3D Reconstruction in an Illumination Dome

Lindsay W. MacDonald
Dept of Civil, Environmental and
Geomatic Engineering
University College London
UK
lindsay.macdonald@ucl.ac.uk

Erica Nocerino
Laboratoire des Sciences de
l'information et des systèmes)
Aix-Marseille Université
France
nocerino@fbk.eu

Stuart Robson
Dept of Civil, Environmental and
Geomatic Engineering
University College London
UK
s.robson@ucl.ac.uk

Mona Hess
Institute of Archaeology, Heritage
Sciences and Art History
Otto-Friedrich University of Bamberg
Germany
mona.hess@uni-bamberg.de

The illumination dome provides a stable environment for photography of cultural heritage objects with multi-directional lighting. Photogrammetry for 3D reconstruction, on the other hand, is usually based on an image set taken under ambient lighting by a hand-held camera moved to many positions around the object. This study investigated whether the photogrammetric image sets could be captured entirely within the dome and compared their accuracy with laser scan data.

Photogrammetry. Structure from motion. Dense point cloud. 3D digitisation. Photometric stereo. Surface normals.

1. INTRODUCTION

The illumination dome at UCL enables sets of images of an object to be captured from a fixed zenithal camera position, with illumination from 64 flash lights at known coordinate positions on the hemisphere (MacDonald et al. 2015). The image sets acquired by this device are used primarily for visualisation of cultural heritage objects by polynomial texture mapping (PTM) (MacDonald and Robson 2010), but they have also proved to be viable for estimation of the surface's angular reflectance distribution function (MacDonald 2014).

The PTM technique provides an ideal means of visualising the surface relief of objects that are predominantly planar with localised detail. Examples are coins, medals, fossils, moulded clay, incised tablets, and impasto paintings. Visualisation through the software viewer provides a compelling illusion of moving a 'virtual torch' above the object, as if it were a real light source being moved above a real 3D surface. But in fact there is no 3D information in the PTM representation, only the coefficients of a biquadratic function of reflected light intensity as a function of angle. The same is true when the image set is fitted by hemispherical harmonics for reflectance transform imaging (RTI).

What is represented is the variation of observed light intensity at each point with angle of incident illumination, not the elevation or gradient of the 3D surface. When 'surface normals' are extracted from PTM for display in the viewer software, they show the direction of the peak of the fitted function, which is not necessarily the same as the geometric vector perpendicular to the surface.

Figure 1: UCL illumination dome for metric photography.

The classic ‘shape from shading’ approach attempts 3D reconstruction from three or more images taken from a single camera position. The problem is that, although high quality surface normals may be extracted, there is insufficient information to determine the scaling factor for the height, and low spatial frequencies are not accurately represented, leading to warping and distortion of the resulting surface. Additional information is always needed for both the lateral scale (dimensions in X and Y) and the depth scale (dimension in Z) which cannot be derived from the projected rays of a static object in the image plane of a static camera (MacDonald 2015). Other problems arise from lens distortion, and the assumption of a Lambertian surface (ignoring gloss and specularity).

Figure 2: Test object – Green faience udjat eye amulet, UCL Petrie Museum UC8504. Dimensions 81x58 mm.

Multiview 3D reconstruction is now well established, from a series of images in which the object is turned relative to the camera. Either the camera is moved by hand around a static object, or the object is turned systematically within the field of view of a static camera, capturing a set of images from many angles of azimuth and elevation (MacDonald et al. 2016). Reconstruction software first detects the corresponding features in each image, and then employs ‘dense matching’ algorithms to fill in the intervening points, producing a point cloud of the object surface. The coordinate values of the points must be scaled, using either a scale bar placed within the scene or from *a priori* knowledge of the dimensions of the object. The disadvantage of the multiview technique is that the point cloud may be noisy, being too dense in some regions, leading to ambiguity, and too sparse in others, causing ‘holes’ in the surface. These may be disguised by triangulation and the fitting of a mesh, but the approximations result in loss of both precision and resolution of surface details.

Ideally, for a flattish surface with relief, one would like to combine the two techniques, to achieve the crispness of the photometric normals and the 3D accuracy of multiview stereo. It would also be good

not to have to remove the camera from its mounting on the dome to take images of the object in a different environment, for example on a tripod or copystand. This study, therefore, investigated whether an ‘intra-dome’ imaging technique would be feasible, keeping the camera fixed in its usual position at the ‘north pole’ of the dome, without changing the lens focus setting, while moving the object systematically in both tilt and orientation inside the illumination dome. It was hoped that the 3D model derived from this image set could inform the 3D reconstruction from the photometric normals obtained from a separate image set of a static object with multi-directional illumination.

For a test object we selected a faience amulet from the Late Period in Ancient Egypt, c.664-332 BC (Figure 2). This had been previously digitised by a 3D colour laser scanner (Hess and Robson 2010) and was featured in the UCL Petrie Museum online 3D gallery (Hess et al. 2016), as shown in Figure 3.

Figure 3: (top) Arius Rover 100 mobile 3D colour laser scanner in use at the UCL Petrie Museum; (bottom) Interactive Web display of the green faience amulet in the online 3D Petrie Museum.

Faience was made from powdered quartz with a vitreous coating, usually producing a translucent glassy surface. It is not pottery, as it contains no clay, but faience is frequently discussed in stylistic surveys of ancient pottery, because objects made in faience are closer to pottery than to ancient

Egyptian glass (Noble 1969). Faience was one of the most favoured materials for amulets, as their shape could easily and inexpensively be moulded. Frequently green and blue shades were chosen, since they were thought to convey the power of life and regeneration. Amulets were carried on the body or worn on the neck or arms, frequently for protection. The magical power of an ancient Egyptian amulet could arise from its shape, decoration, inscription, material, colour, recitations spoken over it, or from other magical acts that were performed with it (Stünkel 2012).

The Eye of Horus (called Udjat or Wadjet) was one of the most common forms of amulet, linked with the Book of the Dead. It represents the healed eye of the god Horus, depicting a combination of a human and a falcon eye (human eye, eyebrow, cosmetic line, and the stylised representations of a falcon's markings underneath its eye), since Horus was often associated with a falcon. In Egyptian mythology the eye of Horus was gouged out by the god Seth, but subsequently restored by the god Thoth. The udjat eye symbolises healing power, regeneration and protection in general (Stünkel 2012). Our chosen test object is double-sided (mirror-image design, with the right-facing form associated with the sun and the left-facing form with the moon), but all imaging in this study was done with the amulet in the 'right-facing' orientation, as shown in Figure 2. A 5mm hole is drilled through the body of the faience, parallel to the object plane, which could have accommodated a string or thong for hanging around the neck or arm (see Figure 17).

For supporting the object and moving it in three dimensions, a 'tilting table' was specially developed, together with a 3D calibration target, as shown in Figure 4. The tilting table consists of an aluminium baseplate of dimensions 100x100 mm, connected by a vertical pillar to a second aluminium plate of the same size. A spring-loaded swivel détente mechanism enables the top plate to be tilted on one axis to angles in the range $\pm 40^\circ$ in increments of 10° . All components were painted matte black and a rubber sheet was affixed to the top plate. A low ledge was fitted on two opposite sides to act as a safeguard to prevent the object from sliding off when the plate was tilted. The 3D calibration target is of the so-called Manhattan design, with 20 vertical rods of various lengths attached to an aluminium baseplate, also of dimensions 100x100 mm. A total of 35 circular retroreflective targets of diameter 1.5 mm were affixed, one on the top of each rod, plus an additional 15 scattered around the baseplate. Four larger targets coded with line segments were placed on the baseplate to assist with automated recognition of the orientation of the target with respect to the camera.

Figure 4: Tilting table and the 'mini-Manhattan' 3D calibration target, with a 15 cm (6 inch) ruler.

2. IMAGE CAPTURE AND PROCESSING

In the course of the study, five image sets were taken by a Nikon D200 camera, mounted on the dome, with a 105mm macro lens set to aperture $f/22$. The focus distance was set to the top surface of the amulet on the tilting mechanism in the horizontal position, and the focus ring of the lens was taped to prevent any change throughout the phases of photography.

Figure 5: (top) Radial distortion of image through 105mm lens; (bottom) distortion vectors over image area.

The first image set was of the mini-Manhattan target, using the four dome flash lights closest to the camera in Tier 5 (approximately 10 degrees off the optical axis) to obtain strong reflections from the target spots. By taking eight orientations of the target at five angles of tilt, 40 images were captured in 8-bit JPG format and were processed by the Vision Measurement System (VMS) software to determine the parameters of the lens distortion model (MacDonald et al. 2016). The overall results (Figure 5) show image shrinkage, with points of the scene mapped to pixel addresses closer to the centre, and points on the diagonals (corners) projected inwards by a relatively greater amount than points on the axes. This produces 'barrel distortion', but the magnitude of 3.3 pixels over the half-width of the image is small in comparison to distortions of typical zoom lenses.

An interesting outcome of the bundle adjustment process in VMS was that the principal distance (PD) of the lens was found to be 129.85 mm, instead of the nominal 105 mm. This may be explained by the lens equation $1/f = 1/u + 1/v$, which for a macro lens at close range gives values of PD longer than one might expect. VMS is effectively calculating v as the focal length plus the lens extension needed to achieve a sharp focus. It is worth noting that the optical construction of the Nikkor 105mm macro causes the effective focal length to vary with focus distance, an effect known as 'focus breathing'. The processing in this case was somewhat compromised by poor edge quality of the small target spots, and targets out of focus at some orientations because of limited depth of field, resulting in a mean target precision of 157 μm .

Figure 6: Albedo and surface normals generated by the bounded regression technique of photometric stereo.

The second set of images was of the amulet lying on the tilting table in the horizontal position, illuminated successively by the 64 flash lights of the dome. These were captured as raw image files (NEF format) and converted by the DCRAW utility to 16-bit linear TIFF files. The inverse of the lens distortion map was applied to all images to correct them geometrically, as if taken through a distortion-free lens. From these corrected images were generated both a PTM file for interactive viewing and images of the albedo and normal vectors (Figure 6), using the 'bounded regression' method described elsewhere (MacDonald 2014). The albedo represents the matte 'body colour', separated from the specular gloss component.

3. PHOTOMETRIC 3D RECONSTRUCTION

The surface normal at every pixel represents the direction of the vector perpendicular to the tangent plane of the surface. From this the gradients in the X and Y axes are easily computed (Figure 7).

Figure 7: Horizontal (left) and vertical (right) gradients computed from the surface normal vectors.

In principle it should be straight-forward to integrate the gradients over the whole surface to obtain the surface elevation, i.e. the height of each point above the ground plane. In practice this is a difficult problem because low spatial frequencies are not properly represented in the sampled data, and the result of a direct integration is frequently distorted with an overall curvature (Figure 8).

Figure 8: Distorted overall form of the 3D reconstruction from direct integration of the surface normals.

Additional information is always needed to calibrate the 3D reconstruction from photometric normals. At minimum this may consist of a small number of coordinate points measured directly from the object with a vertical calliper (MacDonald 2015). An alternative method is to use elevation data derived from the point cloud produced by a 3D scanner (MacDonald et al. 2017). In this case a point cloud was available from the laser scanning of the amulet at the museum (Figure 3). Although its spatial resolution of 10 points/mm (sample pitch of 100 μm) is relatively low, its geometric accuracy is high.

Figure 9: Corresponding details of the amulet eye from the albedo (left) and laser point cloud (right).

The albedo and surface normals (Figure 6) were computed from the second image set at full image resolution, giving a representation with 36 points/mm on the surface of the amulet, i.e. a sample pitch of $27.8\mu\text{m}$. The great improvement in the rendering of fine detail and surface texture relative to the laser scanner is seen in Figure 9.

Figure 10: Height map derived from the laser scanner point cloud: (left) unfilled; (right) filled.

The procedure was to align the 3D point cloud with the horizontal X-Y plane, then to determine the median Z value in each cell on a grid of 10 points per mm. This gave a sparse elevation map for the object, with a pattern of missing points along the sampling path of the scanner beam (Figure 10). An algorithm was written in Matlab to fill each missing point by the mean of its neighbours.

Figure 11: Registration shown by composite image with transformed scanner height in red channel and Z component of photometric normals in green channel.

The scanner elevation map was then registered with the photometric images, which had already been corrected for lens distortion, by rotating and scaling. The composite image in Figure 11 shows the closeness of fit between the two.

Figure 12: Fourier frequency domain maps of $\log(\text{power})$ for (left) scanner height and (right) photometric normals.

Taking the Fourier transform of the two gradient maps produced the $\log(\text{power})$ distributions over the frequency plane, shown in Figure 12. These two were combined by taking the low frequencies (near the central pole) from the laser scanner and the high frequencies (everywhere else) from the camera. Transforming back to the spatial domain gave the 3D form in Figure 13, which lies properly in the horizontal plane, without any distortion.

Figure 13: 3D elevation map of amulet after inverse Fourier transform of combined frequency distributions.

The relationship between the laser surface elevation and the reconstructed 3D surface can be seen from the horizontal cross-section in Figure 14. The height of the latter is everywhere close to that of the laser scanner, but the reconstructed surface has more detail. The mean absolute difference over the whole surface is 0.019 mm. The result was a 3D model with the geometric accuracy of the laser scanner, combined with the fine detail from the photometric surface normals.

Figure 14: Horizontal profile of surface height of laser scan and reconstruction, with detail in blue outline.

4. PHOTOGRAMMETRIC 3D RECONSTRUCTION

A third set of images was taken of the amulet on the tilting table within the dome. In addition to horizontal, the table was tilted to the left to each of the four angles of inclination (10° , 20° , 30° , 40°), at each of which the amulet was rotated to eight angles around its central axis, at increments of approximately 45° . The illumination was from all 28 flash lights in Tiers 3 and 4 of the dome, giving the effect of a 'ring light' that minimised shadows. This resulted in a total of $5 \times 8 = 40$ images.

Figure 15: Masking of background from image: (top left) original image of amulet lying on grey foam; (top right) thresholded image showing outline shadow; (bottom left) edited mask; (bottom right) masked image.

A binary mask was created for each image by thresholding the intensity of the shadow in Matlab, to separate the coloured body of the amulet from the neutral grey foam background, and then pixel editing the mask image to remove unwanted noise (Figure 15). The masked images were processed in Agisoft Photoscan to detect common features and determine the camera orientations, but the results for the lens distortion parameters were anomalous. It seemed that fitting of the distortion model by Photoscan was upset by having the object tilted on one side only, so that the bundle adjustment errors could not be distributed throughout the image. This resulted in a very large offset of the principal point, with -167 pixels in x and -141 in y , producing an asymmetrical distortion pattern (Figure 16).

Figure 16: Lens distortion vectors resulting from fitting by Photoscan of amulet images tilted only to the left.

In an attempt to correct the problem of the asymmetrical lens distortion, a fourth set of images of the amulet was captured for tilt angles of 20° and 40° in the three remaining directions (up, down, right). At each orientation the amulet was turned to eight positions around its centre, at angular increments of approximately 45° . This resulted in a total of $6 \times 8 = 48$ images, each of 3872×2592 pixels, produced in 8-bit JPG format by the camera. To

these were appended the images from the previous set for horizontal and left tilt at 20° and 40° . These 24 images were in 16-bit TIFF format, having been converted by DCRAW from the original NEF files, and so had the raw sensor dimensions of 3900×2616 pixels. Photoscan consequently interpreted these two image subsets as coming from different cameras. Although both had the same lens and the same pixel pitch of $6.10 \mu\text{m}$ in the sensor array, the bundle adjustment led all parameters to differ between the two cameras.

Figure 17: Four images of the amulet on the tilting table, taken within the dome, all at the maximum inclination angle of 40° . The two ends of the hole drilled longitudinally through the amulet body can be seen.

The intention was, while confining the photography within the environment inside the dome, to simulate the images that would be obtained if the camera were moved freely around all sides of a static object in conventional photogrammetric practice. Four views of the amulet, tilted by 40° in each direction, are shown in Figure 17, making the sides of the object visible to the camera.

Figure 18: Image network in Photoscan for 72 images.

The image network generated is shown in Figure 18. The eight images at the pole correspond to the

horizontal directions, then at the eight compass points at each of two co-latitudes (20° and 40°) are four images, corresponding to four directions of tilt. Movements of the object beneath a static camera have been mapped to equivalent movements of a virtual camera around a static object.

No scale bar or ruler was placed alongside the amulet in the images, so for scaling it was necessary to use dimensional information from the original 3D point cloud produced by the laser scanner. Twelve marker points were defined, distributed over the amulet surface (Figure 19). The X,Y,Z coordinates of each corresponding point in the laser scan point cloud were read via the coordinate data picker tool in Pointstream software.

Figure 19: Locations of 12 marker points on amulet.

For each of the 72 images a mask was created as before (see Figure 15), replacing the background by pixel value zero, i.e. black. Then the feature detection/alignment and dense image matching procedures were applied in Photoscan to produce a 3D dense point cloud, containing 5,047,459 points.

The result was curious, because Photoscan, instead of ignoring the masked area around the object, had treated it as part of the object so the result included the mask, giving the impression that the amulet was surrounded by black mastic (Figure 20). The level of the black material was close to that of the amulet's top surface, so that in most places around its perimeter the object's true depth was concealed. After removing all the black points, the amulet point cloud contained 1,627,673 points. Perhaps by using the masking tools provided by Photoscan, this problem would have been avoided.

So here is a conundrum. Photoscan expects that the camera is moving above a static scene in which the object and its background have a fixed relationship. With the object moving on a tilting table beneath a fixed camera there should be a textured background, for example a printed sheet placed on the tilting table under the object, that would move in synchrony with the object. But this was not possible for two reasons: (a) the object

could not have been rotated around its axis: (b) the conservator insisted that the precious ancient amulet could only be in contact with grey museum-grade plastazote foam sheet and not any other material that might cause harm.

Figure 20: (top) Top view of point cloud generated from masked images, showing surrounding black mass; (centre) edge-on view of point cloud; (bottom) close-up detail of right side near the eye.

The results above indicate that preliminary masking of the background area of the image is not a satisfactory approach. To make a good 3D model, it is necessary for the object background to be included in the Photoscan processes of feature identification and dense matching. To ensure that a proper 3D surface could be reconstructed, therefore, a fifth image set was taken with the amulet placed on foam on a marble turntable, on a photographic copystand. Two large fluorescent lights with high-frequency ballasts provided well-diffused illumination. The camera was initially mounted vertically above the turntable, then placed on a tripod to one side. Eight images were taken from above, then sixteen from each of two oblique angles of elevation, still with the lens taped at the same focus setting as in the dome.

The advantage of the turntable was that the light grey marble had a black vein, so provided a natural texture with many features that could be detected by Photoscan. For scaling, 14 marker points on the planar marble surface around the amulet were selected and the coordinates of each measured with callipers. After dense matching the object's top surface and sides were well defined and rose above the ground plane of the turntable (Figure 21). This made it easy subsequently to crop the points representing the turntable by a single cutting plane.

Figure 21: (top) Dense point cloud of marble turntable and amulet, showing locations of 14 markers on turntable (bottom) side view showing elevation above plane.

5. CONCLUSIONS

This study investigated the feasibility of generating 3D models of a cultural heritage object, using only images taken within an illumination dome. A separate dataset was available for reference, namely the point cloud generated by a colour laser scanner. The small tilting table, developed especially for this study, proved to be effective in presenting both the object and a 3D calibration target at a range of orientations and tilt angles. For best results with Photoscan the images should not be masked but should include a textured background, fixed in relation to the object. Also, the tilt angles should be well distributed and not be confined to one side of the image field. We suggest that in future, with automation of the tilting mechanism and a workflow for the image capture and processing stages, this method could facilitate the 3D documentation of museum collections.

6. ACKNOWLEDGEMENT

Sincere thanks to Dr Anna Garnett, Curator of the UCL Petrie Museum, for kindly making available the amulet on several occasions for photography.

7. REFERENCES

- Hess, M., Serpico, M., Amati, G., Pridden, I., Nelson, T. and Robson, S. (2016) Developing 3D imaging programmes – workflow and quality control. *Journal of Computing in Cultural Heritage (JOCCH)*, 9(1):1–11. DOI: [10.1145/2786760](https://doi.org/10.1145/2786760) See also: <http://www.ucl.ac.uk/culture/projects/3d-petrie> and <http://www.ucl.ac.uk/3dpetriemuseum> (retrieved 11 April 2018).
- Hess, M. and Robson, S. (2010) 3D colour imaging for cultural heritage artefacts. *Proc. ISPRS Commission V Symp. 'Close Range Image Measurement Techniques'*, Newcastle. Vol. 38(Part 5), pp.288–292.
- MacDonald, L.W. and Robson, S. (2010) Polynomial texture mapping and 3D representation, *Proc. ISPRS Commission V Symp. 'Close Range Image Measurement Techniques'*, Newcastle, Vol. 38(Part 5), pp.422–427.
- MacDonald, L.W. (2014) Colour and Directionality in Surface Reflectance, *Proc. Conf. on Artificial Intelligence and Simulation of Behaviour (AISB)*, Goldsmiths College, pp.223–229.
- MacDonald, L.W. (2015) Surface Reconstruction from Photometric Normals with Reference Height Measurements, *Proc. SPIE Conf. on Optics for Arts, Architecture and Archaeology V*, Vol. 9527-6, Munich.
- MacDonald, L.W., Hosseiniaveh Ahmadabadian, A. and Robson, S. (2015) Determining the Coordinates of Lamps in an Illumination Dome, *Proc. SPIE Conf. on Videometrics, Range Imaging, and Applications XIII*, Vol. 9528–18, Munich.
- MacDonald, L.W., Toschi, I., Nocerino, E., Hess, M., Remondino, F. and Robson, S. (2016) Accuracy of 3D Reconstruction in an Illumination Dome, *Proc. ISPRS XXIII Congress*, Prague, July.
- MacDonald, L.W., Moitinho de Almeida, V. and Hess, M. (2017) Three-Dimensional Reconstruction of Roman Coins from Photometric Image Sets. *Journal of Electronic Imaging*, 26(1).
- Noble, J.V. (1969). The technique of Egyptian faience. *American Journal of Archaeology*, 73(4):435–439.
- Stünkel, I. (2012) 'Amulets, Egypt'. In *The Encyclopedia of Ancient History*, John Wiley, NY. DOI: [10.1002/9781444338386.wbeah01531](https://doi.org/10.1002/9781444338386.wbeah01531)