Project Proposal: Depth Estimation and Applications

Yifei Wang and Noam Weinberger

Motivation

As automation and artificial intelligence become increasingly popular, so does the need for accurate computer vision and scene rendering. One of the most important aspects of computer vision is depth estimation. This problem has broad applications in areas such as robot vision, human computer interfaces, intelligent visual surveillance, 3D image acquisition, and intelligent driver assistance systems.[1] Depth estimation can also be used for artistic purposes, such as image refocus, bokeh, simulations, and special effects.

Overview

In our project, we will explore and implement image processing algorithms to estimate depth from two corresponding images taken by a dual camera. Once we have a method for generating a depth map of the image, we hope to extend this task further to various possible applications:

- Refocus/Bokeh - With a depth map, we will be able to blur specific depth regions in an image, while leaving other depths alone. This allows us to refocus the image on certain objects or use artistic blurring.
- Fog Simulation - With a depth map, we will be able to realistically simulate fog, smoke, or haze in an originally clear image. Naturally, more distant objects will be more obscured by fog, which is possible once closer and more distant objects are differentiated.
- Object Trajectory Simulation - With a depth map, we will be able to simulate a realistic path an object would take through a scene. For example, if a ball were to roll through the scene at a certain depth, it would pass in front of or behind different objects, depending on their relative ordering in space.
- Absolute Distance Estimation - Though a depth map only shows relative depth of objects, with some prior information, it is possible to estimate absolute depth as well. This might be accomplished if the sizes of some of the objects are known, if the true depths of some of the objects are known, or using some initial calibration of the camera on test images.
- Parallax Simulation - With a depth map, we will be able to show a new perspective of the scene not found in either of the source images. This can be done, to some degree, by extending the small known parallax to a larger one, as if a more shifted image were taken.

Due to time constraints, we do not expect to be able to accomplish all of these tasks, however with mentor guidance we will focus on specific ones and develop them as much as time allows.

---

[1] Lee, Seungwon, et al. "Single image-based depth estimation using dual off-axis color filtered aperture camera." *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013.

Milestones

1. Demosaicing, Image correction: We will process the raw image by implementing image processing pipeline. By 3.3
2. Object recognition and depth estimation: We will identify the objects in the images from dual camera. And by calibrating the displacement of objects, we will be able to get the depth information. By 3.6
3. Refocusing and background effects: We can implement background rendering or refocusing based on depth map we have. By 3.12
4. Other Effect Demonstrations: Object trajectory simulation, Parallax... By 3.12

Related work

There are some work on depth estimation by a variety of methods, for example color filtered dual camera, single camera video clip processing[2] and dual camera on mobile phone[3]. The idea is once we have a few photos from different perspective, we will be able to estimate the depth of the objects in the photo.

Based on the depth information, there are a lot of work implementing photo rendering, such as bokeh effect, foggy scene simulation[4], which basically utilize depth information as a rough 3D model and perform some post processings.

---

[2] Tao, Qiuyan, et al. "Relative depth estimation with an uncalibrated camera for image refocus." *Wireless Communications & Signal Processing (WCSP), 2015 International Conference on*. IEEE, 2015.

[3] Martinello, et al. “Dual Aperture Photography: Image and Depth from a Mobile Camera.” *Computational Photography (ICCP), 2015 IEEE International Conference on.* IEEE, 2015.

[4] Guo, Fan, Jin Tang, and Xiaoming Xiao. "Foggy scene rendering based on transmission map estimation." *International Journal of Computer Games Technology* 2014 (2014): 10.