

MULTIMEDIÁLNÍ A HYPERMEDIÁLNÍ SYSTÉMY

10)
Komprese videa

Petr Lobaz, 23.4.2013

NEKOMPRIMOVANÉ VIDEO

DATOVÝ TOK SD

1 snímek 625/50	720 × 576 × 3 =	1,18 MB
1 snímek 525/60	720 × 480 × 3 =	0,99 MB
1 vteřina 625/50	25 × 1,18 =	29,66 MB
1 vteřina 525/60	30 × 0,99 =	29,66 MB
1 minuta	1500 × 1,18 =	1,73 GB
1 hodina	90000 × 1,18 =	104,28 GB

⇒ datový tok 237 Mbps (HD video 1920p@25: 1,24 Gbps)

- při podvzorkování 4:2:0 datový tok 119 Mbps
- datový tok CD 1,5 Mbps, DVD 11 Mbps, Blu-ray Disc 40 Mbps

MHS – Komprese videa

2/45

KOMPRESSE VIDEO

POŽADAVKY

- kontrola toku dat (zaručený kompresní poměr)
- rychlost komprese/dekomprese
- odolnost vůči chybám
- synchronizace se zvukem (řeší systémová vrstva – kontejner)

CODEC

- compressor/decompressor
- typicky tři vrstvy – video, audio, systém
- systémová vrstva častěji označena jako muxer / demuxer

MHS – Komprese videa

3/45

TYPY KOMPRESSE

BEZTRÁTOVÁ

- samostatně minimální použití
- RLE – formáty FLI/FLC, doplňková metoda
- VLC (Huffmanovo kódování apod.) – doplňková metoda
- LZW – ve formátu GIF 89a
- aritmetické kódování – doplňková metoda
- beztrátová DWT – MJPEG2000

MHS – Komprese videa

4/45

TYPY KOMPRESI

ZTRÁTOVÁ

- vector quantization (VQ)
 - rozdělení obrazu na „bloky“ pixelů
 - tvorba tabulky reprezentativních „bloků“
 - místo bloku kódujeme číslo nejpodobnějšího bloku v tabulce
 - potenciálně silná metoda, problémy s efektivní tvorbou tabulky a jejím uchováváním
 - při jednoduché implementaci velmi rychlá (nenáročná) dekomprese
 - Intel Indeo 3.2, Cinepak, Sorenson Video 1
 - dnes se prakticky nevyužívá

MHS – Kompresie videa

5 / 45

TYPY KOMPRESI

ZTRÁTOVÁ

- rozdělení obrazu na bloky + diskrétní kosinová nebo obdobná transformace + kvantizace koeficientů
 - princip obdobný JPEG
 - dnes zdaleka nejvyužívanější metoda
 - MJPEG, DV, H.261–5, MPEGx, ...
- diskrétní waveletová transformace + kvantizace koeficientů
 - princip obdobný JPEG2000
 - Intel Indeo 5, MJPEG 2000
 - Redcode (a jiné RAW formáty profesionálních kamer)
 - Digital Cinema System (distribuce filmů do kin)

MHS – Kompresie videa

6 / 45

TYPY KOMPRESI

ODSTRANĚNÍ REDUNDANCE

- intra coding
 - odstranění prostorové (spatial) redundance
 - typický obraz ve videu je rozmazaný pohybem, je místy ostrý a místy neostrý
- inter coding
 - odstranění časové (temporal) redundance
 - typické po sobě jdoucí snímky jsou téměř stejné

MHS – Kompresie videa

7 / 45

VYUŽITÍ KOMPRESI

- pro jednosměrný přenos (např. televize)
 - odolnost vůči chybám, škálovatelnost
 - MPEG-2, Windows Media Video, H.264
- obousměrný přenos (např. videokonference)
 - realtime komprese/dekomprese
 - H.261, H.263, MPEG-4
- archivace (např. DVD, Blu-ray)
 - dobrý kompresní poměr
 - MPEG-2, H.264, VC-1
- dočasný formát (např. střih)
 - náhodný přístup do souboru, minimální ztráta
 - DV, MJPEG, AVC-Intra, DNxHD (VC-3)

MHS – Kompresie videa

8 / 45

INTRA KÓDOVÁNÍ

- stejné jako komprese statického obrazu
- výhoda: jednoduchoost, náhodný přístup do videa
⇒ využití pro střih, postprodukci
- nevýhoda: velký datový tok
- parametr „ztrátovost“ lze dynamicky snímek po snímku měnit ⇒ datový tok konstantní za jednotku času
- čím je datový tok variabilnější, tím větší musí být vyrovnávací paměti pro záznam/přehrávání
- někdy je zapotřebí velmi stabilní datový tok
 - konstantní velikost každého snímku
 - konstantní velikost části snímku

MHS – Komprese videa

9 / 45

INTRA KÓDOVÁNÍ

ZÁKLADNÍ POSTUP

- převod obrazu do $Y C_B C_R$
- rozdělení obrazu na makrobloky (MB)
 - příklad: 16×16 pixelů
- aplikace zvolené chroma subsampling (viz tvar MB!)
 - příklad: po 4:2:0 vznikne 16×16 hodnot Y ,
 8×8 hodnot C_B , 8×8 hodnot C_R
- rozdělení makrobloku na bloky
 - příklady: dělení na bloky 8×8 hodnot, vzniknou 4 bloky Y ,
1 blok C_B , 1 blok C_R

MHS – Komprese videa

10 / 45

INTRA KÓDOVÁNÍ

- aplikace DCT na jednotlivé bloky
- kvantizace koeficientů
 - pro Y typicky jiná kvantizační tabulka než pro C_B a C_R
 - kvantizační tabulka se typicky mění mezi snímky i v rámci snímku
- převod bloků do 1D proudu dat
- aplikace bezztrátové kompresní metody, typicky VLC (variable length code) – Huffmanovo kódování, aritmetické kódování

MHS – Komprese videa

11 / 45

INTRA KÓDOVÁNÍ

KONSTANTNÍ DATOVÝ TOK

- nejjednodušší řešení: každý snímek má být průměrně dlouhý X bitů
 - je-li delší, pro další snímek se kvantizace „zhorší“
 - je-li kratší, pro další snímek se kvantizace „zlepší“
 - datový tok průměrně konstantní, nelze vyloučit špičky

MHS – Komprese videa

12 / 45

INTRA KÓDOVÁNÍ

KONSTANTNÍ DATOVÝ TOK

- lepší řízení: každý makroblok má být průměrně dlouhý
Y bitů \Rightarrow stejný princip jako výše aplikovaný na MB
 - nevýhoda: typický obraz má složité (členité) i jednoduché (jednobarevné) MB
 - řešení 1: průchod MB v pseudonáhodném pořadí
 - řešení 2: pro každý MB alokováno Y bitů; bity ze složitých MB se ukládají na nevyužité pozice u jednoduchých MB

INTRA KÓDOVÁNÍ

- kombinace řešení 1 a 2: vzniká značně složitá struktura umístění bitů v makroblocích
- řešení 3: dělení obrazu na superbloky (SB), jejich průchod v pseudonáhodném pořadí: $SB_{1,1}, SB_{2,1}, SB_{3,1}, \dots$
rozdělení SB na MB: $SB_{1,1}, SB_{1,2}, \dots$
přeložení MB z N následných SB:
 $SB_{1,1}, SB_{2,1}, \dots, SB_{N,1}, SB_{2,1}, SB_{2,2}, \dots$

INTRA KÓDOVÁNÍ

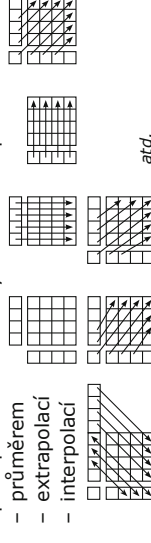
PROKLÁDANÝ OBRAZ

- kódování po půlsnímích nevýhodné, je-li obraz statický
- kódování po snímčích nevýhodné, je-li obraz dynamický
- řešení:
 - pro MB snímku rozhodnout, je-li statický/dynamický
 - statický MB kódovat normálně
 - dynamický MB rozdělit na sudé/liché řádky, kódovat je odděleně, nebo kódovat jejich součet a rozdí

INTRA KÓDOVÁNÍ

PREDIKCE V RÁMCI OBRAZU

- predikce DC koeficientu bloku – lze jej ukládat diferenciálně oproti předchozímu (jako v JPEG)
- predikce vzhledu bloku na základě okolí (H.264):
v okamžiku dekódování bloku jsou již dekódované bloky vlevo a nad aktuálním; lze ukládat jen rozdíly pixelů oproti pixelům odhadnutým z okolních pixelů:



INTER KÓDOVÁNÍ

- následující snímky (F, G) jsou si podobné
⇒ využití rozdílového kódování
- typicky se řeší po makroblocích, stejně jako u intra
⇒ místo makrobloku G_n kódujeme $F_n - G_n$
(rozdílly hodnot pixelů), typicky pomocí DCT
- výhodné u statické scény
- nevýhodné, došlo-li k pohybu (hodnoty $F_n - G_n$ velké)
⇒ místo F_n použijeme jinou část snímku F
- motion compensation - nalezení motion vektoru MV,
který minimalizuje $F_n^{MV} - G_n$



MHS – Kompresce videa

17 / 45

INTER KÓDOVÁNÍ

- dekompresor rekonstruuje snímek F,
po přijetí $F_n^{MV} - G_n$ rekonstruuje snímek G
- bloky ztrátově uložené
– dekompresor rekonstruuje aproximací zF_n , $z(F_n^{MV} - G_n)$
– dekompresor nemůže rekonstruovat zG , protože
nezná F
- řešení: kompresor obsahuje i dekompresor
– ztrátově uloží snímek F, dekomprimuje ho na zF
– snímek G kóduje rozdílově: $zF_n^{MV} - G_n$
- kompresor i dekompresor obsahují paměť pro předchozí
snímek

MHS – Kompresce videa

18 / 45

INTER KÓDOVÁNÍ

MOTION COMPENSATION

- časově nejnáročnější část komprese
- neslouží k identifikaci pohybu, ale k minimalizaci rozdílu
- kvalitnější motion compensation
⇒ menší rozdílový snímek
⇒ lepší kompresní poměr
- vylepšení: převzorkování snímku na vyšší rozlišení
⇒ motion vektor s přesností 1/2 nebo 1/4 pixelu
- kromě dat pixelů ukládáme data motion vektorů
⇒ přesnější motion vektory = menší rozdíly pixelů, více
dat pro motion vektory ⇒ hledání kompromisu
- je-li rozdíl makrobloků velký, kóduje se bez predikce

MHS – Kompresce videa

19 / 45

INTER KÓDOVÁNÍ

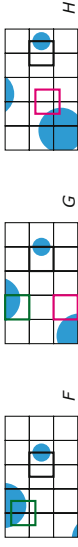
- urychlení motion compensation
– typický pohyb vodorovný – větší hledání vodorovně
– výpočet 1 motion vektoru na makroblok
– výpočet motion vektoru jen pro Y složku
– odhad motion vektoru v nižším rozlišení, pak
zprášení
– motion vektory u okolních makrobloků a předchozího
snímku podobné
⇒ využijeme pro počáteční odhad motion vektoru
⇒ motion vektory kódujeme rozdílově
– využití operace korelace implementované přes FFT
⇒ maxima výsledku dávají odhad motion vektorů

MHS – Kompresce videa

20 / 45

INTER KÓDOVÁNÍ

- problém predikce: makrobloky G_1 , G_2 , G_3



- G_1 – dobrá predikce ze snímku F
- G_2 – predikci ze snímku F nelze udělat
- řešení – obousměrné (bidirectional) kódování
- G_1 – predikce z makrobloku z^{Fw}
- G_3 – predikce z makrobloku z^{Hw}
- G_2 – predikce z makrobloku $\frac{1}{2}(z^{Fw} + z^{Hw})$

INTER KÓDOVÁNÍ

MPEG TERMINOLOGIE

- I snímek – všechny makrobloky bez predikce
- P snímek – některé makrobloky predikované z předchozího snímku I nebo P snímku
- B snímek – některé makrobloky predikované z předchozího či následujícího I nebo P snímku (u H.264 a jiných moderních kodeků i z B snímku)
- značení X_y – snímek č. Y je typu X

INTER KÓDOVÁNÍ

(ZPĚTNÉ) PREDIKTIVNÍ KÓDOVÁNÍ

- sekvence $I_0, P_1, P_2, P_3, P_4, P_5, P_6, \dots$
- výhoda: neustálé vylepšování statického obrazu
- typický P snímek 2–3x menší než I snímek
- v případě chyby poškozen zbytek videa
- nemožnost náhodného přístupu do videa
- nemožnost rozdělit video (není zaručeno, že snímek P_n půjde zakódovat jako I_n ve stejné kvalitě)
- částečné řešení: vynucené vkládání snímku bez predikce – např. sekvence $I_0, P_1, P_2, P_3, \dots, P_{24}, I_{25}, P_{26}, P_{27}, \dots$

INTER KÓDOVÁNÍ

OBOUSMĚRNÉ PREDIKTIVNÍ KÓDOVÁNÍ

- sekvence $I_0, B_1, B_2, P_3, B_4, B_5, P_6, B_7, B_8, I_9, B_{10}, B_{11}, P_{12}, \dots$
- dekodér musí snímky přijímat v jiném pořadí
- dekodér musí obsahovat paměti pro předchozí (F) a následující (H) snímek
- vstup: $I_0, P_3, B_1, B_2, P_6, B_4, B_5, B_7, B_8, I_9, B_{10}, B_{11}, \dots$
- I_0 – dekodovat, uložit do F, zobrazit v čase $t=0$
- P_3 – dekod., uložit do H, naplánovat zobrazení na $t=3$
- B_1 – dekod., neukládat, zobrazit v čase $t=1$
- B_2 – dekod., neukládat, zobrazit v čase $t=2$
- zobrazit H, přesunout H do F
- P_6 – dekod., uložit do H, naplánovat zobrazení na $t=6$

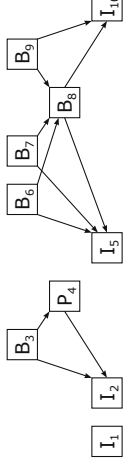
INTER KÓDOVÁNÍ

- **výhoda:** typický B snímek 2–3x menší než P snímek
nevýhoda: data B snímku nezlepší obraz (nepřináší
informaci pro okolní snímky)
- struktura IBBB...BBP typicky nevýhodná
 - velká prodleva (problém u přímých přenosů)
 - obousměrná predikce standardně $\frac{1}{2}(zF_n^{MV} + zH_n^{MV})$
⇒ lepší by bylo $\omega zF_n^{MV} + (1 - \omega)zH_n^{MV}$ $0 < \omega < 1$
⇒ řeší H.264 (MPEG-4 AVC)
⇒ vynikající pro kódování prolínáčky, stmívačky, ...

INTER KÓDOVÁNÍ

VYLEPŠENÉ OBOUSMĚRNÉ PREDIKTIVNÍ KÓDOVÁNÍ

- u klasického obousměrného kódování nesmí B snímek záviset na jiném B snímku – zbytečné omezení řešení (H.264 aj.): B snímek smí záviset na jiném, ale závislosti se nesmí zacyklit



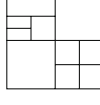
- uložení: $I_1, I_2, P_4, I_5, I_{10}, B_6, B_7, B_8, B_9$

INTER KÓDOVÁNÍ

- GOP (group of pictures) – sekvence na sobě závislých snímků (tj. od I do I snímků)
- sekvence $I_0, B_1, B_2, P_3, B_4, P_5, B_6, B_7, B_8, P_9, B_{10}, B_{11}, P_{12}, \dots$
něže jednoduše rozdělit
- částečná řešení:
 - sekvence (Closed GOP) $I_0, B_1, B_2, P_3, B_4, P_5, B_6, P_7, I_8, B_9, P_{10}, \dots$
– lze bez problémů dělit na hranici GOP
 - sekvence $I_0, B_1, I_2, B_3, I_4, B_5, \dots$
– při rozdělení snadný přepočít B snímků
- struktura GOP stálá nebo proměnná (záleží na kodéru)

DALŠÍ POSTUPY

- dynamická volba bloku: typicky 4×4 až 32×32 px
 - malý blok: přesnější motion compensation, lepší pro členitou část obrazu
 - velký blok: lepší komprese pro plochou část obrazu
- interpolované motion vektory (MV) – místo jednoho MV pro makroblok se uloží pro každý roh jeden
 - ⇒ každý pixel má vlastní interpolovaný MV
- sdílení MV – všechny bloky MB mohou sdílet jeden MV, nebo má každý blok svůj
- bloky (jednotky DCT) vznikají dělením MB na poloviny nebo čtvrtiny



DALŠÍ POSTUPY

- motion compensation na libovolný snímek v GOP
- in-loop deblocking
 - vyhlazováním lze eliminovat hrany na okrajích bloků
 - kodér předpokládá, že dekodér bude deblocking provádět ⇒ motion compensation je vůči vyhlazenému obrazu
- alternativy k DCT s podobnými vlastnostmi, ale založenými na celočíselné aritmetice
- alternativní průchody 2D polem
 - převod do 1D proudu
 - výběr podle toho, kde se očekává koncový šluk nul



DALŠÍ POSTUPY

- do výstupního proudu dat se ukládají kromě kvantizovaných DCT koeficientů také
 - motion vektory,
 - parametry kvantizace
 - typ snímku / makrobloku / bloku
 - typ dělení makrobloku
 - ...
- ⇒ vše se pokud možno kóduje prediktivně (rozdílově), vše se kóduje Huffmanovým nebo aritmetickým kodérem

DALŠÍ POSTUPY

ŠKÁLOVATELNOST

- požadavek: mít několik verzí videa různé kvality (rozlíšení, ztrátovost, snímková frekvence, ...)
- naivní řešení: pro každou kvalitu jedna verze
 - nevýhodné: kvalitnější video musí obsahovat stejné informace jako nekvalitní plus nějaké navíc
- dobré řešení: základní video uloženo v nejhorší kvalitě, pro lepší kvalitu se do proudu dat doplní chybějící informace
- škálovatelnost je přirozená u kodeků založených na diskrétní waveletové transformaci

KOMPRESNÍ SCHÉMATA

Následující snímky představují reálná kompresní schémata různé složitosti a účelu. Cílem není pamatovat si, jaké techniky a parametry dané schéma používá; cílem je získat povědomí o rozsazích parametrů (např. šířka motion compensation) a o typických kombinacích technik uvedených na předchozích snímcích.

H.261

- přenos videa na linkách $p \times 64$ kbit/s (ISDN)
- pro videokonference
- formát CIF nebo QCIF, 4:2:0
- blok 8×8 pixelů, makroblok 2×2 bloky
- pouze I a P snímky, DCT po makroblocích
- kvantizace
 - jeden kvantizační koeficient pro DC koeficienty, jeden pro AC koeficienty – řízení toku dat
 - je-li rozdíl pro P kódování velký, kóduje se MB jako I
- kódování motion vector pomocí DPCM
- VLC pro koeficienty DCT a motion vector
- vynucený intra frame – zotavení z chyby

H.263

- rozšíření H.261, pro videokonference
- datový tok menší i větší než H.261
- další rozlišení a snímkové frekvence
- motion vector
 - přesnost $\frac{1}{2}$ px, bilineární interpolace
 - rozsah –16 až 15,5 px horizontálně i vertikálně
 - predikce MV z makrobloků vlevo, nahoře a vlevo nahoře od (de)kódovaného
- volitelně větší rozsah motion vector, aritmetické kódování, 4 motion vector na makroblok, ...
- podobné MPEG-1
- další specifikace H.263+, H.263++

MPEG-1

- původně pro tok 1,5 Mbit/s
- formát typicky SIF, 4:2:0
 - 352×240 NTSC
 - 352×288 PAL
- činnost po snímcích
- kvalita jako VHS
- VideoCD
- překonaný formát

MPEG-1

- blok 8×8 pixelů, makroblok 2×2 bloky
- motion compensation na makrobloku
 - přesnost motion vector halfpel nebo pel
 - rozsah –512 až 511,5 px (resp. –1024 až 1023 px)
 - makroblok B – 1x nebo 2x motion vector, kompenzace vůči následujícímu, nebo předchozímu i následujícímu snímku

MPEG-1

- motion compensation na makrobloku
- DCT na bloku
- kvantizace
 - pro I (makro)bloky každý koeficient jiný
 - pro P/B všechny koeficienty stejná kvantizace
- struktura „slice“ – řada makrobloků s rozdílově kódovanými motion vektory nebo DC členy
- zigzag scan
- VLC – DCT koeficienty, motion vectors, kvantizační faktory, ...
- formování výstupu
 - zvuk MPEG-1 Audio Layer 1/2/3

MHS – Kompresce videa

37 / 45

MPEG-2

- též H.262, založené na MPEG-1
- vysílání na větší šířce pásma, DVD
- TV 3–15 Mbit/s (typicky 6 Mbit/s), HDTV 15–30 Mbit/s
- obecné kódování

PROFILES, LEVELS

- profil – množina nástrojů
- level – omezení parametrů pro nástroje
- typicky MP@ML (main profile at main level)

MHS – Kompresce videa

38 / 45

MPEG-2

PROFILY

- Simple – bez B snímků
- Main – navíc B snímky
- SNR – navíc SNR škálovatelnost
- Spatial – navíc Spatial škálovatelnost
- High – navíc 4:2:2 chroma subsampling, 11bitové rozlišení DC koeficientu
- Multiview – více pohledů kamer, např. pro stereoskopii

MHS – Kompresce videa

39 / 45

MPEG-2

ÚROVNĚ

Level	High	High1440	Main	Low
max. rozlišení	1920 x 1152	1440 x 1152	720 x 576	352 x 288
max. fps	60	60	30	30
max. px/s	62,7 · 10 ⁶	47,0 · 10 ⁶	10,4 · 10 ⁶	3,0 · 10 ⁶
max. bitrate	80 Mbps	60 Mbps	15 Mbps	4 Mbps
aplikace	film	studio HDTV	studio SDTV	consumer

MHS – Kompresce videa

40 / 45

MPEG-2

- 4:4:4, 4:2:2, 4:2:0 chroma subsampling
- podpora prokládaného i neprokládaného režimu
 - motion compensation mezi snímky, půlsnímky
 - scan DCT koeficientů zigzag nebo vertical



- škálovatelnost – video dělené na několik vrstev
 - SNR – další vrstva zlepšuje kvalitu obrazu
 - Spatial – další vrstva zlepšuje rozlišení
 - Temporal – další vrstva zlepšuje fps

MPEG-4

- multimedialní specifikace
- video – kódování video objects (VO) s libovolným tvarem
 - kódování tvaru
 - kódování textury
 - speciální motion compensation – předpokládáme MV „v rozích“ MB, pro každý pixel odvozen speciální MV
- pro kompresi videa simple profile a advanced simple profile (DivX, Xvid) – VO obdélníkový
- komprese
 - založená na MPEG-2
 - nové schéma AVC (Advanced Video Coding, H.264)

MPEG-4

SIMPLE PROFILE

- advanced intra coding – predikce v intra blocích
- až 4 motion vectors na makroblok
- dlouhé motion vectors
- pouze I a P snímky

ADVANCED SIMPLE PROFILE

- motion vectors s přesností quarterpel
- global motion compensation – speciální režim pro pohyb kamery
- B-VOP (Video Object Plane) – B snímky
- podpora prokládání

H.264

- MPEG-4 part 10 (Advanced Video Coding, AVC), H.264
- makrobloky 16x16
 - intra kódování s predikcí luma na bloku 4x4 (9 režimů) – přechod atd. chroma na bloku 8x8 nebo 16x16 (4 režimy) samotný režim predikovaný
 - inter kódování s motion compensation dělení makrobloku na subbloky pro každý subblok motion vector (quarterpel) motion vectors predikované motion compensation na n obrázcích
- in-loop deblocking filtr

H.264

- dělení makrobloku na bloky 4×4 pixely, integer „DCT“
 - pro intra makrobloky oddělení DC a AC koeficientů
- kvantizace založená na bitových posuvech
- kódování VLC nebo adaptivním aritmetickým kóděrem