CHAPTER

Introduction

1

CHAPTER OUTLINE

Visual communications have a rich history, from the earliest cave painting, through Egyptian hieroglyphs, the invention of paper, moveable type, photography and film, to modern art, television, and "new media." We now capture, store, access, transmit, and receive visual information as part of our everyday lives and we can visualize almost anything, from distant stars to the smallest organism.

Visual information is the primary consumer of communications bandwidth across all broadcast, internet, and surveillance networks. The demand for increased qualities and quantities of visual content is increasing daily and this is creating a major tension between available network capacity and required video bit rate. Network operators,

Communicating Pictures. http://dx.doi.org/10.1016/B978-0-12-405906-1.00001-5
© 2014 Elsevier Ltd. All rights reserved.

content creators, and service providers all need to transmit the highest quality video at the lowest bit rate and this can only be achieved through the exploitation of perceptual redundancy to enable video compression.

This chapter provides a descriptive introduction to image and video compression. We first look at what compression means and why we need it. We then examine the primary drivers for video compression in today's world. Finally we consider the requirements of a video compression system and explain why standards are so important in supporting interoperability.

1.1 Communicating pictures: the need for compression

1.1.1 What is compression?

Before we look at why we need compression, let us consider the fundamental issue of what compression actually does. First let's define compression: there are a number of ways that we can express this, and these are covered in more detail later, but for now let us simply define it as the ratio of the size of (or number of bits in) the original image or video file B_i to that of the file used for storage or transmission B_o

$$\mathrm{CR} = \frac{B_o}{B_i} \quad \text{(unitless)} \tag{1.1}$$

This simple definition ignores a lot about real-time aspects or dynamic channel conditions, but will suffice for now.

Jumping ahead a little to Section 1.1.2, we will see that typical video compression ratio requirements are currently between 100:1 and 200:1, with this increasing to many hundreds or even thousands to one as new, more demanding formats emerge. If we consider Figure 1.1, we can see what that means in terms of a geometric analogy—in order to achieve a compression ratio of 256:1 we must replace the large square on the left by the small one at the bottom right of the figure. There is a common expression "You can't squeeze a quart into a pint pot,"[1] but we can see that the compression process must do exactly that. In fact, for the example of the bottom square in the figure, it must squeeze 128 quarts into a pint pot!

So what does this mean for pictures? Consider a small image, of dimension 10×10 samples (known as picture elements, *pels* or *pixels*), where each pixel is stored as a single byte to produce a 10×10 matrix of byte values. After compression (for a compression ratio of 100:1) we expect that these 100 values will be replaced by just one single byte AND that this is achieved without introducing any noticeable distortion in the image! At first this appears ridiculous, but that is what we need to achieve! Actually it's only what we need to achieve on average—as the only way we could reduce 100 numbers to a single value was if they were all the same (or fitted exactly some other known model)—e.g. the image was of constant luminance.

[1] A quart is a unit of volume (for either the imperial or United States customary units) equal to 2 pints or one quarter of a gallon.

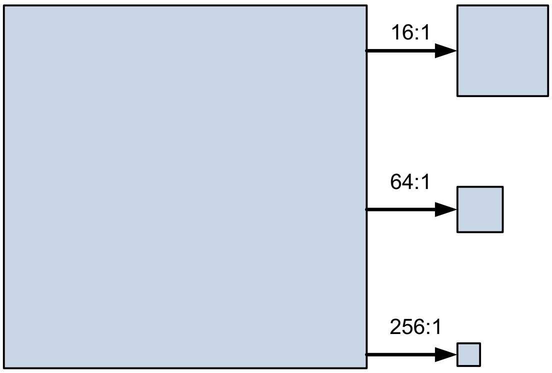

FIGURE 1.1

A geometric interpretation of compression.

In reality most images are not constant valued, but they are highly correlated spatially. Similarly, most video sequences are correlated temporally. As we will see, it is the modeling and exploitation of this spatio-temporal correlation, primarily through transformation and prediction, that enables us to compress images and video so effectively. Consider your Blu-ray disc—most people think of these as extremely high quality. However, they are encoded at around 20 Mbps—a compression ratio of around 60:1 compared to the original movie.[2]

So now let's look in a little more detail at why we need to compress pictures.

1.1.2 Why do we need compression?

Picture formats and bit rate requirements

We will consider picture formats in detail in Chapter 4, but introduce them briefly here in order to obtain an understanding of the requirements of a picture compression system. Pictures are normally acquired as an array of color samples, usually based on combinations of the red, green, and blue primaries. They are then often converted to some other more convenient color space that encodes luminance separately to two color difference signals (see Chapter 4).

Table 1.1 shows typical sampling parameters for a range of common video formats. Without any compression, it can be seen, even for the lower resolution formats, that the bit rate requirements are high—much higher than what is normally provided by today's communication channels. Note that the chrominance signals are encoded at a reduced resolution as indicated by the 4:2:2 and 4:2:0 labels.[3] Also note that two formats are included for the HDTV case (the same could be done for the other formats); for broadcast quality systems, the 4:2:2 format is actually more representative of the

[2]Assumes a 1080p24 4:2:2 format original at 12 bits per pixel. More about formats and color subsampling in Chapter 4.

[3]See Chapter 4 for more details on color spaces and sub-sampling.

Table 1.1 Typical parameters for common digital video formats and their (uncompressed) bit rate requirements.

Format	Spatial sampling ($V \times H$)	Temporal sampling (fps)	Raw bit rate (30 fps, 8/10 bits) (Mbps)
UHDTV (4:2:0) (ITU-R 2020)	Lum: 7680×4320 Chrom: 3840×2160	24, 25, 30, 50, 60,120	14,930*
HDTV (4:2:0) (ITU-R 709)	Lum: 1920×1080 Chrom: 960×540	24, 25, 30, 50, 60	933.1*
HDTV (4:2:2) (ITU-R 709)	Lum: 1920×1080 Chrom: 960×1080	24, 25, 30, 50, 60	1244.2*
SDTV (ITU-R 601)	Lum: 720×576 Chrom: 360×288	25, 30	149.3
CIF	Lum: 352×288 Chrom: 176×144	10–30	36.5
QCIF	Lum: 176×144 Chrom: 88×72	5–30	9.1

**Assumes encoding at 10 bits.*
UHDTV = Ultra High Definition Television; HDTV = High Definition Television; SDTV = Standard Definition Television; CIF = Common Intermediate Format; QCIF = Quarter CIF.

original bit rate as this is what is produced by most high quality cameras. The 4:2:0 format, on the other hand, is that normally employed for transmission after compression.

Finally, it is worth highlighting that the situation is actually worse than that shown in Table 1.1 especially for the new Ultra High Definition Television (UHDTV) standard where higher frame rates and longer wordlengths will normally be used. For example at 120 frames per second (fps) with a 10 bit wordlength for each sample, the raw bit rate increases to 60 Gbps for a single video stream! This will increase even further if 3-D or multiview formats are employed.

Available bandwidth

To see the other side of the inequality, let us now examine the bandwidths available in typical communication channels. Some common communication systems for broadcast and mobile applications are characterized in Table 1.2. This table should however be read with significant caution as it provides the theoretical maximum bit rates under optimum operating conditions. These are rarely, if ever, achieved in practice. Bandwidth limitations are particularly stringent in wireless environments because the usable radio spectrum is limited and the transmission conditions are variable and data loss is commonplace.

The bit rates available to an individual user at the application layer (which is after all what we are interested in) will normally be grossly reduced from the figures

Table 1.2 Theoretical bandwidth characteristics for common communication systems.

Communication system	Maximum bandwidth
3G mobile (UMTS)	384 kbps
4G mobile (4 × 4 LTE)	326 Mbps
Broadband (ADSL2)	24 Mbps
Broadband (VSDL2)	100 Mbps
WiFi (IEEE 802.11n)	600 Mbps
Terrestrial TV (DVB-T2 (8 MHz))	50 Mbps

quoted in Table 1.2. The effective throughput (sometimes referred to as *goodput*) is influenced by a large range of internal and external factors. These include:

- Overheads due to link layer and application layer protocols.
- Network contention and flow control.
- Network congestion and numbers of users.
- Asymmetry between download and upload rates.
- Network channel conditions.
- Hardware and software implementations that do not support all functions needed to achieve optimum throughput.

In particular, as channel conditions deteriorate, modulation and coding schemes will need to be increasingly robust. This will create lower spectral efficiency with increased coding overhead needed in order to maintain a given quality. The number of retransmissions will also inevitably increase as the channel worsens. As an example, DVB-T2 will reduce from 50 Mbps (256QAM @ 5/6 code-rate) to around 7.5 Mbps when channel conditions dictate a change in modulation and coding mode down to 1/2 rate QPSK. Similarly for 802.11n, realistic bandwidths per user can easily reduce well below 10 Mbps. Typical broadband download speeds where the author lives (in a semi-rural area of the UK) are, at the time of writing, 3 Mbps on a good day. 3G download speeds never offer 384 kbps—more frequently they will be less than 100 kbps. In summary, useful bandwidths per user at the application layer will rarely exceed 50% of the theoretical maximum and are more likely to be in the range 10–20%.

On that basis, let us consider a simple example which relates the raw bit rates in Table 1.1 to the realistic bandwidth available. Consider the example of a digital HDTV transmission at 30 fps using DVB-T2, where the average bit rate allowed in the multiplex (per channel) is 15 Mbps. The raw bit rate, assuming a 4:2:2 original at 10 bits, is approximately 1.244 Gbps, while the actual bandwidth available dictates a bit rate of 15 Mbps. This represents a compression ratio of approximately 83:1.

Download sites such as YouTube typically support up to 6 Mbps for HD 1080p[4] format, but more often video downloads will use 360p or 480p (640 × 480 pixels) formats at 30 fps, with a bit rate between 0.5 and 1 Mbps encoded using the H.264/AVC

[4] p stands for progressive transmission—see Chapter 4 for further details.

standard. In this case the raw bit rate, assuming color sub-sampling in 4:2:0 format, will be 110.6 Mbps. As we can see, this is between 100 and 200 times the bit rate supported for transmission.

Example 1.1 (Compression ratio for UHDTV)
Consider the case of 8K UHDTV with the original video in 4:2:2 format (a luminance signal of 7680 × 4320 and two chrominance signals of 3840 × 4320) at 10 bits per sample and a frame rate of 60 fps. Calculate the compression ratio if this video is to be transmitted over a DVB-T2 link with an average bandwidth of 15 Mbps.

Solution. The 4:2:2 color sub-sampling method has, on average, the equivalent of two 10 bit samples for each pixel (see Chapter 4). Thus, in its uncompressed form the bit rate is calculated as follows:

$$\begin{aligned} R &= 7680\ (H) \times 4320\ (V) \times 2\ (\text{samples/pixel}) \times 10\ (\text{bits}) \times 60\ (\text{fps}) \\ &= 39{,}813{,}120{,}000\ \text{b/s} \end{aligned}$$

i.e. a raw bit rate approaching 40 Gbps. Assuming this needs to be transmitted in the same bandwidth as for conventional HDTV (i.e. 15 Mbps) then a compression ratio of 2700:1 would be required!

$$\text{CR} = \frac{39{,}813{,}120{,}000}{15{,}000{,}000} \approx 2700$$

Hopefully this section has been convincing in terms of the need for compression. A tension between user expectations, in terms of quality and ease of access on one hand, and available bandwidth on the other, has existed since the first video transmissions and this has promoted vigorous research in the fields of both coding and networks. Fortunately the advances in communications technology have mirrored those in video compression, enabling the transmission of good (and in most cases very good) quality video that meets user expectations.

In the next section we examine the applications that are currently driving video compression performance as well as those that are likely do so in the future.

1.2 Applications and drivers

1.2.1 Generic drivers

By 2020 it is predicted that the number of network-connected devices will reach 1000 times the world's population; there will be 7 trillion connected devices for 7 billion people [1]. Cisco predict [2] that this will result in 1.3 ZB of global internet traffic by 2016, with over 80% of this being video traffic. The generic drivers for this explosion in video technology are as follows:

- Increased numbers of users with increased expectations of quality and accessibility.
- The ubiquity of user-generated image, video, and multimedia available through social networking and download sites.
- The emergence of new ways of working using distributed applications and environments such as the cloud.
- The demand for rapid and remote access to information, driven by the need for improved productivity, security, and responsiveness.
- Emerging immersive and interactive entertainment formats for film, television, and streaming.

In order to satisfy the emerging demands for accessibility and quality, video communication solutions require a high performance core network, an adaptive wireless infrastructure, robust data transport protocols, and efficient content representations. The interactions between these components in the delivery chain can be thought of as a jigsaw puzzle (Figure 1.2) where complex and dynamically changing pieces need to fit together. Only in combination can these provide a truly end-to-end flexible and robust solution to the delivery challenges of the future. Video compression is a key component in the solution.

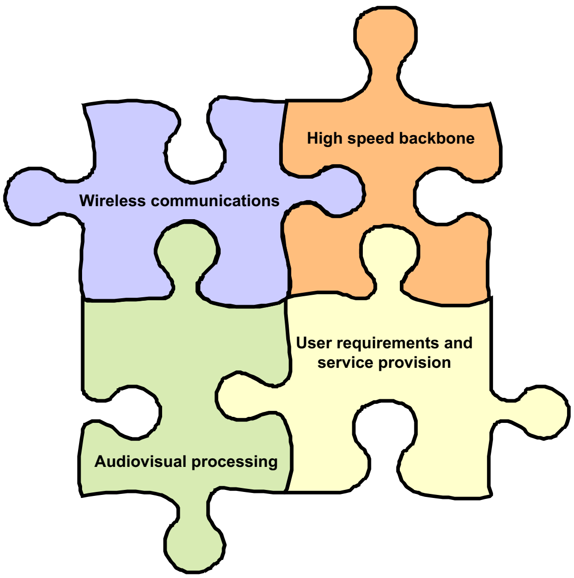

FIGURE 1.2

The multimedia communications jigsaw puzzle.

1.2.2 Application drivers and markets

The range of applications which rely wholly or partially on video technology is enormous; as a consequence, the market for video compression technology is growing rapidly. Some examples are provided below.

Consumer video

Entertainment, personal communications, and social interaction provide the primary applications in consumer video, and these will dominate the video landscape of the future. As mentioned previously, Cisco predict [2] that by 2016, 80% of all internet traffic (1.3 ZB) will be video. According to a recent Frost and Sullivan report, more than 1 billion consumer video devices were sold in 2013 and of these, 59% were smart phones [3]. There has been a massive increase in the consumption and sharing of content on mobile devices and this is likely to be the major driver over the coming years. The Quarterly Global TV Shipment and Forecast Report by NPD DisplaySearch [4] predicts global TV sales to reach 270 million in 2016.

The cinema projector market has been predicted [5] to be $0.8 billion in 2018, with market growth being driven by the shift toward digital cinema technology and the availability of 3-D content. The key drivers in this sector are:

- Broadcast television, digital cinema, and the demand for more immersive content (3-D, multiview, higher resolution, frame rate, and dynamic range).
- Internet streaming, peer to peer distribution, and personal mobile communication systems.
- Social networking, user-generated content, and content-based search and retrieval.
- In-home wireless content distribution systems and gaming.

Business, manufacturing, and automation

Computer vision and visual communications are playing an increasingly important role in business. Particularly, since the perceived security threats associated with air travel have emerged, the demand for higher quality video conferencing and the sharing of visual content have increased. Similarly in the field of automation, vision-based systems are playing a key role in transportation systems and are now underpinning many manufacturing processes, often demanding the storage or distribution of compressed video content. The drivers in this case can be summarized as:

- Video conferencing, tele-working, and other interactive services.
- Publicity, advertising, news, and journalism.
- Design, modeling, and simulation.
- Transport systems, including vehicle guidance, assistance, and protection.
- Automated manufacturing and robotic systems.

Security and surveillance

In an uncertain world, we have become increasingly aware of our safety and security, and video monitoring is playing an increasingly important role in this respect. It is estimated that the market for networked cameras (non-consumer) [6] will grow to around $4.5 billion in 2017 (19% annual growth from 2012 to 2017), with 28 million units shipped in 2017. Aligned with this, the same source predicts an even larger growth in video analytics to $1.2 billion in 2017. The key drivers in this sector are:

- Surveillance of public spaces and high profile events.
- National security.
- Battlefield situational awareness, threat detection, classification, and tracking.
- Emergency services, including police, ambulance, and fire.

Healthcare

The healthcare market is becoming increasingly reliant on imaging methods to aid diagnoses and to monitor the progression of disease states. 3-D imaging modalities such as CT and MRI produce enormous amounts of data for each scan and these need to be stored as efficiently as possible while retaining the highest quality. Video is also becoming increasingly important as a point-of-care technology for monitoring patients in their own homes to reduce costs and limit hospital admissions. The primary healthcare drivers for compression are:

- Point-of-care monitoring.
- Emergency services and remote diagnoses.
- Tele-surgery.
- Medical imaging.

It is clear that all of the above application areas require considerable trade-offs to be made between cost, complexity, robustness, and performance. These issues are addressed further in the following section.

1.3 Requirements and trade-offs in a compression system

1.3.1 The benefits of a digital solution

We live in a world where digital technology dominates. In the context of video compression, digital solutions offer many significant advantages over their analog counterparts. These include:

- Ease of editing, enhancement, special effects, transcoding, and multigeneration processing.
- Ease of indexing, metadata extraction, annotation search, and retrieval.
- Ease of content protection, conditional access, rights management, and encryption.
- Support for scalability and interactivity.
- Source and channel coding can be separated to provide better management of rate, distortion, and robustness to errors.
- Flexible support for new formats such as 3DTV and multiview.

1.3.2 Requirements

The basic requirement of a video compression system can be stated quite simply: we want the highest quality at the lowest bit rate. However, in addition to this, a number of other desirable features can be listed. These include:

- **Robustness to loss:** We want to maintain high quality when signals are transmitted over error-prone channels.
- **Reconfigurability and flexibility:** To support delivery over time-varying channels or heterogeneous networks.
- **Low complexity:** Particularly for low power portable implementations.
- **Low delay:** To support interactivity.
- **Authentication and rights management:** To support conditional access, content ownership verification or to detect tampering.
- **Standardization:** To support interoperability.

1.3.3 Trade-offs

In practice we cannot satisfy all of the above requirements, either due to cost or complexity constraints or because of limited bandwidth or lossy channels. We therefore must compromise and make design trade-offs that provide the best solution given the prevailing constraints. Areas of possible compromise include:

- **Lossy vs lossless compression:** Ideally, we do not want the video to be distorted at all compared to the original version. In practice this is not usually possible as perfectly lossless (reversible) compression places severe limitations on the compression ratio achievable. We must therefore exploit any redundancy in the image or video signal in such a way that it delivers the desired compression with the minimum perceived distortion.
- **Rate vs quality:** In order to compromise between bit rate and quality, we can adjust a number of parameters. These include: frame rate, spatial resolution (luma and chroma), dynamic range, prediction mode, and latency. The influence of these parameters on the final video quality is however highly content dependent. For example, with a fixed camera and a completely static scene, frame rate will have little or no influence on the perceived quality. In contrast, for a high motion scene, a low frame rate will result in unpleasant jerkiness.
- **Complexity and power consumption vs cost and performance:** In general, a more complex video encoder will support more advanced features and thus produce a higher quality result than a lower complexity version. However, more complex architectures invariably are more expensive and will often introduce more delay. In addition, they may not be realizable using a software implementation, possibly demanding custom hardware solutions.
- **Delay vs performance:** Low latency is key in interactive and conversational applications. However, if an application is tolerant to latency in encoding or decoding, then performance gains can be made elsewhere. For example, increased error

resilience features could be incorporated to retransmit data corrupted during transmission or, alternatively, more reference frames could be used to improve motion prediction performance (see Chapters 8 and 9).

1.4 The basics of compression

A simplified block diagram of a video compression system is shown in Figure 1.3. This shows an input digital signal (using for example one of the formats from Table 1.1) being encoded, transmitted, and decoded. Not shown are possible pre-processing stages such as filtering, interlacing, segmenting, color sub-sampling, gamma correction; these are covered in Chapter 4. A full description of the operation of this so-called block-based hybrid motion compensated video encoder is provided in Chapter 9.

1.4.1 Still image encoding

If we ignore the blocks labeled as motion compensation, the diagram in Figure 1.3 describes a still image encoding system, such as that used in JPEG. The intra-frame encoder performs (as its name suggests) coding of the picture without reference to any other frames. This is normally achieved by exploiting spatial redundancy through transform-based decorrelation (Chapters 5, 6, and 9) followed by variable length symbol encoding (VLC) (Chapter 7). The image is then conditioned for transmission using some means of error resilient coding that makes the encoded bitstream more robust to channel errors. Methods for achieving error resilience are described in Chapter 11.

At the decoder, the inverse operations are performed and the original image is reconstructed at the output.

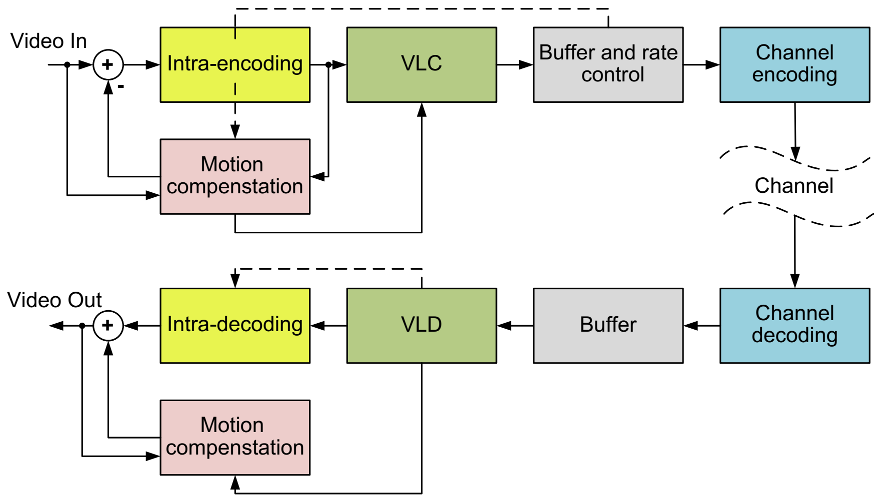

FIGURE 1.3

Simplified high level video compression architecture.

1.4.2 Encoding video

A video signal is simply a sequence of still images, acquired typically at a rate of 24, 25, 30, 50, or 60 fps. Video can of course be encoded as a series of still images using intra-frame methods as described above. However, we can achieve significant gains in coding efficiency if we also exploit the temporal redundancy that exists in most natural video sequences. This is achieved using inter-frame motion prediction as represented by the motion compensation block in Figure 1.3. This block predicts the structure of the incoming video frame based on the contents of previously encoded frames (Chapter 8). The encoding continues as for the intra-frame case, except this time the intra-frame encoder block processes the low energy residual signal remaining after prediction, rather than the original frame.

After variable length encoding, the encoded signal will be buffered prior to transmission. The buffer serves to match the output bit rate to the instantaneous capacity of the channel and this is normally achieved using a rate control algorithm which adjusts coding parameters as described in Chapter 10.

Because of the reliance on both spatial and temporal prediction, compressed video bitstreams are more prone to channel errors than still images, suffering from temporal as well as spatial error propagation. Methods of mitigating this, making the bitstream more robust and correcting or concealing the resulting artifacts are described in Chapter 11.

1.4.3 Measuring visual quality

In the final analysis, we need to assess how good our encoding system is, and the absolute arbiter of this is the human visual system. Because of this, subjective testing methodologies have become an important component in the design and optimization of new compression systems. However, such tests are very time consuming and cannot be used for real-time rate-distortion optimization. Hence there has been a significant body of work reported on the development of objective metrics that provide a realistic estimate of video quality. These are discussed alongside subjective evaluation methods in Chapter 10.

1.5 The need for standards

Standardization of image and video formats and compression methods has been instrumental in the success and universal adoption of video technology. We briefly introduce the motivation for and history of standards here; a more detailed description of the primary features of recent standards is provided in Chapter 12.

1.5.1 Some basic facts about standards

A few basic facts about image and video compression standards are worth mentioning first:

1. Standards are essential for interoperability, enabling material from different sources to be processed and transmitted over a wide range of networks or stored

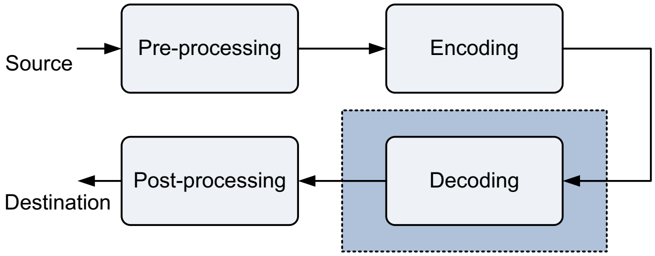

FIGURE 1.4

The scope of standardization.

on a wide range of devices. This interoperability opens up an enormous market for video equipment, which can exploit the advantages of volume manufacturing, while also providing the widest possible range of services for users.

2. Standards define the bitstream format and decoding process, not (for the most part) the encoding process. This is illustrated in Figure 1.4. A standard-compliant encoder is thus one that produces a compliant bitstream and a standard-compliant decoder is one that can decode a standard-compliant bitstream, i.e. a standard specifies the syntax that a bitstream must follow and the corresponding decoding process, not explicitly the encoding process.
3. The real challenge lies in the bitstream generation, i.e. the encoding. This is where manufacturers can differentiate their products in terms of coding efficiency, complexity, or other attributes.
4. Many encoders are standard-compliant, but some are better than others! Compliance is no guarantee of absolute quality.

1.5.2 A brief history of video encoding standards

A chronology of video coding standards is represented in Figure 1.5. This shows how the International Standards Organization (ISO) and the International Telecommunications Union (ITU-T) have worked both independently and in collaboration on various standards. In recent years, most ventures have benefited from close collaborative working.

Study Group SG.XV of the CCITT (now ITU-T) produced the first international video coding standard, H.120, in 1984. H.120 addressed videoconferencing applications at 2.048 Mbps and 1.544 Mbps for 625/50 and 525/60 TV systems respectively. This standard was never a commercial success. H.261 [7] followed this in 1989 with a codec based on a $p \times 64$ kbps ($p = 1 \cdots 30$) targeted at ISDN conferencing applications. This was the first block-based hybrid compression algorithm using a combination of transformation (the Discrete Cosine Transform (DCT)), temporal Differential Pulse Code Modulation (DPCM) and motion compensation. This architecture has stood the test of time as all major video coding standards since have been based on it.

In 1988 the Moving Picture Experts Group (MPEG) was founded, delivering a video coding algorithm targeted at digital storage media at 1.5 Mbs/s in 1992.

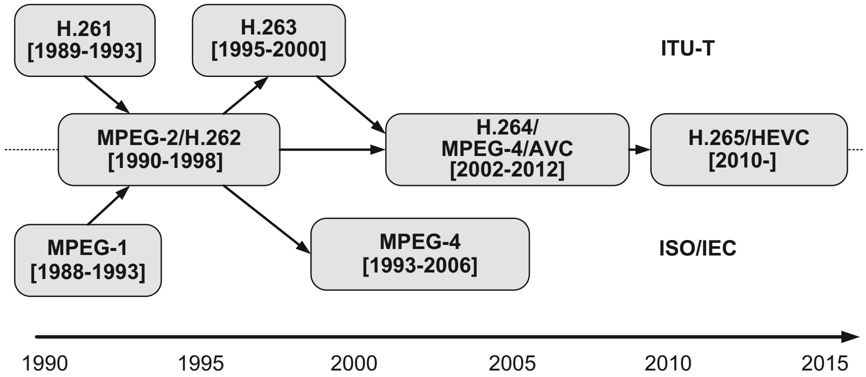

FIGURE 1.5

A chronology of video coding standards from 1990 to the present date.

This was followed in 1994 by MPEG-2 [8], specifically targeted at the emerging digital video broadcasting market. MPEG-2 was instrumental, through its inclusion in all set-top boxes for more than a decade, in truly underpinning the digital broadcasting revolution. A little later in the 1990s ITU-T produced the H.263 standard [9]. This addressed the emerging mobile telephony, internet, and conferencing markets at the time. Although mobile applications were slower than expected to take off, H.263 had a significant impact in conferencing, surveillance, and applications based on the then-new Internet Protocol.

MPEG-4 [10] was a hugely ambitious project that sought to introduce new approaches based on object-based as well as, or instead of, waveform-based methods. It was found to be too complex and only its Advanced Simple Profile (ASP) was used in practice. This formed the basis for the emerging digital camera technology of the time.

Around the same time ITU-T started its work on H.264 and this delivered its standard, in partnership with ISO/IEC, in 2004 [11]. In the same way that MPEG-2 transformed the digital broadcasting landscape, so has H.264/AVC transformed the mobile communications and internet video domains. H.264/AVC is by far the most ubiquitous video coding standard to date. Most recently in 2013, the joint activities of ISO and ITU-T delivered the HEVC standard [12–14], offering bit rate reductions of up to 50% compared with H.264/AVC.

Further details on coding standards are provided in Chapter 12.

1.6 Summary

This chapter has introduced the context and the primary drivers for image and video compression technology in today's world. We have seen that compression exists because of the tensions between, on one hand, the increasing user demands and expectations for increased quality, mobility and immersion, and on the other hand, the limited bandwidth and high traffic volumes that characterize most communications networks. Communicating pictures is a key part of modern life and compression is an essential element in ensuring that they are delivered faithfully, reliably, and in a timely manner.

We have examined the requirements of a compression system and have seen that these represent a compromise between coding rate, picture quality, and implementation complexity. Finally the justification for universal compression standards has been presented, as these are essential in order to enable interoperability across networks, terminals, and storage devices. The standardization process is also a primary driver for continued research in the field and creates a productive interaction between industry and academia.

References

[1] <http://www.wireless-world-research.org/fileadmin/sites/default/files/publications/-Outlook/Outlook4.pdf>.

[2] Cisco Visual Networking Index: Forecast and Methodology, 2011–16 (update 2012–17). <http://www.cisco.com/en/US/netsol/ns827/networking_solutions_sub_solution.html>.

[3] Analysis of the Global Consumer Video Devices Market: Apps, QoE and Transformative User-Experiences Key to Growth in Price-Challenged Industry, Frost & Sullivan Report NC68-01, 2013.

[4] Global TV Shipment and Forecast Report, NPD DisplaySearch, 2013.

[5] Digital Cinema Projectors: A Global Strategic Business Report, Global Industry Analysts Inc., 2012.

[6] Network Camera and Video Analytics Market—Global Forecast, Trend & Analysis—Segmentation by Technology, Function, Resolution, Product & Service Type, System Architecture, Verticals, Application and Geography (2012–2017) Report by marketsandmarkets.com, Report Code: SE 1238, 2012.

[7] International Telecommunication Union-Telecommunication (ITU-T), Recommendation H.261, Video codec for audiovisual services at $p \times 64$ kbit/s, version 1, 1990, version 2, 1993.

[8] ITU-T and ISO/IEC JTC 1, Generic Coding of Moving Pictures and Associated Audio Information—Part 2: Video, ITU-T Rec. H.262 and ISO/IEC 13818-2 (MPEG-2 Video), version 1, 1994.

[9] ITU-T, Video Coding for Low Bitrate Communication, ITU-T Rec. H.263, version 1, 1995, version 2, 1998, version 3, 2000.

[10] ISO/IEC JTC 1, Coding of Audio-Visual Objects—Part 2: Visual, ISO/IEC 14496-2 (MPEG-4 Visual), version 1, 1999, version 2, 2000, version 3, 2004.

[11] ITU-T and ISO/IEC JTC 1, Advanced Video Coding for Generic Audiovisual Services, ITU-T Rec. H.264 and ISO/IEC 14496-10 (AVC), version 1, 2003, version 2, 2004, versions 3, 4, 2005, versions 5, 6, 2006, versions 7, 8, 2007, versions 9, 10, 11, 2009, versions 12, 13, 2010, versions 14, 15, 2011, version 16, 2012.

[12] J.-R. Ohm, G. Sullivan, H. Schwartz, T. Tan, T. Wiegand, Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC), IEEE Transactions on Circuits and Systems for Video Technology 22 (12) (2012) 1669–1684.

[13] G. Sullivan, J.-R. Ohm, W. Han, T. Wiegand, Overview of the high efficiency video coding (HEVC) standard, IEEE Transactions on Circuits and Systems for Video Technology 22 (12) (2012) 1648–1667.
[14] Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 ISO/IEC 23008-2 and ITU-T Recommendation H.265, High Efficiency Video Coding (HEVC), January 2013.