AVC / H.264

— an advanced video coding system for SD and HD broadcasting

Paola Sunna
RAI CRIT

A bitrate of about 270 Mbit/s is needed to transmit uncompressed digital video that accords with ITU-R Rec. BT. 601 (i.e. standard-definition television). Digital HDTV, on the other hand, needs a considerably greater bitrate and – regardless of the modulation scheme adopted – transmission via traditional broadcast channels is impossible without the application of advanced video compression techniques.

This article gives an overview of the current video coding technologies that are suitable for HDTV transmission; in particular, AVC/H.264.

Historical overview

In January 1988, the MPEG (Motion Picture Expert Group) gathered together some experts from the ISO/IEC in order to define a standard system for the encoding of motion pictures. The first standard introduced was **MPEG-1** for applications in the multimedia field and its successor was **MPEG-2** – for broadcasting applications.

The MPEG-2 algorithm performs video compression in the following ways:

- **Lossless compression**
  This is based on exploitation of spatial/temporal redundancy in the pictures (i.e. correlation between adjacent pixels in each frame and between successive fields or frames in time) and on *variable length encoding.*

- **Psycho-visual encoding**
  Here, picture information that is invisible to the human eye is discarded.

- **Lossy compression**
  In this case, certain picture information that is visible to the (trained) human eye is discarded in order to increase the compression factor. This information cannot be reconstructed at the receiver side, resulting in some impairments to the picture quality.

MPEG-2 uses the *Discrete Cosine Transform* (DCT) to reduce spatial redundancy; it actually decreases the spatial correlation between pixels and helps to randomize the statistical distribution of the samples, thus improving the efficiency of the variable length encoding. DCT coefficients strictly relate to the image spectral content and, as the human eye's sensibility is not uniform for all frequencies and angles from the horizontal, the psycho-visual encoding in MPEG-2 is done by means of an appropriate quantization matrix (the coefficients relating to the high frequencies and diagonal directions are quantized more coarsely than the others).

Concerning temporal redundancy reduction, MPEG-2 improves the predictive approach by using a technique called *motion compensation*.

The MPEG-2 standard is structured as *profiles* and *levels* and, for each of them, the maximum bitrate that the decoder must be capable of processing is clearly defined.

After standardization, MPEG-2 came into widespread use (DVB, ATSC, DVD Forum, etc.). MPEG-2, like most other coding schemes, only defines the bitstream syntax and the decoder characteristics – leaving the manufacturers free to implement the algorithm on the encoder side. For these reasons, MPEG-2 coders show different performances, depending on the algorithm type used for the motion compensation, the quantization matrix values and the rate control mechanism – while parameters such as the block dimensions, the entropy encoding and the number of pictures used for the predictions cannot be changed, in order to maintain compatibility with MPEG-2 decoders.

MPEG-4 Part 2

The first successor of MPEG-2 was **MPEG-4 Part 2** published by the ISO in 1999. As in the MPEG-2 case, the encoding efficiency is strictly related to the source material complexity and the encoder implementation. MPEG-4 was defined for applications in the multimedia field at low bitrates but was further extended to broadcast applications. A formal subjective assessment has shown [1] that the encoding efficiency gain of MPEG-4 Part 2, with respect to MPEG-2, is between 15 and 20%. For DVB applications, this efficiency gain is not enough to justify destabilizing the MPEG-2 world – given that MPEG-4 Part 2 is not compatible with MPEG-2.

The AVC/H.264 advanced video compression scheme

In 2001, with the aim of developing a more efficient compression system, the standardization bodies ISO/IEC (MPEG) and the ITU brought their efforts together in the Joint Video Team (JVT), a working group charged with developing a coding system called **Advanced Video Coding** (AVC). In 2003, the AVC system was integrated as **part 10** in the **MPEG-4** ISO/IEC 14496-10 standard and assumed the name **H.264** in the ITU.

In September 2004, the DVB Consortium modified ETSI standard TS 101 154 [2] *(Implementation guidelines for the use of Video and Audio Coding in Broadcasting Applications based on the MPEG-2 Transport Stream)* to also include AVC/H.264.

Available profiles

The AVC/H.264 system does not produce a bitstream compliant with MPEG-2, so its adoption will require the use of new encoders and decoders. The AVC scheme includes different profiles:

❍ **Baseline Profile** – for low-delay end-to-end applications;

❍ **eXtended Profile** – for mobile applications and e-streaming;

❍ **Main Profile** – for broadcasting application at SD (Standard Definition) level.

❍ **High Profile**
   The initial H.264/AVC standard primarily focused on "entertainment-quality" video, based on 8 bits/sample and 4:2:0 chroma sampling. To address the needs of the most demanding applications – such as contribution and distribution of content, studio editing and post processing – a

---

1. Ken McCann: **DVB + MPEG-4 = ?**
   PowerPoint presentation at DVB World 2002.

2. **http://webapp.etsi.org/workprogram/Report_WorkItem.asp?WKI_ID=21401**

<div style="background-color:#f5d792;">

Abbreviations

ATSC	Advanced Television Systems Committee (USA)	**ITU-R**	ITU - Radiocommunication Sector
AVC	(MPEG-4) Advanced Video Coding	**ITU-T**	ITU - Telecommunication Standardization Sector
CAVLC	Context-Adaptive Variable Length Coding	**JVT**	(MPEG/VCEG) Joint Video Team
CABAC	Context-Adaptive Binary Arithmetic Coding	**MPEG**	Moving Picture Experts Group
CIF	Common Intermediate Format	**QCIF**	Quarter Common Intermediate Format
DCT	Discrete Cosine Transform	**SDTV**	Standard-Definition Television
DVB	Digital Video Broadcasting	**SMPTE**	Society of Motion Picture and Television Engineers (USA)
IEC	International Electrotechnical Commission	**SQCIF**	Sub-Quarter Common Intermediate Format
IP	Internet Protocol	**VCEG**	(ITU-T) Video Coding Experts Group
ISO	International Organization for Standardization	**VLC**	Variable-Length Coder/Coding
ITU	International Telecommunication Union	**VoD**	Video-on-Demand

</div>

continuation of the joint project was launched to add new extensions to the capabilities of the original standard. These extensions, originally known as the "professional" extensions, were renamed as the "fidelity range extensions" or *FRExt*. In the process of developing the FRExt amendment, the FRExt project produced a suite of four new profiles collectively called the *High* profiles:

- **The High profile (HP)** – supporting 8-bit video with 4:2:0 sampling, and addressing high-end consumer and other applications that require high-resolution video without a need for extended chroma formats or extended sample accuracy;

- **The High 10 profile (Hi10P)** – supporting 4:2:0 video with up to 10 bits per sample;

- **The High 4:2:2 profile (H422P)** – supporting up to 4:2:2 chroma sampling and up to 10 bits per sample;

- **The High 4:4:4 profile (H444P)** – supporting up to 4:4:4 chroma sampling, up to 12 bits per sample and, additionally, supporting efficient lossless region coding and an integer residual colour transform for coding RGB video while avoiding colour-space transformation errors.

All these profiles support the features of the prior Main profile and, additionally, support an adaptive transform blocksize and perceptual quantization scaling matrices.

The High profile appears certain to be incorporated into the specifications of several important near-term applications, including:

- ❍ The *HD-DVD* video specification of the DVD Forum
- ❍ The *BD-ROM* video specification of the Blu-ray Disc Association, and
- ❍ The *DVB* standards for European broadcast television.

It appears that the High profile may rapidly overtake the Main profile in terms of dominant near-term industry implementation. This is because the High profile adds more coding efficiency to what was previously defined in the Main profile, without adding a significant amount of implementation complexity.

Encoding efficiency: AVC/H.264 versus MPEG-2

According to the Report of the formal Verification Tests on AVC/H.264 [3], the encoding efficiency is clearly superior to MPEG-2. The key advantages of AVC/H.264 are reported below:

1) Motion compensation

---

3. ISO/IEC JTC1/SC29/WG11 MPEG2003/N6231, December 2003 Waikoloa Hawaii, USA.

AVC/H.264 uses blocks of dimension and form that are variable compared with the fixed 16x16 blocks used in MPEG-2. In this way it is possible to achieve an efficiency gain of up to 15%.

The motion vector estimation is more precise in AVC/H.264: down to a ¼ of a pixel in AVC/H.264 against ½ a pixel in MPEG-2. This means it is possible to achieve a gain of up 20%.

AVC/H.264 uses up to 5 frames for the motion estimation against the two frames used in MPEG-2 for the interpolative pictures with a bitrate gain of between 5 and 10%.

2) Spatial redundancy reduction

AVC/H.264 uses an *integer transform* (instead of the DCT used in MPEG-2), thus reducing the influence of rounding errors.

3) Quantization

AVC/H.264 adopts a higher number of quantization levels: 52 against the 31 used in MPEG-2.

4) Entropic encoding

AVC/H.264 uses more complex encoding techniques which are more efficient than the static VLC included in MPEG-2.

5) De-blocking filter

AVC/H.264 uses an adaptive filter with the aim of reducing the blockiness that can seriously degrade the final picture quality in MPEG-2.

The higher efficiency of AVC/H.264 – defined as the bitrate reduction achievable while maintaining the same subjective picture quality – is paid for in terms of increased complexity in both the encoder and the decoder. *Table 1* shows some figures for the increased complexity of the *decoder*.

Table 1
Efficiency/complexity of AVC/H.264 with respect to MPEG-2 *(Source: www.m4if.org)*

Profile	Efficiency with respect to MPEG-2	Increase in decoder complexity
Baseline	About 1.5 times	About 2.5 times
Extended	About 1.75 times	About 3.5 times
Main	About 2 times	About 4 times

An AVC/H.264 *encoder* is eight times more complex than an MPEG-2 encoder but, luckily, solid-state computing power has increased by a factor of 100 since the beginning of MPEG-2 compression, and real-time AVC encoder/decoder chipsets are now under development.

Main competitors to AVC/H.264

In the last few years, many promising proprietary technologies (Windows Media, Real Video, On2, Sorenson etc.) have been developed to distribute multimedia content over narrow-bandwidth transmission channels. These compression systems have been used primarily for low-bitrate video at smaller picture sizes (e.g. SQCIF, QCIF and CIF). However, their scope has more recently been expanded to provide coding efficiency across a wide range of bitrates – from a few kbit/s up to tens of Mbit/s that can cope with SDTV and HDTV.

Windows Media 9 / VC-1

In particular, Microsoft's **Windows Media 9** was a proprietary system until March 2004 when the Redmond company presented a document called "*Proposed SMPTE Standard for Television: VC-9 compressed video Bit-stream format and decoding Process*" to the SMPTE for the standardization of Windows Media 9 technology. The name of the standard was later changed to **VC-1** and Windows Media 9 is now just a software implementation of VC-1.

There are three profiles in the VC-1 recommendation:

❍ **Simple** – which targets low-rate internet streaming and low-complexity applications such as mobile communications, or playback of media in personal digital assistants.

❍ **Main** – which targets high-rate internet applications such as streaming, movie delivery via IP, or TV/VoD over IP.

❍ **Advanced** – which addresses broadcast applications such as digital TV, HD DVD for PC play-back, and HDTV. It is the only VC-1 profile that supports **interlaced** content. In addition, the Advanced profile contains the required syntax elements to transmit its own video bitstreams into generic systems such as the MPEG-2 Transport or Programme Streams (*ISO/IEC 13818-2).*

Licensing issue are still open for VC-1 because MPEG-LA is still trying to establish if there are any VC-1 patent holders and, if so, how many.

Dirac

Another video codec that aims to be competitive with the state-of-the-art codecs discussed here is **Dirac** even if, at the moment, it is in the early stages of development. Dirac uses wavelets, motion compensation and arithmetic coding and has been developed by BBC R&D as a research tool, not a product. An experimental version of the code, written in C++, was released under an Open Source licence agreement in March 2003 and can be found at **http://sourceforge.net/projects/dirac**.

Comparison between AVC/H.264, VC-1 and MPEG-2

Table 2 (on the next page) gives a comparison between the characteristics of MPEG-2, AVC/H.264 and VC-1. In the table, AVC/H.264 and VC-1 seem to be quite similar but, at the time this article was being written (February 2005), no formal data about VC-1 compression efficiency and complexity were available – even if Microsoft claims that the performance of VC-1 is comparable with AVC/H.264, while being less complex to implement.

Conclusions

The MPEG-2 standard, which has been used for Digital TV and in storage applications for more than 10 years, is now suffering competition from new more-efficient video compression systems. In

Paola Sunna was born in Italy in 1971. In 1997 she received a Degree in Electronic Engineering from the *Politecnico* of Turin. Her thesis was on objective video quality assessments in the MPEG-2 domain. Since then, Ms. Paola has been working at RAI CRIT (the Centre for Research and Technological Innovation of the Italian broad-caster RAI); her activities are mainly focused on full lab testing of video quality assessment schemes for broadcasting, webcasting and 3G applications.

Since 2002, Ms Sunna has been the chair of EBU project group B/VIM (Video in Multimedia) that has defined the new subjective methodology for multimedia quality evaluation – SAMVIQ.

Table 2
Main features of MPEG-2, H.264 and VC-1

	MPEG-2	AVC/H.264	VC-1
Profiles	Simple Main High 4:2:2	Baseline Main Extended FRExt	Simple Main Advanced
Input source	Interlaced Progressive	Interlaced Progressive	Interlaced Progressive
Chroma formats	4:2:0, 4:2:2	4:2:0, 4:2:2, 4:4:4	4:2:0
Algorithm type	Hybrid (spatial and temporal redundancy)	Hybrid (spatial and temporal redundancy)	Hybrid (spatial and temporal redundancy)
Encoding structure	Hierarchical Profiles@Levels	Hierarchical Profiles@Levels	Hierarchical Profiles@Levels
Picture type	I, P, B	I, P, B, SP, SI	I, P, B, BI
Macroblock structure	16x16 Y 8x8 C_b 8x8 C_r	16x16 Y 8x8 C_b 8x8 C_r	16x16 Y 8x8 C_b 8x8 C_r
Motion estimation precision	Up to ½ pixel	Up to ¼ pixel	Up to ¼ pixel
***Loop filter* for blocki-ness reduction**	Not implemented	Implemented	Implemented
Transform type	Float (DCT)	Integer	Integer
Entropy encoding	VLC	CAVLC CABAC	VLC

particular, **AVC/H.264** represents the latest development in video coding standards – it typically outperforms MPEG-2 by a factor of two.

A similar performance is achieved by Microsoft Windows Media 9, which is in the process of being adopted by the SMPTE as the **VC-1** video coding standard. However, other compression systems are emerging – also aiming to play a key role in the development of new applications and business opportunities, both within and outside the field of broadcasting (e.g. HDTV, video streaming, Broadband TV, video teleconferencing, medical, scientific, etc.).

Finally, it is not only the coding efficiency that may influence the choice of codec to replace MPEG-2: there are still major Intellectual Property licensing issues to be settled before any sound business plans can be finalised, based on this new generation of more-efficient codecs.