Focus stacking by multi-viewpoint focus bracketing

Yucheng Qiu
Shibaura Inst. Tech.

Daisuke Inagaki
Shibaura Inst. Tech.

Kenji Kohiyama
Keio Univ.

Hiroya Tanaka
Keio Univ.

Takashi Ijiri
Shibaura Inst. Tech.

ABSTRACT

We present an approach to obtain high-quality focus-stacking images. The key idea is to integrate the multi-view structure-from-motion (SfM) algorithm with the focus-stacking process; we carry out focus-bracketing shooting at multiple viewpoints, generate depth maps for all viewpoints by using the SfM algorithm, and compute focus stacking using the depth maps and local sharpness. By using the depth-maps, we successfully achieve focus-stacking results with less artifacts around object boundaries and without halo-artifacts, which was difficult to avoid by using the previous sharpest pixel and pyramid approaches. To illustrate the feasibility of our approach, we performed focus stacking of small objects such as insects and flowers.

CCS CONCEPTS

• **Computing methodologies → Image processing**.

KEYWORDS

Focus stacking, focus bracketing, structure from motion.

ACM Reference Format:
Yucheng Qiu, Daisuke Inagaki, Kenji Kohiyama, Hiroya Tanaka, and Takashi Ijiri. 2019. Focus stacking by multi-viewpoint focus bracketing. In *Proceedings of SA '19 Posters*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3355056.3364592

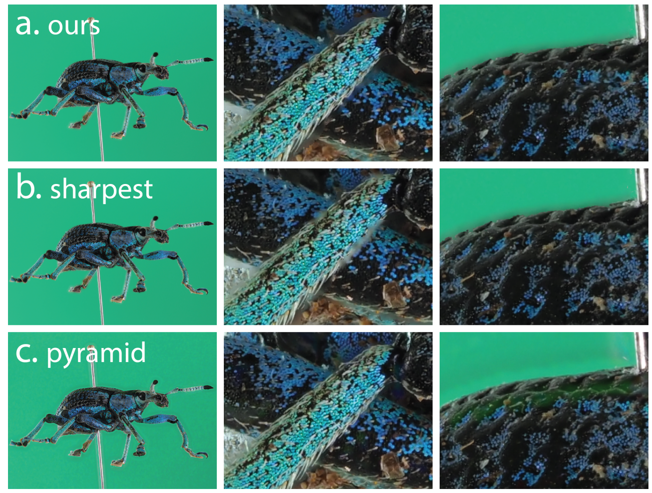

Figure 1: Focus-stacking images generated with (a) our approach, (b) sharpest pixel, and (c) pyramid. In (a), we captured multi-focus-sets at 9 viewpoints.

1 INTRODUCTION

Focus stacking is an image-processing technique that integrates multiple photographs taken with different focus distances into one photograph with a greater depth of field. Some recently available consumer-grade cameras support automatic *focus bracketing*, that takes multiple photographs with varying focuses in a short time, and subsequent focus stacking. Focus stacking is especially useful when taking photographs of small objects, such as insects and flowers, because such objects often require the use of a macro lens with a shallow depth of field.

Existing focus-stacking approaches can be roughly divided into two groups, sharpest pixel and pyramid [Wang and Chang 2011] approaches. Given a set of photographs taken at the same viewpoint and with different focus distances, which we call a multi-focus-set, both approaches first deform all photographs so that the content is aligned pixel-by-pixel. After the alignment, the sharpest-pixel approach selects the source photograph for each pixel of the output image. This approach uses local sharpness, e.g., contrast, as the selection metric. However, this approach often fails to select correct photograph, causing artifacts around the object boundary. As in Figure 1b, the boundary of the insect leg is greatly blurred. Given a multi-focus-set, the pyramid approach computes Laplacian pyramids of all the input photographs, performs maximum intensity projection at each pyramid level to obtain a fused pyramid, then generates the final image by inverse Laplacian pyramid transform. However, this approach often causes halo-artifacts around strong edges. As in Figure 1c right, green color of the background is mixed to the foreground.

This study presents a novel focus-stacking approach that uses depth maps computed from multi-view images. We carry out focus-bracketing shooting at different viewpoints to obtain multiple multi-focus-sets. We then compute a depth map for each viewpoint by adopting the structure from motion (SfM) algorithm. We then select source photographs by considering both the depth maps and local sharpness. By using the depth maps computed using the SfM algorithm, our approach achieves high-quality focus-stacking results with less artifacts around object boundaries (Figure 1a). To illustrate the feasibility of our approach, we provide focus-stacking results for small targets, such as insects and flowers. We also argue that our approach is useful for 3D digitization.

2 OUR APPROACH

Our approach consists of the following four steps.

Step 1. Focus bracketing and alignment. Focus-bracketing shooting is carried out at different viewpoints to obtain multiple multi-focus-sets. After that, we deform the photographs in each set

SA '19 Posters, November 17-20, 2019, Brisbane, QLD, Australia

ACM ISBN 978-1-4503-6943-5/19/11.
https://doi.org/10.1145/3355056.3364592

to align the content. Although a multi-focus-set is taken from the same viewpoint, the alignment is necessary because the angle of view varies with the focus modification. In this study, we used a digital camera, Olympus OM-D E-M1 Mark II, with a macro lens, Olympus M.ZUIKO DIGITAL ED 60mm F2.8 Macro.

Step 2, Initial focus stacking. Focus stacking for each multi-focus-set is carried out independently by using the sharpest-pixel approach. We used the Michelson contrast in a 31×31 window as the sharpness metric. Note that the outputs of this step still contained artifacts, as shown in Figure 1b.

Step 3. Depth-map computation. The SfM algorithm is applied to the focus-stacking images generated in the previous step to obtain depth maps for all sets. We then perform dilation to fill small holes of the depth maps. We used the open source software COLMAP [Schonberger and Frahm 2016] for this process.

Step 4. Final focus stacking. Focus stacking is carried out using the depth maps. We linearly map the depth value into the index of photographs. We assume that the depth values corresponding to the first and last photographs are presented by the user. Note that this linear mapping is specific to our photographing condition. We obtain an index of photographs, i_d, for each pixel. We narrow the search target from the $(i_d - 1) - th$ to $(i_d + 1) - th$ photographs and find a source photograph with the highest Michelson contrast (i.e., sharpest) in a 3×3 window.

3 RESULTS AND FUTURE WORK

To illustrate the feasibility of our approach, we carried out focus stacking of different targets, such as insects (Figure 1a) and flowers (Figure 2 top). Figure 2 compares our approach and the sharpest pixel approach by showing the resulting photographs and ID maps of source photographs. By constructing depth maps from multi-view photographs and selecting pixels in focus by using the depth maps, our approach generated high-quality focus-staking images with smoother ID maps.

Our approach is also useful for 3D digitization. Similar to a previous study[Nguyen et al. 2014], we placed a sample on a rotation stage and took photographs from different viewpoints, i.e., 36 azimuth and 2 elevation angles. At each viewpoint, we carried out focus bracketing to obtain about 35 photographs. We next generated focus-stacking images with our approach. We then reconstructed a textured 3D model from the multiple focus-stacking images by using the commercial software PhotoScan. As in figure 3 and the supporting video, with our approach, it was possible to reconstruct 3D insect models with highly detailed textures.

Limitations and future work. The quality of our focus-stacking depends on the quality of the depth maps. It is difficult to adopt our approach to an object without enough surface textures, because the SfM algorithm often fails to generate accurate depth maps for such object. Our on-going future work is to develop a 3D digitization technique based on the presented focus-stacking approach. We reported our preliminary digitization results in Figure 3. We also plan to improve this approach and conduct a detailed evaluation.

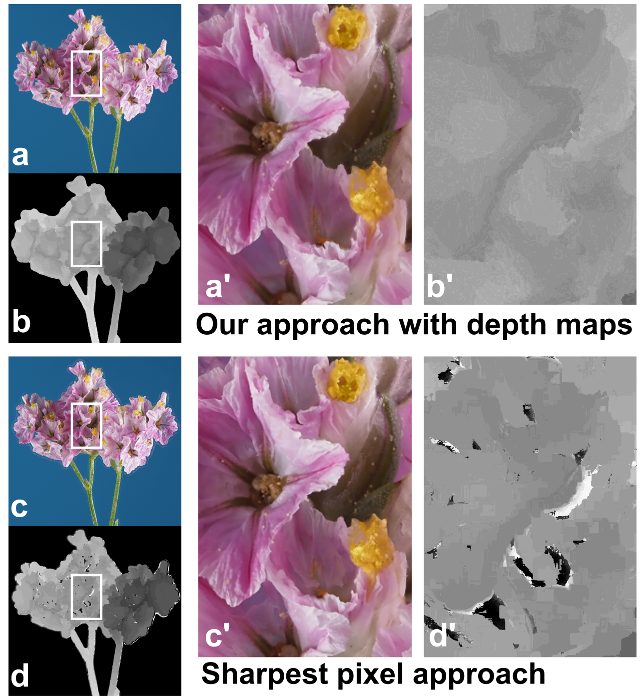

Figure 2: Focus stacking images generated with our approach (top) and the sharpest pixel approach (bottom). We show resulting photographs (a, c), ID maps (b, d) and their enlarged views (a', b', c', d'). In (a), we captured multi-focus-sets at 7 viewpoints.

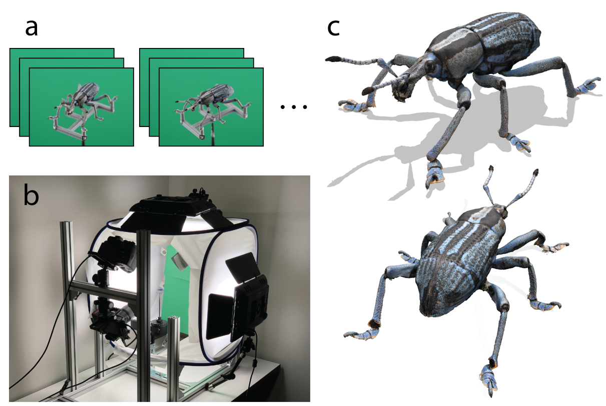

Figure 3: We captured multiple multi-focus-sets (a) by using a rotation stage (b). We then compute focus stacking images and use them to reconstruct a textured 3D model (c).

ACKNOWLEDGMENTS

This work was supported in part by Grant-in-Aid for Scientific Research in Japan (17H01848).

REFERENCES

Chuong V Nguyen, David R Lovell, Matt Adcock, and John La Salle. 2014. Capturing natural-colour 3D models of insects for species discovery and diagnostics. *PloS one* 9, 4 (2014), e94346.

Johannes L Schonberger and Jan-Michael Frahm. 2016. Structure-from-motion revisited. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4104–4113.

Wencheng Wang and Faliang Chang. 2011. A Multi-focus Image Fusion Method Based on Laplacian Pyramid. *JCP* 6, 12 (2011), 2559–2566.