

Improving photometric stereo through per-pixel light vector calculation

Jahanzeb Ahmad
jahanzeb.ahmad@uwe.ac.uk

Jiuai Sun
jiuai2.sun@uwe.ac.uk

Lyndon Smith
lyndon.smith@uwe.ac.uk

Melvyn Smith
melvyn.smith@uwe.ac.uk

Centre for Machine Vision,
Bristol Robotics Laboratory,
University of the West of England,
Bristol, UK.
www.uwe.ac.uk/et/mvl

Abstract

Photometric Stereo can recover dense (at pixel-level) local surface orientation, but the subsequent reconstruction procedures used to obtain 3D shape of the scene are prone to low frequency geometric distortion. This geometric distortion is mainly due to assumption of collimated light source, to overcome the error due to collimated light source we propose a novel calibration process to dynamically calculate light vectors for each pixel with little additional computation cost. We calculate distance of object from camera using the Lambertian Diffused Maxima (LDM) region, from which the corrected light vector per-pixel is derived and the absolute dimensions of the object can subsequently be estimated. Experiments performed on synthetic as well as real data show the proposed approach offers improved performance, achieving a reduction in the estimated surface normal error of almost 2 degrees.

1 Introduction and Related Work

Photometric stereo (PS) has been extensively used in many applications [2, 3, 10, 11, 12, 13, 14, 15], especially for estimating high density local surface orientation in the fields of computer vision and computer graphics. It recovers the surface of an object using several images taken from same view point but under different lighting conditions. However the 3D reconstructions from the recovered surface orientation are prone to low frequency geometric distortion because the real illumination is unable to satisfy the assumed, ideal conditions under which PS works.

Light sources in PS are normally assumed to be at infinite distance from the scene so that a homogeneous and parallel incident light. In reality it is not always possible to produce parallel incident light. Any underestimation or misalignment of the illumination may produce some error during recovery of the surface orientation. For example, a 1% uncertainty in the intensity estimation will cause a 0.5-3.5 degree deviation in the calculated surface normal for a typical three-light source photometric stereo setup [2].

A practical solution is to set the light sources far away from the object [1], so that the light can be approximated as a distant radiation point source. This strategy may help to provide evenly distributed radiance across the object surface, but it sacrifices the majority of the illumination intensity, and correspondingly decreases the signal/noise ratio of the whole system. In addition, such a distant lighting setup usually means a large impractical working space is required. So this approach is only suitable for those light sources able to produce high levels of energy and those applications where a large redundant space is available. In terms of the availability and flexibility of current commercial illumination, the distant illumination solution is often not an optimal choice.

A nearby light source model has been considered as an alternative by Kim [2] and Iwahori [3] to reduce the photometric stereo problem to find a local depth solution using a single non-linear equation. By distributed the light sources symmetrically in a plane perpendicular to camera optical axis, they were able to get a unique solution of non-linear equations. However, selection of initial values for the optimisation process and limitations in the speed for solving non-linear equation are the main problems with this method.

Kozera and Noakes introduced an iterative 2D Leap-Frog algorithm able to solve the noisy and non-distant illumination issue for three light-source photometric stereo [4]. Because distributed illuminators are commercially available, Smith *et al.* approximated two symmetrically distributed nearby point sources as one virtual distant point light source for their dynamic photometric stereo method [5]. Unfortunately, none of these methods lend themselves to a generalized approach.

Varnavas *et al.* [6] implemented parallel CUDA based architecture and computed light vectors at each pixel, so that a changing light direction was taken into account. However in practice the whole surface is not necessarily at the same distance from the light source, especially when the size of the object is comparable to the distance of the light source.

Furthermore PS gives no information concerning the absolute distance of the object from the camera. Another additional imaging modality is normally required for obtaining the range data, for example laser triangulation or stereo vision techniques have been combined with the PS approach [7, 8, 9, 10, 11].

In this paper we present a novel method for calculating the distance of an object using the photometric stereo imaging setup and then use this additional information to improve accuracy of surface normal estimation by calculating per-pixel light direction rather assuming same light direction on every pixel. Using only one camera and four lights without a requirement for any additional hardware, and with only little extra processing cost, the object's distance from the camera is estimated. The object's distance from the camera is estimated by finding the Lambertian Diffused Maxima (LDM), a small patch on the object surface whose normal is pointing towards the light source [12, 13]. From this the estimated distance is used to calculate the light vectors at every image pixel thereby minimizing the error associated with assuming a collimated light source. This allows the photometric stereo method to work with real light sources, on Lambertian surfaces that have at least one small patch with normal vectors pointing directly towards the light.

2 Photometric Stereo

Photometric stereo was first introduced by Woodham in 1980 [14]. It recovers the surface shape of the object or scene by taking several images from the same view point but under different lighting conditions. Light sources are some distance away from the scene with

different directions. Each pixel at the same location within all the images is assumed to correspond to the same object point so there is no need to match features between images.

According to the Lambertian reflectance model the intensity I of light reflected from an object's surface is dependent on the surface albedo ρ and the cosine of the angle of the incident light as described in Equation 1. The cosine of the incident angle can also be referred as dot product of the unit vector of the surface normal \vec{N} and the unit vector of light source direction \vec{L} , as shown in Equation 2.

$$I = \rho \cos(\phi_i) \quad (1)$$

$$I = \rho(\vec{L} \cdot \vec{N}) \quad (2)$$

When more than two images (four images are used in the following work) from same view point are available under different lighting conditions, we have a linear set of Equation 1 and 2 and this can be represented in vector form as shown in Equation 3.

$$\vec{T}(x, y) = \rho(x, y)[L]\vec{N}(x, y) \quad (3)$$

\vec{T} is the vector formed by the four pixels $((I^1(x, y), I^2(x, y), I^3(x, y), I^4(x, y)))$ from four images, $[L]$ is the matrix composed by the light vectors $(\vec{L}^1, \vec{L}^2, \vec{L}^3, \vec{L}^4)$. Where, 1, 2, 3 and 4 is the number with respect to the individual light source direction. $[L]$ is not a square and so not invertible, but the least square method can be used to compute Pseudo-Inverse and local surface gradients $p(x, y)$ and $q(x, y)$, and the local surface normal $N(x, y)$ can be calculated from the Pseudo-Inverse using Equations 4, 5 and 6 where $\vec{M}(x, y) = (m_1(x, y), m_2(x, y), m_3(x, y))$.

$$\vec{M}(x, y) = \rho(x, y)N(x, y) = ([L]^T[L])^{-1}[L]^T\vec{T}(x, y) \quad (4)$$

$$p(x, y) = \frac{m_1(x, y)}{m_3(x, y)}, q(x, y) = \frac{m_2(x, y)}{m_3(x, y)} \quad (5)$$

$$N(x, y) = \frac{p(x, y), q(x, y), 1}{\sqrt{p(x, y)^2 + q(x, y)^2 + 1}} \quad (6)$$

$$\rho(x, y) = \sqrt{m_1^2(x, y) + m_2^2(x, y) + m_3^2(x, y)} \quad (7)$$

3 Proposed Method

By estimating the distance of the object from the camera we can improve the accuracy of the surface normals by calculating the light vector of every pixel based on its distance from the camera and light source. The proposed method is divided into three parts: "Light Source Position estimation", "Object distance estimatio" and "per pixel light direction calculation". Light source position estimation is required only once during the rig calibration process.

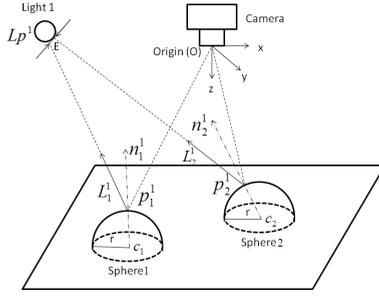


Figure 1: Calibration Setup for Light Position Calculation and Initial (Pseudo) Light Vector Calculation.

3.1 Light source position estimation

The general assumption that the light vector is the same at every point (pixel) is mostly not true in practice, so subsequently we use the intersection of at least two light vectors (calculated at different positions) to obtain the position of light in real world coordinate system. A specular sphere is used to calculate the light vectors at several (we take two as example) different locations in the imaging area. The intersection of these light vectors is taken as the position of the light in the real world coordinate system. The position of light 1 is calculated by finding the intersection point of light vectors \vec{L}_1^1 and \vec{L}_2^1 as shown in Figure 1. \vec{L}_1^1 is the light vector calculated at a sphere surface position p_1^1 by placing the sphere at one random location and \vec{L}_2^1 is the light vector calculated at a sphere surface position p_2^1 by placing the sphere at another random location in the imaging area. To calculate \vec{L}_1^1 and \vec{L}_2^1 Equation 8 is used.

$$\vec{L} = 2(\vec{n} \cdot \vec{d})\vec{n} - \vec{d} \quad (8)$$

Where \vec{d} is reflection direction taken as $(0,0,1)$, \vec{n} is unit surface normal at point p_1^1 or p_2^1 , $\vec{n} = (nx, ny, nz)$, $nx = px - cx$, $ny = py - cy$ and $n_z = \sqrt{(r^2 - n_x^2 - n_y^2)}$, (cx, cy) and (px, py) are the pixel coordinates of the optical centre and the highlight on the sphere respectively, and r is the radius of sphere in the image plane.

The intersection of \vec{L}_1^1 and \vec{L}_2^1 can be calculated using equations 9, 10 and 11 [8]

$$Lp_1^1 = p_1^1 + \left(\frac{(\vec{L}_2^1 \times (p_1^1 - p_2^1)) \cdot (\vec{L}_1^1 \times \vec{L}_2^1)}{(\vec{L}_1^1 \times \vec{L}_2^1) \cdot (\vec{L}_1^1 \times \vec{L}_2^1)} \right) \times \vec{L}_1^1 \quad (9)$$

$$Lp_2^1 = p_2^1 + \left(\frac{(\vec{L}_1^1 \times (p_1^1 - p_2^1)) \cdot (\vec{L}_1^1 \times \vec{L}_2^1)}{(\vec{L}_1^1 \times \vec{L}_2^1) \cdot (\vec{L}_1^1 \times \vec{L}_2^1)} \right) \times \vec{L}_2^1 \quad (10)$$

$$Lp^1 = \frac{Lp_1^1 + Lp_2^1}{2} \quad (11)$$

$$E = |Lp_1^1 - Lp_2^1| \quad (12)$$

Lp^1 is the 3D position of light 1 in the world coordinate system. Lp_1^1 is the point on vector \vec{L}_1^1 closest to \vec{L}_2^1 , Lp_2^1 is the point on vector \vec{L}_2^1 closest to \vec{L}_1^1 , E is the distance between these two points - which can be used to measure the accuracy of the calculation. If E is zero then both light vectors intersect. However, due to error in estimating the light vector, the position of the highlight or sphere centre is not always zero or close to zero. So we use a threshold to establish when the estimated light position is not accurate. In this case the sphere can be positioned in additional places to improve the accuracy.

To calculate the position of light using the above method we need the position of at least two highlights on the sphere surface. As the actual size of the sphere, focal length of the camera and physical pixel size of camera sensor are known, we can find the position of the centre of the sphere in the world coordinate system.

$$c(X, Y, Z) = \left[\frac{-x}{f_x} Z, \frac{-y}{f_y} Z, Z \right] \quad (13)$$

$$Z = \frac{focalLength * sphereActualRadius}{pixelLength * spherePixelRadius} \quad (14)$$

Where Z is the distance of sphere centre from camera in the z direction, f_x and f_y are the focal length in pixels in x and y direction. Once the centre of sphere c is known, the surface normal \vec{n} at point p (highlight pixel position) can be used to calculate p from equation 15.

$$p(X, Y, Z) = c(X, Y, Z) + k * n(X, Y, Z) \quad (15)$$

k is a constant required to calculate p . As p lies on the surface of the sphere so $|p - c|$ should be equal to the sphere radius and by using value of p from equation 15 we can solve the value of k as $|c + k \vec{n} - c| = sphereActualRadius$ and $|\vec{n}| = 1$ so $k = sphereActualRadius$. Once the value of k is calculated, it can be used in equation 15 to calculate the position of the highlight on the sphere surface in real world coordinates; as shown in equation 16.

$$p(X, Y, Z) = c(X, Y, Z) + sphereActualRadius * \vec{n} \quad (16)$$

3.2 Object Distance Estimation

The object distance from the camera is calculated by using the Lambertian Diffused Maxima (LDM), which is calculated by taking the absolute of the dot product between the pseudo light vector and pseudo surface normal, and then applying a threshold; as shown in equation 17. During experimentation we have found that for most cases the threshold is greater then or equal to 0.9.

$$LDM_i = |\vec{N} \cdot \vec{L}^i| > 0.9 \quad (17)$$

\vec{L}^i is a pseudo light vector for light i and \vec{N} is the pseudo surface normal at each pixel. The pseudo light vector \vec{L}^i is calculated during the calibration process by placing the sphere at the centre of the field of view, it is assumed to be same for every pixel. The centre of the LDM gives us the point where the surface normal and the light vector are approximately aligned. Many LDM(s) can exist on the surface of an object but the region with maximum pixel area is considered to be the best choice. Lights are arranged in a square arrangement as shown in Figure 5 and the dot product of the light vectors with surface normals are shown

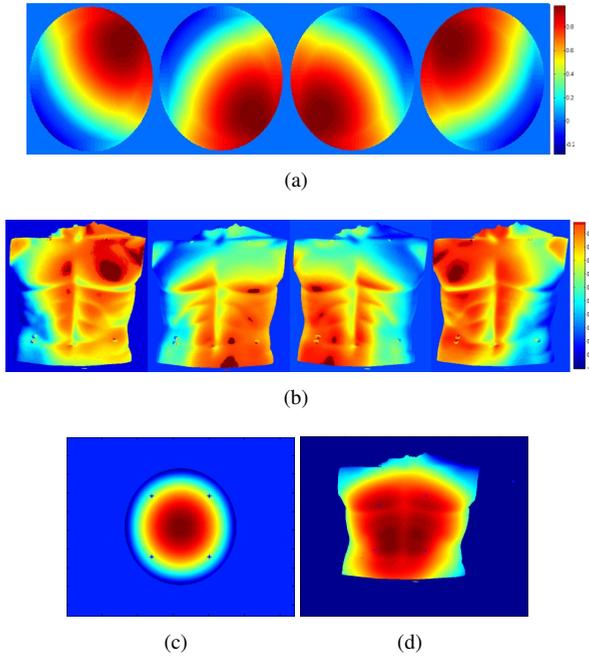


Figure 2: (a) and (b) dot product of image with its light vectors. Diffused maxima regions are in highlighted in dark red colour. (c) and (d) Diffused Maxima Regions centres are plotted on Height Map

in Figure 2. Higher value of dot product means it is more close to diffused maxima. Figure 2 shows the four selected *LDM* centres plotted on a height map of a synthetic sphere and a real human dummy torso.

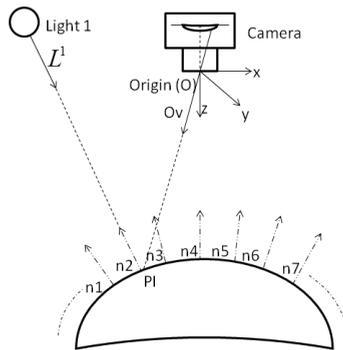


Figure 3: Depth calculation using *LDM* and intersection of vector Ov and L^1 .

Once the *LDM* centre is identified in the image plane, a vector \vec{Ov} can be created from the *LDM* centre to the centre of the lens O . O is also the origin of the world coordinate system as shown in Figure 3 Now by using origin O , position of light L^1 , light vector $\vec{L^1}$ and vector \vec{Ov} , we can determine the intersection point of these two vectors in world coordinates by

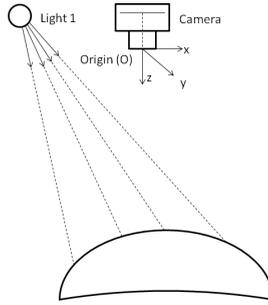


Figure 4: Light vector calculation on each point of object surface.

using equations 9, 10 and 11. The average of the Z coordinate of these points of intersection is the estimated distance of the object from the camera.

3.3 Per pixel light direction calculation

Once the distance of the object is known from the camera, an imaginary plane parallel to the image plane is created. The pseudo height of the object is then defined relative to this plane by adding the reconstructed surface from pseudo normals; so that new light vectors for each pixel point for each light are created. The pseudo height of the object is calculated by integrating [10] the pseudo surface normal N and then scaling the height to compensate for the camera distance.

Traditional photometric stereo assumes that the light direction is same across the whole scene but in reality, particularly where the object has a comparable size to the illumination working distance, it is clear that this varies; as shown in Figure 4. This variation needs to be considered for accurate surface normal calculation because any variations in the illumination position are finally interpreted as uncertainty in recovering the surface normals.

4 Experiments and Results

Experiments were performed on synthetic images as well as with real images. For real images a setup based on a Teledyne DALSA Genie HM1400 1.4 Mega pixel monochrome camera and High power LEDs was designed as shown in Figure 5. A commercial 3dMD [11] system is used to acquire ground truth data as this system has a reported 0.2 mm accuracy in depth measurement. A sphere and dummy human torso are used as the objects.

Figure 6 shows the error (mm) in the calculation of object distance from camera when the initial calibration (pseudo light vectors) of the setup is performed with the specular sphere located approximately at 2000mm from the camera. The $\sim \pm 20$ mm uncertainty is found when the object is moved from 1800mm to 2200mm from the camera. This is relatively high compared to other 3D range finding technologies, however the system can achieve a recovery in pixel level which is not provided by any other 3D imaging systems.

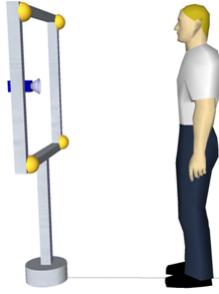


Figure 5: Image acquisition Setup.

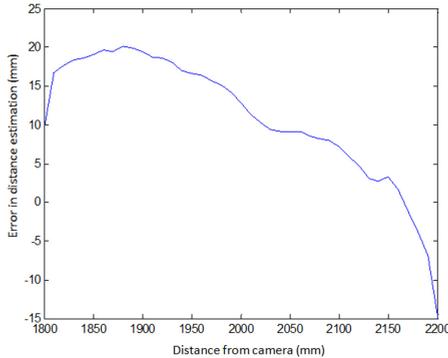


Figure 6: Absolute Error in distance estimation from camera to object.

To test the accuracy of the surface normals acquired from the proposed method we have used Mean Angular Error (MAE) as the measure of accuracy. MAE is calculated by taking the cosine inverse of the dot product of a ground truth surface normal and a calculated surface normal. Table 1 summaries the Mean Angular error calculated from a synthetic as well as real images. Table 1 shows that the mean error in the height calculation of the reconstructed surface is improved around 5-6 mm in height and there is around 2-3 degree improvement in surface normal estimation.

Table 1: Mean Error

	Mean Angular Error in surface normal(degree)		Mean Height Error(mm)	
	Traditional PS	Our Method	Traditional PS	Our Method
Synthetic Sphere	6.53	4.53	14.586	9.108
Polystyrene Sphere	6.72	4.61	15.642	10.714
Human Dummy	6.88	4.86	17.006	11.066

Figure 7(a) shows the surface reconstructed from surface normals obtained from traditional photometric stereo while Figure 7(b) is the surface reconstructed from surface normals obtained from the proposed method by using a Poisson based surface integrator [20]. If we visually compare Figure 7(a) with the ground truth in Figure 7(c) we can easily find low

frequency geometric distortion in addition to high frequency noise. This geometric distortion is due to the fact that photometric stereo in its original form interprets a change in light intensity due to change in light direction as change in surface normal, which is very common in low cost and large field of view photometric stereo imaging setups. In comparison, Figure 7(b) is more flat and closer to the ground truth. This is because the geometric distortion is partially removed by considering the lighting distance from the object surface. The same phenomena can be observed clearly by plotting slices of the surface as shown in Figure 8.

Figure 8 shows slices of a reconstructed surfaces. When comparing proposed method (which estimates distance of object from the camera and calculates light vector for every pixel using distance estimation), with traditional photometric stereo (which assumes the same lighting direction for each pixel), it is clear that the proposed method calculates more accurate surface normals and hence better surface reconstruction.

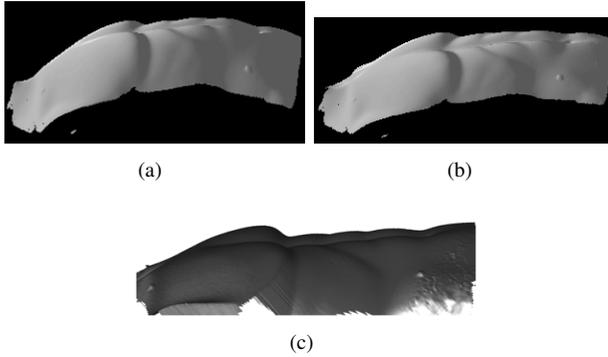


Figure 7: (a) Integrated surface using traditional PS. (b) Integrated surface using Proposed method. (c) Surface scanned from 3dMD as a ground truth.

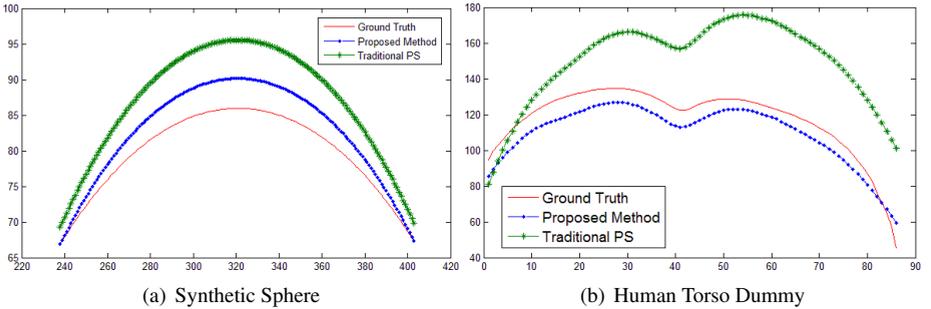


Figure 8: Slices of Integrated Surface

5 Conclusion

In this paper we presented a method to enable light vectors to be calculated dynamically (as an object moves in field of view) for improved photometric stereo 3D surface reconstruction performance. Traditional Photometric Stereo assumes that light vectors at every pixel are same, which is not usually the case in real applications, and especially so where the object size is comparable to object range. The error in estimating the surface normals is highly

dependent on the placement of the calibrated object relative to the camera. By using the proposed method this error is almost constant and independent from the working distance.

References

- [1] www.3dmd.com/3dmdface.html(last accessed 25 Oct 2012).
- [2] J Ackermann, F Langguth, S. Fuhrmann, and M. Goesele. Photometric stereo for outdoor webcams. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 262–269, June 2012. doi: 10.1109/CVPR.2012.6247684.
- [3] Jahanzeb Ahmad, Jiulai Sun, Lyndon Smith, Melvyn Smith, John HENDERSON, and Anirban MAJUMDAR. Novel Photometric Stereo Based Pulmonary Function Testing. In *3rd international conference and exhibition on 3D Body Scanning Technologies, Lugano, Switzerland.*, 2012.
- [4] I Ashdown, P Eng, and Others. Near-Field Photometry-Measuring and Modeling Complex 3-D Light Sources. 1995.
- [5] Peter N. Belhumeur, David J. Kriegman, and Alan L. Yuille. The Bas-Relief Ambiguity. *International Journal of Computer Vision*, 35(1):33–44, November 1999. ISSN 09205691. doi: 10.1023/A:1008154927611.
- [6] Willem de Boer, Joan Lasenby, Jonathan Cameron, Rich Wareham, Shiraz Ahmad, Charlotte Roach, Ward Hills, and Richard Iles. SLP: A Zero-Contact Non-Invasive Method for Pulmonary Function Testing. *Proceedings of the British Machine Vision Conference 2010*, pages 85.1–85.12, 2010. doi: 10.5244/C.24.85.
- [7] Hao Du, Dan Goldman, and Steven Seitz. Binocular Photometric Stereo. In *Proceedings of the British Machine Vision Conference 2011*, pages 84.1–84.11. British Machine Vision Association, 2011. ISBN 1-901725-43-X. doi: 10.5244/C.25.84.
- [8] F Dunn and I Parberry. *3D math primer for graphics and game development*. 2011. ISBN 1556229119.
- [9] Paolo Favaro and Thoma Papadhimetri. A closed-form solution to uncalibrated photometric stereo via diffuse maxima. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pages 821–828. IEEE, June 2012. ISBN 978-1-4673-1228-8. doi: 10.1109/CVPR.2012.6247754.
- [10] Mark F. Hansen, Gary A. Atkinson, Lyndon N. Smith, and Melvyn L. Smith. 3D face reconstructions from photometric stereo using near infrared and visible light. *Computer Vision and Image Understanding*, 114(8):942–951, August 2010. ISSN 10773142. doi: 10.1016/j.cviu.2010.03.001.
- [11] Carlos Hern, George Vogiatzis, and C Hernández. Self-calibrating a real-time monocular 3d facial capture system. *Proceedings International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, 2010.

- [12] Carlos Hernández Esteban, George Vogiatzis, Roberto Cipolla, and C Hernández. Multiview photometric stereo. *IEEE transactions on pattern analysis and machine intelligence*, 30(3):548–54, March 2008. ISSN 0162-8828. doi: 10.1109/TPAMI.2007.70820.
- [13] T Higo, Y Matsushita, N Joshi, and K Ikeuchi. A hand-held photometric stereo camera for 3-D modeling. *2009 IEEE 12th International Conference on Computer Vision*, pages 1234–1241, September 2009. doi: 10.1109/ICCV.2009.5459331.
- [14] Y Iwahori, H Sugie, and N Ishii. Reconstructing shape from shading images under point light source illumination. In *Pattern Recognition, 1990. Proceedings., 10th International Conference on*, volume i, pages 83 –87 vol.1, June 1990. doi: 10.1109/ICPR.1990.118069.
- [15] Andrew Jones, Graham Fyffe, Xueming Yu, Wan-Chun C Ma, Jay Busch, Ryosuke Ichikari, Mark Bolas, and Paul Debevec. Head-Mounted Photometric Stereo for Performance Capture. *2011 Conference for Visual Media Production*, pages 158–164, November 2011. doi: 10.1109/CVMP.2011.24.
- [16] Dong Junyu, G. McGunnigle, Su Liyuan, Fang Yanxia, and Wang Yuliang. Improving photometric stereo with laser sectioning. *12th IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 1748–1754, 2009. doi: 10.1109/ICCVW.2009.5457494.
- [17] B Kim and P Burger. Depth and shape from shading using the photometric stereo method. *CVGIP: Image Understanding*, 54(3):416–427, 1991.
- [18] R Kozera and L Noakes. Noise reduction in photometric stereo with non-distant light sources. *Computer Vision and Graphics*, pages 103–110, 2006.
- [19] Zheng Lu, Yu-Wing Tai, Moshe Ben-Ezra, and Michael S. Brown. A framework for ultra high resolution 3D imaging. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1205–1212, June 2010. doi: 10.1109/CVPR.2010.5539829.
- [20] T Simchony, R Chellappa, and M Shao. Direct analytical methods for solving Poisson equations in computer vision problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):435–446, 1990. ISSN 01628828. doi: 10.1109/34.55103.
- [21] Melvyn L Smith and Lyndon N Smith. Dynamic photometric stereo—a new technique for moving surface analysis. *Image Vision Comput.*, 23(9):841–852, September 2005. ISSN 0262-8856. doi: 10.1016/j.imavis.2005.01.007.
- [22] J Sun, M Smith, L Smith, and A Farooq. Examining the uncertainty of the recovered surface normal in three light photometric stereo. *Image and Vision Computing*, 25(7): 1073–1079, 2007.
- [23] A Varnavas, Vasileios Argyriou, Jeffrey Ng, and A A Bharath. Dense photometric stereo reconstruction on many core GPUs. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, pages 59–65, June 2010. doi: 10.1109/CVPRW.2010.5543152.

- [24] RJ Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):139–144, 1980.
- [25] Chenglei Wu, Yebin Liu, Qionghai Dai, and Bennett Wilburn. Fusing multiview and photometric stereo for 3D reconstruction under uncalibrated illumination. *IEEE transactions on visualization and computer graphics*, 17(8):1082–95, August 2011. ISSN 1941-0506. doi: 10.1109/TVCG.2010.224.