International Journal of Computer Vision 46(1), 5–23, 2002


# 3D Object Recognition in Cluttered Environments by Segment-Based Stereo Vision

YASUSHI SUMI, YOSHIHIRO KAWAI, TAKASHI YOSHIMI AND FUMIAKI TOMITA
*Intelligent Systems Institute, National Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Central 2, Tsukuba, Ibaraki 305-8568, Japan*
y.sumi@aist.go.jp

*Received May 14, 1999; Revised February 20, 2001; Accepted August 1, 2001*

**Abstract.** We propose a new method for 3D object recognition which uses segment-based stereo vision. An object is identified in a cluttered environment and its position and orientation (6 dof) are determined accurately enabling a robot to pick up the object and manipulate it. The object can be of any shape (planar figures, polyhedra, free-form objects) and partially occluded by other objects. Segment-based stereo vision is employed for 3D sensing. Both CAD-based and sensor-based object modeling subsystems are available. Matching is performed by calculating candidates for the object position and orientation using local features, verifying each candidate, and improving the accuracy of the position and orientation by an iteration method. Several experimental results are presented to demonstrate the usefulness of the proposed method.

**Keywords:** 3D object recognition, free-form objects, segment-based stereo vision, 3D object modeling, 3D shape matching, robot vision

## 1. Introduction

This paper addresses the problem of 3D model-based object recognition, that is, identifying a known object in a cluttered environment and determining its position and orientation (6 dof) accurately to enable a robot to pick up the object and manipulate it. Although this issue is a principal goal of machine vision and has a wide scope of potential applications, development of a practical object recognition system has been difficult and challenging. At least the following two basic requirements must be satisfied for general use:

- An object of any shape must be dealt with.
- Objects in any typically occurring environments must be dealt with.

For practical applications, the following are also important:

- Processing time must be short enough for the application tasks.
- Installation cost must be low enough for the application systems.

When designing a recognition system to meet all these requirements, it is very important to decide what type of sensor will be used. We have closely examined segment-based stereo vision for use as the sensor for 3D object recognition systems. Although stereo vision is a typical technique for sensing range from intensity images, it is not often employed in 3D object recognition because it has been considered inadequate for reconstructing the dense and accurate 3D data required for many tasks (Arman and Aggarwal, 1993b). However, stereo vision is suitable for object recognition tasks if it is designed ingeniously. The information required for object recognition is not a dense collection of 3D points, but rather a certain number of accurate and structuralized features. A segment-based stereo