Hierarchical Decomposition of Objects in Line Drawings

Robert Bergevin Martin D. Levine

Computer Vision and Robotics Laboratory
McGill Research Center for Intelligent Machines
McGill University
Montréal, Québec, Canada

Abstract

We describe an algorithm for segmenting a 3-D object into its constituent volumetric parts from a single 2-D line drawing. This algorithm is used at an intermediate stage of a generic object recognition system, the *Primal Access Recognition of Visual Objects* (PARVO) system. The latter is based on *Recognition by Components* (RBC), a theory of human image understanding from the field of psychology. A short description of the RBC theory is presented first. This is followed by an overview of the main aspects of a specific implementation of our PARVO system. A preliminary stage to the segmentation consists of extracting viewpoint-invariant and non-accidental features from this line drawing. The segmentation process itself is described as a hierarchical decomposition of the line drawing on the basis of pairs of high-concavity points in its silhouette. The pairing of these so-called segmentation points rests on (i) an object visibility model describing the generic structures present in any line drawing satisfying our assumptions, and (ii) the feature instances extracted from the specific line drawing analyzed.

Résumé

Nous décrivons un algorithme de segmentation d'un objet tri-dimensionnel en ses parties volumétriques constituantes à partir d'un seul dessin au trait bi-dimensionnel. Cet algorithme est intégré à un système de reconnaissance générique d'objets appellé PARVO (pour *Primal Access Recognition of Visual Objects* ou *Reconnaissance visuelle d'objets par accès primitif*). Ce système est lui-même basé sur une théorie psychologique de la compréhension d'images appellée *Recognition by Components* (RBC). Nous présentons d'abord une courte description de cette théorie suivi d'un survol des principaux aspects de notre implémentation sous la forme du système PARVO. Une étape préliminaire à la segmentation consiste à extraire du dessin des caractéristiques géométriques non-accidentelles et invariantes par rapport au point de vue. Le processus de segmentation lui-même est présenté comme une décomposition hiérarchique du dessin sur la base des paires de concavités situées sur sa silhouette. Un modèle générique des structures visibles du dessin ainsi que les caractéristiques géométriques extraites pour un dessin particulier sont utilisés à cet effet.

1. Introduction

The *Primal Access Recognition of Visual Objects* (PARVO) system is a generic object recognition system. It is based on some ideas expressed in *Recognition by Components* (RBC), a theory of human image understanding developed by I. Biederman [1], a researcher in the field of perceptual psychology. This theory aims at integrating part of the knowledge obtained from psychological and computational vision research into a consistent model describing how objects are visually recognized by humans.

The recognition process modelled by the RBC theory differs significantly from most existing artificial object recognition systems [2, 3, 4, 5] because it addresses the problem of *primal access*, that is, "the first contact of a perceptual input from an isolated, unanticipated object to a representation in memory" [1] (p.32). In other words, RBC models visual recognition of generic objects in unconstrained environments in the sense that no context is used in the analysis of the visual information. The primal access task which is addressed theoretically by RBC and computationally by the PARVO system is very different from what is known as *model-based* recognition in the field of computer vision. In the latter, objects may only be recognized as a result of a detailed inference over a limited set of accurate models. Primal access, on the other hand, relies only on qualitative and coarse models of objects and parts. A typical example of a primal access task is the recognition, by humans, of objects appearing in rapidly displayed photographic slides.

There are five basic principles behind the RBC theory: (i) a *line drawing* is sufficient for the primal access task, (ii) objects are better represented and analysed by decomposing them into their natural *components* (parts), (iii) a *qualitative description* of the components is necessary and sufficient to permit *fast access* to a *large database* of object models, (iv) *non-accidental* instances of *viewpoint-invariant features* in the 2-D line drawing are sufficient to permit fast access to the qualitative model of a 3-D object, (v) primal access for visual object recognition is obtained by *matching* a description of the spatial structure of the object components to an *indexed* database of models in a similar representation.

In this paper, we shall describe the part segmentation module of the PARVO system. The paper begins with a short summary of the RBC theory as reported in [1]. This is followed by an overview of the computational aspects of the PARVO system. Four main aspects are identified. They are: (i) *extraction of features* from a single view line drawing of an object, (ii) *segmentation* of this line drawing in conformity with the part connection structure of the visible object, (iii) *determination* of the *primitive volumetric shapes* each part is an instance of, and (iv) *identification* of the visible object as being one of the many known objects. The emphasis of this paper is on the second aspect. Details of the other aspects will be provided in forthcoming papers.