INCREMENTAL CONSTRUCTION OF 3-D MODELS FROM A SEQUENCE OF FRAMED VIEWS : MATCHING PARTIAL OBJECTS

Shun-en Xie and Thomas W. Calvert

LCCR, School of Computing Science, Simon Fraser University
Burnaby, B.C., V5A 1S6, Canada

ABSTRACT

In the future intelligent mobile robots will be called upon to play many important roles. In many realistic situations, the knowledge of the structure and placement of objects in an environment should be learned rather than built in. Thus the mobile robot must often construct 3-dimensional models for the objects by analysing sensed multiple views.

In this paper, we describe an approach to the incremental construction of 3-D body models in a practical office or warehouse environment by matching planned multiple views. In particular, we discuss the following aspects:

1. the decomposition of a framed view and the construction of partial 3-D descriptions of the view;
2. the matching of partial 3-D descriptions of a view with the built-in model of the robot environment;
3. the matching of partial descriptions of bodies derived from the current framed view with partial models constructed from previous views;
4. the identification of the new information in the current view and the updating of the models;
5. the identification of the unknown parts of the models which are being constructed so that further vantage viewpoints can be planned.

This approach combines such intelligent robot functions as attention, planning, sensing, learning and knowledge rectification. A prototype system for matching and constructing 3-D body models has been implemented and tested with synthesized images using C-PROLOG under Berkeley UNIX on a VAX 11/750.

INTRODUCTION

In the future computer-controlled robots will be called upon to play many important roles in industrial, business and domestic situations. If these robots are to work in complex environments it will be necessary to develop knowledge-based sensory systems. In simple situations, the robot vision system can have built-in models of both the environment and all objects within it; this allows a relatively simple recognition process. In more realistic situations, however, although the geometry of the surrounding environment may be known (i.e. the dimensions of the room, warehouse, etc, in which the robot operates), the type and position of the objects in the environment will generally be unknown. Thus knowledge of the structure and placement of these objects must be learned. To do this the mobile robot must first construct 3-dimensional models for the objects it encounters. It should then be possible to classify these objects by comparing their structural properties with those of generally known classes of objects such as benches, chairs, tables, etc.

In analysing a single framed view of part of a large scene, the problems which will generally stand in the way of constructing the 3-D body models include:

1. partial features;
2. self-occlusion;
3. occlusion;
4. accidental alignment and special alignment;
5. undetermined geometric parameters.

An approach to understanding a scene from image sequences by incrementally constructing body models seems promising. However, even to-day, the information processing load involved in analysing a sequence of images presents a serious technical problem. Dynamic selection of a minimal set of vantage viewpoints and effective selection of only the necessary information will be essential if the burden of computation is to be lightened. Fortunately, a mobile robot, by its nature, offers a good foundation for gathering information from different points of view. Thus combining a vision system with a planner, so that a scene can be analysed from planned multiple views, is both natural and necessary.

In this paper, we describe a system which incrementally constructs 3-D object models of an office or warehouse scene from planned multiple views. In particular, we address the matching and construction of 3-D partial models.

To limit the scope of the immediate research problem the following assumptions have been made:

1. The bodies in the environment are static, rigid, weakly externally visible, and have vertices formed by at most three surfaces. Edges are formed by two surfaces, which can be planar, conical, cylindrical or spherical.

This work was supported in part by grants from the NSERC of Canada. Mr. Xie was supported by the C. D. Nelson Memorial Scholarship.