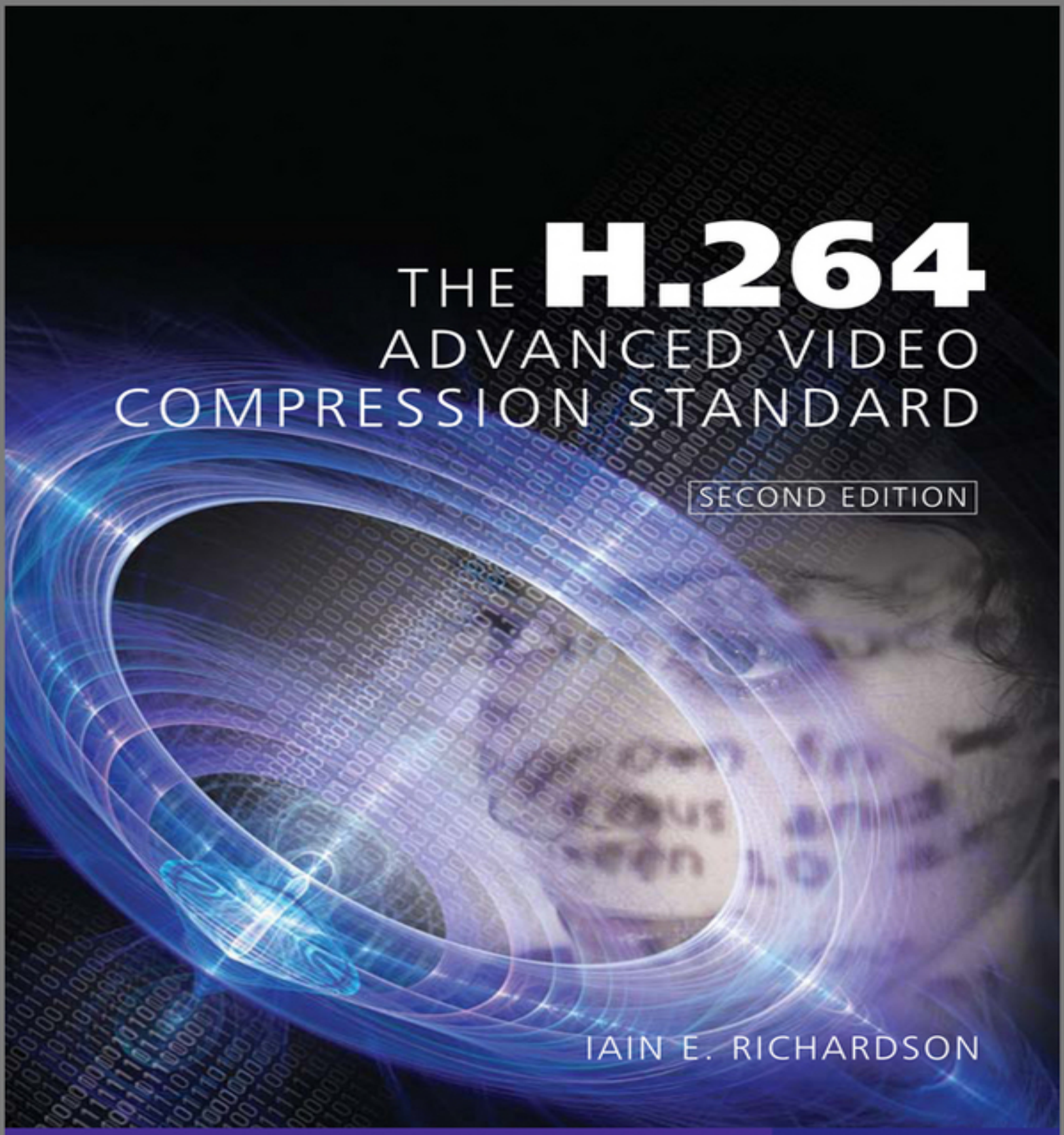
THE H.264 ADVANCED VIDEO COMPRESSION STANDARD

SECOND EDITION



IAIN E. RICHARDSON

WILEY

Contents

THE H.264 ADVANCED VIDEO COMPRESSION STANDARD

Second Edition

Iain E. Richardson
Vcodex Limited, UK

WILEY
A John Wiley and Sons, Ltd., Publication

This edition first published 2010

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.

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

ISBN: 978-0-470-51692-8

To Pat

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

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, MajaBystrom, SamJansen, 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

List of Tables

1

Introduction

1.1 A change of scene

2000:

Most viewers receive analogue television via terrestrial, cable or satellite transmission.

VHS video tapes are the principal medium for recording and playing TV programs, movies, etc.

Cell phones are cell phones, i.e. a mobile handset can only be used to make calls or send SMS messages.

Internet connections are slow, primarily over telephone modems for home users.

Web pages are web pages, with static text, graphics and photos and not much else.

Video calling requires dedicated videoconferencing terminals and expensive leased lines. Video calling over the internet is possible but slow, unreliable and difficult to set up.

Consumer video cameras, camcorders, use tape media, principally analogue tape. Home-made videos generally stay within the home.

2010:

Most viewers receive digital television via terrestrial, cable, satellite or internet, with benefits such as a greater choice of channels, electronic programme guides andhighdefinitionservices.AnalogueTVhasbeenswitchedo