Video Compression
MPEG-4

Catania – 10/04/2008

Arcangelo Bruna

Advanced System Technology

Market's requirements for Video compression standard

- Application's dependent
  - Set Top Boxes (High bit rate)
  - Digital Still Cameras (High / mid / low bit rate)
  - Mobile multimedia (Low / very low bit rate)
  - Video conferences (Very low bit rate)

Video Compression Standards comparison

Video Compression	Market	Video Bitrate	Frame-accurate editing	Scalability	Still Image Mode	Lossless Mode
MPEG-1	Video CD authoring	1.0-1.5 Mbits/s @ 352x240x29.97 fps	No	Low	No	No
MPEG-2	DVD authoring	3.0-100.0 Mbits/s @ 720x480x29.97 fps	No	Low	No	No
MPEG-4	Internet Streaming Mobiles	0.3-1.0 Mbits/s @ 352x240x29.97 fps	No	High (Spatial and temporal)	Yes	No
MJPEG	Video Production	10.0-80.0 Mbits/s @ 720x480x29.97 fps	Yes	Low	Yes	No
DV	Professional Video Production Digital Video Cameras	25.0 Mbits/s @ 720x480x29.97 fps	Yes	Low	No	No
MJPEG2000	Professional Video Production Digital Video Cameras Video/image streaming	2.0-50.0 Mbits/s @ 720x480x29.97 fps	Yes	High (Spatial)	Yes	Yes

Advanced System Technology

3

MPEG 4

Advanced System Technology

4

MPEG – A Movie Compression Standards

- **MPEG is an acronym for "Moving Picture Experts Group". (www.mpeg.org), founded in 1988.**
- **The MPEG 1 was developed for video CD with 25/30 fps and 1.2/1.5 Mbps.**
- **The MPEG 2 allows up to 720x576 pels, 30 fps and 15Mbps (Digital TV and DVD).**
- **MPEG 4 (ISO/IEC 14496) is an object based video compression std designed for low bit rate applications, e.g. wireless communication, …**

MPEG – A Movie Compression Standards

- **MPEG-4 was developed in order to obtain a low/very low bit rate.**
- **It must allow to obtain as good quality as possible reducing all the spatio-temporal redundancies**
- **Lossy algorithms were inserted in order to improve the compression ratio**
- **Object description (i.e. Facial description) were inserted too**

MPEG - 4

- Natural and synthetic visual objects
- Rectangular or arbitrarily shaped pictures
- moving 2D meshes
- animated 3D face models
- texture for synthetic objects

MPEG4: profiles and levels

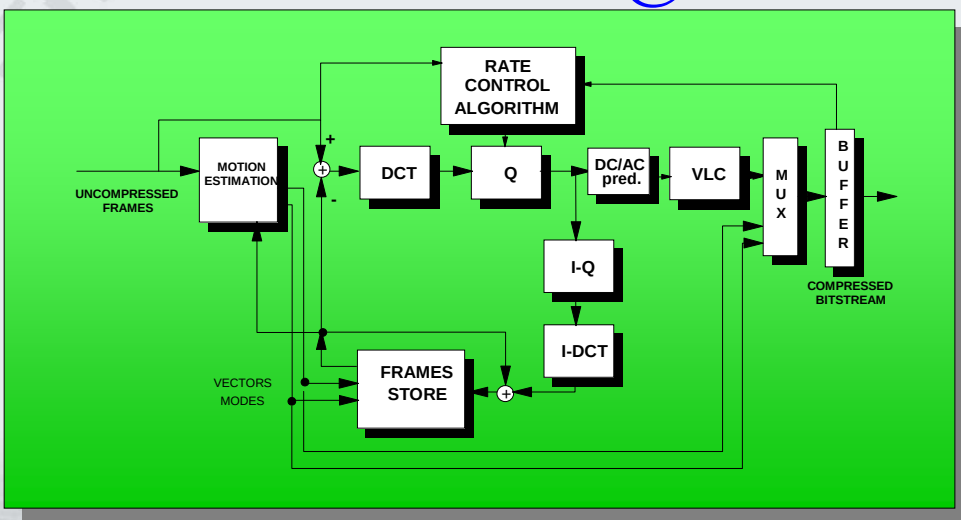
- The MPEG4 allows a lot of different features.
- A “profile” is a defined subset of the entire features.
- A “level” is a defined set of constraints imposed on parameters in the bitstream.

The Simple Profile Level 3 *(SP@L3)* will be explained

Structure of coded visual data

- Visual object sequence
- Video object (VO). It can be “video”, “still”, “face”, “mesh” *(just “video” type in the SP@L3)*
- Video Object Plane (VOP): it is the single frame
- Group of VOP (GOV)
- Macroblock (MB): it is composed by 16x16 pixels
  Note: the 4:2:0 subsampling is used
- Block: it is composed by 8x8 pixels

MPEG-4 – SP@L3



MPEG-4: the algorithms list

- **Motion estimation**
- **Discrete Cosine Transform (DCT/IDCT)**
- **Quantization (Q/I-Q)**
- **DC – AC prediction**
- **Entropy Coding (VLC)**
- **Rate control algorithm**

Motion Estimation

It allows to reduce the temporal redundancy.

Temporal Redundancy

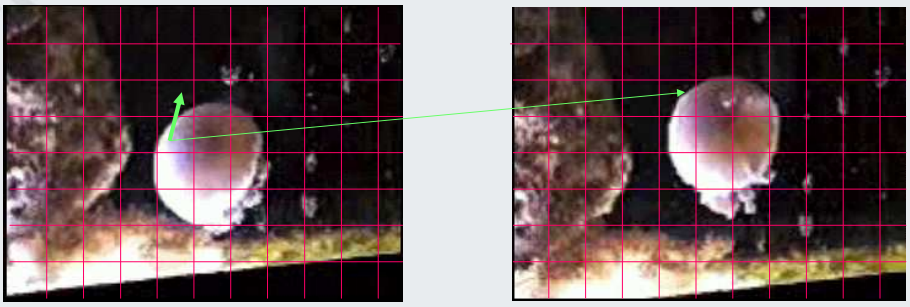
In a typical scene there will be a great deal of similarity between nearby images of the same sequence.

Motion Compensation Approach

Basic idea of ***Motion Compensation***:

- The image is divided in blocks
- For each block we search if there is a 'similar' block in the previous frame.
- If the similarity is not founded the block is encoded without any temporal reference (i.e. as a JPEG block).
- If a similarity is founded just the motion displacement (*motion vector)* and the difference with such block (*prediction error)* is encoded.

Motion Compensation Approach (cont.)



Reference frame

Current frame

Motion Compensation Approach (cont.)

Block Matching--how to find the matching block?

- Matching criteria:
  - In practice we couldn't expect to find the exactly identical matching block, instead we look for close match.
  - Most motion estimation schemes look for minimum Sum Absolute Differences (SAD) between blocks.

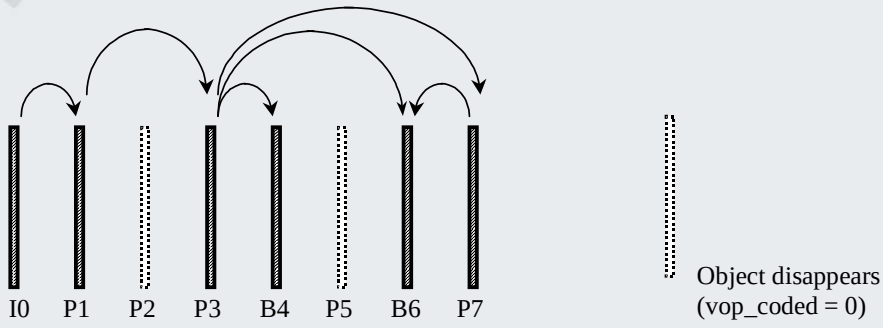
$$SAD = \sum_{n=1}^{N} \left| I_n(x,y) - I_n^{'}(x,y) \right|$$

- Matching block size:
  - How large the matching block will affect coding efficiency
  - Block size used in MPEG : 16×16 pixels

VOP type

- I (intra) frames: the frame is coded without any temporal reference
- P (predicted) frames: the frame is predicted from the previous I or P frame
- B (bidirectional) frames: the frame is predicted from the previous and the next I or P frame
- Not coded: the frame is not coded in the stream

Sequence example



DCT / I-DCT

- Perform the Discrete Cosine Transform and the inverse transformation (identical to the JPEG)
- It allows to compact the image energy in the low frequency coefficients

DCT formulas

$$F(u,v) = \frac{1}{4} C(u)C(v) \left[\sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right]$$

$$f(x,y) = \frac{1}{4} \left[\sum_{u=0}^{7} \sum_{v=0}^{7} C(u)C(v)F(u,v) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right]$$

where:

$$C(u), C(v) = \frac{1}{\sqrt{2}} \quad \text{for } u, v = 0;$$

$$C(u), C(v) = 1 \quad \text{otherwise}$$

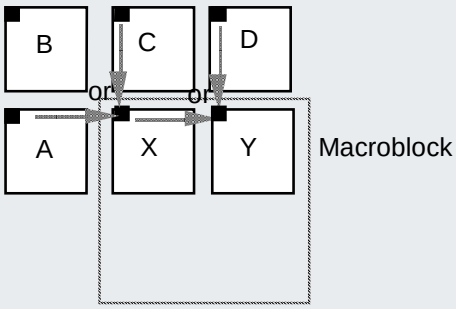
Q / I-Q

- Perform the quantization and the inverse quantization
- Two different quantization can be performed:

  1) Using a weighting matrix (as the JPEG);
  2) Using a fixed quantization value.

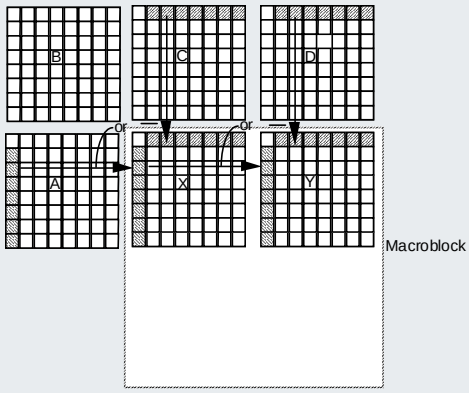
  Both the methods allow to modulate the quantization through a Quantization Parameter (QP). It can be modified at MB level.

DC prediction



- The DC coeff is always predicted by the DC in the previous block or in the block above.

AC prediction



- The first line's AC coeffs or the first column can be predicted by the same block used for the DC prediction

Entropy Coding (VLC)

- A Run-Length / Variable Length encoder is used to compress the quantized data.
- The Huffman codes are imposed by the standard

Rate control algorithm

- It allows to obtain the frame size as close as possible to a target value.
- This feature is very important in channel with constant bandwidth