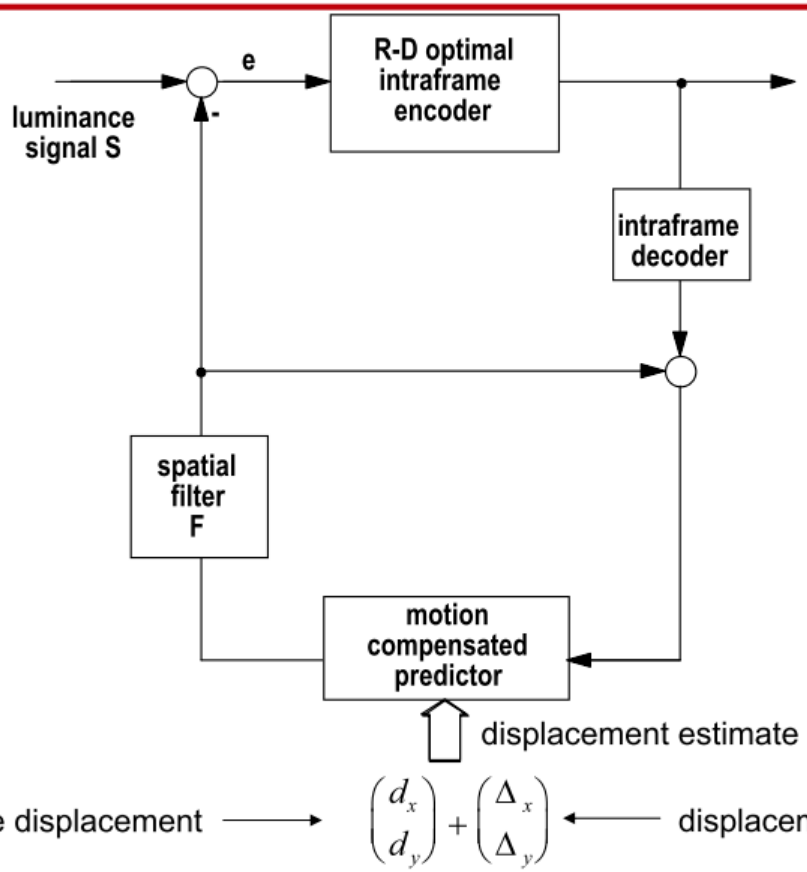
Multimedia communications ECP 610

Omar A. Nasr

omaranasr@ieee.org

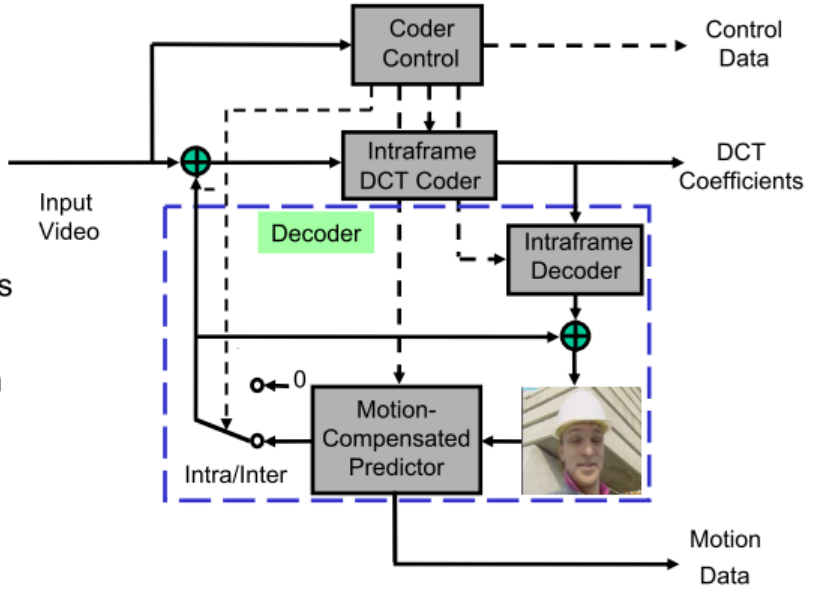
April, 2015

Model of MCP hybrid coder with loop filter



Video coder control

- Encoding decisions
  - Coding modes (intra/inter/motion comp.)
  - Block size
  - Motion vectors
  - Quantizer step size
  - Suppression of DCT coefficients
- Solution
  - Embed rate-constrained motion estimation into mode decision with Lagrangian cost function
  - Couple Lagrange multiplier to quantizer step size
- Difficulties
  - Joint entropy coding of side information
  - Temporal dependencies due to DPCM structure



H.263

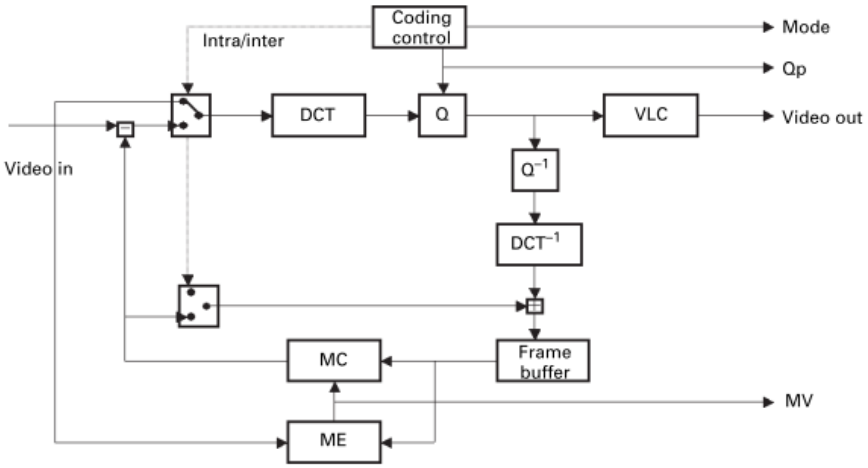


Figure 5.15 A block diagram of an H.263 baseline encoder.

H.263

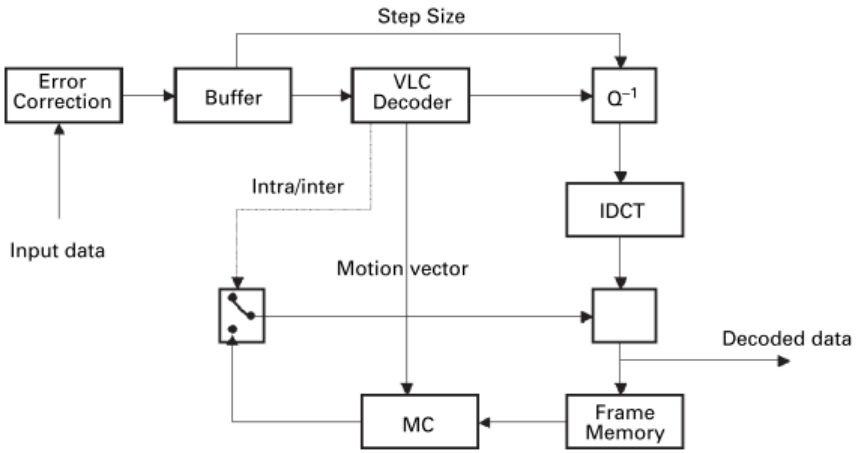
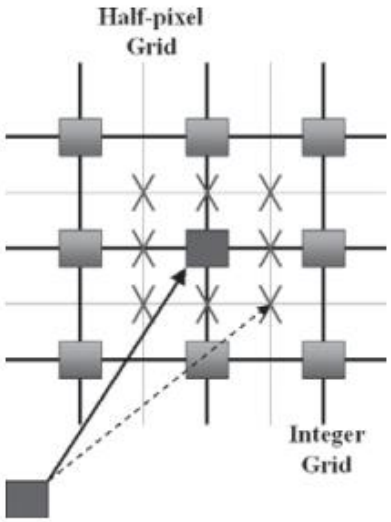


Figure 5.16 A block diagram of an H.263 baseline decoder.

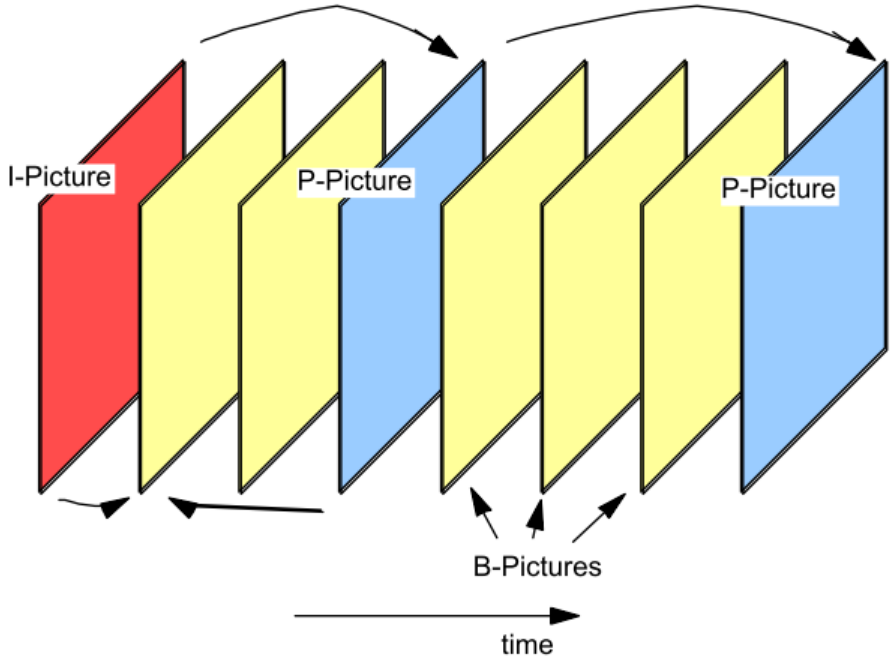
Motion estimation – half pixel grid

- Use “linear interpolation“ to generate the half pixel grid

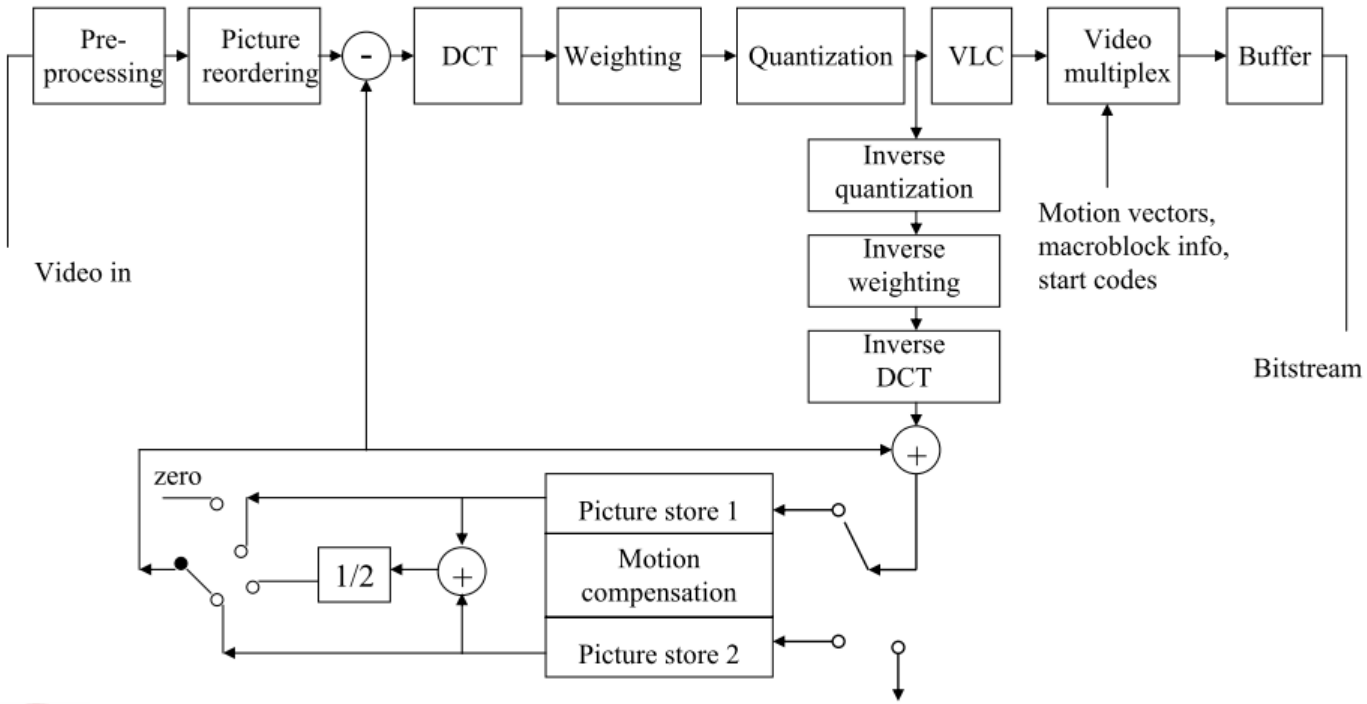


MPEG-1/2: GOP Structure

- "Group of Pictures" = “GOP“, GOP structure is very flexible



MPEG-1 Encoder



MPEG-1: coding of I-pictures

- I-pictures: intraframe coded
- 8x8 DCT
- Arbitrary weighting matrix for coefficients
- Differential coding of DC-coefficients
- Uniform quantization
- Zig-zag-scan, run-level-coding
- Entropy coding
- Unfortunately, not quite JPEG

MPEG-1: coding of P-pictures

- Motion-compensated prediction from an encoded I-picture or P-picture (DPCM)
- Half-pel accuracy of motion compensation, bilinear interpolation
- One displacement vector per macroblock
- Differential coding of displacement vectors
- Coding of prediction error with 8x8-DCT, uniform threshold quantization, zig-zag-scan as in I-pictures

MPEG-1: coding of B-pictures

- Motion-compensated prediction from two consecutive P- or I-pictures
  - **either**
    - only forward prediction (1 vector/macroblock)
  - **or**
    - only backward prediction (1 vector/macroblock)
  - **or**
    - Average of forward and backward prediction = interpolation (2 vectors/macroblock)
- Half-pel accuracy of motion compensation, bilinear interpolation
- Coding of prediction error with 8x8-DCT, uniform quantization, zig-zag-scan as in I-pictures

H.264/AVC Coder

Input Video Signal

Split into Macroblocks 16x16 pixels

Coder Control

Control Data

Transform/ Scal./Quant.

Quant. Transf. coeffs

Decoder

Scaling & Inv. Transform

Entropy Coding

Deblocking Filter

Intra-frame Prediction

Motion-Compensation

Intra/Inter

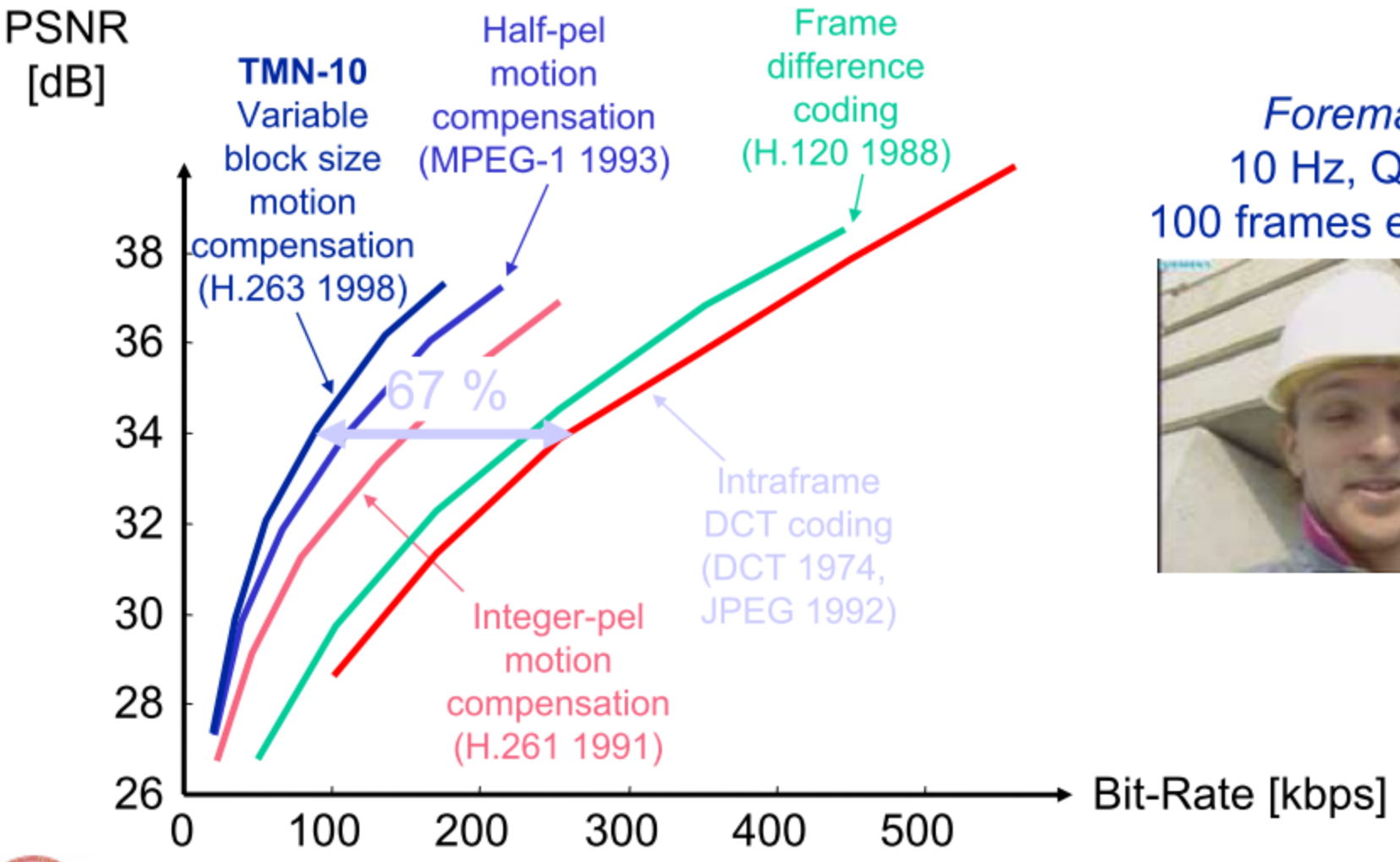
Output Video Signal

Motion Data

Motion Estimation

[source: G. Sullivan, VCEG]

Efficiency of motion-compensated coding



Foreman
10 Hz, QCIF
100 frames encoded

Scalability (SNR)

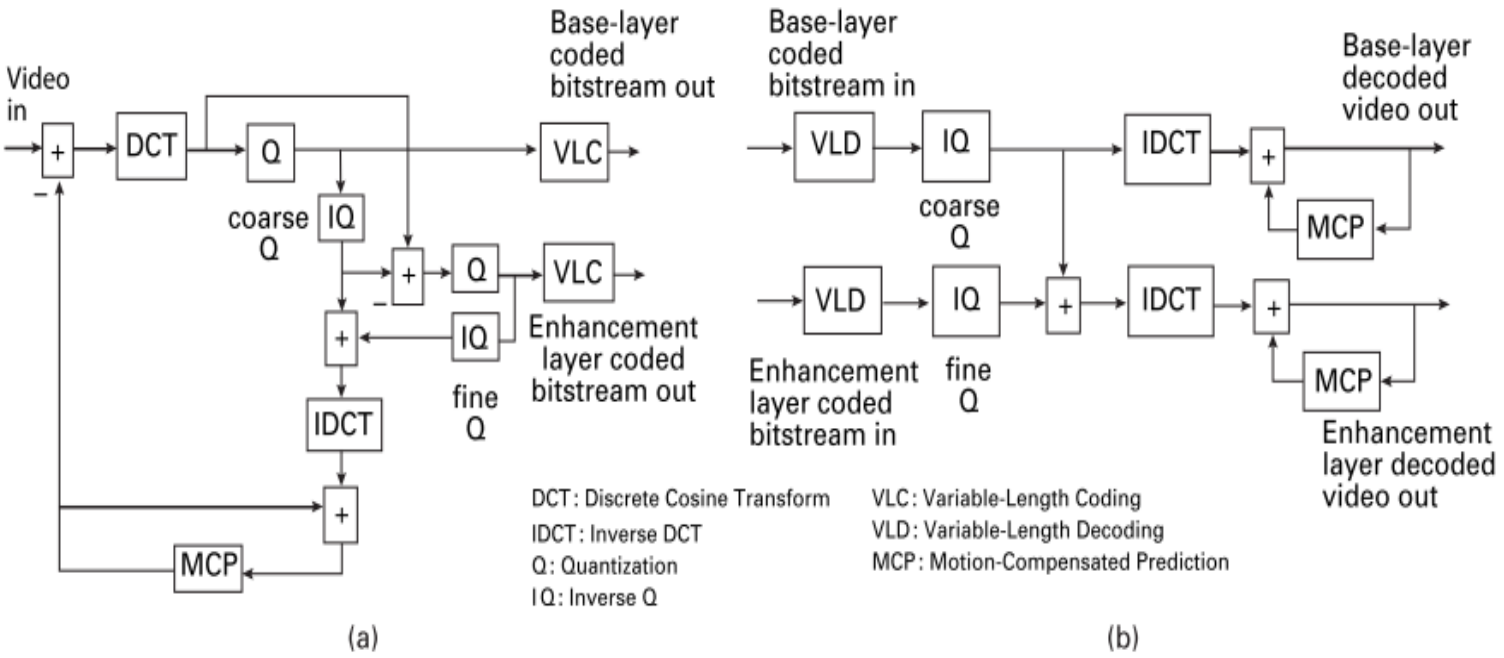


Figure 5.25 MPEG-2 SNR scalability. (a) Both layers encode the video signal at the same spatial resolution with different quantization parameters. (b) Moderate video quality can be achieved by decoding only the base-layer bitstreams while the higher video quality can be achieved by decoding both layers [28] (© Elsevier 2005).

Scalability (Spatial)

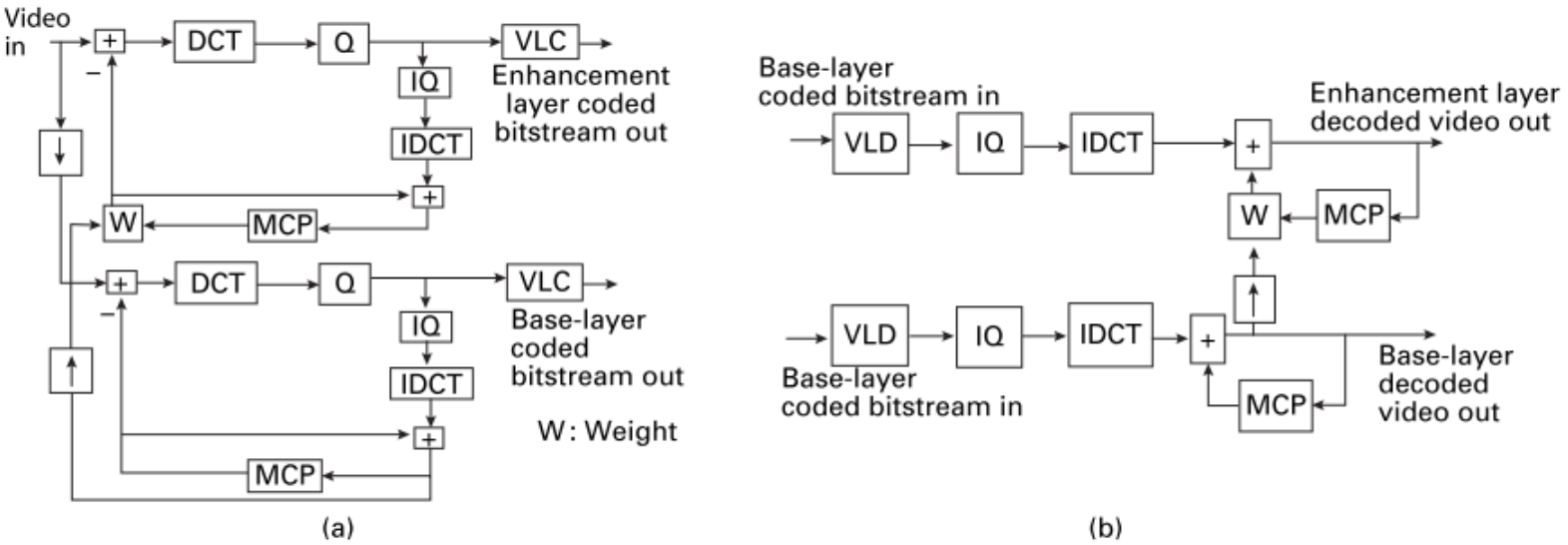


Figure 5.26 MPEG-2 spatial scalability. (a) The downsampled video is used to create a low-resolution base layer. (b) The lower-resolution video is interpolated and weighted before being added to the motion-compensated prediction from the enhancement layer [28] (© Elsevier 2005).

Scalability (Temporal)

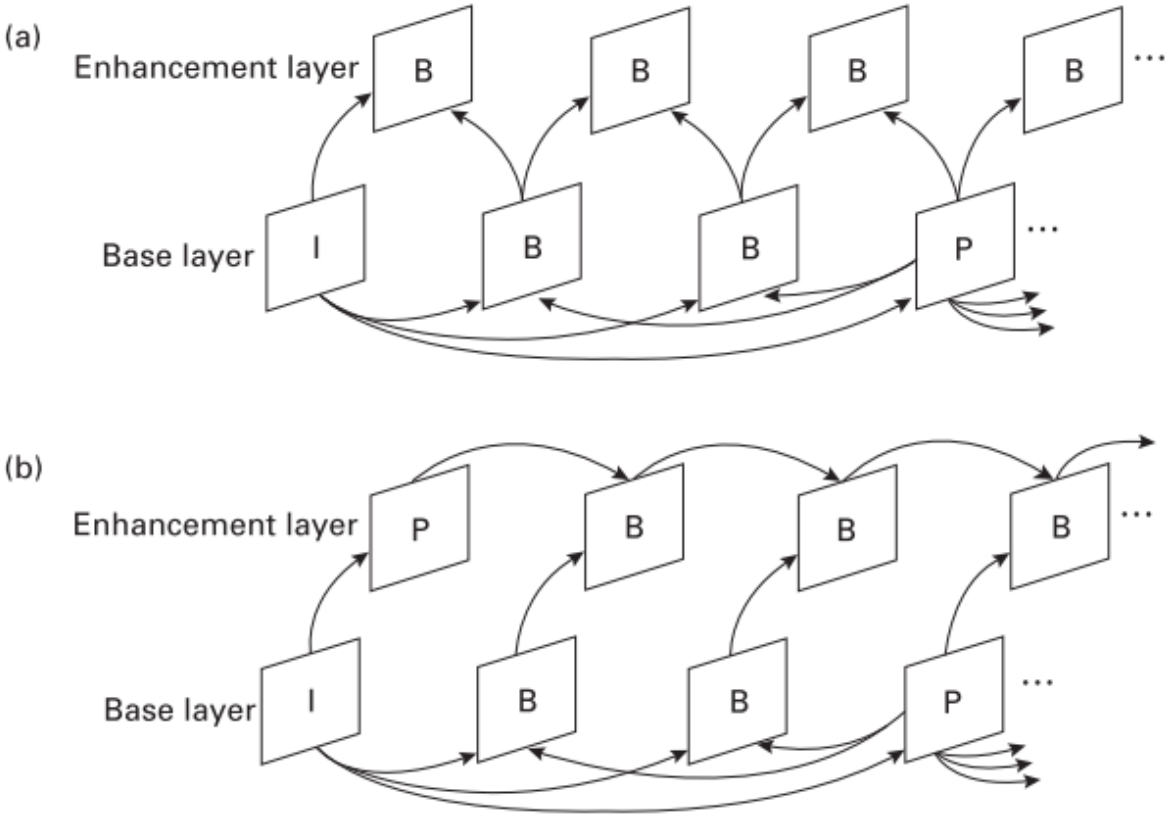


Figure 5.27 MPEG-2 temporal scalability: (a) The base layer is coded at a lower frame rate by dropping some bidirectional frames which are used as the enhancement-layer data for providing better temporal resolution. (b) Temporal scalable coding is used to support stereoscopic video.