

THE H.264 ADVANCED VIDEO COMPRESSION STANDARD

Second Edition

Iain E. Richardson

Vcodex Limited, UK

WILEY

A John Wiley and Sons, Ltd., Publication

Contents

| | |
|---|-------------|
| About the Author | xiii |
| Preface | xv |
| Glossary | xvii |
| List of Figures | xxi |
| List of Tables | xxix |
| 1 Introduction | 1 |
| .1 A change of scene | 1 |
| .2 Driving the change | 4 |
| .3 The role of standards | 4 |
| .4 Why H.264 Advanced Video Coding is important | 4 |
| .5 About this book | 5 |
| .6 Reference | 6 |
| 2 Video formats and quality | 7 |
| 2.1 Introduction | 7 |
| 2.2 Natural video scenes | 7 |
| 2.3 , Capture | 8 |
| 2.3.1 Spatial sampling | 9 |
| 2.3.2 Temporal sampling | 9 |
| 2.3.3 Frames and fields | 11 |
| 2.4 Colour spaces | 12 |
| 2.4.1 RGB | 12 |
| 2.4.2 YCrCb | 13 |
| 2.4.3 YCrCb sampling formats | 14 |
| 2.5 Video formats | 16 |
| 2.5.1 Intermediate formats | 16 |
| 2.5.2 Standard Definition | 17 |
| 2.5.3 High Definition | 18 |
| 2.6 Quality | 19 |
| 2.6.1 Subjective quality measurement | 20 |
| 2.6.2 Objective quality measurement | 21 |

| | | |
|----------|---------------------------------------|-----------|
| 2.7 | Summary | 24 |
| 2.8 | References | 24 |
| 3 | Video coding concepts | 25 |
| 3.1 | Introduction | 25 |
| 3.2 | Video CODEC | 26 |
| 3.3 | Prediction model | 28 |
| 3.3.1 | Temporal prediction | 28 |
| 3.3.2 | Spatial model: intra prediction | 38 |
| 3.4 | Image model | 40 |
| 3.4.1 | Predictive image coding | 41 |
| 3.4.2 | Transform coding | 42 |
| 3.4.3 | Quantization | 50 |
| 3.4.4 | Reordering and zero encoding | 52 |
| 3.5 | Entropy coder | 57 |
| 3.5.1 | Predictive coding | 57 |
| 3.5.2 | Variable-length coding | 58 |
| 3.5.3 | Arithmetic coding | 65 |
| 3.6 | The hybrid DPCM/DCT video CODEC model | 68 |
| 3.7 | Summary | 79 |
| 3.8 | References | 79 |
| 4 | What is H.264? | 81 |
| 4.1 | Introduction | 81 |
| 4.2 | What is H.264? | 81 |
| 4.2.1 | A video compression format | 81 |
| 4.2.2 | An industry standard | 82 |
| 4.2.3 | A toolkit for video compression | 83 |
| 4.2.4 | Better video compression | 83 |
| 4.3 | How does an H.264 codec work? | 83 |
| 4.3.1 | Encoder processes | 85 |
| 4.3.2 | Decoder processes | 89 |
| 4.4 | The H.264/AVC Standard | 91 |
| 4.5 | H.264 Profiles and Levels | 92 |
| 4.6 | The H.264 Syntax | 94 |
| 4.7 | H.264 in practice | 97 |
| 4.7.1 | Performance | 97 |
| 4.7.2 | Applications | 98 |
| 4.8 | Summary | 98 |
| 4.9 | References | 98 |
| 5 | H.264 syntax | 99 |
| 5.1 | Introduction | 99 |
| 5.1.1 | A note about syntax examples | 99 |
| 5.2 | H.264 syntax | 100 |
| 5.3 | Frames, fields and pictures | 101 |

| | | |
|-------|--------------------------------------|------------|
| 5.3.1 | Decoding order | 104 |
| 5.3.2 | Display order | 104 |
| 5.3.3 | Reference picture lists | 106 |
| 5.3.4 | Frame and field coding | 111 |
| 5.4 | NAL unit | 114 |
| 5.5 | Parameter Sets | 115 |
| 5.6 | Slice layer | 117 |
| 5.6.1 | Slice types | 117 |
| 5.6.2 | Slice header | 117 |
| 5.6.3 | Slice data | 118 |
| 5.7 | Macroblock layer | 119 |
| 5.7.1 | Overview | 119 |
| 5.7.2 | The Intra PCM mode | 121 |
| 5.7.3 | Macroblock prediction | 122 |
| 5.7.4 | Residual data | 124 |
| 5.7.5 | Macroblock syntax examples | 127 |
| 5.8 | Summary | 134 |
| 5.9 | References | 135 |
| 6 | H.264 Prediction | 137 |
| 6.1 | Introduction | 137 |
| 6.2 | Macroblock prediction | 137 |
| 6.3 | Intra prediction | 138 |
| 6.3.1 | 4 x 4 luma prediction modes | 143 |
| 6.3.2 | 16 x 16 luma prediction modes | 146 |
| 6.3.3 | Chroma prediction modes | 147 |
| 6.3.4 | 8 x 8 luma prediction, High profiles | 148 |
| 6.3.5 | Signalling intra prediction modes | 148 |
| 6.4 | Inter prediction | 149 |
| 6.4.1 | Reference pictures | 151 |
| 6.4.2 | Interpolating reference pictures | 152 |
| 6.4.3 | Macroblock partitions | 157 |
| 6.4.4 | Motion vector prediction | 158 |
| 6.4.5 | Motion compensated prediction | 162 |
| 6.4.6 | Inter prediction examples | 164 |
| 6.4.7 | Prediction structures | 169 |
| 6.5 | Loop filter | 171 |
| 6.5.1 | Boundary strength | 172 |
| 6.5.2 | Filter decision | 173 |
| 6.5.3 | Filter implementation | 174 |
| 6.5.4 | Loop filter example | 174 |
| 6.6 | Summary | 177 |
| 6.7 | References | 177 |
| 7 | H.264 transform and coding | 179 |
| 7.1 | Introduction | 179 |

| | | |
|----------|--|------------|
| 7.2 | Transforms and quantization | 179 |
| 7.2.1 | The H.264 transforms | 179 |
| 7.2.2 | Transform processes | 180 |
| 7.2.3 | Integer transform and quantization: 4 x 4 blocks | 185 |
| 7.2.4 | Integer transform and quantization: 8 x 8 blocks | 198 |
| 7.2.5 | DC transforms | 203 |
| 7.2.6 | Transform and quantization extensions in the High profiles | 204 |
| 7.3 | Block scan orders | 206 |
| 7.4 | Coding | 207 |
| 7.4.1 | Exp-Golomb Coding | 208 |
| 7.4.2 | Context Adaptive Variable Length Coding, CAVLC | 210 |
| 7.4.3 | Context Adaptive Binary Arithmetic Coding, CABAC | 217 |
| 7.5 | Summary | 220 |
| 7.6 | References | 221 |
| 8 | H.264 conformance, transport and licensing | 223 |
| 8.1 | Introduction | 223 |
| 8.2 | Conforming to the Standard | 223 |
| 8.2.1 | Profiles | 224 |
| 8.2.2 | Levels | 226 |
| 8.2.3 | Hypothetical Reference Decoder | 230 |
| 8.2.4 | Conformance testing | 236 |
| 8.3 | H.264 coding tools for transport support | 237 |
| 8.3.1 | Redundant slices | 237 |
| 8.3.2 | Arbitrary Slice Order (ASO) | 238 |
| 8.3.3 | Slice Groups/Flexible Macroblock Order (FMO) | 238 |
| 8.3.4 | SP and SI slices | 240 |
| 8.3.5 | Data partitioned slices | 243 |
| 8.4 | Transport of H.264 data | 244 |
| 8.4.1 | Encapsulation in RBSPs, NALUs and packets | 244 |
| 8.4.2 | Transport protocols | 245 |
| 8.4.3 | File formats | 247 |
| 8.4.4 | Coding and transport issues | 247 |
| 8.5 | Supplemental Information | 248 |
| 8.5.1 | Supplemental Enhancement Information (SEI) | 248 |
| 8.5.2 | Video Usability Information (VUI) | 248 |
| 8.6 | Licensing H.264/AVC | 248 |
| 8.6.1 | Video coding patents | 250 |
| 8.6.2 | Video coding standards and patents | 252 |
| 8.6.3 | Licensing H.264/AVC patents | 252 |
| 8.7 | Summary | 253 |
| 8.8 | References | 253 |
| 9 | H.264 performance | 255 |
| 9.1 | Introduction | 255 |
| 9.2 | Experimentation with H.264 | 256 |

| | | |
|---------|---|------------|
| 9.2.1 | The JM Reference Software | 256 |
| 9.2.2 | Other software encoders/decoders | 261 |
| 9.2.3 | H.264 stream analysis | 263 |
| 9.3 | Performance comparisons | 265 |
| 9.3.1 | Performance criteria | 265 |
| 9.3.2 | Performance examples: Foreman sequence, QCIF resolution | 265 |
| 9.3.3 | Performance examples: Foreman and Container sequences | 269 |
| 9.3.4 | Performance examples: Inter prediction structures | 271 |
| 9.3.5 | Performance example: H.264 vs. MPEG-4 Visual | 273 |
| 9.4 | Rate control | 274 |
| 9.4.1 | Rate control in the JM reference encoder | 276 |
| 9.5 | Mode selection | 279 |
| 9.5.1 | Rate Distortion Optimized mode selection | 281 |
| 9.6 | Low complexity coding | 283 |
| 9.6.1 | Approximating the cost function | 283 |
| 9.6.2 | Reducing the set of tested modes | 284 |
| 9.6.3 | Early termination | 285 |
| 9.7 | Summary | 285 |
| 9.8 | References | 285 |
| 10 | Extensions and directions | 287 |
| 10.1 | Introduction | 287 |
| 10.2 | Scalable Video Coding | 288 |
| 10.2.1 | Simulcast transmission | 288 |
| 10.2.2 | Scalable transmission | 289 |
| 10.2.3 | Applications of Scalable Video Coding | 290 |
| 10.2.4 | Scalable Video Coding in H.264 | 290 |
| 10.2.5 | Temporal scalability | 292 |
| 10.2.6 | Quality scalability: overview | 294 |
| 10.2.7 | Spatial scalability: overview | 294 |
| 10.2.8 | Spatial scalability in detail | 294 |
| 10.2.9 | Quality scalability in detail | 298 |
| 10.2.10 | Combined scalability | 299 |
| 10.2.11 | SVC performance | 299 |
| 10.3 | Multiview Video Coding | 302 |
| 10.3.1 | H.264 Multiview Video Coding | 304 |
| 10.4 | Configurable Video Coding | 306 |
| 10.4.1 | MPEG Reconfigurable Video Coding | 307 |
| 10.4.2 | Fully Configurable Video Coding | 308 |
| 10.5 | Beyond H.264/AVC | 310 |
| 10.6 | Summary | 310 |
| 10.7 | References | 311 |
| | Index | 313 |