FFT-BASED FILTERING APPROACH TO FUSE PHOTOGRAMMETRY AND PHOTOMETRIC STEREO 3D DATA

A. Karami 1,2*, M. Varshosaz 3, F. Menna 1, F. Remondino 1, T. Luhmann 4

¹ 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy - (akarami, fmenna, remondino)@fbk.eu web: <u>http://3dom.fbk.eu</u>

²University of Trento, Italy - ali.karami@unitn.it

³Geomatics Engineering Faculty, K.N.Toosi University of Technology, Tehran, Iran - varshosazm@kntu.ac.ir

⁴ Institute for Applied Photogrammetry and Geoinformatics, Jade University of Applied Sciences, Oldenburg, Germany luhmann@jade-hs.de

Commission IV, WG IV/3

KEYWORDS: Photogrammetry, Photometric stereo, High-resolution 3D reconstruction, Fast Fourier Transform (FFT), Fusion.

ABSTRACT:

Image-based 3D reconstruction has been successfully employed for micro-measurements and industrial quality control purposes. However, obtaining a highly-detailed and reliable 3D reconstruction and inspection of non-collaborative surfaces is still an open issue. Photometric stereo (PS) offers the high spatial frequencies of the surface, but the low frequency is erroneous due to the mathematical model's assumptions and simplifications on how light interacts with the object surface. Photogrammetry, on the o ther hand, gives precise low-frequency information but fails to utilize high frequencies. As a result, in this research, we present a fusion strategy in Fourier domain to replace the low spatial frequencies of PS with the corresponding photogrammetric frequencies in order to have correct low frequencies while maintaining high frequencies from PS. The proposed method was tested on three different objects. Different cloud-to-cloud comparisons were provided between reference data and the 3D points derived from the proposed method to evaluate high and low frequency information. The obtained 3D findings demonstrated how the proposed methodology generates a high-detail 3D reconstruction of the surface topography (below $20 \,\mu$ m) while maintaining low-frequency information (0.09 μ m on average for three different testing objects) by fusing photogrammetric and PS depth data with the proposed FFT-based method.

1. INTRODUCTION

Image-based three-dimensional (3D) reconstruction is a process of generating the 3D shape of an object from a set of 2D images. In different fields, there is a raising need for very high-resolution 3D information for micro-measurements and quality inspection of object surfaces. Among the available image-based techniques, photogrammetry and PS have always been considered as successful, cost-effective, portable and flexible techniques in many applications, including cultural heritage (Remondino et al., 2016), industrial inspection (Ahmadabadian et al., 2019), and quality control (Karami et al., 2022a). However, each method has its own limitations. For instance, achieving a precise 3D reconstruction of highly reflective and textureless surfaces (Fig. 1) using photogrammetry remains a difficult task. This is because of the sensitivity of the method to the characteristics of the surface texture. Consequently, a noisy result on poorly textured objects is generally generated (Karami et al., 2022b; Ahmadabadian et al., 2019).

On the other hand, photometric stereo (PS) derives a dense field of normals from a sequence of 2D images acquired under various illuminations. It is assumed that all surveyed objects have a Lambertian surface and are lit from at least three distinct directions. PS is a primitive method that has been studied for decades, due to its simple concept and streamlined data collection. Unlike other triangulation-based techniques, PS has the benefit of preserving geometry information at a high frequency. However, its low frequency information is inaccurate as a result of its sensitivity to surface characteristics and the mathematical model used to explain how light interacts with an object's surface resulting in global deformation of the shape (Shi et al., 2018; Li et al., 2020; Karami et al., 2021).



Figure 1. Two examples of non- collaborative objects considered in this research.

^{*} Corresponding author

1.1 Aim of the work

The presented work aims to combine photogrammetry and PS depth maps to get detailed and precise 3D reconstructions of non-collaborative (i.e. shiny, texture-less, translucent, etc.) surfaces. The primary goal of combining two techniques is to overcome the constraints and limitations of one method by leveraging the strengths of the other, allowing the generation of complete and precise 3D reconstruction of optically non-cooperative objects.

In PS depth map, inaccurate low frequencies are normally present due to several not-fulfilled assumptions of the PS mathematical model, such as ideal diffuse reflection with no shadow nor specularities on the surface, parallel illumination direction and orthogonal projection. But high frequency information is preserved with high accuracy regardless of these assumptions. Photogrammetry, on the other hand, fails to exploit high frequencies when the assumption of ideal diffuse reflection with a well-textured surface is not satisfied, despite the fact that its low-frequency information is still reliable. Therefore, the paper proposes to fuse the high spatial frequencies of photometric stereo with low frequencies from photogrammetry in order to have accurate low frequencies and meanwhile keeping high frequencies. To do so, we convert both photogrammetry and PS depth maps to the Fourier domain, decomposing the 2D depth map into sine and cosine components where low and high frequencies could be distinguished/modified more easily. Then, we use a non-linear interpolation to remove inaccurate frequencies while keeping and fusing accurate high and low frequencies.

The rest of the article is structured as follows. Section 2 summarizes prior work on non-collaborative object 3D reconstruction using photogrammetry, PS, and their combination. The proposed integrated technique is described in Section 3. Section 4 presents and discusses the experiments and results of the proposed method. Finally, conclusions are reached are outlined.

2. STATE OF THE ART

In this Section, we are summarizing many research works related to the 3D measurement of optically non-collaborative objects (such as textureless, metallic, and shiny surfaces) into three different categories including multi-view stereo (photogrammetry), PS, and a combination of both approaches.

Nulti-View Stereo (**MVS**) is a photogrammetric-based 3D reconstruction approach offering low-cost data acquisition, automated image processing, and high accuracy in the 3D reconstruction of well-textured objects (Remondino et al., 2014; Ahmadabadian et al., 2017). The MVS method has been used in various areas such as reverse engineering (Menna et al., 2010; Geng and Bidanda, 2017), medicine (De Benedictis et al., 2018; Kim et al., 2018), quality control (Gao et al., 2019), industrial inspection (Rodríguez-Gonzálvez et al., 2017; Karami et al., 2022a) or 3D micromeasurement (Atsushi et al., 2011; Lu and Cai, 2020). However, photogrammetric methods failed or produce noisy results in areas with textureless or repetitive textures, or significant illumination variations across camera stations. To

address these challenges, prior studies focused on improving the visual texture of objects by projecting a known (Hafeez et al., 2020; Menna et al., 2017), random (Hosseininaveh et al.,2015; Ahmadabadian et al.,2017, 2019), or synthetic (Hafeez et al., 2020; Santoši et al., 2019) pattern onto the object, assuming that the surface is Lambertian. In the case of non-Lambertian surfaces with specular reflections or inter-reflection, different approaches such as cross polarisation (Nicolae et al., 2014; Menna et al., 2016), image pre-processing methods (Wallis, 1976; Calantropio et al., 2020), or spraying with powder (Mousavi et al., 2018; Pereira et al.,2019) have also been used to reduce the saturated area on the images. However, covering the reflective surface with white or colored powder might be impractical, especially with delicate cultural heritage measures or real-time 3D inspection. Furthermore, the extra layer may enlarge the object's volume and smooth out 3D roughness and microstructures.

Photometric stereo (PS) aims at recovering the surface normals of a static scene from a set of images captured under different light directions with a fixed camera position (Woodham, 1980; Shi et al., 2018; Abzal et al., 2019). PS can recover a very detailed topography of non-collaborative objects which is impossible using MVS approaches with a few images (Shi et al., 2018; Karami et al., 2021). However, the conventional PS (Woodham, 1980) is based on the ideal Lambertian reflectance model and analyzes per-pixel illumination observation variations. As a result, real-world objects with complicated reflectance geometry cannot fulfill the ideal Lambertian reflectance model, resulting in an inaccurate 3D model with global shape deformation (Shi et al., 2018; Abzal et al., 2018). Several works have been conducted to deal with non-Lambertian surfaces in which non-Lambertian reflectances were treated as outliers, such as techniques based on the RANSAC schemes (Mukaigawa et al., 2007), rank minimization (Wu et al., 2010), or median (Miyazaki et al., 2010). However, they assume that the majority of data meet the Lambertian distributions (Shi et al., 2018; Cao et al., 2022). These approaches need a large number of input images and struggle with complicated non-Lambertian surfaces with board shadowed regions or strong specularity. Some methods (Blinn, 1977; Cook and Torrance, 1982; Georghiades, 2003; Alldrin et al., 2008; Boss et al., 2020) also use different parametric or nonparametric models to model the light behavior when interacting with the non-Lambertian surfaces such as Torrance-Sparrow model (Georghiades, 2003), Cook-Torrance model (Cook and Torrance, 1982), the Ward model (Goldman et al., 2009), or Blinn-Phong model (Blinn, 1977). These advanced reflectance model-based approaches, however, are only applicable to limited types of materials (Shi et al., 2018; Cao et al., 2022). Most recently, learningbased algorithms (Santo et al., 2017; Chen et al., 2020; Samo et al., 2022) capable of approximating highly non-linear mappings have lately been used to handle complicated PS issues. However, the training process requires massive datasets specified for the particular cases, and there is still a gap between real-world and synthetic images, making complicated real-world generalization to objects challenging.

Various researchers tackled the image-based 3D reconstruction of non-collaborative surface 3D

reconstruction through a combination of both photogrammetry and PS. Hernandez et al. (2008) use the silhouettes of the object to get accurate geometry information and combine it with PS to reconstruct 3D details for Lambertian objects. Weber et al. (2004) estimate geometry and reflectance continuously for objects with piece-wise smooth surfaces taking the advantage of multiview PS. Birkbeck et al. (2006) use a variational strategy to cope with non-Lambertian surfaces using a Phong reflectance model assuming that the light sources are calibrated and known. Ahmed et al. (2008) extract normal map and enhance geometry of the final model using calibrated lighting system and multi-view video. More similarly, Vlasic et al. (2008) proposed a high-resolution PS system with a carefully designed lighting structure to estimate the surface normals for moving objects. These works (Joshi et al., 2007; Li al., 2020; Karami et al., 2021) also focus on adding point light sources to traditional stereo, dealing with the assumption of parallel lighting direction and orthogonal projection. More recently, Ren et al., (2021, 2022) combined the PS and sparse 3D points extracted using contact measurements with a CMM to generate highresolution surface topography. Although these systems have tremendously advanced metrology progress, they are limited to highly sophisticated labs and applications with unique metrological demands due to the high cost of the equipment. Other hybrid techniques based on the fusion of

photogrammetry and PS (Park et al., 2016; Logothetis et al., 2019; Li et al., 2020; Karami et al., 2021; Karami et al., 2022a) have recently been proposed, with photogrammetry being used to minimize the global deformation produced by PS.

However, some of these approaches are time-consuming and/or require special and expensive equipment, while some others require either accurate light-source calibration or careful illumination design making it difficult or unattractive for real-time 3D inspection of objects.

Furthermore, most of the approaches apply assumptions and limitations to streamline and make the issue more feasible. As a result, their approach's application window is often too restricted, yielding them unstable and untrustworthy for practical industrial applications that must cope with a diverse set of objects, conditions, and environments. In contrast, our proposed approach (Figure 2) imposes none of these assumptions and constraints and, indeed, explicitly takes the depth map from both photogrammetry and PS and fuse them in frequency domain in which to remove PS low-frequency signals (inaccurate low-frequency; accurate high-frequency) gradually replace with and it corresponding photogrammetric low-frequency signals (accurate low frequency; inaccurate high frequency) to generate a reliable and high-detail 3D reconstruction of non-collaborative surfaces.



Figure 2. General steps - presented in Section 3 - to fuse photogrammetry and PS data using FFT filter. A) acquire the images, B) generate depth using both photogrammetry and PS, C) transfer them to the FFT domain, D) Filter low frequency from PS and high frequency from photogrammetry, E) fuse them using a non-linear blending approach, F) apply inverse FFT to convert the fused depth back to spatial domain and convert the depth map to a point cloud.

3. PROPOSED SOLUTION

Figure 2 shows a general overview of 3D integrated depth maps using the proposed method. The first step is to set up an automated data acquisition system to collect stack of images obtained under different illuminations from different viewpoints (camera stations). Then, after generating the depth maps using both photogrammetry and PS (Karami et al., 2021, 2022a), they are converted to Fourier domain using Fourier transformation decomposing each 2D depth map into its sine and cosine components. In this way it is more convenient to distinguish and modify low and high frequencies. In the Fourier domain, we first create a weighting plane to assign value to each pixel, and then we use a non-linear interpolation to eliminate incorrect frequencies while fusing accurate low and high frequencies. Finally, the generated fused depth maps are transferred back to spatial domain using an inverse Fourier transformation and then converted to 3D point clouds.

3.1 Image acquisition system

The suggested image capturing system (Figure 2A) is a modular system assembled from off-the-shelf optical breadboard components and 3D printed elements that are custom made. The system consists of four basic components: a single digital camera (A Nikon D3X DSLR camera with a resolution of 24 Mpx mounting an AF-S Micro NIKKOR 60mm f/2.8G ED lens), 20 dimmable LED lights attached on four vertical poles, an object table (positioned nearly 50 cm from the objects and camera) to mount the object, and an Arduino microcontroller with electronic circuitry to synchronize the camera and LEDs.

A close-range photogrammetry pipeline was used to calibrate the geometry of the lighting system (light locations with respect to a specified local coordinate system) and estimate the interior and exterior camera parameters. Five rigid square plates with eight imbedded coded targets (Figure 2A) were utilized for this purpose, and they were placed in the middle and corners of the main optical breadboard. Karami et al. (2022a) presented further details regarding the system's structure and the calibration procedure.

To acquire images, once the object is placed on the turntable and the camera is fixed at the first station, the first LED is turned on for half a second, and then the camera takes the image. Once all other LEDs have been used in the sequence, the camera moves on to the following station. The process is repeated for all the viewpoints. Finally, at each viewpoint 20 images are acquired (Figure 2B).

3.2 Fourier transformation

After the image acquisition, photogrammetric depth maps (Figure 2C) are generated using image pair captured from different stations (Karami et al., 2021, 2022a). On the other side, the PS depth maps (Figure 2D) are generated given the stack of 20 images taken only at one viewpoint and the known calibrated light direction (Karami et al., 2021, 2022a). Afterward, both depth maps are transferred to the Fourier (or frequency) domain using FFT. For an image of

size M×N, the two-dimensional Fourier Transform is expressed in Equation (1)

$$F(p,q) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} f(u,v) * e^{-j(2\pi/M)p_m} e^{-j(2\pi/N)q_n}$$

$$p = 0, 1, ..., M-1$$

$$q = 0, 1, ..., N-1$$
(1)

where
$$f(v, u) = a 2D$$
 image in the spatial domain $F(p,q) = a 2D$ image in Fourier domain

It should be noted that the exponential term is the basis function corresponding to each point F(p, q) in the Fourier space.

3.3 Weighting plane

Typically, the Fourier image is shifted in a way that F(0,0) is displayed in the center of the image. An image point's associated frequency rises with distance from the center. Thus, low frequency power is concentrated in a small area at the center of the shifted frequency plane, while high frequency power is displayed as one moves away from the center as shown in Figure 3.



Figure 3. Two examples of photogrammetric (A) and PS (B) depth transformation in Fourier domain for Object 03.

Therefore, we define a Gaussian weighting plane (W) using Equation (2), which assigns each pixel in the Fourier domain a weight between 0 and 1 based on its distance from the plane's center (Fourier image).

$$W = exp^{-\hat{R}^2/_{2*T}}$$
(2)

$$\hat{R} = \frac{R - R_{min}}{R_{max} - R_{min}} \tag{3}$$

where W = Gaussian weighting plane

 \hat{R} = the normalized radial from center of the image T = shifting parameter

This should be noted that Rp (Figure 2E) in the weighting plane, where allocated weight to both depths are equal, can be changed by modifying parameter T in Equation (2). This parameter must be set empirically depending on the dataset.

3.4 Depth fusion in FFT domain

To fuse both depths, a non-linear interpolation (Figure 2E) expressed in Equations (4) is used to generate a fused 2D depth out of the photogrammetry and PS depth maps.

$$FusdF = W * F_{Pho} + (1 - W) * F_{PS} * \frac{Total_{F_{Pho}}}{Total_{F_{PS}}}$$

$$Total_{F_{Pho}} = \sum_{i=1}^{m} \sum_{j=1}^{n} F_{Pho}(i,j)$$

$$Total_{F_{PS}} = \sum_{i=1}^{m} \sum_{j=1}^{n} F_{PS}(i,j)$$
(4)

where F_{F}

 F_{Pho} = photogrammetric depth in Fourier domain F_{PS} = photometric stereo depth in Fourier domain $Total_{F_{Pho}}$ = frequency power for photogrammetry $Total_{F_{PS}}$ = frequency power for photometric stereo

With such interpolation, PS high frequencies are gradually replaced with photogrammetric low frequencies. As one moves away from the center, where the distribution of high frequency power is, the weighting trend progressively tums inverted. The weight assigned to photogrammetric frequency (Pho) drops to 0 while the weight assigned to photometric stereo (PS) rises to 1.

The final step after fusing both depths is to apply inverse Fourier Transform to transfer the generated fused depth to spatial domain (Figure 2F). Successively, the depth is converted to 3D point cloud given the 2D depth map and an RGB image (Pan et al., 2016).

4. EXPERIMENT AND RESULTS

The proposed FFT-based fusion method was tested on three non-collaborative objects (Figure 4a- first column): a gold foiled surface shaped like a Euro coin, a metallic object featuring complex geometry, poor texture, and reflective surface and a two Euro coin.

Three sets of multiple images (ground sample distance – GSD $\approx 37~\mu m$) were taken from different stations, each one under 20 different illuminations using our image acquisition system (Section 3.1).

Two different experiments were carried out to evaluate the accuracy potential of the proposed fusion method. The first test was accomplished to evaluate the accuracy of low-frequencies obtained by proposed method (Section 4.1, Figure 4. The second test was performed to evaluate the potential of the proposed method to exploit the high frequencies (Section 4.2, Figure 6).

4.1 Low frequencies evaluation

To evaluate the accuracy potential of the proposed method in the low-frequency domain, 3D results achieved with PS (Figure 4b), photogrammetry (Figure 4c) and the proposed method (Figure 4d) were geometrically compared using available reference data.

Since the photogrammetric 3D reconstruction's lowfrequency information is still trustworthy, it was utilized as reference data for object 01 (also due to the fact that a laser scanner 3D model was not available).



Figure 4. The result of the cloud-to-cloud comparison with reference data for photometric stereo, photogrammetry and proposed method on three non-collaborative objects.

To this end, a set of images taken from 30 different viewpoints was used. For object 02, a Hexagon AICON Primescan (Hexagon, 2022) with a nominal accuracy of 63 μ m was used to acquire the ground truth data. For object 03, an Evixscan 3D Fine Precision (Evixscan, 2022) with a special resolution of 20 μ m was used. The generated 3D points were aligned to the reference data using the ICP technique (Besl and McKay, 1992) and the RMSEs of the Euclidean distances were measured and compared to analyze low frequency information. This geometric comparison allows to estimate possible global deformations of the recovered 3D shapes.

The negative values of the legend (towards blue color) in Figure 4 indicates that the generated 3D surface is below the reference surface, while the positive values (towards red color) show areas above the reference surface.

The quantitative analysis demonstrates that the proposed fusion method performs noticeably better than PS. For instance, the estimated RMSE of Euclidean distances for PS for objects 01 and 02 are 5.76mm and 2.4mm, respectively. On the other hand, values for the proposed method decreases remarkably for both objects to 0.14mm and 0.09mm, respectively, which was quite close the estimated RMSE for photogrammetry (0.07mm). This analysis demonstrates that the proposed integrated method maintained high frequencies while also improving low spatial frequencies.

4.2 High frequencies evaluation

Object 03 can be a good example to emphasize the proposed fusion method's capability for recovering microstructures on a surface while keeping the low-frequency information since its surface is very reflective, with very detailed structures. A small patch (2.5mm*3mm) on Object 03 (Figure 05a) was selected and measured with a non-contact 3D optical SENSOFAR scanner (Figure 5b) at 0.5μ m resolution. Then, the reference data (Figure 5d) was compared to the same patch/area on the 3D data generated by the proposed method, PS and photogrammetry. To this end, the RMSE of the Euclidean cloud-to-cloud distances between the 3D points on the reconstructed and reference models was computed and compared using CloudCompare software (Figure 6). From the achieved results, it can be seen that the estimated RMSE for proposed method was less than 19 μ m which was quite close to the PS with RMSE 15 μ m. While the estimated RMSE for photogrammetry was almost two times higher (about 35 μ m) than the proposed method. The achieved 3D results indicated how the proposed method generate a highdetails 3D reconstruction of the surface topography quite similar to PS while preserving low-frequency information (Section 4.1) thanks to fusion of photogrammetric and PS depth measurements.

5. CONCLUSIONS

In this paper, we present a depth fusion method in frequency domain to take the advantage of photogrammetry and photometric stereo (PS) for recovering precise 3D reconstruction of non-collaborative objects. To this end, 2D depth maps are converted to the Fourier domain using the Fourier transformation, decomposing the 2D depth map into sine and cosine components, facilitating the differentiation and identification of low and high frequencies. To give a value to each pixel in the Fourier domain, a weighting plane first was established. Then, given the weighting plan and a non-linear interpolation, the wrong frequencies were removed while the accurate low and high frequencies were merged to generate a reliable and precise 3D reconstruction. Three different objects with different surface characteristics were tested and different comparative analyses were accomplished using reference data. The proposed FFT-based fusion method recovered high-resolution details with estimated RMSE below 20 µm quite close to photometric stereo results while inherited the geometric information (low frequency) from photogrammetry with the RMSE of less than $100 \,\mu\text{m}$ on these testing objects.



Figure 5. Reference data for high frequency evaluation. (a) A small area (letter R with dimension of 2.5 mm*3mm) on a two Euro coin (Object 03) was selected and measured. (b) The SENSOFAR scanner with optical resolution of 0.5µm and (c) the 12 overlapping scans to form the entire letter (R). (d) Final 3D model after alignment and co-registration of the scans.





ACKNOWLEDGEMENTS

The authors are thankful Oliver Kahmen (IAPG – Jade University) and Microgeo Srl (Florence, Italy) for their support in collecting the ground truth of the metallic industrial component used in our experiments.

We are grateful to Michele Welponer (3DOM-FBK) for his valuable technical contribution in the experiments related to the depth transformation to 3D point cloud.

REFERENCES

Abzal, A., Saadatseresht, M. and Varshosaz, M., 2018: Development of a novel simplification mask for multi-shot optical scanners. *ISPRS J. Photogramm*, 142, pp.12-20.

Abzal, A., Saadatseresht, M. and Varshosaz, M., 2019: Development of a new dual-pattern fringe projection scanner. *Photogramm. Rec.*, 34(165), pp.63-84.

Ahmed, N., Theobalt, C., Dobrev, P., Seidel, H.P. and Thrun, S., 2008: Robust fusion of dynamic shape and normal capture for high-quality reconstruction of time-varying geometry. *In 2008 IEEE Conference on CVPR* (pp. 1-8). IEEE.

Ahmadabadian, A.H., Yazdan, R., Karami, A., Moradi, M. and Ghorbani, F., 2017: Clustering and selecting vantage images in a low-cost system for 3D reconstruction of texture-less objects. *Measurement*, 99, pp.185-191.

Ahmadabadian, A.H., Karami, A. and Yazdan, R., 2019: An automatic 3D reconstruction system for texture-less objects. *Rob Auton Syst.*, 117, pp. 29-39.

Alldrin, N., Zickler, T. and Kriegman, D., 2008: Photometric stereo with non-parametric and spatially-varying reflectance. *In 2008 IEEE Conference on CVPR* (pp. 1-8). IEEE.

Atsushi, K., Sueyasu, H., Funayama, Y. and Maekawa, T., 2011: System for reconstruction of three-dimensional micro-objects from multiple photographic images. *Computer-Aided Design*, 43(8), pp.1045-1055.

Besl, P.J. and McKay, N.D., 1992: A method for registration of 3-D shapes. *IEEE PAMI*, 14(2).

Birkbeck, N., Cobzas, D., Sturm, P. and Jagersand, M., 2006: Variational shape and reflectance estimation under changing light and viewpoints. *In ECCV* (pp. 536-549). Springer, Berlin, Heidelberg.

Blinn, J.F., 1977: Models of light reflection for computer synthesized pictures. *In Proceedings of the 4th annual conference on Computer graphics and interactive techniques* (pp. 192-198).

Boss, M., Jampani, V., Kim, K., Lensch, H. and Kautz, J., 2020: Two-shot spatially-varying brdf and shape estimation. *In Proceedings of the IEEE/CVF Conference on CVPR* (pp. 3982-3991).

Cao, Y., Ding, B., He, Z., Yang, J., Chen, J., Cao, Y. and Li, X., 2022: Learning inter-and intraframe representations for non-Lambertian photometric stereo. *Opt. Lasers Eng.*, 150, p.106838.

Calantropio, A., Chiabrando, F., Seymour, B., Kovacs, E., Lo, E. and Rissolo, D., 2020: Image pre-processing strategies for enhancing photogrammetric 3d reconstruction of underwater shipwreck datasets. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 43.

Chen, G., Han, K., Shi, B., Matsushita, Y. and Wong, K.Y.K., 2020: Deep photometric stereo for non-lambertian surfaces. *IEEE Trans. Pattern Anal. Mach. Intell*, 44(1), pp.129-142.

Cook, R.L. and Torrance, K.E., 1982: A reflectance model for computer graphics. *ACM Transactions on Graphics* (*ToG*), 1(1), pp.7-24.

De Benedictis, A., Nocerino, E., Menna, F., Remondino, F., Barbareschi, M., Rozzanigo, U., Corsini, F., Olivetti, E., Marras, C.E., Chioffi, F., Avesani P., Sarubbo, S., 2018: Photogrammetry of the Human Brain: A Novel Method for Three-Dimensional Quantitative Exploration of the Structural Connectivity in Neurosurgery and Neurosciences. *World Neurosurgery*, 115:e279-e291.

Evixscan. Evixscan 3D Fine Precision [Online]. Available: <u>https://evixscan3d.com</u> [last access: July 14, 2022].

Gao, W., Haitjema, H., Fang, F.Z., Leach, R.K., Cheung, C.F., Savio, E. and Linares, J.M., 2019: On-machine and in-process surface metrology for precision manufacturing. *CIRP Annals*, 68(2), pp.843-866.

Geng, Z. and Bidanda, B., 2017: Review of reverse engineering systems-current state of the art. *Virtual Phys. Prototyp*, 12(2), pp.161-172.

Georghiades, A.S., 2003: Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. *In Computer Vision, IEEE International Conference on (Vol. 3, pp. 816-816). IEEE Computer Society.*

Goldman, D.B., Curless, B., Hertzmann, A. and Seitz, S.M., 2009: Shape and spatially-varying brdfs from photometric stereo. *IEEE PAMI*, 32(6), pp.1060-1071.

Hafeez, J., Lee, J., Kwon, S., Ha, S., Hur, G. and Lee, S., 2020: Evaluating Feature Extraction Methods with Synthetic Noise Patterns for Image-Based Modelling of Texture-Less Objects. *Remote Sens.*, 12(23), p.3886.

Hernandez, C., Vogiatzis, G. and Cipolla, R., 2008: Multiview photometric stereo. *IEEE PAMI*, 30(3), pp.548-554.

Hexagon. AICON PrimeScan scanner [Online]. Available: <u>https://www.hexagonmi.com</u> [last access: June 29, 2022].

Hosseininaveh, A., Yazdan, R., Karami, A., Moradi, M. and Ghorbani, F., 2015: A Low-cost and portable system for 3D reconstruction of texture-less objects. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 40(1), p.327.

Joshi, N. and Kriegman, D.J., 2007: Shape from varying illumination and viewpoint. *In 2007 IEEE 11th International Conference on Computer Vision* (pp. 1-7). *IEEE*.

Karami, A., Menna, F. and Remondino, F., 2021: Investigating 3d Reconstruction Of Non-Collaborative Surfaces Through Photogrammetry And Photometric Stereo. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 43, pp.519-526.

Karami, A., Menna, F., Remondino, F. and Varshosaz, M., 2022a: Exploiting Light Directionality for Image-Based 3D

Reconstruction of Non-Collaborative Surfaces. *Photogramm. Rec.*, 37(177), pp.111-138.

Karami, A., Battisti, R., Menna, F. and Remondino, F., 2022b: 3d digitization of transparent and glass surfaces: state of the art and analysis of some methods. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 43, pp.695-702.

Kim, A.J., Gu, D., Chandiramani, R., Linjawi, I., Deutsch, I.C.K., Allareddy, V. and Masoud, M.I., 2018: Accuracy and reliability of digital craniofacial measurements using a small-format, handheld 3D camera. *Orthod Craniofac Res.*, 21(3), pp.132-139.

Li, M., Zhou, Z., Wu, Z., Shi, B., Diao, C. and Tan, P., 2020: Multi-view photometric stereo: a robust solution and benchmark dataset for spatially varying isotropic materials. *IEEE Trans. Image Process*, 29, pp.4159-4173.

Logothetis, F., Mecca, R. and Cipolla, R., 2019: A differential volumetric approach to multi-view photometric stereo. *Proc. ICCV*, pp. 1052-1061.

Lu, Z. and Cai, L., 2020: Accurate three-dimensional measurement for small objects based on the thin-lens model. *Applied Optics*, 59(22), pp.6600-6611.

Menna, F., Nocerino, E., Remondino, F., Dellepiane, M., Callieri, M. and Scopigno, R., 2016: 3d digitization of an heritage masterpiece-a critical analysis on quality assessment. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 41.

Menna, F., Nocerino, E., Morabito, D., Farella, E.M., Perini, M. and Remondino, F., 2017: An Open Source Low-Cost Automatic System for Image-Based 3d Digitization. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 42, p.155.

Menna, F. and Troisi, S., 2010: Low cost reverse engineering techniques for 3D modelling of propellers. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 38(Part 5), pp.452-457.

Miyazaki, D., Hara, K. and Ikeuchi, K., 2010: Median photometric stereo as applied to the segonko tumulus and museum objects. *Int. J. Comput*, 86(2), pp.229-242.

Mousavi, V., Khosravi, M., Ahmadi, M., Noori, N., Haghshenas, S., Hosseininaveh, A. and Varshosaz, M., 2018: The performance evaluation of multi-image 3D reconstruction software with different sensors. *Measurement*, 120, pp.1-10.

Mukaigawa, Y., Ishii, Y. and Shakunaga, T., 2007: Analysis of photometric factors based on photometric linearization. JOSA A, 24(10), pp.3326-3334.

Nicolae, C., Nocerino, E., Menna, F. and Remondino, F., 2014: Photogrammetry applied to problematic artefacts. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 40(5), p. 451.

Pan, H., Guan, T., Luo, Y., Duan, L., Tian, Y., Yi, L., Zhao, Y. and Yu, J., 2016: Dense 3D reconstruction combining depth and RGB information. *Neurocomputing*, 175, pp.644-651.

Park, J., Sinha, S.N., Matsushita, Y., Tai, Y.W. and Kweon, I.S., 2016: Robust multi-view photometric stereo using planar mesh parameterization. *IEEE Trans. PAMI*, 39(8), pp. 1591-1604.

Pereira, J.R.M., de Lima e Silva Penz, I. and da Silva, F.P., 2019: Effects of different coating materials on threedimensional optical scanning accuracy. *Adv. Mech. Eng.*, 11(4), p.1687814019842416. Remondino, F., Spera, M.G., Nocerino, E., Menna, F., Nex, F., 2014: State of the art in high density image matching. *Photogramm. Rec.*, Vol. 29(146), pp. 144-166.

Remondino, F., Gaiani, M., Apollonio, F., Ballabeni, A., Ballabeni, M. and Morabito, D., 2016: 3D Documentation of 40 Kilometers of Historical Porticoes - the Challenge. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. XLI-B5, pp. 711-718.

Ren, M., Ren, J., Wang, X., Gao, F., Zhu, L. and Yao, Z., 2020: Multi-scale measurement of high-reflective surfaces by integrating near-field photometric stereo with touch trigger probe. *CIRP Annals*.

Ren, J., Jian, Z., Wang, X., Mingjun, R., Zhu, L. and Jiang, X., 2021: Complex Surface Reconstruction Based on Fusion of Surface Normals and Sparse Depth Measurement. *IEEE Trans. Instrum.*, 70, pp.1-13.

Rodríguez-Gonzálvez, P., Rodríguez-Martín, M., Ramos, L.F. and González-Aguilera, D., 2017: 3D reconstruction methods and quality assessment for visual inspection of welds. *Automation in Construction*, 79, pp.49-58.

Santo, H., Samejima, M., Sugano, Y., Shi, B. and Matsushita, Y., 2017: Deep photometric stereo network. *In Proceedings of the IEEE international conference on computer vision workshops* (pp. 501-509).

Santoši, Ž., Budak, I., Stojaković, V., Šokac, M. and Vukelić, D., 2019: Evaluation of synthetically generated patterns for image-based 3D reconstruction of texture-less objects. *Measurement*, 147, p.106883.

Sarno, F., Kumar, S., Kaya, B., Huang, Z., Ferrari, V. and Van Gool, L., 2022: Neural architecture search for efficient uncalibrated deep photometric stereo. *In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 361-371).

Shi, B., Mo, Z., Wu, Z., Duan, D., Yeung, S.K. and Tan, P., 2018: A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo. *IEEE Trans. Pattern Anal. Mach. Intell.*, 41(2), pp.271-284.

Vlasic, D., Baran, I., Matusik, W. and Popović, J., 2008: Articulated mesh animation from multi-view silhouettes. *In Acm Siggraph 2008 papers* (pp. 1-9).

Wallis, R.H., 1976: An approach for the space variant restoration and enhancement of images. *In Proc. Symposium on Current Mathematical Problems in Image Science*, Nov., 1976.

Weber, M., Blake, A. and Cipolla, R., 2004: Towards a complete dense geometric and photometric reconstruction under varying pose and illumination. *Image Vis. Comput.*, 22(10), pp.787-793.

Woodham, R.J., 1980: Photometric method for determining surface orientation from multiple images. *Opt. Eng.*, 19(1), p.191139.

Wu, L., Ganesh, A., Shi, B., Matsushita, Y., Wang, Y. and Ma, Y., 2010: Robust photometric stereo via low-rank matrix completion and recovery. *In ACCV* (pp. 703-717). Springer, Berlin, Heidelber.