Data Compression
Basics of Video Coding

Dr. Mona Soliman

IT Dept.

DIGITAL VIDEO

- Digital video comprises a series of digital images (*frames*) displayed in rapid succession at a constant rate.
- We measure the rate at which frames are displayed in *frames per second (FPS)*. Since every frame is a digital image it comprises a raster of *pixels*.
- If it has a width of W pixels and a height of H pixels we say that the *frame size* is WxH.
- Pixels have only one property, their color. The *color* of a pixel is represented by a fixed number of bits.

Example :

video can have a duration (T) of 1 hour (3600sec), a frame size of 640x480 (WxH) at a color depth of 24bits and a frame rate of 25fps. This example video has the following properties:

- **pixels per frame** = 640 * 480 = 307,200
- **bits per frame** = 307,200 * 24 = 7,372,800 = 7.37*Mbits*
- **bit rate (BR)** = 7.37 * 25 = 184.25***Mbits/sec***
- **video size (VS)**[3] = 184*Mbits/sec* * 3600sec = 662,400*Mbits* = 82,800*Mbytes* = 82.8***Gbytes***

Uncompressed video

APPLICATION OF VIDEO CODING

- Conversational Video
  - Video conferencing, video telephonic (WIRED/WIRLESS)
- Video streaming
  - Video on demand, HDTV

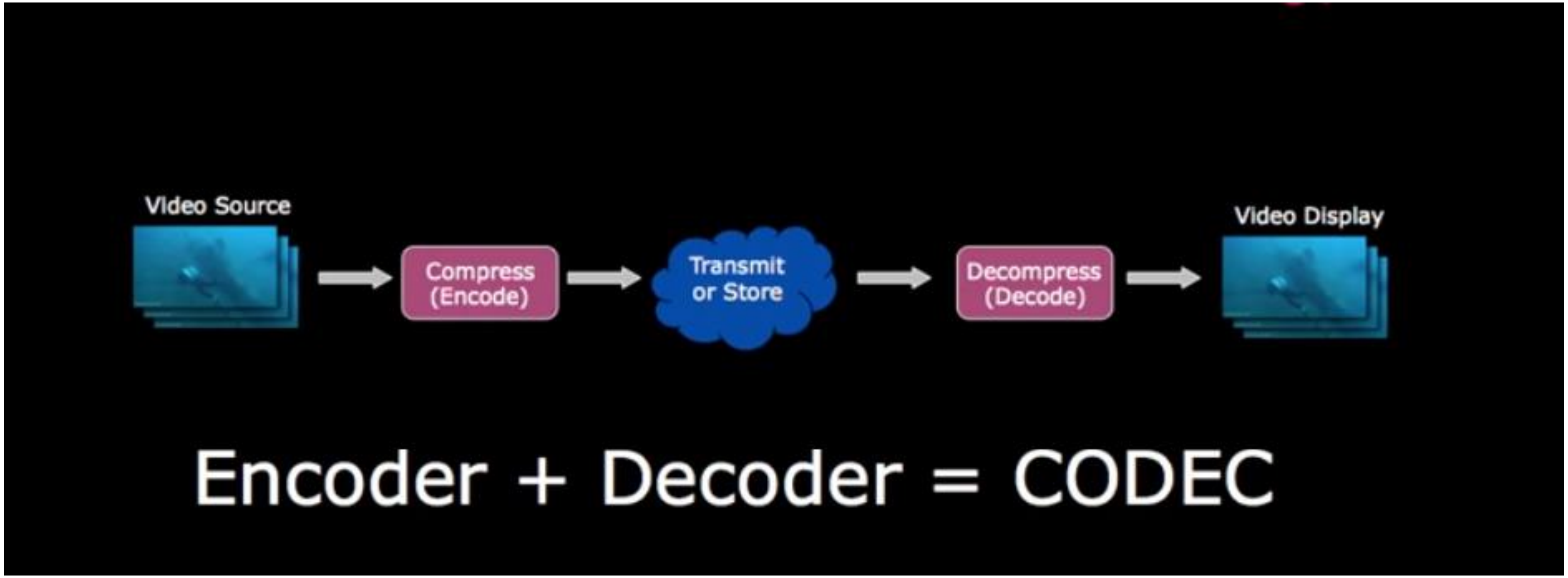
HOW TO COMPRESS A VIDEO ??

- Basic Idea of Video Compression:

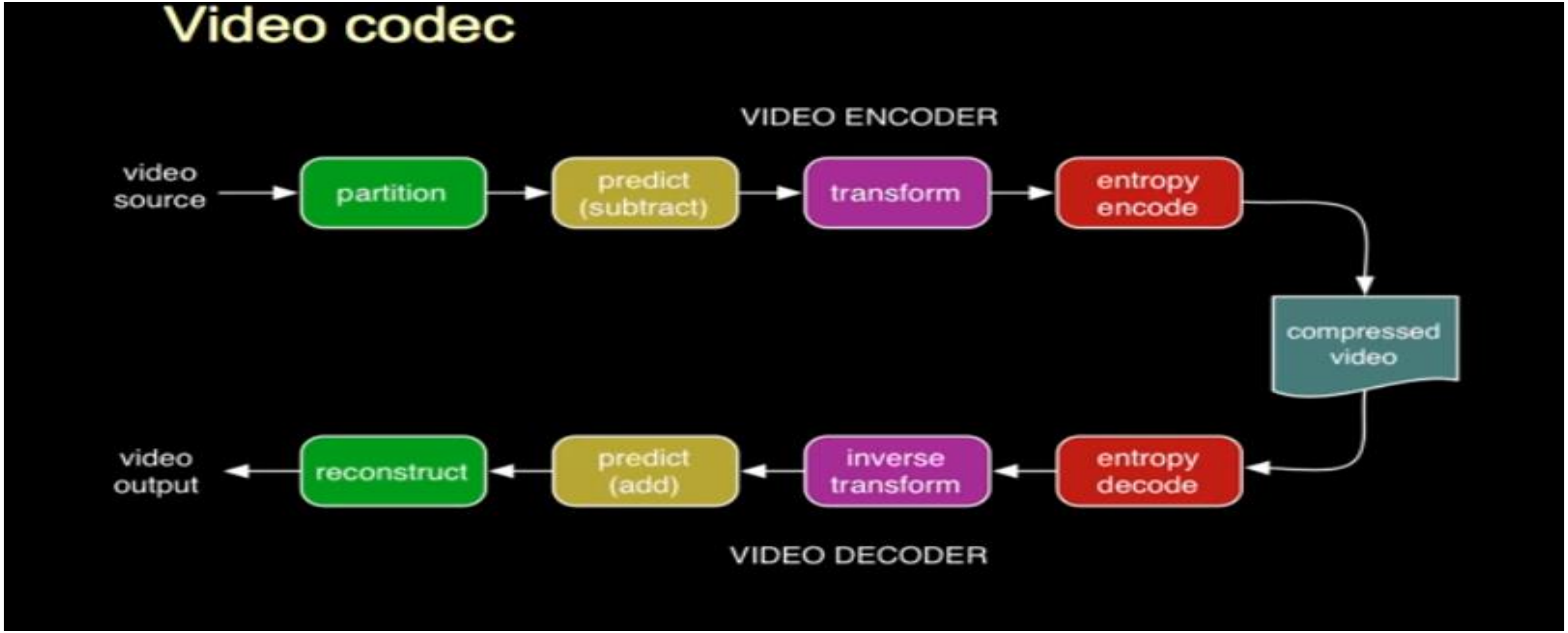
  Exploit the fact that adjacent frames are similar.

- Spatial redundancy removal — intraframe coding (JPEG) NOT ENOUGH BY ITSELF?
- Temporal — greater compression by noting the temporal coherence/incoherence over frames. Essentially we note the difference between frames.
- Spatial and temporal redundancy removal.
  - Temporal redundancy is exploited by predictive coding
  - Spatial redundancy is exploited by transform domain coding

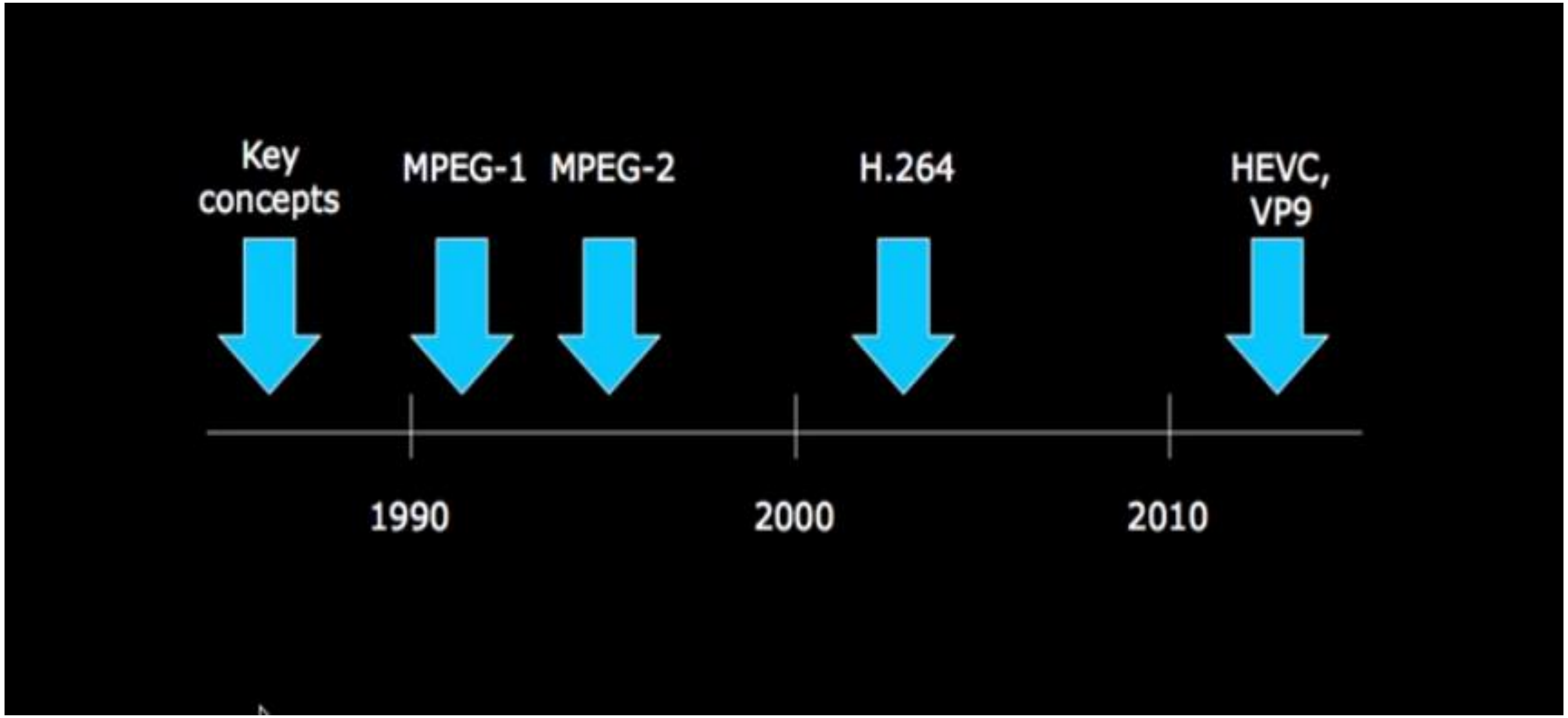
VIDEO ENCODING/DECODING



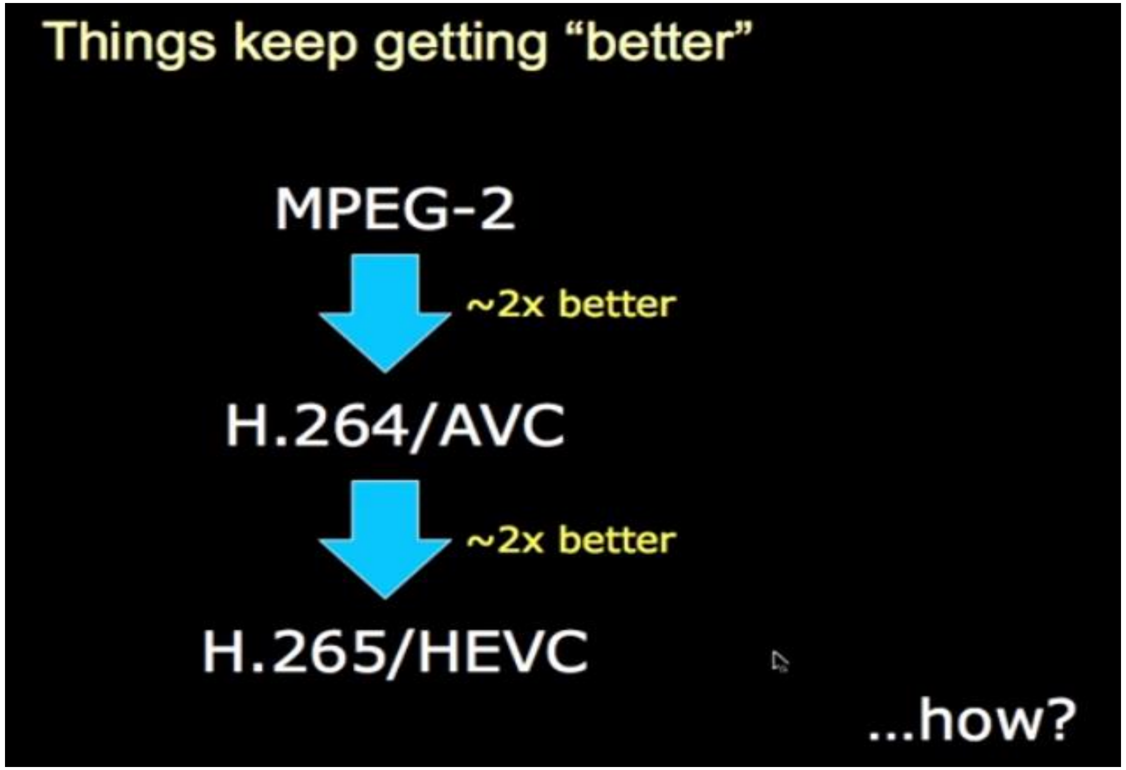
VIDEO CODEC PROCESS

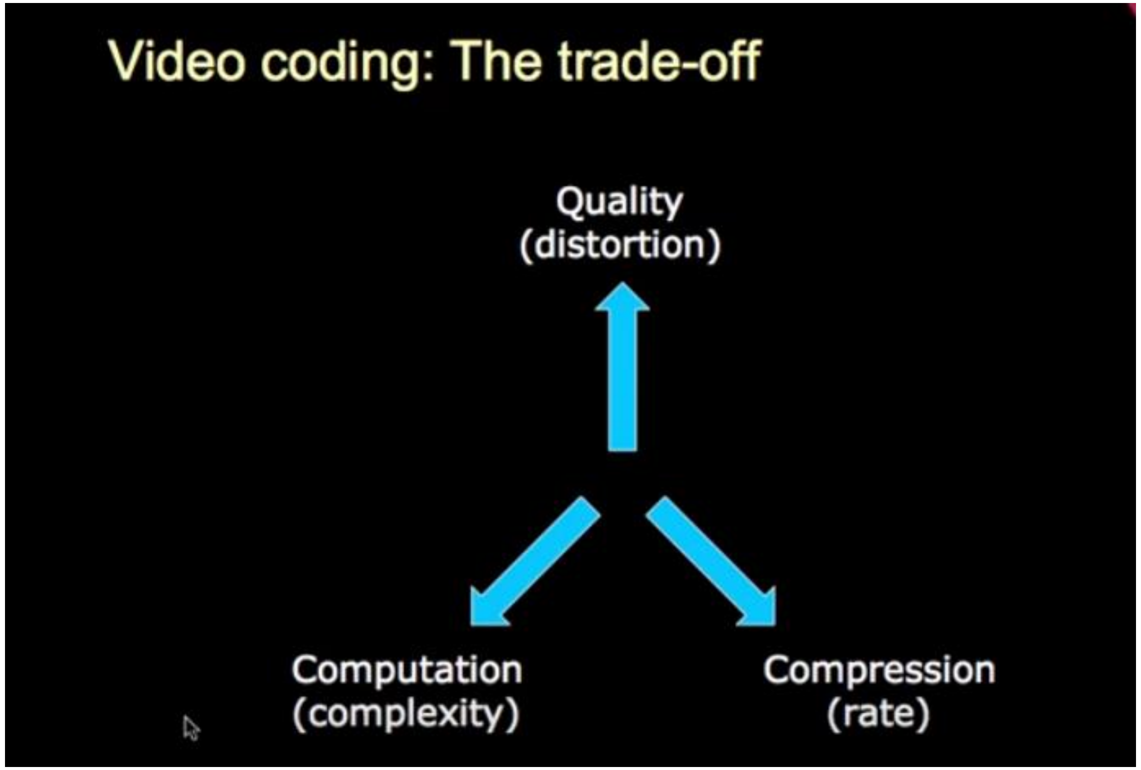


VIDEO CODEC STANDERS

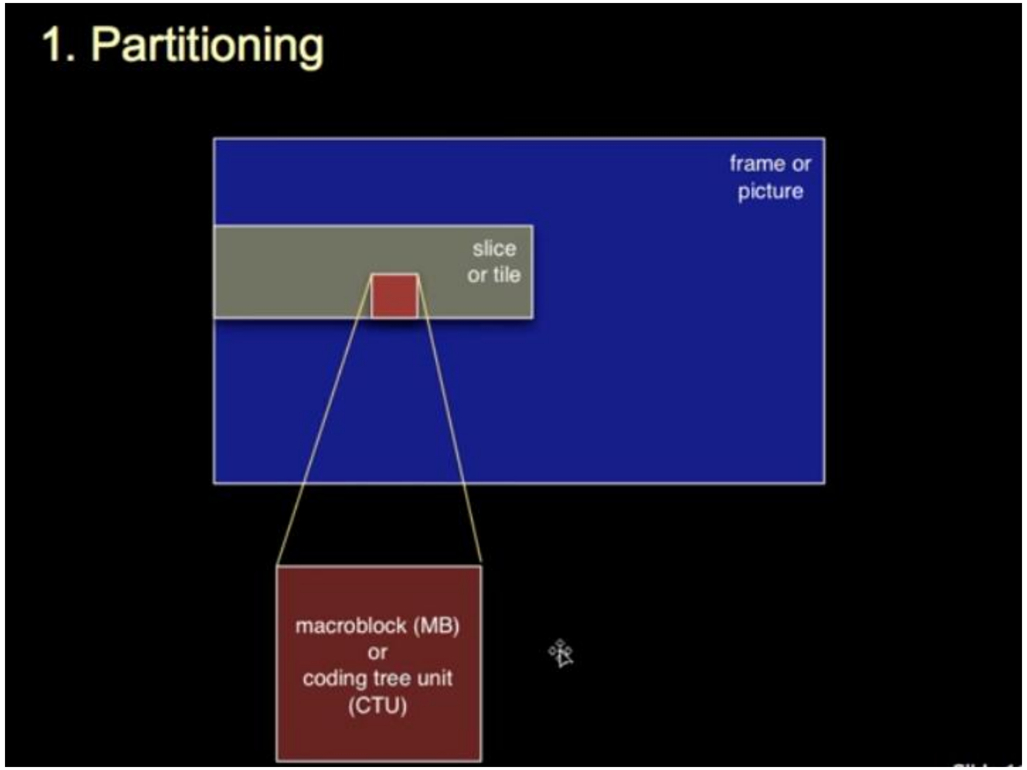


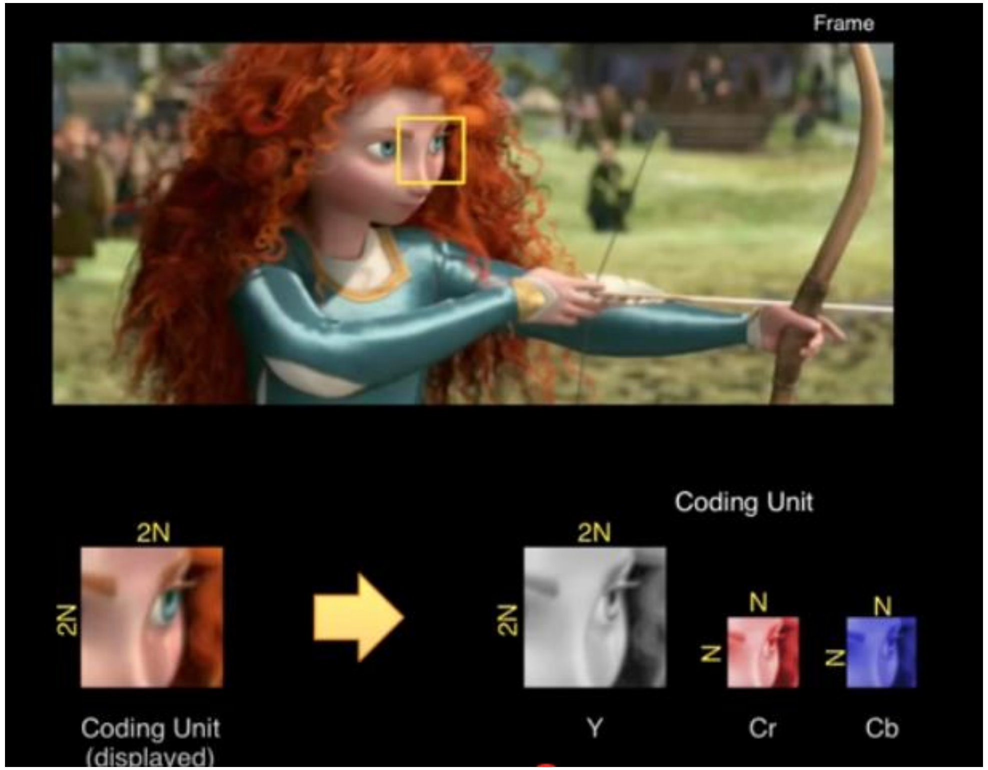
VIDEO CODEC CONT.



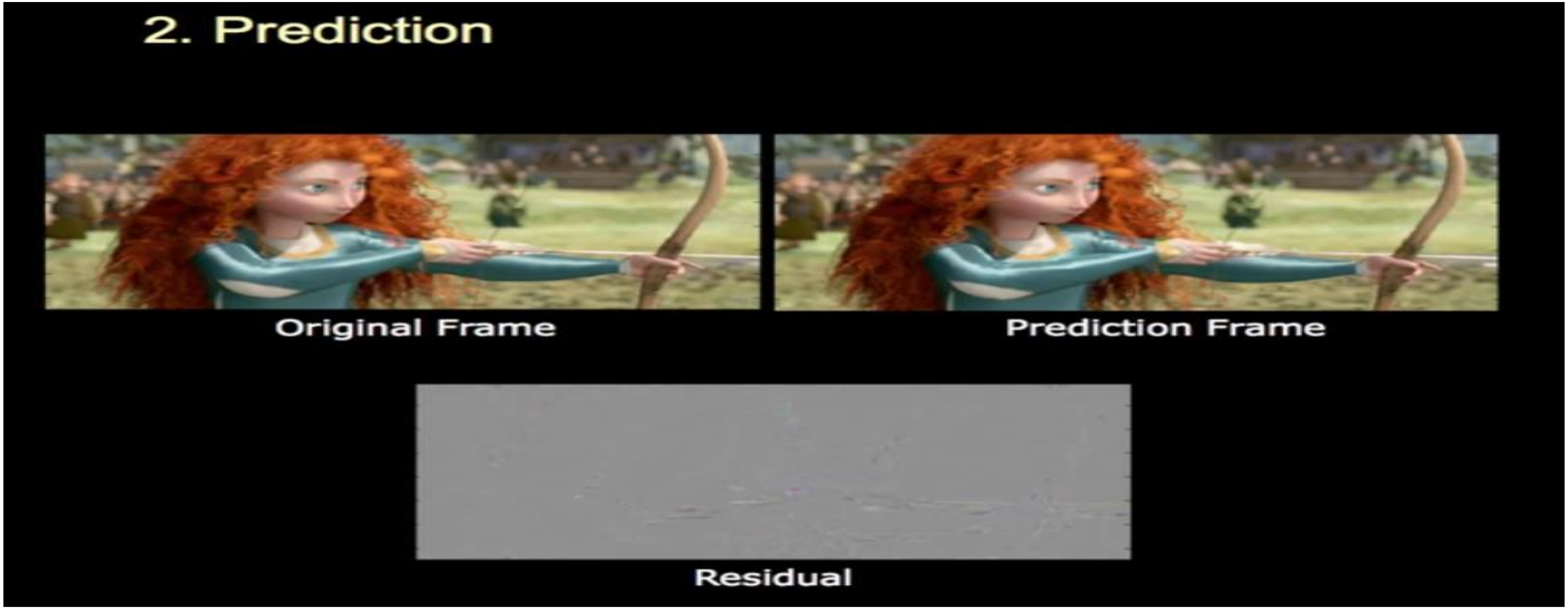


VIDEO CODEC PROCESS
1.PARTITION

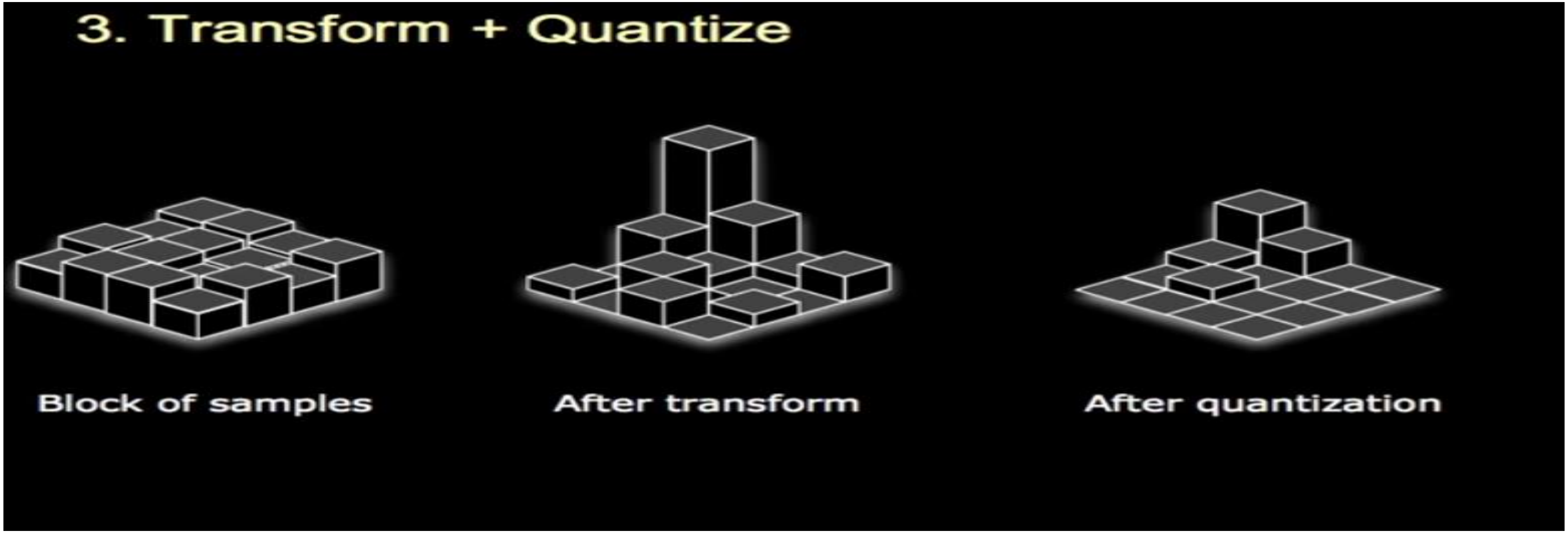




VIDEO CODEC PROCESS
2-PREDECTION

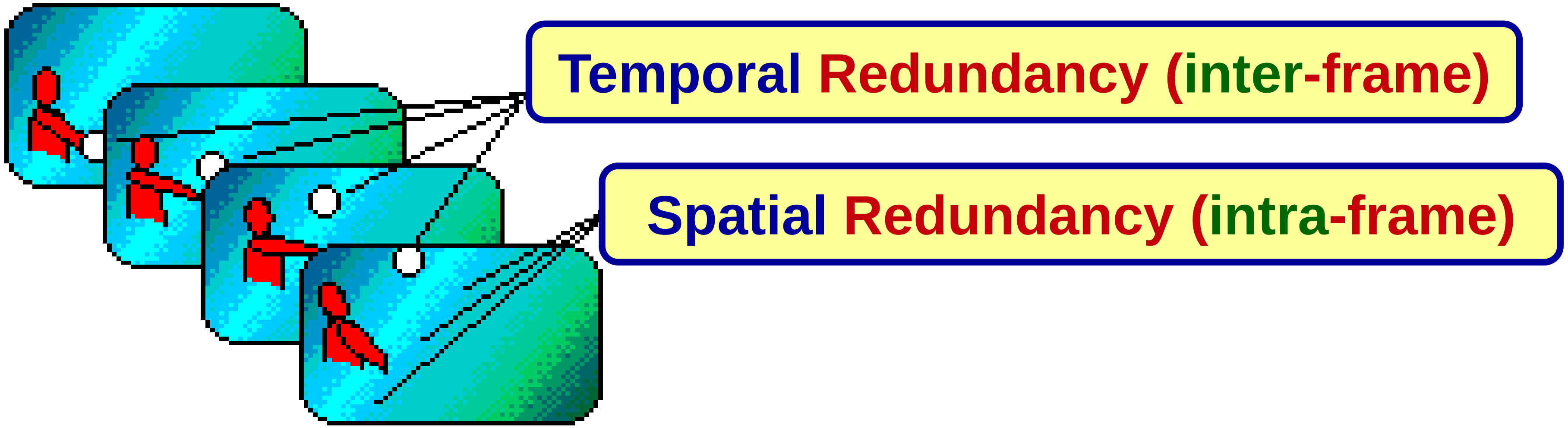


VIDEO CODEC PROCESS
3-TRANSFORMATION

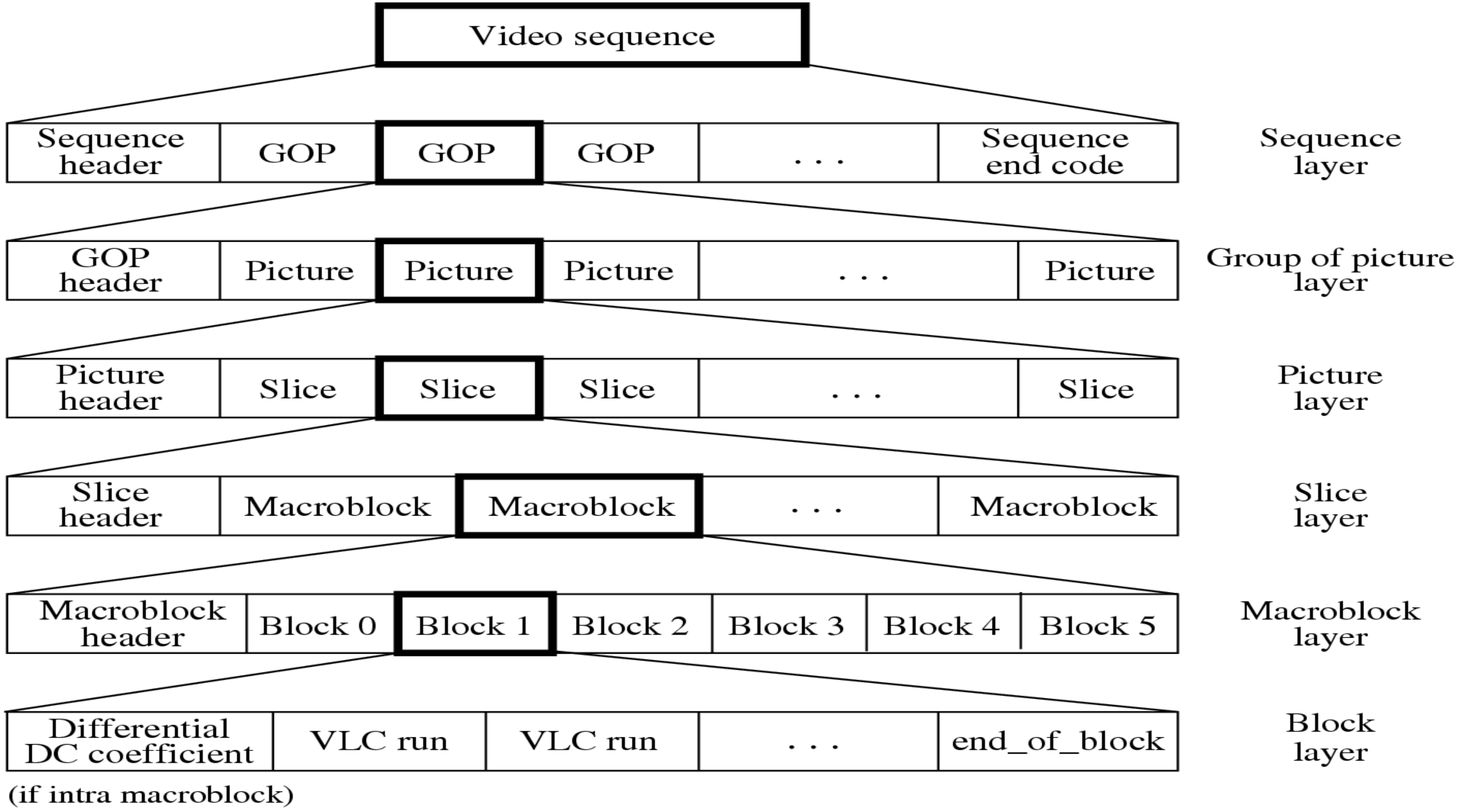


ELIMINATING REDUNDANCIES

- **Spatial Redundancy**
- Pixels are replicated within a single frame of video
- **Temporal Redundancy**
- Consecutive frames of video display images of the same scene



LAYERS OF MPEEG-1 VIDEO SEQUENCE



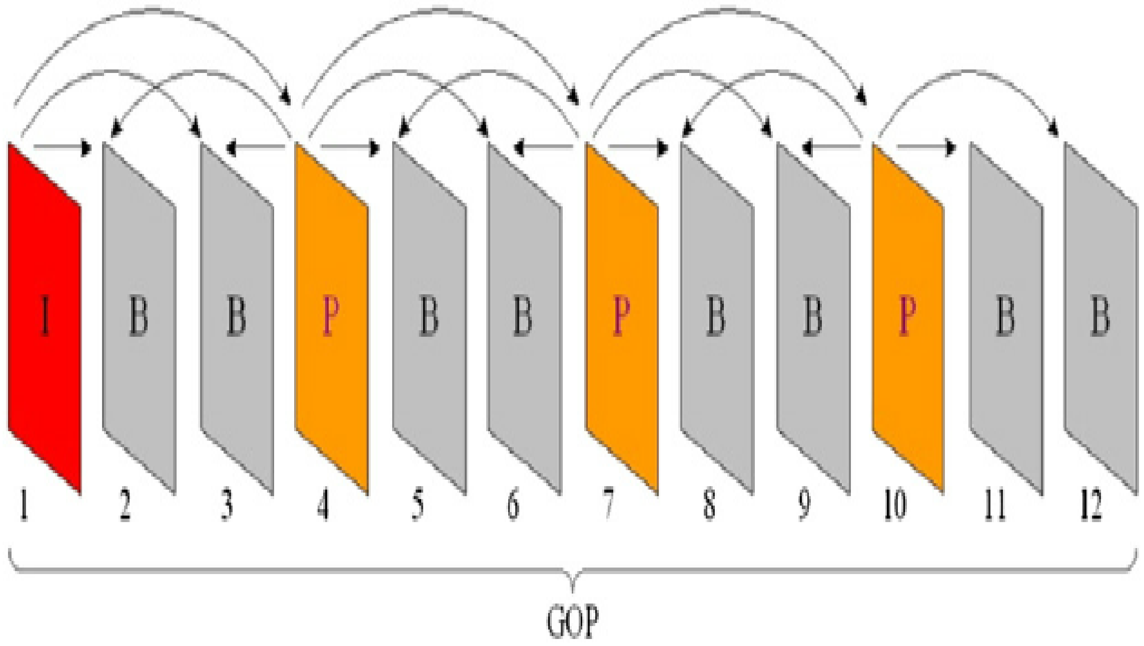
GOP

- GOP - Begins with an "I" frame, followed usually by a number of "P" and "B" frames

  - each GOP is independant: all frames needed for predictions are contained within each GOP

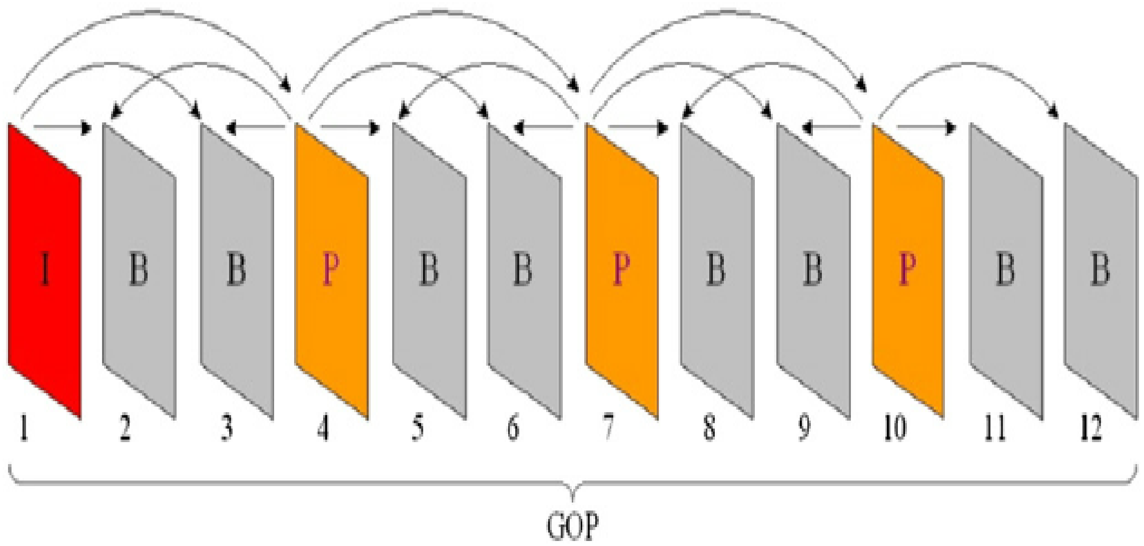
  - GOP's can be as small as a single I frame, or as large as desired, but usually no more than 15 frames in length.

  - the longer the GOP, the more efficient, but less robust the coding

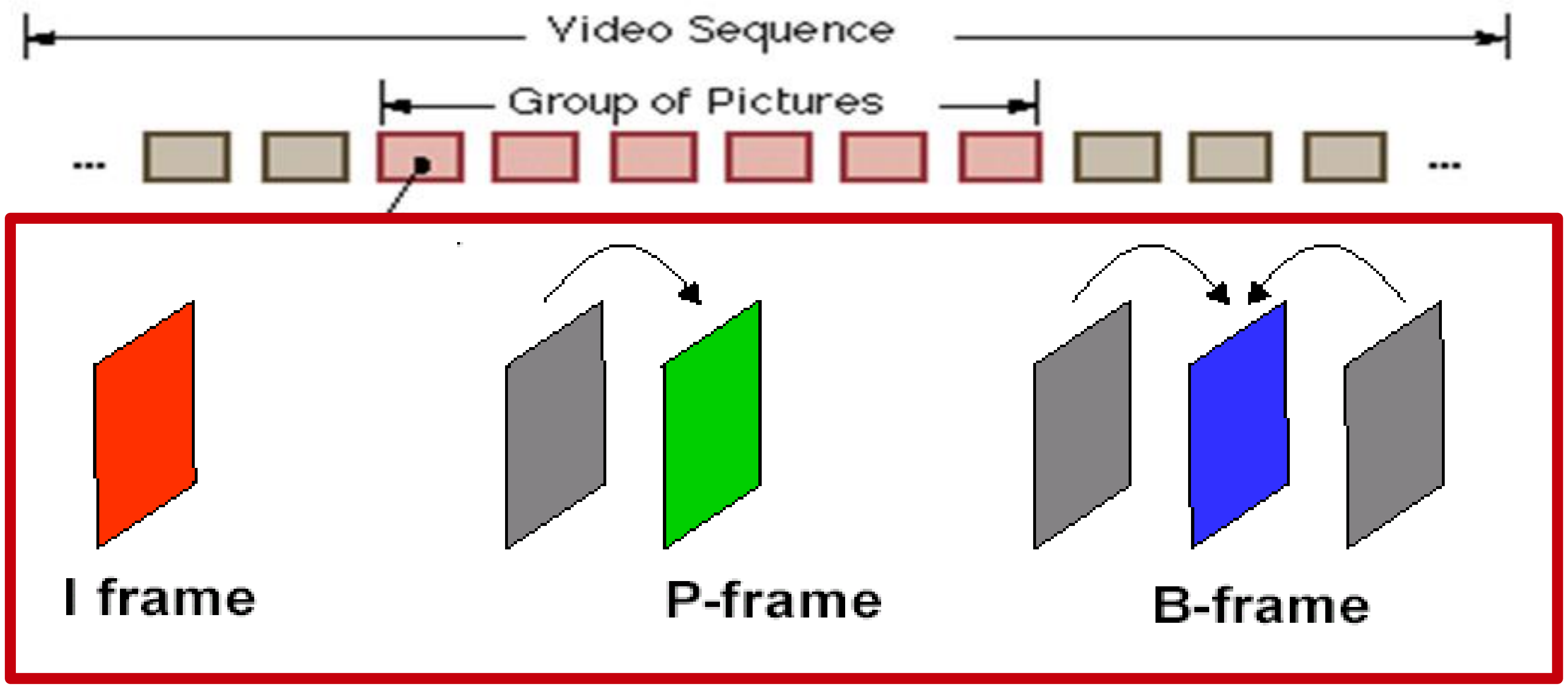


GOP

- The GOP structure is often referred by two numbers (M,N). The first number tells the distance between two anchor frames (I or P). The second one tells the distance between two full images (I-frames): it is the GOP size.
- For the example M=3, N=12, the GOP structure is IBBPBBPBBPBBI.
- For example, in a sequence with pattern IBBBBPBBBBPBBBBI, the GOP size is equal to 15 (length between two I frames) and distance between two anchor frames (**M value**) is 5



MPEG-2 VIDEO SEQUENCE



I-frames: can be reconstructed without any reference to other frames

P-frames: forward predicted from last I-frame or P-frames

B-frames: forward and backward predicted (Bi-directional) from I-frame and P-frames

GROUP OF PICTURE
(I, P, B FRAMES)

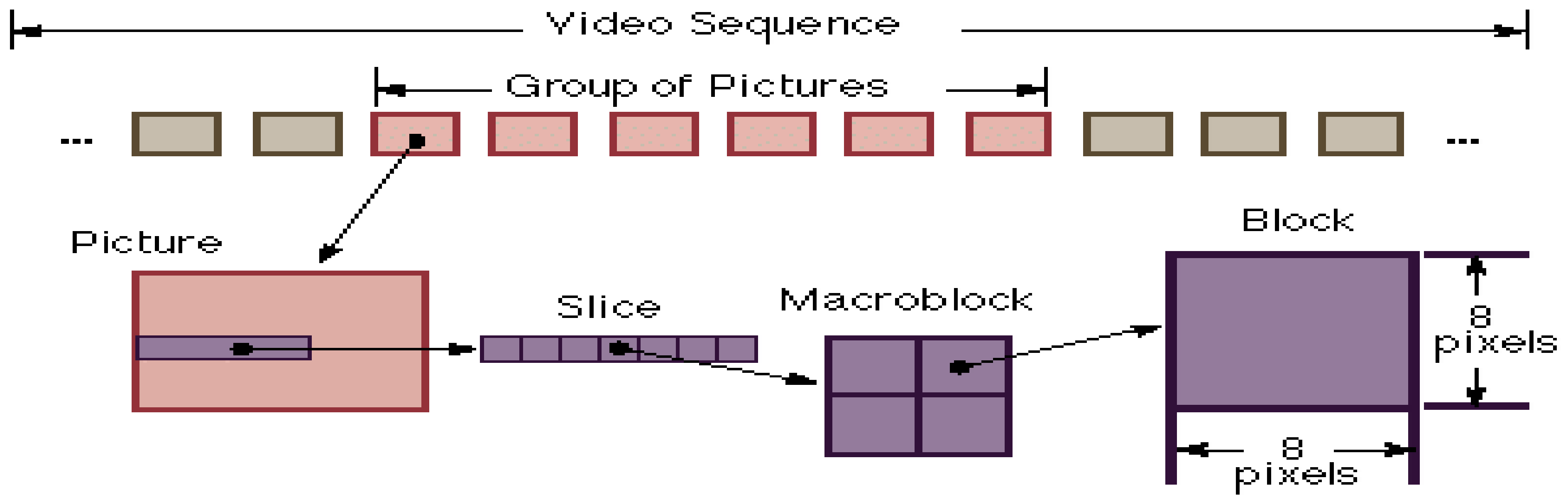
- **I-frame (Intracoded Frame)**

Coded in one frame such as DCT. This type of frame do not need previous frame.(compressed spatially)

- **P-frame (Predictive Frame)**
  - One directional motion prediction from a previous frame
  - The reference can be either I-frame or P-frame
  - Generally referred to as inter-frame
  - Compress temporally relative to I or P frame ,then compressed spatially

- **B-frame (Bi-directional predictive frame)**
  - Bi-directional motion prediction from a previous or future frame
  - The reference can be either I-frame or P-frame
  - Generally referred to as inter-frame
  - Compress temporally relative to preceding I and following I/P frame ,then compressed spatially

MPEG CODING
PARTINING



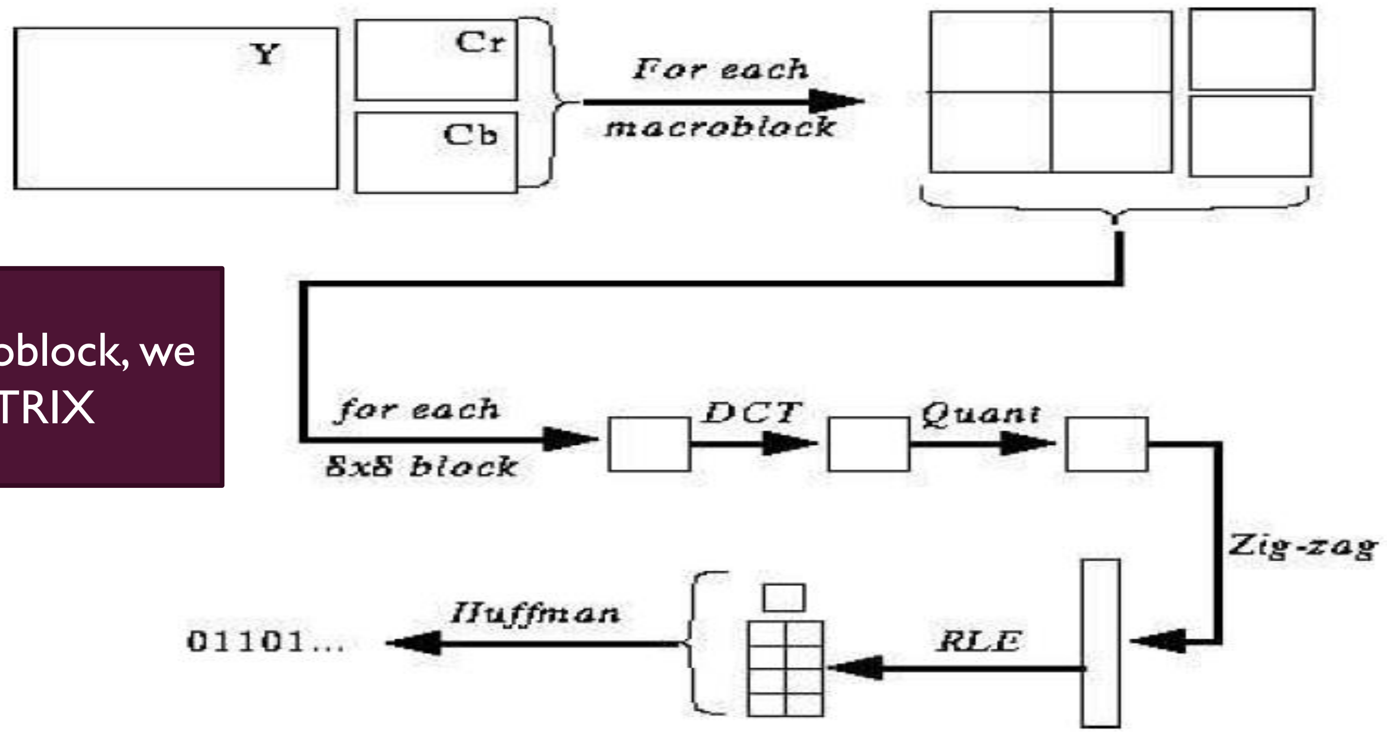
MPEG CODING
I- (I FRAME) INTRADRAME CODING (I FRAME)

- Various lossless and lossy compression techniques use — like JPEG.
- Compression contained only within the current frame.

  “Simpler coding — not enough by itself for high compression”

MPEG CODING
INTRADRAME CODING (I FRAME) CONT.

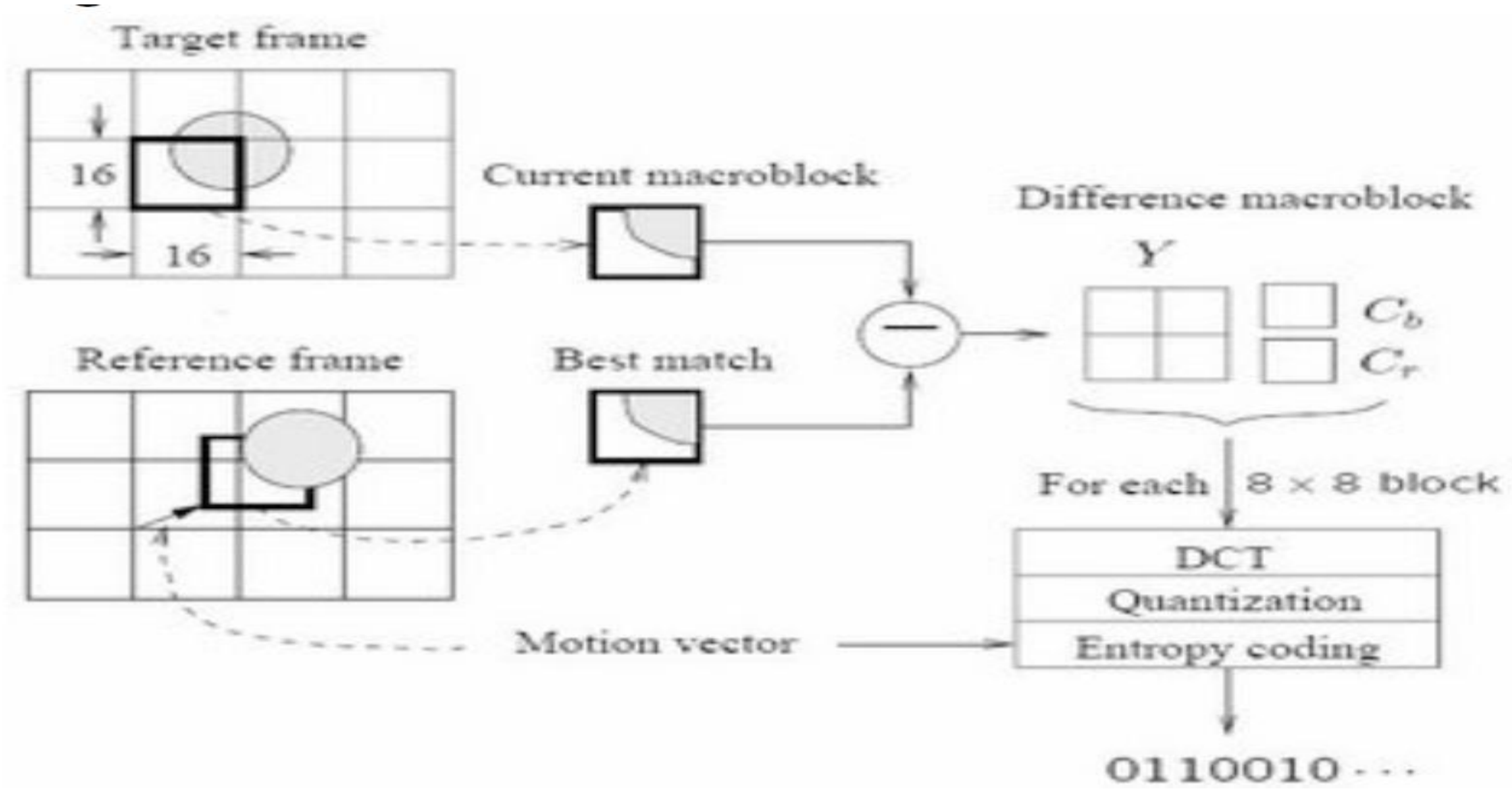
Intra-frame coding is very similar to JPEG:



For each 16x16 macroblock, we have 6 DCT MATRIX

Quantization is by constant value for all DCT coefficients.
I.e., no quantization table as in JPEG.

MPEG CODING
INTER-FRAME (P-FRAME) CODING CONT.

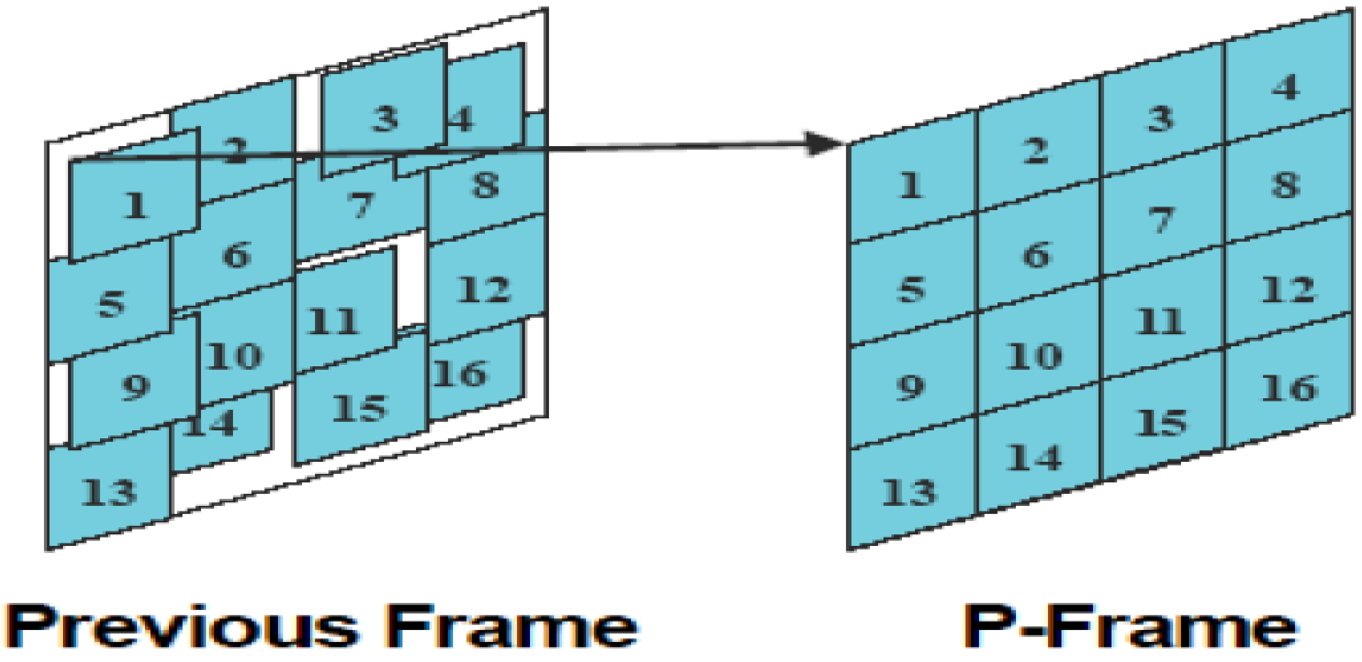


MPEG CODING
2- (P-FRAME) INTER-FRAME (P-FRAME) CODING

BASIC IDEA:

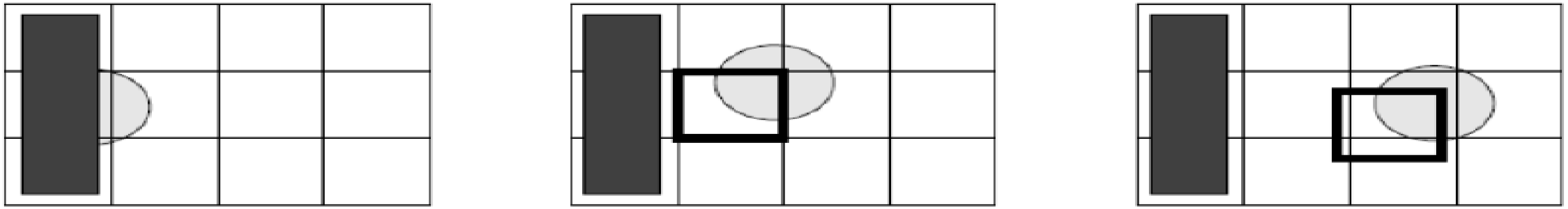
- Most consecutive frames within a sequence are very similar to the frames both before (and after) the frame of interest.
- Aim to exploit this redundancy.
- Use a technique known as block-based motion compensated prediction.
- Need to use motion estimation.
- P frame is encoded either as I frame (JPEG) or with predictive coding (from I frame)
- We select the coding that achieve less storage

MPEG CODING
INTER-FRAME (P-FRAME) CODING CONT.



MPEG CODING
3- (B FRAME) THE NEED FOR A BIDIRECTIONAL SEARCH

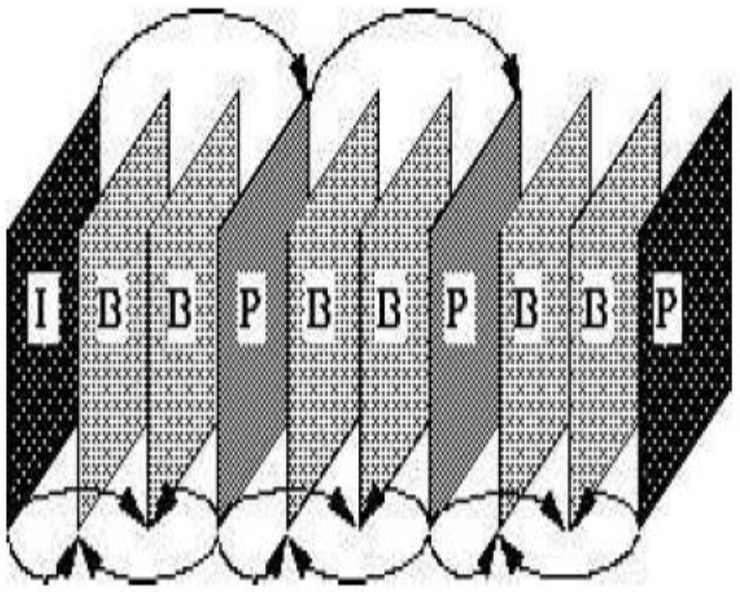
- The problem here is that many macroblocks need information that is **not** in the reference frame.
- For example:



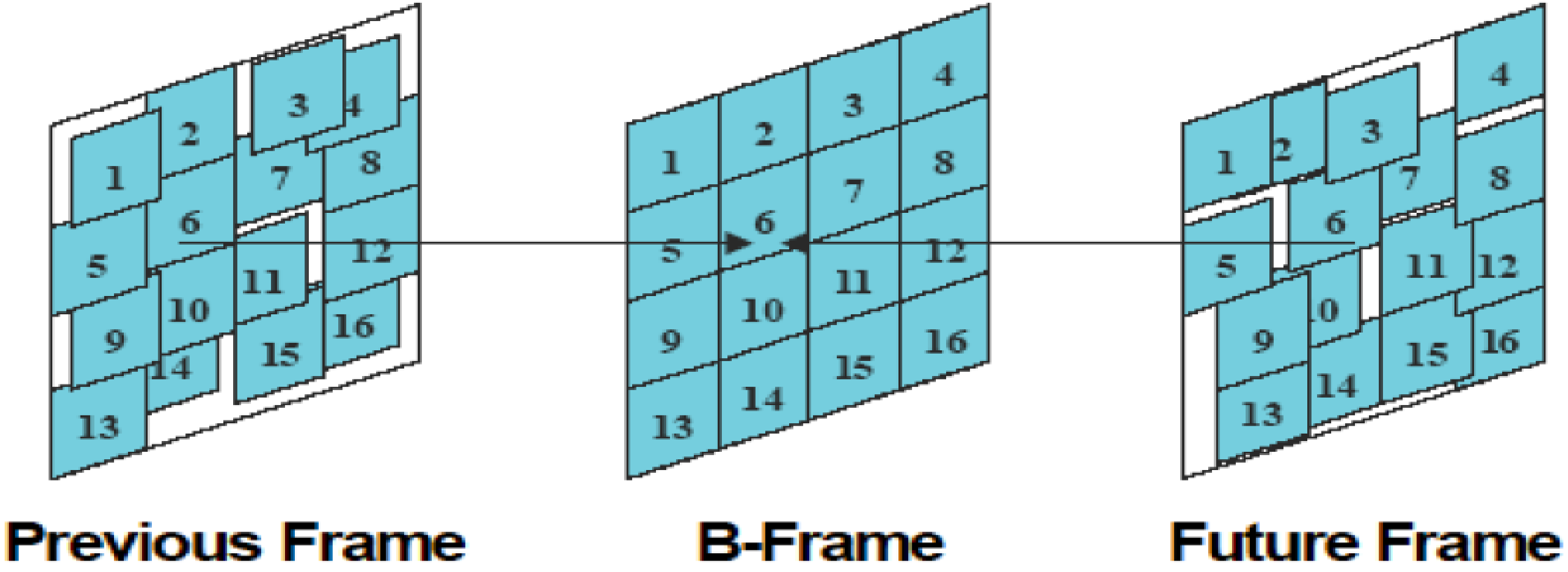
- Occlusion by objects affects differencing
- Difficult to track occluded objects *etc.*
- MPEG uses **forward/backward** interpolated prediction.

MPEG-B FRAME

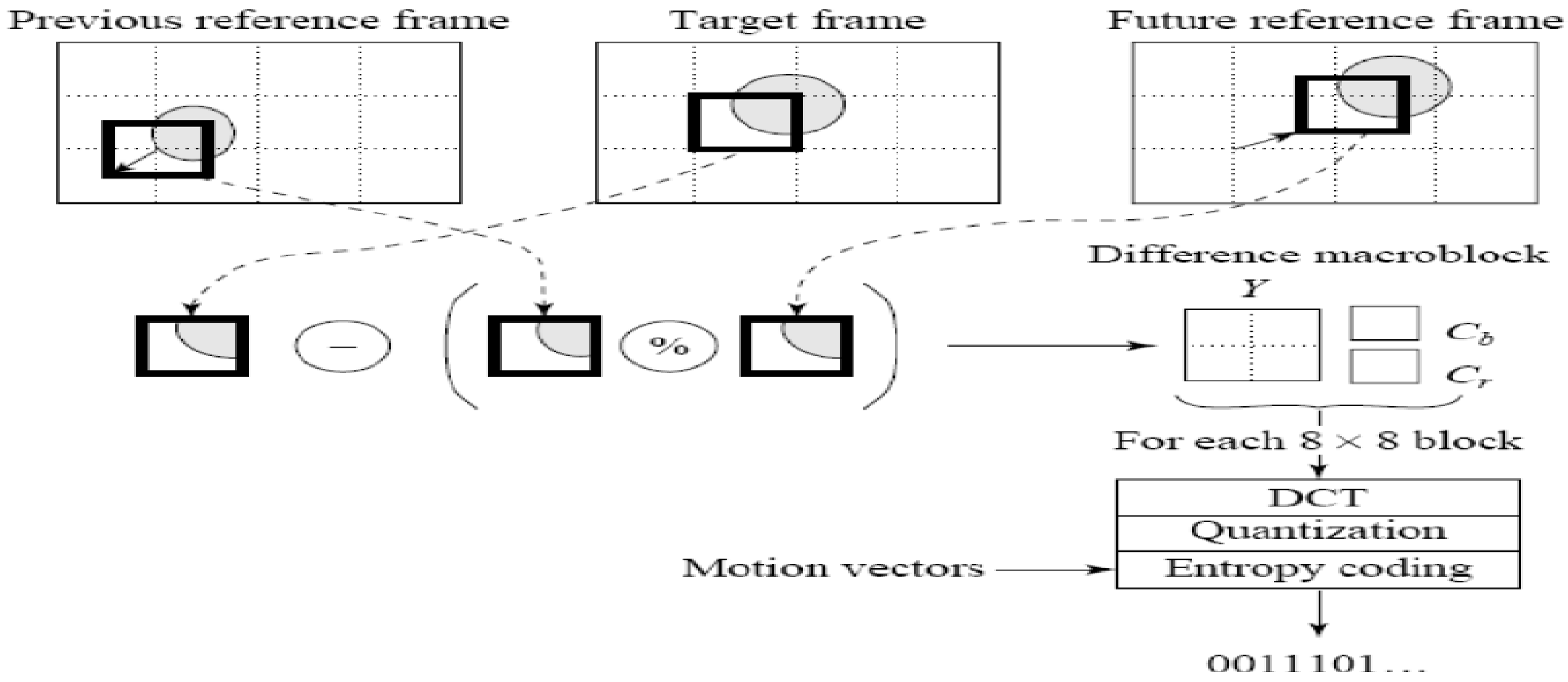
- The MPEG solution is to add a third frame type which is a bidirectional frame, or B-frame
- B-frames search for macroblock in past and future frames.
- Typical pattern is IBBPBBPBB IBBPBBPBB
- B frame is encoded in one of the following cases:
  - *AS I frame (JPEG)*
  - *Backward prediction*
  - *Forward prediction*
  - *Bi-direction prediction*
- We select the coding that achieve less storage
- B pictures are never used as prediction references.



BIDIRECTIONAL PREDICTION



BIDIRECTIONAL PREDICTION (BASED ON PERVIOUS AND FUTURE)



I P B PICTURE REORDERING

- Pictures are coded and decoded in a different order than they are displayed. This is due to bidirectional prediction for B pictures. See example below which illustrates reordering for a 12 picture long GOP.
- Source order and encoder input order:

I(1) B(2) B(3) P(4) B(5) B(6) P(7) B(8) B(9) P(10) B(11) B(12) I(13)

- Encoding order and order in the coded bitstream:

I(1) P(4) B(2) B(3) P(7) B(5) B(6) P(10) B(8) B(9) I(13) B(11) B(12)

- Decoder output order and display order (same as input):

I(1) B(2) B(3) P(4) B(5) B(6) P(7) B(8) B(9) P(10) B(11) B(12) I(13)