Chapter 20
3D Scene Capture and Analysis for Intelligent Robotics

Ray Jarvis
Monash University, Australia

ABSTRACT

The capability of robots to function effectively in the unstructured real world is dominated by the extent to which supporting sensory and computational resources can capture and analyse the 3D working environments within which they are to carry out their tasks. Many device technologies and computational algorithms have been developed over the last 30 years to enable such capabilities. This chapter chronicles a diverse range of such developments, comparing and contrasting them and indicating their various strengths and weaknesses to support intelligent robotic functionality in various domains of application.

INTRODUCTION

Intelligent Robotics is often defined as the melding of perception, reasoning and actuation or, more succinctly, 'sensor informed purposeful actuation' as embodied in a robotic system. Central to this field is the ability to capture and represent 3D scenes (and environments) and to reason over them to plan and execute useful physical action, whether in the domain of robotic manipulators or mobile robots (or both together).

DOI: 10.4018/978-1-61350-326-3.ch020

The ultimate, but currently impossible, ideal in 3D capture is to imagine a 'magic' powder which can be sprayed uniformly and densely over all surfaces (both external and internal) of objects in the scene and then sucked up into a heap with each particle remembering the location, texture and colour of the surface point it was on. Subsequent analysis would then consist of examining this heap for structures enabling the segmenting out of object surfaces, whole distinct objects, shapes and topology, and also, perhaps, the recognition of components against a pre-developed model data base. Elegant representations of extracted information would support efficient formulation of robotic action plans which may involve the collision-free manipulation of selected objects or efficient obstacle avoiding navigation through the environment to nominated goals or for pursuit/avoidance, hiding, rescue etc. operations. There are many ways in which 3D capture can be carried out using a wide variety of sensor technologies, including ultrasound, passive and active stereopsis, time-of-flight laser systems, silhouette, blur, brightness, attenuation, radar and x-ray. Likewise, there are many methodologies, which can be applied in the scene (environmental) analysis phase to answer such questions as the identity, placement, structural juxtaposition and accessibility of objects in the scene. Unfortunately, the 'magic' powder, as described above, does not exist, but many attempts have been made to approximate its capabilities. Many of these will be described in the chapter.

Once a suitable representation of the robot's working environment can be forged, even if only partially, a variety of motion planning algorithms (for trajectories when considering robotic manipulators and paths when considering mobile robots) supporting robotic hand/eye coordination or mobile robot navigation can be applied. In some circumstances 3D capture and planning are intermingled as in the case of a mobile robot simultaneously exploring and navigating its environment. At the opposite extreme of exploration in an initially unknown environment, it is sometimes possible and appropriate to pre-scan a whole working environment (e.g. shopping mall or city square etc.) in sufficient detail as to allow 'cyber navigation' of that space as a preliminary to replicating the plans with real robots in the real environment. Recently available large scale laser scanning technologies make such an approach very attractive when appropriate, since Virtual Reality like explorations can be used as part of the planning process and specific fixed objects and areas of interest annotated by hand with their functionality properties attached. Also, the localizing task can be reliably and accurately carried out by fusing live, on-board sensor data with the pre-collected model data using various scan matching approaches.

Carrying out a robotic task (once a plan has been devised) can, itself, be a special challenge, depending on the complexity (including dynamics) of the situation, the precision to which the robot can be controlled and its pose (location and orientation) determined, the means of locomotion itself (wheels, tracks, propellers, legs etc.), and the properties of the medium (e.g. underwater, in a vacuum etc.). Often plans need to be revised to accommodate new or changed information to compensate for error or when mission goals are modified. The determination of location and orientation (localization) is often an intensely sensor driven process, especially for mobile robot navigation in natural terrain which may also be initially unknown (not previously mapped or explored).

This chapter will cover all of the above subtopics but concentrate on the 3D capture and scene analysis components, covering many device technologies and analysis methodologies, their strengths and weaknesses and application domains in the context of intelligent robotics, providing many pictorial examples and an extensive bibliography to invite more detailed enquiry.

BACKGROUND

There have been many and varied attempts to provide knowledge of the working environment needed to guide robots in carrying out complex tasks in unstructured environments successfully, reliability and efficiently. Generally, the less structured the environment the more intelligence is required to achieve these goals and the richer must be the environmental knowledge to support this intelligence. In structured factory environments (e.g. an assembly line process) robotic actuation can follow precise pre-trained sequences to carry out the required tasks. Precision and speed are the