# Seeing

## The Computational Approach to Biological Vision

## Second Edition

John P. Frisby and James V. Stone

The MIT Press
Cambridge, Massachusetts
London, England

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Adobe InDesign CS3 by John P. Frisby. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Seeing: the computational approach to biological vision / John P. Frisby and James V. Stone—2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-262-51427-9 (pbk.: alk. paper)

1. Visual perception. 2. Optical illusions. 3. Psychophysiology. I. Stone, James V. II. Title.

BF241.F74 2010

152.14—dc22 2009031989

## *Preface to Second Edition*

Writing an introductory text to the vast topic of seeing is rather like painting a picture: both entail a great deal of selection. Not everything can be included, even in brief, and so decisions have to be made about what to put in. In this book we have used as our guide the computational approach to vision. The intellectual foundations of this approach have a long pedigree, extending back at least as far as the work of the pioneers in the field of cybernetics that emerged in the 1950s and 60s. Cybernetics is usually defined as the interdisciplinary study of the structure of regulatory systems, and is hence closely related to control theory and systems theory. One of its hallmarks was a concern with stating clearly the problems that need to be solved. This feature is also the cornerstone of the computational approach to vision, best known from the work of David Marr and his colleagues.

Marr proposed three distinct levels of analysis for the study of seeing: first, the *computational level*, specifying the precise nature of the vision problem at hand and identifying constraints that can be used to solve the problem in principle; second, the *algorithmic level*, giving details of a definite method (an algorithm) for exploiting the these constraints; and third, the *hardware implementation level*, which specifies how to execute the method in a physical system (e.g., the neurons mediating seeing). Note that the computational approach has nothing to do with computers, nor is it to be confused with the field of study called computer vision.

It is a curious fact that the authors of some textbooks on vision and cognitive psychology begin by introducing Marr's three-levels approach, but then largely ignore it in the rest of the book. We also describe Marr's framework early on, but we follow this up by trying to use it systematically in tackling a wide range of vision problems. In each chapter, this approach is applied to a different topic in vision by examining problems that the visual system encounters in interpreting retinal images; the constraints available to solve these problems; the algorithms that can realize the solution; and the implementation of these algorithms in neurons.

Both authors have had considerable experience of using this approach in teaching vision courses to university students. It has led them to dispense with the usual kind of introductory lecture, because abstract ideas (such as *levels of analysis for understanding complex information systems*) tend to "fly right over the heads" of students with no hands-on experience of the problems of vision. Instead, we "plunge in" with some specific topics, and then use these to extract lessons about how the computational approach works out in practice. Hence, one feature of our teaching is that our "introductory" lecture is the final one in the course. We adopt this approach here, in that our final chapter, *Seeing Summarized*, evaluates Marr's approach and discusses it in the context of other approaches. We hope that by the time readers get to that chapter they are in a good position to engage with the debates it contains. If you are a reader who wants that context first, then consider skim-reading it at the outset.

Despite delaying our introduction of general contextual and abstract material to the final chapter, in the first chapter we try hard to make accessible the topic of seeing to the general reader and to students who have not previously attended a vision course. We hope that such readers will find that this book gives them a good general understanding of seeing, despite the fact that much of it is well suited to the needs of advanced undergraduates and postgraduates. In short, this book needs no special background or prior familiarity with basic concepts in vision. A touchstone of our approach is to try to *teach* the subject. For us, this means we cover at length, using many illustrations, topics that we know from long experience can be hard to grasp, (an example is the convolution of a receptive field with an image). We provide at the end of Ch 1 a guide as to what each chapter contains, and all chapters (except the first and last) have a short note at the outset entitled *Why read this chapter?* These are intended to help readers to find a path through the book that suits their expertise and interests.

The upshot of all the foregoing is that we have not set out to write a handbook, or to cover all the topics to which many other texts devote whole chapters. There is no such thing as the "perfect" textbook on seeing and we certainly make no claim in that regard. Readers will, indeed should, consult other texts, and, of course, pursue their particular interests using Web searches (see advice below and our sections at the end of each chapter on *Further Reading*). What we do hope we have achieved is

to provide enough basic coverage of the subject to give readers a sound understanding of core topics, so that they are well equipped to tackle the vast vision literature. And of course, we hope that readers become persuaded that the computational approach has strong merits as a way of thinking about seeing.

Finally, how does this second edition differ from the first? They have much in common, for example, a general emphasis on the computational approach, many illustrations, and an attempt to combine accessibility with rigor. (See Preface to First Edition, below.) That said, this second edition is much more explicit in teaching the computational approach, and it is much longer (roughly three times as many pages). Much of this increase is down to the introduction of new topics, which include mathematical material aimed at advanced students. Of course, we have also brought this second edition up to date. However, although a great deal has happened in the vision literature since the first edition was published in 1979, we have been surprised to realize that much of the first edition remains core material today. Moreover, we believe the general tenor of the computational approach espoused in 1979 in the first edition of *Seeing* still holds valuable lessons for students and vision researchers in the twenty first century.

*Acknowledgements, second edition*
We are greatly indebted to Neville Moray and Tom Stafford, who read and commented on all chapters. David Buckley, Helen Davis, and John Elliott gave detailed helpful advice on numerous chapters. John Mollon read several chapters as part of a publisher's review and his comments and suggestions were extremely useful, as were those from our students. Any errors remaining are of course all our own work. Emily Wilkinson gave us much-needed advice on navigating our way through the perilous seas of publishing. We thank Robert Prior of MIT Press for his support, and for the advice of MIT press staff on various aspects of layout.

JVS would like to thank Raymond Lister for taking the time to introduce to a (then) young psychologist about complexity theory, and Stephen Isard who taught an even younger psychologist how to think in straight lines.

*John P. Frisby and James V. Stone* *October 2009*

***Additional Note*** It has been a pleasure to work with Jim Stone on this second edition. His expertise, particularly using Bayesian methods, has been crucial to updating. We have exchanged 11,053 emails during the 4+ years it has taken to complete the project. And we are still friends. *John Frisby*

## *Preface to the First Edition*

Novelists sometimes say that their characters take over even the best laid plot and determine an ending different from the one originally envisaged. Something of the sort happened in the writing of this book, even though it is not a work of fiction, or at least is not meant to be. I started out to write *The General Reader's Fun Book of Visual Illusions*, with extended captions to give some explanation of why the illusions were as they were. But it proved impossible to say much of interest about them until the scene had been set with some introductory chapters on the fundamental nature of seeing and the basic neural machinery of the visual system. When I had done this, the natural format for the book turned out to be not just a series of illusory figures, but instead a series of chapters dealing with selected topics from the standpoint of how the visual system performs certain jobs. To be sure, pictures of illusions abound—they provide valuable clues about visual mechanisms—but now they complement and illustrate a "story", rather than performing the central guiding role as originally planned. This change makes the book suitable not only for a general readership, but also for an introductory course on vision, especially one in which the teacher wishes to emphasize a combined psychological, physiological, and computational approach. But no attempt has been made to write a conventional textbook. The coverage is anything but encyclopedic, and a course tutor will find many holes that need plugging. But that seems to me no bad thing: why else have a course tutor? More important in my view is to have a text which communicates some fundamental ideas in visual science, as it is being practised today, so laying a foundation upon which the teacher can build.

But despite this change from my expectations and plans, I hope the book is still "fun." It is certainly still intended for the general reader, in that no special knowledge is demanded and all terms are defined as the book proceeds. Even so,