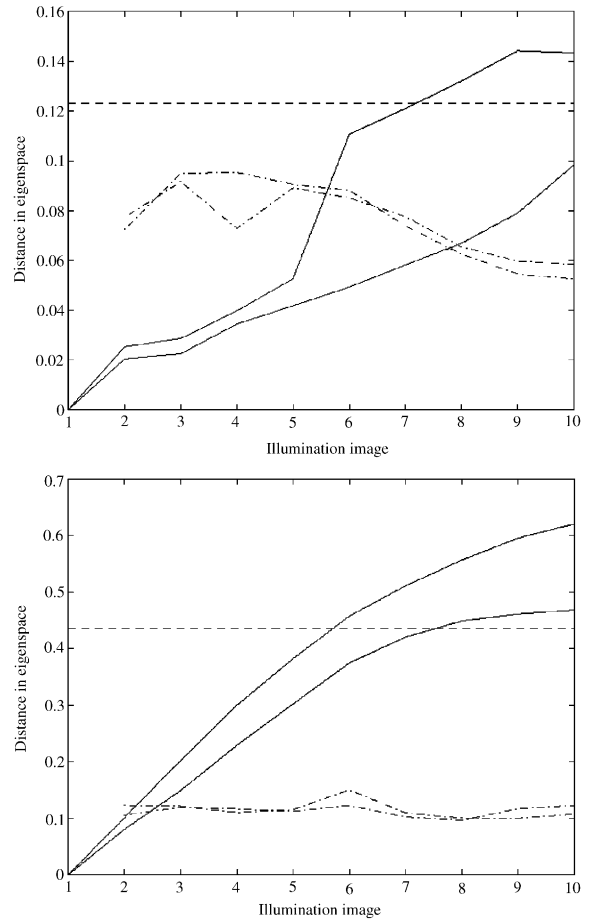


Fig. 4. A simple quantitative analysis of the illumination and re-illumination trajectories of Fig. 3. The top plot is produced by considering the first 3 eigenspace dimensions only, whilst the bottom plot uses 20 dimensions. In each case, the two solid lines plot the distances between the true image and the re-illuminated image in eigenspace for illumination images 1–10 corresponding to illumination directions of  $0-90^\circ$ . The dot-dashed lines are the distances between successive images for each of the true illumination trajectories, whilst the dashed line is the distance between the two viewpoints for  $s = (0, 0, 1)^T$  illumination.

### 5.2. Novel view generation

Generating novel views from shape-from-shading data may be accomplished either by extrapolating from a single view or by interpolating between two or more views. In the first instance, we consider the extrapolation scheme described in Section 4. The procedure adapted here is to extract surface normals from a single example image. The needle-maps are rotated by the desired amount and re-illuminated to provide a synthetic image of the extrapolated pose. Fig. 6 shows illumination trajectories in the parametric eigenspace. Here we use two different views of

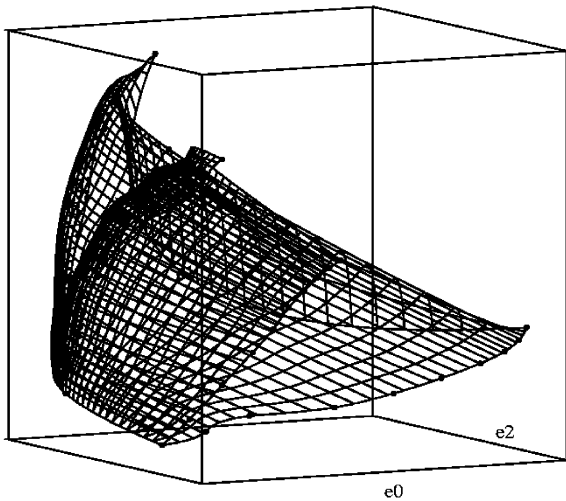


Fig. 5. There are 3 range images of the Bach bust, allowing the generating of a 3D manifold in eigenspace. The 3 views are each illuminated over the range  $0-90^\circ$  to produce the manifold. The smaller manifold enclosed by the larger one is the result of re-illumination from shape-from-shading data.

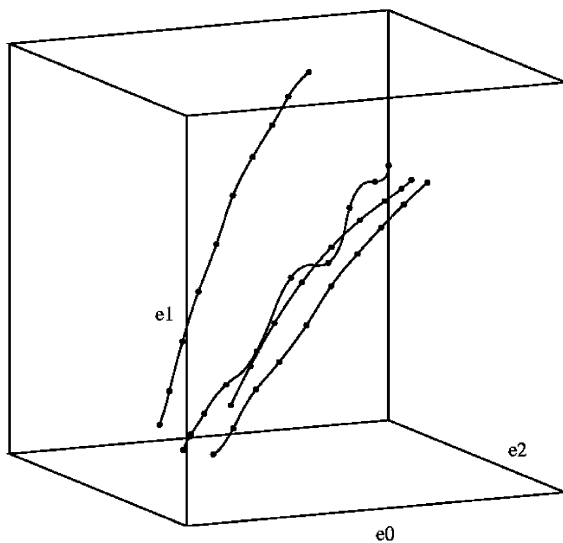


Fig. 6. Taking the 2 views of the Beethoven bust, we apply our shape-from-shading view-extrapolation scheme to generate two intermediate needle-maps, one from each image. Using each of these, we repeat the re-illumination process over the light source direction range  $0-90^\circ$ . The two outer curves represent the true illumination trajectories for the two busts, whilst the middle curves are the re-illumination trajectories from the extrapolated needle-maps.

the Beethoven bust in which the viewing direction changes by  $30^\circ$ . The range images of the two views are shown in the third and fourth panels of Fig. 1. In Fig. 6 the outermost trajectories are obtained by illuminating the range data for the two different views. The two intermediate trajectories

are obtained by extrapolating the two range views by rotation of  $15^\circ$  to obtain an intermediate pose. The extrapolated poses should therefore be co-incident. There are two features to note from this data. First, the two extrapolated illumination trajectories are overlapping. The more erratic of the two is produced by extrapolating the more rotated range image shown in the fourth panel of Fig. 1. The second feature to note is that the extrapolated trajectories appear to approximately interpolate the outer two trajectories.

Finally, we experiment with the simple view interpolation method described in Section 4. Here, we interpolate the surface normals extracted from a pair of illuminated range images in different poses. The interpolated surface normals are then used to produce a synthetic image of the intermediate object pose. Fig. 7 shows the re-illuminated images generated from the interpolated needle-map, and compares them with the illuminated range images. The left-most and right-most images in the top row of the figure show the two illuminated range images used as input to the shape-from-shading algorithm. The centre image is the result of re-illuminating the interpolated needle-map. In each of these three images the light source direction is perpendicular to the image plane. In the subsequent rows of the figure we investigate the effect of varying the light source direction. In each row the left-most and right-most images are the result of illuminating the range images with different light source directions. The centre-image in each row is the result of re-illuminating the interpolated needle-map obtained using the data shown in the top row. In other words, the needle-map interpolation process is applied only once. Given the simplicity of the interpolation scheme employed, these results appear extremely promising. In Fig. 8 we plot the re-illumination trajectory of the interpolated needle-map in comparison to the true illumination trajectories of the two views. The re-illumination trajectory appears to provide a plausible interpolation. In other words, although coarse the synthetic images preserve enough detail to interpolate the appearance manifold.

## 6. Conclusions and outlook

Making the most of the available data is an important issue in appearance-based object recognition. It seems simply impractical to collect and store vast quantities of model views to represent an object under all possible, or even all likely, viewing conditions. In this paper we have begun to demonstrate the potential of shape-from-shading, and by extension other Shape-from-X modules, to increase of the utility of a given image for object recognition. In the case of re-illumination of objects, it seems clear that we can obtain, quickly and easily, a fairly good approximation of how a given object will look under radically different illumination from that under which the original image was obtained. More experimentation is needed to look at ways of incorporating self-shadowing effects into the re-illumination process.

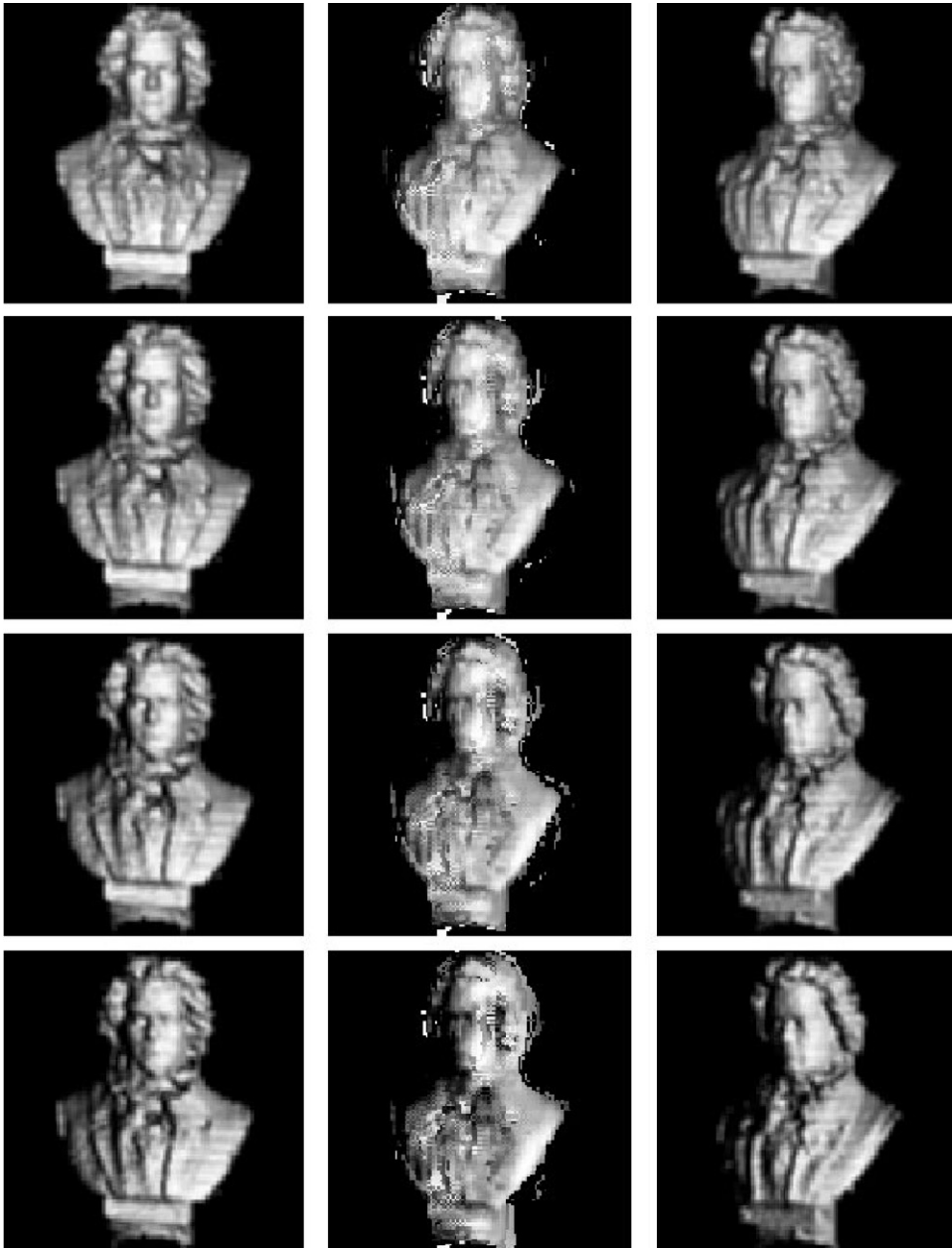


Fig. 7. The two views of the Beethoven bust are shown illuminated over the range of light source directions  $0\text{--}30^\circ$  (left- and right-hand columns). The middle column shows the results of view-interpolation using shape-from-shading, followed by re-illumination over the same range. Although this early attempt at view-interpolation introduces many distorting artifacts, particularly at the boundary of the object, considerable surface detail is retained towards the centre of the object. This detail is relatively stable and accurate under the re-illumination process.

The results presented show that view-extrapolation and view-interpolation can produce coarse synthetic views that are sufficiently good for constructing appearance-based object representations for the purposes of recognition. This is underlined by the well-behaved trajectories obtained by re-illuminating the interpolated needle-maps.

The view-interpolation results may be significantly improved through establishing correspondence between the two model views to be interpolated. Moreover, an interesting prospect is to adapt the linear combination of views method of Ullman [23] to the combination of needle-maps, thus obviating the need for constructing a 3D model at any point.

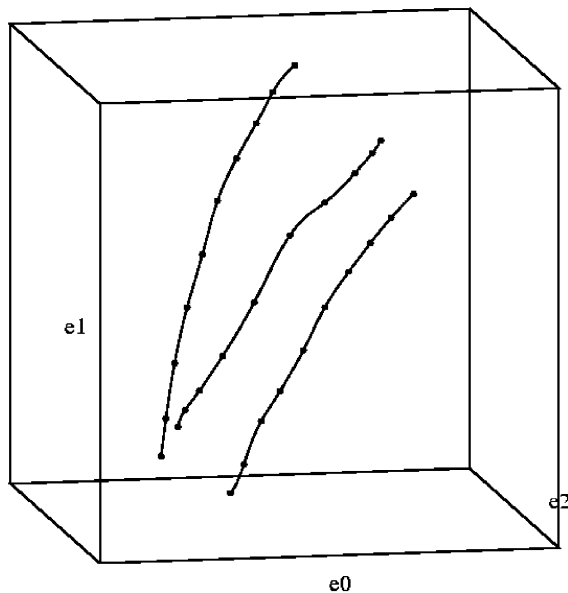


Fig. 8. Taking the 2 views of the Beethoven bust, we apply our shape-from-shading view-interpolation scheme to generate an intermediate needle-map. From this, we repeat the re-illumination process over the light source direction range  $0$ – $90^\circ$ . The two outer curves represent the true illumination trajectories for the two busts, whilst the middle curve is the re-illumination trajectory from the interpolated needle-map.

## References

- [1] H. Murase, S.K. Nayar, Learning object models from appearance, *Proc. AAAI Vol. I* (1993) 836–843.
- [2] S. Ullman, R. Basri, Recognition by linear combinations of models, *IEEE Pattern Anal. Machine Intell.* 13 (10) (1991) 992–1006.
- [3] I. Chakravarty, H. Freeman, Characteristic views as a basis for three-dimensional object recognition, *Proc. SPIE Robot Vision 336* (1982) 37–45.
- [4] R. Wang, H. Freeman, Object recognition based on characteristic view classes, *Proceedings of the International Conference on Pattern Recognition, Vol. I, 1990*, pp. 8–12.
- [5] S. Chen, H. Freeman, Characteristic view modelling of curved-surface solids, *Int. J. Pattern Recognition Artif. Intell.* 10 (1996) 537–560.
- [6] K.W. Bowyer, C.R. Dyer, Aspect graphs: an introduction and survey of recent results, *Int. J. Imaging Systems Technol.* 2 (1990) 315–328.
- [7] M.E. Izquierdo, S. Kruse, Image analysis for 3D modelling, rendering, and virtual view generation, *Journal of Computer Vision and Image Understanding* 71 (2) (1998) 231–253.
- [8] D.P. Greenberg, A framework for realistic image synthesis, *Commun. ACM* 42 (8) (1999) 44–53.
- [9] B. Johansson, View synthesis and 3D reconstruction of piecewise planar scenes using intersection lines between the planes, *Proc. ICCV Vol. I* (1999) 54–59.
- [10] T. Vetter, T. Poggio, Linear object classes and image synthesis from a single example image, *Pattern Anal. Machine Intell.* 19 (7) (1997) 733–742.
- [11] T. Vetter, Synthesis of novel views from a single face image, *Int. J. Comput. Vision* 28 (2) (1998) 103–116.
- [12] D.M. Kennedy, B.F. Buxton, J.H. Gilby, Application of the total least squares procedure to linear view interpolation, *Proc. British Machine Vision Conference, Vol. I* (1999) 305–314.
- [13] K. Sengupta, J. Ohya, Novel scene generation, merging and stitching views using the 2D affine space, *Signal Process: Image Commun.* 14 (1–2) (1998) 39–53.
- [14] S. Avidan, A. Shashua, Novel view synthesis by cascading trilinear tensors, *IEEE Trans. Visualisation Comput. Graphics* 4 (3) (1998) 293–306.
- [15] T.F. Stahovich, R. Davis, H. Shrobe, Generating multiple new designs from a sketch, *Artif. Intell.* 104 (1–2) (1998) 211–264.
- [16] M. Lhuillier, L. Quan, Image interpolation by joint view triangulation, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition II* (1999) 139–145.
- [17] J.S. McVeigh, M.W. Siegel, A.G. Jordan, Intermediate view synthesis considering occluded and ambiguously referenced image regions, *Signal Process: Image Commun.* 9 (1) (1996) 21–28.
- [18] S.M. Seitz, C.R. Dyer, View morphing, *SIGGRAPH* 30 (1996) 21–30.
- [19] R.A. Manning, C.R. Dyer, Interpolating view and scene motion by dynamic view morphing, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition I* (1999) 388–394.
- [20] S.K. Nayar, H. Murase, Dimensionality of illumination manifolds in appearance matching, *ECCV International Workshop on Object Representation in Computer Vision, 1996*, pp. 165–178.
- [21] S.K. Nayar, H. Murase, S.A. Nene, Parametric appearance representation, in: Nayar and Poggio (Eds.) *Early Visual Learning*, Oxford University Press, Oxford, 1996.
- [22] A.S. Georgiades, D.J. Kriegman, P.N. Belhumeur, Illumination cones for recognition under variable illumination: faces, *Proc. IEEE Conf. on Computer Vision and Pattern Recognition I* (1998) 52–58.
- [23] S. Ullman, *High-level Vision*, MIT Press, Cambridge, MA, 1996.
- [24] B.K.P. Horn, Obtaining shape from shading information, in: P.H. Winston (Ed.), *The Psychology of Computer Vision*, McGraw-Hill, New York, 1975, pp. 115–155.
- [25] P.L. Worthington, B. Huet, E.R. Hancock, Appearance based object recognition using shape-from-shading, *Proc. ICPR I* (1998) 412–416.
- [26] P.L. Worthington, E.R. Hancock, New constraints on data-closeness and curvature consistency for shape-from-shading, *IEEE Pattern Anal. Machine Intell.* 21 (1999) 1250–1267.
- [27] M. Bichsel, A.P. Pentland, A simple algorithm for shape from shading, *Proc. CVPR I* (1992) 459–465.
- [28] P.S. Tsai, M. Shah, Shape from shading using linear approximation, *Image Vision Comput.* 12 (8) (1994) 487–498.
- [29] P.L. Worthington, E.R. Hancock, Object recognition using shape-from-shading, *IEEE Pattern Anal. Machine Intell.* 23 (2001) 535–542.
- [30] S. Ullman, Aligning pictorial descriptions: an approach to object recognition *Cognition* 32 (3) (1989) 193–254.
- [31] P.L. Worthington, E.R. Hancock, 3D surface topography from intensity images, *Pattern Recognition* 34 (2001) 823–840.
- [32] Z. Wu, L. Li, A line integration based method for depth recovery from surface normals, *CVGIP* 43 (1988) 53–66.

**About the Author**—PHILIP WORTHINGTON attained BA (Hons) in Engineering and Computer Science from University College, Oxford, in 1996. From 1996–1999 he pursued an EPSRC-funded DPhil in the Computer Vision group at the University of York under the supervision of Professor Edwin Hancock. In 1998, he spent 2 months in Dr Hiroshi Murase’s research group at NTT Basic Research Laboratories in Japan, courtesy of JISTEC and the British Council. In 1999 he was awarded the University of York Gibbs-Plessey award to visit academic and commercial vision groups around the US. Philip was awarded his DPhil in 2000 for his thesis on shape-from-shading for view-based object recognition. Following a spell in industry, he was appointed as a Lecturer in Computer Vision in the Department of Computation, UMIST, in March 2001. He has published several papers on shape-from-shading and view-based object recognition in a range of journals and refereed conference proceedings.

**About the Author**—EDWIN HANCOCK studied Physics as an undergraduate at the University of Durham and graduated with honours in 1977. He remained at Durham to complete a Ph.D. in the area of High Energy Physics in 1981. Following this he worked for ten years as a researcher in the fields of high-energy nuclear physics and pattern recognition at the Rutherford–Appleton Laboratory (now the Central Research Laboratory of the Research Councils). During this period he also held adjunct teaching posts at the University of Surrey and the Open University. In 1991 he moved to the University of York as a lecturer in the Department of Computer Science. He was promoted to Senior Lecturer in 1997 and to Reader in 1998. In 1998 he was appointed to a Chair in Computer Vision.

Professor Hancock now leads a group of some 15 faculty, research staff and Ph.D. students working in the areas of computer vision and pattern recognition. His main research interests are in the use of optimisation and probabilistic methods for high and intermediate level vision. He is also interested in the methodology of structural and statistical pattern recognition. He is currently working on graph-matching, shape-from-X, image data-bases and statistical learning theory. His work has found applications in areas such as radar terrain analysis, seismic section analysis, remote sensing and medical imaging. Professor Hancock has published some 60 journal papers and 200 refereed conference publications. He was awarded the Pattern Recognition Society medal in 1991 for the best paper to be published in the journal *Pattern Recognition*. The journal also awarded him an outstanding paper award in 1997.

Professor Hancock has been a member of the Editorial Boards of the journals *IEEE Transactions on Pattern Analysis and Machine Intelligence*, and, *Pattern Recognition*. He has also been a guest editor for special editions of the journals *Image and Vision Computing and Pattern Recognition*, and he is currently a guest editor of a special edition of *IEEE Transactions on Pattern Analysis and Machine Intelligence* devoted to energy minimization methods in computer vision. He has been on the programme committees for numerous national and international meetings. In 1997 he established a new series of international meetings on energy minimization methods in computer vision and pattern recognition. He was awarded a Fellowship of the International Association for Pattern Recognition in 2000.