Information and Coding

Armando J. Pinho

Departamento de Electrónica, Telecomunicações e Informática

Universidade de Aveiro

ap@ua.pt

Contents

H.261

- H.261 (1990) is a ITU-T video coding standard (Video Codec for Audiovisual Services at p×64 kbit/s) that was developed with the aim of being used
  - In video-phone applications.
  - In video-conference applications.
  - Over ISDN links at $p \times 64$ kbps, $p = 1, \ldots, 30$.
- For example, $p = 1$ (64 kbps) would be appropriated for video-phone, where the video signal was transmitted at 48 kbps and the audio signal at 16 kbps.
- Generally, video-conference required better image quality, implying typically $p \geq 6$ (384 kbps).
- For $p = 30$ we have 1.92 Mbps, which was sufficient for a video quality similar to the old VHS tapes.

H.261

- The encoded stream has the following structure:
  - At the top, the **frame**.
  - Each frame is partitioned into several **groups of blocks**.
  - Each group of blocks is formed of several **macroblocks**.
  - The macroblock is the smallest region that can have a particular coding mode assigned to.
  - The macroblock is composed of four basic **blocks** (a basic block is 8×8) of luminance (Y) and by the corresponding 8×8 chrominance blocks (C_r and C_b).

H.261

- The H.261 uses two compression modes:
  - **Intraframe**: similar to the JPEG compression, i.e., relies on DCT applied to 8×8 blocks of pixels.
  - **Interframe**: temporal prediction (motion compensation), followed by DCT of the prediction residuals.
- Motion compensation (MC) is performed in macroblocks, within a search area of 15 pixels around the macroblock.
- It has 32 quantizers, one of them dedicated to the DC coefficient in intraframe mode (quantization step of 8). The others have quantization steps from 2 to 62.
- Statistical coding is performed with Huffman codes.

MPEG–1

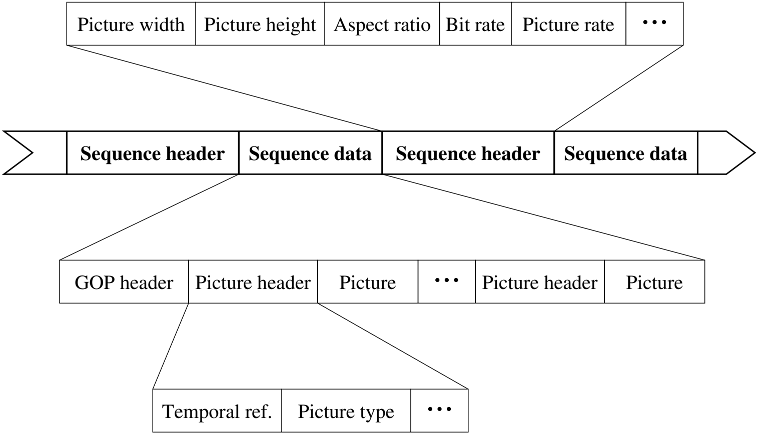
- MPEG–1 (1992) is a ISO/IEC (11172) coding standard that has been developed with the aim of storing video and audio in CD-ROMs.
- Target bitrates were around 1.5 Mbps, which was the bitrate associated to the early CD-ROM readers.
- The main objective of MPEG–1 was to provide means for encoding audio and video for interactive multimedia applications.
- For video segments having a moderate motion content, quality similar to VHS could be attained for MPEG–1 video at 1.2 Mbps.

MPEG–1

- The algorithms used in MPEG–1 are similar to those of H.261, although having some additional characteristics, such as
  - Random access (using type I frames)
  - Fast forward and reverse.
  - Backwards playing.
- Generally, the input is in the CCIR 601 format (576 $\times$ 720, for a 50 fps or 480 $\times$ 720 for 60 fps), and is converted to SIF (Source Input Format) before encoding (luma with 288(240) $\times$ 352 pixels and chroma with 144(120) $\times$ 176 pixels).

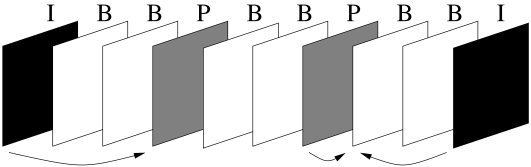
MPEG–1

Organization of the bitstream



MPEG–1

- MPEG–1 allows three types of **frames**:
  - **Type I**: encoding is similar to that of JPEG. These frames serve as entry points for random access.
  - **Type P**: frames encoded in predictive mode, using as reference previous frames of type I or P.
  - **Type B**: frames encoded in predictive mode, using both reference frames from the past and from the future (of type I or P).
- The number of I, P and B frames composing a **group of frames** depends on the application.



MPEG–2

- MPEG–2 (1994) has been developed aiming applications such as
  - Transmission of television signals in standard definition formats (PAL, SECAM, NTSC).
  - High definition television (HDTV).
  - Electronic cinema.
  - Games and high quality multimedia applications.
  - . . .
- Some characteristics of MPEG–2 video:
  - Bitrates up to 100 Mbps.
  - More choices in terms of spatial and temporal resolution.
  - Support for interlaced video (notion of even and odd field).
  - More possibilities for the chrominance sub-sampling.
  - More coding and quantization options.
  - Support for bitstream **scalability**.

H.263

- Initially (1993), the MPEG–4 group started developing a video coding standard for bitrates < 64 kbps, i.e., for **very low bitrates**.
- However, some time after, this line was reformulated into a much more ambitious objective: that of creating a standard for coding audiovisual objects.
- Due to the urgent need for a low bitrate standard (for example, for enabling video over the analog public telephone network or over wireless channels), the work was divided in two phases:
  - One, for the immediate development of a video coding standard for very low bitrates: recommendation H.263 (1995).
  - The other, directed to a more vast set of tools, originated the MPEG–4 standard.

H.263

- Recommendation H.263 specifies an algorithm for video coding, similar to that of H.261, for bitrates of about 22 kbps of a total of 28.8 kbps.
- The main differences between H.261 and H.263 are:
  - New formats available: sub-QCIF, 4CIF and 16CIF, in addition to those already supported by H.261, CIF and QCIF.
  - Possibility of using a motion vector per block as well as one motion vector per macroblock.
  - Half-pixel precision motion estimation and prediction of motion vectors.
  - Arithmetic coding.
  - PB-frames (bi-directional prediction, similar to that used in MPEG).

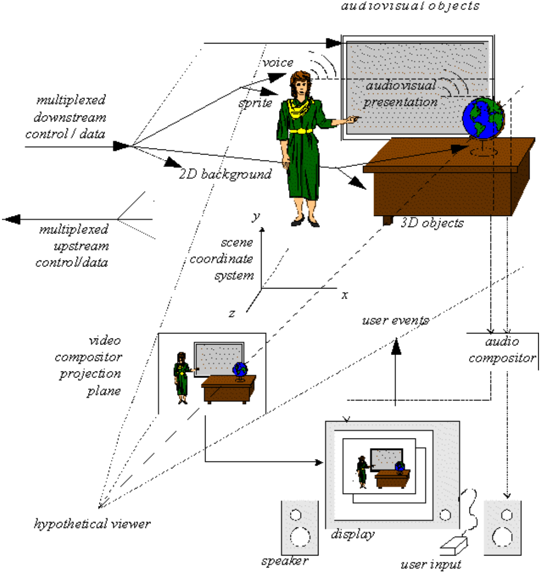
MPEG–4

- MPEG–4 (initial version in 1998) is a ISO/IEC standard providing tools for:
  - **Representing** audio, video or audiovisual data through **media objects** that can be natural (i.e., captured by a microphone or video camera) or synthetic (i.e., computer generated).
  - Describing the **composition** of these objects for creating composed objects and audiovisual scenes.
  - **Multiplexing and synchronizing** the data associated to the media objects, for transmission through the communication channels, providing an appropriate quality of service (QoS) to each object.
  - Enabling the **interaction** of the clients (receptor) with the audiovisual scene.

MPEG–4

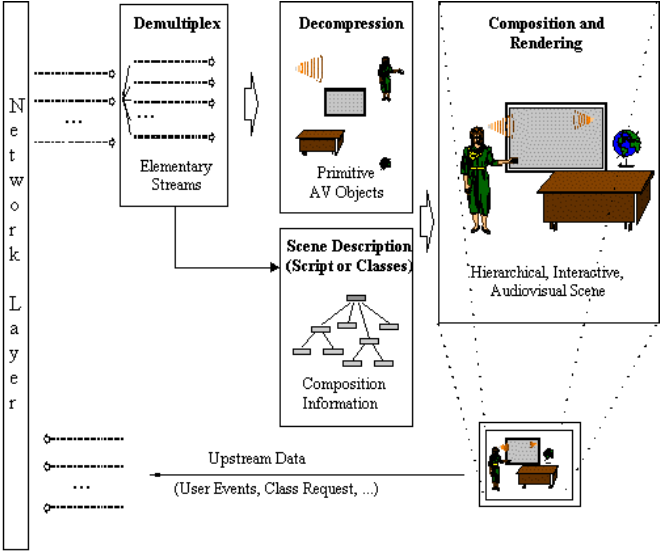
- MPEG–4 defines several primitive objects for representing **natural** and **synthetic** information, as well as **2D** and **3D** data.
- The **audiovisual scenes** are composed of these media objects, hierarchically organized:
  - Images (for example, a fixed background).
  - Video objects (for example, a person talking).
  - Audio objects (for example, the voice of the person, background music, . . .).
  - Text and graphics.
  - Synthetic talking heads and the corresponding text used by the speech synthesizer; animated synthetic bodies.
  - Synthetic sound.
  - . . .

MPEG–4



audiovisual objects
voice
audiovisual
presentation
multiplexed
downstream
control / data
sprite
2D background
3D objects
multiplexed
upstream
control/data
y
scene
coordinate
system
x
z
user events
video
compositor
projection
plane
audio
compositor
hypothetical viewer
display
speaker
user input

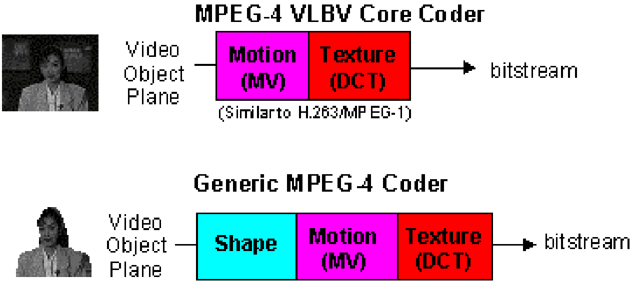
MPEG–4



Network Layer
Demultiplex
...
Elementary Streams
Decompression
Primitive AV Objects
Scene Description (Script or Classes)
Composition Information
Composition and Rendering
Hierarchical, Interactive, Audiovisual Scene
Upstream Data
(User Events, Class Request, ...)
...

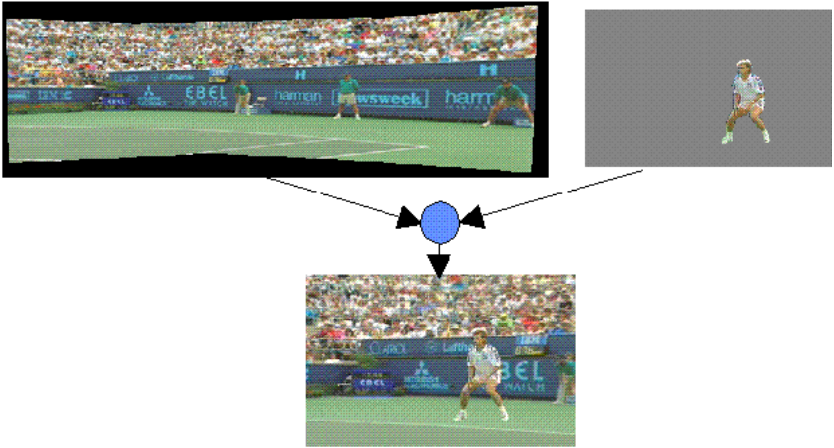
MPEG–4

- Conventional video coding is performed as in MPEG–1/2.
- In **content-based** coding, it is possible to encode regions with arbitrary shape, but, it this case, the shape of the object also needs to be efficiently represented.
- **Shape** is represented using a 8 bit transparency component or a binary mask.



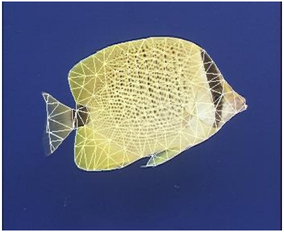
MPEG–4

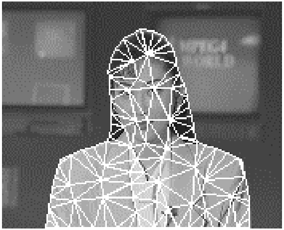
- Sprites:



MPEG–4

- MPEG–4 supports **synthetic visual objects**:
  - Parametric description of **human heads and bodies** (also body animation in Version 2).
  - Parametric description of **static or dynamical meshes** with texture mapping.





  - **Scalable texture coding.**

H.264/AVC

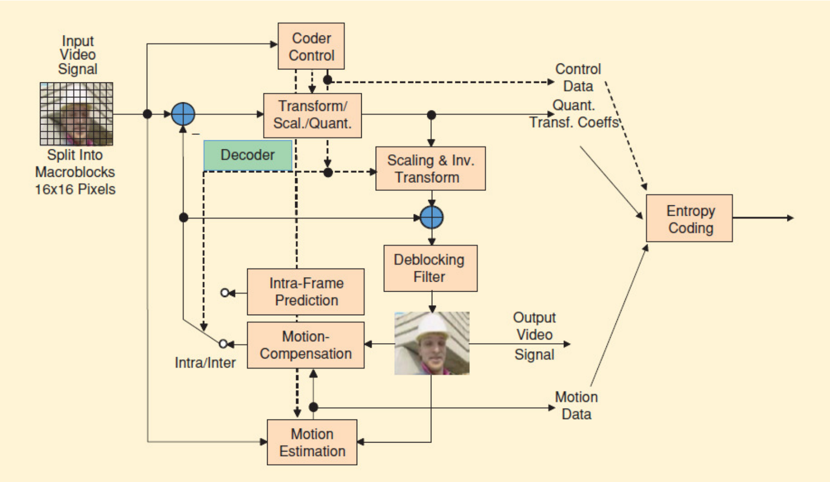
Overview

- H.264/AVC (Advanced Video Coding) was jointly developed by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC MPEG.
- It was finalized in March 2003 and approved by the ITU-T in May 2003.
- H.264/AVC provides gains in compression efficiency of up to 50% over a wide range of bit rates and video resolutions compared to previous standards.
- The decoder complexity is about four times that of MPEG-2 and two times that of MPEG-4 Visual Simple Profile.

H.264/AVC

Overview

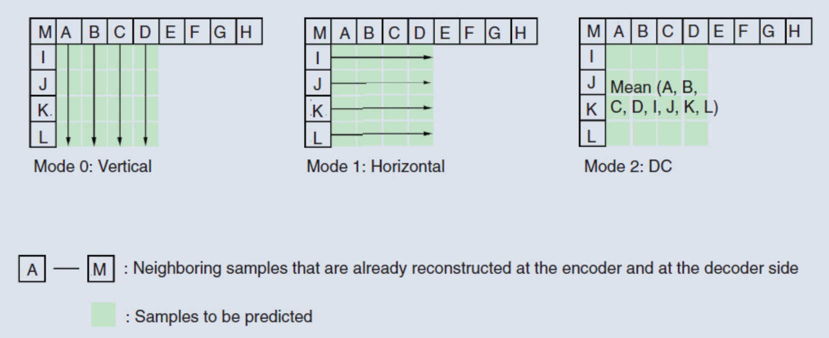
Block diagram of a typical encoding process of H.264/AVC



H.264/AVC

Intra prediction

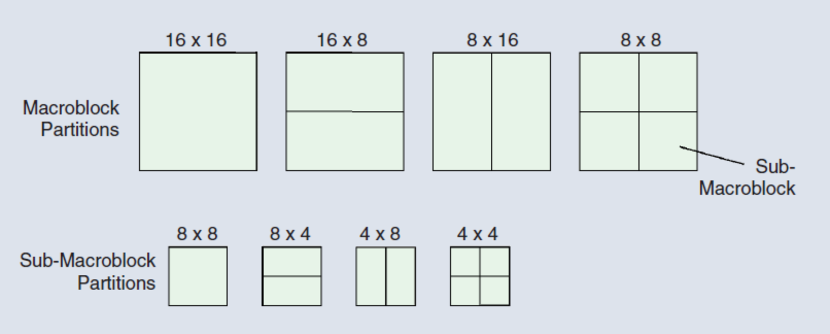
Three out of nine possible intra prediction modes for INTRA_4×4



H.264/AVC

Motion-compensated prediction

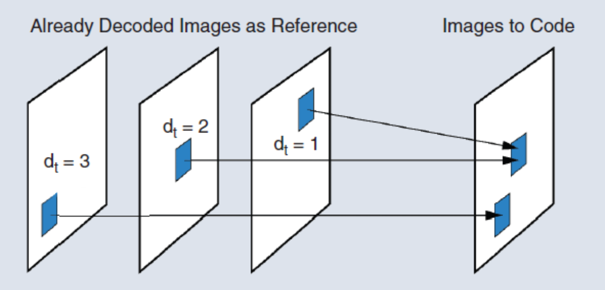
Partition of macroblock/sub-macroblock for motion-compensation



H.264/AVC

Motion-compensated prediction

Motion-compensated prediction with multiple reference images



H.264/AVC

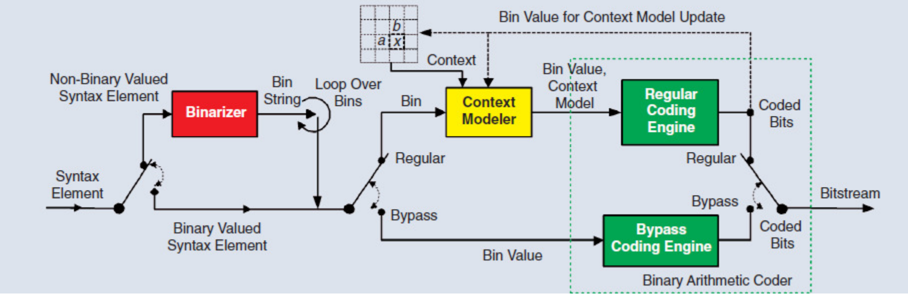
Entropy coding

- H.264/AVC provides two methods for entropy coding:
  - CAVLC, a low-complexity technique based on context-adaptive sets of variable length codes.
  - CABAC, a context-based adaptive binary arithmetic encoder.
- By incorporating context modeling, both methods offer a high degree of adaptation to the underlying source.
- CAVLC relies on 32 different VLCs. For typical coding conditions, it is 2–7% better than conventional codes.
- Typically, CABAC provides bit rate reductions of 5–15% compared to CAVLC.

H.264/AVC

CABAC

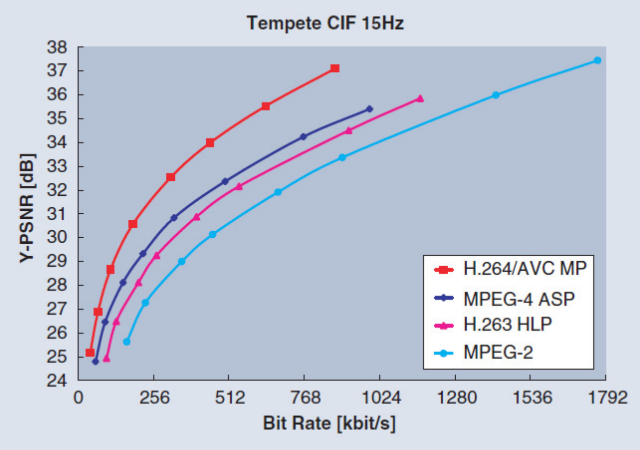
Context-based adaptive binary arithmetic coding



H.264/AVC

Performance

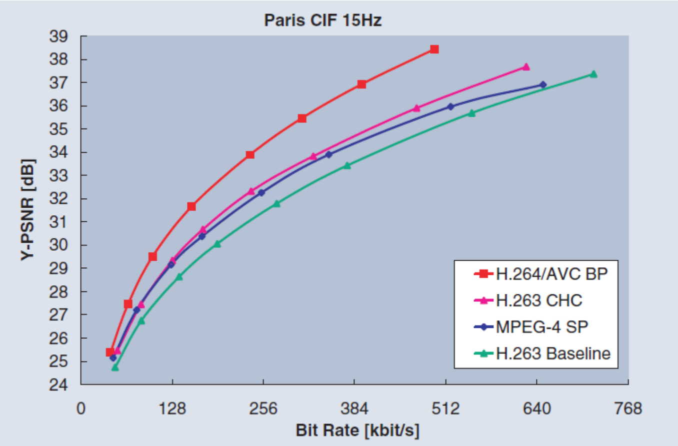
Video streaming application



H.264/AVC

Performance

Video conferencing application



H.265/HEVC

Overview

- H.265/HEVC (High Efficiency Video Coding) was (again) the result of a collaboration between the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC MPEG.
- It is also known as the MPEG-H Part 2 and the first version was finalized in 2013.
- H.265/HEVC can provide gains in compression efficiency of about 50%, when compared to H.264/AVC.
- This is mostly attained by further exploring existing techniques, but at a cost of increasing the complexity of the encoder.
- As with H.264/AVC, H.265/HEVC is dependent of a considerable number of patents, which is preventing its wide use...

H.265/HEVC

Block diagram