
MPEG standardy

**© 1998-2001 Josef Pelikán
KSVI MFF UK Praha**

e-mail: Josef.Pelikan@mff.cuni.cz

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

Historie

- ◆ **Moving Pictures Experts Group** (od 1988)
 - komise expertů při ISO
- ➔ standard **MPEG - 1** (ISO 11172, 1991)
 - audiovizuální media až do 1.5 Mbit/s (disk, CD)
- ➔ standard **MPEG - 2** (ISO 13818, H.262, 1994)
 - širší použití (až 100 Mbit/s), prokládání, hierarchie
- ➔ aktivita **MPEG - 4** (konec 1998)
 - velmi úzké přenosové kanály (8 až 64 kbit/s)

Aplikace MPEG

- ➔ **asymetrické** (časově kritická je jen dekomprese)
 - videotex, multimedia, hry, reklama, ...
- ➔ **symetrické** (komprese i dekomprese v reálném čase)
 - video-mail, video-telefon, videokonference, digitální televizní vysílání, ...
- **standards MPEG** definují pouze **formáty dat** a **dekódovací metody** (nikoliv metody kódování)

Požadavky

- ➔ přímý přístup k libovolnému snímku, možnost střihu a editace
- ➔ rychlý posuv s vyhledáváním (“>>“, “<<“)
- ➔ zpětný chod (“<“)
- ➔ synchronizace: audio \Leftrightarrow video
- ➔ malá citlivost k chybám přenosu
- ➔ malé zpoždění (“latence”, kvůli on-line aplikacím)
- ➔ flexibilita (rozlišení, tvar pixelu, frekvence)
- ➔ dostupnost (HW dekodér technologií r. 1990)

Principy komprese

- ◆ **kompensace pohybu**
 - makrobloky velikosti 16×16
- ◆ **různá vzorkovací schémata (v YC_bC_r)**
 - barevné složky se mohou přenášet v menší časové nebo prostorové kvalitě (schémata 4:2:0, 4:2:2)
- ◆ **DCT následovaná kvantováním koeficientů**
 - DCT bloky velikosti 8×8

MPEG-1: typická rozlišení

snímková frekvence (Hz)	30 (NTSC)	25 (PAL)
CCIR 601	(vzorkování 4:2:2)	
složka Y	720 × 480	720 × 576
složky C _b , C _r	360 × 480	360 × 576
SIF (Source Input Format)	(4:2:0)	
složka Y	360 × 240	360 × 288
složky C _b , C _r	180 × 120	180 × 144
Významná oblast SIF	(4:2:0)	
složka Y	352 × 240	352 × 288
složky C _b , C _r	176 × 120	176 × 144

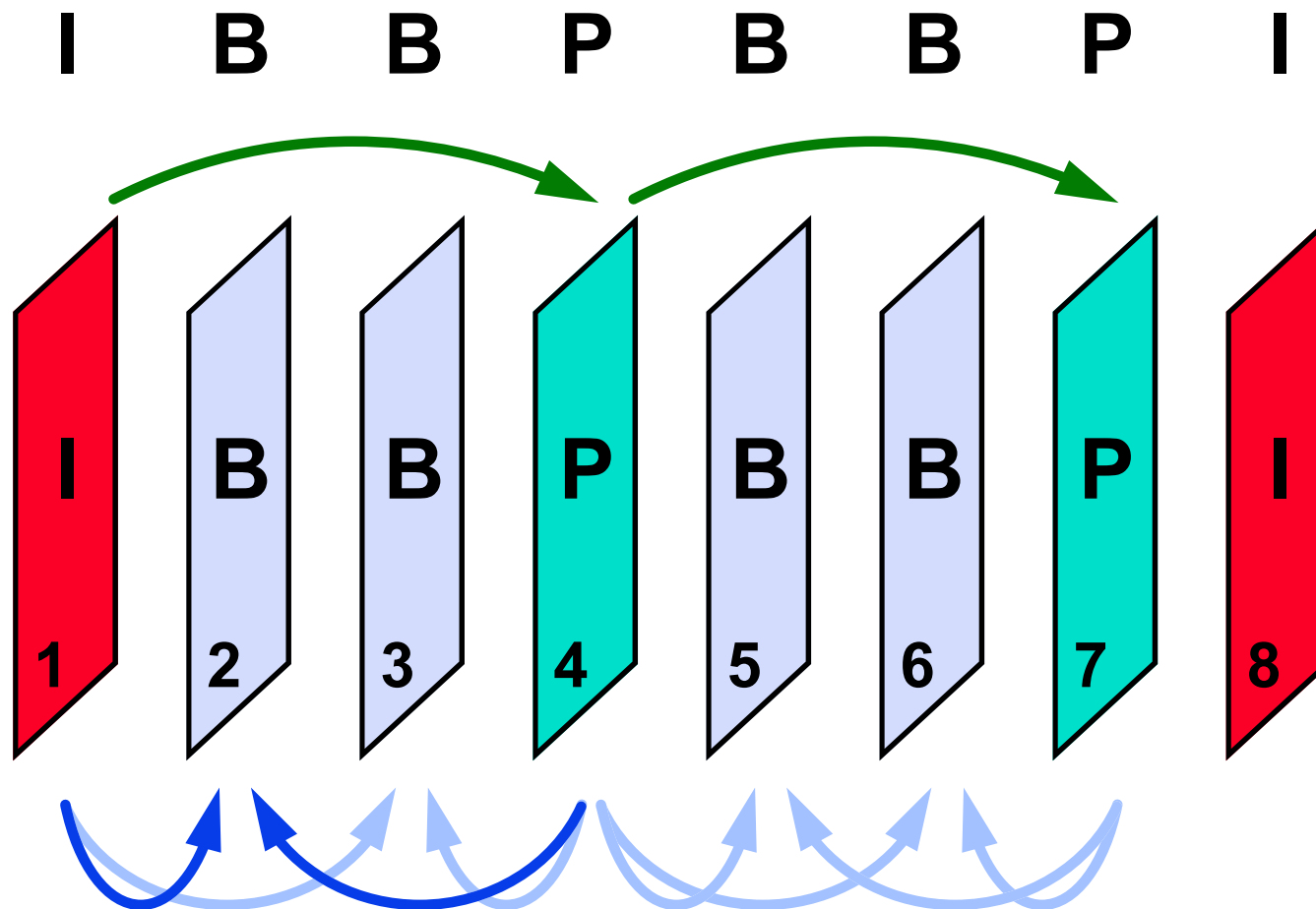
MPEG-1: typická komprese

- ◆ **snímek velikosti 352×240 pixelů**
 - významná oblast SIF
- ◆ **barevná hloubka 12 bitů/pixel**
 - 8 bitů/kanál při vzorkování 4:2:0
- ◆ **snímková frekvence 30 Hz**
 - NTSC (USA, Japonsko)
- ➔ **komprese z 30.4 Mbit/s na 1.15 Mbit/s (26:1)**
 - výsledná obrazová kvalita je nejméně tak dobrá jako u analogového systému VHS

Typy snímků

- ◆ **I (“intra”)**
 - nezávisí na jiných snímcích
- ◆ **P (“predicted”) - predikovaný**
 - kompenzace pohybu z předchozího I nebo P snímku
- ◆ **B (“bidirectionally predicted”) - interpolovaný**
 - kompenzace pohybu z předchozího a budoucího referenčního (I nebo P) snímku
- ◆ **D (“DC-coded”) - jen v MPEG-1**

Predikce snímků



Tok podle typu snímku

- ➔ **SIF** (352×240) při frekvenci **30 Hz**
 - celkový výsledný tok: 1.15 Mbit/s
- ➔ **vzor** opakování: **I P BB P BB P BB P BB P BB**
 - 16 snímků v jedné skupině (GOP)

typ snímku	průměrná velikost	frekv.	tok
I	156 kbit	1.88	292 kbit/s
P	62 kbit	9.38	581 kbit/s
B	15 kbit	18.75	281 kbit/s

Typy makrobloků (16×16)

➔ **“intra”**

– bez predikce, nezávisí na jiných snímcích

➔ **“forward”** - dopředná kompenzace pohybu

– kompenzace pohybu z předchozího I nebo P snímku

➔ **“backward”** - zpětná kompenzace pohybu

– kompenzace pohybu z budoucího I nebo P snímku

➔ **“inter(polated)”** - obousměrná kompenzace p.

– interpolace z předchozího i budoucího ref. snímku

Predikce makrobloků

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

bez predikce

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

dopředu

“motion vector”

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

zpět

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

obousměrně

Kódování DCT bloků (8×8)

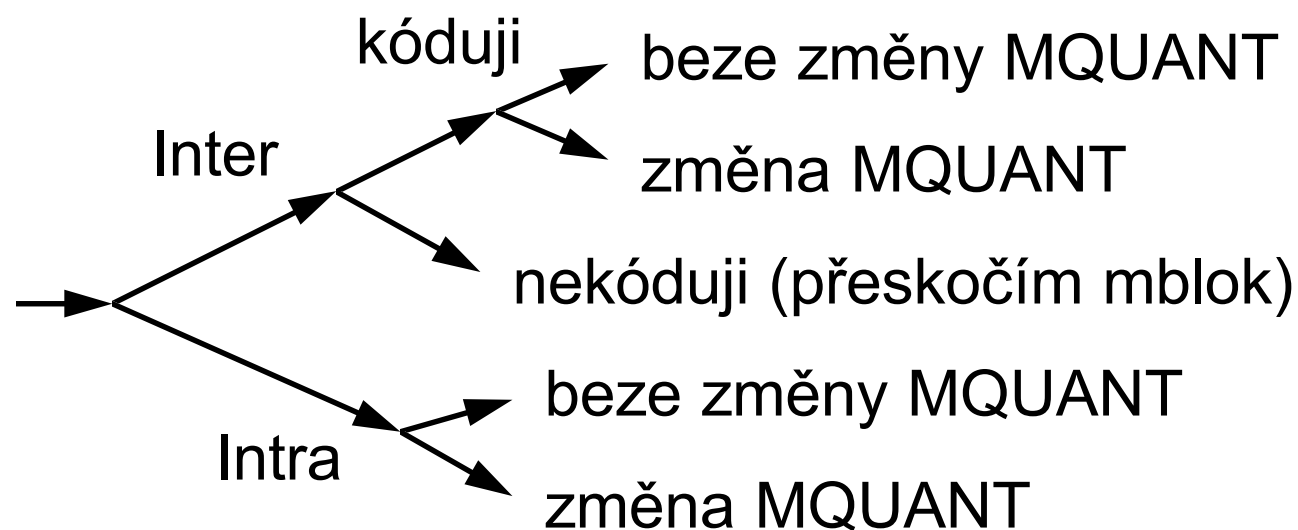
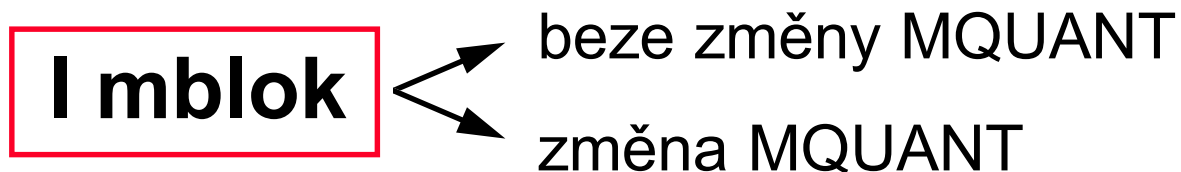
- ◆ **DCT** se provádí na blocích velikosti 8×8
- ◆ **lineární kvantování**
 - kvantovací tabulky se adaptují škálováním a přepínají (mezi “intra” a “inter” bloky)
- ◆ **cik-cak průchod**
- ◆ **Huffmanovo kódování**
 - kódují se dvojice [**počet nul, hodnota**], méně časté dvojice jsou zakódovány po složkách

Kvantování DCT koeficientů

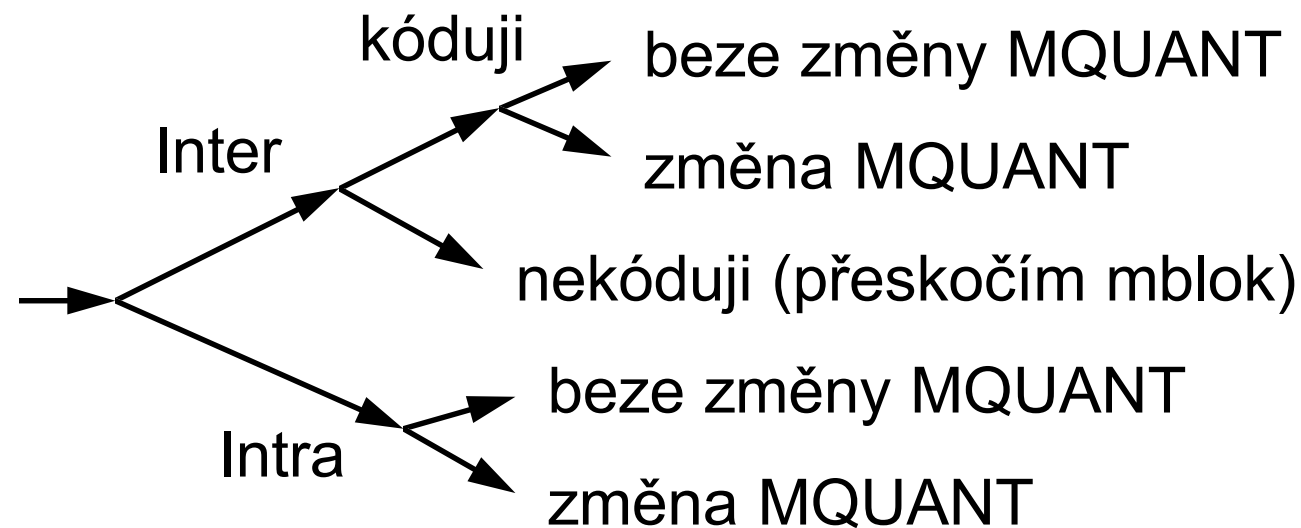
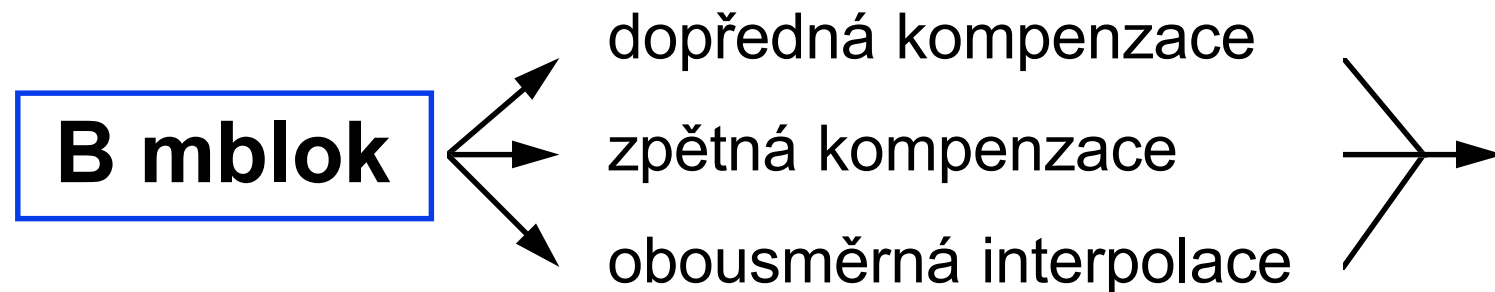
- ◆ **v hlavičce MPEG sekvence** lze definovat dvě explicitní kvantovací tabulky
 - implicitní tabulky - **intra**: JPEG pro kanál Y a 50%,
inter: samé 16
- ◆ **adaptace** na úrovni makrobloku
 - faktor **MQUANT** (1 až 31) - jen pro AC koeficienty

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

Kódování I- a P-makrobloků



Kódování B-makrobloků



Příklad distribuce makrobloků

typ snímku	typ makrobloku				
	I	P	B	zero MV	skip
I	3.300				
P	897	8.587		5.128	568
B	60	7.356	22.845		429

Příklad distribuce DCT bloků

typ snímku	typ makrobloku			
	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%)	

Příklad zatížení MPEG dekodéru

funkce	zátěž (%)
Dekodér hlaviček MPEG souboru	0.44
Huffmanův dekodér a dekvantovač	19.00
Inverzní DCT (8×8)	22.10
Kompenzace pohybu	38.64
Barevná transformace a zobrazení	19.82

Struktura MPEG souboru

1 “sequence”

- globální parametry (rozlišení, frekvence snímků, max. bitový tok, spotřeba paměti, kvantizační tabulky, ..)

2 “group of pictures” (GOP)

- přímý přístup, (nezávislá) kódovací jednotka

3 “picture”

- základní jednotka sekvence (referenční číslo, typ snímku, synchronizační data)

Struktura MPEG souboru

4 “slice”

– zotavení z chyb, adaptace kvantovačů (MQQUANT)

5 “macroblock”

– jednotka kompenzace pohybu (typ, pohybové vektory, maska kódovaných bloků, MQQUANT)

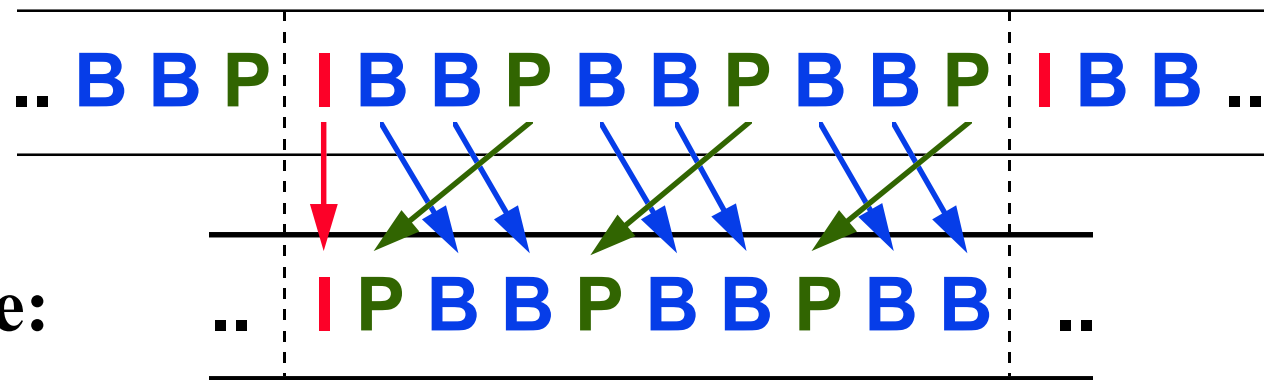
6 “block”

– jednotka DCT transformace (Huffmanův kód po “cik-cak” průchodu zkvantovanými DCT koeficienty)

Skupina snímků (GOP)

- ◆ musí obsahovat **alespoň jeden I-snímek**
- ◆ **začíná** B-snímekem nebo I-snímekem
 - jestliže úvodní B-snímky nezávisí na předchozím referenčním snímku, jde o tzv. **uzavřenou skupinu**
- ◆ **končí** P-snímekem nebo I-snímekem

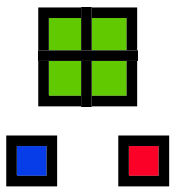
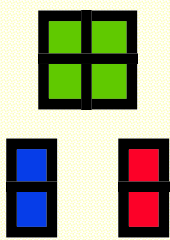
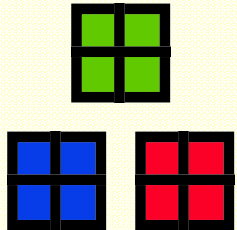
Příklad:



Standard MPEG - 2

- ➔ **bez omezení bitového toku (10 Mbit/s)**
- ➔ **různé stupně kvality a kompresního poměru**
 - profily (“profiles”) a úrovně (“levels”), různé typy barevného vzorkování (4:2:0 až 4:4:4)
- ➔ **podpora prokládání pro digitální TV vysílání**
 - jeden snímek se skládá ze dvou polí (horní a dolní)
- ➔ **rozšiřující vrstvy (“scalable syntax”)**
 - různé dekodéry si mohou vybírat odpovídající kvalitu (bitový tok) - např. podle výstup. rozlišení

MPEG-2: vzorkování barev

snímková frekvence (Hz)		30 (NTSC)	25 (PAL)
→ 4:2:0			
složka Y		720 × 480	720 × 576
složky C _b , C _r		360 × 240	360 × 288
→ 4:2:2			(CCIR 601)
složka Y		720 × 480	720 × 576
složky C _b , C _r		360 × 480	360 × 576
→ 4:4:4			
složka Y		720 × 480	720 × 576
složky C _b , C _r		720 × 480	720 × 576

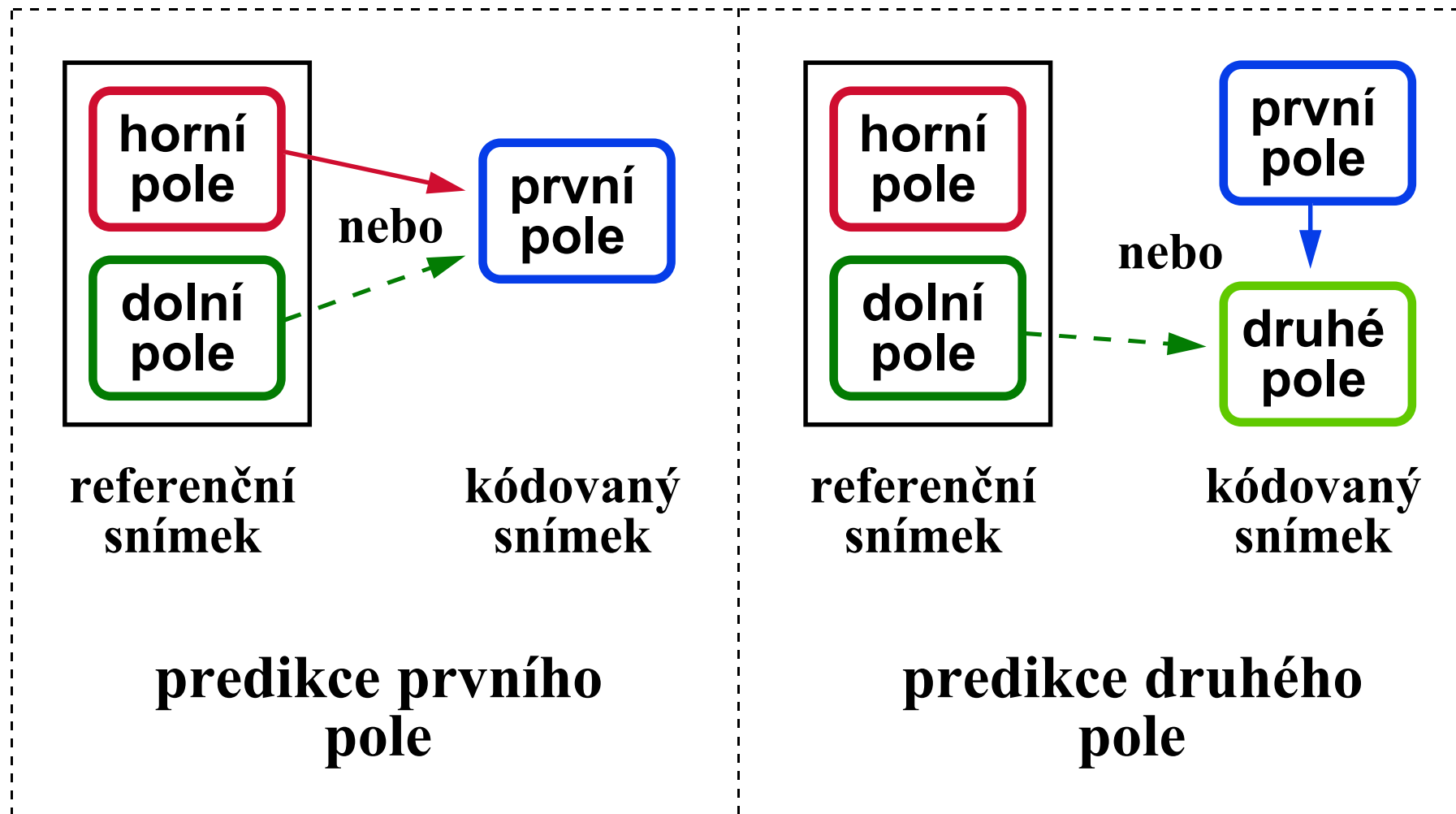
MPEG-2: rozšiřující vrstvy

- ➔ **rozdělení dat** (“data partitioning”)
 - např. rozdělení na kritickou část (hlavičky) a zbytek
- ➔ **rozdělení podle kvality** (“SNR scalability”)
 - stejné rozlišení, různá kvalita (DCT zjemnění)
- ➔ **prostorová hierarchie** (“spatial scalability”)
 - různá rozlišení (viz hierarchický režim JPEG)
- ➔ **časová hierarchie** (“temporal scalability”)
 - stejné rozlišení i kvalita, různé snímkové frekvence

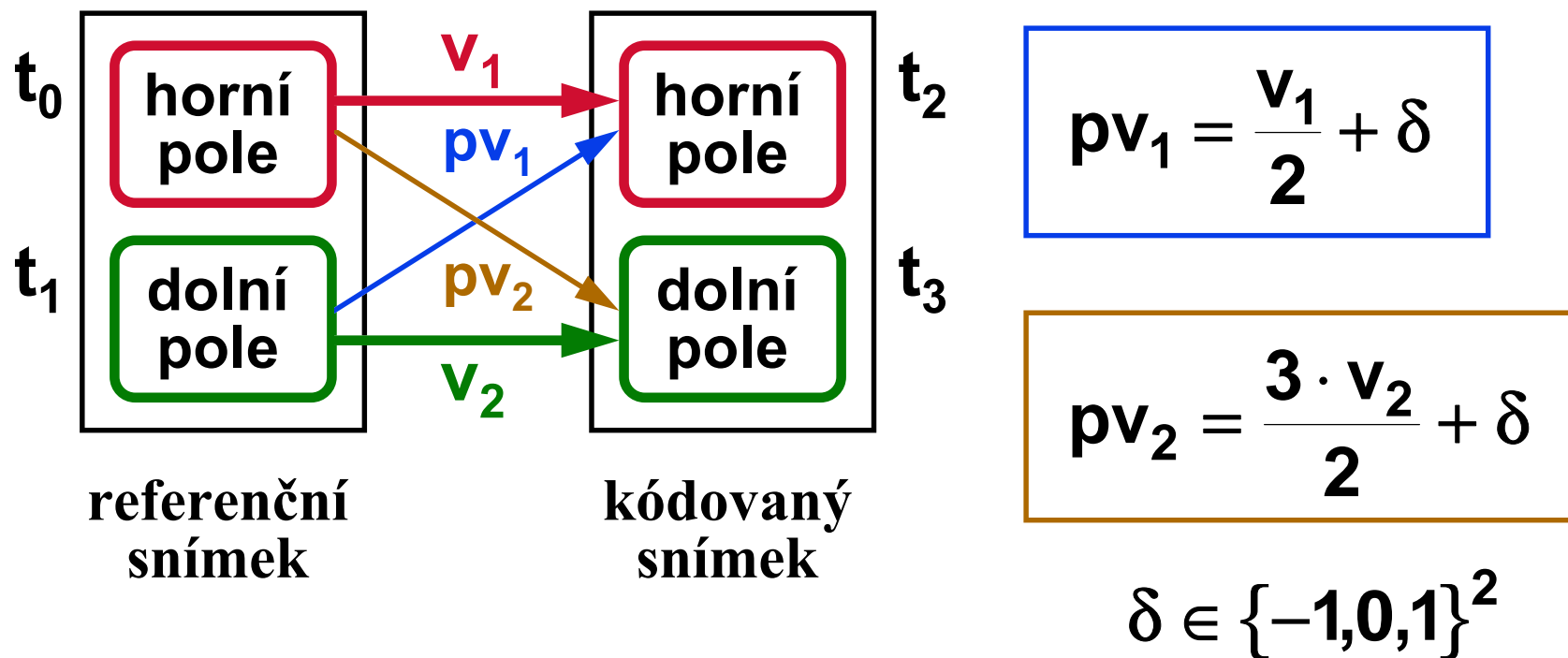
Prokládání (“interlacing”)

- **jasový signál** může být prokládán nebo přenášen postupně ve dvou polích (horní a dolní)
 - na úrovni makrobloků
- **barevné signály** v makroblocích jsou vždy prokládané
- ➔ **další režimy predikce**
 - predikce mezi jednotlivými poli (**I-I**, **I-P**, **P-P**, **B-B**)
- ➔ **další režim kompenzace pohybu** (“dual-prime prediction”)

Predikce mezi poli



“dual-prime” kompenzace



Přenáší se pouze v_1, v_2, δ

MPEG-2: další drobnosti

- ➔ všechny **pohybové vektory** mají půl-pixelovou přesnost
- ➔ každý snímek může mít jiné **kvantovací tabulky**
 - dvě sady tabulek - pro kanál **Y** a pro barevné kanály C_b, C_r
- ➔ **kompence pohybu “16×8”** (při prokládání)
 - pohybový vektor pro každé pole zvlášť (2 nebo 4 vektory / makroblok)

Profily a úrovně v MPEG-2

- čtyři různé **profily** (Simple, Main, Main+, Next)
 - kvalita, zpětná kompatibilita, rozšiřující vlastnosti
 - **Simple**: omezený Main (bez B-predikce)
 - **Main**: standardní profil (bez rozšiřujících vrstev)
 - **Main+**: první profil s rozšiřujícími vrstvami
 - **Next**: všechna hierarchická rozšíření (TV vysílání)
- čtyři různé **úrovně** (Low, Main, High-1440, High)
 - rozlišení snímků (až do 1920×1152 / 60 Hz)
 - **Low**: až SIF, **Main**: až CCIR 601, **High-1440**: až 1440×1152 (HDTV), **High**: až 1920×1152 (HDTV)

Základní literatura

- **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**

Konec

Další informace:

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