Dynamic 3D Surface Reconstruction Using a Hand-Held Camera

Hao Fan*[‡], Lin Qi*, Junyu Dong*, Gongfa Li[†] and Hui Yu[‡]

*College of Information Science and Engineering, Ocean University of China, Qingdao, China, Email: fanhao@stu.ouc.edu.cn, {qilin, dongjunyu}@ouc.edu.cn [†]Key Laboratory of Metallurgical Equipment and Control Technology, Wuhan University of Science and Technology, Wuhan, China,

Email: ligongfa@wust.edu.cn

[‡]School of Creative Technologies, University of Portsmouth, Portsmouth, UK,

Email: hui.yu@port.ac.uk

Abstract—This paper proposes a dynamic 3D reconstruction method for recovering a surface shape from a set of images that are captured by a hand-held camera. A light source is attached to the camera as a photometric constraint. Thus, we can effectively calculate photometric stereo using the relative moving camera. The key contributions of our work are a robust pixel matching method to build effective correspondences between images for normal estimation, and an optimization method to correct the deviation in the recovered surface shape that is caused by the nonideal illumination in a close-range lighting condition. Specially we correct the recovered shape by adding an interpolation surface that is estimated using sparse control points from the structure from motion. The effectiveness of our method is verified on real datasets with a digital camera and a smartphone.

Index Terms—3D reconstruction, photometric stereo, shape from motion and shading

I. INTRODUCTION

High-quality three-dimensional (3D) shape reconstruction from motion is essential for many practical applications, such as medical image processing, obstacles avoidance of unmanned vehicles, Virtual Reality (VR) and Augmented Reality (AR). However, existing systems based on methods, such as structure from motion (SfM) and structured light scanning, still suffer from the resolution or applicability problems. It is necessary to propose a new solution for recovering surface shape from motion with more details even using a common camera.

Photometric stereo is a well-established 3D reconstruction technique [1] which uses the shading cue and can provide very detailed surface [2]. Thus, photometric stereo is applied to enhance the 3D reconstruction performance. However, traditional photometric stereo assumes images are taken from the same viewpoint but under different illumination directions, that increases the difficulty to use this technique in dynamic environments. To alleviate the problems, researchers have proposed many photometric stereo variations by incorporating geometrical methods [3]–[5]. However, these solutions still follow some assumptions such as the uniform collimated illumination from a distant light source. This limits their applications. Recently, depth sensors were applied in the dynamic photometric stereo system [6]. But depth sensors can hardly be adopted in some scenarios.

In this paper, we propose a novel dynamic photometric stereo method only using a hand-held camera with an attached light source to recover surface shape in a relatively unconstrained scenario. When the camera moves around an object, the illumination changes accordingly, which adds a photometric constraint to make it possible to solve the surface normal using photometric stereo.

Given a captured sequence of the surface, We first use SfM [7] to obtain the camera position of each frame, and produce a sparse 3D point cloud simultaneously. Then A pixel matching strategy is proposed to adjust the view and lighting to the selected reference view, that enables us to use the photometric stereo to recover the surface normal and shape. Due to the unideal illumination condition and imprecise pixel matches, the reconstructed surface is probably not accurate enough. Therefore, we propose to combine the estimated sparse point cloud to add an interpolation deviation surface to compensate for the reconstructed surface.

Our method effectively relaxes the assumptions of photometric stereo and simplifies the equipment requirements, while guaranteeing the reconstruction accuracy. The contributions are shown as follows:

- Robust pixel matches for the image sequence of a moving surface. A novel matching strategy is proposed, which realizes the accurate global match by interpolating matched feature points in rectified images.
- High-quality dense 3D reconstruction from motion. A dense reconstruction is integrated from the gradient field estimated by photometric stereo and corrected with sparse control points from SfM.

The rest of the paper is organized as follows: we first review the related work in Section II. In Section III, we describe our approach in detail. Experimental results are presented in Section IV followed with the conclusions.

Thanks to the International Science & Technology Cooperation Program of China (ISTCP) (No. 2014DFA10410), the EPSRC through project 4D Facial Sensing and Modelling (EP/N025849/1), and the Open Fund of the Key Laboratory for Metallurgical Equipment and Control of Ministry of Education in Wuhan University of Science and Technology (grant number 2017B05).

II. RELATED WORK

Shape recovery from motion has received increasing attention in recent years. SfM is the most popular technique which reconstructs a 3D structure from a series of images taken from different viewpoints [8], [9]. SfM usually combines with Multi-view Stereo for a dense 3D reconstruction. Although SfM excels in recovering large-scale structures, it often fails to capture high-frequency details [10]. Recently, researchers blend photometric clues such as photometric stereo with the geometrical methods for recovering a fine-scale dense shape.

Photometric stereo usually assumes that the photos are captured under the same view with different light directions. That makes photometric stereo difficult to be applied in a dynamic condition. Some works have solved the problem by merging geometric and photometric methods in one common frame with two separate datasets [11]–[13]. But they can not recover surface from consecutive frames. Some works fixed a light source to the camera and estimated the changing view and lighting to recover the 3D surface for each frame. Maki et al. [14] and Simakov et al. [15] added the photometric clues in multi-view stereo to produce a relative constant intensity for 3D reconstruction. They do not make full use of photometric clues to compute the surface normals, which can produce more details.

Therefore, recent works estimated the normal and depth for each frame simultaneously. The key technical challenges in these work are to (1) establish correspondences over the entire image set for normal estimation and (2) merge the normal and the geometrical based position for a high-quality dense 3D reconstruction. For a pixel-wise matching, Zhang et al. [3] present an "optical flow" based technique to recover the pixel correspondences and normals, but it requires numerous images from a dense video sequence. More works solved the pixel correspondences through a rough estimated dense depth map, that can be projected to each image with the known projection matrices of each camera. Lim et al. [4] estimated the depth map by interpolating a sparse point cloud from SfM. Joshi et al. [5] solved the depth map with the Multi-view Stereo technique by minimizing a multi-ocular photometric cost function. Zuo et al. [6] estimated the depth map using a depth sensor. Based on the pixel-wise correspondences among images, photometric stereo produces the normal estimation.

Another challenge occurs in shape recovery from normals. Because of the complicated lighting situation and imperfect matching, the reconstructed surface still has deviations. Zhang et al. [3] and Lim et al. [4] employed an iterative produce to refine the shape with correcting the bas-relief ambiguity [16]. In reality, the deviation is not only the bas-relief ambiguity. Joshi et al. [5] and Higo et al. [17] improved the accuracy of the depth map based on the work of Nehab et al. [10], which efficiently combines the geometrical based positions and the normals from photometric stereo. Zuo et al. [6] employed Zhang's method [18] to enhance the surface geometry with structural details, which can further handle the shadow areas. In these works, the surface shape is further optimized, but the optimization methods rely on a fine dense depth, which is not easy to obtain in some conditions.

Compared to the prior works, we do not estimate a depth map. Our method recovers the pixel correspondences via the interpolation of feature points in every two images and optimizes the produced photometric shape by merging a sparse point cloud from SfM [19]. Our work can effectively reduce the deviations in dynamic photometric stereo, improve computation efficiency, and allow high-quality reconstruction even using a simple hand-held camera.

III. RECONSTRUCTION ALGORITHM

The goal of our work is to recover a surface shape from a set of images captured by a relative moving hand-held camera. A light source is attached to the camera and always turns on when capturing images. With the fixed light, different views have different illumination directions, that give photometric constraint to efficiently estimate surface normals. Thus, as shown in Fig. 1, we transform the dynamic photometric stereo issue to the form of traditional photometric stereo for normal estimation.



Fig. 1. Dynamic photometric stereo.

Algorithm 1 illustrates the flow of the proposed method. After calibrating the camera [20], we photograph the target object from different views. Then a sparse point cloud and the extrinsic parameters for each frame are computed using SfM and refined with Bundle Adjustment [7]. Based on the camera intrinsic and extrinsic parameters, we establish the mapping relationship between the sparse point cloud and the image coordinate of each frame. Then we pick one frame as the reference view and propose a robust match strategy to build the pixel correspondences between the reference view and the other views. Then images from different views are aligned to use photometric stereo to recover the surface normal in the reference camera coordinate system. Finally, the shape is computed and refined by merging the sparse points from SfM.

A. Dynamic photometric stereo

This section formulates the dynamic photometric stereo. This work takes shadows and highlights as outliers that are removed from calculation [21]. In the camera coordinate system, the measured intensity i for a single pixel p satisfies:

$$i(p) = \rho(p)\mathbf{n}(p) \cdot \mathbf{l}_c,\tag{1}$$

Algorithm 1 Dynamic 3D reconstruction algorithm.

- 1: Calibrate the Camera and Lighting
- Calibrate the camera intrinsics and the lighting direction. 2: Estimate Camera Projection Matrices

Using Structure from Motion/Bundle Adjustment, recover the camera projection matrices for each frame and obtain a sparse point cloud.

Based on the camera projection matrices, build the correspondences between the sparse point cloud and the image coordinates of each frame.

- 3: Estimate Pixel Correspondences and Normal Map Select one frame to be the reference view, find the pixel correspondences between the reference view and the other views, and compute the normal using photometric stereo.
- 4: **Compute Final Surface** Recover the final surface by fusing the normal and the depth of the sparse point cloud.

where $\rho(p)$ is the reflectance albedo of the point p, $\mathbf{n}(p)$ is the unknown unit surface normal, and \mathbf{l}_c denotes the known unit illumination direction.

In a fixed world-space coordinate system, the correspondence between world-space points and image coordinates is established with the camera projection matrices. For dynamic photometric stereo, one frame o is selected as the reference frame. Then the measurement in the k^{th} view for the same pixel p can be defined as:

$$i(p)_k = \rho(p)(\mathbf{R}_k \mathbf{R}_o^{-1} \mathbf{n}_o(p)) \cdot \mathbf{l}_c, \qquad (2)$$

where \mathbf{R}_o and \mathbf{R}_k are the rotation matrix for the selected view and the measured view k from the world coordinate system to each camera coordinate system. In another word, when we turn the k^{th} view into the selected view, only the lighting changes:

$$i(p)_k = \rho(p)\mathbf{n}_o(p) \cdot (\mathbf{R}_k \mathbf{R}_o^{-1} \mathbf{l}_c).$$
(3)

We define the lighting in the k^{th} view as $\mathbf{l}_k = (\mathbf{R}_k \mathbf{R}_o^{-1} \mathbf{l}_c)$. Then by joining the k measurements for the same pixel together, we have:

$$\mathbf{E}(p) = \rho(p)\mathbf{n}_o\mathbf{L},\tag{4}$$

where the k^{th} column of the matrix $\mathbf{E}(\mathbf{E} \in \mathbb{R}^{1 \times K})$ represents the image intensity in the k^{th} view, and the k^{th} column of $\mathbf{L}(\mathbf{L} \in \mathbb{R}^{3 \times K})$ is the k^{th} illumination direction \mathbf{l}_k .

The columns of L should be non-coplanar to make L be nonsingular and invertible. And a minimum of K = 3 views is required. Then we can estimate \mathbf{n}_o and $\rho(p)$ by solving the linear system.

The above description shows how to recover normals using dynamic photometric stereo. However, there are two key issues that need to be solved. One is how to establish the image correspondences between different views, and another is how to integrate the normals to the final surface with correcting the deviation caused by the close-range nonuniform lighting [19], [22]. The solutions are discussed in following sections.

B. Pixel matching

In this section, we present a method to establish the pixel correspondences between the selected reference frame and the other frames.

In our algorithm, we first use SfM and Bundle Adjustment to compute camera projection matrices for each frame [7], that include rotation matrices $R_0 \sim R_N$ and translation matrices $T_0 \sim T_N$. Then we select one frame as the reference frame and warp the other frames into the reference frame with the camera projection matrices.

Pixel matching between the reference frame and one key frame is a binocular stereo issue. We rectify each two frames for a better matching using Fusiello's method [23]. For the pixel-level match, we first search reliable match by detecting SURF features, then use these reliable matches as control points to deform the image into the warped image to meet a dense match. In detail, we filter the matches only with the same y coordinates, which are then taken as reliable matches.

In the deformation process, we only need to deform the x coordinate in the rectified reference frame to find the matches in the other frame. The match relation is in form of the disparity, which related to the irregular surface depth is nonlinear. Thus, we propose to use thin plate spline interpolation (TPS) to realize the deformation [24]. This approach works well for a continuous surface, that meets the target of our research. The approach produces a function that passes through the selected matches while minimizing the bending energy functional

$$E(f) = \int_{\mathbb{R}^2} |D^2 f|^2 dx dy.$$
⁽⁵⁾

And the interpolation function is of the form

$$f(x,y) = \sum_{i} w_i G(d_i(x,y))) + b_0 + b_x x + b_y y, \qquad (6)$$

where $d_i(x, y)$ is the distance between the *i*-th control point and (x, y), $G(x) \equiv x^2 \log x$ and (w_i, b_0, b_x, b_y) are coefficients.

We compute the function coefficients as described in [25], and then establish pixel correspondences from the reference image to other key images. Finally, we transform all frames into the reference view. Figure 2 shows the matching result of a Chinese leaf.



Fig. 2. Images of a Chinese leaf. (a) The input images with different views. (b) The matched images under the reference view of image 3.

C. Surface reconstruction

Based on the pixel correspondences between images, we estimate surface normals using photometric stereo. The surface shape is then integrated from the normal field using the masked weighted least-squares approach (MWLS) [26]. The integration method can provide a piecewise-smooth surface on an arbitrary connected domain.

However, the recovered surface integrated from the original normal still has the low-frequency bias, which is caused by the unideal illumination condition when using a hand-held setup. Fortunately, the low-frequency bias has rules, which can be approximated to a smooth surface [22]. Inspired by Horovitz's method [19], We add an interpolation surface to fit the low-frequency deviation using the extending thin plate spline interpolation algorithm [24].

In our algorithm, the surface is reconstructed as the sum of two surfaces. One is the original surface height integrated from the gradients. The other is a smooth interpolation surface that compensates for the bias. The interpolation surface uses the values of control points that are the difference between the true control point depths from SfM and the corresponding depth values related to the original surface height at the control point locations. Nevertheless, these two datasets are estimated in different coordinate systems with a different scale. So we need to transform these two datasets to the same coordinate system – the reference camera coordinate system.

We use the method in [22] to implement the transformation of the height estimated by photometric stereo from pixel size to real size, and obtain the real depths of the target surface. After transforming the two datasets to the same coordinate system, the difference values (v) of control points can be packed as:

$$v_r = g_r - b_r. ag{7}$$

where g is the depth value from SfM which is taken as ground truth, b is the depth value estimated from photometric stereo, and they are all defined in the reference camera coordinate system.

Finally, we use smoothed thin-plate spline interpolation (STPS) to find a smooth surface that passes through all control points. This method has a strong anti-interference ability and can fit the low-frequency deviation surface well. Compared to TPS, STPS adds a smoothing parameter for regularization [27]. The chosen function f does not necessarily exactly interpolate all the control points but minimizing

$$E(f) = \sum_{i=1}^{m} |f(X_i) - y_i|^2 + \lambda \int_{\mathbb{R}^2} |D^2 f(X)|^2 dX, \qquad (8)$$

where λ is the smoothing parameter and is a chosen prior, (X_i, y_i) is control spaced tabulated data (X is the image coordinate in our algorithm), and m is the number of control points ($m \leq 20$ in our algorithm). And the interpolation function f about control depth values v is of the form

$$f(x,y) = \sum_{i} w_i G(d_i(x,y))) + b_0 + b_x x + b_y y, \qquad (9)$$



Fig. 3. The experimental setup including a camera and a LED video light source.

where $d_i(x, y)$ is the distance between the *i*-th control point and points of (x, y), $G(x) \equiv x^2 \log x$ and (w_i, b_0, b_x, b_y) are coefficients. The coefficients are computed as described in [25].

The final surface Z_c is corrected by adding the estimated deviation surface from the original surface produced by dynamic photometric stereo. With the camera projection matrix **M**, we can get a dense surface **P** for each pixel **p** of the reference frame in the world coordinate system

$$\mathbf{P} = Z_c \mathbf{M}^{-1} \mathbf{p}. \tag{10}$$

IV. EXPERIMENTAL RESULTS

Two types of cameras are used as our prototype system: one is a digital camera with an attached LED video light, and the other is a smartphone. A light source is fixed to the camera with a known lighting direction. For the digital camera, the light source is right above the camera with a distance of 200mm (Fig. 3). During the capturing, the light source is always turn on. The target objects are usually right ahead of the camera and captured in the center of the image.

We have validated our method on real datasets. We respectively use a digital camera (Kodak) and a smartphone (Honor V9 Play) to take photos of target objects in a room only with natural light.

A. Results of using a digital camera

We first used a digital camera (Kodak) to recover the shape of a Chinese leaf. In this experiment, the distance between the camera and objects was around 300mm. The lighting angle along the optical axis was 40 degrees. The camera's resolution was set to 640×480 . Figure 4-a shows the four input images of a Chinese leaf from different views, and image 1 was chosen as the reference image. By using our pixel match strategy, all the other images were transformed to the reference view (Fig. 4-b). Figure 4 displays our final results on the recovered 3D shape and surface normal map. The shape of the Chinese leaf fits the ground truth without the low-frequency noise. We can clearly observe the texture of the leaf and some abnormal objects on the leaf.

We also tested our method on a human face model. In this experiment, the distance between the camera and objects along the optical axis was around 800mm. The light was fixed to the camera with a slant angle of 15 degrees along the optical



Fig. 4. Results of the Chinese leaf captured by a digital camera. (a) The input images. (b) The matched images under the reference view of image 1. (C) Sparse points computed by SfM. (d) The estimated normal map. (e) The final 3D shape of the Chinese leaf recovered by our method.



Fig. 5. Results on a human face captured by a digital camera. (a) The input face images. (b) The matched images under the reference view of image 1. (c) Sparse points computed by SfM. (d) The normal map estimated by our method. (e) The final 3D face shape recovered by our method.

axis. The camera's resolution was set to 2048×1536 . Figure 5 shows the recovered face model by our method. We effectively remove the low-frequency deviation and obtain a good face shape with details.

B. Results of using a smart phone

We have also used a smartphone (Honor V9 Play) to recover the 3D model of a Chinese leaf and a rag doll. In these experiments, the camera's resolution was set to 612×816 . The flashlight was always on when capturing the target objects. In this way, the light direction is fixed to 0 degrees along the optical axis. The distance between the camera and objects along the optical axis was around 300mm. Figure 2-a and Figure 6-a show the input images of the Chinese leaf and the rag doll. As shown in Fig. 2-b and Fig. 6-b, our method obtains a good match result. Figure 7 and Figure 8 display the final recovered shape and normal of the Chinese leaf and the doll. The textures in the leaf and the rag doll are clearly observed. In comparison with the digital camera, the smartphone has a smaller work distance, but it also can reconstruct the surface shape with good details.

V. CONCLUSION

In this paper, we present a novel approach to recover surface details using a moving hand-held camera with an attached



Fig. 6. The real dataset of a rag doll captured by a smartphone. (a) The input images with different views. (b) The matched images under the reference view of image 3.

light. We extend the classic photometric stereo from handling static objects to relative dynamic objects. A robust pixel match strategy is proposed to establish the correspondences between the reference image and other key images. Then, we estimate the surface normal and correct the deviation in the integrated surface shape. We have validated our method on real datasets for good performances on surface recovery.

Our method is more practical than those in previous work. Nevertheless, there are some limitations. The proposed pixel



Fig. 7. Results on the Chinese leaf captured by a smartphone. (a) The recovered normal map by our method. (b) The final 3D shape recovered by our method.



Fig. 8. Results on the rag doll captured by a smartphone. (a) The recovered normal map by our method. (b) The final 3D shape recovered by our method.

match approach failed in images under partial occlusion and the pixel match is not good enough. Therefore, we are going to improve the match strategy for better normal and shape estimation, while we would like to implement all these procedures in a full 3D model. In the future work, we will optimize our method to run in real time for practical applications, e.g., using in Augmented Reality or in 3D object detection and recognition.

REFERENCES

- Robert J Woodham, "Photometric method for determining surface orientation from multiple images," *Optical engineering*, vol. 19, no. 1, pp. 191139, 1980.
- [2] Carlos Hernandez Esteban, George Vogiatzis, and Roberto Cipolla, "Multiview photometric stereo," *IEEE Transactions on Pattern Analysis* and Machine Intelligence, vol. 30, no. 3, pp. 548–554, 2008.
- [3] Li Zhang et al., "Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multiview stereo," in *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on.* IEEE, 2003, pp. 618–625.
- [4] Jongwoo Lim, Jeffrey Ho, Ming-Hsuan Yang, and David Kriegman, "Passive photometric stereo from motion," in *Computer Vision*, 2005. *ICCV 2005. Tenth IEEE International Conference on*. IEEE, 2005, vol. 2, pp. 1635–1642.
- [5] Neel Joshi and David J Kriegman, "Shape from varying illumination and viewpoint," in *Computer Vision*, 2007. ICCV 2007. IEEE 11th International Conference on. IEEE, 2007, pp. 1–7.
- [6] Xinxin Zuo, Sen Wang, Jiangbin Zheng, and Ruigang Yang, "Detailed surface geometry and albedo recovery from rgb-d video under natural illumination," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3133–3142.

- [7] Changchang Wu, "Towards linear-time incremental structure from motion," in 3D Vision-3DV 2013, 2013 International conference on. IEEE, 2013, pp. 127–134.
- [8] Johannes L Schonberger and Jan-Michael Frahm, "Structure-frommotion revisited," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 4104–4113.
- [9] Bill Triggs, Philip F McLauchlan, Richard I Hartley, and Andrew W Fitzgibbon, "Bundle adjustment – a modern synthesis," in *International* workshop on vision algorithms. Springer, 1999, pp. 298–372.
- [10] Diego Nehab, Szymon Rusinkiewicz, James Davis, and Ravi Ramamoorthi, "Efficiently combining positions and normals for precise 3d geometry," in ACM transactions on graphics (TOG). ACM, 2005, vol. 24, pp. 536–543.
- [11] Zhenglong Zhou, Zhe Wu, and Ping Tan, "Multi-view photometric stereo with spatially varying isotropic materials," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on.* IEEE, 2013, pp. 1482–1489.
- [12] Martin Grochulla and Thorsten Thormählen, "Combining photometric normals and multi-view stereo for 3d reconstruction," in *Proceedings* of the 12th European Conference on Visual Media Production. ACM, 2015, p. 10.
- [13] Reza Sabzevari, Alessio Del Bue, and Vittorio Murino, "Structure from motion and photometric stereo for dense 3d shape recovery," in *International Conference on Image Analysis and Processing*. Springer, 2011, pp. 660–669.
- [14] Atsuto Maki, Mutsumi Watanabe, and Charles Wiles, "Geotensity: Combining motion and lighting for 3d surface reconstruction," *International Journal of Computer Vision*, vol. 48, no. 2, pp. 75–90, 2002.
- [15] Denis Simakov, Darya Frolova, and Ronen Basri, "Dense shape reconstruction of a moving object under arbitrary, unknown lighting.," in *ICCV*, 2003, vol. 3, pp. 1202–1209.
- [16] Peter N Belhumeur, David J Kriegman, and Alan L Yuille, "The basrelief ambiguity," *International journal of computer vision*, vol. 35, no. 1, pp. 33–44, 1999.
- [17] Tomoaki Higo, Yasuyuki Matsushita, Neel Joshi, and Katsushi Ikeuchi, "A hand-held photometric stereo camera for 3-d modeling," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 1234–1241.
- [18] Qing Zhang, Mao Ye, Ruigang Yang, Yasuyuki Matsushita, Bennett Wilburn, and Huimin Yu, "Edge-preserving photometric stereo via depth fusion," in *Computer Vision and Pattern Recognition (CVPR)*, 2012 *IEEE Conference on*. IEEE, 2012, pp. 2472–2479.
- [19] Itsik Horovitz and Nahum Kiryati, "Depth from gradient fields and control points: Bias correction in photometric stereo," *Image and Vision Computing*, vol. 22, no. 9, pp. 681–694, 2004.
- [20] Zhengyou Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [21] Boxin Shi, Zhe Wu, Zhipeng Mo, Dinglong Duan, Sai-Kit Yeung, and Ping Tan, "A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo," in *Proceedings of the IEEE Conference* on Computer Vision and Pattern Recognition, 2016, pp. 3707–3716.
- [22] Hao Fan, Lin Qi, Nan Wang, Junyu Dong, Yijun Chen, and Hui Yu, "Deviation correction method for close-range photometric stereo with nonuniform illumination," *Optical Engineering*, vol. 56, no. 10, pp. 103102, 2017.
- [23] Andrea Fusiello, Emanuele Trucco, and Alessandro Verri, "A compact algorithm for rectification of stereo pairs," *Machine Vision and Applications*, vol. 12, no. 1, pp. 16–22, 2000.
- [24] David Eberly, "Thin plate splines," *Geometric Tools Inc*, vol. 2002, pp. 116, 2002.
- [25] Fred L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 11, no. 6, pp. 567–585, 1989.
- [26] Yvain Quéau and Jean-Denis Durou, "Edge-preserving integration of a normal field: Weighted least-squares, tv and 11 approaches," in *International Conference on Scale Space and Variational Methods in Computer Vision.* Springer, 2015, pp. 576–588.
- [27] Grace Wahba, Spline models for observational data, vol. 59, Siam, 1990.