Paper Reading Report-03

Shreya Chawla
u7195872

Abstract

*This is my reading report for the paper titled: "***PS-FCN: A Flexible Learning Framework for Photometric Stereo***", authored by Guanying Chen (The University of Hong Kong) et al, and published in IEEE ECCV 2018.*

All ENGN8501 submissions will be subject to ANU's TurnitIn plagiarism check, against both the original paper, internet resources, as well as all other students' submissions. So please make sure that you must write your own reports, and declare the following originality statement:

I, Shreya Chawla, hereby confirm that I am the sole author of this report and that I have compiled it in my own words.

1. Problem Statement

This paper attempts to address the problem of photometric stereo (PS) for non-Lambertian surfaces. Photometric stereo is a computer vision technique for estimating the surface normals of objects by observing the subject under different lighting conditions. Unlike the barely existing pure Lambertian for which the bidirectional reflectance distribution function (BRDF) is a constant vector, the estimation of normal matrix N by non-Lambertian photometric stereo is a more complex and common problem.

It has several applications including shape estimation, depth estimation, information transfer, and relative position estimation of 3D objects, which are relevant in robotics, autonomous car, information extraction from from aerial surveys etc. and hence is a significant problem [7].

The current leading techniques are [4], [5] for calibrated and [6] for uncalibrated rendering. However, [6] performs at par with [2] proposed in 2019 by the same authors as this paper (LCNet).

2. Summary of the paper's main contributions

[1] introduced PS-FCN network, a flexible deep fully convolutional net, which takes an arbitrary number of images of a stationary object captured under different illumination directions with a fixed digital camera as input to predict a normal map of the object in a fast feed-forward pass. This learning based method is flexible regarding the requirement of knowledge of light directions (instead of using an identical set of lights at train and test time).

After being extensively trained on synthetic data, it is claimed to perform well on real datasets as well. The authors claim that their method outperforms existing methods for calibrated PS and promising results for uncalibrated PS in 2018 when it was published.

3. Method and Experiment

The pipeline comprises of three parts - a shared-weight feature extractor, an order-agnostic fusion layer, and a normal regression network. The proposed method fuses spatial feature maps with max-pooling layers along the photometric dimension to infer the normal map. It is trained on two synthetic datasets - Blobby shape and Sculpture shape and the MERL BRDF dataset. To handle multiple number of inputs, a muti-branch Siamese network is introduced supervised by Cosine similarity loss. Mean angular error (MAE) in degree is used to measure the accuracy of estimated normal maps.

Several experiments were carried out. They claim max-pooling to have high correlation with response of the surface to a certain light direction, which has been experimented on. Effects of input number and training data are studied. The results were better when number of inputs during training was close to that during testing. When trained on more complex and realistic training data, the performance improved significantly. The robustness of PS-FCN is tested for regions with cast shadows. The model was tested on benchmark datasets like DiLiGenT showing that the average MAE results of [1] are an improvement from previous techniques. The uncalibrated photometric stereo extension PS-FCN called UPS-FCN is also quantitatively compared against existing techniques. The robustness of model is tested on real datasets qualitatively.

4. Critical Analysis

4.1. Significance of the paper's contributions

Outlier rejection methods require huge image dataset and are unable to handle objects with dense non-Lambertian surfaces. Sophisticated reflectance models require solving difficult optimization problems and could only handle materials of limited classes. Exemplar based methods requiring light calibration and longer processing time. DSPN, a learning based method requires pre-defined set of light direction for both training and testing.

Hence, this paper [1] is an improvement from the previous methods. Several existing techniques like pooling layer are used in a novel manner making the contributions non-trivial and incremental. It is the first paper that does not require the usage of light direction for predicting N.

4.2. Validity of the authors' main claims

The experiments and results show that the model outperforms the current techniques as stated. Although, most of the quantitative and qualitative comparisons of PS-FCN and UPS-FCN were made against techniques which are very old. Hence, comparison with more recent techniques is required to solidify their claim. The qualitative comparisons prove the effectiveness of the method. The use of pooling layer, performance on real data, and other claims are justified through experiments.

4.3. Limitation and weaknesses

PS-FCN performes better on materials with a light color than those with a dark color. This might be due to the image intensities of objects with a dark color are mostly very small hence not being aggregated as the most salient features for normal inference by max-pooling. To extract more useful features of objects with dark materials, average pooling could be used.

Another limitation is that PS-FCN requires accurate lightings as input. The NENet [2] is trained with discretized lightings estimated by the LCNet and shows a more robust behavior over noise in the lightings is a possible solution.

Calibration is a tedious process. Though UPS-FCN is introduced, it can be further improved [3] [4].

4.4. Extension and future work

The pipeline for uncalibrated dataset could be further explained and can be compared with existing methods for uncalibrated data. An extension to handle surfaces with SVBRDFs could be made. The model can be examined for images with non-uniform lighting and for objects in the wild as an extension.

I would like to extend this paper by introducing noise and more complex structures in dataset to make it more robust [2]. It can be used to perform a 3D scan of objects for e-commerce markets.

4.5. Is the paper stimulating or inspiring?

This paper is an interesting read as the methodology outperforms previous calibrated photometric stereo methods. It can be easily extended for uncalibrated photometric stereo making it flexible. The mean MAE of UPS-FCN is better than existing methods. The several experiments on synthetic as well as diverse real datasets indicates its effectiveness. Visualization of fused feature map provided insights to what is encoded by model. Although with more appropriate structuring of paper, it could be even more simulating as in [3].

4.6. Conclusion and personal reflection

To conclude, this paper presents a novel state of the art technique to handle calibrated and uncalibrated non-lambertian photometric stereo, which does not require pre-defined set of light directions during training and testing. It can handle input images in order agnostic manner unlike its predecessors.

If I were to work on this problem, I would make use of inherent spacial and photometric context in the image data [5].

References

[1] Guanying Chen, Kai Han, and Kwan-Yee K. Wong. *PS-FCN: A Flexible Learning Framework for Photometric Stereo*. In Proc. IEEE European Conf. on Comput. Vis. (ECCV), 2018. 1, 2

[2] Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita, and Kwan-Yee K. Wong. *Self-calibrating deep photometric stereo networks*. In Proc. IEEE/CVF Conf. on Comput. Vis. Pattern Recognit. (CVPR), 2019. 1, 2

[3] Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita, and Kwan-Yee K. Wong. *What Is Learned in Deep Uncalibrated Photometric Stereo?*. In Proc. IEEE European Conference on Computer Vision (ECCV), 2020. 2

[4] Fotios Logothetis, Ignas Budvytis, Roberto Mecca1, and Roberto Cipolla. *PX-NET: Simple and Efficient Pixel-Wise Training of Photometric Stereo Networks* In CoRR, volume abs 2008.04933, 2020. 1, 2

[5] David Honzátko, Engin Türetken, Pascal Fua, and L. Andrea Dunbar. *Leveraging Spatial and Photometric Context for Calibrated Non-Lambertian Photometric Stereo* In CoRR, volumne abs 2103.12106, 2021. 1, 2

[6] Berk Kaya, Suryansh Kumar, Carlos Oliveira, Vittorio Ferrari, and Luc Van Gool. *Uncalibrated Neural Inverse Rendering for Photometric Stereo of General Surfaces*. In Proc. IEEE/CVF Conf. on Comput. Vis. Pattern Recognit. (CVPR), 2021. 1

[7] *Photometric Stereo* Accessed on: Sept. 16, 2021. [Online]. Available: https://www.sciencedirect.com/topics /computer-science/photometric-stereo 1