A Streamlined Photometric Stereo Framework for Cultural Heritage

Chia-Kai Yeh[1]([✉]), Nathan Matsuda[1], Xiang Huang[3], Fengqiang Li[1],
Marc Walton[2], and Oliver Cossairt[1]

[1] Department of EECS, Northwestern University, Evanston, USA
chiakaiyeh2017@u.northwestern.edu
[2] Northwestern University/Art Institute of Chicago Center for Scientific
Studies in the Art (NU-ACCESS), Evanston, USA
[3] Argonne National Laboratory, Lemont, USA

Abstract. In this paper, we propose a streamlined framework of robust $3D$ acquisition for cultural heritage using both photometric stereo and photogrammetric information. An uncalibrated photometric stereo setup is augmented by a synchronized secondary witness camera co-located with a point light source. By recovering the witness camera's position for each exposure with photogrammetry techniques, we estimate the precise $3D$ location of the light source relative to the photometric stereo camera. We have shown a significant improvement in both light source position estimation and normal map recovery compared to previous uncalibrated photometric stereo techniques. In addition, with the new configuration we propose, we benefit from improved surface shape recovery by jointly incorporating corrected photometric stereo surface normals and a sparse $3D$ point cloud from photogrammetry.

Keywords: Photometric stereo · Reflectance transformation imaging · Near light position calibration · Photogrammetry · $3D$ surface shape reconstruction

1 Introduction

Computational Imaging techniques have been widely used for art history analysis and cultural heritage research in the last decade. Digital imaging technologies empower conservation scientists by revealing more information about works of art, helping to better preserve and protect their history for future generations. Accurate, automatic $3D$ surface recovery using only commodity cameras is particularly important for a number of applications in cultural heritage research. Since artifacts of historical significance are often located in public spaces or museums without the possibility of relocation to a laboratory environment, art conservators require $3D$ shape acquisition techniques that are portable, inexpensive, non-destructive, and fast, in order to uncover previously unknown information about artist techniques and materials. Two commonly used techniques that
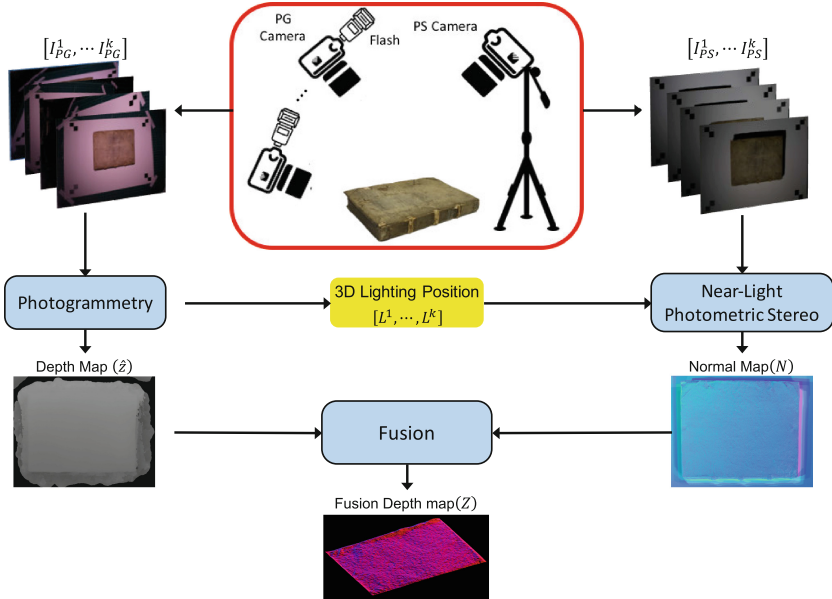
Fig. 1. Overview of the Streamlined photometric stereo framework for cultural heritage: We use photogrammetry to find the $3D$ light positions $[L^1, \cdots, L^k]$ relative to a stationary photometric stereo (PS) camera. The estimated $3D$ light positions then allow us to compute accurate surface normal N from the PS camera. We fuse the computed normal map with a depth map $\hat{z}$, computed using photogrammetry, to generate globally accurate $3D$ shapes Z with high-quality micro surface details.

fit these requirements are Reflectance Transformation Imaging (RTI) and Photogrammetry (PG).

RTI is a visualization technique that allows users to probe the appearance of an artwork under arbitrary illumination conditions computationally, in a post-processing step. RTIs are created from multiple photographs of the object captured by a camera with fixed position and varying illumination. Researchers use RTI to virtually re-light an object under arbitrary illumination conditions. Computational relighting can reveal fine details of the subject's $3D$ surface, for instance when strong raking light is used to visualize the surface appearance. However, because RTI is merely a visualization technique, it provides no direct access to depth information. Photometric stereo (PS) is a well established research topic in computer vision which estimates surface normal from a set of photographs taken with a fixed camera position and multiple known lighting directions. Intensity values in the captured images are modeled as a function of lighting angle, surface normal, and material reflectance. By inverting this model, PS techniques recover surface normal, which can then be integrated to produce $3D$ Surface shape.

Alternatively, Photogrammetry (PG) uses images taken at different camera positions, using triangulation to compute $3D$ surface shape. Using Structure from Motion (SfM) techniques, feature points common between/among multiple views can be used to jointly solve for both $3D$ location of the points and the corresponding camera positions. The resulting $3D$ result is globally accurate but not dense, since $3D$ information is computed only for each feature point, but not each pixel. These sparse $3D$ points can be interpolated to generate a low-resolution $3D$ mesh model of the object.

PS and PG techniques have been explored extensively in the literature, but still have fundamental drawbacks. For example, accurate PS normal output usually requires pre-calibrated lighting positions. In typical setups, this is achieved using either lighting with a fixed calibrated $3D$ geometry (e.g. a lighting dome), or by placing a reflective sphere in the scene to estimate incident lighting directions. $3D$ light position can be accurately pre-calibrated using a lighting dome, but this custom hardware solution is often inaccessible and sometimes impractical. A reflective sphere can accurately measure distant lighting, but produces significant errors when light sources violate the far light condition and are actually located near the object (e.g. within 4 times the size of the object), typical of many PS capture setups [13]. PG techniques do not require controllable lighting, but do require a high number of identifiable correspondence points in order to produce high resolution surface output, precluding the possibility of capturing low-texture or single-material objects frequently found in a wide variety of natural scenes. Furthermore, at large standoff distances, depth precision for PG methods is relatively coarse while PS solutions are capable of capturing highly detailed depth features.

In this paper, we present a robust $3D$ shape recovery capture framework for cultural heritage as shown in Fig. 1. Throughout the remainder of the paper we will refer to these two cameras as the PS camera, capturing reflectance information from a fixed position, and the PG camera, affixed to the light source and capturing scene structure for photogrammetry from multiple views of the object. The PG camera images are processed using existing SfM algorithms to recover the camera position for each frame, and thus the lighting positions for the PS camera as well. Using these computed $3D$ lighting positions we then produce an accurate PS normal map. Because we have generated a point cloud from the PG algorithm as well, we can fuse this sparse $3D$ information with the PS normal map to produce a $3D$ surface with both the fine surface detail typical of PS techniques and the absolute depth accuracy typical of PG techniques. The technique introduces minimal complexity beyond a conventional photometric stereo capture setup, yet can be used to significantly improve the accuracy of $3D$ surface reconstructions.

The specific contributions of this work are:

– **A simple, robust** $3D$ **capture system:** We present a simple system for the free-form photometric stereo capture system using just two camera with wireless synchronize triggers and a on-camera ring light. We show that our

system simplifies reflectance capture and results in more accurate $3D$ surface reconstruction.

– **More accurate light position estimation:** Previous techniques estimate $3D$ light position directly from images from radiometric measurements [13], which are easily corrupted by shadows and specularities. In contrast, our light position estimation is based on geometric triangulation using SfM, and is therefore largely independent of scene reflectance and illumination.
– **Improved near-light PS surface recovery:** Traditional PS techniques assume infinitely distant light sources. Under this assumption, the lighting direction can be calibrated by placing a mirror ball in the scene. Our approach removes this far light assumption and eliminates the need for a lighting calibration object. Instead, $3D$ light position is estimated using a PG camera attached to the light source. We show that by accurately measuring the $3D$ location of the light sources, we can recover more accurate $3D$ surface shapes when using a PS setup that violates the far light assumption.
– **Large scale, high precision $3D$ reconstructions:** We show experimentally that our setup can be used to generate large field of view $3D$ shape reconstructions with high precision. This is done by fusing the fine details from dense normal estimation using PS, with the sparse $3D$ point clouds from our PG camera.

2   Previous Work

2.1   Reflectance Transformation Imaging

Reflectance transformation imaging is widely popular among art conservators through the use of the CHI RTI Builder and Viewer software suites [1]. RTI, originally known as Polynomial Texture Mapping (PTM), was first proposed by Malzbender [15] as a way to use a polynomial basis function for computational relighting. Later, the hemispherical harmonics (HSH) version [8] was introduced to reduce the directional bias in computational relighting results. Palma *et al.* [17] estimated normal from PTM RTIs by fitting the pixel intensity to a local bi-quadratic function of light angles and then setting the derivative to zero, which has the effect of finding the direction of the brightest pixel. Conservators use the CHI software to interactively explore image relighting and normal maps in the RTI Viewer, and also export those images offline for further research.

2.2   Photometric Stereo

In the original photometric stereo formulation introduced by Horn [12], light sources are assumed infinitely distant, the camera is orthographic, and the object surface is Lambertian and convex (*i.e.* no shadows or inter-reflections). Subsequent research has sought to generalize the technique for more practical camera, surface and lighting models. Belhumeur *et al.* [6] discovered that with an orthographic camera model and uncalibrated lighting, the object's surface can be

uniquely determined to within a bas-relief ambiguity. Papadhimitri and Favaro *et al.* [18] recently pointed out that this ambiguity is resolved under the perspective camera model. Several researchers have also sought to relax the Lambertian reflectance assumption and incorporate effects such as specular highlights and shadows. New techniques have been introduced based on non-Lambertian reflectance models [5,10,11], or sophisticated statistical methods to automatically filter non-lambertian effects [14,26,27]. However, less attention has been paid to relaxing assumptions on the lighting model. Several researchers [13,19,24] recently investigated removing the far-light assumption to improve the accuracy of photometric stereo. Others consider non-isotropic illuminations [20]. Ackermann *et al.* [2] recently gave a more comprehensive surveys on earlier and recent photometric stereo techniques.

2.3   Photogrammetry

Developed in the 1990s, this technique has its origins in the computer vision community and the development of automatic feature-matching algorithms from the previous decade. To determine the $3D$ location of points within a scene, traditional photogrammetry methods require the $3D$ location and pose of the cameras, or the $3D$ location of a series of control points to be known. Later, Structure-from-Motion (SfM) relaxed this requirement, simultaneously reconstructing camera pose and scene geometry through the automatic identification of matching features in multiple images [22,23].

2.4   Combining Photometric Stereo and Photogrammetry

Although PS provides relatively accurate surface normal, it is still challenging to reconstruct a globally accurate surface shape. Some work has aimed to combine PG and PS techniques, such as the multi-view photometric stereo method by Hernandez *et al.* [9], which used RANSAC to estimate the light sources position and reconstruct $3D$ surfaces of Lambertian objects. For calibrated light sources, Birkbeck *et al.* [7] employed a variational method to estimate the surface and handle specular reflections using a Phong reflectance model. Ahmed *et al.* [4] used calibrated illumination and multi-view video to capture normal fields and improve the geometry templates. Wu *et al.* [25] performed a spherical harmonic lighting approximation to combine multi-view photometric stereo. Sabzevariuse *et al.* [21] used the $3D$ metric information computed with SfM from a set of $2D$ landmarks to solve for the bas-relief ambiguity for dense PS surface estimation. All of these algorithms require really critical environment constraint either accurate light-source calibration under far light model or careful illumination design. Nehab *et al.*'s [16] hybrid reconstruction algorithm focused on leveraging Poisson system to combine depths and normal information. Their fusion algorithm produces high quality reconstruction of $3D$ surfaces with a given parametric surface.

Our method relaxes the hardware setup constraints relative to these prior methods. To our knowledge ours is the first system to work on fusion between
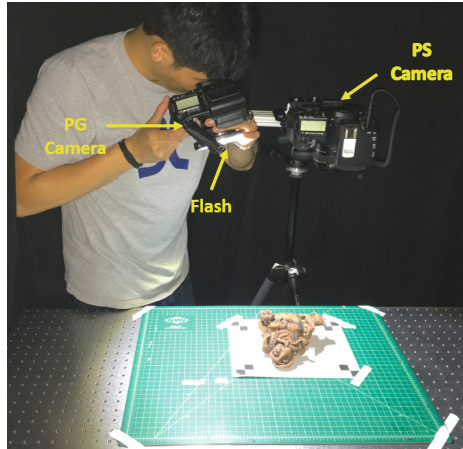
Fig. 2. Capture setup: We use two Canon $5D$ Mark III cameras with $50\,\text{mm}$ prime lens. The PS camera is placed about $0.5\,\text{m}$ away from the object.

near-light PS model and PG. Besides having more accurate light position and surface normal estimates, our method also can leverage the surface estimate obtained using photogrammetry. By fusing PS and PG results we can produce an improved $3D$ surface that retains the advantages of both PS and PG techniques.

3   Our Streamlined Photometric Stereo Framework

3.1   Hardware Setup

Our system setup consists of two Canon 5D Mark III DSLR cameras with $50\,\text{mm}$ Canon Prime lenses. One of these, the PS camera was affixed to a tripod above the imaging area. A Polaroid 18 Super Bright Macro SMD LED Ring Light was mounted to the PG camera lens. Both cameras were attached to a PocketWizard FLex TT5 wireless trigger system to ensure synchronized exposures. Lastly a printed set of corner fiducial makers were affixed to the imaging area to provide a means to scale the PS and PG image sets to match the physical distances between the markers.

3.2   Framework Work Flow

We begin by capturing an image at each of k different PG camera positions (see Fig. 2). A ring light is placed around the lens of the PG camera so that the centroid location of the illumination coincides with the optical center of the lens. The PG camera captures a set of images $[I_{PG}^1, ..., I_{PG}^k]$ of the scene from a unique viewing location. The PS camera also captures k images $[I_{PS}^1, ..., I_{PS}^k]$, but from a fixed position. For the PG camera, illumination is always aligned with the camera axis. For the PG camera, a diversity of illumination directions is captures. The

PG images $[I_{PG}^1, ..., I_{PG}^k]$ are input into an off-the-shelf photogrammetry software Agisoft PhotoScan [3], which outputs the camera centers corresponding to the $3D$ light source positions $[L^1, ..., L^k]$. In addition, the software compute a sparse point cloud estimate of the objects $\hat{z}$. An image from the PS camera is input together with the PG camera images so that the extrinsic parameters from all cameras are determined in a unified global coordinate frame. Note that our PG images do not all have the same lighting and contain specularity and shadows under different lighting environment, none of which is ideal for typical passive multi-view stereo matching. However, we have sufficiently dense views under similar-enough lighting for the matching algorithm to find enough matching features between the images to reconstruct a photogrammetry model which is accurate to within a few millimeters.

Next, the $3D$ light positions $[L^1, ..., L^k]$ are used as input into a PS algorithm to accurately recover normal and albedo based on the spatially-varying incident lighting position at each point in the scene. To accomplish this, we solve a least squares problem to iteratively solve for the albedo a and normal N, given captured images $[I_{PS}^1, ..., I_{PS}^k]$ and corresponding 3D light positions $[L^1, ..., L^k]$, similar to the work by Papadhimitri *et al.* [19].

Finally, the PS algorithm generates a normal map $N = (n_x, n_y, n_z)$ for each pixel in the image. I, The relationship between the estimated normal and the depth map z is then $(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}) = (p, q)$, where $(p, q) \triangleq (-\frac{n_x}{n_z}, -\frac{n_y}{n_z})$. The PG algorithm produces a depth map $\hat{z}$ of the scene only for a sparse subset of pixels. We assume $\hat{z}$ is transformed to the PS camera frame using the extrinsic parameters computed from the PG/SfM software. We then recover the PS-PG fused depth z_i for each pixel i by solving the following least squares problem:

$$\text{minimize} \sum_{i \in I} \left\| \begin{bmatrix} \partial z_i / \partial x_i \\ \partial z_i / \partial y_i \end{bmatrix} - \begin{bmatrix} p_i \\ q_i \end{bmatrix} \right\|_2^2 + \lambda \sum_{i \in \hat{I}} (z_i - \hat{z}_i)^2 \tag{1}$$
$$= \|\nabla Z - \Gamma\|_2^2 + \lambda \|MZ - \hat{Z}\|_2^2,$$

where Z, $\hat{Z}$, and Γ are the lexicographically vectorized versions of z_i, $\hat{z}_i$, and (p_i, q_i), ∇ is the gradient matrix, M is a binary selection matrix that only selects the pixels that have valid PG depths, and λ is the parameter depends on the confidence of PG depth.

Note that this formulation does not rely on any linear constraints or statistical priors; it is simply a weighted least-squares approach that attempts to satisfy, on average, the conditions observed by both the PS and PG recovery techniques. A wide variety of variations on this optimization could be employed depending on the type of object and intended usage of the recovered surface, but a detailed analysis of such possibilities is beyond the scope of this paper. We simply aim to demonstrate that the combination of both sets of simultaneously captured data, even with a rudimentary approach to optimization, characterizes the surface significantly better than either approach alone.

Table 1. Measure value *v.s.* **ground truth for three light positions P1, P2, and P3:** The first row shows the ground truth distance between the PG camera optical center and the $3D$ location of light sources P1, P2, and P3. The second and third rows show ΔL values for our technique and that of Huang *et al.* [13], respectively. The ΔL values reported are the distance between the estimated $3D$ position of the PG camera's optical center, and the ground truth $3D$ position, averaged over five measurements. The fourth and fifth rows report the standard deviation of the distance between the estimated and ground truth $3D$ location of the PG camera.

	P1	P2	P3
Ground truth distance (mm)	869.98	896.28	756.12
ΔL (Our method) (mm)	1.25	9.09	8.27
ΔL (Huang *et al.* [13]) (mm)	242.46	239.76	216.11
SD δ (Our method) (mm)	0.26	0.23	0.09
SD δ (Huang *et al.* [13]) (mm)	2.09	4.18	2.09

4   Experiments and Results

4.1   Light Position Estimation

First, we evaluated the accuracy and stability for our PG camera-based method for light position estimation. In order to compare to known physical lighting positions, we affixed the tripod mount of the PG camera onto an optical mounting post, which we then inserted sequentially into optical post holders at known locations on an optical table. Though we do not consider this manual procedure sufficient to provide ground truth data, the sub-millimeter tolerances of the machined optical table and mounting posts can demonstrate the extent to which the recovered lighting positions can be relied upon.

In Table 1, we repeated the three fixed lighting positions 5 times, which resulted in an average error relative to our measured positions of less than 10 mm, or well under 1 % error. The standard deviation of these values was less than 1 mm, indicating good repeatability of the technique. Compared to Huang *et al.* [13] using image intensity to estimate the lighting position, our approach using PG/SfM has more accurate lighting position estimation for a near-light photometric stereo model.

4.2   Normal Map Accuracy

To confirm that PG lighting position estimation produces a more accurate PS normal map, we compare normal map recovery for a sphere using our method, the near-light model in Huang *et al.* [13], a conventional distant-light PS model, and ground truth.

Figure 3 shows the X-component of the normal map sampled through the center of a sphere for the ground truth, conventional distant-light PS model, near-light model from Huang *et al.* [13], and our PG light estimation. Our method clearly demonstrates increased fidelity in normal map estimation.
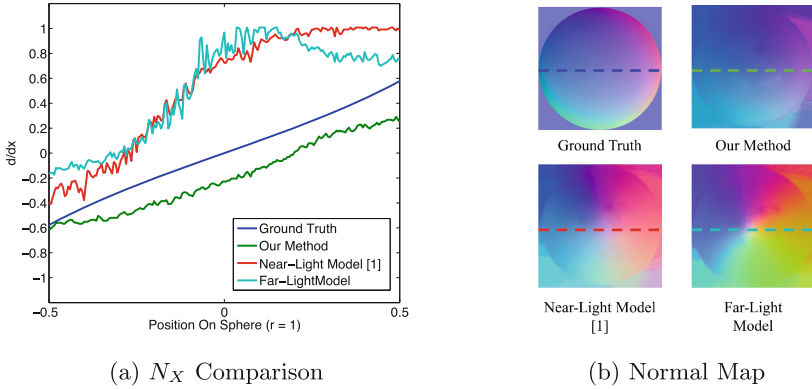
(a) N_X Comparison   (b) Normal Map

Fig. 3. Normal Map Accuracy for a sphere: Comparison between the x component of the estimated normal map for a sphere. The ground truth (shown in blue) normal for the sphere closely resembles a line (the gradient of a parabola is exactly a line). The normal estimate computed using the far light assumption (shown in cyan) and the uncalibrated photometric stereo method from Huang et al. [13] (shown in red) both produce significant errors. Our method (shown in green) accurately estimates $3D$ light position, and therefore produces the most accurate $3D$ normal. (Color figure online)

This method is a unique use case for PG techniques in surface reconstruction because it can be applied to texture less objects that would normally be a failure case for PG. So long as there is sufficient correspondence features to perform bundle adjustment somewhere in the PG camera field of view, our technique will produce accurate lighting positions, and thus more accurate normal maps, regardless of the amount of texture in the target object.

4.3   Fusion Surface Reconstruction

When objects have enough surface texture for the PG algorithm alone to produce a sparse point cloud, we can leverage this data for a more globally accurate surface reconstruction. Surface shape recovery remains a significant challenge for all PS techniques since small errors in normal recovery will produce incorrect geometry upon integration, and the absolute position of the surface can never be recovered. The formulation in Eq. 1 retains the fine surface detail recovered by PS and the gross geometric shape recovered by PG.

We chose to test the visual fidelity of surface fusion reconstructions using a cultural heritage object from our University's rare book collection, an object representative of the intended use case for this technique. Shown in Fig. 4, this 16th century reprinting of Hesiod's 'Works and Days', was covered with a reused parchment from an early manuscript that was scraped down to remove the letters from the top surface. Small ridges on the surface are aligned with the direction of the scraping motion. We hope to observe these abrasions in the context of the largely flat overall surface geometry. PS techniques alone will not retain the

course flatness but will reveal the small ridges, while PG techniques alone will retain the flat surface but will not resolve the ridges at all. This object is thus an example of a surfaces our PS and PG fusion technique is well suited to recover.

Fig. 4. Test object: a 16*th* century book covered with reused parchment. Small surface abrasions on the surface are of interest to historians.

The λ parameter in Eq. 1 was set to 0.15, a value found experimentally that retained surface detail while preventing the large-scale PS errors to propagate into the final output.

In Fig. 5 we show side-by-side comparisons between the full surface and an inset revealing small details. The top row contains a reference image from the PS data set - the full book surface on the left, followed by the pink inset region expanded on the right. These regions are used in subsequent rows, where surface reconstructions are depicted in orthographic renders using a white Lambertian material and raking angle lights to highlight surface variation in blue along the y-axis and red along the x-axis. The 2*nd* row shows the surface output from the photometric stereo algorithm, which despite recovering small surface details exhibits extreme geometric errors which would significantly limit any object analysis based on the surface height. The 3*rd* row shows the PG surface mesh output from Agisoft Photoscan. The PG results correctly recover the general flatness of the object, but lose all fine surface detail. Finally in the bottom row we show our optimized PS+PG fusion results. We retain both the overall flat shape of the book surface while recovering the small wrinkles and abrasions present on the surface of the book.

In order to test our framework in more general settings, we have tested our method on several additional objects with complex geometry and fine surface detail. As shown in Fig. 6, our framework produces accurate 3*D* reconstructions that maintaining both global accuracy and high precision. The results are far superior to 3*D* reconstructions using either PS or PG alone.
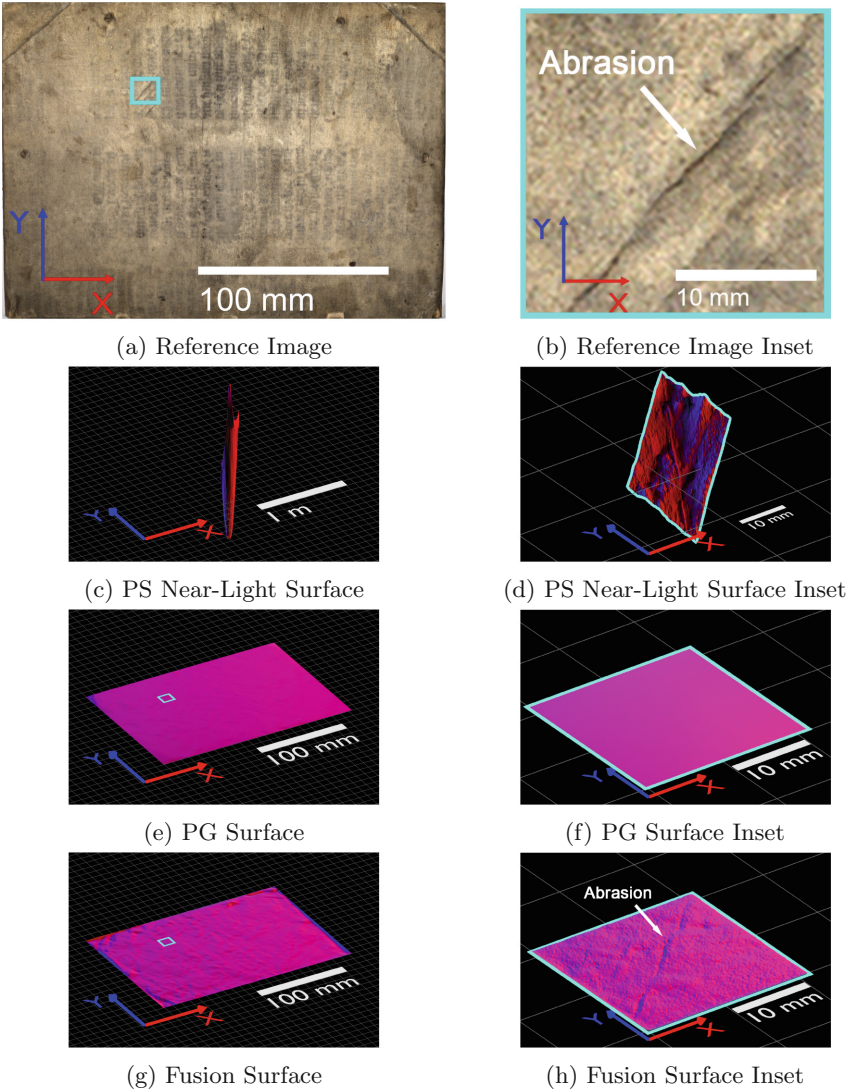
(a) Reference Image

(b) Reference Image Inset

(c) PS Near-Light Surface

(d) PS Near-Light Surface Inset

(e) PG Surface

(f) PG Surface Inset

(g) Fusion Surface

(h) Fusion Surface Inset

Fig. 5. Reconstruction Results: Comparison of reconstruction methods on a $16th$ century book shown in (a), and hi-resolution inset (b), corresponding to the outlined region to the left. After surface recovery, these results are depicted in orthographic perspective and illuminated by a red directional light along the x-axis and a blue directional light along the y-axis to reveal surface details without exaggerating the scale of the z-axis. The PS reconstructions using the method from [13], shown in (c) and (d), exhibit severe global geometry errors due to lack of absolute reference points (the scale in these images were reduced to accommodate the extreme range of z-axis values). PG output from Agisoft Photoscan is shown in (e) and (f). Our fusion results, produced by optimizing the surface for consistency with both PS and PG results are shown in (g), (h). Note that the fusion results exhibit a balance of course geometric accuracy (a flat book surface) while retaining small surface variations. (Color figure online)
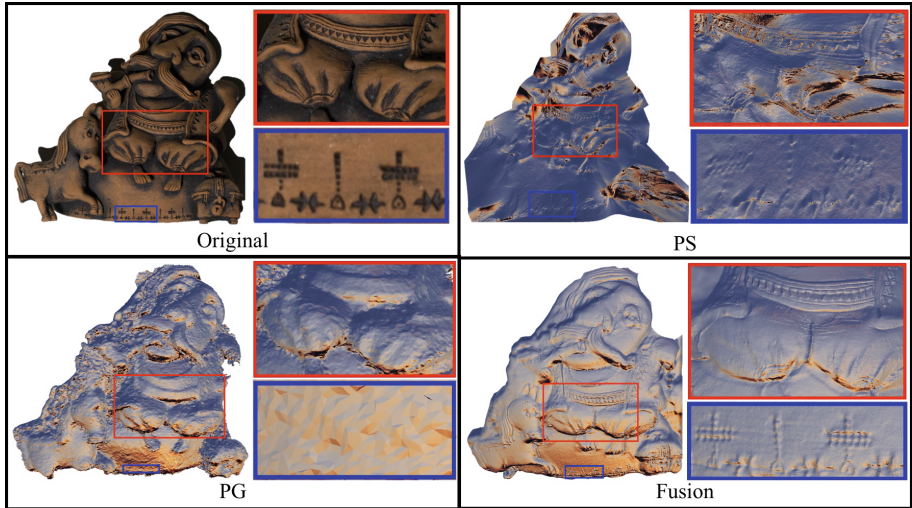
(a) Hutsul Ceramics Ukraine Terracotta Sculpture
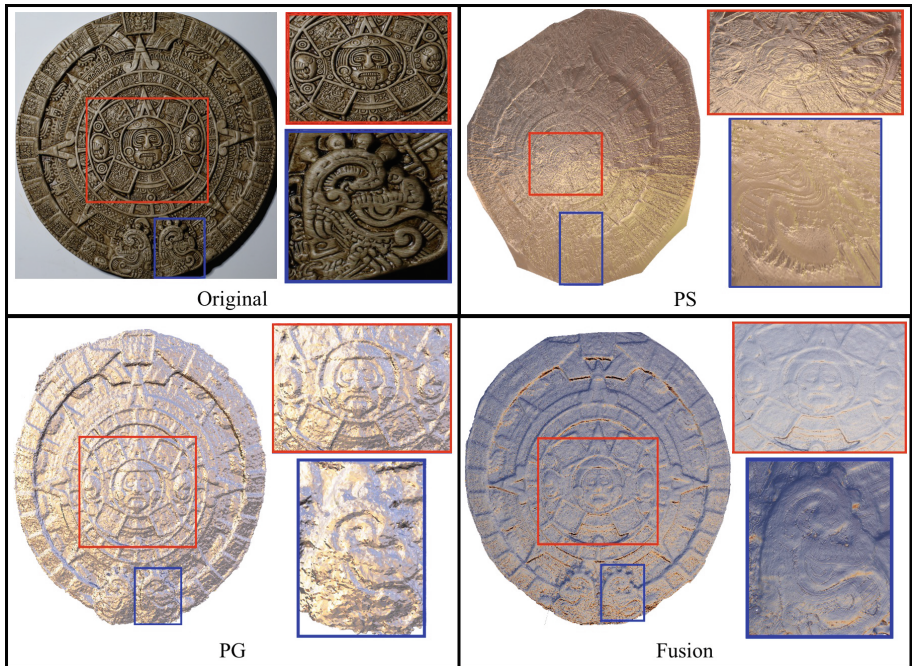
(b) Chinese Jade BI Carvings

Fig. 6. Experimental Results using our Framework: We tested our framework on several objects with complex geometry and fine surface detail. These objects demonstrate that our system produces a good balance between global geometric accuracy and micro surface details. $3D$ reconstruction results using only photometric stereo (PS), and photogrammetry (PG) are shown for comparison. Our fusion results clearly demonstrate superior $3D$ reconstruction quality.

5   Conclusion and Future Work

We have presented a new technique using a PG camera attached to a flash light source to estimate $3D$ lighting positions for more robust photometric stereo $3D$ surface reconstructions. The resulting light position estimates are more accurate than conventional far-light directional estimates or near-light position estimates, and consequently produce more accurate normal maps. We also demonstrate that the PG surface information can be fused with the PS normal map output for surface reconstruction that retains both the fine details from PS and accurate global geometry from PG. We have demonstrated how to use a simple setup to acquire high quality $3D$ reconstruction results of several cultural heritage objects. Our initial results also give rise to another question: if fusion between poor normal recovery and good PG data produces a reasonable result, is the improved PS performance by accurate light position estimation even necessary? Further analysis is necessary to conclusively compare our results to fusion results that do not attempt to improve PS performance, but we believe that at the very least better input data from PS will not perform worse than other fusion methods, and is likely in most cases to perform better. We hope our method will empower conservators and conservation scientists with new tools for simple, inexpensive, $3D$ acquisition of cultural heritage artifacts. It is our belief that doing so will open the doors to new applications in monitoring the deterioration of objects and help inform new methods of damage prevention and preservation.

There are several possible directions for future work. Photometric stereo, technically, is a fix-view $2.5D$ reconstruction method that could not deal with the scene with lots of depth changes. In the future, we are interested in merging multi-view information to account for artifacts that photometric stereo can create and produce a high quality surface detail model. On the other hand, our light source estimation method could be extended to non-point or non-isotropic light sources, an extension applicable to nearly all real-world use cases. By performing PG camera pose estimation on both the PS camera and the PG camera, the full surface of a convex or more complicated surface shape may be recovered. From a systems perspective, the PG camera and flash component could be miniaturized (e.g. replaced with a point-and-shoot camera) to allow for greater freedom by the operator and quicker overall capture times. Two or more of these camera/flash units could be synchronized and processed to capture bidirectional reflectance information and ultimately used to recover more sophisticated material characterization jointly with surface shape. We are also interested in investigating more sophisticated PS algorithms that can handle difficult cases such as shadows and non-lambertian reflectance. Last but not least, although it is quite difficult to have a real ground truth to benchmark a $3D$ reconstruction system, still we would like to compare our framework with the state of art $3D$ acquisition method on cultural heritage application in the nearly future.

References

1. Cultural heritage imaging: reflectance transformation imaging (rti) (2013). http://culturalheritageimaging.org/Technologies/RTI/index.html
2. Ackermann, J., Goesele, M.: A survey of photometric stereo techniques. Found. Trends Comput. Graph. Vis. **9**(3–4), 149–254 (2015). http://dx.doi.org/10.1561/0600000065
3. Agisoft LLC: Agisoft PhotoScan. http://www.agisoft.com
4. Ahmed, N., Theobalt, C., Dobrev, P., Peter Seidel, H., Thrun, S.: Robust fusion of dynamic shape and normal capture for high-quality reconstruction of time-varying geometry. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8 (2008)
5. Alldrin, N., Zickler, T., Kriegman, D.: Photometric stereo with non-parametric and spatially-varying reflectance. In: 26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR (2008)
6. Belhumeur, P.N., Kriegman, D.J., Yuille, A.L.: The bas-relief ambiguity. IJCV **35**(1), 33–44 (1999)
7. Birkbeck, N., Cobzas, D., Sturm, P., Jagersand, M.: Variational shape and reflectance estimation under changing light and viewpoints. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006, Part I. LNCS, vol. 3951, pp. 536–549. Springer, Heidelberg (2006)
8. Elhabian, S.Y., Rara, H., Farag, A.A.: Towards accurate and efficient representation of image irradiance of convex-Lambertian objects under unknown near lighting. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1732–1737 (2011)
9. Esteban, C.H., Vogiatzis, G., Cipolla, R.: Multiview photometric stereo. IEEE Trans. Pattern Anal. Mach. Intell. **30**(3), 548–554 (2008)
10. Goldman, D.B., Curless, B., Hertzmann, A., Seitz, S.M.: Shape and spatially-varying BRDFs from photometric stereo. IEEE Trans. Pattern Anal. Mach. Intell. **32**(6), 1060–1071 (2010)
11. Hertzmann, A., Seitz, S.: Shape and materials by example: a photometric stereo approach. In: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 1–8 (2003)
12. Horn, B.K.P.: Obtaining shape from shading information. In: The Psychology of Computer Vision, pp. 115–155 (1975)
13. Huang, X., Walton, M., Bearman, G., Cossairt, O.: Near light correction for image relighting and 3D shape recovery. In: Guidi, G., Scopigno, R., Brunet, P. (eds.) International Congress on Digital Heritage - Theme 2 - Computer Graphics and Interaction. IEEE (2015)
14. Ikehata, S., Wipf, D., Matsushita, Y., Aizawa, K.: Robust photometric stereo using sparse regression. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 318–325 (2012)

15. Malzbender, T., Gelb, D., Wolters, H.: Polynomial texture maps. In: Proceedings of SIGGRAPH 2001. Annual Conference Series, pp. 519–528. ACM Press, New York (2001)

16. Nehab, D., Rusinkiewicz, S., Davis, J., Ramamoorthi, R.: Efficiently combining positions and normals for precise 3d geometry. ACM Trans. Graph. **24**(3), 536–543 (2005). http://doi.acm.org/10.1145/1073204.1073226

17. Palma, G., Corsini, M., Cignoni, P., Scopigno, R., Mudge, M.: Dynamic shading enhancement for reflectance transformation imaging. J. Comput. Cult. Heritage **3**(2), 1–20 (2010)

18. Papadhimitri, T., Favaro, P.: A new perspective on uncalibrated photometric stereo. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1474–1481 (2013)

19. Papadhimitri, T., Favaro, P., Bern, U.: Uncalibrated near-light photometric stereo. In: Proceedings of the British Machine Vision Conference, pp. 1–12 (2014)

20. Quéau, Y., Durou, J.D.: Some illumination models for industrial applications of photometric stereo. In: QCAV (2015)

21. Sabzevari, R., Del Bue, A., Murino, V.: Structure from motion and photometric stereo for dense 3D shape recovery. In: Maino, G., Foresti, G.L. (eds.) ICIAP 2011, Part I. LNCS, vol. 6978, pp. 660–669. Springer, Heidelberg (2011)

22. Snavely, K.N.: Scene reconstruction and visualization from internet photo collections. Ph.D. thesis, Seattle, WA, USA (2009)

23. Snavely, N., Seitz, S.M., Szeliski, R.: Skeletal graphs for efficient structure from motion. In: Proceedings of Computer Vision and Pattern Recognition (2008)

24. Wetzler, A., Kimmel, R., Bruckstein, A.M., Mecca, R.: Close-range photometric stereo with point light sources. In: 2014 2nd International Conference on 3D Vision, pp. 115–122 (2014)

25. Wu, C., Liu, Y., Dai, Q., Member, S., Wilburn, B.: Fusing multiview and photometric stereo for 3d reconstruction under uncalibrated illumination. Trans. Vis. Comput. Graph. **17**(8), 1082–1095 (2011)

26. Wu, L., Ganesh, A., Shi, B., Matsushita, Y., Wang, Y., Ma, Y.: Robust photometric stereo via low-rank matrix completion and recovery. In: Kimmel, R., Klette, R., Sugimoto, A. (eds.) ACCV 2010. LNCS, vol. 6494, pp. 703–717. Springer, Heidelberg (2011). doi:10.1007/978-3-642-19318-7_55

27. Zhang, M.: Robust surface normal estimation via greedy sparse regression. Ph.D. thesis (2014)