13 Imaging Geometry

13.1 Introduction

Visual perception is undoubtedly one of the most precious and trustworthy sense organs of human beings for understanding the environment. In the recent past, researchers have added this invaluable faculty to machines in addition to their intelligence. Machine vision refers to the viewing or sensing of the environment by the computer, allowing it to synthesize information from the imagery of the concerned scene, analyze it, and finally carry out various interpretations or make decisions. Computer vision employs 3D imaging techniques, which differ remarkably from classical imaging in such a way that it recovers the depth information, i.e. the third dimension, by various techniques. In fact, it is a formidable task to emulate the human visual system in machines, since it requires a detail understanding of the imaging process. It finds extensive application in navigation and path planning of mobile robots, tracking and targeting in air missiles and defense systems, the manufacturing environment, disposal of toxic waste in nuclear power plants, etc. For accomplishing this task the robot has to be equipped with cameras to obtain visual information about its neighborhood.

13.2 Necessity for 3D Reconstruction

While programming the mobile robot to carry out navigational tasks autonomously, it is always necessary to explore the surroundings. Recently the problem of exploring an unknown environment has received considerable attention from the computer science and AI community. Before exploration, it is assumed that the environment is populated with a polygonal obstacle and the robot has to determine its position with respect to a global frame of reference. In practice, it is quite difficult to estimate precisely the position of a mobile robot with respect to an arbitrary frame of

reference. Therefore, in robotic systems an odometry system is used to determine the robot's global position but this suffers from the problem of cumulative errors as the robot moves further from its starting position. In order to avoid the cumulative errors, many localized systems require installing a set of beacons at known locations in the robot's workspace. But this is not viable for robots working in indoor environments. To do some complicated task, such as a pick and placement job in a flexible manufacturing environment, searching for leaking barrels of toxic waste in nuclear plants, surface moving where the surface is not planar and familiar, vision-based 3D exploration is necessary. However 3D scene recovery from 2D planar images remains a challenging task; even today, it is overwhelmed with problems.

13.3 Building Perception

Perception, as discussed in Chap. 1, constructs higher-level knowledge from relatively lower level data or knowledge. Generally the noise-free information is stored in LTM by the state of acquisition. The state of perception employs reasoning tools on the information recorded in LTM and thus derives new rules for subsequent planning and coordination problems. A mobile robot constructs its surrounding map by sensing information around it and preprocessing that information at the state of acquisition. In a 2D planning problem, the boundaries of the obstacles are generally sensed by ultrasonic sensors or laser range-finders. It is always assumed that the top of the obstacles is at a level higher than the mounting point of the sensors. A 3D planning problem, on the other hand, requires keeping track of the obstacle surfaces and their heights as well. To extend the 3D information, generally additional cameras are employed. These cameras are mounted on a pan-tilt platform, which is fixed with the mobile robot. When more than one camera are used for determining the third dimension of the obstacles, it is called stereo vision.

For constructing a 2D world map, the robot has to move around each obstacle. Starting from a given location the robot moves around each obstacle, until all obstacles are visited. A two-dimensional world map for the robot is then built up with the visited obstacles. If the sensory information recorded in the Long Term Memory (LTM) is not completely free from noise, then noise has to be eliminated first.