ACTIVE OBJECT RECOGNITION THROUGH NEXT VIEW PLANNING

by

SUMANTRA DUTTA ROY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

Submitted

in fulfillment of the requirements of the degree of

Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY, DELHI

INDIA

AUGUST 2000

Certificate

This is to certify that the thesis titled **"Active Object Recognition through Next View Planning"** being submitted by Sumantra Dutta Roy to the Department of Computer Science and Engineering, Indian Institute of Technology, Delhi, for the award of the degree of Doctor of Philosophy, is a record of bona-fide research work carried out by him under our guidance and supervision. In our opinion, the thesis has reached the standards fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted to any other university or institute for the award of any degree or diploma.

Dr. S. Banerjee
Professor
Department of
Computer Science & Engineering
Indian Institute of Technology
New Delhi - 110 016

Dr. S. Chaudhury
Professor
Department of
Electrical Engineering
Indian Institute of Technology
New Delhi - 110 016

To my family

Acknowledgements

There are some debts which cannot be paid back. A student owes one of this kind to his or her supervisors. Hence, it is a futile attempt on my part to express my deep sense of gratitude to Prof. S. Banerjee and Prof. S. Chaudhury. Their deep insight into problems, imaginative ideas, zest for work and rigour have been a source of inspiration and motivation for me. They have been very generous with their time, and have been a constant source of support and encouragement.

I am extremely grateful to members of the SRC – Prof. H. M. Gupta, Prof. S. N. Maheshwari, Prof. K. K. Biswas, Dr. P. K. Kalra and Prof. Shashi Kumar, for their helpful suggestions and advice. I must mention Prof. J. B. Srivastava for all his help, advice and constant encouragement.

I am indebted to Dr. S. B. Pathak and Dr. A. K. Jain of the I. I. T. Hospital, without whose efforts and constant care, I would not have been able to see the light of this day.

A very special thanks to former and present Research Scholars for all their help, academic, and otherwise – Dr. Neena Pahuja, Dr. A. R. Naseer, Dr. R. N. Moorthy, Dr. R. P. Rustagi, Dr. Rekha Goel, Dr. Ragini Choudhury, Dr. N. Rajpal, Dr. R. Garg; and Mr. R. S. Jadon, Ms. Shoma Chatterjee, Mr. D. V. L. N. Somayajulu, Mr. M. A. Sarwat, Mr. B. G. Prasad, Ms. Mona Mathur, Mr. H. Ghosh, Mr. A. S. Mandal, Mr. S. C. Jain, Ms. K. A. Nagamani, Mr. A. Rajawat, Mr. S. Baswana, Mr. M. K. Jain, and Mr. R. M. Khandekar. I have been very fortunate in having very friendly and helpful people in all the laboratories I have worked in. I am particularly

grateful to the Technical staff of these laboratories – Mr. K. R. Kaushik, Mr. D. Jaitly, Mr. F. Singh, Mr. G. Singh, Mr. J. S. Rawat, Mr. P. C. Prasad, and Mr. G. Narain for bearing with the strangest of my demands, and at the oddest hours. I will be failing in my duty if I do not acknowledge the contribution of some very cheerful and cooperative staff in the office, laboratories, and the libraries.

My words fail me when it comes to my family members - my parents Prof. S. C. Dutta Roy and Dr. Sudipta Dutta Roy, brother Shoubhik, and furry wet-nosed Lucy. This thesis is dedicated to them.

New Delhi

Sumantra Dutta Roy

Abstract

Most model-based 3-D object recognition systems use information from a single view of an object. However, a single view may not contain sufficient features to recognize it unambiguously. Further, two objects may have all views in common with respect to a given feature set, and may be distinguished only through a sequence of views. A further complication arises when in an image, we do not have a complete view of an object. In this thesis, we propose two new on-line schemes for the recognition of an isolated 3-D object using reactive next view planning. The first one is based on aspect graphs, and the second is based on parts, and inner camera invariants. Both use an uncalibrated camera and simple features. We briefly describe them as follows:

Aspect Graph-based reactive object recognition An aspect constitutes a set of equivalent views of an object, with respect to a set of features. An aspect graph has nodes for an aspect, and links between adjacent aspects. We propose an integrated scheme for constructing aspect graphs from noisy feature detectors, and using them for recognizing isolated 3-D objects. Our aspect graph construction scheme accounts for errors in raw aspect data. Our system handles feature detection errors not only in the aspect graph construction process, but also in the object recognition stage, both of which use the same feature detectors. We propose a novel object recognition algorithm which uses the output of the aspect graph construction algorithm. It is not dependent on any specific feature set. The object recognition algorithm uses a probabilistic hypothesis generation mechanism. Our hierarchical knowledge representation scheme fa-

cilitates recognition and the planning process. The planning process is reactive - it utilizes the current observation and past history for identifying a sequence of moves to disambiguate between similar objects. We show results of aspect graph construction on more than 100 sets of noisy aspect data. The results of over 100 recognition experiments using the same noisy feature detectors, on two sets of models, show the effectiveness of our proposed scheme.

Part-based reactive object recognition For situations in which a complete view of a 3-D object may not be visible, we propose a new reactive, planning-based recognition algorithm which uses probabilistic reasoning. This algorithm is based on parts of an object and their relationships. We use a novel method of complete 3-D pose estimation using image-based measurements which are invariant to the internal parameters of a camera. While the earlier part of our work assumes an orthographic camera, we have formulated inner camera invariants for the more general projective case. For this problem also, we formulate a hierarchical knowledge representation scheme. The results of a large number of experiments with 2-D parts show the effectiveness of our approach in recognizing fairly complex 3-D objects, and localizing the camera with respect to the objects.

We demonstrate that by using reactive next view planning in conjunction with suitable representation schemes, it is possible to recognize 3-D objects with reasonably complex shapes, using simple features.

Contents