A Probabilistic Approach for Object Recognition in a Real 3-D Office Environment

Michael Wünstel
Universität Bremen
Fachbereich Mathematik und Informatik
Postfach 330 440
28334 Bremen, Germany
wuenstel@informatik.uni-bremen.de

Thomas Röfer
Universität Bremen
Fachbereich Mathematik und Informatik
Postfach 330 440
28334 Bremen, Germany
roefer@informatik.uni-bremen.de

ABSTRACT

The scenario used focuses on object recognition in an office environment scene with the goal of classifying office equipment that is located on a table. The recognition system operates on three-dimensional point-clouds of objects on a loosely covered table where no previous information about the precise position of the table is given. As the point-clouds do not cover the complete objects and the data is noisy, especially for smaller objects a robust detection of special features is difficult.

The workflow employed is a three step process: In a first step the table plane is detected and the point clouds of the objects are extracted from the surface. In the second step an object-oriented bounding-box is calculated to get the geometric dimensions, i.e. the properties measured. During a learning phase these simple features are used to calculate the parameters of Bayesian networks. The trained networks are used in the third step, i.e. the classification step. The dimensions of an unknown object form the input for a Bayesian network that yields the most probable object type.

Keywords

Object Recognition, Cognitive Vision, Bayesian Network, Laser Range Data.

1. INTRODUCTION

There are many strategies for object recognition tasks that can be distinguished, e.g., concerning their perception process, namely object recognition strategies operating on 2-D density images or on 3-D range images. Range data has been used for object recognition for over two decades [1]. In this paper an approach is presented, in which the identification is done using just the dimensions of the objects together with general scene knowledge. This means on the one hand that certain objects correlate with certain dimensions, and on the other hand their presence correlates with the given scene. The recognition process is based on these simple but robust features as the extraction of three-dimensional special features is difficult due to occlusions, data noise, and the lack of measuring points especially for small objects.

Posters proceedings ISBN 80-86943-04-6
WSCG'2006, January 30-February 3, 2006
Plzen, Czech Republic.

2. System

A laser range finder that is mounted on a pan-tilt unit scans the scene. During a vertical scan of the pan-tilt unit several layers of scan points are collected and afterwards combined to form a three dimensional depth image. Afterwards the table plane is detected and the objects lying on it are segmented. In the next step an object oriented bounding-box is calculated that yields the dimension measures.

Sensor

As the laser sensor is mounted on a pan-tilt unit, it provides depth information from a viewpoint that is about 1.35 m above the ground (see Figure 1). The system has a horizontal resolution of 0.5° and the scanner is tilted in 0.5° respectively 1° steps.

Figure 1: Table scene with scanning equipment

Segmentation of Objects

As there is no specific information about the table, the exact position of the table top has to be determined separately. This information gives the vertical base level for the object segmentation step. For the plane detection the normal of every point is calculated [2] and connected regions are determined. The object segments are then calculated by a 3-D point density algorithm. This can be done from different positions (see Figure 2).

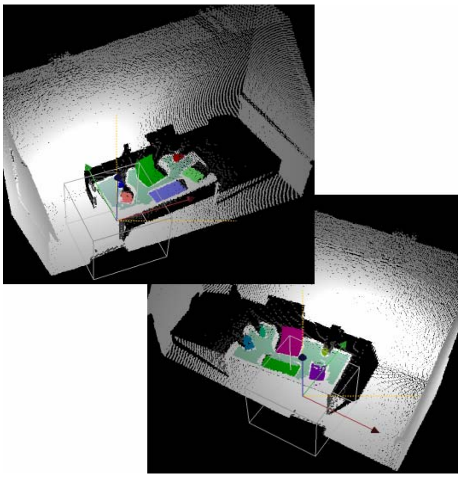

Figure 2: Detected table top and segmented objects from two different positions

Geometric Dimensions

The bounding-box that gives the measurement of an object is calculated iteratively using an optimization function that describes the quality of the actual bounding-box depending on two terms. The first (point-based) term describes the sum of the distances of the points to the bounding-box. The second (side-based) term describes the fitness for a subset of points to each separate side. Beginning with the axes-oriented bounding-box this function is minimized using the downhill simplex method (see Figure 3).

3. Recognition

The key method for the recognition process is the use of Bayesian networks:

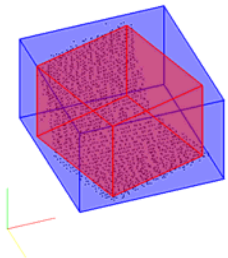

Figure 3: Initial axes-oriented bounding-box (red) and resulting object-oriented bounding-box (blue) for a notebook

Learning Phase

The dimensions of an object category are described by naive continuous Bayesian networks. Every object class corresponds to a single Bayesian Network that is parameterized by its mean and variance values. These values are determined analytically using the values from a training set.

Classification

Given the dimensions of an object, the grade of membership to each class is determined. Therefore the probability density function of each object class, resulting from the training phase, is integrated over an interval around the given evidence. The object is classified to the class with the highest probability value.

4. Results

The eight object classes used for the experiments are the following ones: book, bottle, coffeepot, keyboard, mug, notebook, phone, and hole-punch. The training set consists of ten different well-segmented shots of theses objects, the test set of five. For half of the objects the classification was always correct, for the other half up to two misclassifications occurred.

5. ACKNOWLEDGMENTS

This project is supported by the Deutsche Forschungsgemeinschaft, DFG, through the inter-disciplinary Transregional Collaborative Research Center “Spatial Cognition: Reasoning, Action, Interaction”.

Parts of this work go back to the work of Andreas Meyer. Other parts were developed with support of Christian Kanthak. Thanks for their effort.

6. REFERENCES

[1] Paul J. Besl and Ramesh C. Jain. Three-dimensional object recognition. ACM Computing Surveys, 17(1):75–145, March 1985.

[2] I. Stamos and P. Allen. 3-D Model Construction using Range and Image Data. CVPR-00, pages 531–536, 2000. IEEE.