

Robust Shape Recovery using a Directional Ring Difference Filter

Khurram Ashfaq, Muhammad Tariq Mahmood*

Future Convergence Engineering, School of Computer Science and Engineering, Korea
University of Technology and Education, Cheonan, Korea.
khurram@koreatech.ac.kr, tariq@koreatech.ac.kr*

Abstract

In this paper we propose a Focus Measure (FM) named Directional Ring Difference Filter (DRDF) that is robust to noise in reconstructing a depth map for Shape from Focus (SFF). To estimate the shape of the object, our method extracts multiple vectors in different directions from state-of-the-art Ring Difference Filter and use them as separate filters to construct the shape of the object. We conducted experiments on synthetic datasets that demonstrated superiority of our method in terms of both accuracy and noise handling capabilities when compared to RDF.

I. Introduction

3D shape reconstruction from 2D images is a fundamental task in computer vision that has garnered significant attention from researchers due to its numerous applications in fields such as robotics and computer graphics. Among the various techniques such as the use of light field [1] or stereo systems that have been proposed to tackle this problem, shape from focus (SFF) [2] has demonstrated particularly promising results.

SFF is an approach to extract the depth map of the scene from a stack of images that have been captured using different focus settings of the camera. The core idea behind SFF is that the shape of an object can be inferred from an image of that object that is in focus, based on the thin-lens law, which describes the relationship between the shape of an object or the position of the object relative to the camera, and the focus of the camera. In its pipeline, the first step involves generate a focal stack of the scene by capturing a series of images from the same viewpoint while adjusting the focus of the camera. Next, a focus measure operator (FM) such as [3],[4] is applied to each image in the stack to determine the sharpness of each pixel, resulting in a focus volume (FV) that tells the focus information for each pixel. Finally, the depth map is extracted from the FV using a "winner takes it all" approach, where the image in the stack with the highest focus value for each pixel is used as a representative of that pixel in the depth map.

SFF has several attractive features, including ease of implementation, compact size, and low cost. Additionally, the growing interest in depth sensing for mobile devices has made SFF an appealing choice for platforms which often have limited computational resources and power constraints [5].

II. Method

Let the input image sequence be represented by $I_z(x, y)_c$ where x , y and z denote the indices for width, height and

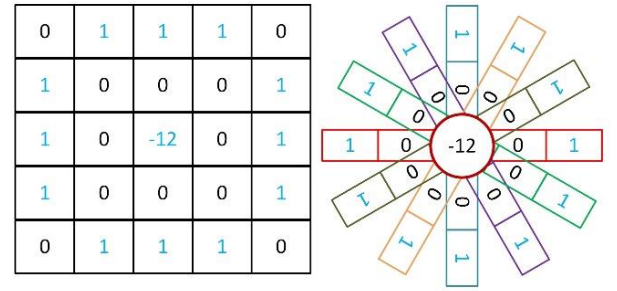


Figure 1: An RDF and its equivalent DRDF

image number respectively, and $c \in \{r, g, b\}$ denotes the color channel of a vector-valued image. The ranges of x , y and z are $(1, \dots, X)$, $(1, \dots, Y)$ and $(1, \dots, Z)$, respectively, such that there are Z vector valued images each of size $X \times Y$ pixels and every image has three color channels. To compute the focus volume for each pixel, we convolve the images in the sequence with directional filters using the following equation:

$$F_z(x, y) = \sum_c \sum_i |I_z(x, y)_c \otimes h_i| \quad (1)$$

where h_i denotes the directional kernel and \otimes represents the convolution operator. Finally, the depth map is computed as:

$$d(x, y) = \operatorname{argmax}_z (F_z(x, y)) \quad (2)$$

Our method Directional Ring Difference Filter (DRDF) is an improvement on the state-of-the-art Ring Difference Filter (RDF) [6]. Though RDF has shown high quality results in computing depth maps, it suffers from response cancellation problem (RCP). The RCP occurs when the response of an image to a two-dimensional Laplacian filter ($|\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}|$), is calculated, and the responses in opposite directions cancel each other out, reducing the effectiveness of the focus measure. To overcome this problem, a modified Laplacian focus measure was proposed ($|\frac{\partial^2}{\partial x^2}| + |\frac{\partial^2}{\partial y^2}|$) in which the responses from the x and y directions are calculated

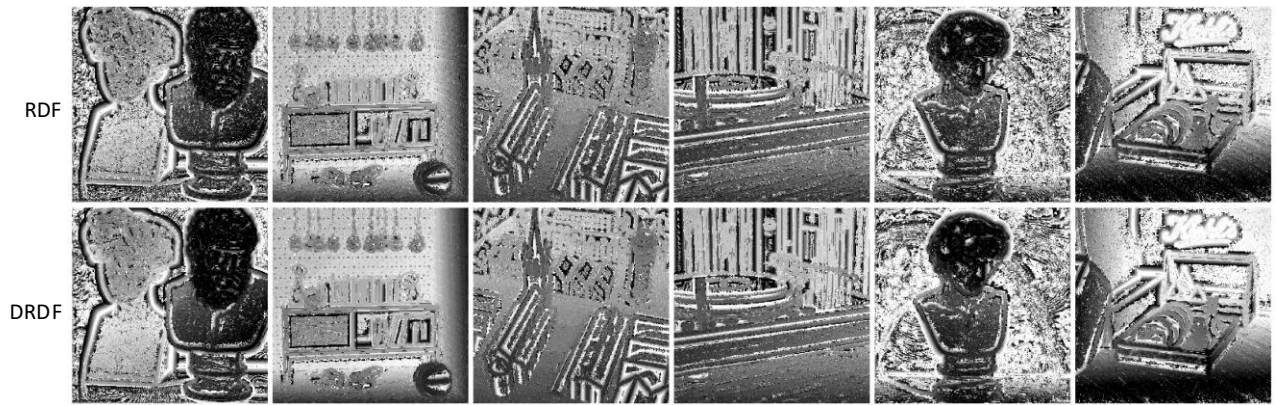


Figure 2: Depth maps of synthetic datasets with the first row representing RDF and second row representing DRDF

separately and the focus measure is taken as the sum of these responses [2]. Similarly, the RDF is spread in all directions, which can lead to the RCP significantly affecting the focus measure. To address this issue, our proposed DRDF takes the different vectors of the RDF in different directions and uses them as separate filters to be convolved with the image stack. By breaking the 2D RDF into different 1D kernels in n directions, our method is able to avoid the RCP and achieve improved accuracy and robustness in computing depth maps. A simple RDF of dimensions 5×5 and its equivalent DRDF at angles 0° , 30° , 60° , 90° , 120° and 150° is shown in (Fig. 1).

III. Results

In this study, we aimed to verify the effectiveness of our proposed method for generating depth maps of a scene. To do this, we conducted a detailed analysis of the performance of RDF and DRDF on 14 synthetic datasets, each containing 30 images. Our analysis included both quantitative and qualitative evaluations. As shown in (Fig. 2), we generated depth maps using both RDF and DRDF, with RDF having a size of 5×5 and its equivalent DRDF. In analyzing the results, we observed that both RDF and DRDF performed well on datasets with relatively plain backgrounds or with minimal background variations. However, DRDF demonstrated a clear advantage over RDF in terms of producing depth maps with less noise.

In addition to producing lower levels of noise, DRDF also outperformed RDF in all datasets in terms of quantitative measures such as root mean square and correlation with ground truth as reference. These results indicate that our proposed method, DRDF, is more effective at generating accurate depth maps and is more robust to noise compared to the state-of-the-art method, RDF.

IV. Conclusion

This paper proposes a new focus measure that can generate more accurate depth maps of a scene. The RCP that was deteriorating the quality of results in state-of-the-art RDF was countered by DRDF that extracts multiple filters in different directions of the same RDF to give more realistic depth map that is robust to noise. Experiments on synthetic datasets showed that DRDF gave better results which were close to the ground truth as compared to RDF.

ACKNOWLEDGMENT

This work is supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT: Ministry of Science and ICT) (2022R1F1A1071452).

REFERENCES

- [1] Raytrix. (2012). 3D Light Field Camera Technology. [Online]. Available: <https://www.raytrix.de/>
- [2] Nayar, S. K., & Nakagawa, Y. (1994). Shape from focus. *IEEE Transactions on Pattern analysis and machine intelligence*, 16(8), 824–831.
- [3] Yang, G., & Nelson, B. J. (2003, October). Wavelet-based autofocusing and unsupervised segmentation of microscopic images. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)* (Cat. No. 03CH37453) (Vol. 3, pp. 2143–2148). IEEE.
- [4] Sakurikar, P., & Narayanan, P. J. (2017). Composite focus measure for high quality depth maps. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 1614–1622).
- [5] Suwajanakorn, S., Hernandez, C., & Seitz, S. M. (2015). Depth from focus with your mobile phone. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3497–3506).
- [6] Jeon, H. G., Surh, J., Im, S., & Kweon, I. S. (2019). Ring difference filter for fast and noise robust depth from focus. *IEEE Transactions on Image Processing*, 29, 1045–1060.