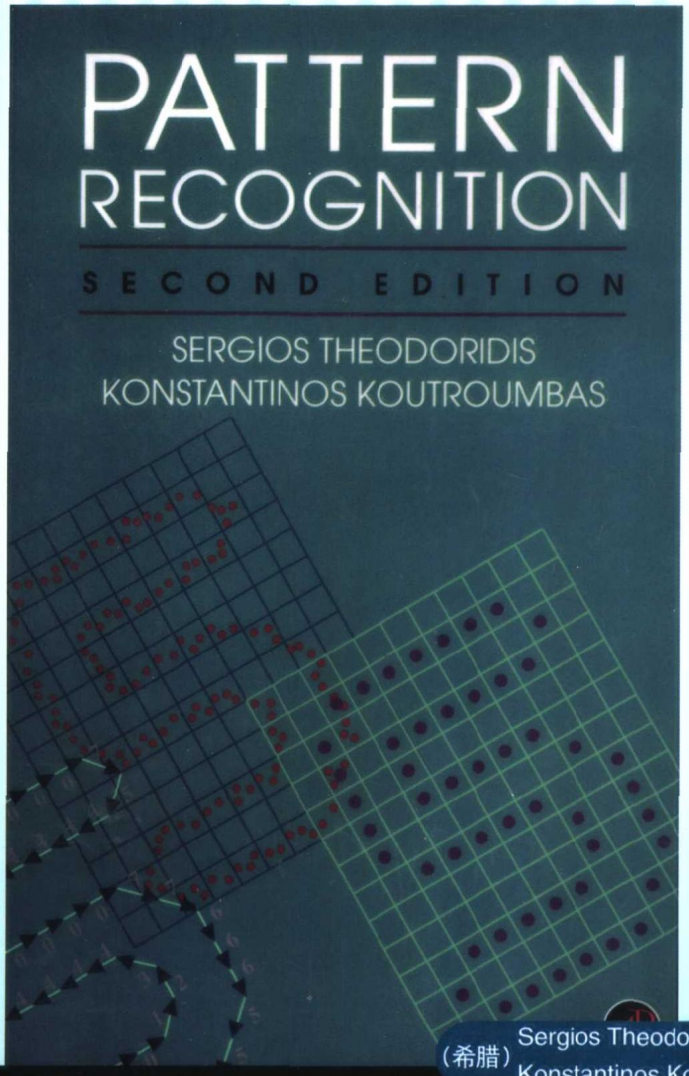
HZ BOOKS
华章教育

经 典 原 版 书 库

模式识别

（英文版 · 第2版）



（希腊）Sergios Theodoridis
Konstantinos Koutroumbas 著

机械工业出版社
China Machine Press

ELSEVIER

O235
3

经典原版书库

模式识别

(英文版·第2版)

Pattern Recognition

(Second Edition)

江苏工业学院图书馆
藏书章

Sergios Theodoridis
(希腊) Konstantinos Koutroumbas 著

机械工业出版社
China Machine Press

Z106833
北京信息工程学院图书馆

Sergios Theodoridis and Konstantinos Koutroumbas: Pattern Recognition, Second Edition (ISBN: 0-12-685875-6).

Original English language edition copyright © 2003 by Elsevier Science (USA).

All rights reserved.

Authorized English reprints copyright © 2003 by Elsevier (Singapore) Pte Ltd.

本书仅限于中华人民共和国境内（不包括中国香港、澳门特别行政区和中国台湾）销售发行。

版权所有，侵权必究。

本书版权登记号：图字：01-2003-5834

图书在版编目（CIP）数据

模式识别（英文版·第2版）/（希腊）西奥多里蒂斯（Theodoridis, S.）等著. -北京：机械工业出版社，2003.9

（经典原版书库）

书名原文：Pattern Recognition, Second Edition

ISBN 7-111-12767-6

I. 模… II. 西… III. 模式识别-英文 IV. TP391.4

中国版本图书馆CIP数据核字（2003）第068445号

机械工业出版社（北京市西城区百万庄大街22号 邮政编码 100037）

责任编辑：迟振春

北京昌平奔腾印刷厂印刷·新华书店北京发行所发行

2003年9月第1版第1次印刷

787mm × 1092mm 1/16 · 44.25印张

印数：0 001-3 000册

定价：69.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

本社购书热线电话（010）68326294

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭橥了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall，Addison-Wesley，McGraw-Hill，Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum，Stroustrup，Kernighan，Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及庋藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专诚为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师们服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国

家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业的教学度身订造的。其中许多教材均已为M. I. T.，Stanford，U.C. Berkeley，C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程，而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下，读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证，但我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下：

电子邮件：hzedu@hzbook.com
联系电话：（010）68995264
联系地址：北京市西城区百万庄南街1号
邮政编码：100037

此为试读，需要完整PDF请访问：www.ertongbook.com

专家指导委员会

（按姓氏笔画顺序）

尤晋元	王　珊	冯博琴	史忠植	史美林
石教英	吕　建	孙玉芳	吴世忠	吴时霖
张立昂	李伟琴	李师贤	李建中	杨冬青
邵维忠	陆丽娜	陆鑫达	陈向群	周伯生
周克定	周傲英	孟小峰	岳丽华	范　明
郑国梁	施伯乐	钟玉琢	唐世渭	袁崇义
高传善	梅　宏	程　旭	程时端	谢希仁
裘宗燕	戴　葵			

PREFACE

This book is the outgrowth of our teaching advanced undergraduate and graduate courses over the past 20 years. These courses have been taught to different audiences, including students in electrical and electronics engineering, computer engineering, computer science and informatics, as well as to an interdisciplinary audience of a graduate course on automation. This experience led us to make the book as self-contained as possible and to address students with different backgrounds. As prerequisitive knowledge the reader requires only basic calculus, elementary linear algebra, and some probability theory basics. A number of mathematical tools, such as probability and statistics as well as constrained optimization, needed by various chapters, are treated in four Appendices. The book is designed to serve as a text for advanced undergraduate and graduate students, and it can be used for either a one- or a two-semester course. Furthermore, it is intended to be used as a self-study and reference book for research and for the practicing scientist/engineer. This latter audience was also our second incentive for writing this book, due to the involvement of our group in a number of projects related to pattern recognition.

The philosophy of the book is to present various pattern recognition tasks in a unified way, including image analysis, speech processing, and communication applications. Despite their differences, these areas do share common features and their study can only benefit from a unified approach. Each chapter of the book starts with the basics and moves progressively to more advanced topics and reviews up-to-date techniques. A number of problems and computer exercises are given at the end of each chapter and a solutions manual is available from the publisher. Furthermore, a number of demonstrations based on MATLAB are available via the web at the book's site, http://www.di.uoa.gr/~stpatrec.

Our intention is to update the site regularly with more and/or improved versions of these demonstrations. Suggestions are always welcome. Also at this web site, a page will be available for typos, which are unavoidable, despite frequent careful reading. The authors would appreciate readers notifying them about any typos found.

This book would have not be written without the constant support and help from a number of colleagues and students throughout the years. We are especially indebted to Prof. K. Berberidis, Dr. E. Kofidis, Prof. A. Liavas, Dr. A. Rontogiannis, Dr. A. Pikrakis, Dr. Gezerlis and Dr. K. Georgoulakis. The constant support provided by Dr. I. Kopsinis from the early stages up to the final stage, with those long nights, has been invaluable. The book improved a great deal after the careful reading and the serious comments and suggestions of Prof. G. Moustakides, Prof. V. Digalakis, Prof. T. Adali, Prof. M. Zervakis, Prof. D. Cavouras, Prof. A. Böhm, Prof. G. Glentis, Prof. E. Koutsoupias, Prof. V. Zissimopoulos, Prof. A. Likas, Dr. A. Vassiliou, Dr. N. Vassilas, Dr. V. Drakopoulos, Dr. S. Hatzispyros. We are greatly indebted to these colleagues for their time and their constructive criticisms. Our collaboration and friendship with Prof. N. Kalouptsidis have been a source of constant inspiration for all these years. We are both deeply indebted to him.

Last but not least, K. Koutroumbas would like to thank Sophia for her tolerance and support and S. Theodoridis would like to thank Despina, Eva, and Eleni, his joyful and supportive "harem."

CONTENTS

Preface xv

CHAPTER 1 INTRODUCTION 1

1.1 Is Pattern Recognition Important? 1
1.2 Features, Feature Vectors, and Classifiers 3
1.3 Supervised Versus Unsupervised Pattern Recognition 6
1.4 Outline of the Book 8

CHAPTER 2 CLASSIFIERS BASED ON BAYES DECISION THEORY 13

2.1 Introduction 13
2.2 Bayes Decision Theory 13
2.3 Discriminant Functions and Decision Surfaces 19
2.4 Bayesian Classification for Normal Distributions 20
2.5 Estimation of Unknown Probability Density Functions 27
2.5.1 Maximum Likelihood Parameter Estimation 28
2.5.2 Maximum a Posteriori Probability Estimation 31
2.5.3 Bayesian Inference 32
2.5.4 Maximum Entropy Estimation 34
2.5.5 Mixture Models 35
2.5.6 Nonparametric Estimation 39
2.6 The Nearest Neighbor Rule 44

CHAPTER 3 LINEAR CLASSIFIERS 55

3.1 Introduction 55
3.2 Linear Discriminant Functions and Decision Hyperplanes 55
3.3 The Perceptron Algorithm 57

3.4 Least Squares Methods 65
3.4.1 Mean Square Error Estimation 65
3.4.2 Stochastic Approximation and the LMS Algorithm 68
3.4.3 Sum of Error Squares Estimation 70
3.5 Mean Square Estimation Revisited 72
3.5.1 Mean Square Error Regression 72
3.5.2 MSE Estimates Posterior Class Probabilities 73
3.5.3 The Bias–Variance Dilemma 76
3.6 Support Vector Machines 77
3.6.1 Separable Classes 77
3.6.2 Nonseparable Classes 82

CHAPTER 4 NONLINEAR CLASSIFIERS 93

4.1 Introduction 93
4.2 The XOR Problem 93
4.3 The Two-Layer Perceptron 94
4.3.1 Classification Capabilities of the Two-Layer Perceptron 98
4.4 Three-Layer Perceptrons 101
4.5 Algorithms Based on Exact Classification of the Training Set 102
4.6 The Backpropagation Algorithm 104
4.7 Variations on the Backpropagation Theme 112
4.8 The Cost Function Choice 115
4.9 Choice of the Network Size 118
4.10 A Simulation Example 124
4.11 Networks With Weight Sharing 126
4.12 Generalized Linear Classifiers 127
4.13 Capacity of the l-Dimensional Space in Linear Dichotomies 129
4.14 Polynomial Classifiers 131
4.15 Radial Basis Function Networks 133
4.16 Universal Approximators 137
4.17 Support Vector Machines: The Nonlinear Case 139
4.18 Decision Trees 143
4.18.1 Set of Questions 146
4.18.2 Splitting Criterion 146
4.18.3 Stop-Splitting Rule 147
4.18.4 Class Assignment Rule 147
4.19 Discussion 150

CHAPTER 5 FEATURE SELECTION 163

5.1 Introduction 163
5.2 Preprocessing 164
5.2.1 Outlier Removal 164
5.2.2 Data Normalization 165
5.2.3 Missing Data 165
5.3 Feature Selection Based on Statistical Hypothesis Testing 165
5.3.1 Hypothesis Testing Basics 166
5.3.2 Application of the t-Test in Feature Selection 171
5.4 The Receiver Operating Characteristics CROC Curve 173
5.5 Class Separability Measures 174
5.5.1 Divergence 174
5.5.2 Chernoff Bound and Bhattacharyya Distance 177
5.5.3 Scatter Matrices 179
5.6 Feature Subset Selection 181
5.6.1 Scalar Feature Selection 182
5.6.2 Feature Vector Selection 183
5.7 Optimal Feature Generation 187
5.8 Neural Networks and Feature Generation/Selection 191
5.9 A Hint on the Vapnik–Chernovenkis Learning Theory 193

CHAPTER 6 FEATURE GENERATION I: LINEAR TRANSFORMS 207

6.1 Introduction 207
6.2 Basis Vectors and Images 208
6.3 The Karhunen–Loève Transform 210
6.4 The Singular Value Decomposition 215
6.5 Independent Component Analysis 219
6.5.1 ICA Based on Second- and Fourth-Order Cumulants 221
6.5.2 ICA Based on Mutual Information 222
6.5.3 An ICA Simulation Example 226
6.6 The Discrete Fourier Transform (DFT) 226
6.6.1 One-Dimensional DFT 227
6.6.2 Two-Dimensional DFT 229
6.7 The Discrete Cosine and Sine Transforms 230
6.8 The Hadamard Transform 231
6.9 The Haar Transform 233

此为试读,需要完整PDF请访问: www.ertongbook.com

6.10 The Haar Expansion Revisited 235
6.11 Discrete Time Wavelet Transform (DTWT) 239
6.12 The Multiresolution Interpretation 249
6.13 Wavelet Packets 252
6.14 A Look at Two-Dimensional Generalizations 252
6.15 Applications 255

CHAPTER 7 FEATURE GENERATION II 269

7.1 Introduction 269
7.2 Regional Features 270
7.2.1 Features for Texture Characterization 270
7.2.2 Local Linear Transforms for Texture Feature Extraction 279
7.2.3 Moments 281
7.2.4 Parametric Models 286
7.3 Features for Shape and Size Characterization 294
7.3.1 Fourier Features 295
7.3.2 Chain Codes 298
7.3.3 Moment-Based Features 301
7.3.4 Geometric Features 302
7.4 A Glimpse at Fractals 303
7.4.1 Self-Similarity and Fractal Dimension 303
7.4.2 Fractional Brownian Motion 306

CHAPTER 8 TEMPLATE MATCHING 321

8.1 Introduction 321
8.2 Measures Based on Optimal Path Searching Techniques 322
8.2.1 Bellman's Optimality Principle and Dynamic Programming 324
8.2.2 The Edit Distance 325
8.2.3 Dynamic Time Warping in Speech Recognition 329
8.3 Measures Based on Correlations 337
8.4 Deformable Template Models 343

CHAPTER 9 CONTEXT-DEPENDENT CLASSIFICATION 351

9.1 Introduction 351
9.2 The Bayes Classifier 351
9.3 Markov Chain Models 352
9.4 The Viterbi Algorithm 353

9.5 Channel Equalization 356
9.6 Hidden Markov Models 361
9.7 Training Markov Models via Neural Networks 373
9.8 A discussion of Markov Random Fields 375

CHAPTER 10 SYSTEM EVALUATION 385

10.1 Introduction 385
10.2 Error Counting Approach 385
10.3 Exploiting the Finite Size of the Data Set 387
10.4 A Case Study From Medical Imaging 390

CHAPTER 11 CLUSTERING: BASIC CONCEPTS 397

11.1 Introduction 397
11.1.1 Applications of Cluster Analysis 400
11.1.2 Types of Features 401
11.1.3 Definitions of Clustering 402
11.2 Proximity Measures 404
11.2.1 Definitions 404
11.2.2 Proximity Measures between Two Points 407
11.2.3 Proximity Functions between a Point and a Set 418
11.2.4 Proximity Functions between Two Sets 423

CHAPTER 12 CLUSTERING ALGORITHMS I: SEQUENTIAL ALGORITHMS 429

12.1 Introduction 429
12.1.1 Number of Possible Clusterings 429
12.2 Categories of Clustering Algorithms 431
12.3 Sequential Clustering Algorithms 433
12.3.1 Estimation of the Number of Clusters 435
12.4 A Modification of BSAS 437
12.5 A Two-Threshold Sequential Scheme 438
12.6 Refinement Stages 441
12.7 Neural Network Implementation 443
12.7.1 Description of the Architecture 443
12.7.2 Implementation of the BSAS Algorithm 444

CHAPTER 13 CLUSTERING ALGORITHMS II: HIERARCHICAL ALGORITHMS 449

13.1 Introduction 449

13.2 Agglomerative Algorithms 450
13.2.1 Definition of Some Useful Quantities 451
13.2.2 Agglomerative Algorithms Based on Matrix Theory 453
13.2.3 Monotonicity and Crossover 461
13.2.4 Implementational Issues 464
13.2.5 Agglomerative Algorithms Based on Graph Theory 464
13.2.6 Ties in the Proximity Matrix 474
13.3 The Cophenetic Matrix 476
13.4 Divisive Algorithms 478
13.5 Choice of the Best Number of Clusters 480

CHAPTER 14 CLUSTERING ALGORITHMS III: SCHEMES BASED ON FUNCTION OPTIMIZATION 489

14.1 Introduction 489
14.2 Mixture Decomposition Schemes 491
14.2.1 Compact and Hyperellipsoidal Clusters 493
14.2.2 A Geometrical Interpretation 497
14.3 Fuzzy Clustering Algorithms 500
14.3.1 Point Representatives 505
14.3.2 Quadric Surfaces as Representatives 507
14.3.3 Hyperplane Representatives 517
14.3.4 Combining Quadric and Hyperplane Representatives 519
14.3.5 A Geometrical Interpretation 521
14.3.6 Convergence Aspects of the Fuzzy Clustering Algorithms 522
14.3.7 Alternating Cluster Estimation 522
14.4 Possibilistic Clustering 522
14.4.1 The Mode-Seeking Property 526
14.4.2 An Alternative Possibilistic Scheme 529
14.5 Hard Clustering Algorithms 529
14.5.1 The Isodata or k-Means or c-Means Algorithm 531
14.6 Vector Quantization 533

CHAPTER 15 CLUSTERING ALGORITHMS IV 545

15.1 Introduction 545
15.2 Clustering Algorithms Based on Graph Theory 545
15.2.1 Minimum Spanning Tree Algorithms 546

15.2.2 Algorithms Based on Regions of Influence 549
15.2.3 Algorithms Based on Directed Trees 550
15.3 Competitive Learning Algorithms 552
15.3.1 Basic Competitive Learning Algorithm 554
15.3.2 Leaky Learning Algorithm 556
15.3.3 Conscientious Competitive Learning Algorithms 556
15.3.4 Competitive Learning–Like Algorithms Associated with Cost Functions 558
15.3.5 Self-Organizing Maps 559
15.3.6 Supervised Learning Vector Quantization 560
15.4 Branch and Bound Clustering Algorithms 561
15.5 Binary Morphology Clustering Algorithms (BMCAs) 564
15.5.1 Discretization 564
15.5.2 Morphological Operations 565
15.5.3 Determination of the Clusters in a Discrete Binary Set 568
15.5.4 Assignment of Feature Vectors to Clusters 570
15.5.5 The Algorithmic Scheme 571
15.6 Boundary Detection Algorithms 573
15.7 Valley-Seeking Clustering Algorithms 576
15.8 Clustering Via Cost Optimization (Revisited) 578
15.8.1 Simulated Annealing 579
15.8.2 Deterministic Annealing 580
15.9 Clustering Using Genetic Algorithms 582
15.10 Other Clustering Algorithms 583

CHAPTER 16 CLUSTER VALIDITY 591

16.1 Introduction 591
16.2 Hypothesis Testing Revisited 592
16.3 Hypothesis Testing in Cluster Validity 594
16.3.1 External Criteria 596
16.3.2 Internal Criteria 602
16.4 Relative Criteria 605
16.4.1 Hard Clustering 608
16.4.2 Fuzzy Clustering 614
16.5 Validity of Individual Clusters 621
16.5.1 External Criteria 621
16.5.2 Internal Criteria 622
16.6 Clustering Tendency 624
16.6.1 Tests for Spatial Randomness 628

Appendix A
Hints from Probability and Statistics 643

Appendix B
Linear Algebra Basics 655

Appendix C
Cost Function Optimization 659

Appendix D
Basic Definitions from Linear Systems Theory 677

Index 681

CHAPTER 1

INTRODUCTION

1.1 IS PATTERN RECOGNITION IMPORTANT?

Pattern recognition is the scientific discipline whose goal is the classification of *objects* into a number of categories or *classes*. Depending on the application, these objects can be images or signal waveforms or any type of measurements that need to be classified. We will refer to these objects using the generic term *patterns*. Pattern recognition has a long history, but before the 1960s it was mostly the output of theoretical research in the area of statistics. As with everything else, the advent of computers increased the demand for practical applications of pattern recognition, which in turn set new demands for further theoretical developments. As our society evolves from the industrial to its postindustrial phase, automation in industrial production and the need for information handling and retrieval are becoming increasingly important. This trend has pushed pattern recognition to the high edge of today's engineering applications and research. Pattern recognition is an integral part in most *machine intelligence* systems built for decision making.

Machine vision is an area in which pattern recognition is of importance. A machine vision system captures images via a camera and analyzes them to produce descriptions of what is imaged. A typical application of a machine vision system is in the manufacturing industry, either for automated visual inspection or for automation in the assembly line. For example, in inspection, manufactured objects on a moving conveyor may pass the inspection station, where the camera stands, and it has to be ascertained whether there is a defect. Thus, images have to be analyzed on line, and a pattern recognition system has to classify the objects into the "defect" or "non-defect" class. After that, an action has to be taken, such as to reject the offending parts. In an assembly line, different objects must be located and "recognized," that is, classified in one of a number of classes known *a priori*. Examples are the "screwdriver class," the "German key class," and so forth in a tools' manufacturing unit. Then a robot arm can place the objects in the right place.

Character (letter or number) recognition is another important area of pattern recognition, with major implications in automation and information handling. Optical character recognition (OCR) systems are already commercially available and more or less familiar to all of us. An OCR system has a "front end" device

此为试读,需要完整PDF请访问: www.ertongbook.com