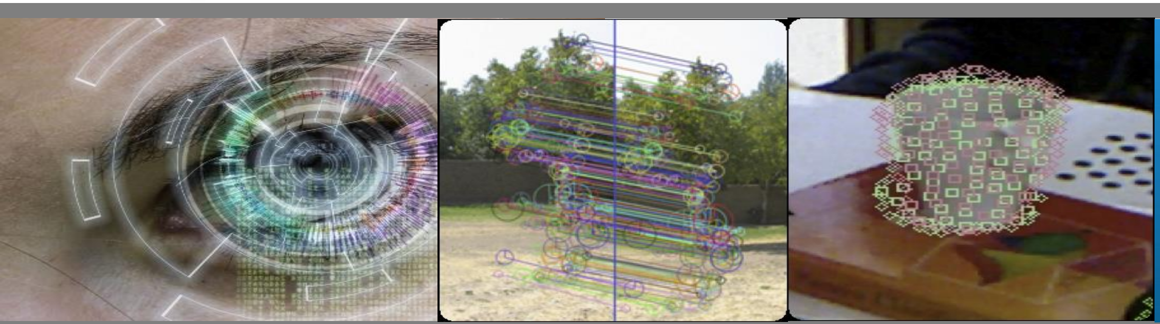
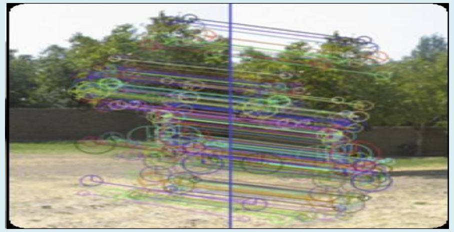
Advanced Multimedia Signal Processing
(#4: Video Model)



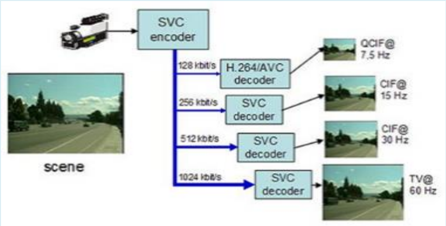
2017. Spring

Prof. Byung-Gyu Kim

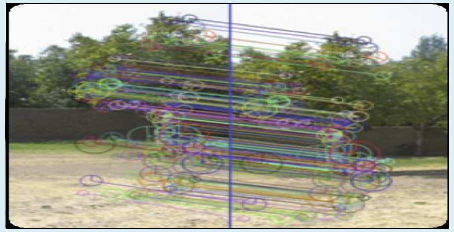
Dept. of IT Engineering, Sookmyung Women's University

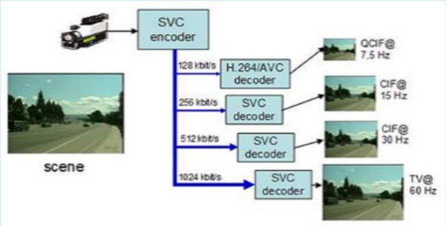



H.265
HEVC
High Effeciency Video Coding


SVC encoder
128 kbit/s
H.264/AVC decoder
256 kbit/s
SVC decoder
512 kbit/s
SVC decoder
1024 kbit/s
SVC decoder
QCIF@ 7.5 Hz
CIF@ 15 Hz
CIF@ 30 Hz
TV@ 60 Hz
scene

Contents

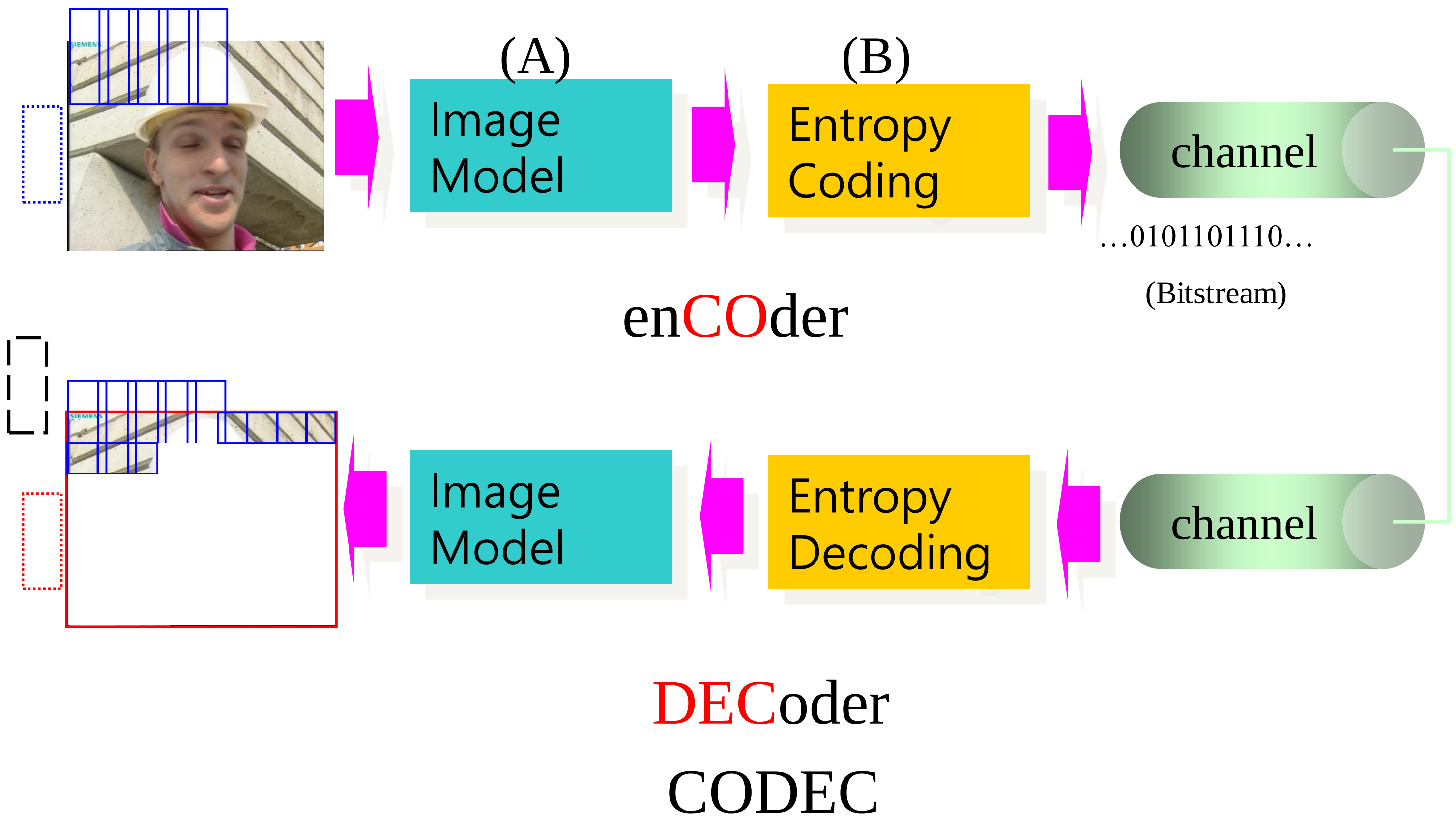




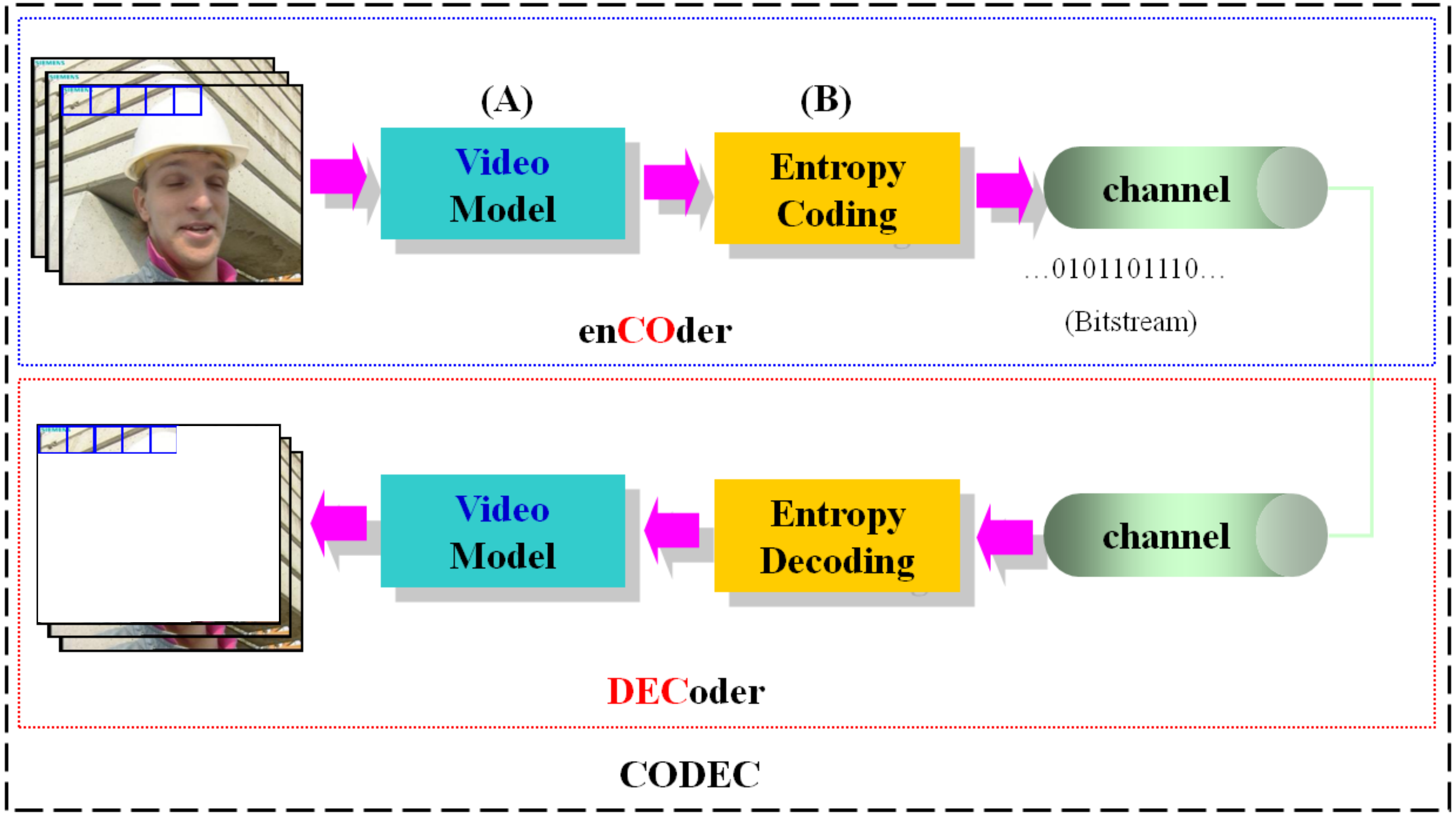
Contents

- **Video Model**
- **Detailed Processes for Video Coding**
- **Video Data Format**

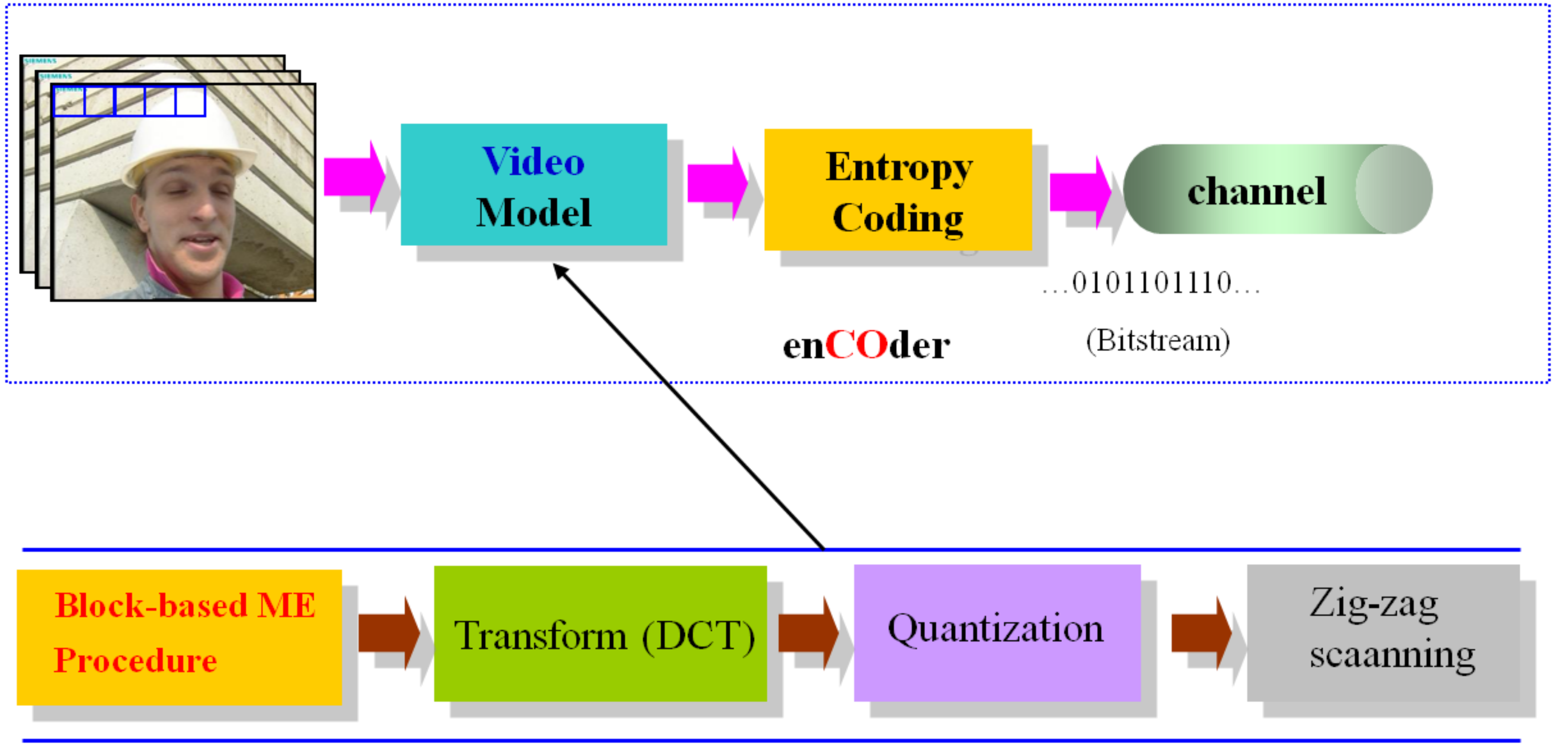
❖ Image Coding Model: Block unit based processing



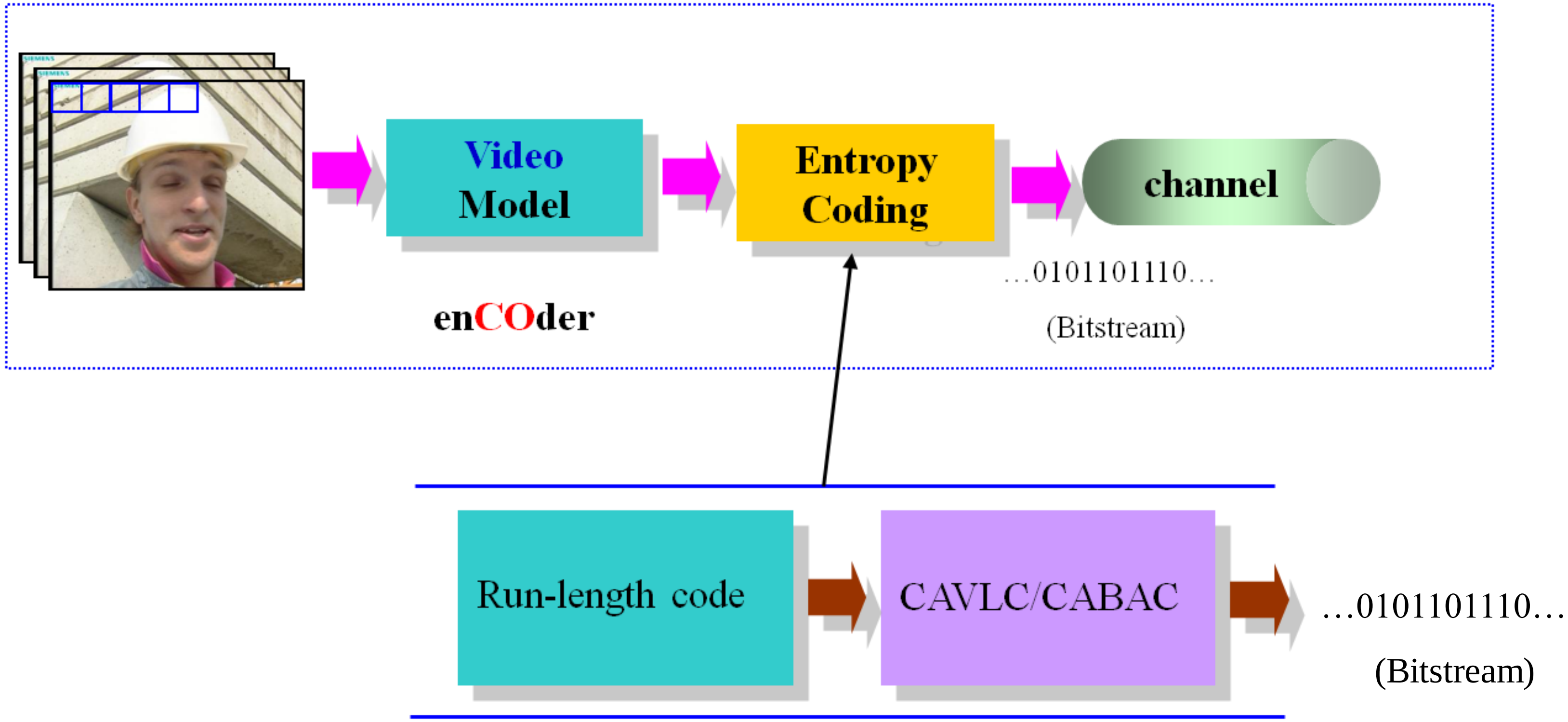
❖ Video Coding Model: Block unit based processing

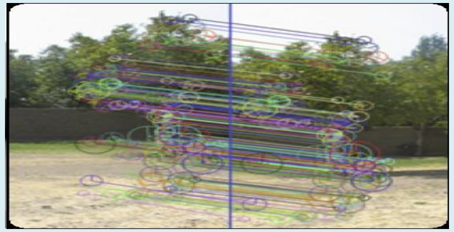


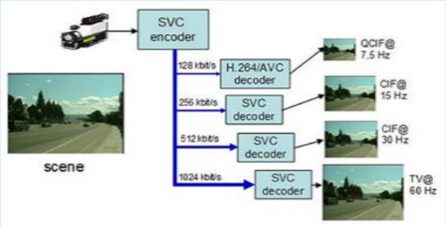
❖ Video Model (A):



- Entropy Coding (B):







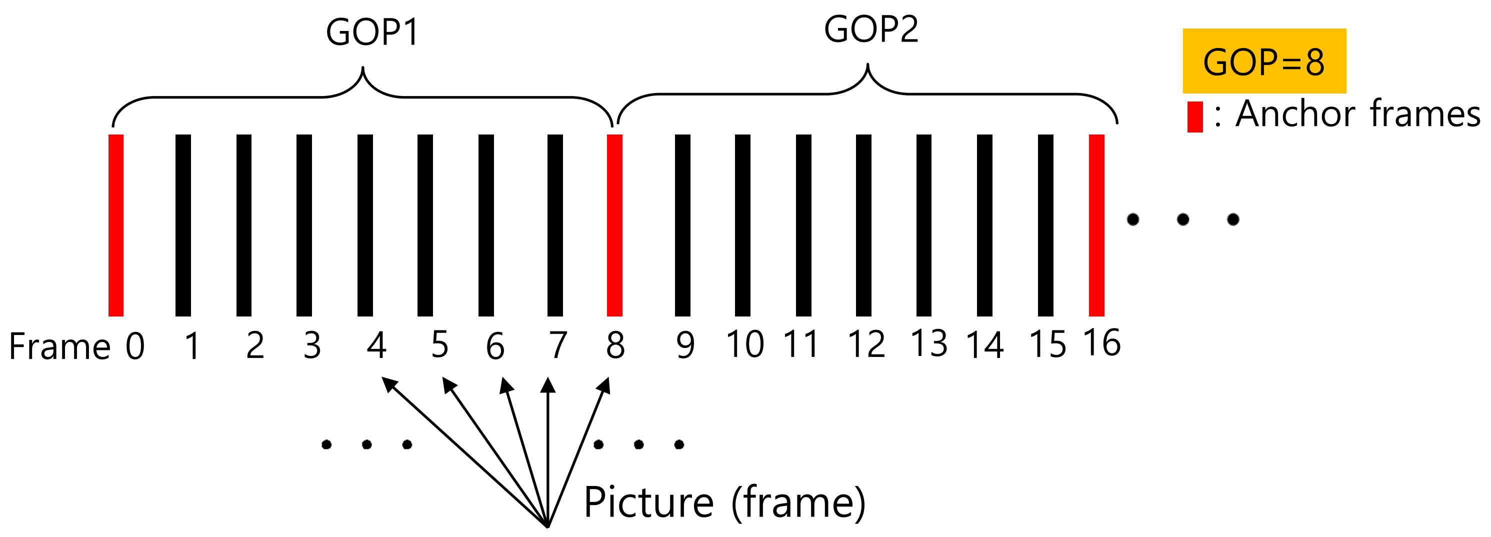
Contents

- Video Model
- Detailed Processes for Video Coding
- Video Data Format

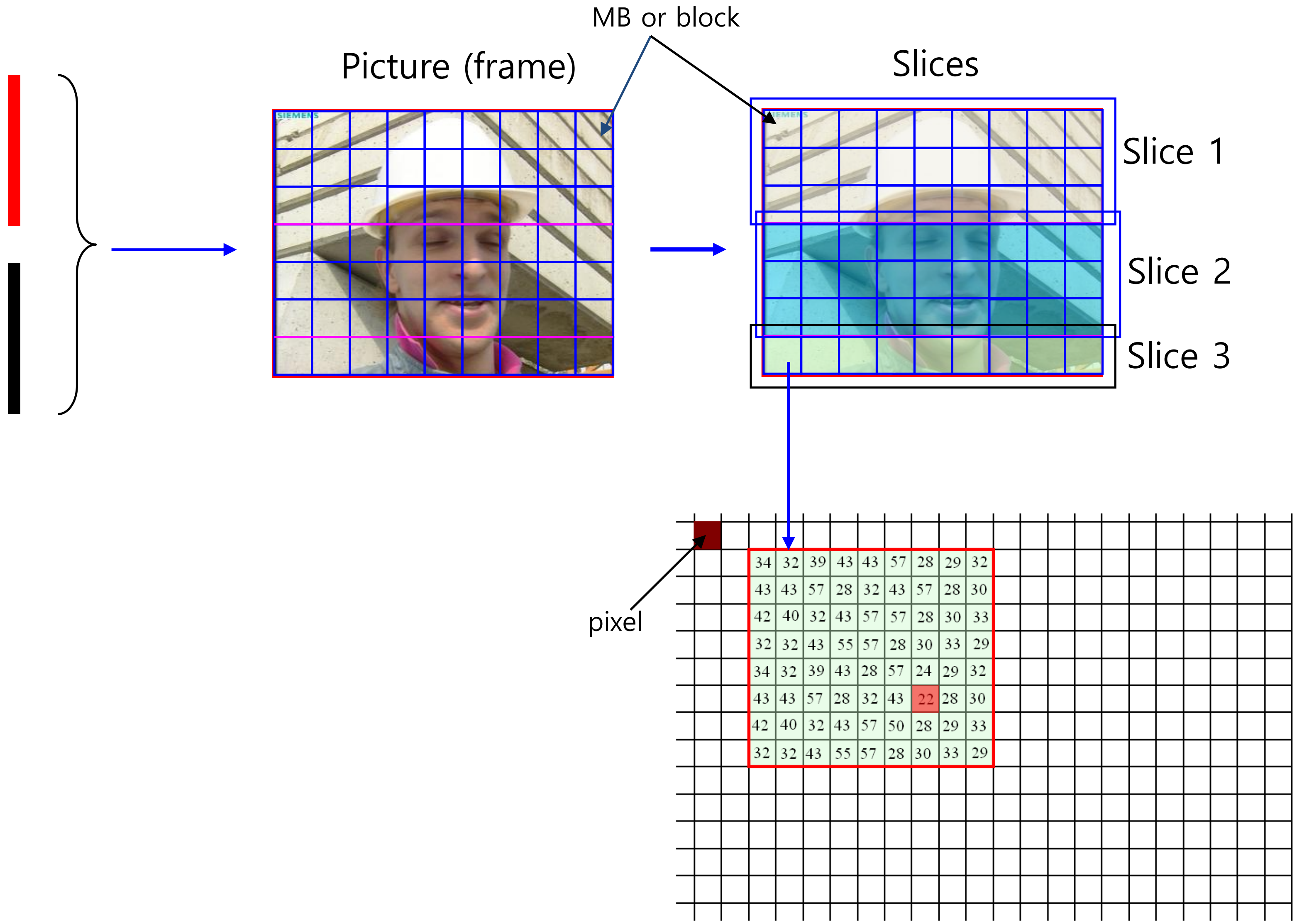
Detailed Processes for Video Coding (0)

❖ Processing Hierarchy of Video Encoding Scheme

- Video sequence: all pictures (frames)-many GOPs.
  - – Fixed length of GOP size in a video.
  - • Group of Picture (GOP): some successive pictures.
    - » Based on **Anchor frames**
  - – Picture (Frame): group of several slices.
    - » Slice: group of successive MBs.
      - - Macro-block (MB): sub-blocks or pixels.

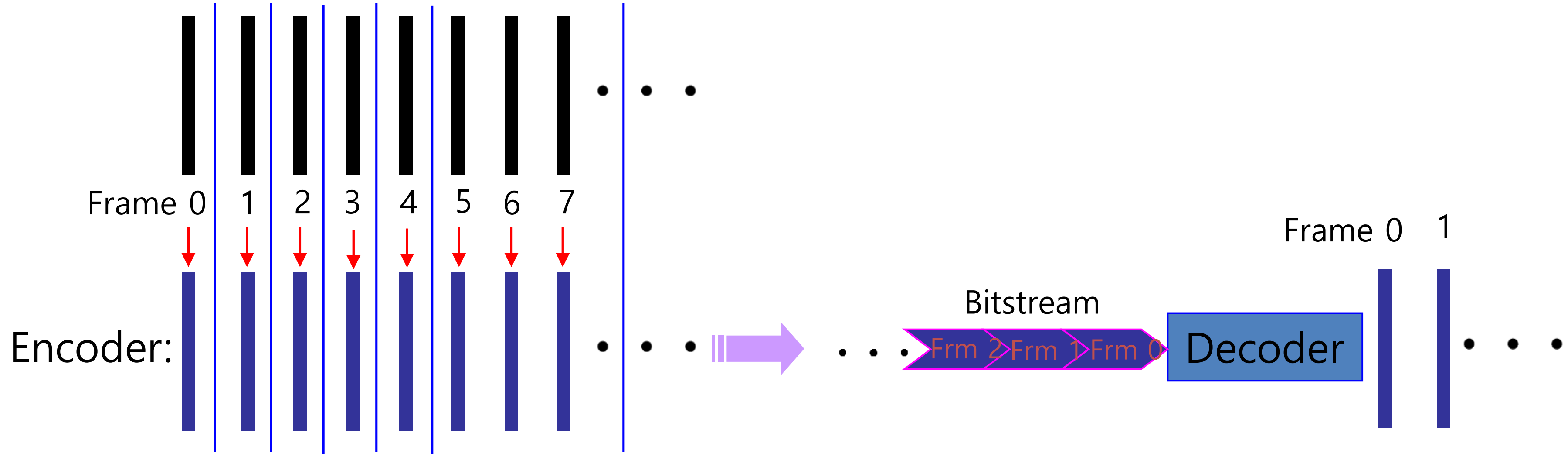


Detailed Processes for Video Coding (1)



❖ Encoding-Decoding Structures (frame unit)

- Video Coding Using *Still Image Coding* (like JPEG)
  - *The order of Frame unit*

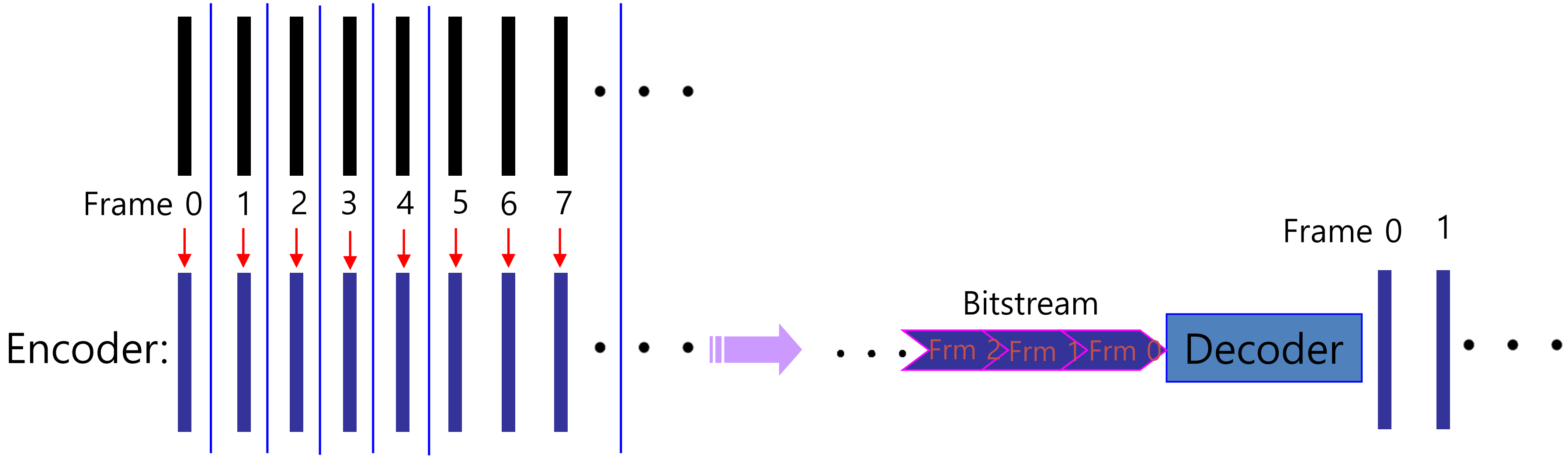


  - *The order of Block unit in frame* at Encoder/Decoder.

Encoder   Decoder

❖ Encoding-Decoding Structures (frame unit)

- Video Coding Using *Still Image Coding* (like JPEG)
  - *The order of Frame unit*



  - *The order of Block unit in frame* at Encoder/Decoder.

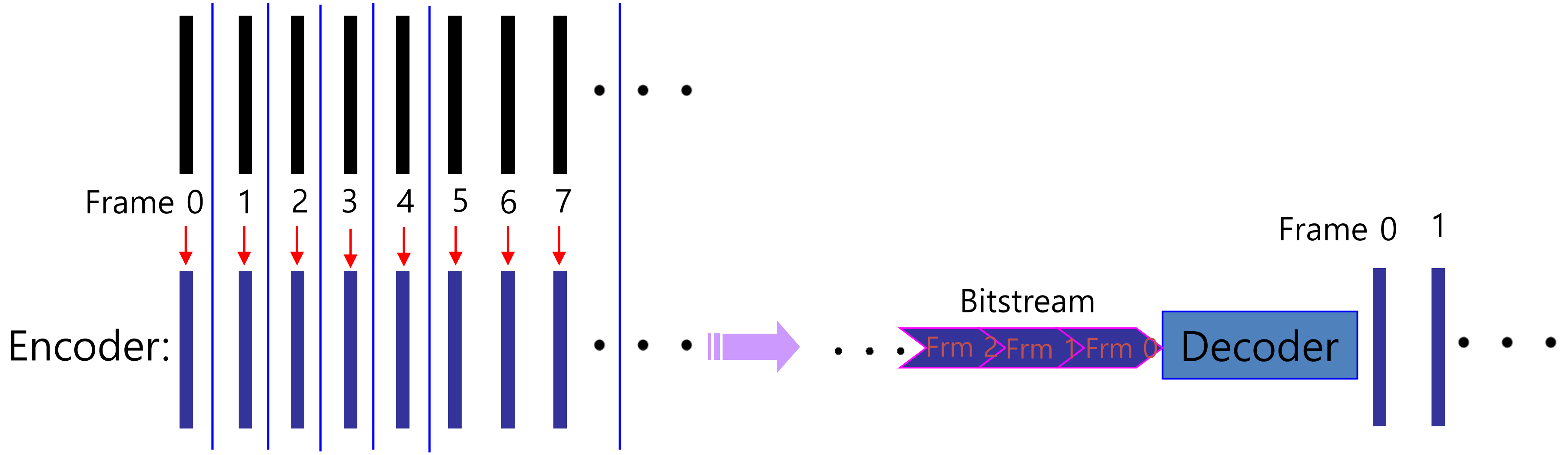
Encoder 

 Decoder

Detailed Processes for Video Coding (2)

❖ Encoding-Decoding Structures (frame unit)

- Video Coding Using *Still Image Coding* (like JPEG)
  - *The order of Frame unit*



  - *The order of Block unit in frame* at Encoder/Decoder.

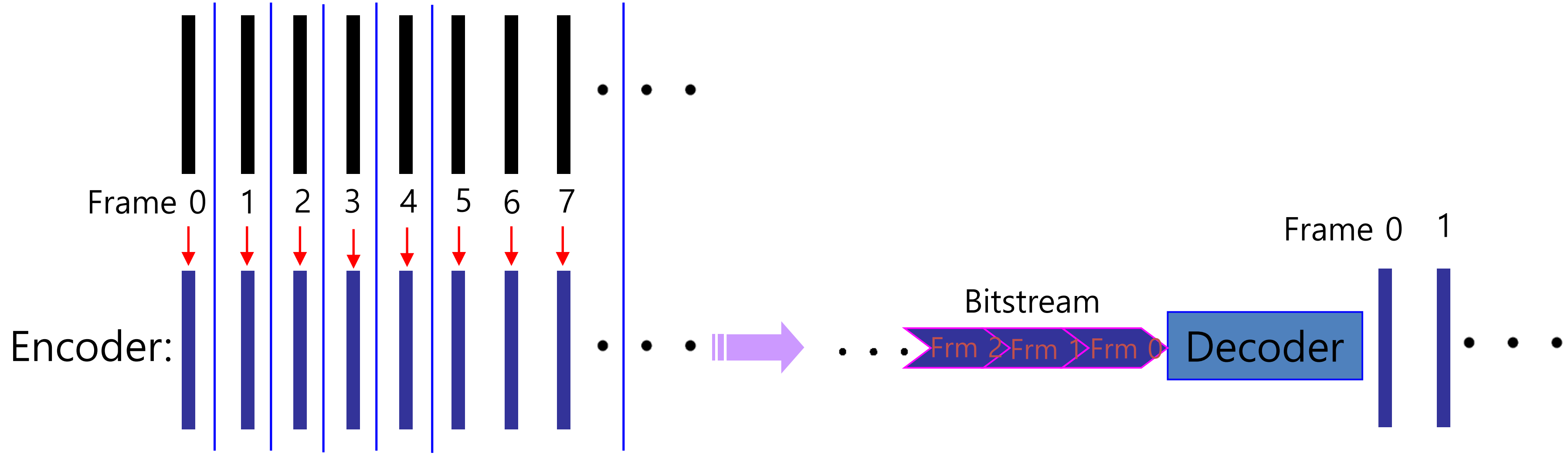
Encoder

Decoder

Detailed Processes for Video Coding (2)

❖ Encoding-Decoding Structures (frame unit)

- Video Coding Using *Still Image Coding* (like JPEG)
  - *The order of Frame unit*



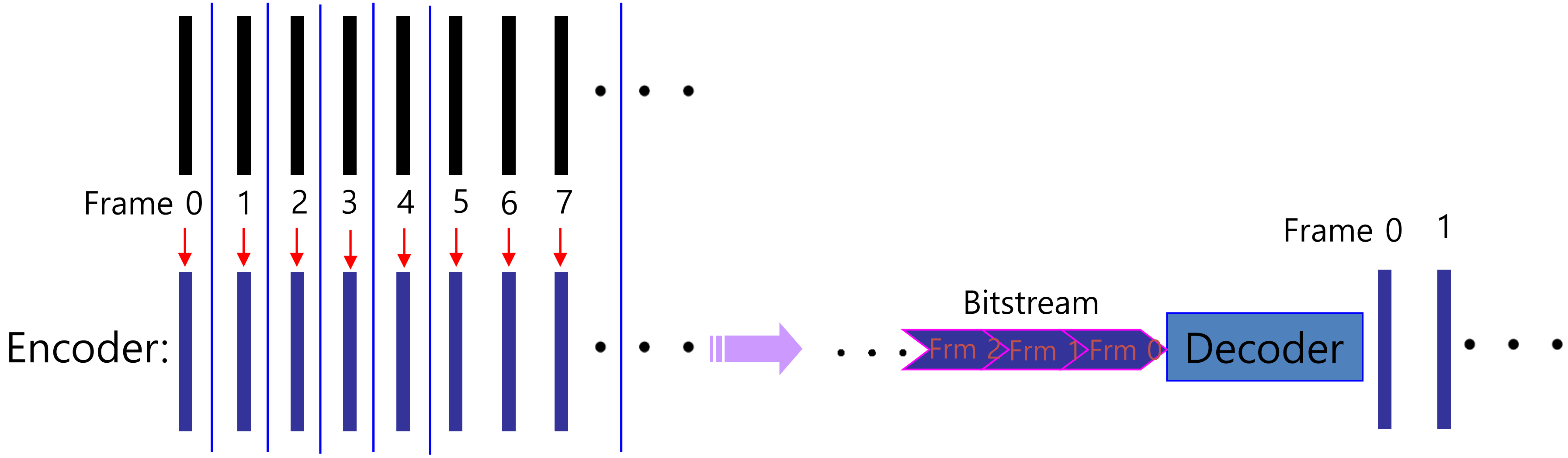
  - *The order of Block unit in frame* at Encoder/Decoder

Encoder

Decoder

❖ Encoding-Decoding Structures (frame unit)

- Video Coding Using *Still Image Coding* (like JPEG)
  - *The order of Frame unit*



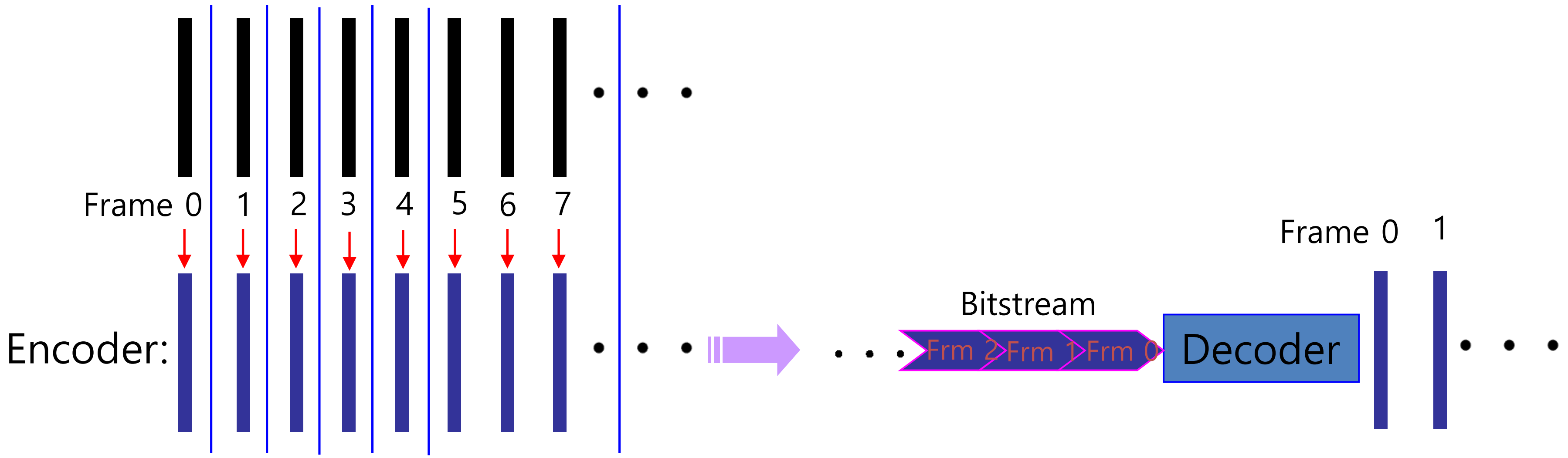
  - *The order of Block unit in frame* at Encoder/Decoder.

Encoder

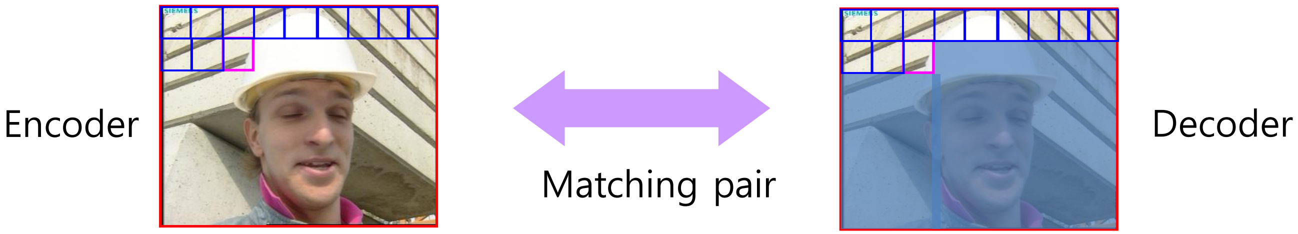
Decoder

❖ Encoding-Decoding Structures (1) (frame unit)

- Video Coding Using *Still Image Coding* (like JPEG)
  - *The order of Frame unit*

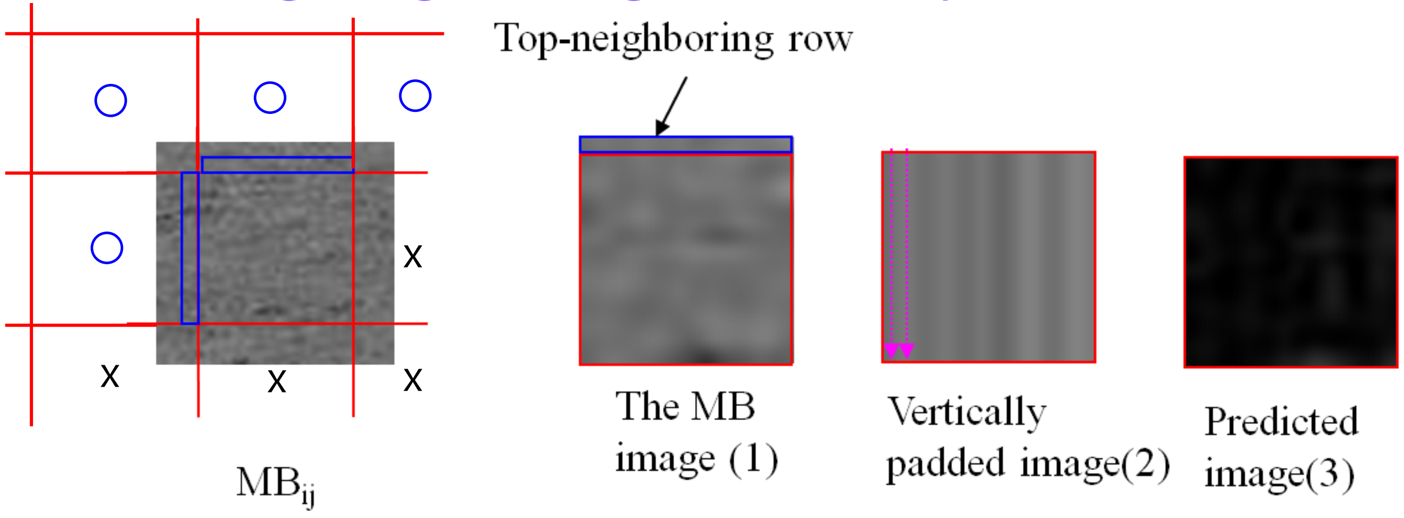


  - *The order of Block unit in frame* at Encoder/Decoder.



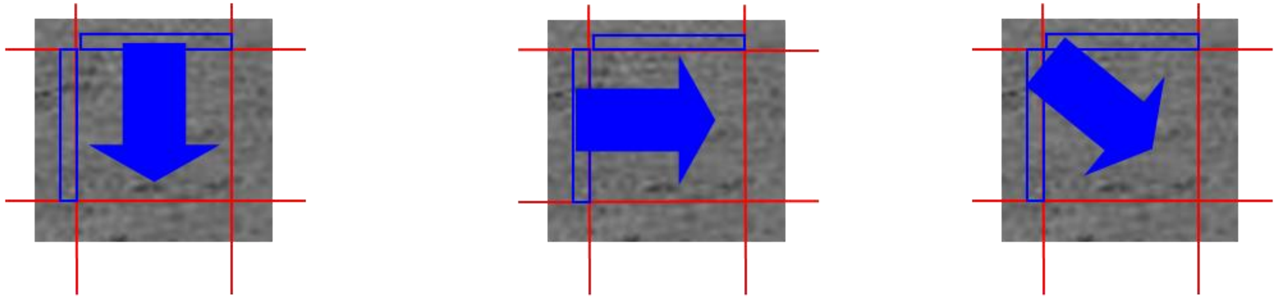
- Q) In still image coding, Is there any way to make *the residual (prediction) image* like ME process?
  - We can use only one (current) frame!

  - *How about using neighboring blocks or pixels?*



VICL
Visual Information Computing Lab

- *Similar way, some directional prediction are available as shown in the following figure.*



Vertical Horizontal Diagonal

- ***Residual (prediction) signal***
  - Data padding image using the top or/and left-neighboring pixels.
  - Residual image(3) : the original MB(1) – Padding image (2)
- DCT using this residual (prediction) image.

We can see that the predicted image is very low-frequency centric image(3). This can give us good property in DCT procedure for higher compression gain.

❖ The Concept of "Intra/Inter"

- Intra- (~안에) Prediction / Frame / Slice / Coding

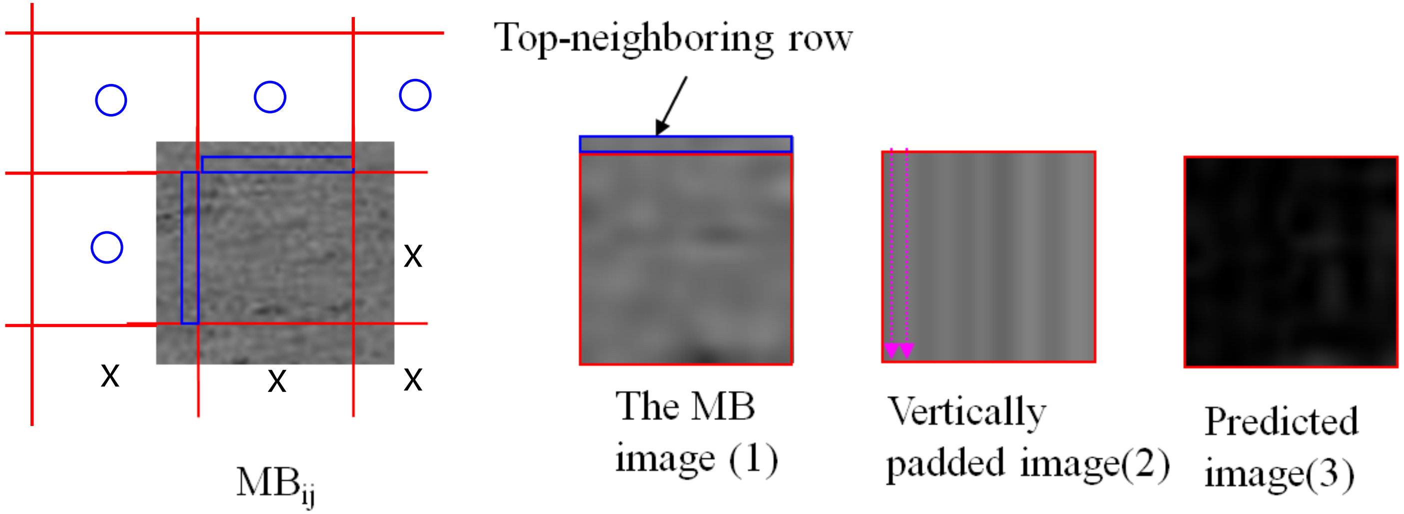
Spatial information

- Inter- (~사이에) Prediction / Frame / Slice / Coding

Frame (t-1) Frame t

Temporal information (using **ME**)

❖ Intra-prediction: to make residual (prediction) image using only neighboring blocks (pixels).



- Intra-frame (slice): I frame
  - A frame (slice) that is composed of all Intra-predicted blocks.
- Intra-coding:

- Inter-prediction: to make residual (prediction) image using neighboring blocks (pixels) and the previous frame (ME).

Frame (t-1) Frame t

  - Inter-frame (slice)
    - A frame (slice) that is composed of all Intra-/Inter-predicted blocks.
  - Inter-coding:

Inter-predicted block is dominant…!

❖Conventional Block-based Video Coding

- **Block-based ME (Inter-prediction)/Intra-prediction**

U

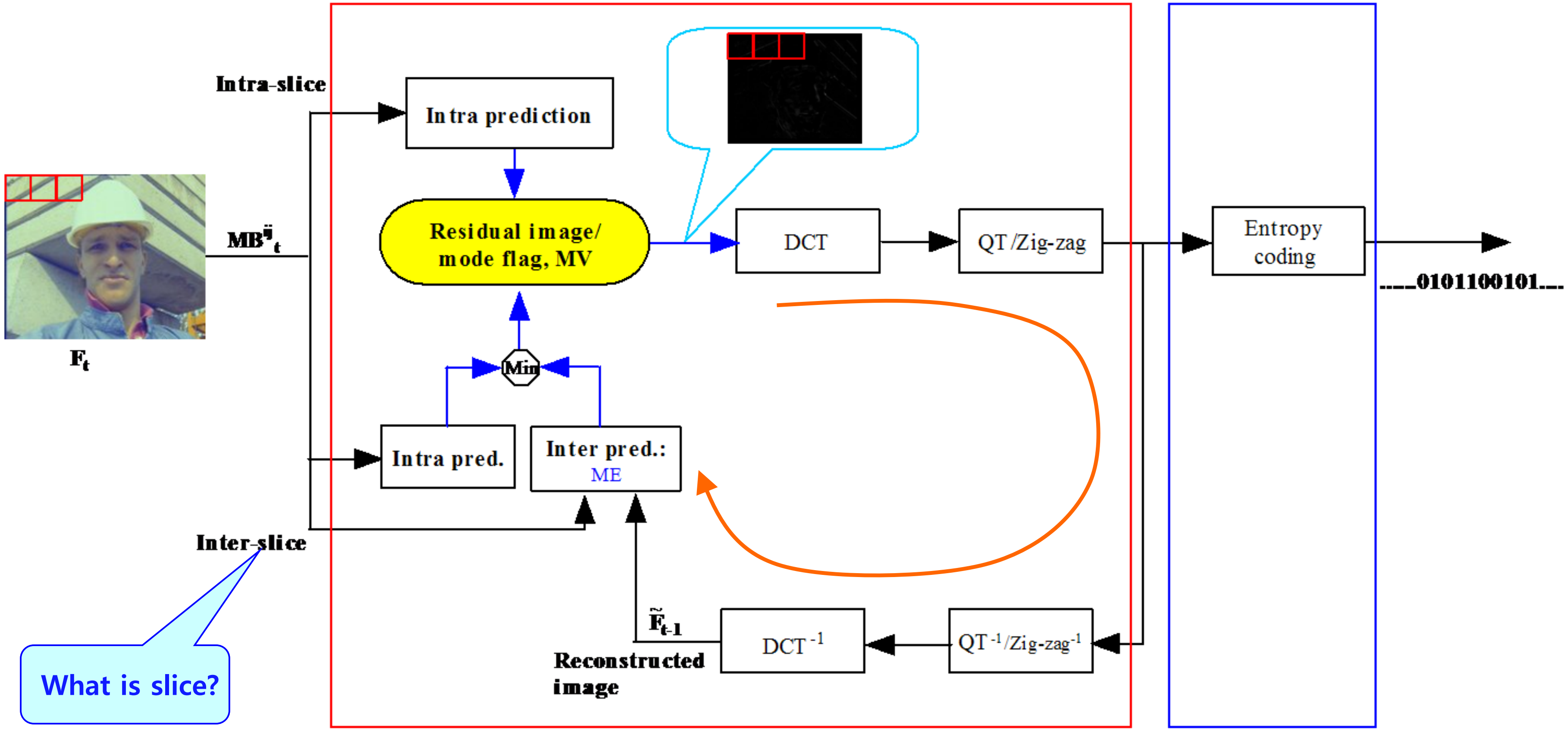
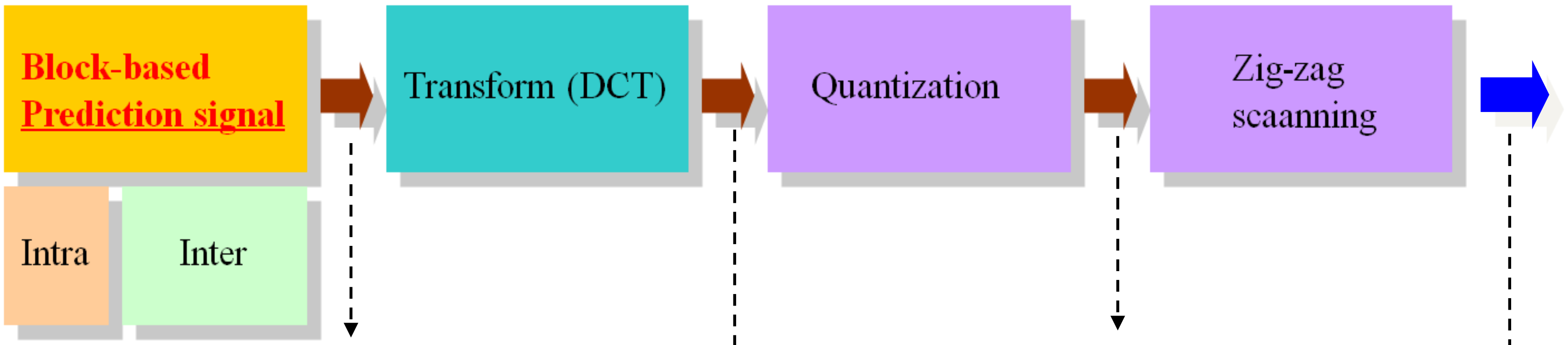
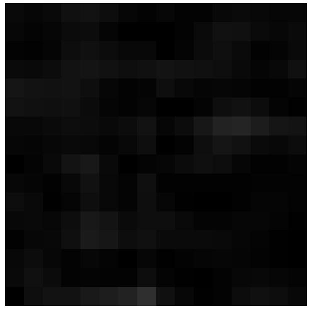


Image model (A)　　Entropy coder (B)

- Video Model (A)



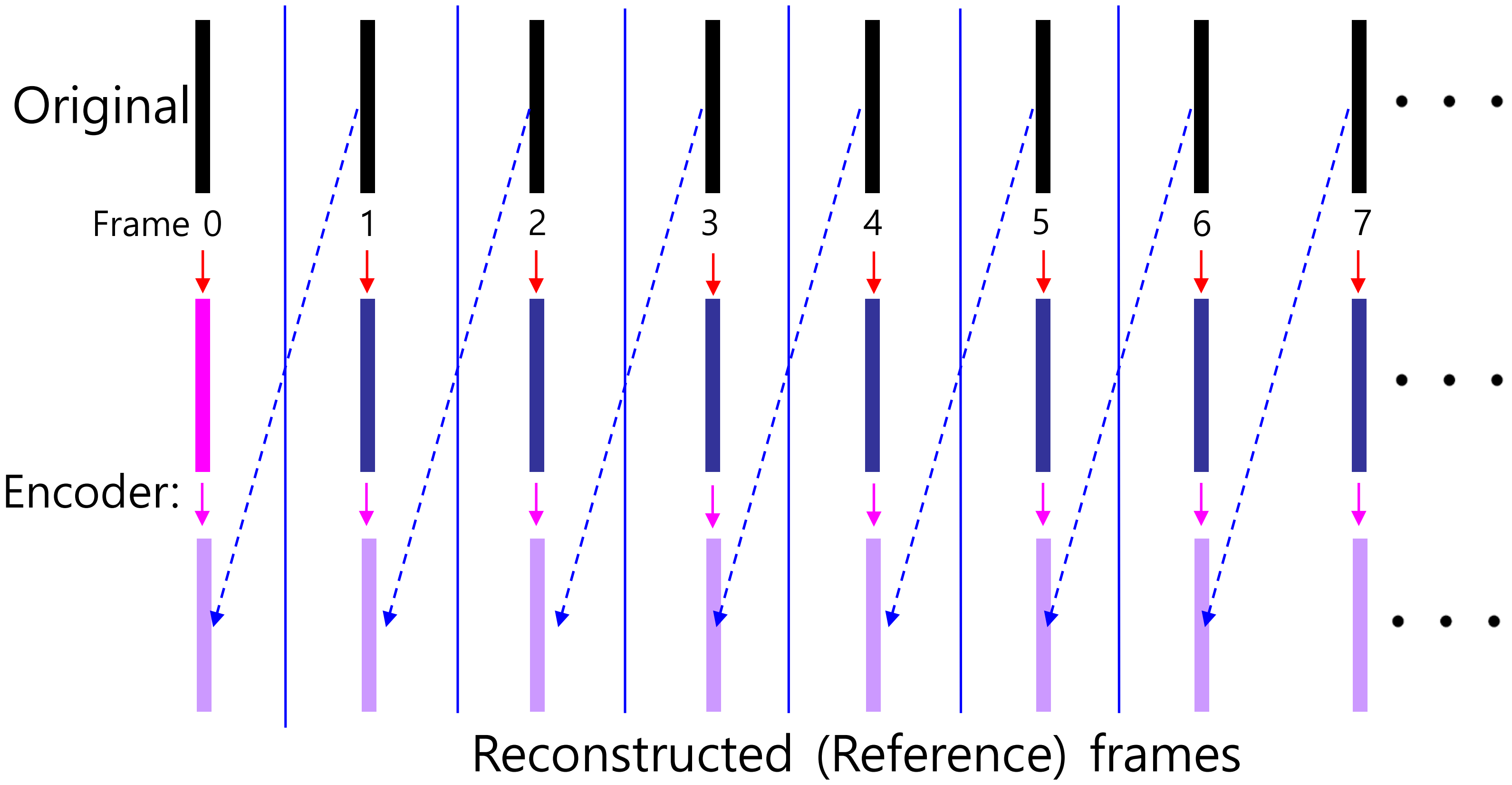


84	34	4	0	0	4	0	0
54	0	0	0	0	0	0	0
15	0	14	0	0	0	0	0
0	0	0	0	0	0	0	0
0	7	0	0	0	0	0	0
0	0	0	0	0	5	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

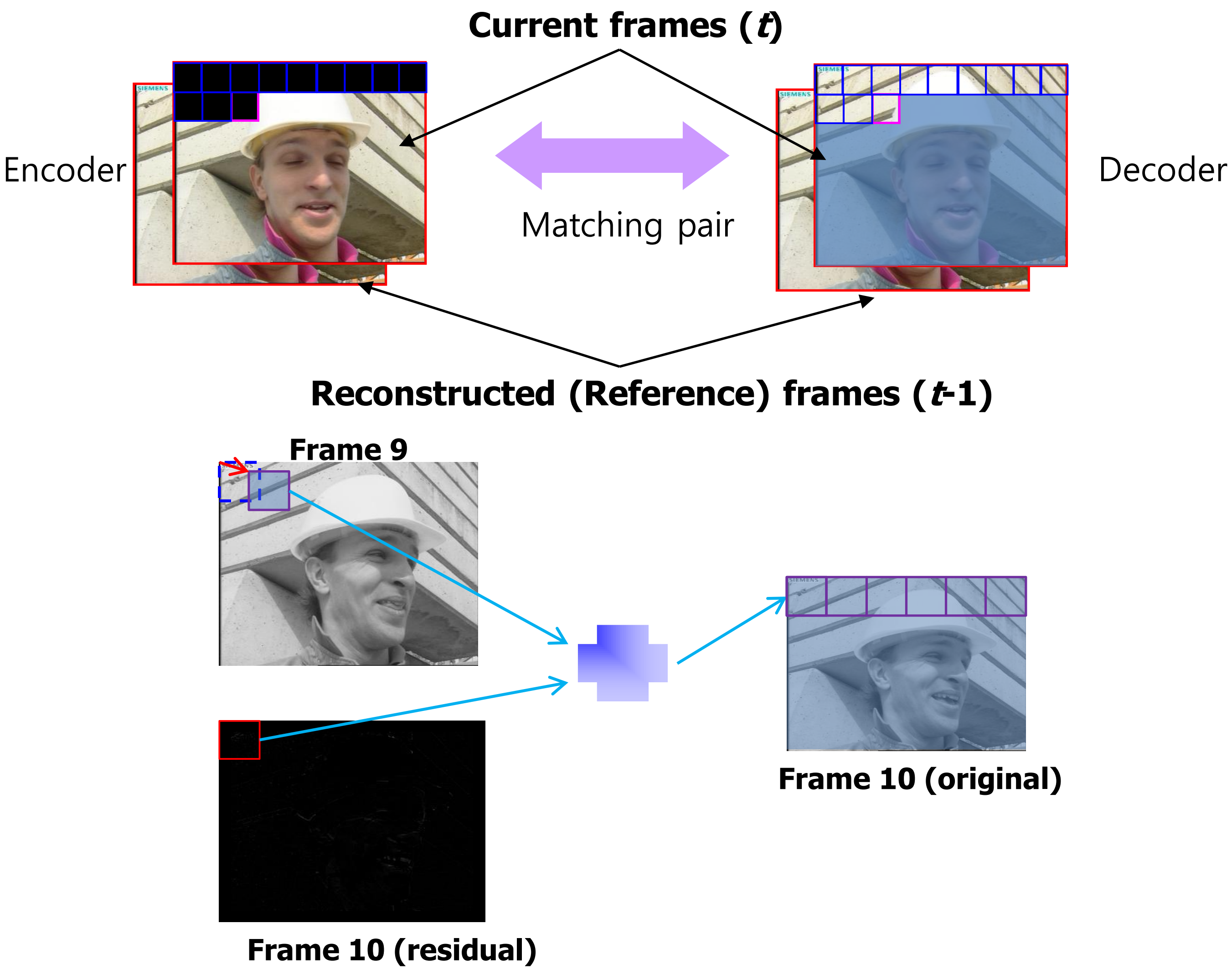
20	5	0	0	0	0	0	0
8	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

$$\overbrace{20,5,8,1,0,0,0,0,0 \,..,0}^{\text{64 coeffs.(symbols)}}$$

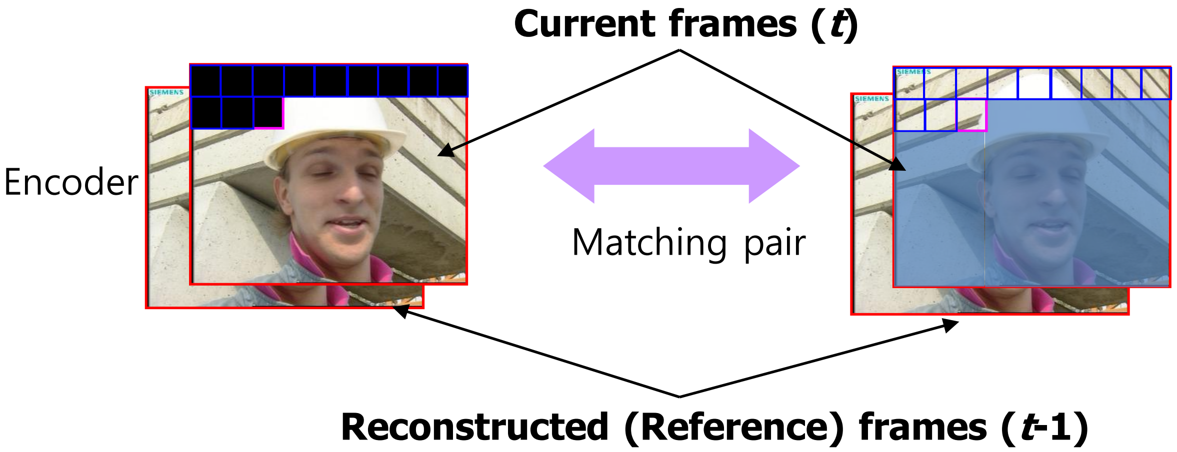
- Encoding-Decoding Structures (2)
  - Video Coding Using *Inter-prediction (ME)*
  - *The order of Frame unit*

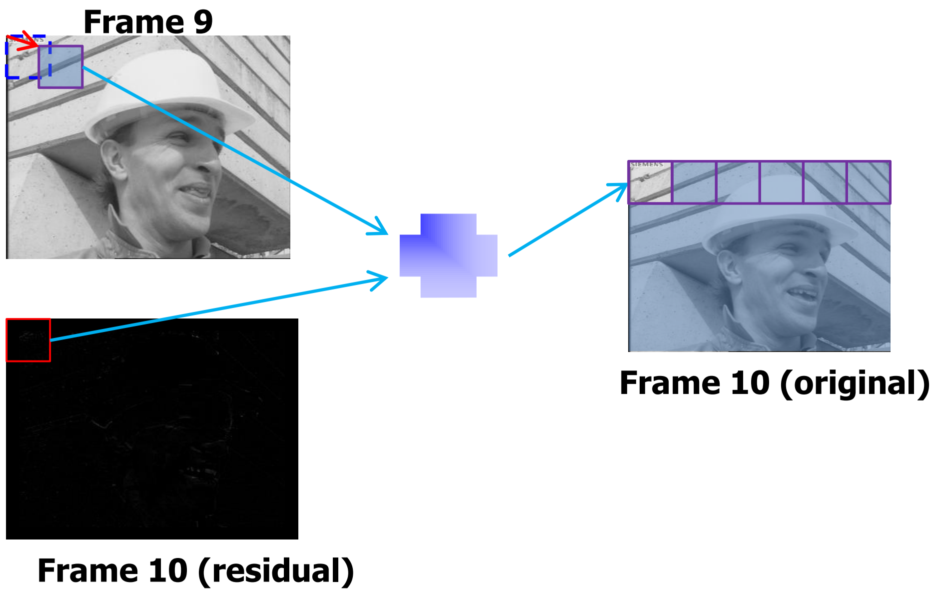


❖ *The order of Block unit in frame* at Encoder/Decoder.

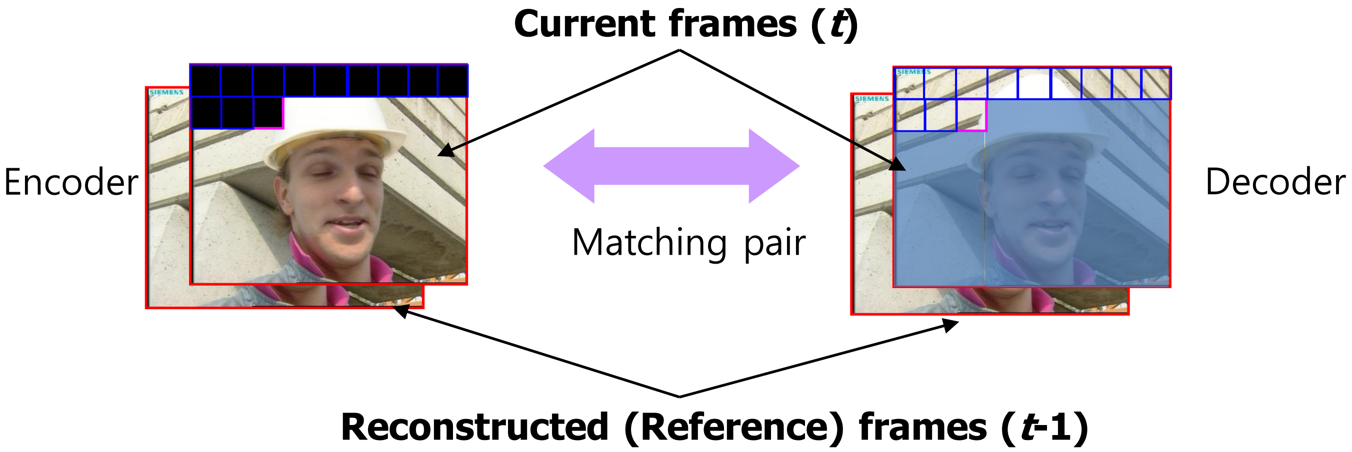


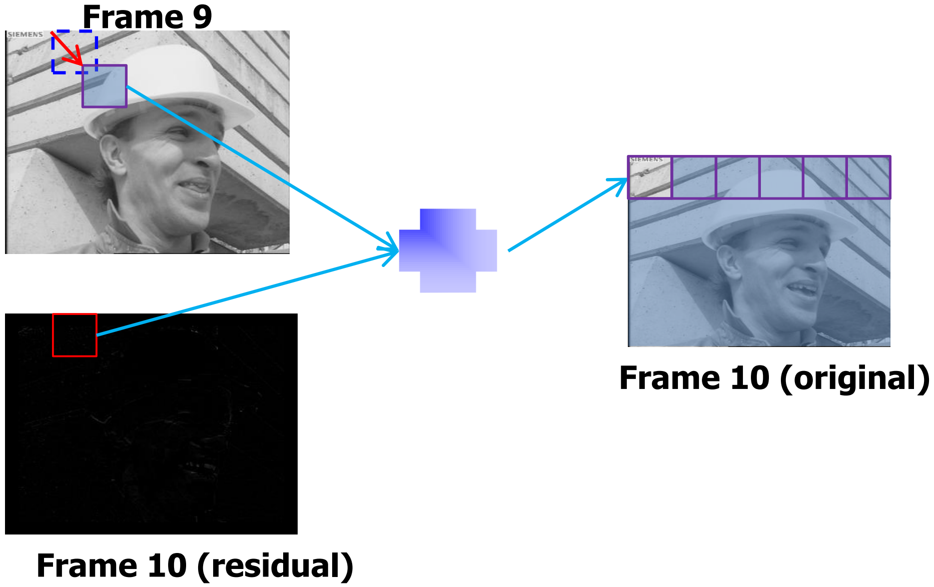
❖ *The order of Block unit in frame* at Encoder/Decoder.



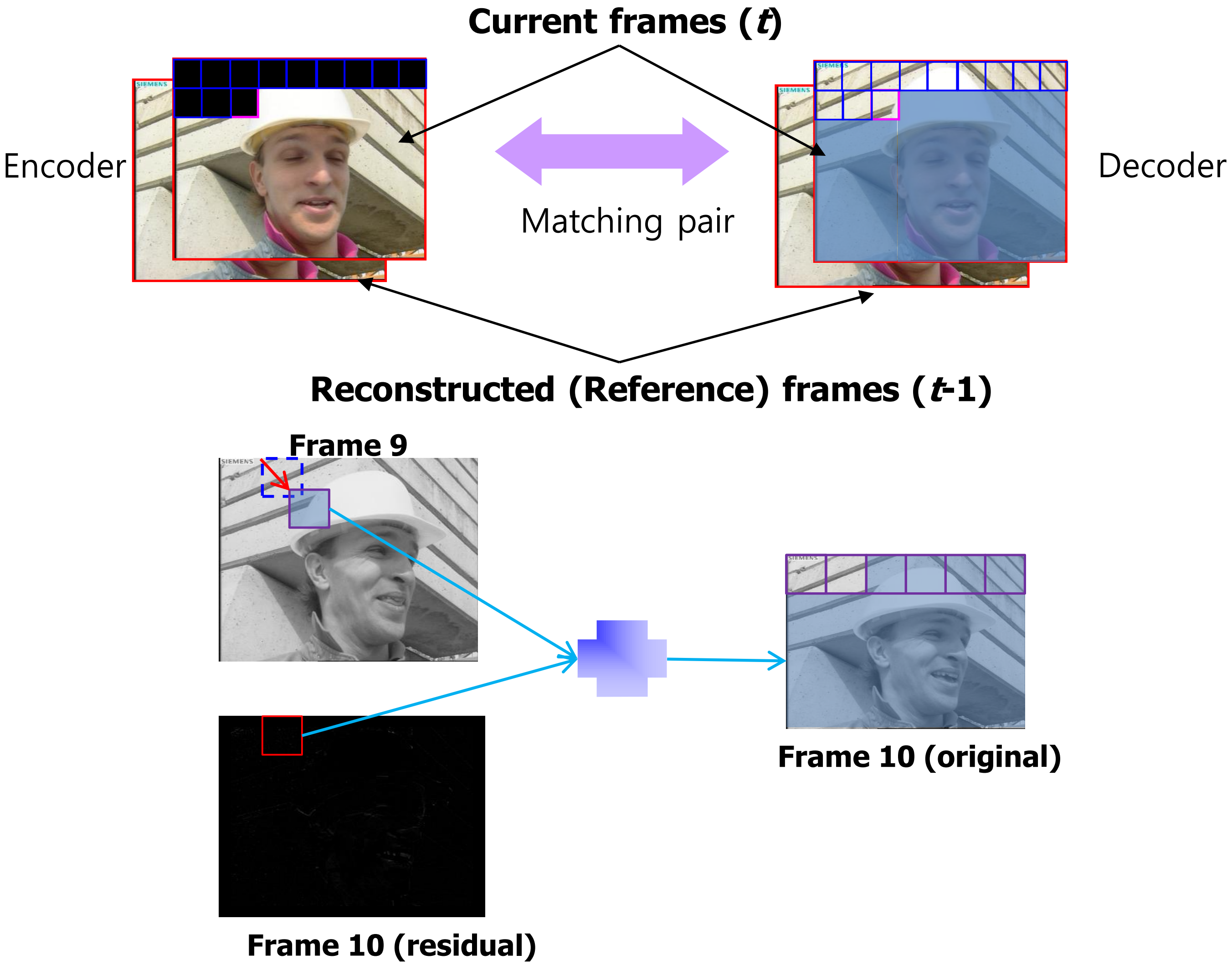


❖ *The order of Block unit in frame* at Encoder/Decoder.

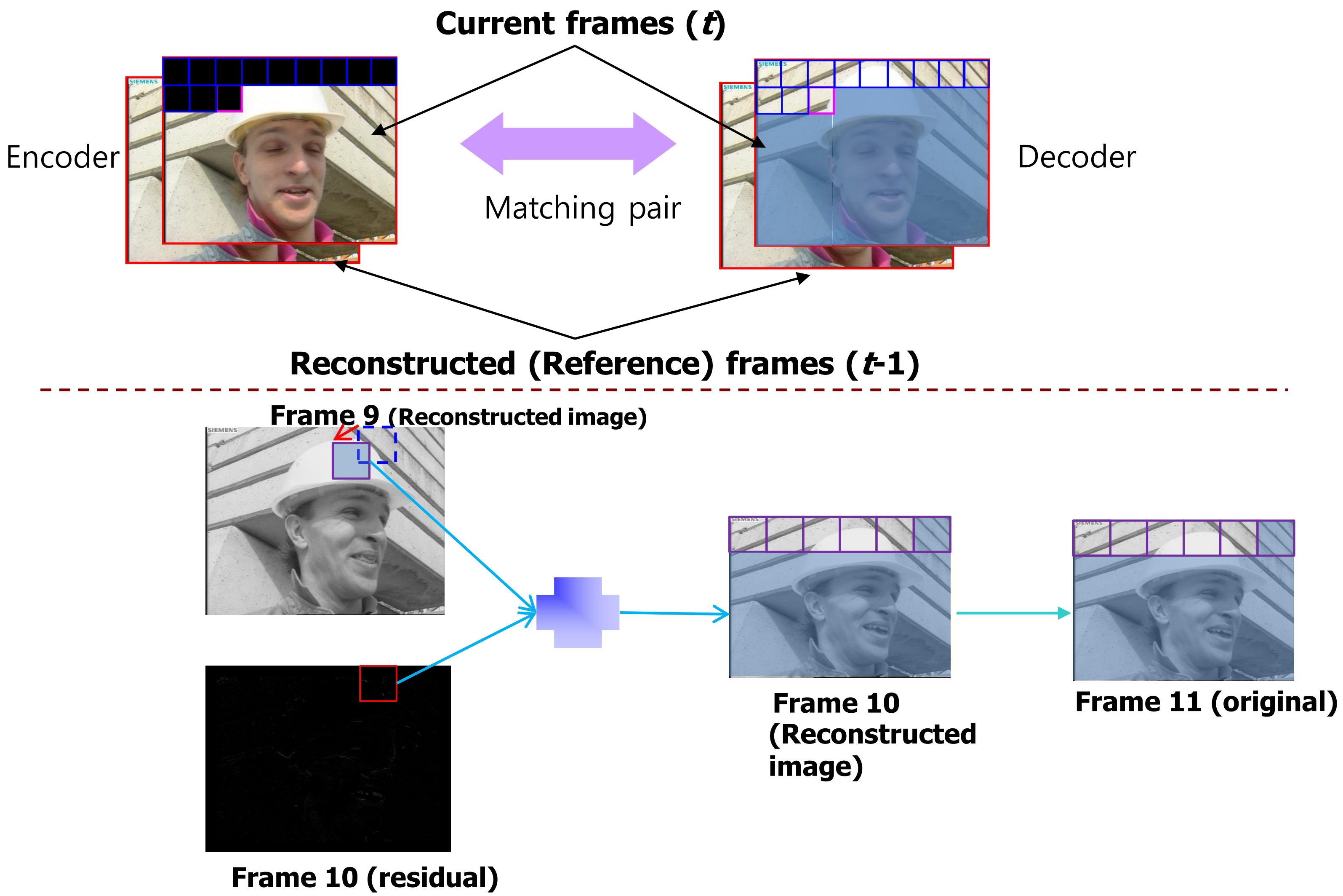




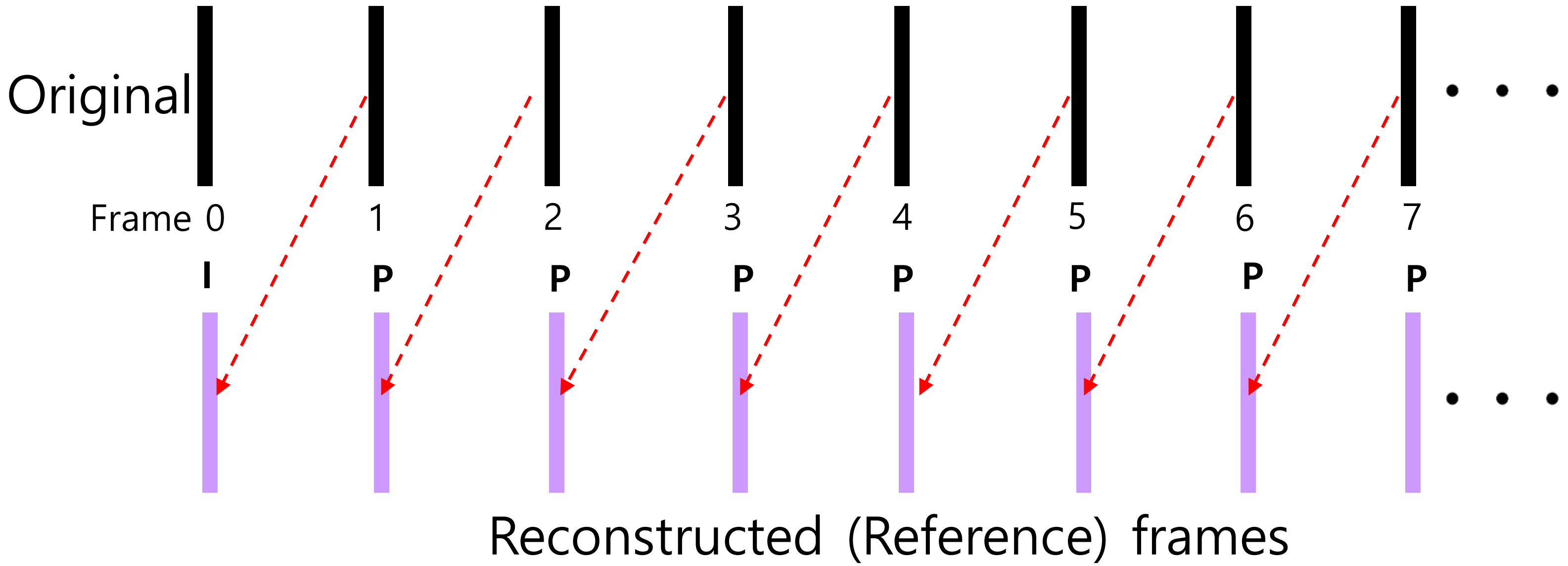
❖ *The order of Block unit in frame* at Encoder/Decoder.



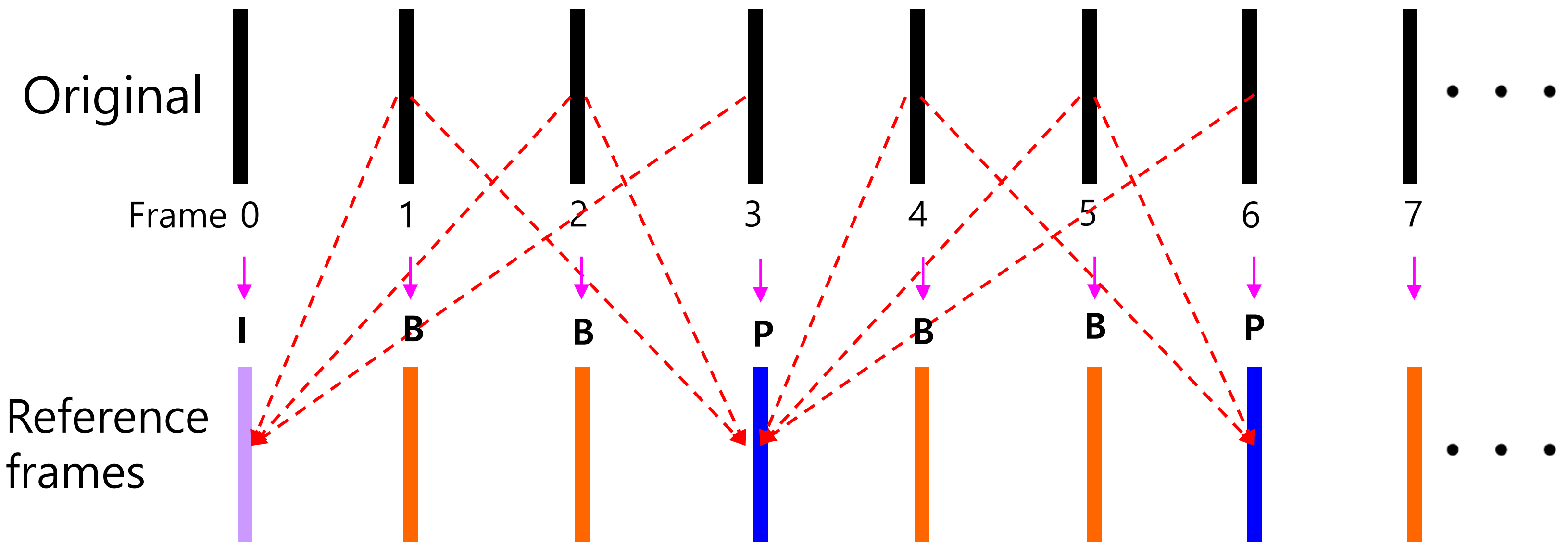
❖ *The order of Block unit in frame* at Encoder/Decoder.



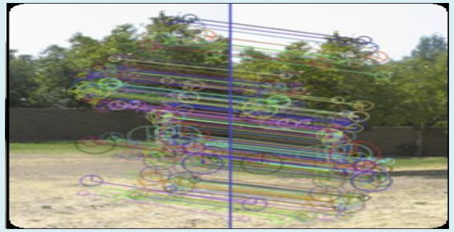
- Inter-frames: P/B frames
  - P frame: Forward direction ME (using (t-1) reference)

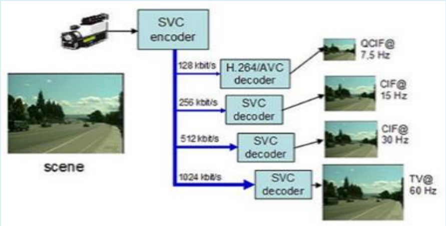


- B frame: Forward and backward direction ME (using (t-1) and (t+1) references).



- Group of Picture (GOP): some successive pictures.
  - Distance between **Anchor frames (I or P frames)**





Contents

- Video Model
- Detailed Processes for Video Coding
- Video Data Format

Video (Color) Data Format (0) (Text 2.4.2, p. 15~)

❖ YCbCr Sampling

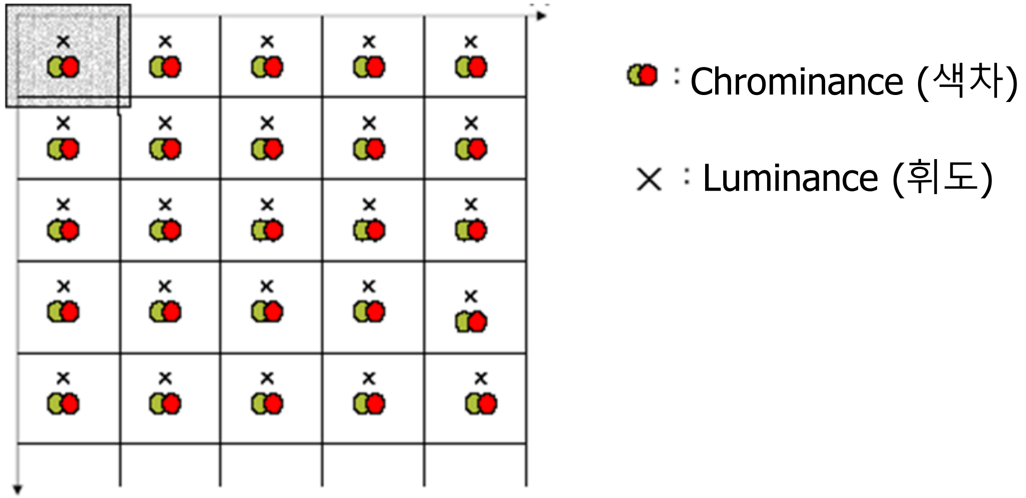
- Y: Luminance (휘도->명암 성분)
- Cb, Cr: Chrominance (색상->칼라 성분)

Gray image
(흑백영상)

- *The eye is more sensitive to Luminance, less sensitive to Chrominance components.*
- 4:4:4 format, 4:2:2 format, 4:2:0 format
  - **SIF (Source Intermediate Format) (352x240)**
  - **HDTV (High Definition Television) formats**
    - 720p (1280x720), 1080i (1920x1080), 1080p(1920x1080)
  - **CIF (Common Intermediate format) (352x288)**
    - videoconferencing applications
  - **QCIF (Quarter CIF) (176x144)**
    - video telephony applications

❖ Types of Color Representation

- 4:4:4 (Y:Cb:Cr)
  - A pixel is represented by 4 Y samples, 4 Cb and 4 Cr samples for each pixel.



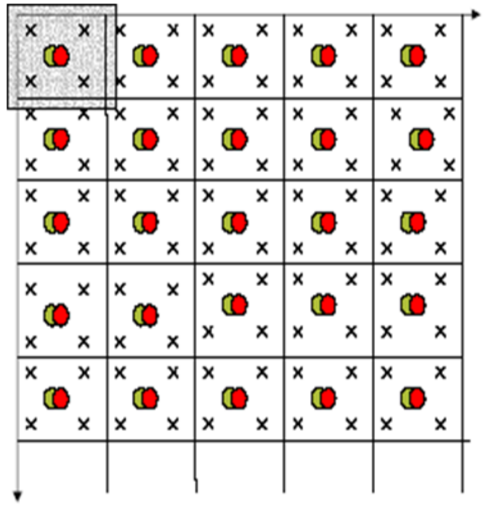
Video (Color) Data Format (2)

❖ 4:2:2 (Y:Cb:Cr)

- A 2x2 block of pixels in an image is represented by 4 Y samples, 2 Cb and 2 Cr samples for each pixel (but all 4 pixels **share** 2Cb and 2Cr samples).

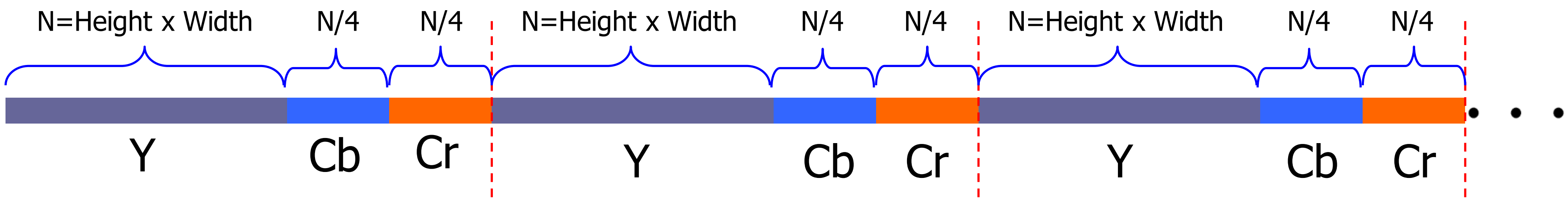
❖ 4:2:0 (4:1:1) (Y:Cb:Cr)

- A 2x2 block of pixels in an image is represented by 4 Y samples, 1 Cb and 1 Cr samples for each pixel (but all 4 pixels **share** a Cb sample and a Cr sample).



❖ Actual Video Data Structure (4:2:0)

- Data type: BYTE (unsigned char).
- Value : A Integer value in [0 ~ 255].
- There is no header information.



❖ Detailed Processes for Image/Video Coding

- Block-based Video Coding
  - Inter-prediction (signal)
  - Inter frames (P and B frame)
- Encoding-Decoding Structures
  - Still Image Coding
  - Video Coding

- Video (Color) Data Format
  - YCbCr Sampling.
  - Types of 4:4:4, 4:2:2, 4:2:0.
  - Actual video data structure.

HomeWork #1 (Due: ~ April 3, Mon.)

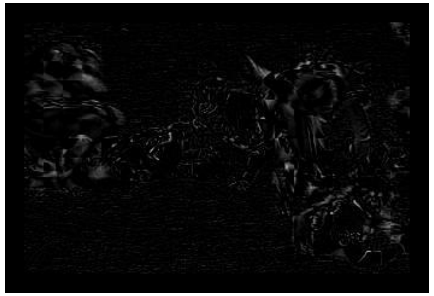
❖ Implementation of ME module

- With the given raw video data, execute the motion estimation (ME) MB(16x16) by MB and then save their residual data into new raw video file (Y component).

Frame 9

Frame 10

Frame 10 (Residual image)



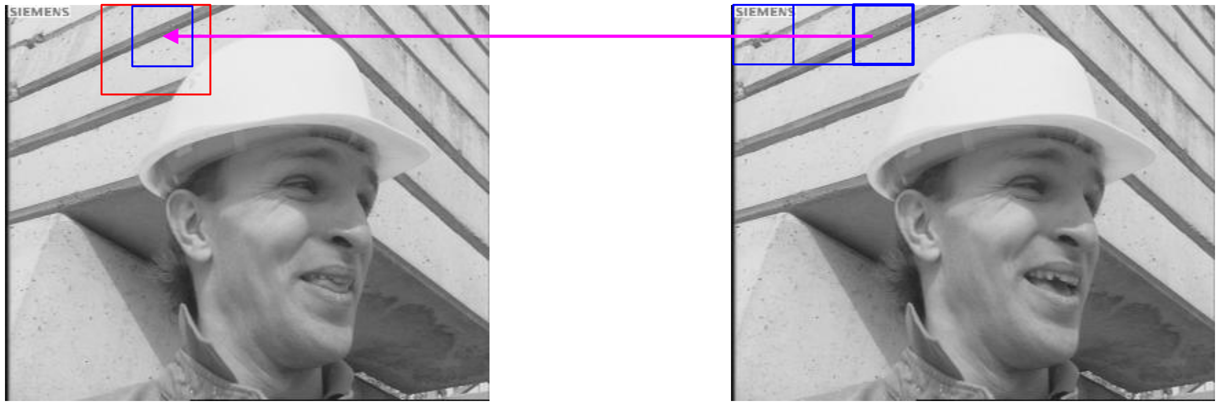
Residual image by *ME* Procedure

HomeWork #1 (Due: ~ April 3, Mon.)

❖ To do so,

- You should get Y-components from raw video data (color).
- Also you should define the search range how large area you want to search.
- You must load successive two frames. From the second frame, get a MB in progressive order, then take a search procedure (MV estimation) pixel by pixel as you have learned. If you find the best matching point, its residual data can be saved into the current MB position.

- Block-based Motion Estimation (ME)
  - Block-based processing:
    - » Macro-block (16 x16) based design.
    - » Block (8x8) can be used as application.
    - » *Progressive* block scanning.
  - *What is the Motion Estimation (ME)?*
    - » Procedure to find the best matching area in the previous picture for the current block (or Macro-block).
    - » A *search range* is generally given or defined by user.



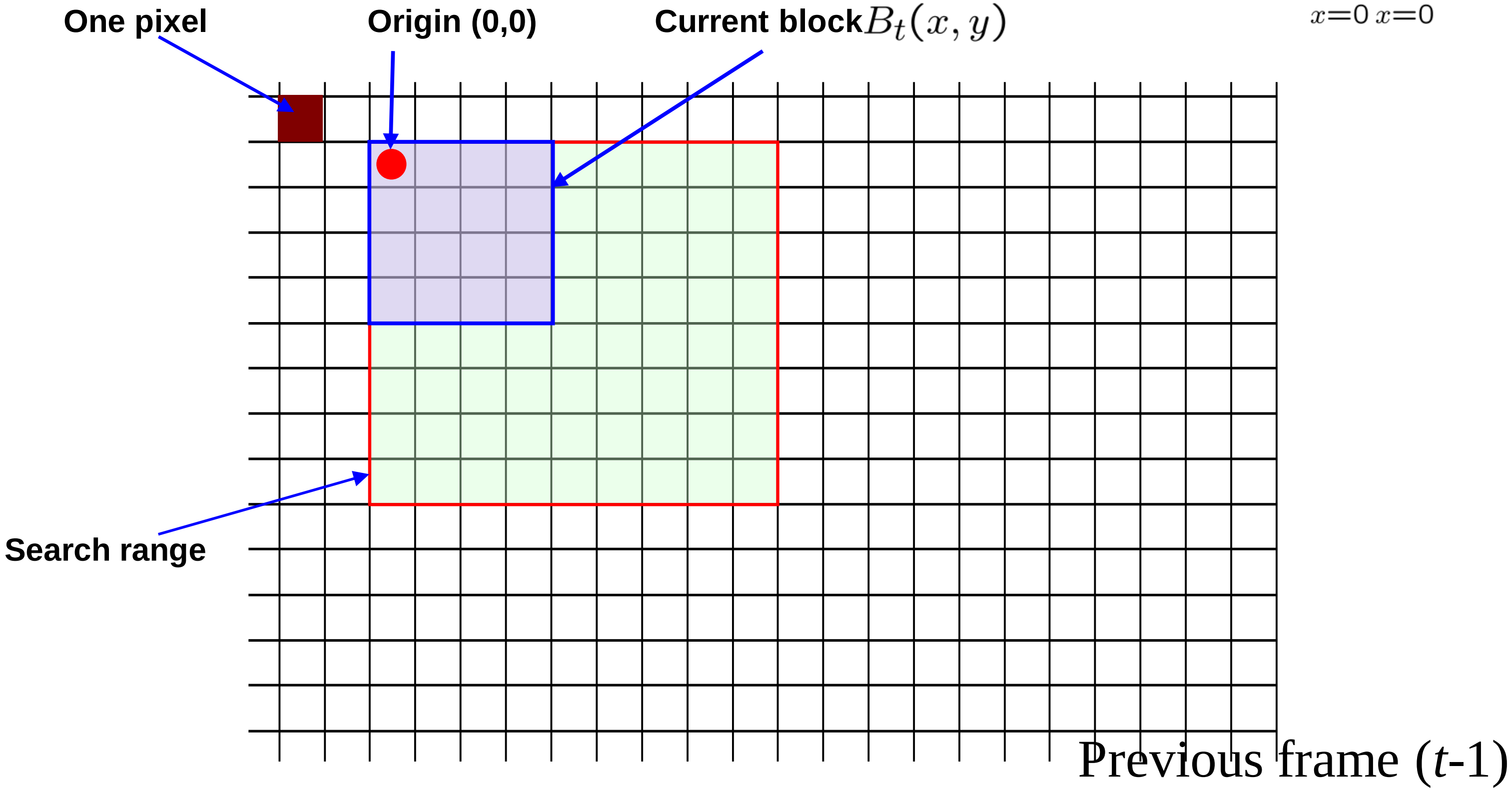
Frame (t-1) Frame t

VICL Visual Information Computing Lab

Basic Processes for Motion Estimation

- *The best matching point*:
  - How to define?
  - Which process/find?

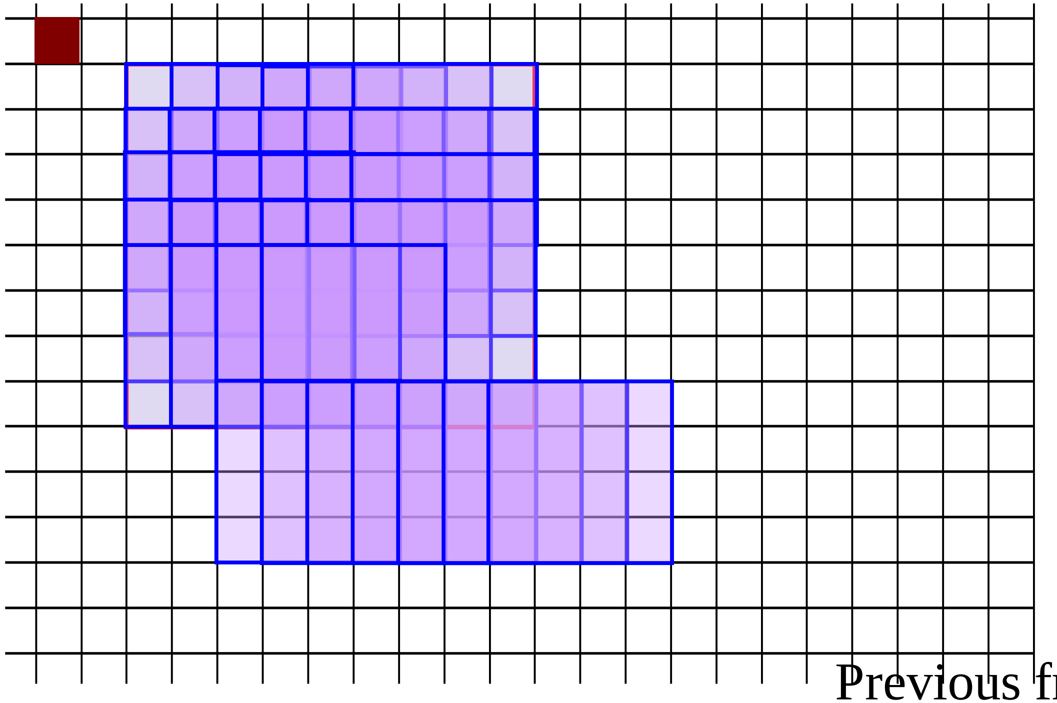
$$SAD(i,j) = \sum_{x=0}^{x=4} \sum_{x=0}^{x=4} |B_t(x,y) - B_{t-1}(x,y)| \quad (1)$$



Basic Processes for Motion Estimation

- *The best matching point*:
  - Using *SAD* equation, *SAD* values are computed at every search (pixel) points.

$$SAD(i,j) = \sum_{x=0}^{x=4} \sum_{x=0}^{x=4} |B_t(x,y) - B_{t-1}(x,y)| \quad (1)$$

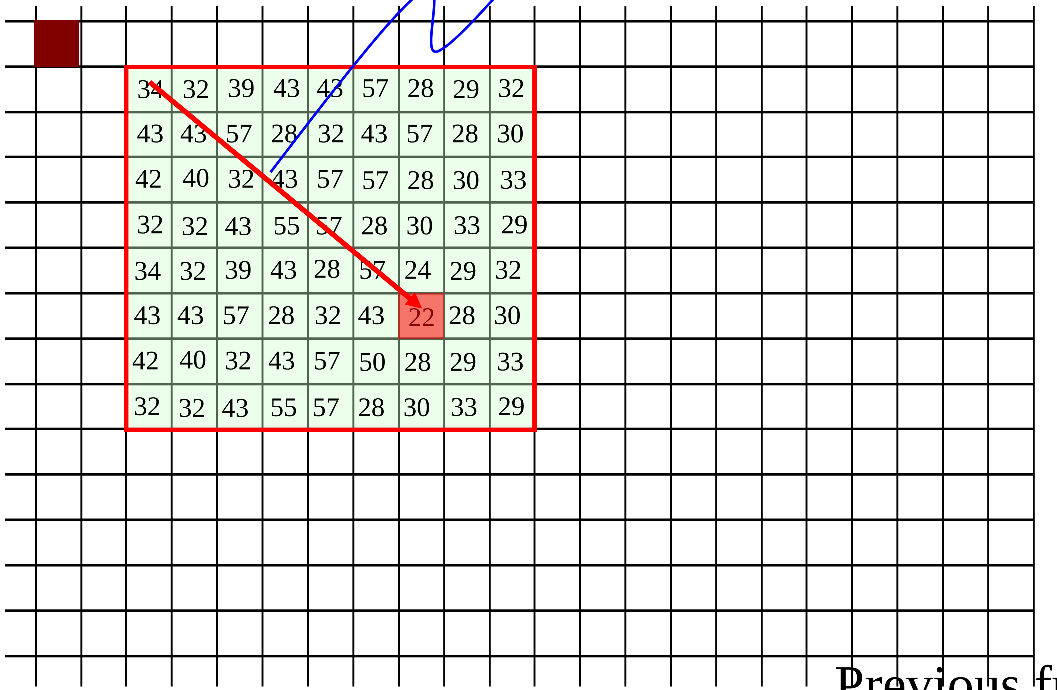


Previous frame (t-1)

- *The best matching point*:
  - Find the minimum distortion (error) point.
  - ***Displacement from the origin to the best point.***

$$SAD(i,j) = \sum_{x=0}^{x=4} \sum_{x=0}^{x=4} |B_t(x,y) - B_{t-1}(x,y)| \quad (1)$$

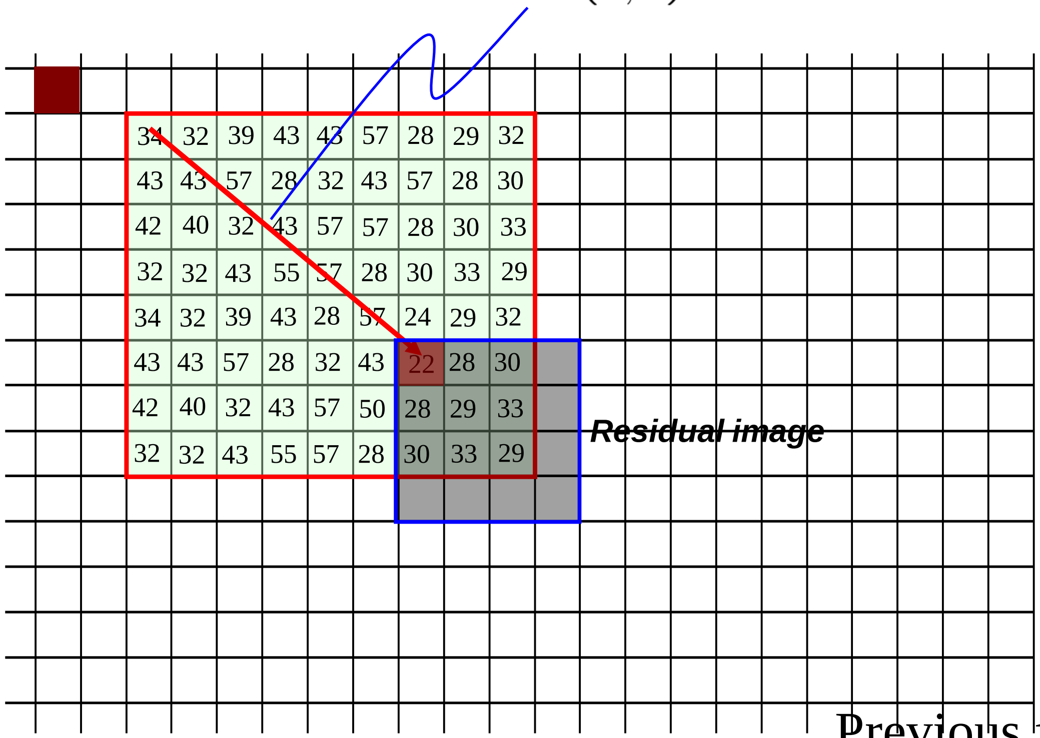
$$MV = (6,6)$$



Previous frame (*t*-1)

$$SAD(i,j) = \sum_{x=0}^{x=4} \sum_{x=0}^{x=4} |B_t(x,y) - B_{t-1}(x,y)| \quad (1)$$

$MV = (6, 5)$



Previous frame (t-1)

Thank you for your attention.!!!

QnA

http://vicl.sookmyung.ac.kr