Chapter 2
3D Reconstruction Using Multiple View Stereo and a Brief Introduction to Kinect

Brojeshwar Bhowmick
IIT Delhi and Innovation Lab, Tata Consultancy Services, India

ABSTRACT

This chapter deals with the methodology of 3D reconstruction, both sparse and dense. The basic properties of the projective geometry and the camera models are introduced to understand the preliminaries about the subject. A more detail can be found in the book (Hartley & Zisserman, 2000). The sparse reconstruction deals with reconstructing 3D points for few image points. There are gaps in the reconstructed 3D points. Dense reconstruction tries to fill up gaps and make the density of the reconstruction higher. Estimation of correspondences is an integral part of multiview reconstruction and the author will discuss the point correspondences among images here. Finally the author will introduce the Microsoft Kinect, a divice which directly capture 3D information in realtime, and will show how to enhance the Kinect point cloud using vision framework.

INTRODUCTION

3D reconstruction from multiple images is an active area of research for past two decades and has come up with enormous stable algorithms in state of the art. In this chapter we mainly deal with basic geometry involved in establishing relations between images and will define camera models. The image formation and relation among images are established in projective geometric framework. A 3D point could be obtained from two or more 2D points in projective space or in Euclidean space depending on the camera models. If the cameras are calibrated in Euclidian spaces, the reconstruction (getting 3D point) obeys Euclidean geometry laws else the camera model and 3D structure are in projective space. Each of these concepts will be dealt in detail in the following sections. Then, we will focus on the methodologies for point correspondence across image set, which is an important aspect in reconstruction. Once these correspondences are found out we will discuss the method of obtaining the 3D points through triangulation. Then we will present a generalized framework called Bundle Adjustment which generates the camera and 3D structure simultaneously. We will also be dealing with different methods of dense reconstruction using the sparse reconstruc-

DOI: 10.4018/978-1-4666-4558-5.ch002

tion as initial input. The dense reconstruction takes care of the overall geometry in finer scale. This will complete the basic overview on reconstruction from images. With the latest release of Kinect device by Microsoft, there is a scope of vision based work using this device. One of the problems with the Kinect is that it provides the point cloud in VGA resolution. There is a scope of enhancing the resolution of the Kinect using two or more high definition cameras, which will be discussed further.

FUNDAMENTALS OF PROJECTIVE GEOMETRY AND PROJECTIVE CAMERA MODELS

Why Projective Geometry?

Image formation is a perspective transformation from 3D world scene to 2D plane and most of the Euclidean geometric property is lost due to this transformation. One of the most familiar properties is that parallel lines which never meet, or some fancy way of saying they meets at infinity in Euclidean space, doesn't hold true after perspective transformation in image as shown in Figure 1. This figure shows two parallel rail lines meet at a point called vanishing point. This phenomenon distinguishes projective geometry from Euclidean geometry. The notion of distance and parallelism is destroyed in projective transformation. When the notion of distance is removed from Euclidean geometry keeping only parallelism, the structure becomes affine. Further, removing the preservation of parallelism property gives rise to projective structure. A more comprehensive detail can be found in (Hartley & Zisserman, 2000).

Perspective Viewing

Figure 1(a) shows the perspective phenomenon, where parallel lines meet at a point called the vanishing point. In perspective drawing an image point is formed at the intersection of the ray emanating from the object through the camera center, with the image plane as shown in Figure 1(b). This also shows that the spacing between the world points are not in sync with the spacing in between the image points. Although Figure 1(b) shows the construction of image under perspective projection, it is difficult to get a similar construction for the image of an image. Figure 1(c) shows that the line S is being imaged through optical center O_S in line s, and this line s is again projected through O_σ. But there is no optical center which is the intersection of the lines joining points of S and σ. A more sophisticated method is required to explain this, and that general method is called projective transformation (Mundy & Zisserman, 1992). A chain of perspective transformations may not be perspective projections but always projective transformation and any projective transformations can be decomposed into two or more perspective transformations (Mundy & Zisserman, 1992).

Properties of Projective Plane

In Euclidean plane, distances are well defined and also parallel lines meet at infinity. If we remove the notion of distance and only preserve the parallelism, then the structure formed, is called affine plane. Here, rectangle can be transformed into arbitrary parallelogram keeping lines parallel. Further, if we remove the notion of parallel lines then the structure is called projective structure, where any two lines meet at a point. In such cases all parallel lines also meet at some point which we call the ideal point. So, in an affine plane, if we introduce these ideal points and consider them as a part of the geometric structure, we get projective plane. In a projective plane following two properties hold,

- **Property 1:** Two distinct points produce a unique line.

Related Content