Wireless Communication Systems
@CS.NCTU

Lecture 9: MPEG-2

Instructor: Kate Ching-Ju Lin (林靖茹)

Chap. 11.3 of "Fundamentals of Multimedia"
Some reference from http://media.ee.ntu.edu.tw/courses/dvt/15F/

Outline

- **Introduction**
- MPEG-2 profiles and levels
- MPEG-2 syntax
- Motion Estimation
- Compression and quantization
- Scalability
- Others

MPEG-2 Video Coding Standard

- Primarily for coding interlaced video at **4 - 9 Mb/s** for digital broadcast TV and high quality digital storage media (DVD); also for HDTV, cable/satellite TV, video services over broadband networks, and high-quality videoconferencing
- Started late 1990 after completion of technical work of MPEG-1
- Committee draft for video part achieved Nov. '93
- Standard specifies only bitstream syntax and decoding process
  - Do not specify encoding process

MPEG-2 Standards

- ISO-IEC/JTC1/SC29/WG11, ITU-T ATM Video Coding Experts Group
  - ISO/IEC 13818
    - 1) Systems
    - 2) Video
    - 3) Audio
    - 4) Conformance Testing
    - 5) Simulation Software
    - 6) Digital Storage Media Control Commands (DSMCC)
    - 7) Non-Backward Compatible Audio
    - 8) 10-bit Video
    - 9) Real-Time Interface
    - 10) DSMCC Conformance
- ITU-T H.262: MPEG-2 Video

Requirements

- Picture quality - ITU-R 601 **interlaced** video with high-quality at 4-9 Mb/s
- **Random access**/channel switching in limit time
- VCR functions
- **Delay**: low delay mode using Simple Profile for visual communications
- **Error resilience**: intra-mv, data-partitioning, layered (scalable) coding
- Allow higher chroma resolution - e.g. 4:2:2 and 4:4:4
- **Scalability**: Data partition, SNR scalability, spatial scalability, temporal scalability, hybrid scalability (up to 3 layers)
- **Compatibility**: decodes MPEG-1 bit-stream, base layer may be decoded by MPEG-1 decoder
- Flexible video formats and frame rates
- Low-cost decoder and reasonable-cost encoder

MPEG-2 Coding & Compression

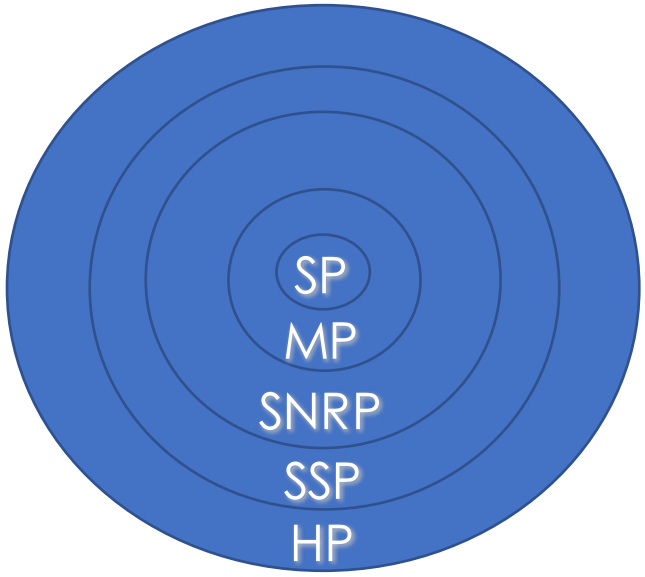
- **Interlace scan** and **progressive scan**
- Color subsampling -- 4:2:0, **4:2:2**, **4:4:4**
- Motion compensation -- prediction and prime vectors
- Quantization -- more flexible for changing matrix and scale factor
- Profile & Level -- A profile is a defined subset of the entire bit stream syntax. Within profile, a level is defined as a set of constraints imposed on the parameters of the bit stream
- Scalability
  - Data partitioning -- for two channels available applications
  - SNR scalability
  - Spatial Scalability
  - Temporal Scalability

Outline

- Introduction
- **MPEG-2 profiles and levels**
- MPEG-2 syntax
- Motion estimation
- Compression and quantization
- Scalability
- Others

Profiles and Applications

- Each profile supports groups of features for an application area
- Simple Profile: low-delay videoconferencing
- **Main Profile**: most important, for general applications
- SNR Profile: multiple grades of quality
- Spatially Scalable Profile: multiple grades of quality and resolution
- High Profile: multiple grades of quality, resolution, and chroma format



New profiles:

- 4:2:2 profile
- Multiview profile

Profiles and Levels

- MPEG-2 defines coding parameters by profiles and levels
  - Profile: Defines sub-set of syntax (functionality)
  - Level: Defines set of constraints (size)
- Most useful one
  → Main Profile @ Main Level (MP@ML)
  - Profile frame format: I, P, B
  - Chrominance format: 4:2:0
  - Scalability: No
  - 720x480, 30 f/sec, or 720x576, 25 f/sec
  - Compressed bit rate: no more than 15Mbit/s

Profiles and Levels

Level	Profile				
	Simple 4:2:0	Main 4:2:0	SNR Scalable 4:2:0	Spatially Scalable 4:2:0	High 4:2:0 or 4:2:2
High 1920x1152 (60 frames/s)		80 Mbit/s			100 Mbit/s for 3 layers
High-1440 1440x1152 (60 frames/s)		60 Mbit/s		60 Mbit/s for 3 layers	80 Mbit/s for 3 layers
Main 720x576 (30 frames/s)	15 Mbit/s	15 Mbit/s	15 Mbit/s for 2 layers		20 Mbit/s for 3 layers
Low 352x288 (30 frames/s)		4 Mbit/s	4 Mbit/s for 2 layers		

* numbers in the table are maximum allowed

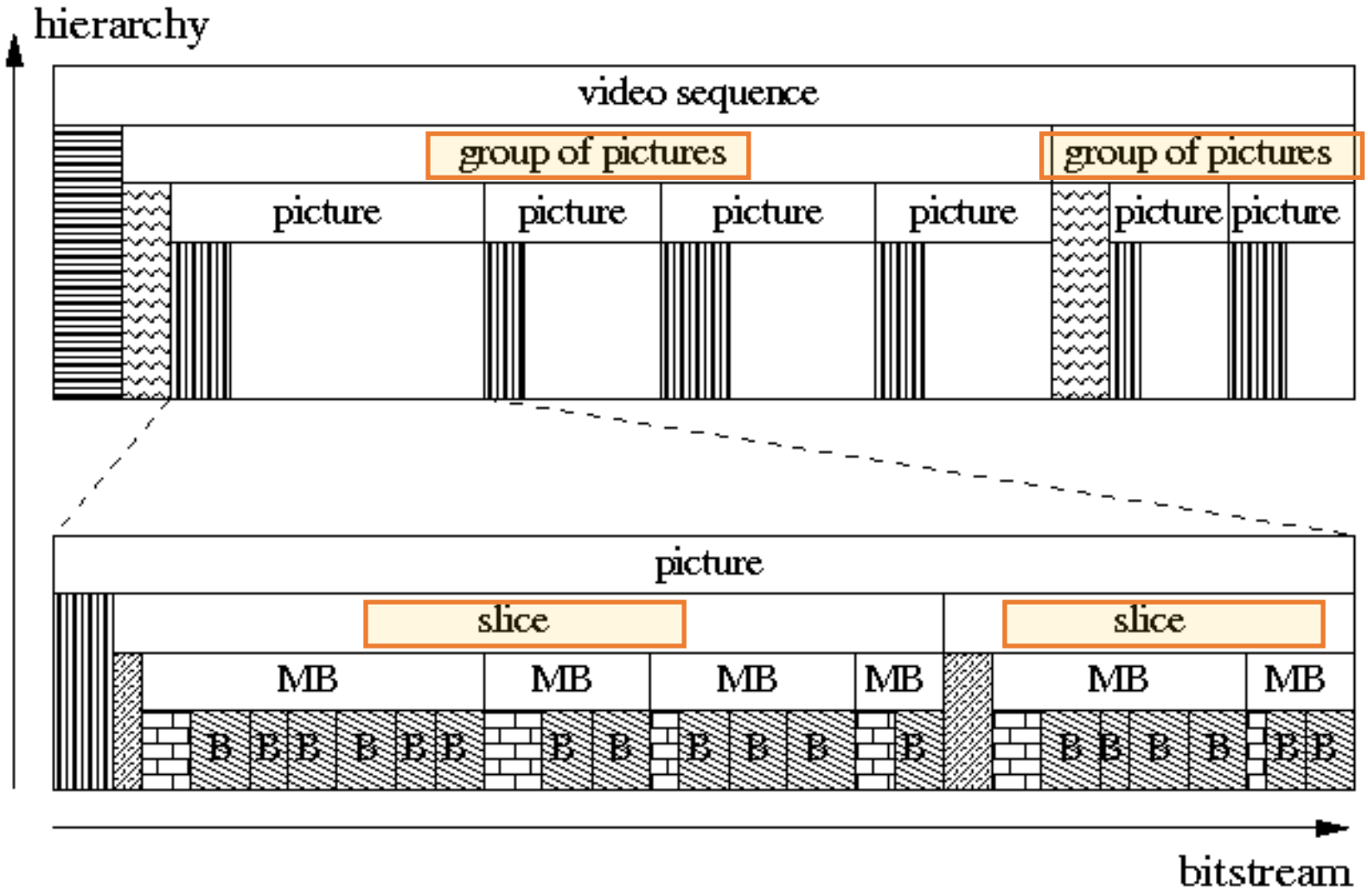
MPEG-2 Resolutions and Formats

- Picture sizes extension up to 16k x16k
  - 720 x480 ~ TV resolution
- Support picture rates:
  - 23.98, 24, 25, 29.97, 30, 50, 59.94, 60
- Support both format
  - progressive and interlaced
- Support sampling formats:
  - 4:2:0, 4:2:2, and 4:4:4

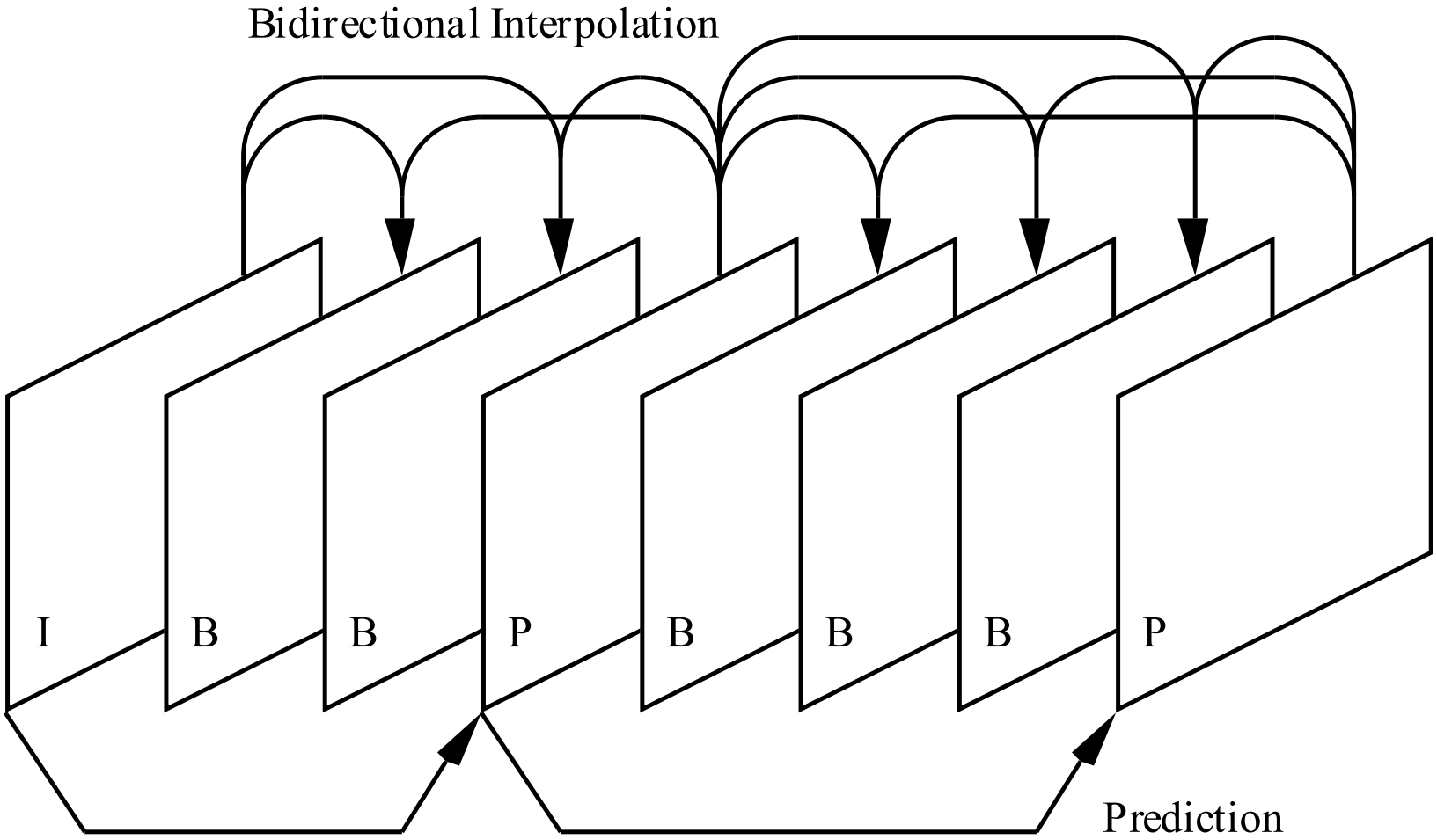
Outline

- Introduction
- MPEG-2 profiles and levels
- **MPEG-2 syntax**
- Motion Estimation
- Compression and quantization
- Scalability
- Others

MPEG-2 Video Stream Syntax



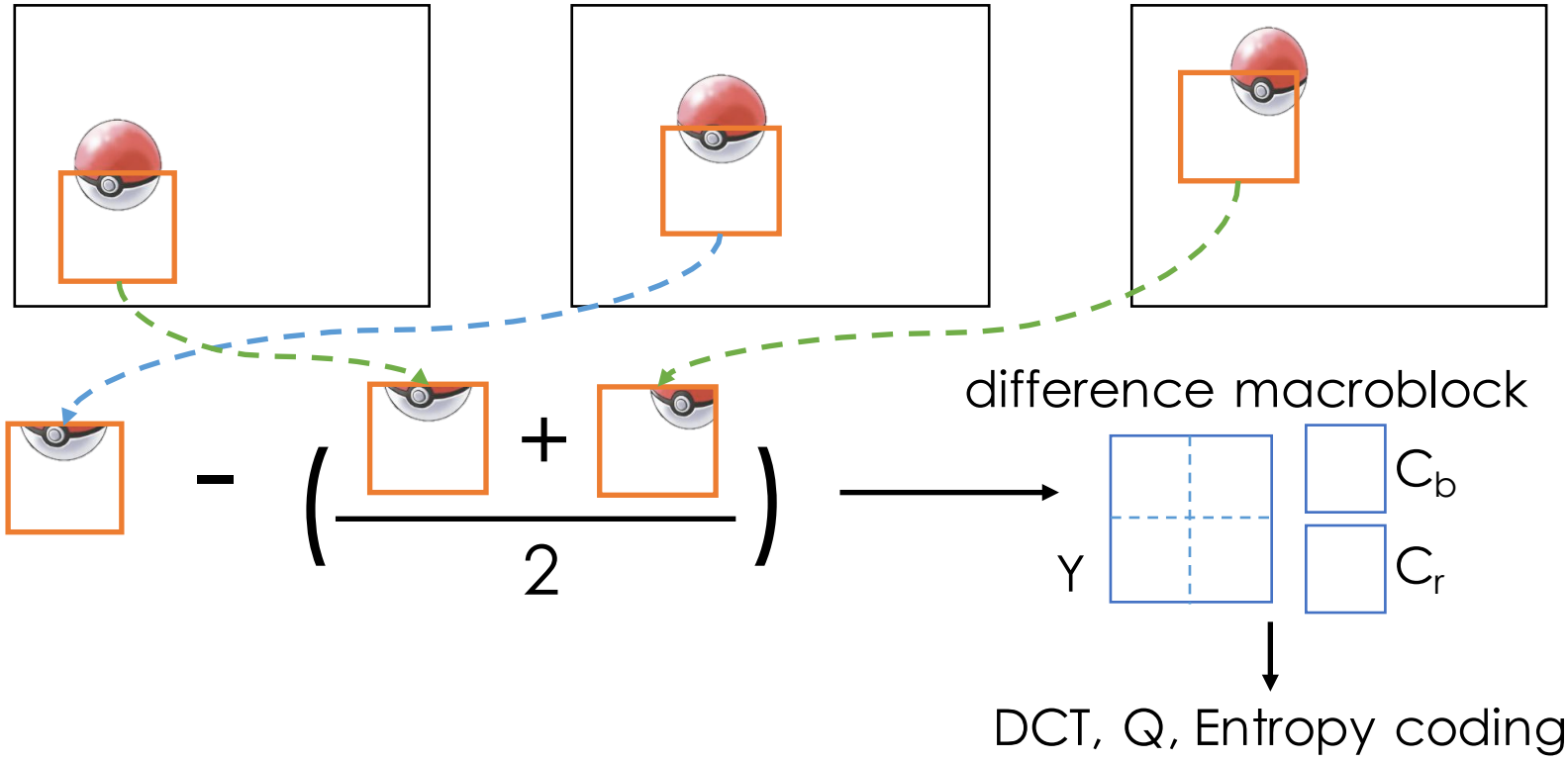
Structure of GOP



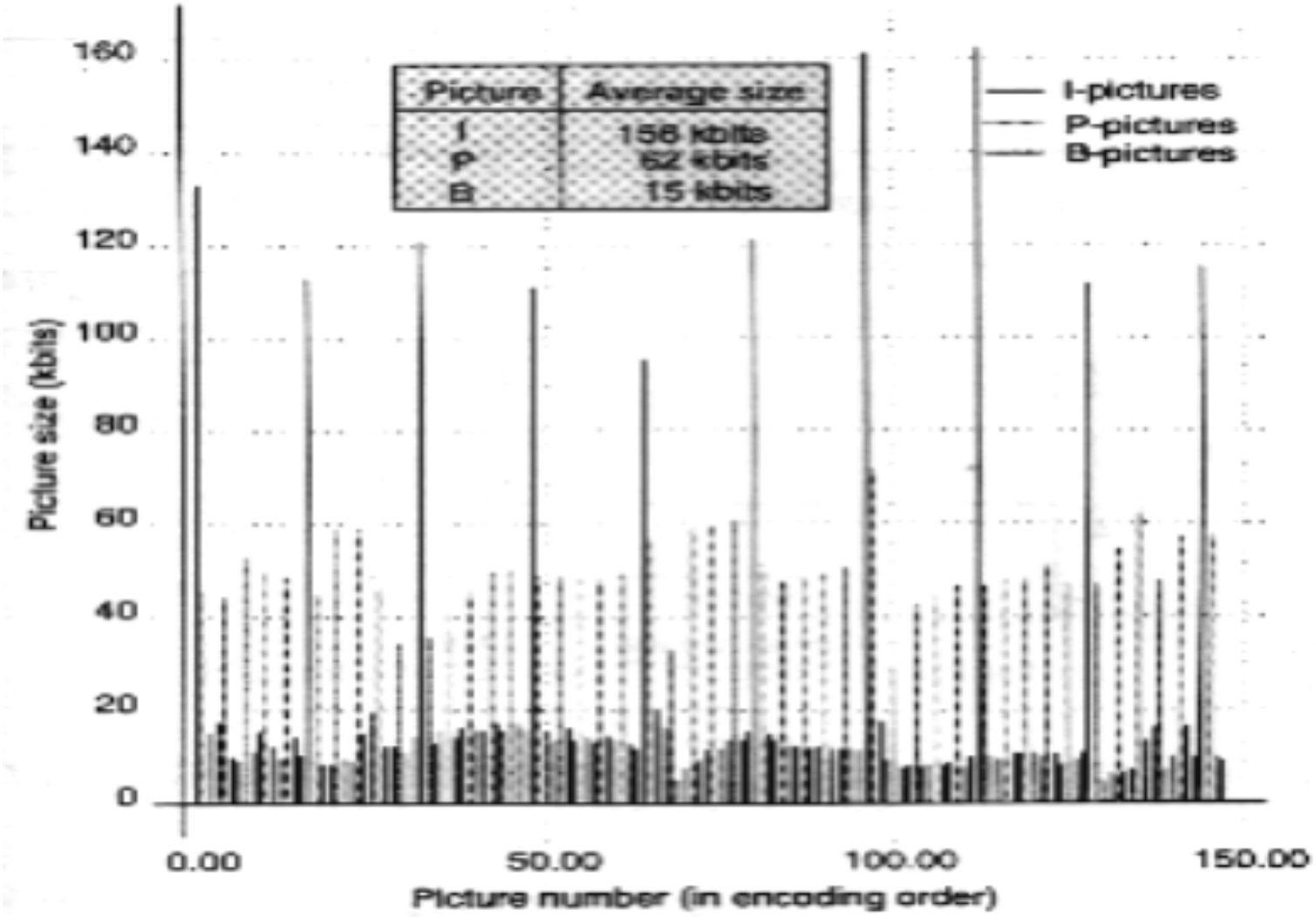
- P (Predicted frame): reference could be I or P
- B (Bi-predicted frame): bidirectional prediction

B-Frame

- Also used in H.263 and MPEG-4
- Bidirectional motion compensation
  - Two motion vector: one from backward prediction, and the other from forward prediction
  - Reference pictures can be I or P frame

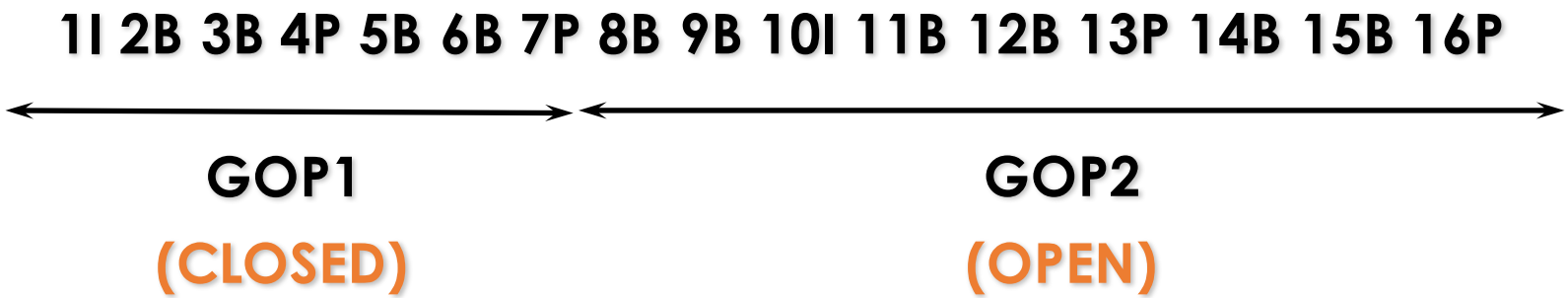


Picture Structure

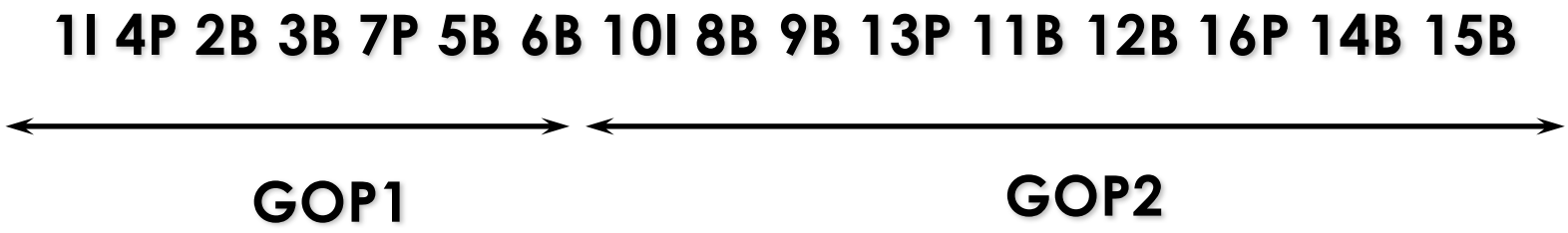


Frame Reordering

Display Order:

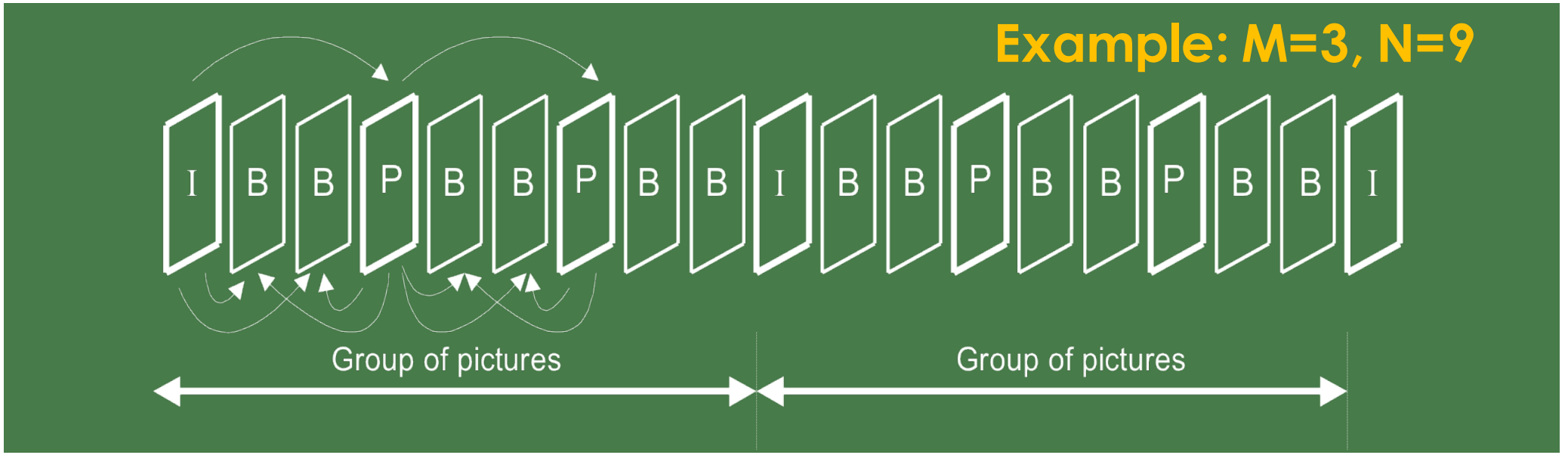


Bitstream (Encoding/Decoding) Order:



- Close: reference bounded in a GOP
- Open: reference across a neighboring GOP
- Reordering delay: a number of consecutive B-frames should wait for a future reference I- or P-frame

Group of Pictures (GOP)



- *N* = number of pictures in a GOP
- *M* = prediction distance (M-1 in-between B-pictures)
- Tradeoff on N and *M*

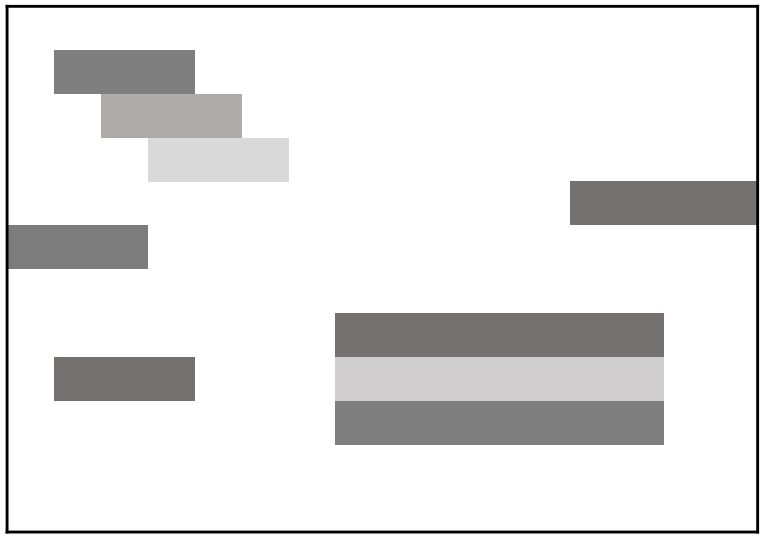
I = Intra-Picture coding, allow **random access**, for reference

P = Predictive coding, **causal prediction only**, can be referenced

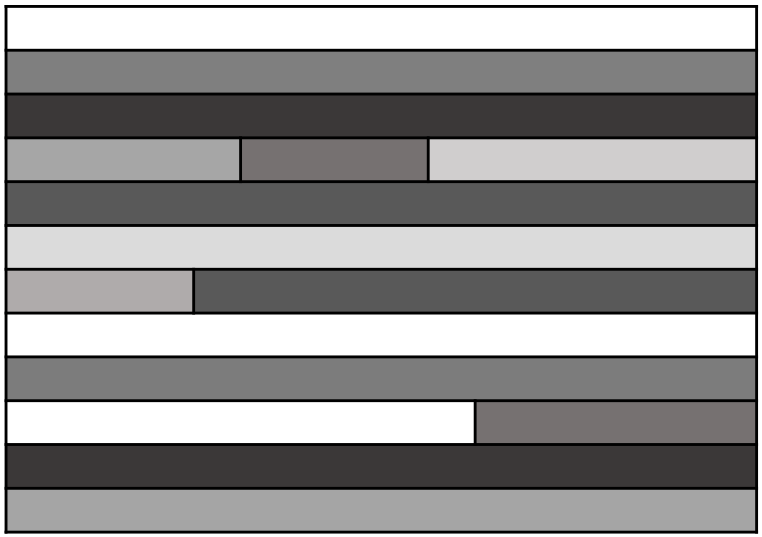
B = Bi-directional coding, **noncausal prediction**, never referenced

Slices

- Unlike GOBs in H.261, an MPEG-1frame can be divided into one or more non-overlapping slides
- Or, each slice is a series of an arbitrary number of consecutive macroblocks
- Two types of slices



Non-constraint slice



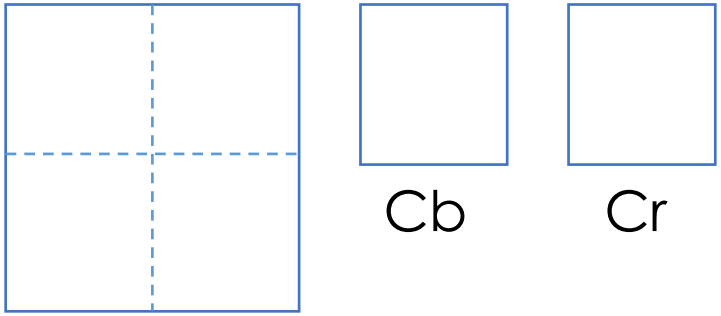
Restricted slice

Slices: Properties

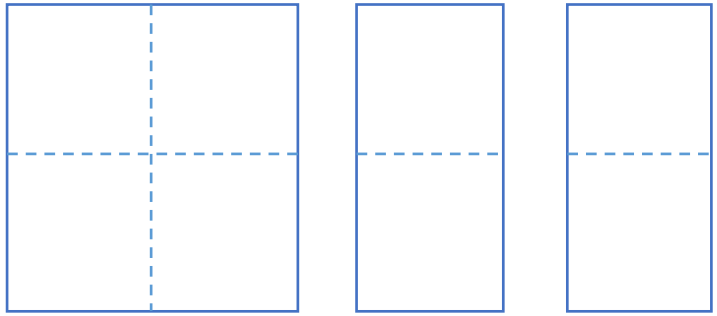
- Some macroblock may not belong to any slice (non-constraint slice)
- Each slice is coded independently
  - e.g., each slice can have different quantization scale factors
- The position of slices may change from picture to picture
- The first and last macroblock of a slice shall be in the same horizontal row of macroblocks
- No prediction across slice is allowed
  - If the bitstream contains an error, the decoder can skip to the start of the next slice → **prevent error propagation**

Macroblock Structure

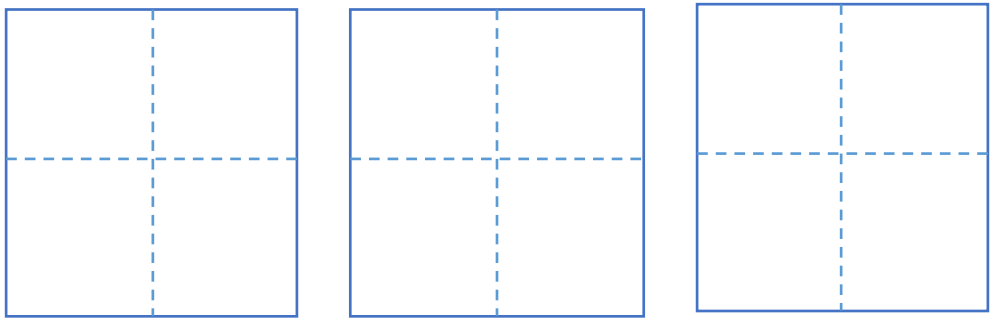
4:2:0



4:2:2

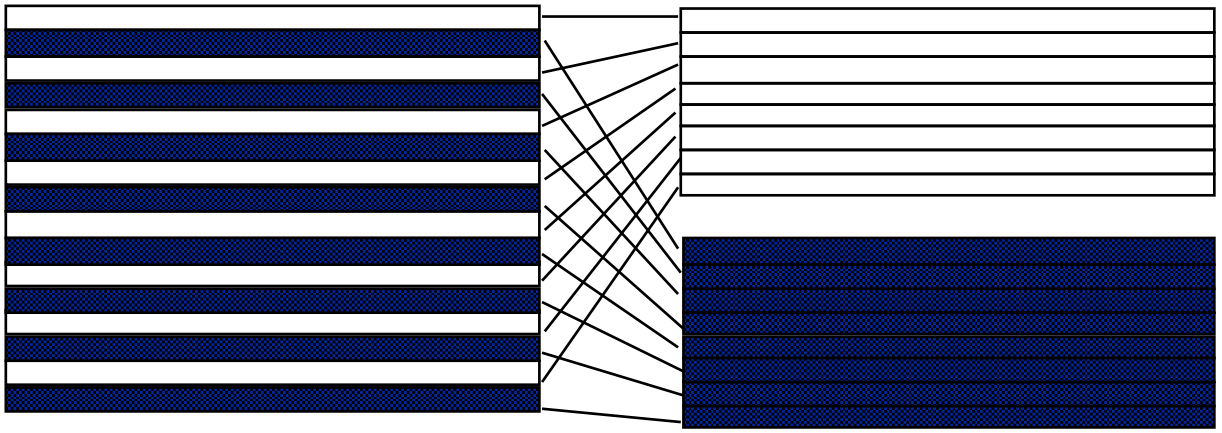


4:4:4



Interlaced Video Coding

- MPEG-2 supports **progressive** and **interlaced** video
  - Interlaced video is used for digital TV and HDTV
  - In interlaced video, each frame consists of two fields
- Frame picture
  - All scanlines from both fields are interleaved to form a frame
  - A frame is divided into 16 x 16 macroblocks for coding
- Field picture
  - A field treated as a separate picture
  - A field picture is divided into 16 x 16 macroblocks for coding



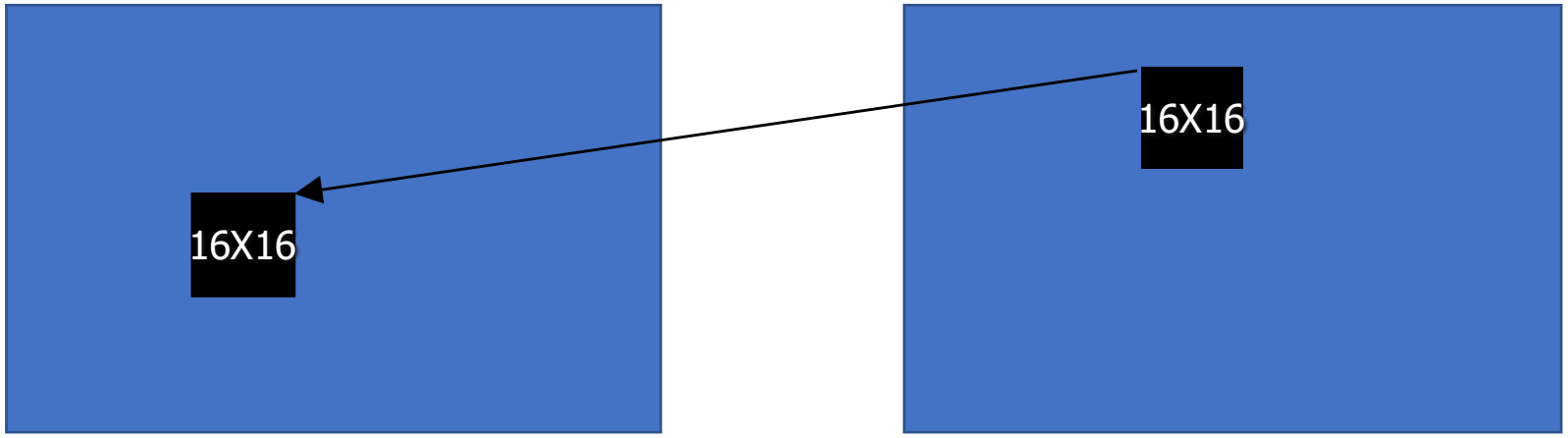
Outline

- Introduction
- MPEG-2 profiles and levels
- MPEG-2 syntax
- **Motion Estimation**
- Compression and quantization
- Scalability
- Others

Prediction Modes for Frame/Field Pictures

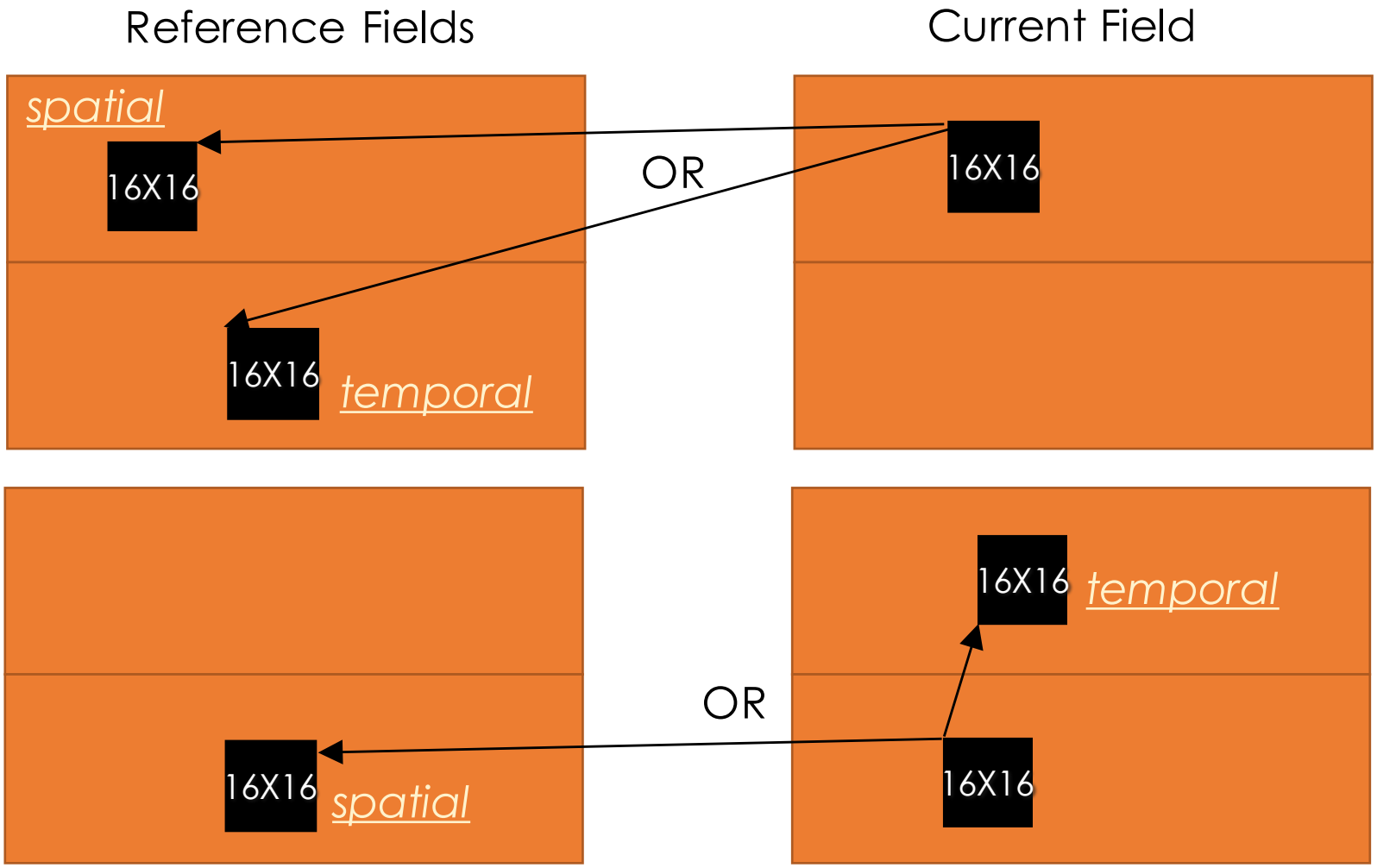
- **Frame prediction** for **frame pictures**
- **Field prediction** for **field pictures**
- **Field prediction** for **frame pictures**
- Dual-prime for P-pictures (either frame/field)
- 16x8 MC for **field pictures**

Frame Prediction for Frame Pictures



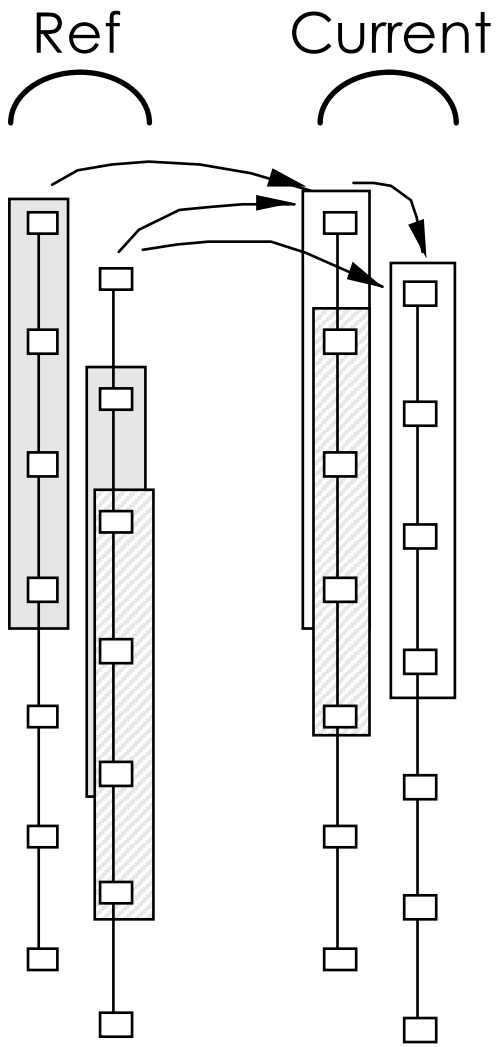
- Same with H.261, MPEG-1

Field Prediction for Field Pictures

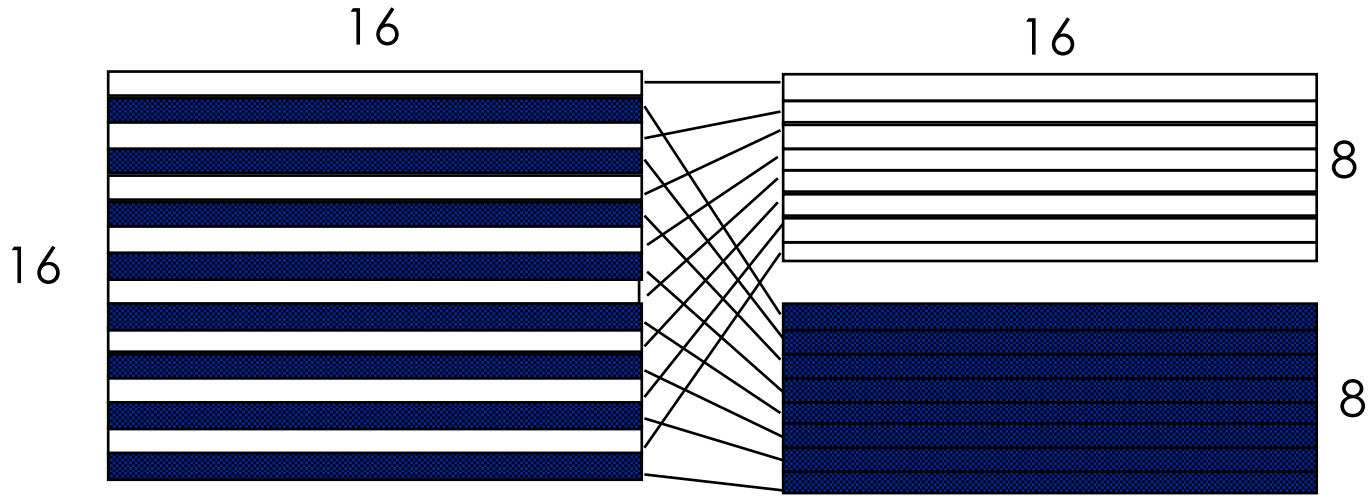


- Up to four motion vectors can be generated for each MB in a B-frame

Field Prediction for Field Pictures

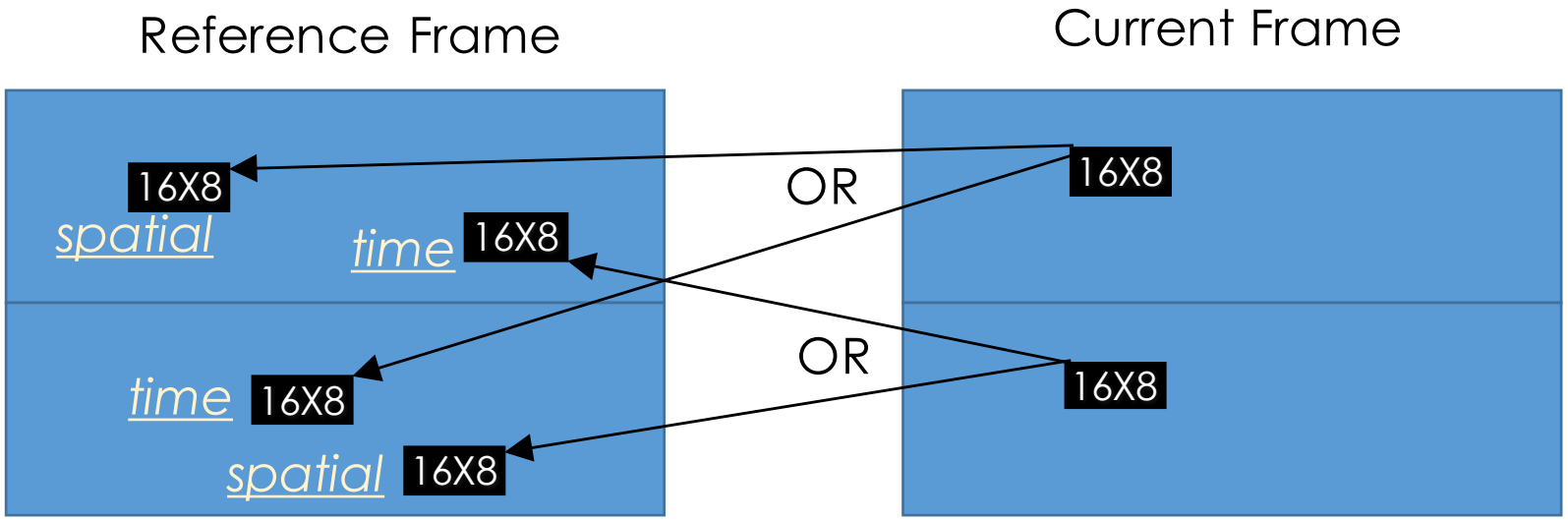


Field Prediction for Frame Pictures



- 16 x 8 ME & MC
  - since a 16 x 16 macrkblock from a frame picture is splits to two 16 x 8 parts
- Prediction from either field of the previous frame
- Good for fast motion

Field Prediction for Frame Pictures



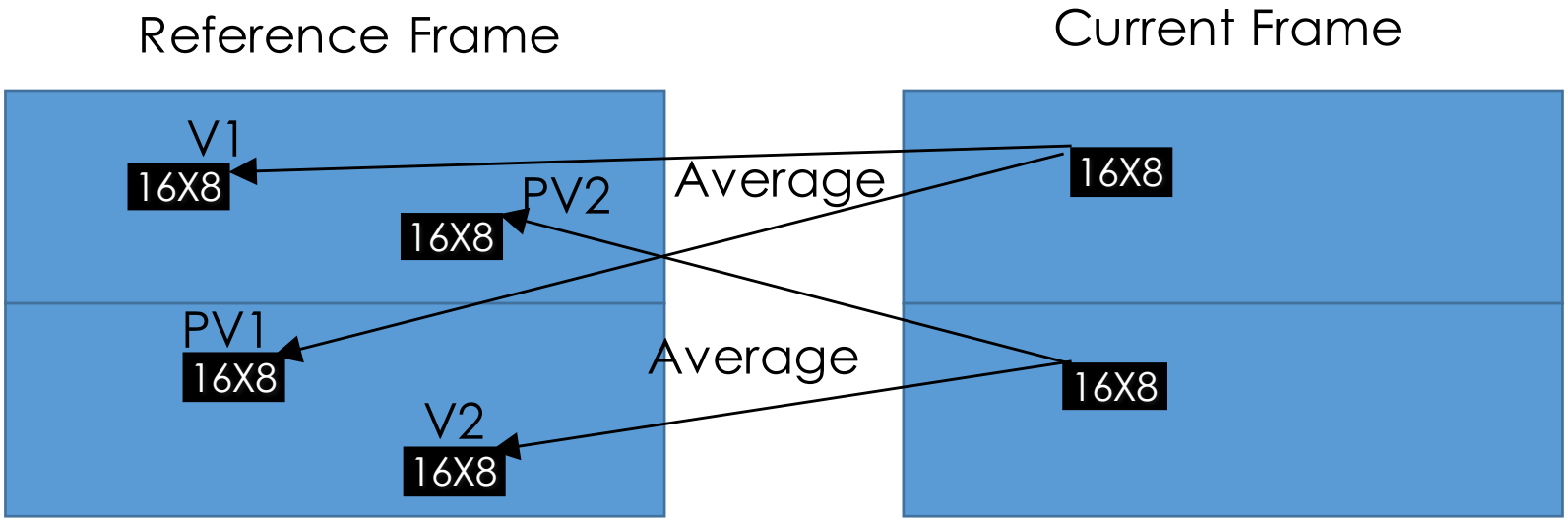
- Why the bottom field is not predicted by the top field in the current frame?
  - Since they belong to the same original frame and hence are taken at the same time
- But, if so, why need field prediction
  - Increase number of prediction choices

Dual-Prime for P-Pictures

- Transmit one vector per MB for P-pictures
- Two preliminary predictions are computed, which are then averaged together to form the final prediction
- The first preliminary prediction is identical to Filed Prediction, except that the reference pels must all come from the previously coded fields having the same parity (top or bottom)
- The second preliminary prediction is derived using computed motion vectors plus a small differential motion vector correction. Reference pels are taken from the opposite field as the first preliminary prediction
- Reference pels which are obtained using the transmitted motion vectors are taken from one field for field-pictures and from two fields for frame-pictures
- Can be applied for both frame pictures and field picture!

Dual-Prime for P-Pictures

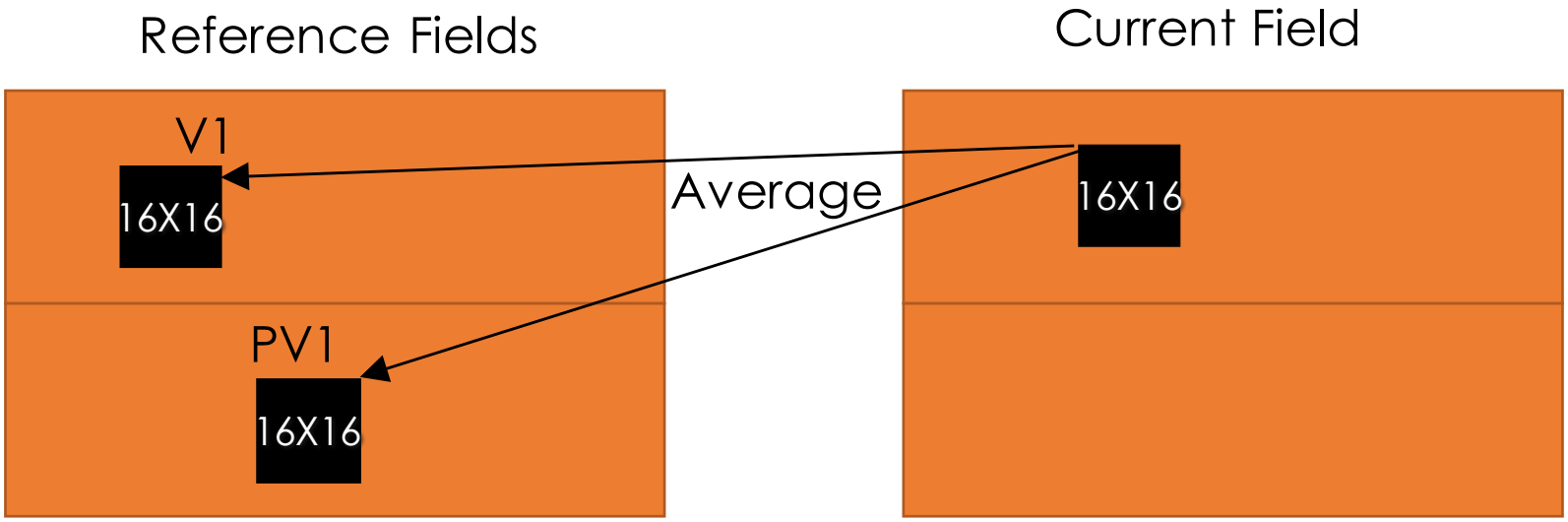
For Frame Pictures



- V: actual motion vector
- PV: motion vector derived from V

Dual-Prime for P-Pictures

For Field Pictures



- V: actual motion vector
- PV: motion vector derived from V

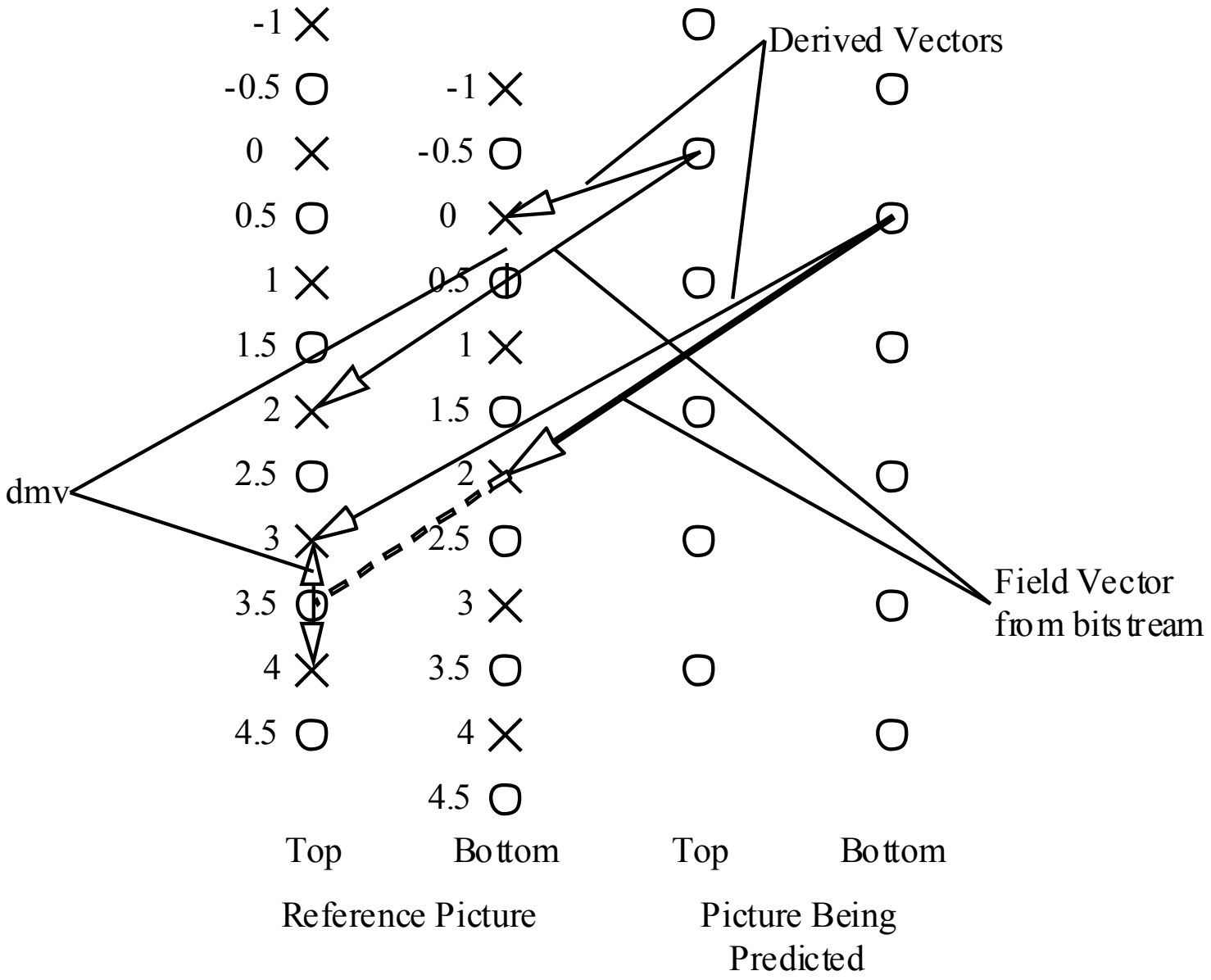
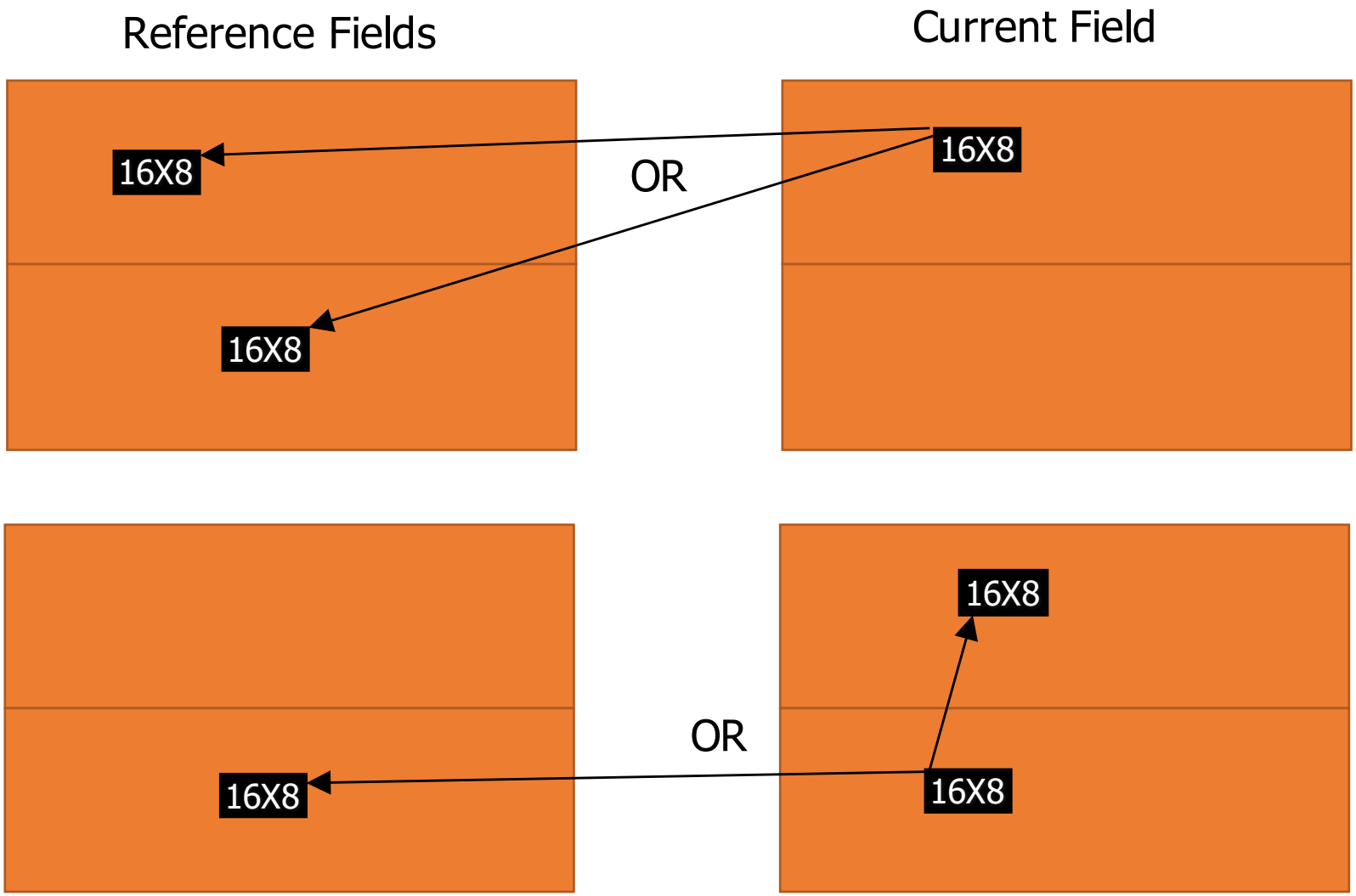


Figure 7-12. Scaling of motion vectors for dual prime prediction

16x8 MC for Field Pictures

- Splits the field picture MB into upper and lower half
- P-picture has 2 MV
- B-picture has 2 or 4 MV
- Useful for field pictures that contain a lot of irregular motion

16x8 MC for Field Pictures



Comparison

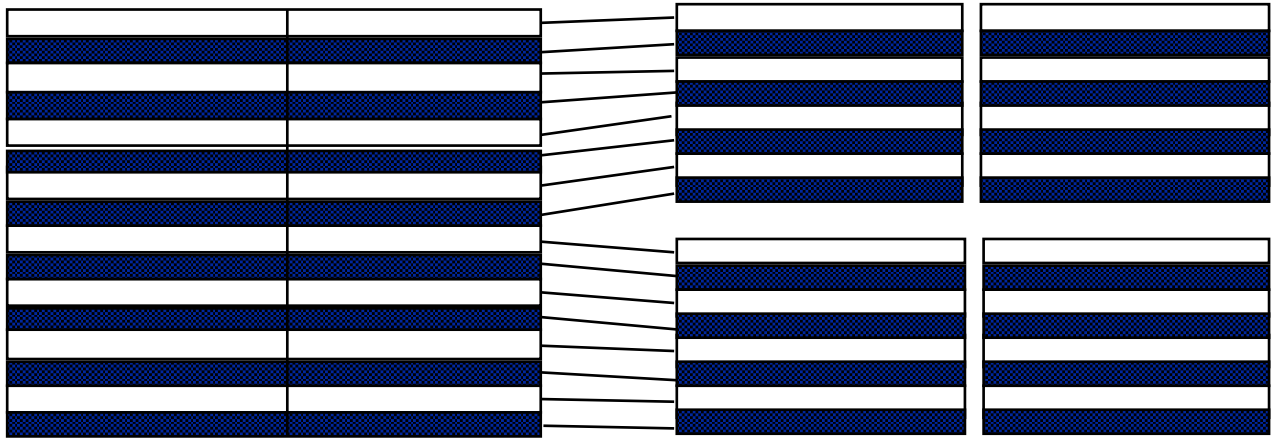
Motion Compensation Mode	Used in Field Pictures?	Used in Frame Pictures?
Frame Prediction for Frame Pictures	No	Yes
Field Prediction for Field Pictures	Yes	No
Field Prediction for Frame Pictures	No	Yes
Dual-Prime for P-Pictures	Yes	Yes
16x8 MC for Field Pictures	Yes	No

Outline

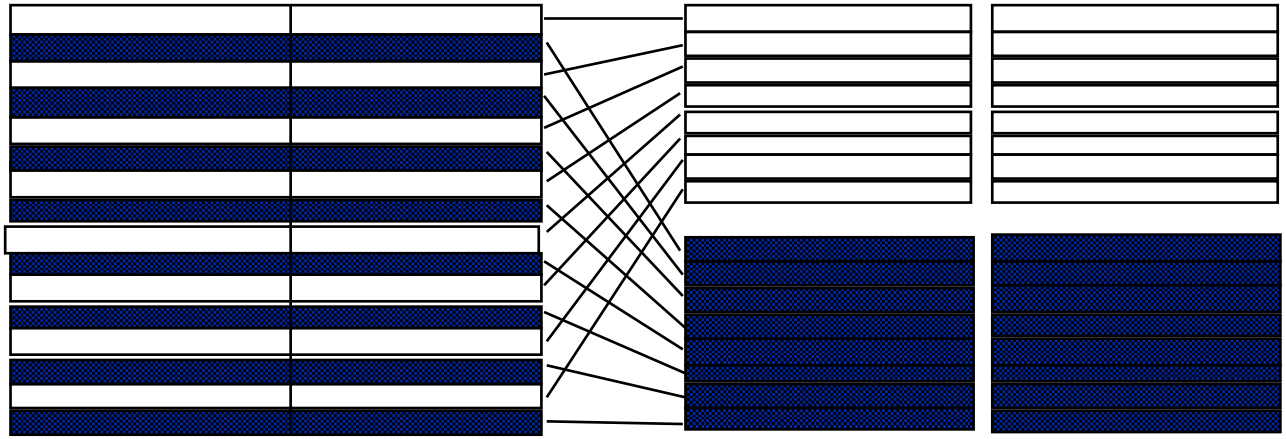
- Introduction
- MPEG-2 profiles and levels
- MPEG-2 syntax
- Motion Estimation
- **Compression and quantization**
- Scalability
- Others

Frame/Field DCT

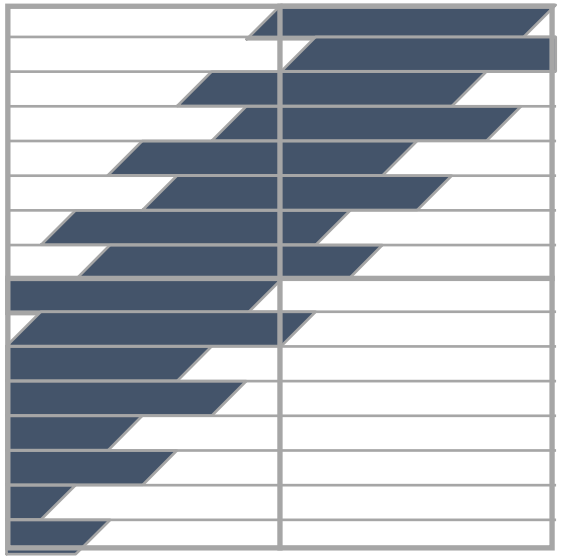
Frame format



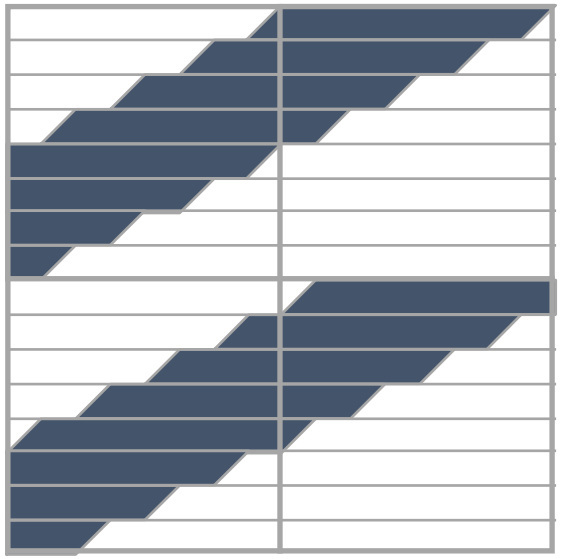
Field format



Frame/Field DCT



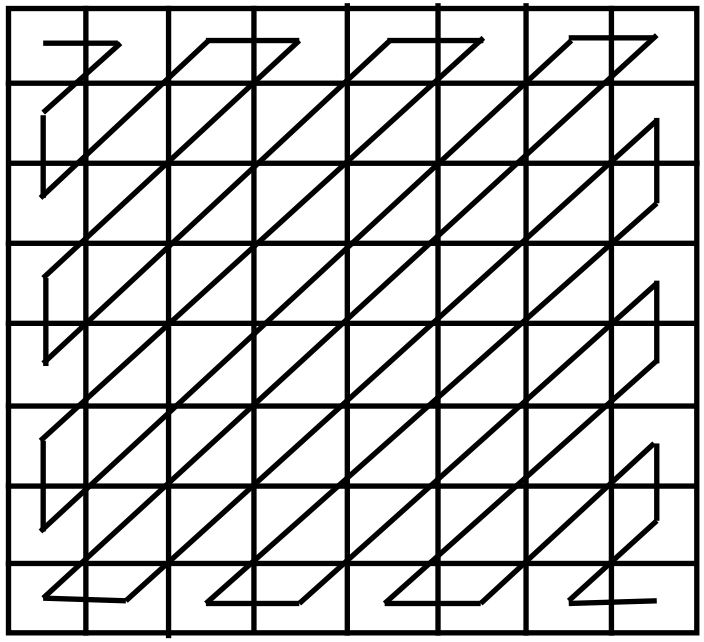
Frame blocks



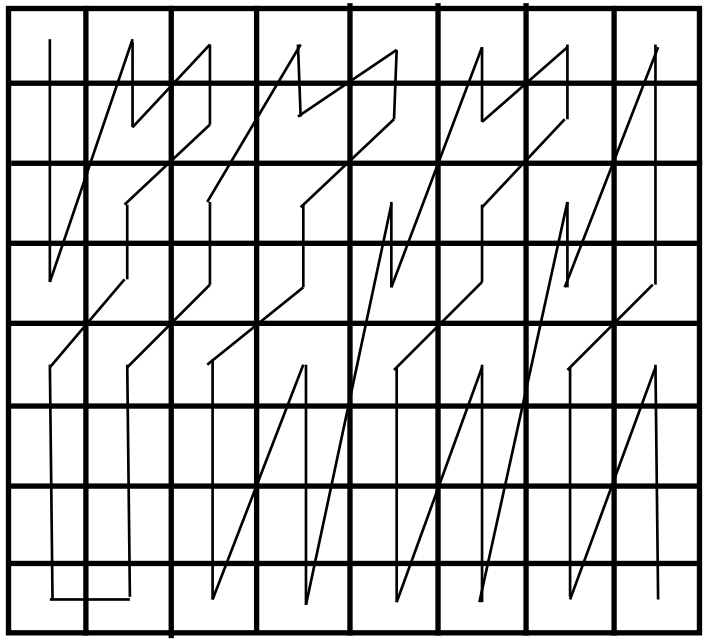
Field blocks

Progressive/Interlaced Scan

Zigzag (progressive)



Alternate (interlaced)



- In interlaced video, rows are subsampled
- Vertical high-frequency coefficients might be slightly larger

Outline

- Introduction
- MPEG-2 profiles and levels
- MPEG-2 syntax
- Motion Estimation
- Compression and quantization
- **Scalability**
- Others

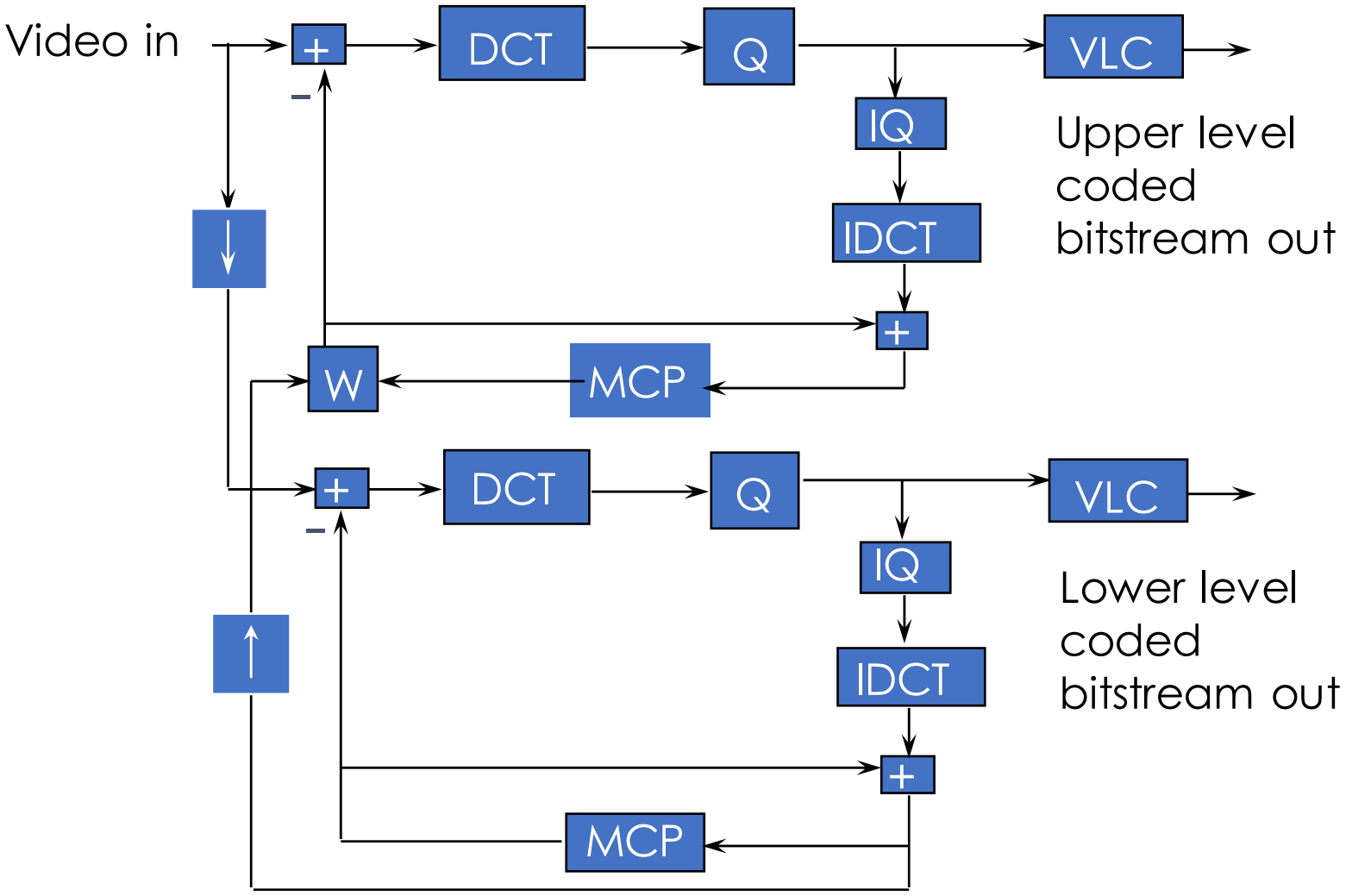
MPEG-2 Scalability

- First propose scalability
  - Scalable coding, also known as layered coding
  - A base layer, and one ore more enhancement layer
- Base layer
  - Independently encoded, transmitted and decoded
- Enhancement layer
  - Depends on the base layer or the previous enhancement layer
- Why layered coding?
  - Flexible for different channel capacity
  - Adapting to variable-bitrate (VBR) channel (e.g., bandwidth fluctuation)
  - Coping with noisy channel

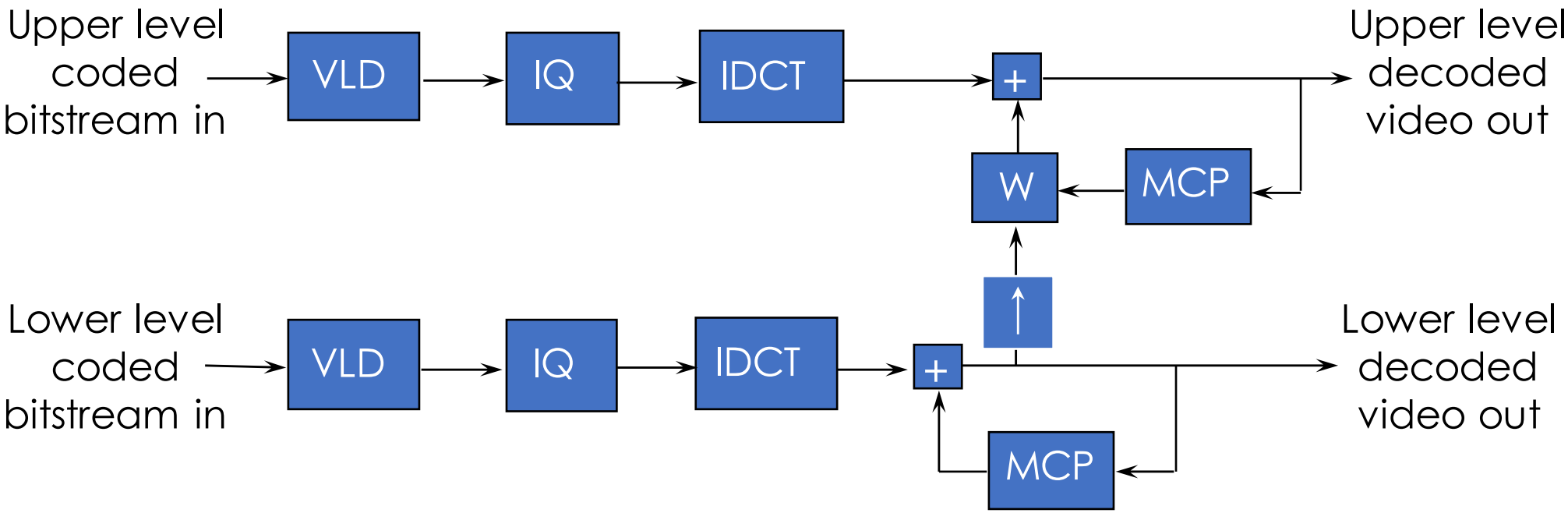
Types of Scalability

- Spatial scalability
- SNR scalability
- Temporal scalability
- Hybrid scalability
- Data partitioning
  - Quantized DCT coefficient are split into partitions

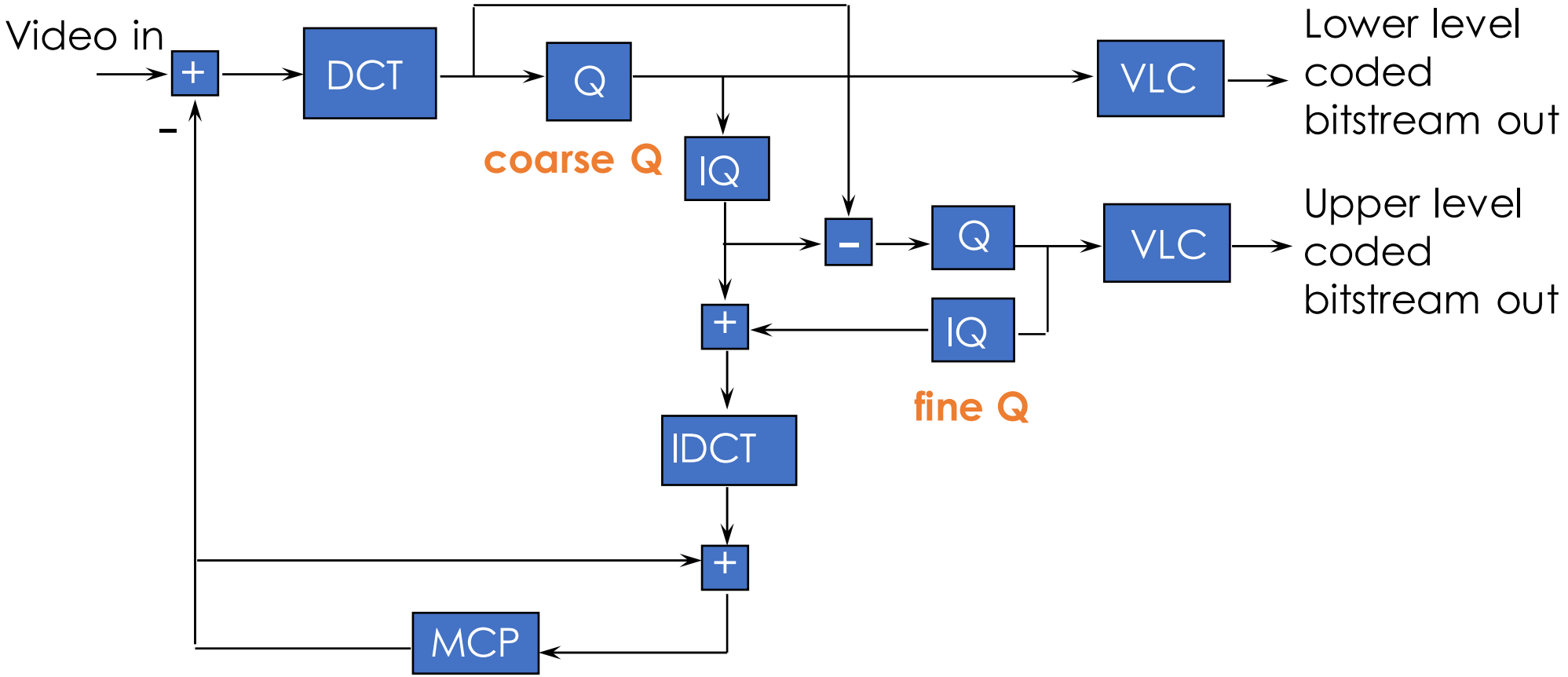
Spatial Scalable Encoder



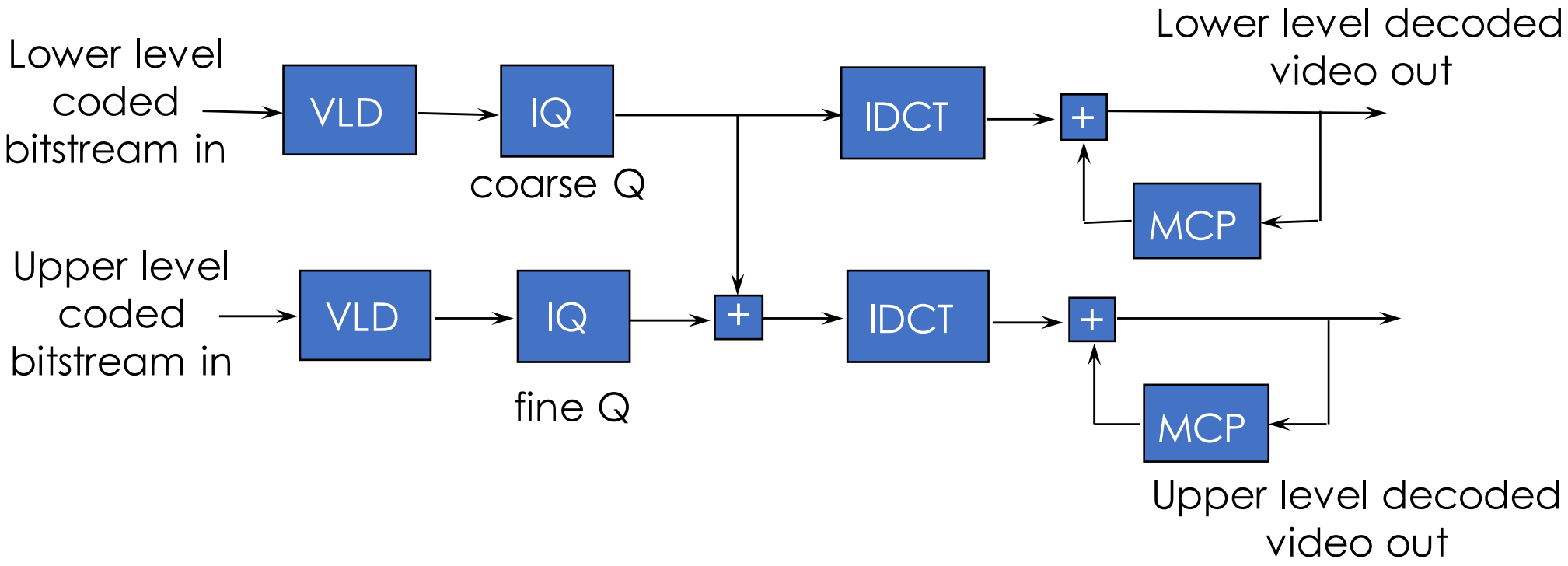
Spatial Scalable Decoder



SNR Scalable Encoder



SNR Scalable Decoder



Temporal and Hybrid Scalability

- Temporal Scalability
  - Leverage the structure of I-P-B frames
- Hybrid Scalability
  - Spatial and temporal scalability
  - SNR and spatial hybrid scalability
  - SNR and temporal hybrid scalability

Error Resilience

- Slice structure
- Concealment motion vectors
- Data partition
- SNR scalability
- Spatial scalability
- Temporal scalability
- Hybrid scalability
- Intra pictures
- Intra slices

MPEG-2 Error Concealment

- Typical way
  - Replace with skipped macroblock
- MPEG-2
  - Intra pictures may optionally contain coded motion vectors (intra-mv)

Outline

- Introduction
- MPEG-2 profiles and levels
- MPEG-2 syntax
- Motion Estimation
- Compression and quantization
- Scalability
- **Others**

MPEG-2 Systems

- Main function of system
  - to provide a means of combing, or multiplexing, several types of multimedia information into one stream
- Methods for multiplexing
  - Time Division Multiplexing (TDM)
  - Packet multiplexing

Two Types of MPEG-2 Streams

- **Program Stream**
  - Is MPEG-1 like and intended for error-free media
  - Use a modify function syntax to support new function
  - Typically employ long and variable -length packets for software based process
  - Error-free environment (compressed data are stored on a disk)
- **Transport stream**
  - Differs from MPEG-1
  - Offers robustness necessary for noisy channels as well as the ability to include multiple programs in a single stream
  - Suitable for delivering compressed video & audio over error-prone channels