

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Bakalářská práce

Metody komprese digitálního videa

Plzeň, 2011

Jiří Loudil

Prohlášení

Prohlašuji, že jsem bakalářskou práci vypracoval samostatně a výhradně s použitím citovaných pramenů.

V Plzni dne 3. června 2011 Jiří Loudil,

Abstrakt

Compression Methods for Digital Video

There are many various methods used for digital video compression. This thesis compares a set consisting of existing widespread and developing video compression methods. The comparison consists of one subjective and two objective digital video quality metrics. Also the compression speed and the output file size are analyzed. The objective measuring was performed by 'common software tool set' and the subjective measurements were done in compliance with a modified version of commonly the used 'SCACJ' method. The results in general showed a very good performance of the 'H.264' based codec across the objective and subjective metrics.

Poděkování

Tímto bych chtěl poděkovat vedoucímu bakalářské práce Ing. Josefu Kohoutovi, Ph.D. za cenné rady, připomínky a vedení práce.

Obsah

1	Úvod	4
2	Existující metody digitální video komprese	5
2.1	MPEG-2 Part 2	5
2.2	H.264/MPEG-4 Part 10	10
2.3	DivX Codec	15
2.4	Xvid	18
2.5	WMV9/VC-1	20
2.6	Dirac	23
2.7	Theora	26
3	Teoretické shrnutí vlastností kompresních metod	28
3.1	MPEG-2 Part 2	28
3.2	H.264/MPEG-4 Part 10	28
3.3	DivX Codec	28
3.4	Xvid	28
3.5	WMV9/VC-1	29
3.6	Dirac	29
3.7	Theora	29
4	Zjišťování kvality video komprese	30
4.1	Úvod do metod měření kvality ztrátové video komprese	30
4.1.1	Metriky pro objektivní hodnocení kvality video komprese . . .	30
4.1.2	Metody pro subjektivní hodnocení kvality video komprese . .	31
4.2	Měřicí nástroje	33
4.2.1	MSU Video Quality Measurement Tool	33
4.2.2	tiny_ssim	34
4.2.3	MSU Perceptual Video Quality Tool	34
5	Popis realizace měření kvality video komprese	35
5.1	Implementace enkodérů	35
5.2	Rozlišení a datové toky testovacích souborů	36
5.2.1	Rozlišení	36
5.2.2	Datové toky	36
5.3	Testovací video data – samplly	38
6	Objektivní část měření	41
6.1	Použitá metodika	41
6.2	MSSIM	43

6.3	PSNR	45
6.4	Chrominanční složky	47
6.5	Výsledná velikost	49
6.6	Doby běhu	51
7	Subjektivní část měření	52
7.1	Použitá metodika	52
7.2	Výsledky prvního kola experimentů	53
7.3	Výsledky druhého kola experimentů	55
8	Praktické shrnutí výsledků	56
8.1	MPEG-2 Part 2	56
8.2	H.264/MPEG-4 Part 10	56
8.3	Xvid	56
8.4	WMV9/VC-1	56
8.5	Dirac	56
8.6	Theora	57
9	Závěr	58

1 Úvod

V dnešním multimediálním světě je v sěži myslitelné, aby byla digitální video data díky svému datovému objemu uchovávána a zpracovávána výhradně ve své nekomprimované podobě. Digitální video data je tedy třeba po zachycení podrobit procesu komprese. Pro svoji často ztrátovou povahu dosahuje tato komprese velice dobrých kompresních poměrů, avšak výsledný video soubor tak přirozeně ztrácí část informace. Metody minimalizace této informační ztráty a jejího dopadu na lidskými smysly vnímatelnou kvalitu obrazu jsou předmětem snah při navrhování ztrátových digitálních video kompresních metod. Přičemž lze provádět porovnání míry ztráty informace na základě rozdílů v obrazové kvalitě mezi původním a výsledným video souborem.

Ztrátové metody video komprese se liší především svým dosažitelným kompresním poměrem, složitostí, kvalitou výstupního video souboru a také oblastí aplikačního nasazení. Především implementace kodeků z rodiny standardů *MPEG*¹ (*MPEG-1 Part 2*, *MPEG-2 Part 2*, *MPEG-4 Part 2 ASP*) a *H.260*² (*H.261*, *H.263*, *H.264*) se staly velice rozšířenými díky svým vhodným vlastnostem pro standardizaci v optických nosičích a síťových službách.

Tato práce rozebírá pomocí objektivního a subjektivního hlediska kvalitu video komprese na skupině silně rozšířených i méně zavedených ztrátových kompresních metod. Také jsou brány v potaz systémové nároky při enkódovacím procesu a výsledná velikost video souboru. Pro srovnání byly vybrány tyto metody komprese: *MPEG-2 Part 2*, *H.264*, *DivX Codec*, *Xvid*, *VC-1*, *Dirac* a *Theora*.

Záměr pro vytvoření této práce je příprava prvotního zmapování existujících metod video komprese pro další využití.

Kapitola číslo 2 stručně popisuje principy zmíněných kompresních metod. Na tuto kapitolu navazuje další s číslem 3, kde jsou shrnuty teoretické předpoklady sestavené na základě předešlého popisu metod. Kapitola 4 objasňuje metriky využívané k posuzování kvality a také použité softwarové nástroje. Popis použitých kodeků a testovacích dat je obsažen v kapitole 5. Výsledky objektivní části měření jsou uvedeny v kapitole 6 a výsledky subjektivní části shrnuje kapitola 7. Souhrnné zhodnocení obsahuje kapitola 8. Konečně kapitola s číslem 9 obsahuje závěr celé práce.

¹*MPEG home page* [online]. c2009, [cit. 2011-17-04]. Dostupné z: <<http://mpeg.chiariglione.org/>>.

²*Audiovisual and multimedia systems* [online]. 30.3.2011, [cit. 2011-17-04]. Dostupné z: <<http://www.itu.int/rec/T-REC-H/en>>.

2 Existující metody digitální video komprese

2.1 MPEG-2 Part 2

Jde o standard představený skupinou *Motion Pictures Experts Group*³ v roce 1995. Byl vyvinut jako nástupce MPEG-1, jež byl zaměřen na kvalitativní úroveň videa blízkou VHS záznamu. MPEG-2 Part 2 přináší oproti tomuto předchůdci nové možnosti kvality.

Taktéž vyřešil jeho největší nedostatky a to především:

- přinesl podporu a standardizaci pro kompresi prokládaného videa;
- zavedl možnost profilů a levelů pro různá nasazení;
- je navržen pro řádově vyšší datové toky a podporuje VBR (variable bit rate);
- podporuje i jiné formáty, než „4:2:0“.

MPEG-2 Part 2 patří v současnosti k nejrozšířenějším standardům a tvoří pomyslnou základní hranici kvality video enkodérů. MPEG-2 má majoritní zastoupení v DVD Video a SVCD nosičích a taktéž v digitálním televizním vysílání (např. rodina televizních standardů DVB⁴). Vyskytuje se v omezené míře i na nosičích Blu-ray Disc nebo HD DVD.

Digitální video signál (zdroj např. videokamera) v klasickém barevném modelu RGB (Red Green Blue) je po gama korekci kódován do modelu $Y'CbCr$, kde Y' představuje jasovou složku a Cb s Cr rozdílové chrominační složky vzhledem ke složce Y' . Někdy také bývá model $Y'CbCr$ označován jako YUV.

Důvodem převodu je fakt, lidské oko předně reaguje na změny jasové složky obrazu⁵. U RGB modelu jsou všechny složky rovnocenné, čímž vzniká redundance. Protože velikost složek Cb a Cr (chrominační) lze subsamplingem výrazně snížit bez zřetelného zhoršení kvality, čím se zmenší velikost souboru a datový tok.

Rodina formátů MPEG využívá nejčastěji subsampling ve formátu „4:2:0“⁶. Tudíž na první řádek 4 referenčních bodů – zvaných vzorky - připadají dva vzorky chrominačních Cb , Cr složek a žádný další přidáný vzorek Cb , Cr složek na druhý řádek.

³MPEG.ORG - MPEG Home [online]. c1998-2010, [cit. 2010-10-05]. Dostupné z: <<http://www.mpeg.org/>>.

⁴DVB - Digital Video Broadcasting - Standards & Technology [online]. c2003, [cit. 2010-10-05]. Dostupné z: <<http://www.dvb.org/technology/>>

⁵LIVINGSTONE, Margaret. *Vision and Art: The Biology of Seeing*. 2002. New York. ISBN 0-8109-0406-3.

⁶TUDOR, P.N. *MPEG-2 video compression* [online]. 17.8.1997, [cit. 2010-10-05]. Dostupné z: <http://www.bbc.co.uk/rd/pubs/papers/paper_14/paper_14.shtml>

Neboli Cb a Cr složky tvoří jen čtvrtinu počtu vzorků v poměru k počtu vzorků složky Y⁷.

Matice 8×8 Y', Cb nebo Cr vzorků se označuje jako *blok* a čtyři Y' bloky (16×16) s náležícími Cb, Cr bloky tvoří jeden *makroblok*. Konkrétní uspořádání makrobloku záleží na použitém formátu, např. při použití formátu „4:2:0“ připadá na čtyři Y' bloky, jeden Cr a jeden Cb blok. Vlastní snímky jsou tvořeny uspořádáním těchto makrobloků do více slice (v počtu jeden až všechny makrobloky snímku).

MPEG-2 specifikuje rozdělení snímků na tři kategorie: *I-snímky* (intra-coded frames), *P-snímky* (predictive-coded frames) a *B-snímky* (bidirectionally-predictive-coded frames).

Kategorizování těchto snímků je hlavním důvodem pro snížení datového toku pomocí predikce.

I-snímek: Jde o snímek, který neobsahuje žádnou informaci o zbývajících snímcích. Jeho obsah je tvořen informací vzniklou intra-frame kompresí. Využívá se jako klíčový snímek v GOP (Group Of Pictures). Využíván jako referenční při mezisnímkové kompresi. Jeho četnost je jedním z kritérií výsledné kvality video složky.

P-snímek: Obsahuje rozdílové informace od předchozího I-snímku nebo P-snímku. Může být využíván jako referenční pro další snímky. V makroblocích tohoto druhu snímku se využívá především dopřednou predikci. V tomto případě nese makroblok informaci v podobě jednoho pohybového vektoru. Ovšem každý z makrobloků může být využit stejně jako makroblok I-snímku – intra-frame zkomprimován. Obecně dosahují výrazně vyšších kompresních poměrů, než I-snímky.

B-snímek: Přechovává informace o rozdílu mezi ním, předchozím i následujícím I-snímkiem nebo P-snímkiem. Jeho makrobloky mohou být použity pro dopřednou, zpětnou, obousměrnou predikci (čili interpolaci) nebo intra-frame zkomprimovány. Dopředná a zpětná predikce ukládá jeden pohybový vektor. Metoda interpolace přiřazuje makrobloku dva pohybové vektory. Tento typ snímku nelze použít jako referenční. Vykazuje ze všech druhů snímků nejvyšší stupeň komprese⁸. Vlastní GOP je charakterizován hodnotami M a N, kde M je počet snímků mezi P-snímky (včetně P-snímku) a N značí celkový počet snímků v GOP. Skupiny GOP tvoří MPEG-2 sekvence. Ideální GOP vypadá např.: I B B P B B P B B I.

⁷MPEG-FAQ: *multimedia compression [1/9]* [online]. c2010, [cit. 2010-05-10]. Dostupné z: <<http://www.faqs.org/faqs/mpeg-faq/part1/>>

⁸LÍŠKA, Dušan. *Kódování pro DVB-T: MPEG-2, nebo MPEG-4 AVC? - DigiZone.cz* [online]. 26.6.2007, [cit. 2010-10-12]. Dostupné z: <<http://www.digizone.cz/clanky/kodovani-pro-dvb-t-mpeg-2-nebo-mpeg-4-avc/>>

Mezislímkové prediktivní kódování (*inter-frame*) P-snímků a B-snímků je založené na metodě detekce a predikce pohybu makrobloků. Základní myšlenka je porovnávání všech prvků, nalézajících se v makrobloku, vůči obsahu makrobloků specifikované množiny makrobloků zdrojového snímku. Při porovnávání makrobloků se specifikuje cenovou funkcí počet rozdílných bodů. Jako referenční makroblok je vybrán ten, s nejmenším počtem odlišných prvků makrobloku – dle srovnávacího kritéria - a je vytvořen příslušný pohybový vektor. Pro vyhledání souhlasného makrobloku se používá řada srovnávacích algoritmů (block matching), které mají za úkol minimalizaci cenové funkce (např. tříkrokový, logaritmický, konjugované vyhledávání, MPDC-LSA, ...)⁹. MPEG-2 pracuje s přesností predikce $\frac{1}{2}$ (HPEL) obrazového bodu¹⁰, dosahuje jí pomocí lineární interpolace sousedních bodů. Při nesplnění srovnávacího kritéria (nenalezení podobného makrobloku) se použije makroblok zdrojového snímku, čímž nedojde k žádné úspoře.

Jako další (*intra-frame*) snížení datového toku snímků – mezibodové redundance – slouží *dvourozměrná DCT-II* (Discrete Cosine Transform Type II), která se provede nad obrazovými bloky o velikosti 8×8 bodů a převede tak signálové hodnoty na frekvenční. Blok 8×8 s 8 bity na bod se tedy zakóduje do 8×8 matice spektrálních koeficientů.

K razantnímu zmenšení velikosti koeficientů je matice kvantizována dynamickými koeficienty (v závislosti na obsazení výstupním bufferu), kde se odstraní lidským smyslem méně viditelné, především vysoké, frekvence obrazu – dojde ke ztrátě informace. Takto dojde k využití prostorové korelace sousedních bodů.

Pro efektivní linearizaci koeficientů je použit diagonální „cik-cak“ („zigzag“ scan) algoritmus, kterým získáme nenulové koeficienty jako první v pořadí. Při použití prokládaných snímků je využita mírně pozměněná varianta tohoto algoritmu. Následně je aplikována bezztrátová RLE (Run Length Encoding). Závěrečnou fází *intra-frame* kódování je aplikace VLC (Variable-length Code) komprese¹¹.

Budoucí P-snímek či B-snímek přichází jako zdrojový po makroblocích na vstup. Zde je provedena detekce a kompenzace pohybu s vytvořením pohybových vektorů vůči referenčnímu snímku. Nalezený referenční makroblok je následně odečten od predikovaného makrobloku, výsledkem je predikční chyba (prediction error). Dále se chyba transformuje použitím 2D DCT-II, kvantizuje Q, jde do inverzní kvantizace Q^{-1} a odtud ještě po provedení inverzní 2D DCT⁻¹ sečtena s referenčním makrob-

⁹JAIN, A.K. *Fundamentals of Digital Image Processing*. New Jersey. 1989. 592 s. ISBN 0133361659.

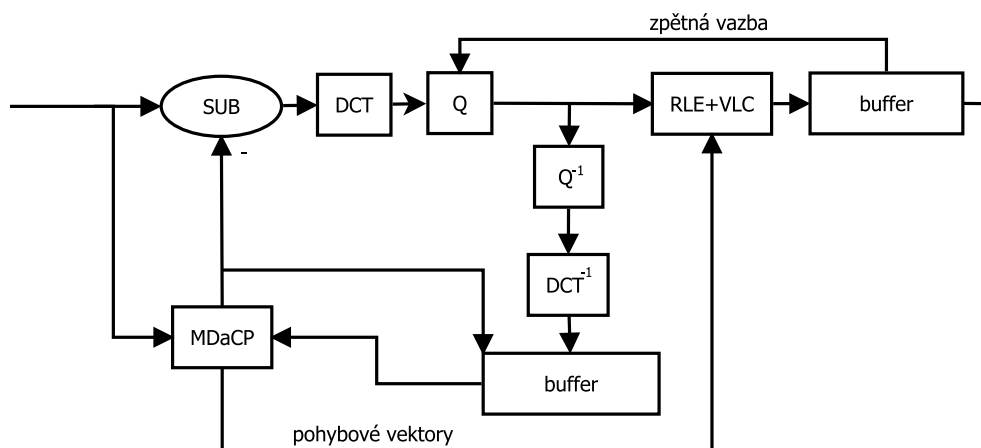
¹⁰LÍŠKA, Dušan. *Kódování ...*, s. 6.

¹¹TUDOR, P.N. *MPEG-2 video ...*, s. 5.

lokem. Tím dostaneme celistvý makroblok, který je uložen na snímkový buffer pro další využití. Kvantizovaná predikční chyba je RLE a VLC zakódována a s pohybovými vektory odeslána na výstupní buffer. Pohybový vektor je odečten typicky od vektoru příslušícího prvnímu makrobloku ve slice a následně RLE zakódován¹².

Pro budoucí I-snímek platí postup: 2D DCT-II transformace, Q , Q^{-1} , 2D DCT^{-1} , uložení na obrazový buffer a výstupní zakódování¹³, jediný rozdíl tvoří oddělené kódování DC a AC koeficientů¹⁴. Při enkódování musí vstupní soubor, stejně jako u MPEG-1, splňovat podmínku dělitelnosti obou rozměrů 16ti beze zbytku. Tímto se liší od drtivé většiny ostatních kodeků, kde stačí dělitelnost beze zbytku 4mi.

Zjednodušené schéma enkódování MPEG-2 Part2 je zobrazeno na obrázku 2.1.



Obrázek 2.1: Zjednodušené schéma enkódování MPEG-2 Part 2.

Nejběžnější profily a levely MPEG-2¹⁵ viz tabulka 2.1.

¹²Tamtéž, s. 5.

¹³Tamtéž, s. 5.

¹⁴SALOMON, D. - MOTTA, G. *Handbook of Data Compression*. 5th edition, 2010. New York. 1340 s. ISBN 978-1-84882-902-2.

¹⁵MULLEN, Steve. *Profiles and Levels* [online]. 1.7.2004, [cit. 2010-10-13]. Dostupné z: <http://digitalcontentproducer.com/videoencodvd/revfeat/video_profiles_levels/>.

rozlišení	max. framerate	formát	max. datový tok
176×144	15	4:2:0	96 Kb/s
352×288 nebo 320×240	15 nebo 24	4:2:0	384 Kb/s
352×288	30	4:2:0	4 Mb/s
720×480	30	4:2:0	50 Mb/s
1440×1080i nebo 1280×720p	30	4:2:0	60 Mb/s
1920×1080i nebo 1280×720p	30 nebo 60	4:2:0	80 Mb/s
720×480	30	4:2:2	50 Mb/s
1440×1080i nebo 1280×720p	30 nebo 60	4:2:2	80 Mb/s
1920×1080i nebo 1280×720p	30 nebo 60	4:2:2	300 Mb/s

Tabulka 2.1: Profily a levely MPEG-2 Part 2.

2.2 H.264/MPEG-4 Part 10

Někdy též nazývaný jako MPEG-4 AVC (Advanced Video Coding) je standard prvotně představený *Moving Picture Experts Group*¹⁶ a *Video Coding Experts Group*¹⁷ v roce 2003. Byl vícekrát modifikován pro zvýšení dosažitelné flexibility (např.: FExt, SVC¹⁸, MVC¹⁹, ...). Standard H.264 si klade mimo jiné za cíl razantně snížit datový tok při stejné, nebo lepší kvalitě, než-li umožňují jeho předchůdci – MPEG4 Part 2, MPEG-2 Part 2 nebo H.263.

Patří mezi nejrozšířenější standardy s velice širokým spektrem nasazení. Je využíván jako jeden z povinných ve standardech pro Blu-ray Disc a HD DVD. Taktéž je široce zastoupen v digitálním televizním vysílání, video-konferencích, streamovaných službách (např. YouTube.com od roku 2007²⁰), nebo v zabezpečovací technice. Jeho velice rozšířenou Open Source implementací je *x264*²¹ kodek.

H.264 se svými extenzemi přináší velice obsáhlý soubor profilů a levelů, které přináší značně odlišné možnosti nasazení. Základní jsou: Baseline, Main a Extended profil. FExt rozšíření přineslo velice populární High profil.

Základní filosofii převzal H.264 z rodiny MPEG a jejich možnosti značně rozšířil, ovšem hlavní částí zůstávají zachovány:

- mezisnímkové prediktivní kódování pro omezení dynamické redundance
- omezení mezibodové redundance tvořené korelací sousedních bodů
- predikční chyba tvořena rozdílem referenčního a predikovaného snímku
- omezení redundance predikční chyby
- výstupní bezztrátové kódování
- tvorba pohybových vektorů²².

¹⁶MPEG ..., s. 5.

¹⁷Video Coding Experts Group - Wikipedia, the free encyclopedia [online]. 18.5.2010, [cit. 2010-10-16]. Dostupné z: <http://en.wikipedia.org/wiki/Video_Coding_Experts_Group>.

¹⁸SCHWARZ, Heiko – WIEGAND, Thomas. IP Internet [online]. [cit. 2010-10-17]. Dostupné z: <http://ip.hhi.de/imagecom_G1/sauce/>.

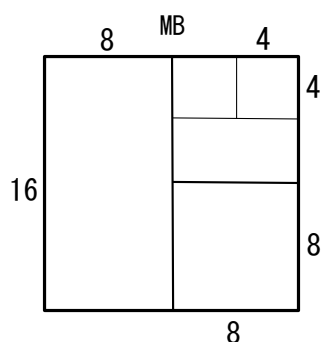
¹⁹SMOLIC, Aljoscha. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION [online]. leden 2008, [cit. 2010-10-17]. Dostupné z: <<http://mpeg.chiariglione.org/technologies/mpeg-4/mp04-mvc/index.htm>>.

²⁰YouTube API Blog: Flash and the HTML5 <video> tag [online]. 27.7.2010, [cit. 2010-10-17]. Dostupné z: <<http://apiblog.youtube.com/2010/06/flash-and-html5-tag.html>>.

²¹x264.nl [online]. [cit. 2010-10-17]. Dostupné z: <<http://x264.nl/>>.

²²H.264 video fastest mpeg 4 encoder high-definition DVD advanced video coding [online]. c2003-2009, [cit. 2010-10-17]. Dostupné z: <http://www.pixeltools.com/h264_paper.html>.

H.264 přináší nové možnosti dělení rozměrů jasových makrobloků a to: 16×16 , 16×8 , 8×16 , 8×8 , 8×4 , 4×8 a 4×4 referenčních bodů. 16×16 makroblok lze dále dělit na *makro-partition* o 16×16 , 16×8 , 8×16 nebo 8×8 referenčních bodech. Při 8×8 ji lze dále dělit na *sub-makroblok*: 8×8 , 8×4 , 4×8 a 4×4 referenčních bodů. Velikosti bloků chrominačních složek jsou úměrné použitému formátu. Toto dílčí dělení je celé využíváno při *VBSMC* (Variable Block-Size Motion Compensation), kde makroblok může být dělen na části, které jsou následně predikovány (viz obrázek 2.2). Každá taková podmnožina makrobloku má svůj vlastní pohybový vektor (maximum 32). Toto víceokrové dělení přináší možnost přesnější predikce jemných částí snímku²³.



Obrázek 2.2: Možnosti dělení makrobloku H.264.

Zpracovávaná přesnost je $\frac{1}{4}$ bodu při provádění predikce u jasového makrobloku. Využívá k tomu šestinásobný filtr a zaokrouhlení. Predikce pohybových vektorů, která čerpá z existující korelace mezi vektory sousedících bloků a samotným predikovaným vektorem, tak může přispět ke úspoře bitů díky kódování diferencí vektorů²⁴.

Výchozím modelem je u H.264 Y'CbCr ve formátu „4:2:0“, FExt přináší rozšíření o „4:0:0“, „4:2:2“, „4:4:4“, zavádí také nový model YCoCg (luma, chroma orange, chroma green) a jeho bezetrátovou variantu YCoCg-R, nebo umožní i použití RGB modelu²⁵.

Závažnou změnou oproti předchůdcům je zavedení více referenčních snímků pro kompenzaci pohybu než-li dvou. Maximální počet je stanoven na 16 snímků. Navíc, nově, i B-snímek může být použit jako referenční. Referenční snímky určené k predikci jsou udržovány v DPB (Decoded Picture Buffer) a enkodér i dekodér udržují

²³RICHARDSON, Iain E. G. *H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia*. 1st edition, 2003. Chichester. 320 s. ISBN 978-0-470-84837-1.

²⁴Tamtéž, s. 11.

²⁵SULLIVAN, Gary J. – TOPIWALA, Pankaj – LUTHRA, Ajay. *The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions*. [online]. srpen 2004, [cit. 2010-10-17]. Dostupné z: <<http://www.fastvdo.com/spie04/spie04-h2640verviewPaper.pdf>>.

jejich informace v seznamech list 0 a list 1. Větší počet referenčních snímků se pozitivně projevuje především u dlouhotrvajících monotóních posloupností snímků²⁶. Další změnou je použití vážené predikce (weighted prediction), která umožňuje při predikci započítávat relativní vzdálenost mezi referenčními snímky²⁷.

Jeden snímek může být, na rozdíl od MPEG-2, tvořen slice různých typů. Čili v jediném snímku mohou být I-slice, P-slice i B-slice.

Nově jsou slice děleny na jednotlivé typy, podobně jako snímky u MPEG-2 a to na: I-slice, P-slice, B-slice, SP-slice (Switching P) a SI-slice (Switching I). Přičemž poslední dva jmenované typy jsou dostupné pouze při použití Extended profilu.

SP-slice/SP-snímek je určen především pro snížení datového toku při použití přepínání mezi proudy nebo pro zprostředkování náhodného přístupu. Jeho výhodou je možnost predikce následujícího P-snímku proudu B, v okamžiku přepnutí z proudu A a do proudu B, z aktuálního SP-snímku proudu A a B. Jeho použitím odpadá vynucené zařazení I-snímku do proudu B. SI-slice/SI-snímek se liší od SP-slice tím, že využívá intra-predikce z referenčního snímku namísto inter-predikce u SP-snímku. Může být využit pro přepnutí do proudu bez nutné korelace s původním proudem²⁸.

Intra-frame kódování snímku se provádí nad bloky 4×4 , 8×8 nebo 16×16 bodů²⁹. Intra-predikce využívá korelace blízkých bloků a probíhá směrovou predikcí bloku ze sousedních, již zakódovaných bloků. Pro každou velikost použitého bloku lze využít množství různých predikčních metod³⁰. Pro 4×4 body je možno použít 9 směrových metod a pro 16×16 bodů metody 4. Tímto postupem by mělo být dosaženo větší snížení mezibodové redundance, než-li u pouhé AC predikce a následné transformace, jako např. MPEG-4 Part 2.

Pro transformaci je v tomto standardu využívána nová ortogonální celočíselná transformace (Exact-match Integer Spatial Block Transform), která přímo vychází z 2D DCT-II³¹ a nebo Hadmardova transformace. Použití celočíselné transformace odstraňuje rozdíly v zaokrouhlování při různých implementacích. Použití závisí na

²⁶Key Compression Tools for H.264 Video – Picture, Coding, Slices, Reference, Based, and Size [online]. c2010, [cit. 2010-10-17]. Dostupné z: <<http://encyclopedia.jrank.org/articles/pages/6801/Key-Compression-Tools-for-H-264-Video.html>>.

²⁷RIABTSEV, Shevach. [Mp4-tech] [MP4-Tech]Weighted Prediction in H.264 [online]. 6.8.2009, [cit. 2010-10-17]. Dostupné z: <<http://lists.mpegif.org/pipermail/mp4-tech/2009-September/009022.html>>.

²⁸RICHARDSON, Iain E. G. *H.264 and ...*, s. 11.

²⁹SRINIVASAN, Ravi. *H.264 High Profile: Codec for Broadcast & Professional Video Application* [online]. c2010, [cit. 2010-10-18]. Dostupné z: <<http://www.design-reuse.com/articles/20776/h-264-high-profile-codec-video.html>>.

³⁰AVC-Intra (H.264 Intra) Compression [online]. 7.8.2007, [cit. 2010-10-18]. Dostupné z: <ftp://ftp.panasonic.com/pub/Panasonic/Drivers/PBTS/papers/WP_AVC-Intra.pdf>.

³¹H.264/MPEG-4 AVC – Wikipedia, the free encyclopedia [online]. 15.10.2010, [cit. 2010-10-18]. Dostupné z: <http://en.wikipedia.org/wiki/H.264/MPEG-4_AVC>.

typu vstupních bloků. H.264 využívá Hadmardovu transformaci pro bloky 4×4 DC koeficientů složky Y' při použití 16×16 intra predikce, kterou aplikuje až po celočíselném transformování všech sub-makrobloků 4×4 . Taktéž Hadmardovu transformaci pro 2×2 DC koeficientů chrominačních složek, které byly před tím již transformovány použitím celočíselné transformace. Pouze celočíselná transformace je využita pro všechny ostatní 4×4 sub-makrobloky po pohybové predikci, nebo při intra kódování³². Pro High profil lze využít i transformaci nad bloky o rozměru 8×8 bodů (použití pro oblasti snímku s vysokou korelací okolních bodů a 4×4 pro ostatní oblasti – možná redukce „ringing“ artefaktů³³). High profil přináší kvantizační škálovací matice, které na rozdíl od uniformních, přinášejí lepší rozložení koeficientů mezi frekvence a měli by tak zlepšovat psychovizuální dojem výsledného obrazu. Vztah velikosti kroků kvantizačních koeficientů a kvantizačního parametru je logaritmický, na rozdíl od lineárního u MPEG-2 Part 2.

Oproti předchůdcům taktéž zavádí povinnou integraci deblokovacího filtru (in-loop deblocking filter), který je včleněn do dekódovací i enkódovací smyčky. To způsobuje, že s jeho výsledky se již promítají do obsahu referenčních makrobloků určených pro predikci. Jeho přítomnost má důsledky v lepší subjektivní kvalitě výstupního videa.

Jako výstupní bezztrátové kódování je u H.264 při Baseline nebo Extended profilu použito *CAVLC* (Context-Adaptive Variable-Length Coding), pro Main a vyšší je použito účinnější, ovšem výpočetně náročnější, *CABAC* (Context-Adaptive Binary Arithmetic Coding). V případě, že nejsou známy o kódovaných datech žádné statistické údaje je použito Exponential-Golomb kódování.

H.264 také přináší nové prvky vylepšující odolnost dat vůči chybám. Například metody FMO (Flexible Macroblock Ordering) a ASO (Arbitrary Slice Ordering). Principiálně tyto metody rozdělí podle různých schémat všechny slice ve snímku do více slice-groups. Vhodnou volbou rozmístění slice-groups lze dosáhnout stavu, kdy je možné opravit informaci nesenou ztraceným slice interpolací z okolních slice.

Standard celkově obsahuje 16 jednotlivých levelů a 17 profilů³⁴. Přehled čtyř nejběžnějších profilů obsažených v H.264 je obsažen v tabulce 2.2.

³²RICHARDSON, Iain E. G. *H.264 and ...*, s. 11.

³³*Key Compression ...*, s. 12.

³⁴SWAIN, Robert. *H.264 Profiles and Levels* / robert.swain [online]. [cit. 2010-10-18]. Dostupné z: <<http://rob.opendot.cl/index.php/useful-stuff/h264-profiles-and-levels/>>.

profil	Baseline	Extended	Main	High
B-slice	Ne	Ano	Ano	Ano
SI a SP slice	Ne	Ano	Ne	Ne
(FMO) Flexible Macroblock Ordering	Ano	Ano	Ne	Ne
(ASO) Arbitrary Slice Ordering	Ano	Ano	Ne	Ne
CABAC výstupní kódování	Ne	Ne	Ano	Ano
8×8 – 4×4 adaptabilní transformace	Ne	Ne	Ne	Ano
kvantizační škálující matice	Ne	Ne	Ne	Ano
černobílý formát (4:0:0)	Ne	Ne	Ne	Ano
výchozí použitý formát	4:2:0	4:2:0	4:2:0	4:2:0
výchozí barevná hloubka	8	8	8	8

Tabulka 2.2: H.264 profily a jejich vlastnosti.

2.3 DivX Codec

Známý pod názvem DivX je jeden ze dvou kodeků obsažených v balíku DivX produkovaného společností *DivX, Inc.*³⁵ Jeho současná podoba (od roku 2001) je implementací MPEG-4 ASP profilu *Moving Picture Experts Group*³⁶ standardu MPEG-4 Part 2. Historický předchůdce je kodek *DivX ;-)*, který byl postaven na základě specifikace MPEG-4 Version 3.

Hlavní motivací vývoje tohoto kodeku byla potřeba zmenšit datový objem DVD Video na úroveň jednoho či dvou CD-R nosičů (650 – 1400 MiB) při nejvyšší možné zachované kvalitě obrazu.

DivX Codec je spolu s Xvid jedním z nejrozšířenějších kodeků, především díky masivnímu rozšíření nelegálních kopií DVD Video nosičů označované jako „ripy“. Důležitou součástí jeho penetrace trhu je také množství HW, který podporuje jeho dekódování. Jde především o implementace v DVD přehrávačích, Blu-ray přehrávačích a herních konzolách (např. Sony Playstation 3³⁷).

Přesný popis funkčnosti proprietárního DivX Codec kodeku tvoří know-how společnosti DivX, Inc. (např. zahrnutá psychovizuální vylepšení obrazu).

Dle standardu MPEG-4 ASP je sekvence snímků označována jako video objekty – VOS (Visual Object Sequence) a skládá se z jednotlivých VO (Video Objects), které v zásadě nemusí mít pravidelný tvar a nemusí se nalézat jen v jednom „snímku“. Pro jeden statický snímek streamu je použito označení VOP (Video Object Pane)³⁸.

DivX nabízí tyto hlavní vlastnosti, které vyhovují standardu MPEG-4 ASP³⁹:

- *B-VOP* – obdobná funkce jako B-snímky MPEG-2 Part 2;
- podporu prokládaného videa – též příbuzné MPEG-2 Part 2;
- *QPEL* (Quarter Pixel Motion Search Precision) – přesnost na 1/4 bodu při provádění pohybové predikce, dosahuje se interpolací z okolních bodů;
- možnost použití různých vlastních (MPEG custom matrix) kvantizačních matic;

³⁵ *DivX Company / DivX.com* [online]. c2010, [cit. 2010-10-23]. Dostupné z: <<http://www.divx.com/en/company>>.

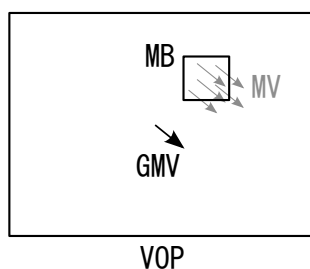
³⁶ *MPEG . . .*, s. 5.

³⁷ *Firmware Update (v2.10) – PlayStation Blog* [online]. 17.12.2007, [cit. 2010-10-23]. Dostupné z: <<http://blog.us.playstation.com/2007/12/17/firmware-update-v210/>>.

³⁸ RICHARDSON, Iain E. G. *H.264 and . . .*, s. 11.

³⁹ *MPEG Industry Forum* [online]. c2007, [cit. 2010-10-25]. Dostupné z: <<http://www.m4if.org/resources/profiles/visualtools.php>>.

- adaptivní kvantizace – možnost použití různých kvantizačních parametrů i na úrovni makrobloků (např. pro tmavé části snímku). Známe též jako *lumi masking*;
- *GMC* (Global Motion Compensation) – detekce pohybu velkých částí obrazu; v případě zjištění stejného pohybu velkých částí je použit malý počet globálních pohybových vektorů (GMV) a v dekodéru se podle nich bilineární interpolací dekódují pohybové vektory (MV) popisující pohyb jednotlivých makrobloků. DivX Codec využívá pouze jediný takovýto GMV⁴⁰ (viz obrázek 2.3).



Obrázek 2.3: Ilustrace využití *GMC* pomocí globálního vektoru v DivX Codec.

Proces kódování videa je velice podobný celé rodině MPEG standardů. Podobně jako v případě H.264 lze Y' makrobloky 16×16 dělit na bloky o 8×8 bodech (4MV)⁴¹, čímž dostáváme čtyři pohybové vektory na makroblok. Toto jemnější dělení je výhodné pro snížení datového toku, díky většímu počtu shodných bloků. Výchozí barevný model je u DivX použit klasický Y'CbCr se složkami ve formátu „4:2:0“.

Pro intra-kódování a reziduální kódování je použita 2D DCT-II transformace nad bloky 8×8 bodů⁴². Stejně jako H.264 obsahuje DivX Codec metodu intra-predikce. Ta ovšem predikuje pouze obsah DC koeficientu nebo první řádky a sloupce AC koeficientů ze sousedních bloků⁴³. Následná kvantizace koeficientů je provedena s možností použití H.263, H.263 Optimized, MPEG-2 nebo vlastní kvantizační matice. Výstup je linearizován diagonálním „cik-cak“ algoritmem, RLE a VLC (stejně jako v MPEG-2 Part 2) bezeztrátově zakódován⁴⁴.

⁴⁰*XviD FAQ: Explanation of many advanced options of XviD codec* [online]. c2001-2010, [cit. 2010-10-25]. Dostupné z: <http://www.gromkov.com/faq/conversion/xvid_options.html>.

⁴¹LÍŠKA, Dušan. *Kódování . . .*, s. 6.

⁴²*MPEG-4 description* [online]. březen 2002, [cit. 2010-10-25]. Dostupné z: <<http://mpeg.chiariglione.org/standards/mpeg-4/mpeg-4.htm>>.

⁴³RICHARDSON, Iain E. G. *H.264 and . . .*, s. 11.

⁴⁴Tamtéž, s. 11.

B-VOP poskytují v DivX kodeku dvě volby a to: ASC (Adaptive Single Consecutive) a AMC (Adaptive Multi Consecutive)⁴⁵. ASC poskytuje omezení počtu B-VOP v rychle se střídajících scénách a tím zlepšení kompresního poměr. Naproti tomu AMC by měl poskytnout plynulejší střídání za cenu většího datového objemu.

DivX Codec taktéž poskytuje kontrolu nad klíčovými snímky (I-VOP) díky volbám: *Max Keyframe Interval* a *Keyframe Threshold*⁴⁶. První z nich definuje interval mezi klíčovými snímky a druhý ovládá vkládání prázdných snímků, pokud počet makrobloků nenalezených motion search algoritmem ve snímku překročí zadaný limit.

Pro inter-predikci je využíván základní block matching nad makrobloky o rozměru 16×16 , nebo bloky 8×8 bodů. Referenční blok je po nalezení odečten od vstupního bloku. Tato predikční chyba je následně 2D DCT-II transformována, kvantifikována, v reverzní formě uložena do bufferu a výstupně zakódována⁴⁷.

Profily používané DivX Codec specifikací⁴⁸ jsou zobrazeny v tabulce 2.3.

profil	High Definition	Home Theater	Portable
max. rozlišení	1280×720×30 fps	720×480×30 fps	352×240×30 fps
	720×576×25 fps		
průměrný datový tok	8000 Kb/s	4000 Kb/s	768 Kb/s
max. datový tok	32000 Kb/s	8000 Kb/s	4000 Kb/s
B-VOP	Ano	Ano	Ano
podpora prokádání	Ano	Ano	Ne

Tabulka 2.3: Používané profily DivX Codec.

⁴⁵*DivX 6 Setup Guide - Page 3 of 4 | Articles - (DivX/XviD) Digital Digest* [online]. 20.6.2006 [cit. 2010-10-25]. Dostupné z: <http://www.divx-digest.com/articles/divx6_setup_page3.html>.

⁴⁶Tamtéž, s. 17.

⁴⁷RICHARDSON, Iain E. G. *H.264 and ...*, s. 11.

⁴⁸*DivX Labs* [online]. c2010, [cit. 2010-10-25]. Dostupné z: <<http://labs.divx.com/>>.

2.4 Xvid

Kodek Xvid je open-source implementací MPEG-4 SP a MPEG-4 ASP profilů *Moving Picture Experts Group*⁴⁹ standardu MPEG-4 Part 2. V současnosti je šířen pod licencí GNU GPL⁵⁰.

Vznik Xvid kodeku lze odvozovat od projektu OpenDivX a jeho vznik jako reakci na odstěpení a následný vznik proprietárního kodeku DivX.

Díky jeho kompatibilitě se standardem MPEG-4 Part 2 lze ve většině případů dekodovat standardní Xvid video na HW zařízeních, která podporují dekodování DivX Codec.

Využívá vlastností standardu MPEG-4 Part 2⁵¹: B-VOP, GMC, QPEL, výběr pro použití H.263, MPEG nebo vlastní kvantizační. Dále podporu adaptivní kvantizace (lumi masking), mřížkové kvantizace, která optimalizuje 2D DCT-II koeficienty pro efektivnější výstupní kompresi matice a chroma pohybovou predikci (Chroma ME), která dovolí zahrnout při pohybové predikci nejen Y' složku, ale i CbCr složky⁵².

Kodek využívá řadu profilů a levelů, oboje kopírují standard MPEG-4 Part 2 – SP, ASP a ARTS. Základní princip intra i inter predikce se neliší od DivX Codec. Odlišnosti ve využitelných vlastnostech obou kodeků se mohou měnit s každou novou verzí.

Nejpodstatnější odlišnosti Xvid od DivX Codec⁵³:

- podpora až tří GMV při GMC;
- cartoon mode – vylepšení kvality obrazu při monotónních oblastech s ostrými přechody;
- větší délka B-VOP intervalu (max. 5).

Základní profily používané kodekem Xvid⁵⁴ viz tabulka 2.4.

⁴⁹ *MPEG . . .*, s. 5.

⁵⁰ *Xvid.org: FAQ* [online]. c2010, [cit. 2010-10-26]. Dostupné z: <<http://www.xvid.org/FAQ.42.0.html>>.

⁵¹ *MPEG-4 description* [online]. březen 2002, [cit. 2010-10-26]. Dostupné z: <<http://mpeg.chiariglione.org/standards/mpeg-4/mpeg-4.htm>>.

⁵² *XviD FAQ: Explanation of many advanced options of XviD codec* [online]. c2001-2010, [cit. 2010-10-27]. Dostupné z: <http://www.gromkov.com/faq/conversion/xvid_options.html>.

⁵³ *Difference in Features "Xvid Vs DivX"[Archive] - Doom9's Forum* [online]. 12.12.2005, [cit. 2010-10-28]. Dostupné z: <<http://forum.doom9.org/archive/index.php/t-103943.html>>.

⁵⁴ *XvidSolutions.com: Product Certification* [online]. c2010, [cit. 2010-10-28]. Dostupné z: <<http://www.xvidsolutions.com/Product-Certification.36.0.html>>.

profil	Xvid Mobile	Xvid Home	Xvid Highdef
max. rozlišení	352×240×30fps	720×480×30fps	1280×720×30fps
	352×288×25fps	720×576×25fps	
max. datový tok	1 334 850 b/s	4 854 000 b/s	9 708 400 b/s
max. B-VOP	5	5	5
QPEL	Ano	Ano	Ano
adaptivní kvantizace	Ano	Ano	Ano
mód 4MV	Ano	Ano	Ano
podpora prokládání	Ne	Ano	Ano

Tabulka 2.4: Profily používané Xvid.

2.5 WMV9/VC-1

V roce 2006 byl přijat skupinou *SMPTE* (The Society of Motion Picture and Television Engineers⁵⁵) jako otevřený standard VC-1, pod označením 421M⁵⁶. WMV 9 představuje implementaci z rodiny *Windows Media Video* vyvinutý firmou Microsoft.

VC-1 tvoří jeden ze tří povinných formátů ve standardu Blu-ray Disc a HD DVD, taktéž je prosazován jako hlavní video formát konzole Microsoft Xbox 360⁵⁷. Je tudíž masivně zastoupen i v HW dekódovacích zařízeních, jako jsou např. přehrávače médií obou zmíněných standardů.

V současnosti (2010) představuje tento kodek hlavního konkurenta pro H.264/MPEG-4 Part 10.

Implementace tohoto standardu jsou WMV3, který představuje Simple a Main profil *VC-1* a *WVC1*, který vyhovuje profilu Advanced⁵⁸.

Základní princip komprimace tvoří tradiční predikce pohybu prováděná po blocích a omezení redundance sousedních bodů. VC-1 zavádí v tomto postupu řadu inovací.

Je zde využito pokročilého dělení makrobloku na menší části: makro-partition, bloky a sub-makrobloky. Jako výchozí formát je stanoven „4:2:0“ modelu Y’CbCr.

Ke standardním I-snímům, P-snímům a B-snímům (v případě Main a Advanced profilu), které se funkčně podobají MPEG-4 Part2, jsou zařazeny *S-snímek* (skipped) a *BI-snímek*.

S-snímky se využívají v případech, kdy je referenční snímek téměř totožný s referenčním – obsah S-snímku je nahrazen obsahem referenčního⁵⁹.

BI-snímky jsou náhradou B-snímku při scénách, kdy jinak tradiční B-snímky obsahují velké množství I-makrobloků. BI-snímky přinášejí dodatečnou úsporu datového toku v těchto specifických situacích⁶⁰.

⁵⁵*Home* – *SMPTE.org* [online]. c2010, [cit. 2010-11-8]. Dostupné z: <<http://www.smpete.org/home/>>.

⁵⁶TOMKO, Heather. *SMPTE Releases VC-1 Standard* - *SMPTE.org* [online]. 3.4.2006, [cit. 2010-11-8]. Dostupné z: <http://www.smpete.org/news/pr/view?item_key=a135f13b173a982bb71f1cd3ee4403671fcf2057>.

⁵⁷*VC-1* – *Wikipedia, the free encyclopedia* [online]. 29.10.2010, [cit. 2010-11-8]. Dostupné z: <<http://en.wikipedia.org/wiki/VC-1>>.

⁵⁸LOOMIS, J. – WASSON, M. *VC-1 Technical Overview* [online]. říjen 2007, [cit. 2010-11-8]. Dostupné z: <<http://www.microsoft.com/windows/windowsmedia/howto/articles/vc1techoverview.aspx>>.

⁵⁹SALOMON, D. – MOTTA, G. *Handbook of Data Compression. 5th edition*, 2010. 1340 s. ISBN 978-1-84882-902-2.

⁶⁰Tamtéž, s. 20.

VC-1 využívá adaptivní transformace (Adaptive Block-Size Transform), která adaptivně nastavuje velikost bloků, nad kterými se provádí transformace. Použité rozměry jsou: 8×8 , 8×4 , 4×8 nebo 4×4 . Např. dvakrát 4×8 , nikoli 4×8 a dvakrát 4×4 . Výsledkem by mělo být lepší potlačení „ringing“ efektu⁶¹.

Pro vlastní transformaci je použita 16b celočíselná transformace, která je opět příbuzná s klasickou 2D DCT-II transformací.

Pro pohybovou predikci a kompenzaci pohybu (motion compensation) při interpolaci je volena velikost bloků 16×16 nebo 8×8 bodů, s přesností $\frac{1}{4}$ bodu (QPPEL). Tímto je dosaženo čtyř pohybových vektorů na makroblok. Pro interpolování obsahu bodů používá čtyřnásobný bikubický filtr a dvojnásobný bilinerání.

Obdobně jako v případě MPEG-4 Part 10 zavádí VC-1 pro potlačení artefaktů, vznikajících na přechodech mezi jednotlivými makrobloky deblokovací filtr.

Standard VC-1 též zavádí metody pro zmenšení datového toku při specifických scénách, kde jsou obecně dosahovány horší kompresní poměry. Jde zejména o scény jako je zatmívačka, nebo jiná postupná globální změna jasu. Souhrnné označení je: *Fading Compensation*.

Díky nehomogenitě makrobloků v jednom snímku je zavedena rozdílová kvantizace (DQuant – Differential Quantization). Využívá různé kvantizační koeficienty pro různé makrobloky v rámci jediného snímku⁶². Jde o obdobu rozšířených škálování matic.

Proces intra-kódování u VC-1 začíná transformací bloku o adaptivně zvolené velikosti za použití zmíněné celočíselné transformace. Koeficienty jsou následně kvantizovány a volitelně je použita DC/AC predikce – principiálně velmi podobná MPEG-4 Part 2⁶³.

Dále jsou koeficienty linearizovány pomocí jedné z variant „cik-cak“ algoritmu, které se liší v závislosti na použitém bloku a dosahují tak vyšší efektivity. Na závěr je použita RLE a VLC bezztrátová komprese. VC-1 na rozdíl od H.264 nevyužívá místo VLC výpočetně náročnou CABAC kompresi.

Inter-kódování probíhá podle schématu, který je částečně příbuzné k MPEG-4

⁶¹LOOMIS, J. – WASSON, M. *VC-1* . . . , s. 21.

⁶²SALOMON, D. – MOTTA, G. *Handbook of . . .* , s. 20.

⁶³Tamtéž, s. 20.

Part 10. V referenčním snímku (snímcích) je predikován referenční blok, následně je predikován pohybový vektor ze sousedních bloků, přenáší s pouze RLE zakódovaný rozdíl predikovaného a referenčního vektoru. Referenční blok se odečte od predikovaného a vznikne tak predikční chyba. Ta se následně transformuje, kvantizuje, linearizuje a na závěr RLE a VLC bezztrátově komprimuje.

Nejpoužívanější profily a levely VC-1⁶⁴ jsou obsaženy v tabulce 2.5.

profil	level	max. datový tok	rozlišení
Main	Low	2 Mb/s	320×240×24fps
Main	Medium	10 Mb/s	720×576×25fps
Main	High	20 Mb/s	1920×1080×30fps
Advanced	L0	2 Mb/s	352×288×30fps
Advanced	L1	10 Mb/s	720×576×25fps
Advanced	L2	20 Mb/s	1280×720×30fps
Advanced	L3	45 Mb/s	1920×1080×30fps
Advanced	L4	135 Mb/s	1920×1080×60fps

Tabulka 2.5: Profily používané VC-1.

⁶⁴LOOMIS, J. – WASSON, M. *VC-1...*, s. 21.

2.6 Dirac

Jde o velice mladou rodinu otevřených kodeků, které byly vyvinuty společností *BBC Research*⁶⁵. Její možné uplatnění je různorodé, od využití v mobilních zařízeních až po nejvyšší rozlišení používaná v digitální kinematografii (D-cinema). Především však nasazení sahá do streamovaných médií, DTV a archivace audiovizuálních dat⁶⁶. Výstupní kvalita poskytovaná kodeky Dirac by měla být srovnatelná se standardem H.264⁶⁷.

High-end zástupce z této rodiny označovaný jako *Dirac Pro* byl přijat skupinou SMPTE (The Society of Motion Picture and Television Engineers) na konci roku 2009 jako otevřený standard VC-2⁶⁸. Dirac Pro byl úspěšně nasazen společností BBC při vysílání LOH 2008 v Pekingu⁶⁹.

Softwarové implementace tohoto kodeku jsou: *Schrödinger* a *Dirac* později přejmenovaný na *dirac-research*⁷⁰.

Základní funkční principy Dirac kodeku jsou podobné klasickým kodekům – např. rodině MPEG. Na rozdíl od majority video kodeků, které využívají diskrétní 2D kosinovou transformaci typu II, využívá Dirac ke stejným účelům *diskrétní waveletovou transformaci* (DWT)⁷¹. Tato transformace může využívat různě výpočetně náročných filtrů a dosahovat tak různého poměru ke kvalitě⁷². Stejnou transformaci využívá i standard JPEG2000.

Kvantizační proces je prováděn pomocí uniformní kvantizace s mrtvou zónou (Uniform Dead-Zone Quantisation)⁷³.

Dirac využívá *YCoCg* barevný model, který se skládá z jasové složky a oranžové

⁶⁵ *BBC RD - Homepage* [online]. 13.10.2009, [cit. 2010-11-24]. Dostupné z: <<http://www.bbc.co.uk/rd/>>.

⁶⁶ *BBC R&D - Dirac - Uses of Dirac* [online]. prosinec 2006, [cit. 2010-11-24]. Dostupné z: <<http://www.bbc.co.uk/rd/projects/dirac/uses.shtml>>.

⁶⁷ *BBC R&D ...*, s. 23 .

⁶⁸ *SMITH, Michael. Advancements in Compression and Transcoding: 2008 and Beyond - SMPTE.org* [online]. c2010, [cit. 2010-11-24]. Dostupné z: <http://www.smpete.org/events/smpete_annual_tech/schedule/06wedspm1/>.

⁶⁹ *FAQ - Dirac Wiki* [online]. 2008, [cit. 2010-11-24]. Dostupné z: <<http://diracvideo.org/wiki/index.php/FAQ>>.

⁷⁰ *BBC R&D - Dirac - Implementation behind Dirac* [online]. duben 2007, [cit. 2010-11-24]. Dostupné z: <<http://www.bbc.co.uk/rd/projects/dirac/implementations.shtml>>.

⁷¹ BORER, T. – DAVIES, T. - SURAPARAJU, A. *BBC R&D White paper WHP 124 - Dirac video compression* [online]. září 2005, [cit. 2010-11-25]. Dostupné z: <<http://www.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP124.pdf>>.

⁷² DAVIES, Thomas. *Dirac: Wavelet transform* [online]. 6.2.2008, [cit. 2010-11-26]. Dostupné z: <http://dirac.sourceforge.net/documentation/algorithm/algorithm/wlt_transform.xht>.

⁷³ DAVIES, Thomas. *Dirac: Quantisation* [online]. 6.2.2008, [cit. 2010-11-26]. Dostupné z: <<http://dirac.sourceforge.net/documentation/algorithm/algorithm/quantisation.xht>>.

se zelenou chrominační složkou (luma, chroma orange, chroma green). Jeho předností je zejména možnost bezztrátového převodu mezi ním a RGB⁷⁴. Tento model takéž dosahuje lepších výsledků než Y'CbCr v efektivitě a jeho převod do RGB je výpočetně méně náročný⁷⁵. Standardní Y'CbCr model je standardem Dirac též podporován.

Typy snímků se typicky dělí na intra-snímky a inter-snímky. Oba dva snímky mohou být použity jako referenční, maximální počet referenčních snímků pro predikci jednoho inter-snímku jsou dva snímky⁷⁶. Intra-snímky se označují klasicky jako I-snímky, pro P-snímky je zavedeno označení L1-snímek a pro B-snímky L2-snímek. Taktéž podporuje skipped snímky⁷⁷ – s obdobným účelem jako u VC-1.

Snímek je dělen na makrobloky. Tyto makrobloky jsou zde tvořeny navzájem se překrývajícími bloky, které nemají standardem definovanou velikost, v uspořádání 4x4. Vzájemné překrytí bloků by mělo eliminovat typický „blocking“ efekt. Tyto makrobloky mohou být používány v konfiguracích: 4×84, 2×82, nebo 1×81. Což ústí v jeden, čtyři, nebo šestnáct pohybových vektorů na makroblok. Typ konfigurace je určován adaptivně. Maximální přesnost predikce dosahuje 1/8 bodu⁷⁸. Při pohybové kompenzaci Dirac využívá klasických pohybových vektorů, které přísluší jednotlivým blokům. Taktéž, podobně jako MPEG-4 ASP, může využívat i globální pohybové vektory v GMC (Global Motion Compensation). Podporované GMC akce jsou: zoom, rotace, panoramatické záběry a náklon⁷⁹.

Při intra-kódování makrobloku se využívá DC predikce, jinak je postup stejný jako u klasických kodeků.

K linearizaci koeficientů je využíván klasický diagonální „cik-cak“ algoritmus. Jako výstupní bezztrátové kódování motion vektorů i linearizovaných dat využívá Dirac exp-Golomb a aritmetické kódování⁸⁰.

Kodek Dirac Pro je určen a navržen především pro profesionální použití. Disponuje

⁷⁴Dirac Specification [online]. 23.9.2008, [cit. 2010-11-26]. Dostupné z: <<http://diracvideo.org/download/specification/dirac-spec-latest.pdf>>.

⁷⁵MALVAR, Henrique – SULLIVAN, Gary. *Transform, Scaling & Color Space Impact of Professional Extensions* [online]. 19.5.2003, [cit. 2010-11-26]. Dostupné z: <http://wftp3.itu.int/av-arch/jvt-site/2003_05_Geneva/JVT-H031.doc>.

⁷⁶BORER, T. - DAVIES, T. - SURAPARAJU, A. *BBC R&D ...*, s. 21.

⁷⁷DAVIES, Thomas. *Dirac: Temporal prediction structures* [online]. 6.2.2008, [cit. 2010-11-26]. Dostupné z: <http://dirac.sourceforge.net/documentation/algorithm/algorithm/temporal_prediction_structures.htm>.

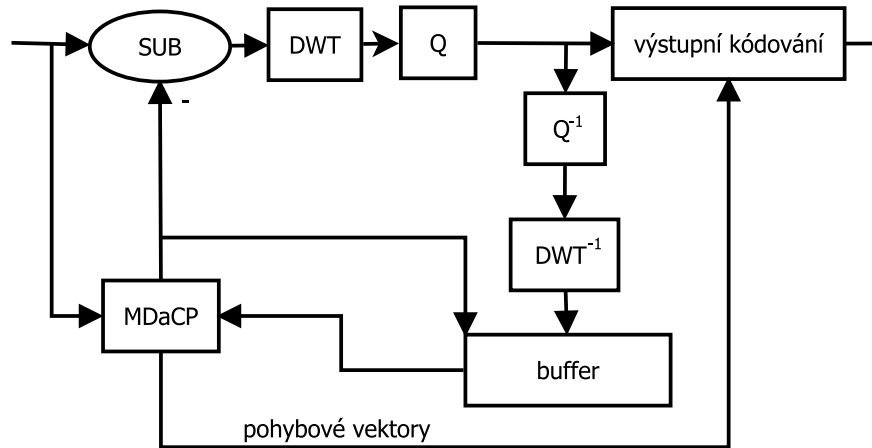
⁷⁸BORER, T. - DAVIES, T. - SURAPARAJU, A. *BBC R&D ...*, s. 21.

⁷⁹DIRAC *The video compression family using open technology* [online]. 2.4.2007, [cit. 2010-11-26]. Dostupné z: <<http://www.bbc.co.uk/rd/projects/dirac/press-downloads/dirac-lores.pdf>>.

⁸⁰Tamtéž, s. 24.

pouze mechanismem intra-kódování – pro snadnou editaci a využívá jen exp-Golomb jako výstupní kódování⁸¹.

Zjednodušené schéma enkódování dle standardu Dirac⁸² je zobrazena na obrázku 2.4.



Obrázek 2.4: Zjednodušené schéma enkódování Dirac.

⁸¹Tamtéž, s. 24.

⁸²*Dirac Specification* [online]. 23.9.2008, [cit. 2010-11-26]. Dostupné z: <<http://diracvideo.org/download/specification/dirac-spec-latest.pdf>>.

2.7 Theora

Jde o otevřený video kodek vyvíjený společností *Xiph.org Foundation*⁸³ na základě kodeku VP3 vytvořeného *On2 Technologies*⁸⁴. Specifikace jeho formátu byla dokončena v roce 2004. Jeho cílem je především poskytnout otevřenou alternativu ke standardu MPEG-4⁸⁵.

Nasazení Theora kodeku je variabilní od mobilních zařízení, přes online služby až po použití v HD zařízeních. Díky podpoře streamovaného nasazení byla Theora zařazena jako jeden z podporovaných video kodeků do specifikace standardu HTML5 (W3C HTML5)⁸⁶. Taktéž byl Theora kodek vybrán jako jediný přípustný video kodek pro video soubory využívané společností Wikimedia⁸⁷.

Jeho implementací je *libtheora*, která je poskytována a vyvíjena taktéž prostřednictvím *Xiph.org Foundation*.

Theora, tak jako mnohé další kodeky, využívá pro omezení dynamické a mezibodové redundance klasické postupy, typické například pro rodinu standardů MPEG.

Jako základní barevný model zde slouží standardní Y'CbCr s formáty „4:2:0“, „4:2:2“ a „4:4:4“⁸⁸. Druhy snímků jsou v tomto kodeku omezeny na dva. Ty jsou označovány jako Inter snímky a Intra snímky. První typ svým určením odpovídá klasickým P-snímům a druhý pak I-snímům⁸⁹. Theora tedy neobsahuje ekvivalentní typ snímku k B-snímům z rodiny MPEG.

Velikost jednoho obrazového bloku Theora stanovuje na 8×8 obrazového bodu. Vlastní snímky jsou tvořeny super-bloky, které jsou složeny z bloků v uspořádání 4x4 bloku. Makroblok o rozměrech 16x16 bodů je tvořen bloky uspořádaných do matice 2x2. Chrominační složky jsou zde poté uspořádány dle použitého formátu⁹⁰.

⁸³ *Xiph.org* [online]. c2010, [cit. 2010-11-30]. Dostupné z: <<http://xiph.org/>>.

⁸⁴ *Home / On2 Technologies* [online]. c2010, [cit. 2011-2-28]. Dostupné z: <<http://www.on2.com/>>.

⁸⁵ *Theora.org :: FAQ - Theora, video for everyone* [online]. c2009, [cit. 2010-11-30]. Dostupné z: <<http://www.theora.org/faq/>>.

⁸⁶ HICKSON, Ian. *HTML5* [online]. 16.11.2010, [cit. 2010-11-30]. Dostupné z: <<http://dev.w3.org/html5/spec/Overview.html#video>>.

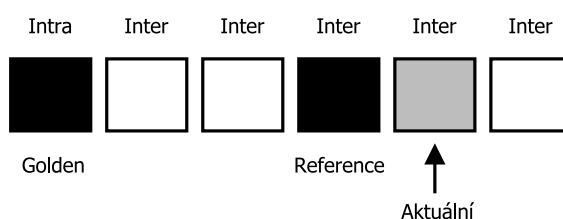
⁸⁷ *Commons:File types - Wikimedia Commons* [online]. 20.11.2010, [cit. 2010-11-30]. Dostupné z: <http://commons.wikimedia.org/wiki/Commons:File_types>.

⁸⁸ *Theora - XiphWiki* [online]. 7.7.2010, [cit. 2010-11-30]. Dostupné z: <<http://wiki.xiph.org/Theora>>.

⁸⁹ Tamtéž, s. 26.

⁹⁰ *Theora specification* [online]. 23.9.2010, [cit. 2010-11-30]. Dostupné z: <<http://www.theora.org/doc/Theora.pdf>>.

Při inter-predikci je využíváno dělení snímku na makrobloky a obdobně, jako v případě MPEG-4 ASP, taktéž na bloky. Tím lze dosáhnout čtyř pohybových vektorů na makroblok (4MV). Theora využívá dvojí druh referenčních snímků, které jsou označovány jako *Golden* a *Reference*. Pro jeden makroblok lze ovšem využít pouze jediný⁹¹. Jako referenční snímek pro následné provedení block matchingu lze tedy zvolit předcházející snímek, nebo nejbližší předešlý Intra snímek (Golden frame)⁹² (viz obrázek 2.5). Pohybovou predikci Theora provádí s přesností ½ bodu (HPeL).



Obrázek 2.5: Proud snímků za použití kodeku Theora.

Transformace probíhá pomocí standardní 2D DCT-II nad s bloky o rozměrech 8×8 bodu s 8 bity na bod. Následně je provedena kvantizace některou z mnoha kvantizačních matic. Theora taktéž, podobně jako MPEG-4 AVC, podporuje adaptivní kvantizaci a deblokovací filtr (in-loop deblocking filter) pro vylepšení psychovizuálního dojmu výsledných video dat⁹³. Dostupné jsou také vlastní kvantizační matice (custom matrices). DC koeficienty jsou, na rozdíl od AC, zakódovány rozdílově od sousedních. Výstup je poté linearizován pomocí diagonálního „cik-cak“ algoritmu. Následně je linearizovaný výstup RLE a VLC bezztrátově zakódován jednou z 80-ti dostupných kódovacích tabulek.

⁹¹Tamtéž, s. 26.

⁹²Tamtéž, s. 26.

⁹³*Theora - Xiph Wiki* [online]. 7.7.2010, [cit. 2010-11-30]. Dostupné z: <<http://wiki.xiph.org/Theora>>.

3 Teoretické shrnutí vlastností kompresních metod

3.1 MPEG-2 Part 2

Měl by zaručovat standardní kvalitu výstupu ovšem s relativně malým kompresním poměrem v porovnání s jeho následníky – minimálně rodina MPEG-4, VC-1 a Theora – kteří disponují značným množstvím přidaných vlastností. Například: pouze fixní velikost bloku, nepokročilé výstupní kódování, pouze HPEL, absence škálujících matic, neexistence vícenásobných blokových referencí, atd. Na druhou stranu by zátěž systému při enkódování určitého konkrétního datového toku měla být relativně nízká v porovnání se zbytkem enkódovacích metod.

3.2 H.264/MPEG-4 Part 10

Má teoretické předpoklady se stát nejefektivnějším enkodérem díky největšímu množství volitelných vlastností a rozšíření (např.: velké množství rozměrů „bloků“ při transformaci i pohybové predikci, pokročilá intra-predikce, možnost velice kvalitního výstupního kódování, vícenásobné reference, vlastní a škálující matice, deblokovací filtr, atd.). Otázkou však zůstává promítnutí těchto vlastností do zátěže systému při enkódování. Je tedy neznámý rozdíl časových intervalů potřebný k enkódingu v porovnání s ostatními video kodeky a především s hlavním konkurentem v podobě VC-1.

3.3 DivX Codec

Dle množství použitelných vlastností by měl v kompresním poměru překonat MPEG-2 Part 2 (např.: dělení makrobloku na bloky, použití vlastních a škálujících matic, QPEL, atd.). Díky podpoře GMC by měl být efektivní při plynulých záběrech (paning). Zátěž pro systém by teoreticky měla být o málo vyšší, než MPEG-2 Part 2.

3.4 Xvid

Tento kodek by měl mít téměř shodnou výstupní kvalitu a efektivitu jako DivX Codec. Má ovšem předpoklady jej předčit v monotónně stylizovaných scénách (cartoon mode) a jiných plynulých záběrech, kde lze využít pokročilejší GMC. Pro jeho použití hovoří jeho neproprietární vlastnosti, narozdíl od DivX Codec.

3.5 WMV9/VC-1

Zde lze předpokládat, že jeho efektivita bude dosahovat kvalit H.264. VC-1 totiž disponuje množstvím voleb při výběru velikosti „bloku“ pro transformace a pohybovou predikci. Stejně tak disponuje např.: D-quant, vylepšeným výstupním kódováním, deblokovacím filtrem, QPEL přesností, atp. Zajímavý je předpoklad o dobré efektivitě při plynulých přechodech díky Fading Compensation. Důležitým avšak neznámým aspektem zůstává zatížení systému při enkódování a dekodování. Opět především při srovnání s H.264.

3.6 Dirac

Zdá se, jako jediný zástupce waveletové komprese, značně nepředvídatelný jak v oblasti kvality obrazu, tak i v náročnosti na systémové zdroje. Nicméně, překrývající se části bloků by měly odstraňovat artefakty vznikající na rozhraní bloků a nadstandardní přesnost predikce až osminy bodu může zapříčinit dobré výsledky při nemonotónních scénách. Ke zmíněnému nepředvídatelnému chování může dopomáhat i relativně velice čerstvá implementace v porovnání s ostatními kodeky.

3.7 Theora

Theora disponuje předpoklady k obdobnému chování jako DivX a Xvid kodeky díky společným znakům jako: hrubší dělení makrobloku na jednotlivé bloky, použití škálujících matic, atd. Navíc zde lze najít velký počet kvantizačních matic a množství voleb výstupního kódování. Ovšem absence B-snímků a maximální HPEL přesnost zanechávají faktor efektivity a náročnosti jako těžko předvídatelné. Proti dobré efektivitě a předvídatelnosti hovoří i relativně krátká doba uplynutá od vzniku tohoto kodeku.

4 Zjišťování kvality video komprese

4.1 Úvod do metod měření kvality ztrátové video komprese

Všechny používané metody k posuzování ztrátové video komprese lze v zásadě rozdělit do těchto dvou skupin:

- Metody pro subjektivní hodnocení kvality – Zde hrají důležitou roli experimentátoři (lidští pozorovatelé), způsoby pozorování a přehrávání video souborů a také samotné metody následného vyhodnocení pozorování.
- Metody/metriky pro objektivní hodnocení kvality – U těchto metod se jedná o strojově prováděné měření a to vyhodnocení rozdílů obrazových bodů na základě matematických předpokladů.

Hlavní nevýhodou první jmenované skupiny je relativně velká náročnost na lidské zdroje, větší časová náročnost a také nemožnost exaktního opakování měření. Tyto nedostatky do jisté míry kompenzují metody řadící se do skupiny objektivních. Zde ovšem hrozí naprosté odloučení získaných výsledků od reálného nasazení. Proto je nejvýhodnější použít subjektivní metriky jako dodatečnou kontrolu metod objektivních⁹⁴.

4.1.1 Metriky pro objektivní hodnocení kvality video komprese

MSE (Mean Squared Error) Lze vyjádřit vztahem:

$$MSE = \frac{1}{M \cdot N} \sum_{i=1}^M \sum_{j=1}^N \|x(i, j) - y(i, j)\|^2 \quad (4.1)$$

Kde M a N jsou rozměry původního snímku x a enkódovaného snímku y . MSE se obvykle stanovuje pro každou složku obrazu zvlášť.

PSNR (Peak Signal-to-Noise Ratio) Jde o poměr maximálních energií signálu a šumu snímku. Nejčastěji je uváděno jako:

$$PSNR = 20 \cdot \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (4.2)$$

Kde MAX_I odpovídá $2^B - 1$ a B představuje počet bitů na jeden kanál. Čili MAX_I je rovno maximální hodnotě, kterou může obrazový bod ve snímku nabývat.

⁹⁴FLIEGEL, Karel. *Kvalita obrazu a modelování HVS* [online]. 20.11.2008, [cit. 2011-02-25]. Dostupné z: <<http://radio.feld.cvut.cz/courses/w2ZKT/materialy.php?akce=dlf&zdroj=vpm&fkey=7&xtgt=2f686f6d652f53657276696365732f>>.

SSIM (Structural Similarity Index) Tato metrika je navržena tak, aby lépe reflektovala reálné nasazení, než-li předchozí dvě⁹⁵.

Metrika pracuje s částmi snímků (výřezy) x a y o rozměrech $N \times N$. Lze ji vyjádřit jako:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1) \cdot (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1) \cdot (\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4.3)$$

$C_i = (K_i L)^2$, kde koeficient L odpovídá MAX_I používanému v PSNR a konstanty K jsou voleny jako $K \ll 1$ (nejčastěji: $K_1 = 0,01$ a $K_2 = 0,03$)⁹⁶. μ_i je použito pro vážený průměr, σ_i^2 zastupuje rozptyl a σ_{xy} kovarianci výřezů x a y .

MSSIM (Mean Structure Similitary Index) Jde o metriku vycházející ze SSIM, avšak aplikovanou na celý snímek⁹⁷.

$$MSSIM(A, B) = \frac{1}{N_x \cdot N_y} \sum_{i,j=1}^{N_x, N_y} SSIM(x, y) \quad (4.4)$$

Kde jde o snímky A a B tvořené N_x a N_y výřezy.

4.1.2 Metody pro subjektivní hodnocení kvality video komprese

DSIS (Double Stimulus Impairment Scale) ⁹⁸ V pevně stanovených časových okamžicích jsou skupině pozorovatelů předkládány testovací videa. Vždy jde o pořadí referenční a poté nekódované video. Po skončení každé ukázky pozorovatel hodnotí kvalitu na stupnici 0 – 100.

SS (Single Stimulus) ⁹⁹ Zde jsou pozorovatelům předkládány pouze testovací – enkódovaná – videa, bez předešlé reference s originálním videem. Po skončení každé ukázky je kvalita hodnocena na stupnici 1 – 5.

SCACJ (Stimulus Comparison Adjectival Categorical Judgement) ¹⁰⁰

Tato metoda je postavena na souběžném pozorování dvou nekódovaných video sou-

⁹⁵WANG, Zhou – BOVIK, Alan C. – SHEIKH, Hamid R. – SIMONCELLI, Eero P. *The SSIM Index for Image Quality Assessment* [online]. 11.2.2011, [cit. 2011-02-25]. Dostupné z: <<https://ece.uwaterloo.ca/~z70wang/research/ssim/>>.

⁹⁶Structural similarity - Wikipedia, the free encyclopedia [online]. 31.9.2010, [cit. 2011-02-25]. Dostupné z: <http://en.wikipedia.org/wiki/Structural_similarity>.

⁹⁷FLIEGEL, Karel. *Kvalita ...*, s. 30.

⁹⁸VERĀT, Ivo. *Televizní, rozhlasová a multimediální technika - hodnocení kvality videa a obrazu* - [online]. 5.11.2010, [cit. 2011-02-26]. Dostupné z: <http://stag.zcu.cz/fel/kae/TRM/rok_2010/vertat_4/TRM_Vertat_4_2010.pdf>.

⁹⁹Tamtéž, s. 31.

¹⁰⁰MSU Perceptual Video Quality: Subjective video quality methods information [online]. 28.4.2009, [cit. 2011-03-01]. Dostupné z: <http://compression.ru/video/quality_measure/subjective_metrics_info_en.html>.

borů. Po dokončení každé ukázky pozorovatelé hodnotí kvalitu jednoho videa ve vztahu k druhému na stupnici -3 – 3. Znamka 0 znamená shodnost, -3 horší kvalitu a 3 lepší kvalitu.

4.2 Měřicí nástroje

4.2.1 MSU Video Quality Measurement Tool

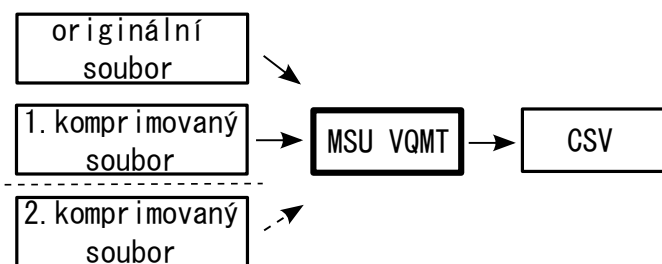
Jde o softwarový nástroj vyvinutý *MSU Graphics & Media Lab*¹⁰¹ pro potřeby snadného provádění objektivního hodnocení kvality video komprese.

Nástroj lze konfigurovat pro zpracování jednoho komprimovaného videa proti nekomprimovanému nebo dvou komprimovaných vůči originálnímu nekomprimovanému.

Jako vstup je využíváno služeb poskytovaných editačním programem *AviSynth*¹⁰².

To znamená, že pro úspěšné načtení jakéhokoli video souboru je nutné, aby byl otevřen a načten pomocí *AviSynth*. Nejsnazší cestou je přímo napsat AVS skript, který následně předložit MSU VQMT ke zpracování.

MSU nástroj nabízí využití prakticky všech objektivních metrik – od MSE až po MSSIM. Jako výstupní MSU VQMT¹⁰³ používá strukturovaný datový formát CSV¹⁰⁴ do kterého zaznamenává hodnoty vypočtené dle zvolené metriky pro každý snímek video souboru¹⁰⁵ (viz obrázek 4.1). Taktéž dovoluje záznam rozdílového video souboru ve formátu AVI.



Obrázek 4.1: Diagram činnosti MSU VQMT.

Hodnoty z výstupního CSV souboru lze dále využít k vizualizaci výsledku snímek po snímku a nebo výpočtu reprezentativního globálního výsledku měření (např. modus, či medián výstupních změřených hodnot). Použita byla FREE verze MSU VQMT 2.7.3.

¹⁰¹ *Main / Graphics and Media Lab* [online]. [cit. 2011-03-01]. Dostupné z: <<http://graphics.cs.msu.ru/>>.

¹⁰² *Main Page - Avisynth* [online]. 12.1.2011, [cit. 2011-03-01]. Dostupné z: <<http://avisynth.org/>>.

¹⁰³ *MSU Video Quality Measurement Tool (PSNR, MSE, VQM, SSIM)* [online]. 12.11.2010, [cit. 2011-02-28]. Dostupné z: <http://compression.ru/video/quality_measure/video_measurement_tool_en.html>.

¹⁰⁴ *Comma-separated values - Wikipedia, the free encyclopedia* [online], 17.1.2011, [cit. 2011-03-01]. Dostupné z: http://en.wikipedia.org/wiki/Comma-separated_values>.

¹⁰⁵ *MSU Video...*, s. 33.

4.2.2 tiny_ssim

Tato velice jednoduchá utilita, uvolněná L. Merrittem¹⁰⁶, dokáže změřit MSSIM (označuje ji jako SSIM) a PSNR metriky pro všechny složky barevného modelu YUV. Vstupem jsou zdrojové a referenční video data ve formátu YUV, měření tedy probíhá nezávisle na vstupním kodeku.

4.2.3 MSU Perceptual Video Quality Tool

MSU PVQT¹⁰⁷ je relativně komplexní nástroj používaný pro měření, ohodnocení a zpracování různých subjektivních metod hodnocení kvality. Skládá se z oddělené konfigurační části pojmenované *manager* a přehrávací části *player*. Poskytuje plnou podporu pro SCACJ metodiku a tudíž dva souběžně promítané obrazy umístěné horizontálně vedle sebe. Jako vstupní lze použít videosoubory v AVI kontejneru, tak AVS skripty. Výstupní log soubor je ukládán formě CSV.

¹⁰⁶MERRITT, Loren. *[MEncoder-users] Codec-independant objective quality measurement* [online]. 5.8.2006, [cit. 2011-03-25]. Dostupné z: <<http://lists.mplayerhq.hu/pipermail/mencoder-users/2006-August/003801.html>>.

¹⁰⁷MSU Perceptual Video Quality Tool [online]. c2008, [cit. 2011-05-08]. Dostupné z: <http://compression.ru/video/quality_measure/perceptual_video_quality_tool_en.html>.

5 Popis realizace měření kvality video komprese

5.1 Implementace enkodérů

Jednotlivé kodeky pro experimentální účely byly vybrány na základě jejich rozšířenosti a vlastní reálném nasazení v praxi. Konkrétní zvolené verze enkodérů odpovídají nejnovějším možným stabilním verzím. Pro svou podobnost byla dána přednost otevřenému kodeku Xvid.

Byly tedy vybrány tyto kodeky:

- **MPEG-2 Part 2: HC Encoder 0.25**¹⁰⁸;
- **Xvid: Xvid 1.03**¹⁰⁹;
- **H.264/MPEG-4 Part 10: x264**, 32b build, rev 1913¹¹⁰;
- **VC-1: Microsoft Expression Encoder 4**; v4.0.1651.0¹¹¹;
- **Dirac: ffmpeg-r26400**¹¹²;
- **Theora: ffmpeg2theora-0.27**¹¹³.

Pro potřeby MSU VQMT byly vytvořeny samostatné AviSynth AVS skripty, které načítají video soubory za použití převážně *FFMS2*¹¹⁴ (FFmpegSource) AVS pluginu a případně aplikují AVS filtry.

U každé kompresní metody bylo využito komprimování s dvojitým průběhem. Pro každý z použitých enkodérů je dostupné přesné nastavení v příslušném souboru – většinou jde o dávkové soubory. Veškeré tyto soubory s použitým nastavením jsou k dispozici v DVD příloze.

Pozn.: veškeré Dirac video soubory dekodované pomocí FFmpeg + libschroedinger obsahovaly tři duplicitní snímky (druhý snímek byl 3x zkopírován). Toto muselo být před měřením zkorigováno za použití nástroje *VirtualDub*.

¹⁰⁸ *HC Encoder 0.26 beta - VideoHelp.com Downloads* [online]. 5.3.2011, [cit. 2011-03-6]. Dostupné z: <www.videohelp.com/tools/HC>.

¹⁰⁹ *Xvid.org: Downloads* [online]. c2010, [cit. 2011-03-6]. Dostupné z: <<http://www.xvid.org/Downloads.15.0.html>>.

¹¹⁰ *x264.nl*, ..., s. 10.

¹¹¹ *Microsoft Expression Encoder 4 Pro | H.264 Encoding | VC-1 Codec | Streaming Silverlight Video | IIS Smooth Streaming | Real Time Encoder | Microsoft Expression 4* [online]. c2011, [cit. 2011-03-6]. Dostupné z: <http://www.microsoft.com/expression/products/EncoderPro_Overview.aspx>.

¹¹² *FFmpeg* [online]. 24.2.2011, [cit. 2011-03-6]. Dostupné z: <<http://www.ffmpeg.org/>>.

¹¹³ *Theora.org :: downloads - Theora, video for everyone* [online]. c1994 - 2011, [cit. 2011-03-6]. Dostupné z: <<http://www.theora.org/downloads>>.

¹¹⁴ *ffmpegsources - Project Hosting on Google Code* [online]. c2010, [cit. 2011-03-6]. Dostupné z: <<http://code.google.com/p/ffmpegsources/>>.

5.2 Rozlišení a datové toky testovacích souborů

Pro výběr testovaných rozlišení byly brány v potaz především ohledy na nasazení v praxi.

5.2.1 Rozlišení

Mobilní Bylo vybráno rozlišení WVGA – 852×480 nebo 848×480 bodů (~400 000 bodů)¹¹⁵ použité s 25 snímků za sekundu. Tento počet obrazových bodů řádově odpovídá rozlišení, které je široce využíváno v mobilních zařízeních současné generace¹¹⁶. V případě 848×480 jde o poměr stran pseudo-16:9, který je rozšířený v praxi a plní požadavek na dělitelnost obou rozměrů 16ti beze zbytku.

HD 1080p je nejběžnější normou pro HDTV¹¹⁷ digitální televizní vysílání a je v současnosti hlavním prosazovaným rozlišením pro domácí zobrazovací zařízení – FullHD. Bohužel, zmíněný nástroj MSU VQMT dovoluje provádět testy s HD video soubory pouze ve své zpoplatněné PRO verzi. Proto bylo nutno se uchýlit k redukci tohoto rozlišení až na 1024×576 bodů (~590 000 bodů) při 25 snímcích za sekundu. Toto rozlišení ovšem plně nereflektuje skutečnost nasazení v HD oblastech. Jinou dostupnou metodou by mohlo být vycentrované 1:1 ořezání 1080p souborů na menší rozlišení, avšak zde by se díky rozdílnému obsahu středu snímků kompletně vytratila vypovídací hodnota srovnání mezi oběma rozlišeními – mobilním a „HD“.

5.2.2 Datové toky

Pro mobilní zařízení je v současnosti pro bezdrátové datové spojení využíváno standardu *3G/IMT-2000* (International Mobile Telecommunications-2000)¹¹⁸. Na území ČR by pokrytí signálem tohoto spojení mělo být takřka kompletní, na rozdíl od ostatních podobných přenosových technik. Pro přenos online streamovaných videí je tedy využíván tento standard. 3G vyžaduje ve svých implementacích splnění minimální přenosové rychlosti 200 Kb/s¹¹⁹ a proto byla tato hodnota vybrána jako zástupce nejnižšího typického mobilního datového toku. Jako doplňující hodnota

¹¹⁵ *List of common resolutions - Wikipedia, the free encyclopedia* [online]. 5.3.2011, [cit. 2011-03-7]. Dostupné z: <http://en.wikipedia.org/wiki/List_of_common_resolutions>.

¹¹⁶ HTC Desire HD, HTC EVO 4G, Nokia N900, Sony Ericsson Xperia X2, Nexus One, atd.

¹¹⁷ *1080p - Wikipedia, the free encyclopedia* [online]. 6.3.2011, [cit. 2011-03-7]. Dostupné z: <<http://en.wikipedia.org/wiki/1080p>>.

¹¹⁸ *What really is a Third Generation (3G)(3G) Mobile Technology* [online]. 17.7.2008, [cit. 2011-03-9]. Dostupné z: <http://www.itu.int/ITU-D/imt-2000/Documents/IMT2000/What_really_3G.pdf>.

¹¹⁹ FRANCA, Bruce. *U.S. IMT-2000 Implementation* [online]. 27.6.2008, [cit. 2011-03-9]. Dostupné z: <<http://wireless.fcc.gov/outreach/presentations/blacksburg2008/BB%20Deployment%20-%20Fallon.pdf>>.

500 Kib/s.

Jako vysoký datový tok v mobilních zařízeních byla vybrána hodnota 1,5 Mb/s. Tato hodnota reprezentuje rozumné maximum, které je možné bez problémů offline přehrát na většině ze současné generace mobilních telefonů¹²⁰. Jako doplňující 2 Mb/s, kterou zvládají přehrát kvalitnější přístroje.

Pro vysoké rozlišení bylo vybrán datový tok 23Mib/s. Protože ten přibližně odpovídá průměrnému datovému toku používaného na komerčních Blu-ray Disc nosičích¹²¹, které jsou v současnosti nejrozšířenějším médiem pro šíření HD obsahu. I kvůli nutnosti učinit kompromis v rozlišení jsem zařadil datový tok 9,8Mb/s, který představuje maximální datový tok pro DVD-Video¹²² jako doplňující.

¹²⁰iPhone 1500Kb/s, HTC HD2 5000Kb/s, HTC Desire 1500Kb/s, Samsung Spica i5700 700Kb/s

¹²¹*Blu-ray Movie Bitrates Here - Blu-ray Forum* [online]. 6.12.2009, [cit. 2011-03-9]. Dostupné z: <<http://forum.blu-ray.com/blu-ray-movies-north-america/3338-blu-ray-movie-bitrates-here.html>>.

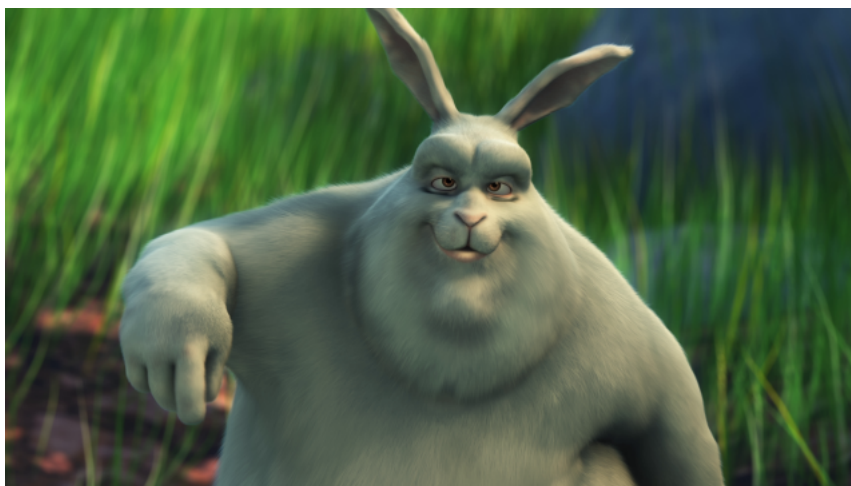
¹²²*What is DVD? - VideoHelp.com* [online]. c1999-2010, [cit. 2011-03-18]. Dostupné z: <<http://www.videohelp.com/dvd>>.

5.3 Testovací video data – samplý

Hledání testovacích dat se muselo podřídit požadavkům, které klade zadání práce. To znamená předem nezkomprimovaný obsah a také dodržení charakteru obsažených scén, které bylo určeno zmíněným zadáním.

Pro jednotlivé charakteristické scény byly vybrány tato vhodná testovací video data, která jsou různými skupinami prakticky využívána pro srovnávací účely:

- **animované video** – animovaný amatérský krátkometrážní film *Big Buck Bunny*¹²³. Originál: 1920×1080p, 24 fps, formát snímků 24b RGB, soubor png. Převáděno na: 1920×1080p, 25 fps, formát YUV 8b „4:2:0“, kontejner avi. Délka využití části: 350 snímků. Ilustrační snímek viz obrázek 5.1.



Obrázek 5.1: Snímek z testovací scény *BBB*.

- **statická kamera a statická scéna** – Video Quality Experts Group testovací data *SnowMnt_8bit*¹²⁴. Originál: 1920×1080p, 30 fps, formát 8b YUV „4:2:2“, kontejner avi. Převáděno na: 1920×1080p, 25 fps, formát YUV 8b „4:2:0“, kontejner avi. Délka využití části: 254 snímků. Ilustrační snímek viz obrázek 5.2.
- **statická kamera a dynamická scéna** – Video Quality Experts Group testovací data *RushFieldCuts_8bit*¹²⁵. Originál: 1920×1080p, 30 fps, formát 8b YUV „4:2:2“, kontejner avi. Převáděno na: 1920×1080p, 25 fps, formát YUV 8b „4:2:0“, kontejner avi. Délka využití části: 196 snímků. Ilustrační snímek je vyjádřen na obrázku 5.3.

¹²³*Index of /BBB* [online]. 24.5.2008, [cit. 2011-03-14]. Dostupné z: <<http://media.xiph.org/BBB/>>.

¹²⁴ftp://vqeg.its.bldrdoc.gov/HDTV/NTIA_source/ [online]. 20.2.2008, [cit. 2011-03-14]. Dostupné z: <ftp://vqeg.its.bldrdoc.gov/HDTV/NTIA_source/>.

¹²⁵Tamtéž, s. 38.



Obrázek 5.2: Snímek z testovací scény *SnowMnt*.



Obrázek 5.3: Snímek z testovací scény *RushFieldCuts*.

- **dynamická kamera a dynamická scéna** – Technische Universität München testovací data *Tractor*¹²⁶. Originál: 1920×1080p, 25 fps, formát 8b YUV „4:2:0“, soubor yuv. Převáděno na: 1920×1080p, 25 fps, formát YUV 8b „4:2:0“, kontejner avi. Délka využití části: 280 snímků. Ilustrační snímek viz obrázek 5.4.

Ideální řešení by představovala tyto video data získaná přímo v obou požadovaných rozlišení – 848×480 a 1024×576 bodů. Díky neexistenci všech dat v tomto testovacím souboru nativně v těchto rozlišení bylo nutné přistoupit k resamplování z originálního rozlišení na nižší – downsamplingu. Toto řešení se nutně musí vyskytovat i v praxi. Jako downsamplovací filtr byl vybrán čtyřnásobný Lanczos filtr

¹²⁶ ftp://ftp.ldv.e-technik.tu-muenchen.de/dist/test_sequences/1080p/ [online]. 29.11.2002, [cit. 2011-03-14]. Dostupné z: <ftp://ftp.ldv.e-technik.tu-muenchen.de/dist/test_sequences/1080p/>.



Obrázek 5.4: Snímek z testovací scény *Tractor*.

*Lanczos4*¹²⁷, který poskytuje velice kvalitní zachování detailů¹²⁸ a je široce využíván.

¹²⁷*Resize - Avisynth* [online]. 17.12.2010, [cit. 2011-03-21]. Dostupné z: <<http://avisynth.org/mediawiki/Resize#LanczosResize>>.

¹²⁸HÖHNE, Hermann. *AVISynth resizing filter comparasion* [online]. 26.7.2007, [cit. 2011-03-21]. Dostupné z: <<http://hermidownloads.craqstar.de/videoresizefiltercomparasion/>>.

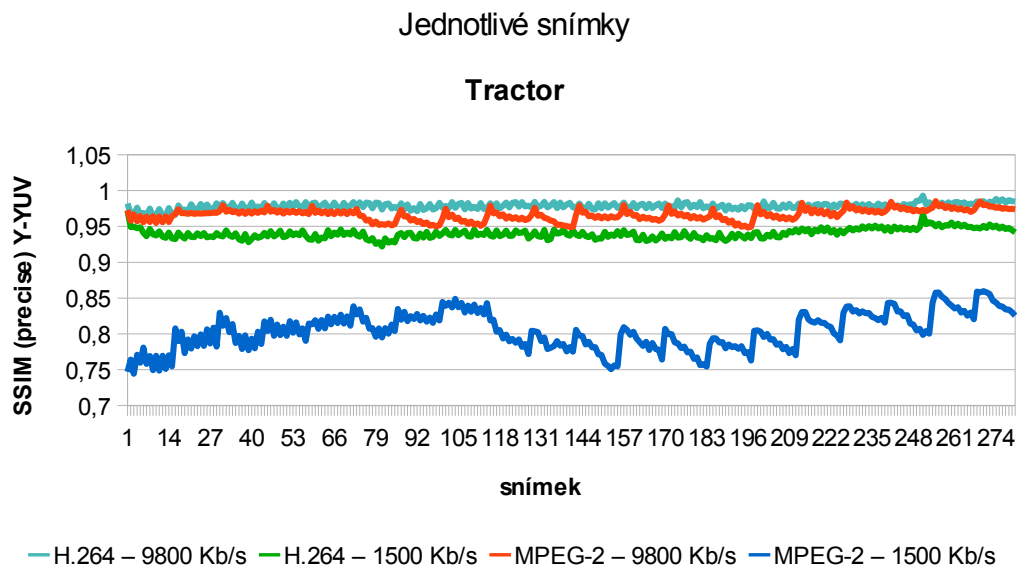
6 Objektivní část měření

6.1 Použitá metodika

Jako objektivní metriky pro měření kvality komprese byla vybrána metrika PSNR, která je velice široce využívána. V nástroji MSU VQMT byla zvolena základní varianta pojmenovaná jako *PSNR*. Pro získání lepší korelace s lidským vnímáním byla zařazena metrika MSSIM. Nástroj MSU VQMT označuje MSSIM jako SSIM, avšak provádí ji korektně nad celým snímkem. Měření probíhalo za použití varianty *SSIM (precise)*.

Soubor dat získaných z měření lze vyjádřit pro jednotlivé snímky, avšak v tomto případě lze ztěžít udržet míru přehlednosti. Zde následuje ilustrativní ukázka pro pouhé dva datové toky a dva kodeky namísto 6 datových toků a 6 kodeků – obrázek 6.1.

Namísto toho bylo výhodnější zvolit charakterizaci souboru dat pomocí jeho arit-



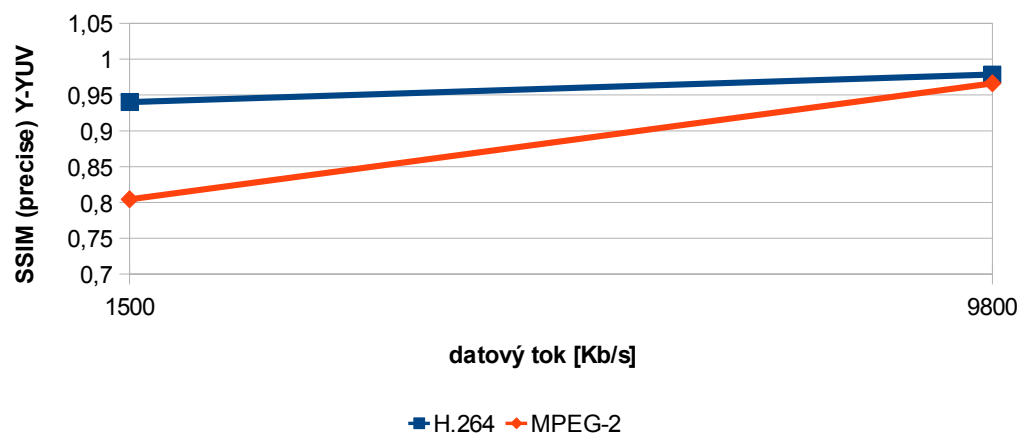
Obrázek 6.1: MSSIM nad luminační složkou – jednotlivé snímky.

metického průměru. Ten je používán jako výstup v případě MSU VQMT, tak i například v enkodéru x264¹²⁹. Následuje grafické vyjádření pro stejnou situaci s použitím průměrných hodnot SSIM (viz obrázek 6.2).

¹²⁹x264.nl, ..., s. 10.

Průměrné hodnoty

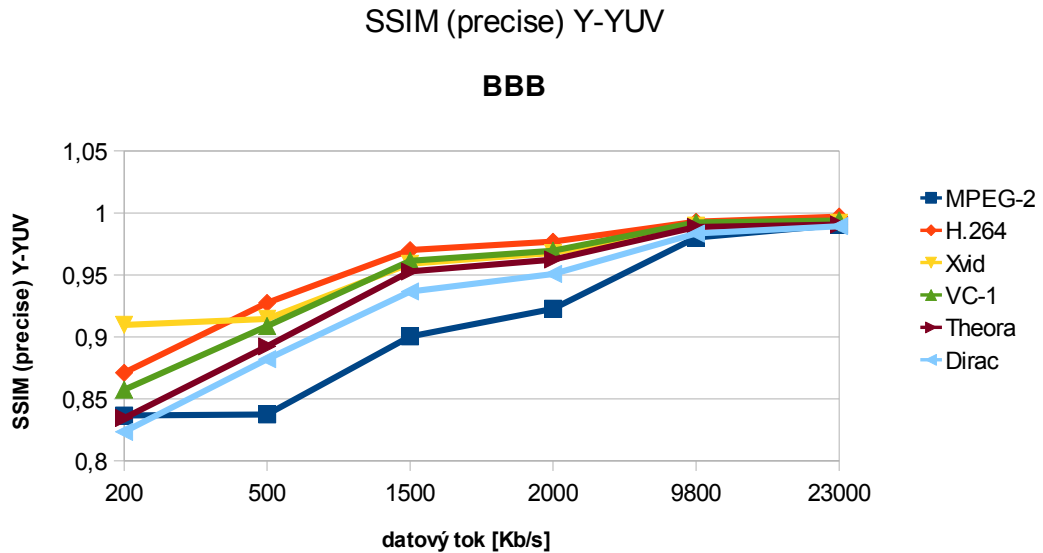
Tractor



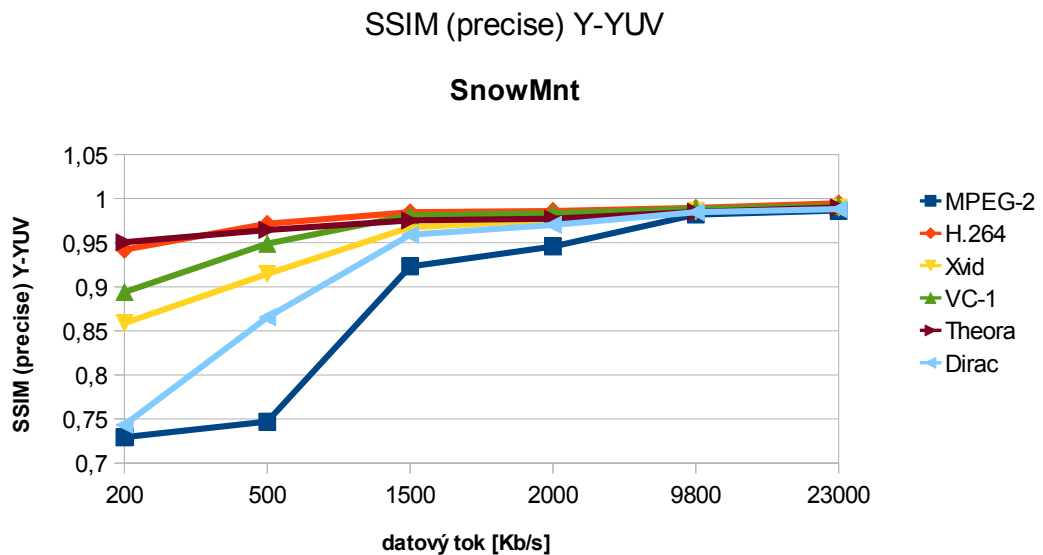
Obrázek 6.2: MSSIM nad luminační složkou – průměrné hodnoty.

6.2 MSSIM

Grafická reprezentace výsledků měření pro jednotlivé testovací soubory prováděné pomocí objektivní metriky MSSIM nad luminační složkou YUV modelu. Provedené pomocí nástroje MSU VQMT.

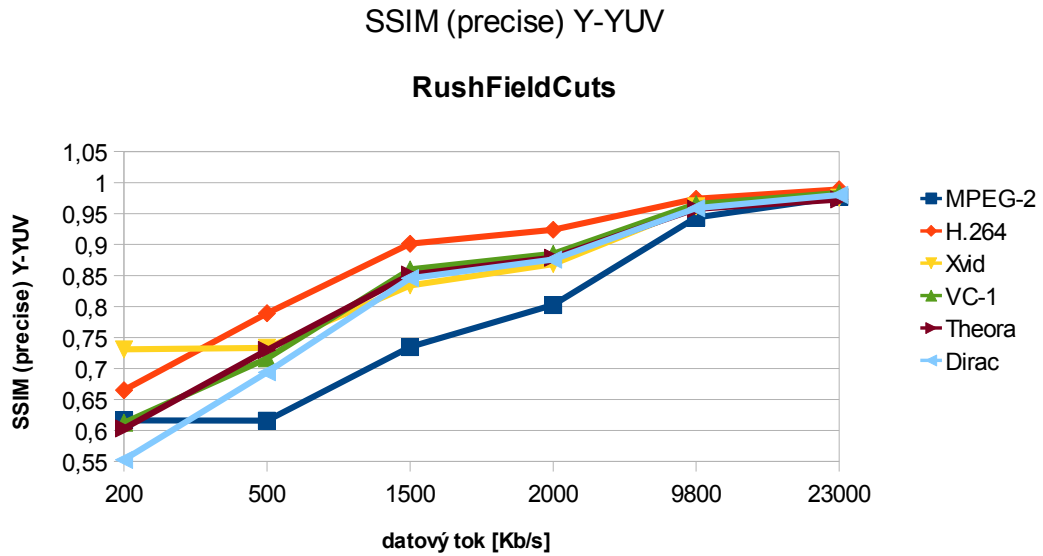


Obrázek 6.3: MSSIM nad luminační složkou pro scénu *BBB*.

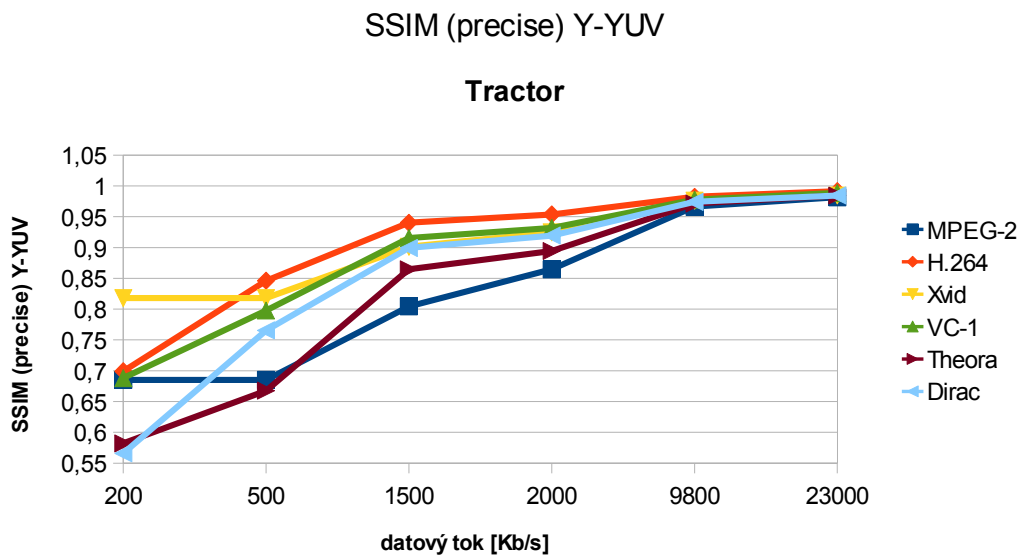


Obrázek 6.4: MSSIM nad luminační složkou pro scénu *SnowMnt*.

Toto měření prokázalo, že rozdíly mezi jednotlivými metodami komprese se rychle stírají při vyšších datových tocích. MPEG-2 Part 2 a Dirac se dají označit za nejhorší kodeky při posuzování touto metrikou – nejhorší výsledky podávají ve



Obrázek 6.5: MSSIM nad luminační složkou pro scénu *RushFieldCuts*.

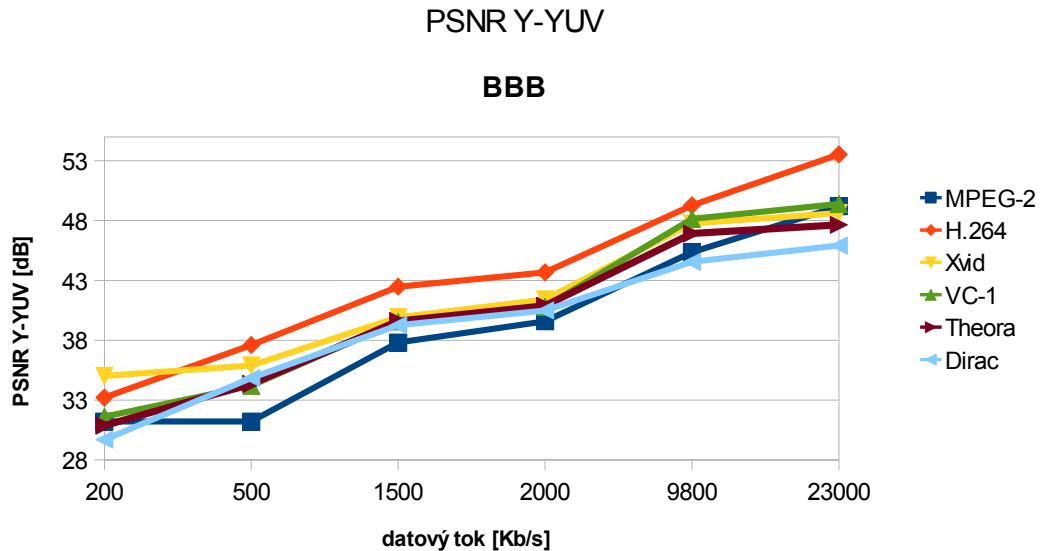


Obrázek 6.6: MSSIM nad luminační složkou pro scénu *Tractor*.

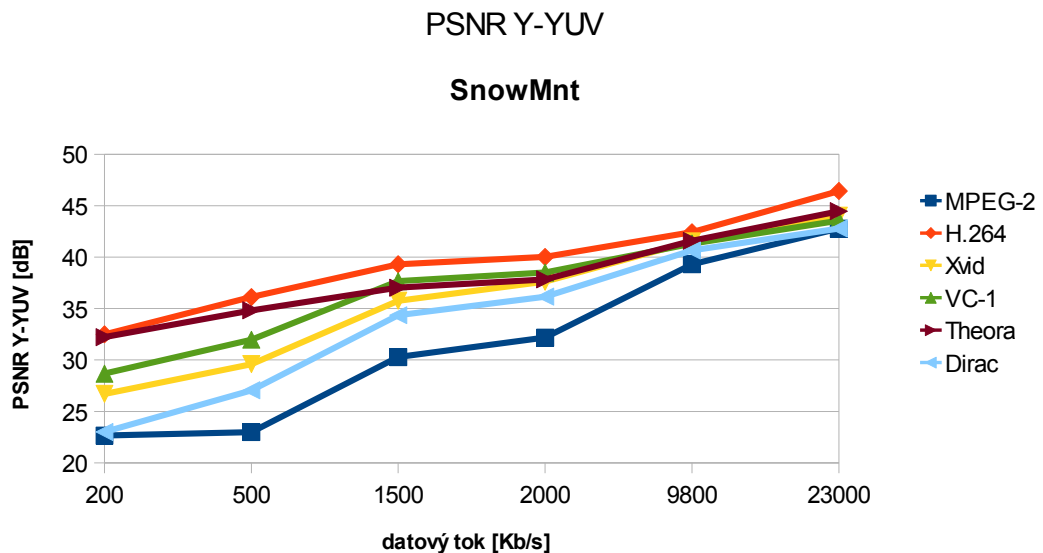
statické scéně (viz obrázek 6.4). Zejména vezmeme-li v potaz zkrslení výsledků dané nesprávným datovým tokem, který ukázalo pozdější porovnání velikostí (viz kapitola 6.5. Velice dobré výsledky ukázala Theora při aplikaci na statickou scénu (viz obrázek 6.4). Celkově nejlepší dle MSSIM metriky skončil kodek sandardu H.264, který vykazoval nejlepší výsledky prakticky ve všech datových tocích a scénách.

6.3 PSNR

Výsledky měření pro jednotlivé testovací soubory prováděné stejným postupem, ovšem pomocí objektivní metriky PSNR taktéž nad luminační složkou. Provedené pomocí nástroje MSU VQMT (viz kapitola 4.2.1).

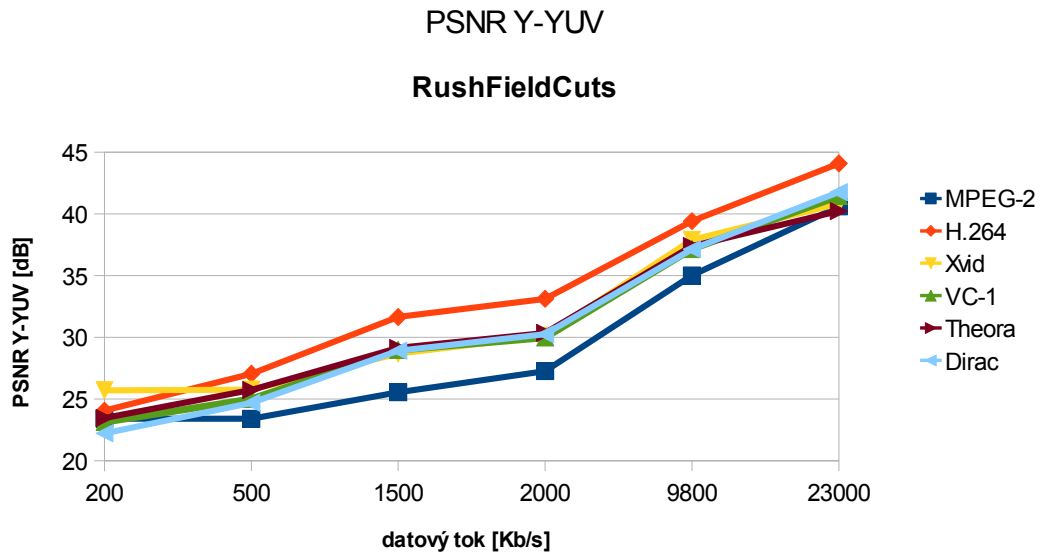


Obrázek 6.7: PSNR nad luminační složkou pro scénu *BBB*.

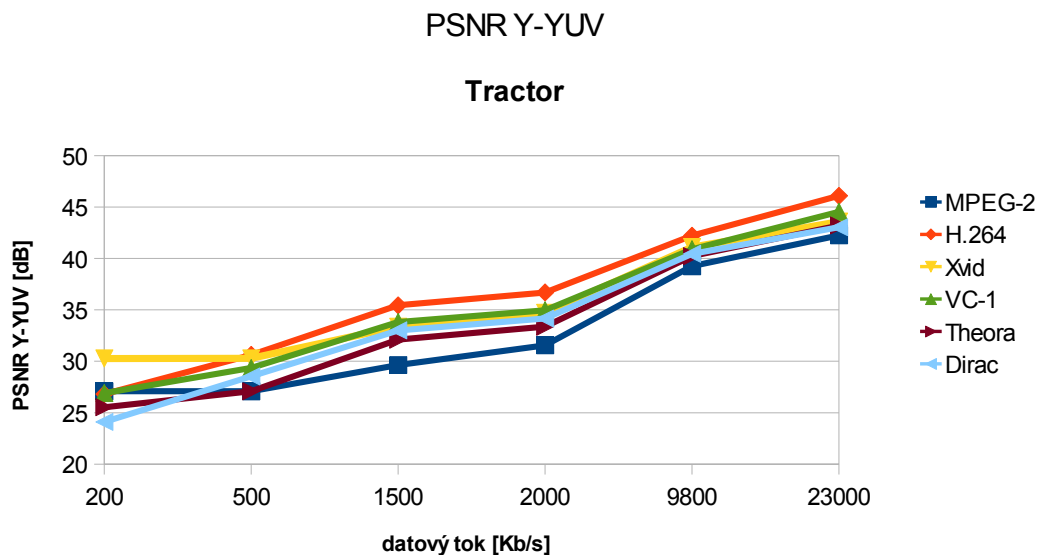


Obrázek 6.8: PSNR nad luminační složkou pro scénu *SnowMnt*.

I v této metrice dominuje po korekci zkreslených výsledků (viz kapitola 6.5) kodek standardu H.264 napříč všemi datovými toky. MPEG-2 Part 2 i zde obsazuje ve srovnání poslední místo. Největší problémy společně s Dirac má MPEG-2 Part 2



Obrázek 6.9: PSNR nad luminační složkou pro scénu *RushFieldCuts*.

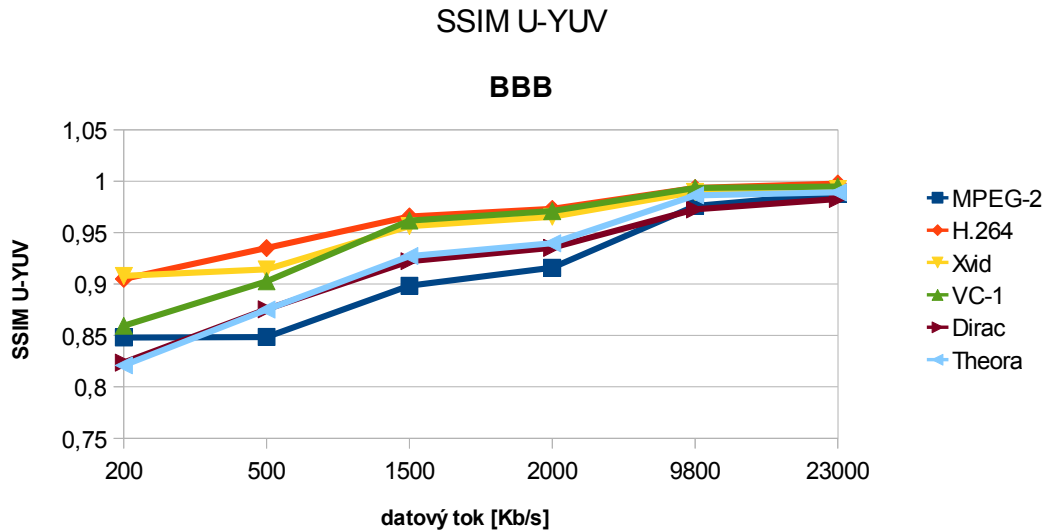


Obrázek 6.10: PSNR nad luminační složkou pro scénu *Tractor*.

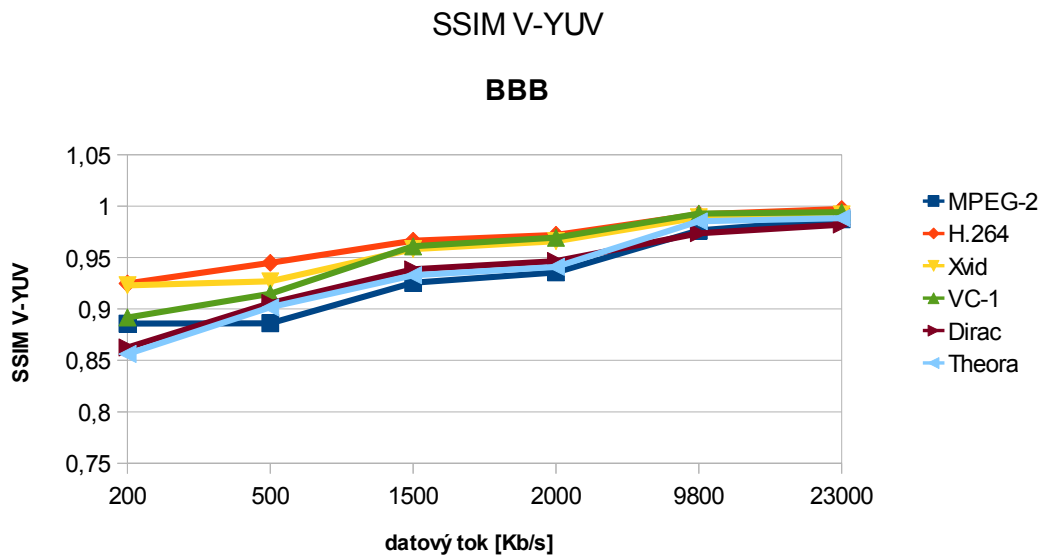
se statickou scénou (viz obrázek 6.8), kdy především při nejnižších datových tocích zdaleka nedosahuje výkonů ostatních kodeků. Také zde ukázala dobré výsledky při statické scéně metoda Theora, která se dokázala vyrovnat H.264 a VC-1.

6.4 Chrominanční složky

Za použití zmíněné utility *tiny_ssim* (viz kapitola 4.2.2) byly pro ilustraci naměřeny hodnoty chrominančních složek dle obou zkoumaných objektivních metrik. Testovací scéna *BBB* byla vybrána na základě největšího počtu barev a také obrazových stříhů.

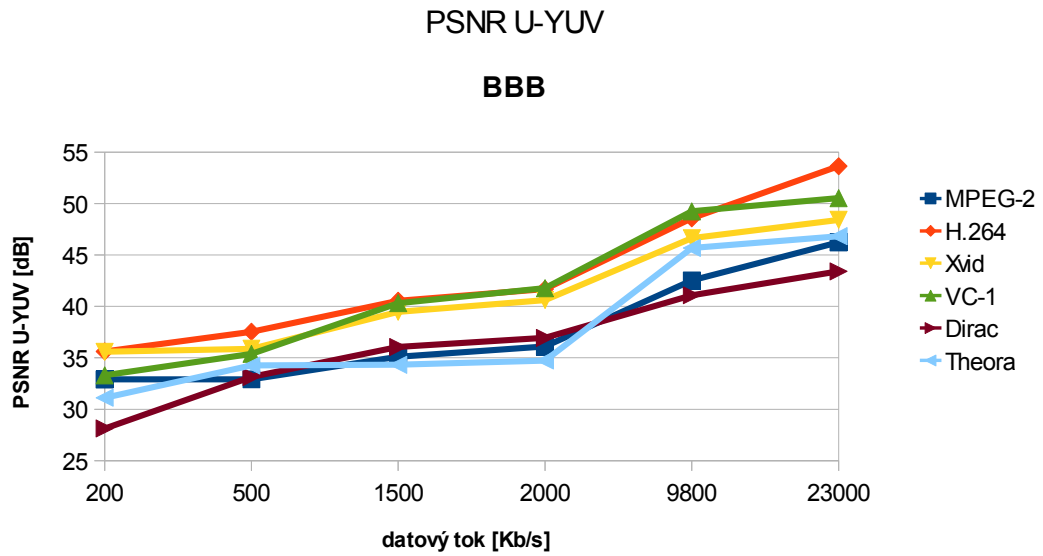


Obrázek 6.11: MSSIM nad chrominanční složkou U-YUV.

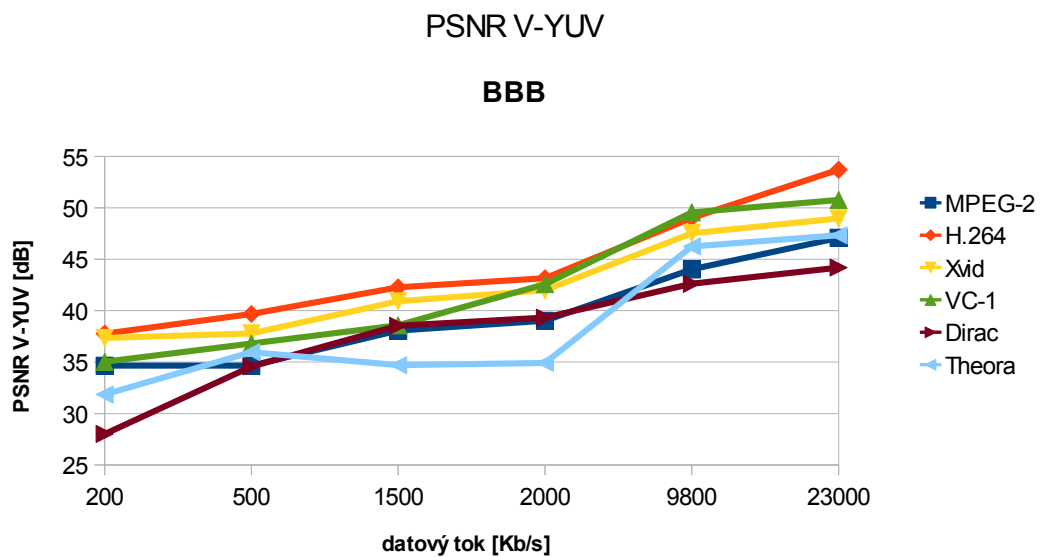


Obrázek 6.12: MSSIM nad chrominanční složkou V-YUV.

Z výsledků lze vyvodit, že obě chrominanční složky jsou zpracovány všemi enkodéry obdobně. Jako jediná význačná odchylka se jeví horší hodnoty PSNR u V-YUV při použití kodeku standardu VC-1 a datovém toku 1500Kb/s (viz obrázek 6.14).



Obrázek 6.13: PSNR nad chrominanční složkou U-YUV.



Obrázek 6.14: PSNR nad chrominanční složkou V-YUV.

Avšak toto zhoršení není zcela průkazné, protože dle metriky MSSIM se složka V-YUV při 1500Kb/s ve VC-1 chová zcela standardně (viz obrázek 6.12).

6.5 Výsledná velikost

Výsledné velikosti enkódovaných samplů jsou vyneseny do tabulek. Rozdíly velikostí dané nestejnými kontejnery se pohybují do 30KB. Xvid a Theora mají horní hranici datového toku pod hodnotou 23Mb/s (20480Kb/s respektive 16778Kb/s) nepodávají tudíž relevantní výsledky v této testované oblasti.

velikost [KB]	200 Kb	500 Kb	1500 Kb	2000 Kb	9800 Kb	23000 Kb
MPEG-2	1246	1248	2304	2982	12554	35922
H.264	333	820	2437	3198	16041	38647
Xvid	944	999	2558	3415	16242	20683
VC-1	492	865	2567	3419	16723	20239
Dirac	349	873	2617	3478	16759	39232
Theora	515	931	2834	3775	15541	17937

Tabulka 6.1: Velikosti – testovací scéna *BBB*.

velikost [KB]	200 Kb	500 Kb	1500 Kb	2000 Kb	9800 Kb	23000 Kb
MPEG-2	472	522	1618	1903	11071	23697
H.264	238	619	1928	2518	12044	27670
Xvid	301	386	1411	2226	12888	21319
VC-1	346	694	1825	2380	11715	21465
Dirac	297	649	1912	2544	12411	28549
Theora	261	627	1868	2484	12096	19319

Tabulka 6.2: Velikosti – testovací scéna *SnowMnt*.

velikost [KB]	200 Kb	500 Kb	1500 Kb	2000 Kb	9800 Kb	23000 Kb
MPEG-2	711	710	1161	1490	8012	21021
H.264	197	472	1405	1878	9139	22810
Xvid	683	685	1169	1567	9411	17863
VC-1	264	468	1403	1887	9208	21387
Dirac	230	598	1479	1960	9557	22073
Theora	227	490	1443	1953	9669	16106

Tabulka 6.3: Velikosti – testovací scéna *RushFieldCuts*.

velikost [KB]	200 Kb	500 Kb	1500 Kb	2000 Kb	9800 Kb	23000 Kb
MPEG-2	1203	1203	1661	2189	12292	25478
H.264	248	687	2046	2713	13305	31432
Xvid	999	1000	1796	2421	13393	25381
VC-1	385	682	2036	2715	13223	30692
Dirac	290	692	2078	2764	13503	31646
Theora	723	706	2130	2838	13765	23141

Tabulka 6.4: Velikosti – testovací scéna *Tractor*.

Z výše uvedených hodnot měření velikosti komprimovaných testovacích dat můžeme vyvodit tyto poznatky. Předně Xvid a MPEG-2 Part 2 špatně zvládají dodržování datového toku při obou nejnižších nastavení. To objasňuje jejich výsledky při měření MSSIM a PSNR metrikami v kapitole 6.2, respektive 6.3.

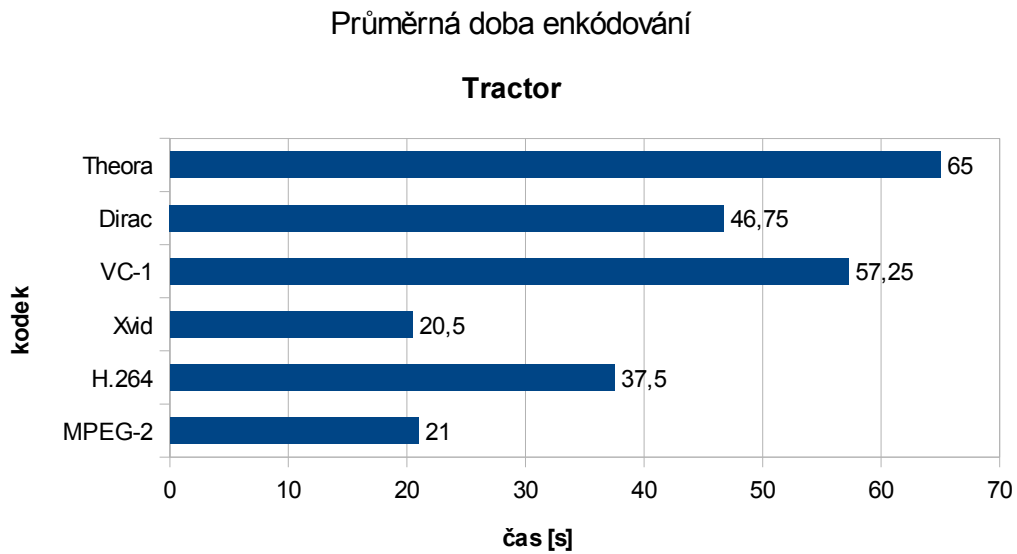
Jako nejvýhodnější a nejkonzistentější se jeví metoda H.264. MPEG-2 Part 2 produkuje nejmenší výsledné video soubory, avšak taktéž zaostává v naměřené objektivní kvalitě. Na závěr datové toky 1500Kb/s a 2000Kb/s nabízí v kombinaci s Xvid dobrý poměr výsledné velikosti a kvality dle obou použitých objektivních metrik.

6.6 Doby běhu

Jako ukázkou dokumentující vytížení systému a s tím spojenou dobu běhu enkódování byla změřena průměrná doba enkódování nejnižšího a nejvyššího datového toku. Pro tento účel byl vybrán testovací soubor *Tractor* jakožto zástupce nejvariabilnější scény. Měření probíhalo s pomocí služeb měření času poskytovaných OS (nástroj *time*), logovacích souborů enkodérů a klasických stopek v případě Xvid.

Většina použitých enkodérů byla designována jako vícevláknová, což bylo možné zjistit z vytížení procesoru. Pouze *ffmpeg2theora* uměl využít jen jedno vlákno.

Měření probíhalo na stroji vybaveném CPU *Intel C2Q 9300*(rev. M1)@(333×7.5) MHz, pevným diskem s 7200 ot./min (*Seagate ST31000528AS*), operační paměť 2×2GB (*A-DATA Vitesta Extreme Edition DDR2-800*@[400 MHz, 5-6-6-20, 2T]) a plně aktualizovaným operačním systémem *MS Windows 7 Professional ENG x86-64*.



Obrázek 6.15: Průměrné doby běhu.

7 Subjektivní část měření

7.1 Použitá metodika

Jako určitý kontrolní mechanismus pro měření provedená čistě objektivními metrikami (viz kapitola 6) byla zařazena tato subjektivní část. Jako základ použité metody byla použita *SCACJ*¹³⁰. Díky množství scén a použitých kodeků bylo nutné redukovat jak počet datových toků, tak pozměnit i použitou metodu měření. U metody měření bylo použito redukované schéma bodování na: lepší, horší a stejné. Tyto bodování svých preferencí experimentátoři vyjadřovali na připravené dotazníky. Datový tok byl vybrán 1500 Kb/s, který dle použitých objektivních metrik vykazoval znatelné rozdíly napříč použitými kodeky a zároveň se zde nevyskytovaly problémy s video souborem.

Systém pro jednotlivé iterace srovnávacího procesu byl zvolen následovný. Jako první byly třikrát pseudo-náhodně vygenerovány nestejně dvojice kodeků. Každému z těchto souborů tří dvojic byla přiřazena jedna skupina experimentátorů. Pro všechny scény v jednom souboru jsou dvojice stejné, ale vyhodnocení je oddělené pro každou z nich. Po vyhodnocení bodových zisků udělených experimentátory jsou z každého souboru z každé dvojice určeny tři lépe ohodnocené metody komprese. Porovnáním těchto metod mezi všemi třemi skupinami dostaneme odděleně pro každou scénu konečnou trojici nejlépe ohodnocených kodeků. V případě shodného hodnocení bylo přistoupeno ke klasickému „roztřelu“.

Pro provedení porovnání byl použit nástroj MSU PVQT¹³¹ (blíže poznámka 107), který umožňuje komfortní souběžné přehrávání dvou snímků. V ostatních případech bylo použito klasického přehrávače *VideoLAN*¹³². Každá projekce obsahovala jednotlivé snímky náhodně rozmístěné, aby nedošlo k ovlivnění experimentátorů.

Celkově se experimentů zúčastnilo 24 experimentátorů a nejpočetnější věkové zastoupení v této skupině se pohybovalo okolo 20 let.

¹³⁰ *MSU Perceptual ...*, s. 31.

¹³¹ *MSU Perceptual ...*, s. 34.

¹³² *VideoLAN - Official page for VLC media player, the Open Source video framework!* [online]. 15.4.2011, [cit. 2011-05-08]. Dostupné z: <www.videolan.org/vlc/>.

7.2 Výsledky prvního kola experimentů

Tato část měla za úkol určit lépe hodnocený kodek z každé dvojice.

První soubor experimentů byl proveden dne 3.5.2011 s 10 experimentátory.

scéna	BBB		SnowMnt		RushFieldCuts		Tractor	
kodek	MPEG-2	H.264	MPEG-2	H.264	MPEG-2	H.264	MPEG-2	H.264
body	1	9	8	6	0	10	1	9
kodek	Dirac	VC-1	Dirac	VC-1	Dirac	VC-1	Dirac	VC-1
body	4	10	7	8	1	10	7	8
kodek	Theora	Xvid	Theora	Xvid	Theora	Xvid	Theora	Xvid
body	1	10	7	7	8	6	4	8

Tabulka 7.1: Soubor experimentů – 1.

Druhý soubor experimentů byl proveden dne 7.5.2011 s 5 experimentátory.

scéna	BBB		SnowMnt		RushFieldCuts	
kodek	Xvid	VC-1	Xvid	VC-1	Xvid	VC-1
body	2	4	3	4	0	5
kodek	MPEG-2	Dirac	MPEG-2	Dirac	MPEG-2	Dirac
body	0	5	2	3	1	5
kodek	H.264	Theora	H.264	Theora	H.264	Theora
body	5	3	3	3	5	1

Tabulka 7.2: Soubor experimentů – 2a.

scéna	Tractor	
kodek	Xvid	VC-1
body	5	3
kodek	MPEG-2	Dirac
body	0	5
kodek	H.264	Theora
body	5	0

Tabulka 7.3: Soubor experimentů – 2b.

Třetí soubor experimentů byl proveden dne 8.5.2011 se 6 experimentátory.

Z výsledků prvního kola experimentů byly pro každou scénu zřejmě tři nejlépe hodnocené kompresní metody. Pro scénu *BBB* to byly: H.264, Xvid a VC-1. Pro scénu *RushFieldCuts*: H.264, VC-1 a Theora. Jako nejlepší pro scénu *Tractor* byly hodnoceny metody: H.264, Xvid a VC-1. V případě scény *SnowMnt* bylo nutno provést „rozstřel“ mezi metodami MPEG-2 Part 2, Theora a Xvid. Z tohoto upřesnění

scéna	BBB		SnowMnt		RushFieldCuts		Tractor	
kodek	H.264	Dirac	H.264	Dirac	H.264	Dirac	H.264	Dirac
body	5	1	2	4	3	3	3	3
kodek	MPEG-2	Xvid	MPEG-2	Xvid	MPEG-2	Xvid	MPEG-2	Xvid
body	1	5	1	5	0	6	0	6
kodek	Theora	VC-1	Theora	VC-1	Theora	VC-1	Theora	VC-1
body	2	4	4	2	0	6	6	0

Tabulka 7.4: Soubor experimentů – 3.

vyšla nejlépe ohodnocena Theora. Tudíž byla zařazena do následujícího kola experimentů spolu s Dirac a VC-1.

Překvapujícím prvkem bylo hodnocení kodeku Dirac v případě statické scény. Při použití objektivních metod se totiž v tomto případě jevil jako značně zaostávající (např. obrázek 6.4). S odstupem nejhorší byl experimentátory označen kodek standardu MPEG-2 Part 2, který oproti ostatním kodekům naprosto propadl v preferencích experimentátorů (viz tabulka 7.4). Jedinou oblastí, kde se částečně vyrovnal konkurentům byla statická scéna.

7.3 Výsledky druhého kola experimentů

Druhé kolo experimentů bylo provedeno s již zmíněnými kodeky, které se umístily na prvních třech místech dle preferencí experimentátorů.

Čtvrtý soubor experimentů byl proveden dne 8.5.2011 se 6 experimentátory.

scéna	BBB			SnowMnt		
kodek	H.264	Xvid	VC-1	VC-1	Dirac	Theora
body	2	2	2	1	4	1
scéna	RushFieldCuts			Tractor		
kodek	H.264	Theora	VC-1	H.264	Xvid	VC-1
body	2	3	1	3	2	1

Tabulka 7.5: Soubor experimentů – 4.

Pátý soubor experimentů byl proveden dne 10.5.2011 se 3 experimentátory.

scéna	BBB			SnowMnt		
kodek	H.264	Xvid	VC-1	VC-1	Dirac	Theora
body	2	1	0	0	0	3
scéna	RushFieldCuts			Tractor		
kodek	H.264	Theora	VC-1	H.264	Xvid	VC-1
body	2	0	1	3	0	0

Tabulka 7.6: Soubor experimentů – 5.

Poslední soubor experimentů byl proveden dne 13.5.2011 se 7 experimentátory.

scéna	BBB			SnowMnt		
kodek	H.264	Xvid	VC-1	VC-1	Dirac	Theora
body	4	2	1	1	0	6
scéna	RushFieldCuts			Tractor		
kodek	H.264	Theora	VC-1	H.264	Xvid	VC-1
body	6	1	0	5	1	1

Tabulka 7.7: Soubor experimentů – 6.

Tyto výsledky ukazují, že ve scénách, ve kterých byla zastoupena metoda H.264 – *BBB*, *RushFieldCuts* a *Tractor* – zvítězila. Tato skutečnost plně potvrzuje výsledky dosažené za pomoci objektivních metrik v kapitole 6. Ve statické scéně *SnowMnt* podává nejlepší výsledky kodek Theora. Tato skutečnost je v souladu s měřením za pomoci objektivních metrik, především pak v případě MSSIM (např. obrázek 6.4).

8 Praktické shrnutí výsledků

8.1 MPEG-2 Part 2

Skončil jako poslední napříč většinou hodnocení. Jeho teoretické předpoklady jej stavěly do této pozice, avšak rozdíl oproti ostatním metodám byl mnohdy skutečně propastný (např. obrázek 6.8). Také použitý kodek nedokázal správně dodržet datový tok při jeho nejnižších testovaných hodnotách (viz kapitola 6.5). Silnou devizou tohoto kodeku byla rychlost komprese, která byla druhá nejrychlejší.

8.2 H.264/MPEG-4 Part 10

Jde o nejlépe hodnocenou metodu z testované množiny, především v objektivní části. Svého hlavního konkurenta VC-1 dokázal porazit ve většině případů. V subjektivní části byl krom statické scény hodnocen velice dobře. Často byly v těchto případech veškeré preference přiřazeny této metodě. Použitá implementace tohoto standardu je abnormálně optimalizovaná, proto bylo překvapivě dosaženo velice rychlého času komprese. Tuto skutečnost by bylo vhodné ověřit na dalších implementacích, které jsou však často komerčního charakteru.

8.3 Xvid

Tato metoda ve všech měření prokázala, že je řádově lepší než MPEG-2 Part 2, v několika případech byla dokonce soupeřem pro H.264 (např. obrázek 6.6). Zátěž systému byla ještě nižší než v případě MPEG-2 Part 2, což jí řadí na první místo v rychlosti komprese (viz obrázek 6.15). Problematickou částí však bylo nesprávné komprimování při nejnižším testovaném datovém toku.

8.4 WMV9/VC-1

H.264 nedokázal překonat jak v objektivní tak subjektivní části a jednoznačně se tak jevil jako horší z této dvojice. Jeho výsledky ho však řadí výše než zástupce MPEG-4 Part 2 ASP Xvid. Jako jednoznačný zápor se též jeví značná zátěž systému při enkódování vzhledem k neexistenci většího množství VC-1 enkodérů.

8.5 Dirac

Waveletový kodek dle standardu Dirac podle objektivních metrik zaujal předposlední pozici, když svými výsledky překonal pouze MPEG-2 Part 2. Navzdory podprůměrným výsledkům v objektivní části byl ovšem v subjektivní části vybrán do

druhého kola v případě statické scény (viz kapitola 7.3). Zřejmě díky své nevyzrállosti působila použitá knihovna *libschroedinger* značné problémy. Odmítala například spolupracovat s FFmpeg a při dekodování byl počáteční símek nekolikrát zkopírován.

8.6 Theora

Kodek založený na standardu Theora podával výsledky v průměru blízké Xvid, k čemuž ho předurčovaly jeho teoretické předpoklady. Jeho vhodnost pro statické scény se prokázala v objektivních i subjektivních testech, kde v případě subjektivního testování předstihla i H.264 – kapitola 7.3. Jako značný problém se ovšem jeví velké zatížení systému. Použitá oficiální implementace enkodéru je totiž pouze jednovláknová a proto byla doba komprese z posuzovaných metod jednoznačně nejdelší (viz obrázek 6.15).

9 Závěr

V obou částech srovnání provedeného v rámci této práce se jevila jako nejlepší kompresní metoda postavená na standardu *H.264*. Tato metoda se umístila na prvním místě ve všech zkoumaných scénách a takřka všech datových tocích. Pouze v případě statické scény byla jako nejlepší subjektivně hodnocena metoda *Theora*. Naopak nejhorší výsledky podával kodek postavený na *MPEG-2 Part 2*, následovaný standardem *Dirac*. Tyto výsledky se ve většině shodují s teoretickými předpoklady, které jsou shrnuté v příslušné kapitole 3. Podrobnosti o výsledcích srovnání obsahuje předešlá kapitola s číslem 8.

Veškeré použité komprimační metody byly prostudovány a posouzeny z teoretického hlediska. Metody posuzování kvality zkomprimovaného digitálního videa byly zpracovány a ty nejvhodnější byly v příslušné části též prakticky použity. Také byla vytvořena testovací množina obsahující nekomprimované charakterické scény, které byly následně podrobeny komprimaci. Na závěr byla posouzena kvalita provedené komprese za použití vybraných metod. Všechny cíle práce byly tedy tímto vypracováním splněny v souladu se zadáním.

Jelikož komprese digitálního videa je neustále se vyvíjející obor, nelze v jednom krátkém časovém období obsáhnout veškeré významné nebo progresivní metody. Tuto práci by tudíž bylo zcela jistě možné v budoucnu rozšířit. Především doplnit další zajímavé kompresní metody. V úvahu by například připadal otevřený kodek *On2 VP8*¹³³, který má velice nadějně vyhlídky na rozšíření díky prosazování formátu WebM.

¹³³*The WebM Project | Documentation* [online]. c2011, [cit. 2011-05-25]. Dostupné z: <<http://www.webmproject.org/code/specs/>>.

Reference

- [5] LIVINGSTONE, Margaret. *Vision and Art: The Biology of Seeing*. 2002. New York. ISBN 0-8109-0406-3.
- [6] TUDOR, P.N. *MPEG-2 video compression* [online]. 17.8.1997, [cit. 2010-10-05]. Dostupné z: <http://www.bbc.co.uk/rd/pubs/papers/paper_14/paper_14.shtml>.
- [9] JAIN, A.K. *Fundamentals of Digital Image Processing*. 1989. 592 s. ISBN 0133361659.
- [14] SALOMON, D. – MOTTA, G. *Handbook of Data Compression*. 5th edition, 2010. 1340 s. ISBN 978-1-84882-902-2.
- [23] RICHARDSON, Iain E. G. *H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia*. 1st edition, 2003. 320 s. ISBN 978-0-470-84837-1.
- [25] SULLIVAN, Gary J. – TOPIWALA, Pankaj – LUTHRA, Ajay. *The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions*. [online]. srpen 2004, [cit. 2010-10-17]. Dostupné z: <<http://www.fastvdo.com/spie04/spie04-h264overviewPaper.pdf>>.
- [34] SWAIN, Robert. *H.264 Profiles and Levels* | robert.swain [online]. [cit. 2010-10-18]. Dostupné z: <<http://rob.opendot.cl/index.php/useful-stuff/h264-profiles-and-levels/>>.
- [56] TOMKO, Heather. *SMPTE Releases VC-1 Standard - SMPTE.org* [online]. 3.4.2006, [cit. 2010-11-8]. Dostupné z: <http://www.smpte.org/news/pr/view?item_key=a135f13b173a982bb71f1cd3ee4403671fcf2057>.
- [61] LOOMIS, J. – WASSON, M. *VC-1 Technical Overview* [online]. říjen 2007, [cit. 2010-11-8]. Dostupné z: <<http://www.microsoft.com/windows/windowsmedia/howto/articles/vc1techoverview.aspx>>.
- [71] BORER, T. – DAVIES, T. - SURAPARAJU, A. *BBC R&D White paper WHP 124 - Dirac video compression* [online]. září 2005, [cit. 2010-11-25]. Dostupné z: <<http://www.bbc.co.uk/rd/pubs/whp/whp-pdf-files/WHP124.pdf>>.
- [72] DAVIES, Thomas. *Dirac: Wavelet transform* [online]. 6.2.2008, [cit. 2010-11-26]. Dostupné z: <http://dirac.sourceforge.net/documentation/algorithm/algorithm/wlt_transform.xht>.
- [94] FLIEGEL, Karel. *Kvalita obrazu a modelování HVS* [online]. 20.11.2008, [cit. 2011-02-25]. Dostupné z: <<http://radio.feld.cvut.cz/courses/w2ZKT/materialy.php?akce=dlf&zdroj=vpm&fkey=7&xtgt=2f686f6d652f53657276696365732f>>.
- [95] WANG, Zhou – BOVIK, Alan C. – SHEIKH, Hamid R. – SIMONCELLI, Eero P. *The SSIM Index for Image Quality Assessment* [online]. 11.2.2011, [cit. 2011-02-25]. Dostupné z: <<https://ece.uwaterloo.ca/~z70wang/research/ssim/>>.

- [98] VEŘÁT, Ivo. *Televizní, rozhlasová a multimediální technika - hodnocení kvality videa a obrazu* - [online]. 5.11.2010, [cit. 2011-02-26]. Dostupné z: <http://stag.zcu.cz/fel/kae/TRM/rok_2010/vertat_4/TRM_Vertat_4_2010.pdf>.
- [128] HÖHNE, Hermann. *AVISynth resizing filter comparasion* [online]. 26.7.2007, [cit. 2011-03-21]. Dostupné z: <<http://hermidownloads.craqstar.de/videoresizefiltercomparasion/>>.

Seznam tabulek

2.1	Profily a levely MPEG-2 Part 2.	9
2.2	H.264 profily a jejich vlastnosti.	14
2.3	Používané profily DivX Codec.	17
2.4	Profily používané Xvid.	19
2.5	Profily používané VC-1.	22
6.1	Velikosti – testovací scéna <i>BBB</i>	49
6.2	Velikosti – testovací scéna <i>SnowMnt</i>	49
6.3	Velikosti – testovací scéna <i>RushFieldCuts</i>	49
6.4	Velikosti – testovací scéna <i>Tractor</i>	49
7.1	Soubor experimentů – 1.	53
7.2	Soubor experimentů – 2a.	53
7.3	Soubor experimentů – 2b.	53
7.4	Soubor experimentů – 3.	54
7.5	Soubor experimentů – 4.	55
7.6	Soubor experimentů – 5.	55
7.7	Soubor experimentů – 6.	55

Seznam obrázků

2.1	Zjednodušené schéma enkódování MPEG-2 Part 2.	8
2.2	Možnosti dělení makrobloku H.264.	11
2.3	Ilustrace využití <i>GMC</i> pomocí globálního vektoru v DivX Codec. . .	16
2.4	Zjednodušené schéma enkódování Dirac.	25
2.5	Proud snímků za použití kodeku Theora.	27
4.1	Diagram činnosti MSU VQMT.	33
5.1	Snímek z testovací scény <i>BBB</i>	38
5.2	Snímek z testovací scény <i>SnowMnt</i>	39
5.3	Snímek z testovací scény <i>RushFieldCuts</i>	39
5.4	Snímek z testovací scény <i>Tractor</i>	40
6.1	MSSIM nad luminační složkou – jednotlivé snímky.	41
6.2	MSSIM nad luminační složkou – průměrné hodnoty.	42
6.3	MSSIM nad luminační složkou pro scénu <i>BBB</i>	43
6.4	MSSIM nad luminační složkou pro scénu <i>SnowMnt</i>	43
6.5	MSSIM nad luminační složkou pro scénu <i>RushFieldCuts</i>	44
6.6	MSSIM nad luminační složkou pro scénu <i>Tractor</i>	44
6.7	PSNR nad luminační složkou pro scénu <i>BBB</i>	45
6.8	PSNR nad luminační složkou pro scénu <i>SnowMnt</i>	45
6.9	PSNR nad luminační složkou pro scénu <i>RushFieldCuts</i>	46
6.10	PSNR nad luminační složkou pro scénu <i>Tractor</i>	46
6.11	MSSIM nad chrominanční složkou U-YUV.	47
6.12	MSSIM nad chrominanční složkou V-YUV.	47
6.13	PSNR nad chrominanční složkou U-YUV.	48
6.14	PSNR nad chrominanční složkou V-YUV.	48
6.15	Průměrné doby běhu.	51

Seznam použitých zkratek

3G	3rd generation mobile telecommunications - označení pro skupinu standardů pro třetí generaci mobilních komunikačních sítí.
4MV	4 Motion Vectors – typ pohybové kompenzace, kdy je mekroblok rozložen na čtyři bloky s příslušujícími pohybovými vektory.
AC	DCT koeficient se souřadnicemi jinými než [0;0].
AMC	Adaptive Multiple Consecutive – technika vylepšující plynulost obrazu zařazováním B-snímků v DivX Codec.
ASC	Adaptive Single Consecutive – technika omezující četnost B-snímků v „rychlých“ scénách v DivX Codec.
ASO	Arbitrary Slice Ordering – metoda řazení slice pro lepší odolnost proti chybám přenosu; používané u H.264.
ASP	Advanced Simple Profile – jeden z profilů standardu MPEG-4 Part 2/MPEG-4 Visual.
AVC	Advanced Video Coding – používaná zkratka pro videokodek označovaný jako standard H.264/MPEG-4 Part 10.
AVI	Audio Video Interleave – rozšíření multimediální kontejner.
AVS	AviSynth – rozříšený nástroj pro editaci video souborů.
B-VOP	Bidirectional-Visual Object Pane – MPEG-4 Part 2 ekvivalent k B-snímku.
CABAC	Context-Adaptive Binary Arithmetic Coding – metoda bezztrátového kódování; efektivní a výpočetně méně náročná varianta používaná u H.264.
CAVLC	Context-Adaptive Variable-Length Coding – metoda bezztrátového kódování; efektivnější varianta VLC.
Cb	modrý chrominanční/barevný komponent.
Cg	zelený chrominanční/barevný komponent.
Co	oranžový chrominanční/barevný komponent.
Cr	červený chrominanční/barevný komponent.
CSV	Comma-Separated Values – základní souborový formát; založený na oddělování datových záznamů čárkami.
DC	DCT koeficient se souřadnicemi [0;0].
DCT	Discrete Cosine Transform – diskrétní kosinová transformace; nejčastěji typu II.
DCT-II	Discrete Cosine Transform Type II – diskrétní kosinová transformace typu II; často označována pouze jako „DCT“.
DPB	Decoded Picture Buffer – buffer určený k udržování referenčních snímků při enkódování.

DSIS	Double Stimulus Impairment Scale – metoda subjektivního měření kvality videa; založené na opakovaném porovnávání předlohy a reference.
DTV	Digital Television – souhrnné označení pro digitální přenos televizního signálu.
DVB	Digital Video Broadcasting – skupina standardů pro digitální způsob přenosu televizního vysílání.
DVD	Digital Versatile Disc/Digital Video Disc – formát digitálního optického datového nosiče.
FFmpeg	otevřená knihovna pro práci s mnoha různými druhy audiovizuálních dat.
FFMS2	FFmpegSource – AviSynth knihovna založená na FFmpeg.
FMO	Flexible Macroblock Ordering – metoda řazení slice pro lepší odolnost proti chybám přenosu; používané u H.264.
FRExt	Fidelity Range Extensions – rozšíření standardu H.264/MPEG-4 Part 10, které přineslo další možnosti enkódování ve vysokých rozlišeních a datových tocích (např. High profil, „4:4:4“, podpora pro 3D).
FullHD	plné HD rozlišení; nejčastěji 1080p.
GMC	Global Motion Compensation – globální pohybová kompenzace; využívaná např. MPEG-4 Part 2.
GMV	Global Motion Vector – globální pohybový vektor; využívá jej např. MPEG-4 Part 2.
GNU	GNU's Not Unix – projekt vyvíjející svobodný software unixového typu.
GNU GPL	GNU General Public License – svobodná licence pro softwarové produkty; původně určena pro projekt GNU.
GOP	Group of Pictures – sekvence snímků mezi dvěma intra-snímků.
HD	High-definition – obecné označení pro vysoké rozlišení; typicky 720p a vyšší.
HD DVD	High-Definition DVD – formát optického datového nosiče určeného pro video data o vysokém rozlišení.
HDTV	High-definition Television – formát televizního vysílání ve vysokém rozlišení; počet řádků obrazu 720 nebo 1080 při 23,976 fps a vyšším.
HPEL	Half Pixel Motion Estimation – metody vyhledávání makrobloku s přesností půl bodu; používané např. v MPEG-2 Part 2.
HTML5	HyperText Markup Language 5 – W3C rozšiřující standard jazyka HTML.
I-VOP	Intra-Visual Object Pane – MPEG-4 Part 2 ekvivalent k I-snímku.
JPEG2000	obrazový komprimační standard založený na waveletové transformaci.
luma	zkratka pro luminační/jasovou složku.
MPEG	Motion Picture Experts Group – pracovní skupina vyvíjející standardy používané pro kódování audiovizuálních informací.

MSE	Mean Squared Error – střední kvadratická chyba; jde o metodu měření chyby pedikce a následné skutečnosti.
MSSIM	Mean Structure Similitary Index – střední hodnota indexu strukturální podobnosti; metoda podobná SSIM, avšak aplikovaná na celý snímek.
MSU	Moscow State University – Lomonosovova univerzita v Moskvě.
MV	Motion Vector – pohybový vektor; využívaný při mezinímkové kompresi.
MVC	Multiview Video Coding – rozšíření standardu H.264/MPEG-4 Part 10, které přineslo rozšířené možnosti enkódování více datových proudů určených pro 3D zobrazení.
PSNR	Peak Signal-to-Noise Ratio - špičkový poměr signálu a šumu; vztah mezi maximy energie signálu a šumu.
QPEL	Quarter Pixel Motion Estimation – metody vyhledávání makrobloku s přesností půl bodu; používané např. v MPEG-4 ASP.
RGB	barevný model; označení z anglického Red-Green-Blue.
RLE	Run-Length Encoding – kódování délkou běhu; metoda bezztrátové komprese.
SCAJC	Stimulus Comparison Adjectival Categorical Judgement – metoda subjektivního měření kvality videa; založené na souběžném porovnávání dvou video souborů.
SMPTE	The Society of Motion Picture and Television Engineers – mezinárodní organizace udržující standardy audiovizálního průmyslu.
SS	Single Stimulus – metoda subjektivního měření kvality videa; založené na porovnání více předloh k jedné referenci za sebou.
SSIM	Structural Similarity Index – index strukturální podobnosti; metoda porovnání a vyjádření podobnosti výřezů dvou obrazů.
SVC	Scalable Video Coding – rozšíření standardu H.264/MPEG-4 Part 10, které přineslo rozšířené možnosti enkódování a editace více datových proudů.
SVCD	Super Video Compact Disc – alternativní audiovizuální formát k DVD-Video.
VBR	Variable Bit Rate – proměnlivý datový tok; využívá se v audiovizuální technice.
VBSMC	Variable Block-Size Motion Compensation – pohybová kompenzace s dynamickou velikostí bloku.
VC-1	audiovizuální standard známý též jako SMPTE 421M.
VC-2	standard skupiny SMPTE pro profesionální bezztrátový kodek Dirac Pro.
VHS	Video Home System – standard pro domácí video na videokazetách.
VLC	Variable-Length Code – rodina typů bezztrátové komprese; zástupce např. Huffmanovo kódování.

VO	Visual Object – označení visuálního objektu používaného v MPEG-4 Part 2, který nemusí být nutné pravidelného tvaru.
VOP	Visual Object Pane – MPEG-4 Part 2 označení snímku.
VOS	Visual Object Sequence – MPEG-4 Part 2 ekvivalent k GOP.
VQMT	Video Quality Measurement Tool – softwarový nástroj pro vyhodnocování objektivní kvality video souborů.
WMV	Windows Media Video – proprietární video formát vyvíjený společností Microsoft.
WVGA	Wide Video Graphics Array – označení používané pro širokoúhlé VGA rozlišení.
Y	jasová složka obrazu v rodině YUV modelů; někdy bývá též nazývaná jako luma.
YCbCr	barevný model využívaný v audiovizuální technice; Y – jasová složka, Cb – modrý chrominanční/barevný komponent, Cr – červený chrominanční/barevný komponent.
YCoCg	barevný model využívaný v audiovizuální technice; Y – jasová složka, Co – oranžový chrominanční/barevný komponent, Cg – zelený chrominanční/barevný komponent.
YUV	barevný model využívaný v audiovizuální technice; Y – luminační/jasová složka, U a V – chrominanční/barevné složky.