# *ELEC 691X/498X – Broadcast Signal Transmission*
# *Fall 2015*

**Instructor:** Dr. Reza Soleymani, Office: EV-5.125, Telephone: 848-2424 ext.: 4103.

**Office Hours:** Wednesday, Thursday, 14:00 – 15:00

**Time:** Tuesday, 2:45 to 5:30

**Room:** H 411

# Lecture 3:
# Video Compression

In this lecture we cover the following topics:

- Intra-frame Compression in MPEG.
- Inter-frame Compression in MPEG.
- MPEG Profiles.
- MPEG Levels

# Lecture 3:
# MPEG: History

Moving Picture Experts Group (MPEG) is a working group organized by ISO and IEC. The idea of creating this joint working group was initially suggested first by NTT. The first meeting of MPEG was held in May 1988 in Ottawa. MPEG has currently over 350 members. Following is a list of standards generated by the MPEG working group

- MPEG-1 (1993): This is mainly limited to 1.5 Mbps and is intended for Video CD. While its video compression is now being replaced by newer generations of the standard, its audio compression technique has survived as mp3 (Audio Level III) format.
- MPEG-2 (1995): This format supports broadcast quality TV and can transport interlacing and HD video. It has been adopted by most broadcasting standards such as DVB, ATSC, cable TV as well as DVD disc standard.
- MPEG-3: MPEG-3 dealt mainly with scalability and multi-resolution compression. As such it was added to MPEG-2 and was never become a standalone standard.

# Lecture 3:
# MPEG: History

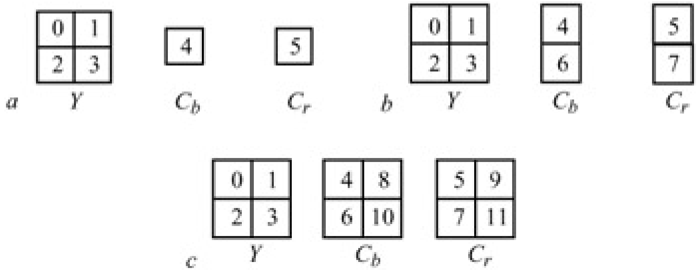
- MPEG-4 (1998): MPEG-4 incorporates further coding tools to make the encoding more efficient. They include:
  - MPEG-4 Part 2 (or Simple and Advanced Simple Profile) and
  - MPEG-4 AVC (or MPEG-4 Part 10 or H.264). MPEG-4 AVC may be used on HD DVD and Blu-ray Discs, along with VC-1 and MPEG-2.
- HEVC (2013): High Efficiency Video Coding (HEVC) or H.265 is a successor to H.264/MPEG-4 AVC . It has not yet been implemented fully. So, most of its efficiency has not yet been fully realized. A full fledge hardware implementation will require 50% less bandwidth compared to H.264.
- In addition there are several standards such as MPEG-7, MPEG-21, MPEG-A, MPEG-B, etc. that provide interface for different environments or tools for intellectual management tools, etc.

# *Lecture 3:*
# *MPEG: Compression*

- As a three dimensional source (2 spatial dimensions and time as the third dimension), a video signal can not only be compressed spatially, but also in time domain. The spatial compression is the same as what we saw in the previous lecture for picture coding using JPEG. It is called **intra-frame** compression. Intra-frame compression takes advantage of the correlation that exists between the different samples of the images (pixels).
- Temporal aspect of video compression, called **inter-frame** coding uses the fact that the different frames of a video source are correlated and uses this correlation in order to remove the redundancy between different frames. In MPEG standard there are three types of frames: I-frame, P-frame and B-frame.

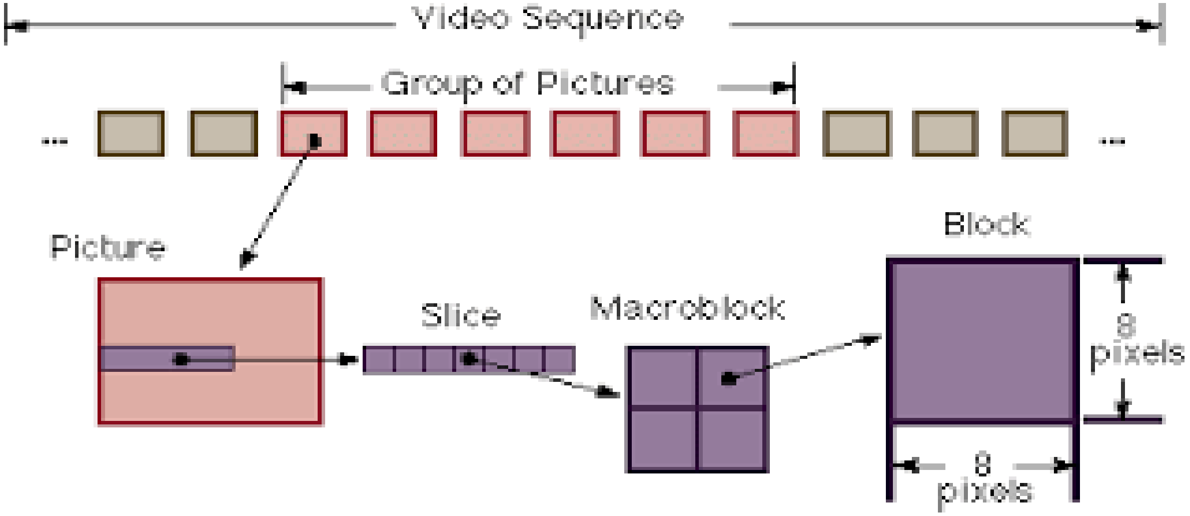
# Lecture 3:
# MPEG: Intra-frame Compression

- We saw in the last lecture that the transform block of a JPEG encoder is an 8 by 8 block of pixels. Since in a picture or a frame of video we may use less chroma samples than the lumina samples, i.e., using 4:2:0 or 4:2:2 instead of 4:4:4, the processing unit in JPEG and MPEG is a **macroblock**. A macroblock is a 16 by 16 bock of pixel or in another word, a set of 4 ( 2x2) basic 8x8 blocks. Below you see the arrangement of blocks for 4:2:0, 4:2:2 and 4:4:4:



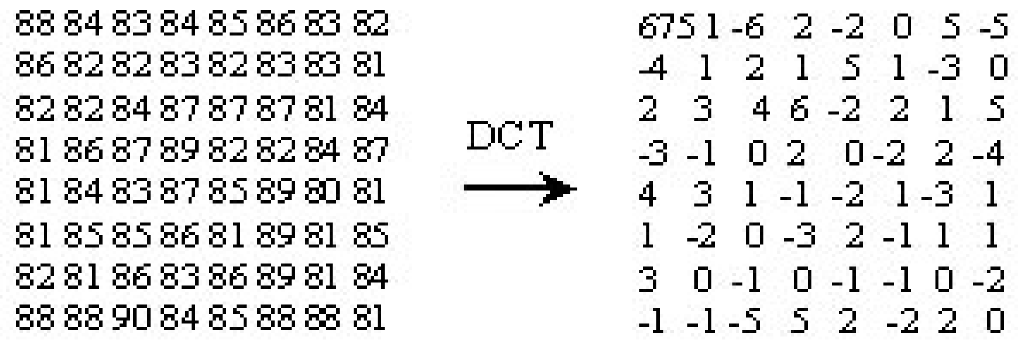
# Lecture 3:
# MPEG: Intra-frame Compression

- Let's now start from some content (a video sequence) that we wish to compress. First we divide it into several group of pictures (GOP) that from which we later will try to remove redundancy. Then we slice each picture into several macroblocks which in turn will be divided into four transform blocks.



# *Lecture 3:*
# *MPEG: Intra-frame Compression*

- As we discussed last week, each block is then scanned in a zig-zag manner to generate an array of 64 correlated values.
- The above 64 numbers will be then transformed into 64 transform (frequency) domain values.



| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 88 | 84 | 83 | 84 | 85 | 86 | 83 | 82 |
| 86 | 82 | 82 | 83 | 82 | 83 | 83 | 81 |
| 82 | 82 | 84 | 87 | 87 | 87 | 81 | 84 |
| 81 | 86 | 87 | 89 | 82 | 82 | 84 | 87 |
| 81 | 84 | 83 | 87 | 85 | 89 | 80 | 81 |
| 81 | 85 | 85 | 86 | 81 | 89 | 81 | 85 |
| 82 | 81 | 86 | 83 | 86 | 89 | 81 | 84 |
| 88 | 88 | 90 | 84 | 85 | 88 | 88 | 81 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 675 | 1 | -6 | 2 | -2 | 0 | 5 | -5 |
| -4 | 1 | 2 | 1 | 5 | 1 | -3 | 0 |
| 2 | 3 | 4 | 6 | -2 | 2 | 1 | 5 |
| -3 | -1 | 0 | 2 | 0 | -2 | 2 | -4 |
| 4 | 3 | 1 | -1 | -2 | 1 | -3 | 1 |
| 1 | -2 | 0 | -3 | 2 | -1 | 1 | 1 |
| 3 | 0 | -1 | 0 | -1 | -1 | 0 | -2 |
| -1 | -1 | -5 | 5 | 2 | -2 | 2 | 0 |

- The value on the left top corner (the DC coefficient) is the most prominent frequency domain value. This is the mean value of the pixels in the block. They show high correlation from block to block. The DC components are coded using DPCM.

# Lecture 3:
# MPEG: Intra-frame Compression

- The transform domain coefficients are divided component-wise by a quantization matrix such as:

$$\begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

- The entries of this matrix determines how much we compress the picture. Note that the higher the values towards the left-bottom are the more higher frequencies are suppressed.

# Lecture 3:
# MPEG: Intra-frame Compression

- As an example, assume that we have the following frequency domain matrix:

$$\begin{bmatrix} -415 & -33 & -58 & 35 & 58 & -51 & -15 & -12 \\ 5 & -34 & 49 & 18 & 27 & 1 & -5 & 3 \\ -46 & 14 & 80 & -35 & -50 & 19 & 7 & -18 \\ -53 & 21 & 34 & -20 & 2 & 34 & 36 & 12 \\ 9 & -2 & 9 & -5 & -32 & -15 & 45 & 37 \\ -8 & 15 & -16 & 7 & -8 & 11 & 4 & 7 \\ 19 & -28 & -2 & -26 & -2 & 7 & -44 & -21 \\ 18 & 25 & -12 & -44 & 35 & 48 & -37 & -3 \end{bmatrix}$$

Dividing it by the quantization matrix in the previous slide, we get:

$$\begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -3 & 4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -4 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

# Lecture 3:
# MPEG: Intra-frame Compression

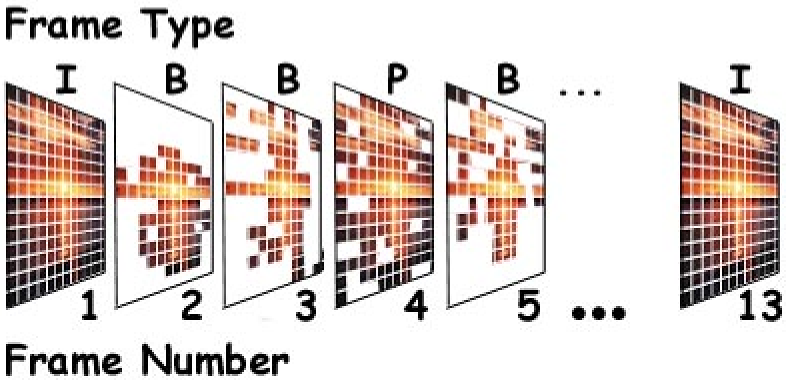
- The remaining (non-zero) AC coefficients are then coded using a Variable Length Code (VLC) also known as entropy encoder or Huffman code.
- We can see that given the small number of significant frequency coefficients as well as their small dynamic range results in a great reduction in the number of bits required to represent a frame.
- In the next few slides we talk about further rate reduction taking into consideration inter-frame correlation.

# Lecture 3:
# MPEG: Inter-frame Compression

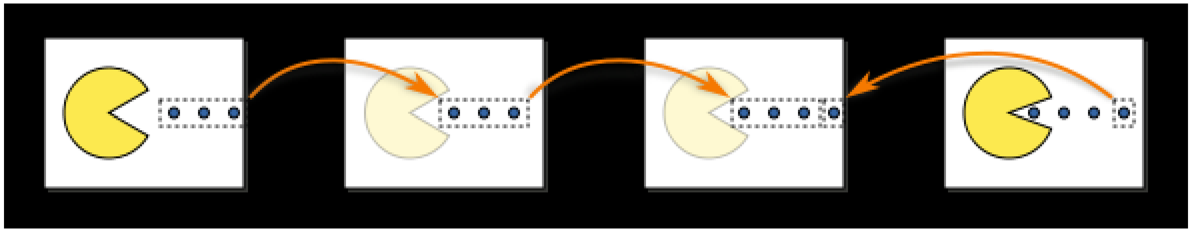
- Inter-frame compression is achieved using a technique called **Motion Compensation**. Motion compensation takes advantage of the fact that in a video (or a movie)for many frames the only difference between a frame and the next is mainly due to the movement of camera or an object in the frame.
- The technique used in video compression is called BMC (Block Motion Compensation). The frames are divided into blocks of pixels. For example, in MPEG into 16x16 macroblocks. Each block is predicted from a block of equal size in a reference block. The blocks are then shifted to the position of the predicted block. This shift is represent by a motion vector (MC). Since there is correlation between adjacent blocks, one can encode only the difference between motion vectors.

# Lecture 3:
# MPEG: Inter-frame Compression

- **I-frames** (Intra-coded frames) are the ones least compressible. They are the frames only gone through spatial (intra-frame) compression.
- A **P-frame** (predictive frame) is one using data from previous frames to decompress and is more compressible than an I-frame.



- **B-frames** (Bi-predictive frames)can use both previous and future frames and get compressed more than P-frames.



# *Lecture 3: MPEG: Compression*

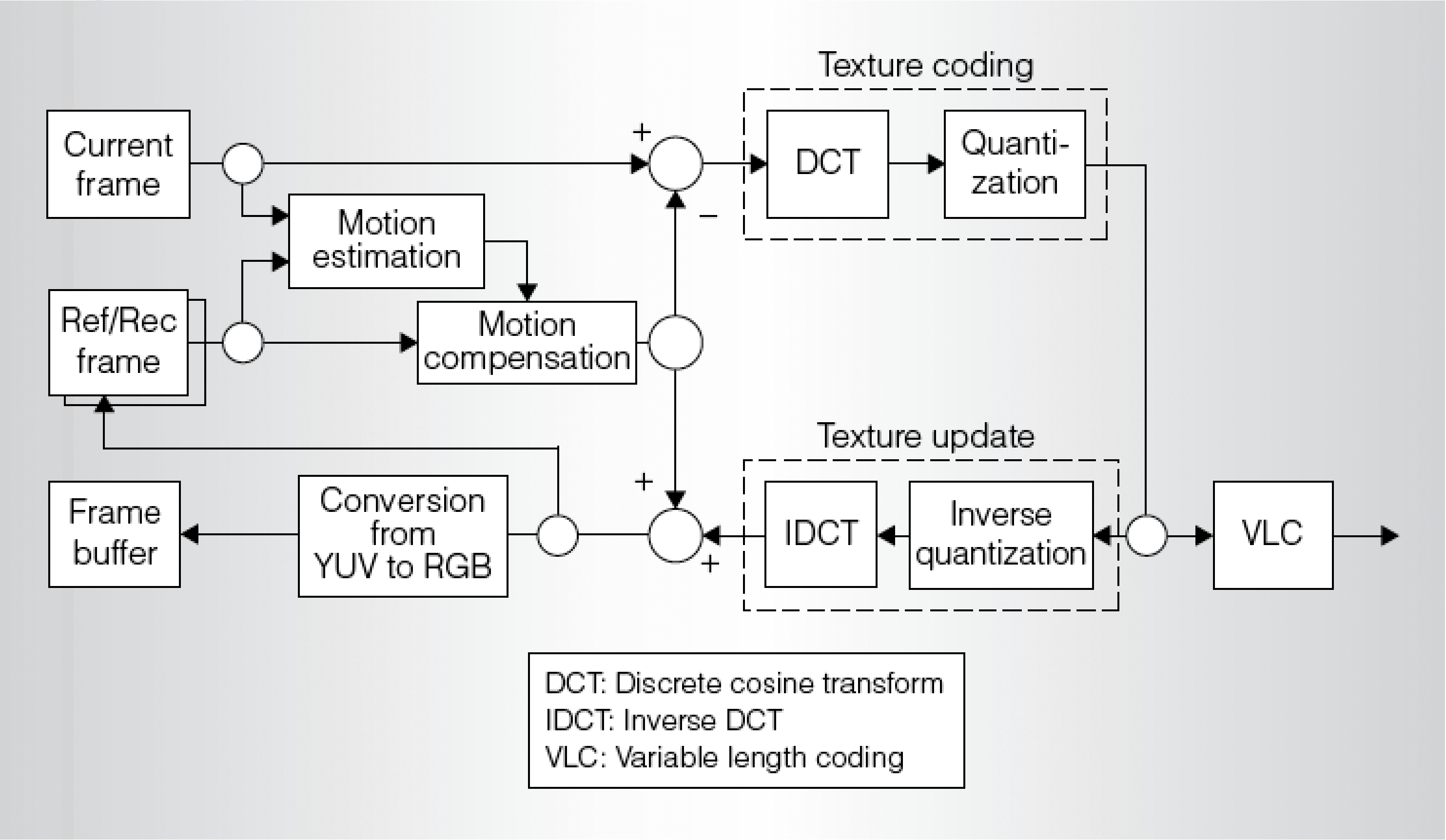
- Examples of frame sequences:

  - I B B P B B I B B P B B I
  - I I I I I I I I I I I I I
  - I P I P I P I P I P I P I
  - I P I P I I I P I P I I I

Exercise 3.1: Find the decoding delay for each of the above examples if the frame rate is 60 frames per second.

# Lecture 3:
# MPEG: Encoder structure

- The bock diagram of a typical MPEG encoder:



# Lecture 3:
# MPEG: Profiles

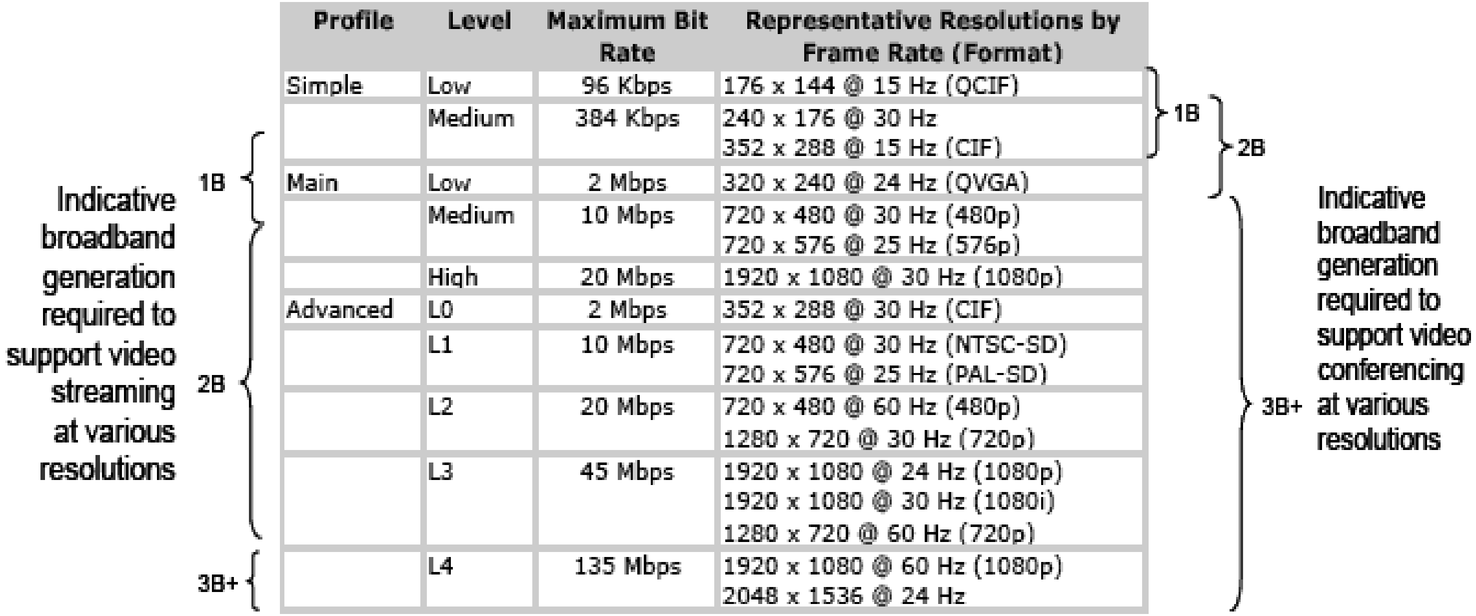
- A **Profile** is a set of parameters defined by the standard for different applications. Following is a list of some of the MPEG profiles
- **Baseline Profile** (BP): This profile is also intended for low cost applications. But it has some extra features making it suitable for low cost applications requiring data loss robustness.
- **Main Profile** (MP): This profile is intended for standard definition TV when using MPEG-4.
- **High Profile** (HiP): This is the primary profile for broadcast and disc storage applications, particularly for high-definition television.
- **High 10 Profile** (Hi10P): This profile builds on top of the High Profile, adding support for up to 10 bits per sample of decoded picture precision.
- **High 4:2:2 Profile** (Hi422P): This profile targets professional applications that use interlaced video, this profile has support for the 4:2:2 chroma subsampling format and uses up to 10 bits per sample.
- **High 4:4:4 Predictive** Profile (Hi444PP): This profile supports up to 4:4:4 chroma sampling, up to 14 bits per sample.

# Lecture 3: MPEG: Profiles

| | BP | MP | HiP | Hi10P | Hi422P | Hi444PP |
|---|---|---|---|---|---|---|
| **Bit depth** | 8 | 8 | 8 | 8 to 10 | 8 to 10 | 8 to 14 |
| **Chroma formats** | 4:2:0 | 4:2:0 | 4:2:0 | 4:2:0 | 4:2:0/4:2:2 | 4:2:0/4:2:2/ 4:4:4 |
| **Flexible macroblock ordering (FMO)** | Yes | No | No | No | No | No |
| **Arbitrary slice ordering (ASO)** | Yes | No | No | No | No | No |
| **Redundant slices (RS)** | Yes | No | No | No | No | No |
| **Interlaced coding (PicAFF, MBAFF)** | No | Yes | Yes | Yes | Yes | Yes |
| **B slices** | No | Yes | Yes | Yes | Yes | Yes |
| **CABAC entropy coding** | No | Yes | Yes | Yes | Yes | Yes |
| **4:0:0 (Monochrome)** | No | No | Yes | Yes | Yes | Yes |
| **8×8 vs. 4×4 transform adaptivity** | No | No | Yes | Yes | Yes | Yes |
| **Quantization scaling matrices** | No | No | Yes | Yes | Yes | Yes |
| **Separate Cb and Cr QP control** | No | No | Yes | Yes | Yes | Yes |
| **Separate color plane coding** | No | No | No | No | No | Yes |
| **Predictive lossless coding** | No | No | No | No | No | Yes |

# Lecture 3: MPEG: Levels

- A **Level** is a specified set of constraints that indicate a degree of required decoder performance for a profile. A level within a profile, for example, specifies the maximum picture resolution, frame rate, and bit rate that a decoder may use.



| Profile | Level | Maximum Bit Rate | Representative Resolutions by Frame Rate (Format) |
|---|---|---|---|
| Simple | Low | 96 Kbps | 176 x 144 @ 15 Hz (QCIF) |
|  | Medium | 384 Kbps | 240 x 176 @ 30 Hz<br>352 x 288 @ 15 Hz (CIF) |
| Main | Low | 2 Mbps | 320 x 240 @ 24 Hz (QVGA) |
|  | Medium | 10 Mbps | 720 x 480 @ 30 Hz (480p)<br>720 x 576 @ 25 Hz (576p) |
|  | High | 20 Mbps | 1920 x 1080 @ 30 Hz (1080p) |
| Advanced | L0 | 2 Mbps | 352 x 288 @ 30 Hz (CIF) |
|  | L1 | 10 Mbps | 720 x 480 @ 30 Hz (NTSC-SD)<br>720 x 576 @ 25 Hz (PAL-SD) |
|  | L2 | 20 Mbps | 720 x 480 @ 60 Hz (480p)<br>1280 x 720 @ 30 Hz (720p) |
|  | L3 | 45 Mbps | 1920 x 1080 @ 24 Hz (1080p)<br>1920 x 1080 @ 30 Hz (1080i)<br>1280 x 720 @ 60 Hz (720p) |
|  | L4 | 135 Mbps | 1920 x 1080 @ 60 Hz (1080p)<br>2048 x 1536 @ 24 Hz |

Indicative broadband generation required to support video streaming at various resolutions

Indicative broadband generation required to support video conferencing at various resolutions