Depth from Defocus via Active Quasi-random Point Projections

Avery Ma   University of Waterloo, ON, Canada
Francis Li   University of Waterloo, ON, Canada
Alexander Wong   University of Waterloo, ON, Canada

Abstract

Depth sensing has many practical applications in vision-related tasks. While many different depth measurement techniques exist and depth camera technologies are constantly being advanced, active depth sensing still rely on specialized hardware that are highly complex and costly. Motivated by this, we present a novel technique for inferring depth measurements via depth from defocus using active quasi-random point projection patterns. A quasi-random point projection pattern is projected onto the scene of interest, and each projection point in the image captured by a camera is analysed using a calibration model to estimate the depth at that point. The proposed method has a relatively simple setup, consisting of a camera and a projector, and enables depth inference from a single capture. Furthermore, the use of a quasi-random projection pattern can allow us to leverage compressive sensing theory to produce full depth maps in future applications. Experimental results show the proposed system has strong potential for enabling active depth sensing in a simple, efficient manner.

1 Introduction

Depth sensing is a fundamental low level computer vision task that allows for 3D geometric measurements of the world. Depth information is crucial for many applications including 3D reconstruction, action recognition, human-computer interfacing, etc. As such, depth cameras have received much attention both academically and in industry with constant advancements to depth camera technology.

Traditionally, active depth sensing techniques include methods such as projector-camera structured light systems, laser scanners, ToF cameras, and IR projected point patterns such as the Microsoft Kinect [1]. While recent efforts have lowered the cost of depth camera sensors, depth sensing still requires specialized complex hardware that are not compact. Motivated by these challenges, in this paper we present a system for inferring sparse depth measurements from a single camera capture by leveraging active quasi-random point projections and camera defocus. The proposed method of inferring sparse depth has a relatively simple setup, thus can potentially lead to very compact and low cost active depth sensing systems. The proposed method involves projecting a quasi-random point pattern onto the scene of interest. The quasi-random point pattern can for instance be generated by using a light source with a point pattern mask, resulting in much simpler and cheaper hardware configuration. The detected projected points on the imaged scene are analyzed to determine their point spread functions (PSF), which is then are used to estimate depth via a calibrated model.

2 Related Work

Depth from defocus (DFD) methods generally estimate depth by analyzing the difference in blurriness of two images captured at different focal lengths [2, 3, 4], with different methods using different filters for determining the measure of blur. A major drawback to such DFD methods is the unreliable detection of blur, especially in untextured areas of the image. This problem is mitigated in active depth sensing, where an optical projection is used to find correspondences for triangulation. A review of structured light patterns for depth measurement is provided by Salvi *et al.* [5]. Therefore, using active projection patterns does not depend on the objects in the scene and is also effective in untextured regions. On the other hand, active depth sensing systems suffer from occlusion and require complex hardware. As such, we are motivated in the proposed system to takes advantage of the benefits from both DFD and active depth sensing methods to design a system that has a simple setup yet reliable in the depth measurements.

The concept of using DFD using active projections have also been explored in literature. Pentland *et al.* [6] used evenly spaced line projections to determine depth from line spread. This simple method is able to create low resolution depth maps. Nayar *et al.* [7] used a dual sensor plane with optimized projection and camera

setup to produce a dense depth map and reduce front/back focal ambiguity. Ghita *et al.* [8] used a dense projected pattern with a tuned local operator for finding the relationship between blur and depth. Moreno *et al.* [9] used an evenly spaced point pattern with defocus to obtain an approximate depth map used for automatic image refocusing. These methods use a high density projection pattern which require either a projector or more specialized calibrated hardware, neither of which are required by the method presented here.

3 Proposed Method

The proposed system can be described as follows. A quasi-random point pattern is projected onto the scene, which is then captured by a camera. The camera's focus is fixed such that the degree of focus of each point in the quasi-random point pattern as it appears in the captured image is dependent on the depth of the surface. The relationship between depth and the point spread function (PSF) of the projected point pattern is characterized in a calibration model, and used to infer the depth at each point to produce sparse depth measurements. A one-time calibration step is required to learn the calibration model.

3.1 Calibration

The purpose of the calibration procedure is to characterize the relationship between PSF and depth in a calibration model. When out of focus, a projected point will appear blurred, with the degree of blurriness correlated with the depth of the scene at that point, which we can model using a PSF. The blur effect of a projected point is visualized in Figure 1. Here, the PSF of the projected dot as captured in the image can be approximately modelled by a 2-D Gaussian [10]:

$$h(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

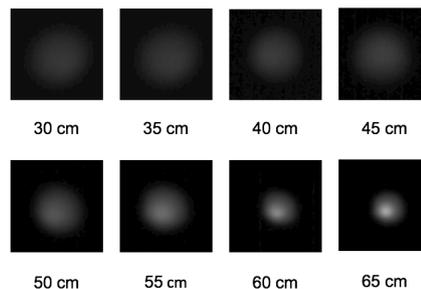where σ is the standard deviation of the PSF.

Fig. 1: Projected points on an object at various distances (30cm to 60cm) away from the setup, as captured by the camera.

Therefore, to learn the calibration model, a point pattern is projected onto various planes with known distances away from the projector-camera setup. The measured data points corresponding to standard deviation vs. depth can then be used to construct the calibration model characterizing the relationship between standard deviation of the PSF and the distance away from setup. To obtain a continuous curve for the calibration model, regression with a third order polynomial function is used to fit the data points. A visualization of the calibration model is shown in Figure 2.

3.2 Sparse Depth Estimation Pipeline

With the calibrated model, the proposed system can then be used to estimate sparse depth of the scene. To this end, the proposed depth recovery method can be divided into 4 main stages outlined in Figure 3 and described as follows.
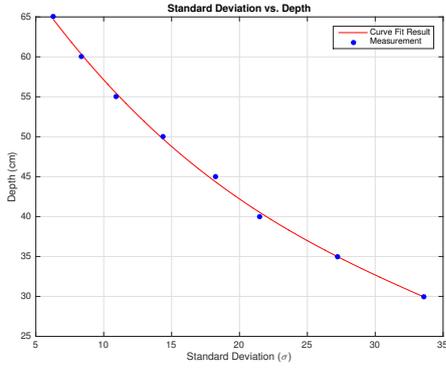
$$\mu = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}}{N} \tag{3}$$

Figure 5 shows a point spread function computed for an example projected point.

Recovery of Sparse Depth Map Based on the standard deviation of the PSF of a projected point, the calibration model can then be used to infer the depth corresponding to that projected point. By performing this on all projected points in the quasi-random point projection pattern, the sparse depth map can be obtained.

4 Experimental Setup

The main goal of this current realization of the proposed technique is to build a compact and portable system to obtain depth information of the scene. For this purpose, the scene is imaged using a Point Grey Chameleon3 USB Camera (resolution: 1280×1024) with a 30mm prime lens and the quasi-random point pattern is projected using a BENQ MH630 Digital Projector (resolution: 1440×900).

The experimental setup is described in Figure 6. The camera is mounted on a tripod, 40 cm behind the projector. The camera lens is in-line with the projection lens. The focal plane of both projector and the camera are on a plane 80 cm away from the projector.

5 Results

The test scene was the edge of a rectangular box (two slanted planar surfaces 90°of each other), and the quasi-random point density is approximately 0.05% of the projector resolution.

Active Quasi-random Point Projection: A quasi-random point pattern is projected onto the same. In the current system, a Poisson Disc Sampling method was utilized to generate the quasi-random point pattern such that the random points are tightly packed together, but no closer than a specified minimum distance. Compared to other random sampling methods, Poisson-Disc method significantly reduces the chances of having overlaps between blurred projected points, which would result in erroneous depth recovery.

Point Localization: After the projected point pattern has been captured by the camera, Otsu's method is used to obtain a binary map consisting of regions corresponding to the projected points [11]. The centroid of each region is computed and the distance between the closest pair of centroid is used to define the regions of interest of the projected points. The identified projected points and their corresponding regions are shown in Figure 4.

Point Spread Function Estimation of Projected Points: Based on the detected regions of interest of the projected points, the standard deviation of the point spread function of each projected point is estimated by using an intensity-weighted approach, where the number of samples at each location is approximated by multiplying the pixel intensity by a factor of 10,000. Based on these generated sample populations, the sample mean and sample standard deviation can be computed as follows:

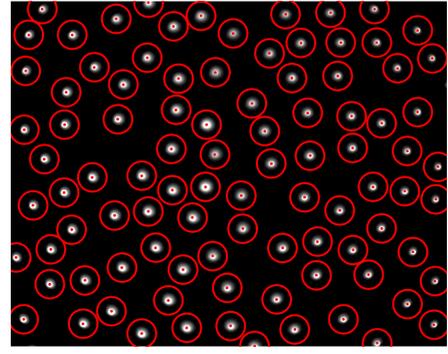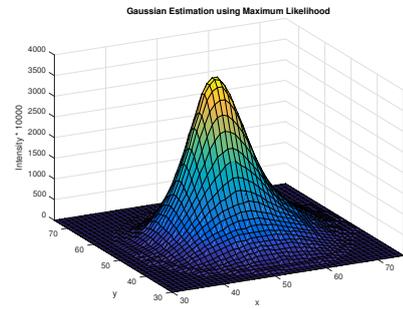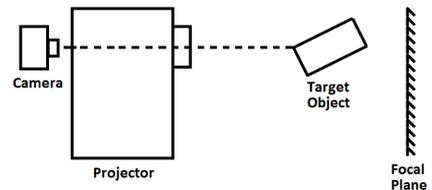$$\sigma = \sqrt{\frac{\sum_{i=1}^{m} \sum_{j=1}^{n} (x_{ij} - \mu)^2}{N - 1}} \tag{2}$$
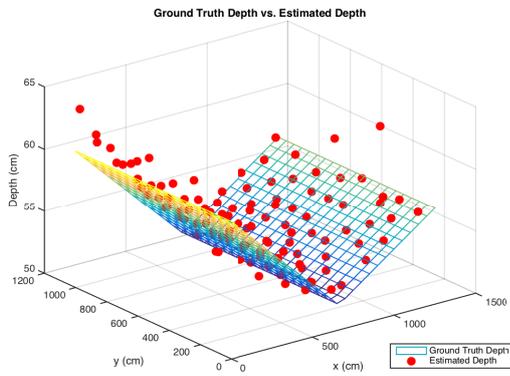
Fig. 7: Sparse depth map of the scene with ground truth surfaces

Ground truth surfaces were constructed from known measurement points. Figure 7 shows the depth recovery of the object compared to the ground truth surfaces. To evaluate the performance of the pipeline, the correlation coefficient was computed between the estimated depth and the ground truth depth at corresponding locations.
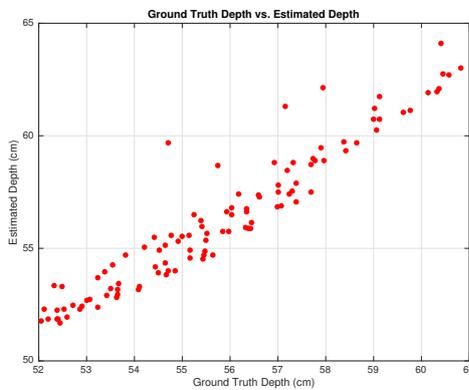

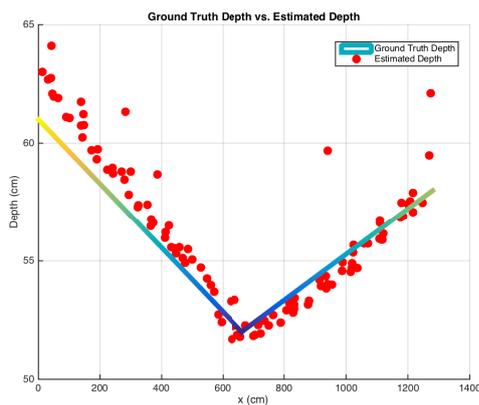Fig. 8: Ground truth depth vs. estimated depth

Fig. 9: Top view of the reconstructed sparse depth map

The correlation coefficient was computed to be **0.9533**, which demonstrates that a strong positive correlation between the ground truth depth and the estimated depth. This strong correlation between the ground truth depth map and the estimated depth can also be observed in the sample point visualization shown in Figure 8.

Figure 9 shows the top view of the sparse depth recovery of the test scene. It is worth mentioning that when the projected points appear close to the edge of the captured image, part of the projected points end up being cut off from the image, thus leading to inaccurate estimates of the standard deviation of the PSF of these cut-off projected points. This is clearly observed at the left and right edges of the reconstructed sparse map of the scene, where

the points stray far from the ground truth depth locations. Nevertheless, it can be observed that the basic geometry of the target object is well captured using the proposed technique.

6 Conclusion

This work presented a sparse depth measurement system based on active quasi-random point projection and depth from defocus. Preliminary results demonstrated promising results for inferring depth data of the scene. The main advantage of the proposed method is its simplicity in hardware and computation, requiring merely a camera and projected dot pattern.

There are many additional future directions that can be explored. Different quasi-random patterns other than Poisson Disc patterns can be explored to determine if improvements in depth accuracy can be achieved. As well, different point spread function models can be examined to account for lighting and object texture. In addition, the density of the active quasi-random projection patterns can be dynamically adjusted to generate denser measurements in areas of more detailed structure.

Acknowledgments

References

[1] Microsoft. (2016) Kinect hardware (online). https://developer.microsoft.com/en-us/windows/kinect/hardware. [Online]. Available: https://developer.microsoft.com/en-us/windows/kinect/hardware

[2] Y. Xiong and S. A. Shafer, "Depth from focusing and defocusing," in *Computer Vision and Pattern Recognition, 1993. Proceedings CVPR'93., 1993 IEEE Computer Society Conference on*. IEEE, 1993, pp. 68–73.

[3] A. Pentland, T. Darrell, M. Turk, and W. Huang, "A simple, real-time range camera," in *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR'89., IEEE Computer Society Conference on*. IEEE, 1989, pp. 256–261.

[4] M. Watanabe and S. K. Nayar, "Rational filters for passive depth from defocus," *International Journal of Computer Vision*, vol. 27, no. 3, pp. 203–225, 1998.

[5] J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern recognition*, vol. 37, no. 4, pp. 827–849, 2004.

[6] A. Pentland, S. Scherock, T. Darrell, and B. Girod, "Simple range cameras based on focal error," *JOSA A*, vol. 11, no. 11, pp. 2925–2934, 1994.

[7] S. K. Nayar, M. Watanabe, and M. Noguchi, "Real-time focus range sensor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 12, pp. 1186–1198, 1996.

[8] O. Ghita, P. F. Whelan, and J. Mallon, "Computational approach for depth from defocus," *Journal of Electronic Imaging*, vol. 14, no. 2, pp. 023 021–023 021, 2005.

[9] F. Moreno-Noguer, P. N. Belhumeur, and S. K. Nayar, "Active refocusing of images and videos," *ACM Transactions On Graphics (TOG)*, vol. 26, no. 3, p. 67, 2007.

[10] A. Mennucci and S. Soatto, "On observing shape from defocused images," in *Image Analysis and Processing, 1999. Proceedings. International Conference on*. IEEE, 1999, pp. 550–555.

[11] L. Jianzhuang, L. Wenqing, and T. Yupeng, "Automatic thresholding of gray-level pictures using two-dimension otsu method," in *Circuits and Systems, 1991. Conference Proceedings, China., 1991 International Conference on*. IEEE, 1991, pp. 325–327.