MULTIRESOLUTION MODELING AND MATCHING USING SALIENT FEATURES

Z.N. Li
School of Computing Science
Simon Fraser University
Burnaby, B.C. V5A 1S6

Q. Song
Department of Electrical Engineering and Computer Science
University of Wisconsin-Milwaukee
Milwaukee, WI 53201

ABSTRACT

Automatic recognition of industrial parts is vital to industrial automation and robotics. In this paper, we describe a program that recognizes 2D objects using salient features. The objects used to illustrate our results are ordinary room keys from a set of test images. Multiresolution images in a pyramid-like structure are used. At a coarse level the (standard) Hough transform is applied for hypothesizing orientations and locations of the objects. At lower levels in the pyramid the generalized Hough transform is applied for the matching of finer salient features. Methods for hierarchical modeling and hierarchical matching are presented. Control strategies in this hierarchical environment are also discussed.

KEYWORDS Hierarchical modeling, Hough transforms, Multiresolution, Object recognition, Pyramid, Salient features

1. Introduction

Automatic recognition of industrial parts is vital to industrial automation and robotics. In this paper, we describe a program that recognizes 2D objects using salient features. The objects used to illustrate our results are ordinary room keys from a set of test images. Multiresolution images in a pyramid-like structure are used. At a coarse level the (standard) Hough transform is applied for hypothesizing orientations and locations of the objects. At lower levels in the pyramid the generalized Hough transform is applied for the matching of finer salient features. Methods for hierarchical modeling and hierarchical matching are presented. Control strategies in this hierarchical environment are also discussed.

Model-based vision systems have been extensively studied by many researchers [1]. Among the many systems developed, Perkins [2] reported a model-based vision system for recognizing industrial parts. "Concurves" of parts are formed from "chain" data and then matched to models. Brooks [3] described model-based 3D interpretations of 2D images in the ACRONYM system. The system can recognize simple 3D objects such as electrical motors.

Hierarchical object recognition was pioneered by Kelly in his 1971 work of recognizing human faces [4]. Edges of faces are first located in a reduced resolution image. The edges found are then used as a "plan" for processing the original picture. In 1975, Tanimoto and Pavlidis [5] proposed a pyramid structure that supports multiresolution and hierarchical representation of image data. The basic pyramid structure has n levels, each level k ($0 \le k < n$) has $2^k \times 2^k$ nodes. Each node at level k is linked to its 13 neighbors, i.e., 1 parent, 8 siblings and 4 children. The connectivity between layers of nodes in the pyramid structure makes the hierarchical data flow (bottom-up and/or top-down) very efficient.

Recently, Neveu et al. [6] described a program for 2D object (room keys) recognition using multiresolution models. Objects are not represented by connected "chains" of points such as ribbons, skeletons, or concurves. Instead, simple sets of unconnected edge points are adopted in modeling. The program uses generalized Hough transform for all its matching processes. "The coarsest scale boundary description of the object is stored as the root node of the model graph and is matched first with the corresponding resolution level of an input image. The result of this match are used to hypothesize a list of approximate positions and orientations of the object." The program then uses higher resolution models to match finer boundary features. It decides the detection of objects according to the accumulated weights. The hierarchical modeling and the coarse-to-fine matching have been shown to speed up the recognition of room keys significantly. One of the advantages of using the generalized Hough transform is that the matching process is less sensitive to noisy and broken contours. Neveu's program follows a simple top-down search pattern. Its success relies heavily on the high level hypotheses. Since the top level image resolution is very low (16×16 in [6]), the generalized hough transform may not be able to yield good hypotheses. Moreover, a simple breadth-first search strategy is used to traverse the model graph composed of many features of interest for the object. It seems that an application of salient features would considerably reduce the search space for the domain of their objects (keys, bolts, washers).

We have been studying object recognition in the pyramid environment using key (salient) features [7, 8]. The application of salient features appears to be even more powerful for the recognition of 2D industrial parts which often