

# Formáty videa

David Bařina

31. března 2017

# Obsah

- 1 Komprese videa
- 2 Bezeztrátové formáty
- 3 Ztrátové formáty
- 4 Kontejnery
- 5 Shrnutí

# Multimédia



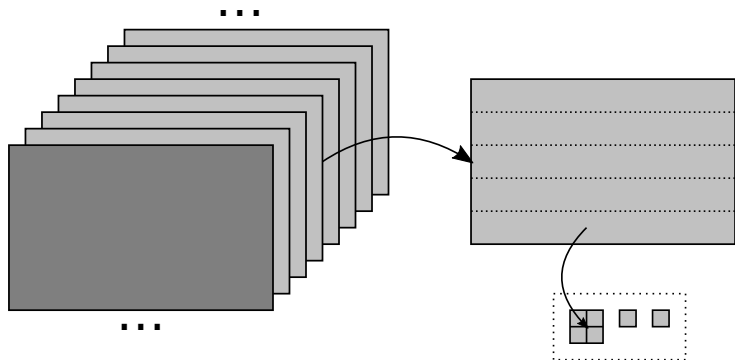
- multimédia:
  - text, zvuk, statický obraz, **video**, metainformace, . . .
- potřeba:
  - ▶ získávat (kamera),
  - ▶ ukládat (pevný disk, komprese),
  - ▶ vyhledávat (podle popisu),
  - ▶ přehrávat,
  - ▶ upravovat (střih videa), . . .
- ukládání: kontejner + kodeky

# Pojmy

- kontejner
  - ▶ datové proudy/stopy (audio, video, titulky, kapitoly)
  - ▶ např. AVI, Matroska, MPEG-TS
- formát videa
  - ▶ ztrátová, bezztrátová komprese
  - ▶ např. H.264, MPEG-4 ASP
- kodek
  - ▶ kompresor + dekompresor
  - ▶ např. x264, DivX, libavcodec
- FourCC (4CC, four character code)
  - ▶ AVI
  - ▶ např. DIVX označuje DivX
- přenosová rychlost (bitrate) [bps]
- bitrate na segment:  
CBR (konstantní), VBR (fixní kvalita), ABR (dlouhodobý průměr)
- postprocessing (deblocking)

# Video

- 3D signál  $(x, y, t)$ , barevný  $(Y, C_b, C_r)$
- sekvence (v čase podobných) snímků
- hierarchie:  
video, skupiny snímků, snímky, řezy, makrobloky, bloky, vzorky



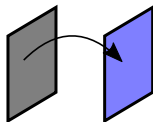
- typ (klíčový, rozdílový)
- snímková frekvence [fps]
- časová razítka
  - ▶ decoding time stamp (DTS)
  - ▶ presentation time stamp (PTS)
- index (dle času, čísla snímku)

# Typy snímků

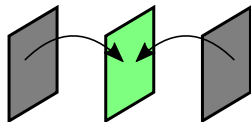
- klíčový / intra / I snímek
  - ▶ prostorová souvislost a redundance
- rozdílový / inter / P / B snímek
  - ▶ časová redundance
  - ▶ B-snímek problematický
  - ▶ framedrop u B-snímků
- jiné snímky
  - ▶ D-snímek
  - ▶ golden a altref snímky
- změna pořadí (frame reordering)
  - ▶ B-snímek až po referenčních



I



P



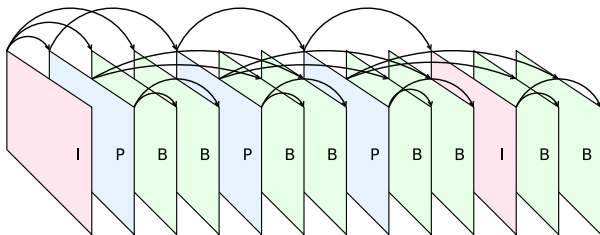
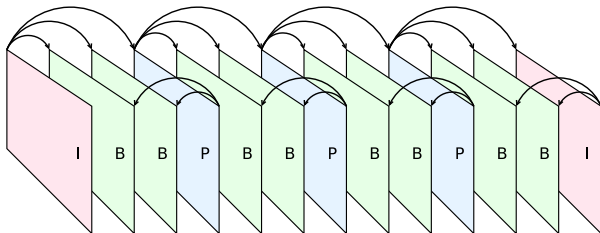
B

# Skupina snímků

- GOP (*group of pictures*)
- sekvence po sobě jdoucích snímků
- začíná I-snímkem
- otevřený vs. uzavřený GOP
- jakýkoli snímek dekódovatelný ze své GOP (uzavřený)
- odolnost proti chybám
- velikost GOP
  - ▶ seekování déle trvá
  - ▶ náročnost na paměť
  - ▶ vyšší kompresní poměr
  - ▶ typicky 12, max. 18

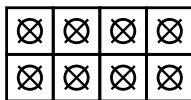


# Skupina snímků

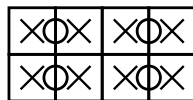


# Podvzorkování

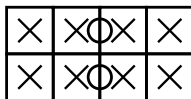
- $Y$  vs.  $C_b$ ,  $C_r$
- člověkem nepozorovatelné
- $J:a:b$ , 4:4:4, 4:2:2, 4:1:1, 4:2:0
- centroidy



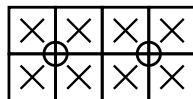
4:4:4



4:2:2



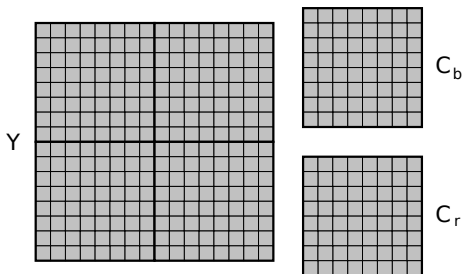
4:1:1



4:2:0

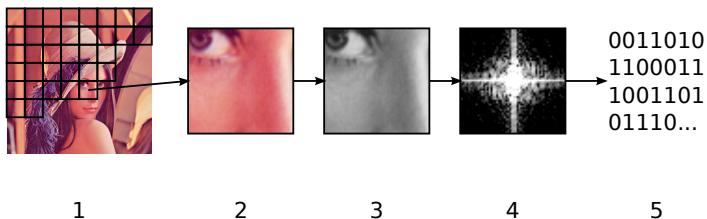
# Bloky, makrobloky

- blok =  $8 \times 8$  vzorků, DCT
- v novějších standardech bloky i menší (4–16)
- makroblok = 6 bloků (4:2:0)  
 $2 \times 2 Y + 1 C_b + 1 C_r = 6$   
 $16 \times 16 + 8 \times 8 + 8 \times 8$
- makrobloky mohou být děleny na oblasti (H.264)
- základní jednotka kompenzace pohybu

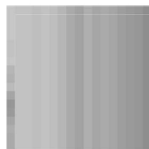
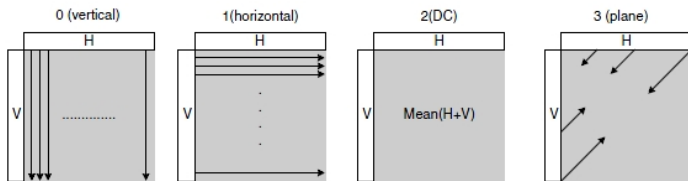


# Intra-snímkové kódování

- prostorová redundance
- jako JPEG
- makrobloky, případně části makrobloků
- predikce (několik režimů)
- transformace (DCT, DWT)
- kvantizace (adaptivní  $q \cdot Q$ ), zigzag
- kódování (RLE-0, VLC)



# Predikce makrobloku



0 (vertical), SAE = 3985



1 (horizontal), SAE = 5097



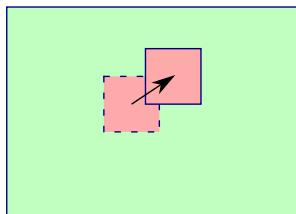
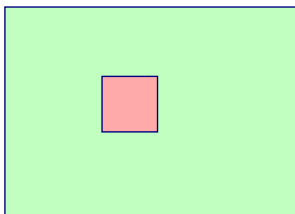
2 (DC), SAE = 4991



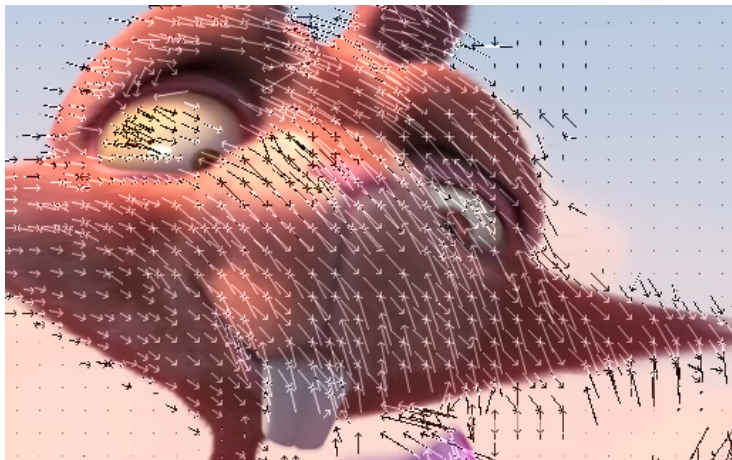
3 (plane), SAE = 2539

# Odhad a kompenzace pohybu

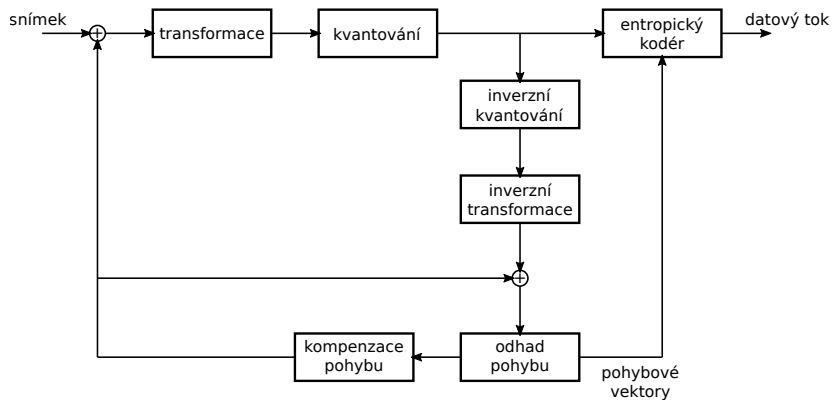
- časová redundance
- makrobloky, případně části makrobloků
- odhad pohybu (ME)
- pohybový vektor (MV) – přesnost (half-pel, Qpel), rozdílově
- kompenzace pohybu (MC)
  - ▶ globální (GMC) – afinní transformace
  - ▶ bloková (BMC) – pohybový vektor
  - ▶ s přesahujícími bloky (OBMC) – výpočetní složitost



# Pohybové vektory



# Hybridní kodér





# Prokládání

- v jednom okamžiku jen sudé nebo liché řádky
- při pohybu viditelné artefakty



- odstranění prokládání (algoritmy)
- digitální obraz, MPEG
  - ▶ liché a sudé řádky ve zvláštních snímcích (field)
  - ▶ liché a sudé řádky v jednom snímku prokládaně (frame)

# Standardy

- organizace: ISO/IEC, ITU-T (CCITT), SMPTE
- standardizace: syntaxe datového toku + dekodér
- profily (algoritmy)
  - ▶ enkodér
  - ▶ dekodér
- levely (technická omezení)



# Nekomprimované (RAW) snímky

- FourCC: 0x00000000
- nekomprimovaný DIB, nejčastěji BGR24

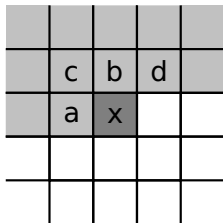
# Microsoft RLE

- FourCC: `mrle`
- MSDN: 4, 8 bpp; FFmpeg: 16, 24, 32
- dvojice bajtů A, B

A	B	význam
0	0–2	escape kód (konec obrázku)
0	3–255	dalších B bajtů kopírovat
1–255	0–255	dalších A bajtů hodnotu B

# Huffyuv

- rychlý, GPL, fork Lagarith
- FourCC: HFYU
- prediktor + entropický kodér
- prediktory: levý  $a$ , rovina  $p$ , medián  $\text{med}(a, b, p)$



$$p = a + b - c$$

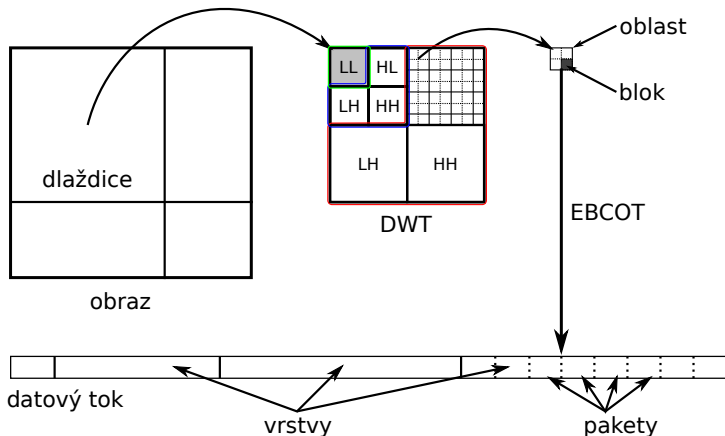
- barevný model:  $R - G, G, B - G$
- každý kanál vlastní Huffmanovu tabulku (statická)
- Lagarith: prediktor (medián), RLE, aritmetické kódování, null frames

## Motion JPEG

- není standard, nekompatibilita, kamery, mnoho FourCC
- I-snímky, JPEG
- existuje Motion JPEG 2000

# JPEG 2000

- klíčové snímky, JPEG 2000
- profesionální použití (digitální kinematografie)

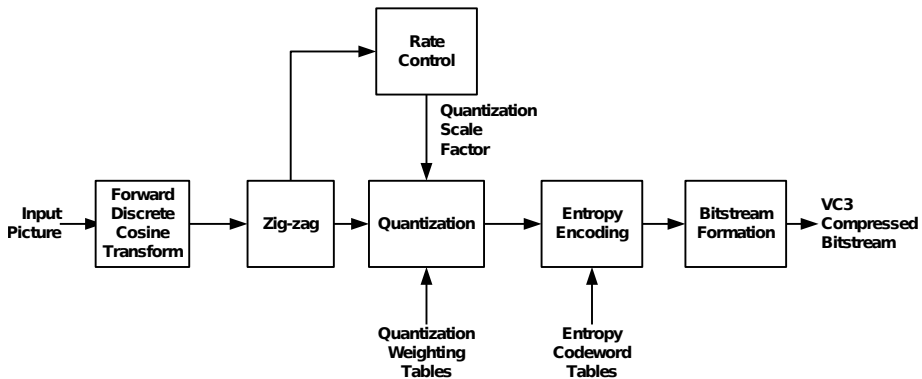




## VC-3 (DNxHD)

- americká společnost Avid, v roce 2008 standardizace SMPTE
- jako JPEG, velmi krátká specifikace
- profesionální použití (televizní průmysl), často v MXF
- 4:2:2 (8bitové YCbCr, 10bitové YCbCr) a 4:4:4 (10bitové RGB)
- progresivní i prokládané video
- rozdíly proti JPEGu
  - ▶ koeficient DC není kvantován
  - ▶ kvantizační a VLC tabulky pevně specifikovány ve standardu
  - ▶ kvantizační tabulka je vynásobena součinitelem (na makroblok)
  - ▶ podpora kontrolního součtu CRC ke každému snímku

# VC-3 (DNxHD)



- krátká specifikace (cca 1000 řádků), 2004, 4CC VP30, VP31
- YCbCr 4:2:0,  $8 \times 8$  DCT, kvantizace, RLE-0, zig-zag, DC rozdílově, Huffmanovo kódování
- bloková kompenzace pohybu (BMC), golden(I)/inter(P)-snímky
- fragment = blok  $8 \times 8$   
superblok =  $4 \times 4 = 16$  fragmentů (nezávisle kanály)  
makroblok = bloky  $2 \times 2 Y + 1 U + 1 V = 6$  (přes kanály)
- bloky v superbloku po Hilbertově křivce
- golden-snímky: jako JPEG
- inter-snímky: na každý makroblok 8 módů  
(I-snímek, rozdíl od předchozího nebo I-snímku, pohybový vektor)

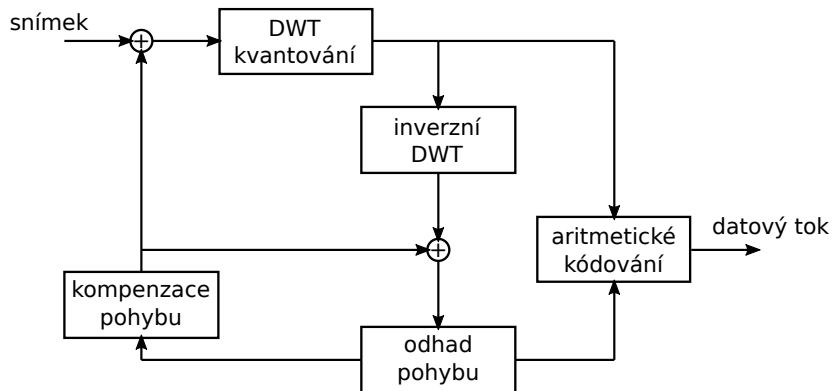
# Theora



- pro Ogg, Xiph.org Foundation, FourCC theo
- vychází z VP3
- I/P-snímky (ne B-snímky)
- $8 \times 8$  DCT, podvzorkování C, BMC, max. 8 bitů na kanál
- libtheora

# Dirac

- ztrátový, open source, BBC, FourCC: BBCD
- modely YCC, prokládání, podvzorkování 4:4:4, 4:2:2, 4:2:0
- DWT, na výběr 7 vlnek
- implementace
  - ▶ Schrödinger v C (optimalizovaný)
  - ▶ dirac-research (původně Dirac) v C++
- GOP, I/P/B-snímky, kompenzace pohybu (pohybové vektory)
- kontextový arit. kodér nebo exp-Golombovy kódy
- podmnožina jako VC-2 (Dirac Pro) – pouze I-snímky
- není blokový efekt jako u MPEG
- specifikace 134 stran, 2008



- Google (On2), podobný H.264
- implementace libvpx, podpora v prohlížečích (HTML5)
- přesná DCT, WHT (Walsh–Hadamard transform)
- pouze 8bitový YUV 4:2:0
- každý snímek na makrobloky ( $16 \times 16$  Y,  $8 \times 8$  U a V pixelů), makrobloky v rastrovém průchodu
- makrobloky na  $4 \times 4$  podbloky (16 Y, 4 U, 4 V), podbloky v rastrovém pořadí
- DCT/WHT na  $4 \times 4$ 
  - ▶ DCT pro podbloky 16Y, 4U, 4V
  - ▶ WHT pro DC koeficienty z  $4 \times 4$  DCT bloků Y
- není změna pořadí snímků (frame reordering)
- boolean entropy coder = varianta arit. kodéru
- specifikace pomocí úseků zdrojových kódů v C; 303 stran, ale čtivá

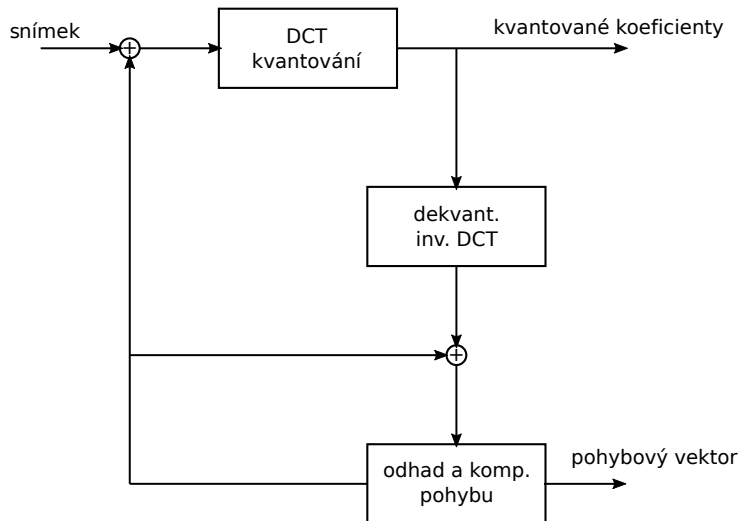
- pouze I(intra,key)/P(inter)-snímky, žádné B-snímky
- P-snímky z kteréhokoli předchozího snímku GOP
- zavádí golden a altref snímky: I-snímek je vždy golden a altref
- jakýkoli P-snímek může nahradit poslední golden nebo altref snímek
- makrobloky P-snímků jsou predikovány z jednoho snímku: předchozího, posledního golden, altref, dekódované části současného
- I(intra)-predikce: predikce makrobloku z již dekódovaného řádku nad a sloupce vlevo; 4 metody
- P(inter)-predikce: pohybové vektory ( $Q_{pel}$ ), pohybové vektory pro chromatické podbloky se spočtou jako průměr vektorů ze 4 odpovídajících Y podbloků
- po dekódování snímku následuje ještě filtrace (loop filter) – odstraní blokové artefakty (nezaměňovat s postprocesingem)



# H.261

- ITU-T, 1988, 29 stran
- CIF, QCIF, YCbCr 4:2:0
- makroblok ( $16 \times 16$  Y +  $8 \times 8$  Cb +  $8 \times 8$  Cr)
- $8 \times 8$  DCT, kvantizace, zigzag, EC (RLE-0 + VLC)
- inter/intra-snímky (inter-snímky: časová redundance)
- kompenzace pohybu (MC)
  - ▶ jeden vektor na makroblok
  - ▶ z předchozího snímku
  - ▶ pohybové vektory (pouze integer + možný blur)
- loop filter

# H.261



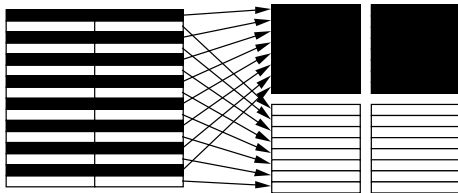
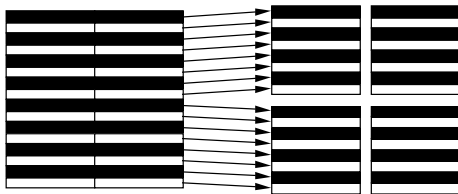
## MPEG-1 část 2

- ISO/IEC 11172-2, 1993, vychází z H.261 a JPEG
- VCD, (SVCD), (DVD)
- YCbCr 4:2:0
- I/P/B/D-snímky
- GOP (15–18)
- I-snímky jako JPEG
- P-snímky oproti I/P-snímkům
- pohybové vektory pro každý makroblok (half-pel)
- P-snímky 1 vektor / makroblok; B-snímky 2 vektory
- DTS (decoding time stamps) v kontejneru
- D-snímky DC koeficienty
- bloky  $8 \times 8$ , makrobloky  $16 \times 16 = 6$  bloků (4Y + Cb + Cr)

## MPEG-2 část 2 / H.262

- ISO/IEC 13818-2, 1995, vylepšuje MPEG-1, ale složitější
- prokládané (interlaced) video
- podvzorkování 4:2:2, 4:2:0, 4:4:4
- I/P/B-snímky (není D)
- makrobloky, 1 pohybový vektor (MV) na makroblok
- standard 243 stran
  - ▶ profily (algoritmy, formáty obrazu)
  - ▶ levely (bitrate, velikost snímku)
  - ▶ např. SP nemá B-snímky
- SVCD, DVD, HD DVD, BD, digitální TV (DVB-T)
- VBR

# MPEG-2 část 2 / H.262



## MPEG-4 část 2

- ISO/IEC 14496-2, 1999, založen na H.263
- YCbCr 4:2:0, I/P/B-snímky, DCT, MV
- není loop filter
- standard: profily, 517 stran, komplexní (3D tvar)
  - ▶ SP: bez B-snímků
  - ▶ ASP: B-snímky, prokládání, Qpel, GMC (global motion compensation)
- Qpel
- GMC
- mnoho implementací (DivX, Xvid, 3ivx, FFmpeg), mnoho 4CC

## MPEG-4 část 10 / H.264

- ISO/IEC 14496-10, ITU-T H.264, AVC (Advanced Video Coding)
- 2003, mnoho nových vlastností
- 4CC AVC1, H264
- 4:0:0 (monochromatický), 4:2:0, 4:2:2, 4:4:4
- BD, HD DVD, digitální TV (DVB – HD programy)
- I/P/B-snímky, DCT a WHT (Walshova–Hadamardova transformace)
- WHT pro DC koeficienty
- loop filtr
- standard 680 stran

## MPEG-4 část 10 / H.264

- bloky transformace různých velikostí
- pro predikci P/B-snímku až 16 referenčních snímků (namísto 1/2)
- kompenzace pohybu pro bloky  $4 \times 4$  až  $16 \times 16$
- makroblok rozdělen na oblasti
- více pohybových vektorů (MV) na jeden makroblok (oblasti)
- vážená predikce (mezi dvěma ref. snímky)
- pro predikce I-snímku predikce z hran okolních bloků
- entropické kodéry: Exp-Golomb, CABAC a CAVLC
- implementace: x264, FFmpeg, DivX, CoreAVC



## Další formáty

- H.265 / HEVC  
High Efficiency Video Coding, ISO/IEC 23008-2, MPEG-H část 2  
ve vývoji, nástupce H.264
- VP9
- WMV + VC-1
- ...

# AVI (Audio Video Interleave)

- přípona .avi
- společnost Microsoft, od roku 1992
- implementace RIFF (Resource Interchange File Format)
- k identifikace kodeku FourCC (např. DIVX = DivX)
- není možné uložit titulky, kapitoly, menu (třetí strany)
- původní omezení 1 GB
- index podle čísla snímku (ne času) na konci souboru (nevhodný ke streamování)
- široká podpora
- nemá podporu B-snímků

# Matroska

- přípona .mkv
- otevřený formát, oznámen 2002
- podporuje vše, neomezený počet stop
- založen na EBML (Extensible Binary Meta Language)
- kodek určen CodeCID (řetězec)
  - ▶ např. V\_MPEG4/ISO/AVC = H.264
  - ▶ typ V\_MS/VFW/FOURCC zaobaluje FourCC
  - ▶ <http://haali.su/mkv/codecs.pdf>
- WebM (přípona .webm) je profil formátu Matroska
  - ▶ Google, 2010, určen pro web (HTML5)
  - ▶ kodeky pouze VP8 + Vorbis

# MP4 (MPEG-4 Part 14)

- přípona .mp4
- standard ISO, od roku 2003 (verze 2)
- založen na MPEG-4 Part 12, který je založen na QuickTime
- k identifikaci objektů (také kodeků) čtyřznakové identifikátory
  - ▶ např. avc1 pro MPEG-4 AVC
  - ▶ code-points, <http://www.mp4ra.org/codecs.html>
- podporuje libovolné kodeky (privátní a registrované), z registrovaných je široce podporovaných pouze pár

# ASF (Advanced Systems Format)

- přípony .asf, .wma, .wmv
- Microsoft
- objekty (také použité kodeky) identifikovány GUID
  - ▶ zaobaluje FourCC, např.  
`DEFINE_GUID(CLSID_XVID, mmioFOURCC('D','I','V','X')), ...`
- teoreticky libovolný kodek
- index přiřazen klíčovým snímkům, může být podle času, čísla snímku nebo ref. čas. razítka

# QuickTime, Ogg

## QuickTime

- přípona .mov
- společnost Apple, také QuickTime File Format (QTFF)
- je na něm založen MPEG-4 Part 12

## Ogg

- přípony .ogv, .oga, .ogx, .ogg
- otevřený formát
- součást projektu Ogg, zastřešené Xiph.Org Foundation
- dříve ve specifikaci HTML5
- hack OGM (Ogg Media)

# FLV, 3GP

## Flash Video

- přípony .flv, .f4v
- společnost Adobe Systems (původně Macromedia)
- dvě verze: FLV (SWF) a F4V (MPEG-4 Part 12)
- omezená podpora kodeků

## 3GP, 3G2

- přípony .3gp, .3g2
- založeny na MPEG-4 Part 12
- pro mobilní zařízení
- omezená podpora kodeků

# MPEG

## MPEG-PS

- ISO, specifikován v MPEG-1 Part 1 a MPEG-2 Part 1
- VOB, EVO (s omezeními) – DVD-Video a HD DVD
- omezená podpora kodeků + privátní proudy
- ES (audio, video), PES (pakety),  
PS (kombinuje více PES – např. audio a video)

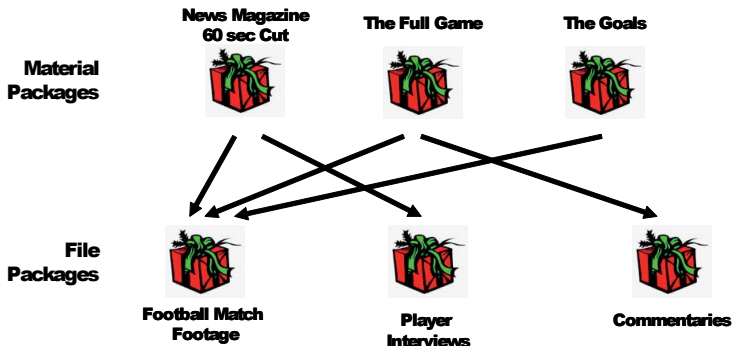
## MPEG-TS

- ISO, 1995, specifikován v MPEG-2 Part 1
- může nést více PS (Program Streams)
- omezená podpora kodeků + privátní proudy
- televizní vysílání
- TS obaluje více PES



# MXF (Material eXchange Format)

- televizní průmysl
- schéma KLV (Key-Length-Value)



# Shrnutí

- pojmy: formát, kodek, kontejner, FourCC, snímek, GOP, makroblok
- metody: intra-kódování, inter-kódování, hybridní kódér
- formáty videa:  
MS RLE, Huffiyuv,  
VP3/Theora, Dirac, VP8, H.261, MPEG-{1,2,4}, H.264
- kontejnerové formáty:  
AVI, MKV, MP4 (ISO, 3GP, MOV), MPEG-{PS,TS}