# Multimedia Software Systems
# CS4551

Video Compression – ISO Standards (MPEG)



## ISO Standards

- MPEG-1(movies on CD-ROM)
- MPEG-2(digital television, movies on DVD)

- MPEG was established in 1988.
- MPEG-1 was approved in 1991.
- MPEG-2 was approved in 1994.

# ISO MPEG-1 Video

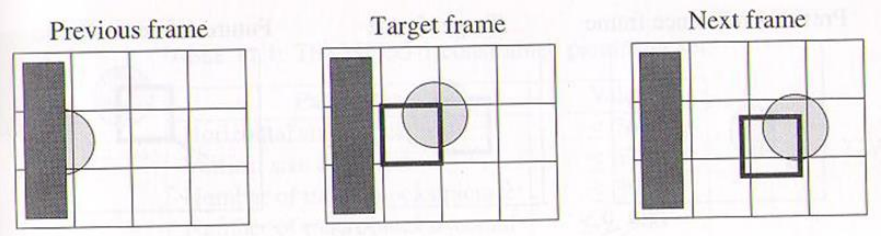
- MPEG-1: true multimedia standard with specifications for coding and transmission of audio, video and data streams in a series of synchronized, mixed packets
- Driving focus: storage of multimedia content on digital storage media (eg. CDs or VCDs)
- Quality: VHS VCR-like video quality and CD audio quality
- Picture format: SIF (Source Input Format)
  - 352x288, 25 fps for PAL (Phase Alternating Line) video
  - 354x240, 30 fps for NTSC (National Television System Committee) video
  - 4:2:0 Chroma sampling



# Motion Compensation in MPEG-1

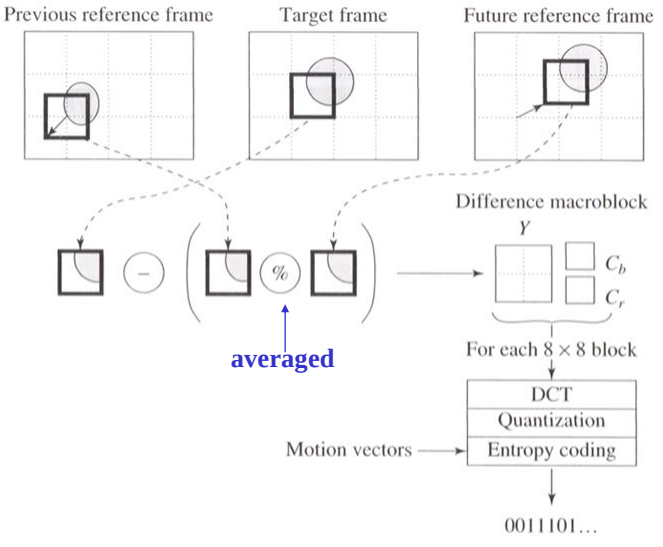
- Coding mechanism similar to H.26x
- Three types of frames:
  - *I-frames* (coded in intra-frame mode)
  - *P-frames* (coded with motion compensation using as reference a previous *I* or *P* frame)
  - *B-frames* (coded with bidirectional motion compensation based on a previous and a later *I* or *P* frame)

# Need for Bidirectional Search




The need for bidirectional search.

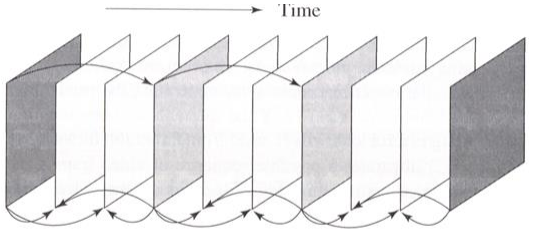





B-frame coding based on bidirectional motion compensation.

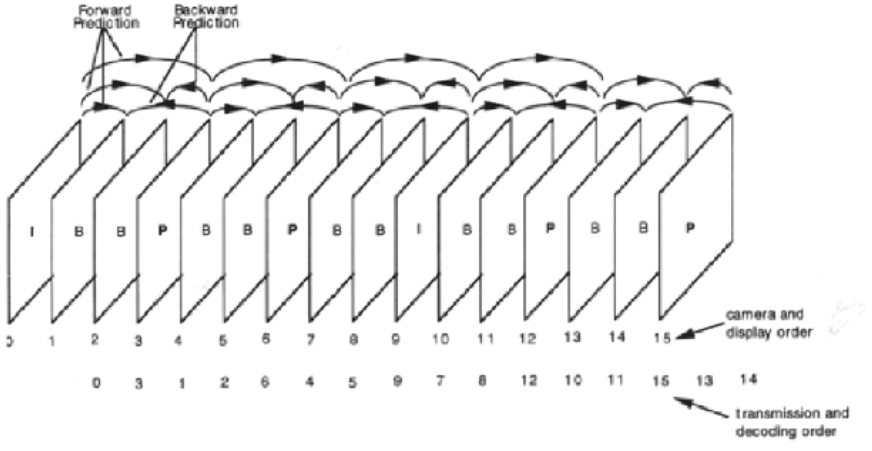
# Motion Compensation in MPEG-1

- For B-frame encoding/decoding, it needs succeeding P- or I-frame, the actual coding and transmission order is different from the display order. => Inevitable delay and need for buffering.






# Acquisition/Transmission Order



# Example Question

- Suppose we encode a sequence of 23 frames using MPEG. Each frame is encoded in I(intra frame), P (forward prediction) or B (bi-directional prediction) mode according to the following order: I P P P B B P I P P P I P P B B B B P I P I P
  - Derive the correct transmission and decoding order. *Answer) 1 2 3 4 7 5 6 8 9 10 11 12 13 14 19 15 16 17 18 20 21 22 23 with respect to the indices of the original sequence starting from 1.*
  - What is the minimum delay (in terms of frame period) due to bi-directional encoding in this case?

    *Answer) 4 frames. The 15th frame in the sequence (a B-frame) cannot be encoded until the 19th frame in the sequence has been encoded.*



# ISO MPEG-1 Video

- Differences from H.261
  - Quantization (allow non-constant quantization interval for the block)
  - Half pixel accuracy
  - Supports larger gaps between I and P frames
  - Bitstream format allows random access

# ISO MPEG-1 Video

- Quantization in MPEG-1

Default quantization table ($Q_1$) for intra-coding.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 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 |

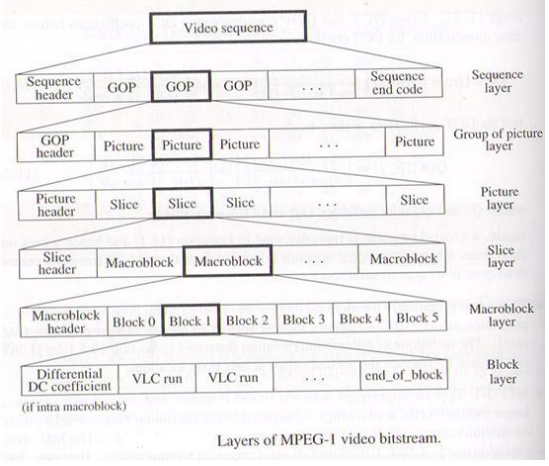
Default quantization table ($Q_2$) for inter-coding.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 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 |

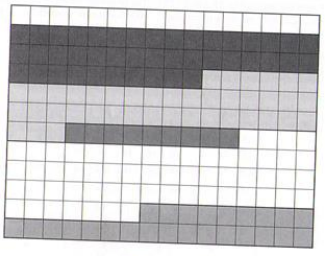


# ISO MPEG-1 Video

- MPEG-1 Video Bitstream




Layers of MPEG-1 video bitstream.



Slices in an MPEG-1 picture.

# ISO MPEG-2

- MPEG-2 is for high-quality video at a bit rate of more than *4Mbps*. It was originally designed as a standard for digital broadcast TV.
- Over time, the MPEG-2 video standard was expanded to include high-resolution video (such as HDTV), as well as *hierarchical* (or *scalable* or *layered*) video coding.
- MPEG-2 has gained wide acceptance beyond broadcasting digital TV over satellite or cable networks.
- It is also adopted for DVDs.



# ISO MPEG-2

- MPEG-2 defined 7 profiles aimed at different applications such as low-delay video conferencing, scalable video, or HDTV.
- 7 profiles: Simple, Main, SNR Scalable, Spatially Scalable, High, 4:2:2 and Multiview.
- Within each profile, up to four levels are defined.

# ISO MPEG-2

Profiles and Levels in MPEG-2.

| Level | Simple profile | Main profile | SNR scalable profile | Spatially scalable profile | High profile | 4:2:2 profile | Multiview profile |
|---|---|---|---|---|---|---|---|
| High | | * | | | * | | |
| High 1440 | | * | | * | * | | |
| Main | * | * | * | | * | * | * |
| Low | | * | * | | | | |

Four levels in the main profile of MPEG-2.

| Level | Maximum resolution | Maximum fps | Maximum pixels/sec | Maximum coded data rate (Mbps) | Application |
|---|---|---|---|---|---|
| High | 1,920 × 1,152 | 60 | $62.7 \times 10^6$ | 80 | Film production |
| High 1440 | 1,440 × 1,152 | 60 | $47.0 \times 10^6$ | 60 | Consumer HDTV |
| Main | 720 × 576 | 30 | $10.4 \times 10^6$ | 15 | Studio TV |
| Low | 352 × 288 | 30 | $3.0 \times 10^6$ | 4 | Consumer tape equivalent |

• DVD video specification uses only a restricted form of the MPEG-2 Main profile at the Main and Low levels



# MPEG-1 vs. MPEG-2

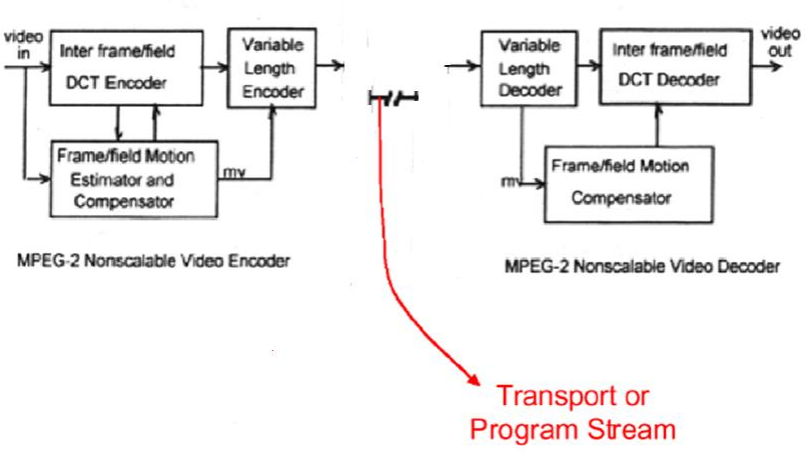
- Compression based on DCT-based transform coding and motion-prediction.
- Differences from MPEG-1
  - Support of interlaced video whereas MPEG-1 supports only non-interlaced (progressive) video. MPEG-2 is adopted by digital broadcast TV such as HDTV, it must support interlaced video.
  - Scalability – make a *single* video stream scalable to various bitrates.
  - Support of 4:2:2 and 4:4:4 chroma sampling
  - More flexible video formats
  - …

# MPEG-2 – Scalability

- The coded representation (bit-stream) is generated in such a way that decoders of various complexities are able to decode video of different resolution/quality from the same bit stream
  - If the bitstream is truly scalable, decoders of different complexities can coexist: *inexpensive decoders* would be expected to decode only small portions of the same bitstream producing basic quality pictures, while *more sophisticated and expensive decoders* will produce higher quality pictures
- For applications in networks with multi-quality video services and windowed video on computer workstations.



# MPEG-2 – Standard Non-Scalable Encoder

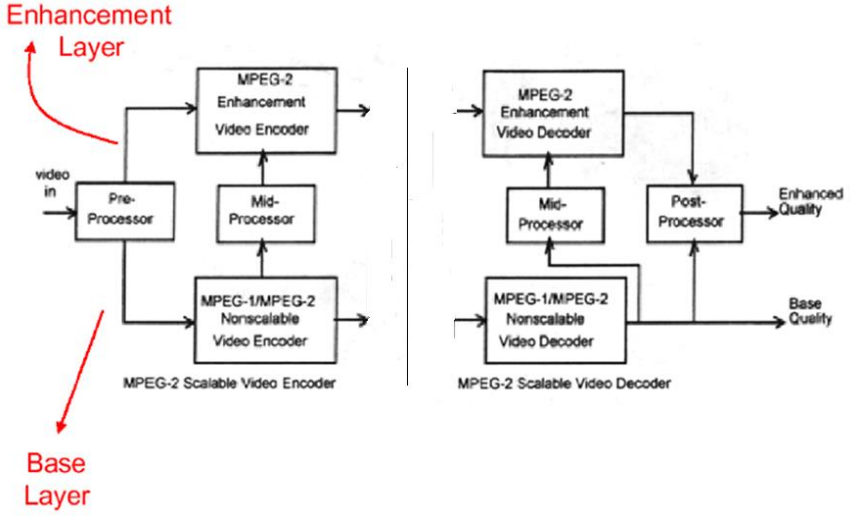


# MPEG-2 – Scalabilities

- MPEG-2 scalable coding is also known as layered coding.
- Input video goes through a pre-processor producing a *base layer* signal and *one or more enhancement layers.*
  - *Base layer* is decoded independently by a standard MPEG-2 non-scalable video encoder with a coarse quantization, which result in fewer bits and a relatively low-quality video.
  - *Enhancement layer* is encoded *with respect to the base layer* by a MPEG-2 *enhancement encoder.*



# 2 Layer Scalable MPEG-2 Video Encoder



# MPEG-2 – Scalability

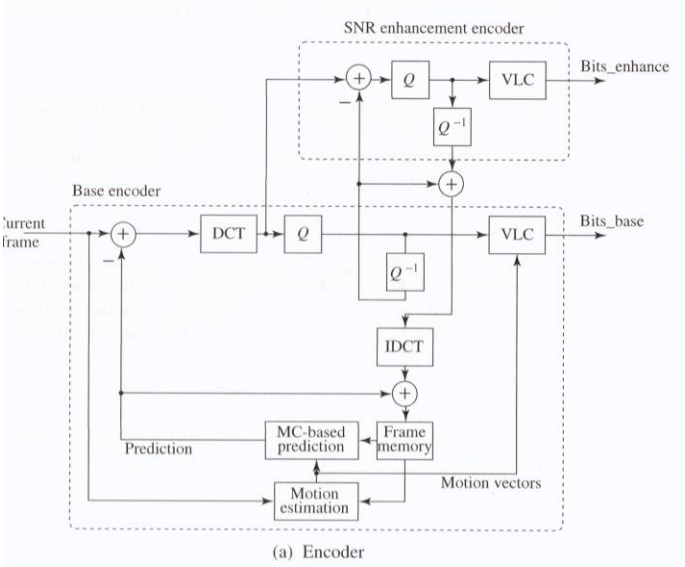
- MPEG-2 supports the following scalabilities
  - SNR scalability
  - Spatial Scalability (like pyramid)
  - Temporal Scalability (higher frame rate)
  - Hybrid scalability (combines any two of the above three)
  - Data partitioning (like Spectral selection)
- Scalability is very useful for
  - Very different bitrates: different network link speeds (fast/slow)
  - Variable-bit rate channel: bitrate of the channel deteriorates sometime
  - Noisy connections: send the base layer via a less-noisy channel
  - Progressive transmission: coarse to finer image



# MPEG-2 – SNR Scalability

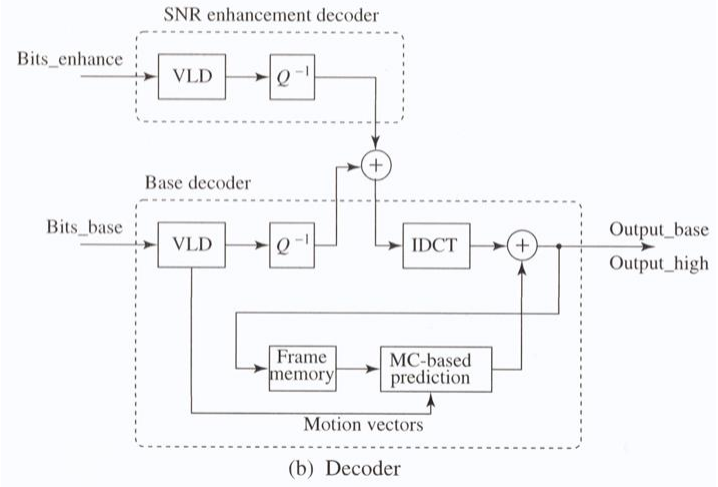
- SNR scalable encoder generates two layers
  - Base layer bit stream called `Bits_base`
    - Coarse quantization of the DCT coefficients is employed. -> Fewer bits
  - Enhancement layer bit stream called `Bits_enhance`
    - Generate a DCT coefficient refinement by computing quantization error.
    - Quantization error = original DCT – de-quantization (quantized DCT in the base layer)

# SNR Scalability




(a) Encoder

# SNR Scalability




(b) Decoder

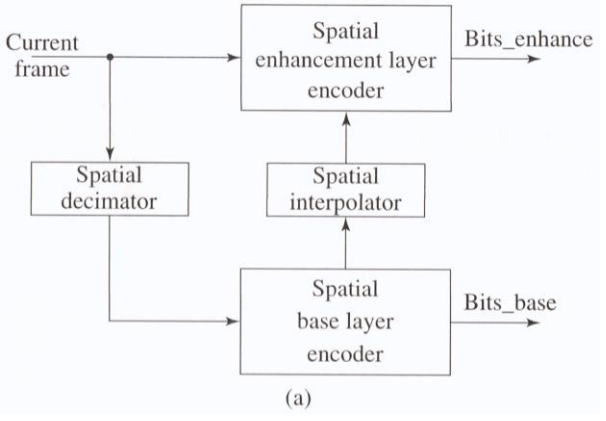
MPEG-2 SNR scalability: (a) encoder; (b) decoder.

# MPEG-2 – Spatial Scalability

- Spatial Scalability
  - Base layer
    - Generate a bitstream of reduced-resolution frames
  - Enhancement layer(s)
    - Generate a bitstream of the details of higher resolution frames
  - Frames at the original resolution is generated by combining two layers.

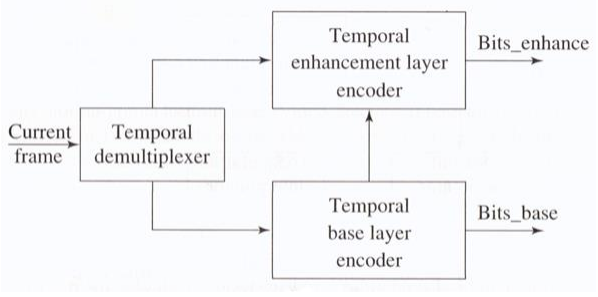


# MPEG-2 – Spatial Scalability



# MPEG-2 – Temporal Scalability

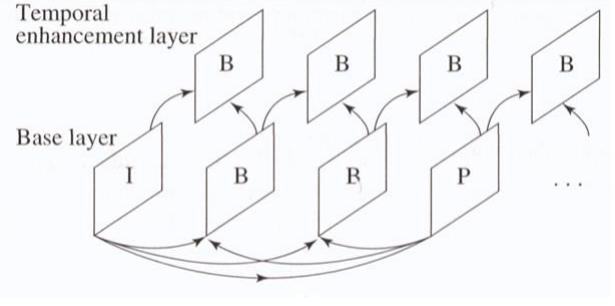
- Temporal Scalability (a) - Both the base layer and enhancement layer are generated at a reduced frame rate with the same spatial resolution as in the input video. To restore the original video, it combines two layers.





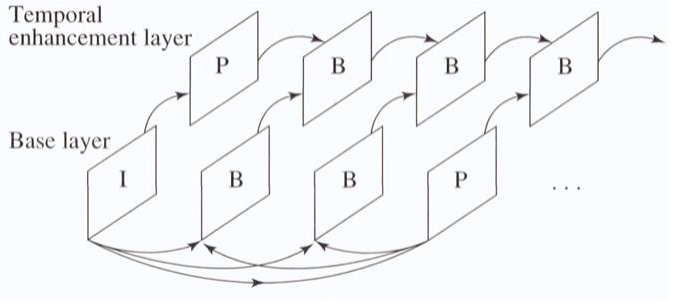

# MPEG-2 – Temporal Scalability

- Temporal Scalability (b) - The prediction of matching macroblocks at the enhancement layer by *interlayer motion-compensated prediction*



# MPEG-2 – Temporal Scalability

- Temporal Scalability (c) - The prediction of matching macroblocks at the enhancement layer by *combined motion-compensation prediction* and *interlayer motion compensation prediction*






# MPEG-2 – Other Scalability

- Hybrid scalability (combines any two of the above three)
  - SNR and Spatial
  - Spatial and temporal
  - SNR and temporal
- Data partitioning (similar to Spectral selection)
  - Base layer contains lower frequency DCT coefficients
  - Enhancement layers contain higher frequency DCT coefficients