

## POLYNOMIAL TEXTURE MAPPING AND 3D REPRESENTATIONS

Lindsay MacDonald<sup>1</sup> and Stuart Robson<sup>2</sup>

<sup>1</sup>Faculty of Media, London College of Communication, Elephant & Castle, London SE1 6SB

<sup>2</sup>Department of Civil, Environmental and Geomatic Engineering, UCL, Gower Street, London WC1E 6BT  
l.macdonald@lcc.arts.ac.uk

Commission V, WG V/1

**KEY WORDS:** Surface normals, artefacts, photometric stereo, polynomial texture mapping, 3D laser scanner.

### ABSTRACT

The results of a multi-image photographic technique are compared with those obtained by close-range scanning of the same object by a 3D colour laser scanner. The 3D representation achieved by the polynomial texture map (PTM) is compared with the extraction of normals from the same image set by photometric stereo. It is shown that the median of each of the X,Y,Z distributions of normal components for a large number of 3-light combinations is a good estimator of the true normal at each point. Using an incised Egyptian stone tablet, normals derived from these two techniques are compared with those generated by a 3D laser scanner. It is shown that the photometric stereo method produces the highest quality rendering of the surface.

### 1. SHAPE FROM SHADING

A normal  $\mathbf{N}$  to a surface  $S$  at point  $P$  is a vector perpendicular to the tangent plane touching the surface at  $P$ . For a plane given by the equation  $ax + by + cz = d$ , the vector  $(a,b,c)$  is a normal. If the surface is defined as the set of points satisfying  $S(x, y, z) = 0$ , then a normal at a point  $(x, y, z)$  on the surface is given by the gradient formed by the partial first derivatives with respect to each of the three axes:

$$\mathbf{N} = \nabla f(x, y, z) = \left[ \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, \frac{\partial S}{\partial z} \right]^T = \frac{\partial S}{\partial x} \hat{\mathbf{i}} + \frac{\partial S}{\partial y} \hat{\mathbf{j}} + \frac{\partial S}{\partial z} \hat{\mathbf{k}} \quad (1)$$

where  $\hat{\mathbf{i}}, \hat{\mathbf{j}}, \hat{\mathbf{k}}$  are the unit normal vectors along the  $x, y, z$  axes respectively. In the case where the surface is defined by  $z = S(x, y)$  the normal is:

$$\mathbf{N} = \left[ \frac{\partial S}{\partial x}, \frac{\partial S}{\partial y}, -1 \right]^T = [p, q, -1]^T \quad (2)$$

where  $p, q$  are the first derivatives of  $S(x, y)$  with respect to  $x$  and  $y$ , and the vector  $\mathbf{N}$  is an outward surface normal at  $(x, y)$ . The term outward refers to the direction of the normal with respect to the viewer. By convention, the viewer is assumed to be along the negative  $Z$  axis in relation to the surface being viewed. The surface normal  $(0,0,-1)$  points directly at the viewer and is orthogonal to the image plane. The distance between the viewer and the object is assumed to be great enough in relation to the object's size that the projection onto the imaging plane can be approximated by an orthographic projection. The normalised (unitary) normal vector  $\mathbf{n}$  is defined as:

$$\mathbf{n} = \frac{\mathbf{N}}{\|\mathbf{N}\|} = \frac{1}{\sqrt{p^2+q^2+1}} [p, q, -1]^T \quad (3)$$

Instead of gradients  $p, q$  the surface slant  $\phi$  and tilt  $\theta$  are commonly used, related to the surface normal by:

$$(n_x, n_y, n_z) = (\sin\phi \cos\theta, \sin\phi \sin\theta, \cos\phi) \quad (4)$$

For a plane given by the vector equation  $\mathbf{r} = \mathbf{t} + \alpha\mathbf{u} + \beta\mathbf{v}$ , where  $\mathbf{t}$  is a vector to get from the origin onto the plane and  $\mathbf{u}, \mathbf{v}$  are non-parallel vectors lying on the plane, the normal is given by  $\mathbf{N} = \mathbf{u} \times \mathbf{v}$ , i.e. the cross product of the vectors. There are two normals to every surface, one facing outward and the other inward, but for rendering the surfaces of objects the outward normal is usually chosen.

In shape from shading (SFS) algorithms, given a grey level image, the aim is to recover the light source and the surface shape at each pixel in the image. The reflectance from a surface can be used to infer the surface normal at each point and hence

to determine the 3-D shape of an object surface. Christensen and Shapiro (1994) specified the assumptions about the scene and the object as:

- All parts of the surface under consideration have the same physical properties (colour, reflectance, roughness, etc.), which are known or estimated *a priori*;
- Every light source is distant, so the intensity fall-off across the surface can be neglected. The position, shape, and intensity of the light source is known;
- The directions to each light source must differ significantly between the images, otherwise the shape information in the two images is redundant. On the other hand, if the light source positions are too different, the part of the surface that is illuminated in both images will be small;
- The object does not receive light reflected from other objects (or from other parts of the same object).

For a Lambertian surface, from which the incident light is scattered equally in all directions, the luminance of the reflected light is given by the vector dot product:

$$L_r = \rho \mathbf{L}_i \cdot \mathbf{n} = \rho |L_i| \cos\alpha \quad (5)$$

where  $L_r$  is the luminance of the diffusely reflected light (with no angular dependence),  $\rho$  is the maximum surface reflectance (or albedo),  $\mathbf{L}_i$  is the incident light vector,  $\mathbf{n}$  is the unit normal of the surface, and  $\alpha$  is the angle between  $\mathbf{L}_i$  and  $\mathbf{n}$ .

Because the normal vector has three components, at least three equations are needed to solve the system. This can be achieved by illuminating the surface for successive images from three different lighting directions with incident light vectors  $\mathbf{L}^1, \mathbf{L}^2$  and  $\mathbf{L}^3$ . This system can be written as:

$$I^k = \rho \mathbf{L}^k \cdot \mathbf{n} \quad (6)$$

where  $k=1,2,3$  are the three lighting directions. The three observed intensity values  $I^k$  of the reflected light can be stacked to form the 3x1 intensity vector  $\mathbf{I} = (I^1, I^2, I^3)$ , and the incident light vectors can be stacked row-wise giving the 3x3 lighting matrix  $\mathbf{L} = (\mathbf{L}^1, \mathbf{L}^2, \mathbf{L}^3)$ . Then Eq. (6) can be rewritten as:

$$\mathbf{I} = \rho \mathbf{L} \cdot \mathbf{n} \quad (7)$$

If the lighting directions  $\mathbf{L}^k$  are not coplanar, the matrix  $\mathbf{L}$  can be inverted, giving:

$$\rho \mathbf{n} = \mathbf{L}^{-1} \mathbf{I} \quad (8)$$

Since the normal vector  $\mathbf{n}$  is unitary, both its direction and albedo (modulus) can be recovered.

Woodham (1980) was the first to show the viability of SFS. He assumed the surfaces to be uniform, matte, convex and smooth with continuous first and second partial derivatives. The surfaces of real objects, however, have a surface microstructure that perturbs the angles of reflection, and also varying degrees of gloss that produces a specular component of reflectance. The effect is to introduce noise into the computation and a consequent scattering of the directions of normals from adjacent pixels. Coleman and Jain (1982) adapted the technique to include surfaces varying in both specularity and visual texture, which they treated as resulting from non-uniform surface albedo. They showed how specular reflections can be detected and removed by including a fourth source of illumination (Fig. 1). They calculated the overall deviation  $R_{dev}$  of the reflectances  $R_i$  from the mean  $\mu$  by:

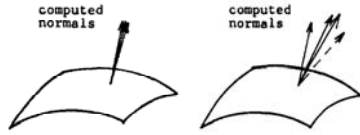


Figure 1. Computed normals from four light sources incident on (left) a matte surface, and (right) a surface exhibiting specularity (Coleman and Jain, 1982).

$$R_{dev} = \frac{\sum_{i=1}^4 (R_i - \mu)}{4 * R_{min}} \quad (9)$$

where  $R_{min}$  is the minimum of the four values. Before finalising the calculation of the surface normal at each point, a threshold value  $R_t$  is chosen representing the largest amount of reflectance deviation allowed before specularity occurs. To eliminate this specular component, the surface normal is averaged from the other three intensity values which have the smallest reflectance factor.

Rushmeier and Bernadini (1999) employed a simple rig with five tungsten lamps, one near the camera axis and four separated approximately  $45^\circ$  from the axis. At each pixel, after correction for variation in lamp intensity, the five values were sorted into order and the highest and lowest discarded. If the three remaining values were all non-zero then the normal and relative reflectance were calculated. Plata *et al* (2009) used an alternative setup in which both camera and object rotated with respect to a stationary light source. They showed that four-source photometric stereo could be applied effectively for recovering both shape and texture in RGB colour.

## 2. POLYNOMIAL TEXTURE MAPPING (PTM)

Malzbender *et al* (2001) introduced the polynomial texture map (PTM), a novel image-based relighting technique for visualising the appearance of a surface under a spatially variable source of illumination. It takes a set of images captured from a fixed camera position, with each image lit by a point source at a known coordinate position. The algorithm fits a parametric polynomial function to the set of intensities at each pixel location. The interactive viewer software uses the cursor position, representing the coordinates of a 'virtual light source', to generate the intensity of each pixel as if it had been illuminated from that direction. The effect is of a 'virtual torch' moving over a static 3D object surface, although there is no inherent 3D representation of the surface.

Malzbender was motivated by models of bidirectional texture function (BTF), but simplified the procedure by holding the exitant direction constant, i.e. with the reflected angle always toward the fixed camera position. The pixel intensity is a function of the angular coordinates  $(\Theta_i, \Phi_i)$  of the incident light source and two spatial variables  $(u, v)$ :

$$I = I_{r,g,b}(\Theta_i, \Phi_i, u, v) \quad (10)$$

By not including dependence on the exitant direction, he sacrificed the ability to capture view-dependent effects such as specularity, but retained the ability to represent arbitrary geometric shadowing and diffuse shading effects across a surface. On the assumption of a Lambertian surface, only the luminance of each pixel varies with light source direction and the chromaticity is taken to be constant. This enables separability of the reconstruction function, with a constant colour per pixel modulated by an angle-dependent luminance factor  $L(u, v)$ :

$$I = L(\Theta_i, \Phi_i, u, v)R(u, v) \quad (11)$$

and similarly for  $G(u, v)$  and  $B(u, v)$ . The dependence of luminance on light direction is modelled by a biquadratic function:

$$L(u, v; I_u, I_v) = a_0(u, v)I_u^2 + a_1(u, v)I_v^2 + a_2(u, v)I_u I_v + a_3(u, v)I_u + a_4(u, v)I_v + a_5(u, v) \quad (12)$$

where  $(I_u, I_v)$  are projections of the normalised light vector into the local texture coordinate system  $(u, v)$  and  $L$  is the resultant surface luminance. Given  $n+1$  images, the best fit at each pixel is computed using singular value decomposition (SVD) to solve the system of equations for  $a_0$ - $a_5$ :

$$\begin{bmatrix} I_{u0}^2 & I_{v0}^2 & I_{u0}I_{v0} & I_{u0} & I_{v0} & 1 \\ I_{u1}^2 & I_{v1}^2 & I_{u1}I_{v1} & I_{u1} & I_{v1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ I_{un}^2 & I_{vn}^2 & I_{un}I_{vn} & I_{un} & I_{vn} & 1 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_5 \end{bmatrix} = \begin{bmatrix} L_0 \\ L_1 \\ \vdots \\ L_n \end{bmatrix} \quad (13)$$

Thus a separate set of coefficients ( $a_0$ - $a_5$ ) is fitted to the image data for each pixel and stored as a spatial map referred to as a Polynomial Texture Map (PTM). The PTM has the same spatial resolution as each of the original images, but has a low resolution in the angular space of incident illumination, because the  $n$  directions of the image set are approximated by only 6 coefficients at each pixel.

Malzbender later designated the function in Eq. (11) as a Unidirectional Texture Function (UTF). By sacrificing the generality of viewing angle of a BTF, the UTF has the advantages of a compact texture representation well matched to the rendering process, which can also be directly employed for synthesis. PTMs derived in this way can be then used in place of conventional texture maps and applied to 3D objects, providing interactive and realistic control of lighting effects, such as shading, self-shadowing, inter-reflections and surface scattering (Hel-Or *et al*, 2003).

An attractive application of PTMs is the representation of ancient artefacts, such as inscriptions on early clay tablets. The interactive control of lighting conditions enables greatly enhanced perception of the surface structure compared to static photographs of the artefacts, thereby enhancing the legibility of surface relief and inscriptions. In a study on the paleontological illustration of fossils, Hammer *et al* (2002) found that PTM gave better results than laser scanning for specimens with very low surface relief. They noted that spatial resolution was compromised by computation of geometric surface normals from laser point cloud data, because of the convolution with a kernel having a spatial extent, whereas for PTM the normal estimation for each pixel is performed independently.

The PTM technique was applied at the National Gallery in London to make detailed visual representations of the surface structure of paintings. Twenty-four tungsten lamps were mounted onto an open framework in three tiers of eight lamps each. The camera was mounted at the top of the framework, pointing down to the painting on the floor. The lamps were turned on and off manually for each image in the sequence to be captured. With the variable 'virtual light' in the PTM viewer more features were visible than could be seen by raking light

from one direction alone, enabling the study of features in the painting resulting from ageing, such as craquelure and distortion of the support. Comparing PTM renderings made before and after physical changes to the painting facilitated examination of alterations in its texture and shape (Padfield *et al*, 2005).

We have implemented an apparatus for capturing sets of images from a fixed viewpoint. An acrylic hemispherical dome of diameter 1030 mm was fitted with 64 flash lights, arranged in five tiers (Fig. 2). The control electronics enables any combination of the lights to be selected and synchronised with a Nikon D200 digital camera mounted at the 'north pole' above an object placed on the horizontal baseboard in the 'equatorial' plane. By this means, multiple pixel-registered colour images of the object may be captured in sequence, each illuminated from a different direction. The coordinates of the flash lights were determined by a geometric calibration procedure based on shadows cast by a vertical pin placed at the centre of baseboard.



Figure 2. Hemispherical acrylic dome with 64 flash lamps.

Processing of the complete ensemble of 64 images generates the PTM representation, fitting a two-dimensional function of the variation of intensity with illumination angle for every pixel in the image. Display of the PTM file through the interactive viewer enables visualisation of the surface as if illuminated by a point source at any position within the hemisphere (the 'virtual torch' metaphor), throwing surface detail into relief when the light is placed at a low raking angle. Analysis of the 64-image set enables the surface normal to be extracted at each position, from which surface geometry can be approximated by the 'shape from shading' technique.

### 3. PHOTOGRAPHIC IMAGE CAPTURE

A small tablet from the Petrie Museum at UCL, a fragment of greywacke (fine sandstone) from the Egyptian 12<sup>th</sup> Dynasty, c. 1900 BC, provided a good basis for evaluation of the imaging process. The object is irregular, approximately 10 cm in diameter, probably from a table top. The inscription is quite shallow with a depth varying between 1.3 and 2 mm. Under direct lighting from above, i.e. along the axis of the lens, as would typically be obtained from a camera with an integrated flash, the inscribed characters are barely visible. As the angle of illumination increases, however, their dimensionality becomes well defined (Fig. 3). For the sixteen flash lamps in the lowest tier, the angle of incidence of approximately 5° to the surface produces raking light with dramatic shadows.

The Nikon D200 camera with a 17.5mm zoom lens was mounted in its fixed position at the top of the dome, at a focussing distance of approximately 700mm from the object surface. The images were captured at ISO 100 with the lens set

at  $f/4$  aperture and 55mm focal length. In this geometry the image samples the surface of the tablet at a linear resolution of 15.3 pixels per mm, i.e. each pixel represents 65  $\mu\text{m}$  on the surface.



Figure 3. A fragment of an Egyptian tablet c.1900 BC in the Petrie Museum at UCL (item UC35501). The five images were all taken from the same camera viewpoint, with flash illumination from each tier of the dome, corresponding to angles of incidence of: (top row) 80°, 60°, 40° (bottom row) 20°, 5°.

The PTM representation was generated by processing the complete ensemble of 64 images, using the utility program 'PTMfitter' supplied by HP ([www.hpl.hp.com/research/ptm/](http://www.hpl.hp.com/research/ptm/)) The algorithm, developed by Malzbender *et al* (2001), uses singular value decomposition (SVD) to fit a two-dimensional function to the variation of intensity with illumination angle for every pixel (Eq. 13). Each pixel in the output file is represented by 6 luminance coefficients and three RGB colour values, all encoded as unsigned 8-bit bytes. Hence the size of the PTM file, with 9 bytes per pixel, is three times the size of an RGB TIFF file with 3 bytes per pixel, but the representation of the object is much richer because it incorporates information about the surface relief. The PTM algorithm was also implemented in Matlab and checked to ensure that the generated PTM file was identical to that produced by the utility.

Display of the PTM file through the interactive viewer (Fig. 4) enables visualisation of the surface as if illuminated by a point source at any position within the hemisphere. It is as if a 'virtual torch' were being moved above the object. The primary window of the viewer (left) shows the reconstruction of the object, lit from the current lighting direction. The circular area (upper right) is a user interface widget, representing all the positions of the hemisphere, from which the lighting direction can be controlled. The program provides a second function whereby placing the cursor over the object displays in the circular area the directional luminance distribution for the pixel selected.

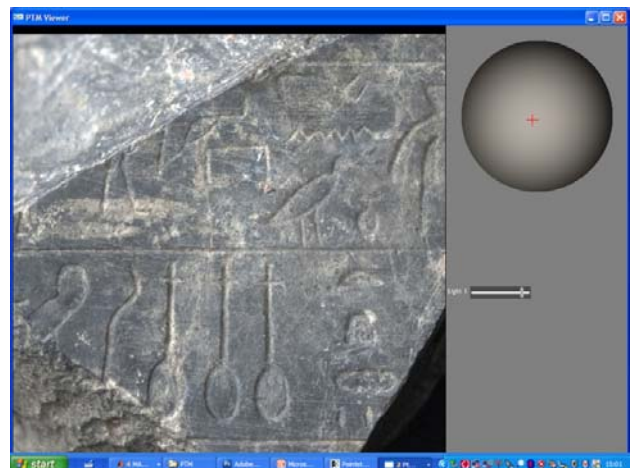


Figure 4. Screen shot of the PTM viewer utility.

#### 4. GENERATING SURFACE NORMALS

An essential step in producing a 3D representation is to calculate the surface normals at each position of the image. The normals are readily extracted from the PTM representation because the six coefficients stored for each pixel already contain the directional luminance information. An estimate of the normal at each pixel can be extracted from the PTM by setting  $\frac{\partial L}{\partial u} = \frac{\partial L}{\partial v} = 0$  to solve for the maximum of the biquadratic in Eq. (12). This represents the angle of illumination at which the reflected luminance is maximum, bisected by the surface normal  $\mathbf{N}$ :

$$\mathbf{N} = (I_u, I_v, \sqrt{1 - I_u^2 - I_v^2}) \quad (14)$$

where:

$$I_u = \frac{a_2 a_4 - 2a_1 a_3}{4a_0 a_1 - a_2^2} \quad I_v = \frac{a_2 a_3 - 2a_0 a_4}{4a_0 a_1 - a_2^2} \quad (15)$$



Figure 5. Images of normals in the X,Y and Z directions extracted from the PTM representation.

The normals  $N_x$  and  $N_y$  are in the range  $[-1, +1]$  whereas  $N_z$  is always positive in the range  $[0, 1]$ . For convenient storage and display of the normals as an image map, they are scaled to 16-bit pixel values  $P_x$ ,  $P_y$  and  $P_z$ , in which the zero values for  $N_x$  and  $N_y$  are represented by 32767 (mid-grey):

$$\begin{aligned} P_x &= \text{int}(32767 * (N_x + 1)) \\ P_y &= \text{int}(32767 * (N_y + 1)) \\ P_z &= \text{int}(65535 * N_z) \end{aligned} \quad (16)$$

The resulting normals for the Egyptian tablet (Fig. 5) show clearly the directionality of the recessed inscription. Slopes upwards to the right, which have negative  $N_x$ , are darker than the mid-grey of the horizontal surface. Conversely slopes downwards to the right, which have positive  $N_x$ , are lighter. Thus the image of  $N_x$  shows the predominantly vertical features of the surface. Similarly the image of  $N_y$  shows the predominantly horizontal features. The value of  $N_z$  is close to 1 over the whole surface except where the broken parts fall away at the sides, and carries little visual information.



Figure 6. (left) Detail of 150x150 pixels of an image from lamp 34 (Tier 3). Normals were computed from the green channel for three-lamp combination 17-36-58: (centre) X normal; (right) Y.

The normals for the same object were calculated by the three-light photometric stereo method (Eq. 8), using a subset of three of the 64 images captured in the dome, corresponding to lamps in Tiers 2, 3, 4. The image data was linearised through a lookup table with the inverse 8-bit tone curve of the camera. The quantisation errors arising from this step and other aspects of the 8-bit encoding are evident in the noisiness of the result (Fig.

6 centre) for a region near the centre of the tablet. The normals at each pixel were computed from the green channel of the image. There are various possible reasons for the noise (variability) seen in the computed normals:

- Coordinates of three lamps are not linearly independent;
- Errors in the XYZ coordinates of the lamps;
- Non-point source (non-zero length of flash lamp);
- Variable intensity of the flashes (three sources different);
- Non-uniform reflectance of surface as a function of angle;
- Scattered (flare) light adds to the direct illumination;
- Non-uniform illumination distribution across image area;
- Physical movement of camera relative to object;
- Non-linear signal from the camera;
- Noise in the camera (amplified by the matrix inversion);
- Quantising errors in the RGB image (3x8 bits per pixel);
- Compression errors from the JPEG image file encoding.

The effect of the relative geometry of the lamps was explored by analysing the normals computed from all  $16*16*12 = 3072$  three-lamp combinations of the lamps in Tiers 2, 3 and 4. Histograms (Fig. 7) of the  $N_x$ ,  $N_y$  and  $N_z$  normal values for one pixel reveal the distributions to be very widely spread but with a well-defined central peak. The mean was calculated for each distribution and then outliers greater than one standard deviation distant were removed and the mean recalculated.

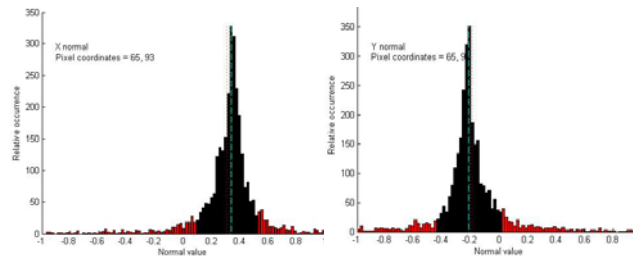


Figure 7. Histograms of X and Y normals in 100 bins, computed for the pixel at coordinate (65,93), for 3072 combinations of three lamps. The black bars are within  $\pm 1$  stdev of the mean of the distribution, the red bars are outliers. The cyan dashed line is the median.

It was found that in all cases the revised mean (after the removal of outliers, corresponding to the bars shown in red in Fig. 7) moved closer to the median of the original distribution. The median of the distribution of each normal component can therefore be regarded as a good approximation of the normal.

To optimise the quality of the normals, the coordinate positions of the three lamps should fulfil two criteria in relation to each other and the object:

1. The distance between the plane through the lamps and the centre of the object should be maximised. If the plane through the lamps passed through the reconstruction point on the object surface, the normals could not be resolved;
2. The lamps should be as far from collinear as possible, in order to give the greatest differentiation of the orthogonal X and Y components on the object surface. This is equivalent to maximising the area of the triangle with vertices at the coordinates of each of the three lamps.

These relationships can be visualised in Fig. 8, in which the triangle connects one lamp in each of Tiers 2, 3 and 4 of the dome. Criterion 1 requires the triangle to be as nearly parallel as possible to the baseboard (ground plane), whereas Criterion 2 requires the area of the triangle (red) to be maximised.

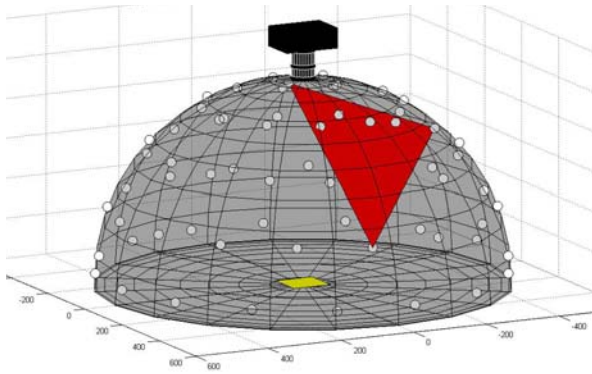


Figure 8. Schematic layout of 64 lamps (white circles) on the surface of the hemispherical dome, showing the camera at the ‘north pole’, the object (yellow) in the equatorial plane, and a triangular section (red) of a plane intersecting the coordinates of three lamps in Tiers 2, 3 and 4.

The first criterion can be expressed mathematically by the equation of a plane passing through three points  $x_1, x_2, x_3$ :

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = \begin{vmatrix} x - x_1 & y - y_1 & z - z_1 \\ x_2 - x_1 & y_2 - y_1 & z_2 - z_1 \\ x_3 - x_1 & y_3 - y_1 & z_3 - z_1 \end{vmatrix} = 0 \quad (16)$$

This is equivalent to the standard form:

$$ax + by + cz = d = 0$$

where the coefficients  $a, b, c, d$  are defined by the determinants:

$$d = \begin{vmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{vmatrix}, \quad a = \begin{vmatrix} 1 & x_2 & x_3 \\ 1 & y_2 & y_3 \\ 1 & z_2 & z_3 \end{vmatrix},$$

$$b = \begin{vmatrix} x_1 & 1 & x_3 \\ y_1 & 1 & y_3 \\ z_1 & 1 & z_3 \end{vmatrix}, \quad c = \begin{vmatrix} x_1 & x_2 & 1 \\ y_1 & y_2 & 1 \\ z_1 & z_2 & 1 \end{vmatrix} \quad (18)$$

The three components of the normal  $n$  to the plane are given by:

$$n_x = \frac{a}{r}, \quad n_y = \frac{b}{r}, \quad n_z = \frac{c}{r} \quad (19)$$

where  $r = \sqrt{a^2 + b^2 + c^2}$ , and the perpendicular distance from the origin to the plane is:

$$D = \frac{d}{r} \quad (20)$$

The second criterion can be expressed in a different way. Three points  $x_1, x_2, x_3$  are defined to be collinear if they lie on a single straight line, in which case the ratios of distances satisfy:

$$x_2 - x_1 : y_2 - y_1 : z_2 - z_1 = x_3 - x_1 : y_3 - y_1 : z_3 - z_1 \quad (21)$$

A more tractable condition is obtained by noting that the area of the triangle determined by three points will be zero if and only if the points are collinear:

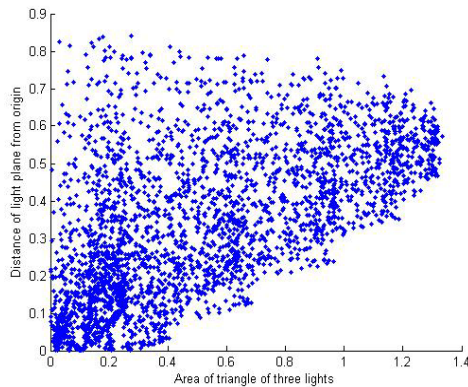


Figure 9. Distance vs area for 3072 combinations of three lamps.

$$A = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0 \quad (22)$$

Thus the objective is to maximise  $D$  from Eq. (20) and  $A$  from Eq. (22). The plane of the three lamps is perpendicular to the plane of the object if  $n_z = 0$ , and parallel if  $n_z = 1$ . The scatter plot of area  $A$  vs distance  $D$  (Fig. 9) shows a wide spread, with a correlation coefficient  $R=0.58$ , verifying their relative semi-independence as measures of the suitability of a given combination of three lamps. These values, normalised to the radius of the hemisphere, were calculated for each of the 3072 lamp combinations and plotted against the difference between the calculated normal and the median for each of the three normal components (Fig. 10).

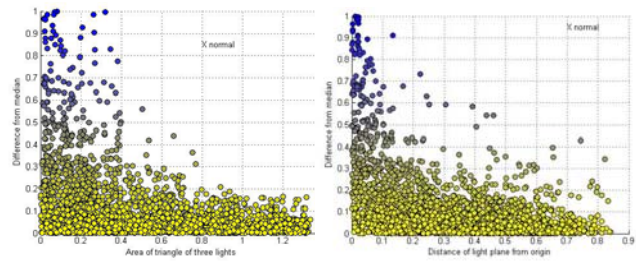


Figure 10. Distributions of differences between calculated and median values for X normals for pixel (65,93) for 3072 combinations of three lights vs distance from origin of intersecting plane (left) and vs area of triangular patch with lights at its vertices (right). The univariate colour coding corresponds to the absolute size of the differences, from yellow (best) to blue (worst).

For all three components of the normal there is a clear dependence of the error on the area of the triangle, with the great majority of erroneous points (each corresponding to an unsuitable geometric configuration of three lamps) falling below the abscissa value of 0.5. Excluding these points left 579 lamp combinations (18.8% of the total), which reduced the mean (absolute) error for both X and Y normals to 0.08.

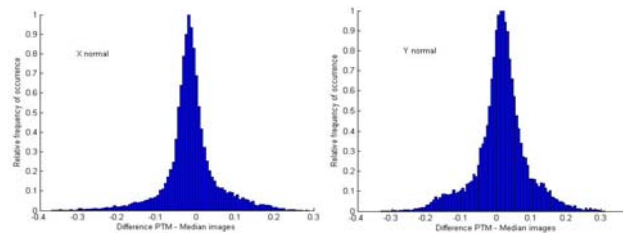


Figure 11. Distribution of differences in X normals between PTM and median images.

Comparing the normals extracted from the PTM file with the normals generated by the photometric stereo method revealed that there is much more detail in the latter. The histograms of differences in the X and Y normals (Fig. 11) show a broad distribution, with a mean absolute difference of 0.05.

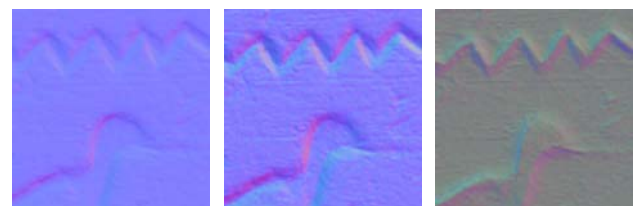


Figure 12. Normals derived from PTM (left) and photometric stereo (centre), and their differences (right).

The effect on the image is clear from comparison of the pseudo-colour images of the normals (Fig. 12) derived from PTM and from the median, for a detail of 150x150 pixels. The PTM normals lack both contrast (i.e. gentler slopes) and high frequencies (i.e. fine detail smoothed out). The difference image (Fig. 12 right) shows that the differences are greatest in the regions of maximum gradient.

### 5. NORMALS FROM 3D SCANNER

The tablet was scanned by a 3D colour laser scanner at a resolution of 10 lines/mm, i.e. a sampling pitch of 100  $\mu\text{m}$ . The data was exported as a text file, in which each point is represented by one line of 9 numerical fields encoded as ASCII text. Each line contains the X,Y,Z point coordinates, R,G,B colour values, and  $N_x$ ,  $N_y$ , and  $N_z$  point normal values. The point data was projected onto the nearest locations in a two-dimensional image array with a resolution of 10 pixels/mm, ignoring variations in the Z coordinate. The resulting RGB image and the corresponding normal image are shown in Fig. 13.

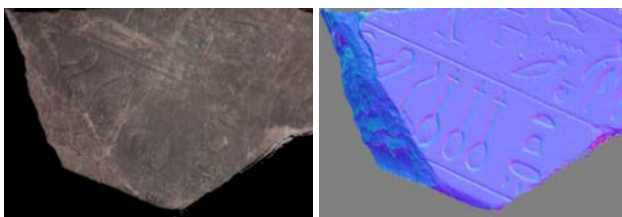


Figure 13. (left) 2D image constructed from 3D scan data on a 100  $\mu\text{m}$  grid; (right) corresponding image of 2D normals, based on data generated by the scanner software.

Side-by-side comparison of a corresponding region (Fig. 14) shows clearly the superiority of the photometric stereo technique for this object. It is not only the higher sampling rate (15.5 pixels/mm for the camera instead of 10 pixels/mm for the laser scanner) that gives greater clarity of the details. The MTF of the scanner is poorer in the vertical direction, because of jitter in the galvanometer positioning mechanism of the laser beam.

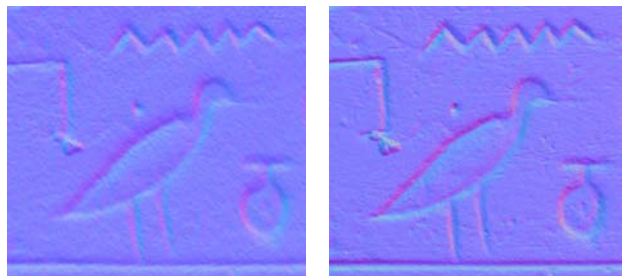


Figure 14. Normals of an area of size 22.3 by 20.4 mm generated by (left) 3D scan data, image of 223x204 pixels; (right) photometric stereo from photographic image set, image of 342x320 pixels. The image is reproduced here at approximately 1.7x magnification of life size.

The ability of the camera and 3D scanner to resolve spatial detail was evaluated by the slanted edge method (Burns, 2000), with the results shown in Fig. 14. The camera has similar spatial frequency response in both horizontal and vertical directions, falling smoothly to a low value at the Nyquist frequency of 7.6 cycles/mm. The scanner showed very different behaviour in the two directions, having a poor response in the vertical axis (along the laser scan direction), but an excessively high response causing aliasing in the horizontal axis (direction of traverse of the carriage).

In conclusion, this study has shown that surface normals extracted from the PTM image data structure are of lower quality than surface normals computed by photometric stereo

methods from a subset of the same images used to make the PTM. The directional accuracy of the normals is compromised by the PTM biquadratic function's smoothing over the range of illumination angles in the hemisphere. For surfaces derived from laser scanning, a comparable loss of resolution of surface detail is caused by the smoothing effect of regression over the point cloud and triangulation of rendering meshes. The great advantage of the laser scanner, of course, is that the surface height data (Z coordinate) is directly available, whereas for photometric methods it must be generated from the normals.

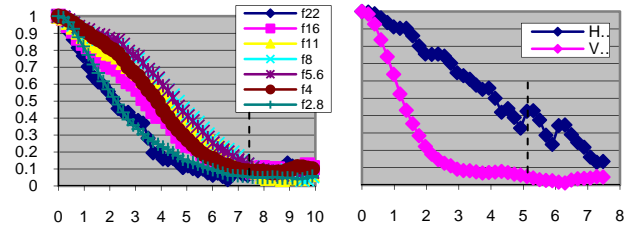


Figure 14. Spatial frequency response in cycles/mm on the object surface of (left) the Nikon D200 camera with zoom lens set at 55mm; (right) the 3D laser scanner. The vertical dotted line indicates the Nyquist frequency for each.

### REFERENCES

- Burns PD (2000) Slanted-Edge MTF for Digital Camera and Scanner Analysis, *Proc. IS&T PICS Conf.*, 135-138.
- Christensen PH and Shapiro LG (1994) Three dimensional shape from colour photometric stereo, *Intl. Journal of Computer Vision*, **13**(2) 213-227.
- Coleman EN and Jain R (1982) Obtaining 3-dimensional shape of textured and specular surfaces using four source photometry, *Computer Graphics and Image Processing* **18**(4) 309-328.
- Hammer Ø, Bengtson S, Malzbender T and Gelb D (2002) Imaging Fossils Using Reflectance Transformation and Interactive Manipulation of Virtual Light Sources, *Palaeontologia Electronica* **5**(4), 9pp.
- Hel-Or Y, Malzbender T and Gelb D (2003) Synthesis of Reflectance Function Textures from Examples, Technical Report HPL-2003-16 (R.1), HP Laboratories Palo Alto.
- Malzbender, T., Gelb, D. and Wolters, H. (2001), 'Polynomial Texture Maps', *Proc. ACM SIGGRAPH*, **28**, 519-528.
- Padfield J, Saunders D and Malzbender T (2005) Polynomial texture mapping: a new tool for examining the surface of paintings, *ICOM 14<sup>th</sup> Triennial Meeting*, Vol. 1, 504-510.
- Plata C, Nieves JL, Valero EM and Romero J (2009) Trichromatic red-green-blue camera used for recovery of albedo and reflectance of rough-textured surfaces under different illumination conditions, *Applied Optics*, **48**(19) 3643-3653.
- Rushmeier H and Bernardini F (1999) Computing Consistent Normals and Colors from Photometric Data, *Proc. IEEE Second Intl. Conf. on 3-D Imaging and Modeling (3DIM '99)*, 10pp.
- Soucy M, Godin G, Baribeau R, Blais F and Rioux M (1996) Sensors and Algorithms for the Construction of Digital 3-D Colour Models of Real Objects, *Proc. IEEE Intl. Conf. on Image Processing (ICIP)*, Lausanne, 409-412.
- Woodham RJ (1980) Photometric method for determining surface orientation from multiple images, *Optical Engineering*, **19**(1) 139-144.

### ACKNOWLEDGEMENTS

The authors would like to acknowledge the Petrie Museum for making the Egyptian tablet available, and Mona Hess for kind assistance with laser scanning and photographic imaging.