Network & Media Laboratory

Fundamentals of Multimedia

Authored by Ze-Nian Li Mark S. Drew

Lecturer： Lu Dongming

10. MPEG Video Coding 1

MPEG-1 and MPEG-2

Overview

- **MPEG**
  - **Moving Pictures Experts Group, established in 1998 to delivery digital video and audio**
- **Membership**
  - **From 25 experts in 1988, to more than 350 experts from about 200 companies**

Content

- MPEG-1
  - Motion Compensation in MPEG-1
  - Other Major Differences from H.261
  - MPEG-1 Video Bitstream
- MPEG-2
  - Introduction
  - Interlaced Video
  - MPEG-2 Scalabilities
  - Other Major Differences from MPEG-1

Network & Media Laboratory

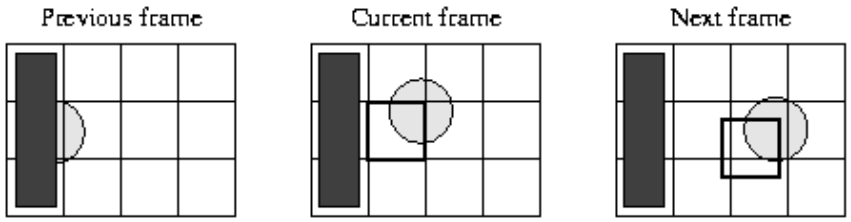
1. MPEG-1

Introduction of MPEG-1

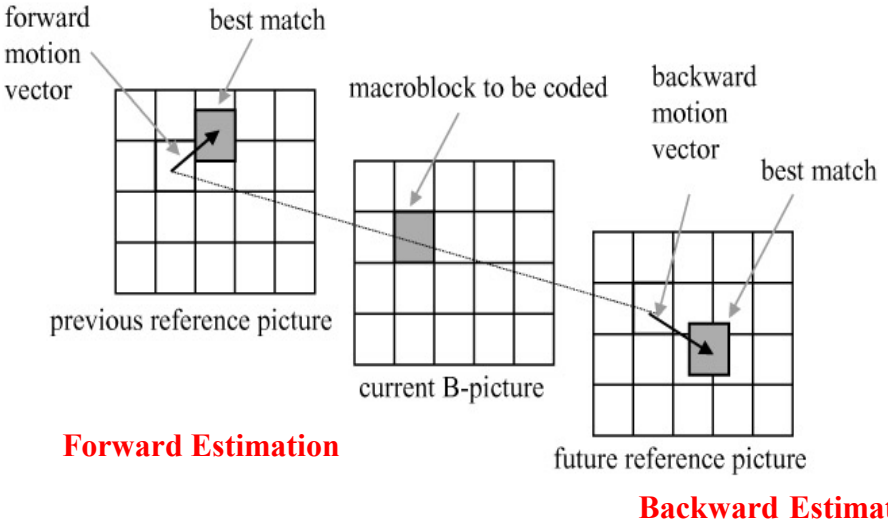
- 1991 ISO/IEC
  - Coding of moving pictures and associated audio for Digital storage media at up to about 1.5Mbit/s
  - CDs and VCDs; 1.2M for Video and 256K for audio
  - 5 parts: Systems, video, Audio, Conformance, software
- Adopt the CCIR601 digital TV format – SIF （Source Input Format）
  - Support only non-interlaced Video
  - 352*240 for NTSC at 30 fps
  - 352*288 for PAL at 25 fps
  - 4:2:0 chroma subsampling

1.1 Motion Compensation in MPEG-1

- H.261, predicted by previous frame
  - Forward Prediction

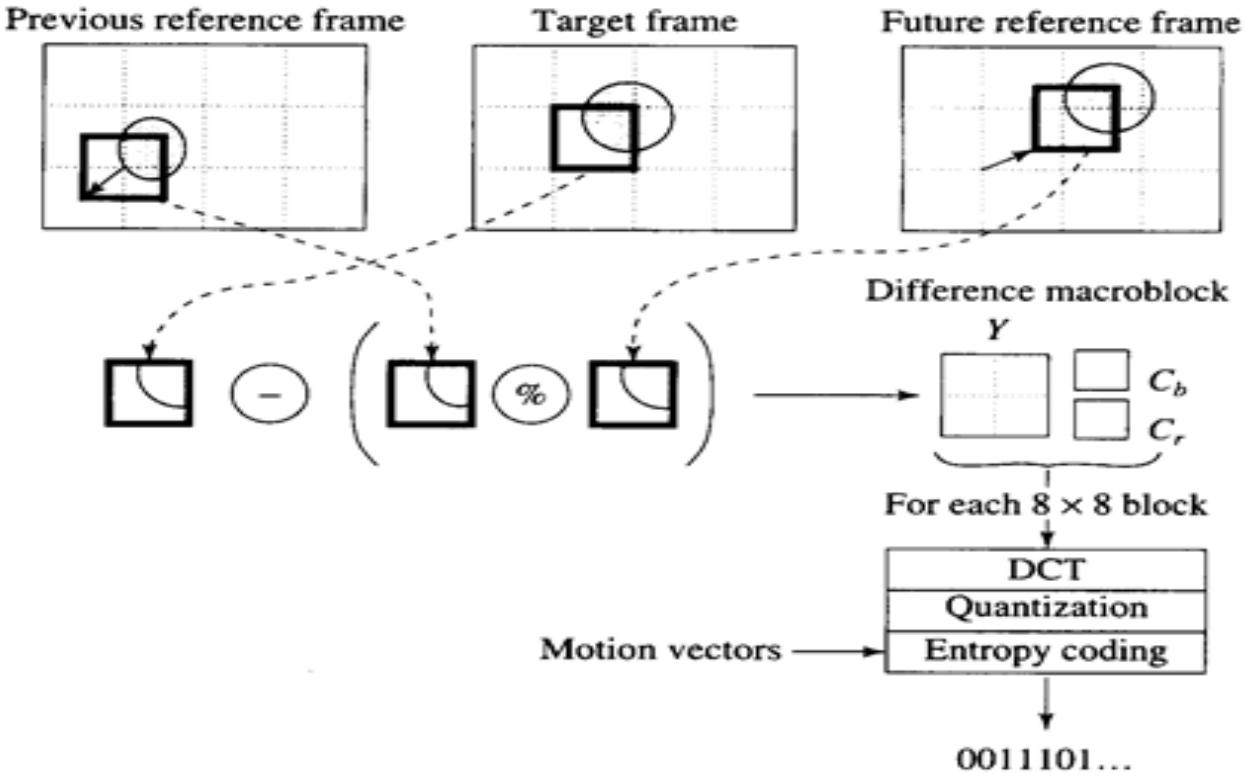


- MPEG-1 introduces B frames
  - Bidirectional Motion Compensation



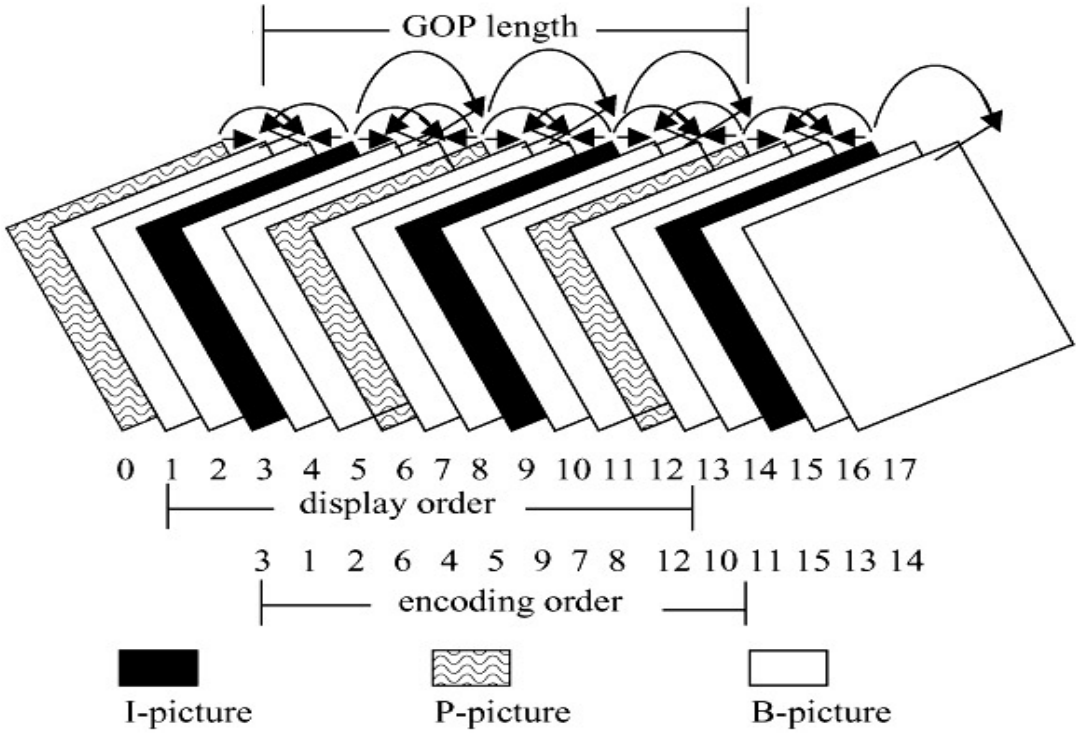
1.1 Motion Compensation in MPEG-1

- Each macro-block from B frame has two motion vectors



1.1 Motion Compensation in MPEG-1

- B frame dependents on its succeeding P or I frame, so the play order is different from the coding order.



1.2 Other Major Differences from H.261

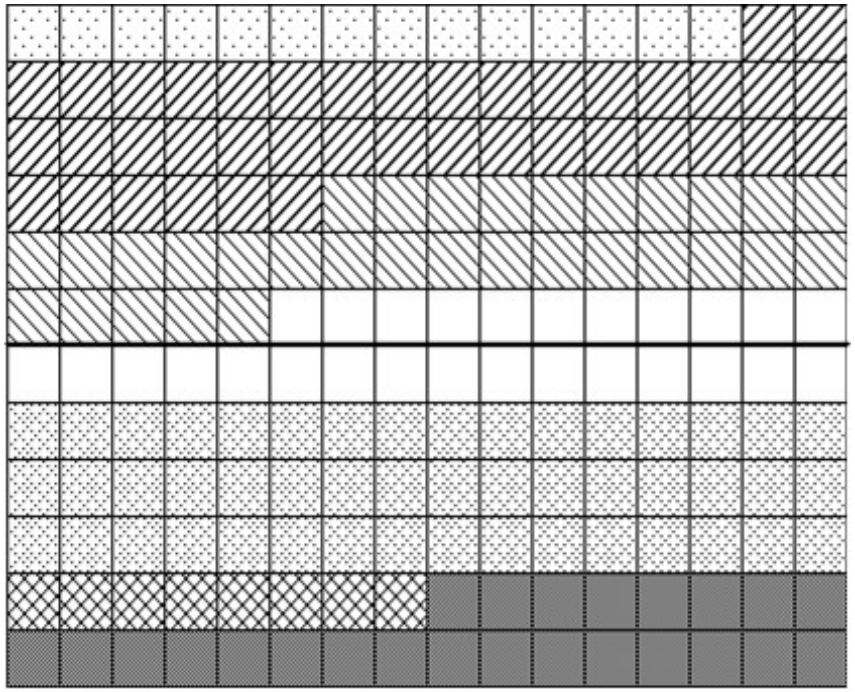
- **H.261**
  - **Supports only CIF(352×288), QCIF (176×144)**
- **MPEG-1**
  - **Supports SIF(352×240 for NTSC ,352×288 for PAL),**
  - **Allows specification of other formats**

parameter	value
horizontal size	<=768
vertical size	<=576
macroblocks per image	<=396
macroblocks per second	<=9900
frame rate	<=30fps
bitrates	<=1,856kbps

- **H.261 has no GOP, Frame (Picture) is the highest layer.**

1.2 Other Major Differences from H.261

- MPEG-1 picture can be divided into one or more slices, each slice in coded independently.



Slices in an MPEG-1 picture

1.2 Other Major Differences from H.261

- **MPEG-1 uses different quantization tables for its intra-coding and inter-coding.**
- **The quantization numbers for intra-coding vary within a macroblock, which is different from H.261**

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

Intra Quantization Table

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

Inter Quantization Table

1.2 Other Major Differences from H.261

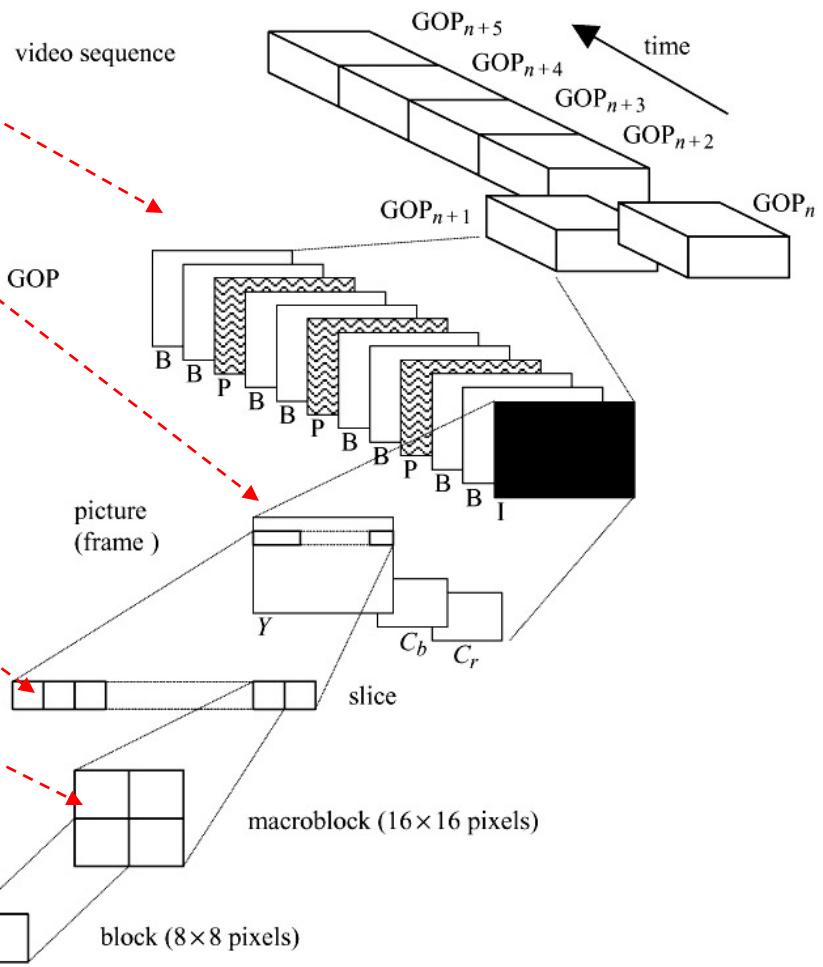
- MPEG-1 allows motion vectors to be of 0.5pixel.
- MPEG-1 supports larger gaps between I and P frames and larger motion-vector search range.
- MPEG-1 bitstream allows random access, accomplished by GOP layer.

Type	Size	Compression
I	18kB	7:1
P	6kB	20:1
B	2.5kB	50:1
Average	4.8kB	27:1

Typical compression performance of MPEG-1 frames

1.3 MPEG-1 Video Bitstream

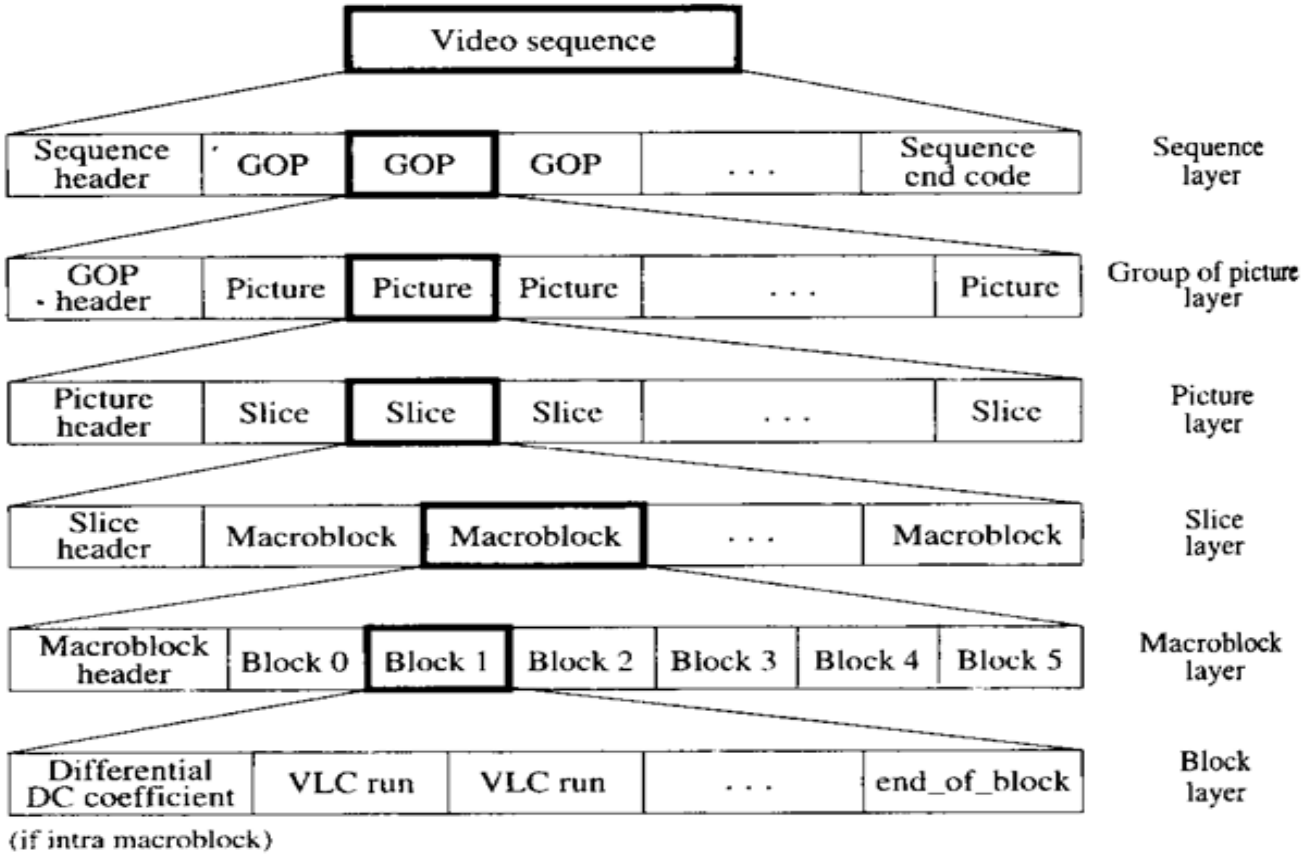
- **GOP Layer** : a set of frames between two adjacent I frames
- **Frame**: can be I、P、or B type
- **Slice**: GOB (Group of Macroblock)
- **Macroblock**: 16*16 macroblock, subsample at 4:2:0, 4 Y blocks, 1 U block, 1 V block
- **Block**: unit for DCT and quantization



1.3 MPEG-1 Video Bitstream

- Sequence layer
- Group of pictures (GOPs) layer
- Picture layer:
  - I-,P-,B-,and an uncommon type, D-picture (DC coded)
- Slice layer
- Macroblock layer
- Block layer

1.3 MPEG-1 Video Bitstream



Layers of MPEG-1 video bitstream

2. MPEG-2

2.1 Overview

- Started in 1990, for higher-quality video at a bitrate of more than 4 Mbps.
- Meets the compression and bitrate requirements of digital TV/HDTV
- Different resolutions, different compression complexities.
- Gained wide acceptance： terrestrial, satellite, cable network
- Other applications： Interactive TV, DVDs (digital video discs or digital versatile discs)

2.1 Overview

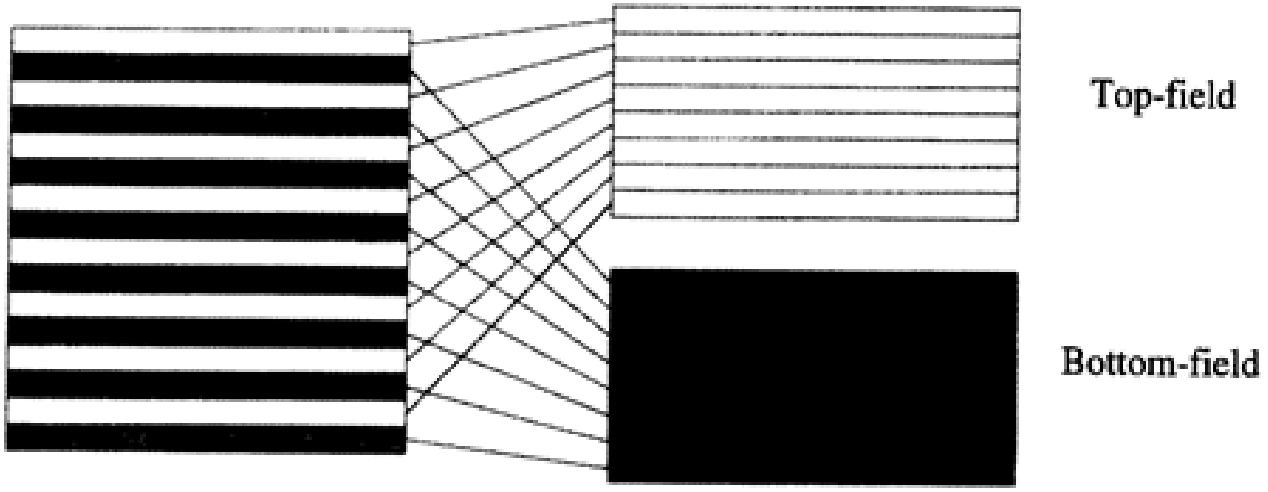
- **MPEG-2 defined 7 profiles, aimed at different applications, up to 4 levels defined in each profile.**
  - **Simple, Main, SNR Scalable, Spatially Scalable, High, 4:2:2 and Multiview**

Level	Maximum	Maximum fps	Maximum Pixels×10^6/sec	Maximum coded Data rate(Mbps)	Application
High	19201152	60	62.7	80	Film production
High1440	1440×1152	60	47.0	60	Consumer HDTV
Main	720×576	30	10.4	15	Studio TV
Low	352×288	30	3.0	4	Consumer tape equivalent

Four levels in the main profile of MPEG-2

2.2 Supporting Interlaced Video

- MPEG-2 supports SIF (352 × 288 × 25 or 30), HDTV(1920 x 1250 x 60) as input video format.
- Most of them are interlaced videos, each frame consists of two fields, each field is treated as a separate picture

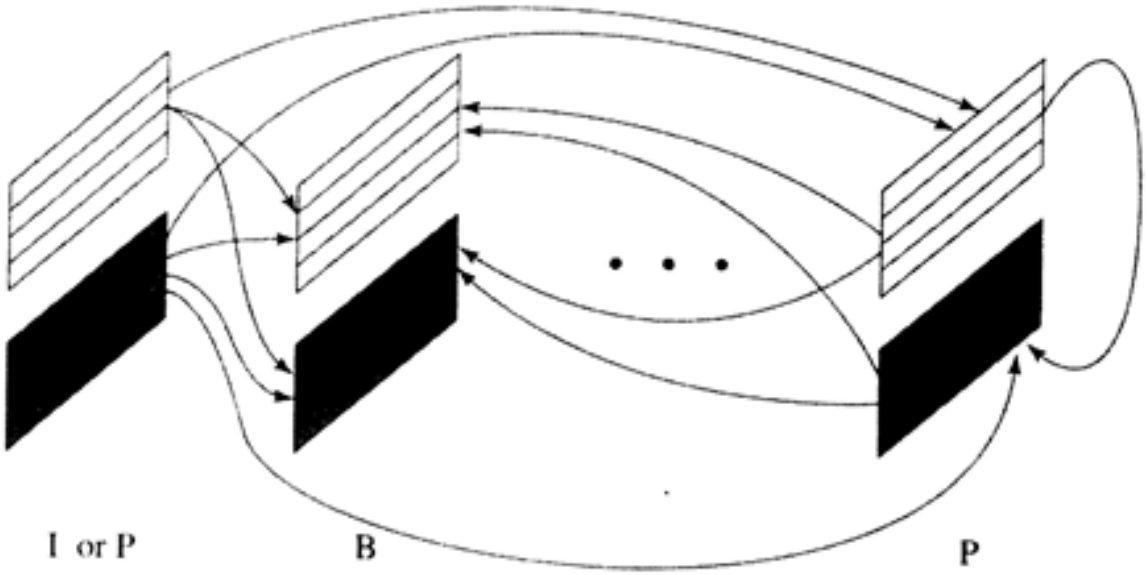


2.2 Supporting Interlaced Video

- **Five Modes of predictions**
  - **Frame prediction for frame-pictures**
    - **Identical to Mpeg1 motion compensation**
    - **Works well for videos containing slow and moderate object**
  - **Field prediction for field-pictures**
    - **See figure in the next page**
  - **Field prediction for frame-pictures**
    - **Treat top field and bottom field separately**
  - **16× 8 MC for field-pictures**
    - **Good for motion is rapid and irregular**
  - **Dual-prime for P-pictures**
    - **MV is used to derive a calculated motion vector CV**

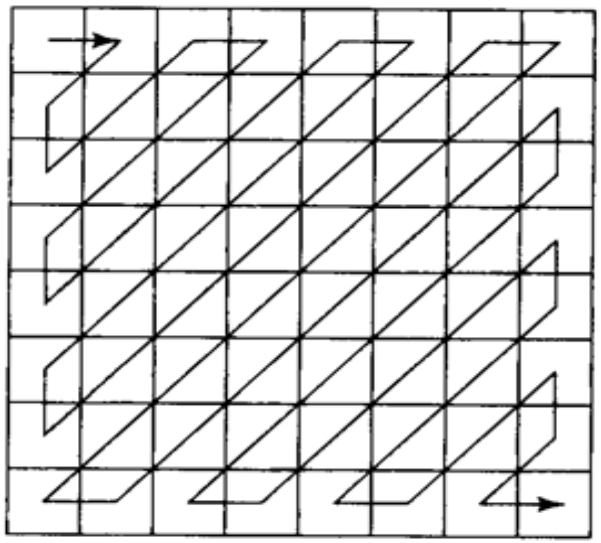
2.2 Supporting Interlaced Video

- **Field prediction for field-pictures**

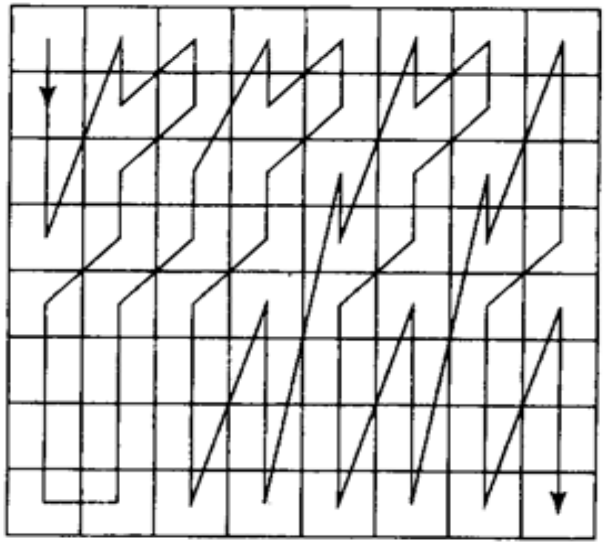


2.2 Supporting Interlaced Video

- **Alternate scan and Field DCT**
  - **Applicable only to frame-pictures in interlaced video.**
  - **Alternate scan can improve the PSNR up to 0.3dB.**



Zigzag



Alternative scans

2.3 MPEG-2 Scalabilities

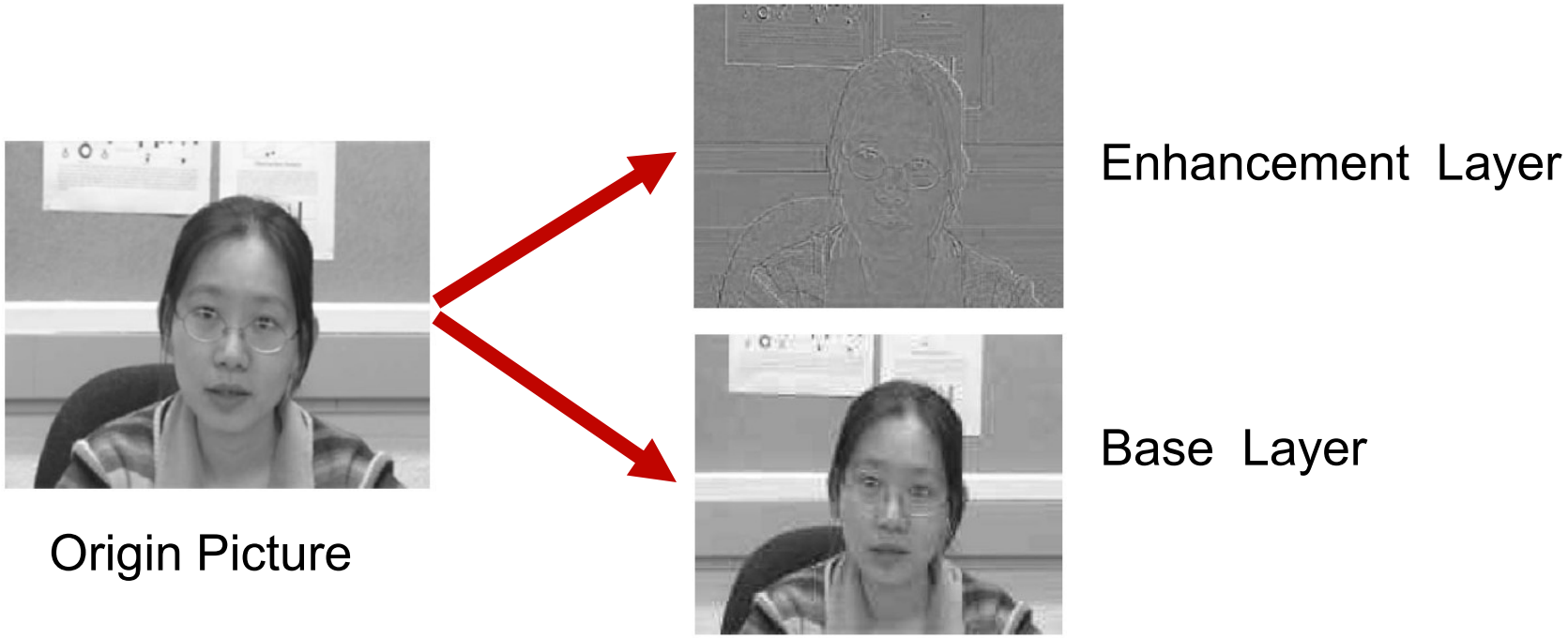
- Scalability is important for different applications
  - Defines a base layer and several enhancement layers
  - Transmitted with very different characteristics
  - Suitable for variable-bitrate (VBR) channels
  - Bitstreams are sent first to give users basic view of the video

2.3 MPEG-2 Scalabilities

- **Typical Methods**
  - **SNR scalabilities**
  - **Spatial scalabilities**
  - **Temporal scalabilities**
  - **Data partitioning**

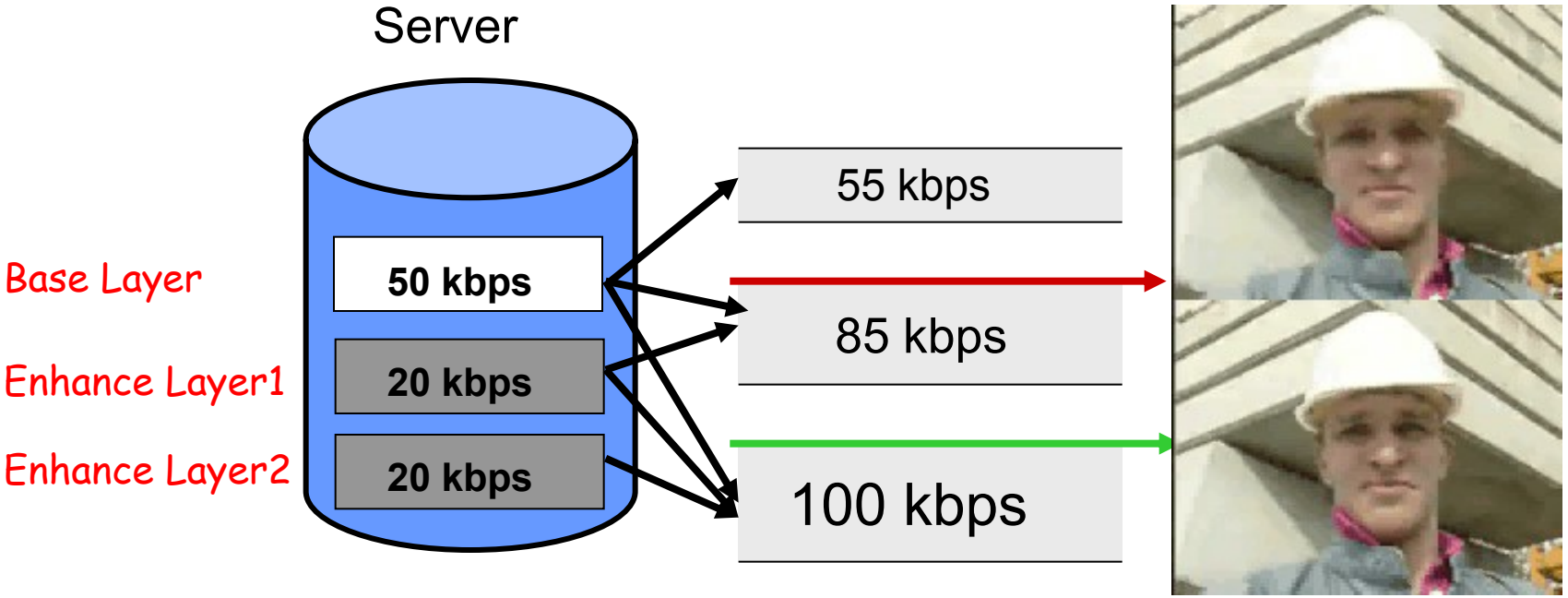
2.3 MPEG-2 Scalabilities

- **Scalable Coding**: **The basic layer provide basic video quality, enhancement layers provide better qualities.**
- **The base layer is encoded independently, enhancement layer depends on the base layer or the previous enhancement layer.**



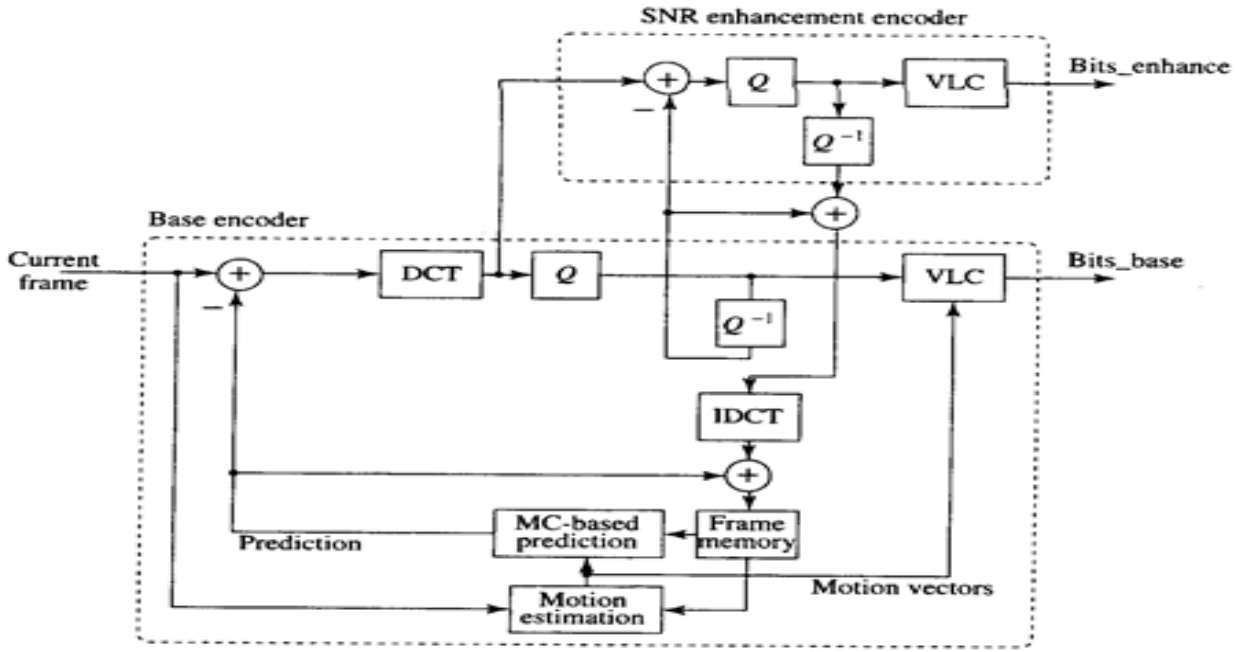
2.3 MPEG-2 Scalabilities

- Send basic layer only over insufficient band
- Basic and enhancement layers for broad-band so that receivers can obtain better quality.



2.3 MPEG-2 Scalabilities

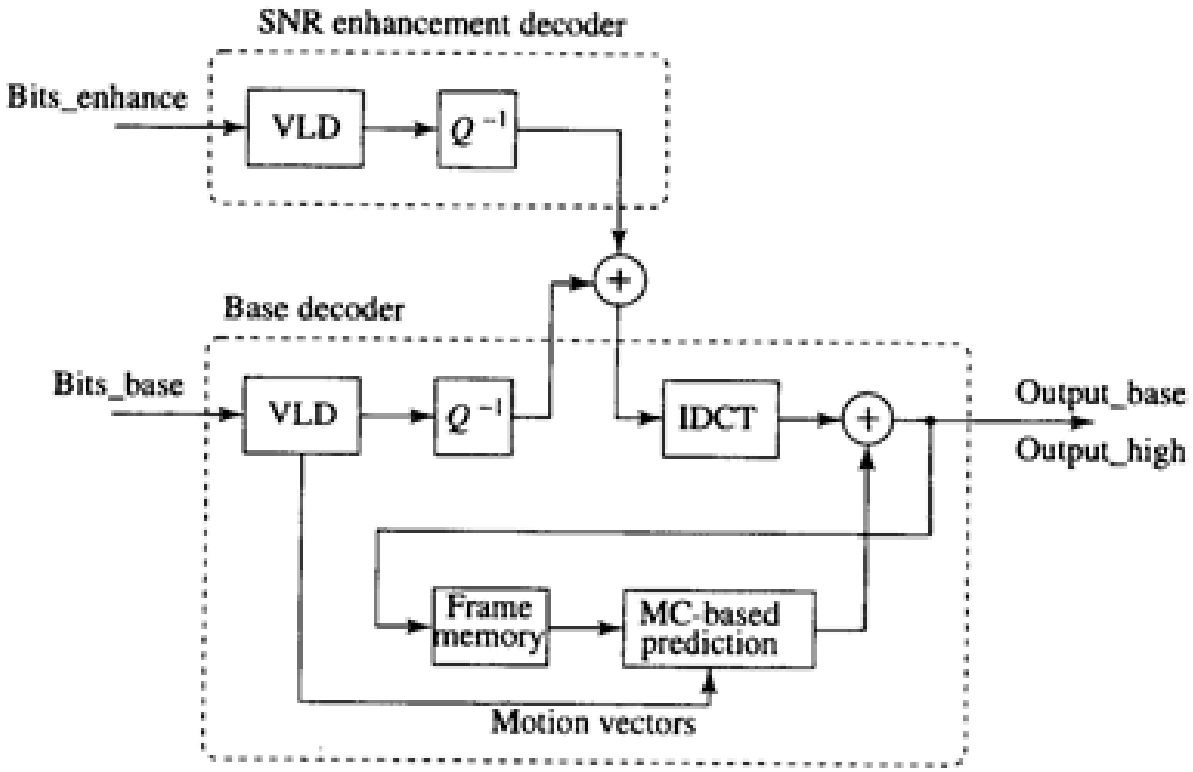
- **SNR Scalability: A coarse quantization of the DCT coefficients is employed at the base layer, difference between the video reconstructed from base layer and the original forms enhancement layer.**



SNR scalabilities (encoder)

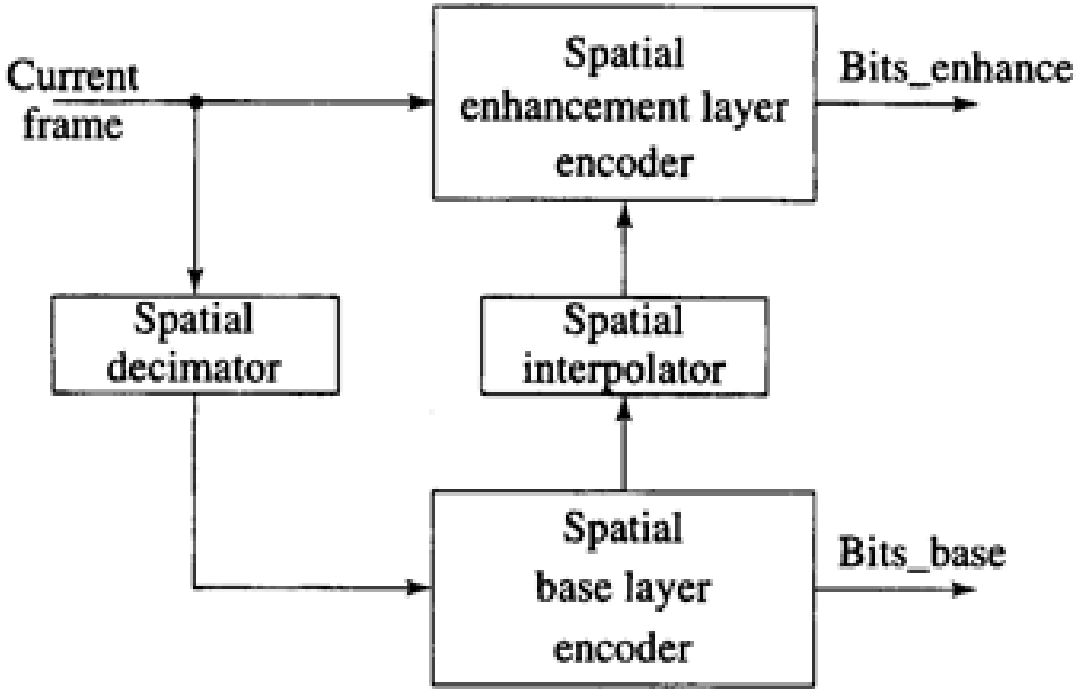
2.3 MPEG-2 Scalabilities

- SNR scalabilities (decoder)



2.3 MPEG-2 Scalabilities

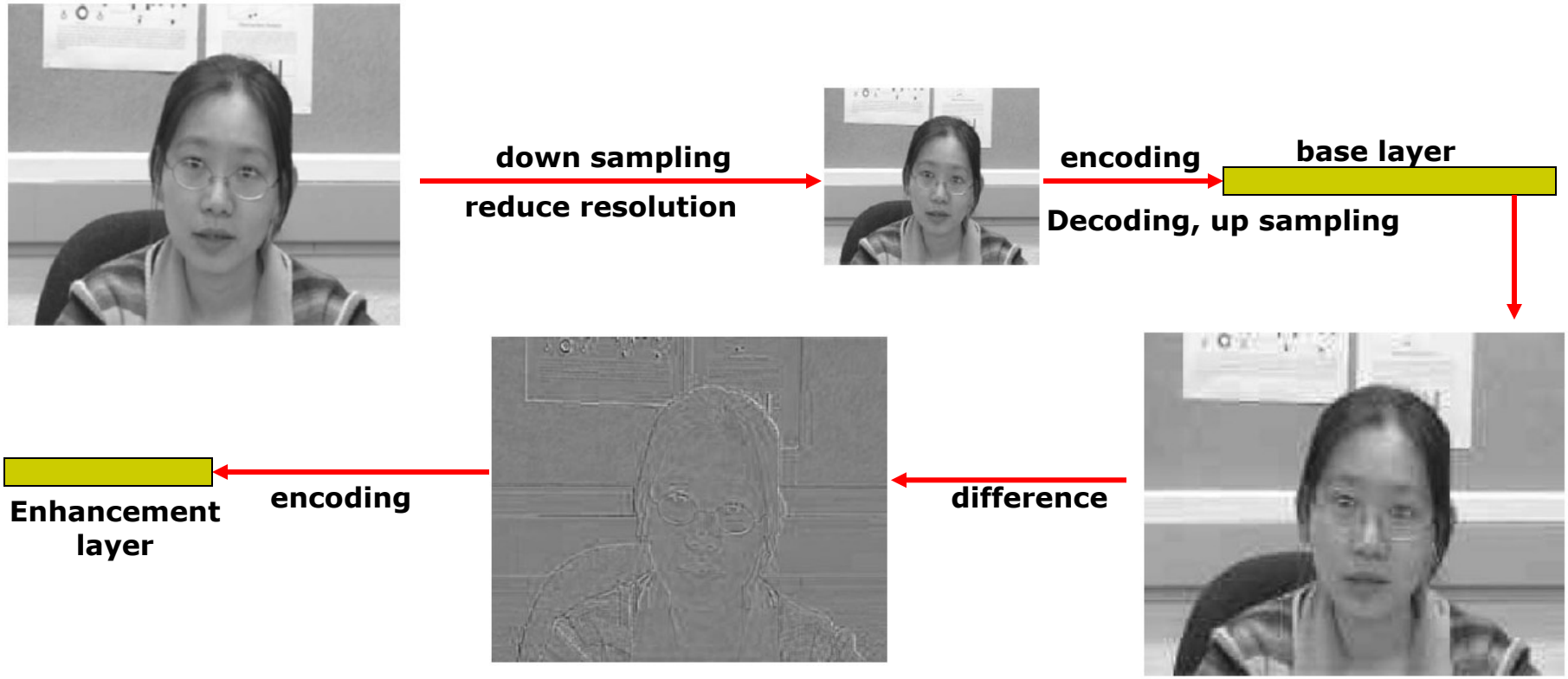
- **Spatial Scalability**



(a) Block Diagram

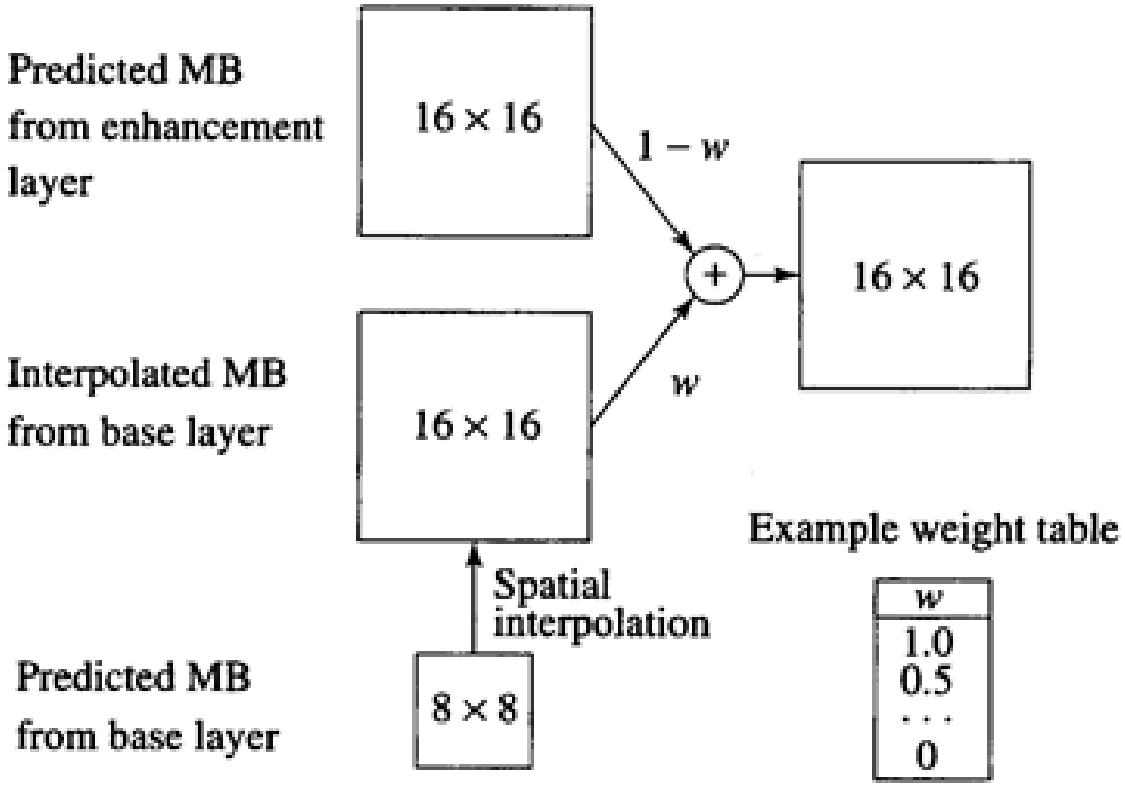
2.3 MPEG-2 Scalabilities

Spatial Scalability



2.3 MPEG-2 Scalabilities

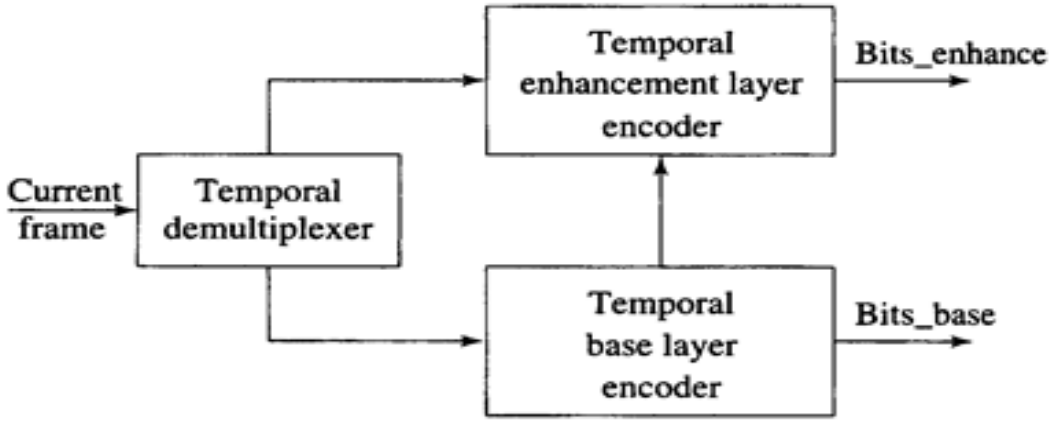
- **Spatial Scalability**



(b) Combining temporal and spatial predictions for encoding at enhancement layer

2.3 MPEG-2 Scalabilities

- **Temporal scalabilities**
  - **Base and enhancement layers at reduced frame rate**
  - **Prediction of matching macroblocks at the enhancement layer can be obtained in two ways:**
    - **interlayer motion-compensated prediction**
    - **combined motion-compensated prediction and interlayer motion-compensated prediction**



2.3 MPEG-2 Scalabilities

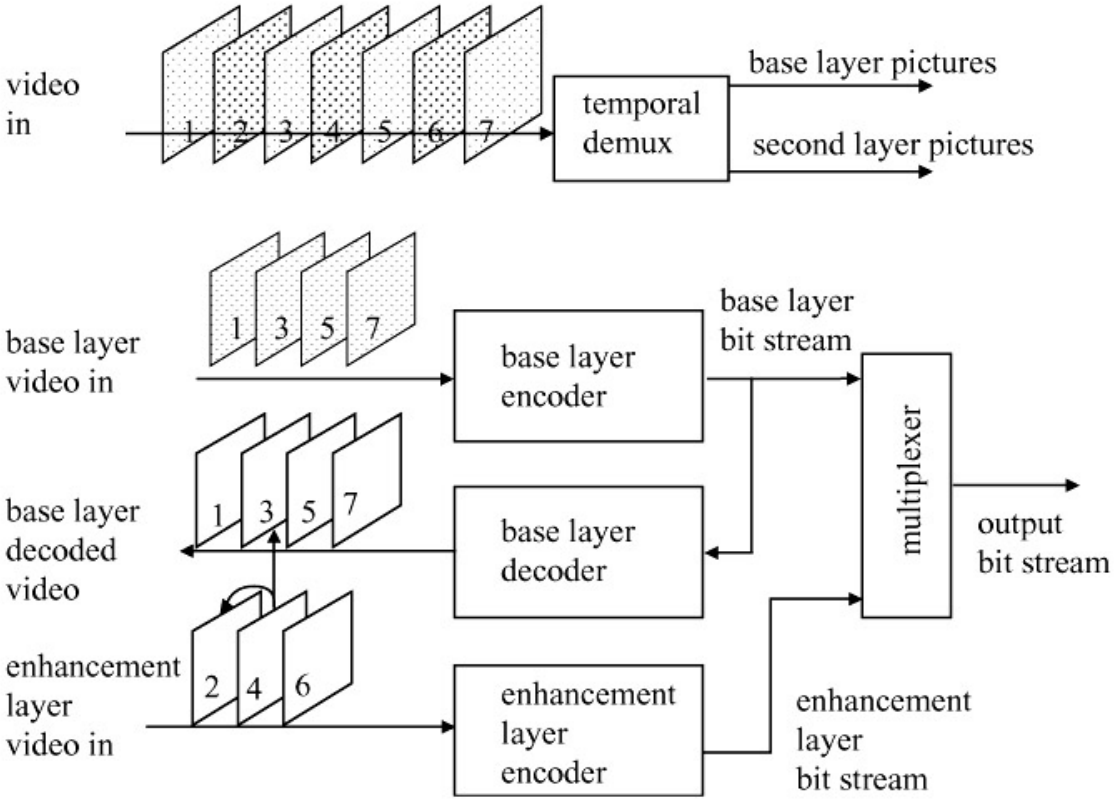
- **Pictures from base layer and enhancement layers have the same spatial resolution as the input.**

Temporal scalability

1、3、5、7 etc. odd frames as base layer, even frames as enhancement layer.

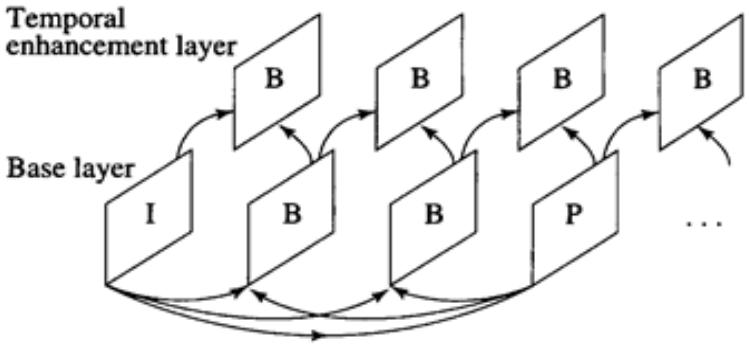
Base layer encoded as ordinary video, usually has I and P frame types.

Frames in enhance layer can take frames at base layer or its own layer as references

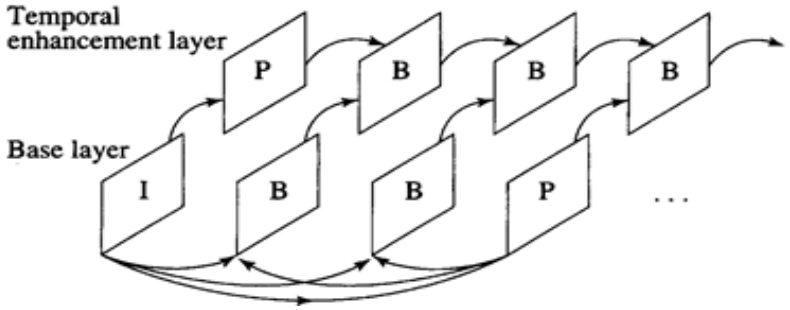


2.3 MPEG-2 Scalabilities

- **Interlayer motion-compensated prediction**

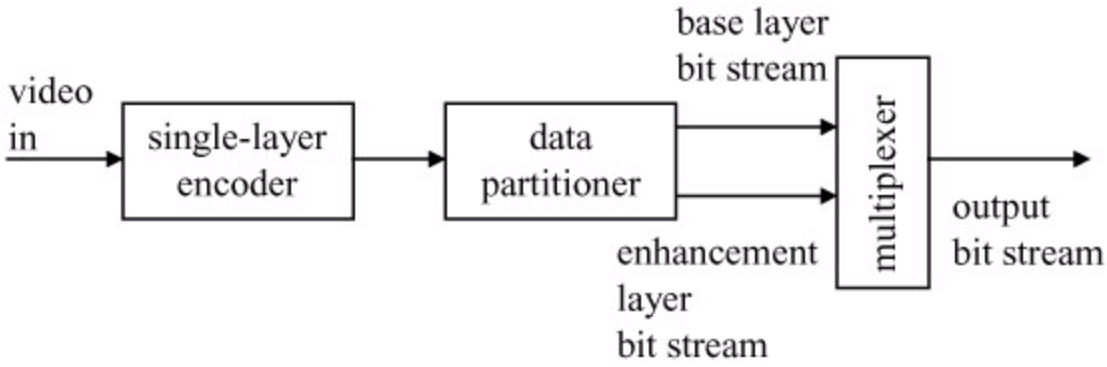


- **Combined motion-compensated prediction and interlayer motion-compensated prediction**



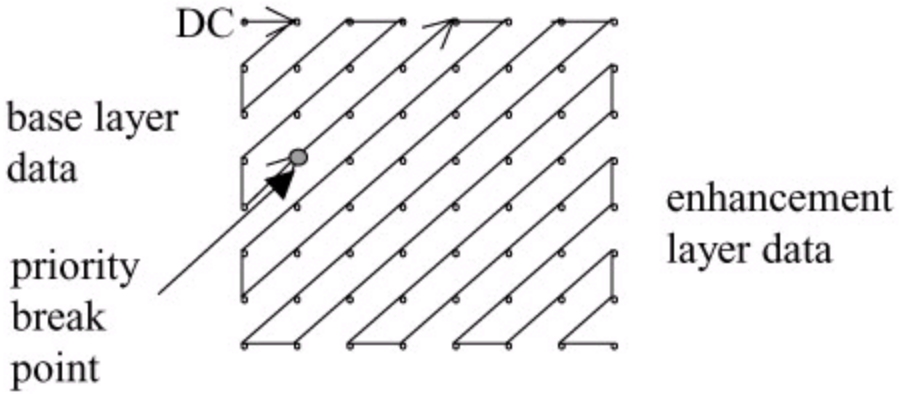
2.3 MPEG-2 Scalabilities

- **Strictly speaking, data partition doesn't conduct scalable coding. it divided video into two partitions according to their importance.**



Base layer: lower-frequency DCT coefficients, motion vectors, data head and other important information.

Enhancement layer: high-frequency DCT coefficients



Data partitioning

- Illustration of data partitioning

Base Layer

Base Layer + Enhancement Layer

2.4 Other Major difference from MPEG-1

- Better resilience to errors
- Support of 4:2:2 and 4:4:4 chroma subsampling
- Nonlinear restricted structure

Possible nonlinear scale in MPEG-2

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
$scale_i$	1	2	3	4	5	6	7	8	10	12	14	16	18	20	22	24
i	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	
$scale_i$	28	32	36	40	44	48	52	56	64	72	80	88	96	104	112	

- More restricted slice structure
- More flexible video formats

The end!

Thanks !