

Deep Learning Based Photometric Stereo from Many Images and Under Unknown Illumination

Hossein Malekmohamadi

Institute of Artificial Intelligence, School of Computer Science and Informatics

De Montfort University, Leicester LE1 9BH, UK

Email: hossein.malekmohamadi@dmu.ac.uk

Abstract—Shape from X is an interesting area of research in computer vision community. This topic is divided into passive and active methods. Example of passive methods is shape from texture, shape from defocus and shape from the silhouette. For active methods, the important categories are shape from shading and photometric stereo. In shape from shading, the cue for shape reconstruction is shading which is the relation between intensity and shape. In this case, only one image is considered. In photometric stereo, where multiple vantage points exist, 3D reconstruction considers multiple images (at least three). Photometric stereo on its own can be categorised depending on pre-existing information of illumination directions, illumination intensities, Lambertian surfaces or non-Lambertian surfaces. This paper presents a method employing deep learning for photometric stereo where lighting and surface conditions are unknown. The proposed method is applied to a public dataset. Based on the experimental results, this method outperforms currently existing techniques.

Keywords: Photometric Stereo, Non-Lambertian, Deep Learning, Surface Normals, Regressive Deep Learning, Uncalibrated.

I. INTRODUCTION

In computer vision community, there are volumetric or surface data types. Volumetric data could be voxel grids or density. Surface data types are point clouds and range maps. One important way to acquire range is the optical method. This is on its own could be divided into passive and active scenarios. For passive mode, methods like shape from motion, stereo, shading, texture and focus exist. For active mode, important methods are stereo with projected texture, active depth from defocus, photometric stereo and time of flight cameras. Advantages of these methods are they are non-invasive, fast and inexpensive. They have some disadvantages such as sensitivity to transparency, specularities and inter-reflections. Amongst the active modes, photometric stereo has the advantage of employing 2D images from different lighting conditions captured with a fixed orthographic camera. This makes photometric stereo an inexpensive method while it can provide many details of surface geometry.

There are numerous methods in photometric stereo by assuming several conditions for illumination and surface properties. The photometric stereo problem could be tackled algebraically or by using machine learning techniques as well. In this paper, a method based on deep learning is presented for photometric stereo. There is no constraining assumption on

the photometric stereo as unknown illumination and surface diffuseness have been used in this paper. More importantly, human in the loop is discarded as all available images are used to reconstruct in 3D not only some selected images or by imposing some algebraic constraints.

The rest of this paper is organised as follows: Section 2 is dedicated to the background and literature review on photometric stereo. In Section 3, principles of photometric stereo are described. In Section 4, experimental results based on deep learning are presented and discussed. Finally, Section 5 concludes this paper.

II. BACKGROUND

Photometric stereo has numerous applications in analysing surface properties. Amongst them are skin analysis [1], facial recognition [2], [3], micro-geometry capture [4], industrial quality inspection [5] and material type identification [6] such as paper type identification [7]. One of the advantages of photometric stereo is that it can be implemented with a very cheap 2D camera and illumination set-up. Furthermore, it can provide many details of the surface structures. A method that is able to work with unknown lighting conditions and unknown surfaces will obtain images of a scene from internet and by applying photometric stereo it can deliver 3D surface properties of the scene.

Classic photometric stereo [8] considers varying light directions while capturing multiple images with a fixed orthographic camera. It is shown that this technique can provide enough information to determine surface orientations. The assumption in [8] is for Lambertian surfaces. In other words, the surface's luminance is isotropic, and the luminous intensity obeys Lambert's cosine law. Photometric stereo has many variants. Depending on the pre-existing information on lighting directions, lighting intensities and obeying Lambert law it can be categories into: (1) Lambertian calibrated, (2) Lambertian uncalibrated, (3) Non-Lambertian calibrated and (4) Non-Lambertian and uncalibrated.

In this paper, the dataset introduced in [9] is used for the experiments as shown in Figure 1. In [9], the authors survey methods in photometric stereo emphasising on non-Lambertian and uncalibrated methods. They also introduce the *DiLiGenT* photometric stereo image dataset which is used

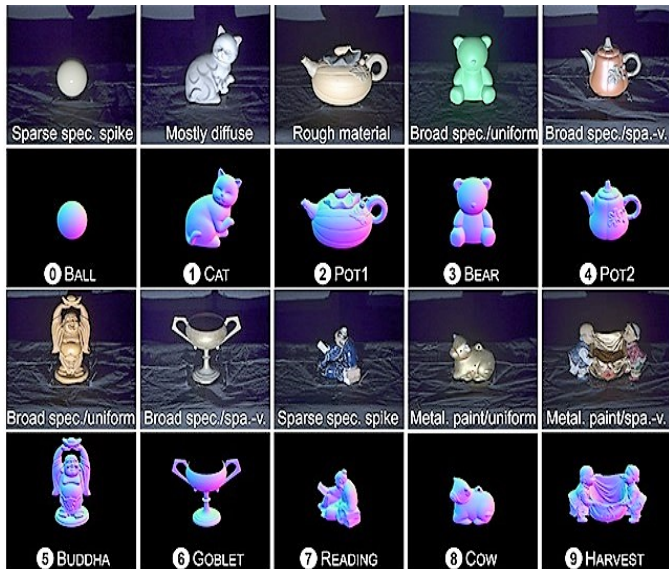


Figure 1: The dataset used in this paper and this image are taken from [9]. Rows 1 and 3 show objects and their material types and rows 2 and 4 shows ground truth normals.

in this paper. Based on this dataset, authors in [9] quantitatively evaluate state-of-the-art photometric stereo methods for general non-Lambertian materials and unknown lightings with respect to the provided ground truth data.

Authors of [10] present a robust photometric stereo method for various non-Lambertian corruptions such as specularities, shadows, and image noise. They construct a constrained sparse regression problem that enforces both Lambertian, rank-3 structure and sparse, additive corruptions. This method is categorised in non-Lambertian calibrated. In [11], a bi-polynomial reflectance model that can precisely represent the low-frequency component of reflectance is presented. This bi-polynomial model can accurately solve inverse problems by effectively discarding the high-frequency component while retaining nonlinear variations in the low-frequency part. This model is useful for estimating reflectance and shape of an object in the non-Lambertian calibrated scenario [11].

Authors of [12] describe a photometric stereo method designed for surfaces with spatially-varying bidirectional reflectance distribution functions (BRDFs). Their method builds on the observation that most objects are composed of a small number of fundamental materials by constraining each pixel to be representable by a combination of at most two such materials. Their method uses previous information by constraining conditions on lighting [12]. In [13], a photometric stereo method that is purely pixelwise and handles general isotropic surfaces is presented. Authors in [13] construct a constrained bivariate regression problem where the regression function is approximated by smooth, bivariate Bernstein polynomials. The unknown normal vector was separated from the unknown reflectance function by considering the inverse representation of the image formation process, and then they could accurately compute the unknown surface normals by solving a simple and efficient quadratic programming problem [13]. This method

falls into non-Lambertian calibrated method while they put a constraint in solving the photometric stereo equations. The authors of [14] show how to perform photometric stereo, assuming that all lights in a scene are isotropic and distant from the object but otherwise unconstrained. Lighting in each image may be an unknown and arbitrary combination of diffuse, point and extended sources. Their work is based on recent results showing that for Lambertian objects, general lighting conditions can be represented using low order spherical harmonics [14]. As a result, their method falls into Lambertian surfaces.

Authors of [15] propose a new prior on the albedo distribution of natural images. The entropy of the distribution should be minimised. This prior is justified by the fact that many objects in the real-world are composed of a small finite set of albedo values [15]. In [16] a self-calibrating photometric stereo method is proposed. From a set of images taken from a fixed viewpoint under unknown illuminations, the method in [16] automatically determines a radiometric response function and resolves the generalised bas-relief ambiguity for estimating accurate surface normals and albedos. Authors in [16] have developed a complete auto-calibration method for photometric stereo. In [17] a method to identify local diffuse reflectance maxima is proposed, where the normal vector is coincident with the lighting vector. Authors in [18] consider an additive bivariate specular reflection to solve the generalised bas-relief ambiguity. The methods in [16] and [17] assume Lambertian surfaces for objects and then they select illumination vectors. Their solution to photometric stereo GBR ambiguity is based on putting assumption on being Lambertian surface and solve the photometric stereo equation either as a linear problem or by brute force search. All these methods need a human in the loop for selection and assumptions.

Neural networks are used in clustering through unsupervised learning, classification through supervised learning, or regression. In other words, they help group unlabeled data, categorise labelled data or predict continuous values. Regression neural networks map one set of continuous inputs to another set of continuous outputs. For photometric stereo case, the target is surface normals x , y and z components to be continuously predicted. The method introduced in [19] shows how to reconstruct the shape of the specular surface by learning the mapping between three image irradiances observed under the illumination from three lighting directions and the corresponding surface gradient. This method reconstructs the surface gradient distribution after determining the values of reflectance parameters of a test object using two-step neural network which consist of one to extract two gradient parameters from three image irradiances and its opposite one [19]. In this method, the authors use three images to train a neural network. The process of selecting three images of many possible images is the limitation of this model. In [20], a neural-network-based photometric stereo approach for 3D surface reconstruction is proposed. The neural network inputs are the pixel values of the 2D images to be reconstructed. The normal vectors of the surface can then be obtained from the neural network after

learning.

Deep learning is a machine learning method based on learning data representations, as opposed to task-specific algorithms. Learning can be supervised, semi-supervised or unsupervised. Some representations are almost based on interpretation of information processing and communication patterns in a biological nervous system. Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks have been applied to many applications including computer vision, where they have produced results comparable to and in some cases superior to human experts. Authors of [21] present a photometric stereo method based on deep learning. Their method takes reflectance observations under varying light directions and infers the corresponding surface normal per pixel. Authors of [21] normalised input images to the corresponding light intensities which is pre-existing information.

III. PHOTOMETRIC STEREO PRINCIPLES

For photometric stereo, we need to record 2D images with a fixed orthographic camera and varying illumination. As shown in Figure 2, one 2D image is recorded per each illumination source. Whilst the target is to estimate surface normals, the angle between look vector and illumination vector is known. However, the angle between the surface normal and reflectance is unknown.

Image intensity (I) can be written as a function of the light vector (direction and intensity), surface normal (unknown) and $\rho()$ which is the diffuse albedo. This is shown as:

$$I = \max \{ \rho(n, l) \circ N^T L, 0 \}$$

For every image, we record its intensity while surface normal distribution is unknown. For photometric stereo reconstruction, we need at least three images. Some current methods put a constraint on $\rho()$ to solve the photometric stereo problem. Suggestions are that $\rho()$ is constant [8], bi-polynomial [11], sparse matrix [10], [22], isotropic depending on some of the angles (As shown in Figure 2) [23]–[25] or a Bernstein polynomial depending on light and look vectors [13]. In the basic photometric stereo problem, we have only three images while in reality, many images can exist. Consequently, we have several equations for each recorded image intensity:

$$I_1 = \max \{ \rho(n, l_1) \circ N^T L_1, 0 \}$$

$$I_2 = \max \{ \rho(n, l_2) \circ N^T L_2, 0 \}$$

$$I_3 = \max \{ \rho(n, l_3) \circ N^T L_3, 0 \}$$

.....

.....

.....

$$I_m = \max \{ \rho(n, l_m) \circ N^T L_m, 0 \}$$

In the above equations, m is the number of illumination

sources (images per object). Each image has a resolution of $W \times H$. Consequently, image intensities are vectors of $WH \times 1$, normal vectors are $WH \times 3$ and light is a 3×1 vector. Considering that each object is captured with least 3 images (in this dataset 96 or in the formula m) and we are only keeping positive values (p non-zero elements), we can create a stack of image vectors per each object as:

$$J_{p \times m} = \rho(n, l) \circ N_{p \times 3}^T L_{3 \times m}$$

Please note that the behaviour of the function $\rho()$ and surface normal vectors N are unknown. We may put some constraints on $\rho()$ like in [8], [10], [11], [13], [22]–[25]. If light intensities and directions are known, by pseudo-inverting the above equations we can estimate surface normal vectors. As an example, if $\rho()$ is the constant μ and light intensities are known then by multiplying the inverse light vector we have:

$$J_{p \times m} \times L_{m \times 3}^\dagger = \mu \circ N_{p \times 3}^T L_{3 \times m} \times L_{m \times 3}^\dagger$$

Please note that L^\dagger is pseudo-inverse of L (if exists) and:

$$L_{3 \times m} \times L_{m \times 3}^\dagger = I_{3 \times 3}$$

Finally, we achieve the scaled surface normal vector as:

$$J_{p \times 3}^* = \mu \circ N_{p \times 3}^T$$

One algebraic assumption to find the pseudo-inverse vector is to use only 3 images out of our many images. In this case we have the classic photometric stereo problem. In other words, we have to select 3 suitable images for reconstruction. As an example, putting 3 independent LED light sources will do this job if they sufficient spatial distance. Please also note that we are simplifying photometric stereo in multiple levels by (1) putting a constraint on albedo, (2) pseudo-inverting light vectors and (3) limiting number of images.

Onward in this paper, the effect of pseudo-inverting the light vector is called *Calibrated* or *Known lighting* and if this is not used then the method is *Uncalibrated* or *Unknown lighting*. The rationale behind this is that calculating inverse light vectors is the same as applying some pre-existing and known information. However, the method in this paper falls into the latter. In other words: (1) we use all images provided for each object, (2) we do not use illumination properties to model photometric stereo and (3) we do not put any constraints on the diffuse albedo. The target is to find surface normal vector x, y and z components.

The dataset used in this paper is taken from [9]. In total, we have 10 objects: *Ball, Bear, Buddha, Cat, Cow, Goblet, Harvest, Pot1, Pot2* and *Reading*. Each object has 96 images at 612×512 resolution. In terms of surface shapes, we have the simple sphere *Ball*, smoothly curved surfaces *Bear, Cat, Goblet* and *Cow*. Smooth surfaces with local details

are *Pot1* and *Pot2*. Surfaces with complicated geometry are *Buddha* and *Reading*, and delicate shapes with concave parts are *Harvest*. Ground truth normals and light properties have been provided. However, in this paper, we do not use illumination properties as our method deals with photometric stereo in unknown lighting and unknown surface conditions. In the next section, the method to train and test the deep learning algorithm is described. It proves that the technique in this paper is robust and accurate for most shapes and surfaces.

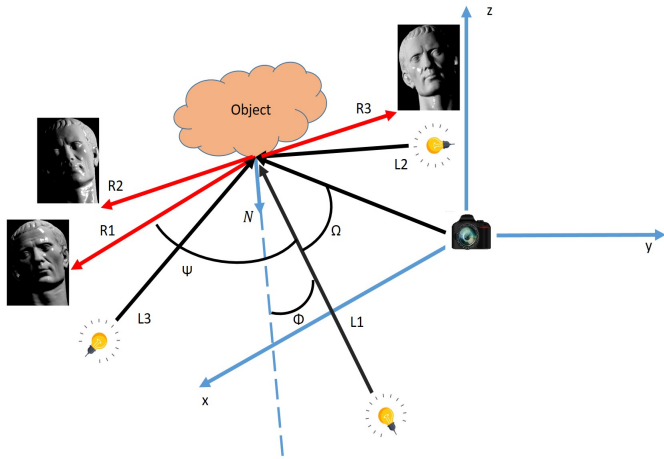


Figure 2: Photometric Stereo Principles Shown for one Illumination Source: Many light sources (at least three) exist in photometric stereo. The light vector L , surface normal N , angle between look vector and light Ω , the angle between the surface normal and light θ and the angle between light and reflectance ψ are shown.

IV. DEEP LEARNING BASED ESTIMATION

In this section, the method to train and test the deep learning based neural network is described. A feed-forward artificial neural network model, also known as the deep neural network or multi-layer perceptron (MLP), is the most common type of deep neural network. Several other types are convolutional neural networks (CNN) and recurrent neural networks (RNN). MLPs work well on tabular data. However, if we have image data, then CNNs are a great choice. If we have sequential data like text or audio, then RNNs can perform better. In this paper, TensorFlow [26] is used to train and test the deep neural network.

As stated in the previous section, we have 10 objects and for every object 96 images have been recorded. Each object image has a different number of positive values that we keep. The deep learning method in this paper is based on a multi-layer feed-forward artificial neural network that is trained with stochastic gradient descent using back-propagation. The network can contain a large number of hidden layers consisting of neurons ([10,10] in this paper) with rectifier activation function. L1 regularisation which can add stability and improve generalisation is not used. L2 regularisation that adds stability and improve generalisation is not used either. Input layer dropout ratio is 0. Stopping metric to use for early stopping is deviance and adaptive learning rate time decay factor is 0.99. Adaptive learning rate smoothing factor is 10^{-8} .

We have created 3 independent regressive deep neural networks to estimate x , y and z of surface normals. Each network has [10,10] hidden layers. Our dataset has 377976 samples of 96 features. We kept 75% data for training and the 25% for testing. The target is to imitate ground truth normals without previous information on illumination and surface properties. For objects, *Ball* and *Bear* the mean angular errors are 7.68 and 9.28 respectively as shown in Table I. The resulting images are shown in Figure 3. For objects, *Buddha* and *Cat* the mean angular errors are 14.55 and 7.20 respectively. The resulting image is shown in Figure 3. For objects, *Cow* and *Goblet* the mean angular errors are 15.28 and 12.81 respectively. For objects *Harvest* and *Pot1*, the mean angular errors are 21.82 and 8.58. Finally, for objects, *Pot2* and *Reading* the mean angular errors are 10.67 and 18.14 respectively. Training of deep regression network for x channel took 13.761 seconds, for y channel took 13.216 seconds and 13.374 seconds for z . The deep learning algorithm was implemented on a laptop with i7-6700HQ CPU, 16.0 GB RAM and NVIDIA GeForce GTX 950M. As shown in Table I the method introduced in this paper can outperform existing ones in 7 out of 10 investigated objects. Angular errors for each object are shown in Figure 4. The average error for our method is 12.60 whereas for the nearest technique is 16.67. For other three methods as shown in Table I, the mean angular error of our method is almost half or less than half of the rest.

Object *Cat* is mostly diffuse where our method outperforms the rest. *Pot1* has a rough surface where our method has the lowest mean angular error as shown in Table I. For surfaces with strong and sparse specular spikes like *Ball* or *Reading*, the presented method in this paper has the lowest error for *Reading* as shown in Table I. For objects with broad and soft specular lobes on a uniform like *Bear* and *Buddha*, the method in [18] has the lowest angular error. The performance of our method on spatially-varying materials like *Pot2* and *Goblet* shows its robustness as shown in Table I. *Cow* is an object with metallic paint on mostly uniform whereas *Harvest* is spatially-varying surface. For both, our method has the lowest error as shown in Table I.

Other advantages of the method in this paper are that it operates without any assumption on lighting conditions, surface properties or algebraic constraints on the photometric stereo equation. The technique presented in this paper can be extended to constrained scenarios considering illumination. As an example, we can train the deep neural network given image intensities and light vectors as input while they estimate surface normals. This suggestion will fall into *Calibrated* categories tackling photometric stereo as discussed earlier. Another possible scenario is by pseudo-inverting light as stated above in the previous section, we convert *Uncalibrated* to *Calibrated*. In this case, feature vector size to be fed into neural networks reduce from a $p \times m$ to $p \times 3$. Finally, this technique enables us to calculate input variable importances in such a way that we can identify images contribute more to 3D reconstruction as shown in Table II.

Table I: Mean angular errors for the method introduced in this paper compared with state of the art methods. Dominant method for each object has been highlighted. The method in this paper outperforms in 7 out of 10 objects with good results for the rest 3 objects. The average mean angular error for our method is much less than the rest.

Method	Ball	Bear	Buddha	Cat	Cow	Goblet	Harvest	Pot1	Pot2	Reading	Average
Alldrin et al. [15]	7.27	16.81	32.81	31.45	54.72	46.54	61.7	18.37	49.16	53.65	37.25
Shi et al. [16]	8.9	11.98	15.54	19.84	22.73	48.79	73.86	16.68	50.68	26.93	29.59
Papadimitri et al. [17]	4.77	9.07	14.92	9.54	19.53	29.93	29.21	9.51	15.9	24.18	16.67
Wu et al. [18]	4.39	6.42	13.19	36.55	19.75	20.57	55.51	9.39	14.52	58.96	23.92
Method in this paper	7.68	9.28	14.55	7.20	15.28	12.81	21.82	8.58	10.67	18.14	12.60

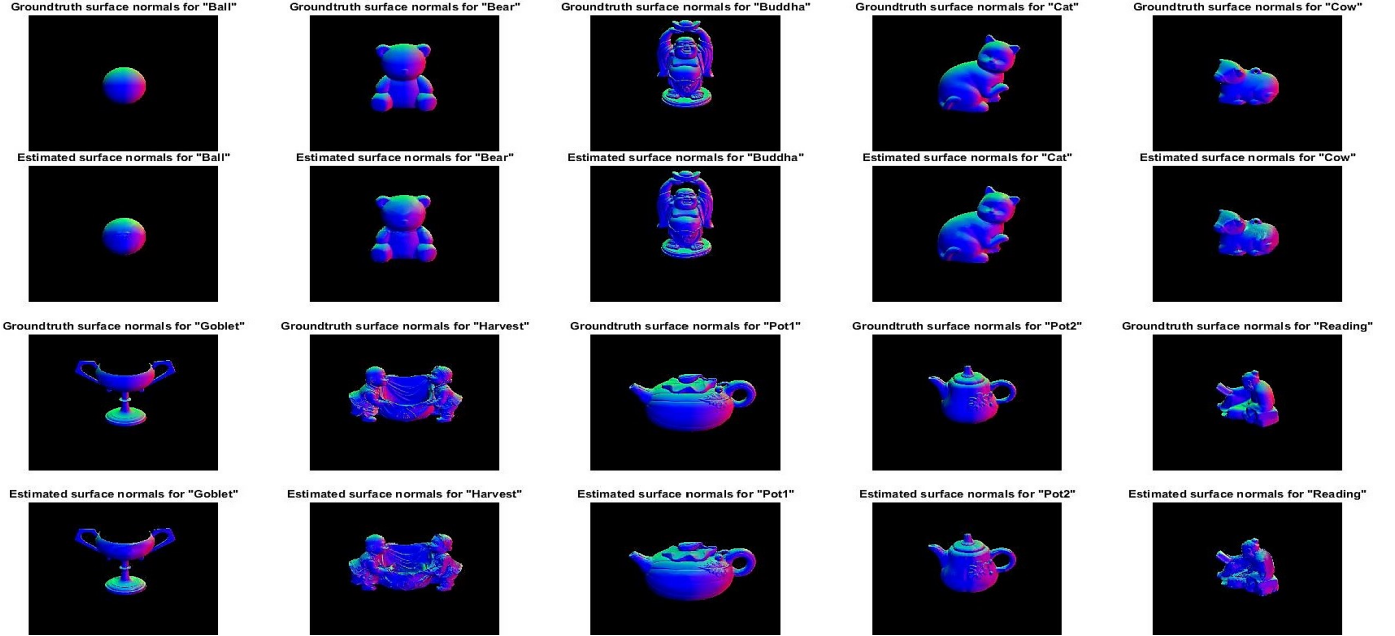


Figure 3: Estimated vs. groundtruth normals for all objects.

V. CONCLUSION

In this paper, a method using deep regression neural networks has been applied to the photometric stereo problem. This method considers only the information from captured images as input to the learning algorithm. In other words, information regarding illumination or surface properties has not been used in this paper. In photometric stereo categorisation, the method presented in this paper is uncalibrated and for non-Lambertian surfaces. Test results on a public dataset show that this method can outperform currently existing methods. Other advantages include speed, no human in the loop and it can sort used images during reconstruction. This means that given m number of images for an object, after learning and validation, we can sort images based on their importance. This is useful in terms of designing illumination factors to reduce capturing time. Future works include reintroducing human in the loop in the photometric stereo problem by either algebraically constraining the problem or by considering illumination properties. Either of these approaches will convert the photometric stereo problem to a calibrated case. One application of the method introduced in this paper is it can be applied to current internet images to produce 3D surface normals of different scenes.

REFERENCES

- [1] J. Sun, M. Smith, L. Smith, L. Coutts, R. Dabis, C. Harland, and J. Bamber, "Reflectance of human skin using colour photometric stereo: with particular application to pigmented lesion analysis," *Skin research and technology*, vol. 14, no. 2, pp. 173–179, 2008.
- [2] S. K. Zhou, G. Aggarwal, R. Chellappa, and D. W. Jacobs, "Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 230–245, 2007.
- [3] S. Zafeiriou, G. A. Atkinson, M. F. Hansen, W. A. Smith, V. Argyriou, M. Petrou, M. L. Smith, and L. N. Smith, "Face recognition and verification using photometric stereo: The photoface database and a comprehensive evaluation," *IEEE transactions on information forensics and security*, vol. 8, no. 1, pp. 121–135, 2013.
- [4] M. K. Johnson, F. Cole, A. Raj, and E. H. Adelson, "Microgeometry capture using an elastomeric sensor," in *ACM Transactions on Graphics (TOG)*, vol. 30, no. 4. ACM, 2011, p. 46.
- [5] A. R. Farooq, M. L. Smith, L. N. Smith, and S. Midha, "Dynamic photometric stereo for on line quality control of ceramic tiles," *Computers in industry*, vol. 56, no. 8, pp. 918–934, 2005.
- [6] A. Hertzmann and S. M. Seitz, "Example-based photometric stereo: Shape reconstruction with general, varying brdfs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1254–1264, 2005.
- [7] H. Malekmohamadi, K. Emrith, S. Pollard, G. Adams, M. Smith, and S. Simske, "Paper type classification based on a new 3d surface texture measure," *Electronics Letters*, vol. 50, no. 8, pp. 596–598, 2014.

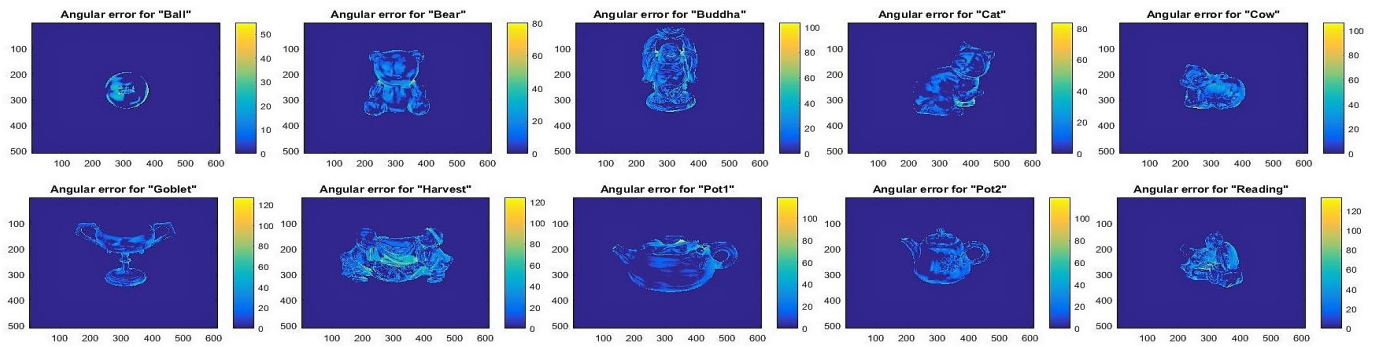


Figure 4: Angular error for all objects. Image resolutions are 612×512 and the legends give an idea of heat map for the error.

Table II: Importance of input images to estimate x , y and z : only the first 10 images (out of 96) are shown for each component. The results are for all 10 objects in this paper.

Component	Image	Relative Importance	percentage
X	147	1	0.0211
	146	0.9433	0.0199
	191	0.9077	0.0191
	183	0.8084	0.0171
	145	0.7862	0.0166
	148	0.7841	0.0165
	192	0.7717	0.0163
	190	0.7655	0.0161
	141	0.7243	0.0153
	142	0.7231	0.0153
Y	148	1	0.0224
	154	0.829	0.0186
	140	0.8109	0.0182
	150	0.7939	0.0178
	151	0.7552	0.0169
	115	0.6926	0.0155
	123	0.6896	0.0154
	114	0.6866	0.0154
	111	0.6748	0.0151
	165	0.6683	0.015
Z	148	1	0.0186
	140	0.929	0.0173
	157	0.8896	0.0165
	154	0.8865	0.0165
	165	0.8736	0.0162
	192	0.8519	0.0158
	152	0.8514	0.0158
	14	0.7835	0.0146
	123	0.7689	0.0143
	111	0.7544	0.014

- [8] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical engineering*, vol. 19, no. 1, p. 191139, 1980.
- [9] B. Shi, Z. Wu, Z. Mo, D. Duan, S.-K. Yeung, and P. Tan, "A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 3707–3716.
- [10] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 318–325.
- [11] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Bi-polynomial modeling of low-frequency reflectances," *IEEE transactions on pattern analysis and machine intelligence*, vol. 36, no. 6, pp. 1078–1091, 2014.
- [12] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying brdfs from photometric stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 6, pp. 1060–1071, 2010.
- [13] S. Ikehata and K. Aizawa, "Photometric stereo using constrained bivariate regression for general isotropic surfaces," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2179–2186.
- [14] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *International Journal of computer vision*, vol. 72, no. 3, pp. 239–257, 2007.
- [15] N. G. Alldrin, S. P. Mallick, and D. J. Kriegman, "Resolving the generalized bas-relief ambiguity by entropy minimization," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–7.
- [16] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan, "Self-calibrating photometric stereo," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 1118–1125.
- [17] T. Papadimitri and P. Favaro, "A closed-form, consistent and robust solution to uncalibrated photometric stereo via local diffuse reflectance maxima," *International journal of computer vision*, vol. 107, no. 2, pp. 139–154, 2014.
- [18] Z. Wu and P. Tan, "Calibrating photometric stereo by holistic reflectance symmetry analysis," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1498–1505.
- [19] Y. Iwahori, R. J. Woodham, H. Tanaka, and N. Ishii, "Neural network to reconstruct specular surface shape from its three shading images," in *Neural Networks, 1993. IJCNN'93-Nagoya. Proceedings of 1993 International Joint Conference on*, vol. 2. IEEE, 1993, pp. 1181–1184.
- [20] W.-C. Cheng, "Neural-network-based photometric stereo for 3d surface reconstruction," in *Neural Networks, 2006. IJCNN'06. International Joint Conference on*. IEEE, 2006, pp. 404–410.
- [21] H. Santo, M. Samejima, Y. Sugano, B. Shi, and Y. Matsushita, "Deep photometric stereo network," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 501–509.
- [22] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," in *Asian Conference on Computer Vision*. Springer, 2010, pp. 703–717.
- [23] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances," *Computer Vision—ECCV 2012*, pp. 455–468, 2012.
- [24] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.
- [25] T. Higo, Y. Matsushita, and K. Ikeuchi, "Consensus photometric stereo," in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 1157–1164.
- [26] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015, software available from tensorflow.org. [Online]. Available: <https://www.tensorflow.org/>