JPEG and MPEG-2: a brief overview

O. Le Meur
olemeur@irisa.fr

Univ. of Rennes 1
http://www.irisa.fr/temics/staff/lemeur/

January 3, 2011
JPEG and MPEG-2: a brief overview

1. Introduction

2. JPEG

3. What does MPEG-2 standard tell us?
Introduction

What does MPEG-2 standard tell us?
Motivations - compression standards for still pictures

Motivations

Still color picture 800x300 pixels, 24 bits/pixel: 6.3 Mbits/image (modem 56k → 3min 39s !!).

Tremendous progress made in a number of advanced technologies leading to powerful computers and computations. Pictures, widely utilized in our daily lives. Possibility to compress data information due to the statistical redundancy of information and to the psychovisual redundancy.
What does MPEG-2 stand tell us?

What is the goal of a compression standard?

**Definition (Standard)**

A format that has been approved by a recognized standards organization or is accepted as a de facto standard by the industry.

For the compression standards, there are two organizations: ITU-T-VCEG and ISO MPEG. A standard only specifies bitstream syntax and decoding process.

The goal is to create the best video compression standards for targeted applications.

CfE: Call for Evidence; CfP: Call for Proposal.

Image Compression standards

- **JPEG** (Joint Photographic Experts Group), ISO/IEC 10918-1, ITU-T Recommendation T.81, (1992);

- **JPEG2000** (Joint Photographic Experts Group) ISO/IEC 15444-1, (2004);

- **GIF** (Graphics Interchange Format) two standards GIF87a and GIF89a (LZW algo);

- **JBIG** (Joint Bi-level Image experts Group) is a format for bi-level (black/white) image compression. International Standard, IS 14492 (1999) ISO/IEC 1359, (1999);
Introduction

JPEG

What does MPEG-2 standard tell us?

1. Introduction

2. JPEG
   - Presentation/general scheme of JPEG (baseline)
   - Input (YUV 420, 8x8 grid)
   - DCT
   - Quantization
   - Coding of the coefficients AC
   - Coding of the coefficients DC
   - Example
   - Summary

3. What does MPEG-2 standard tell us?
JPEG: Joint Photographic Experts Group

ISO/IEC JTC1/SC29/WG10

ISO: International Organization for Standardization
IEC: International Electrotechnical Commission
JTC: Joint ISO/IEC Technical Committee

- Widely used for ... a lot of applications.

Two profiles: JPEG (Baseline) and JPEG-LS (Lossless).
General scheme of JPEG (baseline)

**General layout of the coding scheme**

- **Image**
- **Transform coding**
- **Quantization**
- **Entropy Encoder**
- **Binary stream**

**Block diagram of JPEG Baseline**

- **Image**
- **8x8 DCT**
- **Q**
- **DC index**
- **AC index**
- **Quantization tables**
- **Predictive coding**
- **VLC**
- **Bits**
- **DC Huffman table**
- **AC Huffman table**
- **Zig-Zag Scan**
- **RLE**
- **VLC**
- **Bits**
**Introduction**

**JPEG**

What does MPEG-2 standard tell us?

**Input format**

**Block diagram of JPEG Baseline**

- Input format
- JPEG
- Presentation/general scheme of JPEG (baseline)
- Input (YUV 420, 8x8 grid)
- DCT
- Quantization
- Coding of the coefficients AC
- Coding of the coefficients DC
- Example
- Summary

---

**Color space is** $YUV$:

\[
\begin{bmatrix}
Y \\
U \\
V
\end{bmatrix} =
\begin{bmatrix}
0.299 & 0.587 & 0.114 \\
-0.1687 & -0.3313 & 0.5 \\
0.5 & -0.4187 & -0.813
\end{bmatrix}
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}
\]

---

(a) Source
(b) Y
(c) U (Blue)
(d) V (Red)
Input format

Format 4:2:0 (Human eyes are less sensitive to the chrominance than to the luminance):
Introduction

What does MPEG-2 standard tell us?

Block grid

Presentation/general scheme of JPEG (baseline)
Input (YUV 420, 8x8 grid)
DCT
Quantization
Coding of the coefficients AC
Coding of the coefficients DC
Example
Summary

Block diagram of JPEG Baseline

Image → 8x8 DCT → Q → Predictive coding → VLC → Bits

DC index
AC index
Quantization tables

Zig-Zag Scan → RLE → VLC → Bits

DC Huffman table
AC Huffman table

8x8 block
Padding
**DCT**

**Block diagram of JPEG Baseline**

- **Image** → 8x8 DCT → Quantization tables → Predictive coding → VLC → Bits
- DC index → DC Huffman table
- AC index → AC Huffman table → Zig-Zag Scan → RLE → VLC → Bits

**Reminder:**

\[ I = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \langle I, \varphi_{n,m} \rangle \varphi_{n,m} \]

\[ B = \{ \varphi_{n,m} \}_{0 \leq n < N, 0 \leq m < M} \]

an orthogonal basis.

**DCT 8x8**

For a given block NxN:

\[ \varphi_{n,m}(x, y) = \lambda(n)\lambda(m) \frac{2}{N} \cos \left( \frac{(2x+1)n}{2N} \right) \cos \left( \frac{(2y+1)m}{2N} \right) \]

\[ \lambda(t) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } t = 0 \\ 1 & \text{otherwise} \end{cases} \]
DCT and inverse DCT

The DCT of a block \( I \) of size \( N \times N \) is defined by:

\[
DCT(I)(n, m) = \frac{2}{N} \lambda(n) \lambda(m) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \cos \frac{\pi(2i+1)n}{2N} \cos \frac{\pi(2j+1)m}{2N} I(i, j)
\]

\[\Leftrightarrow \langle \varphi_{n,m}, I \rangle\]

The inverse DCT is defined by:

\[
I(x, y) = \frac{2}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \lambda(n) \lambda(m) \cos \frac{\pi(2x+1)n}{2N} \cos \frac{\pi(2y+1)m}{2N} DCT(I)(n, m).
\]

Two dimensional DCT basis. The source data (8x8) is transformed to a linear combination of these 64 frequency squares.
DCT

Example

Original  Normalized histogram  DCT  Dist. DCT coeff.

Input (YUV 420, 8x8 grid)
DCT
Quantization
Coding of the coefficients AC
Coding of the coefficients DC
Example
Summary
Introduction
JPEG
What does MPEG-2 standard tell us?

Quantization

Block diagram of JPEG Baseline

Image → 8x8 DCT → Q → Predictive coding → VLC

DC index → AC index → Zig-Zag Scan → RLE → VLC

Quantization tables

DC Huffman table
AC Huffman table

Bits

Quantization

\[ F_Q(u, v) = \left\lfloor \frac{F(u, v) + Q(u, v)}{2} \right\rfloor \]

with, \( F \) the unquantized DCT coefficients, \( \lfloor . \rfloor \) the floor function, \( Q \) the typical quantization matrix, as specified in the original JPEG Standard.
Quantization

Quantization table $Q$

Luminance Table:

$$QL = \begin{bmatrix}
16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\
12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\
14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\
14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\
18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\
24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\
49 & 64 & 78 & 87 & 103 & 121 & 120 & 77 \\
92 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \\
\end{bmatrix}$$

Chrominance Table:

$$QC = \begin{bmatrix}
17 & 18 & 24 & 47 & 99 & 99 & 99 & 99 \\
18 & 21 & 26 & 66 & 99 & 99 & 99 & 99 \\
24 & 26 & 56 & 99 & 99 & 99 & 99 & 99 \\
\end{bmatrix}$$

Deduced from CSF (Contrast Sensitivity Function). CSF gives the sensitivity of observers to luminance-varying sinusoidal gratings of different spatial frequencies. $CSF = \frac{1}{TH}$, $TH$ minimal amplitude to detect the target.

Notice that we can use a quality factor to scale the quantization (tradeoff between quality and compression ratio).
Quantization

Example (Extracted from Digital Images Compression Techniques, Majid Rabbani and Paul W. Jones)

\[
f = \begin{bmatrix}
139 & 144 & 149 & 153 & 155 & 155 & 155 \\
144 & 151 & 153 & 156 & 159 & 156 & 156 \\
150 & 155 & 160 & 163 & 158 & 156 & 156 \\
159 & 161 & 162 & 160 & 160 & 159 & 159 \\
159 & 160 & 161 & 162 & 162 & 155 & 155 \\
161 & 161 & 161 & 160 & 157 & 157 & 157 \\
162 & 162 & 161 & 163 & 162 & 157 & 157 \\
162 & 162 & 161 & 163 & 158 & 158 & 158 
\end{bmatrix}
\]

\[
F = \begin{bmatrix}
1260 & -1 & -12 & -5 & 2 & -2 & -3 & 1 \\
-23 & -17 & -6 & -3 & -3 & 0 & 0 & -1 \\
-11 & -9 & -2 & 2 & 0 & -1 & -1 & 0 \\
-7 & -2 & 0 & 1 & 1 & 0 & 0 & 0 \\
-1 & -1 & 1 & 2 & 0 & -1 & 1 & 1 \\
2 & 0 & 2 & 0 & -1 & 1 & 1 & -1 \\
-1 & 0 & 0 & -1 & 0 & 2 & 1 & -1 \\
-3 & 2 & -4 & -2 & 2 & 1 & -1 & 0 
\end{bmatrix}
\]
Example (Extracted from Digital Images Compression Techniques, Majid Rabbani and Paul W. Jones)

**Transform block**

\[
F = \begin{bmatrix}
1260 & -1 & -12 & -5 & 2 & -2 & -3 & 1 \\
-23 & -17 & -6 & -3 & -3 & 0 & 0 & -1 \\
-11 & -9 & -2 & 2 & 0 & -1 & -1 & 0 \\
-7 & -2 & 0 & 1 & 1 & 0 & 0 & 0 \\
-1 & -1 & 1 & 2 & 0 & -1 & 1 & 1 \\
2 & 0 & 2 & 0 & -1 & 1 & 1 & -1 \\
-1 & 0 & 0 & -1 & 0 & 2 & 1 & -1 \\
-3 & 2 & -4 & -2 & 2 & 1 & -1 & 0 \\
\end{bmatrix}
\]

**Typical quantization matrix**

\[
Q_L = \begin{bmatrix}
16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\
12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\
14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\
14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\
18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\
24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\
49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\
2 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \\
\end{bmatrix}
\]

**Transform quantized block**

\[
F_Q = \begin{bmatrix}
79 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
-2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]
Coding of the coefficients AC

**Block diagram of JPEG Baseline**

- **Image** → **8x8 DCT** → **Q** → **Predictive coding** → **VLC** → **Bits**
- **DC index** -> **DC Huffman table**
- **AC index** -> **AC Huffman table**
- **Zig-Zag Scan** → **RLE** → **VLC** → **Bits**
- **Quantization tables**

**Zig-Zag Scan**

\[ F_Q = \begin{bmatrix}
XX & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
-2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix} \]

The zig-zag scan gives: 0,-2,-1,-1,-1,0,0,-1,EOB. EOB stands for End Of Block (all subsequent coefficients are zero).
Coding of the coefficients AC

Block diagram of JPEG Baseline

![Block diagram of JPEG Baseline]

RLE - Run Length Coding

As many of quantized AC coefficients become zero, they can be very efficiently encoded by exploiting the run of zeros. The run-length of zeros are identified by the non-zero coefficients. A sequence of identical symbols is called a run. Each run is represented by a single codeword. The sequence 0,-2,-1,-1,-1,0,0,-1,EOB of the previous example is thus encoded:

\[(1,-2);(0,-1);(0,-1);(0,-1);(2,-1);EOB.\]
Coding of the coefficients DC (zero frequency)

DPCM - predictive coding

\[
DPCM(t_1) = DC_{t_1} - DC_{t_0}
\]

To take into account the spatial redundancy of the average luminance.
Introduction

What does MPEG-2 standard tell us?

Example

Example (JPEG compression)

(a) $CR = 2.6 : 1$

(b) $CR = 23 : 1$

(c) $CR = 144 : 1$

$CR = \text{Compression rate} = \frac{\text{Nb bits of input}}{\text{Nb bits of output}}$

What does MPEG-2 standard tell us?

Presentation/general scheme of JPEG (baseline)
- Input (YUV 420, 8x8 grid)
- DCT
- Quantization
- Coding of the coefficients AC
- Coding of the coefficients DC
- Example

Summary

General layout of the JPEG coding scheme

- Still color picture coding;
- RGB to YUV color-space conversion;
- 2D DCT on block 8x8;
- Quantization by table quantization (Quality factor);
- RLE and Huffman code the non-zero quantized DCT coefficients.
Introduction

What does MPEG-2 standard tell us?

- Motion-compensated hybrid coding
- Hierarchical syntax - types of pictures - GOP
- Input format (4:2:0, progressive, interlaced)
- DCT, Quantization, Zig-zag scan
- Macroblock coding and MC Prediction
- Bit-stream structure
MPEG = Moving Picture Experts Group
Part of the International Standards Organization (ISO)

Targeted applications

MPEG-2 is dedicated for digital storage media and broadcast. The targeted bit rate is in the range 1 to 20 Mbps:

- 1 to 6 Mbps: digital television broadcastion (SD);
- 5 to 8 Mbps: DVD video;
- 10 to 20 Mbps: digital television broadcastion (HD).

Brief history

MPEG-2 (ISO/IEC 13818):

- 13818-1: Systems;
- 13818-2: Video;
- 13818-3: Audio;
- 13818-4: Conformance;
- 13818-5: Software;
- 13818-6: Digital Storage Media;
- 13818-7: Non-Backward Compatible Audio;
- ...

MPEG-2 = MPEG-1 (ISO/IEC 11172)
+ Interlace Tools (Field picture, DCT, prediction...)
+ Profiles & Levels
Motion-compensated hybrid coding scheme

Typical MPEG Encoder Structure

- Input
- Prediction error
- Predicted image
- Null
- MC prediction
- Motion Estimation

DCT

Q

Prediction choice

Frame Memory 1

Frame Memory 2

Reconstructed image

Motion vectors

Huffman

Open to invention and proprietary techniques

DCT

Q

Q⁻¹

DCT⁻¹

Binary stream
What does MPEG-2 standard tell us?

- JPEG
- Motion-compensated hybrid coding
- Hierarchical syntax - types of pictures - GOP
- Input format (4:2:0, progressive, interlaced)
- DCT, Quantization, Zig-zag scan
- Macroblock coding and MC Prediction
- Bit-stream structure

Hierarchical syntax

MPEG structure

- Video sequence
- Group of pictures
- Slice
- Picture
- Macroblock 16x16
- Block 8x8
Slices

Definition (Slice)

A slice is composed of an arbitrary number of consecutive macroblocks.

- The first and last macroblocks of a slice shall not be skipped macroblocks;
- Every slice shall contain at least one macroblock;
- Slices shall not overlap;
- The position of slices may change from picture to picture;
- The first and last macroblock of a slice shall be in the same horizontal row of macroblocks.
Introduction

JPEG

What does MPEG-2 standard tell us?

Macroblock

Three macroblock structures:

(a) 420

(b) 422

(c) 444
Types of pictures

- **I, P, and B pictures**
  - Intra picture (I): intra-frame spatial DCT;
  - Predicted picture (P): DCT with forward prediction (residual coding);
  - Bi-directional picture (B): DCT with bi-directional prediction (residual coding).

These three types of pictures are used to form the GOP (Group of Pictures).
Group of Pictures

**GOP N-M**

N is the I picture interval and M is the anchor picture interval (M-1 B pictures between anchor pictures).

- A GOP must contain a I picture;
- B pictures must be located between anchor pictures (I or P);
- A GOP must start with a I picture in **coding order**;
- A GOP must start with a I or B picture and must end with an I or P picture in **display order**.

**Example (GOP in coding order)**

<table>
<thead>
<tr>
<th>GOP 1-1</th>
<th>I</th>
<th>I</th>
<th>I</th>
<th>I</th>
<th>I</th>
<th>I</th>
<th>I</th>
<th>I</th>
<th>I</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOP 6-2</td>
<td>I</td>
<td>B</td>
<td>P</td>
<td>B</td>
<td>P</td>
<td>B</td>
<td>I</td>
<td>B</td>
<td>P</td>
</tr>
<tr>
<td>GOP 12-3</td>
<td>I</td>
<td>B</td>
<td>B</td>
<td>P</td>
<td>B</td>
<td>B</td>
<td>P</td>
<td>B</td>
<td>B</td>
</tr>
</tbody>
</table>
Display and coding order

- **Display order**: input of the encoder and output of the decoder;
- **Coding order**: output of the encoder and input of the decoder.

### Display order (input encoder)

- $B_1$, $B_2$, $I_1$, $B_3$, $B_4$, $P_1$, $B_5$, $B_6$, $P_2$, $B_7$, $B_8$, $P_3$, $B_9$

### Coding order (output encoder)

- $I_1$, $B_1$, $B_2$, $P_1$, $B_3$, $B_4$, $P_2$, $B_5$, $B_6$, $P_3$, $B_7$, $B_8$

### Display order (output decoder)

- $B_1$, $B_2$, $I_1$, $B_3$, $B_4$, $P_1$, $B_5$, $B_6$, $P_2$, $B_7$, $B_8$
YUV 4:2:0

- Color space YUV 4:4:4:
  \[
  \begin{bmatrix}
  Y \\
  U \\
  V
  \end{bmatrix} =
  \begin{bmatrix}
  0.299 & 0.587 & 0.114 \\
  -0.1687 & -0.3313 & 0.5 \\
  0.5 & -0.4187 & -0.813
  \end{bmatrix}
  \begin{bmatrix}
  R \\
  G \\
  B
  \end{bmatrix}
  \]

- Format 4:2:0 (Human eyes are less sensitive to the chrominance than to the luminance):
  - (a) Source
  - (b) Y
  - (c) U (Blue)
  - (d) V (Red)
Progressive vs Interlaced

- **Progressive**: one time instant is required to acquire the picture;
- **Interlaced**: two time instants are required to acquire the picture (Two fields, field period = frame period / 2).

Interlaced format
Interlaced format

Progressive vs Interlaced

Interlaced scanning is a way to save bit rate (almost invisible for us due to the vision persistence and our critical flicker frequency).

- Interlaced scanning provides a high vertical resolution for still scenes (similar to progressive format);
- Artifacts for moving areas.

$\tau$ $\tau + 20 \text{ ms}$
DCT

Orthogonal transform of 8x8 pixel block into 8x8 frequency coefficient matrix

The DCT of a block \( I \) of size \( N \times N \) is defined by:

\[
DCT(I)(n, m) = \frac{2}{N} \lambda(n)\lambda(m) \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \cos \left( \frac{\pi(2i+1)n}{2N} \right) \cos \left( \frac{\pi(2j+1)m}{2N} \right) I(i, j)
\]

The inverse DCT is defined by:

\[
I(x, y) = \frac{2}{N} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \lambda(n)\lambda(m) \cos \left( \frac{\pi(2x+1)n}{2N} \right) \cos \left( \frac{\pi(2y+1)m}{2N} \right) DCT(I)(n, m).
\]

\[
\lambda(t) = \begin{cases} 
\frac{1}{\sqrt{2}} & \text{if } t = 0 \\
1 & \text{otherwise}
\end{cases}
\]

Two dimensional DCT basis. The source data (8x8) is transformed to a linear combination of these 64 frequency squares.
Introduction

JPEG

What does MPEG-2 standard tell us?

DCT

Motion-compensated hybrid coding
Hierarchical syntax - types of pictures - GOP
Input format (4:2:0, progressive, interlaced)
DCT, Quantization, Zig-zag scan
Macroblock coding and MC Prediction
Bit-stream structure

Example

Original  Normalized histogram  DCT  Dist. DCT coeff.

Original  Normalized histogram  DCT  Dist. DCT coeff.

DC coeff.  AC coeff.
Frame or field DCT?

Frame DCT is the transformation mode used in MPEG-1 as illustrated below. Field DCT is applied on the field (see below):

You have the choice between these two different DCT. What is the general idea to be efficient?
Quantization

Definition (Scalar quantization)

\[ Q : \mathcal{X} \quad \longrightarrow \quad C = \{y_i, i = 1, 2, \ldots N\} \]

\[ x \quad Q(x) = y_i \]

- \( N \) is the number of quantization level;
- \( \mathcal{X} \) discret;
- \( C \) is always discret (codebook, dictionary);
- \( \text{card}(\mathcal{X}) > \text{card}(C) \);
- As \( x \neq Q(x) \), we will lost some information (lossy compression).
Quantization

Two different quantizer scale

- Linear quantizer scale (qscaletype=0)
- Non-linear quantizer scale (qscaletype=1)
Quantization

Quantization matrix

- Possibility to use quantization matrix to weight the quantized coefficients.
- This matrix is based on the Contrast Sensitivity Function (coarser quantization of high spatial frequencies without visual annoyance).
- Default matrices are specified by the standard. Therefore, it is not necessary to send them. Not the case for proprietary matrix.

Example (Illustration of our visual sensitivity (CSF))

Psychophysi experiments seek to determine whether the subject can detect a stimulus.
Quantization

Default quantization matrix

Default matrix for intra coding:

\[
Q_I = \begin{bmatrix}
8 & 16 & 19 & 22 & 26 & 27 & 29 & 34 \\
16 & 16 & 22 & 24 & 27 & 29 & 34 & 37 \\
19 & 22 & 26 & 27 & 29 & 34 & 34 & 38 \\
22 & 22 & 26 & 27 & 29 & 34 & 37 & 40 \\
22 & 26 & 27 & 29 & 32 & 35 & 40 & 48 \\
26 & 27 & 29 & 32 & 35 & 40 & 48 & 58 \\
26 & 27 & 29 & 34 & 38 & 46 & 56 & 69 \\
27 & 29 & 35 & 38 & 46 & 56 & 69 & 83 \\
\end{bmatrix}
\]

Non intra matrix:

\[
Q_{NI} = \begin{bmatrix}
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
16 & 16 & 16 & 16 & 16 & 16 & 16 & 16 \\
\end{bmatrix}
\]

Non-intra quantization Matrix (MPEG-2 Test Model 5)

\[
Q_{NI} = \begin{bmatrix}
16 & 17 & 18 & 19 & 20 & 21 & 22 & 23 \\
17 & 18 & 19 & 20 & 21 & 22 & 23 & 24 \\
18 & 19 & 20 & 21 & 22 & 23 & 24 & 25 \\
19 & 20 & 21 & 22 & 23 & 24 & 25 & 27 \\
20 & 21 & 22 & 23 & 25 & 26 & 27 & 28 \\
21 & 22 & 23 & 24 & 26 & 27 & 28 & 30 \\
22 & 23 & 24 & 26 & 27 & 28 & 30 & 31 \\
23 & 24 & 25 & 27 & 28 & 30 & 31 & 33 \\
\end{bmatrix}
\]
Two scan patterns are used in MPEG-2

- Normal zig-zag scan as defined in MPEG-1 (below on the left);
- Alternate zig-zag scan. This new scan pattern can be used when the frame DCT is applied on an interlaced video. There are more DCT coefficients for the highest vertical frequencies.
Macroblock coding

I pictures
- All macroblocks are encoded in INTRA.

P pictures
- Macroblocks can be encoded in INTRA;
- Macroblocks can be encoded in INTER (with a forward prediction);

B pictures
- Macroblocks can be encoded in INTRA (fallback mode);
- Macroblocks can be encoded in INTER:
  - forward prediction;
  - backward prediction;
  - bidir prediction.
Motion estimation and Motion compensation (MC).

An image of the sequence is defined by $I(x, y, t)$. $(x, y)$ represents the spatial coordinates whereas $t$ represents the time.

**Fundamental assumption**

The image intensity is conserved along trajectories !!!

$$I(x, y, t) = I(x + \delta_x, y + \delta_y, t + \delta_t)$$

Classification:

- **Feature / Region Matching**: the motion field is estimated by correlating features (edge, intensity...) from one frame to another (Block Matching, Phase correlation...);

- **Gradient-based methods**: the motion field is estimated by using spatio-temporal gradients of the image intensity distribution (Pel-recursive method, the Horn-Schunck algorithm...).
Motion models

Motion model uses in MPEG-2 is 2D Translation (2 parameters, this model dealing only with the translation is used in video coding (works quite well because motion between consecutive frames is rather small)).

\[
\begin{bmatrix}
u \\
v
\end{bmatrix} = \begin{bmatrix}
x \\
y
\end{bmatrix} + \begin{bmatrix}
dx \\
dy
\end{bmatrix}
\]

Rotation, scaling and deformation are not taken into account to evaluate the displacement.
Backward prediction

- A backward-predicted macroblock depends on decoded pixels from the immediately following anchor picture;
- This mode can only be used to encode macroblocks in B pictures.
What does MPEG-2 standard tell us?

Motion-compensated hybrid coding
Hierarchical syntax - types of pictures - GOP
Input format (4:2:0, progressive, interlaced)
DCT, Quantization, Zig-zag scan
Macroblock coding and MC Prediction
Bit-stream structure

MC Prediction

Bi-directional prediction

- A bi-directionally-predicted macroblock depends on decoded pixels from the anchor pictures immediately following and immediately preceding;
- This mode can only be used to encode macroblocks in B pictures.
### MB modes

#### I picture

- **Intra:**
  - DCT frame
  - DCT field

#### P picture

- **Forward frame:**
  - DCT frame
  - DCT field
- **Forward field:**
  - DCT frame
  - DCT field
- **Intra:**
  - DCT frame
  - DCT field
- **NoMC:**
  - Skip MB if and only if the quantized coefficients are null and the motion vector is null.
What does MPEG-2 standard tell us?

**MB modes**

**B picture**

- **Forward frame:**
  - → DCT frame
  - → DCT field

- **Forward field:**
  - → DCT frame
  - → DCT field

- **Backward frame:**
  - → DCT frame
  - → DCT field

- **Backward field:**
  - → DCT frame
  - → DCT field

- **Bidir frame:**
  - → DCT frame
  - → DCT field

- **Bidir field:**
  - → DCT frame
  - → DCT field

- **Intra:**
  - → DCT frame
  - → DCT field

- **Skip MB:**
  - the quantized coefficients are null;
  - the motion vector is null.