ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Multimedia Technology

Section # 13: Video coding: MPEG-1/2

Instructor: George Xylomenos

Department: Informatics

Contents

- MPEG-1
- Frame groups
- Data stream
- MPEG-2
- Scalable coding
- Program and Transport streams

MPEG-1

Class: Multimedia Technology, **Section # 13:** Video coding: MPEG-1/2

Instructor: George Xylomenos, **Department:** Informatics

What is MPEG? (1 of 2)

- MPEG: Motion Pictures Experts Group
  - Similar name (and approach) to JPEG
  - Goal: digital video and audio
  - Media distribution or transport
    - No need for worldwide compatibility
    - Distribution needs random access
  - MPEG-1 (1990): CD-ROM, T1
  - MPEG-2 (1996): DVD, DTV
  - MPEG-4 (two “phases”): HDTV, mobile, interaction

What is MPEG? (2 of 2)

- Based on previous standards
  - JPEG: intraframe coding
  - H.261: interframe coding
- MPEG system: set of standards
  - Audio coding (MP3, AAC, …)
  - Video coding
  - Data multiplexing

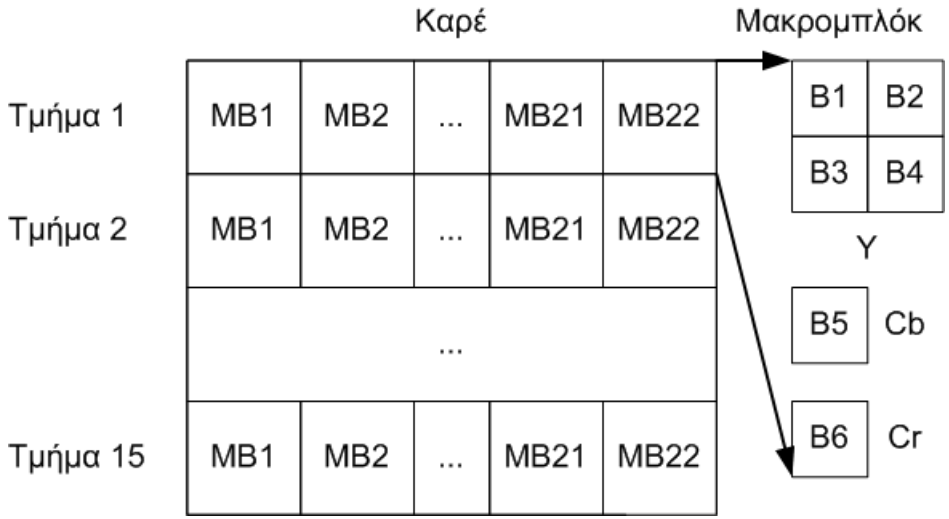
MPEG-1 goals (1 of 2)

- Bitrate: 1.2-1.5 Mbps
  - 1x CD-ROM (1.2 Mbps) or T-1 line (1.5 Mbps)
- Image format
  - YCbCr color, (4:1:1 4:2:0) subsampling
  - 8 bits per pixel per component
- 14 pixel aspect ratios
  - 1:1 is square pixels (as in computers)
  - Different ratios for 4:3 and 16:9 screens

MPEG-1 goals (2 of 2)

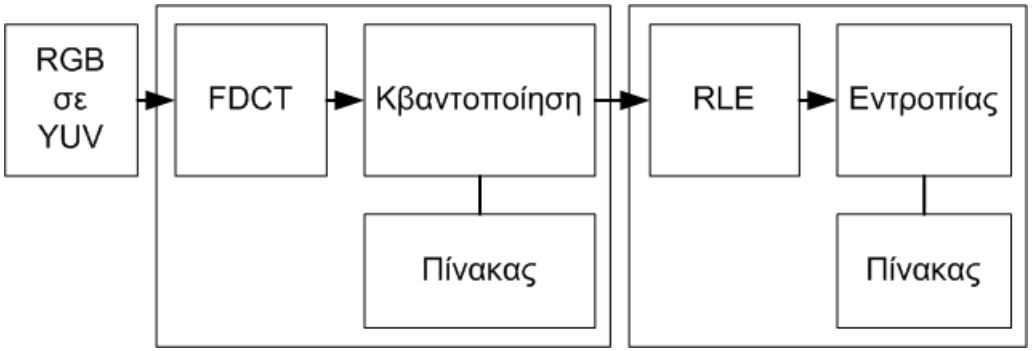
- Standard Interchange Format (SIF)
  - NTSC
    - 352 x 240 (luminance)
    - 176 x 120 (chrominance)
  - PAL/SECAM
    - 352 x 288 (luminanc)
    - 176 x 144 (chrominance)
  - Frame rates from 23.97 Hz to 60 Hz

MPEG-1 frame structure



- Macroblock (MB)
  - 16 x 16 luminance, 8 x 8 chrominance
- Slices: horizontal MB strips
  - 352 x 240 (NTSC): 15 slices x 22 MBs
  - 352 x 288 (PAL/SECAM): 18 slices x 22 MBs

I-Frames (1 of 2)

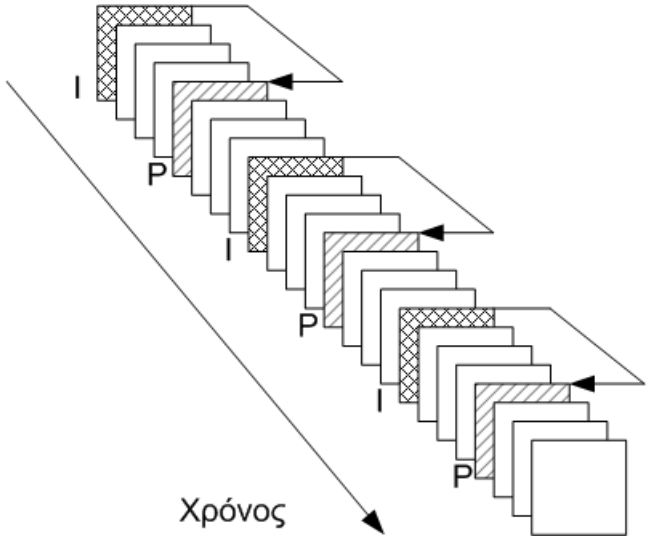


- I-frames: intracoded
  - Coded without references to other frames
  - Ideal for random access
  - Nearly the same as baseline JPEG
    - Minor changes for simplification

I-Frames (2 of 2)

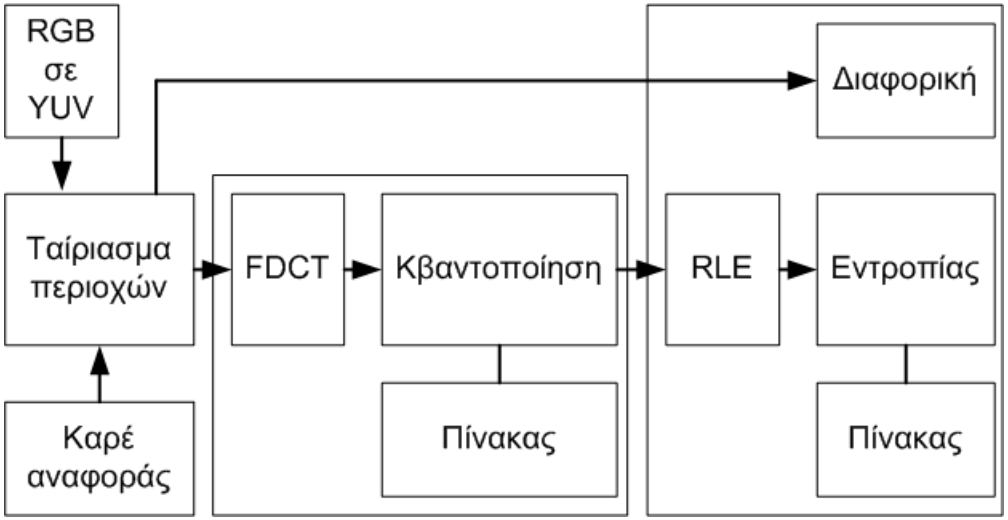
- Input: Block (8x8)
  - DCT transform
  - Coefficient quantization
  - DPCM for DC coefficients
  - Zig-zag sequence for AC coefficients
  - Run Length Encoding
  - Huffman-like entropy coding

P-Frames (1 of 2)



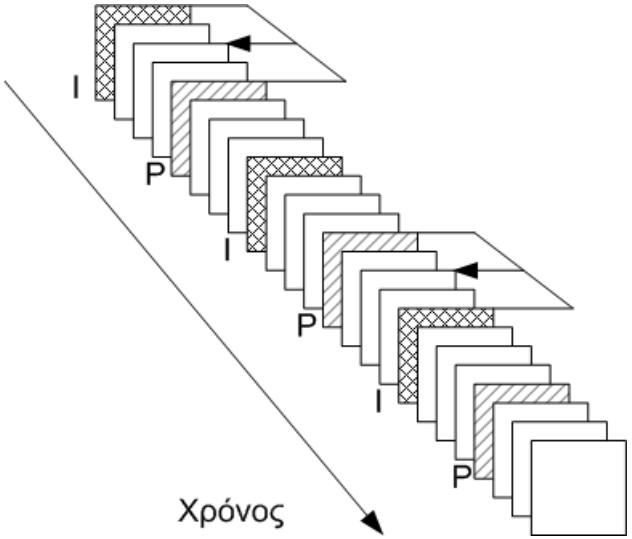
- P-frames
  - Motion compensation relative to reference
  - Reference: preceding I-frame or P-frame
  - Location of most similar area
    - Half pixel resolution
  - MBs may also be intracoded when no good prediction found

P-Frames (2 of 2)



- Motion vector
  - DPCM coded between consecutive MBs
- Block differences: DCT
  - Same coding for DC/AC coefficients

B-Frames



- B-frames
  - Prediction with two reference frames
    - One or two motion vectors
  - Same encoding as with P-frames
  - Based on periodic creation of I- and P-frames

D-Frames

- D-frames
  - Intracoded
  - Lower quality than I-frames
    - Only DC coefficient coded
  - Coded in addition to I-frames
  - Not used for prediction
  - Intended for fast forward/rewind
  - Dropped in newer MPEG versions

Bit rate

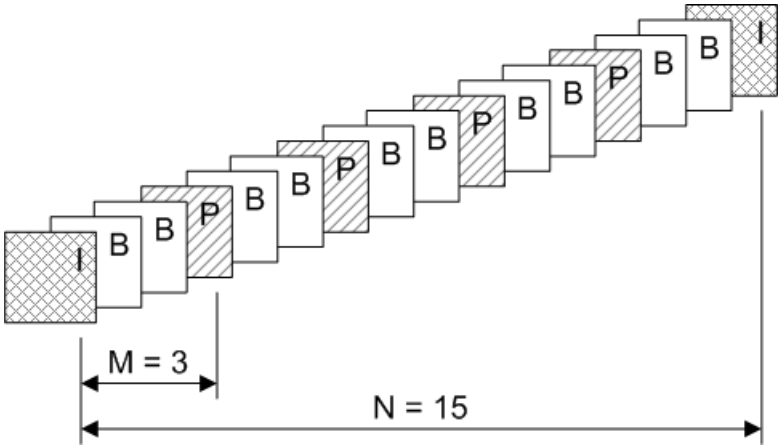
- Achieving a fixed bit rate
  - Quantization scale factor
    - Similar idea to H.261
    - Integer between 1 and 31
  - Used to multiply quantizers
    - In I-frames we have a quantizer matrix (as in JPEG)
    - In P/B-frames we have a single quantizer (as in H.261)
  - Monitors buffer occupancy
  - Transmitted whenever it changes value

Frame groups

Class: Multimedia Technology, **Section # 13:** Video coding: MPEG-1/2

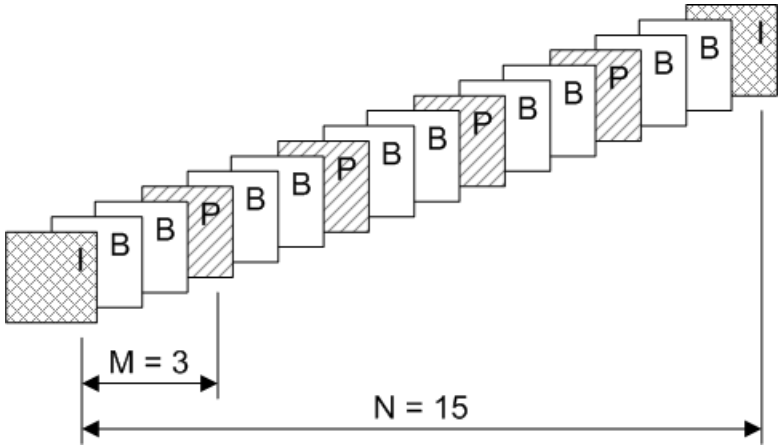
Instructor: George Xylomenos, **Department:** Informatics

Group of pictures



- Group of Pictures (GOP)
  - Starts with an I-frame
  - Ends right before the next I-frame
  - Includes all intermediate P- and B-frames
  - N: minimum random access unit

Group of frames



- Group of frames
  - Starts with an I- or a P-frame
  - Ends right before the next P- or I-frame
  - M: minimum decoding unit

Frame buffering (1 of 2)

- How long is each frame buffered?
  - I-frame: until next P-frame
  - P-frame: until next P- or I-frame
  - B-frame: no buffering
    - Never used as reference frames
    - Need next P- or I-frame for decoding

Frame buffering (2 of 2)

- Display order vs decoding order
  - I-frames do not reference other frames
  - P-frames reference previous I or P-frame
  - But B-frames also reference future frames!
  - Before encoding a B-frame, we need to wait
    - First, we encode the next I or P-frame
    - Then, we encode all the intermedia B-frames
    - Decoder must also start with next I or P-frame

Frame sequences

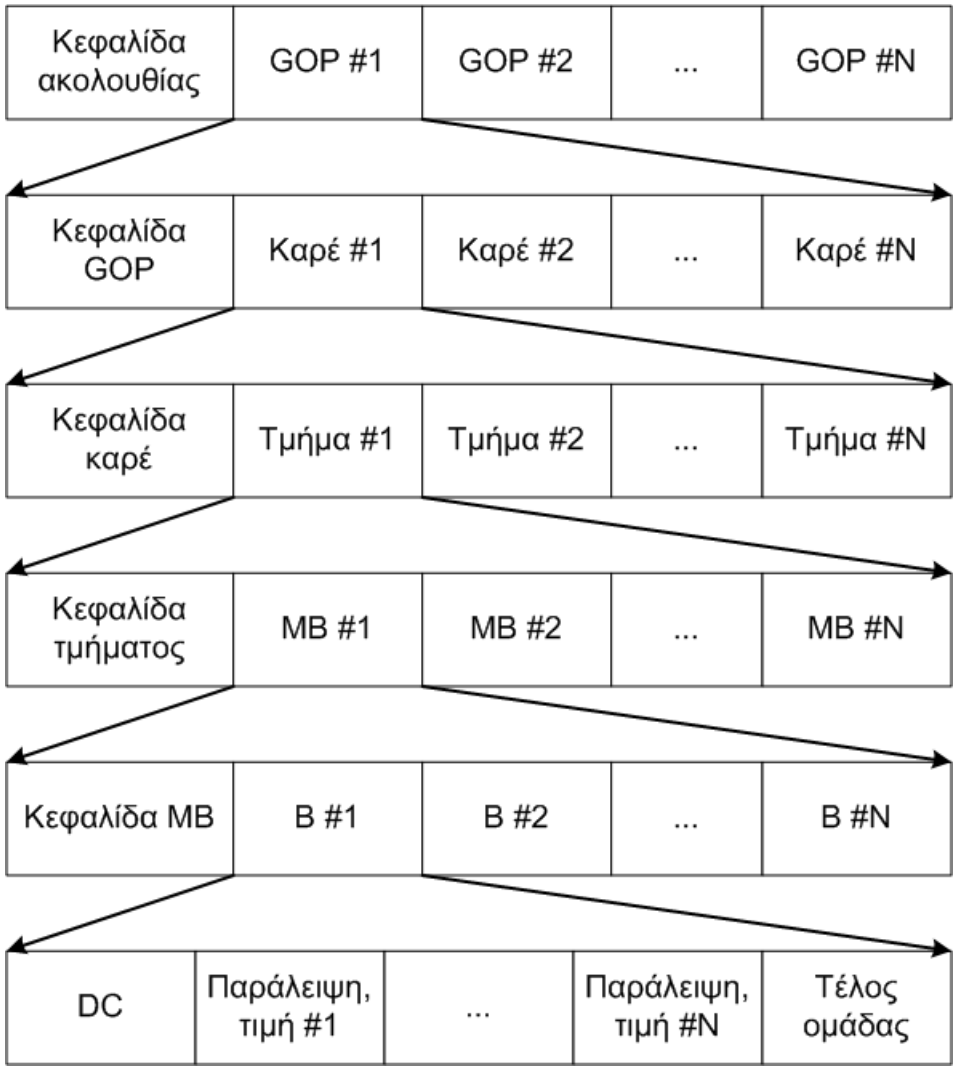
- Common frame sequences
  - PAL
    - Display “IBBPBBPBBIBBPBBPBB…”
    - (De)coding “IPBBPBBIBBPBBPBB…”
  - NTSC
    - Display “IBBPBBPBBPBBIBBPBBPBBPBB …”
    - (De) coding “IPBBPBBPBBIBBPBBPBBPBB …”
  - Transmission in (de)coding order

Data streams

Class: Multimedia Technology, **Section # 13:** Video coding: MPEG-1/2

Instructor: George Xylomenos, **Department:** Informatics

Stream organization



Data layers (1 of 2)

- Video bitstream consists of 6 layers
- Video sequence layer: buffering!
  - Corresponds to a video clip
  - States bitrate and buffering requirements
  - Image resolution, pixel aspect ratio
- GOP layer: random access
  - At least one I-frame from which to begin decoding
  - Timestamp of first frame
- H.261 does not have such layers

Data layers (2 of 2)

- Image layer: frame
  - Sequence numbers to identify frames
- Slice layer: slices for each frame
  - Special sync code in the beginning
    - After an error, we skip to the next sync code
- Macroblock level: MB details
- Block level: block details
- These are similar to H.261

Multiplexing (1 of 2)

- Multiplexing of audio and video
  - Each medium's bit stream divided into packets
    - Typical size: 2048 byte (CD-ROM sector)
    - Other sizes possible
  - Timestamps for synchronization (90 kHz)
  - Packetized elementary stream (PES)
    - Derived from encoder's Elementary Streams (ES)
    - An MP3 bit stream is the audio ES

Multiplexing (2 of 2)

- Multiplexing of audio and video
  - The PES's are multiplexed into packs
  - Each pack has a timestamp
    - Allows decoder to synchronize
  - Heading in the first packet
    - Maximum bitrate, PES types included
  - Hard to access in a network setting
    - Pack must be received from its start (for the heading)

MPEG-1 vs H.261

- H.261 is much simpler
  - Slightly older (1988 vs 1990)
  - But: created for a single app (conferencing)
    - No need for random access
    - Need for worldwide interoperability
- MPEG-1 can be used for different apps
  - Transmission or distribution of video
  - Many compatibility options

MPEG-2

Class: Multimedia Technology, **Section # 13:** Video coding: MPEG-1/2

Instructor: George Xylomenos, **Department:** Informatics

MPEG-2 goals

- Goals
  - High resolution for HDTV
  - Used in DVD video
  - Bitrates from 4 Mbps to 100 Mbps
  - Appropriate for storage and transmission
- Why there is no MPEG-3?
  - Initially, MPEG-3 was intended for HDTV
  - But HDTV was added into MPEG-2

Differences with MPEG-1

- Supports interlaced scanning
  - MPEG-1 was created for computers only
- 4:2:0 or 4:2:2 subsampling
  - 4:2:2 used inside the studio
- Frame rates up to 60 Hz
- No D-frames (redundant)
- New coding tools

Profiles and levels (1 of 3)

- Profiles vs levels
  - Profile: supports a class of apps
    - Describes a coding algorithm
  - Level: provides parameters
    - Subsampling, resolution, bit rate
- Levels
  - Low: analog TV quality
  - Main: DVD, digital video
  - High: HDTV in two variants

Profiles and levels (2 of 3)

- Simple profile: similar to MPEG-1
- Main profile: DVD, digital TV
- Scalable profiles
  - Low resolution mobile TV
  - Fixed HDTV
- High profile
  - Both SNR and spatial scalability
  - 4:2:2 subsampling

Profiles and levels (3 of 3)

Profile	Simple	Main	SNR	Spatial	High
B-frames	N	Y	Y	Y	Y
Subsampling	4:1:1	4:1:1	4:1:1	4:1:1	4:1:1,4:2:2
Scalability	N	N	SNR	Spatial	Both
Level					
High (1920x1152)		<80 Mbps			<100 Mbps
High-1440 (1440x1152)		<60 Mbps		<60 Mbps	<80 Mbps
Main (720x576)	<15 Mbps	<15 Mbps	<15 Mbps		<20 Mbps
Low (352x288)		<4 Mbps	<4 Mbps		

- Predefined combinations

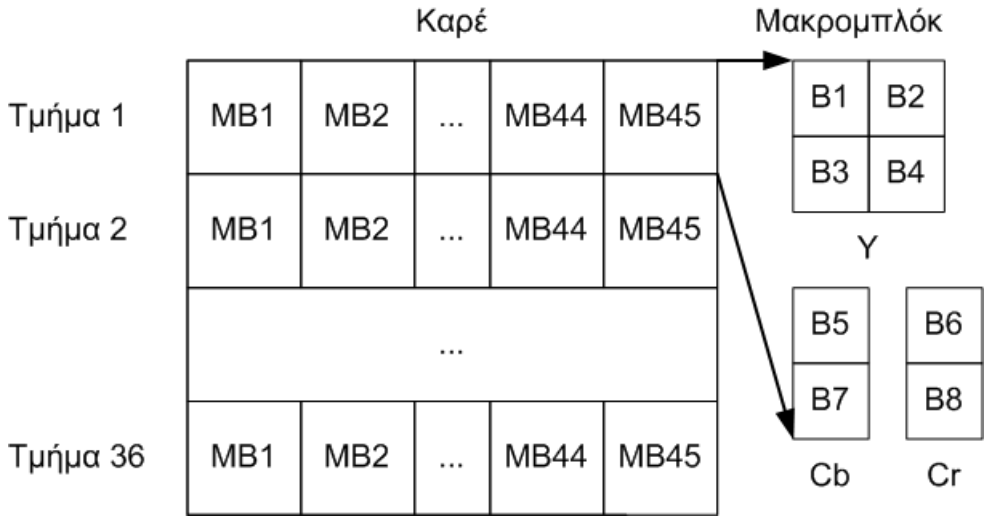
Interleaved modes (1 of 2)

- Interleaved scanning
  - Every frame is divided into two fields
  - Odd field: odd lines of the frame (1,3,5,…)
  - Even field: even lines of the frame (2,4,6,…)
- Forming macroblocks (I-frames)
  - Field mode: different MBs per field
  - Frame mode: MBs per frame

Interleaved modes (2 of 2)

- Macroblock coding (P- and B-frames)
  - What is used for prediction?
  - Frame mode
    - Next/previous frames
    - Next/previous fields
      - We first break the frame into fields
  - Field mode
    - Next/previous fields

MPEG-2 image structure



- Example: 720x576 resolution (DVD PAL)
  - 36 slices per frame
  - 45 MBs per slice
  - 6 blocks per MB in (4:1:1) or 8 blocks in (4:2:2)

Scalable coding

Class: Multimedia Technology, **Section # 13:** Video coding: MPEG-1/2

Instructor: George Xylomenos, **Department:** Informatics

Scalability (1 of 4)

- Why scalable coding?
  - Support heterogeneous channels
  - Support variable bit rate channels
    - Always send base layer
    - Add enhancement layers when possible
  - Support noisy channels
    - Send base layer with error protection
    - Add enhancement layers without error protection

Scalability (2 of 4)

- Spatial scalability
  - Spatially scalable profile
    - Similar to hierarchical JPEG
  - Pyramid encoding of frame
    - Usually, based on subsampling by powers of 2
  - Encode subsampled base layer
  - Magnify base layer to next resolution
  - Encode difference as next layer

Scalability (3 of 4)

- SNR scalability
  - SNR scalable profile
    - Similar to progressive JPEG
  - Variable quantization of DCT coefficients
    - First use a large quantizer
    - Reverse the quantization
    - Calculate different from original
    - Quantize difference with smaller quantizer

Scalability (4 of 4)

- Time scalability
  - Reduce frame rate at each layer
  - Example: 15 fps + 15 fps
  - The base layer is independently encoded
    - Motion compensation based only on its frames
  - The next layer has two options
    - Encode only referencing the base layer
    - Encode referencing the base and enhancement layers

Program and Transport streams

Class: Multimedia Technology, **Section # 13:** Video coding: MPEG-1/2

Instructor: George Xylomenos, **Department:** Informatics

Data stream (1 of 3)

- Stream multiplexing in MPEG-2
  - We start with the PES of MPEG-1
- Program Stream (PS)
  - Suitable for storage (e.g., DVD)
    - Error-free environment (DVD has tons of error correction!)
  - Similar to MPEG-1 streams
    - One program (video) per stream
  - Large variable length packets
  - Common timing base for all media

Data stream (2 of 3)

- Transport stream (TS)
  - PES plus independent timing bases
  - Allows multiple programs per stream
    - Table assigns PES to programs
    - Essentially, many channels in a stream
  - Suitable for transmission (e.g., HDTV)
    - Error-susceptible environments
    - Fits multiple channels in 6 or 8 MHz

Data stream (3 of 3)

- Transport Stream (continued)
  - Fixed length packets (184+4 bytes)
    - Originally for ATM networks (48 bytes/cell)
  - Simpler error correction
  - Low latency multiplexing
- End-to-end latency < 1 sec
  - Too high for interactive applications
  - Good enough for channel switching

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

End of Section # 13

Class: Multimedia Technology, **Section # 13:** Video coding: MPEG-1/2

Instructor: George Xylomenos, **Department:** Informatics