

Chapter 8

Projective Geometry for 3D Modeling of Objects

Rimon Elias
German University in Cairo, Egypt

ABSTRACT

This chapter surveys many fundamental aspects of projective geometry that have been used extensively in computer vision literature. In particular, it discusses the role of this branch of geometry in reconstructing basic entities (e.g., 3D points, 3D lines, and planes) in 3D space from multiple images. The chapter presents the notation of different elements. It investigates the geometrical relationships when one or two cameras are observing the scene creating single-view and two-view geometry. In other words, camera parameters in terms of locations and orientations, with respect to 3D space and with respect to other cameras, create relationships. This chapter discusses these relationships and expresses them mathematically. Finally, different approaches to deal with the existence of noise or inaccuracy in general are presented.

INTRODUCTION

Techniques are developed to reconstruct objects/surfaces in 3D space. These techniques use groups of images taken by cameras. Variations of the problem include 3D reconstruction from uncalibrated monocular image sequence (Azevedo, Tavares, & Vaz, 2009, Fitzgibbon, Cross, & Zisserman, 1998, Pollefeys, Koch, Vergauwen, & Gool, 1998); 3D reconstruction from calibrated monocular image sequence (Nguyen & Hanajik,

1995); and 3D reconstruction from stereo images. This later case includes pairs of images taken at the same time by two cameras or at two different instants by one camera provided that the scene is static. In many cases, the solution is divided into two steps (Zhang, 1995). These steps are:

1. Extracting and matching features between corresponding images; and
2. Determining structure from corresponding features.

DOI: 10.4018/978-1-4666-3994-2.ch008

Projective geometry (O. Faugeras, 1996, R. I. Hartley & Zisserman, 2004, Elias, 2009b) plays a key role in solving different problems in computer vision in general and determining structure from corresponding features in particular. Indeed, using the notation of this branch of geometry simplifies the complicated relationships among different spaces. When a camera takes different shots (i.e., perspective projections) for the same object or scene, geometrical relationships can be established among images. Projective geometry can determine these relationships mathematically in an efficient and easier way.

Consider the situation where two cameras observing a 3D space point of unknown coordinates. In this case, two point projections are formed onto two images. If the locations of these points as well as the locations of the cameras are known, two rays starting at the cameras and passing through the projections may be constructed to intersect in space at the 3D location of the point. This 3D reconstruction process can be handled easily in projective geometry (Beardsley, Zisserman, & Murray, 1997).

Epipolar geometry (O. Faugeras, 1996) is emerged in case of stereo vision with a number of useful relationships and matrices. The fundamental matrix (Luong & Faugeras, 1995) is an example of these matrices that can be used to limit the search for correspondences. When these correspondences are known, different objects can be reconstructed in 3D space.

In this chapter, we will discuss the basics of projective geometry (Elias, 2009a) and the notation of different entities (e.g., points, lines) expressed in 2D as well as 3D spaces. We will also discuss the mathematics of perspective projection when one or two cameras are used to view the same object/scene (i.e., single-view and two-view geometry). Having those main points explained, we will explain how to use projective geometry in order to reconstruct points, lines and planes in 3D space given their projections onto different images.

New directions of the 3D reconstruction problem include modeling from massive collections of images (e.g., images for famous places obtained from the Web) (Snavely, Seitz, & Szeliski, 2008, 2006a, 2006b) and learning 3D scene structure using clues from a single image (Saxena, Sun, & Ng, 2007). However, such approaches are beyond the scope of this chapter.

NOTATION

In this section, the notations used to describe points, lines, planes and directions in the rest of the chapter are mentioned. We indicate a 2D point by a bold lowercase letter (e.g., **p**) while a 3D point is indicated by a bold uppercase letter (e.g., **P**). A matrix of any dimensions is represented by a fixed-size letter (e.g., **P**). A scalar variable is represented by an italic letter (e.g., *p*). Finally, corresponding points in different scene views are indicated using the prime notation (e.g., **p**, **p'**).

2D Points

A 2D point may be represented in homogeneous or inhomogeneous coordinates. The inhomogeneous point **p** is identified by a 2D vector $[x, y]^T$ representing the Cartesian coordinates of the point. The same point **p** is represented in homogeneous coordinates as a 3D vector $[x, y, 1]^T$. In general, the third term may not be equal to 1 (e.g., $[x, y, s]^T$). In such cases, the Cartesian coordinates of the point can be estimated by dividing all the terms

by the third one (e.g., $\left[\frac{x}{s}, \frac{y}{s}, 1 \right]^T$).

3D Points

Similar to 2D points, a point in 3D space can be represented in both inhomogeneous and homogeneous coordinates. The inhomogeneous coordinates of a 3D point **P** identify the Cartesian

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/projective-geometry-modeling-objects/77542

Related Content

Video Content-Based Retrieval Techniques

Waleed E. Farag and Hussein Abdel-Wahab (2004). *Multimedia Systems and Content-Based Image Retrieval* (pp. 114-155).

www.irma-international.org/chapter/video-content-based-retrieval-techniques/27057

Research on Pre-Processing Methods for License Plate Recognition

Weifang Zhai, Terry Gao and Juan Feng (2021). *International Journal of Computer Vision and Image Processing* (pp. 47-79).

www.irma-international.org/article/research-on-pre-processing-methods-for-license-plate-recognition/270876

A Hybrid Multimodal Medical Image Fusion Technique for CT and MRI Brain Images

Leena Chandrashekar and Sreedevi A. (2018). *International Journal of Computer Vision and Image Processing* (pp. 1-15).

www.irma-international.org/article/a-hybrid-multimodal-medical-image-fusion-technique-for-ct-and-mri-brain-images/212373

Foreign Body Shape Classification Using Intuitionistic Fuzzy Rule Based Approach on Pediatric Radiography Images

Vasumathy M. and Mythili T. (2020). *International Journal of Computer Vision and Image Processing* (pp. 45-57).

www.irma-international.org/article/foreign-body-shape-classification-using-intuitionistic-fuzzy-rule-based-approach-on-pediatric-radiography-images/252233

An Overview of Image and Video Segmentation in the Last 40 Years

Yu-Jin Zhang (2006). *Advances in Image and Video Segmentation* (pp. 1-16).

www.irma-international.org/chapter/overview-image-video-segmentation-last/4834