

IR Shape From Shading Enhanced RGBD for 3D Scanning

Nolan Lunscher
Georges Younes
David Abou Chacra
Henry Leopold
Jeremy Pinto
John Zelek

University of Waterloo, ON, Canada
University of Waterloo, ON, Canada
University of Waterloo, ON, Canada
University of Waterloo, ON, Canada
University of Waterloo, ON, Canada
University of Waterloo, ON, Canada

Abstract

RGBD Cameras such as the Microsoft Kinect that can quickly provide usable depth maps have become very affordable, and thus very popular and abundant in recent years. Beyond gaming, RGBD cameras can have numerous applications, including their use in affordable 3D scanners. These cameras however are limited in their ability to capture finer details. We explore the use of additional 3D reconstruction algorithms to enhance the depth maps produced from RGBD cameras, allowing them to capture more detail.

1 Introduction

RGBD cameras have become very affordable and accessible, with a number of models available on the market today. One of the most notable RGBD camera models is the Kinect, which has sold millions world wide. The Kinect uses a Structured Light system with an Infrared (IR) camera and projector pair for depth measurement, along with an RGB camera [1]. More recently, Intel has released a line of RGBD cameras including the Intel Realsense F200, R200 and SR300. The F200 and SR300 cameras work similar to the Kinect system, utilizing Structured Light, with an IR camera and projector. The R200 camera utilizes stereo IR cameras to measure scene depth, with an IR projector used to provide additional texture to the scene. All of the Intel Realsense cameras also have an RGB camera similar to the Kinect, for the capture of color information.



Fig. 1: An Intel Realsense F200 RGBD Camera.

These cameras are very cheap and an easily accessible, making them an attractive offering as a platform for creating affordable 3D scanners. The main limitation with off the shelf RGBD cameras is however that they are known to produce noisy depth maps, and in the case of the Kinect, also suffer from quantization error [1].

2 RGBD 3D Scanning

While these cameras produce noisy depth maps, algorithms such as KinectFusion [2], can be used to scan a scene and produce accurate scene models by combining multiple overlapping view points. KinectFusion works by moving a Kinect camera around a scene, while the camera records the RGBD data and in real time produces an overall scene model.

In work by Zhe Wu et. al. [3], an RGBD-M scanner was created, consisting of an IR camera and multiple IR light sources used in combination with an RGBD camera. This system supplements KinectFusions scene geometry, with the ability to capture more accurate BRDF information in indoor scenes. In work by Gyeongmin Choe et. al. [4], the initial geometry from KinectFusion is enhanced using a Near IR Light Shading Model, to recover finer surface details.

These scanning algorithms are great for scanning stationary scenes, but are less ideal for scanning an object such as a person, without requiring them to remain very stationary. For scanning objects with the potential of movement, techniques for enhancing RGBD from a smaller number of overlapping images and viewpoints is more desirable.

3 RGBD Enhancement

The limited quality of the depth maps produced by off the shelf RGBD cameras have made them non ideal for 3D capture in applications that require more precision and ability to detect smaller details from single viewpoints. In order to tackle this problem, a number of techniques have been developed, that combine other 3D reconstruction algorithms with the outputs of the RGBD cameras, to produce enhanced results.

In work by Faraj Alhwarin et. al. [5], two Asus Xtion RGBD cameras are used together in a stereo setup. This system incorporates an additional depth map calculated from the two IR cameras in stereo, which is combined with the initial depth map to produce an enhanced result, that can account for objects that appear transparent to typical structured light cameras. Achuta Kadambi et. al [6] used a depth map created using Shape from Polarization to improve the quality of the initial RGBD camera depth map, which has the added benefit of being robust to materials and lighting conditions.

In work by Mohammadul Haque et. al. [7], a Kinect camera is used in combination with a Shape from Shading algorithm. This system utilizes an additional image from the IR camera, and with a separate uniform IR light source to acquire normal estimations, that are fused with the initial Kinect depth map, to produce improved results. Roy Or-El et. al. [8], used a near field lighting model, that explicitly accounts for object albedos and specularities to enhance depth maps, utilizing an addition image from the IR camera and a uniform IR light source. In [9], RGBD-Fusion, a Shape from shading algorithm, was used to enhance the depth map results. In this case, using the color RGB camera to capture the shading information under natural lighting, utilizing the initial depth map to account for multiple light sources.

4 Approach

When scanning objects with the potential for movement, all view points or separate images need to be in sync if they are to be fused together without creating artifacts. Due to this, utilizing data from RGBD cameras that can be captured all at once is preferred over methods requiring multiple capture steps. Additionally, IR camera images are preferred over color RGB images for scanning in indoor settings without controlled lighting, as the narrow IR band filters out most undesirable ambient light [4].

In this work we explore the use of the RGBD-Fusion algorithm [9], utilizing the IR camera image rather than the RGB image to enhance the depth map created by an Intel Realsense F200 camera, shown in Figure 1. The IR image we use for Shape from Shading depth estimation is taken at the same time as the image used for the cameras initial Structured Light depth estimation. This means that the IR image we use, contains the projected Structured Light pattern (non-uniform light). An example of an image taken with the

F200's IR camera, in the presence of the IR projector's pattern is shown in Figure 2.



Fig. 2: The IR image used to enhance the depth map of an RGBD camera using the RGBD-Fusion algorithm [9].

5 Results

Preliminary results are shown Figure 3, which shows a noticeable improvement in the quality of the depth map. The projected IR pattern as a lighting source does not seem to produce much in the way of negative effects in the resulting depth map.

6 Conclusion

RGBD camera depth maps can be enhanced in a number of different ways, but these generally require additional hardware or a more complex capture procedure. IR cameras as opposed to RGB cameras for Shape from Shading, allow for controlled lighting in typical naturally lit indoor environments, which is more practical for real world applications. Enhancing RGBD camera depth maps utilizing images that are already required by the cameras initial depth map allows for faster data capture, without additional overhead.

References

- [1] Ryan Fanello, S., Rhemann, C., Tankovich, V., Kowdle, A., Orts Escolano, S., Kim, D. and Izadi, S. HyperDepth: Learning Depth From Structured Light Without Matching. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2016).
- [2] Newcombe, R.A., Izadi, S., Hilliges, O., Molyneaux, D., Kim, D., Davison, A.J., Kohi, P., Shotton, J., Hodges, S. and Fitzgibbon, A. KinectFusion: Real-time dense surface mapping and tracking. Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on (2011).
- [3] Wu, Z., Yeung, S.K. and Tan, P. Towards Building an RGBD-M Scanner. arXiv preprint arXiv:1603.03875 (2016).
- [4] Choe, G., Park, J., Tai, Y.W. and Kweon, I.S. Refining Geometry from Depth Sensors using IR Shading Images. arXiv preprint arXiv:1608.05204 (2016).
- [5] Alhwarin, F., Ferrein, A. and Scholl, I. IR stereo kinect: improving depth images by combining structured light with IR stereo. Pacific Rim International Conference on Artificial Intelligence (2014).
- [6] Kadambi, A., Taamazyan, V., Shi, B. and Raskar, R. Polarized 3d: High-quality depth sensing with polarization cues. In Proceedings of the IEEE International Conference on Computer Vision (2015).

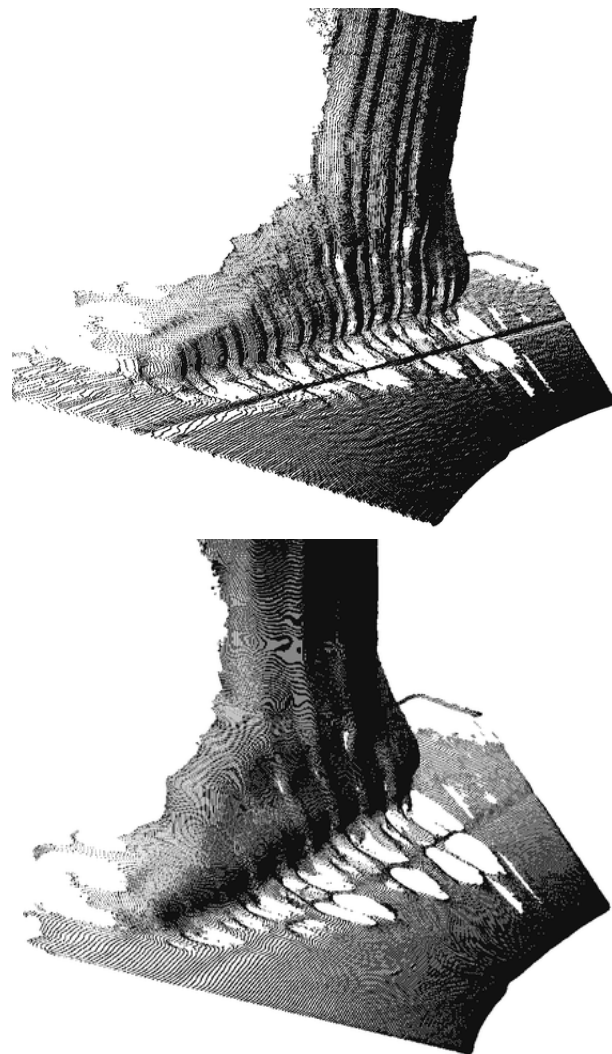


Fig. 3: Top: The point cloud depth map from an Intel Realsense F200 RGBD camera. Bottom: The enhanced point cloud depth map, after using the RGBD-Fusion algorithm [9] with an IR image.

- [7] Haque, M., Chatterjee, A., and Madhav Govindu, V. High quality photometric reconstruction using a depth camera. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2014).
- [8] El, R.O., Hershkovitz, R., Wetzler, A., Rosman, G., Bruckstein, A.M. and Kimmel, R. Real-Time Depth Refinement for Specular Objects. arXiv preprint arXiv:1511.08886 (2015).
- [9] Or-EI, R., Rosman, G., Wetzler, A., Kimmel, R. and Bruckstein, A.M. Rgbd-fusion: Real-time high precision depth recovery. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2015).