Barème and correction

First Name: Last Name:

## Exam - Master VICO - UE Multimedia Communication
Video Coding Part

October 2022 – duration 1h30 – answer **directly on the exam subject**

Authorized documents: handwritten and personal notes.
Laptop and phone are not authorized.

Indicative scale (/ 20): Exercise part on 12 points. Questions part on 8 points.

8   *12*

Exercice – DPCM

Figure 1 shows a DPCM (Difference Pulse Code Modulation) scheme applied to a grey level image compression, where $X(i, j)$ is the current pixel to encode, $Xp(i, j)$ is its prediction, $E(i, j)$ is the prediction error, $Eq(i, j)$ is the quantized prediction error, and $Xr(i, j)$ is the reconstructed pixel obtained after decoding. Note that the $Xr(i, j)$ are used in order to predict.

CODEUR :

DECODEUR :

Figure 1 : DPCM coder-decoder.
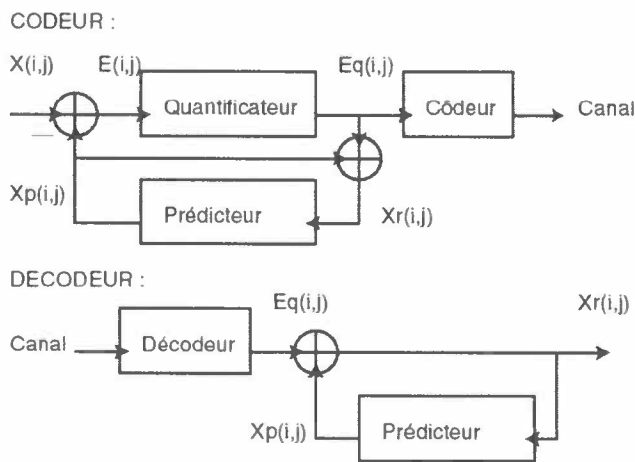
B	C
A	X

Figure 2: Definition used by for the prediction.

Knowing that a zigzag is used in order to scan the image, the points definitions used by the predictor are given at the Figure 2 where X is the pixel to predict. The used prediction function is $P(X)=A$, and we consider that the 1st column pixels have been already transmitted without error by using another procedure.
The used quantizer is described at the Figure 3.

Questions :

Why we have to consider that the 1st column pixel have been transmitted by using a different procedure?

→ No past   ②

Fill the Figure 4 tables that correspond to the coding of a small image.

⟶ ④

If a natural code is used in order to transmit the Eq, how many bits per Eq are necessary? Explain. Note that with a natural code, all the generated codewords have the same size in bits.

②

$16\ E_q \Rightarrow 2^4 = 16$

$4\ bits/E_q$

1/4

X

20	20	40	40
20	40	40	60
20	40	40	80

Xp $= A$

	20	18	38
	20	40	38
	20	40	38	Ⓐ

E $= X - X_p$

	0	22	2
	20	0	22
	20	0	42	Ⓐ

Eq $= Q(E)$

	-2	20	3
	20	-2	20
	20	-2	38	Ⓐ

E(i, j)	Eq(i, j)
] -255, -113]	-126
] -113, -89]	-101
] -89, -66]	-77
] -66, -48]	-55
] -48, -28]	-37
] -28, -13]	-19
] -13, -4]	-7
] -4, 2]	-2
] 2, 5]	3
] 5, 14]	8
] 14, 29]	20
] 29, 47]	38
] 47, 67]	56
] 67, 90]	78
] 90, 114]	102
] 114, 255]	127

Figure 3 : The quantizer.

Xr $= E_q + X_p$

20	18	38	41
20	40	38	58
20	40	38	76	Ⓐ

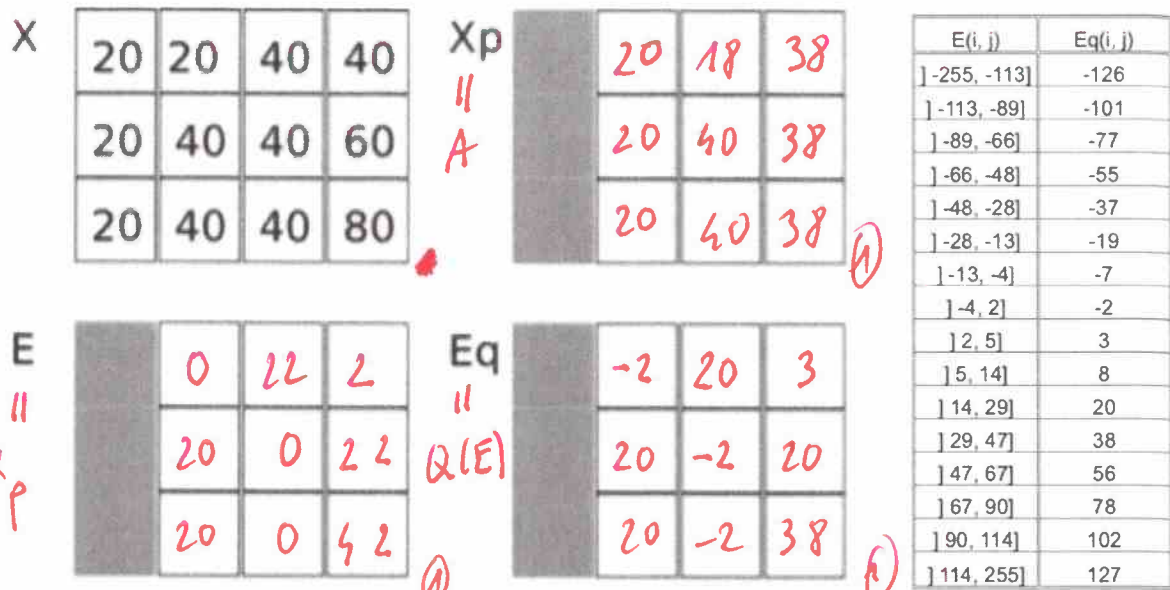*Figure 4 : Different steps of the DPCM coding of a small image.*

Questions about video coding

How many bits per second (*bpp*) do you have in this raw video :
1080 50p (namely: spatial resolution 1080×1920 pixels, progressive format, 50 frames/s), 4:2:2 format (see also the Figure 5 where each component is respectively coded on 8 bits)? Only give the details of the literal formula of the computation.

$$1080 \times 1920 \times 50 \times 16 \quad \text{bits/s}$$

$$4:2:2 \implies \frac{32 \text{ bits}}{2 \text{ pixels}} \implies 16 \text{ bits/pixel} \quad Ⓐ$$

An exhaustive search is used for motion estimation. If the size of the target block *b* is 8×8 pixels, and if the size of the search window *B* is 16×16 pixels, how many blocks from *B* will be compared to *b*? Only give the details of the literal formula of the computation.

$$(8+1)^2 \text{ blocks} \quad Ⓐ$$

The Figure 6 illustrates the ME (Motion Estimation) module of VCDemo where the "Vectra" video sequence have been processed. Answer the questions:

- What information type is represented in the top-left sub-image?

Original frame

- What information type is represented in the top-right sub-image? Explain.

FD : Frame Difference

- What information type is represented in the bottom-left sub-image? Explain

MV : Motion Vectors

- What information type is represented in the bottom-right sub-image? Explain.

MCFD: Motion Compensated Frame Difference

- Why does the table on the right display the column « Vector entropy »? Explain the importance of this information in video coding context.

Estimation of the MV costs (in bit ?) because MV have to be coded.

- How is the bottom-right sub-image content when the ME has been very efficient? Explain.

We obtain an uniform gray Frame (No difference)

The Figure 7 illustrates the architecture of a hybrid encoder. For a given block coded by using an inter mode by MCFD, what are the 2 types of information which have to be coded and transmitted in order that the decoder can reconstruct the block?

- 1st type of information:

MV

- 2nd type of information:

Prediction Error

The Figure 7 illustrates the architecture of a hybrid encoder. What the main conceptual differences between the intra prediction coding mode and the inter-images prediction mode?
Note that the inter prediction mode is based on motion estimation and compensation.

Intra production ⟹ based on the intra frame

Inter prediction ⟹ based on the inter frames

Let be the following GOP used to encode a video sequence by using MPEG1 standard: $I_1, B_1, B_2, B_3, P_1, B_4, B_5, B_6, P_2, B_7, B_8, B_9, I_2, \ldots$

Which frames have to be successively decoded in order to display **only** the B_5 frame?

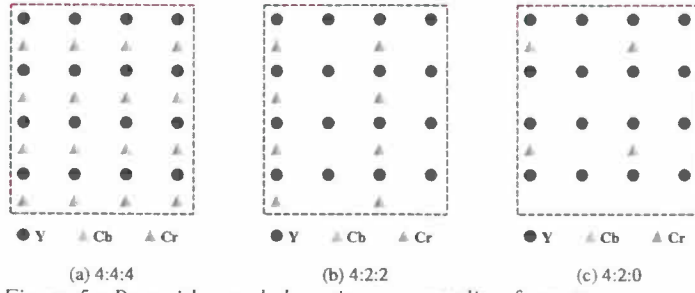$I_1 \longrightarrow P_1 \longrightarrow P_2 \longrightarrow B_5$

decoding

①

(a) 4:4:4   (b) 4:2:2   (c) 4:2:0

● Y   ▲ Cb   ▲ Cr

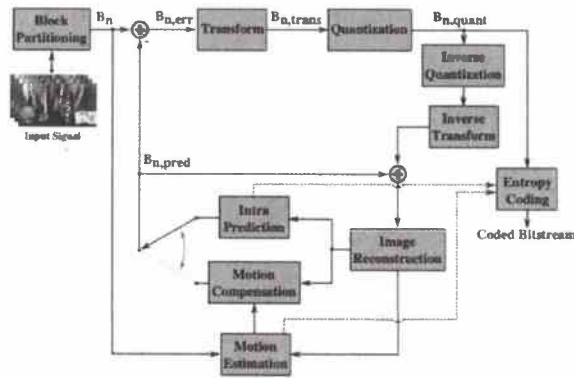*Figure 5 : Raw video and chrominance sampling formats.*
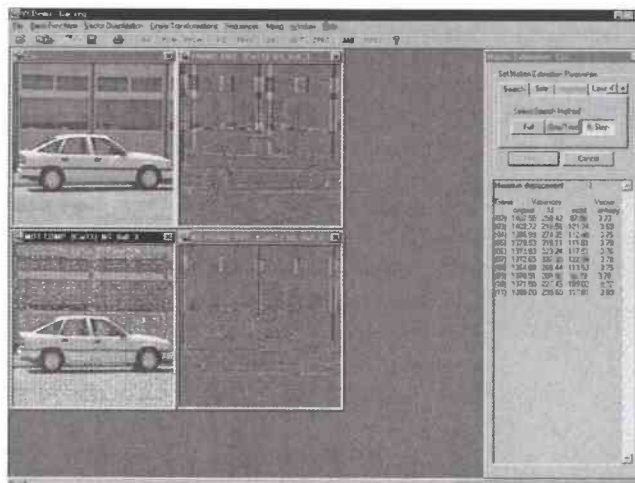

Figure 7 : Architecture of a hybrid encoder.

Figure 6 : ME « Motion Estimation » module of VCDemo.