HZ BOOKS
华章教育
CENGAGE Learning

经 典 原 版 书 库

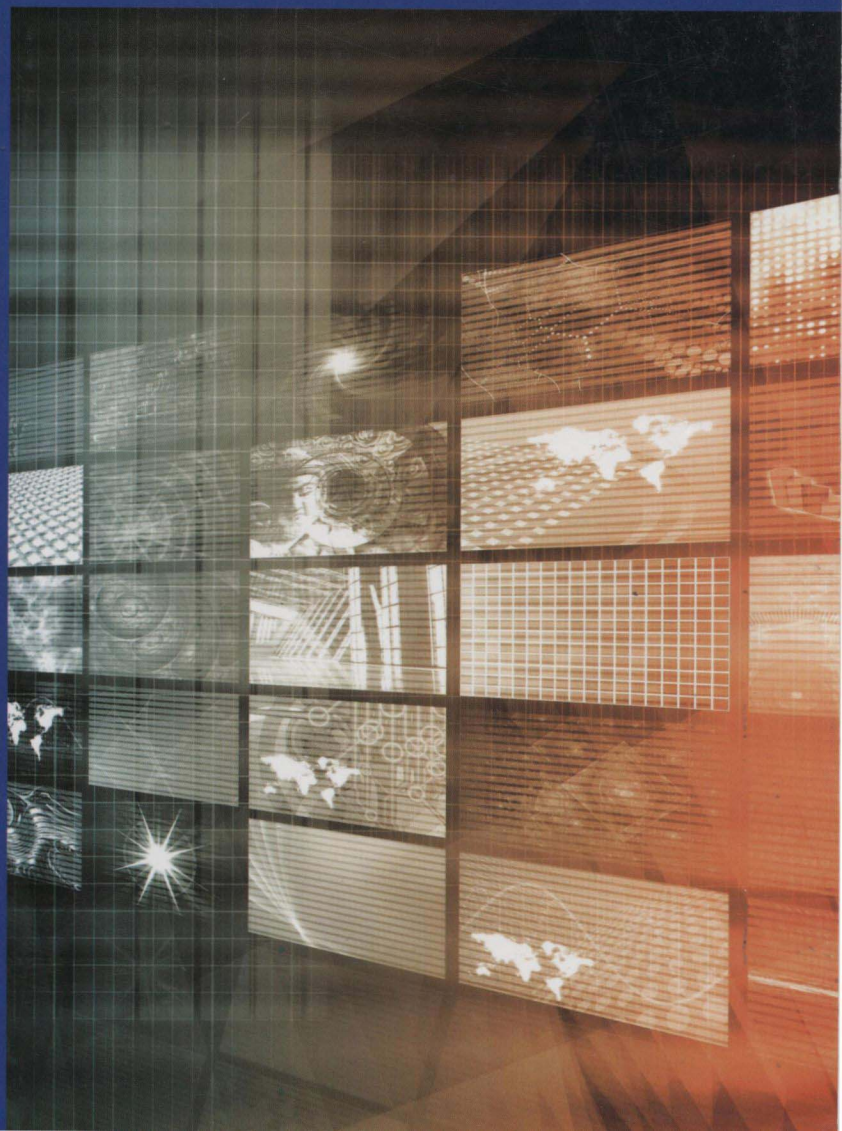
多媒体系统

算法、标准和工业实践

[美] 保劳格·哈沃达尔 (Parag Havaldar) 杰拉德·梅迪奥尼 (Gérard Medioni) 著

Multimedia Systems
Algorithms, Standards, and Industry Practices, China Student Edition

（英文精编版）



机械工业出版社
China Machine Press

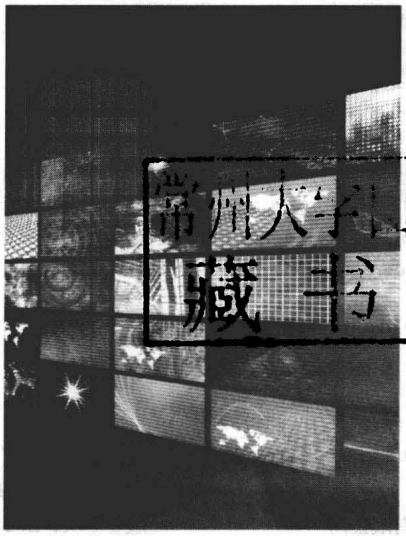
经典原版书库

多媒体系统

算法、标准和工业实践

Multimedia Systems

Algorithms, Standards, and Industry Practices, China Student Edition



常州大学图书馆
藏书章

[美] 保劳格·哈瓦达尔 杰拉德·梅迪奥尼 著
(Parag Havaldar) (Gerard Medioni)

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

多媒体系统：算法、标准和工业实践（英文精编版）/（美）哈沃达尔（Havaldar, P.），（美）梅迪奥尼（Medioni, G.）著．—北京：机械工业出版社，2015.4
（经典原版书库）
书名原文：Multimedia Systems: Algorithms, Standards, and Industry Practices, China Student Edition

ISBN 978-7-111-49929-9

I. 多… II. ①哈… ②梅… III. 多媒体技术－英文 IV. TP37

中国版本图书馆 CIP 数据核字（2015）第 073978 号

本书版权登记号：图字：01-2013-6798

Parag Havaldar, Gerard Medioni
Multimedia Systems: Algorithms, Standards, and Industry Practices, China Student Edition
ISBN: 978-1-305-02733-6
Copyright © 2015 Cengage Learning Asia Pte Ltd.
Original edition published by Cengage Learning. All Rights reserved.

China Machine Press is authorized by Cengage Learning to publish and distribute exclusively this custom reprint edition. This edition is authorized for sale in the People's Republic of China only (excluding Hong Kong, Macao SAR and Taiwan). Unauthorized export of this edition is a violation of the Copyright Act. No part of this publication may be reproduced or distributed by any means, or stored in a database or retrieval system, without the prior written permission of the publisher.

Cengage Learning Asia Pte. Ltd.
151 Lorong Chuan, #02-08 New Tech Park, Singapore 556741

本书原版由圣智学习出版公司出版。版权所有，盗印必究。

此客户定制影印版由圣智学习出版公司授权机械工业出版社独家出版发行。此版本仅限在中华人民共和国境内（不包括中国香港、澳门特别行政区及中国台湾）销售。未经授权的本书出口将被视为违反版权法的行为。未经出版者预先书面许可，不得以任何方式复制或发行本书的任何部分。

本书封面贴有 Cengage Learning 防伪标签，无标签者不得销售。

出版发行：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　　责任校对：董纪丽
印　　刷：蓟城市京瑞印刷有限公司　　版　　次：2015 年 6 月第 1 版第 1 次印刷
开　　本：170mm×242mm　1/16　　印　　张：29.25（含 1 印张彩插）
书　　号：ISBN 978-7-111-49929-9　　定　　价：79.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991　88361066　　投稿热线：（010）88379604
购书热线：（010）68326294　88379649　68995259　　读者信箱：hzjsj@hzbook.com

版权所有 · 侵权必究
封底无防伪标均为盗版
本书法律顾问：北京大成律师事务所　韩光 / 邹晓东

出版者的话

文艺复兴以来，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；也正是这样的优势，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，还揭示了学术的源变，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和发展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到积极的推动作用，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公司较早意识到“出版要为教育服务”。自1998年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与Pearson，McGraw-Hill，Elsevier，MIT，John Wiley & Sons，Cengage等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出Andrew S. Tanenbaum，Bjarne Stroustrup，Brain W. Kernighan，Dennis Ritchie，Jim Gray，Afred V. Aho，John E. Hopcroft，Jeffrey D. Ullman，Abraham Silberschatz，William Stallings，Donald E. Knuth，John L. Hennessy，Larry L. Peterson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、研究及珍藏。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相助，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，有的还专门为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，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。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

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。随着计算机科学与技术专业学科建设的不断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将步入一个新的阶段，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。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：

华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：（010）88379604

联系地址：北京市西城区百万庄南街1号

邮政编码：100037

华章科技图书出版中心

The successful completion of any large project needs devotion, discipline, and sacrifice.

To my parents for their love and the values they instilled in me.

To my family, teachers, friends, and well wishers for their support.

To my students for their feedback and invaluable discussions.

But little did I know whose sacrifice it really was:

To my children Veebha and Shreya—

for the weekends I could not take you hiking; for the afternoons that you waited for me to play; for the evenings I did not take you swimming, bicycling, or skating; and for the nights when I couldn't be beside you when you went to bed.

To my wife Chandrani—

without whose understanding, support, and love, this book was just not possible.

Parag Havaldar

此为试读,需要完整PDF请访问: www.ertongbook.com

PREFACE

SCOPE AND RATIONAL FOR ANOTHER BOOK IN MULTIMEDIA

Multimedia is now a broad "umbrella" that innovatively combines different fields of research and industry to produce practical solutions that are used on a wide scale today. Some of these fields are signal processing, imaging and color science, video and audio analysis, 2D/3D graphics, information theory, compression, networking, databases, watermarking, encryption, mobile terminals, and user interfaces. Research in each field is progressing, and our need to consume digital information has been forever changing. This has resulted in novel multimedia applications and faster dissemination of information that is constantly making our life more convenient when it comes to communication, entertainment, learning, interacting, and so on.

There are many books that address the progress of each of these above-mentioned fields individually. And although there exist books that deal with multimedia systems, most of them have been rather weighted and biased toward explaining only one or a few aspects of multimedia as a whole. For instance, many multimedia books target only the networking and distributing aspects, or only the compression and storage aspects. There is no comprehensive textbook that puts all these concepts coherently together, explaining each area sufficiently enough to understand the problems, solutions, technologies, and standards that can ultimately be used to create broad end-to-end applications in this ever-evolving field.

This book intends to serve that purpose by bringing together the different aspects of a modern multimedia pipeline from content creation, compression, distribution, and consumption on different end terminals. This book is borne out of teachings that the author has been carrying out at the University of Southern California,

feedback from students, and, more important, the author's perspectives gained from working in the industry. We discuss the issues involved in architecting an end-to-end multimedia pipeline and give plenty of examples from the industry, including digital television, IPTV, mobile deployments, Digital Rights Management solutions, digital cinema pipelines, and so on. We also provide lots of practical questions and programming assignments, which are really projects, to augment the student's understanding of the text.

TARGET AUDIENCE AND PREREQUISITES

The content, explanations, and exercises in this book have been designed for senior-level undergraduate students or graduate students in engineering and technical art disciplines.

We do not expect that you have taken courses in all of the engineering fields mentioned and most of the explanations do not assume this. However, it will be helpful to your overall understanding of the multimedia field if you are familiar with one or more of these fields. With regard to exercises, all of them have a rating between 1 and 10, with 1 being very easy, needing only a few moments to answer, whereas a 10 might turn out to be a weekend project. The programming exercises at the end of every chapter should give you hands-on experience with relevant aspects of multimedia. Although we do not assume that you are an expert programmer, you will need to know basic programming or scripting to attempt these. We also provide starter code in C++ and Java, with many sample data sets for the programming exercises. These can be found on the publisher's Web site, *www.cengage.com*, under the Student Downloads section of the book's catalog page.

ORGANIZATION

We start with an introductory chapter that takes the reader from a naïve perspective of multimedia and provides a more concrete definition of multimedia explaining the history, evolution, and current status of what multimedia is today. It also explains the three-part organization of the chapters, each with many visual figures, exercises, programming assignments (programming starter code available at *www.cengage.com*), along with solutions to selected exercises. A complete solution set is available to instructors via the Instructor Downloads section of *www.cengage.com*.

The first part of the book deals with authoring, where we explain relevant issues involved in capturing, representing, and creating content. We start with the digitization process for all media types, explaining the theoretical and practical details, issues in rendering on various display/sound devices, working of cameras, and formats of different media types. Also described are paradigms used by commercial authoring tools in the creation of rich multimedia content for different consumers on high-bandwidth digital networks to low-bandwidth mobile networks. This part also explains each media type (text, images, video, audio, graphics) from its simplistic individual

aspects to more complex content formed by the combinations, such as surround sound, spatial audio, THX, composite, and component video.

The second part is devoted to the data economics of storage and transmission of multimedia content. It first gives an overview of compression, which discusses theoretical and practical limits of information compression as well as explains a taxonomy of algorithms/standards in lossless and lossy coding. The succeeding chapters in this section discuss how these generic compression ideas are purposefully used to minimize perceptual distortion in compressing each media type (images, video, audio, and graphics) with good illustration of examples through every stage of the algorithms. Also discussed are the prominent ISO and ITU compression standards, which include JPEG, JPEG2000, MPEG-1, MPEG-2, H.264, CELP, MP3, MIDI, Topological Surgery used in MPEG-4, X3D, Java3D, and so on. In many cases, we also discuss the bit stream organization and syntax in these relevant standards. We also give examples of user interfaces and parameters exposed by industry-grade compressors for video and graphics.

The third part of the book pays attention to more recent trends and applications in multimedia. Here, we show the paradigm shift in content description/distribution using the MPEG-4 standard compared with the earlier MPEG standards. It shows examples of how MPEG-4 works, where it is currently used, and what might be possible with the standard. One chapter is also devoted to multimedia databases, where we explain the role of semantic queries and the complication involved in formulating and processing semantic queries when compared with queries in standard text databases. We show how solutions in this area have proposed the use of metadata and depict standards that use metadata such as MPEG-7, TV-Anytime, Dublin Core, MXF, DMS-1, and so on. One future requirement of media consumption will be the creation of frameworks that can seamlessly exchange multimedia data with different networks. With many different kinds and brands of commercial networks becoming commonplace today—cell phone networks, Internet, digital cable networks—we discuss the current progress in MPEG-21 to create such a framework where seamless and commercial exchange can be possible.

Finally, the last chapter concludes with a summary of the content covered in the book as well as content that was left out deliberately or because of the changing progress in the field. We also illustrate the impact that multimedia information consumption and dissemination has had in our industry and society and provide a perspective on where the multimedia industry is likely to move in future.

TEACHING

The book contains more than enough material for a one-semester course on multimedia and can also be used for a two-semester course depending on the depth that an instructor might want to go into each topic. The book has been organized into three parts that progress sequentially. However, depending on the student comfort level and exposure to prerequisites, each part could be taught individually. In the

Computer Science Department at the University of Southern California, the authors have been teaching a one-semester course to graduate students majoring in computer science, electrical engineering, industrial engineering, arts, and sciences. The first half of the course has normally covered material in varying detail from the first two parts of the book. The second half of the course has covered selected chapters from the third parts, depending on overall student interests and the new technologies that had then been in the news.

ACKNOWLEDGEMENTS

A project succeeds because of the contribution of many people. I'd like to thank students of CSCI 576 at the University of Southern California. There are very many who have gone through the course and are now part of the multimedia industry. Every semester has brought forth engaging discussions, novel ideas, and invaluable feedback that have played a pivotal role in the structure of this textbook. I would like to thank the reviewers who helped shape the contents of the manuscript with their numerous suggestions and corrections: Michael Hennessy, University of Oregon; Chung-wei Lee, Auburn University; Roberto Manduchi, University of California, Santa Cruz; Gloria Melara, California State University, Northridge; Refaat Mohamed, Western Kentucky University; Jane Ritter, University of Oregon; and Matthew Turk, University of California, Santa Barbara. Finally, Green Pen Quality Assurance provided technical editing for each chapter. This text will be more useful to students because of their combined efforts.

I would like to thank all the professionals at Cengage Learning, especially Amy Jollymore, Alyssa Pratt, and Jennifer Feltri for their efforts and organization throughout the project. The quality of the textbook would not have been possible without their meticulous oversight and timely management during the production process.

I want to thank friends and family who made it possible for me to undertake and complete this book. In life, the balance of work, family time, parental responsibilities, health, and fun is very critical. This is especially so when you are simultaneously engaged in a large project such as writing a book. My parents instilled values in me that have helped me keep that balance. They have long waited for the completion of the textbook. My daughters Shreya and Veebha have missed me many times when I disappeared in their presence with my laptop and thoughts. And last, but the very most, I could not have done this without the understanding, the support, and the smiles of my wife Chandrani.

We are happy to answer any questions about the book, receive corrections, engage in discussions regarding this evolving field, and provide additional features to help readers have a fruitful experience with the textbook. Please feel free to contact Parag Havaldar directly at havaldar@usc.edu.

CONTENTS

Preface v

CHAPTER 1

Introduction to Multimedia—Past, Present, and Future 1

1 Multimedia: Historical Perspective 2

2 Multimedia Data and Multimedia Systems 4

- 2.1 Inherent Qualities of Multimedia Data 4
- 2.2 Different Media Types Used Today 6
- 2.3 Classification of Multimedia Systems 8

3 A Multimedia System Today 9

4 The Multimedia Revolution 11

5 A Possible Future 13

6 Map of This Book 14

7 How to Approach the Exercises 15

PART 1 Multimedia Content Creation

CHAPTER 2

Digital Data Acquisition 17

1 Analog and Digital Signals 18

2 Analog-to-Digital Conversion 19

- 2.1 Sampling 19
- 2.2 Quantization 20
- 2.3 Bit Rate 23

3 Signals and Systems 24

- 3.1 Linear Time Invariant Systems 25
- 3.2 Fundamental Results in Linear Time Invariant Systems 25
- 3.3 Useful Signals 26
- 3.4 The Fourier Transform 26

4 Sampling Theorem and Aliasing 28

- 4.1 Aliasing in Spatial Domains 30
- 4.2 Aliasing in the Temporal Domain 30
- 4.3 Moiré Patterns and Aliasing 30

5 Filtering 33

- 5.1 Digital Filters 33
- 5.2 Filtering in 1D 35
- 5.3 Filtering in 2D 35
- 5.4 Subsampling 38

6 Fourier Theory 39

7 Exercises 44

Programming Assignments 47

CHAPTER 3

Media Representation and Media Formats 51

1 Digital Images 51

- 1.1 Digital Representation of Images 52
- 1.2 Aspect Ratios 55
- 1.3 Digital Image Formats 55

此为试读,需要完整PDF请访问: www.ertongbook.com

2 Digital Video 60
2.1 Representation of Digital Video 60
2.2 Analog Video and Television 61
2.3 Types of Video Signals 64
2.4 YUV Subsampling Schemes 65
2.5 Digital Video Formats 67
3 Digital Audio 69
3.1 Digital Representation of Audio 69
3.2 Surround Sound 70
3.3 Spatial Audio 71
3.4 Commonly Used Audio Formats 72
4 Graphics 73
5 Exercises 77
Programming Assignments 80

CHAPTER 4
Color Theory 81

1 The Color Problem 81
1.1 History of Color and Light 82
1.2 Human Color Sensing 84
1.3 Human Color Perception 85
2 Trichromacity Theory 86
2.1 Cone Response 87
2.2 The Tristimulus Vector 88
3 Color Calibration 90
3.1 Color Cameras 90
3.2 Rendering Devices 92
3.3 The Calibration Process 93
3.4 CIE Standard and Color-Matching Functions 94
4 Color Spaces 95
4.1 The CIE XYZ Color Space 96
4.2 RGB Color Space 97
4.3 CMY or CMYK Color Space 98
4.4 YUV Color Space 99
4.5 HSV Color Space 101

4.6 Uniform Color Spaces 102
4.7 Device Dependence of Color Spaces 103
5 Gamma Correction and Monitor Calibration 104
6 Exercises 105
Programming Assignments 108

CHAPTER 5

Multimedia Authoring 111

1 Examples of Multimedia 112
2 Requirements for Multimedia Authoring Tools 117
3 Intramedia Processing 118
3.1 Intramedia Issues Related to Images 119
3.2 Intramedia Issues Related to Video 119
3.3 Intramedia Issues Related to Audio 122
3.4 Intramedia Issues Related to 2D/3D Graphics 122
4 Intermedia Processing 124
4.1 Spatial Placement Control 125
4.2 Temporal Control 126
4.3 Interactivity Setup 127
5 Multimedia Authoring Paradigms and User Interfaces 127
5.1 Timeline 128
5.2 Scripting 129
5.3 Flow Control 131
5.4 Cards 131
6 Role of User Interfaces 132
6.1 User Interfaces on Mobile Devices 132
6.2 Multiple Devices as User Interfaces 133
7 Device-Independent Content Authoring 134
8 Distributed Authoring and Versioning 136
9 Multimedia Services and Content Management 137
10 Asset Management 138
11 Exercises 139
Programming Assignments 141

COLOR INSERT

PART 2 Multimedia Compression

CHAPTER 6

Overview of Compression 145

1 The Need for Compression 146

2 Basics of Information Theory 147

2.1 Information Theory Definitions 148

2.2 Information Representation 151

2.3 Entropy 151

2.4 Efficiency 153

3 A Taxonomy of Compression 154

3.1 Compression Metrics 155

3.2 Rate Distortion 155

4 Lossless Compression 156

4.1 Run Length Encoding 157

4.2 Repetition Suppression 157

4.3 Pattern Substitution 158

4.4 Huffman Coding 160

4.5 Arithmetic Coding 161

5 Lossy Compression 164

5.1 Differential PCM 165

5.2 Vector Quantization 166

5.3 Transform Coding 169

5.4 Subband Coding 172

5.5 Hybrid Compression Techniques 173

6 Practical Issues Related to Compression Systems 175

6.1 Encoder Speed and Complexity 175

6.2 Rate Control 176

6.3 Symmetric and Asymmetric Compression 176

6.4 Adaptive and Nonadaptive Compression 177

7 Exercises 177

Programming Assignments 184

CHAPTER 7

Media Compression: Images 187

1 Redundancy and Relevancy of Image Data 189

2 Classes of Image Compression Techniques 190

3 Lossless Image Coding 191

3.1 Image Coding Based on Run Length 192

3.2 Dictionary-Based Image Coding (GIF, PNG) 192

3.3 Prediction-Based Coding 192

4 Transform Image Coding 193

4.1 DCT Image Coding and the JPEG Standard 194

4.2 JPEG Bit Stream 198

4.3 Drawbacks of JPEG 200

5 Wavelet Based Coding (JPEG 2000) 201

5.1 The Preprocessing Step 202

5.2 The Discrete Wavelet Transform 203

5.3 JPEG 2000 Versus JPEG 205

6 Fractal Image Coding 207

6.1 Fractals 208

6.2 Fractal Block Coding 209

6.3 The Future of Fractal Image Compression 210

7 Transmission Issues in Compressed Images 210

7.1 Progressive Transmission Using DCTs in JPEG 211

7.2 Progressive Transmission Using Wavelets in JPEG 2000 213

8 The Discrete Cosine Transform 213

9 Exercises 216

Programming Assignments 221

CHAPTER 8

Media Compression: Video 223

1 General Theory of Video Compression 224

1.1 Temporal Redundancy 227

1.2 Block-Based Frame Prediction 228

1.3 Computing Motion Vectors 231
1.4 Size of Macroblocks 233
1.5 Open Loop versus Closed Loop Motion Compensation 235

2 Types of Predictions 236
2.1 I Frames 237
2.2 P Frames 238
2.3 B Frames 238
2.4 Multiframe Prediction 240
2.5 Video Structure—Group of Pictures 242

3 Complexity of Motion Compensation 243
3.1 Sequential or Brute Force Search 244
3.2 Logarithmic Search 245
3.3 Hierarchical Search 246

4 Video-Coding Standards 247
4.1 H.261 248
4.2 H.263 248
4.3 MPEG-1 248
4.4 MPEG-2 249
4.5 MPEG-4—VOP and Object Base Coding, SP and ASP 251
4.6 H.264 or MPEG-4—AVC 252

5 VBR Encoding, CBR Encoding, and Rate Control 254

6 A Commercial Encoder 256

7 Exercises 258

Programming Assignments 265

CHAPTER 9

Media Compression: Audio 269

1 The Need for Audio Compression 270

2 Audio-Compression Theory 271

3 Audio as a Waveform 273
3.1 DPCM and Entropy Coding 273
3.2 Delta Modulation 274
3.3 ADPCM 275
3.4 Logarithmic Quantization Scales—A-law and μ law 275

4 Audio Compression Using Psychoacoustics 276
4.1 Anatomy of the Ear 277
4.2 Frequency Domain Limits 277
4.3 Time Domain Limits 278
4.4 Masking or Hiding 278
4.5 Perceptual Encoder 281
5 Model-Based Audio Compression 283
6 Audio Compression Using Event Lists 285
6.1 Structured Representations and Synthesis Methodologies 286
6.2 Advantage of Structured Audio 287
7 Audio Coding Standards 287
7.1 MPEG-1 288
7.2 MPEG-2 291
7.3 Dolby AC-2 and AC-3 292
7.4 MPEG-4 294
7.5 ITU G.711 294
7.6 ITU G.722 295
7.7 ITU G.721, ITU G.726, and ITU G.727 295
7.8 ITU G.723 and ITU G.729 295
7.9 ITU G.728 295
7.10 MIDI 296
8 Exercises 297
Programming Assignments 300

CHAPTER 10

Media Compression: Graphics 301

1 The Need for Graphics Compression 303
2 2D Graphics Objects 305
2.1 Points 305
2.2 Regions 305
2.3 Curves 305
3 3D Graphics Objects 306
3.1 Polygonal Descriptions 307
3.2 Patch-Based Descriptions 308
3.3 Constructive Solid Geometry 308

此为试读,需要完整PDF请访问: www.ertongbook.com