3D Reconstruction from Camera Images

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

**Abstract**—These instructions give you guidelines for preparing papers for IEEE Computer Society Transactions. Use this document as a template if you are using Microsoft Word 6.0 or later. Otherwise, use this document as an instruction set. Please note that use of IEEE Computer Society templates is meant to assist authors in correctly formatting manuscripts for final submission and does not guarantee how the final paper will be formatted by IEEE Computer Society staff. This template may be used for initial submissions; however, please consult the author submission guidelines for formatting instructions as most journals prefer single column format for peer review. An abstract should be 100 to 200 words for regular papers, no more than 50 words for short papers and comments, and should clearly state the nature and significance of the paper. Abstracts *must not* include mathematical expressions or bibliographic references. Please note that abstracts are formatted as left justified in our editing template (as shown here).

—————————— ◆ ——————————

# 1 Introduction

T

HE reconstruction of 3D models is an active area of research in computer vision and is highly relevant to industry and society at large. The point of 3D construction is to capture the shape and appearance of real objects. Some examples of useful applications of 3D reconstruction include use in medical imgaing, robotic navigation, augmented reality and movie production. In this project we will explore the methods required to create 3D reconstructions.

# 2 Project Goals

Our goal is to produce a visually appealing 3D reconstruction of an object using camera images taken of this obejct from various locations 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. Several assumptions need to be made in this method such as the object having uniform reflectance and albedo which may not be true in practice.

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 frontal 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 [3] took successive images of an object at different levels of focus. These focus levels are then interpolated to get accurate depth estimation.

*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. This deformed line was then used to estimate the 3D location of each point the beam touch.

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 this 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 point 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 dataset 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. As well as the dense 3D model avalible for compareation. (What does the last sentence mean?)

**4.2 Image preprocessing**

Most images taken by camera suffer from distortions (i.e. straight lines become curved) so the aim of this task is to alter the images in a way that makes them more amenable to analysis. In this task we will use camera calibration in OpenCV to counteract these negatives effects. Find key points in each image using SIFT. (What does the last sentence mean?)

**4.3 Feature extraction**

Optical flow and K-nearest- neighbors will be used for feature extraction. We decided to use this technique over SIFT or SURF because it is superior in efficiency and matching features.

**4.4 Build point cloud**

We aim to build a point-cloud out of the features extacted in the previous step. We will use the Virtual SFM library to create a point cloud from the extracted features.

Alternatively, we can use the Point Cloud Library (PCL) in OpenCV along with MVS dataset to construct a point cloud.

**4.4 Modify point cloud**

After building the point cloud we will modify it to improve visual appearance in the final step. We will fill in gaps in the model and give it a smooth overall appearance. To accomplish this, we will use algorithms (i.e. fast triangulation) to connect points, smoothing and normal estimation using Moving Least Squares.

**4.4 Show final 3D model**

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

**5 Plan**

**Week 9:**

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

Tasks:

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

**Week 10:**

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

Tasks:

1. Research algorithms for building point clouds (Jingran Cheng)
2. Implent chosen algorithm (Kane Walter)
3. Test implementation (Jinshi Yang)

**Week 11:**

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

Tasks:

1. Research algorithms for altering the point cloud (Kane Walter)
2. Implement chosen algorithm (Jingshi Yang)
3. 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.