

MPEG standards

© 1998-2016 Josef Pelikán
CGG MFF UK Praha

pepca@cgg.mff.cuni.cz

<http://cgg.mff.cuni.cz/~pepca/>



History

- ◆ **Moving Pictures Experts Group** (from 1988)
 - board of experts with ISO
- ➔ standard: **MPEG-1** (ISO 11172, 1991)
 - audiovisual media up to 1.5 Mbit/s (disc, CD-ROM)
- ➔ standard: **MPEG-2** (ISO 13818, H.262, 1994)
 - general use (up to 100 Mbit/s), interlacing, hierarchy
- ➔ further activity: **MPEG-4** (ISO 14496, 1999-2014+)
 - narrow transmission channels (8 to 64 kbit/s)
 - more efficient compression, HD video, ..



MPEG applications

- ➔ **asymmetric** (only decoding is time-critical)
 - videotex, multimedia, games, advertisements, ...
- ➔ **symmetric** (encoding and decoding in real-time)
 - video-mail, video-telephony, videoconferencing, digital TV broadcasting (live studios), ...
- **MPEG standards** are defining **data formats and decoding methods** only (not encoding)



Requirements

- direct access to arbitrary frame (video editing, cutting)
- fast forward with search („>>“, „<<“)
- reverse replay („<“)
- synchro: audio \Leftrightarrow video
- low sensitivity to transmission errors
- low latency (interactive applications)
- flexibility (resolution, pixel aspect ratio, fps)
- affordability (HW decoders with 1990's technology)

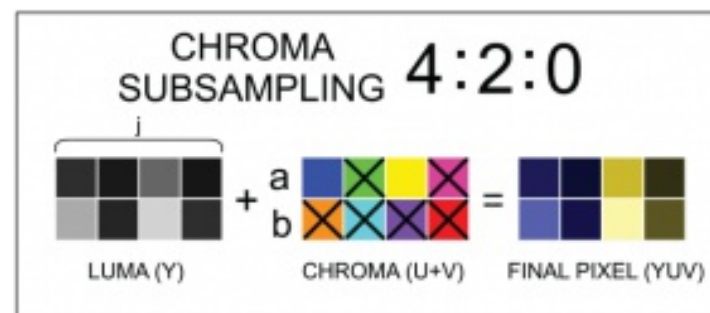
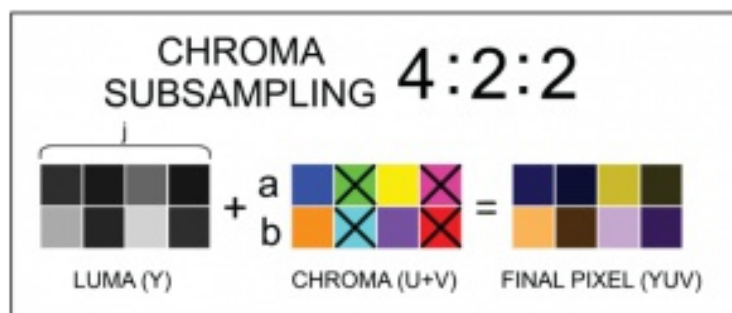
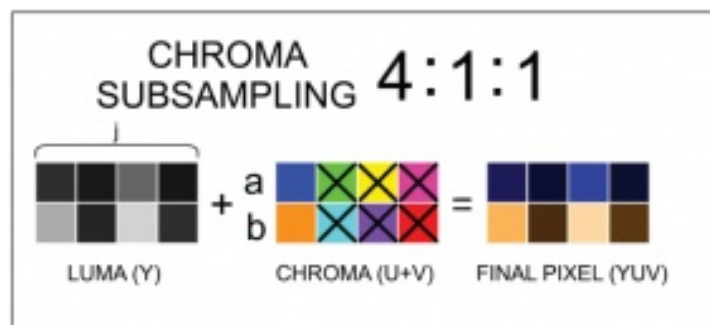
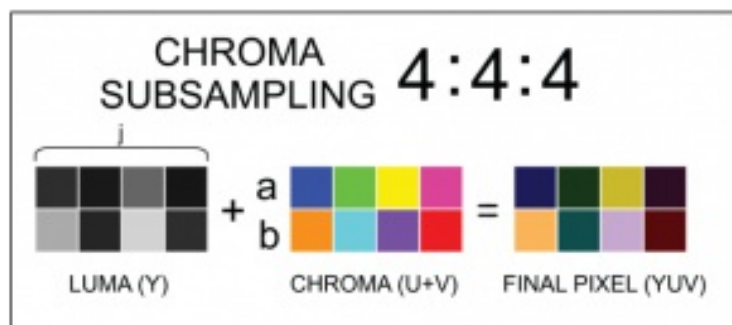


Compression principles

- ◆ **motion compensation**
 - macroblocks 16×16 px
- ◆ various **color sampling schemas** (YC_bC_r)
 - chroma components usually encoded in less quality
- ◆ **DCT** followed by **quantizing** (à la JPEG-1)
 - DCT blocks 8×8 px



Chroma subsampling



CCIR 601 (HW standard)

JPEG-1 classic

(Image © Videomaker, Inc.)



MPEG-1: typical formats

Frame frequency (Hz)	30 (NTSC)	25 (PAL)
CCIR 601		(sampling 4:2:2)
comp Y	720 × 480	720 × 576
comp C _b , C _r	360 × 480	360 × 576
SIF (Source Input Format)		(4:2:0)
comp Y	360 × 240	360 × 288
comp C _b , C _r	180 × 120	180 × 144
Significant region SIF		(4:2:0)
comp Y	352 × 240	352 × 288
comp C _b , C _r	176 × 120	176 × 144



MPEG-1: compression examples

- ◆ **frame size 352×240 px**
 - significant SIF
- ◆ **color depth 12 bits/pixel**
 - 8 bits/channel using 4:2:0 sampling
- ◆ **frame rate 30 Hz**
 - NTSC (USA, Japan)
- ➔ **compression from 30.4 Mbit/s to 1.15 Mbit/s (26:1)**
 - Reconstruction quality is comparable to analog VHS

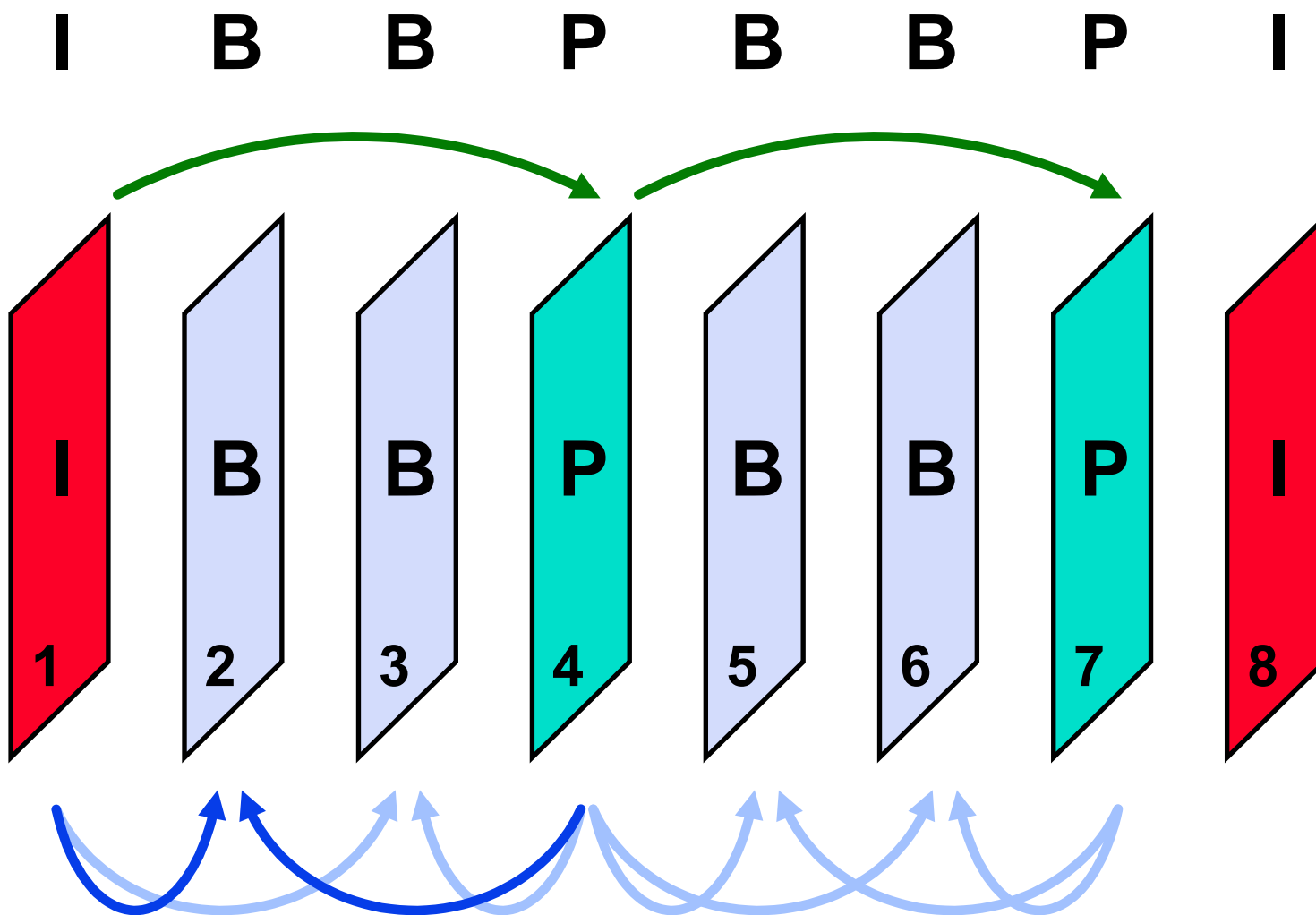


Frame types

- ◆ **I („intra”)**
 - independent, key frame
- ◆ **P („predicted”)**
 - motion compensation from previous I or P frame
- ◆ **B („bidirectionally predicted”) - interpolated**
 - motion compensation both from previous and future reference frames (I or P)
- ◆ **D („DC-coded”) - obsolete (only in MPEG-1)**



Frame prediction





Bitrate by frame type

- **SIF** (352×240), frequency **30 fps**
 - total bitrate: 1.15 Mbit/s
- **pattern: I P BB P BB P BB P BB P BB**
 - 16 frames in one group (GOP)

Frame type	Average frame size	Freq.	Bitrate
I	156 kbit	1.88	292 kbit/s
P	62 kbit	9.38	581 kbit/s
B	15 kbit	18.75	281 kbit/s



Macroblock types (16×16)

- ➔ **„intra“**
 - no prediction = independent
- ➔ **„forward“** - forward motion compensation
 - motion compensation from previous I or P frame
- ➔ **„backward“** - backward motion compensation
 - motion compensation from future I or P frame
- ➔ **„inter(polated)“** - bidirectional motion comp.
 - prediction source is interpolated from both previous and future reference fr.



Macroblock prediction

$$\underline{\text{Intra}_i(\mathbf{x})} = 0$$

w/o prediction

$$\underline{\text{Forw}_i(\mathbf{x})} = I_{i-k}(\mathbf{x} + \underline{\text{MV}_{-k}})$$

forward

“motion vector”

$$\underline{\text{Back}_i(\mathbf{x})} = I_{i+l}(\mathbf{x} + \text{MV}_{+l})$$

back

$$\underline{\text{Inter}_i(\mathbf{x})} = \frac{I_{i-k}(\mathbf{x} + \text{MV}_{-k}) + I_{i+l}(\mathbf{x} + \text{MV}_{+l})}{2}$$

bidirectional



DCT block coding (8×8)

- ◆ **Discrete Cosine Transform** on 8×8 px blocks
- ◆ **linear quantizers**
 - quantizer tables can be adapted by scaling and switching („intra” vs. „inter” blocks)
- ◆ **zig-zag pass**
- ◆ **Huffman coding**
 - tuples [**zeroes, value**], less frequent tuples are divided into components



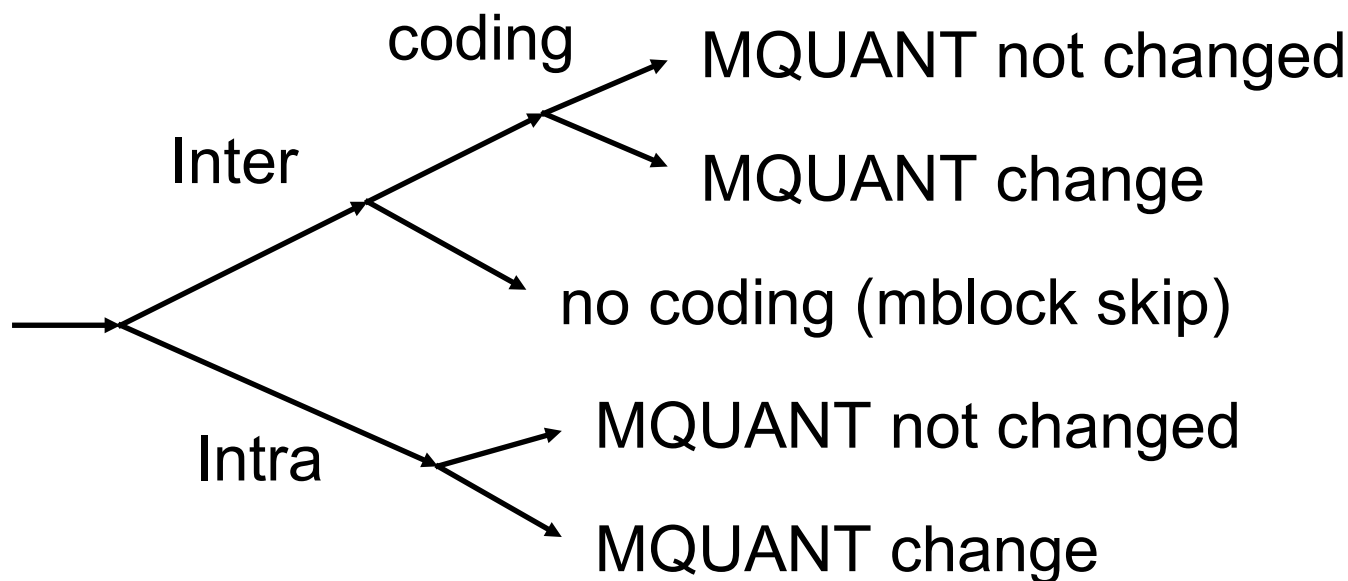
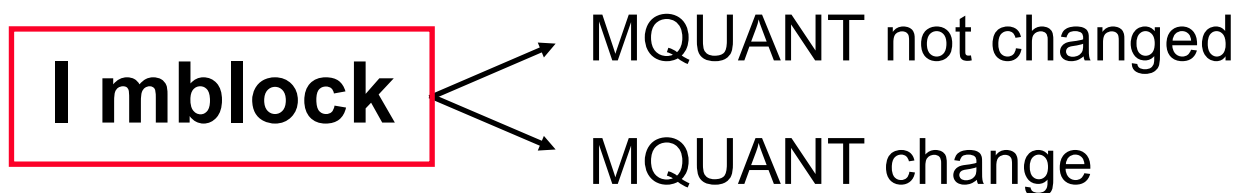
Quantizing of DCT coefficients

- ◆ **MPEG header** can contain two explicit quantizer tables
 - implicit tables - **intra**: JPEG for Y channel and 50%, **inter**: all 16-s
- ◆ **adaptation** on macroblock level
 - **MQQUANT** factor (1 to 31) - for AC coefficients only

$$F^Q(u, v) = \text{round} \left(\frac{8 \cdot F(u, v)}{\text{MQQUANT} \cdot Q(u, v)} \right)$$

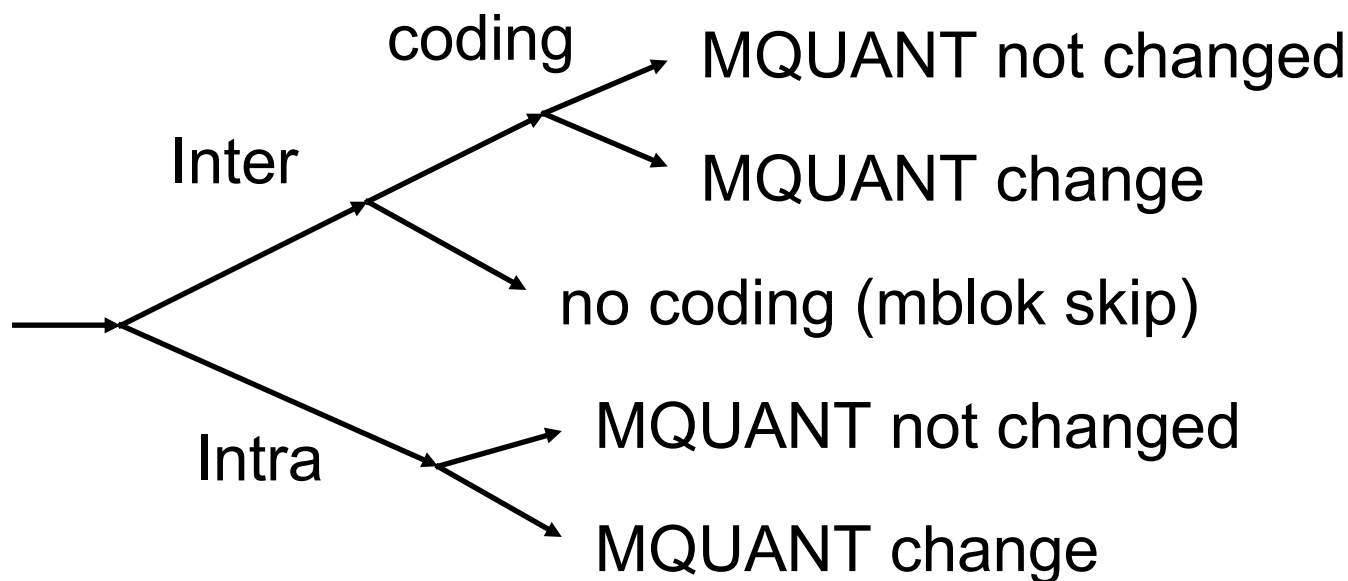
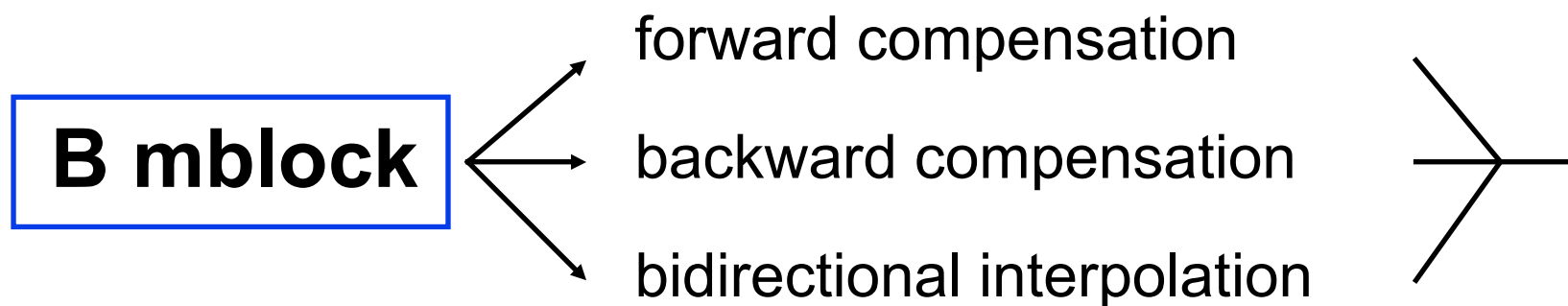


I- and P-macroblock coding





B-macroblock coding





Macroblock distribution example

frame type	macroblock type				
	I	P	B	zero MV	skip
I	3.300				
P	897	8.587		5.128	568
B	60	7.356	22.845		429



DCT block distribution example

frame type	macroblock type			
	I	P	B	zero MV
I	19.800			
P	5.382	30.730 (60%)		18.146 (59%)
B	360	8.176 (19%)	18.853 (14%)	



Example of MPEG decoder load

function	load (%)
MPEG header & stream decoder	0.44
Huffman decoder & dequantizer	19.00
Inverse DCT (8×8)	22.10
Motion compensation	38.64
Color transform & rendering	19.82



MPEG file structure (hierarchy)

① „sequence”

- global parameters (resolution, fps, max. bitrate, memory requirements, quantizer tables, ..)

② „group of pictures” (GOP)

- direct access, (independent) decoding unit

③ „picture”

- basic element of the sequence (reference number, frame type, synchro data)



MPEG file structure (hierarchy)

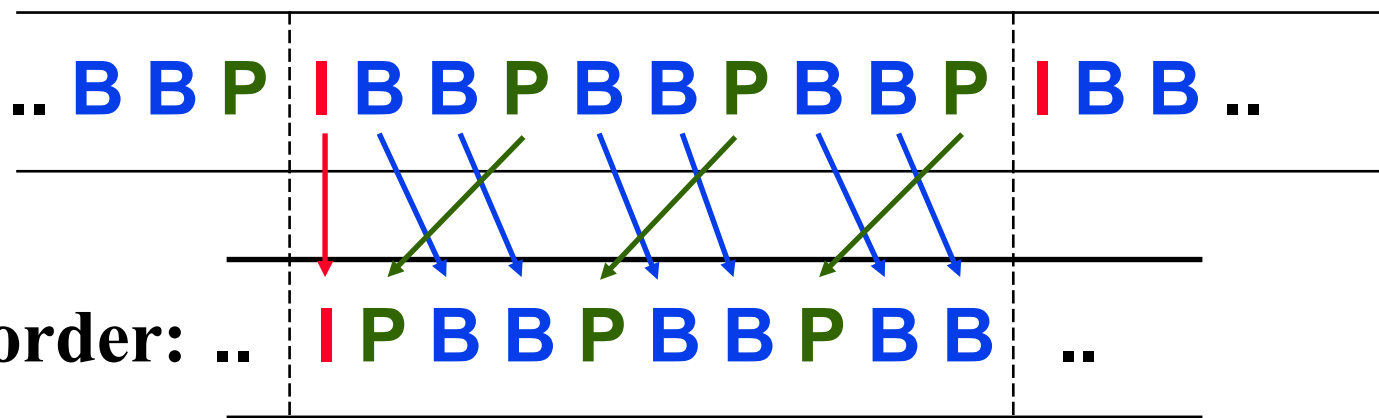
- ④ **„slice”**
 - error recovery, quantizer adaptation (MQQUANT)
- ⑤ **„macroblock”**
 - motion compensation unit (type, motion vector[s], DCT black bitmask, MQQUANT)
- ⑥ **„block”**
 - DCT transform unit (Huffman code after “zig-zag” pass through quantized DCT coefficients)



Group of pictures (GOP)

- ◆ at least one **I-frame** is mandatory
- ◆ **begins** with a B-frame or an I-frame
 - if starting B-frame[s] are not dependent on previous reference frame, this GOP is **„closed group“**
- ◆ **ends** with a P-frame or I-frame

Example:



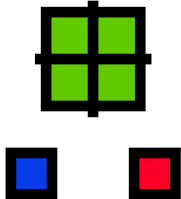
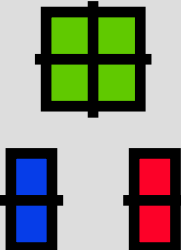
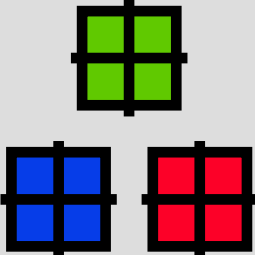


MPEG-2 standard

- ➔ **w/o bitrate restrictions** (10 Mbit/s or more)
- ➔ different levels of **quality** and **compression ratio**
 - „profiles” and „levels”, various chroma sampling schemes (4:2:0 to 4:4:4)
- ➔ **interlacing support** for digital TV broadcasting
 - one **frame** can be divided into two **fields** (upper and lower)
- ➔ **extension layers** („scalable syntax”)
 - decoder (client) can choose quality (bitrate) – based on display resolution, reception quality..



MPEG-2: chroma sampling

Frame rate (Hz)		30 (NTSC)	25 (PAL)
→ 4:2:0			
comp Y		720 × 480	720 × 576
comp C _b , C _r		360 × 240	360 × 288
→ 4:2:2			(CCIR 601)
comp Y		720 × 480	720 × 576
comp C _b , C _r		360 × 480	360 × 576
→ 4:4:4			
comp Y		720 × 480	720 × 576
comp C _b , C _r		720 × 480	720 × 576



MPEG-2: extension layers

- ➔ **„data partitioning”**
 - e.g. critical part (headers) vs. the rest
- ➔ **„SNR scalability”**
 - equal resolution, different quality (DCT refinement)
- ➔ **„spatial scalability”**
 - different resolutions (see hierarchical JPEG mode)
- ➔ **„temporal scalability”**
 - equal resolution and quality, different frame-rates

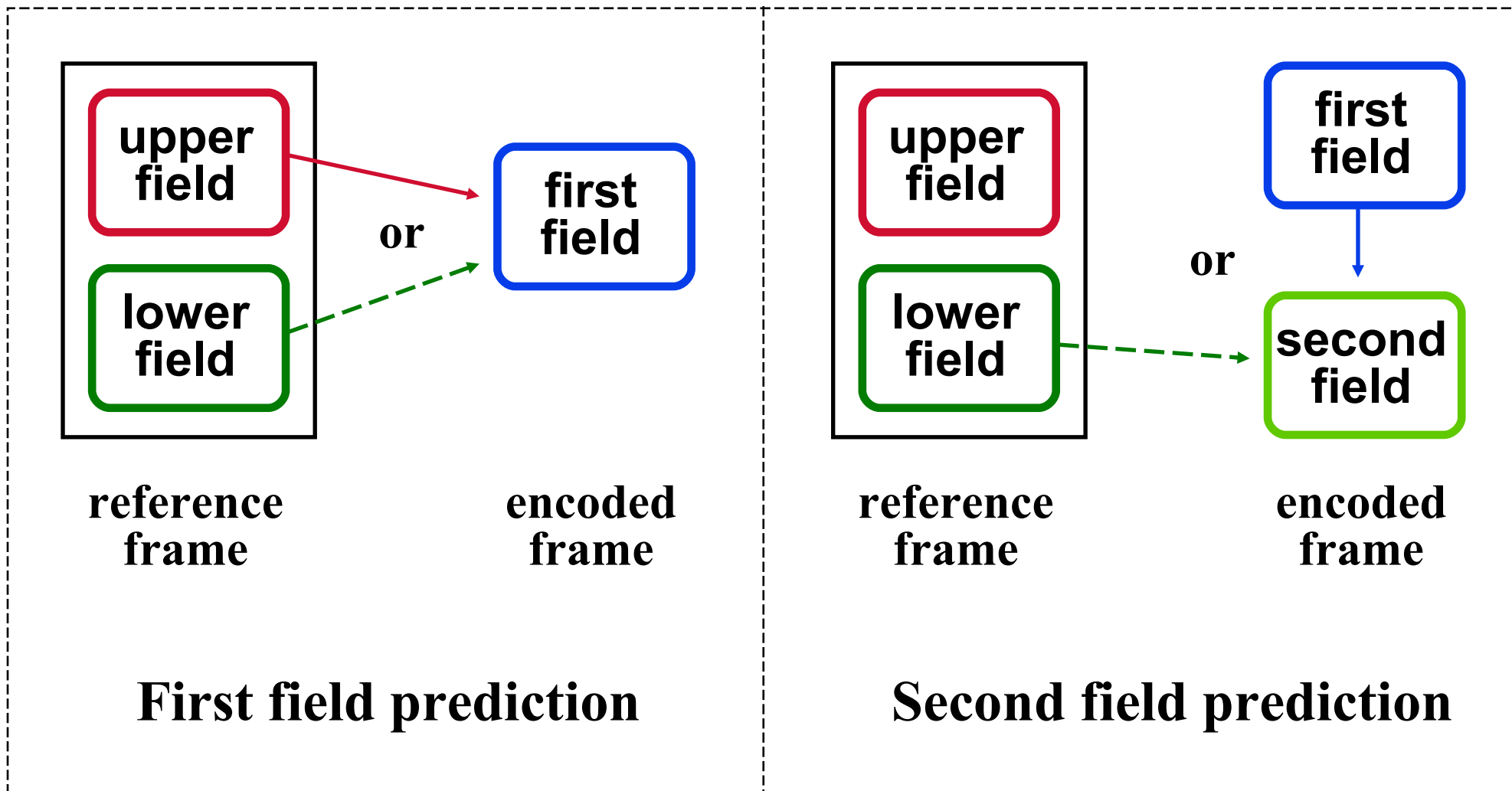


Interlacing

- **luminance (Y)** can be interlaced or transmitted in two fields (upper & lower)
 - on macroblock level
- **chromacity** values are always interlaced in macroblocks
- ➔ **more prediction modes**
 - inter-field predictions (**I-I**, **I-P**, **P-P**, **B-B**)
- ➔ **new motion compensation mode**
 - „dual-prime prediction”

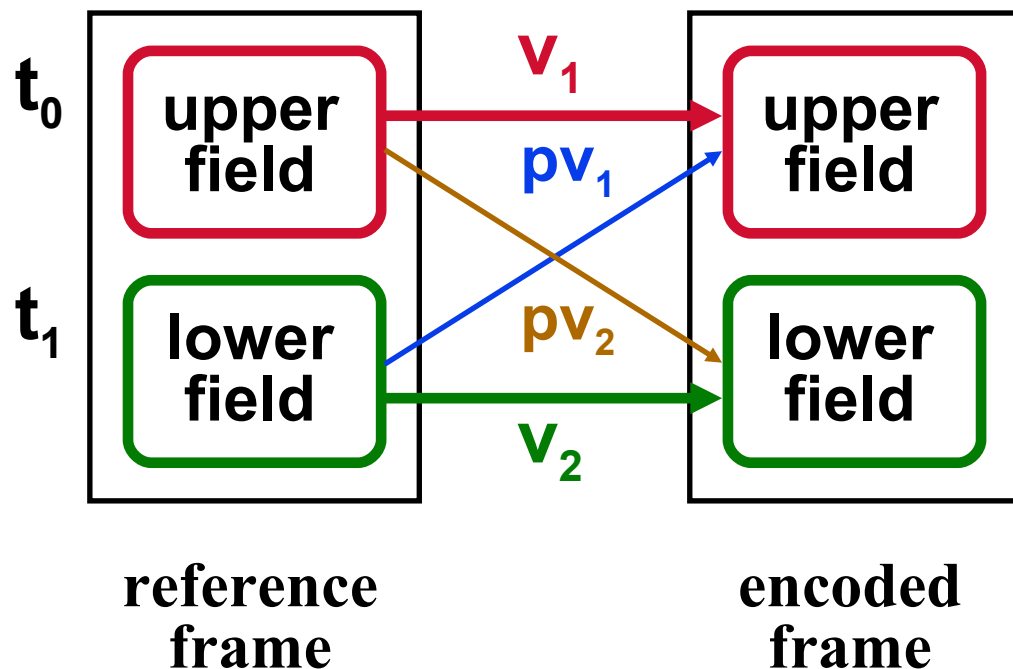


Inter-field prediction





„Dual-prime” compensation



$$pv_1 = \frac{v_1}{2} + \delta$$

$$pv_2 = \frac{3 \cdot v_2}{2} + \delta$$

$$\delta \in \{-1, 0, 1\}^2$$

Only v_1, v_2, δ are transmitted..



MPEG-2: more details

- every **motion vector** has half-pixel precision
- every frame can have its own **quantizer tables**
 - two table sets – for luminance **Y** and for chroma channels **C_b**, **C_r**
- **motion compensation „16×8”** (for interlacing)
 - separate motion vectors for each field (2 or 4 vectors per macroblock)



Profiles and levels in MPEG-2

- four **profiles** (Simple, Main, Main+, Next)
 - quality, backward compatibility, extensions
 - **Simple**: restricted Main (w/o B-prediction)
 - **Main**: standard profile (w/o further extensions)
 - **Main+**: 1st profile with extended features
 - **Next**: all hierarchical extensions (TV broadcasting)
- four **levels** (Low, Main, High-1440, High)
 - frame resolution (up to 1920×1152 px / 60 Hz)
 - **Low**: up to SIF, **Main**: up to CCIR 601, **High-1440**: up to 1440×1152 (HDTV), **High**: up to 1920×1152 (HDTV)



Basic resources

- **ISO/IEC JTC1 CD 11172: *Coding of moving pictures and associated audio for digital storage media up to 1.5 Mbits/s*, ISO 1992**
- **ISO/IEC JTC1 CD 13818: *Generic coding of moving pictures and associated audio*, ISO'94**
- **V. Bhaskaran, K. Konstantinides: *Image and Video Compression Standards, Algorithms and Architectures*, Kluwer Academic Publishers, Boston 1995, 161-194**



The End

More information:

ed. by H.-M. Hang, J. Woods: *Handbook of Visual Communications*, Academic Press, San Diego
1995, 384-392