Basic Concepts of Digital Video

Luca Guarnera, Ph.D.

Research Fellow

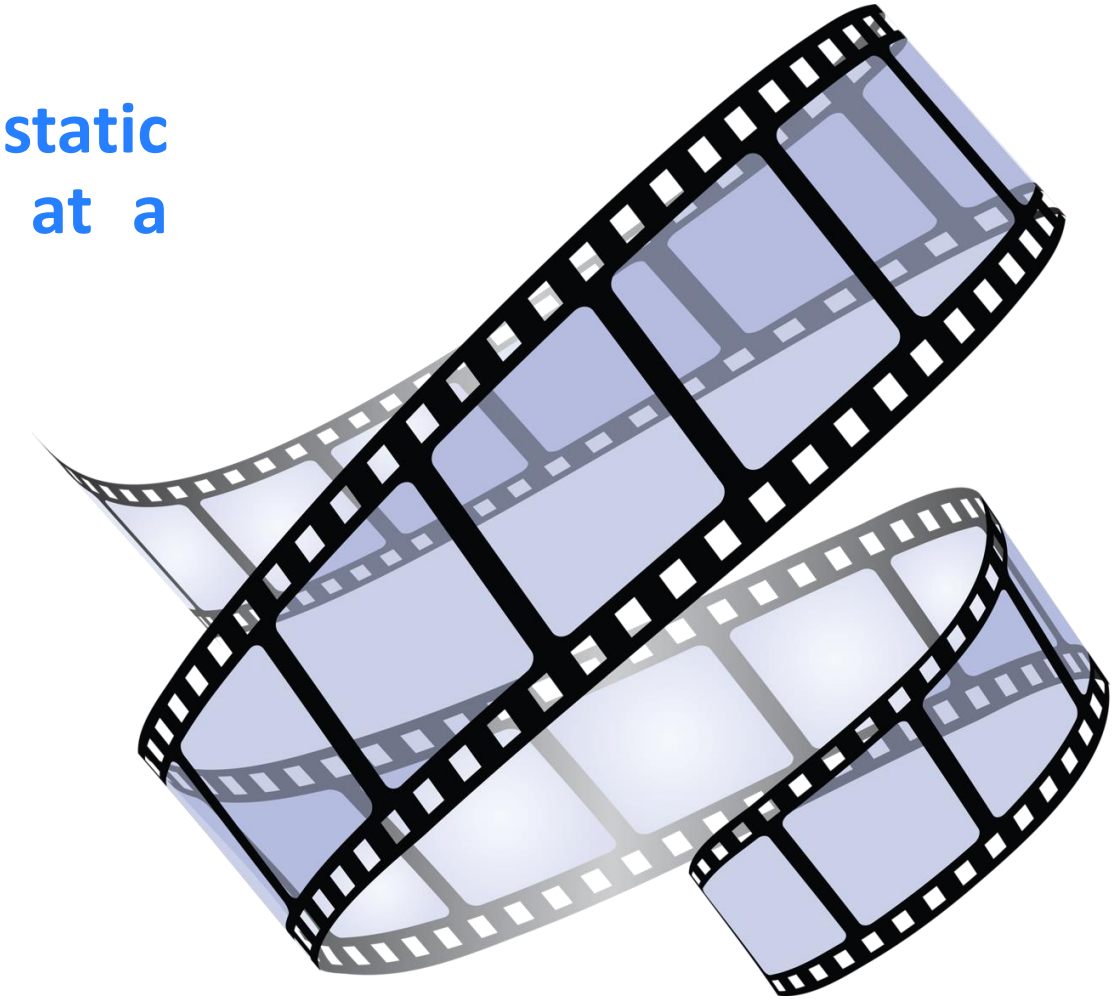
luca.guarnera@unict.it

University of Catania

Dipartimento di Economia e Impresa

Digital video: fundamentals

- A digital video consists of a **sequence of static images that are displayed sequentially at a certain temporal frequency**.



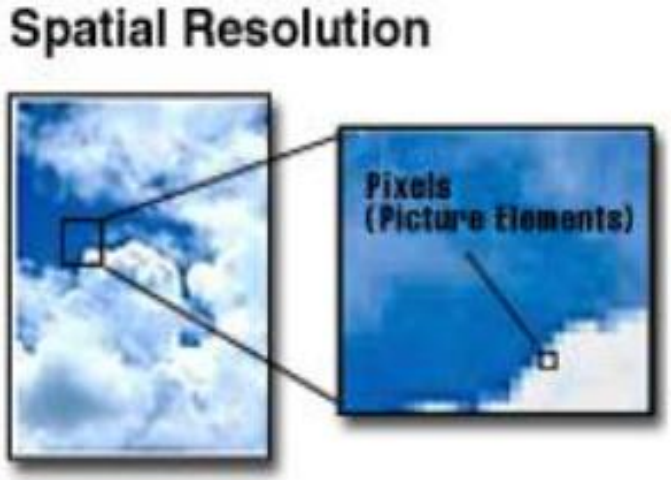
- Video as a discrete signal
  - Temporal sampling of the scene
  - At each instant the scene is "photographed"

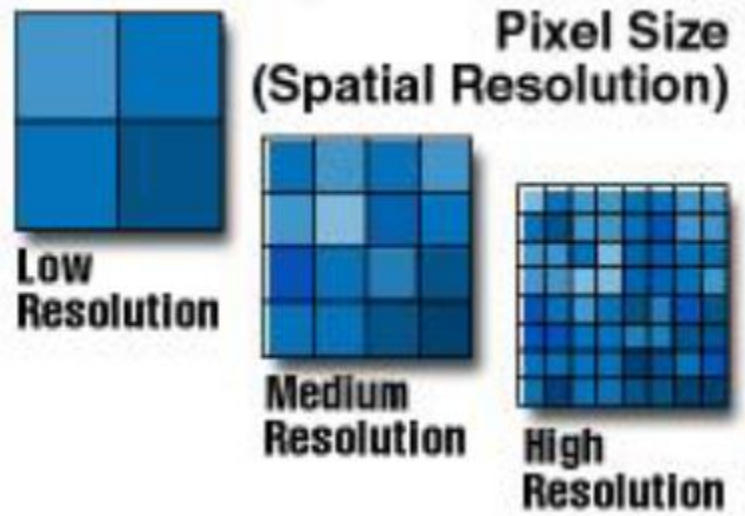
- Video sequence
  - **Frames** (frames)

Spatial Resolution

Definition

Spatial resolution refers to the specific number of information points (pixels - Picture Element) in an 'image.



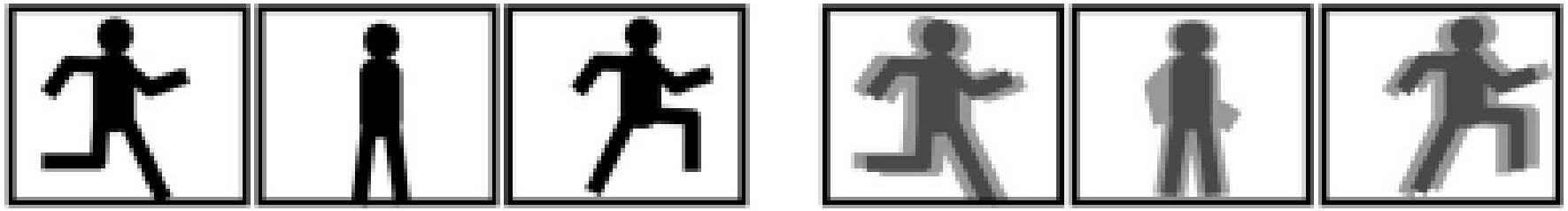


Temporal Resolution

Definition

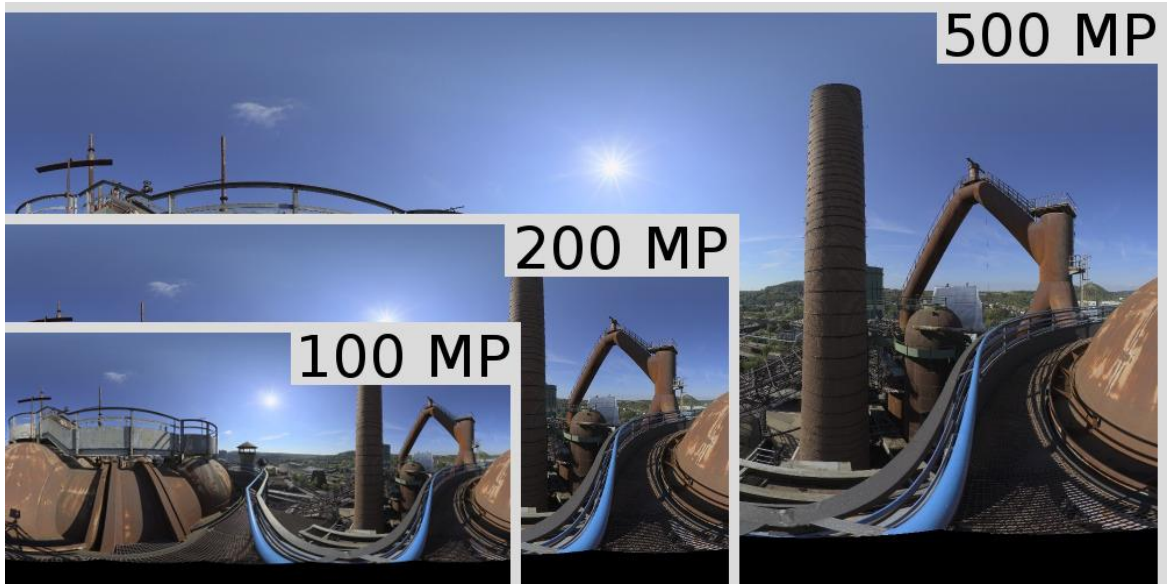
Image rate, also called **frame rate**, is the number of images per unit time that are displayed.

- To achieve the illusion of a 'moving image the minimum frame rate is about 10 frames per second.



Resolution
MP “MegaPixel”

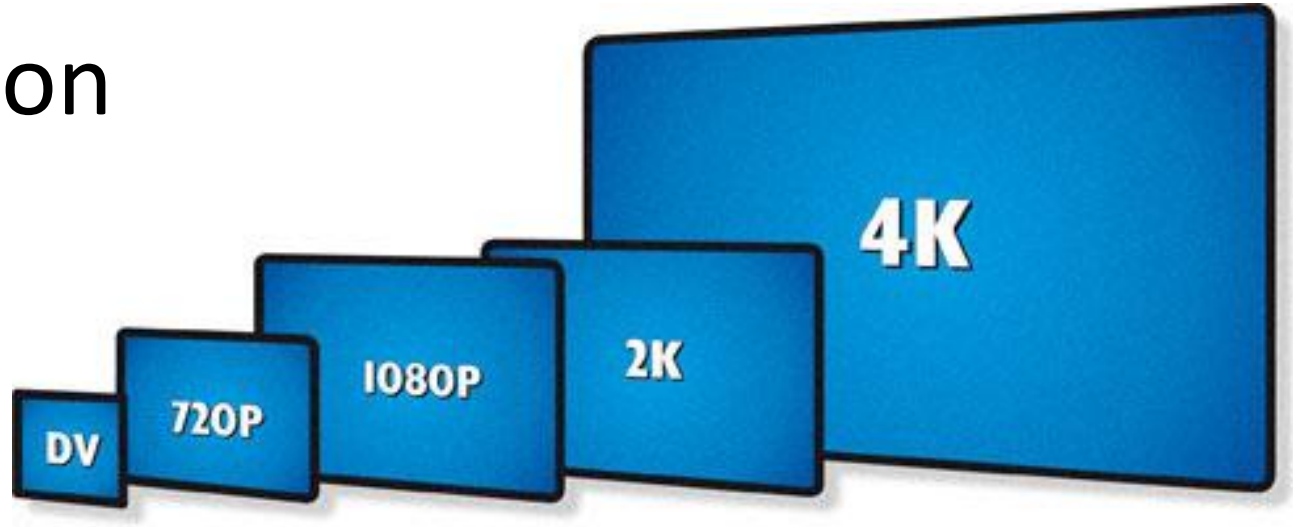
- Unit of measurement that is equivalent to 1 million pixels



Calculating the MPs of a device

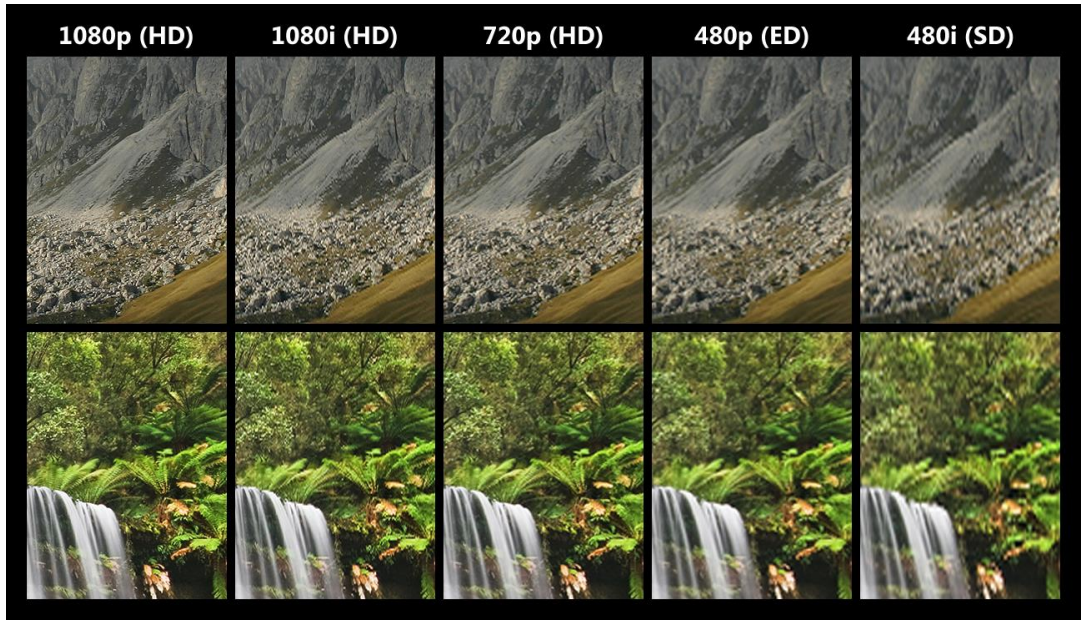
$$\frac{Max_{horizontal\ resolution} * Max_{veritcal\ resolution}}{1.000.000}$$

Format resolution



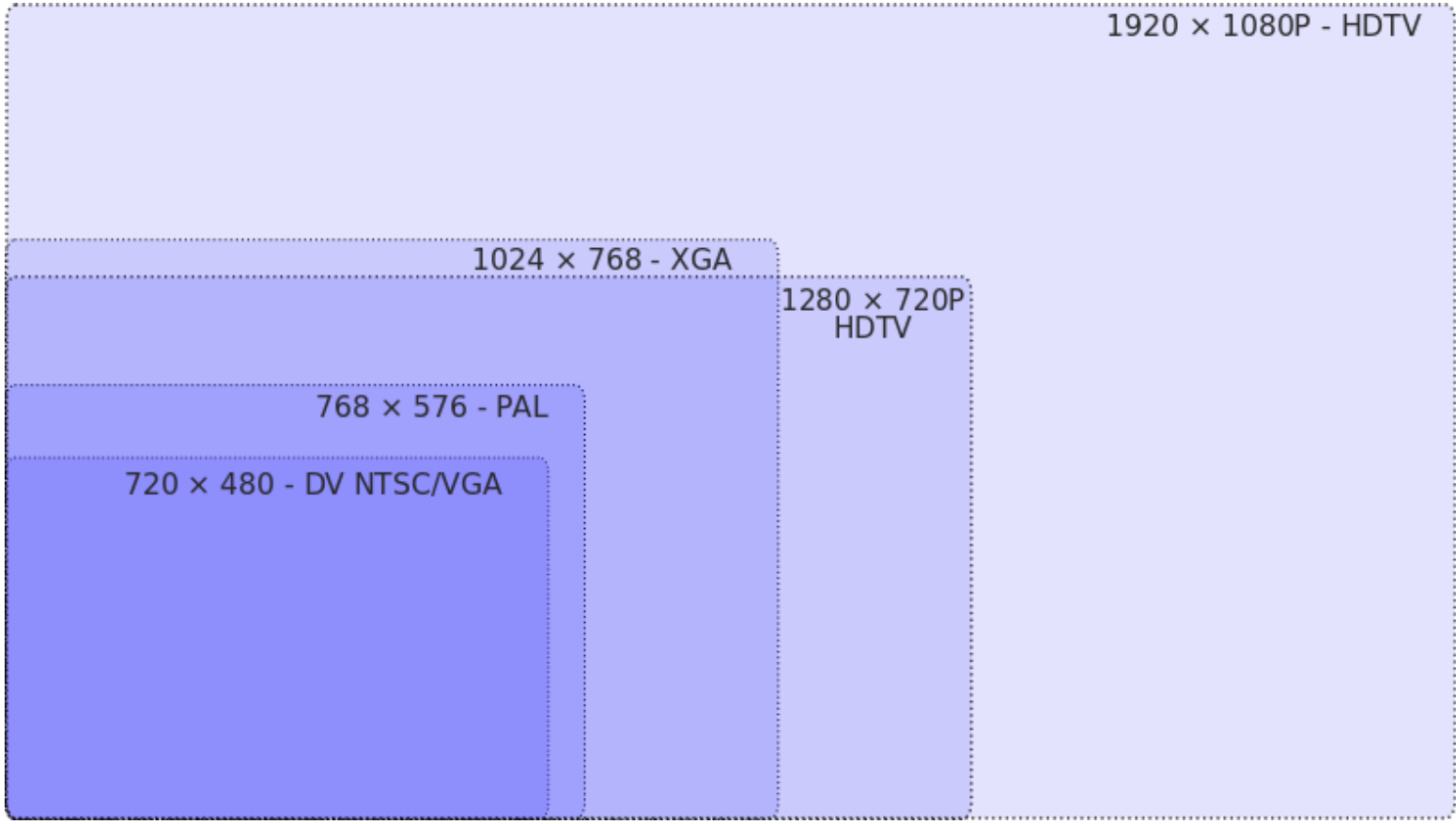
4 formats

- **Half resolution** (540p): 960x540 pixel
- **HD ready** (720p): 1280x720 pixel
- **1080i**: 1920x1080 pixel (interlacciato)
- **Full HD** (1080p): 1920x1080

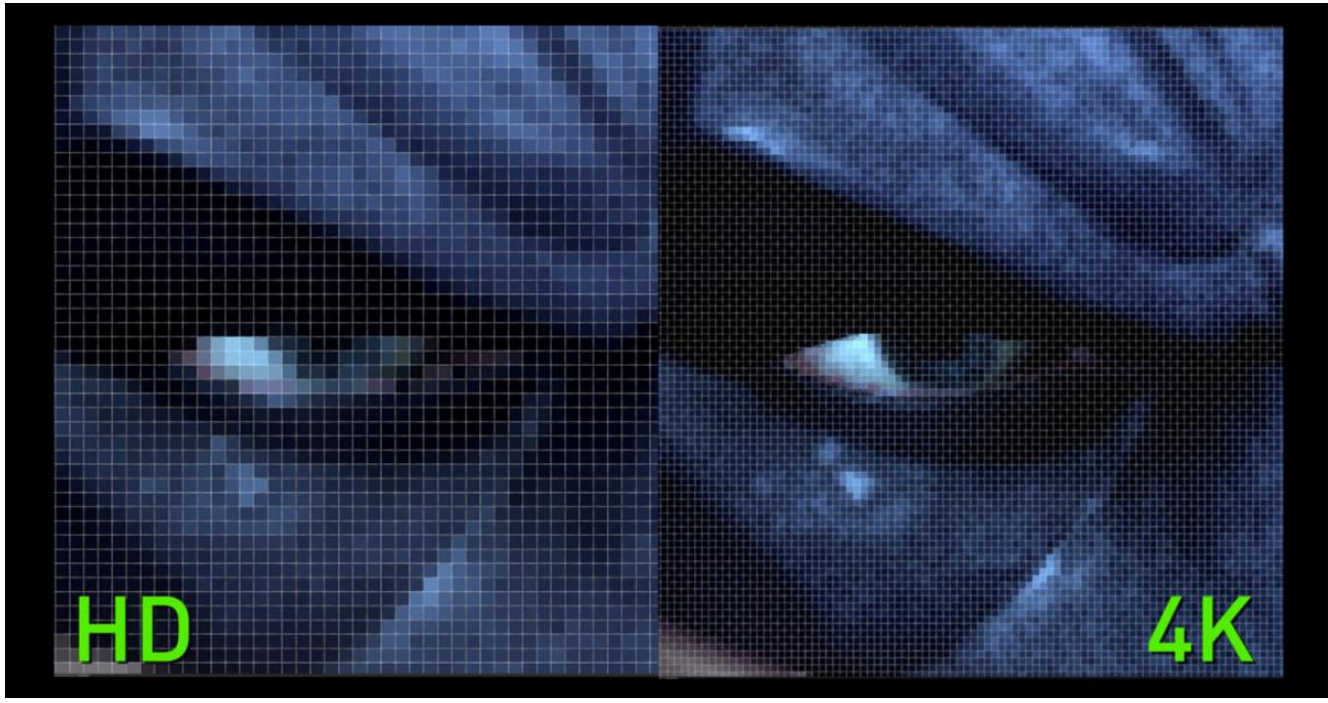


Future

- **Super High Definition (SHD), called 4K**
  - 3840x2160 pixel resolution (4 times a FullHD)
  - Present a few videos on YouTube
- **Ultra High Definition TeleVision (UHDTV)**
  - Resolution 7680x4320 pixels (16x a FullHD)

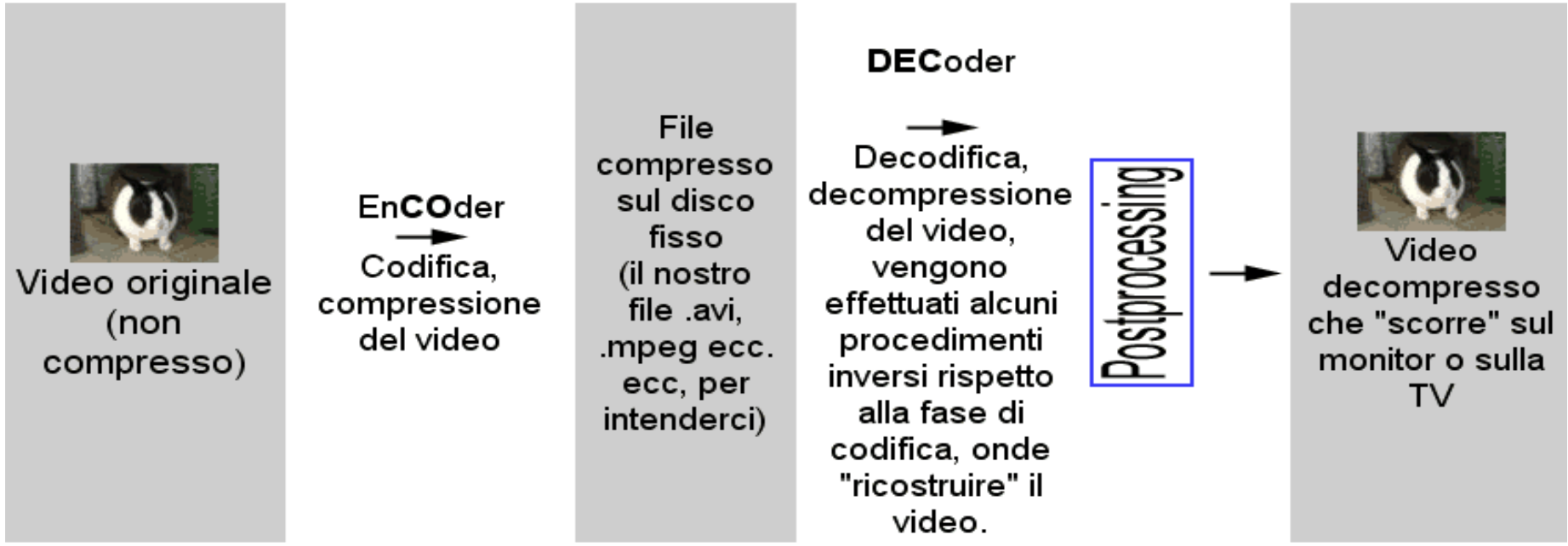


Comparison of format resolution



Video Compression

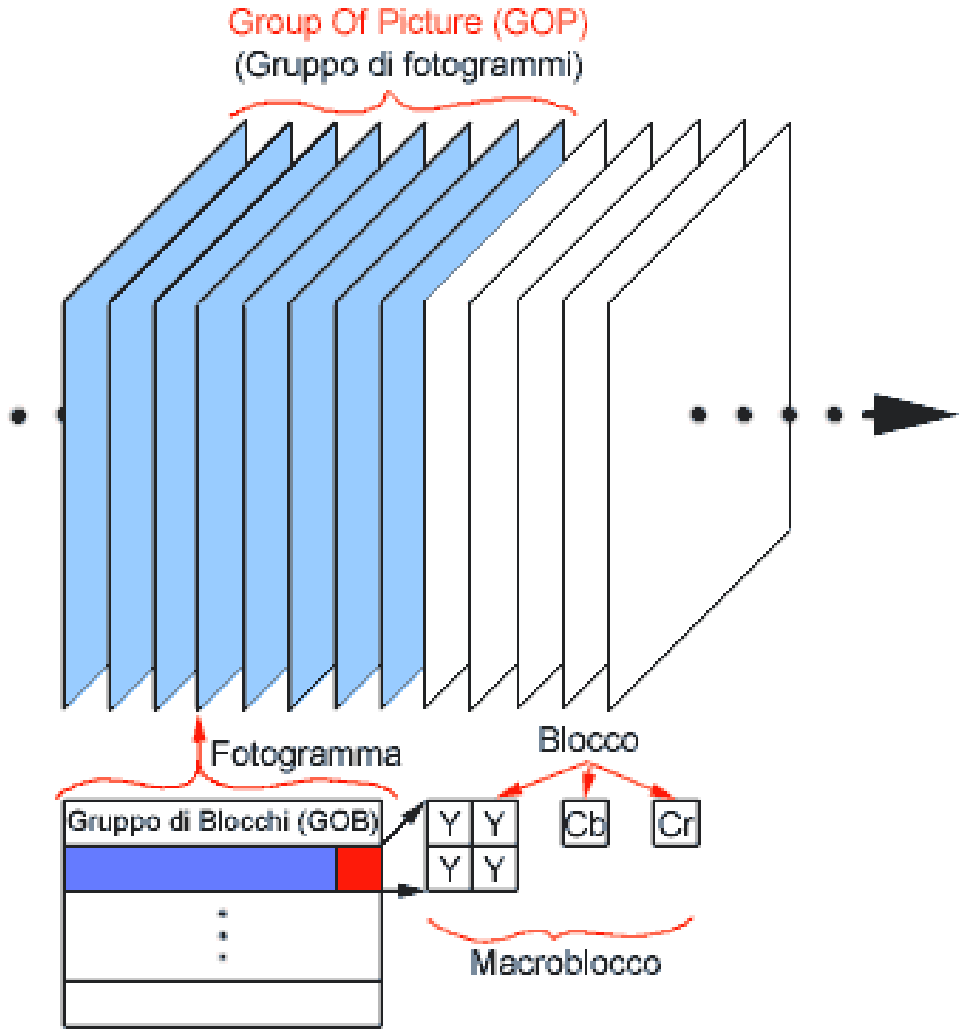
A video **CoDec** (Co-Dec = Coder/Decoder) is software consisting of two parts: the **enCoder** that **compresses the sequence of images** (video) by storing it in a file and a **Decoder** needed to **decompress the sequence and be able to view it again**.



Video

A video consists of a **sequence of images** that follow one another in rapid sequence.

So, when you **compress a video stream you are essentially compressing images**.



How is it possible to compress a video?

Techniques are used that exploit some inherent characteristics of the video itself, in combination with the characteristics of the human visual system.

In particular, it is possible to compress a video signal by attacking:

- **spatial redundancy and temporal redundancy**
- **the characteristics of the human visual system**

How is it possible to compress a video?

In particular, it is possible to compress a video signal:

1. **Removing the statistical redundancy (repetition) contained in a video and retaining only the information that is actually useful; a "less correlated" representation of images is sought, eliminating "repetitions."**
    1. It can be shown that adjacent, neighboring pixels within the same image exhibit very similar characteristics with regard to color and brightness;
    2. intra-frames encoding deals with removing this repetitiveness also called spatial redundancy within the same

How is it possible to compress a video?

- **There is also a clear correlation not only between pixels in the same frame, but also between pixels in adjacent frames**:
  - one frame and its two neighbors (the next and the previous) are often nearly identical (exceptions are situations where there are scene changes);
  - this **temporal redundancy** between neighboring frames, which exploits their minimal differences, is dealt with by inter-frames coding.

How is it possible to compress a video?

2. **Taking advantage of some peculiarities of the human visual system: the poor sensitivity of the eye to high video frequencies especially when dealing with moving images.**
   - It is possible to "cut out," throw away, some information especially relative to the high frequencies of an image without introducing visible artifacts. In fact, the human visual system is unable to perceive variations in the details of highly jagged figures;
     - it is very difficult to realize a loss of detail in the foliage of some moving trees;
     - much easier, on the other hand, to notice even the smallest variation in color or brightness in the blue of a clear, serene sky in the background of a video.

Encoding Formats

- The format is a kind of box that contains the codec and integrates it with the system.

- The codec (COmpressor-DECompressor) is software that tells the computer with which mathematical operations it should manipulate images to compress them and which ones to perform to display the compressed ones.

- There are many codecs in existence as opposed to formats.

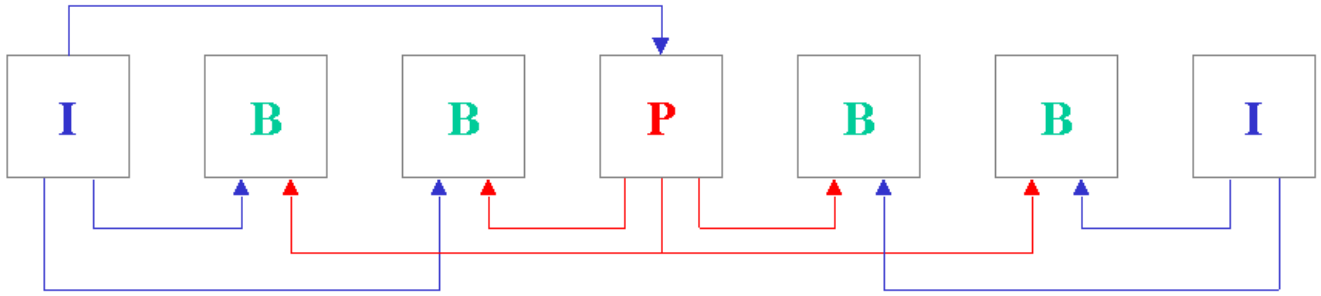
Encoding Formats

- List of major formats:
  - .avi - .mpeg e mpeg2 - .hdv - .mpeg4 o .divx - .wmv - .mov - .flv - .3gp - .asf - .flm - .real media audio/video - .dv

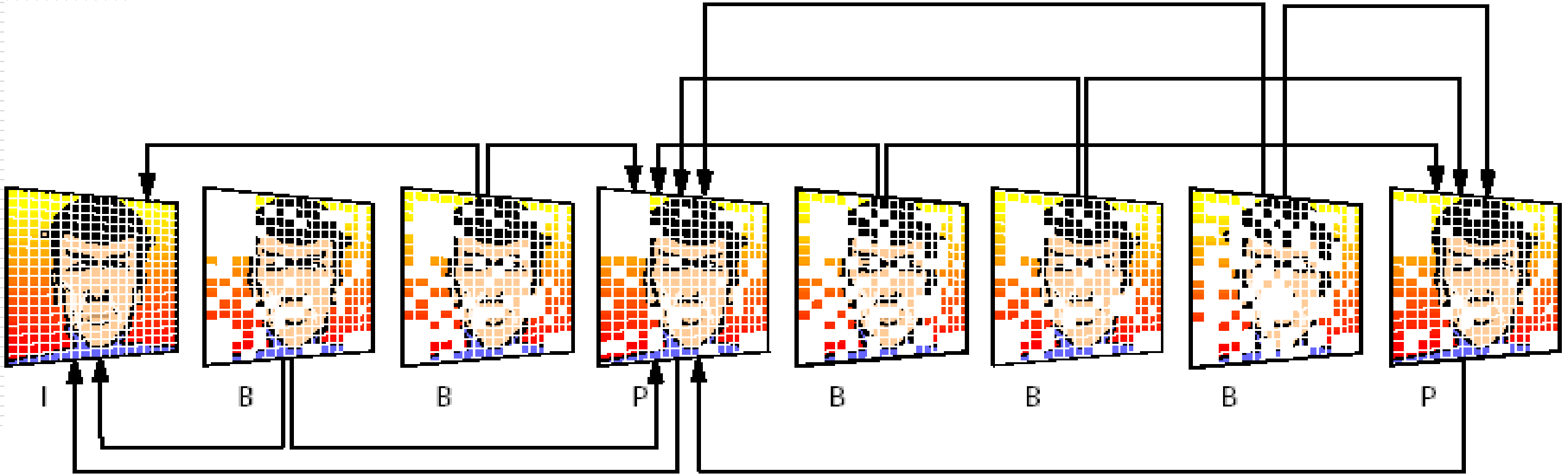
Frames I/P/B

MPEG standards provide for the classification of frames into three types: I, B, P .

- **I frame** is a completely independent video frame.
- **P frame** (predictive frame) is based on a previous I frame.
- **B frame** (bi-directional frame) consists of information derived from both I frame and P frame (also successive) through interpolation .



Frames I/P/B – Esempio



Intra-Frames

- Type I frames, also called **Intra-Frames or Key-Frames**, are frames that are encoded using the information contained in the frame itself and contain no reference or information about adjacent frames;

- They are basically compressed in the same way as a single image, in the same way as when an image is saved in **JPEG format**.

- No temporal compression (i.e., compression that also takes into account subsequent and/or preceding frames) is applied to these frames.

Intra-Frames

- Can be generated by an encoder to create a random access point
- To enable a decoder to start decoding correctly, starting from scratch at that position in the video
- Typically require more bits to encode than other types of frames They are used as a reference for decoding other images

Intra-Frames

- Typically, keyframes are inserted by the codec whenever there is a sudden change between two successive images.

- If a maximum interval is also specified between one keyframe and the next, the codec will necessarily insert a keyframe even if it is not strictly necessary.

Compression

- The term **data compression** refers to the data processing technique that, implemented by means of appropriate algorithms, allows the reduction of the amount of bits required for the digital representation of a piece of information.

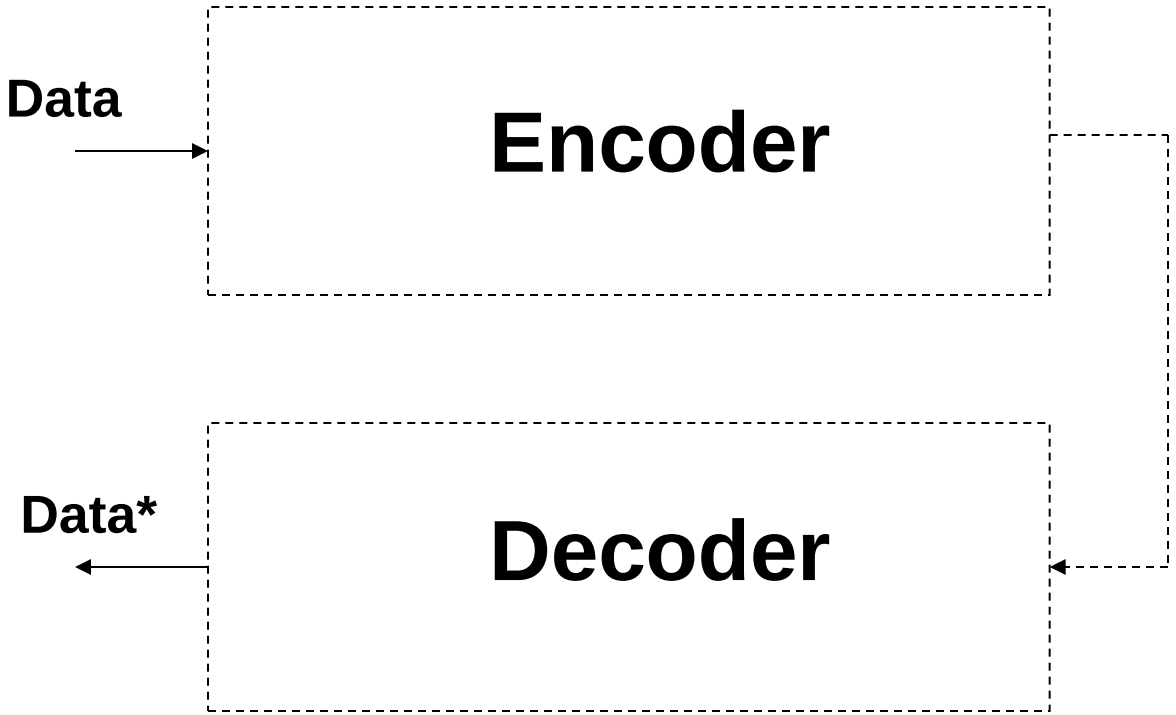
- Compression reduces the size of a file and thus the space required to store it.

- Since different amounts of data can be used to represent the same amount of information, representations that contain irrelevant or repeated information contain so-called **redundant data**.

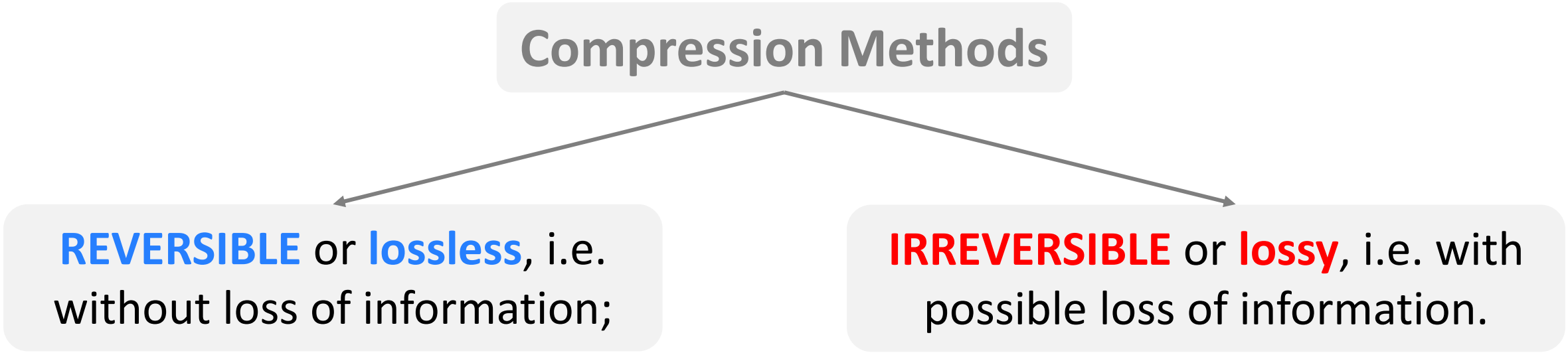
Compression Algorithm

A compression algorithm is a technique that removes redundant information from data and saves memory

Brief overview of JPEG compression



Classification of compression methods



Lossless Compression

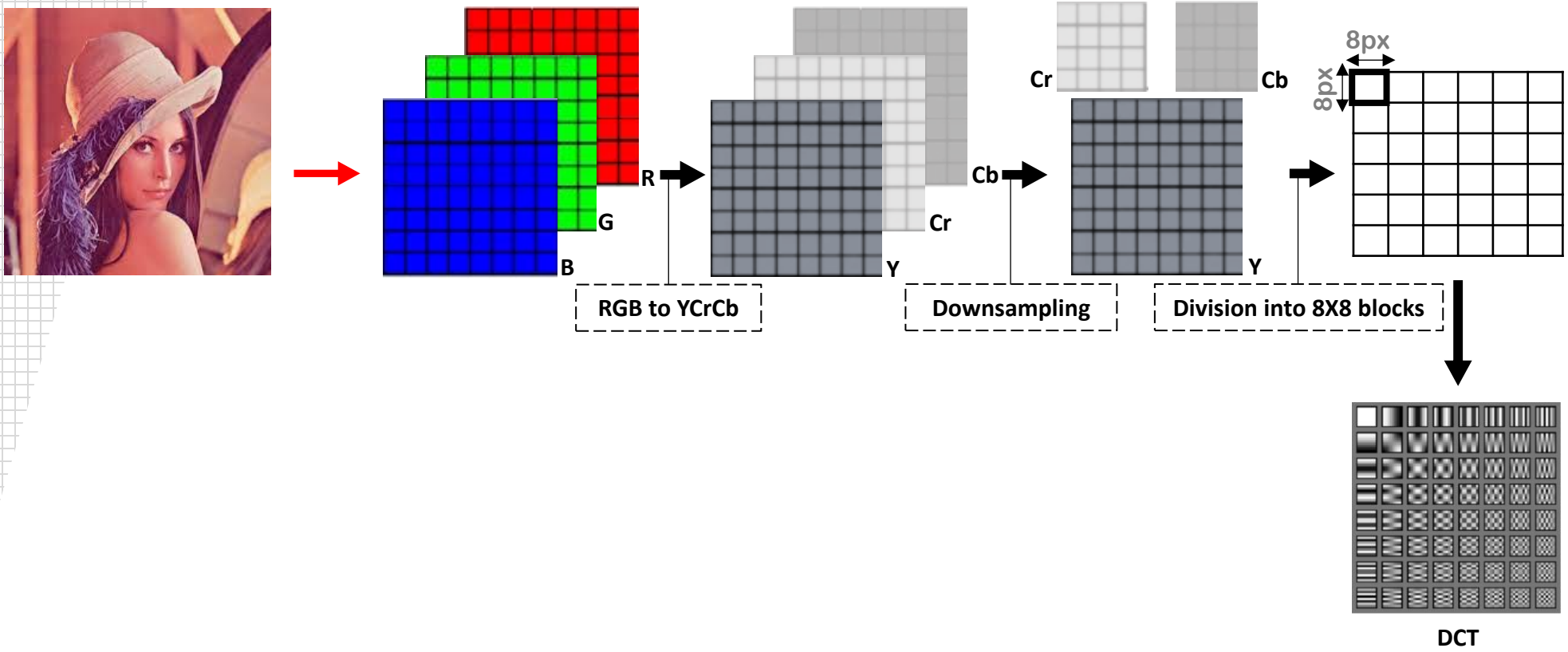
- **Definition**: Reduces file size without any data loss; the original file can be perfectly reconstructed from the compressed file.
- **Result**: The compressed file is identical to the original when decompressed.
- **Applications**: Used for text, software, and high-quality media where data integrity is crucial. Examples include:
  - PNG (images)
  - FLAC (audio)
  - ZIP (data archiving)
- **Pros**:
  - No quality degradation.
  - Reversible compression.
- **Cons**:
  - Typically achieves less compression compared to lossy methods.
  - Larger file sizes compared to lossy compression.

Lossy Compression

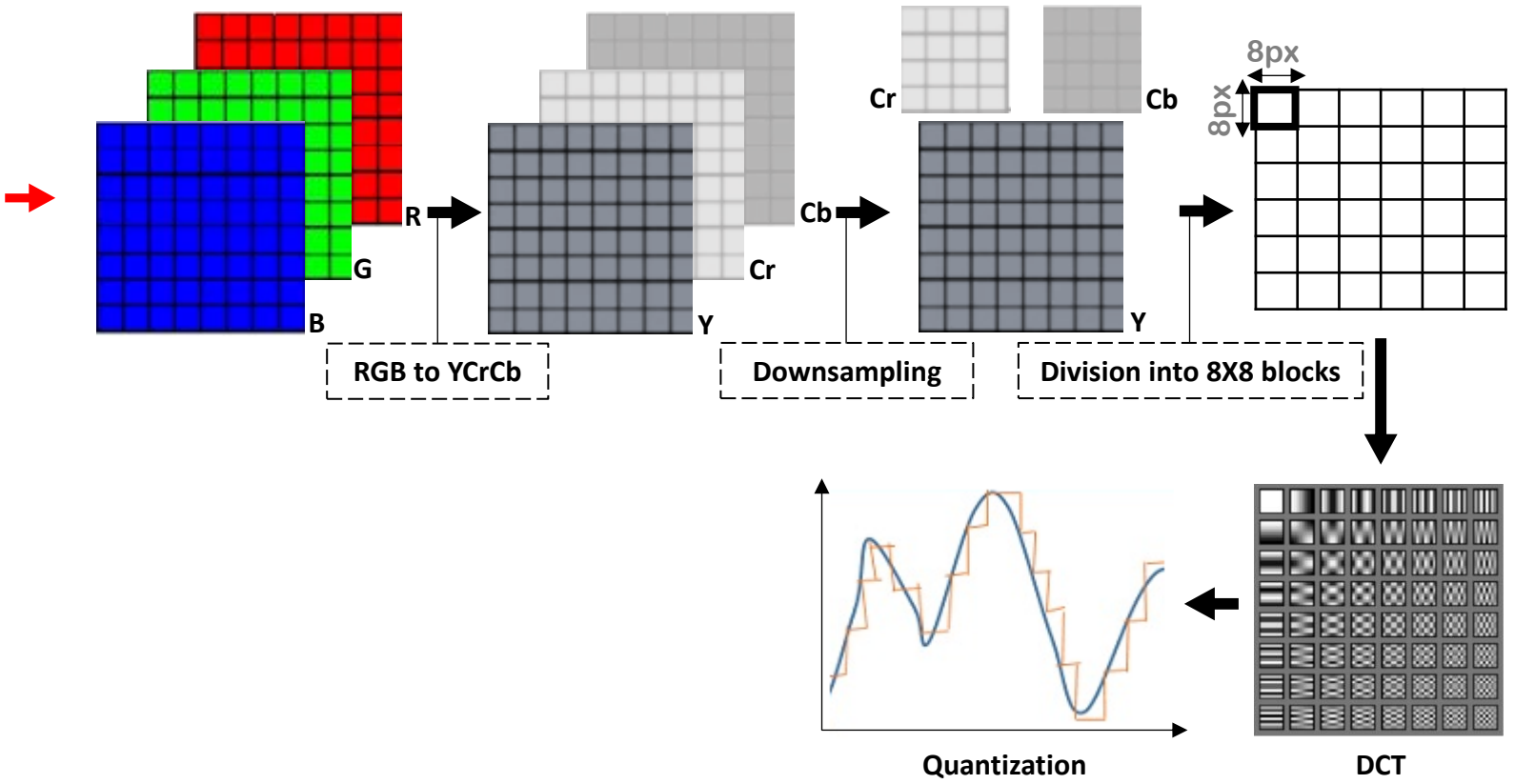
- **Definition**: Reduces file size by permanently eliminating some data deemed less important.
- **Result**: The compressed file is smaller, but it is not identical to the original.
- **Applications**: Used for media files like images, audio, and video where perfect fidelity is less critical. For example:
    - JPEG (images)
    - MP3 (audio)
    - MP4 (video)

- **Pros**:
    - Significant file size reduction.
    - Suitable for streaming or storage with limited capacity.
- **Cons**:
    - Quality loss, which may become noticeable if over-compressed.
    - Cannot restore the original file after compression.

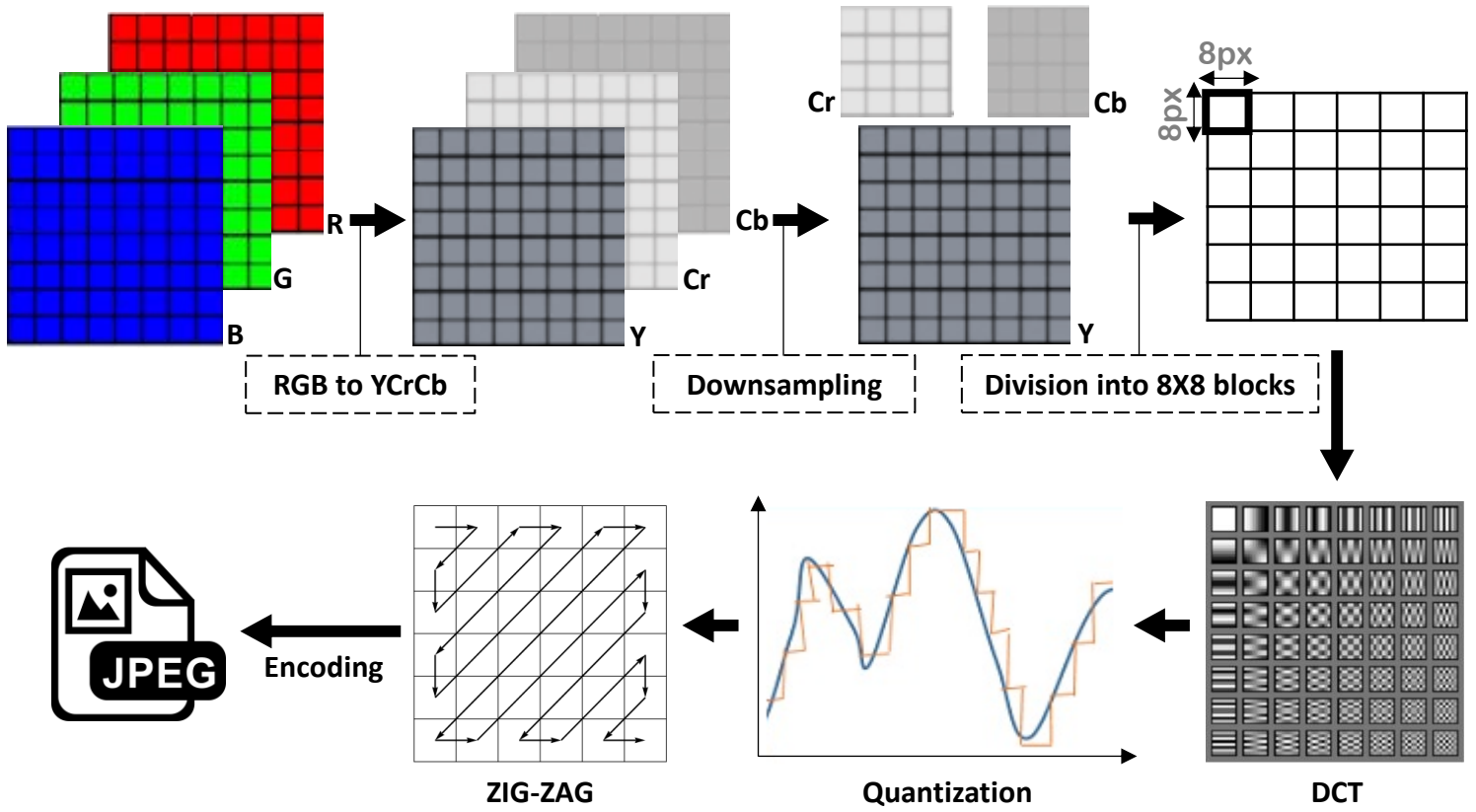
Brief overview of JPEG compression



Brief overview of JPEG compression



Brief overview of JPEG compression



P-Frames

- The P-Frame, (Predicted frames) is encoded using information acquired based on the frame preceding it, whether it is an I-Frame or a P-Frame.

- Each macroblock of 16x16 pixels in a P-Frame can be encoded independently (as in the case of the I-Frame) or it can be compensated, that is, balanced using information from the previous frame.

- Using the similarities between successive frames, the P-Frames turn out to be smaller than the corresponding I-Frames.

P-Frames

- It can contain the frame data displacements (motion vector) with respect to the frame on which it depends

- A frame of type P contains the position information (X',Y') in the current frame in which a block that had coordinates (X,Y) in the previous one moved (Motion Estimation/Compensation).

The disadvantage of using this type of frame is at the decoding phase; in fact, it is necessary to "reconstruct" each P frame before it can be displayed, and to do this one must always start from the P frame following the last key frame.

B-Frames

- Bidirectional Predicted, Bidirectional-dependent frames
- Requires previous decoding of other frames before being decoded
- Basically, B-frames are Bidirectional in the sense that they refer to both what precedes them and what follows them
- It can contain the frame data the displacements (motion vector) with respect to the frame on which it depends

Example

Suppose we have 4 frames to compress. The first of these will necessarily be a key-frame (I-Frame), while we want the next two to be B-Frames (which are generally 1/4 the size of the corresponding P-Frame). The last one must necessarily be a P-Frame, since the B-Frames need after them something to be derived from. In sequence we would have:

n° fotogramma	1	2	3	4
tipo di fotogramma:	I	B	B	P

Frames will be stored within the movie like this:

n° fotogramma	1	4	2	3
tipo di fotogramma:	I	P	B	B

Homework

Write a python program (using opencv) that:

- **Method 1:** Reads a video frame by frame and saves only the even-numbered frames in a folder.
- **Method 2:** Apply in the even frames the negative function and to the odd frames any other operation. Show (by code) and save the final video
- **Method 3:** Extract the I, P and B frames and save them in 3 different folders

Basic Concepts of Digital Video

Luca Guarnera, Ph.D.

Research Fellow

luca.guarnera@unict.it

University of Catania

Dipartimento di Economia e Impresa