• Review•

Summary study of data-driven photometric stereo methods

Qian ZHENG[1],　Boxin SHI[2 *],　Gang PAN[3]

1. *School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore*

2. *National Engineering Laboratory for Video Technology, Department of CS, Peking University, Beijing, China*

3. *State Key Lab of CAD&CG, College of Computer Science and Technology, Zhejiang University, Hangzhou, China*

* Corresponding author，　shiboxin@pku.edu.cn

Abstract Background　A photometric stereo method aims to recover the surface normal of a 3D object observed under varying light directions. It is an ill-defined problem because the general reflectance properties of the surface are unknown. **Methods**　This paper reviews existing data-driven methods, with a focus on their technical insights into the photometric stereo problem. We divide these methods into two categories, per-pixel and all-pixel, according to how they process an image. We discuss the differences and relationships between these methods from the perspective of inputs, networks, and data, which are key factors in designing a deep learning approach. **Results**　We demonstrate the performance of the models using a popular benchmark dataset. **Conclusions**　Data-driven photometric stereo methods have shown that they possess a superior performance advantage over traditional methods. However, these methods suffer from various limitations, such as limited generalization capability. Finally, this study suggests directions for future research.

Keywords Photometric stereo; Data-driven methods; Non-Lambertian reflectance

1 Introduction

Modern 3D computer stereo vision methods, including geometric (*e.g.*, binocular[1] and multi-view stereo[2]) and photometric approaches [3], have produced faithful 3D reconstructions from a set of images. Photometric methods are capable of reproducing the fine details of the surface at a superior resolution for highly accurate 3D shape reconstruction[4]. Despite its long history in computer vision[3], photometric stereo (PS) is still a fundamentally challenging research problem due to the unknown reflectance and global illumination effects of real-world objects[5]. Traditional methods address these difficulties by modeling non-Lambertian reflectance using a Bidirectional Reflectance Distribution Function (BRDF) (*e.g.*, analytical[6] or empirical[7] BRDF representations) and considering global illumination effects as outliers (*e.g.*, Sparse Bayesian

Learning [8]). However, such hand-crafted reflectance models are generally only effective for limited categories of reflectance[9].

Inspired by the powerful modeling capacity of deep neural networks for various computer vision tasks (*e.g*., light estimation[10], stereo vision[1]), researchers have investigated to develop practical reflectance models through data-driven approaches to solve the problem of photometric stereo. DPSN (Deep Photometric Stereo Network)[9] was the first attempt to address non-Lambertian reflectance using deep learning technologies. This approach requires that the testing probe shares the same pre-defined set of light directions as the training data, which limits its generalization. Therefore, a new model has to be retrained to test data with different lighting conditions. CNN-PS (Convolutional Neural Network based Photometric Stereo)[11], PS-FCN (Photometric Stereo using Fully Convolutional Network)[12], and IRPS (neural Inverse Rendering for general reflectance Photometric Stereo)[13] relax this constraint so that data with order-agnostic light directions can be tested. LMPS (Learning to Minify Photometric Stereo)[14] and SPLINE-Net (Sparse Photometric stereo though Lighting Interpolation and Normal Estimation)[15] further consider a small number of lighting conditions, which helps reduce the complexity of the data capture process. SDPS (Self-calibration Deep Photometric Stereo)[16] assumes uncalibrated lighting and achieves state-of-the-art performance. Moreover, Outdoor-PS (single day Outdoor Photometric Stereo)[17] applies data-driven photometric stereo methods to outdoor scenarios (*i.e.*, a partly cloudy or sunny day).

This paper reviews eight recent attempts to use data-driven methods to solve the problem of photometric stereo, based on our tutorial[1] and course[2] in the latest conferences. For a comprehensive discussion of non-learning based photometric stereo methods, we refer readers to survey papers[5,18,19].

2 Data-Driven Photometric Stereo Methods

A recent survey paper[5] divides traditional non-Lambertian photometric stereo methods into outlier rejection based methods[8,20], analytic BRDF modeling based methods[6], and empirical BRDF modeling based methods[7,21,22] according to the reflectance model adopted. These methods can also be categorized as per-pixel[6,7,8,21,22] or all-pixel[20] according to how they process the input images, *i.e.*, using either observed intensities for a single pixel or the whole image. We follow this simple strategy and divide data-driven methods into per-pixel [9,11,14,15] and all-pixel methods[12,13,16,17]. Figure 1 shows frameworks of a per-pixel [9] and an all-pixel method[12]. Besides the inputs and networks, Figure 1 also illustrates a training dataset[9] and a testing dataset[5] used by a data-driven method.

Data-driven photometric stereo methods aim to optimize a neural network $f(\cdot)$ such that

$$N = f(I),$$

where the input I can be either L observed intensities for a specific pixel or L observed images under L light directions[3]. The output surface normal N is accordingly represented by either a 3-dimensional vector or a
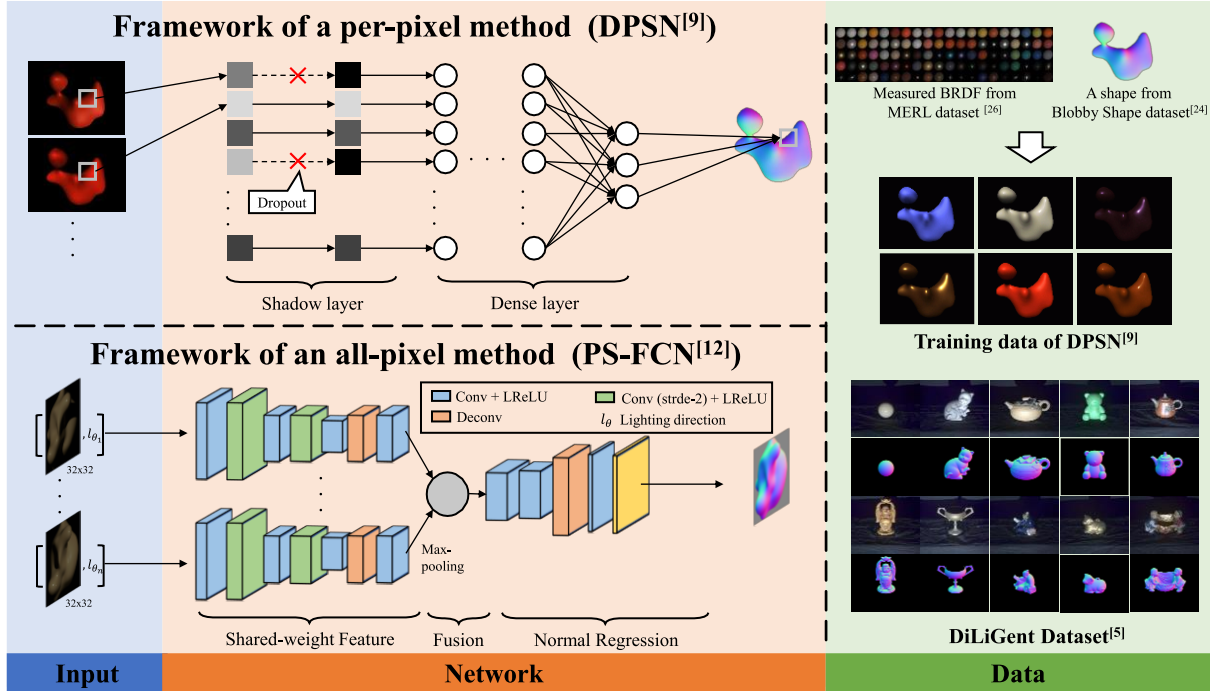
Figure 1 Frameworks of a per-pixel and an all-pixel method (left), examples of a training dataset[9] and a testing dataset[5] (right).

map with the same resolution as the input images, and $f(\cdot)$ is optimized through a training dataset. The following discussion avoids mathematical notation as much as possible and focuses on insights in the aspects of inputs, networks, and data.

2.1 Inputs

Per-pixel methods[9,11,14,15] take observed intensities as the input and output a surface normal for a single pixel, while all-pixel methods[12,13,16,17] directly take observed images or patches (multi-pixels) as the input and output a surface normal map with the same resolution as the input. This difference indicates that per-pixel methods aim to fit an accurate reflectance model for each pixel while all-pixel methods focus on extracting accurate surface normal maps from various appearances.

Per-pixel observation as an intensity profile. An intensity profile is a L-dimension vector that orders elements in observed intensities by the indices of light directions. As information about the light directions is not fully used during network training, methods with this input (*i.e.*, DPSN[9]) assume that the light directions in both the testing and training data are the same. This strong assumption limits the generalization, which means that a new model has to be retrained to test the data when the lighting conditions are different.

Per-pixel observation as an observation map. The observation map[11] has been proposed to overcome the shortcoming described above (Figure 2). It rearranges observation intensities according to light directions by directly encoding light directions as 2D coordinates and accordingly projecting observation intensities onto a 2D space. CNN-PS[11], LMPS[14], and SPLINE-Net[15] adopt this data structure to obtain inputs to their

neural networks. Since the information about light directions is fully retained, CNN-PS[11], LMPS[14], and SPLINE-Net[15] can handle inputs with order-agnostic lightings.
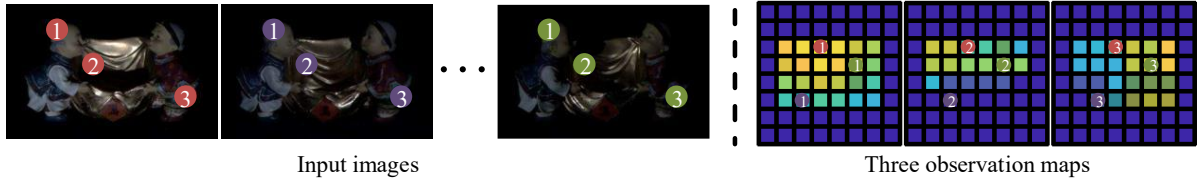

Input images   Three observation maps

Figure 2 The illustration of observation maps (from[11]). Left: different numbers represent different pixels and different colors indicate different light directions. Right: 2D coordinates of pixels on observation maps (representing intensities in input images) are determined by the projection of a light direction from a 3D space (hemisphere) to a plane.

All-pixel observation using the whole image. The input of Outdoor-PS[17] consists of L 16×16 image patches. These patches are ordered according to timestamps in a day (different light directions can be observed in different timestamps). Therefore, the light directions are not fully used during network training and Outdoor-PS[17] suffers from a similar limitation as that of DPSN[9] *i.e.*, a new model has to be retrained to test data when the lighting conditions are different.

All-pixel observation using patches. Other all-pixel methods use L whole images[16] as well as their corresponding light directions[12,13] as the input. To test data with order-agnostic lightings, these methods either leverage a sharing weights scheme[12,16] or impose an unsupervised manner[13]. In addition to classical photometric stereo, several recent advances in multispectral photometric stereo have also used an all-pixel method, For example, Antensteiner *et al.*[23] used the whole image as the input and estimated a surface normal through a U-Net[24], and Ju *et al.*[25] first adopted image patches to estimate a coarse surface normal map and then refined it using a per-pixel method. Multispectral photometric stereo involves two additional challenges not present in classical photometric stereo[3]: ambiguity brought by the decomposition of spectra and an extremely small number of observations (*i.e.*, often 3). Therefore, the following discussion focuses on photometric stereo that does not rely on spectral constraints.

2.2 Networks

We describe the overall architectures and specific designs of these data-driven photometric stereo methods in this section. For more details, refer to relevant papers[9,11,12,13,14,15,16,17].

Overall architectures. Except for DPSN[9], which leverages a classical deep neural network (DNN) architecture (consisting of an input layer, hidden layers, and an output layer), other data-driven methods impose a convolutional neural network (CNN) architecture. This is because both observation maps and natural images share the property of spatial continuity (Figure 2). In per-pixel methods[9,11,14,15], variations of DenseNet [23] are utilized to generate a 3-dimensional surface normal because DenseNet[26] is expected to strengthen feature propagation and encourage the reuse of features for low dimension (*i.e.*, 3) outputs[26]. In all-pixel methods[12,13,16,17], the idea of sharing weights among different modules is used because this design benefits from aggregating features extracted from multiple observations (*e.g.*, PS-FCN[12] in Figure 1) or enriching the features extracted from a single observation[13,16]. The architectures of simple setups (*i.e.*, known and large number of lightings[9,11,12,17]) are less complicated, while those of difficult ones (*i.e.*, PS

with a small number of light directions[15], unknown lightings[16], or in an unsupervised manner[13]) contain two sub-networks for joint optimization (*e.g.*, SDPS[16] in Figure 3).
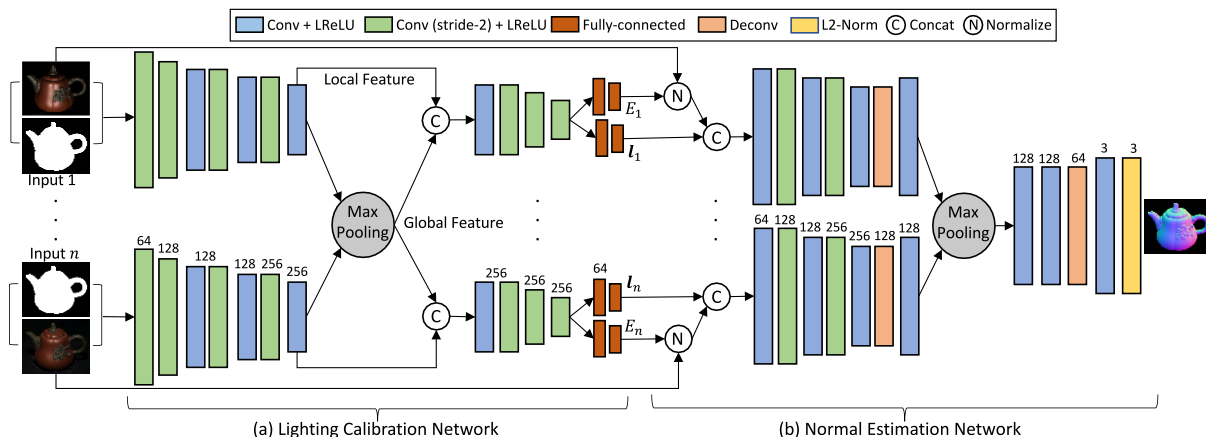

Figure 3 The framework of SDPS[16], which consists of a lighting calibration network and a normal estimation network.
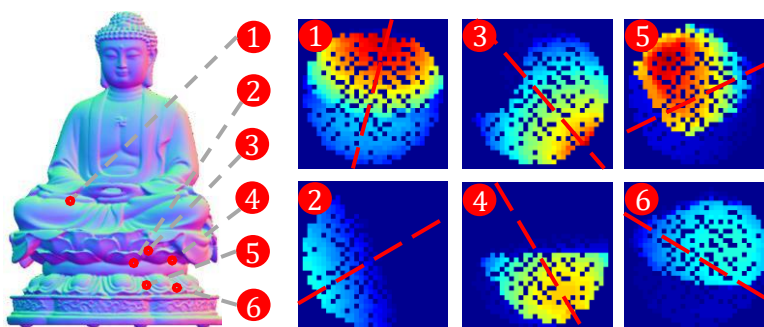

Figure 4 Illustration of symmetric (①②) and asymmetric patterns (③④⑤⑥) on observation maps (from[15]). Red lines represent the symmetric axis.

Specific designs. Despite cues that are implicitly learned from training data, data-driven methods design specific modules or loss functions to improve the robustness of photometric stereo. Per-pixel methods explicitly leverage the general reflectance properties (*i.e.*, isotropic BRDF and global illumination effects), while all-pixel methods focus on effectively regressing appearances to shapes in an end-to-end scheme.

- **Isotropic BRDF.** The consideration of isotropy can be used to narrow down the solution space during the training of neural networks. CNN-PS[11] shows that the observation map is rotational pseudo-invariant to the surface normal based on which additional data are augmented for training (by rotating the observation map and its surface normal simultaneously). SPLINE-Net[15] further demonstrates that an ideal observation map (*i.e.*, without global illumination effects) exhibits a symmetric pattern and proposes a symmetric loss function (Figure 4).

- **Global illumination effects.** Taking global illumination effects such as cast shadow and inter-reflection into account helps provide robust estimation of real-world data. To simulate shadows cast during training, DPSN[9] adopts the dropout operation, and LMPS[14] annotates the observation map by requiring some of the intensities to be zero. SPLINE-Net[15] shows that the global illumination effects can break the symmetry of an observation map (Figure 4) and introduces an asymmetric loss function.

- **End-to-end.** All-pixel methods pay more attention to the transfer of information among different modules from a high level view. These methods concatenate all observed images[13,16,17] and/or aggregate all features (from all images)[12,16] to extract shape information since all images contain the same surface normal map. Some of these methods further extract reflectance[13] (or lighting[16]) and shape information from a single image, as reflectance (or lighting) information is jointly determined by the single observation and the shape.

2.3 Data

Since the data used in outdoor PS methods[17] are quite different from those used for indoor methods, this section focuses on training and testing datasets used in indoor PS methods.

Training datasets. It is difficult to capture large scale data with the ground truth of surface normal. Therefore, most[4] data-driven photometric stereo methods synthesize data for training. Table 1 details the shapes, materials, light configurations, and rendering engines that have been used when synthesizing training data for different methods. As can be observed, 3D models come from the Blobby [27] or Sculpture Shape datasets[28] or the internet. Surface materials are approximated using MERL BRDF[29] or Disney's principled BSDF[30]. Light directions are fixed (the same as those of the testing data), uniformly sampled, or randomly sampled. Mitsuba[31] or Cycles[32] (*e.g.*, Outdoor-PS[17]) engines are employed for rendering. Table 1 also displays the number of shapes, materials, light directions, and images used by each method. It shows that per-pixel methods[9,11,14,15] require a small number of shape models (*i.e.*, smaller than 15) while all-pixel methods[12,16] impose a much larger number of shape models (*i.e.*, about 42K). This observation is consistent with the discussion in Section 2.1, which explained that per-pixel methods aim to fit an accurate reflectance model for each pixel while all-pixel methods focus on extracting accurate surface normal maps from various appearances.

Table 1   Details of training data regarding shape, material, light, and images for different data-driven photometric stereo methods

	Shape (Number)	Material (Number)	Light (Number)	Image (Number)
DPSN[9]	Blobby[27] (8)	MERL[29] (100)	Fixed (96)	Mitsuba[31] (76800)
CNN-PS[11], SPLINE-Net[15]	Internet (15)	Disney[30] (~15000)	Uniform (1300)	Cycles[32] (19500)
LMPS[14]	Blobby[27] (9)	MERL[29] (100)	Random (144)	Mitsuba[31] (10368)
PS-FCN[12], SDPS[16]	Blobby[27] & Sculpture[28] (~42K)	MERL[16] (100)	Random (64)	Mitsuba[31] (~5.4M)

Testing datasets. The DiLiGenT dataset[5] is the most widely used real-world dataset for evaluation. It consists of 10 different objects with different scales of non-Lambertian reflectance (Figure 1). Each object is illuminated and photographed under 96 different lighting directions. The ground truth of surface normal maps is also provided. Table 2 displays the quantitative results from these data-driven methods as well as a traditional photometric stereo method ST14[22] that achieves state-of-the-art performance. Note that the inputs of these methods are quite different, *i.e.*, DPSN[9], CNN-PS[11], PS-FCN[12], IRPS[13], ST14[22] take 96 images

---

[4] Note that IRPS[13] does not require any synthetic training data because it is an unsupervised method.

with known light directions as inputs, SPLINE-Net[15] takes 10 random images, LMPS[14] takes 10 optimal images, and SDPS[16] takes 96 images without lighting information. As can be observed, results from deep learning based methods (*i.e.*, DPSN[9], CNN-PS[11], PS-FCN[12], IRPS[13]) achieved much better performance as

Table 2 Quantitative comparisons in terms of angular error (degree) on the DiLiGenT dataset[5], green to red means small to large errors

	Ball	Bear	Buddha	Cat	Cow	Goblet	Harvest	Pot1	Pot2	Reading	Average
DPSN[9]	2.0	6.3	12.7	6.5	8.0	11.3	16.9	7.0	7.9	15.5	9.4
CNN-PS[11]	2.2	4.1	7.9	4.6	8.0	7.3	14.0	5.4	6.0	12.6	7.2
PS-FCN[12]	2.8	7.6	7.9	6.2	7.3	8.6	15.9	7.1	7.3	13.3	8.4
IRPS[13]	1.5	5.8	10.4	5.4	6.3	11.5	22.6	6.1	7.8	11.0	8.8
SPLINE-Net[15]	5.0	6.0	10.0	7.5	8.8	10.4	19.1	8.8	11.8	16.1	10.4
LMPS[14]	4.0	8.7	11.4	6.7	10.2	10.5	17.3	7.3	9.7	14.4	10.0
SDPS[16]	2.8	6.9	9.0	8.1	8.5	11.9	17.4	8.1	7.5	14.9	9.5
ST14[22]	1.7	6.1	10.6	6.1	13.9	10.1	25.4	6.5	8.8	13.6	10.3

Table 3 A summary of data-driven photometric stereo methods

	Method	Input	Architecture	General BRDF	End-to-end	Limitation
DPSN[9]	Per-pixel	Intensities	DNN	✓	✗	Constrained lightings
CNN-PS[11]	Per-pixel	Intensities	CNN	✓	✗	Sensitivity to global illumination effects
SPLINE-Net[15]	Per-pixel	Intensities	CNN	✓	✗	Sensitivity to global illumination effects
LMPS[14]	Per-pixel	Intensities	CNN	✓	✗	Constrained lightings
IRPS[13]	All-pixel	Images	CNN	✗	✓	Expensive computation
PS-FCN[12]	All-pixel	Images	CNN	✗	✓	Uniform materials
SDPS[16]	All-pixel	Images	CNN	✗	✓	Uniform materials
Outdoor-PS[17]	All-pixel	Images	CNN	✗	✓	Constrained lightings & uniform materials

compared with traditional methods when the same inputs were used. Moreover, deep learning based methods achieved comparable or better results as compared with traditional methods when the setting was more difficult (*i.e.*, a small number of inputs[14,15] or uncalibrated lightings[16]).5

Besides the quantitative evaluation, the visual quality was also evaluated using real data such as GOUDRD&APPLE[33] and Light Stage Data Gallery[34]. In addition, there are several synthetic datasets that can be used for validation[9,11,12,15,16,17]. These synthetic data are generally rendered using the same methods used for the training data.

3 Discussion

Despite the state-of-the-art performance achieved by data-driven methods, they suffer from limitations such as expensive computation for testing[13], sensitivity to global illumination effects[11,15], constrained light directions[9,14,17], and uniform material surfaces[12,16,17]. A brief summary of these methods and their

---

5 Please refer https://sites.google.com/site/photometricstereodata/single?authuser=0 for more results of traditional methods.

limitations is provided in Table 3. As can be observed, per-pixel methods pay more attention to the modeling of general BRDF, and they are robust to non-uniform distributions of surface materials; however, they perform less optimally for regions with global illumination effects since the shape information is not explicitly considered. All-pixel methods are commonly trained using various shapes with a uniform material for each shape. Therefore, they produce accurate results for regions with shadows or inter-reflection, while they are less accurate for objects with non-uniform materials. Based on the discussion above, we suggest the following directions for future research:

- **Combination.** This study demonstrates the unique characteristics of per-pixel and all-pixel methods. Mutually combining these two types of approaches to further improve performance will provide another research area to explore.

- **Lambertian reflectance.** A recent work[15] suggests that deep learning based methods generally produce unsatisfied results for Lambertian reflectance that are not comparable with the baseline method[3]. Therefore, reflectance that is less similar to DiLiGenT[5] or MERL[29] should be considered in a data-driven method to avoid overfitting.

- **Practicality.** As data-driven methods have achieved promising performance in a lab environment, leveraging deep learning techniques to solve the photometric stereo problem in a more under-constrained scenario (*e.g.*, Outdoor-PS[17]) should be considered.

References

1 Kendall A, Martirosyan H, Dasgupta S, Henry P, Kennedy R, Bachrach A, Bry A. End-to-end learning of geometry and context for deep stereo regression. In: Proceedings of the IEEE International Conference on Computer Vision, 2017, 66-75

DOI:10.1109/ICCV.2017.17

2 Furukawa Y, Ponce J. Accurate, dense, and robust multiview stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(8): 1362–1376

DOI:10.1109/tpami.2009.161

3 Woodham R J. Photometric method for determining surface orientation from multiple images. Optical Engineering, 1980, 19(1): 191139

DOI:10.1117/12.7972479

4 Park J, Sinha S N, Matsushita Y, Tai Y W, Kweon I S. Robust multiview photometric stereo using planar mesh parameterization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(8): 1591–1604

DOI:10.1109/tpami.2016.2608944

5 Shi B X, Wu Z, Mo Z P, Duan D L, Yeung S K, Tan P. A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA. IEEE, 2016

DOI:10.1109/cvpr.2016.403

6 Chen L X, Zheng Y Q, Shi B X, Subpa-Asa A, Sato I. A microfacet-based model for photometric stereo with general isotropic reflectance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019

DOI:10.1109/tpami.2019.2927909

7 Ikehata S, Aizawa K. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. IEEE, 2014

DOI:10.1109/cvpr.2014.280

8 Ikehata S, Wipf D, Matsushita Y, Aizawa K. Robust photometric stereo using sparse regression. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI. IEEE, 2012

DOI:10.1109/cvpr.2012.6247691

9 Santo H, Samejima M, Sugano Y, Shi B X, Matsushita Y. Deep photometric stereo network. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW). Venice. IEEE, 2017

DOI:10.1109/iccvw.2017.66

10 Garon M, Sunkavalli K, Hadap S, Carr N, Lalonde J F. Fast spatially-varying indoor lighting estimation. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, CA, USA. IEEE, 2019

DOI:10.1109/cvpr.2019.00707

11 Ikehata S. CNN-PS: CNN-based photometric stereo for general non-convex surfaces// Computer Vision – ECCV 2018. Cham: Springer International Publishing, 2018: 3–19

DOI:10.1007/978-3-030-01267-0_1

12 Chen G Y, Han K, Wong K Y K. PS-FCN: A flexible learning framework for photometric stereo// Computer Vision – ECCV 2018. Cham: Springer International Publishing, 2018: 3–19

DOI:10.1007/978-3-030-01240-3_1

13 Taniai T, Maehara T. Neural inverse rendering for general reflectance photometric stereo. In: International Conference on Machine Learning, 2018

14 Li J X, Robles-Kelly A, You S D, Matsushita Y. Learning to minify photometric stereo. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA. IEEE, 2019

DOI:10.1109/cvpr.2019.00775

15 Zheng Q, Jia Y M, Shi B X, Jiang X D, Duan L Y, Kot A. SPLINE-net: sparse photometric stereo through lighting interpolation and normal estimation networks. In: 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea (South). IEEE, 2019

DOI:10.1109/iccv.2019.00864

16 Chen G Y, Han K, Shi B X, Matsushita Y, Wong K Y K K. Self-calibrating deep photometric stereo networks. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA, USA. IEEE, 2019

DOI:10.1109/cvpr.2019.00894

17 Hold-Geoffroy Y, Gotardo P, Lalonde J F. Single day outdoor photometric stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019

DOI:10.1109/TPAMI.2019.2962693

18 Ackermann J, Goesele M. A survey of photometric stereo techniques. Foundations and Trends® in Computer Graphics and Vision, 2015, 9(3/4): 149–254

DOI:10.1561/0600000065

19 Herbort S, Wöhler C. An introduction to image-based 3D surface reconstruction and a survey of photometric stereo methods. 3D Research, 2011, 2(3): 4

DOI:10.1007/3dres.03(2011)4

20 Wu L, Ganesh A, Shi B X, Matsushita Y, Wang Y T, Ma Y. Robust photometric stereo via low-rank matrix completion and recovery// Computer Vision – ACCV 2010. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011: 703–717

DOI:10.1007/978-3-642-19318-7_55

21 Zheng Q, Kumar A, Shi B X, Pan G. Numerical reflectance compensation for non-lambertian photometric stereo. IEEE Transactions on Image Processing, 2019, 28(7): 3177–3191

DOI:10.1109/tip.2019.2894963

22 Shi B X, Tan P, Matsushita Y, Ikeuchi K. Bi-polynomial modeling of low-frequency reflectances. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(6): 1078–1091

DOI:10.1109/tpami.2013.196

23 Antensteiner D, Stolc S, Soukup D. Single image multi-spectral photometric stereo using a split u-shaped CNN. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019

24 Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation// Lecture Notes in Computer Science. Cham: Springer International Publishing, 2015: 234–241

DOI:10.1007/978-3-319-24574-4_28

25 Ju Y K, Dong X H, Wang Y Y, Qi L, Dong J Y. A dual-cue network for multispectral photometric stereo. Pattern Recognition, 2020, 100: 107162

DOI:10.1016/j.patcog.2019.107162

26 Huang G, Liu Z, van der Maaten L, Weinberger K Q. Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI. IEEE, 2017

DOI:10.1109/cvpr.2017.243

27 Johnson M K, Adelson E H. Shape estimation in natural illumination. In: CVPR 2011. Colorado Springs, CO, USA. IEEE, 2011

DOI:10.1109/cvpr.2011.5995510

28 Wiles O, Zisserman A. SilNet: single- and multi-view reconstruction by learning from silhouettes. In: Proceedings of the British Machine Vision Conference 2017, London, UK. British Machine Vision Association, 2017

DOI:10.5244/c.31.99

29 Matusik W, Pfister H, Brand M, McMillan L. A data-driven reflectance model. In: ACM SIGGRAPH 2003 Papers on - SIGGRAPH. San Diego, California. New York, USA: ACM Press, 2003 DOI:10.1145/1201775.882343

30 Burley B, Studios W D. Physically-based shading at Disney, part of practical physically based shading in film and game production. In: Proceedings of ACM SIGGRAPH Courses, 2012

31 Jakob, W.Mitsuba renderer, 2010

32 Cycles. https://www.cycles-renderer.org/

33 Alldrin N, Zickler T, Kriegman D. Photometric stereo with non-parametric and spatially-varying reflectance. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, USA. IEEE, 2008 DOI:10.1109/cvpr.2008.4587656

34 Einarsson P, Chabert C F, Jones A, Ma W C, Lamond B, Hawkins T, Bolas M, Sylwan S, Debevec P. Relighting human locomotion with flowed reflectance fields. In: Proceedings of the Eurographics Conference on Rendering Techniques, 2006, 183-194