

Binocular Photometric Stereo

Hao Du^{1,3}
duhao@cs.washington.edu

Dan B Goldman²
dgoldman@adobe.com

Steven M. Seitz^{1,3}
seitz@cs.washington.edu

¹ University of Washington
Seattle, WA, USA

² Adobe Systems
Seattle, WA, USA

³ Google Inc.
USA

Abstract

This paper considers the problem of computing scene depth from a stereo pair of cameras under a sequence of illumination directions. By integrating parallax and shading cues, we obtain both metric depth and fine surface details. Casting this problem into the filter flow framework [14], enables a convex formulation of the problem, and thus a globally optimal solution. We demonstrate high quality, continuous depth maps on a range of examples.

1 Introduction

Binocular stereo methods yield relatively coarse shape reconstructions (Fig. 1(a)). This lack of geometric detail is intrinsic to the parallax cue and the fact that images are discrete—if the disparity range is 10 pixels, you have 10 depth values (sub-pixel interpolation provides limited improvement). An additional limitation is that smooth untextured regions are hard to reconstruct. In contrast, *photometric stereo* [9][10] methods produce beautifully-detailed models (Fig. 1(b)), even in smooth untextured regions, due to their ability to directly estimate *continuous-valued* surface normals. Although these normals are defined on an integer pixel grid, the fact that they are continuous-valued rather than discrete results in the preservation of very fine details in photometric stereo results, compared to binocular stereo. A key weakness of photometric stereo, however, is the *lack of metric shape information*; i.e., it is not possible to compute the depth of the scene or the relative depth of two objects.

This paper demonstrates that it is possible to achieve the best of both worlds—fine details and metric depth — by adding a second camera to the traditional photometric stereo setup. Our system’s input consists of a stereo sequence (a synchronized pair of image sequences from two cameras) of a fixed object under a sequence of different illumination directions. Such a sequence can be produced, for example, by waving a light source around an object captured from a stereo rig. The output is a *continuous-valued depth map*. Furthermore, we introduce a novel convex formulation for this purpose, which can be globally optimized using well-known methods. The approach is simple to implement and outperforms the state-of-the-art for both stereo and photometric stereo methods.

Although there is a small literature on combining shading and parallax cues for shape reconstruction [12][13][15], these methods have failed to replace pure stereo and photometric

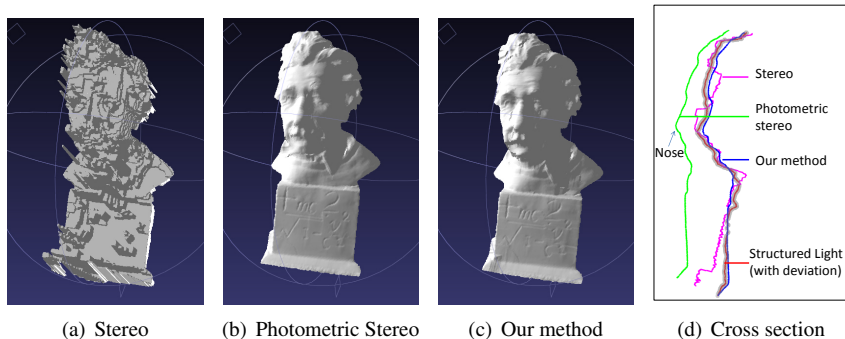


Figure 1: Reconstructions using binocular stereo, photometric stereo and our method.

stereo in practice, due in part to added complexity, restricted operating range, and/or difficulty of deployment. The main idea, common to most of these methods, is that the surface normals obtained through shading cues provide a constraint on disparity values obtained through parallax/motion cues. However, the relation between disparity and surface normals is nonlinear, and therefore challenging to impose in an optimization framework. We provide the first convex formulation for this problem.

Our approach casts the binocular photometric stereo problem in the *filter flow* framework, recently introduced by Seitz and Baker [16]. Rather than solving for depth explicitly, each depth value is represented as the centroid of a 1D convolution filter kernel. There is a kernel for each pixel in the reference image (e.g., the left camera view), and the collection of these kernels represents a space-variant convolution filter. Depth computation is reformulated as inverse filtering, i.e., solving for the filter kernels that transform the left camera image sequence into the right. The key insight is that the relationship between surface normals and depths can be expressed *linearly* in this framework, and solved via a single linear program.

Our convex formulation does not enforce compactness, a non-convex constraint that was necessary in previous work using filter flow for optical flow problems [16]. Compared to optical flow where compactness plays an important role, binocular photometric stereo is much better constrained by virtue of more input image data and we’ve found that adding compactness yields only a small improvement in metric accuracy.

2 Related Work

There are several ways to combine the information from binocular stereo and photometric stereo. Lee et al. [17] measure some sparse sample points of the shape using binocular stereo, and deform the shape obtained from photometric stereo to conform with these sparse measurements. Nehab et al. [18] use a laser-scanned shape to rectify the low frequency component of normals from photometric stereo, and combine the rectified normals with the laser-scanned shape to solve for an optimized reconstruction. Beeler et al. [9] apply photometric constraint in a separate refinement step for facial geometry capture. These methods, as well as [9] [10], require an initially computed 3D shape, either from laser or stereo.

Multiview photometric stereo methods [6] [13] [4] optimize the shape using the photometric normals as well as the visual hull. These methods do not (extensively) use cues from surface appearance for depth estimation, and generally rely on many more than two views to resolve ambiguous matches.

Ikeda [8] runs photometric stereo using two color-separated illuminations. Their goal of using stereopsis was to help reduce the number of images (illuminations) required in order to achieve fast capture.

Kong et al. [10] use orientation-consistency to find correspondence but do not use normal information for shape reconstruction. Zickler et al.'s work on binocular Helmholtz stereo [20] exploits normal information for correspondence, but only in a 1D scanline-by-scanline basis. It is also limited to two lightsource positions, limiting the accuracy of binocular normal recovery.

3 Binocular Photometric Stereo

We propose a binocular photometric stereo setup in which a second camera is added to the traditional photometric stereo system. The scene is assumed to be stationary, the camera pair is fixed, and the distant illumination varies between successive views.

3.1 The Problem

The problem of *binocular photometric stereo* can be formulated as solving for a continuous-valued depth map that best complies with both the parallax and photometric stereo cues: the intensities match between binocular correspondences and the depths satisfy the photometrically acquired normals.

We assume the cameras have been rectified so that parallax is strictly horizontal. Let f be the focal length, b the baseline, and $d^{\mathbf{u}}$ the disparity for pixel $\mathbf{u} = (u, v)$ on one image (say the left image). The 3D position $P^{\mathbf{u}} = (X^{\mathbf{u}}, Y^{\mathbf{u}}, Z^{\mathbf{u}})$ is given by

$$Z^{\mathbf{u}} = \frac{fb}{d^{\mathbf{u}}} \quad X^{\mathbf{u}} = \frac{u}{f}Z^{\mathbf{u}} \quad Y^{\mathbf{u}} = \frac{v}{f}Z^{\mathbf{u}} \quad (1)$$

Following previous derivations of perspective photometric stereo [14] [15], denote the tangents (along u and v directions) of the surface corresponding to pixel \mathbf{u} by $T_u^{\mathbf{u}}, T_v^{\mathbf{u}}$:

$$T_u^{\mathbf{u}} = \frac{\partial P^{\mathbf{u}}}{\partial u} = \left[-\frac{1}{f}(u \frac{\partial Z^{\mathbf{u}}}{\partial u} + Z^{\mathbf{u}}) ; -\frac{1}{f}v \frac{\partial Z^{\mathbf{u}}}{\partial u} ; \frac{\partial Z^{\mathbf{u}}}{\partial u} \right] \quad (2)$$

$$T_v^{\mathbf{u}} = \frac{\partial P^{\mathbf{u}}}{\partial v} = \left[-\frac{1}{f}u \frac{\partial Z^{\mathbf{u}}}{\partial v} ; -\frac{1}{f}(v \frac{\partial Z^{\mathbf{u}}}{\partial v} + Z^{\mathbf{u}}) ; \frac{\partial Z^{\mathbf{u}}}{\partial v} \right]. \quad (3)$$

The fundamental cue of binocular stereo is brightness consistency: the assumption that scene points appear the same brightness in both views. The fundamental principle of surface reconstruction from normal is that the surface tangents at all positions should be perpendicular to their normals. These yield a minimization of an objective function with the following two weighted terms (using L1-norm),

$$\|I_l(u, v) - I_r(u + d^{\mathbf{u}}, v)\|_1 \quad (4)$$

$$+\lambda \left(\left\| T_u^{\mathbf{u}}(P^{\mathbf{u}}) \cdot N^{\mathbf{u}} \right\|_1 + \left\| T_v^{\mathbf{u}}(P^{\mathbf{u}}) \cdot N^{\mathbf{u}} \right\|_1 \right), \quad (5)$$

in which disparities are the only variables. These terms are difficult to optimize, since $T_u^{\mathbf{u}}$ and $T_v^{\mathbf{u}}$ are nonlinearly related to $d^{\mathbf{u}}$. We must resolve the following two questions: First, since there is no closed form representation for I_l and I_r , how can the term (4) be made convex; second, since $T_u^{\mathbf{u}}$ and $T_v^{\mathbf{u}}$ are non-linear to $d^{\mathbf{u}}$, how can the term (5) be made convex?

3.2 The Filter Flow Formulation

To address the non-closed-form and non-linear issues as discussed above, we cast the binocular photometric stereo problem into the filter flow framework, recently introduced by Seitz and Baker [16]. With the aid of a small approximation, the problem can be formulated as a single convex optimization in the filter flow framework, and solved using linear programming.

3.2.1 Data Objective

Consider the i^{th} rectified image pair, as illustrated in Fig. 2(a). Each pixel \mathbf{u} in the left image with intensity $I_l^i(u, v)$ corresponds to a 1D filter $M^{\mathbf{u}}$ that, when applied to the image pixels $I_r^i(u + j, v)$ on the right image, produces the value $I_l^i(u, v)$ matching pixel \mathbf{u} , i.e.,

$$I_l^i(u, v) = \sum_j M_j^{\mathbf{u}} I_r^i(u + j, v). \quad (6)$$

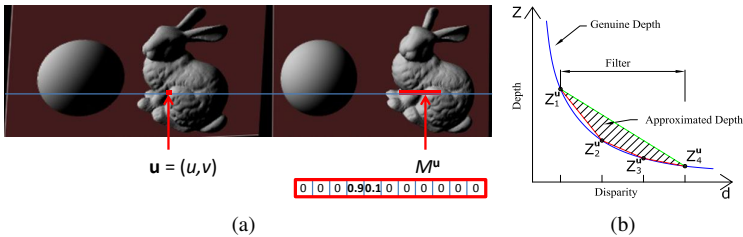


Figure 2: (a) The principle of filter flow for stereo. (b) The depth approximation.

As shown by an example filter $M^{\mathbf{u}}$ in Fig. 2(a), if the filter represents a shift of $d^{\mathbf{u}}$ (disparity) pixels, there is one entry of value 1 (integer disparity) or two neighboring entries that sum to 1 (subpixel disparity) at the $d^{\mathbf{u}}$ 'th entry of $M^{\mathbf{u}}$, while other entries in $M^{\mathbf{u}}$ are all zero.

We use the *centroid* of a filter to represent the continuous disparity, defined as,

$$d^{\mathbf{u}} = \sum_j j M_j^{\mathbf{u}}. \quad (7)$$

To regularize the filters, we enforce *Non-negative* and *Sum-to-one* constraints:

$$M_j^{\mathbf{u}} \geq 0 \forall j \quad , \quad \sum_j M_j^{\mathbf{u}} = 1, \quad (\text{POS-M, SUM1-M})$$

Seitz and Baker [16] also introduced a compactness constraint to encourage narrow filters corresponding to simple offsets, but found this constraint was not always necessary when other constraints were present. As described in Section 4.5, we've found compactness to make only a minor difference for binocular photometric stereo, where the combination of texture and normal constraints are sufficient to regularize the solution.

Substituting Eq. (7) into Eq. (4) and summing over all pairs of images, we can now linearly represent the objective function of binocular stereo on filter entries as follows, noted by *Data Objective* (DO),

$$\sum_i \sum_{\mathbf{u}} \left\| I_l^i(u, v) - \sum_j M_j^{\mathbf{u}} I_r^i(u + j, v) \right\|_1. \quad (\text{DO})$$

3.2.2 Normal Objective

In this section we describe an approximation that linearly represents the objective function of photometric stereo (Eq. (5)) using filter entries in the filter flow framework.

Since disparity $d^{\mathbf{u}}$ is linear with respect to filter entries $M_j^{\mathbf{u}}$ (Eq. (7)), and the objective (Eq. (5)) is linear with respect to depth $Z^{\mathbf{u}}$ (Eq. (2) and (3)), the only remaining challenge is that the relationship between depth and disparity is nonlinear (Eq. (1)).

However, consider the trivial case of a compact filter in which only two adjacent entries have nonzero weight. In this case, each entry corresponds to a fixed disparity and therefore a fixed depth, and we can approximate the depth by linearly interpolating the depths corresponding to these two disparities.

More generally, for arbitrary filters, each filter entry $M_j^{\mathbf{u}}$ corresponds to an element-disparity j and an element-depth $Z_j^{\mathbf{u}} = (fb)/j$. By averaging these depths weighted by filter entries, we get a linear approximation of the depth, denoted by $\hat{Z}^{\mathbf{u}}$, for pixel \mathbf{u} ,

$$\hat{Z}^{\mathbf{u}} = \sum_j M_j^{\mathbf{u}} Z_j^{\mathbf{u}} = \sum_j M_j^{\mathbf{u}} \frac{fb}{j}. \quad (8)$$

Substituting the depth $Z^{\mathbf{u}}$ in Eq. (2)(3) by the approximated depth $\hat{Z}^{\mathbf{u}}$ from Eq. (8), we get approximated tangents $\hat{\tau}_u^{\mathbf{u}}$ and $\hat{\tau}_v^{\mathbf{u}}$ with components $\hat{Z}^{\mathbf{u}}$, $\frac{\partial \hat{Z}^{\mathbf{u}}}{\partial u}$ and $\frac{\partial \hat{Z}^{\mathbf{u}}}{\partial v}$, where, from Eq. (8) we have the linear relationship

$$\frac{\partial \hat{Z}^{\mathbf{u}}}{\partial u} = \frac{\partial \hat{Z}^{(u,v)}}{\partial u} = fb \left(\sum_j \frac{M_j^{(u+1,v)}}{j} - \sum_j \frac{M_j^{(u,v)}}{j} \right) \quad \frac{\partial \hat{Z}^{\mathbf{u}}}{\partial v} = \frac{\partial \hat{Z}^{(u,v)}}{\partial v} = fb \left(\sum_j \frac{M_j^{(u,v+1)}}{j} - \sum_j \frac{M_j^{(u,v)}}{j} \right). \quad (9)$$

Using the approximated depth from Eq. (8) and approximated tangents derived from Eqs. (9) in the objective Eq. (5), we obtain our *Normal Objective* (NO),

$$\sum_{\mathbf{u}} \left\| \hat{\tau}_u^{\mathbf{u}}(P^{\mathbf{u}}) \cdot N^{\mathbf{u}} \right\|_1 + \left\| \hat{\tau}_v^{\mathbf{u}}(P^{\mathbf{u}}) \cdot N^{\mathbf{u}} \right\|_1. \quad (\text{NO})$$

According to Eqs. (2, 3, 8, 9), the Normal Objective (NO) is linear with respect to filter entries.

For filters with one nonzero entry this approximation is exact, but for general non-compact filters it is a convex combination of the depths corresponding to nonzero filter entries. Fig. 2(b) illustrates the idea of this approximation. The true depth lies on the blue curve according to Eq. (1), i.e. a $f(x) = 1/x$ function. In the filter flow approximation, each filter entry corresponds to an element-depth, noted by $Z_j^{\mathbf{u}}$ ($j = 1, 2, 3, 4$ here). Under the Non-negative and Sum-to-one (POS-M, SUM1-M) constraints, the approximated depth can lie anywhere in the shaded region. The closest approximation is along the red line, which happens when the filter is compact (either has a 1 entry or the summation of neighboring two entries equals to 1). The red line can be made closer to the true depth by increasing the resolution of the filter. The worst approximation is along the green line, which happens when all entries but the two at the sides are 0.

3.2.3 The Optimization and 3D Reconstruction

We optimize the weighted sum of Data Objective (DO) and Normal Objective (NO) subject to the Non-Negative (POS-M) and Sum-To-One (SUM1-M) constraints. The optimization is convex, and a global minimum can be found using linear programming.

Once an optimal solution to the entire filter flow is found, we use Eq. (8) to reconstruct the depths and use the projective geometry Eq. (1) to recover the 3D positions.

4 Experimental Results

We evaluate the performance of our method for binocular photometric stereo on both synthetic and real captured data. Visual and numerical comparisons with traditional binocular stereo and traditional photometric stereo demonstrate that our method achieves significantly better results than either algorithm individually. We also show a comparison with a recently developed method of Nehab et al. [24] originally designed to combine photometric normals with a shape acquired through laser scan, which can also be applied to our problem by substituting a binocular stereo reconstruction for their laser scans. Results show that our method is comparable for easy cases such as curved objects with strong correspondence cues, but our method does better with weak correspondence cues such as a planar textureless surface.

We estimate normals using the traditional Lambertian photometric stereo method [9] with 9 – 11 input images for each scene. In our implementation of filter flow, we set all the filters to have the same size and offset such that the corresponding disparities are able to cover the maximum and minimum depth of the scene. For the optimization, we use MOSEK’s [2] interior point solver.

We evaluated both AdaptingBP method [10] (a top-ranked Middlebury algorithm) and the FilterFlow method [6] (by optimizing the Data Objective (DO) and a Centroid Smoothness Objective (MRF1-M)) to obtain the comparative binocular stereo results. When comparing to the method of Nehab et al., we provide these binocular results in place of the laser scanned models used in the original paper. We found that the binocular reconstructions and their errors using AdaptingBP and FilterFlow are largely comparable. Figure 1(a) shows result of applying AdaptingBP. Figure 3(a) 3(b) show results of applying filter flow. Table 1 shows reconstruction errors of applying the two stereo methods.

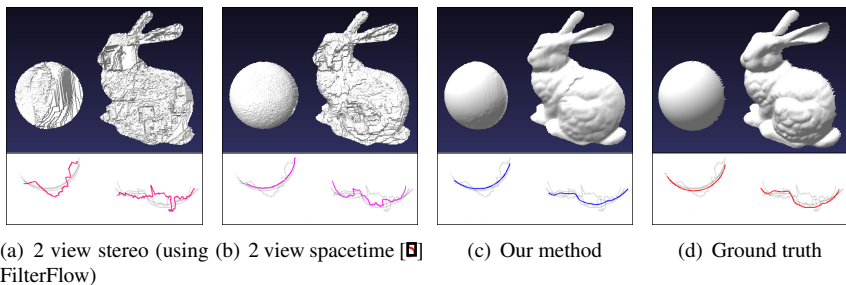


Figure 3: Comparison of methods on a synthetic scene.

4.1 Comparison to Binocular Stereo

A sample rectified image pair of a synthetic bunny and a sphere is shown in Fig. 2(a). Fig. 3 shows the results produced by different methods on this dataset. For each method we show a 3D rendered view and a 2D cross section cutting horizontally through the middle of the sphere and bunny.

Fig.3(a) and 3(b) are the binocular stereo reconstructions using one image pair, and multiple image pairs with changing illumination (spacetime stereo [6]) respectively. Using multiple image pairs with changing illuminations improves stereo: The body of the reconstructed bunny appears flat in the single-pair reconstruction but curved – albeit noisy – in the multiple-pair reconstruction; the sphere has significant errors in the single-pair reconstruction.

tion but is well-approximated in the multiple-pair reconstruction. However, both single-pair and multiple-pair binocular reconstruction distort the small-scale geometric detail. Our binocular photometric stereo method combines the binocular correspondence and photometric normal cues, producing much better results 3(c) than individually applying binocular stereo. The ground truth is shown in Fig.3(d).

4.2 Comparison to Nehab et al.

Nehab et al. recently developed a method [L4] that combines normals from photometric stereo and positions from a laser scanned shape to achieve enhanced reconstruction. They make a rectified normal map by combining the low frequencies from the scanned positions and high frequencies from the photometric normals, and fuse the normal map with the scanned shape. Their method can be used in our setting by treating a binocular stereo reconstruction (of much lower quality than a laser scan) as the input positions.

Rendered reconstructions of the bust of Einstein demonstrate that photometric stereo (Fig. 1(b)), the method of Nehab et al. (Fig. 4(b)) and our method (Fig. 4(a)) all produce visually clean reconstructions, which, in contrast to the low-quality binocular reconstruction Fig. 1(a) on this textureless object, recreate fine surface detail.

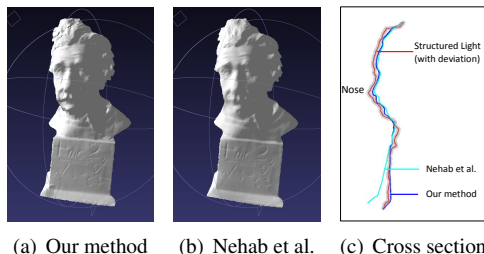


Figure 4: The reconstruction of a bust of Einstein.

Fig. 4(c) shows a 2D cross section vertically cut through the middle of the bust. The red curve with an error-window is the structured light reconstruction as reference. The cyan curve shows the result using the method of Nehab et al. [L4], and the blue curve shows the reconstruction using our method. Our reconstruction is closer to ground truth than Nehab et al. A comparison of numerical error is provided later in this section.

The method of Nehab et al. is susceptible to large errors when the input position data is inaccurate. In this case we only have low-quality binocular stereo reconstruction as our input, which is especially poor in flat textureless areas where correspondence cues are weak (e.g. the bottom box of the Einstein bust). Our method using correspondence and normal cues can operate effectively in these low-texture areas.

4.3 Comparison to Photometric Stereo

Fig. 5 shows reconstructions of three separate objects: a sphere, horse and buddha. In the scene, the sphere is put closer to the camera than the horse and buddha. Fig.5(a) is the reconstruction using photometric stereo. The surface contains nice details but the layout of these objects does not reflect the correct depth we see in the structured light reconstruction Fig.5(c), because photometric stereo lacks metric depth information. Fig.5(b) is produced

by our method, which reflects both metric depth and photometric normals. Fig.5(d) is the cross section view made by a horizontal 2D plane.

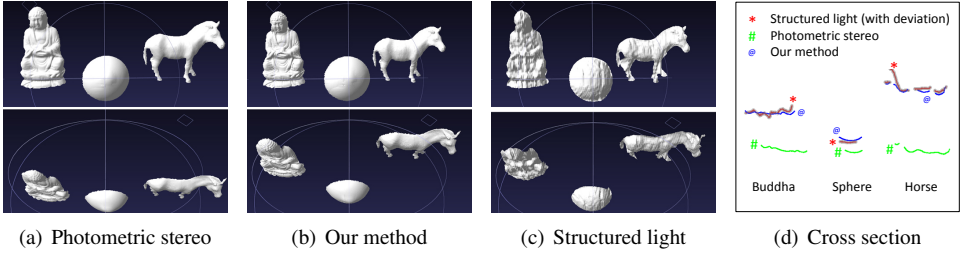


Figure 5: The reconstruction of three disconnected objects.

Fig. 6(a) is one captured image of the dinosaur. From this point of view, there exists a large depth discontinuity between its body and right hand. Fig.6(b) is the reconstruction using photometric stereo, which contains nice surface details but as seen from the top view (Fig.6(f)), the right hand is actually at the completely wrong position. Fig.6(c) is the reconstruction using our method, which reflects detailed surface normals as well as metric depths. The reconstruction by our method in the top view, Fig.6(f), shows that the position of the right hand is correctly recovered. In addition, there exists a large distance between the photometric stereo reconstruction and our reconstruction as shown in Fig.6(f), because metric depths are missing from the photometric reconstruction. This is also reflected in the 2D cross section view 6(e) which shows a horizontal cut through the middle of the object.

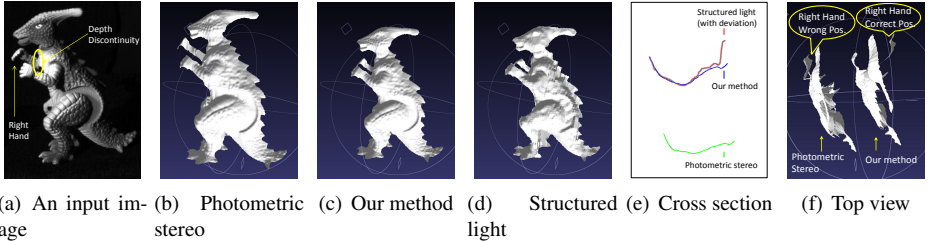


Figure 6: The reconstruction of a dinosaur.

4.4 Reconstruction Errors

For the experiments shown above, we compare the reconstruction errors among binocular stereo, photometric stereo, Nehab et al. [14] and our method. The ground truth of the synthetic scene and the structured light reconstruction of real scenes (accurate to an error of ± 0.04 unit length of the calibration chessboard pattern) are used as the reference shapes for the error calculation. Following the error evaluation scheme of [14], we compute the reconstruction error by first throwing out 10% scene points in the reconstructed shape that are of largest distances to the referencing shape, and use the maximum distance among the remaining scene points. For our method and Nehab et al., we choose the same parameters for each method to run through the datasets. The best ratio (that minimizes the average reconstruction errors) to weight the position and normal objectives for Nehab et al. is found to be $1/10^5$ (chosen among $1/10$, $1/10^2$, ..., $1/10^7$, and the best ratio $1/\lambda$ to weight the correspondence and normal objectives for our method is $1/2000$ (chosen among $1/500$, $1/1000$, $1/2000$, $1/3000$ and $1/5000$). The errors are listed in Table 1.

	Photometric st.	(Binocular st. with FilterFlow)		(Binocular st. with AdaptingBP)		Our method
		Binocular st.	Nehab et al.	Binocular st.	Nehab et al.	
Bunny	2.425	0.197	0.208	0.270	0.217	0.199
Einstein	1.680	0.490	0.345	1.272	0.463	0.184
Three Objects	2.906	0.410	0.231	0.776	0.209	0.253
Dinosaur	1.546	0.381	0.174	0.368	0.225	0.170

Table 1: The reconstruction errors (measured in the unit length).

Not surprisingly, photometric stereo has large errors, owing to the lack of metric depth. Binocular stereo does better, (in one particular case it produces even a smaller error than our method), but the numerical scores do not reflect the poor surface normals which produce noisy renderings (see figures). Using binocular stereo and photometric stereo together, Nehab et al and our method both produce better reconstruction errors in general, while our method significantly outperform Nehab et al’s for the Einstein case which contains a flat, textureless surface – see the 3D renders shown above in this section.

4.5 Extension: Applying Compactness Objective

As reported by Seitz and Baker [16], adding a term that encourages filters to be *compact* results in superior results for flow problems, at the expense of making the problem non-convex. We evaluated adding both compactness and soft-compactness terms from [16] to our binocular photometric stereo implementation. The former has an integer-depth bias and produced ridging artifacts (Fig. 7(a)). The latter behaves better, but also results in a few striation lines through the reconstruction (Fig. 7(b)) as compared to the reconstruction without any compactness terms (Fig. 4(a)). Some possible causes for the reduced visual quality include a) bias towards integer disparities; b) inaccurate normals at depth discontinuities; and c) violation of our image formation assumptions such as shadows, reflections, and non-Lambertian reflectance. Both compactness terms make the problem non-convex, require iterative optimization, and dramatically increase solution time (by a factor of 3-5 or more).



(a) Compactness (b) Soft compact. (c) Cross section

	Our method	Compactness	Soft-CP
Bunny	0.199	0.191	0.112
Einstein	0.184	0.182	0.136
Three Objects	0.253	0.244	0.138
Dinosaur	0.170	0.170	0.083

Figure 7: Comparison of reconstruction and metric errors (in unit length) with/without compactness terms.

The table in Figure 7 shows the comparison of reconstruction errors using our method (DO:NO = 1:2000), with compactness (DO:NO:CO = 1:2000:1), and with soft-compactness (DO:NO:SCO:W = 1:2000:1:2), where DO, NO, CO, SCO and W are the parameters for Data Objective, Normal Objective, Compactness, Soft-Compactness and the window size for Soft-Compactness respectively.

The overall conclusion is that adding compactness is probably not worthwhile in general, but should be considered in applications where small improvements in metric depth are more important than visual fidelity.

5 Conclusion

In this paper, we propose binocular photometric stereo, i.e. adding a second camera to the traditional photometric stereo setting. The reconstruction is modeled using filter flow, which linearly represents the disparity and correspondence cues and linearly approximates the depth and normal cues, such that the problem is solved within a single convex optimization. We demonstrate that, utilizing the information from both worlds, binocular photometric stereo is able to produce a reconstruction with high quality surface details and metric depth.

Acknowledgements

This work was supported in part by National Science Foundation grants IIS-0963657 and IIS-0811878, Adobe, Intel, Google, Microsoft, and the Animation Research Labs.

References

- [1] Daniel G. Aliaga and Yi Xu. A self-calibrating method for photogeometric acquisition of 3d objects. *IEEE Trans. on PAMI*, 32:747–754, April 2010.
- [2] Mosek ApS. The mosek optimization software. <http://www.mosek.com/>.
- [3] Thabo Beeler, Bernd Bickel, Paul Beardsley, Bob Sumner, and Markus Gross. High-quality single-shot capture of facial geometry. *SIGGRAPH*, 29(4), July 2010.
- [4] Neil Birkbeck, Dana Cobzas, Peter Sturm, and Martin Jagersand. Variational shape and reflectance estimation under changing light and viewpoints. In *ECCV*, 2006.
- [5] J Davis, D Nehab, R Ramamoorthi, and Rusinkiewicz S. Spacetime stereo: A unifying framework for depth from triangulation. *IEEE Trans. on PAMI*, 27:296–302, February 2005.
- [6] C Hernandez, G Vogiatzis, and R Cipolla. Multiview photometric stereo. *IEEE Trans. on PAMI*, 30:548–554, March 2008.
- [7] A Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *IEEE Trans. on PAMI*, 27:1254–1264, August 2005.
- [8] O Ikeda. Shape reconstruction from two color images using photometric stereo combined with segmentation and stereopsis. In *IEEE Conf. on Advanced Video and Signal based Surveillance (AVSS)*, 2005.
- [9] N. Joshi and D.J. Kriegman. Shape from varying illumination and viewpoint. In *ICCV*, 2007.
- [10] A. Klaus, M. Sormann, and K. Karner. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure. In *ICPR*, 2006.
- [11] Hui Kong, Pengfei Xu, and Earn Khwang Teoh. Binocular uncalibrated photometric stereo. *Lecture notes in computer science*, 4291:283–292, November 2006.

- [12] Simon Lee and Michael Brady. Integrating stereo and photometric stereo to monitor the development of glaucoma. *Image and Vision Computing*, 9:39–44, February 1991.
- [13] J. Lim, J. Ho, M.H. Yang, and D. Kriegman. Passive photometric stereo from motion. In *ICCV*, 2005.
- [14] Diego Nehab, Szymon Rusinkiewicz, James Davis, and Ravi Ramamoorthi. Efficiently combining positions and normals for precise 3D geometry. *SIGGRAPH*, 24(3), August 2005.
- [15] S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *CVPR*, 2006.
- [16] Steven M Seitz and S. Baker. Filter flow. In *ICCV*, 2009.
- [17] Ariel Tankus and Nahum Kiryati. Photometric stereo under perspective projection. In *ICCV*, 2005.
- [18] Daniel Vlasic, Pieter Peers, Ilya Baran, Paul Debevec, Jovan Popović, Szymon Rusinkiewicz, and Wojciech Matusik. Dynamic shape capture using multi-view photometric stereo. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, 28(5), December 2009.
- [19] Robert J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19:139–144, 1980.
- [20] Li Zhang, Brian Curless, Aaron Hertzmann, and Steven M. Seitz. Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multi-view stereo. In *The 9th IEEE International Conference on Computer Vision*, pages 618–625, Oct. 2003.
- [21] Todd Zickler, Jeffrey Ho, David J. Kriegman, Jean Ponce, and Peter N. Belhumeur. Binocular helmholtz stereopsis. In *ICCV*, 2003.