Compression for Multimedia

Providing a thorough theoretical understanding of lossy compression techniques for image, video, speech, and audio compression, this book also covers the key features of each system, as well as practical applications, implementation issues, and design trade-offs. It presents comparisons of multimedia standards in terms of achieving known theoretical limits, whilst common and distinguishing features of the existing standards are explained and related to the background theory. There is detailed coverage of such topics as the H.264 video coding standard, low-complexity code-based vector quantizers, and the Blahut rate-distortion algorithm. Examples based on real multimedia data are also included, together with end-of-chapter problems to test understanding; algorithms that allow the reader to represent speech and audio signals efficiently; and an appendix on the basics of lossless coding. With an excellent balance of theory and practice, this book is ideal for undergraduate and graduate students, and is also a useful reference for practitioners.

Irina Bocharova is an Associate Professor at the Saint Petersburg State University of Information Technologies, Mechanics, and Optics. She has published over 50 technical papers and is the co-inventor of seven US patents in speech, video, and audio coding. Her current research interests include convolutional codes, communication systems, source coding and its applications to speech, audio, and image coding.
Compression for Multimedia

IRINA BOCHAROVA

St Petersburg State University of Information Technologies, Mechanics, and Optics
To
my teachers
## Contents

*Preface*  
page ix

1. **Introduction**  
   
2. **Analog to digital conversion**  
   2.1 Analog and digital signals  
   2.2 Scalar quantization  
   2.3 Vector quantization  

3. **Elements of rate-distortion theory**  
   3.1 Rate-distortion function  
   3.2 The Blahut algorithm  
   3.3 The Shannon lower bound and high-resolution approximations  
   3.4 Comparison of quantization procedures  
   3.5 Characteristics of digital speech, audio, image, and video signals  

4. **Scalar quantization with memory**  
   4.1 Discrete-time filters  
   4.2 Linear predictive coding  

5. **Transform coding**  
   5.1 Properties of transforms  
   5.2 The Karhunen–Loeve transform  
   5.3 The discrete Fourier transform  
   5.4 The discrete cosine transform  

6. **Filter banks and wavelet filtering**  
   6.1 Linear filtering as a linear transform  
   6.2 “Moving average” filtering as a linear transform  
   6.3 “Moving difference” filtering as a linear transform  
   6.4 The Haar filter bank
**Contents**

6.5 Wavelet transform and wavelet filter bank 121
6.6 Hierarchical wavelet filtering as a linear transform. Properties of wavelet filtering 126
6.7 Historical background 132
6.8 Application of wavelet filtering to image compression 132
6.9 Embedded zerotree coding and set partitioning in hierarchical trees 135

7 Speech coding: techniques and standards 141
7.1 Direct sample-by-sample quantization: Standard G.711 145
7.2 ADPCM coders: Standards G.726, G.727 147
7.3 Linear prediction analysis-by-synthesis coders (LPC-AS) 152
7.4 LPC vocoders. MELP standard 162
7.5 An algorithm for computing linear spectral parameters 165

8 Image coding standards 171
8.1 Coding of bi-level fax images: JBIG Standard 172
8.2 Coding of halftone images: JPEG Standard 177
8.3 JPEG-LS 186
8.4 Standard JPEG-2000 188

9 Video-coding standards 197
9.1 Motion compensation method 197
9.2 Overview of video-coding standards 200

10 Audio-coding standards 223
10.1 Basics of perceptual coding 223
10.2 Overview of audio standards 227

A Lossless-coding techniques 238
A.1 Symbol-by-symbol lossless coding 238
A.2 Lossless block coding 241
A.3 Arithmetic coding 243
A.4 Decoding issues 246
A.5 Context lossless coding for sources with memory 249
A.6 QM-coder 251
A.7 Monotonic codes 254
A.8 Binarization 257
A.9 Adaptive coding 258

References 261
Index 265
Preface

Compression for Multimedia was primarily developed as class notes for my course on techniques for compression of data, speech, music, pictures, and video that I have been teaching for more than 10 years at the University of Aerospace Instrumentation, St Petersburg.

During spring 2005 I worked at Lund University as the Lise Meitner Visiting Professor. I have used part of this time to thoroughly revise and substantially extend my previous notes, resulting in the present version.

I would also like to mention that this task could not have been fulfilled without support. Above all, I am indebted to my colleague and husband Boris Kudryashov. Without our collaboration I would not have reached my view of how various compression techniques could be developed and should be taught. Boris’ help in solving many TEX problems was invaluable. Special thanks go to Grigory Tenengolts who supported our research and development of practical methods for multimedia compression. Finally, I am grateful to Rolf Johannesson who proposed me as a Lise Meitner Visiting Professor and, needless to say, to the Engineering faculty of Lund University who made his recommendation become true! Rolf also suggested that I should give an undergraduate course on compression for multimedia at Lund University, develop these notes, and eventually publish them as a book. Thanks!

Lund, April 2009

Irina Bocharova