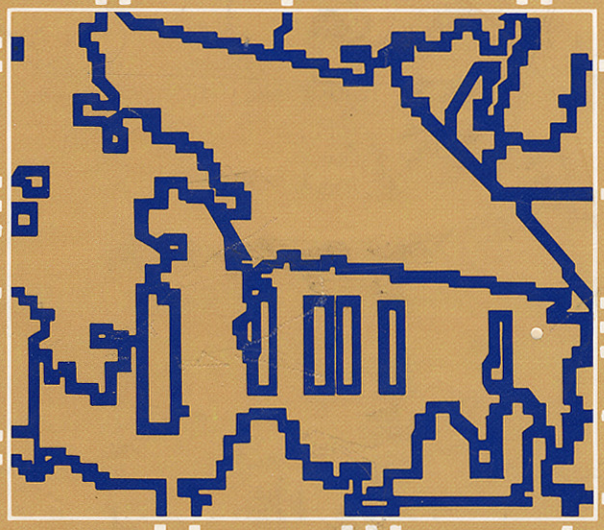
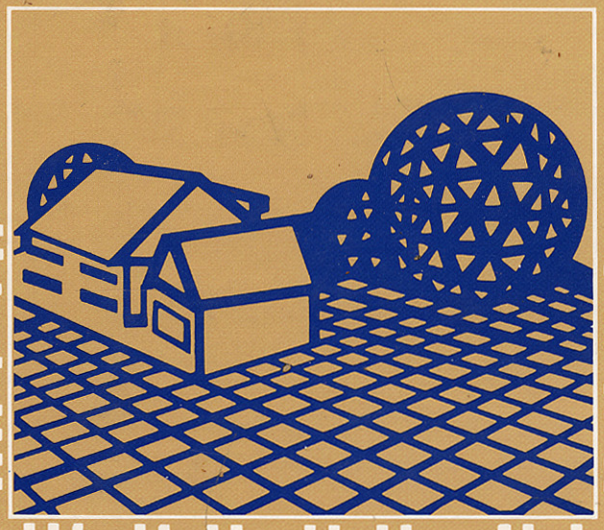
COMPUTER VISION

DANA H. BALLARD • CHRISTOPHER M. BROWN





COMPUTER VISION

Dana H. Ballard
Christopher M. Brown

Department of Computer Science
University of Rochester
Rochester, New York

PRENTICE-HALL, INC., Englewood Cliffs, New Jersey 07632

Library of Congress Cataloging in Publication Data

BALLARD, DANA HARRY.
Computer vision.

Bibliography: p.
Includes index.
1. Image processing. I. Brown, Christopher M.
II. Title.
TA1632.B34 621.38'0414 81-20974
ISBN 0-13-165316-4 AACR2

Cover design by Robin Breite

Printed in the United States of America

10 9 8 7 6 5 4 3 2

ISBN 0-13-165316-4

PRENTICE-HALL INTERNATIONAL, INC., *London*
PRENTICE-HALL OF AUSTRALIA PTY. LIMITED, *Sydney*
PRENTICE-HALL OF CANADA, LTD., *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*

Preface

The dream of intelligent automata goes back to antiquity; its first major articulation in the context of digital computers was by Turing around 1950. Since then, this dream has been pursued primarily by workers in the field of *artificial intelligence,* whose goal is to endow computers with information-processing capabilities comparable to those of biological organisms. From the outset, one of the goals of artificial intelligence has been to equip machines with the capability of dealing with sensory inputs.

Computer vision is the construction of explicit, meaningful descriptions of physical objects from images. Image understanding is very different from image processing, which studies image-to-image transformations, not explicit description building. Descriptions are a prerequisite for recognizing, manipulating, and thinking about objects.

We perceive a world of coherent three-dimensional objects with many invariant properties. Objectively, the incoming visual data do not exhibit corresponding coherence or invariance; they contain much irrelevant or even misleading variation. Somehow our visual system, from the retinal to cognitive levels, understands, or imposes order on, chaotic visual input. It does so by using *intrinsic information* that may reliably be extracted from the input, and also through assumptions and *knowledge* that are applied at various levels in visual processing.

The challenge of computer vision is one of *explicitness.* Exactly what information about scenes can be extracted from an image using only very basic assumptions about physics and optics? Explicitly, what computations must be performed? Then, at what stage must domain-dependent, prior knowledge about the world be incorporated into the understanding process? How are world models and knowledge represented and used? This book is about the representations and mechanisms that allow image information and prior knowledge to interact in image understanding.

Computer vision is a relatively new and fast-growing field. The first experiments were conducted in the late 1950s, and many of the essential concepts

have been developed during the last five years. With this rapid growth, crucial ideas have arisen in disparate areas such as artificial intelligence, psychology, computer graphics, and image processing. Our intent is to assemble a selection of this material in a form that will serve both as a senior/graduate-level academic text and as a useful reference to those building vision systems. This book has a strong artificial intelligence flavor, and we hope this will provoke thought. We believe that both the intrinsic image information and the internal model of the world are important in successful vision systems.

The book is organized into four parts, based on descriptions of objects at four different levels of abstraction.

1. Generalized images—images and image-like entities.
2. Segmented images—images organized into subimages that are likely to correspond to "interesting objects."
3. Geometric structures—quantitative models of image and world structures.
4. Relational structures—complex symbolic descriptions of image and world structures.

The parts follow a progression of increasing abstractness. Although the four parts are most naturally studied in succession, they are not tightly interdependent. Part I is a prerequisite for Part II, but Parts III and IV can be read independently.

Parts of the book assume some mathematical and computing background (calculus, linear algebra, data structures, numerical methods). However, throughout the book mathematical rigor takes a backseat to concepts. Our intent is to transmit a set of ideas about a new field to the widest possible audience.

In one book it is impossible to do justice to the scope and depth of prior work in computer vision. Further, we realize that in a fast-developing field, the rapid influx of new ideas will continue. We hope that our readers will be challenged to think, criticize, read further, and quickly go beyond the confines of this volume.

D. H. Ballard
C. M. Brown

Acknowledgments

Jerry Feldman and Herb Voelcker (and through them the University of Rochester) provided many resources for this work. One of the most important was a capable and forgiving staff (secretarial, technical, and administrative). For massive text editing, valuable advice, and good humor we are especially grateful to Rose Peet. Peggy Meeker, Jill Orioli, and Beth Zimmerman all helped at various stages.

Several colleagues made suggestions on early drafts: thanks to James Allen, Norm Badler, Larry Davis, Takeo Kanade, John Kender, Daryl Lawton, Joseph O'Rourke, Ari Requicha, Ed Riseman, Azriel Rosenfeld, Mike Schneier, Ken Sloan, Steve Tanimoto, Marty Tenenbaum, and Steve Zucker.

Graduate students helped in many different ways: thanks especially to Michel Denber, Alan Frisch, Lydia Hrechanyk, Mark Kahrs, Keith Lantz, Joe Maleson, Lee Moore, Mark Peairs, Don Perlis, Rick Rashid, Dan Russell, Dan Sabbah, Bob Schudy, Peter Selfridge, Uri Shani, and Bob Tilove. Bernhard Stuth deserves special mention for much careful and critical reading.

Finally, thanks go to Jane Ballard, mostly for standing steadfast through the cycles of elation and depression and for numerous engineering-to-English translations.

As Pat Winston put it: "A willingness to help is not an implied endorsement." The aid of others was invaluable, but we alone are responsible for the opinions, technical details, and faults of this book.

Funding assistance was provided by the Sloan Foundation under Grant 78-4-15, by the National Institutes of Health under Grant HL21253, and by the Defense Advanced Research Projects Agency under Grant N00014-78-C-0164.

The authors wish to credit the following sources for figures and tables. For complete citations given here in abbreviated form (as "from . . ." or "after . . ."), refer to the appropriate chapter-end references.

Fig. 1.2 from Shani, U., "A 3-D model-driven system for the recognition of abdominal anatomy from CT scans," TR77, Dept. of Computer Science, University of Rochester, May 1980.

Fig. 1.4 courtesy of Allen Hanson and Ed Riseman, COINS Research Project, University of Massachusetts, Amherst, MA.
Fig. 2.4 after Horn and Sjoberg, 1978.
Figs. 2.5, 2.9, 2.10, 3.2, 3.6, and 3.7 courtesy of Bill Lampeter.
Fig. 2.7a painting by Louis Condax; courtesy of Eastman Kodak Company and the Optical Society of America.
Fig. 2.8a courtesy of D. Greenberg and G. Joblove, Cornell Program of Computer Graphics.
Fig. 2.8b courtesy of Tom Check.
Table 2.3 after Gonzalez and Wintz, 1977.
Fig. 2.18 courtesy of EROS Data Center, Sioux Falls, SD.
Figs. 2.19 and 2.20 from Herrick, C.N., *Television Theory and Servicing: Black/White and Color,* 2nd Ed. Reston, VA: Reston, 1976.
Figs. 2.21, 2.22, 2.23, and 2.24 courtesy of Michel Denber.
Fig. 2.25 from Popplestone et al., 1975.
Fig. 2.26 courtesy of Production Automation Project, University of Rochester.
Fig. 2.27 from Waag and Gramiak, 1976.
Fig. 3.1 courtesy of Marty Tenenbaum.
Fig. 3.8 after Horn, 1974.
Figs. 3.14 and 3.15 after Frei and Chen, 1977.
Figs. 3.17 and 3.18 from Zucker, S.W. and R.A. Hummel, "An optimal 3-D edge operator," *IEEE Trans. PAMI 3,* May 1981, pp. 324-331.
Fig. 3.19 curves are based on data in Abdou, 1978.
Figs. 3.20, 3.21, and 3.22 from Prager, J.M., "Extracting and labeling boundary segments in natural scenes," *IEEE Tans. PAMI 12,* 1, January 1980. © 1980 IEEE.
Figs. 3.23, 3.28, 3.29, and 3.30 courtesy of Berthold Horn.
Figs. 3.24 and 3.26 from Marr, D. and T. Poggio, "Cooperative computation of stereo disparity," *Science,* Vol. 194, 1976, pp. 283-287. © 1976 by the American Association for the Advancement of Science.
Fig. 3.31 from Woodham, R.J., "Photometric stereo: A reflectance map technique for determining surface orientation from image intensity," *Proc. SPIE,* Vol. 155, August 1978.
Figs. 3.33 and 3.34 after Horn and Schunck, 1980.
Fig. 3.37 from Tanimoto, S. and T. Pavlidis, "A hierarchical data structure for picture processing," *CGIP 4,* 2, June 1975, pp. 104–119.
Fig. 4.6 from Kimme et al., 1975.
Figs. 4.7 and 4.16 from Ballard and Sklansky, 1976.
Fig. 4.9 courtesy of Dana Ballard and Ken Sloan.
Figs. 4.12 and 4.13 from Ramer, U., "Extraction of line structures from photgraphs of curved objects," *CGIP 4,* 2, June 1975, pp. 81–103.
Fig. 4.14 courtesy of Jim Lester, Tufts/New England Medical Center.
Fig. 4.17 from Chien, Y.P. and K.S. Fu, "A decision function method for boundary detection," *CGIP 3,* 2, June 1974, pp. 125–140.
Fig. 5.3 from Ohlander, R., K. Price, and D.R. Reddy, "Picture segmentation using a recursive region splitting method," *CGIP 8,* 3, December 1979.
Fig. 5.4 courtesy of Sam Kapilivsky.
Figs. 6.1, 11.16, and A1.13 courtesy of Chris Brown.
Fig. 6.3 courtesy of Joe Maleson and John Kender.
Fig. 6.4 from Connors, 1979. Texture images by Phil Brodatz, in Brodatz, *Textures.* New York: Dover, 1966.
Fig. 6.9 texture image by Phil Brodatz, in Brodatz, *Textures.* New York: Dover, 1966.
Figs. 6.11, 6.12, and 6.13 from Lu, S.Y. and K.S. Fu, "A syntactic approach to texture analysis," *CGIP 7,* 3, June 1978, pp. 303-330.

Fig. 6.14 from Jayaramamurthy, S.N., "Multilevel array grammars for generating texture scenes," *Proc. PRIP,* August 1979, pp. 391-398. © 1979 IEEE.
Fig. 6.20 from Laws, 1980.
Figs. 6.21 and 6.22 from Maleson et al., 1977.
Fig. 6.23 courtesy of Joe Maleson.
Figs. 7.1 and 7.3 courtesy of Daryl Lawton.
Fig. 7.2 after Prager, 1979.
Figs. 7.4 and 7.5 from Clocksin, W.F., "Computer prediction of visual thresholds for surface slant and edge detection from optical flow fields," Ph.D. dissertation, University of Edinburgh, 1980.
Fig. 7.7 courtesy of Steve Barnard and Bill Thompson.
Figs. 7.8 and 7.9 from Rashid, 1980.
Fig. 7.10 courtesy of Joseph O'Rourke.
Figs. 7.11 and 7.12 after Aggarwal and Duda, 1975.
Fig. 7.13 courtesy of Hans-Hellmut Nagel.
Fig. 8.1d after Requicha, 1977.
Figs. 8.2, 8.3, 8.21a, 8.22, and 8.26 after Pavlidis, 1977.
Figs. 8.10, 8.11, 9.6, and 9.16 courtesy of Uri Shani.
Figs. 8.12, 8.13, 8.14, 8.15, and 8.16 from Ballard, 1981.
Fig. 8.21 b from Preston, K., Jr., M.J.B. Duff, S. Levialdi, P.E. Norgren, and J-i. Toriwaki, "Basics of cellular logic with some applications in medical image processing," *Proc. IEEE,* Vol. 67, No. 5, May 1979, pp. 826–856.
Figs. 8.25, 9.8, 9.9, 9.10, and 11.3 courtesy of Robert Schudy.
Fig. 8.29 after Bribiesca and Guzman, 1979.
Figs. 9.1, 9.18, 9.19, and 9.27 courtesy of Ari Requicha.
Fig. 9.2 from Requicha, A.A.G., "Representations for rigid solids: theory, methods, systems," *Computer Surveys 12,* 4, December 1980.
Fig. 9.3 courtesy of Lydia Hrechanyk.
Figs. 9.4 and 9.5 after Baumgart, 1972.
Fig. 9.7 courtesy of Peter Selfridge.
Fig. 9.11 after Requicha, 1980.
Figs. 9.14 and 9.15b from Agin, G.J. and T.O. Binford, "Computer description of curved objects," *IEEE Trans. on Computers 25,* 1, April 1976.
Fig. 9.15a courtesy of Gerald Agin.
Fig. 9.17 courtesy of A. Christensen; published as frontispiece of *ACM SIGGRAPH 80 Proceedings.*
Fig. 9.20 from Marr and Nishihara, 1978.
Fig. 9.21 after Tilove, 1980.
Fig. 9.22b courtesy of Gene Hartquist.
Figs. 9.24, 9.25, and 9.26 from Lee and Requicha, 1980.
Figs. 9.28a, 9.29, 9.30, 9.31, 9.32, 9.35, and 9.37 and Table 9.1 from Brown, C. and R. Popplestone, "Cases in scene analysis," in *Pattern Recognition,* ed. B.G. Batchelor. New York: Plenum, 1978.
Fig. 9.28b from Guzman, A., "Decomposition of a visual scene into three-dimensional bodies," in *Automatic Interpretation and Classification of Images,* A. Grasseli, ed., New York: Academic Press, 1969.
Fig. 9.28c from Waltz, D., "Understanding line drawing of scenes with shadows," in *The Psychology of Computer Vision,* ed. P.H. Winston. New York: McGraw-Hill, 1975.
Fig. 9.28d after Turner, 1974.
Figs. 9.33, 9.38, 9.40, 9.42, 9.43, and 9.44 after Mackworth, 1973.

Figs. 9.39, 9.45, 9.46, and 9.47 and Table 9.2 after Kanade, 1978.
Figs. 10.2 and A2.1 courtesy of Dana Ballard.
Figs. 10.16, 10.17, and 10.18 after Russell, 1979.
Fig. 11.5 after Fischler and Elschlager, 1973.
Fig. 11.8 after Ambler et al., 1975.
Fig. 11.10 from Winston, P.H., "Learning structural descriptions from examples," in *The Psychology of Computer Vision,* ed. P.H. Winston. New York: McGraw-Hill, 1975.
Fig. 11.11 from Nevatia, 1974.
Fig. 11.12 after Nevatia, 1974.
Fig. 11.17 after Barrow and Popplestone, 1971.
Fig. 11.18 from Davis, L.S., "Shape matching using relaxation techniques," *IEEE Trans. PAMI 1,* 4, January 1979, pp. 60–72.
Figs. 12.4 and 12.5 from Sloan and Bajcsy, 1979.
Fig. 12.6 after Barrow and Tenenbaum, 1976.
Fig. 12.8 after Freuder, 1978.
Fig. 12.10 from Rosenfeld, A.R., A. Hummel, and S.W. Zucker, "Scene labeling by relaxation operations," *IEEE Trans. SMC 6,* 6, June 1976, p. 420.
Figs. 12.11, 12.12, 12.13, 12.14, and 12.15 after Hinton, 1979.
Fig. 13.3 courtesy of Aaron Sloman.
Figs. 13.6, 13.7, and 13.8 from Garvey, 1976.
Fig. A1.11 after Duda and Hart, 1973.
Figs. A2.2 and A2.3 from Hanson, A.R. and E.M. Riseman, "VISIONS: A computer system for interpreting scenes," in *Computer Vision Systems,* ed. A.R. Hanson and E.M. Riseman. New York: Academic Press, 1978.

Mnemonics for Proceedings and Special Collections Cited in the References

CGIP

Computer Graphics and Image Processing

COMPSAC

IEEE Computer Society's 3rd International Computer Software and Applications Conference, Chicago, November 1979.

CVS

Hanson, A. R. and E. M. Riseman (Eds.). *Computer Vision Systems.* New York: Academic Press, 1978.

DARPA IU

Defense Advanced Research Projects Agency Image Understanding Workshop, Minneapolis, MN, April 1977.

Defense Advanced Research Projects Agency Image Understanding Workshop, Palo Alto, CA, October 1977.

Defense Advanced Research Projects Agency Image Understanding Workshop, Cambridge, MA, May 1978.

Defense Advanced Research Projects Agency Image Understanding Workshop, Carnegie-Mellon University, Pittsburgh, PA, November 1978.

Defense Advanced Research Projects Agency Image Understanding Workshop, University of Maryland, College Park, MD, April 1980.

IJCAI

2nd International Joint Conference on Artificial Intelligence, Imperial College, London, September 1971.

4th International Joint Conference on Artificial Intelligence, Tbilisi, Georgia, USSR, September 1975.

5th International Joint Conference on Artificial Intelligence, MIT, Cambridge, MA, August 1977.

6th International Joint Conference on Artificial Intelligence, Tokyo, August 1979.

IJCPR

2nd International Joint Conference on Pattern Recognition, Copenhagen, August 1974.

3rd International Joint Conference on Pattern Recognition, Coronado, CA, November 1976.

4th International Joint Conference on Pattern Recognition, Kyoto, November 1978.

5th International Joint Conference on Pattern Recognition, Miami Beach, FL, December 1980.

MI4

Meltzer, B. and D. Michie (Eds.). *Machine Intelligence 4.* Edinburgh: Edinburgh University Press, 1969.

MI5

Meltzer, B. and D. Michie (Eds.). *Machine Intelligence 5.* Edinburgh: Edinburgh University Press, 1970.

MI6

Meltzer, B. and D. Michie (Eds.). *Machine Intelligence 6.* Edinburgh: Edinburgh University Press, 1971.

MI7

Meltzer, B. and D. Michie (Eds.). *Machine Intelligence 7.* Edinburgh: Edinburgh University Press, 1972.

PCV

Winston, P. H. (Ed.). *The Psychology of Computer Vision.* New York: McGraw-Hill, 1975.

PRIP

IEEE Computer Society Conference on Pattern Recognition and Image Processing, Chicago, August 1979.

COMPUTER VISION

DANA H. BALLARD ▪ CHRISTOPHER M. BROWN

What information about scenes can be extracted from an image using only basic assumptions about physics and optics?

How are images segmented into meaningful parts?

At what stage must domain-dependent, prior knowledge about the world be incorporated into the understanding process?

How are world models and conceptual knowledge represented and used?

These and many other questions, inherent in this relatively new and fast-growing field, are explored and answered in **Computer Vision** by **Dana H. Ballard** and **Christopher M. Brown.** The authors assemble crucial material from many disciplines including artificial intelligence, psychology, computer graphics, and image processing to form a practical text and reference for anyone involved in building vision systems.

Ballard and Brown write in their preface, "**Computer Vision** has a strong artificial intelligence flavor, and we hope this will provoke thought. The text shows how both intrinsic image information and internal models of the world are important in successful vision systems."

Divided into four parts, **Computer Vision** offers descriptions of objects at four levels of abstraction:

- Generalized images—images and image-like entities;
- Segmented images—images organized into sub-images that are likely to correspond to "interesting objects";
- Geometrical structures—quantitative models of image and world structures;
- Relational structures—complex symbolic descriptions of image and world structures.

PRENTICE-HALL, INC., Englewood Cliffs, N.J. 07632

ISBN 0-13-165316-4