TURING 图灵原版计算机科学系列

ELSEVIER 爱思唯尔

Machine Vision

Theory, Algorithms, Practicalities

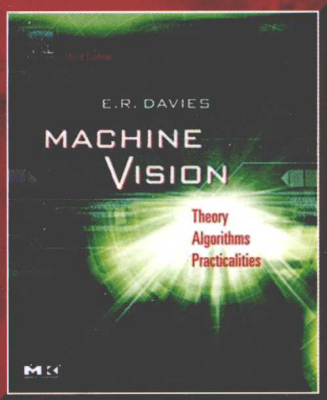
Third Edition

机器视觉

理论、算法与实践

（英文版·第3版）

[英] E. R. Davies 著



人民邮电出版社
POSTS & TELECOM PRESS

TP302.7/Y1
2009.

TURING 图灵原版计算机科学系列

Machine Vision

Theory, Algorithms, Practicalities

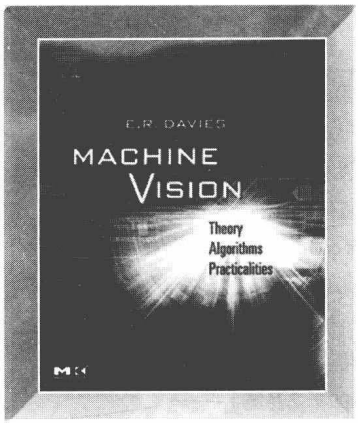
Third Edition

机器视觉

理论、算法与实践

(英文版·第3版)

[英] E. R. Davies 著



人民邮电出版社
北京

图书在版编目（CIP）数据

机器视觉：理论、算法与实践＝Machine Vision: Theory, Algorithms, Practicalities：第3版；英文/（英）戴维斯著. —北京：人民邮电出版社，2009.2
（图灵原版计算机科学系列）
ISBN 978-7-115-19549-4

I. 机… II. 戴… III. 计算机视觉－教材－英文
IV. TP302.7

中国版本图书馆CIP数据核字（2008）第204823号

内 容 提 要

本书是机器视觉课程的理想教材，作者清晰、系统地阐述了机器视觉的基本概念，介绍理论的基本元素的同时强调算法和实用设计的约束。书中阐述各个主题时，既阐述了基本算法，又介绍了数学工具。此外，本书还使用案例演示具体技术的应用，并阐明设计现实机器视觉系统的关键约束。

本书适合作为高等院校计算机及电子工程相关专业研究生的教材，更是从事机器视觉、计算机视觉和机器人领域研究的人员不可多得的技术参考书。

图灵原版计算机科学系列

机器视觉：理论、算法与实践（英文版·第3版）

◆ 著 [英] E. R. Davies
责任编辑 杨海玲

◆ 人民邮电出版社出版发行 北京市崇文区夕照寺街14号
邮编 100061 电子函件 315@ptpress.com.cn
网址 http://www.ptpress.com.cn
北京顺义振华印刷厂印刷

◆ 开本：800×1000 1/16
印张：60.5 彩插：2
字数：1162千字 2009年2月第1版
印数：1-2 000册 2009年2月北京第1次印刷

著作权合同登记号 图字：01-2008-5831号
ISBN 978-7-115-19549-4/TP

定价：128.00元

读者服务热线：（010）88593802 印装质量热线：（010）67129223
反盗版热线：（010）67171154

版权声明

Machine Vision: Theory, Algorithms, Practicalities, Third Edition by E. R. Davies, ISBN: 0-12-206093-8.
Copyright © 2005 by Elsevier. All rights reserved.
Authorized English language reprint edition published by the Proprietor.
ISBN: 978-981-272-319-2
Copyright © 2009 by Elsevier (Singapore) Pte Ltd. All rights reserved.

Elsevier (Singapore) Pte Ltd.
3 Killiney Road
#08-01 Winsland House I
Singapore 239519
Tel: (65)6349-0200
Fax: (65)6733-1817

First Published 2009
2009年初版

Printed in China by POSTS & TELECOM PRESS under special arrangement with Elsevier (Singapore) Pte Ltd. This edition is authorized for sale in China only, excluding Hong Kong SAR and Taiwan. Unauthorized export of this edition is a violation of the Copyright Act. Violation of this Law is subject to Civil and Criminal Penalties.

本书英文影印版由Elsevier (Singapore) Pte Ltd. 授权人民邮电出版社在中华人民共和国境内（不包括香港特别行政区和台湾地区）出版发行。未经许可之出口，视为违反著作权法，将受法律之制裁。

Dedication

This book is dedicated to my family.

To my late mother, Mary Davies, to record her never-failing love and devotion.

To my late father, Arthur Granville Davies, who passed on to me his appreciation of the beauties of mathematics and science.

To my wife, Joan, for love, patience, support, and inspiration.

To my children, Elizabeth, Sarah, and Marion, the music in my life.

此为试读,需要完整PDF请访问: www.ertongbook.com

Praise for

Machine Vision

THEORY, ALGORITHMS, PRACTICALITIES

3RD EDITION

"This book brings together the analytic aspects of image processing with the practicalities of applying the techniques in an industrial setting. It is excellent grounding for a machine vision researcher."

—*John Billingsley, University of Southern Queensland*

"The book in its previous incarnations has established its place as a unique repository of detailed analysis of important image processing and computer vision algorithms. This edition builds on these strengths and adds material to guide the reader's understanding of the latest developments in the field. The result is a comprehensive up-to-date reference text."

—*Farzin Deravi, University of Kent*

"This book is an essential reference for anyone developing techniques for machine vision analysis, including systems for industrial inspection, biomedical analysis, and much more. It comes from a long-term practitioner and is packed with the fundamental techniques required to build and prototype methods to test their applicability to the problem at hand."

—*Majid Mirmehdi, University of Bristol*

"The book contains a large number of experimental design and evaluation procedures that are of keen interest to industrial application engineers of machine vision."

—*William Wee, University of Cincinnati*

Foreword

An important focus of advances in mechatronics and robotics is the addition of sensory inputs to systems with increasing "intelligence." Without doubt, sight *is* the "sense of choice." In everyday life, whether driving a car or threading a needle, we depend first on sight. The addition of visual perception to machines promises the greatest improvement and at the same time presents the greatest challenge.

Until relatively recently, the volume of data in the images that make up a video stream has been a serious deterrent to progress. A single frame of very modest resolution might occupy a quarter of a megabyte, so the task of handling thirty or more such frames per second requires substantial computer resources.

Fortunately, the computer and communications industries' investment in entertainment has helped address this challenge. The transmission and processing of video signals are an easy justification for selling the consumer increased computing speed and bandwidth. A digital camera, capable of video capture, has already become a fashion accessory as part of a mobile phone. As a result, video signals have become more accessible to the serious engineer. But the task of acquiring a visual image is just the tip of the iceberg.

While generating sounds and pictures is a well-defined process (speech generation is a standard "accessibility" feature of Windows), the inverse task of recognizing connected speech is still at an unfinished state, a quarter of a century later, as any user of "dictation" software will attest. Still, analyzing sound is not even in the same league with analyzing images, particularly when they are of real-world situations rather than staged pieces with synthetic backgrounds and artificial lighting.

The task is essentially one of data reduction. From the many megabytes of the image stream, the required output might be a simple "All wheel nuts are in place" or "This tomato is ripe." But images tend to be noisy, objects that look sharp to the eye can have broken edges, boundaries can be fuzzy, and straight lines can be illusory. The task of image analysis demands a wealth of background know-how and mathematical analytic tools.

Roy Davies has been developing that rich background for well over two decades. At the time of the UK Robotics Initiative, in the 1980s, Roy had formed a relationship with the company United Biscuits. We fellow researchers might well have been amused by the task of ensuring that the blob of jam on a "Jaffacake" had been placed centrally beneath the enrobing chocolate. However, the funding of

the then expensive vision acquisition and analysis equipment, together with the spur of a practical target of real economic value, gave Roy a head start that made us all envious.

Grounded in that research is the realization that human image analysis is a many-layered process. It starts with simple graphical processing, of the sort that our eyes perform without our conscious awareness. Contrasts are enhanced; changes are emphasized and brought to our attention. Next we start to code the image in terms of lines, the curves of the horizon or of a face, boundaries between one region and another that a child might draw with a crayon. Then we have to "understand" the shape – is it a broken biscuit, or is one biscuit partly hidden by another? If we are comparing a succession of images, has something moved? What action should we take?

Machine vision must follow a similar, multilayered path. Roy has captured the essentials with clear, well-illustrated examples. He demonstrates filters that by convolution can smooth or sharpen an image. He shows us how to wield the tool of the Hough transform for locating lines and boundaries that are made indistinct by noise. He throws in the third dimension with stereo analysis, structured lighting, and optical flow. At every step, however, he drags us back to the world of reality by considering some practical task. Then, with software examples, he challenges us to try it for ourselves.

The easy access to digital cameras, video cameras, and streaming video in all its forms has promoted a tidal wave of would-be applications. But an evolving and substantial methodology is still essential to underpin the "art" of image processing. With this latest edition of his book, Roy continues to surf the crest of that wave.

John Billingsley
University of Southern Queensland

Preface

Preface to the third edition

The first edition came out in 1990, and was welcomed by many researchers and practitioners. However, in the space of 14 years the subject has moved on at a seemingly accelerating rate, and topics which hardly deserved a mention in the first edition had to be considered very seriously for the second and more recently for the third edition. It seemed particularly important to bring in significant amounts of new material on mathematical morphology, 3-D vision, invariance, motion analysis, artificial neural networks, texture analysis, X-ray inspection and foreign object detection, and robust statistics. There are thus new chapters or appendices on these topics, and these have been carefully integrated with the existing material. The greater proportion of this new material has been included in Parts 3 and 4. So great has been the growth in work on 3-D vision and motion—coupled with numerous applications on (for example) surveillance and vehicle guidance—that the original single chapter on 3-D vision has had to be expanded into the set of *six* chapters on 3-D vision and motion forming Part 3. In addition, Part 4 encompasses an increased range of chapters whose aim is to cover both applications and all the components needed for design of real-time visual pattern recognition systems. In fact, it is difficult to design a logical ordering for all the chapters—particularly in Part 4—as the topics interact with each other at a variety of different levels—theory, algorithms, methodologies, practicalities, design constraints, and so on. However, this should not matter, as the reader will be exposed to the essential richness of the subject, and his/her studies should be amply rewarded by increased understanding and capability.

A typical final-year undergraduate course on vision for Electronic Engineering or Computer Science students might include much of the work of Chapters 1 to 10 and 14 to 16, and a selection of sections from other chapters, according to requirements. For MSc or PhD research students, a suitable lecture course might go on to cover Part 3 in depth, including some chapters in Part 4 and also the appendix on robust statistics,[1] with many practical exercises being undertaken on

1 The importance of this appendix should not be underestimated once one gets onto serious work, though this will probably be outside the restrictive environment of an undergraduate syllabus.

an image analysis system. Here much will depend on the research program being undertaken by each individual student. At this stage the text will have to be used more as a handbook for research, and, indeed, one of the prime aims of the volume is to act as a handbook for the researcher and practitioner in this important area.

As mentioned above, this book leans heavily on experience I have gained from working with postgraduate students: in particular I would like to express my gratitude to Barry Cook, Mark Edmonds, Simon Barker, Daniel Celano, Darrel Greenhill, Derck Charles, and Mark Sugrue, all of whom have in their own ways helped to shape my view of the subject. In addition, it is a special pleasure to recall very many rewarding discussions with my colleagues Zahid Hussain, Ian Hannah, Dev Patel, David Mason, Mark Bateman, Tieying Lu, Adrian Johnstone, and Piers Plummer; the last two named having been particularly prolific in generating hardware systems for implementing my research group's vision algorithms.

The author owes a debt of gratitude to John Billingsley, Kevin Bowyer, Farzin Deravi, Cornelia Fermuller, Martial Hebert, Majid Mirmehdi, Qiang Ji, Stan Sclaroff, Milan Sonka, Ellen Walker, and William G. Wee, all of whom read a draft of the manuscript and made a great many insightful comments and suggestions. The author is particularly sad to record the passing of Professor Azriel Rosenfeld before he could complete his reading of the last few chapters, having been highly supportive of this volume in its various editions. Finally, I am indebted to Denise Penrose of Elsevier Science for her help and encouragement, without which this third edition might never have been completed.

Supporting Materials:

Morgan Kaufmann's website for the book contains resources to help instructors teach courses using this text. Please check the publishers' website for more information:

www.textbooks.elsevier.com

E.R. DAVIES
Royal Holloway,
University of London

Preface to the first edition

Over the past 30 years or so, machine vision has evolved into a mature subject embracing many topics and applications: these range from automatic (robot) assembly to automatic vehicle guidance, from automatic interpretation of documents to verification of signatures, and from analysis of remotely sensed images to checking of fingerprints and human blood cells; currently, automated

visual inspection is undergoing very substantial growth, necessary improvements in quality, safety and cost-effectiveness being the stimulating factors. With so much ongoing activity, it has become a difficult business for the professional to keep up with the subject and with relevant methodologies: in particular, it is difficult for him to distinguish accidental developments from genuine advances. It is the purpose of this book to provide background in this area.

The book was shaped over a period of 10–12 years, through material I have given on undergraduate and postgraduate courses at London University and contributions to various industrial courses and seminars. At the same time, my own investigations coupled with experience gained while supervising PhD and postdoctoral researchers helped to form the state of mind and knowledge that is now set out here. Certainly it is true to say that if I had had this book 8, 6, 4, or even 2 years ago, it would have been of inestimable value to myself for solving practical problems in machine vision. It is therefore my hope that it will now be of use to others in the same way. Of course, it has tended to follow an emphasis that is my own—and in particular one view of one path towards solving automated visual inspection and other problems associated with the application of vision in industry. At the same time, although there is a specialism here, great care has been taken to bring out general principles—including many applying throughout the field of image analysis. The reader will note the universality of topics such as noise suppression, edge detection, principles of illumination, feature recognition, Bayes' theory, and (nowadays) Hough transforms. However, the generalities lie deeper than this. The book has aimed to make some general observations and messages about the limitations, constraints, and tradeoffs to which vision algorithms are subject. Thus there are themes about the effects of noise, occlusion, distortion, and the need for built-in forms of robustness (as distinct from less successful *ad hoc* varieties and those added on as an afterthought); there are also themes about accuracy, systematic design, and the matching of algorithms and architectures. Finally, there are the problems of setting up lighting schemes which must be addressed in complete systems, yet which receive scant attention in most books on image processing and analysis. These remarks will indicate that the text is intended to be read at various levels—a factor that should make it of more lasting value than might initially be supposed from a quick perusal of the Contents.

Of course, writing a text such as this presents a great difficulty in that it is necessary to be highly selective: space simply does not allow everything in a subject of this nature and maturity to be dealt with adequately between two covers. One solution might be to dash rapidly through the whole area mentioning everything that comes to mind, but leaving the reader unable to understand anything in detail or to *achieve* anything having read the book. However, in a practical subject of this nature this seemed to me a rather worthless extreme. It is just possible that the emphasis has now veered too much in the opposite direction, by coming down to practicalities (detailed algorithms, details of lighting schemes, and so on):

此为试读,需要完整PDF请访问: www.ertongbook.com

individual readers will have to judge this for themselves. On the other hand, an author has to be true to himself and my view is that it is better for a reader or student to have mastered a coherent series of topics than to have a mish-mash of information that he is later unable to recall with any accuracy. This, then, is my justification for presenting this particular material in this particular way and for reluctantly omitting from detailed discussion such important topics as texture analysis, relaxation methods, motion, and optical flow.

As for the organization of the material, I have tried to make the early part of the book lead into the subject gently, giving enough detailed algorithms (especially in Chapters 2 and 6) to provide a sound feel for the subject—including especially vital, and in their own way quite intricate, topics such as connectedness in binary images. Hence Part 1 provides the lead-in, although it is not always trivial material and indeed some of the latest research ideas have been brought in (e.g., on thresholding techniques and edge detection). Part 2 gives much of the meat of the book. Indeed, the (book) literature of the subject currently has a significant gap in the area of intermediate-level vision; while high-level vision (AI) topics have long caught the researcher's imagination, intermediate-level vision has its own difficulties which are currently being solved with great success (note that the Hough transform, originally developed in 1962, and by many thought to be a very specialist topic of rather esoteric interest, is arguably only now coming into its own). Part 2 and the early chapters of Part 3 aim to make this clear, while Part 4 gives reasons why this particular transform has become so useful. As a whole, Part 3 aims to demonstrate some of the practical applications of the basic work covered earlier in the book, and to discuss some of the principles underlying implementation: it is here that chapters on lighting and hardware systems will be found. As there is a limit to what can be covered in the space available, there is a corresponding emphasis on the theory underpinning practicalities. Probably this is a vital feature, since there are many applications of vision both in industry and elsewhere, yet listing them and their intricacies risks dwelling on interminable detail, which some might find insipid; furthermore, detail has a tendency to date rather rapidly. Although the book could not cover 3-D vision in full (this topic would easily consume a whole volume in its own right), a careful overview of this complex mathematical and highly important subject seemed vital. It is therefore no accident that Chapter 16 is the longest in the book. Finally, Part 4 asks questions about the limitations and constraints of vision algorithms and answers them by drawing on information and experience from earlier chapters. It is tempting to call the last chapter the Conclusion. However, in such a dynamic subject area any such temptation has to be resisted, although it has still been possible to draw a good number of lessons on the nature and current state of the subject. Clearly, this chapter presents a personal view but I hope it is one that readers will find interesting and useful.

Acknowledgments

The author would like to credit the following sources for permission to reproduce tables, figures and extracts of text from earlier publications:

The Committee of the Alvey Vision Club for permission to reprint portions of the following paper as text in Chapter 15 and as Figs. 15.1, 15.2, 15.6:

Davies, E.R. (1988g).

CEP Consultants Ltd (Edinburgh) for permission to reprint portions of the following paper as text in Chapter 19:

Davies, E.R. (1987a).

Portions of the following papers from *Pattern Recognition and Signal Processing* as text in Chapters 3, 5, 9–15, 23, 24 and tables 5.1, 5.2, 5.3 and as Figs. 3.7, 3.9, 3.11, 3.12, 3.13, 3,17, 3,18, 3.19, 3.20, 3.21, 3.22, 5.2, 5.3, 5.4, 5.8, 9.1, 9.2, 9.3, 9.4, 10.4–10.14, 11.1, 11.3–11.11, 12.5–12.11, 13.1–13.5, 14.1–14.3, 14.6, 14.7, 14.8, 15.8, 23.2, 23.5:

Davies, E.R. (1986a,c, 1987c,d,g,k,l, 1988b,c,e,f, 1989 a–d).

Portions of the following papers from Image and Vision Computing as text in Chapter 5, as Tables 5.4, 5.5, 5.6 and as Figs. 3.33, 5.1, 5.5, 5.6, 5.7:

Davies, E.R. (1984b, 1987e)

Figures 5, 9; and about 10% of the text in the following paper:

Davies, E.R., Bateman, M., Mason, D.R., Chambers, J. and Ridgway, C. (2003). "Design of efficient line segment detectors for cereal grain inspection." Pattern Recogn. Lett., 24, nos. 1–3, 421–436.

Figure 4 of the following article:

Davies, E.R. (1987) "Visual inspection, automatic (robotics)", in *Encyclopedia of Physical Science and Technology*, Vol. 14, Academic Press, pp. 360–377.

About 4% of the text in the following paper:

Davies, E.R. (2003). "An analysis of the geometric distortions produced by median and related image processing filters." *Advances in Imaging and Electron Physics*, 126, Academic Press, pp. 93–193.

The above five items reprinted with permission from ***Elsevier***.

IEE for permission to reprint portions of the following papers from the IEE Proceedings and Colloquium Digests as text in Chapters 14, 22, 25, and as Figs. 14.13–14.15, 22.2–22.6, 25.13:

Davies, E.R. (1985b, 1988a)
Davies, E.R. and Johnstone, A.I.C. (1989)
Greenhill, D. and Davies, E.R. (1994b)

and

Figures 1–5; and about 44% of the text in the following paper:

Davies, E.R. (1997). Principles and design graphs for obtaining uniform illumination in automated visual inspection. Proc. 6th IEE Int. Conf. on *Image Processing and its Applications*, Dublin (14–17 July), IEE Conf. Publication no. 443, pp. 161–165.

Table 1; and about 95% of the text in the following paper:

Davies, E.R. (1997). Algorithms for inspection: constraints, tradeoffs and the design process. IEE Digest no. 1997/041, Colloquium on *Industrial Inspection*, IEE (10 Feb.), pp. 6/1–5.

Figure 1; and about 35% of the text in the following paper:

Davies, E.R., Bateman, M., Chambers, J. and Ridgway, C. (1998). Hybrid non-linear filters for locating speckled contaminants in grain. IEE Digest no. 1998/284, Colloquium on *Non-Linear Signal and Image Processing*, IEE (22 May), pp. 12/1–5.

Table 1; Figures 4–8; and about 15% of the text in the following paper:

Davies, E.R. (1999). Image distortions produced by mean, median and mode filters. IEE Proc. Vision Image Signal Processing, **146**, no. 5, 279–285.

Figure 1; and about 45% of the text in the following paper:

Davies, E.R. (2000). Obtaining optimum signal from set of directional template masks. Electronics Lett., **36**, no. 15, 1271–1272.

Figures 1–3; and about 90% of the text in the following paper:

Davies, E.R. (2000). Resolution of problem with use of closing for texture segmentation. Electronics Lett., **36**, no. 20, 1694–1696.

IEEE for permission to reprint portions of the following paper as text in Chapter 14 and as Figs. 3.5, 3.6, 3.8, 3.12, 14.4, 14.5:

Davies, E.R. (1984a, 1986d).

IFS Publications Ltd for permission to reprint portions of the following paper as text in Chapter 22 and as Figs. 10.1, 10.2, 22.7:

Davies, E.R. (1984c).

The Council of the Institution of Mechanical Engineers for permission to reprint portions of the following paper as Tables 28.1, 28.2:

Davies. E.R. and Johnstone, A.I.C. (1986).

MCB University Press (Emerald Group) for permission to reproduce Plate 1 of the following paper as Fig. 22.8:

Patel, D., Davies, E.R. and Hannah, I. (1995).

Pergamon (Elsevier Science Ltd) for permission to reprint portions of the following articles as text in Chapters 6, 7 and as Figs. 3.16, 6.7, 6.24:

Davies, E.R. and Plummer, A.P.N. (1981).
Davies, E.R. (1982).

Royal Swedish Academy of Engineering Sciences for permission to reprint portions of the following paper as text in Chapter 22 and as Fig. 7.4:

Davies, E.R. (1987f).

World Scientific for permission to reprint:

Figures 2.7(a, b), 3.4–3.6, 10.4, 12.1–12.3, 14.1; plus about 3 pp. of text extracted from p. 49 et seq; about 1 p. of text extracted from p. 132 et seq; about 2 pp. of text extracted from p. 157 et seq; about 3 pp. of text extracted from p. 187 et seq; about 7 pp. of text extracted from p. 207 et seq; about 3 pp. of text extracted from p. 219 et seq, in the following book:

Davies, E.R. (2000). *Image Processing for the Food Industry.*

in all cases excluding text which originated from earlier publications

The Royal Photographic Society for permission to reprint:

Table 1; figures 1, 2, 7, 9; and about 40% of the text in the following paper:

> Davies, E.R. (2000). A generalized model of the geometric distortions produced by rank-order filters. Imaging Science Journal, **48**, no. 3, 121–130.

Figures 4(b,e,h), 5(b), 14(a,b,c) in the following paper:

> Charles, D. and Davies, E.R. (2004). Mode filters and their effectiveness for processing colour images. Imaging Science Journal, **52**, no. 1, 3–25.

EURASIP for permission to reprint:

Figures 1–7; and about 65% of the text in the following paper:

> Davies, E.R. (1998). Rapid location of convex objects in digital images. Proc. EUSIPCO'98, Rhodes, Greece (8–11 Sept.), pp. 589–592.

Figure 1; and about 30% of the text in the following paper:

> Davies, E.R., Mason, D.R., Bateman, M., Chambers, J. and Ridgway, C. (1998). Linear feature detectors and their application to cereal inspection. Proc. EUSIPCO'98, Rhodes, Greece (8–11 Sept.), pp. 2561–2564.

CRC Press for permission to reprint:

Table 1; Figures 8, 10, 12, 15–17; and about 15% of the text in the following chapter:

> Davies (2003): Chapter 7 in Cho, H. (ed.) (2003). *Opto-Mechatronic Systems Handbook*.

Springer-Verlag for permission to reprint:

about 4% of the text in the following chapter:

> Davies, E.R. (2003). Design of object location algorithms and their use for food and cereals inspection. Chapter 15 in Graves, M. and Batchelor, B.G. (eds.) (2003). Machine Vision Techniques for Inspecting Natural Products, 393–420.

Unicom Seminars Ltd. for permission to reprint:

Portions of the following paper as text in Chapter 22 and as Figs. 9.2, 10.3, 12.3, 14.15, 15.9:

> Davies, E.R. (1988h).

Royal Holloway, University of London for permission to reprint:

Extracts from the following examination questions, written by E.R. Davies:

EL333/01/2, 4–6; EL333/98/2; EL333/99/2, 3, 5, 6; EL385/97/2; PH4760/04/1–5; PH5330/98/3, 5; PH5330/03/1–5.

University of London for permission to reprint:

Extracts from the following examination questions, written by E.R. Davies:

PH385/92/2, 3; PH385/93/1–3; PH385/94/1–4; PH385/95/4; PH385/96/3, 6; PH433/94/3, 5; PH433/96/2, 5.

此为试读,需要完整PDF请访问: www.ertongbook.com