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 and high definition services. Analogue TV has been switched off in many countries. Many TV programmes can be watched via the internet.

DVDs are the principal medium for playing pre-recorded movies and TV programs. Many alternatives exist, most of them digital, including internet movie downloading (legal and not-so-legal), hard-disk recording and playback and a

The H.264 Advanced Video Compression Standard, Second Edition Iain E. Richardson

variety of digital media formats. High definition DVDs, Blu-Ray Disks, are increasing in popularity.

Cell phones function as cameras, web browsers, email clients, navigation systems, organizers and social networking devices. Occasionally they are used to make calls.

Home internet access speeds continue to get faster via broadband and mobile connections, enabling widespread use of video-based web applications.

Web pages are applications, movie players, games, shopping carts, bank tellers, social networks, etc, with content that changes dynamically.

Video calling over the internet is commonplace with applications such as Skype and iChat. Quality is still variable but continues to improve.

Consumer video cameras use hard disk or flash memory card media. Editing, uploading and internet sharing of home videos is widespread.

A whole range of illegal activities has been born – DVD piracy, movie sharing via the internet, recording and sharing of assaults, etc.

Video footage of breaking news items such as the Chilean earthquake is more likely to come from a cell phone than a TV camera.

All these changes in a ten-year period signify a small revolution in the way we create, share and watch moving images. Many factors have contributed to the shift towards digital video – commercial factors, legislation, social changes and technological advances. From the technology viewpoint, these factors include better communications infrastructure, with widespread, relatively inexpensive access to broadband networks, 3G mobile networks, cheap and effective wireless local networks and higher-capacity carrier transmission systems; increasingly sophisticated devices, with a bewildering array of capabilities packed into a lightweight cellular handset; and the development of easy-to-use applications for recording, editing, sharing and viewing video material. This book will focus on one technical aspect that is key to the widespread adoption of digital video technology – video compression.

Video compression or video encoding is the process of reducing the amount of data required to represent a digital video signal, prior to transmission or storage. The complementary operation, decompression or decoding, recovers a digital video signal from a compressed representation, prior to display. Digital video data tends to take up a large amount of storage or transmission capacity and so video encoding and decoding, or **video coding**, is essential for any application in which storage capacity or transmission bandwidth is constrained. Almost all consumer applications for digital video fall into this category, for example:

- Digital television broadcasting: TV programmes are coded prior to transmission over a limited-bandwidth terrestrial, satellite or cable channel (Figure 1.1).
- Internet video streaming: Video is coded and stored on a server. The coded video is transmitted (streamed) over the internet, decoded on a client and displayed (Figure 1.1).

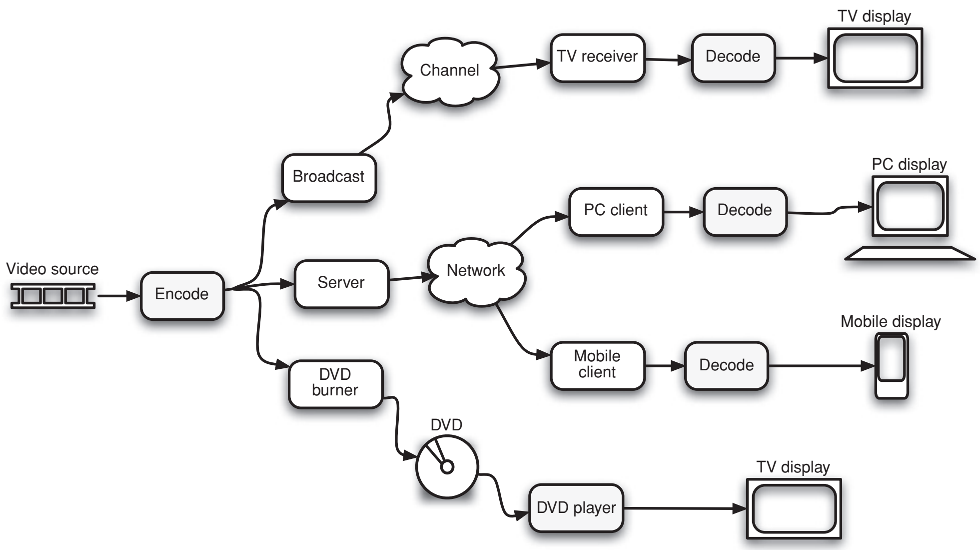

Figure 1.1 Video coding scenarios, one-way

- ◦ Mobile video streaming: As above, but the coded video is transmitted over a mobile network such as GPRS or 3G (Figure 1.1).
- ◦ DVD video: Source video is coded and stored on a DVD or other storage medium. A DVD player reads the disk and decodes video for display (Figure 1.1).
- ◦ Video calling: Each participant includes an encoder and a decoder (Figure 1.2). Video from a camera is encoded and transmitted across a network, decoded and displayed. This occurs in two directions simultaneously.

Each of these examples includes an encoder, which compresses or encodes an input video signal into a coded bitstream, and a decoder, which decompresses or decodes the coded bitstream to produce an output video signal. The encoder or decoder is often built in to a device such as a video camera or a DVD player.

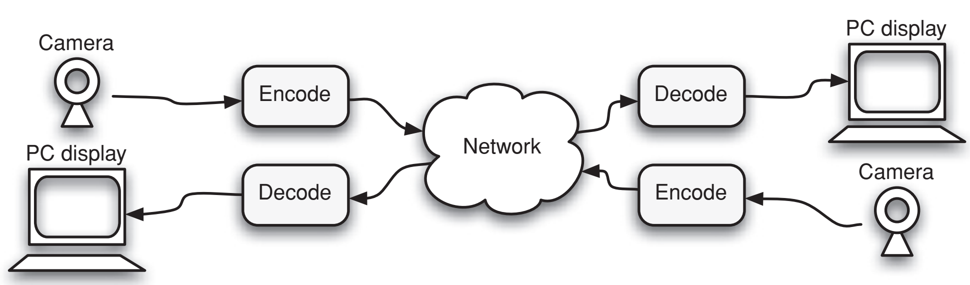

Figure 1.2 Video coding scenario, two-way

1.2 Driving the change

The consumer applications discussed above represent very large markets. The revenues involved in digital TV broadcasting and DVD distribution are substantial. Effective video coding is an essential component of these applications and can make the difference between the success or failure of a business model. A TV broadcasting company that can pack a larger number of high-quality TV channels into the available transmission bandwidth has a market edge over its competitors. Consumers are increasingly discerning about the quality and performance of video-based products and there is therefore a strong incentive for continuous improvement in video coding technology. Even though processor speeds and network bandwidths continue to increase, a better video codec results in a better product and therefore a more competitive product. This drive to improve video compression technology has led to significant investment in video coding research and development over the last 15-20 years and to rapid, continuous advances in the state of the art.

1.3 The role of standards

Many different techniques for video coding have been proposed and researched. Hundreds of research papers are published each year describing new and innovative compression techniques. In contrast to this wide range of innovations, commercial video coding applications tend to use a limited number of standardized techniques for video compression. Standardized video coding formats have a number of potential benefits compared with non-standard, proprietary formats:

- Standards simplify inter-operability between encoders and decoders from different manufacturers. This is important in applications where each 'end' of the system may be produced by a different company, e.g. the company that records a DVD is typically not the same as the company that manufactures a DVD player.
- Standards make it possible to build platforms that incorporate video, in which many different applications such as video codecs, audio codecs, transport protocols, security and rights management, interact in well-defined and consistent ways.
- Many video coding techniques are patented and therefore there is a risk that a particular video codec implementation may infringe patent(s). The techniques and algorithms required to implement a standard are well-defined and the cost of licensing patents that cover these techniques, i.e. licensing the right to use the technology embodied in the patents, can be clearly defined.

Despite recent debates about the benefits of royalty-free codecs versus industry standard video codecs [i], video coding standards are very important to a number of major industries. With the ubiquitous presence of technologies such as DVD/Blu-Ray, digital television, internet video and mobile video, the dominance of video coding standards is set to continue for some time to come.

1.4 Why H.264 Advanced Video Coding is important

This book is about a standard, jointly published by the International Telecommunications Union (ITU) and the International Standards Organisation (ISO) and known by several names:

'H.264', 'MPEG-4 Part 10' and 'Advanced Video Coding'. The standard itself is a document over 550 pages long and filled with highly technical definitions and descriptions. Developed by a team consisting of hundreds of video compression experts, the Joint Video Team, a collaborative effort between the Moving Picture Experts Group (MPEG) and the Video Coding Experts Group (VCEG), this document is the culmination of many man-years' work. It is almost impossible to read and understand without an in-depth knowledge of video coding.

Why write a book about this document? Whilst the standard itself is arguably only accessible to an insider expert, H.264/AVC has huge significance to the broadcast, internet, consumer electronics, mobile and security industries, amongst others. H.264/AVC is the latest in a series of standards published by the ITU and ISO. It describes and defines a method of coding video that can give better performance than any of the preceding standards. H.264 makes it possible to compress video into a smaller space, which means that a compressed video clip takes up less transmission bandwidth and/or less storage space compared to older codecs. A combination of market expansion, technology advances and increased user expectation is driving demand for better, higher quality digital video. For example:

- TV companies are delivering more content in High Definition. Most new television sets can display HD pictures. Customers who pay a premium for High Definition content expect correspondingly high image quality.
- An ever-increasing army of users are uploading and downloading videos using sites such as YouTube. Viewers expect rapid download times and high resolution.
- Recording and sharing videos using mobile handsets is increasingly commonplace.
- Internet video calls, whilst still variable in quality, are easier to make and more widely used than ever.
- The original DVD-Video format, capable of supporting only a single movie in Standard Definition seems increasingly limited.

In each case, better video compression is the key to delivering more, higher-quality video in a cost effective way. H.264 compression makes it possible to transmit HD television over a limited-capacity broadcast channel, to record hours of video on a Flash memory card and to deliver massive numbers of video streams over an already busy internet.

The benefits of H.264/AVC come at a price. The standard is complex and therefore challenging to the engineer or designer who has to develop, program or interface with an H.264 codec. H.264 has more options and parameters – more 'control knobs' – than any previous standard codec. Getting the controls and parameters 'right' for a particular application is not an easy task. Get it right and H.264 will deliver high compression performance; get it wrong and the result is poor-quality pictures and/or poor bandwidth efficiency. Computationally expensive, an H.264 coder can lead to slow coding and decoding times or rapid battery drain on handheld devices. Finally, H.264/AVC, whilst a published industry standard, is not free to use. Commercial implementations are subject to licence fees and the intellectual property position in itself is complicated.

1.5 About this book

The aim of this book is to de-mystify H.264 and its complexities. H.264/AVC will be a key component of the digital media industry for some time to come. A better understanding of the technology behind the standard and of the inter-relationships of its many component parts should make it possible to get the most out of this powerful tool.

This book is organized as follows.

Chapter 2 explains the concepts of digital video and covers source formats and visual quality measures.

Chapter 3 introduces video compression and the functions found in a typical video codec, such as H.264/AVC and other block-based video compression codecs.

Chapter 4 gives a high-level overview of H.264/AVC at a relatively non-technical level.

Chapters 5, 6 and 7 cover the standard itself in detail. **Chapter** 5 deals with the H.264/AVC syntax, i.e. the construction of an H.264 bitstream) including picture formats and picture management. **Chapter 6** describes the prediction methods supported by the standard, intra and inter prediction. **Chapter 7** explains the residual coding processes, i.e. transform and quantization and symbol coding.

Chapter 8 deals with issues closely related to the main standard – storage and network transport of H.264 data, conformance or how to ensure compatibility with H.264 and licensing, including the background and details of the intellectual property licence associated with H.264 implementations.

Chapter 9 examines the implementation and performance of H.264. It explains how to experiment with H.264, the effect of H.264 parameters on performance, implementation challenges and performance optimization.

Chapter 10 covers extensions to H.264/AVC, in particular the Scalable and Multiview Video Coding extensions that have been published since the completion of the H.264 standard. It examines possible future developments, including Reconfigurable Video Coding, a more flexible way of specifying and implementing video codecs, and possible successors to H.264, currently being examined by the standards groups.

Readers of my earlier book, "H.264 and MPEG-4 Video Compression", may be interested to know that Chapters 4–10 are largely or completely new material.

1.6 Reference

i. Ian Hickson, 'Codecs for <audio> and <video>', HTML5 specification discussion, http://lists.whatwg.org/htdig.cgi/whatwg-whatwg.org/2009-June/020620.html, accessed August 2009.