Outdoor Photometric Stereo

Lap-Fai Yu [1*] Sai-Kit Yeung [2] Yu-Wing Tai [3] Demetri Terzopoulos [1] Tony F. Chan [4]

[1] University of California, Los Angeles [2] Singapore University of Technology and Design
[3] Korea Advanced Institute of Science and Technology [4] Hong Kong University of Science and Technology

Abstract

We introduce a framework for outdoor photometric stereo utilizing natural environmental illumination. Our framework extends beyond existing photometric stereo methods intended for laboratory environments to encompass robust outdoor operation in the real world. In this paper, we motivate our framework, describe the components of its processing pipeline, and assess its performance in synthetic experiments as well as in natural experiments including objects in outdoor environments with complex real-world illuminations.

1. Introduction

Photometric stereo is a technique for inferring 3D surface shape from pixel intensities in ordinary 2D images [23, 11]. In the conventional photometric stereo setting, multiple images of an object are captured under different illumination conditions. By fitting different reflectance models to the observed pixel intensities, pixel-dense surface normals are estimated [7, 22, 20, 3, 4, 21, 14, 29, 6]. The focus in the majority of prior photometric stereo work was to tackle technical challenges such as reflectance model assumptions, self-shadowing, specularities, noise and outliers, and so on; accordingly, experiments were normally conducted in well-controlled laboratory environments, usually dark rooms with single localized light sources, which facilitated empirical evaluation while minimizing uncertainties.

The objective of our work is to free photometric stereo from its laboratory confines and help make it a useful tool for computational photography. To this end, we develop an outdoor photometric stereo framework that can deal with real-world objects subject to natural illumination conditions. Figure 1 illustrates the application of our framework to a horse statuette in a sunlit outdoor environment. By analyzing the variation of pixel intensities among a set of input images (top row) acquired under different environmental illuminations whose associated environment maps are acquired using a mirror sphere (middle row), our algorithms can estimate the surface normals using our generalized reflectance model and reconstruct the 3D shape of the horse statuette surface (bottom row). To deal with outliers such as shadows, highlights and/or small misalignment errors across input images, we apply low rank matrix completion [13] to preprocess the input images. Our shape estimation approach is then formulated as an optimization method which alternates between normal estimation and estimation of environmental illumination that contributes to pixel intensities. Finally, total variation regularization [18, 5] is applied to refine the estimated normals by reducing ambiguities and noise in preparation for 3D surface reconstruction.

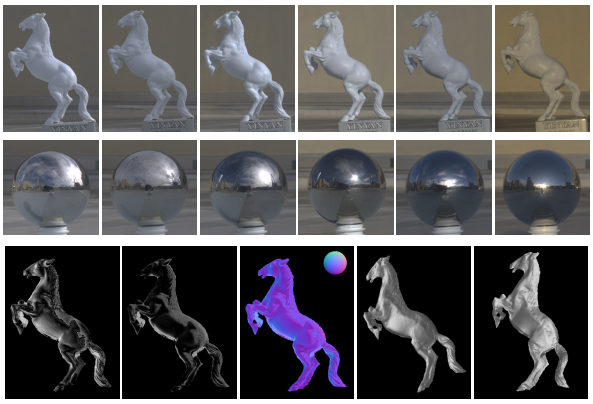


Figure 1. HORSE (SUNLIGHT). Top row: Captured input images. Middle row: Environmental illumination captured in a mirror sphere. Note the variance among the input images under different illumination conditions. Bottom row: The left two images show the normal maps $\mathbf{n}$ displayed as $\mathbf{n} \cdot \mathbf{l}$ with $\mathbf{l} = [-1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]^T$ and $\mathbf{l} = [1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]^T$, respectively. The third image shows the color-coded normal map. The fourth and the fifth images show two different views of the reconstructed 3D surface of the horse.

After reviewing related prior work (Section 2), we will present the generalized reflectance model employed in our outdoor photometric stereo framework along with the key steps of its processing pipeline, including our acquisition

*Part of the work reported herein was done while Lap-Fai Yu was a visiting student at SUTD.

Method	Surface Assumption	Calibration Object	Capturing Environment	#Images
Ours	Lambertian; varying albedo	mirror sphere	natural illumination	6-10 (theoretical: 4)
[12]	Lambertian; uniform albedo (painted to match calibration sphere)	same-material sphere	natural illumination	1
[4]	Lambertian; varying albedo	none	dark room general unknown lighting, fixed intensity	32-64 (theoretical:27)
[1]	mixtures of 2-3 fundamental materials	none	natural illumination	30-50 from more than 20k over a year
[16]	isotropic BRDF	mirror sphere	natural illumination	1

Table 1. Comparisons between our work and previous works.

system, preprocessing step, main algorithm, and postprocessing step that yields high-quality pixel-dense surface normal estimation in natural environments (Sections 3–4). We will also discuss implementation considerations relevant to these steps. Then, we will demonstrate the effectiveness of our framework through a variety of experiments including synthetic data, different background scenes, indoor scenes with different combinations of light sources, and outdoor scenes with varying sunlight (Section 5). Finally, we will present our conclusions and suggest avenues for future work (Section 6).

2. Related Work

There is a substantial amount of literature on photometric stereo, with representative papers including [23, 24, 11, 7, 20, 3, 28, 27, 4, 25, 19, 21, 14, 29, 6].

Much of this prior work is based on the Lambertian surface reflectance model, which requires at least three illumination directions to solve the surface normal estimation problem [23, 24]. For more than three input images, the surface normal vector at every pixel may be obtained using least squares fitting [28, 27], robust low rank minimization [25], or subspace clustering [21]. Several authors have relaxed the Lambertian model assumption; for instance, Tagare and deFigueiredo [22] employed an *m*-lobed reflective map, Solomon and Ikeuchi [20] used the Torrance-Sparrow reflectance model, Hertzmann and Seitz [10] used a reference object of the same material and a known shape to compute surface normals through analogy, Nayar *et al.* [15] used a hybrid reflectance model (Torrance-Sparrow and Beckmann-Spizzichino), Goldman *et al.* [9] optimized the shape and the BRDFs alternatively by assuming a set of basis materials according to the isotropic Ward model, and Yeung *et al.* [29] applied orientation consistency to estimate normals for transparent objects.

To our knowledge, there are only a handful of papers that consider general/natural illumination conditions. Johnson and Adelson [12] describe a shape-from-shading algorithm under natural illumination; however, their requirement of a calibration sphere of the same material BRDF as the captured object limits its practicality. Oxholm and Nishino [16] relax this restriction by using a mirror sphere to calibrate the illumination, while Yu *et al.* [30] utilize information obtained from depth cameras to constrain the problem. Since only a single input image is used, such shape-from-shading methods share some common limitations, including the fact that the estimated surface normals can easily be corrupted by outliers. Basri *et al.* [4] used low-order spherical harmonics to model general illumination, akin to the environment mapping representation [17]. Their model includes 27 variables and thus requires significantly more input images compared to conventional photometric stereo. The prior work that is most relevant to ours is by Ackermann *et al.* [1] who captured over twenty thousand outdoor webcam images throughout the year. The robustness of their photometric stereo method can be attributed to the large amount of data and a smart data selection process. However, acquiring this much image data is not an easy task.

We offer an approach to outdoor photometric stereo that strikes an attractive balance between the amount of image data required (about 6 to 10 input images) and the type of calibration object needed (a mirror sphere). Compared to previous methods, our method is both practical and accurate. Table 1 compares our work to the aforementioned efforts.

3. Environment Light Photometric Stereo

In this section, we describe the basic model of our environment light photometric stereo and our image data acquisition process.

3.1. Basic model

In the Lambertian surface model (assuming a linear camera response function), the intensity of a pixel I depends on the surface albedo ρ, the illumination direction $\mathbf{l}$, and the surface normal $\mathbf{n}$ according to

$$I(x) = \rho(x)\mathbf{l}(x) \cdot \mathbf{n}(x), \tag{1}$$

where x is the image coordinate. The common photometric stereo setting presumes a distant, directional light source. Thus, $\mathbf{l}$ is spatially invariant and it can easily be estimated

Figure 2. Left: Our simple setup for data acquisition. A mirror sphere is placed near the object. Middle: From the image of the mirror sphere, we estimate the illumination environment map using the method in [8]. Right: For each pixel, we must estimate the illumination directions that contribute to the pixel intensity, which depends on the orientation of the associated surface normal.

from the input images [19] or using a calibration object [28]. To solve the surface normal $\mathbf{n}$ in (1), we capture multiple images each taken with a different illumination direction. Hence, we obtain more observations than unknowns in (1) and $\mathbf{n}$ can be solved effectively using methods presented in [23, 28, 25, 21].

When we have multiple directional light sources, we can extend (1) by summing the contribution of each light source to the pixel's intensity, as follows:

$$I(x) = \rho(x) \sum_{i=1}^{K} c_i \mathbf{l}_i \cdot \mathbf{n}(x), \tag{2}$$

where K is the number of light sources in the scene, and c_i is the strength of light source i with direction $\mathbf{l}_i$. In the case where the illumination comes from all directions, we can describe the image intensity using (2) with K tending to infinity; i.e., an integral. Note that in (2) the number of unknowns for $\mathbf{n}$ remains the same as in (1) and (2) remains a linear equation when ρ and $\mathbf{l}_i$ are known.

3.2. Data acquisition

In the conventional photometric stereo setting, images are acquired in a darkroom and the light source directions can be well calibrated with fixed light sources. By contrast, we wish to use natural environmental illumination. Hence, a major component in our environment light photometric stereo approach is the estimation of the light source directions $\mathbf{l}_i$ in (2).

Figure 2 shows our experimental setup for data acquisition. We use a mirror sphere to capture the strength of the incoming light from all directions in the form of an environment map, a common method in graphics rendering for representing the effect of distant light sources illuminating object surfaces [8], and adopt it for surface normal estimation in our environment light photometric stereo. We put the mirror sphere near the object of interest and capture images of both.

Once we have acquired the environment map, we sample a number of directions in the illumination hemisphere using an icosahedron with sub-division [2], and average the illumination environment map over these directions; i.e., according to (2),

$$I(x) = \frac{\rho(x)}{K} \sum_{i=1}^{K} c_i \mathbf{l}_i \cdot \mathbf{n}(x), \tag{3}$$

where K is now the number of directions. In our implementation, we sample 2562 directions in the environment map to approximate the illumination. Note that when $\mathbf{l}_i \cdot \mathbf{n}(x) \leq 0$, then $\mathbf{l}_i$ does not contribute to the image intensity $I(x)$ in (3). Hence, we also need to estimate the lighting directions (Figure 2 (Right)) that contribute to the pixel intensity, as described in the next section.

4. Normal Estimation Algorithm

We will now describe our surface normal estimation algorithm. We first present our preprocessing step based on low-rank matrix completion. Then, we present our method for normal and light contribution refinement. Finally, we describe how to include total variation regularization [18, 5] to postprocess surface normals using spatial support.

4.1. Preprocessing via low-rank matrix completion

Our input data is subject to different sources of error—e.g., shadows, highlights, and even pixel misalignment across different image captures—which can affect the performance of our algorithm. To deal with these errors, we adopt the low-rank matrix completion technique [13, 25]. We assemble our n input images I^j for $j = 1, \cdots, n$ into a data matrix

$$\mathbf{D} = [\text{vec}(I^1) \cdots \text{vec}(I^n)], \tag{4}$$

where $\text{vec}(I^j) = [I^j(1), \cdots, I^j(m)]^T$ is the vectorized input image and m is the number of pixels in the object mask. Since our environment light model in (2) is linear and the dimension of $\mathbf{n}$ is 3, the rank of matrix $\mathbf{D}$ is in principle at most 3. However, due to the various errors, we observe in practice that the rank of $\mathbf{D}$ is greater than 3. As observed in [25], the errors related to photometric stereo are usually sparse; hence, we can isolate them by formulating the problem as a matrix rank minimization:

$$\min_{\mathbf{A},\mathbf{E}} ||\mathbf{A}||_* + \lambda ||\mathbf{E}||_1 \quad \text{such that} \quad \mathbf{D} = \mathbf{A} + \mathbf{E}, \tag{5}$$

where $\mathbf{A}$ is a rank 3 matrix, $\mathbf{E}$ is the matrix of error residuals, $||\cdot||_*$ and $||\cdot||_1$ are the nuclear norm and L_1-norm, respectively, and $\lambda > 0$ is a weighting parameter. We use the Accelerated Proximal Gradient [13] to solve (5). The clean low-rank matrix $\mathbf{A}$, computed separately for each color channel, will be used as input data in the subsequent steps.

The above preprocessing step offers us three major advantages: First, by isolating the errors, such as specular

highlights and shadows, our normal estimation algorithm is robust. Second, although we fix our captured object and camera physically, small misalignment errors during image data capture are inevitable and the preprocessing step makes our method robust to small misalignment errors by repairing the misaligned pixels with proximal values, which ensures the rank 3 property of matrix $\mathbf{A}$. Third, the low-rank matrix completion allows us to relax the strict requirement of a Lambertian surface reflectance model in comparing with the method in [12], as long as the non-Lambertian aspect can be factorized into the sparse residual matrix $\mathbf{E}$. Figure 3 compares results without and with low-rank matrix completion.

4.2. Normal refinement using least squares

We will now present our method for normal refinement. We will first assume that we know the lighting directions that contribute to the intensity of a pixel. In the next subsection, we will describe how to refine the contribution of each lighting direction given the surface normal. Hence, the steps for normal and lighting direction refinement will be performed in an alternating optimization fashion.

In order to deal with surface albedo, we follow the procedures in [28] to choose a denominator image I^d and to estimate the surface normals from ratio images

$$\frac{I^j}{I^d} = \frac{\sum_i^K c_i^j \mathbf{l}_i^j \cdot \mathbf{n}}{\sum_i^K c_i^d \mathbf{l}_i^d \cdot \mathbf{n}} \tag{6}$$

where I^j, $j = 1, \cdots, n-1$ are the input after low-rank matrix completion. From (6), we can re-write (3) as

$$\mathbf{An} = \mathbf{0} \tag{7}$$

where $\mathbf{A} = [I^j \sum_i^K c_i^d \mathbf{l}_i^d - I^d \sum_i^K c_i^j \mathbf{l}_i^j]$. The least square solution of $\mathbf{n}$ can be obtained by singular value decomposition (SVD), which explicitly enforces $||\mathbf{n}|| = 1$.

4.3. Illumination contribution refinement

Given the estimated normal direction, we now wish to refine the contribution of illumination direction that affects the pixel intensity. Without self-occlusion, this can be achieved by fitting a hemisphere of directions such that $\mathbf{l} \cdot \mathbf{n} > 0$ (Figure 2 (Right)). We propose a simple heuristic method to evaluate self-occlusion: If the normal direction between neighboring pixels forms a concave shape and the current normal direction is closer to $[0, 0, 1]^T$, it is likely that the incoming light from the neighboring directions is being occluded. Hence, we give a smaller weight to the light from that projected lighting direction. We evaluate self-occlusion for all directions within the local neighborhood and finally obtain a weighted mask for the contribution of illumination directions, which represents the relative contributions of the light from different directions.

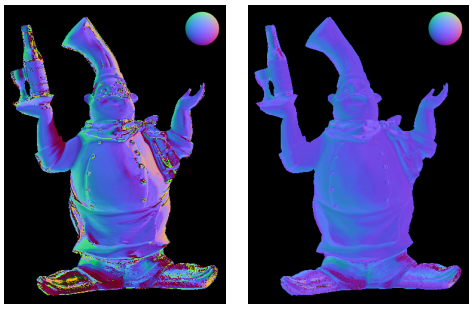


Figure 3. Self-comparisons of our results without (left) and with (right) low-rank matrix completion.

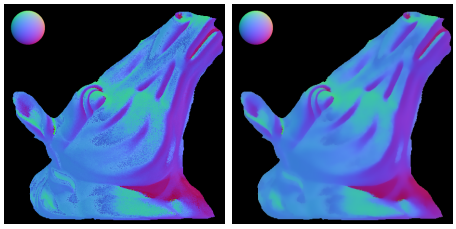


Figure 4. Self-comparisons of our results without (left) and with (right) TV regularization.

We initialize the normal direction and the corresponding hemisphere of environment illumination using exhaustive search, which minimizes the errors from the input images. The exhaustive search algorithm generally provides a good initialization, but it is slow if the search space is large. Therefore, we sample only 42 different normal directions for the exhaustive search initialization; i.e., an icosahedron subdivided once. This yields a good balance between accuracy and efficiency for our alternating optimization (AO) approach.

4.4. Spatial refinement using TV regularization

Thus far, our normal estimation method processes each pixel individually. As demonstrated in several previous works [28, 9, 19], spatial regularization is useful in error correction as well as in improving the overall accuracy of the estimated surface normals. In our postprocessing step, we employ L_1-norm vectorial total variation (TV) regularization to refine the estimated surface normals $\mathbf{n}^*$ obtained from the previous section, as follows:

$$\mathbf{n}^\star = \arg\min_{\mathbf{n}} ||\mathbf{n}^* - \mathbf{n}||^2 + \lambda |\nabla_\Omega \mathbf{n}|, \tag{8}$$

where $\nabla_\Omega \mathbf{n}$ is the vectorial first derivative of $\mathbf{n}$ defined over a local neighborhood in Ω and $\lambda = 0.1$ is the regularization weight (see [5] for the details).

Figure 7 shows intermediate results during the AO iterations and Figure 4 compares the result without and with

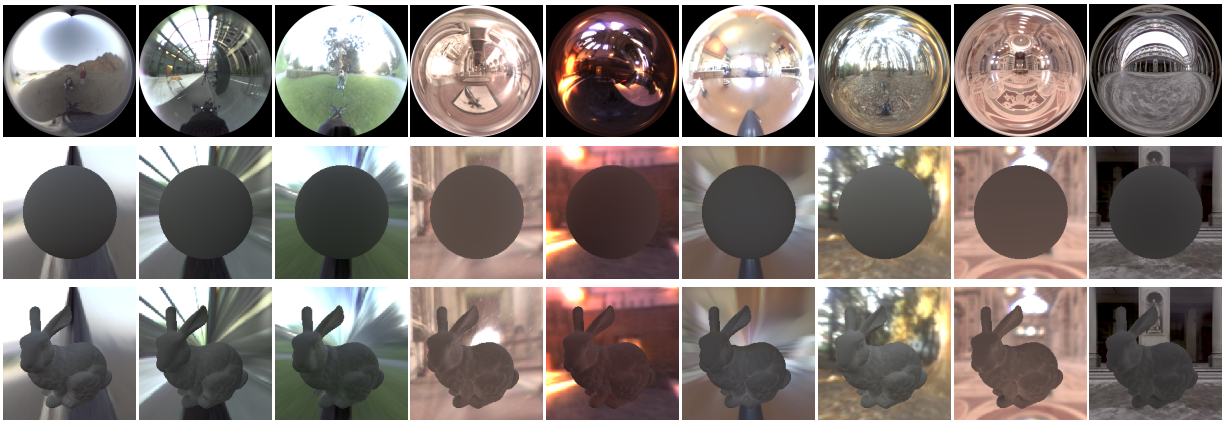


Figure 5. Input environments and images for the synthetic examples SPHERE and BUNNY.

spatial refinement with TV regularization. The TV regularization postprocessing step produces our final normal estimation results $\mathbf{n}^\star$, and we use the technique from [26] to reconstruct the 3D surface from $\mathbf{n}^\star$.

5. Experimental Results

We will now validate the efficacy of our proposed method in experiments using both synthetic and real objects.

5.1. Quantitative evaluation with synthetic images

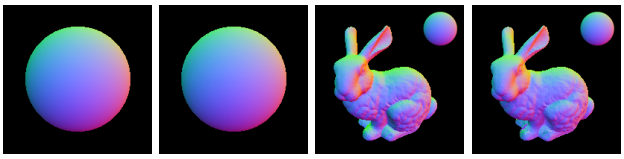
Our first experiment evaluates uses synthetic input images for which ground truth normal maps are available, and we analyze the effect of the number of input images and the convergence of our AO.

Two synthetic examples SPHERE and (Stanford) BUNNY were used for quantitative evaluation. We use the environment maps from [8] to render the synthetic input images as shown in Figure 5. We show the color coded ground truth normal maps and estimated normal maps in Figure 6 for qualitative comparison. Our approach faithfully estimates the surface normals, which closely approximate the ground truth normals with RMS error of 0.0099 and 0.1051 for the SPHERE and BUNNY, respectively.

To evaluate the robustness of our method, we plot the RMS error of the estimated normals with different numbers of input images in Figure 7. As expected, the RMS error decreases as the number of images increases, with less significant decreases forthcoming after more than 5 input images. Figure 7 plots the RMS error against the number of iterations, which shows that our approach converges in 4 to 5 iterations for both examples.

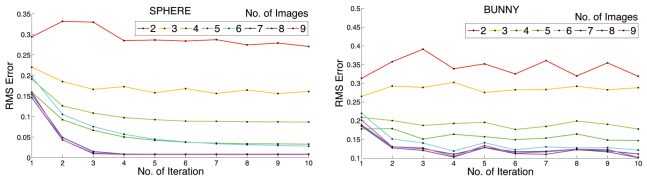
5.2. Qualitative evaluation with real images

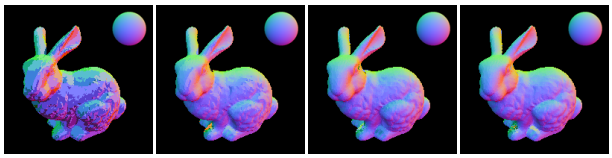
Next, we evaluate our framework on various real world examples under different illumination conditions, including different background scenes, an indoor scene with different indirect light sources, and an outdoor scene with sunlight direction varying throughout a day. The running times are indicated in Table 2.



(a) Ground truth (b) Result (c) Ground truth (d) Result

Figure 6. Comparison between ground truth and normal maps obtained using nine environments. The results are obtained after four iterations of the AO process.





Iteration 0 Iteration 1 Iteration 2 Iteration 3

Figure 7. Convergence analysis of our alternating optimization framework. Top: RMS error of the estimated normals versus the number of iterations, using different numbers of input images. Bottom: qualitative illustration of our intermediate results for BUNNY. Iteration 0 is the result after the exhaustive search initialization.

	#images	size	time (sec)
SPHERE	9	200 × 200	40.3
BUNNY	9	200 × 200	27.9
COUPLE	10	300 × 420	450.0
MOTHER&BABY	10	230 × 500	334.0
HORSEHEAD (INDOOR)	7	600 × 600	575.7
CHEF (INDOOR)	7	371 × 514	468.7
SHOE (INDOOR)	7	496 × 234	309.2
HORSE (SUNLIGHT)	8	260 × 347	140.8
CHEF (SUNLIGHT)	8	371 × 503	466.9

Table 2. The running times of our Matlab R2009b implementation were measured on a 3.33GHz Intel Xeon PC.

Figure 8. Input environments and images for the real examples COUPLE and MOTHER&BABY

Different background scenes. Our first real-world experiment closely mimics the synthetic experiments, by capturing the object in different illumination environments. We used 10 input images for the examples COUPLE and MOTHER&BABY. The input images are shown in Figure 8 and Figure 11 depicts the results. The reconstructed normal maps and surfaces appear faithful. We show the $\mathbf{l} \cdot \mathbf{n}$ images under two different lighting conditions to reveal the shading, and we show the images of the real objects alongside the reconstructed surfaces captured from a similar viewpoint. The corresponding closeup view illustrates the details preserved in the reconstructed surfaces. For example, the arms and legs of the COUPLE are clearly estimated and we can clearly see that the mother is holding her baby in MOTHER&BABY.

Indoor scene with different illumination conditions. Next, we consider in an indoor scene with different illumination conditions, by turning on/off different light sources in a room. Note the presence of ambient light and the use of indirect light sources (e.g., table lamp, floor lamp). We captured illumination environments and images for SHOE (INDOOR), HORSEHEAD (INDOOR) and CHEF (INDOOR), some of which are shown in Figure 9. Figure 11 shows the results. As can be seen from the closeup image, our results are very good under these conditions and subtle details such as the textures of the shoe are faithfully reconstructed.

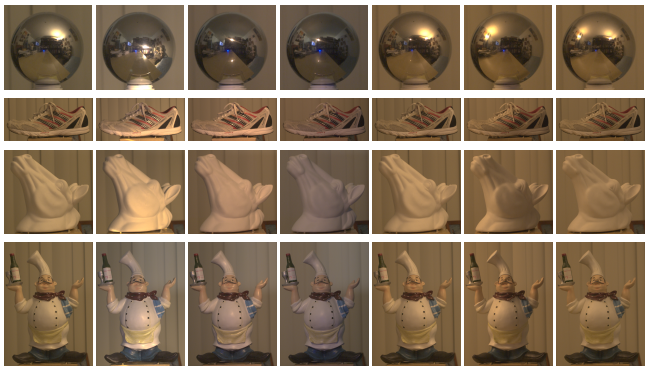


Figure 9. Input environments and images for the real examples SHOE (INDOOR), HORSEHEAD (INDOOR) and CHEF (INDOOR)

Figure 10. Input environments and images for the real examples CHEF (SUNLIGHT).

Outdoor scene with moving sunlight. Our final experiments were conducted in an outdoor environment using sunlight for reconstruction, which is the main goal of this project. We captured images of the HORSE (SUNLIGHT) and CHEF (SUNLIGHT) objects every hour from 10am to 5pm, obtaining eight input images per object. The input images and results are shown in Figure 1, Figure 10, and Figure 11. The results of CHEF (INDOOR) are shown alongside to facilitate comparisons. We find that although the results for the outdoor environment are reasonably good, they are not as good as those for the indoor environment. Part of the reason is the relatively modest variation in sunlight as the sun moves along its trajectory, whereas the light sources in the indoor environment are well distributed in different directions.

6. Conclusion and Future Work

We have presented a photometric stereo framework which employs natural environmental illumination, demonstrating the feasibility of practical outdoor photometric stereo. Featuring a simple setup for data capture with an optimization framework for dense object surface normal estimation, our system achieves high quality normal estimation even for complex indoor and outdoor scenes with natural illumination. By exploiting low-rank matrix completion

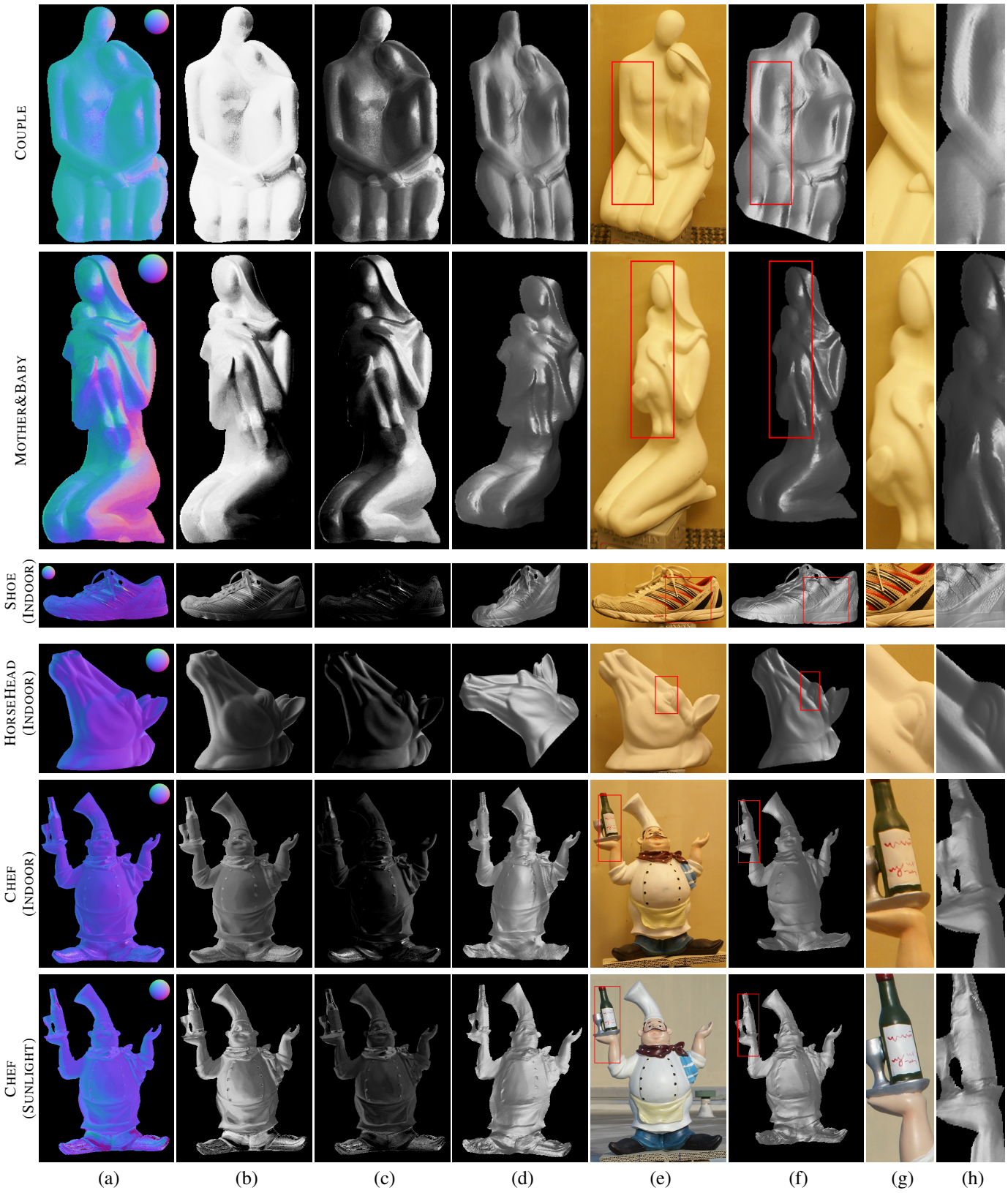


Figure 11. Real-world results. (a) Color-coded normal map. (b–c) Normal map shaded by $\mathbf{l} \cdot \mathbf{n}$ with $\mathbf{l} = [-1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]^T$ and $\mathbf{l} = [1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]^T$, respectively. (d) Novel view of the reconstructed surface. (e-h) Closeup view comparing the reconstructed surface with the real object from a similar viewpoint.

and total variation regularization techniques, our framework is robust to small object misalignment, shadows, and highlights. We believe that our framework has effectively mitigated the limitations of conventional photometric stereo algorithms, among them the need for a controlled environment for image data capture. In the future, we plan to extend our framework to incorporate non-Lambertian surface reflectance models. We also aim to integrate our framework with multi-view structure-from-motion algorithms in order to reconstruct high quality models of 3D objects in their entirety.

Acknowledgements

The research reported herein was partially supported by Singapore University of Technology and Design (SUTD) StartUp Grant ISTD 2011 016, by SUTD-MIT International Design Centre (IDC) Research Grant IDSF1200110OH, and by the National Research Foundation (NRF: 2012-0003359) of Korea funded by the Ministry of Education, Science and Technology. We thank Ka-Keung Lau for his kind assistance with equipment. We also thank Teresa Wan for narrating the demonstration video.

References

[1] J. Ackermann, F. Langguth, S. Fuhrmann, and M. Goesele. Photometric stereo for outdoor webcams. In *CVPR*, pages 262–269, 2012.

[2] D. Ballard and C. Brown. Computer vision. In *Prentice Hall*, 1982.

[3] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. on PAMI*, 25(10):1239–1252, October 2003.

[4] R. Basri, D. W. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *IJCV*, 72(3):239–257, 2007.

[5] X. Bresson and T. F. Chan. Fast dual minimization of the vectorial total variation norm and applications to color image processing. Technical report, UCLA CAM Report, 2008.

[6] M. Chandraker, J. Bai, and R. Ramamoorthi. A theory of differential photometric stereo for unknown BRDFs. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2505–2512, 2011.

[7] E. Coleman, Jr. and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *CGIP*, 18(4):309–328, April 1982.

[8] P. E. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *ACM SIGGRAPH*, pages 189–198, 1998.

[9] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying brdfs from photometric stereo. *IEEE Trans. on PAMI*, 32(6):1060–1071, 2010.

[10] A. Hertzmann and S. Seitz. Shape and materials by example: a photometric stereo approach. In *CVPR*, pages I: 533–540, 2003.

[11] B. Horn. *Robot Vision*. McGraw-Hill, 1986.

[12] M. K. Johnson and E. H. Adelson. Shape estimation in natural illumination. In *CVPR*, pages 2553–2560, 2011.

[13] Z. Lin, A. Ganesh, J. Wright, L. Wu, M. Chen, and Y. Ma. Fast convex optimization algorithms for exact recovery of a corrupted low-rank matrix. Technical Report UILU-ENG-09-2214, UIUC, 2009.

[14] Z. Lu, Y.-W. Tai, M. Ben-Ezra, and M. S. Brown. A framework for ultra high resolution 3d imaging. In *CVPR*, 2010.

[15] S. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. on Robotics and Automation*, 6(4):418–431, 1990.

[16] G. Oxholm and K. Nishino. Shape and reflectance from natural illumination. In *ECCV (1)*, pages 528–541, 2012.

[17] R. Ramamoorthi and P. Hanrahan. An efficient representation for irradiance environment maps. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2001)*, 20(3):497–500, 2001.

[18] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[19] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. In *CVPR*, 2010.

[20] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Trans. on PAMI*, 18(4):449–454, April 1996.

[21] K. Sunkavalli, T. Zickler, and H. Pfister. Visibility subspaces: Uncalibrated photometric stereo with shadows. In *ECCV*, pages 251–264, 2010.

[22] H. Tagare and R. deFigueiredo. A theory of photometric stereo for a class of diffuse non-lambertian surfaces. *IEEE Trans. on PAMI*, 13(2):133–152, February 1991.

[23] R. Woodham. Photometric method for determining surface orientation from multiple images. *Opt. Eng.*, 19(1):139–144, January 1980.

[24] R. Woodham. Gradient and curvature from the photometric-stereo method, including local confidence estimation. *JOSA-A*, 11(11):3050–3068, November 1994.

[25] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *ACCV*, pages 703–717, 2010.

[26] T.-P. Wu, J. Sun, C. Tang, and H. Shum. Interactive normal reconstruction from a single image. *ACM Trans. Graph.*, 27(5), 2008.

[27] T.-P. Wu and C.-K. Tang. Photometric stereo via expectation maximization. *IEEE Trans. on PAMI*, 32(3):546–560, 2010.

[28] T.-P. Wu, K.-L. Tang, C.-K. Tang, and T.-T. Wong. Dense photometric stereo: A markov random field approach. *IEEE Trans. on PAMI*, 28(11):1830–1846, 2006.

[29] S.-K. Yeung, T.-P. Wu, C.-K. Tang, T. F. Chan, and S. Osher. Adequate reconstruction of transparent objects on a shoestring budget. In *CVPR*, pages 2513–2520, 2011.

[30] L.-F. Yu, S.-K. Yeung, Y.-W. Tai, and S. Lin. Shading-based shape refinement of rgb-d images. In *CVPR*, 2013.