Polyhedral Environment in Stereo Views: Representation and Extraction

Ronald Chung *Andrew Arengo*

Department of Mechanical and Automation Engineering
The Chinese University of Hong Kong, Shatin, Hong Kong
E-mail: rchung@cuhk.edu.hk

Abstract

A representation of polyhedral environment in a stereo pair of images is proposed and its three-dimensional recovery is presented. The polyhedral environment is represented as a number of image-to-image mappings in the form of matrices, each corresponding to a planar surface in the environment. Unlike a mere depth map, such a representation is segmented, in the sense that different surfaces in the environment correspond to different matrices and are separated. A mechanism is proposed to recover the representation even for scene that is not densely featured. Experimental results on typical polyhedral scenes are given.

1 Introduction

Reconstructing three-dimensional (3-D) information of a scene from its images, to the extent that surfaces and objects in it can be separated, is a basic goal of computer vision. Stereo vision is an important and well-studied vision cue for that. Surveys on stereo vision work over the years can be found in [1, 3].

To solve the reconstruction problem, the most important and the only scene-independent clue that stereo vision can exploit is the epipolar geometry. For any feature in one image, there exists in the other image a line named as the epipolar line on which the correspondence of the feature must be located. So long as the spatial relationship of the two cameras is known, the epipolar line is predictable, and the originally 2-D search in an image for stereo correspondence becomes a 1-D search along the corresponding epipolar line. However, even with such a useful constraint, matching features across the stereo images still has ambiguity along the epipolar lines. The classical approach of resolving the ambiguity is to rely on two assumptions: the surface-continuity or the surface-smoothness assumption, and the feature-ordering assumption. The first assumes that the scene is continuous or smooth, and the second assumes that features in the two images follow the same left-to-right order along the epipolar lines.

While the classical approach does give satisfactory results over smooth scenes, it has limited performance towards scenes with occlusions, for the two assumptions are not valid across occlusion boundaries. There have been extensions to the approach in the literature, which typically disable the two assumptions locally at selected places in the scene. However, the performance has been limited, owing to the mere fact that the disabling decisions are local.

In this paper, a representation of a polyhedral environment that is pictured in a stereo pair of images, like the building structures in the aerial views of an urban city or the corridor in the stereo views of inside a building, is proposed. The representation consists of a number of image-to-image mappings in the form of matrices, each corresponding to a planar surface in the environment. Unlike a mere depth map, the representation is segmented, in the sense that different surfaces in the environment correspond to different matrices and are separated. The essence of the representation is that it can be recovered through a simple mechanism even for environment that is not densely featured. For scene like a corridor, the representation is recovered not only with sparse depth estimates over the edges of walls and ceiling and floor, but also with the information that which line segments in the images form a wall or ceiling or floor in the environment. The representation can be used not only for depth recovery as required in autonomous navigation, but also for image transfer to arbitrary viewpoints as required in virtual reality applications.

2 Preliminaries on Homography

An image-to-image mapping was introduced recently by Faugeras [2] for three-view problems, [6, 5], namely the