Visualising an Egyptian Artefact in 3D: Comparing RTI with Laser Scanning

Lindsay W. MacDonald
Photogrammetry, 3D Imaging and Metrology Research Centre
Dept. of Civil, Environmental and Geomatic Engineering, University College London
ucfslwm@ucl.ac.uk

3D digital representations of an ancient Egyptian artefact were compared for their rendering of surface detail. Normals were generated by three methods: (1) point clouds from the Arius 3D colour scanner; (2) reflectance transform imaging (RTI); (3) photometric stereo. The latter two were constructed from sets of 64 digital images taken under directional lighting in a hemispherical dome. Analysis of the 3D surface normals of corresponding sections of each object indicated that the photometric stereo method produced the best resolution of spatial detail.

3D surface normals. Polynomial texture mapping. Reflectance transform imaging. Laser scanning.

1. INTRODUCTION

An object was selected from the UCL Petrie Museum (Fig. 1). This funerary cone, originally embedded above the entrance to a tomb of the Egyptian New Kingdom c.1200 BC, is moulded from clay, and is approximately 10 cm in diameter and nearly circular. The inscription, reading down each column and left-to-right, refers to a man called Nefer-lah and the moon-god Lah and a woman called Hemet-Netjer and the god Amun. The raised characters protrude from the surface (positive relief) to a height of approximately 1 mm.

Figure 1: Funerary cone from Petrie Museum, UC37585

1.1 3D colour laser scanner

The Arius 3D colour laser scanner at UCL (Fig. 2) employs technology developed over the past 25 years at NRC, Canada (Blais 2004). The dimensions of the scanning volume (in mm) are X=890, Y=762, Z=508, where X and Y are in the

base plane and Z is vertically upwards. In operation the three laser beams (red, green, blue) are scanned by a galvanometer on the Y axis while the whole head assembly is moved along the X axis (and also in Y and Z if necessary for the scan trajectory). At a given position of the head, the range of the scanner is a cube of side 60 mm.

Figure 2: Careful placement of the funerary cone in the Arius 3D colour laser scanner

The position of the object surface along the Z axis is measured by laser triangulation, in which one side of the scanning mirror deflects the laser beams across the object (Fig. 3) while the opposite side of the same mirror compensates the return beam's angular movement across the CCD sensor. Changes in the position of the light spot along the Z axis produce net movement across the CCD. The laser spot size is nominally 80 μm and the surface is sampled at 100 μm intervals in normal scanning mode, enabling details of 0.2 mm to be resolved.

Figure 3: White raster line formed by three colour laser beams scanned across the surface of the clay cone

1.2 Hemispherical illumination dome

An apparatus has been developed at UCL 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. 4). 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 using measurements of the length and direction of shadows cast by a vertical pin placed at the centre of baseboard. All images were captured as RAW (.nef file format) then converted to 16-bit TIFF via the utility DCRAW, to ensure linearity of the signal with respect to scene luminance.

Figure 4: Hemispherical acrylic dome with 64 flash lamps

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 characters on the object surface are barely visible. As the angle of illumination increases, however, their

dimensionality becomes well defined (Fig. 5). 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.

Figure 5: The images of the clay funerary cone were all taken from the same camera viewpoint with the same exposure, under flash illumination from five tiers of the dome, corresponding to approximate incident angles of (left to right): 80°, 60°, 40°, 20°, 5°

2. REFLECTANCE TRANSFORM IMAGING (RTI)

Malzbender *et al.* (2001) introduced 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 dependence of luminance on light direction is modelled by a biquadratic function:

$$L(u, v; I_u, I_v) = a_0(u, v)I_u^2 + a_1(u, v)I_v^2 + a_2(u, v)I_u I_v + a_3(u, v)I_u + a_4(u, v)I_v + a_5(u, v) \quad (1)$$

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 representation 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.

Over the past decade the principle of PTM has been generalised into a family of image capture techniques, known as reflection transformation imaging. RTI has become established as a powerful method for acquiring and representing the 3D reflectance properties of an object, and displaying them as an interactive 2D image. From

an initial set of photos of the object, taken under controlled illumination, an approximation of the reflectance function of the objects' surface is calculated for each pixel enabling the rendering and 'relighting' of the object. Contrast enhancement of specular and diffuse components has proved to be very useful for both documentation and visualisation, revealing surface features not visible during physical inspection.

An attractive application of RTIs is the representation of ancient artefacts, such as inscriptions on early clay tablets. The interactive control of lighting conditions in the viewer software enables greatly enhanced perception of the surface structure compared with 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 the spatial resolution was compromised by computation of geometric surface normals from laser point cloud data, because of convolution with a kernel having a spatial extent, whereas for PTM the normal estimation for each pixel is performed independently.

RTI's usefulness has been demonstrated for the surfaces of objects, including coins (Mudge *et al.* 2005), cuneiform tablets (Willems *et al.* 2005), Byzantine glass tesserae (Zányi *et al.* 2007), marble friezes (Dellepiane *et al.* 2006), rock art (Mudge *et al.* 2006), and the Antikythera Mechanism (Freeth *et al.* 2006). Advantages of RTI in the cultural heritage domain include (Happa *et al.* 2009):

- Non-contact acquisition of data;
- Clear representation of 3D shape characteristics;
- Visualisation through interactive viewing tools;
- Better discernment of surface detail than direct physical examination;
- No data loss caused by shadows and specular highlights;
- Simple image processing pipeline;
- Higher resolution on the object surface than obtainable with 3D scanners.

Initially, RTI was applied only to small-sized objects, typically up to 30 cm diameter, that would fit within the image field of a camera mounted in the fixed geometry of an illumination dome. More recently, however, new capture techniques have enabled larger objects to be represented. Instead of a physical illumination dome, the light may be placed in different locations to form a 'virtual illumination dome'. The dimensions of this virtual dome and its light distribution depend on the size of the target object and on the number of light directions needed to sample its reflectance function. Dellepiane *et al.* (2006) employed this technique by placing the lamp successively in pre-

determined positions to create a virtual dome with a diameter in excess of 3 metres, for *in situ* capture of relief surfaces on tombs and sarcophagi.

Mudge *et al.* (2006) developed a highlight-based method (HRTI), recovering the light directions from highlights recorded on one or more glossy spheres placed as calibration targets in the scene. Image analysis enables the direction vector to be determined for each light source (Barbosa *et al.* 2007). HRTI offers greater flexibility in image capture because the light source can be placed in any convenient position and its geometrical coordinates calculated later from information contained within the image.

3. EXTRACTING SURFACE NORMALS

3.1 Normals from 3D scanner

The Arius scanner produced a point cloud for each object, representing the surface topography. 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 , N_z point normal values. To compensate for the slant of the top surface, principal component analysis was applied to determine the plane of best fit. The rotated point data was projected onto the nearest locations in a two-dimensional image array with a resolution of 10 pixels/mm, ignoring variations in Z. The resulting RGB image and corresponding normal image are shown in Fig. 6.

Figure 6: (left) 2D image constructed from 3D scan data on a 100 μm grid; (right) corresponding image of 2D normals, based on data generated by scanner software

3.2 Normals from RTI

An implementation in Matlab fitted the biquadratic function of Eq. (1) at each pixel of the complete ensemble of 64 images to generate the RTI representation. Display of the resulting file through the interactive PTM viewer from HP enables visualisation of the surface as if illuminated by a point source at any position within the hemisphere (the 'virtual torch' metaphor), throwing the surface detail into relief when the light is placed at a low raking angle.

Figure 7: Images of normals in the X and Y directions extracted from the RTI representation, and (right) combined into an RGB pseudo-colour image

The surface normals are readily extracted from the RTI representation because the six coefficients stored for each pixel already contain the directional luminance information (MacDonald & Robson 2010). The resulting normals for both objects (Fig. 7) show clearly the directionality of the raised relief of the moulding on the cone's surface. Edges on the left and lower sides of the raised features have negative normals corresponding to positive slopes, shown as darker than the mid-grey of the horizontal surface. Conversely edges on the right and upper sides have positive slopes and are lighter. Thus the image of X normal (N_x) shows the predominantly vertical features of the surface. Similarly the image of Y normal (N_y) shows the predominantly horizontal features. The value of the Z normal (N_z) is close to 1 over the whole surface and carries little visual information.

Figure 8: Pseudo-colour images of the normals computed from 144 combinations of three lights

3.3 Normals from photometric stereo

Photometric stereo is a sub-class of 'shape from shading' (SFS) algorithms. Given a grey level image, the aim is to recover the light source and

the surface height 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 & Shapiro 1994).

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:

$$L_r = \rho \mathbf{L}_i \cdot \mathbf{n} = \rho |\mathbf{L}_i| \cos \alpha \quad (2)$$

where L_r is the luminance of the diffusely reflected light (with no angular dependence), ρ 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 α is the angle between $\mathbf{L}_i$ and $\mathbf{n}$.

Because the normal vector has three components, three equations are sufficient 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:

$$I^k = \rho \mathbf{L}^k \cdot \mathbf{n} \quad (3)$$

3.4 Lamp triplets

Using the same set of 64 images as for the RTI computation, the photometric stereo method was applied to combinations of three lamps selected from Tiers 2, 3 and 4 of the dome, and the normals calculated for each. Examination of the results shows an obvious wide variation in quality (Fig. 8). Although the majority of the pseudo-colour normal images resemble the image of the RTI normals (Fig. 7), some are significantly different. Possible reasons for the variability seen in the computed normals include:

- Non-uniform reflectance of surface as a function of angle (i.e. non-Lambertian);
- Errors in XYZ coordinates of the lamps;
- Non-point source (the flash lamp of length 8 mm at dome radius of 515 mm subtends an angle of approximately 0.9° at the object surface);
- Variable intensity of the flashes (differing illuminance of the surface by the three sources);
- Ambient and flare light (scattered from other areas) adds to the direct illumination;
- Non-uniform illumination across image area;
- Physical movement of the camera relative to object between exposures;
- Non-linear signal from the camera (pixel value not proportional to luminance);
- Noise in camera (amplified by matrix inversion);
- Coordinates of three lamps not linearly independent (i.e. not spanning 3D space).

The noisiness of the result obtained by computing surface normals from one triplet of lamps is clear in Fig. 9. The normals at each pixel were computed from the green channel of the image for a small area of approximately 1 cm square near the centre of the object. For a Lambertian surface in a perfectly linear system without noise or quantising errors, all three colour channels should give equal normals, because the intensities of a given pixel should be in the same ratio under each light.

Figure 9: (left) Detail of image from lamp 19 (Tier 2), of size 140x140 pixels. Normals were computed from the green channel of this detail, for lamp combination 18-39-52: (centre) X normal; (right) Y normal

The effect of the relative geometry of lamps was explored by analysing the normals computed from all valid three-light combinations of the lamps in Tiers 2, 3 and 4. For each of the 64 images the luminance was calculated as the weighted sum of the *R,G,B* channels, a 5x5 median filter was applied, and a 100x100 pixel section extracted. For each triplet lamp #1 was selected from Tiers 2, 3, and lamps 2, 3 selected from Tiers 3, 4. Of the total possible 32x28x28=25,088 combinations, a subset of 6,780 was selected by ensuring that no lamp was repeated and that the condition number was less than 10 (to prevent a near-singular matrix).

Figure 10: Schematic layout of 64 lamps (white circles) on the surface of hemispherical dome, showing camera at '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 5

Histograms of the N_x , N_y , N_z normal values revealed the distributions to be widely spread but with a well-defined central peak. It was found in all cases that the median could be regarded as a good approximation of the true value of the normal. To optimise the quality of the normals, the coordinate positions of the three lights 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 the three lights, shown in red in Fig. 10.

The difference of each calculated normal from the median is plotted against triangle area in Fig. 11 and shows a wide spread, with the general trend that larger area of the lamp triangle generally gives lower errors. The figure also colour codes the matrix condition number, on a scale from grey (best) to red (worst), which shows that lamp triangles with larger areas are better conditioned.

Figure 11: Difference from median vs area of lamp triangle for 6780 combinations of three lamps

The statistics were calculated for all of the selected 6780 three-lamp combinations across all pixels in the 100x100 image detail (Fig. 12). Two subsets of 301 and 31 of the lamp combinations were selected, by ranking the minimum overall difference from the median (calculated as sum of squares of the differences of X and Y normal components). The optimum set of 31 lamp triplets was then used to calculate the normals for the whole image, using the median of the 31 values at each pixel.

Figure 12: Distribution of Y normal values for one pixel for selected three-lamp combinations: (black) 6780; (yellow) 301; (green) 31 lamps

3.5 Integrated lamp set

Instead of calculating multiple triplet lamp combinations there is a more direct way to calculate the normals. The fundamental equation (3) can be solved for normals $\mathbf{n}$ at each pixel position as:

$$\rho \mathbf{n} = \mathbf{L}^{-1} \mathbf{I} \quad (4)$$

where $\mathbf{L}$ is the 3×64 lamp matrix consisting of the X,Y,Z direction cosines for each of the 64 lamps and $\mathbf{I}$ is the 1×64 vector of linear intensity values. This pseudo-inverse is easily calculated as $\backslash \mathbf{L}$ in Matlab, which performs a regression to find the least-mean-squares fit.

Figure 13: Normals of 20x20 mm area of cone from (top left) Arius 3D scanner; (top right) RTI; (bottom left) PS 64-lamps; (bottom right) PS selected lamp triplets

4. COMPARISON OF RESULTS

Surface normals computed by the four techniques described above are presented in Fig. 13 for a small area of 20x20mm of the funerary cone. The image area for the Arius scanner is 200x200 pixels at 10 pixels/mm and for the camera images it is 300x300 pixels at 15 pixels/mm. Relative to the photometric stereo (PS) results, the Arius normals provide good contrast but lack detail.

Figure 14: Difference images between normals: (left) PS – RTI; (right) PS triplets – PS 64-lamps

The effect on the image is clear from comparison of the pseudo-colour images of the normals derived from RTI (detail of Fig. 7) and from the lamp triplets. The RTI normals are lacking in both contrast (i.e. gentler slopes) and high frequencies (i.e. fine detail smoothed out). The difference image (Fig. 14 left) shows that the differences are greatest in regions of maximum gradient.

Figure 15: Histograms of X normals generated by four different methods

Comparison of the histograms of the normals (Fig. 15) within a 960x960 area of the normals image shows that the PS triplet method gives the broadest spread and the RTI method the narrowest spread. This is consistent with the greater and lesser contrast seen in the respective images (Fig. 13).

Cross-sections through the X normals, i.e. the intensity plot of a single row of the normal image, show the relative contrast of the three versions of the normals derived from the set of 64 images (Fig. 16). It is clear that all three show the same features, and that the contrast of the PS triplet results is greater than the contrast of the PS 64-lamp results, which in turn is greater than the contrast of the RTI results.

Figure 16: Cross-sections through X normal map produced by three methods

Figure 17: The red lines show the cross-section through: (left) X normals map generated by PS triplets method; (right) Depth map (Z coordinate) generated from Arius 3D scanner point cloud

An indication of the physical surface profile is obtained by integrating the normals along the row. The higher contrast of the results of the PS triplet method produces a greater amplitude between peaks and troughs of the profile. Comparing the three profiles with the same cross-section of the 'ground truth' dimensional data from the laser scanner (Fig. 17 right), it is clear that the PS triplet method is nearest to the original. The peak at the right of the PS profile has a height of 0.8 mm compared with 1.1 mm for scanner data (Fig. 18).

Figure 18: (top) Z coordinates of Arius data for cross-section; (bottom) height profile reconstructed from surface normals

The frequency spectrum of the normal maps was determined by taking the 1-D FFT across the width of each row and averaging over all rows. This technique has previously been used for analysis of the characteristic spatial frequency 'signature' of prints and patterned material (MacDonald 2010). The results (Fig. 19) show a peak response in the range 0.15-0.3 cycles/mm, corresponding to slopes in the range 3-6. The laser scanner and PS triplets give similar results for these low frequencies, but the laser scanner response falls away above 1 cycle/mm then flattens out above 3 cycles/mm as it reaches the noise limit. The power at all frequencies from the PS triplets method is greatest and from the RTI method is lowest.

Figure 19: Average 1-D frequency spectrum of X normals; (bottom) zoomed section showing peak

5. CONCLUSION

This study has shown that surface normals extracted from the RTI image data structure are of lower quality than surface normals computed by photometric stereo methods from the same set of images. Their directional accuracy is compromised by the smoothing of the RTI biquadratic function over the hemispherical range of illumination angles. This becomes clear when plotting the actual distribution of the 64 intensities from the lamps with the reconstructed intensities in each of the lamp directions (Fig. 20). The actual distribution (pink) is more irregular than the prediction of the biquadratic function (grey), being smaller at lower angles and larger at medium angles. Also the RTI normal (blue) is closer to vertical than the PS normal (red).

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 3D laser scanner, of course, is that surface height data (Z) is directly available, whereas for photometric methods it must be generated from normals.

Figure 20: 3D plots of the actual intensity for one pixel of the surface illuminated by 64 lamps (pink) vs the intensity reconstructed by the RTI biquadratic function (grey)

Dellepiane *et al.* (2006) analysed the quality of normals derived from PTMs made from 105 light sources, to determine how few lights could be used. They concluded that a subset of 60-70 lamps, evenly spaced over the hemisphere are sufficient to produce a PTM without noticeable degradation. The key factors that affect the quality of the RTI representation are:

- Spatial resolution of the images, i.e. ability to resolve fine details;
- Signal-to-noise ratio and dynamic range, i.e. ability to represent tones, darkest to brightest;
- Number of light sources and how well these sample the hemisphere;
- Fit of basis functions to the actual reflectance distribution function of the surface;
- Spectral resolution, i.e. ability to reconstruct reflectance spectrum from the colour channels.

Substantially better results are obtained from photometric stereo methods, using the same set of images from which the RTI is derived. The technique of taking the median of the distribution of normals obtained from a set of 31 triplets of lamps produced marginally better results than regression over the whole 64-lamp set. Drew *et al.* (2009) noted that better results could be obtained by regression on luminance using Least Median of Squares, automatically identifying outlier pixels as shadows if they are darker than matte and specular highlights if they are brighter than white. By using such a technique optimum surface normals should be expected.

6. REFERENCES

Barbosa J, Sobral JL and Proença AJ (2007) Imaging techniques to simplify the PTM generation of a bas-relief, *Proc. 8th Intl. Symp. Virtual Reality, Archaeology & Cultural Heritage* (VAST'07) 28-31.

Blais F (2004) A review of 20 years of range sensor development, *J. Electronic Imaging*, 13(1) 231-240.

Christensen PH and Shapiro LG (1994) Three-dimensional shape from colour photometric stereo, *Intl. Journal of Computer Vision*, 13(2) 213-227.

Dellepiane M, Corsini M, Callieri M and Scopigno R (2006) High quality PTM acquisition: Reflection transformation imaging for large objects, *Proc. 7th Intl. Symp. on Virtual Reality, Archaeology and Cultural Heritage* (VAST'06).

Drew MS, Hajari N, Hel-Or Y and Malzbender T (2009) Specularity and Shadow Interpolation via Robust Polynomial Texture Maps, *Proc. British Machine Vision Conf. (BMVC'09)*, London.

Freeth T, Bitsakis Y, Moussas X, Seiradakis J, Tselikas A, Mangou H, Zafeiropoulou M, Hadland R, Bate D, Ramsey A, Allen M, Crawley A, Hockley P, Malzbender T, Gelb D, Ambrisco W, Edmunds M (2006) Decoding the ancient Greek astronomical calculator known as the Antikythera Mechanism, *Nature* 444:587-591.

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.

Happa J, Mudge M, Debatista K, Artusi A, Gonçalves A and Chalmers A (2009) Illuminating the Past - State of the Art, *Proc. 10th Intl. Symp. on Virtual Reality, Archaeology and Cultural Heritage*.

Malzbender, T., Gelb, D. and Wolters, H. (2001), 'Polynomial Texture Maps', *Proc. ACM SIGGRAPH*, 28, 519-528.

MacDonald, LW (2010) The Limits of Resolution, *Proc. Conf. on Electronic Visualisation and the Arts (EVA)*, London.

MacDonald LW and Robson S (2010) Polynomial texture mapping and 3D representation, *Proc. ISPRS Commission V Symp. 'Close Range Image Measurement Techniques'*, Newcastle.

Mudge M, Voutaz J-P, Schroer C and Lum M (2005) Reflection transformation imaging and virtual representations of coins from the hospice of the Grand St. Bernard, *Proc. 6th Intl. Symp. on Virtual Reality, Archaeology and Cultural Heritage*.

Mudge M, Malzbender T, Schroer C and Lum M (2006) New reflection transformation imaging methods for rock art and multiple viewpoint display, *Proc. 7th Intl. Symp. on Virtual Reality, Archaeology and Cultural Heritage* (VAST'06).

Ng HS, Wu TP and Tang CK (2010) Surface-from-Gradients without Discrete Integrability Enforcement: a Gaussian Kernel Approach, *IEEE Trans. Patt. Analysis and Machine Intell.*

Willems G, Verbiest F, Moreau W, Hameeuw H, Van Lerberghe K and Van Gool L (2005) Easy and cost-effective cuneiform digitizing, *Proc. 6th Intl. Symp. on Virtual Reality, Archaeology and Cultural Heritage* (VAST'05), 73-80.

Zányi E, Schroer C, Mudge M and Chalmers A (2007) Lighting and byzantine glass tesserae. *Proc. Conf. on Electronic Visualisation and the Arts (EVA)*, London.