

H.264 and MPEG-4 Video Compression

Video Coding for Next-generation Multimedia

Iain E. G. Richardson

The Robert Gordon University, Aberdeen, UK



WILEY

Contents

About the Author	xiii
Foreword	xv
Preface	xix
Glossary	xxi
1 Introduction	1
1.1 The Scene	1
1.2 Video Compression	3
1.3 MPEG-4 and H.264	5
1.4 This Book	6
1.5 References	7
2 Video Formats and Quality	9
2.1 Introduction	9
2.2 Natural Video Scenes	9
2.3 Capture	10
2.3.1 Spatial Sampling	11
2.3.2 Temporal Sampling	11
2.3.3 Frames and Fields	13
2.4 Colour Spaces	13
2.4.1 RGB	14
2.4.2 YCbCr	15
2.4.3 YCbCr Sampling Formats	17
2.5 Video Formats	19
2.6 Quality	20
2.6.1 Subjective Quality Measurement	21
2.6.2 Objective Quality Measurement	22
2.7 Conclusions	24
2.8 References	24

3 Video Coding Concepts	27
3.1 Introduction	27
3.2 Video CODEC	28
3.3 Temporal Model	30
3.3.1 Prediction from the Previous Video Frame	30
3.3.2 Changes due to Motion	30
3.3.3 Block-based Motion Estimation and Compensation	32
3.3.4 Motion Compensated Prediction of a Macroblock	33
3.3.5 Motion Compensation Block Size	34
3.3.6 Sub-pixel Motion Compensation	37
3.3.7 Region-based Motion Compensation	41
3.4 Image model	42
3.4.1 Predictive Image Coding	44
3.4.2 Transform Coding	45
3.4.3 Quantisation	51
3.4.4 Reordering and Zero Encoding	56
3.5 Entropy Coder	61
3.5.1 Predictive Coding	61
3.5.2 Variable-length Coding	62
3.5.3 Arithmetic Coding	69
3.6 The Hybrid DPCM/DCT Video CODEC Model	72
3.7 Conclusions	82
3.8 References	83
4 The MPEG-4 and H.264 Standards	85
4.1 Introduction	85
4.2 Developing the Standards	85
4.2.1 ISO MPEG	86
4.2.2 ITU-T VCEG	87
4.2.3 JVT	87
4.2.4 Development History	88
4.2.5 Deciding the Content of the Standards	88
4.3 Using the Standards	89
4.3.1 What the Standards Cover	90
4.3.2 Decoding the Standards	90
4.3.3 Conforming to the Standards	91
4.4 Overview of MPEG-4 Visual/Part 2	92
4.5 Overview of H.264 / MPEG-4 Part 10	93
4.6 Comparison of MPEG-4 Visual and H.264	94
4.7 Related Standards	95
4.7.1 JPEG and JPEG2000	95
4.7.2 MPEG-1 and MPEG-2	95
4.7.3 H.261 and H.263	96
4.7.4 Other Parts of MPEG-4	97
4.8 Conclusions	97
4.9 References	98

5 MPEG-4 Visual	99
5.1 Introduction	99
5.2 Overview of MPEG-4 Visual (Natural Video Coding)	100
5.2.1 Features	100
5.2.2 Tools, Objects, Profiles and Levels	100
5.2.3 Video Objects	103
5.3 Coding Rectangular Frames	104
5.3.1 Input and Output Video Format	106
5.3.2 The Simple Profile	106
5.3.3 The Advanced Simple Profile	115
5.3.4 The Advanced Real Time Simple Profile	121
5.4 Coding Arbitrary-shaped Regions	122
5.4.1 The Core Profile	124
5.4.2 The Main Profile	133
5.4.3 The Advanced Coding Efficiency Profile	138
5.4.4 The N-bit Profile	141
5.5 Scalable Video Coding	142
5.5.1 Spatial Scalability	142
5.5.2 Temporal Scalability	144
5.5.3 Fine Granular Scalability	145
5.5.4 The Simple Scalable Profile	148
5.5.5 The Core Scalable Profile	148
5.5.6 The Fine Granular Scalability Profile	149
5.6 Texture Coding	149
5.6.1 The Scalable Texture Profile	152
5.6.2 The Advanced Scalable Texture Profile	152
5.7 Coding Studio-quality Video	153
5.7.1 The Simple Studio Profile	153
5.7.2 The Core Studio Profile	155
5.8 Coding Synthetic Visual Scenes	155
5.8.1 Animated 2D and 3D Mesh Coding	155
5.8.2 Face and Body Animation	156
5.9 Conclusions	156
5.10 References	156
6 H.264/MPEG-4 Part 10	159
6.1 Introduction	159
6.1.1 Terminology	159
6.2 The H.264 CODEC	160
6.3 H.264 structure	162
6.3.1 Profiles and Levels	162
6.3.2 Video Format	162
6.3.3 Coded Data Format	163
6.3.4 Reference Pictures	163
6.3.5 Slices	164
6.3.6 Macroblocks	164

6.4	The Baseline Profile	165
6.4.1	Overview	165
6.4.2	Reference Picture Management	166
6.4.3	Slices	167
6.4.4	Macroblock Prediction	169
6.4.5	Inter Prediction	170
6.4.6	Intra Prediction	177
6.4.7	Deblocking Filter	184
6.4.8	Transform and Quantisation	187
6.4.9	4×4 Luma DC Coefficient Transform and Quantisation (16×16 Intra-mode Only)	194
6.4.10	2×2 Chroma DC Coefficient Transform and Quantisation	195
6.4.11	The Complete Transform, Quantisation, Rescaling and Inverse Transform Process	196
6.4.12	Reordering	198
6.4.13	Entropy Coding	198
6.5	The Main Profile	207
6.5.1	B Slices	207
6.5.2	Weighted Prediction	211
6.5.3	Interlaced Video	212
6.5.4	Context-based Adaptive Binary Arithmetic Coding (CABAC)	212
6.6	The Extended Profile	216
6.6.1	SP and SI slices	216
6.6.2	Data Partitioned Slices	220
6.7	Transport of H.264	220
6.8	Conclusions	222
6.9	References	222
7	Design and Performance	225
7.1	Introduction	225
7.2	Functional Design	225
7.2.1	Segmentation	226
7.2.2	Motion Estimation	226
7.2.3	DCT/IDCT	234
7.2.4	Wavelet Transform	238
7.2.5	Quantise/Rescale	238
7.2.6	Entropy Coding	238
7.3	Input and Output	241
7.3.1	Interfacing	241
7.3.2	Pre-processing	242
7.3.3	Post-processing	243
7.4	Performance	246
7.4.1	Criteria	246
7.4.2	Subjective Performance	247
7.4.3	Rate–distortion Performance	251

7.4.4 Computational Performance	254
7.4.5 Performance Optimisation	255
7.5 Rate control	256
7.6 Transport and Storage	262
7.6.1 Transport Mechanisms	262
7.6.2 File Formats	263
7.6.3 Coding and Transport Issues	264
7.7 Conclusions	265
7.8 References	265
8 Applications and Directions	269
8.1 Introduction	269
8.2 Applications	269
8.3 Platforms	270
8.4 Choosing a CODEC	270
8.5 Commercial issues	272
8.5.1 Open Standards?	273
8.5.2 Licensing MPEG-4 Visual and H.264	274
8.5.3 Capturing the Market	274
8.6 Future Directions	275
8.7 Conclusions	276
8.8 References	276
Bibliography	277
Index	279