Towards Sensor-Derived Models of Objects

F.P. Ferrie J. Lagarde P. Whaite

Computer Vision and Robotics laboratory
McGill Research Center for Intelligent Machines
McGill University
Montréal, CANADA

Abstract

In order to cope with an environment, a robot needs to know where things are located, about the size and shape of objects, and ultimately what the objects are. This paper is about computing such information and presents an overview of our current work in building geometric descriptions of objects from sensor data. Its particular focus is a computational methodology which addresses the constituent problems of reconstruction and parts decomposition.

1. Introduction

In order to cope with an environment, a robot needs to know where things are located, about the size and shape of objects, and ultimately what objects are. This paper is about computing such information and presents an overview of our current work in building geometric descriptions of objects from sensor data. The kinds of descriptions we seek are volumetric and part oriented. Objects are represented as conjunctions of volumetric elements, where each element is used to approximate the shape of its corresponding part [Ferrie 86; Ferrie & Levine 85,88; Pentland 86]. The idea is to proceed in bottom-up fashion using only general constraints about the structure of objects in computing these descriptions. Differential geometry can provide these constraints as well as serve as a unifying framework within which to address the constituent problems of reconstruction [Blake & Zisserman 87], inference [Ferrie & Levine 88] and multiple view integration [Ferrie & Levine 87].

In contrast, the current trend in the literature seems to be towards so-called model-based vision where information extracted from sensor data is used to index into a database of object descriptions [Grimson & Lozano-Perez 84; Bhanu 84; Bhanu 87; Bhanu & Ho 87; Bolles & Cain 82]. Presumably, once a description is matched to invariant features of sensor data, the necessary attributes of an object can be found by look-up. But there are some shortcomings with this approach, primarily the difficulty in creating generic object models [Biederman 85] and the complexity incurred by explicitly representing each instance of an object in the domain of observation [Grimson & Lozano-Perez 84]. While model-based vision does indeed work, we believe that there is more to be gained in the long run, both in terms of generality and insight into the larger vision task, by investigating the problem of how basic structure and shape information can be inferred from sensor data.

Our previous work focused on the problem of computing coarse descriptions of objects from intensity and range data [Ferrie & Levine 88]. By trading off loss of fine detail against smoothing, it was shown that it is possible to obtain coarse, but stable descriptions of the form shown in Figure 1b from the data shown in Figure 1a. While adequate for certain tasks, the earlier approach was limited in that it had no provision for evolving descriptions from coarse to fine, and was also limited to measurements from a single sensor. This paper describes an evolution of the earlier methodology which extends the computational framework to remove these limitations. By augmenting the local representation and applying minimization techniques, it is possible to reliably estimate features at finer resolution and combine estimates from different sensors. Recent work by Zucker and his co-workers [Zucker et al. 88] suggests that a similar approach can also be used in the aggregation of local features.

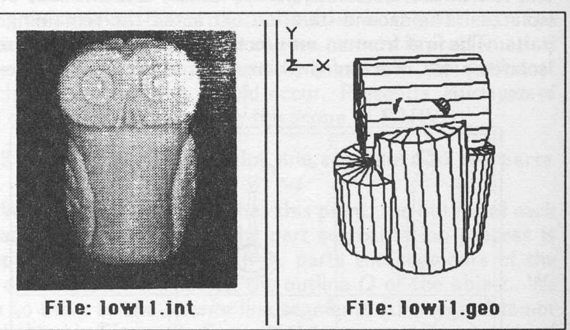

Figure 1a & 1b (a) Image of a stone owl sculpture (b) Rendition computed from (a) using simple volumetric primitives

We begin in Section 2 by outlining a paradigm developed earlier for deriving volumetric descriptions of objects from sensor data [Ferrie & Levine 88]. This is followed in Section 3 by a brief review of the mathematical tools used for computation as well as the representations we use for objects. The principal contribution of this paper follows next in Section 4 and describes the computational approach used to extract surface features and determine the parts decomposition. The testbed used for implementing our models and performing experiments is briefly described in Section 5 with the results of some experiments shown in Section 6.

2. Images to Surfaces to Objects: A Paradigm

Objects can be described at many different levels of abstraction, depending on what properties need to be made