Additional Results for:
Multi-view Photometric Stereo with Spatially Varying Isotropic Materials

Zhenglong Zhou Zhe Wu Ping Tan
National University of Singapore

1. System Pipeline

We provide an block diagram of our system in Figure 1. We capture images from multiple viewpoints. At each viewpoint, we capture photometric stereo images with a moving light source, which can be simply a handheld bulb. We design a robust algorithm to identify iso-depth contours from these images. Further, we apply structure-from-motion [1] to images from different viewpoints to reconstruct a sparse set of 3D points. We then derive a complete 3D shape by propagating the depths of these points along the dense iso-depth contours. This initial shape is further refined according to the method described in [3]. Once the shape is fixed, we estimate a set of basis isotropic BRDFs and their mixing weights at each surface point by the ACLS method [2] to model the surface reflectance.

2. More Examples

An example result from the handheld system is provided in Figure 2. Figure 2 (a) shows a sample input image. This example was captured from 10 viewpoints, which allow us to reconstruct part of its surface. To better visualize the recovered shape, we render it with uniform diffuse shading in (b). Most of the geometry details are successfully captured. (c) is a rendering according to the captured reflectance from the same viewpoint and lighting condition as the input image in (a). To provide a quantitative evaluation on shape capture, we visualize the shape reconstruction error (measured in millimeters) in (d). The larger errors at the surface boundary are due to the insufficient and slanted observations. Overall, the median (and mean) shape error is 0.62 (and 0.96) millimeters. Here, the object diameter is 140 millimeters. This result was a bit noisy because the brushed paint was not precisely isotropic.

Another two examples, 'Cat' and 'Teapot2', are included in Figure 3. Their diameters are 140 and 120 millimeters respectively. These examples have different materials. The 'Cat' is covered by a brushed matte paint. The clay 'Teapot2' has soft and extended highlight. We captured 35 and 30 viewpoints for the 'Cat' and 'Teapot2' examples respectively. Our method consistently performed well on all

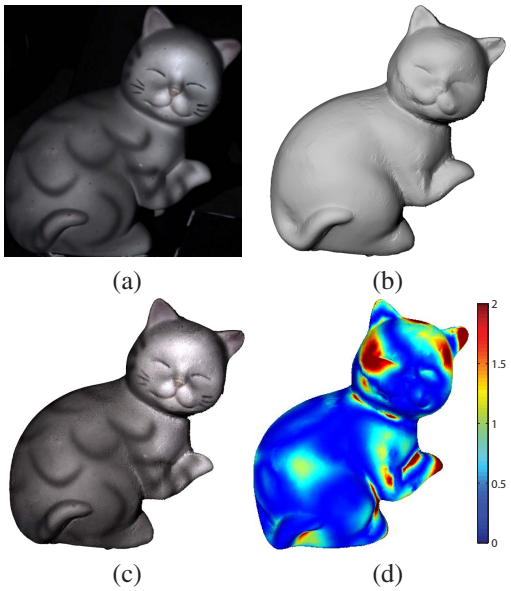

Figure 2. Results from the handheld system. (a) one of the input images, (b) the recovered shape rendered with uniform diffuse shading. (c) a rendering with the recovered reflectance model from the same viewpoint and lighting condition as the image in (a). (d) the color-coded shape error (in millimeters) compared to laser-scanned 'ground truth'.

of them. Their median (or mean) shape reconstruction errors were 0.24 and 0.24 (or 0.53 and 0.66) millimeters respectively. The 'Teapot2' example had relatively larger error at one side, mainly due to the imprecise structure-from-motion reconstruction caused by erroneous feature matching.

3. Intermediate results

Figure 4 shows the reconstructed shape at different stages. Shown on the left are 3D points obtained from multi-view stereo. In the middle are the 3D points (with normal directions) obtained by depth propagation. On the right is the result after Poisson surface reconstruction. The final optimized shape is at Figure 4 (d). Note the face becomes clearly smoother after optimization. Figure 5 visual-

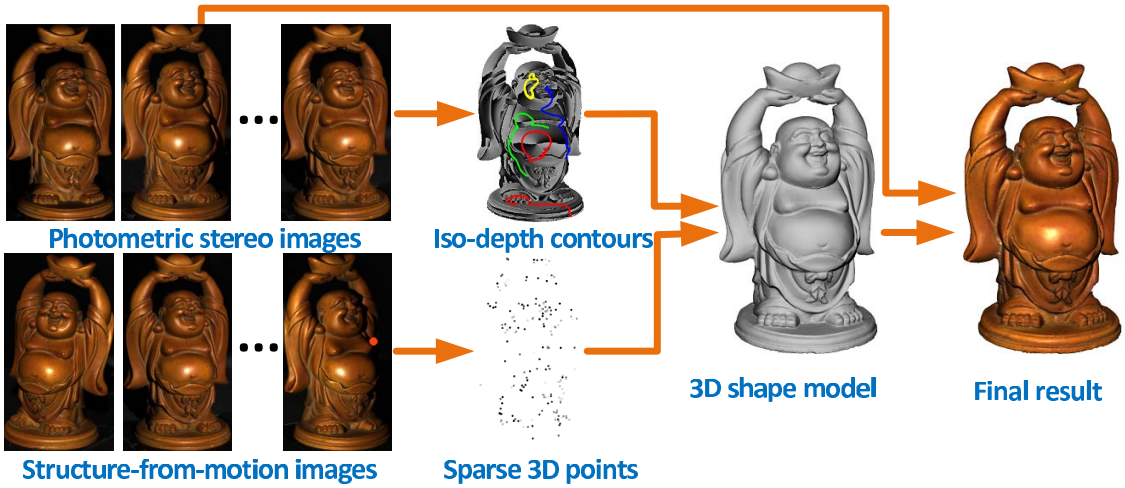

Figure 1. System pipeline. We recover iso-depth contours from photometric stereo images and recover a sparse 3D point cloud by structure-from-motion. In the figure showing iso-depth contours, the gray intensity encodes the estimated azimuth angles, and the colored curves are iso-depth contours. We then propagate the depths of these 3D points along the iso-depth contours to recover the complete 3D shape. Once the shape is fixed, we estimate the spatially varying BRDF from the original input images.

(a) (b) (c) (d)

Figure 3. Results from the ring-light system. From left to right, these figures are arranged in the same way as Figure 2.

izes the BRDF mixture weights and the basis BRDFs. The red and green channels are the normalized mixture weight of the first and second basis BRDFs. Each basis BRDF is applied to render a sphere under front lighting and viewing directions. Most of our examples consist of a shiny and a less shiny basis BRDFs. This can be seen clearly from the 'Cup' and 'Frog' examples.

4. Re-rendering

To provide an intuitive evaluation of our results, we rendered all examples under novel lighting and viewpoint and compared them with captured photographs in Figure 6 and Figure 8. Note these images were not used in our shape and reflectance capture system. The first two rows are results from the handheld system, while the others come from the ring-light system. From top to bottom, the median (or mean) intensity differences for each example were 4.8 6.9 8.3, 5.0, 2.6, 5.7 and 6.0 (or 7.3 12.2 12.5, 6.8, 4.0, 10.8 and 9.5) intensity levels with pixel values between [0,255] respectively. Figure 7 further shows rendering of the captured object under novel environment lighting.

References

[1] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, New York, NY, USA, 2 edition, 2003. 1

[2] J. Lawrence, A. Ben-Artzi, C. DeCoro, W. Matusik, H. Pfister, R. Ramamoorthi, and S. Rusinkiewicz. Inverse shade trees

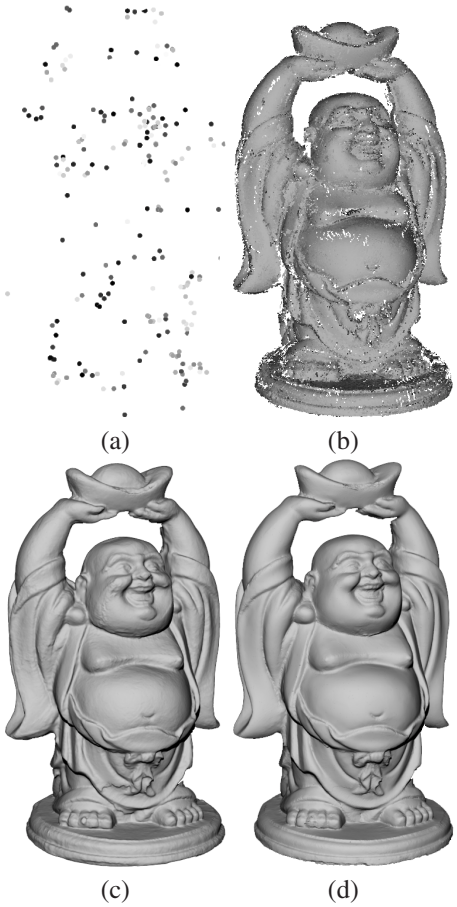

Figure 4. (a) initially reconstructed 3D points; (b) 3D points obtained by depth propagation; (c) initial shape after Poisson surface reconstruction; (d) final result.

for non-parametric material representation and editing. *ACM Trans. Graph.*, 25:735–745, July 2006. 1

[3] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi. Efficiently combining positions and normals for precise 3d geometry. *ACM Trans. Graph.*, 24:536–543, 2005. 1

Figure 5. The normalized BRDF mixture weights are visualized in the different color channels. The corresponding basis BRDFs are used to render a sphere on the right.

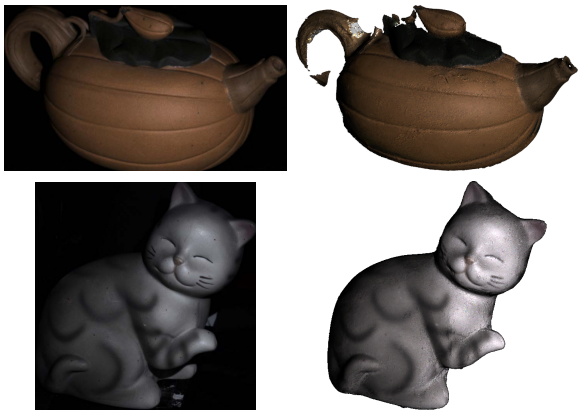

Figure 6. Results from handheld system. Left: reference real photographs. Right: rendering under novel viewpoint and lighting condition.

Figure 7. Rendering under environment lighting.

Figure 8. Results from the ring-light system. Left: reference real photographs. Right: rendering under novel viewpoint and lighting condition.