3D Reconstruction from Camera Images

Cheng, J., Walter, K. & Yang, J.

1 Introduction

THE reconstruction of 3D models is an active area of research in computer vision and is highly relevant to industry and society at large. The aim of 3D re-construction is to capture the shape and appearance of real objects. Some examples of useful applications of 3D reconstruction include use in medical imaging, robotic navigation, augmented reality and movie production. In this project we will explore and implement a syste capable of creating a 3D model of an object using 2D photos of the object as input.

2 Project Goals

Our goal is to produce a visually appealing 3D reconstruction of an object using camera images taken from various vantage points around the object.

3 Literture Survey

There are many ways to infer 3D shape from still images. This literature survey aims to outline several of the more important ways this inferenence can be accomplished that have been documented in the research literature.

Shape from object attribute

One of the classic ways to obtain a 3D model from a still image is to inspect the shading of an object. This method was proposed by Horn [1] in 1975 and works due to the intensity of a region of the object dropping off as the shape moves into the plane of the picture. Essentially, as a shaded region becomes 'darker' we gain information about the change in direction of the surface normal and hence the orientation of the point. Orientaion information about each point in the image gives us a description of the overall 3D shape of the object.

Object texture can also provide information about object shape. The idea behind this method is that textures are made up of instances of pattern elements arranged in some way. Either the distribution of the elements or the elements themselves can be distorted and the new state of the system can be analysed. For example, Lobay used clusters of keypoints (obtained by SIFT) to identify pattern elements on clothes and used several instances of these elements to construct a model of the piece of clothing [2].

Like the analysis of shading on an object surface, we can analyse blurred regions of an object to infer shape. Nayar & Nakagowa took successive images of an object at different levels of focus. These focus levels were then interpolated to get accurate depth estimation [3].

Active Rangefinding

Several methods that measure the distance from the camera to the object have been developed to infer shape.

Curless and Levoy [4] used a laser to sweep across a scene. As the laser passed over objects, the beam was deformed. The deformation of straight beam created by the laser was used with triangulation to estimate the 3D location of each point the beam touched.

A variation on this technique was proposed by Bouget and Perona [5]. Instead of a laser, a stick was used in front of a light that was shining on the object. The shadow produced by the stick was swept over the object and the deformation of the shadow was used in the same way that Curless and Levoy used the laser.

Point based representations

These representations involve describing the surface of an object with surface points, which can be thought of as physical particles with attractive forces and an orientation [6]. Having the points behave in this manner allows for interpolation of partial 3D data to account for 'holes' in the representation while retaining the ability to render the particle system as a continuous surface.

Model Based Representaions

If we know something about the object that we are modelling, we can use this information to obtain a more accurate and detailed 3D model. For example, architecture is generally made of planar shapes that are oriented at 90-degree angles to each other. The system of Zhu and Kanade [7] is a good example of this, as it used several geometic primitives such as planes, cyclinders and cones to describe various architectural features of the scene such as roofs or doors and windows.

4 Problem Decomposition

4.1 Image dataset

We will source our data from the Multi View Stereo datasets available from the EPEL Computer vision lab. These datasets contain a complete set of images along with camera matrics and ground truth data. This dataset also includees a 3D model constructructed from the images and we may use this to test similarity with our result.

4.2 Image preprocessing

Most images taken by camera suffer from distortions (i.e.

straight lines become curved) and may not be in a suitable state for effective feature detection. We will use various techniques (i.e. camera calibration) to counteract camera distortions and apply image processing steps learnt in class (i.e. thresholding, histogram equalization) to get the images ready for feature detection and extraction in the next step.

4.3 Feature extraction

In this step we will apply the Scale Invariant Feature Transform algorithm to camera images of the object from various vantage points. For feature matching/extraction we will use the optical flow and K-nearest- neighbours algorithms.

4.4 Build point cloud

We aim to build a point-cloud out of the features extacted in the previous step. To accomplish this, we will use Multi View Stereo and the Point Cloud Library (PCL) in OpenCV.

4.4 Modify point cloud

In the previous step we built a 'rough' point cloud. In this step we will apply algorithms such as Moving Least Squares and Fast Triangulation to make the model surface continous and fill any 'holes' in the model. This will make the model look more realistic in the final visualization step.

4.4 Show final 3D model

Finally, we will visualise the 3D model using PCL, Meshlab or Visual Toolkit.

5 Plan

Week 9:

We will be focusing on image preprocessing and feature extraction during this week.

Tasks:

- Find appropriate dataset and get it in a form so we can start using it in the project (Jingshi Yang)
- Image preprocessing (Jingran Cheng)
- Feature extraction (Kane Walter)

Week 10:

We will be focusing on building the point cloud during this week.

Tasks:

- Research algorithms for building point clouds (Jingran Cheng)
- Implent chosen algorithm (Kane Walter)
- Test implementation (Jinshi Yang)

Week 11:

We will be focusing on altering the point cloud for better visulisation during this week.

Tasks:

- Research algorithms for altering the point cloud (Kane Walter)
- Implement chosen algorithm (Jingshi Yang)
- Work on visualization (Jingran Cheng)

7 References

[1] R. Szeliski, *Computer vision*. London: Springer, 2011, p. 580.

[2] A. Lobay and D. Forsyth, "Shape from Texture without Boundaries", *International Journal of Computer Vision*, vol. 67, no. 1, pp. 71-91, 2006.

[3] S. Nayar and Y. Nakagawa, "Shape from focus", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 8, pp. 824-831, 1994.

[4] B. Curless and M. Levoy, "Better optical triangulation through spacetime analysis", In *Fifth International Conference on Computer Vision*, pp. 987-994, Cambridge, Massacheusetts, 1995.

[5] J. Bouguet and P. Perona, *International Journal of Computer Vision*, vol. 35, no. 2, pp. 129-149, 1999.

[6] R. Szeliski and D. Tonnesen, "Surface modeling with oriented particle systems", *ACM SIGGRAPH Computer Graphics*, vol. 26, no. 2, pp. 185-194, 1992.

[7] Z. Zhu and T. Kanade, "Modeling and Representations of Large-Scale 3D Scenes", *International Journal of Computer Vision*, vol. 78, no. 2-3, pp. 119-120, 2008.