

Hints on video coding

TLC Network Group
firstname.lastname@polito.it
<http://www.telematica.polito.it/>

Andrea Bianco - TNG group - Politecnico di Torino Computer Networks Design and Management-1

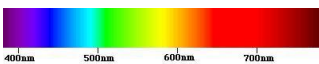
Summary

- Visual perception
- Analog and digital TV
- Image coding: hints on JPEG Standard
- Video coding
 - Motion compensation
 - MPEG: hierarchical data organization
 - MPEG-1
 - MPEG-2: scalability, profiles and levels
 - MPEG-4: content-based coding, sprite coding
 - Synchronization: MPEG-2 systems

Andrea Bianco - TNG group - Politecnico di Torino Computer Networks Design and Management-2

Visual perception

- Human eye is able to capture all wavelengths in the range 250-780 nm

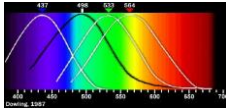


- Eye sensitivity depends on the wavelength: for a given energy level, the radiation is perceptually received as more or less intense depending on λ
- The colour is a function of the wavelength and of the energy that it emits or reflects
- Two receptors: retinal cones and rods. Cones are more sensitive to wavelength, rods to energy

Andrea Bianco - TNG group - Politecnico di Torino Computer Networks Design and Management-3

Colours

- Three families of cones exist, more sensible to short (blue), medium (green) and long (red) wavelengths



Normalized sensitivity of cones (white curves) and rods (black curve) as a function of wavelength

Colours

- All colours that the human eye is able to perceive can be created by mixing three «primary» colours
- Several triples of colours can be used as primary
- Normally red, green and blue are used for the reason outlined above
 - RGB coding (Red, Green, Blue)

Definitions

- **Intensity**: radiated energy per unit area
- **Luminance**: photometric measure. It represents the radiated energy per unit area weighted by a sensitivity function related to human visual perception.
- **Brightness**: absolute value. It is a subjective attribute of visual perception in which a source appears to be radiating or reflecting light.
- **Lightness**: relative perceptive response. Brightness of an area relative to the brightness of a similarly illuminated area that appears to be white (highly transmitting)

Lightness e Brightness

- Intensity and luminance are objective quantities, that can be measured with proper instruments
- Brightness and lightness are subjective quantities: they depend on many factor (among the others the luminance of the environment in which the human eye is) and are different from person to person
- Luminance perception is non linear
 - The lightness of a source whose luminance is 18% of the one of the reference source is roughly 50%

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-7

Luminance and crominance

- R, G and B components of a colour are strongly correlated:
 - This redundancy can be exploited to reduce the amount of information needed to represent a given colour
- Analog TV standard use separate signals for:
 - Luminance
 - Image representation using a grey scale system
 - Crominance
 - Colour information

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-8

Luminance and crominance

- Luminance and crominance are almost non correlated
- Luminance contains info on lightness and brightness,
 - For example defines figure contours
- Since the human eye is particularly sensible to lightness and brightness, the most fundamental image information is concentrated on the luminance

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-9

Analog TV

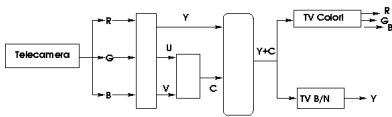
- PAL (Phase Alternate Line) standard is based on YUV coding
 - Y is the luminance, U and V are the two chrominance
- YUV components can be obtained from RGB components via a linear transformation
 - $Y = 0.3R + 0.59G + 0.11B$
 - $U = 0.493 (B - Y)$
 - $V = 0.877 (R - Y)$
- NTSC (National Television System Committee) standard exploits YIQ
- RGB to YIQ transformation is also linear but with different coefficients for I and Q
 - $Y = 0.3R + 0.59G + 0.11B$
 - $I = 0.74 (R - Y) - 0.27 (B - Y)$
 - $Q = 0.48 (R - Y) + 0.41 (B - Y)$

Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management- 10

Analog TV

- Black and white analog TV exploits only the luminance signal, colour TV also the chrominance



Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management- 11

TV standards

| PAL | | NTSC |
|-------|---------------------------|-------|
| 4:3 | Aspect Ratio | 4:3 |
| 625 | Number of lines per frame | 525 |
| 25 | Number of frames/s | 29.97 |
| 8 MHz | Transmission bandwidth | 6 MHz |

Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management- 12

Digital video

- The ITU-R 601 standard defines a digital format for PAL and NTSC
- Both formats have 720 samples per line
 - Corresponding sampling frequency is 13.5 MHz
- Y, U and V components are independently sampled
 - Since U and V are less important, they are sub-sampled with respect to Y with ratios 4:2:2 or 4:1:1
- Using 8 bit to represent each component of each sample, the overall bitrate is
 - $(13.5 + 2 \cdot 6.75) 106 \text{ sample/s} \cdot 8 \text{ bit/sample} = 216 \text{ Mbit/s}$
- More precisely, in NTSC the useful lines (no retracing) are 486 with 720 samples per line
 - $720 \text{ sample/line} \cdot 486 \text{ lines} \cdot 30 \text{ frame/s} \cdot 8 \text{ bit/sample} = 84 \text{ Mbit/s}$ (luminance only)

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 13

Digital video

- HDTV standards exploit up to
 - 1440 or 1920 samples per line
 - 1152 lines per frame
 - 60 frames/s
- Resulting bitrate can easily exceed 1 Gbit/s
 - Only professional studios can store, transmit and elaborate flows at this speed
 - Compression techniques become fundamental
- Videoconferencing standards
 - CIF (H.261): 4:1:1 ratios.
 - 352 sample/line, 288 lines/frame (luminance), 36 Mbit/s
 - QCIF
 - 176 sample/line or 144 lines/frame (luminance), 18 Mbit/s

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 14

Video compression

- Video presents a high level of redundancy
- Statistical redundancy :
 - *Spatial*: adjacent pixels in the same frame are correlated (intra-frame compression)
 - *Temporal*: pixel in the same position in consecutive frames are correlated (inter-frame)
- Perceptive redundancy: related to characteristics and features of human vision system
- Redundancy can be exploited to compress video

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 15

Compression

- Entropic coding
 - Do not exploit info on the source characteristics
 - Huffman algorithm
 - Shorter representation for more likely symbols
 - Run-Length encoding (RLE)
 - exploits correlation among adjacent elements
 - Long sequences of symbols with the same value are coded as pairs (value, number of repetitions)
 - Lossless
 - Level of compression somehow limited

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 16

Compression

- Source coding
 - Predictive
 - exploits correlation among adjacent elements,
 - e.g., the dynamics of the difference is smaller than the dynamics of the original signal (like DPCM)
 - Transform
 - examine the image in a domain in which the redundancy contained in the information can be better highlighted
 - FFT (Fast Fourier Transform) and DCT (Discrete Cosine Transform) highlight the fact that most of the image information is concentrated in low frequency spectrum components

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 17

Compression

- Vectorial: takes a block of data (vector) and maps it to the element that best match it in a pre-defined codebook
 - Block can be mono or bi dimensional
- Layered: the image is hierarchically decomposed in several layers
 - Each layer enhances the image quality of the previously defined layers
 - The decomposition is obtained through sampling at different frequencies or in different sub-bands
- Source coding is often lossy
- Very often hybrid coding techniques are used:
 - Several compression schema are used in series to obtain better performance

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 18

JPEG standard

- Image compression standard approved in 1992 by the “Joint Photographic Experts Group” of ISO
- Lossy coding exploiting the human vision perception to reduce the redundancy
- The compression ratio can be varied depending on the target quality level of the compressed image

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 19

JPEG

- The algorithm operates independently on luminance and chrominance (represented with three different matrices)
 - It may be necessary to exploit the transformation RGB->YUV or RGB->YJQ
- The three matrices are divided in 8x8 blocks
- The DCT transform is applied to each block
 - Linear transformation (lossless)
 - Modifies the representation system of the image
 - Image represented in the frequency domain
- A quantization block is applied to the transform
 - Lossy

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 20

JPEG

- The “continuous” component of the block is stored in the upper left corner of the matrix
- Moving from left to right and from top to bottom the elements of the transformed block represent increasing frequencies
- Low frequency components contain the most significant information on the image
 - They are quantized with a better granularity
- The DC “component” is coded as a difference with respect to the DC component of the previous block

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 21

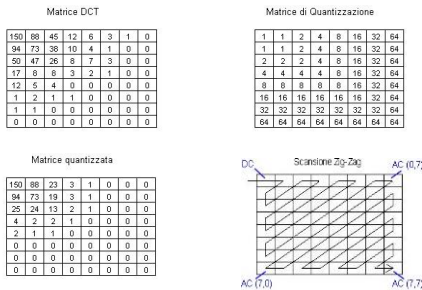
JPEG

- Most of the high frequencies components are negligible or null due to the coarser quantization
- AC coefficient are encoded according to RLE, following a zig-zag order in the matrix, to put in sequence the high frequencies null coefficients
- Finally, the pairs (value, number of repetitions) are coded according to the Huffman method
- The quantization granularity determines the compression ratio and the level of degradation of the compressed image
- Coding and decoding have the same complexity

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-22

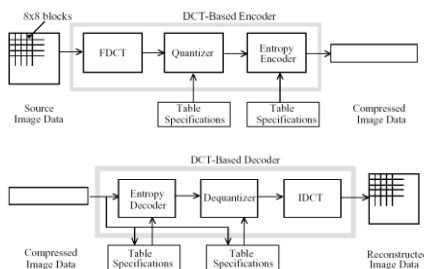
Example



Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-23

Coder and decoder



Da [1]

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-24

MPEG coding

- A video stream is composed by a sequence of images (frame)
- Single frames are compressed according to a scheme similar to JPEG
- Temporal correlation among frames is exploited using techniques such as
 - differential coding and prediction
 - motion compensation (to identify object movement)

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 25

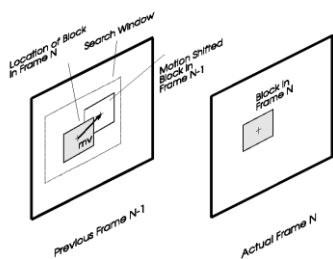
Motion compensation

- Frame N is divided in blocks
- For each block, a motion vector is estimated
 - All blocks in frame (N-1) with adjacent positions to the considered block in frame N are examined to select the most similar one
- The block is coded as the difference with respect to the previous block plus the motion vector
- Works well for translation, not well for zoom, rotation
- Block is not a physical object
- Coding operation is more complex and time consuming than decoding

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 26

Motion compensation)



Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 27

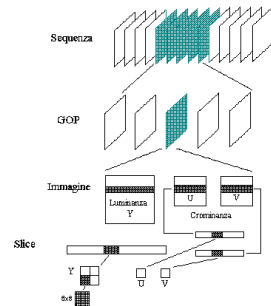
MPEG – Data organization

- Data are hierarchically organized in layers
- Each layer supports a signal processing function and a logic function
- Six layers
 - Sequence
 - Group of Pictures (GOP)
 - Picture
 - Slice
 - Macroblock
 - Block

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-28

MPEG – data organization



Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-29

Sequence and GOP

- The sequence defines the video flow in terms of
 - Frame size, number of frames per second and bitrate
- Within each sequence, GOPs are identified
 - Groups of contiguous, independent, pictures classified as I, P, B:
 - Intra-pictures
 - inter-frame Predicted pictures
 - Bi-directional inter-frame predicted pictures
 - A GOP can contain a variable number of I, I and P, or I, P and B pictures

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-30

Pictures

- I pictures
 - Coded/decoded in isolation, with no reference to other images
 - Can be used as a reference to code P or B pictures
 - Identify the starting point of a GOP
 - Useful to support random access
 - Limit error propagation
 - Compression level limited
- P pictures
 - Coded referring to the nearest I or P picture
 - Can be used as a reference to code P or B pictures
- B pictures
 - Coded referring to two (previous or next) I or P pictures
 - Never used as a reference
- A large number of P and B pictures
 - permit to increase the compression ratio but also coding delay and complexity
 - makes it more difficult the random access
 - makes the flow more sensitive to errors

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 31

Pictures

- Since B images refer also to pictures to be played back later, the visualization order is different from the coding and transmission order
- Example
 - Visualization order: I0 B1 B2 P3 B4 B5 P6
 - Dependencies:
 - I0 -> none
 - P3 -> I0
 - B1 and B2 -> I0 e P3
 - P6 -> P3
 - B4 e B5 -> P3 e P6
 - Coding and transmission order: I0 P3 B1 B2 P6 B4 B5

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 32

Slices

- Slices are portion of images that include an integer (variable) number of macroblocks
- A slice does not contain spatial references to other slices
 - Can be decoded independently (in isolation)
- Exploited for synchronization purposes

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 33

Macroblocks

- The Macroblock (MB) is the fundamental unit for motion prediction
- In MPEG-1 the macroblock size is 16x16 pixels
- Each macroblock can be of type:
 - *Skipped*: it is identical to the MB in the same position in the reference image
 - It is neither coded nor transmitted
 - *Inter*: differentially coded with respect to another MB in the reference image
 - motion vector, values of the difference and quantization levels are transmitted
 - *Intra*: coded in isolation
 - Samples values and quantization levels are transmitted

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 34

Macroblocks

- The MB type is a function of the picture type:
 - Type I: only Intra MB are used
 - Types P and B: can exploit Intra, Inter and Skipped MB
- For B pictures and Inter MB, the prediction can be:
 - *Backward*: the motion vector refer to a MB of a past picture
 - *Forward*: the motion vector refers to a MB of a successive picture
 - *Interpolated*: exploit two motion vectors, one referring to the a past picture one to a following picture; the prediction is computed on the average values of the two MBs

Andrea Bianco – TNG group - Politecnico di Torino

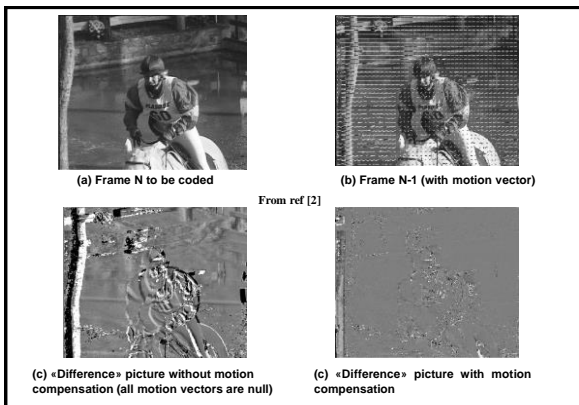
Computer Networks Design and Management- 35

Macroblocks

- The standard does not specify how to compute the motion vectors and the criteria to choose the MB type
- Usually, block matching techniques are used:
 - The algorithm looks for the motion vector that minimizes the energy of the difference between the block to be coded and the one to which the vector refer to
 - If the energy difference is below a pre-defined threshold the Inter MB is chosen (differences are transmitted)
 - Otherwise, an Intra MB is chosen (the whole MB is transmitted coded in isolation)
- Compression more difficult than decompression

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 36



Blocks

- *Blocks* are the fundamental data unit over which the spatial redundancy is applied
- Block size is 8x8 pixels
- Blocks are represented as one luminance matrix and two chrominance matrices
- Chrominances are sub-sampled in ratios 4:1:1. Thus a mB contains 4 luminance and 2 chrominance 8x8 matrices
- The single block coding follow the JPEG standard: DCT transform, quantization, differential coding for the DC component, RLE and entropic compression with zig-zag scanning for AC components

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 38

MPEG-1 standard (1992)

- Defined for VHS quality for bitrate up to 1.5Mbit/s (close to an audio CD bitrate)
- Interlaced pictures are not supported
- The MPEG-1 “constrained parameter set” (set of reference for standard implementation) is:
 - Pixel per line ≤ 768
 - Lines per picture ≤ 576
 - Macroblocks per picture ≤ 396
 - Macroblocks/s ≤ 9900
 - Pictures/s ≤ 30
 - Bitrate ≤ 1.856 Mbit/s

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 39

MPEG-2 standard (1994)

- Defined for digital TV and HDTV, CD storing, terrestrial and satellite broadcasting, interactive retrieval
- Main features:
 - Video quality not worse than PAL/NTSC
 - Support for interlaced pictures
 - Video scalability (the picture quality can be progressively reduced in case of transmission error/losses)
 - Compatible with MPEG-1
 - Definition of Profiles and Levels to ease the interoperability among partial implementation of the standard

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 40

Video scalability

- Data flow is decomposed in a “base” layer and in “enhancement” layers
- Successive layers enhance the offered video quality with respect to the previous layers
- A receiver may decode only a given number of layers, depending on the amount of available resources (display, processor, etc.)
- Video scalability may be SNR, spatial or temporal, depending on the decomposition criteria adopted
- Base layer may get better service (e.g. high priority) in the transmission system

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 41

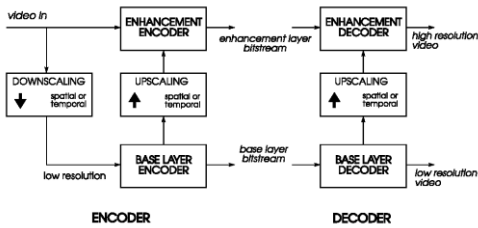
Video scalability

- SNR: base layer uses a coarse granularity for DCT coefficients, enhanced layers use a more fine granularity
- Spatial: changes the picture spatial resolution
 - for example, the base layer subsamples the picture, enhanced layers transport additional pixel information
 - useful to support display of different size
- Temporal: changes the temporal video resolution
 - for example enhanced layers increases the number of images/s
 - useful also for stereoscopic vision (a left and a right channel for the same picture)

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 42

Scalable coder and decoder



From ref [2]

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 43

Profiles and levels

- A profile defines a set of supported algorithms (additional to those of lower profiles)
- A level defines the supported parameter range (picture size, number of pictures/s, bitrate)
- The pair profile-level identifies the decoder supported functionalities
- All decoders should support at least the MAIN profile with level MAIN

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 44

Profiles

| | |
|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| HIGH | All the features of the SPATIAL profile plus support for: <ul style="list-style-type: none"> • Coding with 3 levels of spatial and SNR scalability • Colour coding YUV 4:2:2 |
| SPATIAL scalable | All the features of the SNR profile plus support for: <ul style="list-style-type: none"> • Scalable spatial coding (2 levels) • Colour coding YUV 4:0:0 |
| SNR scalable | All the features of the MAIN profile plus support for: <ul style="list-style-type: none"> • Scalable SNR coding (2 levels) • Colour coding YUV 4:2:0 |
| MAIN | Non scalable coding algorithms plus support for: <ul style="list-style-type: none"> • Interlaced video • Random access • Bi-directional prediction (B-pictures) • Colour coding YUV 4:2:0 |
| SIMPLE | As the MAIN profile but: <ul style="list-style-type: none"> • Does not support bi-directional prediction • Colour coding YUV 4:2:0 |

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 45

Levels

| | HIGH | HIGH 1440 | MAIN | LOW |
|---------------------|------|--------------|------|-----|
| Samples/line | 1920 | 1440 | 720 | 352 |
| Lines/frame | 1152 | 1152 | 576 | 288 |
| Frames/s | 60 | 60 | 30 | 30 |
| Bitrate (Mbit/s) | 80 | 60 | 15 | 4 |

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 46

MPEG-4 standard (1998)

- Objectives
 - Robustness in error prone environment (wireless networks/links, congested links, etc.)
 - High interactivity level, with the possibility of modify and store data in a very flexible way
 - Efficient coding of both natural and syntetich infos
 - Efficient compression, with support for bitrate as low as 64kbit/s
- “Content Based” approach
 - Separately identifies and codes objects in a video stream
 - Video is composed by “putting together” the various objects

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 47

Video objects

- A video object is a sequence of bitmap of any shape
 - “Video Object Planes”
- Shape and position of VOP vary over time
- For every object the transmitted infos are:
 - Shape
 - Trasparency
 - Spatial coordination
 - Scaling and rotation

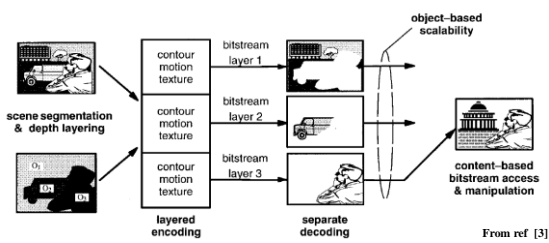
Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 48

Video objects

- Every object is coded on a separate stream
- The receiver can:
 - Decode only some objects in the flow
 - Add new objects
 - Modify the representation parameters of objects
- It is also possible to refer to objects contained in a local library at the receiver

Video objects



Other objects

- Audio objects
 - Sounds produced by the different instruments in an orchestra
 - Voices in a conversation
- Synthetic objects
 - Superimposed text
 - Computer animated objects
 - Faces, human figures, "texture mapped wire-grid"

Structure

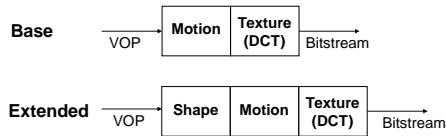
- The standard supports rectangular pictures as MPEG-1 e MPEG-2
- “VLBV Core” is a portion of the standard that defines the real time coding technique of flows:
 - non content-based
 - at very low bitrate (5 – 64 kbit/s)
 - with high error resilience
 - at low delay
 - at low complexity
- “HBV Core” provides the same functionalities but with higher bitrate (few Mbit/s)

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-52

Structure

- Other portions of the standard add content-based functionalities to VLBV e HBV coder and decoder
- In the below example, a VLBV base coder adds content-based infos thanks to a block that defines the VOP shapes

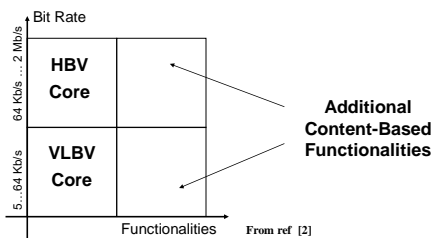


From ref [2]

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-53

Structure



VLBV = Very Low Bitrate Video

HBV = High Bitrate Video

From ref [2]

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-54

Sprite Coding

- *Sprite Coding* is a technique that exploits the presence of static, large size portions of the picture
 - Background or landscape
- The sequence is decomposed in “foreground” and “background” sprite
- For the foreground, all object parameters are transmitted every frame
- For the background, the full bitmap is transmitted only once
 - In the other frames only the motion of the camera framing the background is transmitted

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 55

Sprite Coding

- In the following example,
 - the foreground sprite is the tennis player
 - the background is the field and the audience
- Transmission
 - First, frame 200 containing all background info
 - Later, all the parameters of the foreground and the motion parameter of the background (translation, rotation, zoom...)

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management- 56

Frame 1



Frame 50



Frame 100



Frame 200



From ref [3]

Foreground and background

foreground

flexible 2D-object
with coherent motion



SA-DCT: 12 000 bit/frame
motion: 200 bit/frame
contour: 500 bit/frame
=> ca. 320 kbit/s

background

rigid 3D-background
with global camera motion



SA-DCT: 7000 bit/frame
motion: 140 bit/frame
=> ca.180 kbit/s From ref [3]

Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management- 58

Performance

- Sprite coding technology permits to obtain very high compression ratios with a good quality
- The need to separate foreground and background makes the technique easier to be used in multimedia database, where off-line processing is easy
 - Not perfectly suited for real time broadcasting

Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management- 59

Performance

- Same picture extracted from a sequence coded according to MPEG-1 (left) and MPEG-4 (right) for the same bitrate (1 Mb/s)



From [3]

Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management- 60

Synchronization: MPEG-2 Systems

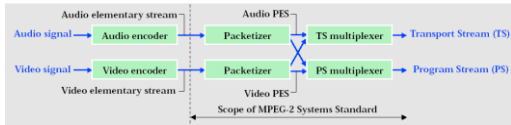
- It is the part of the standard that defines the syntax and the semantics of the bitstream
- Specify how to multiplex several flows on the same bitstream and how to synchronize them during the decoding phase
- The multiplexing criteria (how to multiplex packets generated by different sources) is not specified
- An *Elementary Stream* is the coded flow produced by a single video or audio source

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-61

MPEG-2 Systems

- Once segmented in packets, it is named *Packetized Elementary Stream (PES)*
- PES are multiplexed into a *stream*
- Two types of stream: Program Stream and Transport Stream



Da "Broadcast Technology" no.11, Summer 2002

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-62

Time-stamp

- PES include synchronization time-stamp in the header:
 - SCR (System Clock Reference): provides the time reference for the demultiplexing of PES of a program
 - DTS (Decoding Time Stamp): specify the time instant at which each pictures should be decoded
 - PTS (Presentation Time Stamp): specify the time instant at which each picture should be visualized

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-63

Stream

- Program Stream (MPEG-1 e MPEG-2):
 - Multiplex audio and video source with a common base time, equivalent to a TV program
 - Defined to store info on CDs and DVDs
 - Based on PS Packs packets *PS Packs* of variable size, ranging from 1 to 64 Kbyte
- Transport Stream (solo MPEG-2):
 - Multiplex a given number of programs, each one with its time base
 - Defined for broadcasting TV via cable, satellite, etc.
 - Fixed packet size of 188 byte

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-64

Transport Stream

- Every packet in the stream contains a "Packet ID" (PID) that identifies the elementary stream to which it belongs to
- PID 0 is reserved and transports the info related to the "Program Association Table" (PAT)
- The PAT associates every program contained in the stream to a "Program Map Table" (PMT), specifying the transported PID
- The PMT lists all PID of the elementary stream of the program (audio, video, ...)

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-65

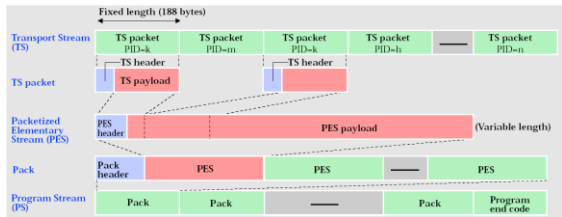
Demultiplexing

- The decoder, to demultiplex program **P**:
 - Extract packets with PID 0 and rebuilds the PAT
 - In the PAT it reads the PID **X** of the packets containing the PMT of program **P**
 - Extracts packets with PID **X** and builds the PMT of program **P**
 - Extracts all packets with one of the PID listed in the PMT (**Y**, **Z**, etc.)

Andrea Bianco – TNG group - Politecnico di Torino

Computer Networks Design and Management-66

Data organization



Da "Broadcast Technology" no.11, Summer 2002

Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management-67

Bibliography

1. G. K. Wallace, "The JPEG Still Picture Compression Standard", IEEE Transactions on Consumer Electronics, Vol. 38, No. 1, Feb. 1992, pp. xviii-xxxiv
2. T. Sikora, "MPEG Digital Video-Coding Standards", IEEE Signal Processing Magazine, Sept. 1997, pp. 82-100
3. P. Kauff, B. Makai, S. Rauthenberg, U. Golz, J. L. P. De Lameilleure, T. Sikora, "Functional Coding of Video Using a Shape-Adaptive DCT Algorithm and an Object-Based Motion Prediction Toolbox", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 7, No. 1, Feb. 1997, pp. 181-196
4. Color FAQ, <http://www.poynton.com/>
5. L. Chiariglione home page, <http://www.chiariglione.org/>
6. MPEG-2 Tutorial, <http://www.breit.com/mpeghtml/MPEGindex.htm>

Andrea Bianco - TNG group - Politecnico di Torino

Computer Networks Design and Management-68