

THE H.264 ADVANCED VIDEO COMPRESSION STANDARD

Second Edition

Iain E. Richardson

Vcodex Limited, UK



A John Wiley and Sons, Ltd., Publication

THE H.264 ADVANCED VIDEO COMPRESSION STANDARD

THE H.264 ADVANCED VIDEO COMPRESSION STANDARD

Second Edition

Iain E. Richardson

Vcodex Limited, UK



A John Wiley and Sons, Ltd., Publication

This edition first published 2010
© 2010, John Wiley & Sons, Ltd

First Edition published in 2003

Registered office

John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloguing-in-Publication Data

Richardson, Iain E. G.

The H.264 advanced video compression standard / Iain E. Richardson. – 2nd ed.

p. cm.

Rev. ed. of: H.264 and MPEG-4 video compression. c2003.

Includes bibliographical references and index.

ISBN 978-0-470-51692-8 (pbk.)

1. Digital video—Standards. 2. Video compression—Standards. 3. MPEG (Video coding standard) 4. Multimedia systems. I. Richardson, Iain E. G. H.264 and MPEG-4 video compression. II. Title.

TK6680.5.R52 2010

006.6'96—dc22

2009054387

A catalogue record for this book is available from the British Library.

ISBN: 978-0-470-51692-8

Typeset in 10/12pt Times by Aptara Inc., New Delhi, India

Printed and bound in Great Britain by CPI Autony Rowe, Chippenham, Wiltshire

To Pat

*Language is living,
but what is most important
goes deeper than words.*

Contents

About the Author	xiii
Preface	xv
Glossary	xvii
List of Figures	xxi
List of Tables	xxix
1 Introduction	1
1.1 A change of scene	1
1.2 Driving the change	4
1.3 The role of standards	4
1.4 Why H.264 Advanced Video Coding is important	4
1.5 About this book	5
1.6 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	4×4 luma prediction modes	143
6.3.2	16×16 luma prediction modes	146
6.3.3	Chroma prediction modes	147
6.3.4	8×8 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	Transform and quantization	179
7.2.1	The H.264 transforms	179
7.2.2	Transform processes	180
7.2.3	Integer transform and quantization: 4×4 blocks	185
7.2.4	Integer transform and quantization: 8×8 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	Experimenting 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

About the Author

Professor Iain Richardson is an internationally known expert on the MPEG and H.264 video compression standards.

The author of *H.264 and MPEG-4 Video Compression*, a widely cited work in the research literature, Professor Richardson has written two further books and over 70 journal and conference papers on video compression. He regularly advises companies on video compression technology, video coding patents and company acquisitions in the video coding industry. Professor Richardson leads an internationally renowned video coding research team, contributes to the MPEG industry standards group and is sought after as an expert witness. Based in Aberdeen, Scotland, he regularly travels to the US and Europe.

Preface

The last decade has seen a quiet revolution in digital video technology. Digital video is everywhere: on our televisions, our DVD and Blu-Ray players, our computers, our music players and our mobile handsets. Only recently, a video image in a web page was an unusual sight. Nowadays, many of us are just as likely to catch the latest news on the web as on the TV. With the explosion of digital video applications, a billion-dollar industry has developed and expanded, with new companies and niche markets emerging, thriving and disappearing faster than anyone can easily track. Video compression is essential to all of these applications and markets, and the H.264 format is considered by many to be the state of the art in video compression.

When I wrote the first edition of this book in 2003, H.264 Advanced Video Compression had just been published as an International Standard and it was hard to predict its impact on industry. Its predecessor, MPEG-4 Visual, had arguably failed to live up to its promise, with only limited adoption in the market. Since 2003, the significant performance improvements that are built into H.264 have made it the clear successor to the older MPEG video standards in many applications, from mobile video to High Definition broadcasting. At the time of writing, the MPEG and VCEG standards committees are debating the possible successor to H.264. It is likely to be several years before a new standard is released, and several years after that before H.264 begins to become obsolete.

This book is intended to be a practical, accessible and unbiased guide to the H.264 video compression standard. As always, I have chosen to explain the details of H.264 in my own way, concentrating on what I feel is important to the engineer, researcher or student who needs a 'way in' to this complex yet important technical subject. This book is not the final word on H.264. By definition, that final word is provided by the standard itself and I advise any serious developer or implementer of H.264 to get hold of a copy of the standard. There is a need for a guidebook to the standard that explains the concepts, tools, benefits and disadvantages of the format, just as a good guidebook helps the tourist to get to know a foreign country and to become more at home there. Some visitors may be disappointed that their favourite subject is not covered in as much depth as they would like. I have made a deliberate choice to cover certain topics such as Scalable and Multiview Video Coding only briefly as they are still, in my view, in the early stages of practical implementation.

My sincere thanks to the many people who have helped to shape this book, including the readers of my earlier books who told me what they liked and what they wanted; the many companies and individuals who have asked me to solve their video compression problems; Kourosh Soroushian for discussions on Hypothetical Reference Decoders; Abharana Bhat,

Maja Bystrom, Sam Jansen, Sampath Kannangara and Yafan Zhao for reading and commenting on draft chapters; Gary Sullivan for many comments, corrections, suggestions and discussions; Nicky, Simone and the editorial team at John Wiley & Sons; and to Pat for reading the manuscript, cracking the whip and making me finish it.

I hope that you find the book useful; more importantly, I hope you enjoy it. Visit my website at www.vcodex.com and *tell me what you think*.

Iain Richardson
Aberdeen, 2010

Glossary

4:2:0 (sampling)	Sampling method: chrominance components have half the horizontal and vertical resolution of luminance component
4:2:2 (sampling)	Sampling method: chrominance components have half the horizontal resolution of luminance component
4:4:4 (sampling)	Sampling method: chrominance components have same resolution as luminance component
access unit	Complete coded frame or field
arithmetic coding	Coding method to reduce redundancy
artefact	Visual distortion in an image
ASO	Arbitrary Slice Order, in which slices may be coded out of raster sequence
block	Region of macroblock
block matching	Motion estimation carried out on rectangular picture areas
blocking	Square or rectangular distortion areas in an image
B slice	Coded slice predicted using bidirectional motion compensation
CABAC	Context-based Adaptive Binary Arithmetic Coding
CAVLC	Context Adaptive Variable Length Coding
chrominance or chroma	Colour difference component
CIF	Common Intermediate Format, a colour image format
CODEC	COder / DECoder pair
Coded Picture Buffer (CPB)	Buffer containing coded frames or fields
colour space	Method of representing colour images
DCT	Discrete Cosine Transform, a mathematical transform and/or its practical approximation(s)
direct prediction	A coding mode in which no motion vector is transmitted
DPCM	Differential Pulse Code Modulation

DSCQS	Double Stimulus Continuous Quality Scale, a scale and method for subjective quality measurement
DWT	Discrete Wavelet Transform
entropy coding	Coding method to reduce redundancy
error concealment	Post-processing of a decoded image to remove or reduce visible error effects
Exp-Golomb or ExpG	Exponential Golomb variable length codes
field	Odd- or even-numbered lines from an interlaced video sequence
FMO	Flexible Macroblock Order, in which macroblocks may be coded out of raster sequence
Full Search	A motion estimation algorithm
Fully Configurable Video Coding	A framework for video coding in which a codec may be completely re-configured during a communication session
GOP	Group of Pictures, a set of coded video images
H.261	A video coding standard
H.263	A video coding standard
H.264	A video coding standard
HDTV	High Definition Television
Huffman coding	Coding method to reduce redundancy
HVS	Human Visual System, the system by which humans perceive and interpret visual images
hybrid (CODEC)	CODEC model featuring motion compensation and transform
Hypothetical Reference Decoder (HRD)	Decoder ‘model’ that may be used to test bitstream conformance
IEC	International Electrotechnical Commission, a standards body
inter (coding)	Coding of video frames using temporal prediction or compensation
interlaced (video)	Video data represented as a series of fields
intra (coding)	Coding of video frames without temporal prediction
I slice	Slice coded without reference to any other frame
ISO	International Standards Organisation, a standards body
ITU	International Telecommunication Union, a standards body
JPEG	Joint Photographic Experts Group, a committee of ISO (also an image coding standard)
latency	Delay through a communication system
Level	A set of conformance parameters (applied to a Profile)

loop filter	Spatial filter placed within encoding or decoding feedback loop
luminance or luma	Monochrome or brightness component
Macroblock	Region of frame coded as a unit (usually 16×16 pixels in the original frame)
Macroblock partition	Region of macroblock with its own motion vector
Macroblock sub-partition	Region of macroblock with its own motion vector
motion compensation	Prediction of a video frame with modelling of motion
motion estimation	Estimation of relative motion between two or more video frames
motion vector	Vector indicating a displaced block or region to be used for motion compensation
MPEG	Motion Picture Experts Group, a committee of ISO/IEC
MPEG-1	A multimedia coding standard
MPEG-2	A multimedia coding standard
MPEG-4	A multimedia coding standard
MVC	Multiview Video Coding, in which multiple views of a scene may be jointly coded
NAL	Network Abstraction Layer
objective quality	Visual quality measured by algorithm(s)
Picture (coded)	Coded (compressed) video frame
P-picture (slice)	Coded picture (or slice) using motion-compensated prediction from one reference frame
profile	A set of functional capabilities (of a video CODEC)
progressive (video)	Video data represented as a series of complete frames
PSNR	Peak Signal to Noise Ratio, an objective quality measure
QCIF	Quarter Common Intermediate Format
quantize	Reduce the precision of a scalar or vector quantity
rate control	Control of bit rate of encoded video signal
rate-distortion	Measure of CODEC performance (distortion at a range of coded bit rates)
RBSP	Raw Byte Sequence Payload
RVC	Reconfigurable Video Coding, a framework for video coding in which a decoder may be constructed from pre-defined Functional Units.
RGB	Red/Green/Blue colour space
ringing (artefacts)	'Ripple'-like artefacts around sharp edges in a decoded image
RTP	Real Time Protocol, a transport protocol for real-time data

scalable coding	Coding a signal into a number of layers
SVC	Scalable Video Coding
SI slice	Intra-coded slice used for switching between coded bitstreams (H.264)
slice	A region of a coded picture
SP slice	Inter-coded slice used for switching between coded bitstreams
statistical redundancy	Redundancy due to the statistical distribution of data
studio quality	Lossless or near-lossless video quality
subjective quality	Visual quality as perceived by human observer(s)
subjective redundancy	Redundancy due to components of the data that are subjectively insignificant
sub-pixel (motion compensation)	Motion-compensated prediction from a reference area that may be formed by interpolating between integer-valued pixel positions
test model	A software model and document that describe a reference implementation of a video coding standard
texture	Image or residual data
tree-structured motion compensation	Motion compensation featuring a flexible hierarchy of partition sizes
VCEG	Video Coding Experts Group, a committee of ITU
VCL	Video Coding Layer
video packet	Coded unit suitable for packetization
VLC	Variable Length Code
VLD	Variable Length Decoder
VLE	Variable Length Encoder
VLSI	Very Large Scale Integrated circuit
VQEG	Video Quality Experts Group
weighted prediction	Motion compensation in which the prediction samples from two references are scaled
YCrCb	Luminance/Red chrominance/Blue chrominance colour space

List of Figures

1.1	Video coding scenarios, one-way	3
1.2	Video coding scenario, two-way	3
2.1	Still image from natural video scene	8
2.2	Spatial and temporal sampling of a video sequence	8
2.3	Image with two sampling grids	9
2.4	Image sampled at coarse resolution (black sampling grid)	10
2.5	Image sampled at finer resolution (grey sampling grid)	10
2.6	Interlaced video sequence	11
2.7	Top field	12
2.8	Bottom field	13
2.9	Red, Green and Blue components of colour image	13
2.10	Cr, Cg and Cb components	14
2.11	4:2:0, 4:2:2 and 4:4:4 sampling patterns (progressive)	15
2.12	Allocation of 4:2:0 samples to top and bottom fields	17
2.13	Video frame sampled at range of resolutions	18
2.14	SD and HD formats	19
2.15	DSCQS testing system	21
2.16	PSNR examples: (a) Original; (b) 30.6dB; (c) 28.3dB	22
2.17	Image with blurred background (PSNR = 27.7dB)	22
3.1	Encoder/Decoder	26
3.2	Spatial and temporal correlation in a video sequence	26
3.3	Video encoder block diagram	27
3.4	Frame 1	29
3.5	Frame 2	29
3.6	Difference	29
3.7	Optical flow	30
3.8	Macroblock (4:2:0)	31
3.9	Motion estimation	32
3.10	Frame 1	33
3.11	Frame 2	33
3.12	Residual : no motion compensation	33
3.13	Residual : 16 × 16 block size	34
3.14	Residual : 8 × 8 block size	34
3.15	Residual : 4 × 4 block size	34

3.16	Close-up of reference region	35
3.17	Reference region interpolated to half-pixel positions	36
3.18	Integer, half-pixel and quarter-pixel motion estimation	36
3.19	Residual : 4×4 blocks, 1/2-pixel compensation	37
3.20	Residual : 4×4 blocks, 1/4-pixel compensation	37
3.21	Motion vector map : 16×16 blocks, integer vectors	38
3.22	Motion vector map : 4×4 blocks, 1/4-pixel vectors	39
3.23	Intra prediction: available samples	39
3.24	Intra prediction: spatial extrapolation	39
3.25	2D autocorrelation function of image	40
3.26	2D autocorrelation function of residual	41
3.27	Spatial prediction (DPCM)	41
3.28	4×4 DCT basis patterns	45
3.29	8×8 DCT basis patterns	45
3.30	Image section showing 4×4 block	46
3.31	Close-up of 4×4 block; DCT coefficients	47
3.32	Block reconstructed from (a) 1, (b) 2, (c) 3, (d) 5 coefficients	47
3.33	Two-dimensional wavelet decomposition process	48
3.34	Image after one level of decomposition	49
3.35	Two-stage wavelet decomposition of image	49
3.36	Five-stage wavelet decomposition of image	50
3.37	Scalar quantizers: linear; non-linear with dead zone	52
3.38	Vector quantization	53
3.39	8×8 DCT coefficient distribution (frame)	53
3.40	Residual field picture	54
3.41	8×8 DCT coefficient distribution (field)	54
3.42	Zigzag scan example : frame block	55
3.43	Zigzag scan example : field block	55
3.44	Wavelet coefficient and 'children'	57
3.45	Motion vector prediction candidates	58
3.46	Generating the Huffman code tree: Sequence 1 motion vectors	60
3.47	Huffman tree for sequence 2 motion vectors	61
3.48	MPEG4 TCOEF VLCs (partial)	64
3.49	Sub-range example	66
3.50	Arithmetic coding example	67
3.51	DPCM/DCT video encoder	69
3.52	DPCM/DCT video decoder	69
3.53	Input frame F_n	71
3.54	Reconstructed reference frame F'_{n-1}	71
3.55	Residual $F_n - F'_{n-1}$: no motion compensation	72
3.56	16×16 motion vectors superimposed on frame	73
3.57	Motion compensated reference frame	73
3.58	Motion compensated residual frame	74
3.59	Original macroblock : luminance	74
3.60	Residual macroblock : luminance	75
3.61	DCT coefficient magnitudes : top-right 8×8 block	76

3.62	Comparison of original and decoded residual blocks	78
3.63	Decoded frame F'_n	79
4.1	The H.264 video coding and decoding process	82
4.2	Video coding: source frames, encoded bitstream, decoded frames	83
4.3	Video codec: high level view	84
4.4	Typical H.264 encoder	84
4.5	Typical H.264 decoder	85
4.6	Prediction: flow diagram	86
4.7	Intra prediction	86
4.8	Original macroblock, intra prediction and residual	86
4.9	Inter prediction	87
4.10	Original macroblock, inter prediction and residual	87
4.11	Forward transform	88
4.12	Quantization example	88
4.13	Example: Block, transform coefficients, quantized coefficients	88
4.14	Rescaling example	89
4.15	Inverse transform: combining weighted basis patterns to create a 4×4 image block	90
4.16	Rescaled coefficients and inverse transform output	90
4.17	Reconstruction flow diagram	91
4.18	Profiles and Levels: example	93
4.19	H.264 syntax : overview	94
4.20	Syntax example: P-macroblock	95
4.21	P-macroblock decoding process	96
4.22	Residual luma and chroma coefficient blocks	97
4.23	A video frame compressed at the same bitrate using MPEG-2 (left), MPEG-4 Visual (centre) and H.264 compression (right)	97
5.1	Syntax overview	100
5.2	Picture handling in H.264, overview	103
5.3	Decoded Picture Buffer and picture orders	104
5.4	Display order: Type 0 example	105
5.5	Display order: Type 1 example	106
5.6	List 0 and List 1 ordering: example	108
5.7	Default reference pictures: example	108
5.8	Reference picture re-ordering syntax, simplified overview	109
5.9	Picture Adaptive Frame Field Coding example	112
5.10	Frame with macroblock pairs	113
5.11	MB pair coded using Macroblock Adaptive Frame Field Coding	113
5.12	Example: Sequence and Picture Parameter Sets	116
5.13	Macroblock layer syntax overview	119
5.14	mb_pred and sub_mb_pred syntax overview	123
5.15	Residual data syntax overview	125
5.16	Block scanning order, 4×4 transform, 4:2:0 sampling	125
5.17	Block scanning order, 8×8 luma transform, 4:2:0 sampling	126
5.18	Residual CAVLC block syntax overview	126
5.19	Residual CABAC block syntax overview	127

5.20	P macroblock, example 2	129
5.21	P macroblock example 4	131
5.22	B macroblock, example 5	133
6.1	Example of macroblock types and prediction sources	138
6.2	Intra prediction: adjacent blocks example	139
6.3	Intra prediction source samples, 4×4 or 8×8 luma blocks	140
6.4	Intra prediction source samples, chroma or 16×16 luma blocks	140
6.5	Example of intra block size choices, CIF, Baseline Profile. Reproduced by permission of Elecard.	141
6.6	QCIF frame with highlighted macroblock	142
6.7	Predicted luma frame formed using H.264 intra prediction	142
6.8	Residual after subtracting intra prediction	143
6.9	4×4 luma block to be predicted	144
6.10	Labelling of prediction samples, 4×4 prediction	144
6.11	4×4 intra prediction modes	144
6.12	Prediction blocks, 4×4 modes 0–8	145
6.13	Intra 16×16 prediction modes	146
6.14	16×16 macroblock	146
6.15	Prediction blocks, intra 16×16 modes 0–3	147
6.16	Intra mode prediction example	148
6.17	P macroblock prediction example	150
6.18	Example of integer and sub-pixel prediction	152
6.19	Current region	153
6.20	4×4 block to be predicted	153
6.21	Reference region	153
6.22	Prediction from integer samples	154
6.23	Reference region, half-pixel interpolated	154
6.24	Prediction from interpolated samples	154
6.25	Interpolation of luma half-pel positions	155
6.26	Interpolation of luma quarter-pel positions	156
6.27	Luma region interpolated to quarter-pel positions	156
6.28	Interpolation of chroma eighth-pel positions	157
6.29	Macroblock partitions and sub-macroblock partitions	158
6.30	Current and neighbouring partitions : same partition sizes	159
6.31	Current and neighbouring partitions : different partition sizes	159
6.32	Scaled motion vector example	160
6.33	Temporal direct motion vector example	161
6.34	Forming a motion compensated prediction	162
6.35	Biprediction example	163
6.36	MBAFF: prediction from corresponding field	164
6.37	P slice showing partition choices. Reproduced by permission of Elecard	165
6.38	B slice showing macroblock modes. Light-shaded circles are skipped macroblocks. Reproduced by permission of Elecard.	166
6.39	Inter prediction example, P slice	167
6.40	Inter prediction example, B slice	168
6.41	Low delay prediction structure	169

6.42	'Classic' Group of Pictures prediction structure	169
6.43	IPPP. . . with multiple reference pictures	170
6.44	Hierarchical GOP structure	171
6.45	Edge filtering order in a macroblock	172
6.46	Pixels adjacent to vertical and horizontal boundaries	172
6.47	16×16 luma macroblock showing block edges	173
6.48	Original frame, 'violin' frame 2	174
6.49	Reconstructed, QP = 36, no filter	175
6.50	Reconstructed, QP = 36, with filter	175
6.51	Reconstructed, QP = 32, no filter	176
6.52	Reconstructed, QP = 32, with filter	176
7.1	Re-scaling and inverse transform	181
7.2	Forward transform and quantization	181
7.3	Luma forward transform: default	182
7.4	Luma inverse transform: default	182
7.5	Luma forward transform: Intra 16×16 mode	183
7.6	Luma inverse transform: Intra 16×16 mode	183
7.7	Luma forward transform: 8×8 transform	183
7.8	Luma inverse transform: 8×8 transform	183
7.9	Chroma forward transform: 4:2:0 macroblock	184
7.10	Chroma inverse transform: 4:2:0 macroblock	184
7.11	Chroma forward transform: 4:2:2 macroblock	184
7.12	Chroma inverse transform: 4:2:2 macroblock	184
7.13	Development of the forward transform and quantization process	185
7.14	Development of the rescaling and inverse transform process	186
7.15	Quantization parameter QP vs. effective quantizer step size, logarithmic y-axis	192
7.16	Frequency dependent quantization, 4×4 block	204
7.17	Progressive scan orders for 4×4 and 8×8 blocks	206
7.18	Field scan orders for 4×4 and 8×8 blocks	207
7.19	CAVLC encoder overview	211
7.20	CABAC coding process overview	218
8.1	Baseline, Constrained Baseline, Extended and Main Profiles	225
8.2	Main and High Profiles	226
8.3	Main and Intra Profiles	227
8.4	Selected Level constraints	228
8.5	Selected display resolutions	229
8.6	H.264 encoder and decoder buffers	230
8.7	Hypothetical Reference Decoder (HRD)	230
8.8	HRD example 1: encoder buffer	232
8.9	HRD example 1: decoder CPB	232
8.10	HRD Example 2: encoder buffer	234
8.11	HRD Example 2: decoder CPB	234
8.12	HRD Example 3: encoder buffer	235
8.13	HRD Example 3: decoder CPB	236
8.14	Bitstream conformance testing	236

8.15	Decoder conformance testing	237
8.16	Arbitrary Slice Order: Example	238
8.17	FMO: Interleaved map, QCIF, 3 slice groups	239
8.18	FMO: Dispersed macroblock map, QCIF, 4 slice groups	239
8.19	FMO: Foreground and Background map, 4 slice groups	240
8.20	FMO: Box-out, Raster and Wipe maps	240
8.21	Switching streams using I-slices	241
8.22	Switching streams using SP-slices	242
8.23	Encoding SP-slice A_2 (simplified)	242
8.24	Encoding SP-slice B_2 (simplified)	242
8.25	Decoding SP-slice A_2 (simplified)	243
8.26	Encoding SP-slice AB_2 (simplified)	243
8.27	Decoding SP-slice AB_2	244
8.28	Fast-forward using SP-slices	244
8.29	Encapsulation of H.264 syntax elements	245
8.30	MPEG-2 Transport Stream	246
8.31	RTP packet structure (simplified)	246
8.32	ISO Media File	247
8.33	Block diagram from US patent 3679821 (redrawn)	251
8.34	Issued US patents including the terms ‘video coding’ or ‘video compression’, 1990–2007. Source: USPTO patent database.	251
9.1	JM software operation	256
9.2	Planar YCbCr file format, 4:2:0 sampling	257
9.3	JM encoder configuration file	258
9.4	JM encoder output display	258
9.5	Original, reconstructed and decoded frames, container.qcif, QP = 32	259
9.6	Section of coded frame, JM encoder (left), $\times 264$ encoder (right)	262
9.7	Screenshot of an H.264 stream analyzer: Baseline Profile frame. Reproduced by permission of Elecard	263
9.8	Screenshot of stream analyzer: Main Profile frame. Courtesy of Elecard	264
9.9	Foreman/QCIF/Basic complexity	266
9.10	Foreman/QCIF/Basic complexity and options	267
9.11	Foreman/QCIF/Basic and Medium Complexity	268
9.12	Foreman/QCIF/Medium complexity with rate control	269
9.13	Foreman, QCIF sequences: coding time	269
9.14	QCIF sequences: rate vs. PSNR	270
9.15	CIF sequences: rate vs. PSNR	270
9.16	CIF sequences: coding time	271
9.17	Rate-distortion comparison of prediction structures	271
9.18	Sample frames from sequences using different prediction structures, coded at 280kbps	272
9.19	Carphone, QCIF: H.264 vs. MPEG-4 Visual	273
9.20	Frame from ‘Foreman’ sequence showing macroblock sizes	274
9.21	Encoder with rate feedback	275
9.22	Bitrate allocation for rate control	275
9.23	Foreman, QCIF, 100 frames: coded bitrate	278

9.24	Foreman, QCIF, 100 frames: QP per frame	278
9.25	Foreman, QCIF, 100 frames: Luma PSNR per frame	279
9.26	Available macroblock prediction modes	280
9.27	Rate and MSE costs for different coding options	281
9.28	Rate, quality and complexity	283
10.1	Multiple streams/simulcast	288
10.2	Multiple streams/scalable	289
10.3	Overview of scalability types	291
10.4	Temporally scalable sequence, 3 layers	292
10.5	Decoding a temporally scalable sequence	293
10.6	Hierarchical prediction structure	293
10.7	Quality Scalability	295
10.8	Spatial scalability, two layers	296
10.9	Medium Grain Quality Scalability	298
10.10	Combined Spatial, Temporal and Quality scalability	299
10.11	Simulcast vs. scalable bitrates	300
10.12	Quality scalability at CIF resolution	301
10.13	Spatial + temporal scalability, CIF \rightarrow 4CIF resolution	301
10.14	Spatial scalability, 720p \rightarrow 1080p resolution	302
10.15	Three views of the same scene	302
10.16	Multiview video: view examples	303
10.17	Multiview video: views and frames	303
10.18	Inter-view prediction of key frames	304
10.19	Inter-view prediction of all frames	305
10.20	Overview of configurable video codec	306
10.21	Reconfigurable Video Coding scenario	307
10.22	Fully Configurable Video Coding framework	309

