RESEARCH ARTICLE | JANUARY 15 2019

Achieving 3D imaging through focus stacking FREE

J. Wlodek; K. J. Gofron; Y. Q. Cai

AIP Conf. Proc. 2054, 050001 (2019)
https://doi.org/10.1063/1.5084619

Achieving 3D Imaging through Focus Stacking

J. Wlodek[1, 2], K.J. Gofron[1, a)] and Y.Q. Cai[1]

[1]*National Synchrotron Light Source II, Brookhaven National Laboratory, Upton, NY 11973, USA.*
[2]*Stony Brook University, Stony Brook, NY 11794, USA.*

[a)] Corresponding author: kgofron@bnl.gov

Abstract. Newly developed 3D reconstruction technology has many uses ranging from facial recognition, self-driving vehicles, and producing accurate models of samples. To utilize 3D reconstruction, a depth map of the sample must be generated from a series of images. This can be done using stereoscopic 3D, or in the case of a fixed camera (such as the 10ID On-Axis Microscope at NSLS-II used for this paper) through focus stacking. This paper will explain how focus stacking can be used to generate in-focus images at very shallow depth of fields, and merge these images to create depth maps and 3D models.

INTRODUCTION

3D reconstruction involves the use of two dimensional images to recreate depth information for a scene. It is often used in robotics and AI to allow for depth perception, and in the creation of high quality 3D models. Two widely used techniques for 3D reconstruction are stereoscopic 3D, also known as stereo vision, as well as photogrammetry. In stereo vision, two cameras are used, and the relative positions of pixels in the two images generated by the cameras provide the depth map. In photogrammetry, photos are taken of the object from all angles, and it is later painstakingly reconstructed from the images. In some use cases, however, neither method proves to be suitable. The best example of this is under a microscope. Most microscopes only have one camera that can only move on one axis for focus adjustment. As a result, stereo 3D, requiring 2 cameras, cannot be used, and photogrammetry, in addition to being time consuming, is also unusable because microscopes are required to remain stationary. Our proposed solution is to utilize a technique known as focus stacking [1], consisting of collecting a set of images of the sample at different focal depths, identifying the area of each image that is in focus, and from there calculating the depth at which each feature lies. The images are then merged into a reconstructed, in-focus image. The depth map of the image is constructed simultaneously, and both the merged image and depth map are used to generate a 3D model.

FOCUS STACKING

The first portion of our project focused on image stacking, which was the preliminary step to the 3D reconstruction we would do later. We began by experimenting with different possible algorithms that would evaluate the relative sharpness of a given pixel. The purpose of this was to identify areas of the image that are perceived as in-focus, or sharp, v.s. areas that were blurry. Our initial tests used gradients to rank pixel sharpness, with high gradients representing a sharp feature, while gradual gradients representing out-of-focus regions, but ultimately, we settled on running a Gaussian blur matrix followed by a Laplacian on the images. Convolving a Gaussian blur matrix on the image blurred it, making out of focus regions even less sharp, while making the in-focus regions stand out. After convolving with the Gaussian kernel, a Laplacian kernel was used as a form of edge detection. Because we had blurred the image in the previous step, in-focus portions of the image appeared as edges, and were detected by the Laplacian. The resulting image after being convolved with the Laplacian gave us a

Proceedings of the 13th International Conference on Synchrotron Radiation Instrumentation – SRI2018
AIP Conf. Proc. 2054, 050001-1–050001-6; https://doi.org/10.1063/1.5084619
Published by AIP Publishing. 978-0-7354-1782-3/$30.00

“sharpness ranking” of each pixel in each image. From there, we iterated over all the pixels in the images, selecting those with the highest value for the merged image. Results for this technique for some sample sets proved to be adequate, but in some cases resulted in poor results. After some further experimenting, we found that the primary issue was the fact that changing focal depth caused the subject of the images in several stacks to appear to warp, move, and zoom. As a result, the stacking did not have a 1:1 pixel match in all cases, which in turn caused inaccurate results. To combat this, we decided that image alignment was necessary, and developed an alignment method that utilizes a Euclidean Transform, a function in OpenCV [2]. Initially, the alignment function used Oriented Fast and Rotated Brief (ORB) feature detection and Brute Force matching, which is the traditional image alignment method. Unfortunately, this proved to be ineffective, as we had not accounted for a key issue. Due to the fact that all of the images were taken at a very shallow depth of focus (intentionally), the key points detected in one image might not be key points at all in another, but rather simply a part of the blurred area. As a result, merging images using this technique provided very poor results. Our next attempt at image alignment used a Euclidean transform, which uses all of the pixels in an image to find a mapping between two images, and finally the merging process produced impressive results. (see Figure 1)

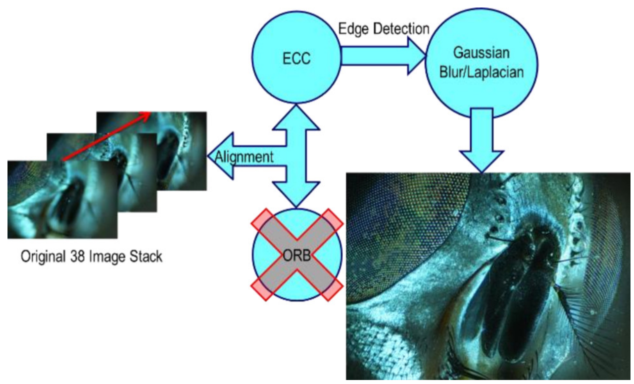

FIGURE 1. Focus stacking process. Horse Fly-5X microscope.

DEPTH MAP GENERATION

Once focus stacking was finished, the next step in the process was to use it to create a depth map. The idea behind extracting depth from a focus stack is intuitive. We used the following equation, in which F_{min} and F_{max} are the focal depths of the camera at the closest and farthest picture respectively, I is the current image in the stack, and N is the number of images in the stack.

$$Depth = I * \frac{(F_{max} - F_{min})}{N} \tag{1}$$

This formula gave us the relative depth of each image with the furthest image as a baseline with a depth of 0, and other images with depths measured in scientific units. A traditional depth map, however, utilizes a scale from 0 to 255 to store depth information (black to white), where a value of 0, or black, represents the furthest point in the image, and a 255 value represents the closest to the viewer. As a result, a conversion rate from scientific units was necessary, and was created with a proportion involving the max focal depth over 255 being set equal to the image focal depth over X. Solving for X gave us the depth in terms of this scale. Once the formulas had been made, the next step was to add depth generation to the focus stacking algorithm already in place. To do this, we simply added a second step in the merging process described previously. Once the “sharpness values” had been assigned, to each pixel, and the pixels were being placed in the merged image, each pixel’s original image was found in the array, the depth of said image was calculated, and the appropriate value was placed into the grayscale depth map. This way, the function that previously returned just the merged image, now returned the merged image along with the depth map of said image. This technique posed some issues, however. Firstly, while some image noise was acceptable in a color image (if points are taken from the wrong image but in the same area the color will be similar and thus not very noticeable), it was very distracting when in a depth map. A single pixel out of place could lead to a random out of place spike that protrudes from the model. As a result, a necessary part of generating the depth map was aggressive denoising and smoothing in order to get optimal results, at the cost of some fine detail. An improved algorithm would consider the damage caused by noise in the depth maps and would perhaps attempt to use bilinear or trilinear filtering to eliminate the “texture” caused by the noise, or perhaps a more robust focus stacking method could be implemented. Another problem we faced when generating the depth map, was that occasionally, the depth map generated would include unnecessary artefacts on the outermost portions of the image. This issue was caused by the alignment process used by the focus stacking algorithm prior to the merging process. If an image must be shifted or rotated in order to align it with the other images in the stack, the result would cut portions of said image out, and fill the voids with black pixels. As a result, when running edge detection, the lines between these black pixels and the rest of the image were very pronounced, and hence appeared in both the merged image and its depth map. The only solution to this problem was to crop portions of the images out, removing the corrupted pixels. Provided the sample in the stack does not move excessively between images, alignment (and as a result cropping), may be unnecessary, however, in almost all our tests, aligned image merging provided superior results.

STACK COLLECTION AUTOMATION

For the On-Axis Microscope at the IXS 10-ID [3] beamline we demonstrated an implementation of the open source EPICS control system library to automate focus stack collection. The code requires the user to first find a start and end point for where the sample is in focus in order to avoid unnecessary photos in the stack. This is necessary because too many photos in the stack will lead to unnecessary depth map noise, as in-focus portions of images will overlap. Too few images in the stack, on the other hand, will lead to information being lost, as portions of the final merged image will remain blurry. Additionally, for the EPICS automation, the user is required to specify a number of images. The software will then calculate the distance the motor will move the camera by subdividing the total distance between the outermost images by the number of images selected. As before, the chosen number cannot be too high, as overlapping in focus areas can cause problems for generating an accurate depth map, while a number too low will miss certain regions, which will lead to parts of the image remaining out of focus. The merged image in Figure 2, has missed regions visible as superstructure bands of the horse fly right eye. Once the variables are inputted, the code splits the distance between the start and end point into the specified number of subdivisions, automatically moves the motor to each of these points and takes a picture. The process was further accelerated when instead of moving the motor to each position, stopping it, and then collecting the image, a running collection process was developed. The motor would now move continuously, and a trigger was sent to the camera at each calculated point.

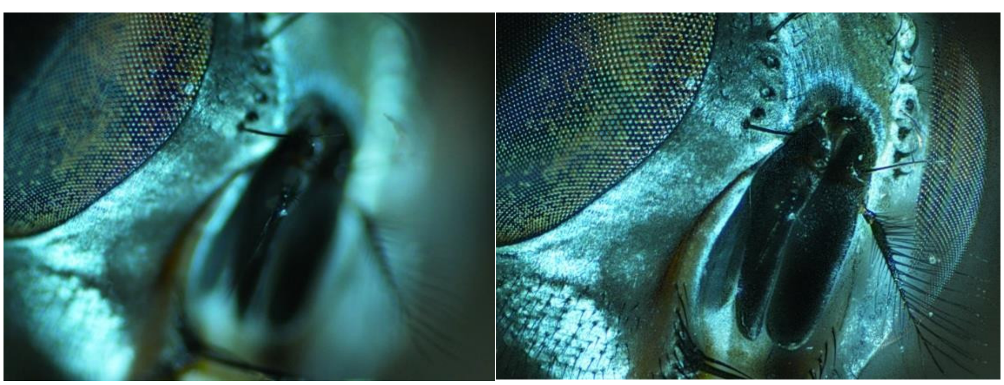

FIGURE 2. Single image from 5X microscope (left panel), and one obtained from Image Stack (right panel).

3D RECONSTRUCTION

After the automated collection process has collected the stack, the images are first passed to the image stacking and depth map generation programs. Once these have completed, a merged image and its corresponding depth map are available. These can be used by scientists to better understand their sample in 3D space, or to simply acquire a better view of the sample. These benefits can be further enhanced through 3D reconstruction using both the merged image and its depth map. Rather than working with the time-consuming process of comparing the depth map with the image, a model can be generated, and rotated freely in three dimensions, despite the stack being originally collected by just one camera. The reconstruction process used in our project was similar to that discussed by Michael Moeller in his paper on "Variational Depth from Focus Reconstruction" [1]. Using a custom OpenGL[4] pipeline implemented with the python wrapper for OpenGL, each pixel in the depth map is rendered as a GL Quad. The coordinates of the Quad in three-dimensional space are generated by converting the relative units in the depth map back into metric units. The Quads are then colored based on the color of their corresponding pixel in the merged image. The model is then rendered onto the screen, and rotated about its center in each direction. This rotation is recorded and saved as a video, so that examination of the model could be accelerated in the future. Both this model and the depth map created by a focus stacking approach for the On-Axis Microscope at NSLS-II enable a unique perspective on samples being studied, allowing for the use of depth information in experiments. They also allow for new and exciting ways for demonstrating samples and scientific findings.

SAMPLE FOCUS STACKS

38 Image Stack of a Horse Fly under a 5x Microscope: We tested our image stacking applications on a Horse Fly positioned under a 5x magnification microscope. We collected a stack of 38 images, which were combined to form one merged image. The difference we see in Figure 2 between one of the images taken from the stack and the merged image is striking, and effectively shows the benefit that such a technique could provide if the user wishes to have a detailed single image of his or her subject.

18 Image Stack of a Mineral under a 5x Microscope: To test the system on a different kind of sample, we examined a mineral under the same 5x microscope as the fly, and once again utilized the image stacking techniques discussed earlier. Additionally, we deployed the 3D imaging application to this stack as well, and it provided some great results. Figure 3 shows the merged image and the depth map obtained from the mineral, and additionally a video of the mineral's three-dimensional OpenGL focus-stacked model is available at: https://www.youtube.com/watch?v=FdpdAdoirwA. [5, 6].

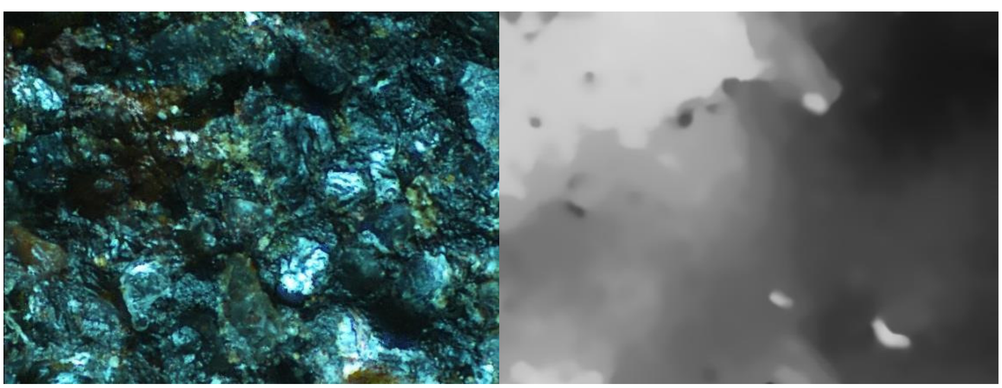

FIGURE 3. Image of a mineral (left panel) and its depth map (right panel).

15 Image Stack of Beetle under On-Axis Microscope: An image stack of 15 images was collected at the 10ID On-Axis Microscope of a beetle. The image stacked results from the images are shown in Figure 4 (left panel). The right panel of Figure 4 shows the depth map obtained from images. Using the depth map a 3D video of the beetle was produced [5, 6].

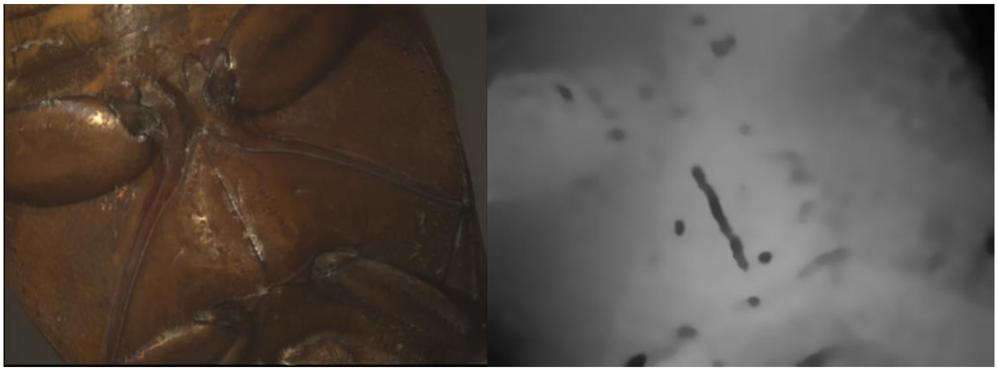

FIGURE 4. Image of a beetle (left panel) and depth map (right panel).

FUTURE IMPROVEMENTS

There are several possible future improvements that could be implemented to improve both the performance and accuracy of the image stacking system. The first of which would be to add additional weightings to the merging process for several factors aside from sharpness. For example, if creating a merged image of a focus stack of a car, a reflection of clouds in the windshield may move from image to image, so instead of merging pixel by pixel, the entire windshield should be lifted from the image. Ultimately movement and lighting are an inherent weakness in the focus stacking approach, given its one camera system, but measurements such as this can help mitigate such weaknesses. Additionally, gamma correction could be implemented prior to merging to ensure that all the images in the stack have the same coloring. Both improvements would also benefit the 3D reconstruction software, as a cleaner model would result in a cleaner depth map, and hence a more effective model. The previously described improvements focus on improving focus stacking accuracy, but there are also improvements that could be made to benefit performance. The primary bottleneck of the software is rendering the three-dimensional model in real time. As a result, a GPU accelerated system would greatly increase performance. Additionally, the use of Vertex Buffer Objects to improve OpenGL performance, and the use of triangles over quads for the model could also be worthwhile. Aside from the rendering, multithreading could be added to increase performance in the most demanding areas.

CONCLUSION

This paper demonstrates the process by which a traditional weakness of fixed position microscopes (shallow depth of focus) can be turned into a strength using focus stacking and three-dimensional reconstruction. In applications where traditional reconstruction methods fail because of the nature of fixed cameras, this focus-stacking approach allows for both greater flexibility and simpler automation.

ACKNOWLEDGEMENTS

We acknowledge Alexei Soares' contribution to the project for discussion and image collection of the biological crystals used as samples in certain tests. We also acknowledge the 10ID beamline team for their cooperation and support for this project. This project was supported in part by the Brookhaven National Laboratory (BNL), the BNL Photon Sciences department under the BNL Supplemental Undergraduate Research Program (SURP), and the Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory

Internships Program (SULI). The work at the National Synchrotron Light Source-II, Brookhaven National Laboratory, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-SC0012704.

APPENDIX

Due to limitations of the printed version, all videos obtained from Image Stacking are provided on the www.youtube.com channel: 3D Microscope. The following are URLs to the two 3D model videos on the channel. https://www.youtube.com/channel/UC-SfBpwDIiuw41_r0qqYkZQ?view_as=subscriber.
https://www.youtube.com/watch?v=FdpdAdoirwA, https://www.youtube.com/watch?v=UFx2EDouO-k.

REFERENCES

1. M. Moeller, M. Benning, C. Schönlieb, D. Cremers, "Variational Depth From Focus Reconstruction." *IEEE Trans. Image Process.*, vol. 24, no. 12, 2015, pp. 5369–5378.,doi:10.1109/tip.2015.2479469.
2. I. Culjak, D. Abram, T. Pribanic, H. Dzapo and M. Cifrek, "A brief introduction to OpenCV," *2012 Proc. 35th Int. Conv. MIPRO*, Opatija, 2012, pp. 1725-1730.
3. K. J. Gofron, J. Wlodek, Y. Q. Cai, "On-Axis 3D Microscope for X-ray Beamlines at NSLS-II", *Proc. 16th Int. Conf. on Accelerator and Large Experimental Control Systems*, Barcelona, 2017, pp. 2048-1053, doi:10.18429/JACoW-ICALEPCS2017-WEBPL02.
4. K. Group, "The Industry's Foundation for High Performance Graphics." OpenGL.org, www.opengl.org/.
5. https://www.youtube.com/watch?v=UFx2EDouO-k.
6. Beetle and a mineral 3D microscope video
https://www.youtube.com/watch?v=FdpdAdoirwA,
https://www.youtube.com/watch?v=UFx2EDouO-k.