

3D Image Reconstruction from Multiple 2D Images

Introduction

The main goal of this project is to prototype a system which reconstructs rudimentary 3D images from a batch of 2D images.

The project will be split into 3 parts: data collection, depth map generation/fusion, and 3D visualization. Images will be obtained off-line. Furthermore, all computations will occur off-line (i.e. there will be no real-time processing elements to this project).

1. Experiment Setup: Data Collection

One of the main goals of the experimental setup is to keep the project within scope of the EE368 course. The experiment will be designed such that one can recreate 3D images from 2D images without deviating much from the themes presented in EE368. Furthermore, the goal of the experiment is to create a scenario in which one can avoid some of the obstacles faced during the 2D-to-3D conversion process.

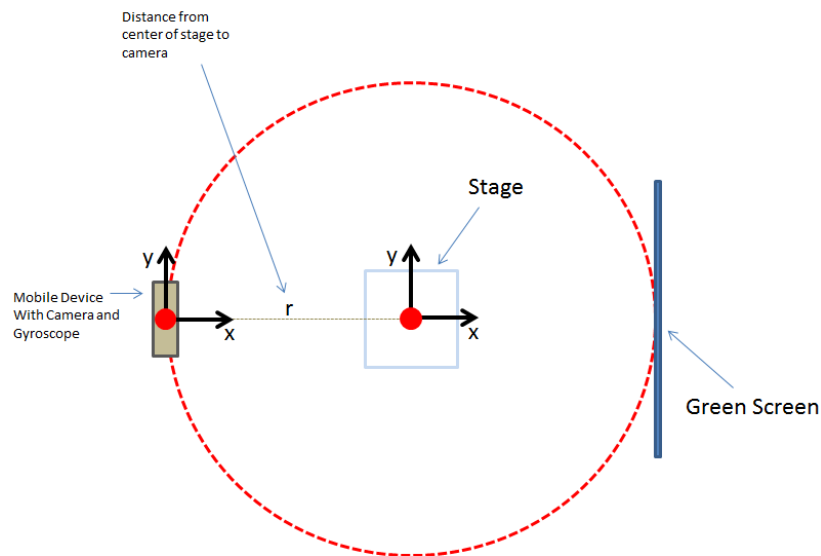


Figure 1: The experimental setup. The setup will consist of: mobile device with camera and gyroscope capabilities, a stage to hold the object to be reconstructed in 3D, and a green screen. The distance between the mobile device and the center of the

stage shall be held constant through the experiment. During the experiment, the camera will be rotated around the stage at constant intervals.

An experiment will be designed to allow for the collection of well-conditioned 2D images. The term 'well-conditioned' shall refer to a scene with the following characteristics: uniform illumination, easily segmented, and straightforward correspondences when referencing multiple images of the same scene.

The experiment will consist of: a mobile device for image collection (SM-G925T), a stage for setting objects to be reconstructed in 3D, and a green screen to aid in segmentation. The distance between the camera and stage will be held constant.

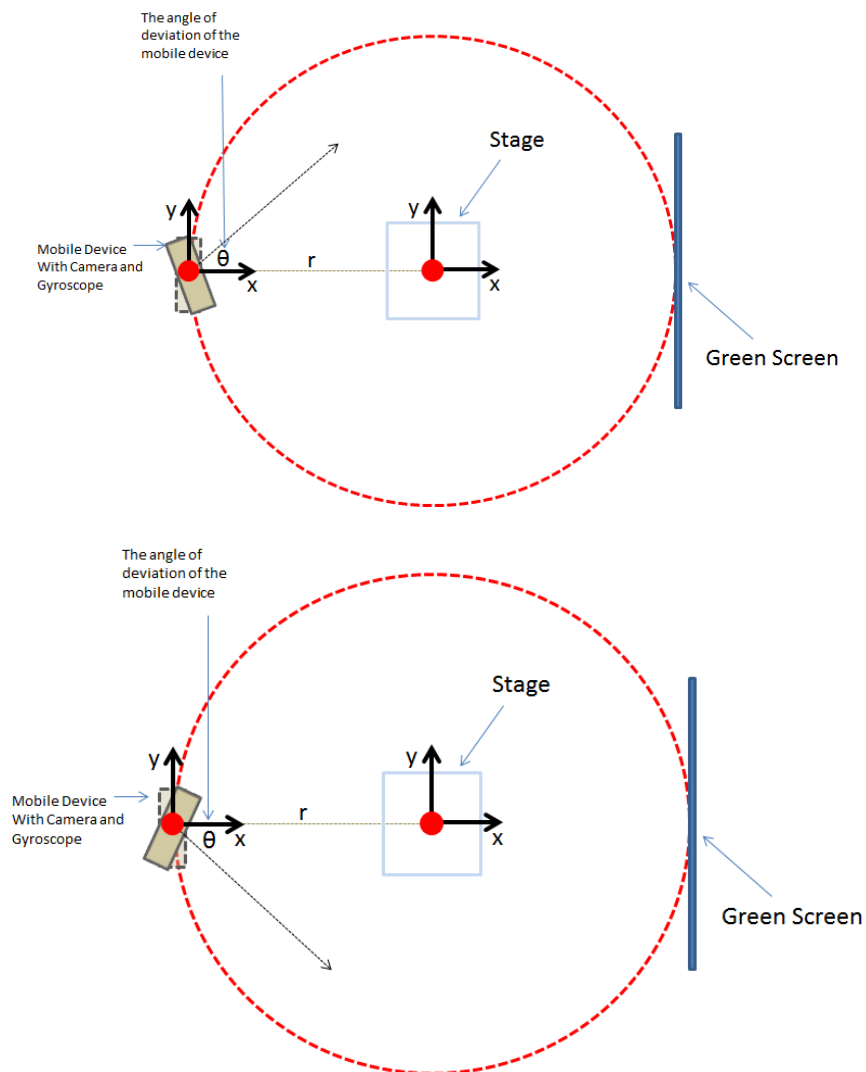


Figure 2: During the experiment, multiple images will be obtained by the mobile device at various angles between the camera and stage. The first image obtained, at a deviation $\theta=0$ will be known as the template image. θ shall be varied at evenly spaced intervals and symmetric about $\theta = 0$ (i.e. symmetric about the template image).

The angle between the camera and scene (θ) will be varied in order to provide the 3D reconstruction algorithm with images of the scene at various perspectives. These images will be used to recreate the depth information of the scene. $\Delta\theta$ will be constant throughout the entire experiment. N images will be collected at $\Delta\theta$. N will be determined empirically such that the images obtained provide enough data to reconstruct a 3D scene with a reasonable amount of processing.

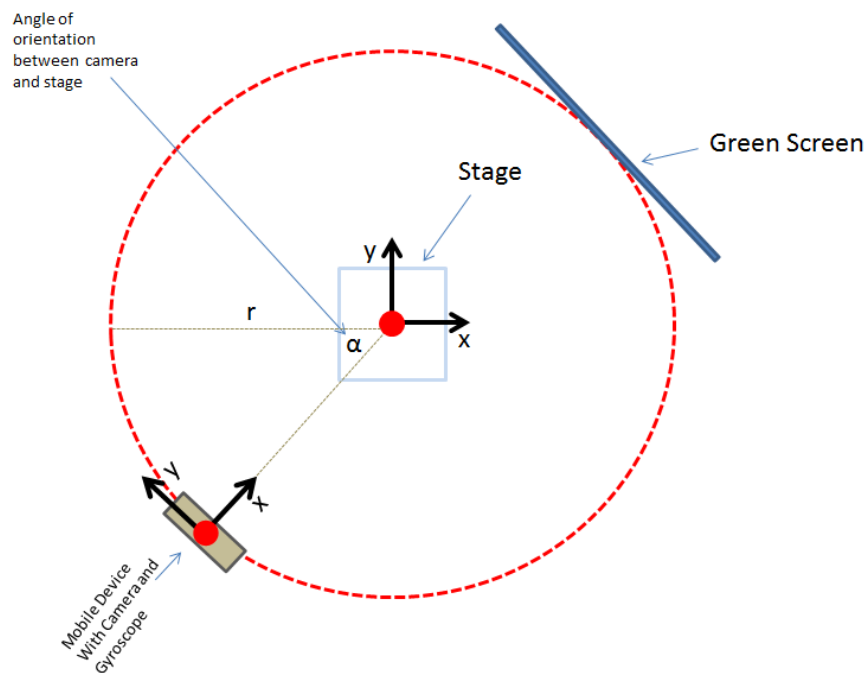


Figure 3: During the experiment, multiple images will be obtained by the mobile device at various points on a circle surrounding the stage. This angle shall be referred to as α .

The camera shall be placed at equally spaced points on the circle surrounding the stage (see Figure 3). That is to say, $\Delta\alpha$ will be constant. The goal of the rotation about the stage at intervals of $\Delta\alpha$ is to provide the 3D reconstruction algorithm with scene information at 360° . The amount of data collected will be determined empirically such that the images obtained provide enough data to reconstruct a 3D scene with a reasonable amount of processing.

2. Generating 3D Information

The second portion of the project will involve generating a depth map using the 2D images collected at each α (i.e. at different points on the circle in Figure 3). It may be possible that only a small set of the images are used during 3D reconstruction.

The depth generation algorithms are roughly classified into three categories which utilize different kinds of depth cues: the binocular, monocular and pictorial depth cues [4]. In this project, pictorial depth cues

will be explored. Other conventional depth estimation methods include: the use of color channels, vanishing points, image warping [3], edge information [4], and object classification[1][5].

The approach to obtain the depth map will involve the following idea: items in the scene closest to the observer will translate much less than objects furthest away from the observer [9] as θ varies. For key-point detection and image fusion, techniques from EE368 will be employed. The goal is to represent the depth map in the form of a 256 bit image such that many of the techniques learned in EE368 can be used to process the depth map images.

3. Visualization

The final portion of the project will involve rendering the 3D data. In this step, the depth information generated in Step 2 will be rendered using OpenGL. GLUT will be used to provide a simple method for a user to observe the rendered 3D object.

References

[1] S. Battiato, S. Curti, M. La Cascia, M. Tortora, E. Scordato. Depth-Map Generation by Image Classification

[2] Kolev, K., Meier, L., Camposeco, F., Saurer, O., Pollefeys, M. Live Metric 3D Reconstruction on Mobile Phones Conference on Computer Vision (ICCV), 2013 IEEE International

[3] Daniel Donatsch, Nico Färber, Matthias Zwicker 3D CONVERSION USING VANISHING POINTS AND IMAGE WARPING

[4] S. Bharathi A. Vasuki. 2D-To-3D Conversion of Images using Edge Information International Conference on Recent Trends in Computational Methods, Communication and Controls

[5] Po Lai-Man Semi-Automatic 2D-to-3D Image Conversion Techniques for Touchscreen Device Applications (Slides)

[6] Ofir Pele. SIFT - The Scale Invariant Feature Transform. International Journal of Computer Vision, 60, 2 (2004), pp. 91-110 (Slides)

[7] Sylvain Paris. Methods for 3D Reconstruction from Multiple Images (Slides)

[8] Richard I. Hartley Euclidean Reconstruction from Uncalibrated Views GE - Corporate Research and Development

[9] 3D Reconstruction from Two 2D Images Ted Shultz and Luis A. Rodriguez