Photometric Stereo with Near Point Lighting: A Solution by Mesh Deformation

Wuyuan Xie, Chengkai Dai, and Charlie C. L. Wang*
Department of Mechanical and Automation Engineering,The Chinese University of Hong Kong

Photometric stereo (PS) [1] estimates a dense field of normals from a set of 2D images captured by a fix camera under different illumination conditions. However, the shape distortions generated by PS has negative effects in the low frequency, since light sources are assumed to be located at infinite positions (i.e., the illumination can be regarded as parallel lighting). As a result, it will make the luminance attenuate sharply to cause images with poor quality. In this paper, we provide a solution to the problem of near point lighting PS (NPL-PS) to overcome this limitation.

Without loss of generality, the input of our NPL-PS approach has k images, $\mathcal{I}_1, \cdots, \mathcal{I}_k$, where each has the same interested region composing of m pixels. The image $\mathcal{I}_i$ is captured under the illumination of the i-th lighting source with its position $\mathbf{p}_i \in \Re^3$ known. In addition, the observed appearance brightness I of a Lambertian object under a lighting direction $\mathbf{l} \in \Re^3$ at a surface point $\mathbf{x} \in \Re^3$ can be described as $I(\mathbf{x}) = \rho \mathbf{n} \cdot \mathbf{l}$, where ρ is the Lambertain reflection albedo and $\mathbf{n} \in \Re^3$ is the surface normal at $\mathbf{x}$. Therefore, akin to the lighting model presented in [3], we employ a NPL model as

$$I_i(\mathbf{x}) = \frac{\rho}{\alpha \|\mathbf{p}_i - \mathbf{x}\|^2} \left(\mathbf{n}(\mathbf{x}) \cdot \frac{\mathbf{p}_i - \mathbf{x}}{\|\mathbf{p}_i - \mathbf{x}\|} \right), \tag{1}$$

where α is the attenuation coefficient, and therefore $(\mathbf{p}_i - \mathbf{x})/\|\mathbf{p}_i - \mathbf{x}\|$ gives the lighting direction $\mathbf{l}_i$ at $\mathbf{x}$.

In our NPL model, the position of light sources can be obtained by a calibration procedure [4] when the relative position between camera and light sources are fixed during the shape acquisition. Even after determining the positions of light sources, Eq.(1) is still non-linear due to the unknown depth value of $\mathbf{x}$ for each pixel in the captured images.

To solve the values of $\mathbf{n}(\mathbf{x})$ and $\mathbf{x}$ in Eq.(1) simultaneously, we propose to use local/global mesh deformation. Our formulation for solving the NPL-PS problem is based on converting each pixel (i, j) in the interested region into a quadrangular facet $f_{i,j}$, the boundary of which is defined by four vertices $\mathbf{v}_{i,j}$, $\mathbf{v}_{i+1,j}$, $\mathbf{v}_{i+1,j+1}$ and $\mathbf{v}_{i,j+1}$. A vertex $\mathbf{v}_{i,j}$ has its x- and y-coordinates fixed and $z_{i,j}$ as an unknown variable to be determined – that is $(o_x + iw, o_y + jh, o_z + z_{i,j})$. Here (o_x, o_y, o_z) specifies the shifting between the image coordinate system and the world coordinate system. The initial values of $z_{i,j}$s can be assigned as $z_{i,j} = 0$ or be given randomly. The collection of facets forms a mesh surface $\mathcal{M}$ with C^0-continuity.

Generally speaking, the strategy of local/global mesh deformation decouples the nonlinear optimization procedure into interlaced steps of *local shaping* and *global blending*. In each iteration, the *local shaping* step is first performed to determine the position and orientation of each facet according to its target normal and its current shape. After that, the *global blending* step is employ to glue all the facets back into a connected mesh surface. Specifically, the *local shaping* is associated with two operations: 1) determining the normal $\mathbf{n}_{i,j}$ of $f_{i,j}$ and 2) rotating the facet to following the orientation of $\mathbf{n}_{i,j}$. The normal $\mathbf{n}_{i,j}$ is determined at the center of $f_{i,j}$, that is $\mathbf{c}_{i,j} = \frac{1}{4}(\mathbf{v}_{i,j} + \mathbf{v}_{i+1,j} + \mathbf{v}_{i+1,j+1} + \mathbf{v}_{i,j+1})$. Thus, for the k-th image, an equation can be obtained from the nonlinear lighting model at $\mathbf{c}_{i,j}$ as

$$(\mathbf{p}_k - \mathbf{c}_{i,j}) \cdot \frac{\rho}{\alpha} \mathbf{n}_{i,j} = I_k(i, j) \|\mathbf{p}_k - \mathbf{c}_{i,j}\|^3 \tag{2}$$

$I_k(i, j)$ denotes the light intensity at the pixel (i, j) in the image $\mathcal{I}_k$. Incorporating all images, the normal vector $\mathbf{n}_{i,j}$ can be obtained by a least-square solution as

$$\mathbf{T}(\frac{\rho}{\alpha}\mathbf{n}_{i,j}) = \sum_k \|\mathbf{p}_k - \mathbf{c}_{i,j}\|^3 I_k(i, j)(\mathbf{p}_k - \mathbf{c}_{i,j}). \tag{3}$$

After knowing $\mathbf{n}_{i,j}$, we shift the position of $f_{i,j}$'s vertices along the z-axis to put them on the plane $\mathcal{P}_{i,j}$ that passes through $\mathbf{c}_{i,j}$ and has the normal $\mathbf{n}_{i,j}$. A vector formed by depth components of the four projected vertices is defined as $\mathbf{p}(f_{i,j}) = \{p_{i,j}, p_{i+1,j}, p_{i+1,j+1}, p_{i,j+1}\}$. Here, the depth values are

$$p_{k,l} = \mathbf{c}^z_{i,j} - \frac{(k - i - \frac{1}{2}) w \mathbf{n}^x_{i,j} + (l - j - \frac{1}{2}) h \mathbf{n}^y_{i,j}}{\mathbf{n}^z_{i,j}} \tag{4}$$

with $k \in \{i, i+1\}$ and $l \in \{j, j+1\}$.

The *global blending* step minimizes a functional to reduce the difference between $\mathbf{z}(f_{i,j})$ and $\mathbf{p}(f_{i,j})$ on each facet $f_{i,j}$, where $\mathbf{z}(f_{i,j})$ is a column vector formed by the depths of $f_{i,j}$'s four vertices $\mathbf{z}(f_{i,j}) = \{z_{i,j}, z_{i+1,j}, z_{i+1,j+1}, z_{i,j+1}\}$, and $\mathbf{p}(f_{i,j})$ is its corresponding vector after applying the local shaping step. The relationship between $\mathbf{z}(f_{i,j})$ and $\mathbf{p}(f_{i,j})$ is relaxed by applying the mean-subtraction technique [2] as

$$\Phi(\{z_{i,j}\}) = \sum_{f_{i,j}} \|\mathbf{N}\mathbf{z}(f_{i,j}) - \mathbf{N}\mathbf{p}(f_{i,j})\|^2, \tag{5}$$

where $\mathbf{N} = \mathbf{I}_{4\times4} - \frac{1}{4}\mathbf{1}_{4\times4}$ with $\mathbf{1}$ being a matrix with all elements equal to 1. It is found that Eq.(5) can be reformulated into a more compact form as

$$\Phi(\{z_{i,j}\}) = \|\mathbf{A}\mathbf{d} - \mathbf{b}\|^2, \tag{6}$$

where $\mathbf{A}$ is a $4m \times n$ matrix derived from $\mathbf{N}\mathbf{z}(f_{i,j})$ and $\mathbf{b}$ is a vector with $4m$ components derived from $\mathbf{N}\mathbf{p}(f_{i,j})$ on a mesh surface $\mathcal{M}$ with m quadrangular facets and n vertices.

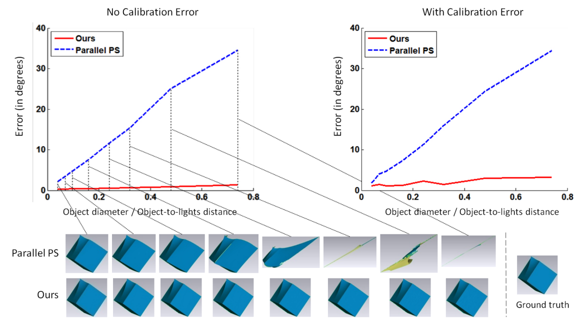

Figure 1: Comparison in the virtual environment by using 1) images taken in the exposure of far point lighting in the parallel PS to 2) images taken in the NPL exposure in our framework. To obtain nearly parallel lighting in case 1), the light sources must be placed far ways from the object to be reconstructed. As a result, the captured images become too dark to be used in reconstruction the shape. This comparison is actually as using the left point on the blue curve to compare with other points on the red curve. Here it assumes a very strong light source. In practice, when a light-source with lower intensity is employed, larger errors will be generated on the reconstruction from parallel PS.

[1] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25 (10):1239–1252, 2003.

[2] S. Bouaziz, M. Deuss, Y. Schwartzburg, T. Weise, and M. Pauly. Shape-up: Shaping discrete geometry with projections. *Compter Graphics Forum*, 31(5):1657–1667, 2012.

[3] T. Higo, Y. Matsushita, and K. Ikeuchi. Consensus photometric stereo. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2010.

[4] W. Zhou and C. Kambhamettu. Estimation of illuminant direction and intensity of multiple light sources. In *Proceedings of the 7th European Conference on Computer Vision-Part IV*, ECCV '02, pages 206–220, 2002.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.