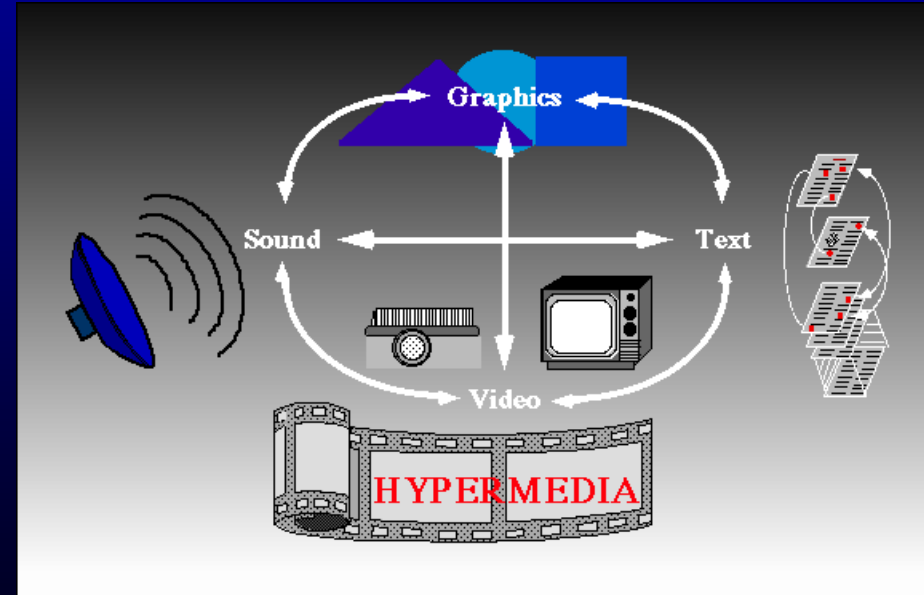
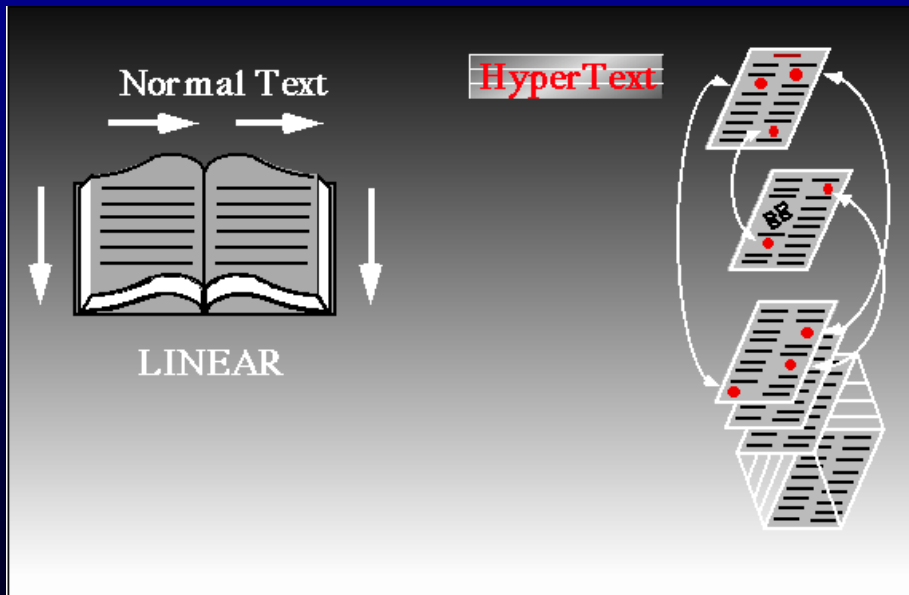


Multimediální technologie v sítích

eva@fi.muni.cz
hopet@ics.muni.cz

Přehled

- multimédia: přináší další způsoby interakce s počítačem – zvuk a video
- hypermédia – nelineární práce s médii – Ted Nelson (1965)



Interakce multimédií a sítí

- kódování do vhodného formátu a jeho dekódování u příjemce
- ošetření chyb přenosu
- parametry Quality of Service (QoS)
 - šířka pásma, zpoždění (latence, RTT), jitter
- synchronizace mezi médii
- klient – server vs. peer-to-peer (P2P) aplikace
- multicast

Aplikace

- WWW
- multimedialní produkty, databáze, hry
- videokonference
- streamování (aka video-on-demand, webcasting či netcasting), interaktivní TV
- virtuální realita
- zpracování videa, distribuované střížny

- ukázky

- WebMuseum, Paris

- <http://www.ibiblio.org/wm/>

- BBC, CNN

- <http://www.bbc.co.uk/>, <http://www.cnn.com>

- Index of Multimedia Information Sources

- <http://viswiz.gmd.de/MultimediaInfo/>

- archívy konferencí Internet2, Canarie

- <http://www.internet2.edu/>, <http://www.canarie.ca/>

- ResearchChannel

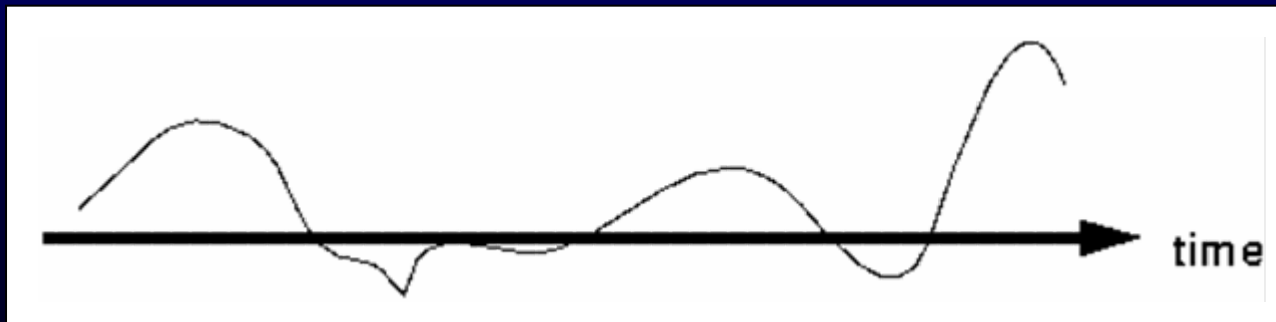
- <http://www.researchchannel.org/>

Audio

- nahrávání/přehrávání
- syntéza zvuku (příp. řeči)
- efekty nad zvukem
- posílání zvuku v digitální formě sítí

Audio – digitalizace

- zvuk
 - spojité tlakové vlny putující vzduchem
 - běžné fyzikální jevy vln: odraz, ohyb, difrakce
- mikrofon – převod zvuku na napěťový analogový signál



- PCM (Pulse Code Modulation)

- *vzorkování* – rozdělení časové osy

- většina lidí nedokáže rozlišit zvuky nad 20 kHz
 - vzorkovací frekvence musí být nejméně dvojnásobek frekvence zaznamenávané (Nyquistův teorém)
 - řeč: 11.025 kHz (telefon: 8 kHz)
zvuk nižší kvality: 22.05 kHz
CD: 44.1 kHz (frekvence >22.01 kHz se odfiltrují před zpracováním)
DAT: 48 kHz
 - vysoce kvalitní záznam: 96 kHz

– *kvantování (quantization)* – rozdělení svislé osy (intenzita – tlak/napětí)

- 1 bit odpovídá přibližně 6 dB (2x hlučnejší)
- CBR vs. VBR
- telefon: 8 bitů (256 úrovní, mono)
CD: 16 bitů (65536 úrovní, stereo)
- je-li V_{max} maximální pozitivní a negativní amplituda signálu a n je počet použitých bitů, pak kvantizační interval q je definován:

$$q = \frac{2V_{max}}{2^n}$$

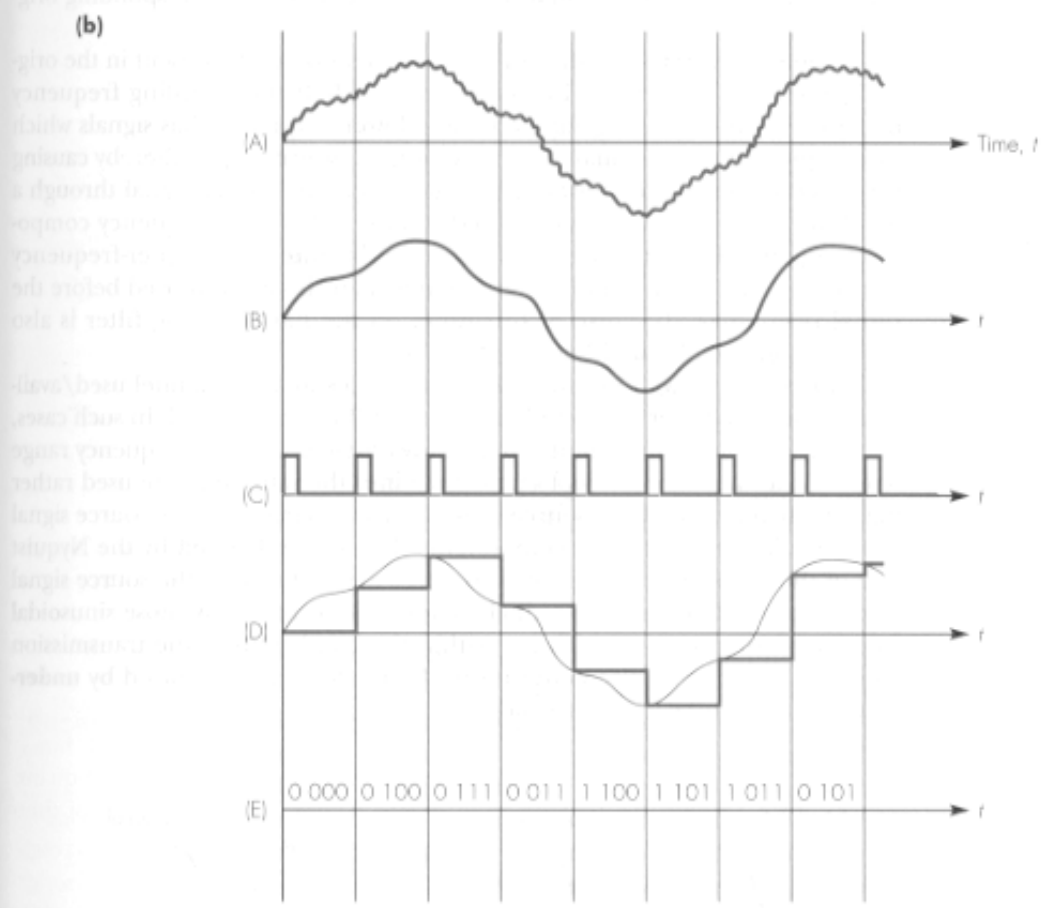
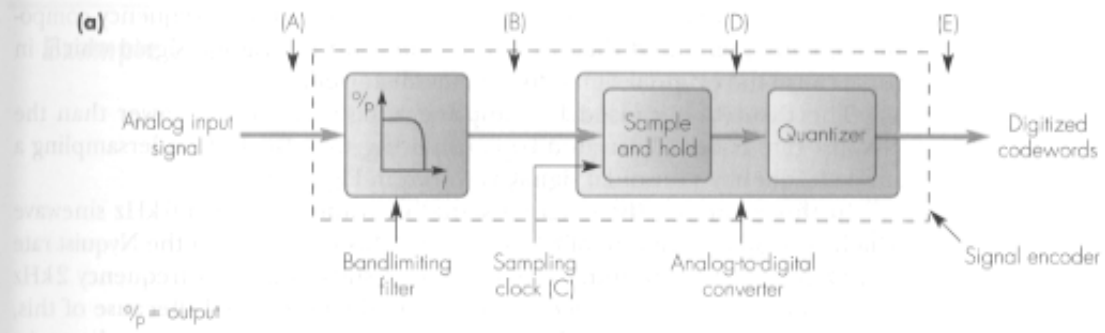
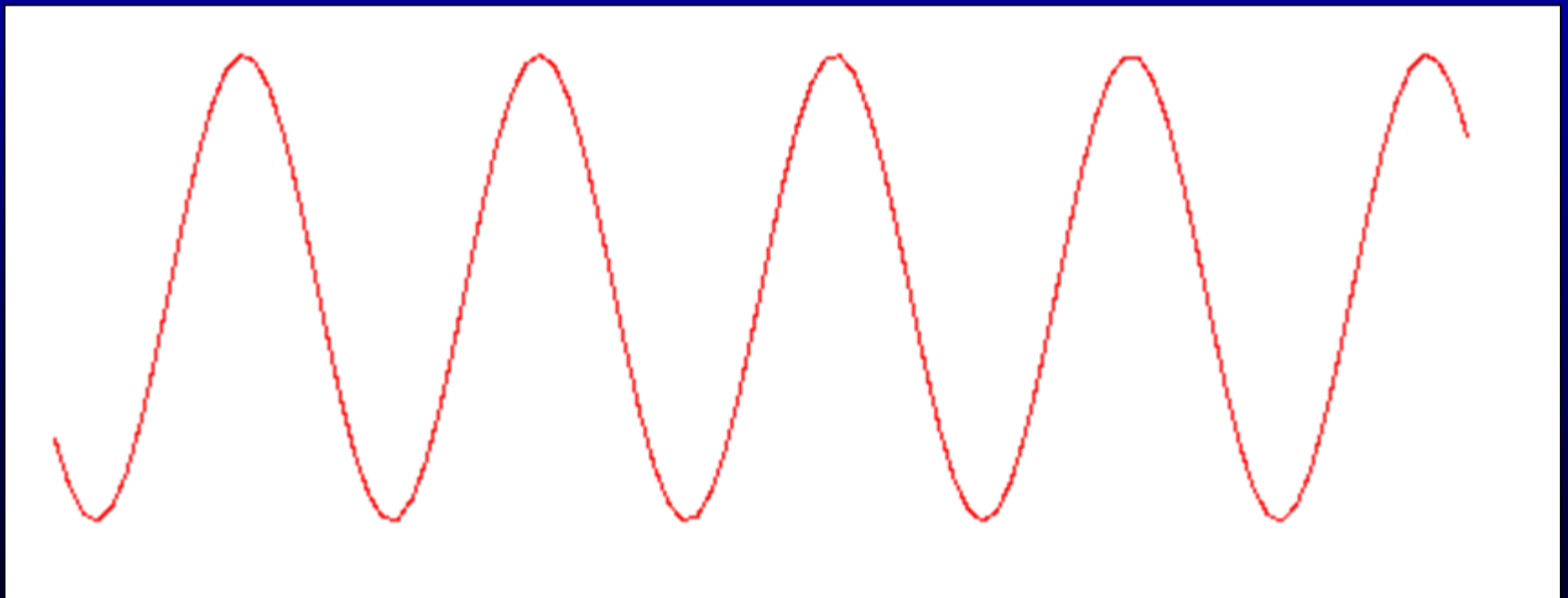
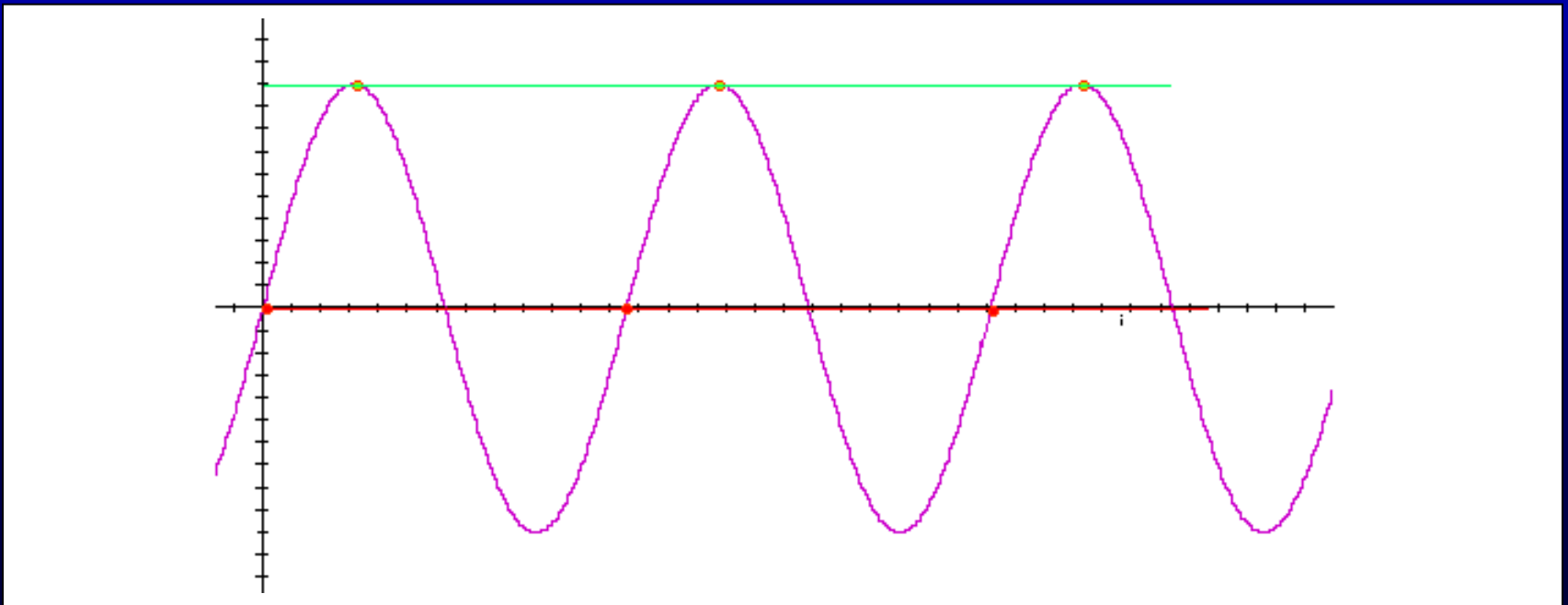


Figure 2.2 Signal encoder design: (a) circuit components; (b) associated waveform set.

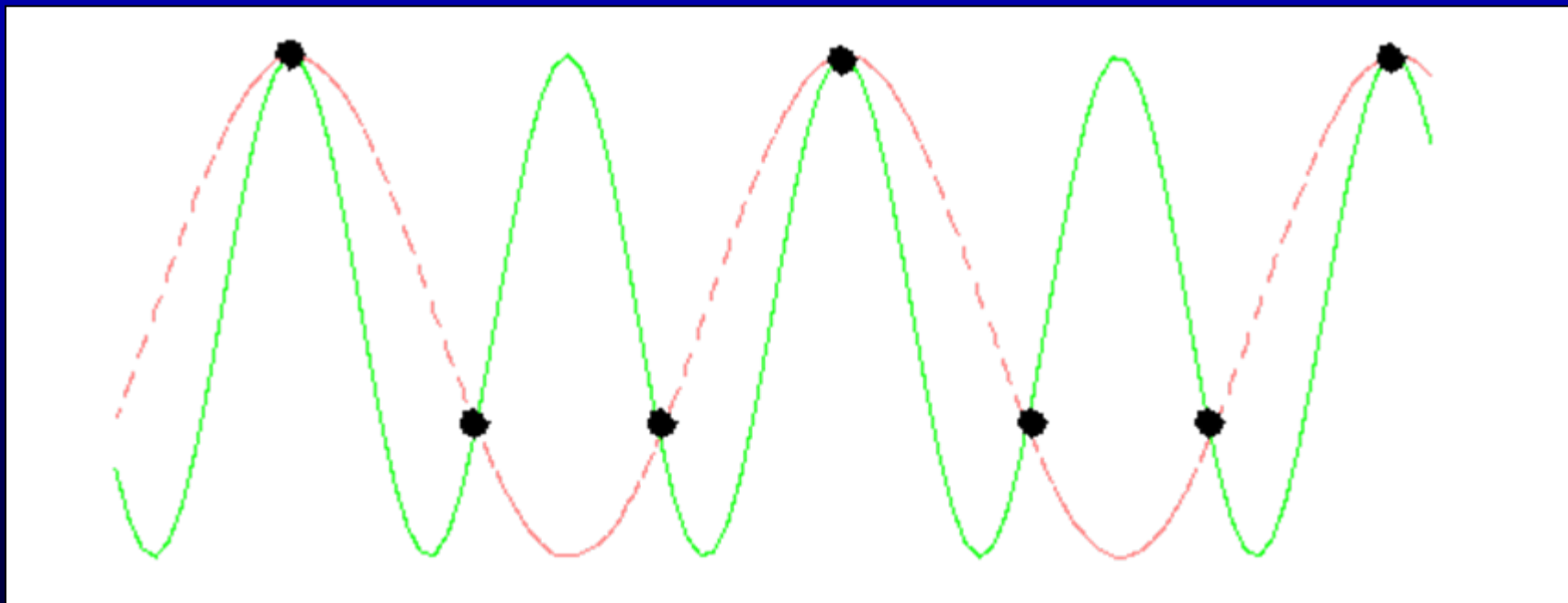
- odbočka k Nyquistovu teorému
– původní signál



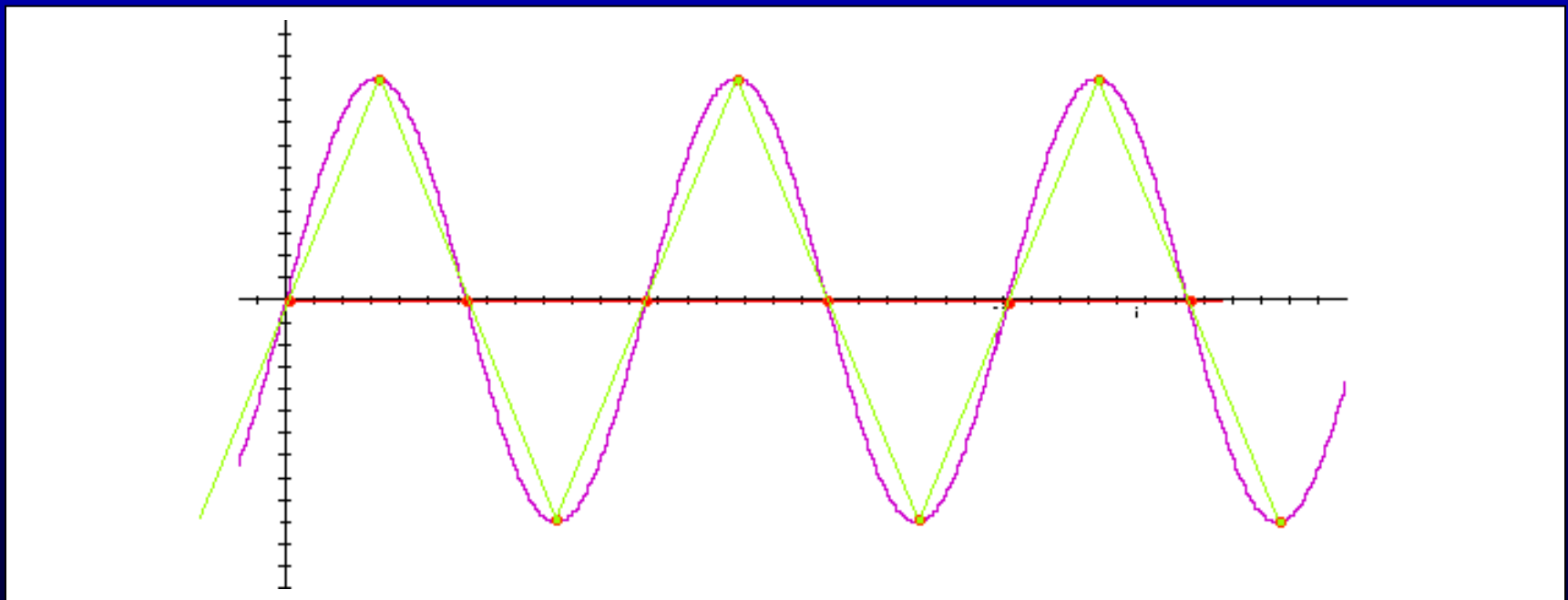
– vzorkování na stejné frekvenci



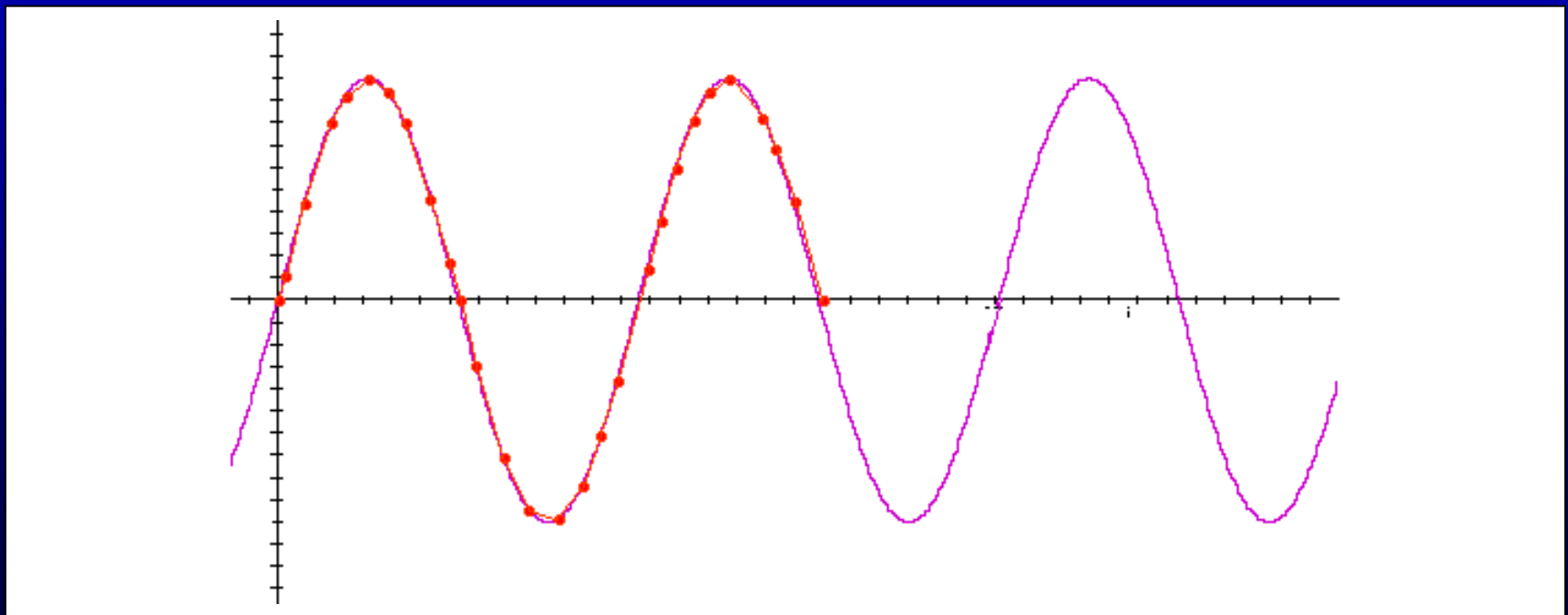
– vzorkování na 1.5-násobné frekvenci



– vzorkování na 2-násobné frekvenci



– vzorkování na ještě vyšší frekvenci



- podvzorkování (undersampling)
 - podvzorkovaná frekvence se „schová“ za jinou frekvenci

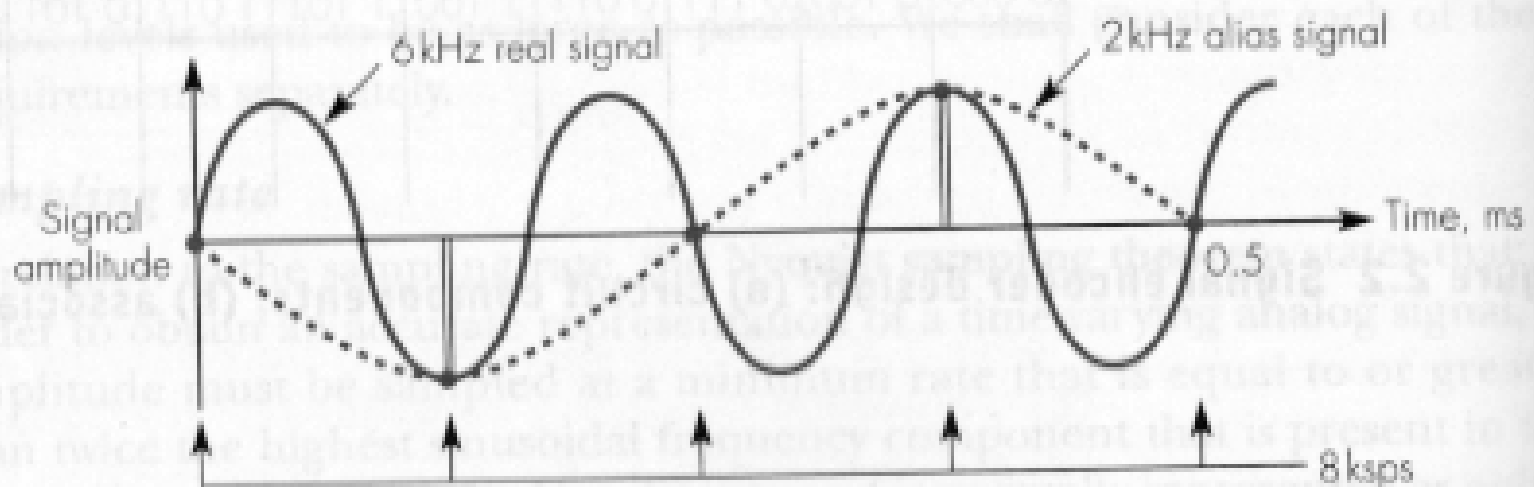


Figure 2.3 Alias signal generation due to undersampling.

- chyby v kvantování
 - chyba v záznamu může být $\pm q/2$
 - několik úrovní intenzity se reprezentuje jednou hodnotou => kvantovací šum
 - vylepšit lze
 - ditheringem
 - nelineárním kvantováním
 - ucho je nelineární člen soustavy zvuk – mozek ;-)
(zhruba logaritmický)
 - zhuštění úrovní v oblasti, kde je to třeba; u zvuku je třeba mít jemněji rozlišenou oblast s nízkými kmitočty, protože ucho je v této oblasti citlivější na šum

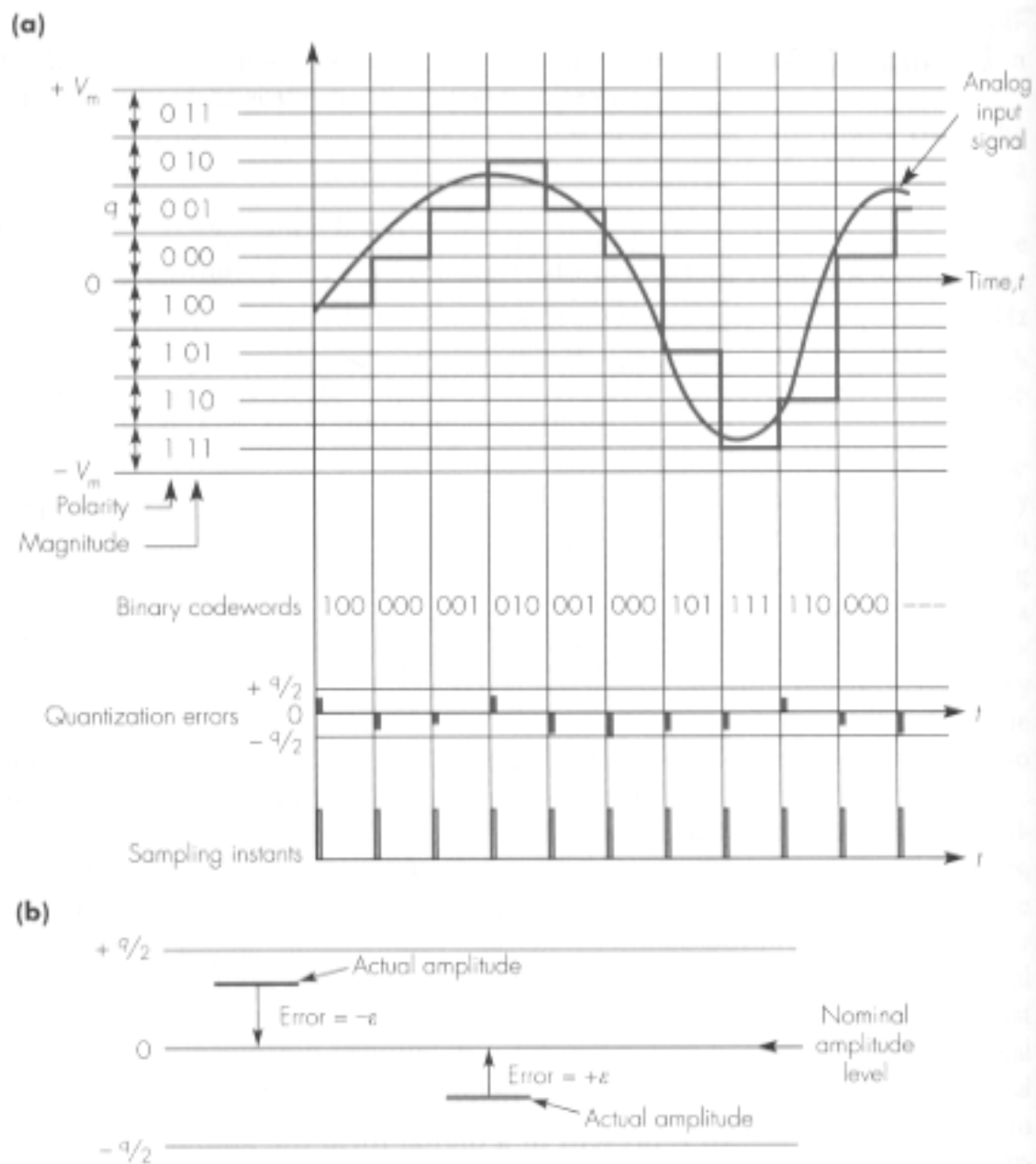


Figure 2.4 Quantization procedure: (a) source of errors; (b) noise polarity.

- šum
 - signal / noise ratio

$$SNR = 10 \log \frac{V_{signal}^2}{V_{noise}^2} = 20 \log \frac{V_{signal}}{V_{noise}}$$

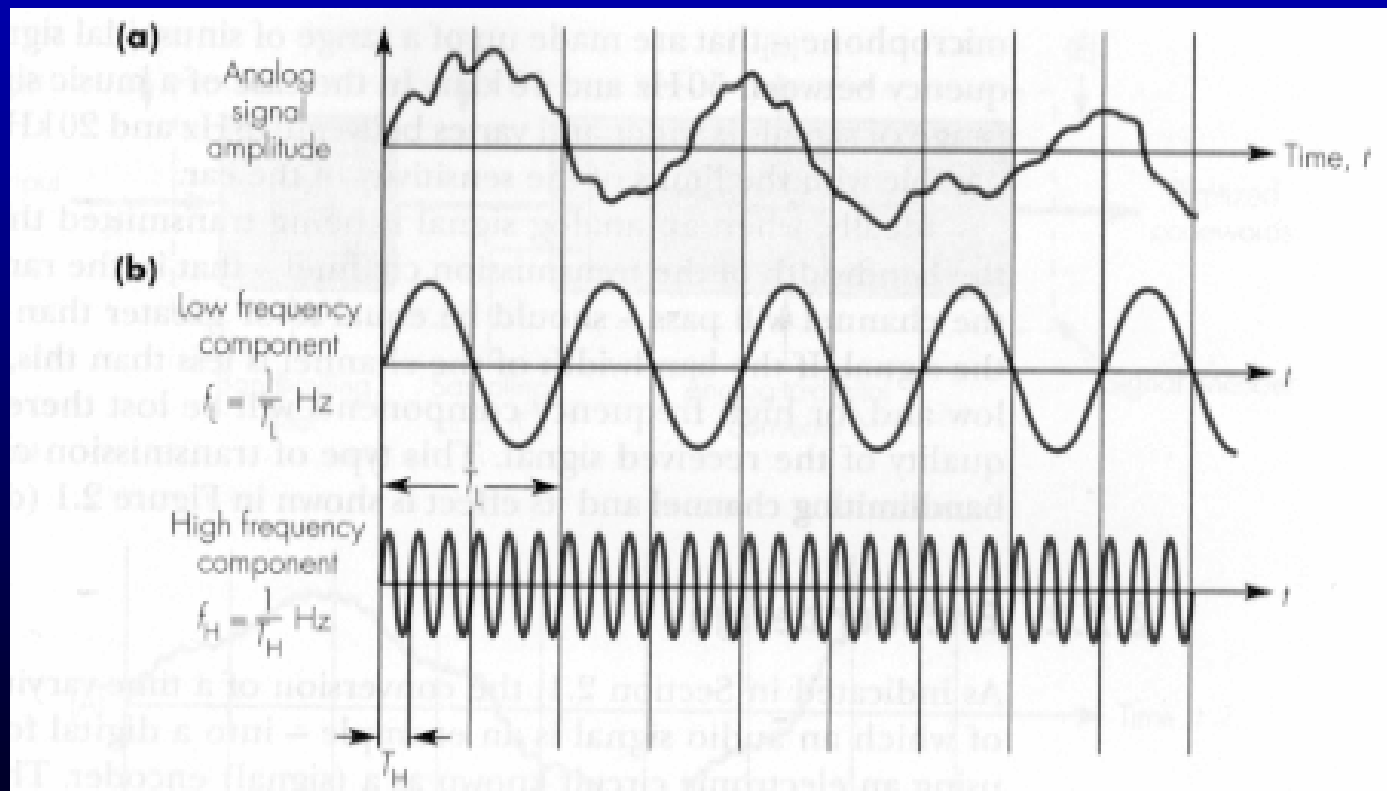
- 8 bitů – SNR 50 dB
- 16 bitů – SNR 98 dB

- příklady

Quality	Sample Rate (KHz)	Bits per Sample	Mono/ Stereo	Data Rate (Uncompressed)	Frequency Band
Telephone	8	8	Mono	8 KBytes/sec	200-3,400 Hz
AM Radio	11.025	8	Mono	11.0 KBytes/sec	
FM Radio	22.050	16	Stereo	88.2 KBytes/sec	
CD	44.1	16	Stereo	176.4 KBytes/sec	20-20,000 Hz
DAT	48	16	Stereo	192.0 KBytes/sec	20-20,000 Hz

Zpracování signálu – FT

- Fourierova analýza



- Fourierova transformace

$$x(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(2\pi k f_0 t) + b_k \sin(2\pi k f_0 t)$$

- na počítači nemůžeme dělat nekonečné součty => diskrétní FT
- můžeme předstírat, že $x(t)$ je periodické se stejnou délkou, jako je délka záznamu
- základní frekvence = 1/délka záznamu
- window function – minimalizace nespojitosti fází na rozhraních segmentů

Zpracování signálu – diskrétní kosinová transformace (DCT)

- DCT používá reálná čísla
- obrázek o $N \times N$ pixelech: $[p_{uv}, 0 \leq u < N, 0 \leq v < N]$
- DCT je pole koeficientů: $[DCT_{uv}, 0 \leq u < N, 0 \leq v < N]$
kde:

$$DCT_{uv} = \frac{1}{\sqrt{2N}} C_u C_v \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p_{xy} \cos \left[\frac{(2x+1)u\pi}{2N} \right] \cos \left[\frac{(2y+1)v\pi}{2N} \right]$$

kde:

$$C_u C_v = \frac{1}{\sqrt{2}} \text{ for } u, v = 0$$

$$C_u C_v = 1 \text{ otherwise}$$

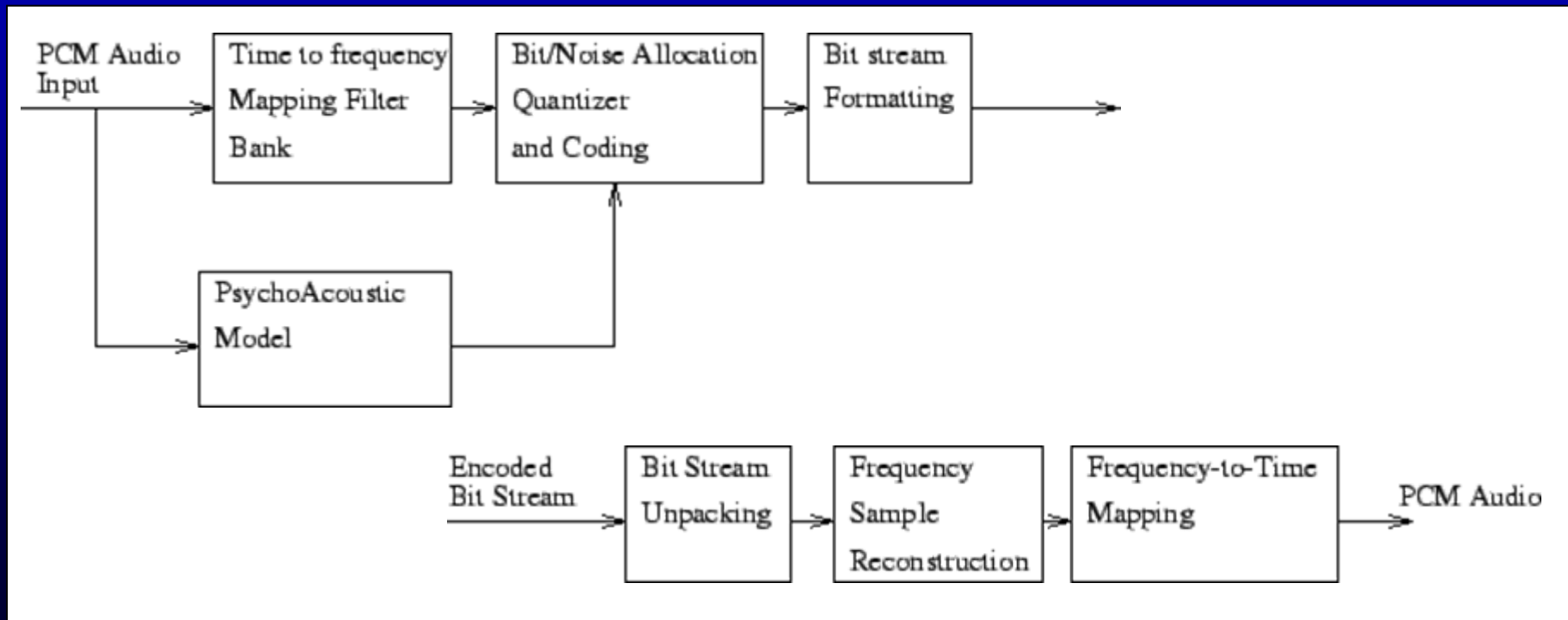
Poznámky k FT a DCT

- složitost:
 - diskrétní FT (DFT) ... $O(n^2)$
 - rychlá FT (FFT) ... $O(n \cdot \log_2 n)$
 - DCT ... $O(n^4)$ resp. $O(n^2)$
- odkazy:
 - <http://www.dataq.com/applicat/articles/an11.htm>
 - <http://www.chipcenter.com/eexpert/bmasta/bmasta001.html>
 - http://bellatrix.isr.umd.edu/TechReports/ISR/1996/PhD_96-9/PhD_96-9.phtml

Formáty audia

- PCM
 - telefon: $1 \cdot 8000 \cdot 8 = 64$ kbps
 - CD kvalita: $2 \cdot 44100 \cdot 16 = 1.41$ Mbps
- DPCM (Differential Pulse Code Modulation)
 - ztrátová komprese, používá 4 bity na záznam rozdílů nezávisle na původním kvantování
- ADPCM (Adaptive Differential Pulse Code Modulation)
 - snaží se o prediktivní analýzu potřebného počtu bitů v závislosti na složitosti signálu
- μ -law (A-law v Evropě)
 - podobné ADPCM, použití v telefonech (kvantování pak odpovídá spíše 13 bitům než 8)

- LPC (Linear Predictive Coding), (Code Excited Linear Predictor) a GSM
 - optimalizace na hlas
- Postup kódování z PCM do MP3 a naopak

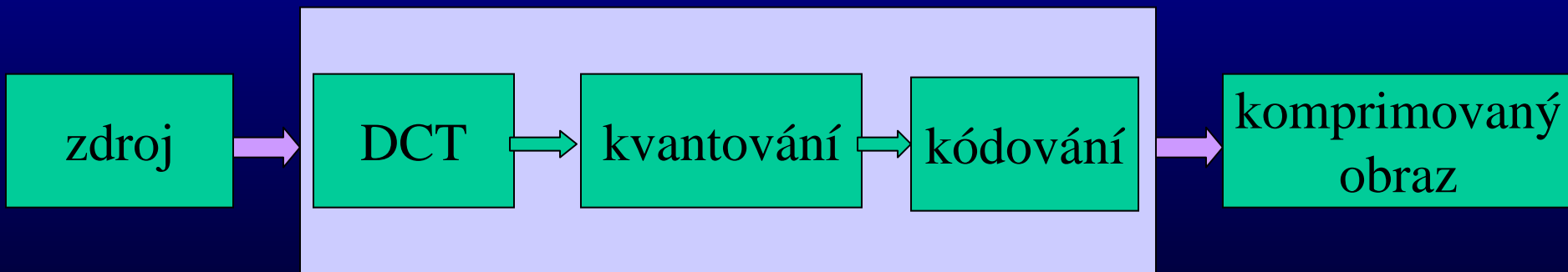


MP3

- frekvenční pásma – subbands
trik MP3 – výběr správných pásem
psychoakustický model
- DCT -> kvantování -> Huffmanovo
kódování
- Layer I – 384 kbps – 4:1
Layer II – 192 kbps – 8:1
Layer III – 128 kbps – 12:1

Video – kódování, komprese

- komprese obrázků
 - Lempel-Ziv, GIF (8 bitový index barev, slovníková LZ komprese; 10:1 až 2:1)
 - JPEG



- JPEG
 - DCT – bezeztrátová
 - DC, AC koeficienty
 - kvantování – tabulky kvantovacích koeficientů
 - kódování
 - RLE pro nulové koeficienty
 - Huffman pro nenulové
 - delta kódování pro DC koeficienty
 - barva
 - možnost redukce 24 bitů na 8 bitů

Jak funguje MPEG

- intraframe vs. interframe komprese
- 3 typy snímků (frames)
 - I frames – intrapicture, referenční snímek
 - P frames – predicted, rozdíl oproti předchozímu snímku
 - B frames – bidirectional predicted, interpolace mezi předcházejícím a následujícím

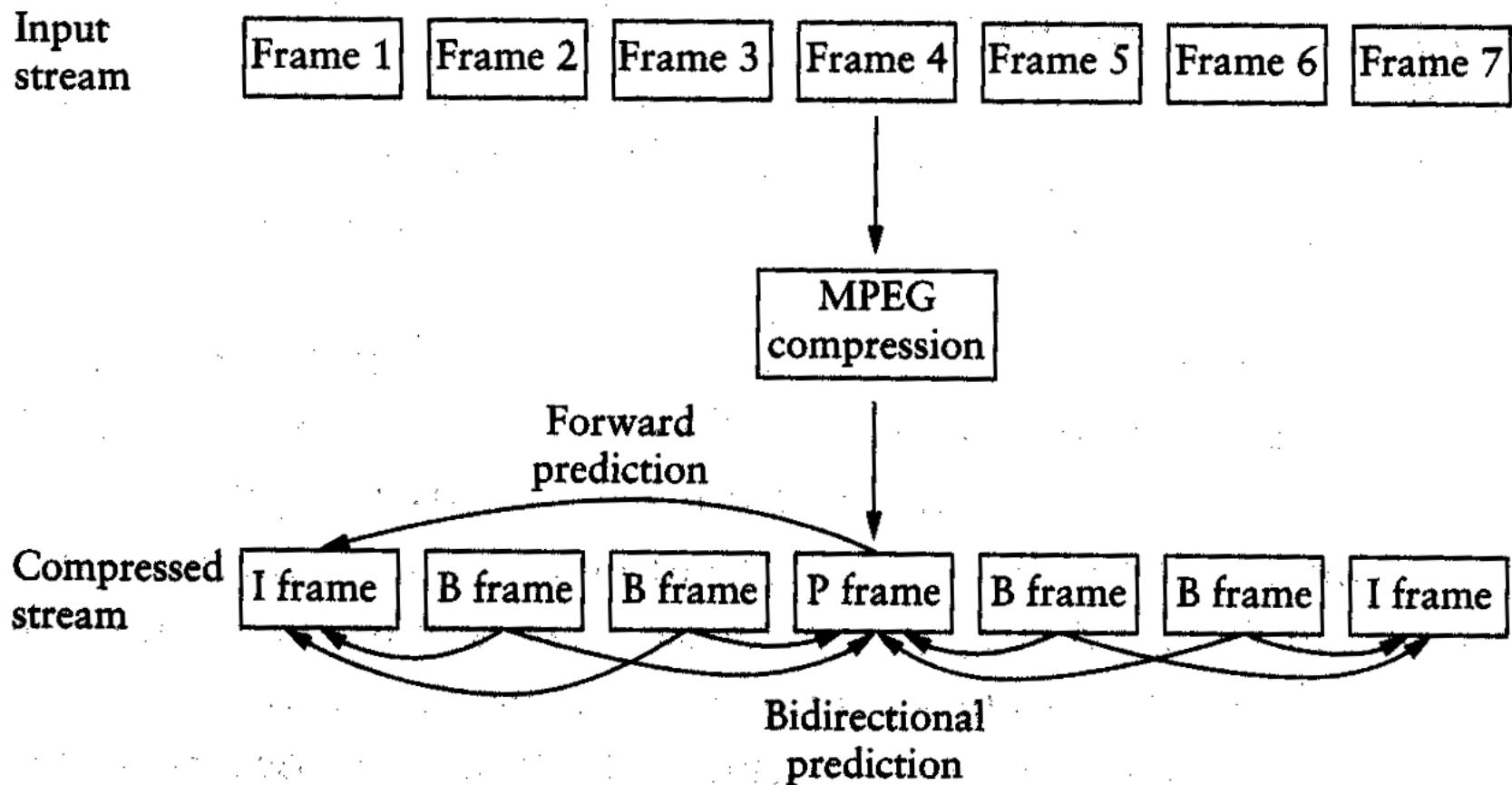


Figure 7.13 Sequence of I, P, and B frames generated by MPEG.

MPEG - komentář

- pro přenos po síti je důležité:
 - snímek P nemá smysl bez snímku I
 - snímek B nemá smysl bez snímků I a P
 - v případě snímků B se snímky neposílají v sekvenčním pořadí
- MPEG nedefinuje poměr I/P/B snímků
 - pouze I snímky ~ MJPEG

- snímky I – makrobloky 16x16
- vyšší citlivost oka na jasovou složku (Y)

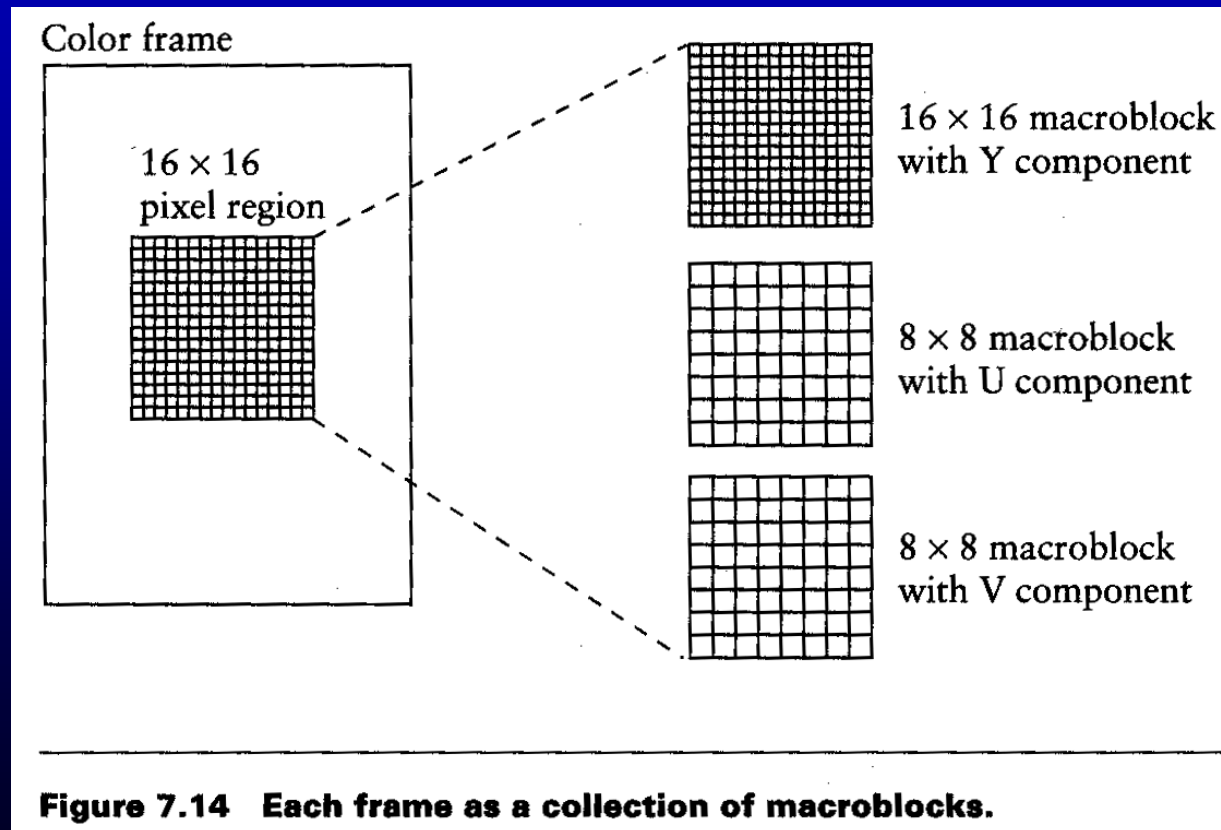
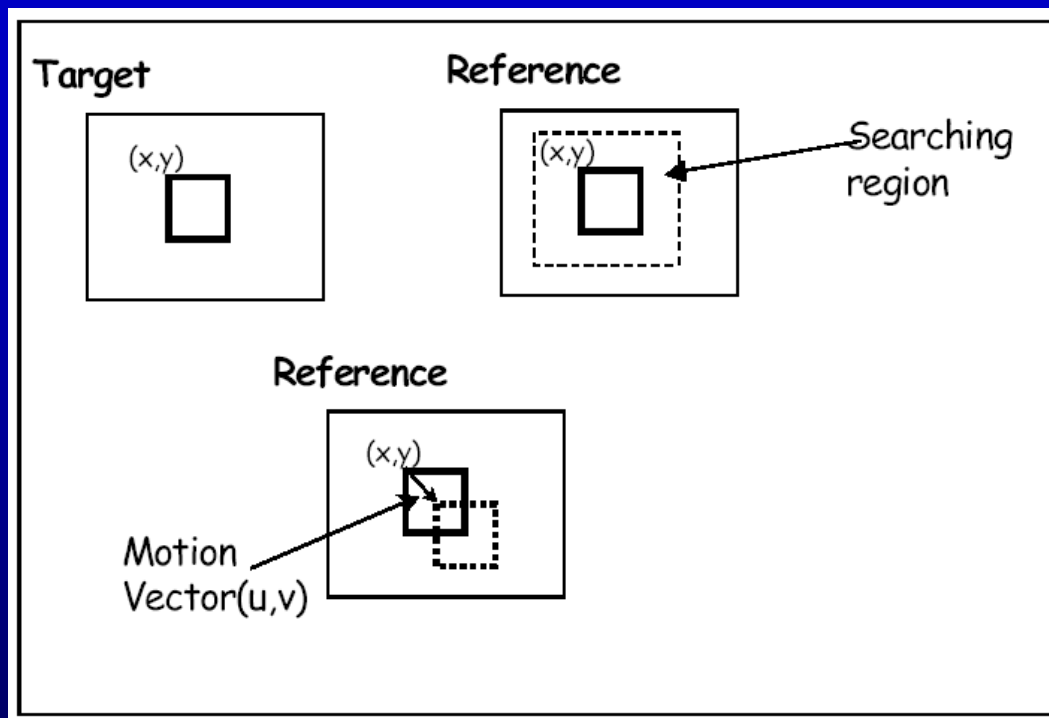


Figure 7.14 Each frame as a collection of macroblocks.

- snímky B a P jsou zpracovány také po makroblocích
- snímky B
 - výjimečně může být makroblok stejný jako v případě I snímku (v případě velké změny)
 - čtveřice
 - souřadnice makrobloku ve snímku
 - motion vector vzhledem k předchozímu referenčnímu snímku
 - motion vector vzhledem k následujícímu referenčnímu snímku
 - δ pro každý pixel – změna pixelu vzhledem k referenčním snímkům (komprese DCT)

$$F_c(x,y) = (F_p(x+x_p,y+y_p) + F_f(x+x_f,y+y_f))/2 + \delta(x,y)$$

- motion vector & motion estimation



- MPEG v reálném čase? S podporou HW ano, bez ní to trochu bolí...

MPEG formáty

- MPEG-1
 - rok 1993
 - 1 – 1.5 Mbps
 - náhodný přístup – 0.5 s
- MPEG-2
 - rok 1995
 - digital TV, DV, DVD
 - až 100 Mbps
 - od 4 Mbps výrazně lepší kvalita než MPEG-1

– MPEG-2 main profile

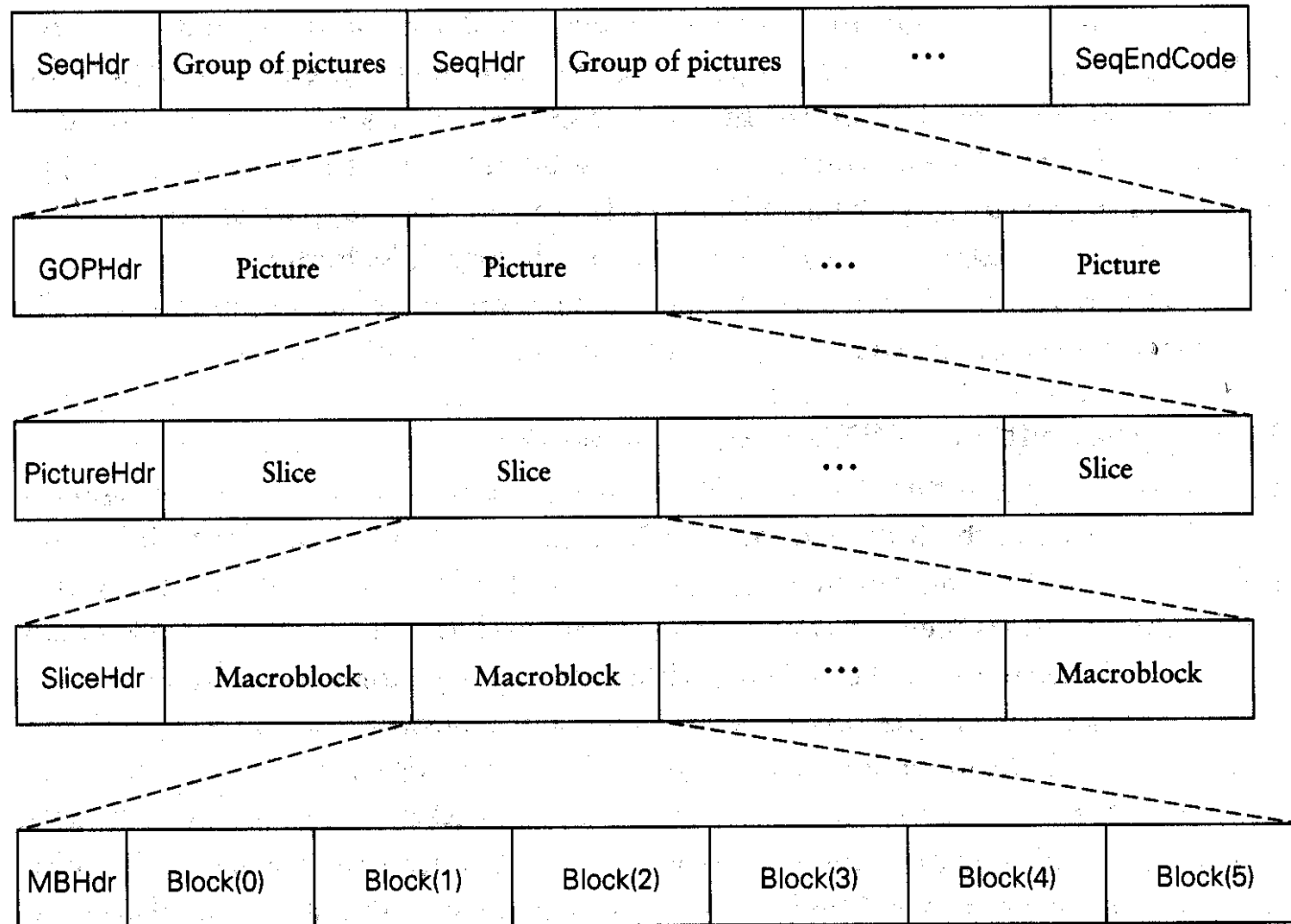


Figure 7.15 Format of an MPEG-compressed video stream.

- CBR vs. VBR (změny BR pro jednotlivé GOP – např. změnou kvantování)
- co je optimální BR v případě multicastu??
=> odpověď: vrstvení dat
- problémy s retransmisí v TCP – nepřijatelně velká latence v případě výpadku
- přenos po UDP – má smysl pečlivě volit rozhraní paketů, aby ztráta paketu „moc nebolela“ => Application Level Framing
- problémy se ztrátami I snímků => diffserv
- problém s latencí:
IBBBBBPBBBBBI
člověk vnímá hranici 100 ms, ale 5x 67ms (15 fps)
= 335 ms!

- kombinace s audiem – interleaving
- optimalizace v případě ucpání sítě
 - vynechání B snímků a přenos audia místo nich, protože člověk je mnohem citlivější na nekvalitní audio než na nekvalitní video

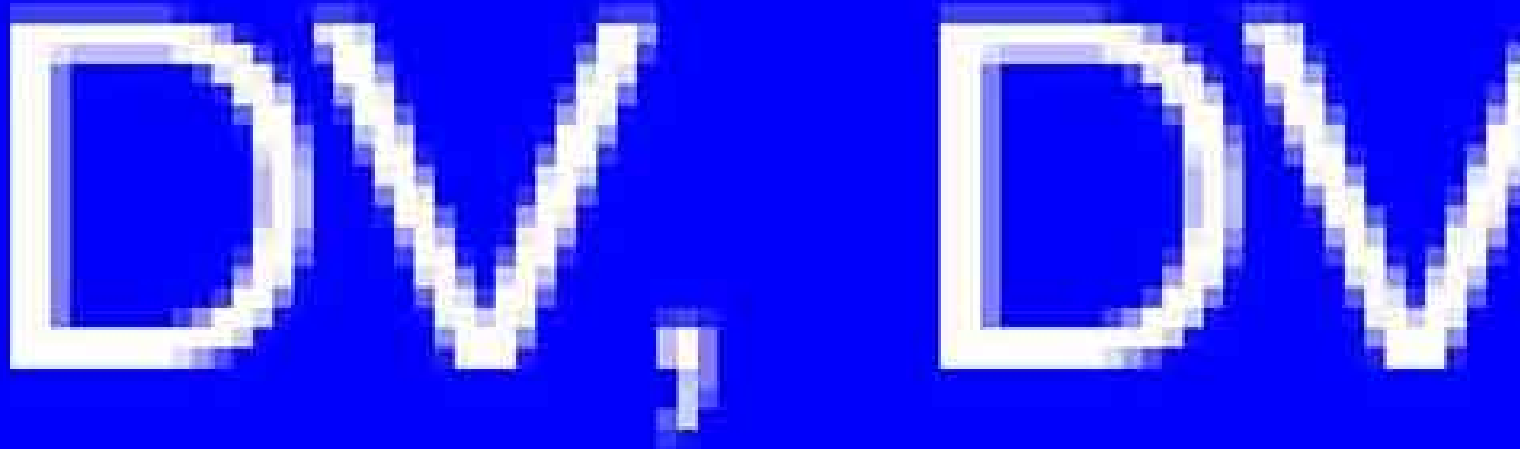
- MPEG-4

- určeno pro malou šířku pásma 10 kbps až 1 Mbps
- důležitý bezchybný přenos
- organizace do objektů, k nimž lze přistupovat samostatně
- hierarchie úrovní detailů (posílány současně)
 - podpora klientů s různou kapacitou zpracování
- MS (Streaming) Media, RealMedia, DivX

DV

- MPEG-2 (like)
 - pečlivě licenčně ošetřeno (no royalties)
- 25 Mbps (DVCPRO-50 50 Mbps)
- kompresní poměr 5:1
- intraframe a interfield komprese
 - frames (snímky) vs. fields (půlsnímky)

DV artefakty - fathering



DV, DV

DV, DVCAM, DVCPRO

DV artefakty - quilting



DV artefakty - motion blocking

