# EBU5303

# Multimedia Fundamentals

## MPEG

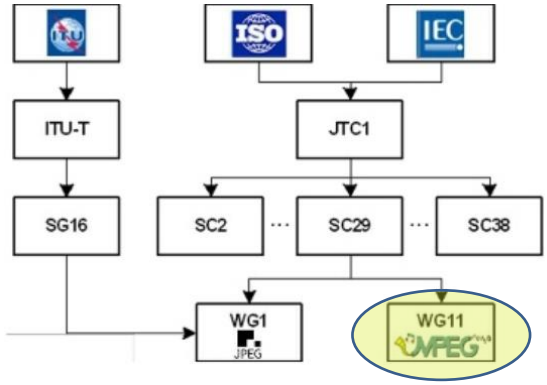


## Agenda

- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..

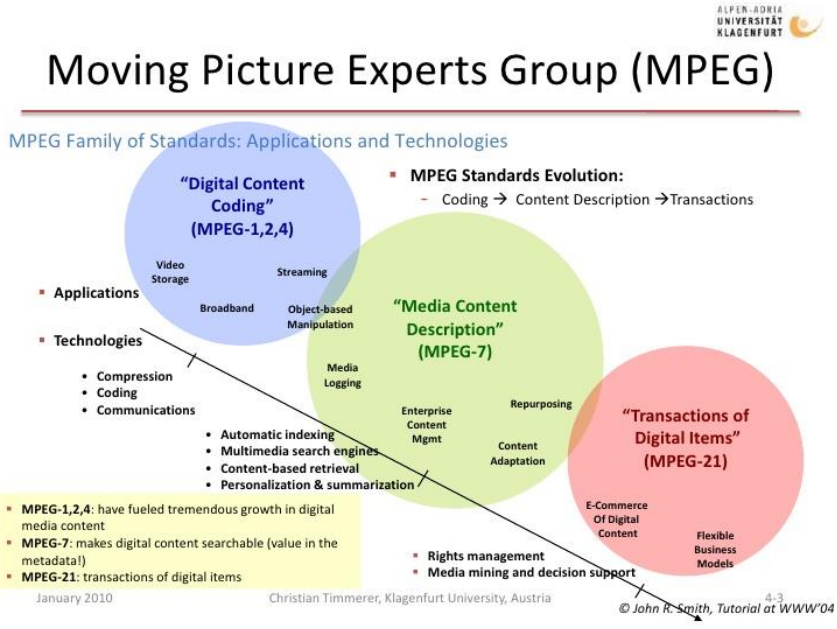
# What is MPEG?

- MPEG = Moving Pictures Experts Group
- Standardised set of algorithms so that the various implementations all produce files in the same format.




EBU5303

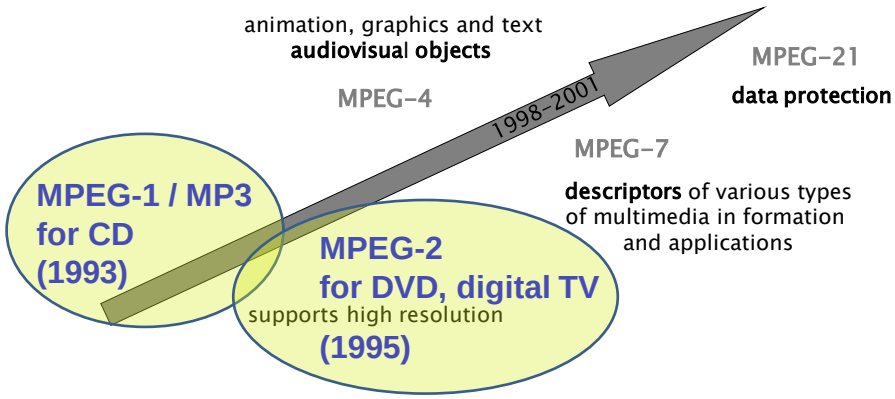
3




EBU5305

4

# MPEG family of standards: MPEG-1, MPEG-2, MPEG-4, MPEG-7 and MPEG-21




5

5

# MPEG-1

- The first MPEG compression standard for audio and video.
- Designed to allow moving pictures and sound to be encoded into the bitrate of a Compact Disc (CD).
- Commonly limited to about 1.5 Mbps (Megabits per second).
- To meet the low bit requirement, MPEG-1 downsamples the images, and uses frame rates of only 24–30 Hz.
- Only supports progressive pictures
- Includes the popular MPEG-1 Audio Layer III (MP3) audio compression format.

EBU5305

6

# MPEG-2

- Transport, video and audio standards for broadcast-quality television.
- Supports interlacing and high definition.
- Chosen as the compression scheme for over-the-air digital television (e.g. DVB), digital satellite TV services (e.g. Dish Network), digital cable television signals, SVCD and DVD Video.



# Agenda

- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..

# Video compression

- Unlike image encoding, video encoding is rarely done in lossless form.
- No storage medium has enough capacity to store a practical sized lossless video file
  - Lossless DVD video - 221 Mbps
  - Compressed DVD video - 4 Mbps
  - 50:1 compression ratio!

9

# Video compression

- Video: a sequence of pictures (or frames)
- JPEG algorithm used for intra-frame coding
  - e.g. moving JPEG (MJPEG) exploits only intra-frame coding (*spatial redundancy*)
- High correlation between successive frames (temporal redundancy)
  - motion estimation and motion compensation: inter-frame coding
  - use a combination of actual frame contents and predicted frame contents

10

# Inter-frame and intra-frame coding

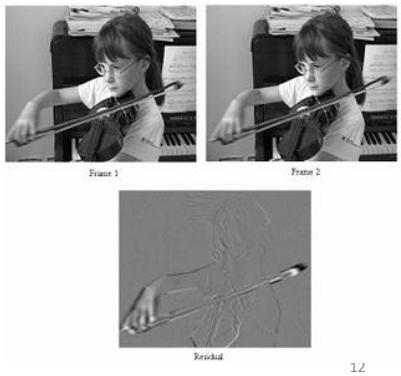
- **Inter-frame** compression uses one or more earlier or later frames in a sequence to compress the current frame.
- **Intra-frame** compression uses only the current frame, which is effectively image compression.
- MPEG 1 & 2 use both inter-frame and intra-frame coding


11

11

# Temporal redundancy

- Temporal redundancy occurs when, in a sequence of images, consecutive images contain similarities.
  - For example, in a sequence of a tennis match, the court does not change; only the players and the ball make significant movements.
- With temporal compression, only the changes from one frame to the next are encoded.




12

12

# Agenda

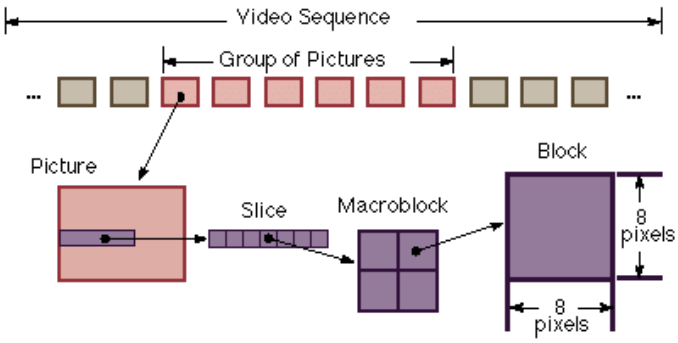
- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..



# Two main types of compression

- Spatial compression (intra-frame coding)
  - Allow fast random access
  - Applied to key frames (e.g. every 12th frame, start of new scene, etc)
    - **I-frames**
- Temporal compression (inter-frame coding)
  - Allow high compression
  - Exploit temporal redundancy between frames
  - Applied to frames occurring between key frames
  - Based on motion detection
    - **P- and B- frames**

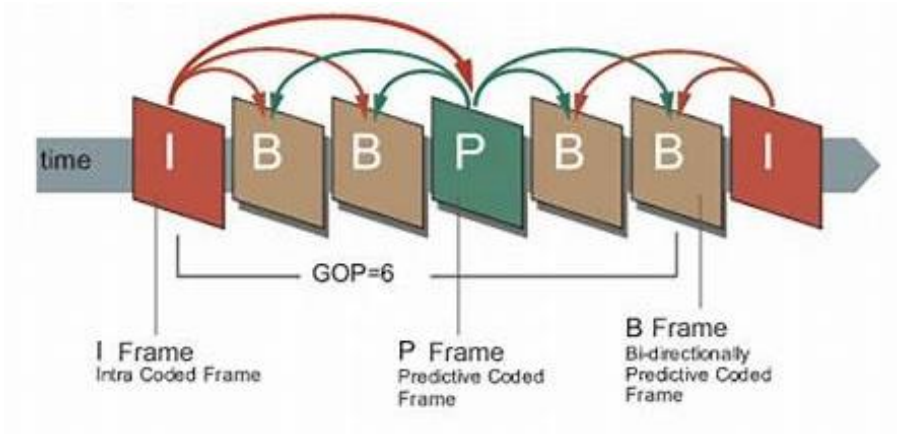
# Structure of Digital Videos






15

# Group of Pictures (GOP)






16

# MPEG video encoding

- input frames are preprocessed
  - color space conversion
  - spatial resolution adjustment (chroma sub-sampling)
- frame types (Intra or Inter) are decided for each frame/picture
- each picture is divided into macroblocks of 16 X 16 pixels
- macroblocks
  - are intracoded for I frames
  - are predictive coded or intracoded for P and B frames
  - are divided into six blocks of 8 X 8 pixels
    - 4 luminance and 2 chrominance
    - DCT is applied to each block → transform coefficients
      - » quantized
      - » zig-zag scanned
      - » variable-length coded

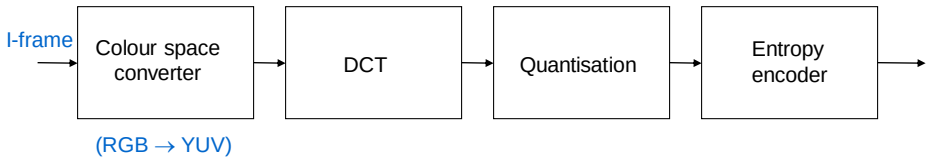


# Agenda

- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..

# I-frames (Intra-coded frames)

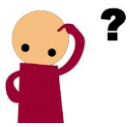
- Self-contained (without reference to other frame)
- Treated as a still image and make use of JPEG
- Compression rate is the lowest
- Points for random access in video stream




19

19

# Question



Which of the following statements describes best intra-frame coding? Choose one answer.

- It uses one earlier frame in a sequence to compress the current frame.
- It uses only the current frame.
- It uses both earlier and later frames in a sequence to compress the current frame.
- It uses one later frame in a sequence to compress the current frame.

20

20

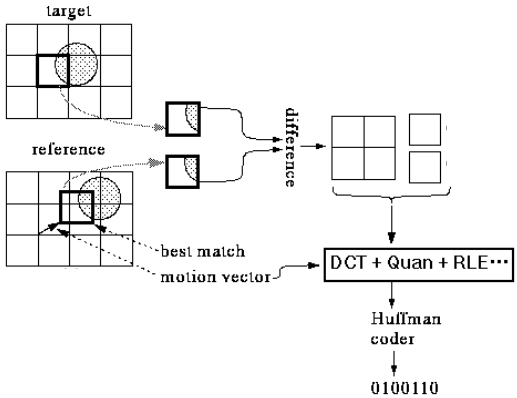
# Agenda

- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..



# P-frames (Predictive-coded frames)

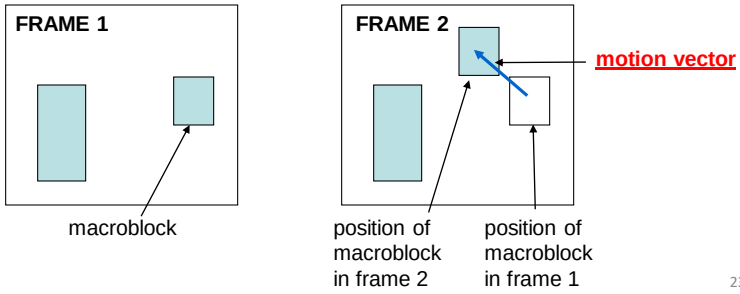
- Use a reference frame for motion estimation
- Hold only the changes in the image from a previous frame
- Compression rate is higher



# Motion estimation

Based on looking for matching macroblocks of data in successive frames, and measuring the difference between matching macroblocks in terms of:

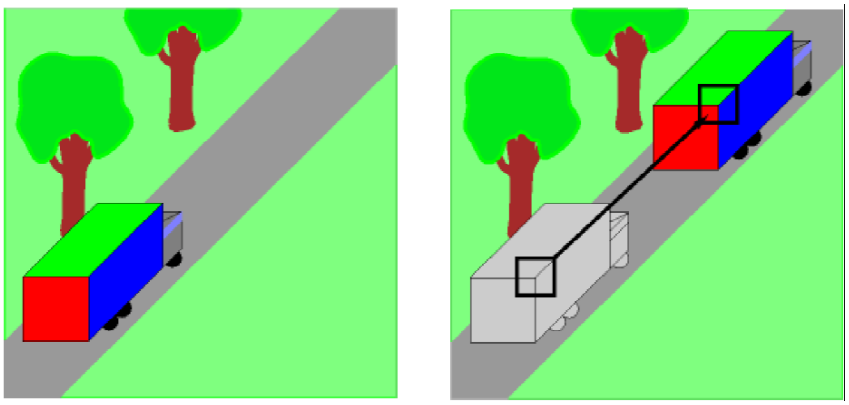
- Position within the frame (expressed as motion vectors)
- Luminance and colour difference (error term)




23

23

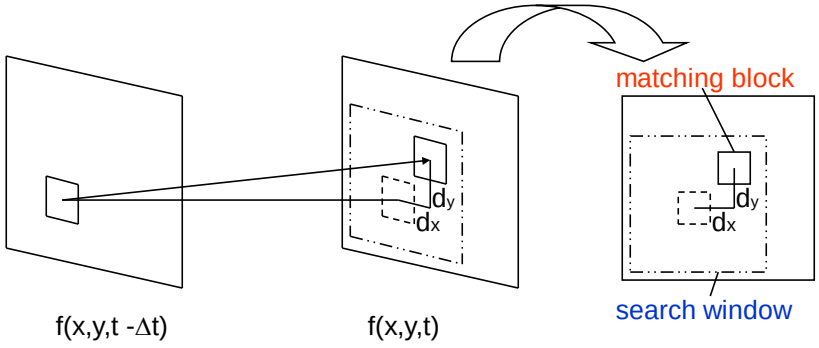
# Motion estimation of macroblocks



24

24

# Motion estimation




$$\min \sum_{B} \left\| f(x,y,t) - f(x-d_x, y-d_y, t-\Delta t) \right\|$$

25

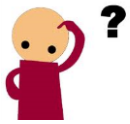
# Motion estimation

- We can't search all possible macro blocks, so we need a smart algorithm
- Matching macroblocks: only if a "close" match can be found
  - Evaluate "closeness" with MSE or other metric
  - If no suitable match found, just encode the macroblock as an I-block
- We don't want too many P or B frames in a row
  - Predictive error will keep propagating until next I frame
  - Delay in decoding



26

# Question



What can you do to optimise motion estimation?





# Motion estimation

When comparing macroblocks, three things could happen:

1. One identical macroblock is found: no need to encode anything
2. No matching macroblock is found: use intra-frame encoding
3. One similar macroblock is found: encode difference values between the two macroblocks and a motion vector

# P-frames

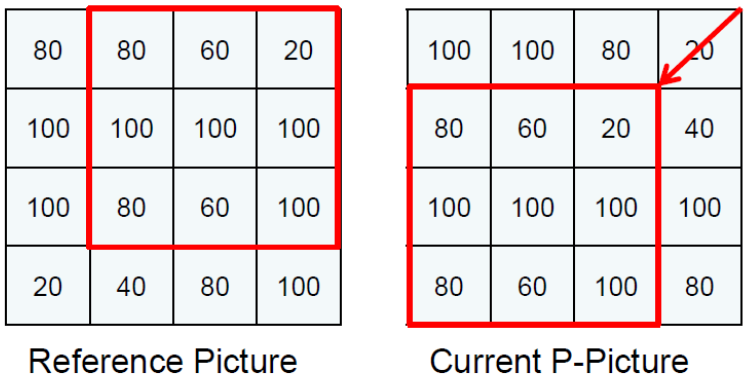
Each macroblock in a P-frame can be encoded either as an I-macroblock or as a P-macroblock.

- An *I-macroblock* is encoded in the same way as a macroblock in an I-frame (i.e. similar to JPEG)
- A *P-macroblock* is encoded as an area of the past reference frame, and an error term. To specify where the macroblock came from in the reference frame, a motion vector is used.
  - If the motion vector is zero, then the macroblock has not moved between frames
  - If the difference values are also zero, then the area does not need to be coded



29

# Motion vector






30

# Error term

| | | |
|---|---|---|
| 100 | 100 | 100 |
| 100 | 80 | 60 |
| 20 | 40 | 80 |

| | | |
|---|---|---|
| 100 | 100 | 100 |
| 80 | 80 | 60 |
| 20 | 20 | 80 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 20 | 0 | 0 |
| 0 | 20 | 0 |

31

# Agenda

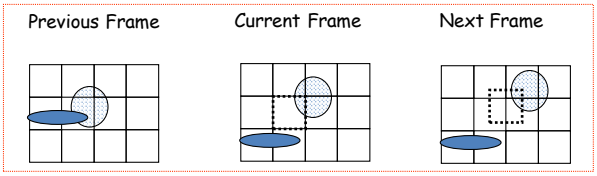
- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..



32

# B-frames (Bi-directionally predictive-encoded frames)

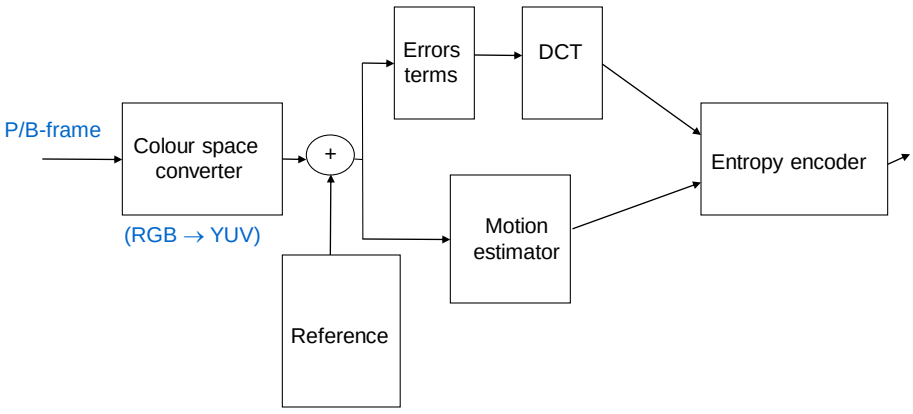
- Prediction limit
  - Some macroblocks may need information that is not present in the *previous reference frame*




  - Such information might be available in a *subsequent frame*
- MPEG uses a third frame type (B-frame)
  - to form a B-frame: search for matching MBs in both past and future frames
  - typical pattern is IBBPBBPBB IBBPBBPBB IBBPBBPBB
  - actual pattern is up to encoder, and need not be regular

33

# P- and B- frames




34

34

# P- and B- frames

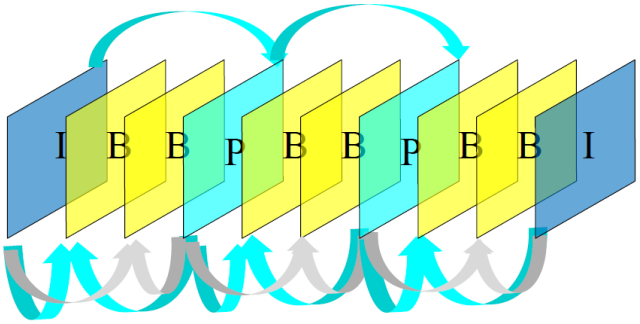
- P-frame contains 1/3 data of I-frame
  - motion vectors
  - error macroblock

- B-frame contains 1/2 to 1/5 data of P-frame
  - least data but most computational
  - delay issue:
    - need to transfer future I or P frames before any dependent B-frames can be processed

35

35

# Group Of Pictures

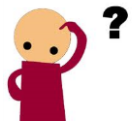



- I frames do not rely on other frames
- P frames rely on previous I and P frames
- B frames rely on previous and future I and P frames
- No frames rely on B frames

36

36

# Question



What can you say about the content of an MPEG video that contains very few I frames?

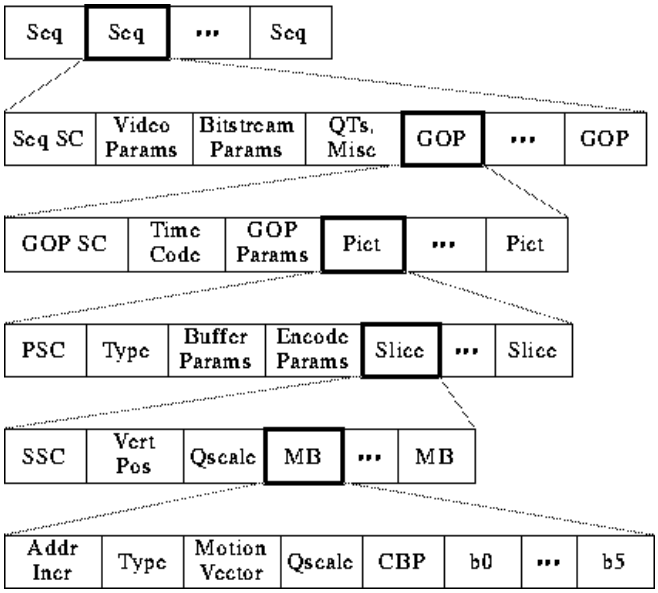




# Bitrate Allocation

- CBR – Constant BitRate
  - Streaming media uses this
  - Easier to implement
- VBR – Variable BitRate
  - DVD's use this
  - Usually requires 2-pass coding
  - Allocate more bits for complex scenes
  - This is worth it, because you assume that you encode once, decode many times

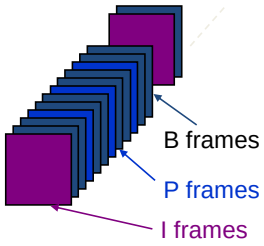
# MPEG Video Bitstream




39

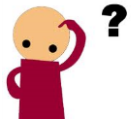
# MPEG 1&2 summary

- Images coded INTRA (I):
  - Random access
  - Error resiliance
- Images coded INTER (P):
  - Prediction from previous decoded image (I, P)
- Images coded BI-INTER (B):
  - Prediction from previous and / or future decoded image (I, P)
  - allow effective prediction of uncovered background (areas of the current picture that were not visible in the past and visible in the future)

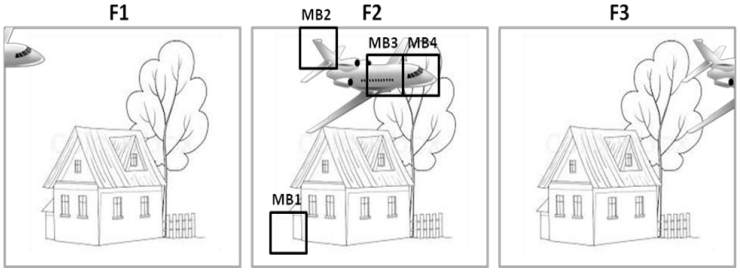



40

# Question



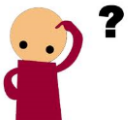
The figure below shows three consecutive frames (F1, F2 and F3) of a video. In the middle frame (F2), four macroblocks (MB1, MB2, MB3 and MB4) are shown. For each of these four macroblocks, briefly explain how it should be encoded in MPEG1 to achieve the highest compression.






41

# Question

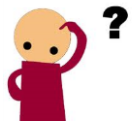


- Explain how the choice of GOP structure affects the size of the MPEG video file.



42

# Question



- The frames of any GOP structure need to be re-ordered before transmission. Why?



# Agenda

- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..

# MPEG4

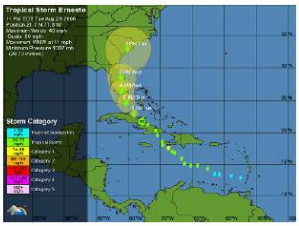
- MPEG 4(1998) provides standardised ways of representing units of aural, visual or audio-visual content, as discrete “media objects”
- These can be of natural of synthetic origin
- For example, they could be recorded with a camera or microphone, or generated with a computer



# Object-based Multimedia

# Object-based Multimedia






# MPEG4

- MPEG-4 can be used in numerous applications: streaming video; delivering 2D still images; controlling animated 3D models; handling two-way video conferences, etc.
- Coding of audio-visual objects:

– Objects can be coded and decoded *independently*

– Object can be flexibly *composed* to create different scenes

– *Natural and synthetic* object are treated in the same way

– Excellent error resilience

# MPEG4

MPEG-4 consists of several standards (parts)

- Part 2 –Advanced Simple Profile (ASP), used by codecs such as DivX, Xvid and Nero Digital
- Part 10 –Advanced Video Coding (AVC), used by the x264 codec, Quicktime 7, BBC iPlayer and Blu-ray Disc
- Part 14 –a multimedia container standard (MP4)



# Agenda

- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..

# Part 10 and H.264 AVC

- Approx. half the bit rate of MPEG2/MPEG4 Part 2
- Has many new features that allow it to compress video much more effectively, e.g.:

– Up to 16 reference frames can be used, unlike previous limit of 2 for B-frames

– Variable block-size motion compensation (VBSMC) with block size from 16 x 16 to 4 x 4, enabling precise segmentation of moving regions.



# H.264 AVC applications

- The AVC encoder is designed for use in a variety of applications: mobile video streaming; video conferencing; internet video; DSL (Digital Subscriber Line) video streaming; Digital HD TV, etc.
- It is a generic standard supporting a wide range of bit rates and a wide range of resolutions.

# QuickTime

- A multimedia framework developed by Apple Computer, with 2 main components:

- The QuickTime framework providing a common set of APIs for encoding and decoding audio and video

- The QuickTime Moving (.mov) file format, an open source media container

- Its open architecture supports many file formats and codecs, including Motion JPEG, MPEG, H.264 and is extensible to support future codecs.
- In February 1998, the ISO standards body gave QuickTime a boost by deciding to use it as the basis for parts of the new MPEG-4 standard.



# AVI

- Audio Video Interleave
- Video for Windows multimedia container format
- A special case of the RIFF (Resources Interchange File Format), defined by Microsoft
- Video for Windows supports several data compression techniques, including RLE, Indeo (Intel), MPEG4 and DivX

# WMV

- Windows Media Video
- Compressed video file format incorporating several proprietary Microsoft codecs
- WMV 9 codec is a competitor to H.264, DivX, etc.
- Adopted as format for Blu-ray discs



# Flash Video

- Container file format used to deliver video over the Internet : FLV or F4V
- Can use multiple codecs: Sorenson Spark, VP6, H.264 video

# HTML5 Video

- A HTML 5 tag introduced for the purpose of playing videos or movies
- Intended to become the new standard way to show video on the web.
- Current draft specification does not specify which video formats browsers should support



# Summary

- MPEG: Standardised set of algorithms
- Video Compression exploits temporal redundancy
- Two main types of compression: spatial and temporal
- Three types of frame:
  - Intra Coded Frames (I)
  - Predictive Coded Frames (P)
  - Bi-directional Predictive Coded Frames (B)
- MPEG 4 encodes media objects
- Other Codecs ..