Università
Ca' Foscari
Venezia

Part of this material has been reproduced (with permission) from the lecture notes of the course "Multimedia systems" held by Prof. Augusto Celentano
http://www.dais.unive.it/~auce/smm2015-16/

Computer Science Applications to Cultural Heritage

Digital video

Filippo Bergamasco (filippo.bergamasco@unive.it)
http://www.dais.unive.it/~bergamasco
DAIS, Ca' Foscari University of Venice
Academic year 2018/2019

Digital video & CH

Similar to images and audio, digital video plays an important role in cultural heritage:

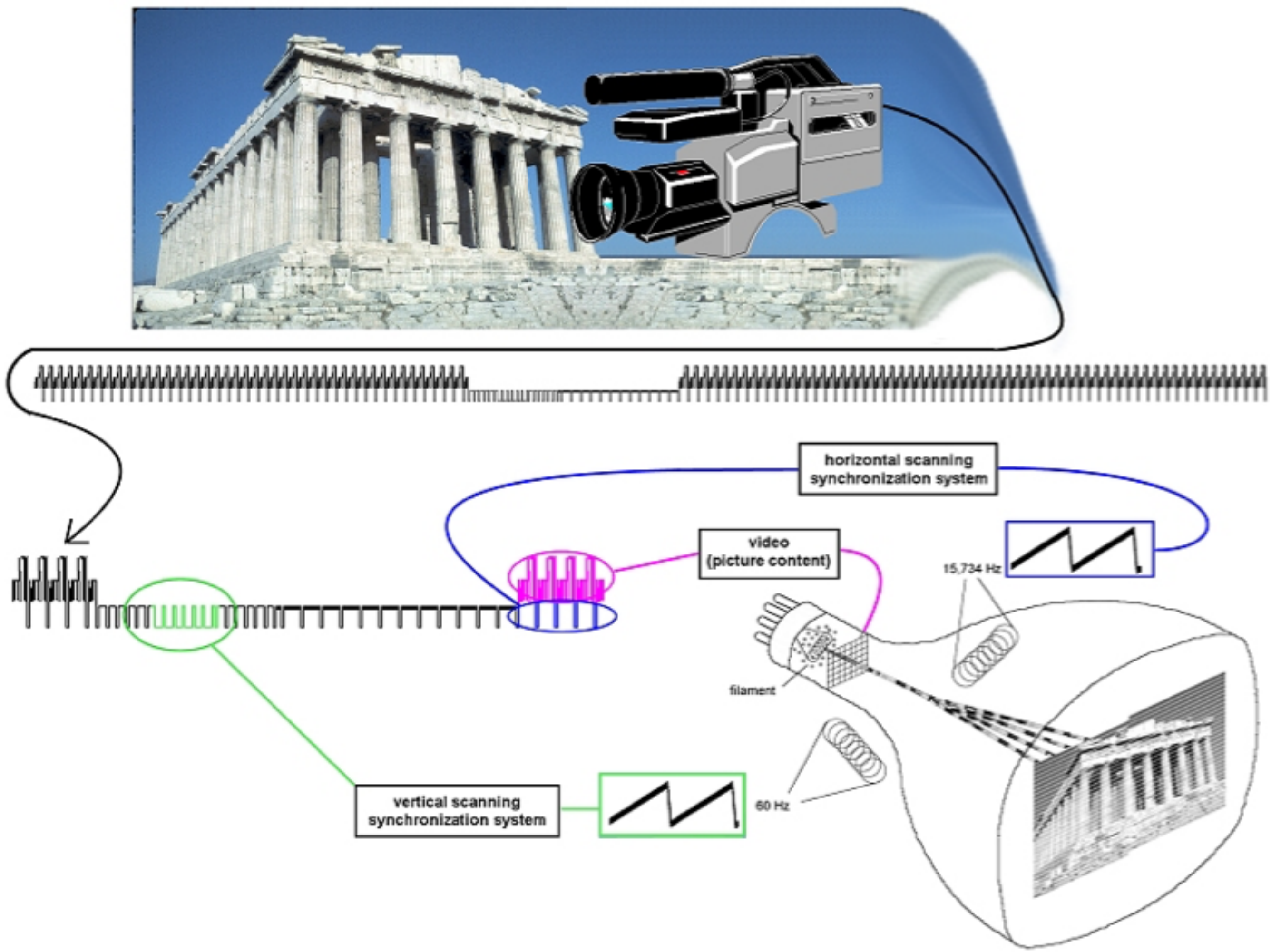
- Preservation of recent movies, documentary, and other form of video contents
- Documentation of restoration processes that are taken in the field

Storing and working with videos is a complex task for the amount of data and the spatial/temporal nature of the content (combines all the issues faced with digital images and audio)

Università Ca' Foscari Venezia

Analog video

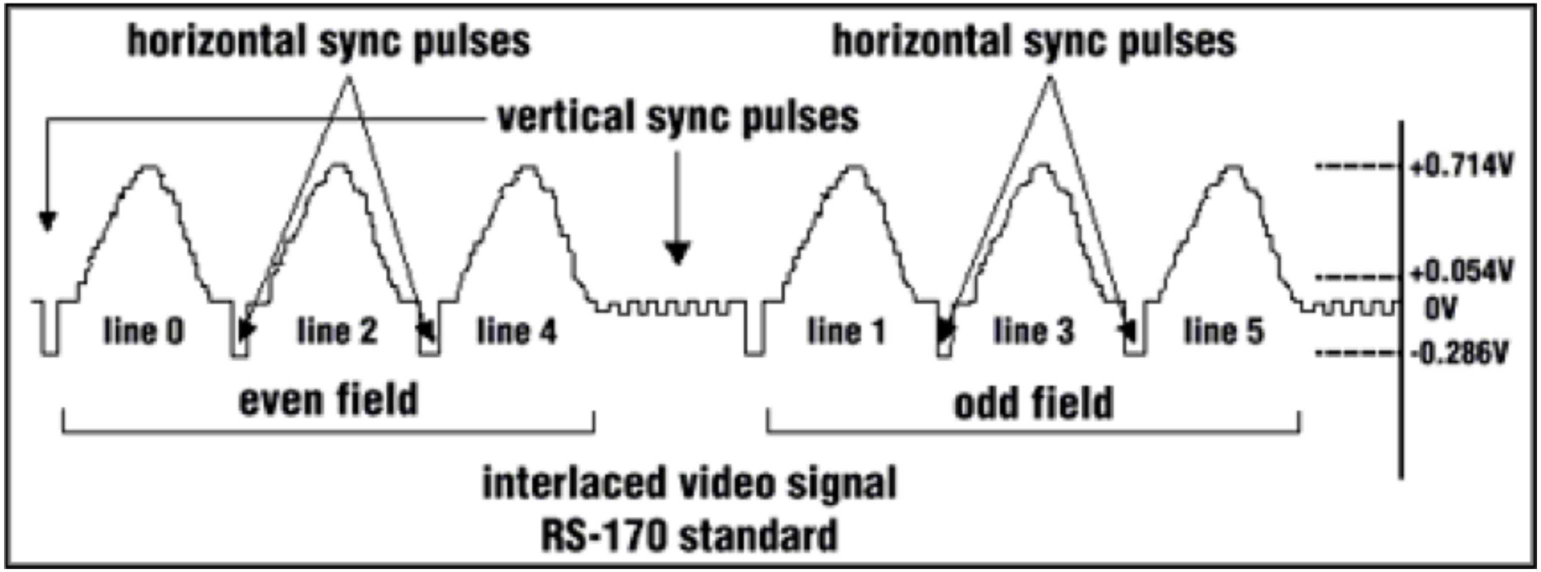
Analog video is encoded from a continuous signal changing over time.



Università Ca' Foscari Venezia

Analog video

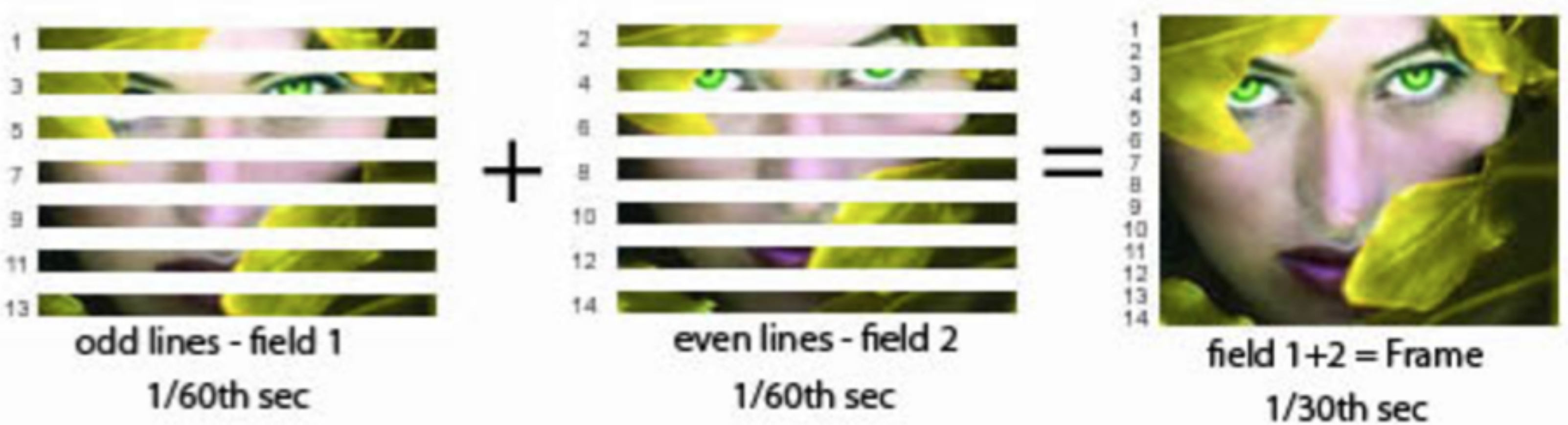
Analog video is encoded from a continuous signal changing over time.



Analog video: interlacing

Images are normally drawn using interlaced scanning

- Each frame is split into two fields, each containing only half lines, thus halving the bandwidth needed
- Even and odd scan lines are drawn two consecutive fields



What happens with fast moving subjects?

Università
Ca' Foscari
Venezia

Analog video: interlacing

Analog video: colors

How to encode color information?

Component video: Each primary channel (red,green,blue) are transmitted as separate signals

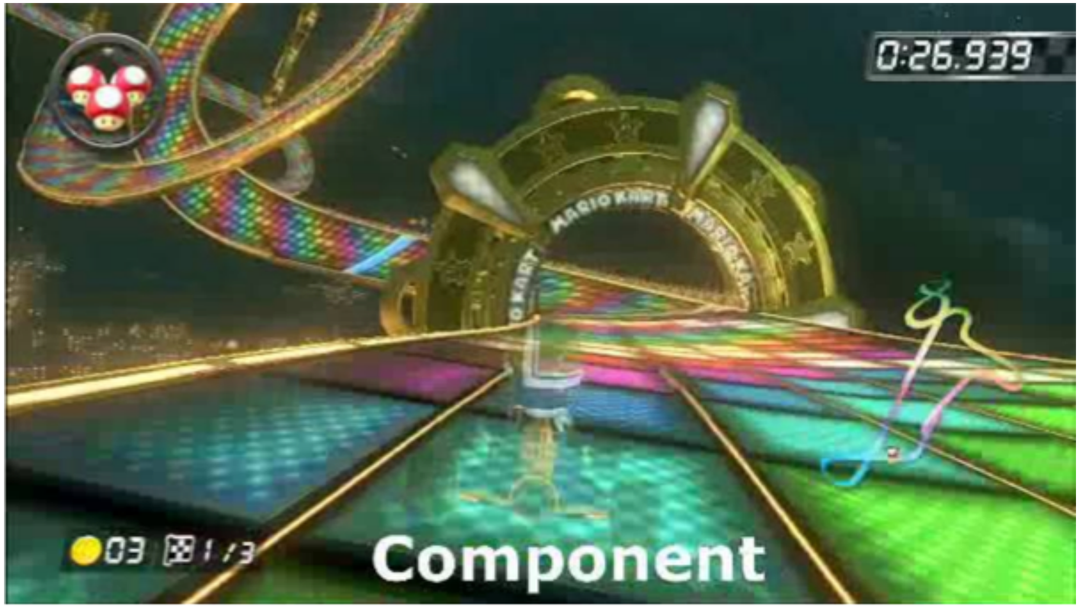
Problem: accurate temporal synchronization to avoid artifacts

Composite video: luminance and chrominance signals are processed separately, mixed on a single carrier and split at destination

Problem: interference between channels

Università
Ca' Foscari
Venezia

Component vs. Composite



Digital video

Analog video signal can in theory be digitized as-is.

Problems?

Can be easily filtered and digitized, but the elaboration of the content is problematic due to the two-dimensional nature of images

Digital video is stored as a sequence of digital images:

- full processing of the single frame content
- direct access to any frame
- no need for additional signals (vsync, etc)

Digital video: properties

Resolution: Number of pixels composing each frame

Interlaced/non-interlaced: Similar to analog video, digital video can store only even/odd rows in consecutive frames

Colour-subsampling: Resolution of chroma information can be reduced wrt. Luminance (like in jpeg images)

Frame-rate: Number of frames per second

Aspect-ratio: the ratio of the width to the height of an image or screen

Compression: usually lossy because uncompressed video is very large

Ex: 25 frames/second, 704x576 pixels, each pixel coded with 2 bytes (luminance + chrominance) requires a bandwidth of about 160Mbps = 20 Mbyte/second

Università Ca' Foscari Venezia

Digital video: standards

Low-definition standards follow some specifications of analog video for compatibility reasons:

	CCIR 601 525/60 NTSC	CCIR 601 625/50 PAL/SECAM	CIF	QCIF
Luminance resolution	720 x 480	720 x 576	352 x 288	176 x 144
Chrominance resolution	360 x 480	360 x 576	176 x 144	88 x 72
Colour Subsampling	4:2:2	4:2:2	4:2:0	4:2:0
Fields/sec	60	50	30	30
Interlaced	Yes	Yes	No	No

Digital video: HD standards

High definition video modes

Video mode	Frame size in pixels (W×H)	Pixels per image[1]	Scanning type	Frame rate (Hz)
720p	1,280×720	921,600	Progressive	23.976, 24, 25, 29.97, 30, 50, 59.94, 60, 72
1080i	1,920×1,080	2,073,600	Interlaced	25 (50 fields/s), 29.97 (59.94 fields/s), 30 (60 fields/s)
1080p	1,920×1,080	2,073,600	Progressive	24 (23.976), 25, 30 (29.97), 50, 60 (59.94)
1440p	2,560×1,440	3,686,400	Progressive	24 (23.976), 25, 30 (29.97), 50, 60 (59.94)

Ultra high definition video modes

Video mode	Frame size in pixels (W×H)	Pixels per image[1]	Scanning type	Frame rate (Hz)
2000	2,048×1,536	3,145,728	Progressive	24
2160p (also known as 4k)	3,840×2,160	8,294,400	Progressive	60, 120
2540p	4,520×2,540	11,480,800	Progressive	
4000p	4,096×3,072	12,582,912	Progressive	
4320p (also known as 8k)	7,680×4,320	33,177,600	Progressive	60, 120

Università Ca' Foscari Venezia

Digital video compression

Video content is so big that compression is mandatory in practice. As usual, compression is based on redundancy elimination:

Intra-frame coding: compress each image frame independently exploiting spatial redundancy (see for example how jpeg compression works)

Inter-frame coding: exploits the temporal redundancy of video data (ie. subsequent frames are usually very similar in the content)

Digital video compression

Frame encoding of a digital video can omit a lot of information based on temporal redundancy

- Unless a scene change occurs, the differences between two consecutive frames are small
- Differential encoding can be used based on the previous frame content
- If the difference is small the compression is efficient

Motion compensation

The differences between two sequential frames is often due to translations of parts of the image:

- the translations may be due to movements of the camera or objects in the scene
- by splitting the image into small segments, corresponding segments in sequential frames due to partial translations can be identified by similarity
- the difference between an actual frame and a frame computed by motion estimation is often very small

This compression technique is called motion estimation compensation (motion compensation)

Università Ca' Foscari Venezia

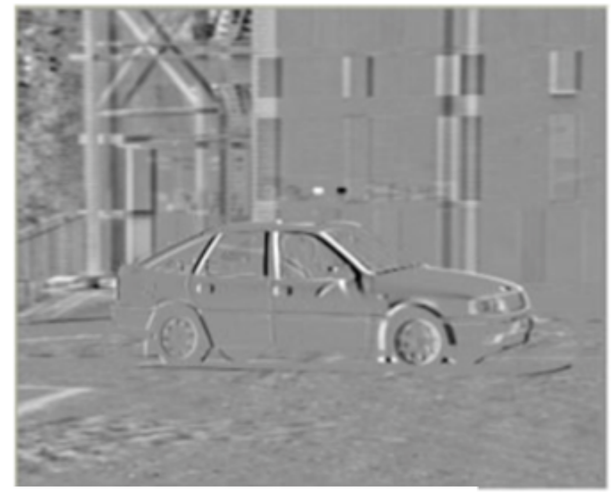
Motion compensation

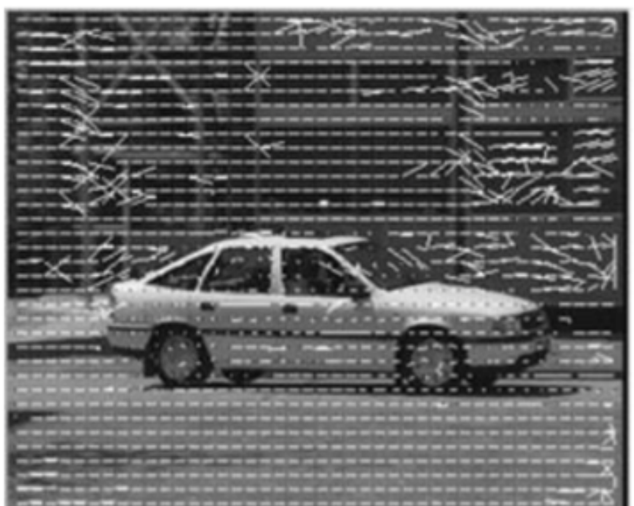
Usually performed in 3 steps:

1. the frame image is split in macroblocks (usually 16x16 pixel, 8x8 pixel if chroma information is subsampled)
2. for each macroblock a motion vector is computed by searching in previous or next frame the most similar block
3. movement (motion vector) and residual differences between each moved block and the original frame are coded

Università
Ca' Foscari
Venezia

Motion compensation





Motion vectors

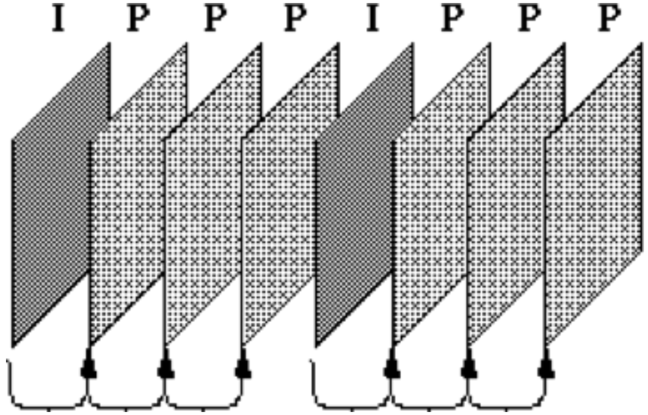
Residual differences

H.261 & H.263 compression

H.261 and H.263 were the first encoding standards based on motion compensation.

Frames are divided into two categories:

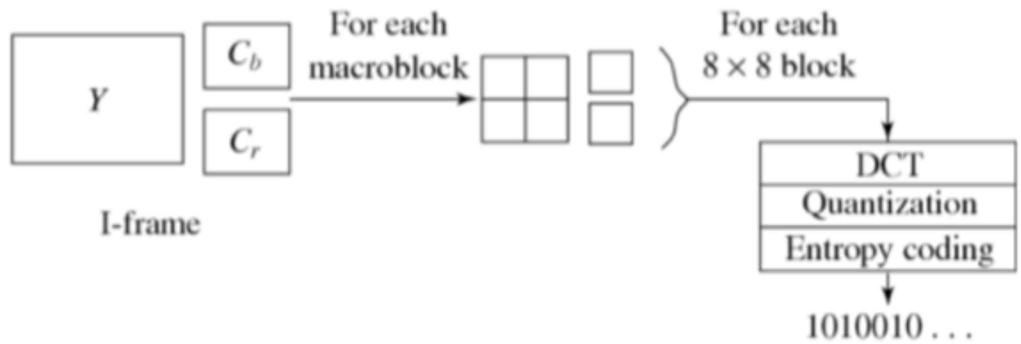
- i-frames (intra-frames) are coded independently with JPEG compression techniques
- p-frames (predicted-frames) are coded by motion compensation from the previous i- or p- frame



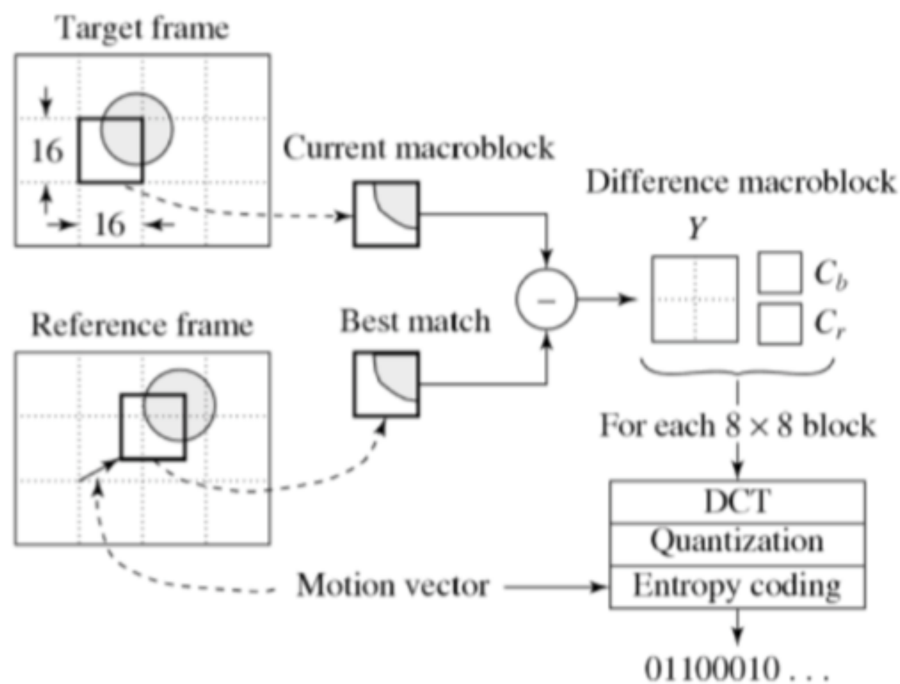
Università
Ca' Foscari
Venezia

H.261 & H.263 compression

i-frame coding



p-frame coding

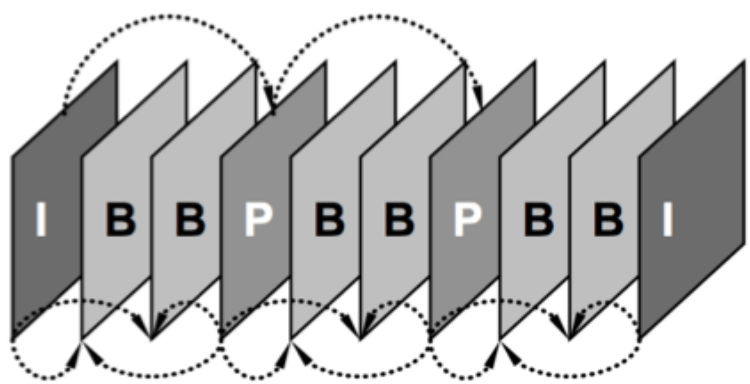


MPEG-1

Università Ca' Foscari Venezia

The MPEG-1 compression standard improves the H.261 and H.263 compression algorithm with a more sophisticated motion prediction scheme:

- **i-frames** are fully coded with a JPEG algorithm
- **p-frames** are encoded according to an estimate from the previous i- or p-frame
- **b-frames** are coded according to two motion estimates from previous and following frames (bidirectional motion estimation)



MPEG-1

I-frames (intra-coded)

- require more storage space
- enable random access to any frame

P-frames (predictively coded)

- minimum differences are computed on the absolute value of the luminance components
- are small but propagate transmission errors

B-frames (bi-directional predictively coded)

- are the most compressed frames but require both previous and sequent frames

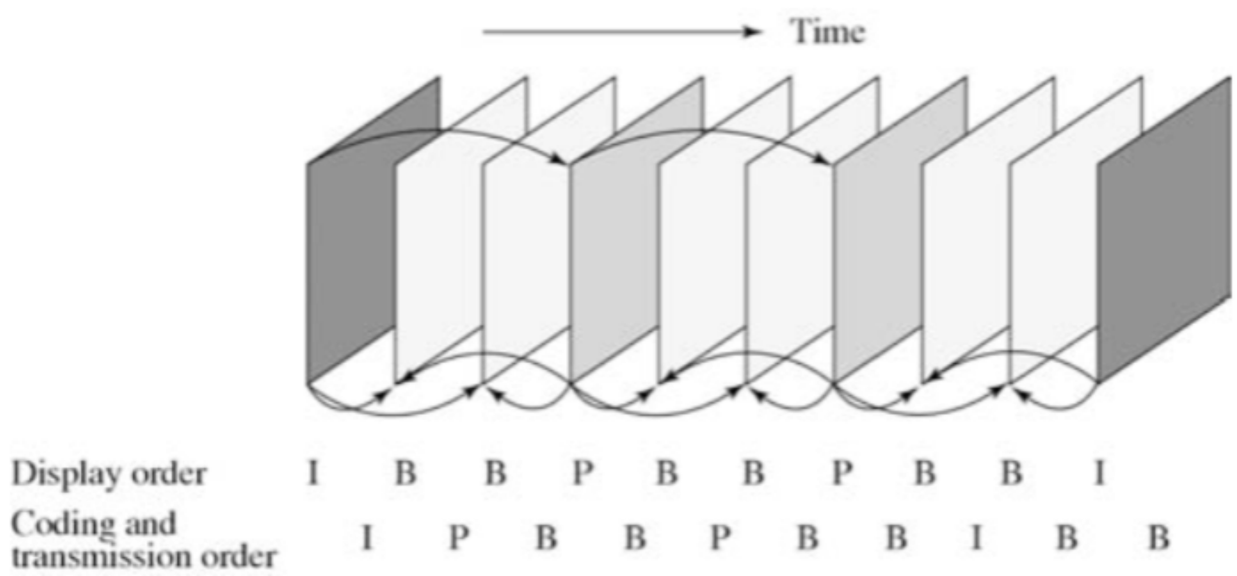
I-frames allow random access to intermediate movie points, but increase the bit-rate

Università Ca' Foscari Venezia

MPEG-1

B-frame encoding minimizes the information when the frame content changes gradually

The dependence of B-frames on the next frame alters the coding and transmission frame sequence over the display frame sequence



H.264

H.264 exploits extensive intra-frame and inter-frame redundancies with **variable size blocks** and **extended frame range:**

- i-frame compression applies a differential coding over the already encoded blocks of the same frame
- p-frames and b-frame compression extends the search range to several previous frames looking for a best match
- macroblocks can be hierarchically decomposed from 16x16 pixels to 4x4 pixels (quad-tree)

An estimate 30-50% compression improvement over MPEG (1 and 2)

Università Ca' Foscari Venezia

Video codecs and formats

A **codec** (coder-decoder) is an algorithm that allows compression and decompression of audio and video data streams according to a model

- e.g., MPEG-1, H263, H.264

A **format container** is a file format that contains and organizes data for different types of audio and video streams, compressed using standard codecs, and other information (e.g., subtitles, chapters)

- e.g., AVI, MPG, VOB

A format container can support multiple codecs, and a codec can be supported by several format containers

Video format container

The choice of a format container is based on 5 properties

- **popularity:** a widely used container can be interpreted by many applications on many platforms
- **overhead:** a container may require additional data that increase the file size without increasing the audio/video quality
- **support for advanced features:** some codecs require more complex data structures that could not be supported by all containers (e.g., time stamps, back propagation)
- **support for different types of content beyond audio and video streams:** chapters, subtitles, metadata
- **support for streaming transmission:** depends on multiple streams structure and interleaving

Popular video containers

Name	Extension	Video codecs	Audio codecs	Streaming?
MPEG	.mpg	Mpeg-1, Mpeg-2	mp3	yes
AVI	.avi	many	many	no
MPEG-4	.mp4	H.264, h.265, etc	mp3,aac	yes
Matroska	.mkv	many	many	yes
DivX	.divx	Mpeg-4 part 2	Mp3, PCM, AC3	yes