

Machine Perception

Ramakant Nevatia

MACHINE PERCEPTION

MACHINE PERCEPTION

RAMAKANT NEVATIA

*Departments of Electrical Engineering and Computer Science
University of Southern California*

Prentice-Hall, Inc.,

Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging in Publication Data

Nevatia, Ramakant.
Machine perception.

Bibliography: p.
Includes index.
I. Pattern recognition systems. I. Title.
TK7882.P3N48 1982 001.53'4 82-9136
ISBN 0-13-541904-2 AACR2

CONTENTS

Editorial/production supervisor: Nancy Milnamow
Cover design: Ray Lundgren
Manufacturing buyer: Joyce Levatino

© 1982 by Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632

All rights reserved. No part of this book
may be reproduced in any form or
by any means without permission in writing
from the publisher.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

ISBN 0-13-541904-2

Prentice-Hall International, Inc., *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte. Ltd., *Singapore*
Whitehall Books Limited, *Wellington, New Zealand*

1. INTRODUCTION	1
1.1 SOME EXAMPLES AND PROBLEMS	2
1.2 AN OVERVIEW AND HISTORY	5
1.3 RELATIONSHIP TO PSYCHOLOGY	9
1.4 PLAN OF THE BOOK	9
2. PATTERN CLASSIFICATION METHODS	12
2.1 DIGITAL REPRESENTATION OF AN IMAGE	12
2.1.1 Connectivity in Digital Images	13
2.2 TEMPLATE MATCHING	15
2.3 PATTERN CLASSIFICATION IN FEATURE SPACE	16
2.4 PERCEPTRONS	18
2.5 SYNTACTICAL APPROACHES	21
2.6 SUMMARY	22

3. SIMPLE POLYHEDRAL SCENES	24
3.1 PERCEPTION OF SIMPLE POLYHEDRAL SOLIDS	24
3.1.1 Extraction of Line Drawings	26
3.1.2 Model Matching	28
3.2 MODEL TRANSFORMATIONS	30
3.2.1 Perspective Transformations	31
3.2.2 Homogeneous Coordinates	32
3.2.3 Geometrical Transformations	34
3.3 FITTING OF MODELS	37
3.3.1 Camera Calibration	39
3.4 SUMMARY	40
4. COMPLEX SCENES OF POLYHEDRA	41
4.1 SEGMENTATION OF POLYHEDRAL SCENES	41
4.2 CLASSIFICATION OF LINES	45
4.2.1 Inclusion of Shadows	49
4.3 GEOMETRICAL CONSTRAINTS FOR POSSIBLE OBJECTS	51
4.3.1 Gradient Spaces and Dual Graphs	52
4.4 DESCRIPTIONS OF OBJECT ASSEMBLIES	55
4.4.1 Computation of Descriptions	56
4.4.2 Learning of Structural Descriptions	58
4.5 SUMMARY	59
5. SHAPE ANALYSIS AND RECOGNITION	61
5.1 REPRESENTATION OF COMPLEX SHAPES	62
5.2 LINE DESCRIPTIONS	63
5.2.1 Storage of Lines	63
5.2.2 Line Approximations	65
5.2.3 Analytical Line Shape Measures	66
5.3 AREA DESCRIPTIONS	67
5.3.1 Simple Shape Measures	67

5.3.2 Analytical Measures	69
5.3.3 A Medial Axis Transform	70
5.4 DESCRIPTIONS OF 3-D OBJECTS	72
5.4.1 Generalized Cones	73
5.4.2 Computation of Generalized Cones	75
5.5 RECOGNITION OF OBJECTS	80
5.5.1 Graph Matching	80
5.5.2 Relaxation Labeling	82
5.5.3 Multilevel Matching	85
5.6 SUMMARY	86
6. PERCEPTION OF BRIGHTNESS AND COLOR	90
6.1 BRIGHTNESS AND COLOR	91
6.1.1 The Human Eye	92
6.1.2 Local Measurements	92
6.2 LATERAL INHIBITION AND LIGHTNESS COMPUTATION	95
6.3 SUMMARY	98
7. EDGE AND CURVE DETECTION	100
7.1 EDGE DETECTION	101
7.1.1 Edge Enhancement and Differentiation	101
7.1.2 Edge Fitting	103
7.1.3 Edge Detection by Template Matching	105
7.1.4 Statistical Edge Detectors	111
7.1.5 Choice of Thresholds	111
7.2 RESULTS OF EDGE DETECTION	112
7.3 LINE AND CURVE DETECTION	116
7.3.1 Hough Transform	116
7.3.2 Graph-Theoretic Techniques	118
7.3.3 Method of Projections	119
7.4 CONTOUR FILLING AND FEEDBACK	120
7.5 COLOR EDGES	123
7.6 SUMMARY	124

8. REGION SEGMENTATION AND TEXTURE ANALYSIS	128
8.1 REGION SEGMENTATION	129
8.1.1 Thresholding and Recursive Segmentation	129
8.1.2 Region Growing	136
8.1.3 Semantically Guided Region Growing	137
8.1.4 Tracing Region Boundaries	138
8.2 EDGE VERSUS REGION SEGMENTATION	139
8.3 TEXTURE ANALYSIS	141
8.3.1 Statistical Texture Measures	144
8.3.2 Structural Texture Descriptions	149
8.3.3 Comparison of Texture Features	151
8.3.4 Texture Segmentation	152
8.4 SUMMARY	153
9. DEPTH MEASUREMENT AND ANALYSIS	158
9.1 STEREO AND MOTION	159
9.1.1 Correspondence Search	161
9.1.2 Global Correspondences	163
9.1.3 Motion Detection and Analysis	165
9.2 ACTIVE RANGING	167
9.2.1 Triangulation Ranging	167
9.2.2 LIDAR Ranging	170
9.3 SEGMENTATION USING RANGE	170
9.3.1 Boundary Detection	171
9.3.2 Detection of Planes and Surfaces	172
9.4 SHAPE FROM SHADING	173
9.4.1 Reflectance Maps	174
9.5 TEXTURE GRADIENTS	177
9.6 CONTOUR ANALYSIS	181
9.7 SUMMARY	183
10. SYSTEMS AND APPLICATIONS	187
10.1 GENERAL SYSTEMS	188
10.2 KNOWLEDGE BASED SYSTEMS	189

10.3 APPLICATIONS	193
10.3.1 Industrial Applications	194
10.3.2 Photo Interpretation and Change Detection	195
10.3.3 Guidance, Navigation, and Scene Registration	196
10.3.4 Medical Applications	197
10.3.5 Hardware Requirements	198
10.4 SUMMARY AND FUTURE	199
Index	203

PREFACE

The field of machine perception concerns the building of machines that sense and interpret their environments. This book is about visual perception. Potential applications for such systems include tasks such as automation of industrial processes of inspection and assembly, automated medical x-ray diagnosis, vehicle guidance and automatic photo-interpretation. Beginning from the analysis of simple polyhedral scenes in the early 1960s, the field has progressed to the point where useful analysis of complex natural and industrial scenes is possible and many practical prototype and commercial systems are available. Because of its immense potential applications, interest in this field has been growing rapidly.

This book is based on my experience in teaching graduate courses in the field at USC for several years. Similar but less detailed material has also been presented at one-week "short courses" intended for industry professionals. This book assumes no previous knowledge of the field and aims to provide a comprehensive knowledge of its methods. It is intended for use as a text for a one-semester graduate or senior-level course and also as a guide for the practicing professional.

Research literature in the field has multiplied but is scattered over many journals, conference proceedings, and research reports. While the active researchers seem to share much the same knowledge of previous work, a newcomer to the field has a difficult job in sorting out the vast literature. This book is aimed at easing this task by providing basic

concepts, details of the major approaches, and a guide to the literature.

The field of visual perception is still maturing. Some aspects of the problems are fairly well understood and have a well developed theory which is described in detail in the text. For other problems, however, comprehensive techniques do not exist, and the literature consists of a large number of methods of limited utility. In such cases I have grouped the techniques by their common themes and described the basic concepts. Varying amounts of detail are given for the specific techniques, based on my judgment of their generality and importance. I have tried to provide fairly complete references. To simplify the logistics of acquiring the pictures, I have provided examples from my own work or those of my students and colleagues, where applicable; similar examples could have been taken from others' work.

This book does not discuss computer programs in detail, but one must be familiar with digital computer programming in order to fully appreciate the difficulties of mechanizing the described processes. The mathematical content of the text is small, and for an overview the mathematical parts can be skipped. However, for a detailed understanding of certain topics, knowledge of a variety of mathematical tools would be helpful. These tools include calculus, analytical geometry, matrix theory and linear algebra, numerical analysis and graph theory. Knowledge of freshman-level physics may also help.

The facilities of the University of Southern California have been essential to the production of this book. Many of the examples were generated using the University's excellent laboratory and photographic facilities. A book in an active research field can hardly be written outside an active research environment. I am grateful for the support of the Defense Advanced Research Projects Agency (DARPA) of the Department of Defense for their support of my work at USC under their Image Understanding Program for many years (under contract numbers F33615-76-C-1203 and F33615-80-C-1080). The DARPA Image Understanding program was initiated by Lt. Col. David Carlstrom and later managed by Lt. Col. Lawrence E. Druffel and Cmdr. Ronald Ohlander.

Valuable comments on the drafts of the text were provided by Drs. Ruzena Bajcsy, Keith Price, and Barry Soroka and by many students at USC, especially David King. Permission of the various authors and publishers for use of their illustrations is gratefully acknowledged (credits are given in the text where used). The manuscript for this book was produced using SCRIBE text formatting system with a Graphics Systems Incorporated photocompositon device. Keith Uncapher made available the use of USC Information Science Institute's Xerox Penguin printer for "debugging" the early versions. Hilda Marti revised many drafts with her excellent typing and formatted

the text. The illustrations were drafted by Doyle Howland, and Ray Schmidt took the photographs.

Ramakant Nevatia
Los Angeles, California