ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Multimedia Technology

Section # 12: Video coding: H.261/3

Instructor: George Xylomenos

Department: Informatics

Contents

- Video coding basics
- H.261
- H.263
- H.263+/++

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

Video coding basics

Class: Multimedia Technology, **Section # 12:** Video coding: H.261/3

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

Redundancy in video (1 of 2)

- Spatial redundancy
  - Similarity between neighboring pixels
  - Examination of a single frame
  - Intraframe coding
- Temporal redundancy
  - Similarity between consecutive frames
    - Comparison of frames
    - Identification of similar areas
  - Interframe coding

Redundancy in video (2 of 2)

- Spatial compression
  - Low delay but not much compression
  - Similar with still image coding
    - Mostly based on JPEG
- Spatial and temporal compression
  - Higher delay due to search of entire frame
    - JPEG + prediction
    - Many more compression opportunities
  - Bit: errors may affect multiple frames!

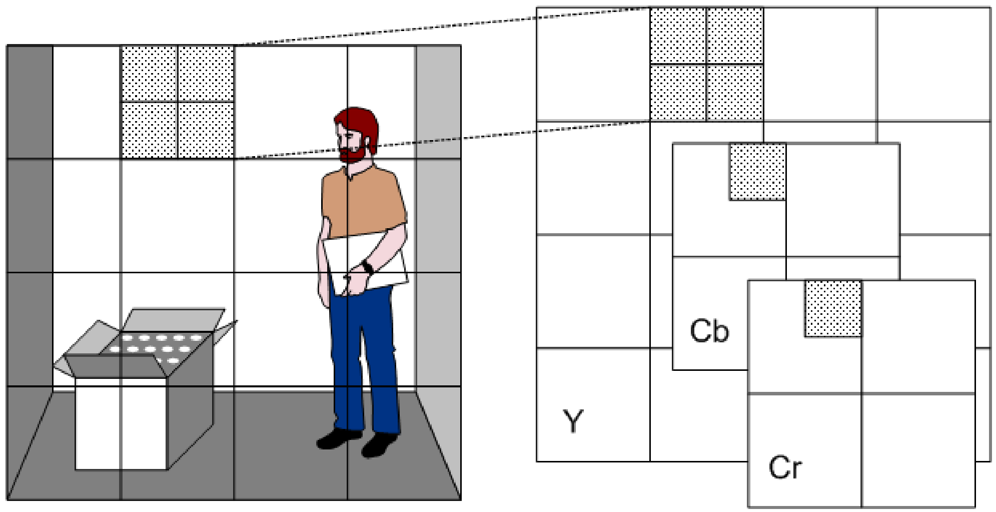
Intraframe coding

- When do we need intraframe?
- Required for first video frame
  - There is nothing to predict from!
- Makes sense when we have scene changes
  - Prediction does not offer much
- Periodic throughout the video
  - Random access
  - Recovery from errors

Interframe coding

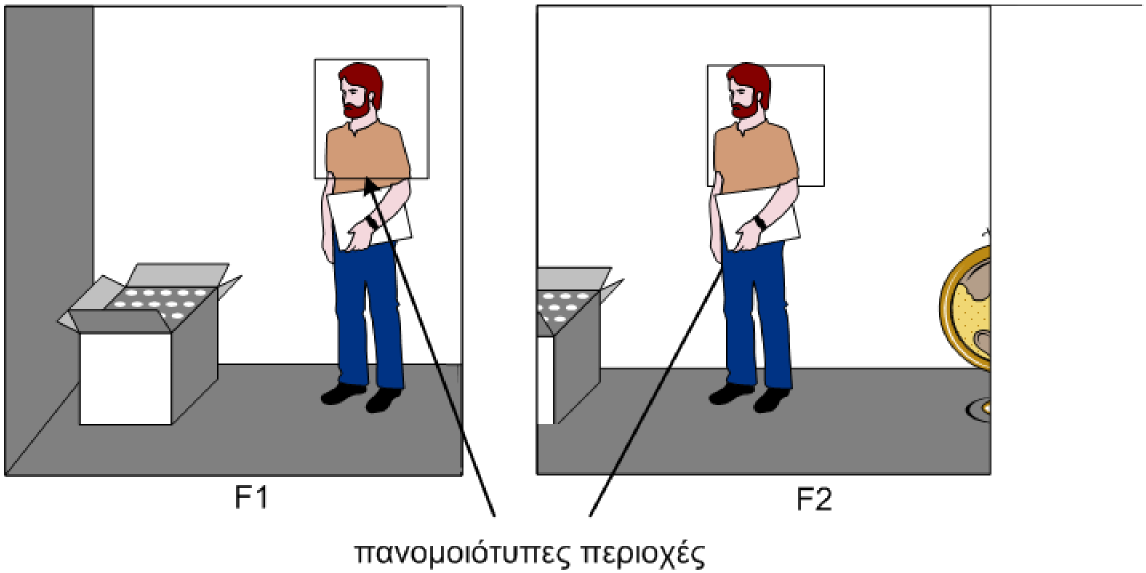
- Compare current frame with reference
  - Identification of similar areas
  - Computationally complex (too many options!)
  - Matches may not be exact
- Motion prediction
  - Motion vector: how much the area moved
- Motion compensation
  - How the area was modified

Blocks and macroblocks



- Video coded in blocks of 8x8 pixels
  - Similar flow to JPEG
- Macroblocks are n blocks (typically 6)
  - Blocks from all components in an area
  - Same as an MCU in JPEG

Frame types (1 of 7)

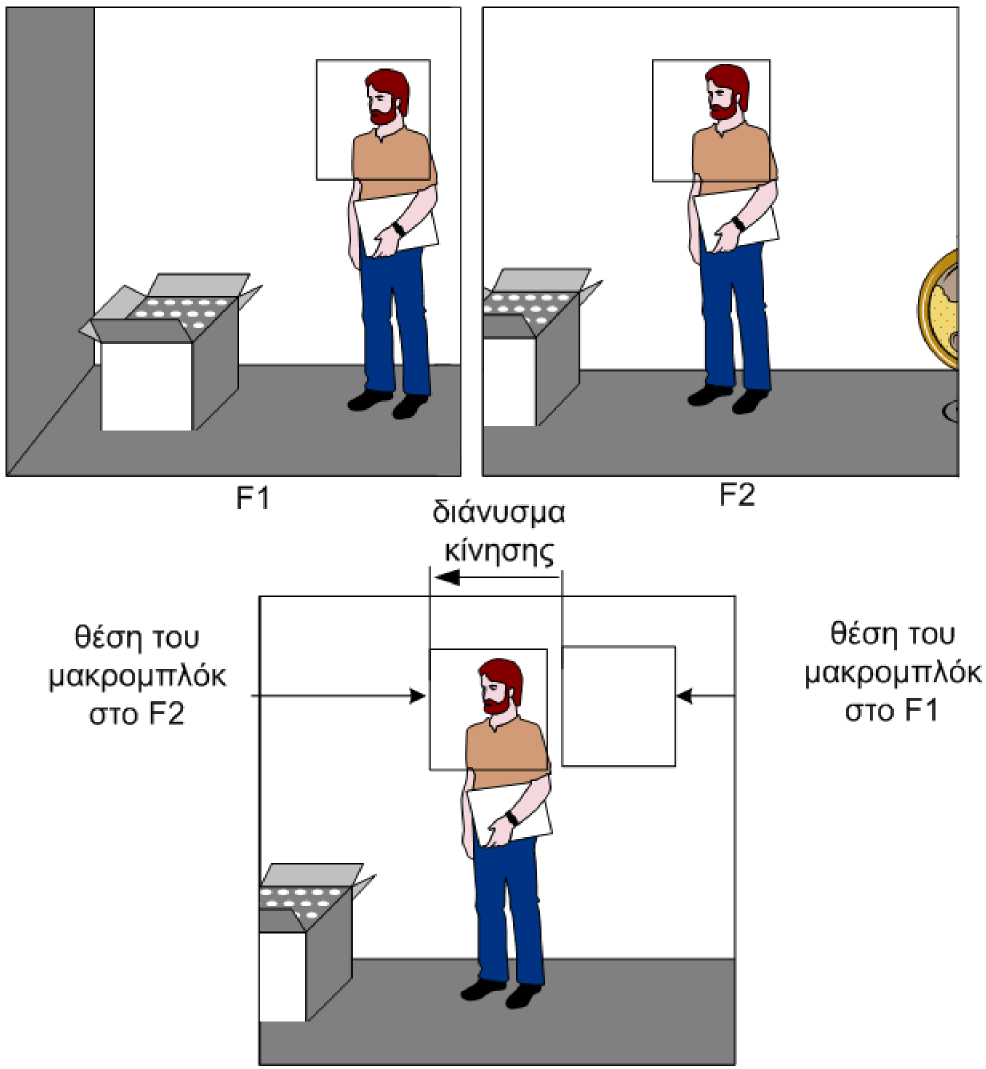


- I-frames: intra-coded frames
  - Coded similar to JPEG
  - Used as <u>reference</u> frames
  - Example: frame F1

Frame types (2 of 7)

- P-frames: intra-coded frames
  - Coded with reference to another frame
  - We look for similar macroblocks
    - All blocks of the macroblock should be similar
  - Motion vector: how the macroblock moved
    - Compared to the similar area in the reference frame
  - Differences between current and reference blocks
    - Differences encoded similar to I-frames

Frame types (3 of 7)



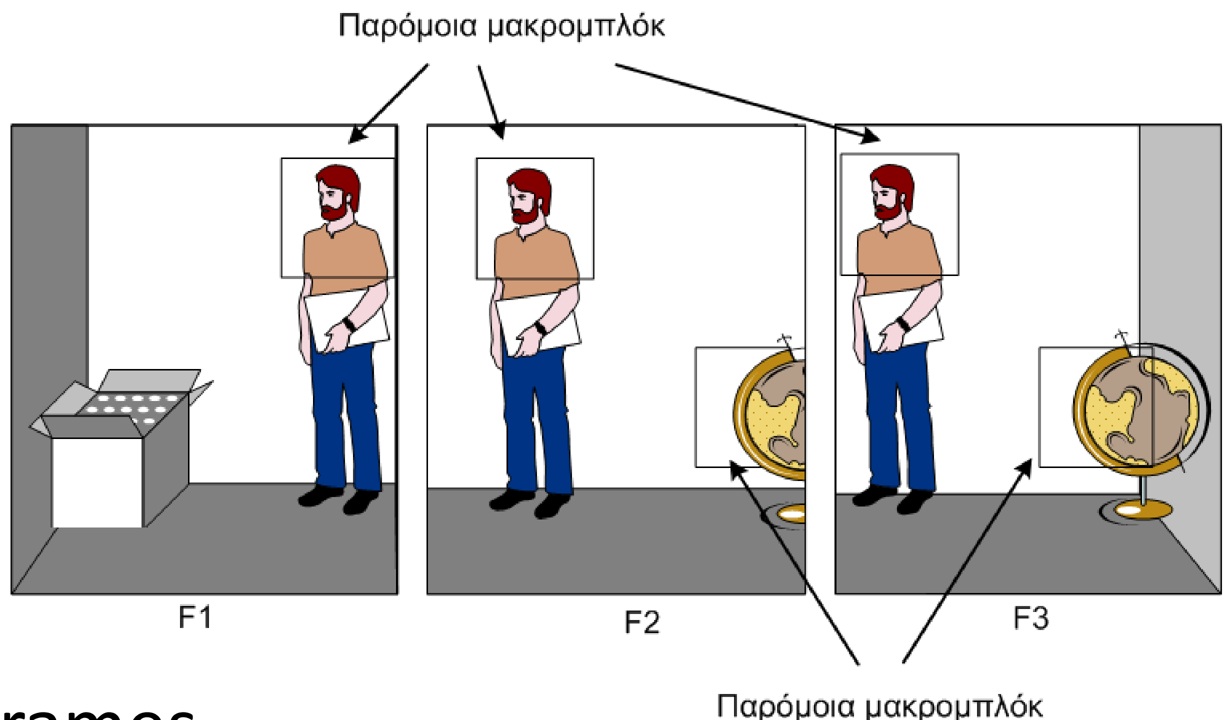
Frame types (4 of 7)

- How are similar areas located?
  - Standards do not say how!
    - They may limit the search area
    - It cannot be larger than the motion vector
  - Motion vectors are expressed in pixels
    - Sometimes, fractions of pixels (see later)
  - How many blocks are up to one pixel away?
    - 9 (-1,0,+1 pixel in each dimension)
    - With more pixels, the number goes up fast

Frame types (5 of 7)

- Many options exist
  - Tradeoffs between performance and efficiency
  - Various tricks to avoid repeating calculations
- A very simple policy
  - Just use the luminance samples
    - Subtract the values of current and reference
    - Sum up the absolute differences
    - Choose the reference with the smallest sum

Frame types (6 of 7)



- B-frames
  - Two reference frames (one earlier, one later)
  - One or two motion vectors
    - If two, interpolate two reference blocks

Frame types (7 of 7)

- Why do we even have B-frames?
  - Say that the sequence is F1, F2, F3
  - Say that F2 has similarities with F3
    - Which do not exist with F1
  - Then, why not predict F3 from F2?
  - The reason is the periodic I-frames
    - We need them for random access
    - Therefore, we can use them as references

H.261

Class: Multimedia Technology, **Section # 12:** Video coding: H.261/3

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

ISDN video conferencing (1 of 3)

- A long time ago (80s) ISDN was designed
  - As successor of the analog PSTN
  - ISDN lines consist of 64 Kbps B channels
    - There are also D channels for signalling
- ISDN Primary Rate (PRI): 30 B channels
  - Trunk line for companies (AUEB had some)
- ISDN Basic Rate (BRI): 2 B channels
  - Residential line (I had one at home)

ISDN video conferencing (2 of 3)

- H.261: video coder for either PRI and BRI
  - Targets p x 64 Kbps lines, where p=1,2,…,30
  - Standardized in 1988
  - Conferencing also needs an audio encoder
- Coding requirements
  - Real-time coding (not offline)
  - Real-time decoding
  - Maximum latency < 150 ms
    - Conferencing is an interactive application

ISDN video conferencing (3 of 3)

- Targets a fixed bit rate
  - ISDN is digital, but circuit switched
- Variable transmission quality
  - Based on variable quantizer
  - Changes quantization step to achieve target
- “Economical” terminals
  - Videotelephones (not computers)

Picture format (1 of 3)

- Has to work worldwide
  - Unlike TV which is inherently local
  - You must be able to call other countries
  - Exactly as in telephony
- Compromise between PAL and NTSC
  - Frame rate from NTSC
  - Resolution from PAL

Picture format (2 of 3)

- Input frame rate: 29.97 fps (NTSC)
- Output frame rate: 29.97, 15 or 7,5 fps
- Progressive scanning
- Luminance (Y), chrominance (Cb and Cr)
  - 4:3 aspect ration
  - 4:1:1 subsampling
  - 8 bits per pixel per component

Picture format (3 of 3)

- Common interchange format (CIF): optional
  - Luminance: 352 x 288
  - Chrominance: 176 x 144
  - Based on PAL/SECAM
- Quarter CIF (QCIF): mandatory
  - Luminance: 176 x 144
  - Chrominance: 88 x 72
  - Cheap terminals and lines could not do CIF

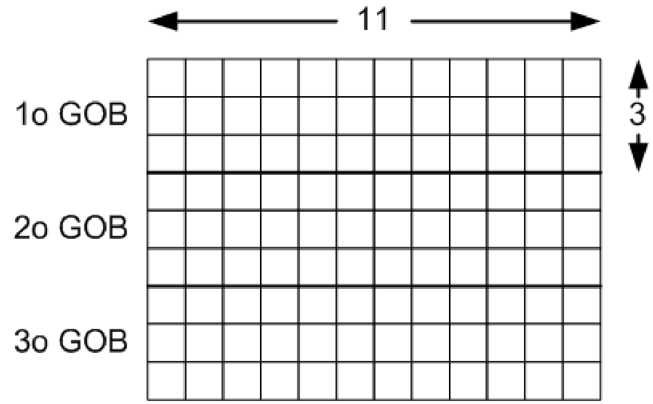
Compression ratio

- Bit rate before compression
  - QCIF at 15 fps: 4.55 Mbps
  - CIF at 30 fps: 36.45 Mbps
- Target bitrate for transmission
  - 1:70 for QCIF at 15 fps in 1 B channel (64 Kbps)
    - And we also need sound!
  - 1:95 for CIF at 30 fps in 6 B channels (384 Kbps)
    - Suitable only for PRI links

Frame structure (1 of 2)

- Block
  - 8x8 pixels in a single component
- Macroblock (MB)
  - 4 Y blocks, 1 block each for Cb and Cr
  - 16x16 pixel area in the luminance compoment
- Group of Blocks (GOB)
  - 3x11=33 MBs
  - Starts with a fixed bit sequence
    - Allows synchronization after bit errors

Frame structure (2 of 2)



- QCIF: three GOBs
  – 3 “rows” x 1 “column”
- CIF: 23 GOBs
  – 6 “rows” x 2 “columns”
- After an error, you lose the rest of the GOB

Intraframe coding

- We start with a DCT
  - For each 8x8 pixel block
- Separate quantization of DC/AC coefficients
  - But with only <u>two</u> quantizers
    - Fixed DC quantizer (8)
    - Variable AC quantizer (2-62)
    - Simpler than the quantizer matrix of JPEG
- Ends with entropy coding (JPEG-like)

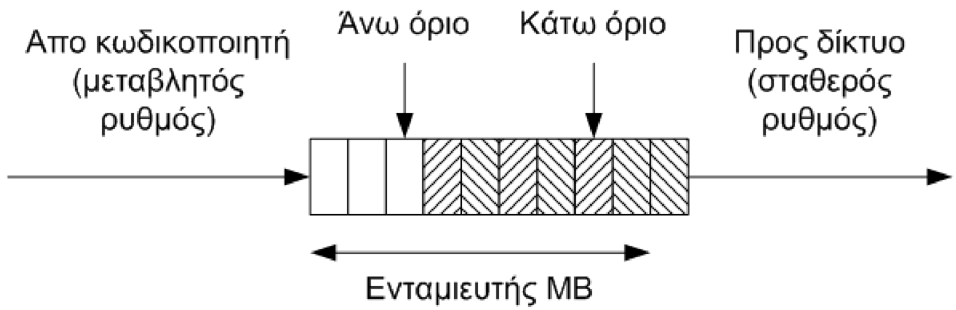
Interframe coding

- Used whenever the encoder choose so
  - No periodic I-frames
  - There is no random access!
- Prediction of each MB
  - Comparison with areas from previous frame
    - And <u>only</u> the previous frame
  - Reference may be I-frame or P-frame
- If not enough similarity?
  - Just use intraframe coding

Motion coding

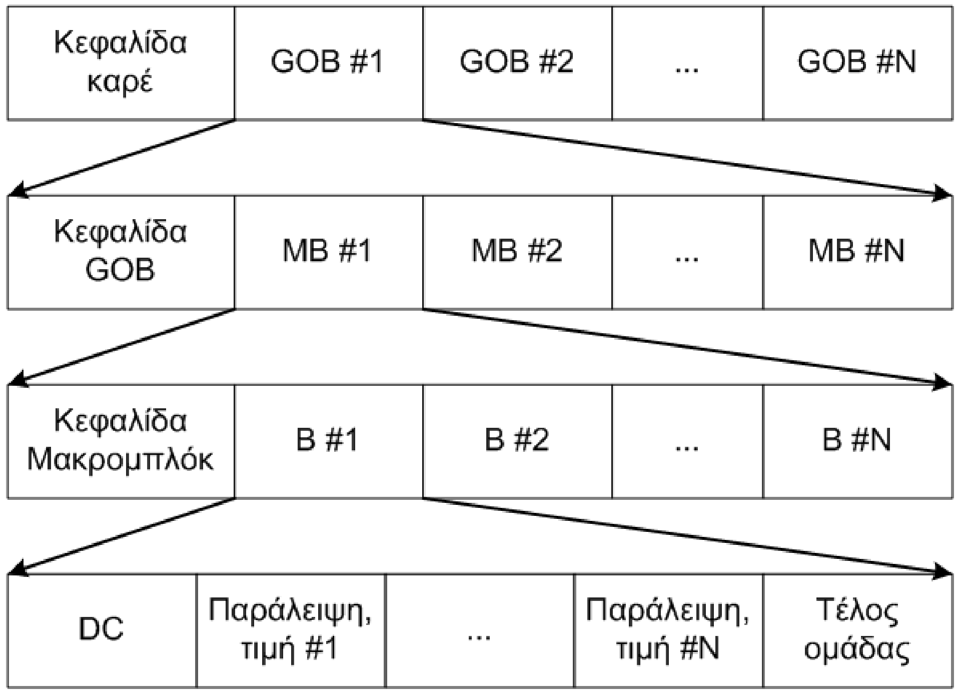
- Motion vectors
  - Displacement from reference area
  - Vectors are entropy coded
- Differences between blocks
  - Starts with differences (DPCM)
  - Then, same as intraframe
  - Very small differences are skipped
    - Higher compression at a cost to quality

Achieving a fixed bit rate



- The quantizer step can be adjusted
  - It applies only to AC coefficients
- Buffer fills -> increase quantizer
- Buffer empties -> decrease quantizer
  - High and low watermark for step change
- Result: fixed bit rate, variable quality

Bit stream (1 of 2)



- Generalization of JPEG
  - The bit stream is divided into frames
  - The frames are divided into GOBs

Bit stream (2 of 2)

- Frame header
  - Sync code, timestamp
- GOB header
  - Sync code, GOB number
  - Quantizer step
- MB header
  - Position in GOB, type (I or P)
  - Quantizer step (for MB only)

H.263

Class: Multimedia Technology, **Section # 12:** Video coding: H.261/3
Instructor: George Xylomenos, **Department:** Informatics

New features (1 of 4)

- H.263: extension of H.261/262 (=MPEG-2)
  - Standardized in 1996
  - Appropriate even for PSTN
    - Supports even lower bitrates
    - Offers improved error resilience
- Additional resolutions
  - SQCIF: 128x96, 15 or 7.5 fps
  - 4CIF and 16CIF for higher resolutions
  - Also supports CIF/QCIF to interoperate with H.261

New features (2 of 4)

- Half pixel accuracy
  - We take two reference areas
  - And interpolate their corresponding values
- Motion vector prediction
  - We use the median of three motion vectors
    - Left, above, above right
  - And transmit the difference from the prediction
  - Saves space when multiple MBs can be predicted

New features (3 of 4)

- Prediction outside the frame edge
  - Reference is an area partially outside the edge
  - We assume a fixed value for the pixels outside
  - Useful in smaller resolutions
- Arithmetic coding of coefficients
  - Rather than Huffman
    - With better compression rates
  - Different statistics for each hierarchy level

New features (4 of 4)

- Bidirectional prediction
  - Similar to B-frames in MPEG-2
  - PB-frames: coding P and B frame together
  - P-frame predicted from previous P or I
  - B-frame predicted from previous and current
  - Better quality for the same bitrate
    - But also requires more memory
    - In 1996 though, we have computers!

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

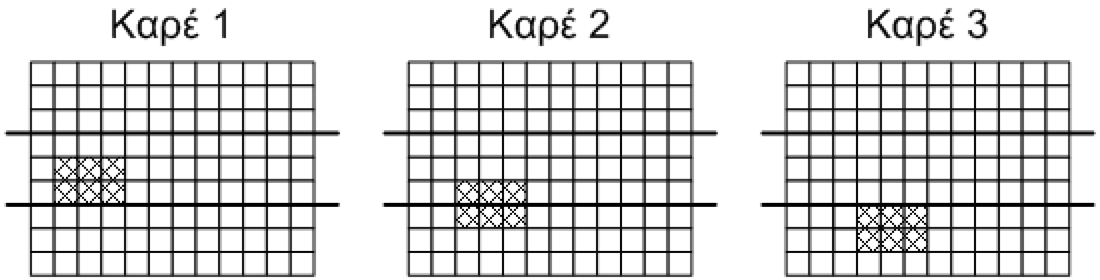
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

H.263+/++

Class: Multimedia Technology, **Section # 12:** Video coding: H.261/3

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

Error recovery (1 of 5)



- Bit errors lead to losing a GOB
  - Decoding fails until next sync code
  - Which appears at the beginning of each GOB
  - These failures also affect future frames
    - Say that we lost the blocks of an entire GOB
    - If these are used as references, the loss persists
  - Bigger problem with PSTN (compared to ISDN)

Error recovery (2 of 5)

- Dealing with errors in H.261
  - Periodically send I-frames
    - I-frames do not rely on previous ones
    - But we do not know when to send them
    - Bitrate is wasted due to no feedback
  - H.263+/++ offer better choices
    - Recovery without full I-frames
  - Standardized in 1988 (+) and 2000 (++)

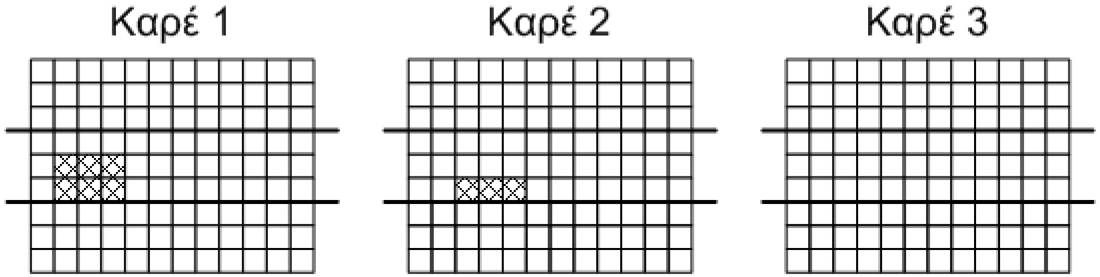
Error recovery (3 of 5)

- Base technique: error tracking
  - The encoder monitors dependencies
    - Which block was predicted from which area
  - The decoder sends feedback on lost GOBs
  - The encoder checks which MBs are affected
    - Their blocks can be intraframe coded
    - We stop error propagation
    - We only use intra when we have to

Error recovery (4 of 5)

- Reference frame selection
  - Reference to older frames
    - Not just the previous one
    - Allows recovery from lost GOBs
    - Need to track dependencies
    - Need to buffer more previous frames
  - Decoder can send two types of feedback
    - Acknowledge correct reception
    - Notify of erroneous reception

Error recovery (5 of 5)



- Independent GOB decoding
  - Each GOB treated separately as a subframe
  - We can only refer to the same previous GOB
  - We lose some compression opportunities
  - But we have simpler dependency tracking
  - And errors do not propagate outside the GOB

End of Section # 12

Class: Multimedia Technology, **Section # 12:** Video coding: H.261/3

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