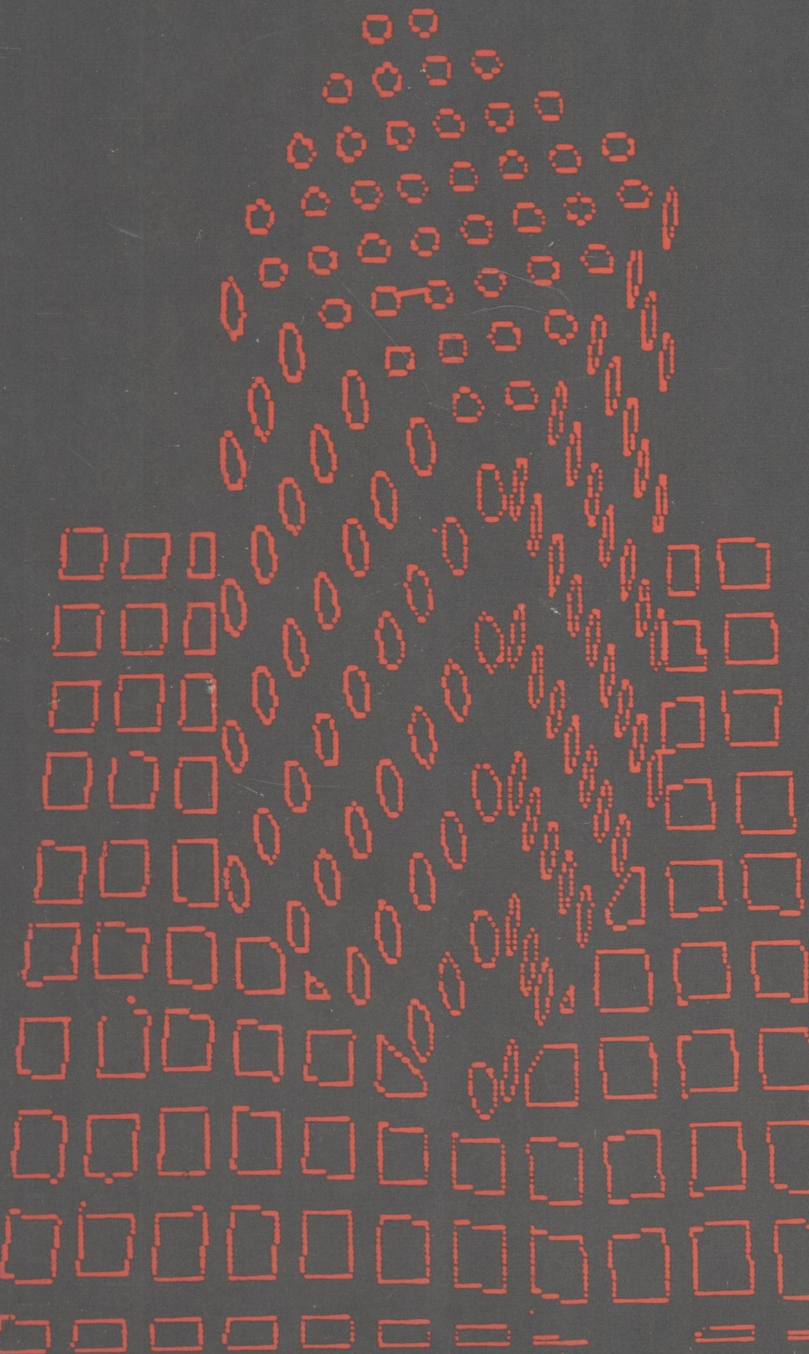
VISION IN MAN AND MACHINE

MARTIN D. LEVINE



VISION IN MAN AND MACHINE

Martin D. Levine

Professor of Electrical Engineering
McGill University, Montreal

Senior Fellow
The Canadian Institute for Advanced Research

McGraw-Hill Book Company

New York St. Louis San Francisco Auckland Bogotá Hamburg
Johannesburg London Madrid Mexico Montreal New Delhi
Panama Paris São Paulo Singapore Sydney Tokyo Toronto

This book was set in Times Roman.
The editors were Sanjeev Rao and James W. Bradley;
the production supervisor was Marietta Breitwieser.
The cover was designed by Joan E. O'Connor.
The drawings were done by J & R Services, Inc.
Halliday Lithograph Corporation was printer and binder.

VISION IN MAN AND MACHINE

1234567890 HALHAL 898765

ISBN 0-07-037446-5

Library of Congress Cataloging in Publication Data

Levine, Martin D., date
Vision in man and machine.

(McGraw-Hill series in electrical engineering.
Computer Engineering)
Includes bibliographies and index.
1. Image processing—Digital techniques. I. Title.
II. Series.
TA1632.L48 1985 001.53'4 84-21827
ISBN 0-07-037446-5

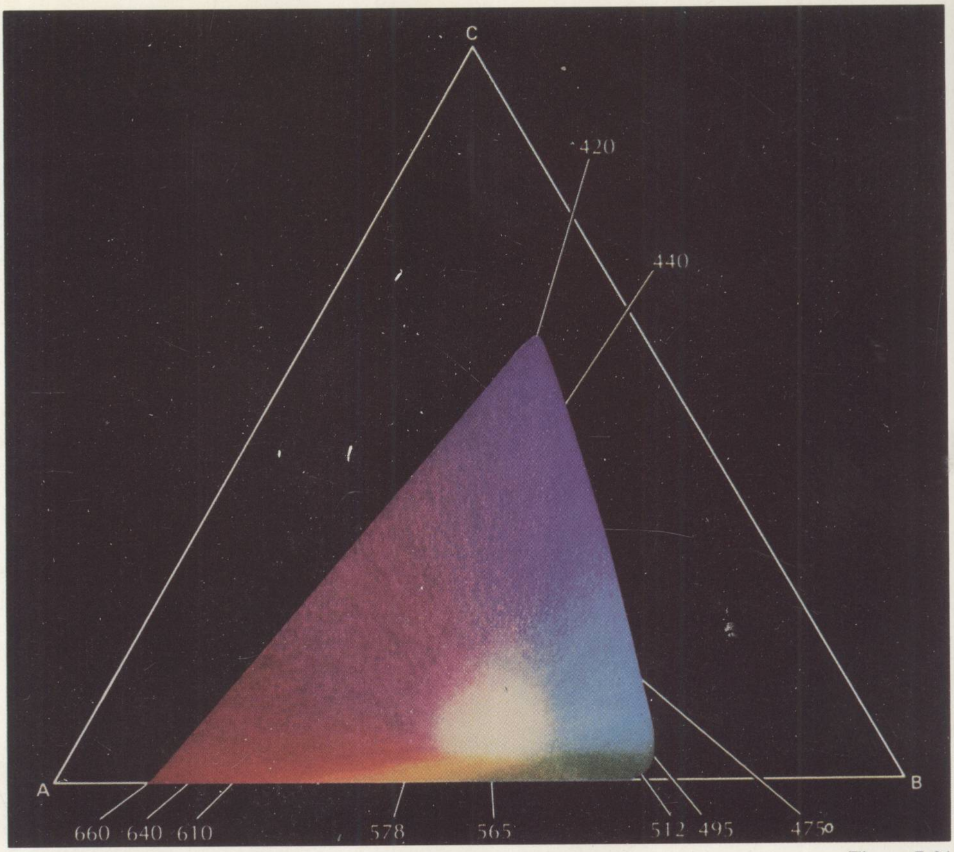


Figure 7.21

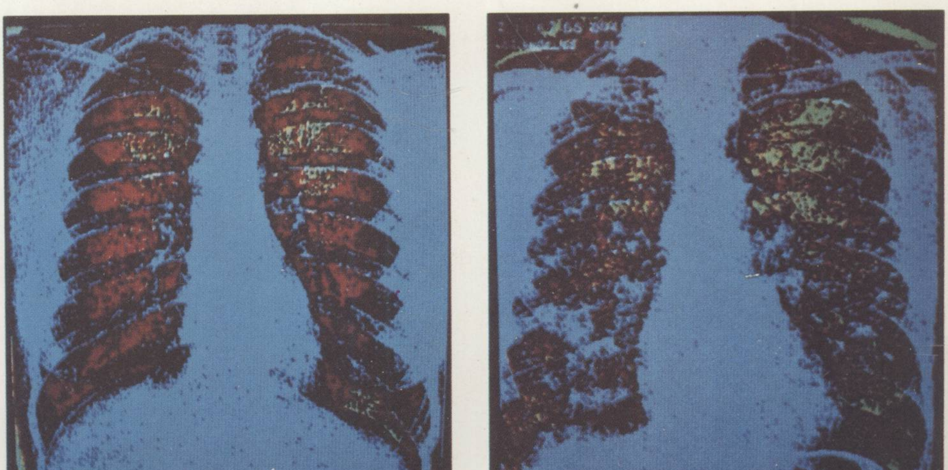
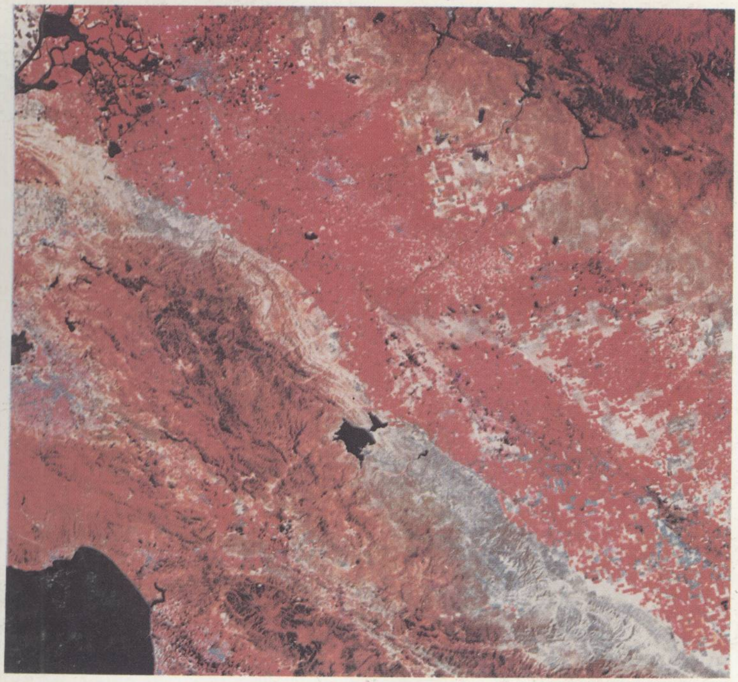
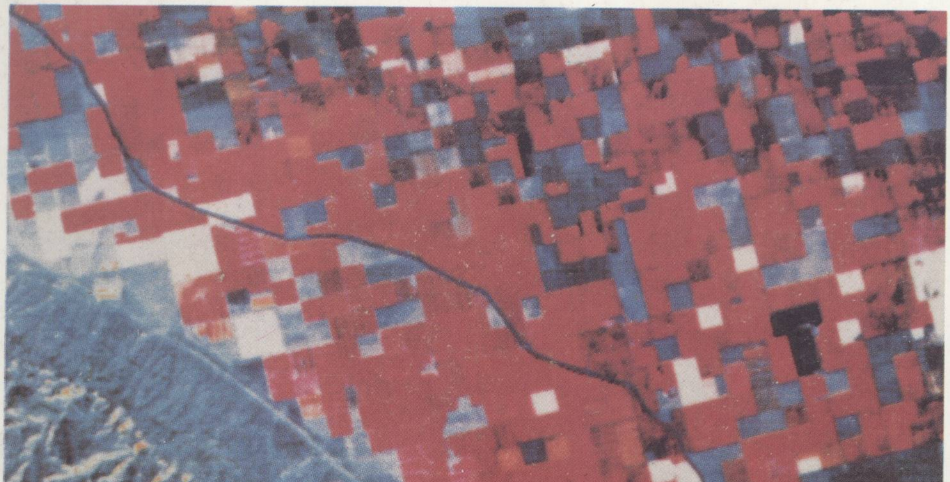


Figure 7.34



(*a*)



(*b*)

Figure 7.37

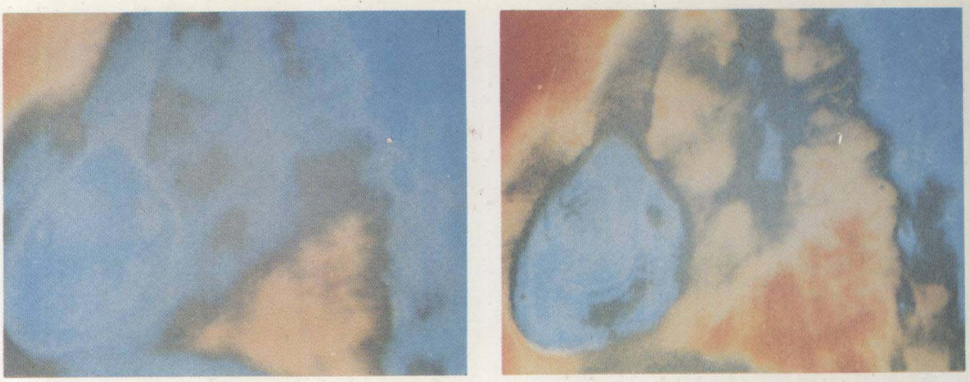
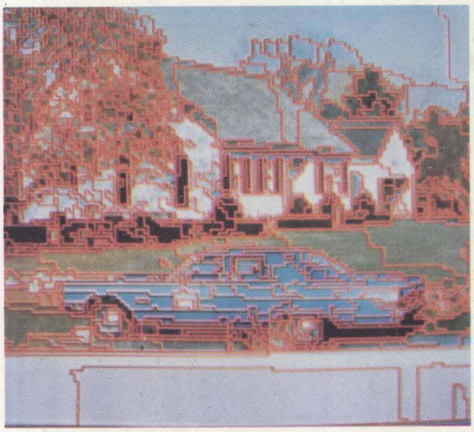
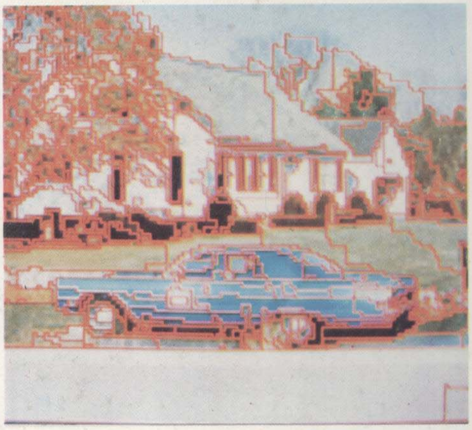


Figure 7.36



(a)



(b)

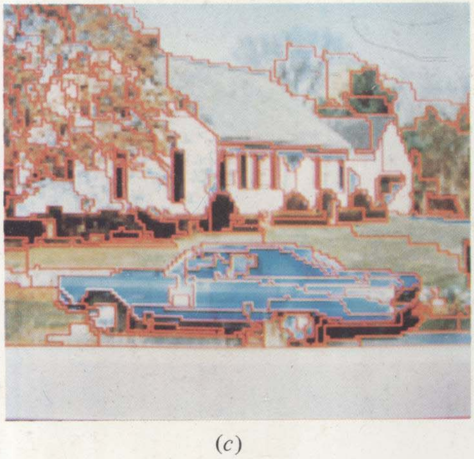


Figure 8.11

(c)

VISION IN MAN AND MACHINE

McGraw-Hill Series in Electrical Engineering

Consulting Editor
Stephen W. Director, Carnegie-Mellon University

Networks and Systems
Communications and Information Theory
Control Theory
Electronics and Electronic Circuits
Power and Energy
Electromagnetics
Computer Engineering
Introductory and Survey
Radio, Television, Radar, and Antennas

Previous Consulting Editors

Ronald M. Bracewell, Colin Cherry, James F. Gibbons, Willis W. Harman, Hubert Heffner, Edward W. Herold, John G. Linvill, Simon Ramo, Ronald A. Rohrer, Anthony E. Siegman, Charles Susskind, Frederick E. Terman, John G. Truxal, Ernst Weber, and John R. Whinnery

Computer Engineering

Consulting Editor
Stephen W. Director, Carnegie-Mellon University

Bell and Newell: *Computer Structures: Readings and Examples*
Garland: *Introduction to Microprocessor System Design*
Gault and Pimmel: *Introduction to Microcomputer-Based Digital Systems*
Givone: *Introduction to Switching Circuit Theory*
Givone and Roesser: *Microprocessors/Microcomputers: An Introduction*
Hamacher, Vranesic, and Zaky: *Computer Organization*
Hayes: *Computer Organization and Architecture*
Kohavi: *Switching and Finite Automata Theory*
Levine: *Vision in Man and Machine*
Peatman: *Design of Digital Systems*
Peatman: *Digital Hardware Design*
Peatman: *Microcomputer Based Design*
Sandige: *Digital Concepts Using Standard Integrated Circuits*
Taub: *Digital Circuits and Microprocessors*
Wiatrowski and House: *Logic Circuits and Microcomputer Systems*

This book is dedicated
to Debbie, Jonathan, and Barbara

. . . it is marvellous in our eyes . . .

Psalm 118, Verse 23

PREFACE

My interest in computer vision began about 1967, being particularly attracted by the possibility of applying these techniques to biomedical imagery. Indeed the first project I worked on involved the automated quantification of human alveolar structure in histological sections. I later spent the 1972–1973 academic year with the Image Analysis Group at the Jet Propulsion Laboratory in Pasadena, California, working on the vision system for a Martian robot. Since then I have been teaching a course on digital picture processing in the Electrical Engineering Department of McGill University in Montreal, Canada. Although this course was at the undergraduate/graduate level, the majority of the students were in fact undergraduates. The two basic prerequisites are introductory courses in computers and linear systems.

The course is the first of a two-stage sequence, in which the initial semester deals with the introductory aspects of the subject and the second follows with the more advanced and research-related topics. One way of characterizing this dichotomization is to say that the first course considers low-level vision while the second concerns itself with high-level vision. In fact, one of the goals of this book is to attempt to delineate these two processes. I suggest that low-level vision essentially deals with the measurement of the attributes of a picture, their organization into a data structure, and possibly some simple aggregative feature descriptions of this data structure. High-level vision is concerned with the interpretation of these data.

Two possible approaches may be taken to using this book as a text in a course on picture processing. The first does not place great emphasis on human and animal vision. It uses this material only for additional reading or project assignments. A typical course might cover the following: Chapter 1; Chapter 2; Sections 5.1, 5.3, 5.4, 5.5; Sections 6.1, 6.4, 6.5, 6.6; Sections 7.1, 7.3, 7.4, 7.5; Sections 8.1, 8.3, 8.4; Chapter 9; Chapter 10. An alternative treatment is to cover Chapters 1 to 7 in detail, and assign reading or projects from Chapters 8, 9, and 10. The Appendix contains an extensive set of both library and computer projects. In addition, throughout the book I have included references and bibliographies to assist both the student and the researcher.

Everyone agrees that the problem of programming a computer to analyze and, what is more, to understand the content of pictures is extremely difficult. My view is that if we are expected to write algorithms to achieve these goals, it is incumbent upon us to know how humans and animals achieve this same function. Therefore one of the major objectives of this book is to show the relevance of biological models to the design of engineering systems. They can give us hints about the trade-offs required and point out potentially interesting technical problems. Obviously the human visual system acts in the role of an "existence proof" by specifying to us at least those problems that we know can be solved. It also indicates the operative constraints, thereby challenging the designer of computer systems to overcome them. Recently, there has been a growing tendency for scientists who work with computer and human visual systems to read and study each other's literature. In my opinion, this interdisciplinary trend warrants encouragement.

The study of "living" visual systems may be considered from the point of view of either neurophysiology or psychology. Neurophysiologists are generally concerned with the behavior of single cells or small groups of cells. Psychologists who are concerned with visual perception generally treat visual processing as if it were achieved by the proverbial engineering "black box." By this is meant that the "hardware" details of the response of a human or animal to visual inputs is generally not of great concern. Perhaps this is not always the case, but we do realize that realistically it is very difficult for us to provide a linkage between a given perceptual event and its neurophysiological correlate.

If we agree that the computer scientist or engineer who will write programs for computer vision should be aware of the current research and results in these two fields, we may then ask which material should one emphasize. Throughout the period I have been teaching this course I have naturally selected certain subjects and details which I thought were both interesting and relevant. During the 1979–1980 academic year I had the opportunity to go on sabbatical at the Computer Science Department of the Hebrew University in Jerusalem, Israel. I took the time to reread some of the literature and also to rethink my ideas about what material was or was not important. One thing is absolutely clear—the amount of literature on the subject in both these areas is overwhelming! I have therefore decided to consider the complete process of computer vision in the biological sense from the point of view of systems modeling. That is, suppose for example that the human visual system were a hardware or computer system made up of elemental electronic components. Also suppose that I were able to make "input-output" measurements on this system. The obvious analogy to neurophysiological experimentation with cells and the psychophysical experimental paradigm can be seen. Then in order for the material to be considered for inclusion in this book, there should exist what I have called "correlative models" in the two fields. Thus I expect that since the psychophysical measurement is in the final analysis the result of the behavior of networks of cells, I should be able to corroborate the activity of the former by measurements related to the latter. How difficult this is to accomplish will immediately become evident to the reader if one contemplates pursuing this same type of analysis with the original computer analogy. Where such models exist, my task has been easy or relatively easy. This approach provides a very broad filter, which tends to exclude a large percentage of the literature in both fields. However, I have not been able to maintain such an uncompromising attitude throughout. In certain cases I have selected material which I thought had the best probability of being corroborated in one field by the other.

Readers may find it presumptuous on my part to be writing about psychology and neurophysiology, two subject areas in which I am admittedly a neophyte. I beg their forgiveness. My only excuse is that this book is primarily intended for computer scientists and engineers and is meant to provide them with some introductory background material about the biological process of vision. These individuals understand and love block diagrams, and I have therefore tried to present the subject using such a systems approach. This is generally not the conventional method of presentation found in most psychology and neurophysiology sources.

Throughout the book I have attempted to maintain a parallelism between vision in man and machine. This is not to suggest that the mechanisms in both cases are similar. I have generally left it to the readers to reach their own conclusions about the similarities and differences since usually these are quite obvious. Given the two primary themes, the book may be conceptually divided into three major parts. The first one involves transforming a given scene or image into a representation internal to the biological or computer system. The second is concerned with measuring the features of this stimulus input. In particular I discuss the attributes of edge content and color. The third section, the most hypothetical in nature from the point of view of living systems, deals with the organization and description of the low-level feature data.

In a certain sense the human visual system acts as an upper bound on our ambitions with regard to writing computer algorithms for picture processing. The more I have studied vision by man, the more fascinating both its intricacy and functionality have become to me. Each new scientific discovery which provides a corroborative model is exciting because generally it also seems to make sense from the point of view of information processing by computer. I hope that I have conveyed some of this enthusiasm to the reader.

Many people have assisted me in writing and preparing this book. At the outset I would like to thank George Nagy for introducing me to the subject of picture processing in the late sixties. He was very instrumental in influencing me to change my research orientation from the field of control theory. More recently, my colleague Steve Zucker, who has a zealous interest in animal vision and human perception, has intensified my own concern for the relationship between human and computer vision.

Several students have helped me by writing programs, preparing data and picture output, and carrying out library research. In this regard, I would like to thank Christian De Keresztes, Margaret Dalziel, Andreas Dill, Cem Eskenazi, Frank Ferrie, Wade Hong, Harold Hubschman, Benjie Kimia, John Lloyd, Gail McCartney, Ahmed Nazif, Samir Shaheen, and Youssry Youssef. Without their assistance this book would not have been possible.

Various versions of the manuscript have been read and criticized, a rather odious task to request of a colleague. Thanks are due Masayoshi Aoki, Jake Aggarwal, Paul Mermelstein, Naftali Minsky, Peter Sander, Dov Rosenfeld, Herbert Freeman, K. S. Fu, R. Rajesy, Azriel Rosenfeld, Arthur Sanderson, and Harry Wechjlek. A special appreciation is also due Peter Sander, who painstakingly obtained the permissions for each of the cited figures and arranged to have them photographed. Apropos of the latter, I would like to thank Rolf Selbach of the Instructional Communications Centre at McGill for taking the many pictures I needed for the figures.

I would also like to acknowledge the aid of Fran Lew, Prema Menon, and Heather Roberts in typing and preparing the manuscript.

Finally, it is my pleasure to give particular credit to Cem Eskenazi and Fran Lew, the former my graduate student and the latter my secretary. Without their intense physical effort, their dedication, and their encouragement I would not have been able to complete the manuscript on schedule. I am truly grateful for their involvement with this book.

Martin D. Levine

VISION IN MAN AND MACHINE

CONTENTS