Edge Detection on Polynomial Texture Maps

Cristian Brognara[1], Massimiliano Corsini[2], Matteo Dellepiane[2], and Andrea Giachetti[1]

[1] Dip. Informatica, Università di Verona, Strada le Grazie 15 - 37134 Verona, Italy
andrea.giachetti@univr.it,
WWW home page:
[2] Visual Computing Laboratory, ISTI-CNR, 56124 Pisa, Italy

Abstract. In this paper we propose a simple method to extract edges from Polynomial Texture Maps (PTM) or other kinds of Reflection Transformation Image (RTI) files. It is based on the idea of following 2D lines where the variation of corresponding 3D normals computed from the PTM coefficients is maximal. Normals are estimated using a photometric stereo approach, derivatives along image axes directions are computed in a multiscale framework providing normal discontinuity and orientation maps and lines are finally extracted using non-maxima suppression and hysteresis thresholds as in Canny's algorithm. In this way it is possible to discover automatically potential structure of interest (inscriptions, small reliefs) on Cultural Heritage artifacts of interest without the necessity of interactively recreating images using different light directions. Experimental results obtained on test data and new PTMs acquired in an archaeological site in the Holy Land with a simple low-end camera, show that the method provides potentially useful results.

Keywords: Polynomial Texture Maps, Cultural Heritage, Edge Detection

1 Introduction

Polynomial Texture Maps [7] are an extremely useful tool for the documentation and the visual analysis of ancient coins, bas-reliefs, paintings and many other Cultural Heritage objects. They are relightable image, i.e. image where the user can modify interactively the lighting conditions. Each pixel contains a bi-quadratic polynomial that encodes an approximate reflectance function of the scene allowing the possibility modify the image given the illumination direction. PTMs of an object can be created from multiple images acquired under different incident light directions, without specific hardware (low-end digital camera provides enough resolution to produce good PTMs, and almost any type of light source can be used). Several applications of this technique have been proposed, mainly in the field of Cultural Heritage [9, 2, 3], where this type of image and other similar ones [5] have demonstrated to be very useful for analysis purposes.

PTM data can be visualized with ad-hoc software packages, that we describe in the next, allowing the interactive relighting of the scene that can be used to

find the best way to discover relevant information. Specific methods to enhance structures using multiple light information have been proposed as well [10].

2 Polynomial Texture Maps: acquisition and interpretation

A PTM can be generated using different light directions sampling. From this set of photos, the coefficients $(a_0, \ldots, a_5)$ of a bi-quadratic polynomial expressing the reflectance properties of the image can be estimated. This polynomial is in the form:

$$\begin{aligned} L(l_u, l_v, x, y) &= a_o(x, y) + a_1(x, y)l_u + a_2(x, y)l_v + \\ &\quad + a_3(x, y)l_u l_v + a_4(x, y)l_u^2 + a_5(x, y)l_v^2 \\ R_f(l_u, l_v, x, y) &= L(l_u, l_v, x, y)R(x, y) \\ G_f(l_u, l_v, x, y) &= L(l_u, l_v, x, y)G(x, y) \\ B_f(l_u, l_v, x, y) &= L(l_u, l_v, x, y)B(x, y) \end{aligned} \tag{1}$$

where (l_u, l_v) is the light direction vector (normalized) projected on the image plane (see Fig. 1). Since this vector is normalized the third component is redundant ($l_z = \sqrt{l_u^2 + l_v^2}$). (R_f, G_f, B_f) is the final color for the given pixel and the given light direction. This type of PTM is called LRGB PTM since the "luminance" term modulates the RGB color channels. In the case of an RGB PTM each channel has its own coefficients, for a total of 18 coefficients per pixel. We refer to $L(l_u, l_v, x, y)$ as "luminance" but, more precisely, this term represents the reflectance functions with the self-shadowing effects embedded. Other basis can be substituted for PTM to better approximate materials characterized by more complex reflectance behavior, such as gold or marble [8].

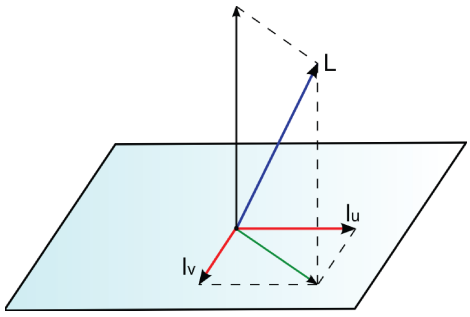

Fig. 1. Projection of light vector on the image plane.

Hewlett-Packard labs, where this technology has been developed, freely distribute software tools for the creation and the interactive visualization of PTM data as well as test data, available at the web site *http://www.hpl.hp.com/research/ptm*.

Using the HP tool PTM files can be created from sets of images acquired with standard cameras, provided that the direction of the light is known. A trick to evaluate this direction directly from the acquired images is to include a reflective sphere near the object allowing the direction estimation (l_u, l_v) from the specular highlight. For a correct creation of PTMs of object of interest, therefore, it is not necessary to have an expensive setup: for the acquisition we made in the archaeological site in the Holy Land, we used simply:

- A low-end digital camera (Fuji FINEPIX HS 20 EXR)
- Two tripods, one for the camera and one for the black reflective ball
- One lamp able to generate approximately uniform light on the size of the objects of interest.

Even if the requirements are simple, the acquisition protocol should be, especially for large objects, carefully planned, in order to produce high quality PTMs. In [2] the problem of optimizing acquisition parameters for medium-large objects is discussed. In our work, we acquire a set of large planar stones. For this case we developed and applied the following protocol:

- Take the measures of the object of interest, find its center and its distance from the ground.
- Put the camera on the tripod at a distance of approximately 5/3 of the side of the object from its center
- Measure aperture and shutter speed under the illumination of the central light. Keep these values fixed for all the photos, in order to have a constant exposure.
- Take 40-60 photos moving the light source in a way to attempt to sample uniformly the hemisphere

In order to calculate a precise illumination function, a critical factor is that the digital camera must not move from one photo to the other. Even a misalignment of a few pixel can produce a bad result, with visible aliasing.

3 Edge Detection on PTMs

The analysis of PTM data is usually done with visualization packages allowing the interactive generation of relighted images that optimally put in evidence the structures of interest in the artifacts. This procedure, however, can be sometime difficult and time consuming, especially for large images and when it is not known in advance what should be searched exactly in the images. It could be useful, therefore, to use some automatic processing tools able to extract relevant information and suggest, where structures of interest are located. Filters for the enhanced visualization of this type of images have been proposed in [10] and implemented in a specific tool (*RTIViewer*) available at the web site *http://culturalheritageimaging.org*.

The visualization enhancement is possible thanks to the information provided by the per-pixel bi-quadratic polynomial. Here, we propose an alternative way

to help the experts in the artifacts' examination, i.e. not to try to enhance the rendering, but to extract specific information from data.

In ancient tablets or stones, structures of interest are human artifacts and inscriptions that are usually composed by a set of lines. The classical image processing tools used to search lines in images are edge detectors, trying to create lines joining pixels with high values of image gradient. Edges should highlight the important structures in the image rejecting noise and negligible information. We can do something similar for the PTM image data: instead of processing relighted images trying to find optimal visualization, we can simply extract edges from the PTM coefficients and use them to interpret more easily the data acquired.

The detection of depth discontinuities from multi-flash images has been proposed in [11], where authors look for pixels where the difference between baseline and specific illumination is high, and use the results to obtain nice non-photorealistic rendering able to clearly separate objects from background. Our approach is different as well as the application: we used, in fact, classical image processing techniques able to follow continuous lines possibly belonging to artifacts or inscriptions by exploiting PTM coefficients instead of by differentiating the multi-flash images.

Edge detection methods working on standard images first compute a color discontinuity map, e.g. the gradient magnitude, then try to follow its directionally maximal lines to trace edge lines. We propose to replace here the gradient magnitude with a measure of the local variation of the normal vectors that can be estimated at pixel locations from the PTM coefficients. In PTM images, in fact, for each pixel the local normal can be estimated considering the light direction that maximizes the value of the polynomial $L(u, v, l_u, l_v)$. The maximum can be obtained by finding the zero of the spatial derivatives as in [7]. By solving the system:

$$\partial L/\partial u = \partial L/\partial v = 0 \tag{2}$$

it is possible to obtain first the projections of surface normal (n_u, n_v)

$$n_u = \frac{a_2 a_4 - 2a_1 a_3}{4a_0 a_1 - a_2^2} \quad n_v = \frac{a_2 a_3 - 2a_0 a_4}{4a_0 a_1 - a_2^2} \tag{3}$$

and then the full normal as

$$\boldsymbol{N} = (n_u, n_v, \sqrt{1 - n_u^2 - n : v^2}) \tag{4}$$

This normal estimation, however, can be not accurate in certain cases, and tends to be oversmoothed [6]. To reduce these problems we propose to compute normals using a typical photometric stereo [12] approach that consists in assuming the material Lambertian, i.e. a pure diffusive material, and setup a linear system in the following way:

$$\begin{gathered} N_u l_{u,1} + N_v l_{v,1} + N_z l_{z,1} = L_1 \\ N_u l_{u,2} + N_v l_{v,2} + N_z l_{z,2} = L_2 \\ \cdots \\ N_u l_{u,n} + N_v l_{v,n} + N_z l_{z,n} = L_n \end{gathered} \tag{5}$$

where (N_x, N_y, N_z) is the unknown normal for the specific pixel, (l_x, i, l_y, i, l_z, i) is a light direction and L_i is the value of the polynomial for that direction. The number of directions here used is 32. The set of directions is obtained by sampling the polar coordinates that identify a direction (elevation and azimuth angles) by step of 20 degree starting from 10 to 70, for the elevation, and subdividing the azimuth in 8 parts. In this way the sampling is not uniform (more samples at the pole), but the relatively high number of samples provides a good estimation of the real surface normal. The estimated normal differs from the real one as the material deviates by a Lambertian reflector. The linear system of equation (5) can be solved by using a Singular Value Decomposition (SVD) approach which consists in decomposing the matrix A into three sub-matrix that can be combined to obtain a robust least square solution of the system.

From the normal vector $\boldsymbol{N}$ we can estimate the normal discontinuity along u and v directions (at different scales). If we consider directional derivatives of normal vectors $\partial\boldsymbol{N}/\partial u, \partial\boldsymbol{N}/\partial v$ we can obtain the intensity of the discontinuity as

$$I = \left(\frac{\partial \boldsymbol{N}}{\partial u}\right)^2 + \left(\frac{\partial \boldsymbol{N}}{\partial v}\right)^2 \tag{6}$$

that can be computed using ad hoc masks at different scales. The direction of this edge can be computed as well as

$$\theta = arctan(-\frac{\partial \boldsymbol{N}}{\partial v} \cdot \boldsymbol{e_2} / \frac{\partial \boldsymbol{N}}{\partial u} \cdot \boldsymbol{e_1}) \tag{7}$$

From the maps I (normalized dividing by its maximum) and θ estimated we can then apply the classical steps of Canny edge detection [1] e.g. non-maxima suppression and hysteresis thresholding. The first step checks if the intensity values are maximal perpendicularly to the edge direction and put to zero all the matrix elements that do not meet this condition. Angles representing edge directions are in the implementation rounded at sampled values of 0, 45, 90 and 135 degrees. The result is a list of candidate edge points that are sorted by decreasing values. Edge lines are finally traced connecting edge candidates with hysteresis thresholding, e.g. using two different thresholds. Edge candidate location with the highest intensity is removed from the list and used to start paths if the corresponding value is larger than the highest threshold, then neigboring candidates (that are as well removed from the list) are joined until their values are higher than the second threshold. The procedure is repeated until all the possible starting points have been removed.

In our implementation we computed normal derivatives using 5-pixel masks, and supported a multiscale estimation of the derivatives/discontinuity as follows: derivatives at different levels of detail are obtained using iterative smoothing and taking derivative masks with increased sampling steps. The final multiscale I and are obtained selecting for each pixel position the value corresponding to the scale maximizing I. Results, as in the usual edge detectors, clearly depend on the parameters chosen (thresholds) and on a possible initial normal smoothing (not necessary however, being the normal maps derived by the PTM coefficients

already smoothed). However, for our test a higher threshold of 0.2-0.3 and a smaller one of about 0.05 were able to provide reasonable results on the tested images.

4 Experimental results

The normal maps computed with the proposed technique appear sufficiently accurate to capture the fine structures of the PTM image.
Fig. 2 shows the results obtained on a sample PTM downloaded from the HP site representing an ancient tablet from the Archaeological Research Collection of the University of Southern California. Fig. 2 A shows a single relighting of the data, Fig. 2 B the normal map obtained with our photometric stereo approach and Fig. 2 the normal discontinuity map I computed with two-scale derivatives. The result shows clearly the inscriptions. It is also interesting to compare the edge map obtained applying non maxima suppression and hysteresis thresholding on the previous map, compared with a similar edge detection performed on a single relighted image. It is possible to see that the classical edge detection as well as the gradient map does not highlight correctly the structures of interest, while the PTM based method does 3

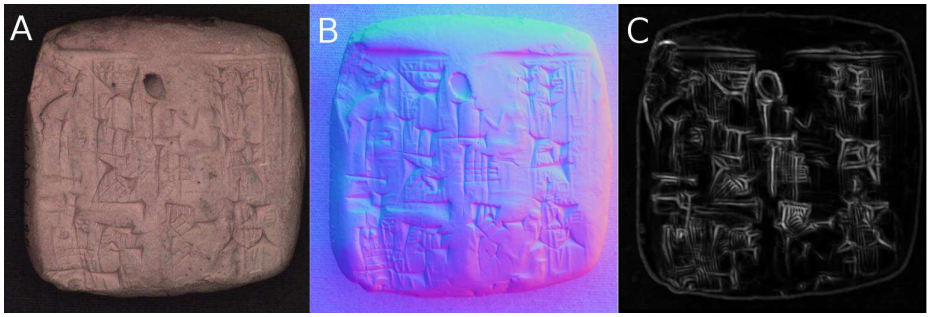

Fig. 2. A: A relighted images from a demo PTM acquisition. B Color coded normal map estimated with the photometric stereo approach. C: Normal discontinuity measure, computed with a 2-scale derivative estimation.

Similar results can be obtained also on our on-field acquisitions on the Holy Land site. Fig. 4 shows an image created by relighting a PTM created with the acquired images. It represents a large stone wall in a tomb, and it is difficult to determine if something interesting can be found in it using the default frontal illumination used.

Looking at small details and changing the light direction something interesting can be, however found. Fig. 5 A shows a central detail on the image of Fig.4, where it is possible to find a written text. To find it automatically, without the

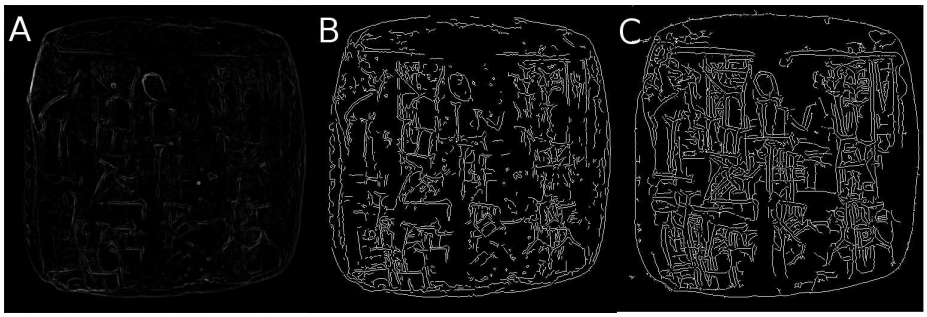

Fig. 3. A: Gradient magnitude computed on a single relighted image. B Corresponding edges detected with the Canny approach. C Edge detection with non maxima suppression and hysteresis thresholding applied to the normal discontinuity map.

necessity of searching manually the optimal light direction, it is possible to try an enhancement filter like those implemented in RTI viewer.

Fig. 5 B shows, for example, the output of the coefficient unsharp masking filter, while Fig.5 C show the result of the static multi-light enhancement, where the inscription is more visible.

Extracting normal discontinuities, however, the fact that an inscription is present in that region becomes clearer. Fig. 6 A shows the normal discontinuity map compute at the finest scale, and Fig. 6 B the extracted edges. It is possible to see lines belonging to a text.

Fig. 4. Relighted PTM acquired on site in an archaeological site in the Holy Land. The quality of the PTMs computed is rather good despite the simple and low-cost acquisition setup.

Fig. 5. Detail of the previous image where it is possible to find an inscription. It is not visible using a default light direction (A) and still not visible using filters like unsharp masking on coefficients (B). The static multi-light enhancement makes the lines of interest partially visible.

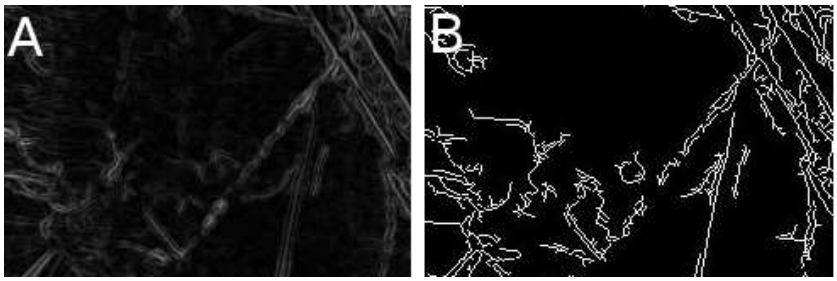

Fig. 6. The normal discontinuity map on the previous detail evidentiates text lines (A). The subsequent edge tracing (B) partially detects the lines of interest.

Fig. 7 shows another example of edge extraction revealing the location of inscriptions. In Fig. 7 A a frontal relighting shows the lid of a tomb, where it is not easy to see where human artifacts are present. Fig. 7 B shows the normal discontinuity map extracted using derivatives at two scales. Fig. 7 C the edge extracted where it is possible to see squares and lines traced in the stone. Fig. 8 show another part of the archaeological site where the normal discontinuity map and the edges reveals artifacts and text fragments.

A possible improvement of the method could consist in storing local information about the normals and estimating the position of lights enhancing the lines of interest that could be combined creating an adaptively multiple source relighted image.

5 Discussion

Reflection Transform Imaging (RTI) techniques encode a per-pixel parametric approximation of the reflectance behavior of the object depicted, allowing the creation of rendered images with varying lighting of the scene of interest. RTIs are becoming a popular tool to study archaeological sites, stones and artifacts. The analysis of the acquired data is usually done interactively by testing the effects of different light sources on the rendered image. Automatic tools able, for example,

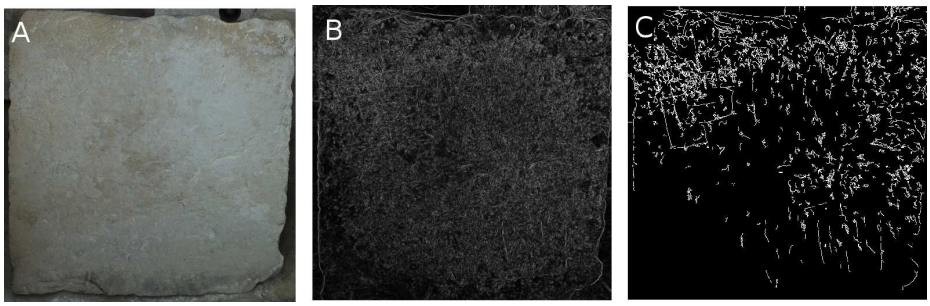

Fig. 7. A: Frontal relighting of a square lid of a tomb. B normal discontinuity map extracted using derivatives at two scales. C: edges extracted after hysteresis thresholding.

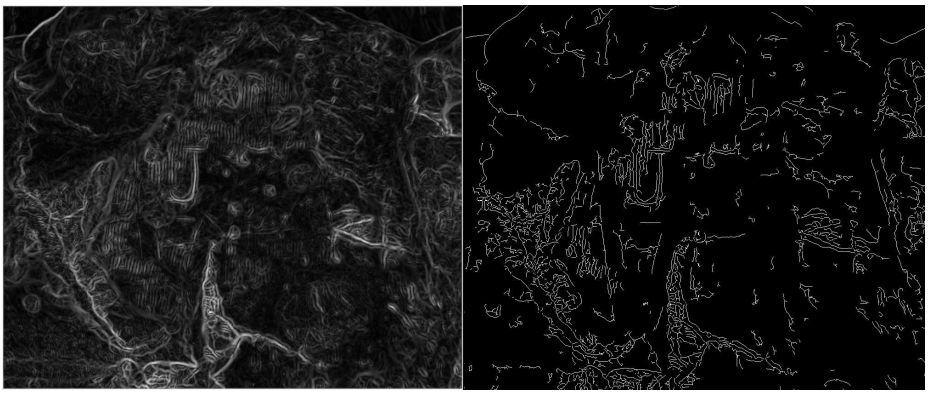

Fig. 8. Left: Normal discontinuity map showing another inscription on a captured stone wall. Right: extracted edge lines.

to identify automatically regions of interests in the images could be extremely important to help archaeological study of large collections of data. In this paper we proposed a simple edge detector applied to PTM data (a type of RTI) that can be used to enhance lines at different scales, possibly corresponding to human artifacts. Computing 3D surface normals discontinuity and tracing discontinuity profiles on the 2D image plane using a classical edge detection approach, we found that is possible to capture relevant information on the acquired data. Normals are computed using a photometric stereo approach on relighted images and not directly derived from the PTM coefficients, allowing a more precise result.

Lines characterizing inscriptions and artifacts have been successfully extracted on PTM acquired by us on an archaeological site with a simple and cheap setup.

The technique has been here applied to Polynomial Texture Maps, but can be applied as well on different types of RTI [4] which provide a better approximation of the per-pixel reflectance function. We expect to obtain, in this case,

improved results. We plan also to extend the idea of automatic processing of RTI data, testing other types of computer vision/pattern recognition methods. For example unsupervised and supervised classification could be applied using the local reflectance coefficients as feature vectors in order to detect regions corresponding to different materials. Pattern recognition tools could be applied as well to recognize characters from lines obtained with the proposed edge detector.

References

1. J Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, June 1986.
2. Matteo Dellepiane, Massimiliano Corsini, Marco Callieri, and Roberto Scopigno. High quality ptm acquisition: Reflection transformation imaging for large objects. In *Proceedings of 7th International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage (VAST2006). The Eurographics Association*, pages 179–186, 2006.
3. Graeme Earl, Kirk Martinez, and Tom Malzbender. Archaeological applications of polynomial texture mapping: analysis, conservation and representation. *Journal of Archaeological Science*, 37(8):2040–2050, 2010.
4. Prabath Gunawardane, Oliver Wang, Steve Scher, James Davis, Ian Rickard, and Tom Malzbender. Optimized image sampling for view and light interpolation. In *VAST 2009: 10th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*. Faculty of ICT, University of Malta, 2009.
5. Jassim Happa, Mark Mudge, Kurt Debattista, Alessandro Artusi, Alexandrino Gonçalves, and Alan Chalmers. Illuminating the past: state of the art. *Virtual reality*, 14(3):155–182, 2010.
6. Lindsay MacDonald and Stuart Robson. Polynomial texture mapping and 3d representations. In *Proc. ISPRS Commission V Symp.Close Range Image Measurement Techniques*, 2010.
7. Tom Malzbender, Dan Gelb, and Hans Wolters. Polynomial texture maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 519–528. ACM, 2001.
8. Mark Mudge, Tom Malzbender, Alan Chalmers, Roberto Scopigno, James Davis, Oliver Wang, Prabath Gunawardane, Michael Ashley, Martin Doerr, Alberto Proenca, and Joao Barbosa. Image-Based Empirical Information Acquisition, Scientific Reliability, and Long-Term Digital Preservation for the Natural Sciences and Cultural Heritage . pages –, Crete, Greece, 2008. Eurographics Association.
9. Joseph Padfield, David Saunders, and Tom Malzbender. Polynomial texture mapping: a new tool for examining the surface of paintings. *ICOM Committee for Conservation*, 1:504–510, 2005.
10. Gianpaolo Palma, Massimiliano Corsini, Paolo Cignoni, Roberto Scopigno, and Mark Mudge. Dynamic shading enhancement for reflectance transformation imaging. *Journal on Computing and Cultural Heritage (JOCCH)*, 3(2):6, 2010.
11. Ramesh Raskar, Kar-Han Tan, Rogerio Feris, Jingyi Yu, and Matthew Turk. Non-photorealistic camera: depth edge detection and stylized rendering using multi-flash imaging. In *ACM Transactions on Graphics (TOG)*, volume 23, pages 679–688. ACM, 2004.
12. Robert J. Woodham. Shape from shading. chapter Photometric method for determining surface orientation from multiple images, pages 513–531. MIT Press, Cambridge, MA, USA, 1989.