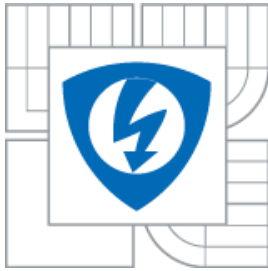




VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH
TECHNOLOGIÍ
ÚSTAV RADIOELEKTRONIKY

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF RADIO ELECTRONICS

MĚŘENÍ KVALITY PRO HEVC

VIDEO QUALITY MEASUREMENT FOR HEVC

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. EVA KLEJMOVÁ

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. MARTIN SLANINA, Ph.D.

BRNO 2014



VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

Ústav radioelektroniky

Diplomová práce

magisterský navazující studijní obor
Elektronika a sdělovací technika

Studentka: Bc. Eva Klejmová
115199

ID:

Ročník: 2

Akademický rok: 2013/2014

NÁZEV TÉMATU:

Měření kvality pro HEVC

POKYNY PRO VYPRACOVÁNÍ:

Seznamte se s principy objektivního měření kvality pro komprimované videosekvence. Prostudujte dostupnou dokumentaci kodeku High Efficiency Video Coding (HEVC) a vytvořte databázi komprimovaných videosekvencí.

Aplikujte různé objektivní kvalitativní metriky na vhodný soubor kódovaných videosekvencí a určete jejich vhodnost pro vybraný kodek s ohledem na výsledky subjektivních testů.

DOPORUČENÁ LITERATURA:

- [1] SULLIVAN, G. J., OHM, J.-R., HAN, W.-J., WIEGAND, T. Overview of the High Efficiency Video Coding (HEVC) Standard. IEEE Transactions on Circuits and Systems for Video Technology, 2012, vol. 22, no. 12, p. 1649 - 1668.
- [2] KRATOCHVÍL, T.; SLANINA, M. Digital Video Image Quality. In Digital Video, Edited by: Floriano De Rango, Chapter 23: Digital Video Image Quality. Vienna: Intech, 2010.

Termín zadání: 10.2.2014

Termín odevzdání: 23.5.2014

Vedoucí práce: Ing. Martin Slanina, Ph.D.

Konzultanti diplomové práce:

doc. Ing. Tomáš Kratochvíl, Ph.D.
Předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Diplomová práce se zabývá popisem základních metrik pro objektivní a subjektivní hodnocení kvality videa a analýzou jejich aplikovatelnosti pro HEVC. Dále obsahuje základní popis komprimačního standardu H.265/HEVC. Hlavním zaměřením práce je vytvoření vhodné databáze komprimovaných videosekvencí. Jsou uvedeny parametry a vlastnosti referenčního enkodéru HM-12. Dále jsou na vytvořenou databázi aplikovány vybrané metody objektivního hodnocení kvality. Součástí práce je návrh metody pro subjektivní hodnocení kvality, aplikace této metody a s tím spojený sběr dat. Výsledná data jsou pak statisticky zpracována a je diskutována jejich korelace s objektivními testy.

KLÍČOVÁ SLOVA

kvalita videa, objektivní hodnocení, subjektivní hodnocení, PSNR, SSIM, VQM, H.265, HEVC

ABSTRACT

This diploma thesis deals with standard objective and subjective video quality assessments and with analysis of their applicability to HEVC. Also basic description of video compression standard H.265/HEVC is presented. The main focus of the thesis is a creation of the database of compressed video sequences. Important parameters and features of the reference encoder HM-12 are discussed. Selected methods of objective video quality assessments are implemented on the created database. A part of this thesis is also a suggestion of method for objective video quality assessment, application of this method and associated data collection. Final data is statistically analyzed and its correlation with objective tests is discussed.

KEYWORDS

video quality, objective video quality, subjective video quality, PSNR, SSIM, VQM, H.265, HEVC

Klejmová, E. Měření kvality pro HEVC. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií. Ústav radioelektroniky, 2014. 46 s., 10 s. příloh. Diplomová práce. Vedoucí práce: Ing. Martin Slanina, Ph.D.

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma Měření kvality pro HEVC jsem vypracovala samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušila autorská práva třetích osob, zejména jsem nezasáhla nedovoleným způsobem do cizích autorských práv osobnostních a/nebo majetkových a jsem si plně vědoma následků porušení ustanovení § 11 a následujících zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne

.....

(podpis autora)

PODĚKOVÁNÍ

Děkuji vedoucímu semestrální práce Ing. Martinu Slaninovi, Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé diplomové práce.

V Brně dne

.....

(podpis autora)

Experimentální část této diplomové práce byla realizována na výzkumné infrastruktuře
vybudované v rámci projektu CZ.1.05/2.1.00/03.0072
Centrum senzorických, informačních a komunikačních systémů (SIX)
operačního programu Výzkum a vývoj pro inovace.

OBSAH

Seznam obrázků	viii
Seznam tabulek	x
Úvod	1
1 Hodnocení kvality obrazu	2
1.1 Objektivní metriky	2
1.1.1 Peak Signal to Noise Ratio (PSNR).....	2
1.1.2 Structural Similarity Index (SSIM).....	3
1.1.3 Video Quality Model (VQM)	4
1.2 Subjektivní metriky.....	7
1.2.1 Metody využívající dvě sekvence.....	8
1.2.2 Metody využívající jednu sekvenci	9
2 H.265/HEVC	11
2.1 Stromová struktura.....	11
2.2 Transformační kódování	12
2.3 Entropické kódování	12
2.4 Intrapredikce	13
2.5 Interpredikce	13
2.6 Rekonstrukční filtr	13
3 Databáze Videosekvencí	14
3.1 Zdrojové sekvence	14
3.2 Časová a prostorová aktivita.....	16
3.3 Kódování sekvencí.....	17
3.4 Zjištěné vlastnosti	18
3.5 Vytvořená databáze.....	22
4 Aplikace objektivních metrik	24
4.1 PSNR	24
4.2 SSIM	25
4.3 General VQM	25

5	Subjektivní hodnocení	30
5.1	Metoda sběru dat.....	30
5.2	Metoda hodnocení.....	32
5.3	Sada testovacích sekvencí.....	32
5.4	Program pro sběr dat.....	33
5.5	Složení panelu hodnotitelů.....	34
5.6	Zpracování dat	35
5.7	Výsledky hodnocení	37
5.7.1	Bitová rychlost.....	37
5.7.2	PSNR	39
5.7.3	SSIM	40
5.7.4	VQM	41
5.7.5	Diferenční stupnice	42
5.7.6	Korelační koeficienty.....	43
5.7.7	Intervaly spolehlivosti CI.....	43
5.7.8	Další zjištěné vlastnosti	45
	Závěr	46
	Literatura	47
	Seznam symbolů, veličin a zkratk	49
	Seznam Příloh	50

SEZNAM OBRÁZKŮ

Obrázek 1-1 Časová struktura DSIS testu [12].....	8
Obrázek 1-2 Hodnotící stupnice pro DSIS metodu [12].....	8
Obrázek 1-3 Hodnotící stupnice pro DSCQS metodu [12].	9
Obrázek 1-4 Časová struktura ACR testu [14].	9
Obrázek 1-5 Hodnotící stupnice pro ACR metodu [14].	10
Obrázek 2-1 Možné velikosti PB [17].	12
Obrázek 2-2 Rozdělení CTB do CB a TB	12
Obrázek 3-1 Spektrální a temporální aktivita použitých zdrojových sekvencí.	17
Obrázek 3-2 Časová náročnost kódování sekvencí v SD rozlišení.	19
Obrázek 3-3 Průměrná relativní odchylka bitové rychlosti u full HD sekvencí.....	20
Obrázek 3-4 Průměrná relativní odchylka bitové rychlosti u HD sekvencí.	20
Obrázek 3-5 Průměrná relativní odchylka bitové rychlosti u SD sekvencí.....	20
Obrázek 3-6 Tři po sobě jdoucí snímky sekvence Ducks_take_off_480p_500kbps.....	21
Obrázek 3-7 Rozdíl zdrojové a kódované sekvence (1300kbps).....	22
Obrázek 4-1 Závislost PSNR na bitové rychlosti pro full HD sekvence.....	26
Obrázek 4-2 Závislost PSNR na bitové rychlosti pro HD sekvence	27
Obrázek 4-3 Závislost PSNR na bitové rychlosti pro SD sekvence.	27
Obrázek 4-4 Závislost SSIM na bitové rychlosti pro full HD sekvence.	28
Obrázek 4-5 Závislost SSIM na bitové rychlosti pro SD sekvence.....	28
Obrázek 4-6 Závislost VQM na bitové rychlosti pro full HD sekvence.	29
Obrázek 4-7 Závislost VQM na bitové rychlosti pro SD sekvence.....	29
Obrázek 5-1 Dotazník v programu pro testování.....	33
Obrázek 5-2 Obrazovka pro hodnocení.	34
Obrázek 5-3 DMOS v závislosti na bitové rychlosti pro rozlišení 720p.....	38
Obrázek 5-4 DMOS v závislosti na bitové rychlosti pro rozlišení 480p.....	38
Obrázek 5-5 DMOS v závislosti na PSNR pro rozlišení 720p.....	39
Obrázek 5-6 DMOS v závislosti na PSNR pro rozlišení 480p.....	39
Obrázek 5-7 DMOS v závislosti na SSIM pro rozlišení 720p.....	40
Obrázek 5-8 DMOS v závislosti na SSIM pro rozlišení 480p.....	40
Obrázek 5-9 DMOS v závislosti na VQM pro rozlišení 720p.....	41
Obrázek 5-10 DMOS v závislosti na VQM pro rozlišení 480p.....	41

Obrázek 5-11 DMOS v závislosti na diferenčním PSNR pro rozlišení 720p.....	42
Obrázek 5-12 DMOS v závislosti na diferenčním PSNR pro rozlišení 480p.....	42
Obrázek 5-13 Velikosti CI hodnot DMOS (vlevo 720p, vpravo 480p).....	44

SEZNAM TABULEK

tabulka 3-1 Parametry původních sekvencí.	14
tabulka 3-2 Parametry a popis full HD zdrojových sekvencí.	15
tabulka 3-3 Rozlišení SD a HD zdrojových sekvencí.	15
tabulka 3-4 Rozdíly mezi PSNR mezi snímky 1-224 a snímky 225-250 sekvence Skyfall_1.	22
tabulka 4-1 Převodní hodnoty PSNR.	25
tabulka 5-1 Parametry použitých přenosných zařízení.	30
tabulka 5-2 Hodnoty PSNR mezi H.265 a H.264 sekvencemi.	31
tabulka 5-3 Procentuální zastoupení jednotlivých skupin respondentů.	35
tabulka 5-4 Hodnoty korelačních koeficientů SROCC.	43
tabulka 5-5 Hodnoty korelačních koeficientů PLCC.	43

ÚVOD

S rozvojem požadavků na kvalitu videosekvencí, spojeným s rostoucími nároky na jejich rozlišení a snímkovou rychlost při zachování nízké bitové rychlosti, musí nutně docházet k rozvoji a vývoji nových metod a standardů užívaných ke kompresi těchto sekvencí. Jeden z relativně nových standardů je i H.265/HEVC, který je v této práci představen. Umožňuje značnou úsporu dat oproti dnes běžně používaným standardům. Jako příklad můžeme jmenovat MPEG-2 nebo MPEG-4/AVC. S vývojem nových metod však dochází i ke změnám ve strukturách kódovaného videa, které se projevují převážně při vysokém stupni komprese. Některé ze zavedených metrik pro objektivní hodnocení kvality však tyto změny nemusí správně zaznamenat a jejich výstupní hodnoty pak nevypovídají nic o skutečné kvalitě obrazu. Proto s tímto vývojem dochází i k dalšímu rozvoji objektivních metrik. Ty mohou být díky větší dostupné výpočetní kapacitě moderní elektroniky čím dál tím sofistikovanější. Nové metody berou v úvahu dostupné znalosti o lidském zraku a využívají model lidského vidění. Jejich výsledky pak více odpovídají subjektivním testům, které poskytují nejuvěrnější informaci o skutečně vnímané kvalitě lidskými pozorovateli. V současné době je kladen důraz na automatizaci procesu a subjektivní metriky začínají sloužit spíše jako kontrola a ověření metrik objektivních.

Cílem práce je seznámit se se základními principy metod, používaných pro hodnocení kvality obrazu, potažmo videa. Dále je to pak vytvoření vhodné databáze videosekvencí, která bude sloužit k testování. Zároveň bude navržena metoda subjektivního hodnocení, které bude společně s objektivními testy na databázi aplikována. Součástí práce je diskuze dosažených výsledků.

Diplomová práce je členěna celkem do pěti kapitol. V první kapitole jsou popsány metody hodnocení kvality videosekvencí. Jsou popsány vybrané objektivní metody s plnou a částečnou referencí. Dále jsou popsány nejběžnější metody používané při subjektivním hodnocení kvality. Jsou zahrnuti zástupci jak metod, které využívají dvě sekvence (referenční a testovanou), tak metod, které používají pouze jednu sekvenci (testovanou). Druhá kapitola se zabývá základní popisem standardu H.265/HEVC. Třetí kapitola se věnuje popisu postupu při vytváření databáze videosekvencí, komprimovaných pomocí kodeku H.265. Je popsán výběr zdrojových snímků a jejich následné kódování pomocí referenční implementace HM-12. Dále jsou popsány vybrané parametry a chování kodéru a vlastnosti vytvořené databáze. Ve čtvrté kapitole jsou na vytvořenou databázi aplikovány vybrané metriky pro hodnocení objektivní kvality obrazu. Jsou shrnuty získané výsledky. Poslední, pátá kapitola se věnuje návrhu metody pro subjektivní testy. Následuje její realizace. V této kapitole je také popsán samotný sběr dat a jejich následné vyhodnocení. V závěru této kapitoly jsou porovnány výsledky získaných subjektivních a objektivních testů.

1 HODNOCENÍ KVALITY OBRAZU

Vzhledem k potřebě jednoznačného zhodnocení kvality obrazu ve videosekvencích byly vyvinuty různé metody a metriky. Obecně je lze podle přístupu rozdělit do dvou skupin. První skupinou jsou objektivní metriky, založené na výpočtech předem vybraných parametrů podle jasně definovaných vztahů. Druhou skupinou jsou subjektivní metody, které využívají hodnocení reálných pozorovatelů. Tento postup je však relativně drahý a časově náročný a proto se stále více uplatňují metody objektivní.

1.1 Objektivní metriky

Jak bylo řečeno, objektivní metriky poskytují popis kvality obrazu, založený na více, či méně složitých algoritmech. Nejjednodušší metody poskytují jen hrubý odhad kvality obrazu. S rozšiřováním poznatků o lidském zraku a vidění obecně dochází ke zlepšování model lidského vidění. Tento matematický popis je tedy čím dál tím sofistikovanější a při jeho použití dostáváme věrohodnější výsledky. Velikou výhodou těchto metrik je jednodušší implementace a menší časová náročnost oproti subjektivním testům. Obecně se dají rozdělit do třech základních skupin. Metody s plnou referencí (full-reference), pro které je potřeba mít k dispozici kompletní originální videosekvenci. Dále metody s částečnou referencí (reduced-reference), kdy k výpočtu musí být k dispozici alespoň vybrané parametry originálního signálu. Poslední skupinou jsou metody bez reference (no-reference) při kterých se vyhodnocuje pouze výstupní zarušená sekvence. Dále je možné rozlišit, zda se u metody uplatňuje model lidského vidění (HVS) či nikoliv. V následujících podkapitolách jsou popsány metody PSNR a SSIM, které jsou plně referenční a metoda VQM, které potřebuje alespoň částečnou referenci.

1.1.1 Peak Signal to Noise Ratio (PSNR)

PSNR patří mezi objektivní, plně referenční pixelově-orientované metriky. Je jednou z nejjednodušších a také nejrozšířenějších metod pro hodnocení kvality obrazů či videosekvencí. Velkou výhodou je malá výpočetní náročnost, díky které je tato metrika hojně využívána i přes značné nedostatky. Je odvozená od střední kvadratické chyby pixelů. Ve výpočtu tedy nejsou vůbec zohledněny vlastnosti lidského zraku a v určitých případech nemusí vykazovat relevantní výsledky. Příkladem může být šum ve vysoce členité části obrazu (např. scéna se skalami), který nemusí být pozorovatelem vnímán, ale metrikou bude vyhodnocen stejně jako šum v jednodušejší oblasti (např. obloha), který bude pozorovatelem vnímán jako velmi rušivý. Přesto tato metrika může být použita pro prvotní odhad kvality. Vzhledem k tomu, že výstupní obraz může mít vysoký dynamický rozsah zarušení, je PSNR udáváno v decibelové míře. Typické hodnoty jsou v rozmezí 20 dB až 50 dB, přičemž hodnoty kolem 50 dB znamenají vysokou obrazovou kvalitu, [1], [2], [3].

Jak bylo řečeno, základem metody je výpočet střední kvadratické chyby MSE. Ta je definována jako střední hodnota druhých mocnin rozdílů dvou šedotónových hodnot pixelů obrazu. Výpočet MSE se provádí dle následujícího vztahu [1]

$$MSE = \frac{1}{TXY} \sum_t \sum_x \sum_y [I(t, x, y) - \tilde{I}(t, x, y)]^2, \quad (1)$$

kde X a Y jsou rozměry snímku, T je počet snímků a I, \tilde{I} jsou hodnoty chrominančních nebo jasových koeficientů. Hodnota PSNR v decibelech je poté dána vztahem,

$$PSNR_{dB} = 10 \cdot \log_{10} \left(\frac{m^2}{MSE} \right), \quad (2)$$

kde m je maximální hodnota pixelu (např. 255 pro 8 bitovou reprezentaci). Ze vzorců je patrné, že je metrika možno chápat jako poměr špičkového výkonu a šumu. Je počítána pro jasovou a každou chrominanční složku zvlášť a nejčastěji se uvádí aritmetický průměr těchto tří složek nebo pouze PSNR jasového signálu Y.[1]

1.1.2 Structural Similarity Index (SSIM)

SSIM je další plně referenční, objektivní metrika pro hodnocení kvality obrazu. Na rozdíl od PSNR však částečně bere v úvahu model lidského vidění (human visual system, HVS). Základním stavebním kamenem byla modifikace MSE tak, aby chyby v obraze byly hodnoceny podle toho, jak moc jsou pro pozorovatele rušivé. SSIM vychází z předpokladu, že lidský zrak je vysoce citlivý na strukturu pozorované scény. Rozdíly struktur referenčního a zkoumaného obrazu pak mohou poskytnout relativně dobrou informaci o tom, jak je zarušený obraz pozorovatelem vnímán. Vzhledem k tomu, že střední jas a kontrast obrazu nejsou na struktuře závislé je každá z těchto složek porovnávána samostatně a poté vhodně váhována. Vztah pro funkci porovnávání jasu vychází z toho, že lidský zrak je citlivější na relativní změny jasu než na absolutní změně (Weber's law), při stejné změně jasu záleží na jasu v pozadí scény. Obdobně je tomu u funkce porovnání kontrastu. Výpočet porovnání struktury se provede jako poslední. U barevných obrazů se standardně počítá pouze pro jasovou složku Y. Index SSIM nabývá hodnot od -1 do 1, kde hodnota 1 označuje identické obrazy. [28], [4]

Definice SSIM je následující [4]:

$$SSIM(x, y) = l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma, \quad (3)$$

kde l je funkce pro porovnání jasu, c je funkce pro porovnání kontrastu a „ s “ je funkce porovnávající strukturu obrazů. Jednotlivé funkce jsou vyjádřeny vztahy [4]:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (4)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (5)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \quad (6)$$

Pro hodnoty $\alpha = \beta = \gamma = 1$ a $C_3 = C_2 / 2$ dostáváme upravený vztah [4]:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1) \cdot (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1) \cdot (\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (7)$$

kde α, β, γ jsou kladné váhovací koeficienty, μ_x, μ_y jsou střední hodnoty intenzity pixelu, σ_x^2, σ_y^2 jsou rozptyly x a y, σ_{xy} jsou kovariance x a y, $C_i = (K_i L)^2$, kde L je maximální hodnota pixelu v obraze a K jsou malé konstanty běžně udávané $K1 = 0,01$ a $K2 = 0,03$. Dle [4] není výsledná hodnota SSIM na těchto konstantách příliš závislá.

1.1.3 Video Quality Model (VQM)

Je objektivní, plně referenční či částečně referenční metrika hodnocení kvality obrazu. Oproti předešlým metodám je výpočetně náročnější, ale dává relevantnější výsledky. Je totiž navržena tak, aby v co nejvyšší možné míře brala v úvahu model lidského vidění (HVS). Existují dva modely, které jsou zahrnuty v ITU-T J.144, ITU-R BT.1683, ITU-T J.244 a ITU-T J.249. Jmenovitě NTIA General Model (VQM) and Full Reference Calibration a Fast Low Bandwidth Model and Reduced Reference Calibration.

NTIA General Model

Tento model, označován VQM_G je optimalizován pro dosažení vysoké korelace mezi objektivním a subjektivním hodnocením kvality obrazu. Byl standardizován ANSI (American National Standards Institute) v dokumentu T1.801.03 [5] a je zahrnut ve dvou ITU doporučeních [6],[7]. Je vhodný k použití pro širokou škálu výstupních bitových rychlostí videosekvencí a s tím související rozsah výsledné kvality (např. rychlosti od 10 kb/s do 45 Mb/s, MPEG-1/2/4, progresivní i prokládaný formát). [8],[9],[6].

Model využívá částečnou referenci ke zhodnocení celkové kvality. Potřebná data jsou extrahována z originálního signálu. Přídavný kanál potom dosahuje velikosti 9,3 % z původní nekomprimované videosekvence a data potřebná pro kalibraci přidávají dalších 4,7 %. Ve standardu jsou zahrnuty kalibrační techniky, které se provedou před samotným výpočtem VQM_G . Tato kalibrace zahrnuje časové a prostorové zarovnání originálního a výstupního videosignálu, výpočet celkového posunu jasu a určení relevantní části obrazu. Proces prostorového a časového zarovnání iterativně určuje vertikální a horizontální posun mezi originální a výstupní sekvencí. Protože při přenosu může být obraz doplněn o oblasti, které nenesou užitečnou obrazovou informaci (např. černé okraje, poškozené řádky na kraji snímku), jsou pomocí výpočtu PVR (processed valid region) tyto pixely vyjmuty z dalšího zpracování. Ke kalibraci posuvu úrovně musí být obrazy alespoň částečně prostorově a časově zarovnané. Pro jasový a oba chrominancní signály je pak posuv počítán zvlášť na malých čtvercových sub-oblastech nebo blocích. Vzhledem k tomu, že kalibrace je nutná převážně v přenosových systémech a během pouhé komprese videa není potřebná je pro cíle této práce nadbytečná a proto je zde uvedena pouze ve zkratce. Více informací je k dispozici například v literatuře [8],[9],[10].

K samotnému výpočtu VQM dochází po kalibraci (pokud je potřebná). Výsledná hodnota je lineární kombinací sedmi nezávislých parametrů, které jsou nastaveny tak, aby v co největší míře odpovídaly HVS. Data jsou získána ze sub-bloků, definovaných v časové a prostorové oblasti, přičemž sub-blok originálního a zarušeného obrazu si musí odpovídat.

Vzhledem k tomu, že se u lidského zraku uplatňuje vliv maskování (zkreslení je méně výrazné v obraze s vyšší časovou a prostorovou aktivitou), je snížení kvality jednotlivých sub-bloků počítáno pro některé parametry počítáno pomocí Euklidovské vzdálenosti. [8]

$$p = \sqrt{(f_o - f_p)^2 + (f_{o2} - f_{p2})^2}, \quad (8)$$

Nicméně většina parametrů používá k výpočtu podíl

$$p = (f_p - f_o) / f_o, \quad (9)$$

nebo logaritmický podíl

$$p = \log_{10} \left(\frac{f_p}{f_o} \right), \quad (10)$$

kde f_o a f_{o2} jsou parametry originálního obrazu a f_p a f_{p2} jsou odpovídající parametry výstupního obrazu.

Tyto vzorce poskytují pozitivní hodnoty, které naznačují zisk a negativní ztráty. Tím, že jsou zisky a ztráty vyhodnocovány zvlášť, dochází ke zvýšení přesnosti algoritmu. Je to způsobeno tím, že lidský zrak reaguje hůře na přídavné zhoršení (např. blokové artefakty vytvářející ostré hrany), než na ztrátu informace (např. rozmazání obrazu).

Další metodou, jak zajistit co nejlepší přiblížení k subjektivním výsledkům, je prahování použitých parametrů. To je umožněno díky tomu, že pozorovatel má tendenci se zaměřovat na lokalizované chyby, které mají potom výrazný vliv na celkové hodnocení kvality. Citlivost metriky na ojedinělé zarušení je potom snížena dle následujícího předpisu:

$$p' = \begin{cases} 0 & \text{pro } p \leq t \\ p - t & \text{pro } p > t \end{cases}, \quad (11)$$

kde p' je nová hodnota parametru, p je původní hodnota parametru a t je práh vhodně zvolený tak aby co nejvíce odpovídal HVS.

Jak bylo zmíněno, *General model* je počítán pomocí sedmi parametrů. Čtyři z nich jsou založeny na změně jasové složky v prostoru obrazu, dva jsou založeny na kombinaci chrominančních složek a poslední je založen na měření kontrastu a pohybu pomocí jasové složky.

Parametr „*si_loss*“ popisuje ztrátu prostorové informace (např. rozmazání). Využívá SI13 filtru [17] s maskou 13x13 pixelů, která je aplikována v horizontálním a vertikálním směru na jasovou složku. Oba směry jsou poté kombinovány vzorcem (8).

Parametr „*hv_loss*“ detekuje změnu uhlů hran jasové složky. Používá výstup SI13 filtru ze kterého vytvoří obraz vertikálních a horizontálních hran a obraz šikmých hran a spočítá se jejich podíl. Originální a zarušený snímek jsou pak porovnány pomocí vzorce (9). Výstup je poté upraven nelineárním filtrem a prahován pomocí předpisu (11).

Parametr „*hv_gain*“ sleduje změny hran z šikmých na vertikální nebo horizontální (např. při výskytu blokových artefaktů). K porovnání snímků používá vzorec (10).

Parametr „*chroma_spread*“ popisuje změny v rozložení obou chrominančních signálů – tedy změnu barvy výstupního obrazu. Počítá střední hodnotu složek v každém sub-bloku, přičemž červená složka je násobena hodnotou 1,5, aby došlo k jejímu zvýraznění. K porovnání originální a zarušené videosekvence je použit vzorec (11).

Parametr „*si_gain*“ měří zlepšení kvality způsobené systémem, který umožňuje zvýšení kontrastu a případně mírné zostření hran. Porovnání vstupního a výstupního obrazu probíhá pomocí vzorce (10). Je jediným parametrem, který umožňuje vylepšení celkové hodnoty VQM.

Parametr „*ct_ati_gain*“ hodnotí množství prostorové a časové aktivity v sub-blocích. Při velké aktivitě nebudou chybné bloky a šum tolik patrné. Výpočet se provádí pouze s jasovou složkou a bere v úvahu vliv kontrastu. Porovnání vstupního a výstupního obrazu je provedeno pomocí vzorce (9).

Parametr „*chroma_extreme*“ je počítán obdobně jako *chroma_spread* s tím rozdílem, že je optimalizován pro detekci velkých chyb v barevnosti v určité oblasti. Tyto chyby vznikají převážně při přenosu dat a jsou pro pozorovatele velmi rušivé.

Výsledný model je pak sestaven dle následujícího vzorce:[8]

$$\begin{aligned}
 VQM = & -0,2097 * si_loss \\
 & + 0,5969 * hv_loss \\
 & + 0,2483 * hv_gain \\
 & + 0,0192 * chroma_spread \\
 & - 2,3416 * si_gain \\
 & + 0,0431 * ct_ati_gain \\
 & + 0,0076 * chroma_extreme.
 \end{aligned}
 \tag{12}$$

Výstupní hodnoty jsou v rozsahu od nuly, kdy je výstupní obraz totožný s originálním, do jedné, kdy je zkruslení velmi markantní. Model byl navržen tak aby výsledky po vynásobení 100 přibližně odpovídaly hodnotám DMOS (Difference Mean Opinion Score) pro pozorovací vzdálenost od 4 do 6-ti násobku výšky obrazu. Pro případ, že by parametr *si_gain* přesunul VQM do záporných hodnot (což by naznačovalo, že výstupní obraz dosahuje vyšší kvality než výstupní), jsou hodnoty menší než nula zaokrouhleny na nulu.

Aby u silně zarušených sekvencí nedocházelo k přílišnému nárůstu hodnoty VQM na hodnotu 1 (a tím i špatné korelaci se subjektivními testy), je zaveden 50-ti procentní práh. Dodatečný výpočet je pak prováděn následovně:

$$\text{pro } VQM > 1,0 \quad VQM = \frac{(1+c) \cdot VQM}{c + VQM}, \quad (13)$$

kde $c = 0,5$. Tím je docíleno že hodnoty se budou pohybovat v intervalu $<0, 1>$, přičemž se při extrémní chybovosti tento interval může posunout mírně nad hodnotu jedna.[8]

Fast Low Bandwidth Model

Jedná se o metodu s částečnou referencí, která ve velké míře vychází z *NTIA General Model*. Byl standardizován ANSI i ITU. K monitorování kvality používá přídavnou referenční informaci o bitové rychlosti menší než 10 kb/s. Tato přídavná informace je přesto dostačující a metrika poskytuje výsledky s dobrou korelací se subjektivními metodami. K výpočtu je využito 8 parametrů, přičemž každý z nich zahrnuje jiný aspekt poškození obrazu. Výstupní hodnotou je pak jejich lineární kombinace. Stejně jako *NTIA General Model* umožňuje kalibraci zarušeného obrazu. Tento model byl uveden jen ve zkratce pro ilustraci a více o něm lze nalézt v literatuře, např. [11].

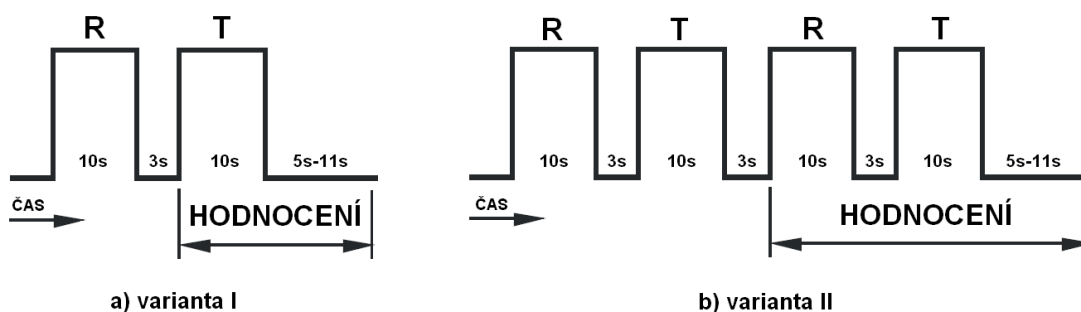
1.2 Subjektivní metriky

Základem subjektivních metrik je hodnocení poskytované respondenty, tedy skutečnými lidskými pozorovateli. Ti jsou vybráni podle předem definovaných požadavků, ve většině případů tak aby tvořili reprezentativní vzorek populace. Dle doporučení ITU-T P.910 je absolutní minimum pro vyhodnocení jsou 4 hodnotitelé. Ze statistického hlediska není příliš významné navyšovat počet pozorovatelů nad číslo 40. Obecně by se testu mělo zúčastnit alespoň 15 osob, které by neměly mít detailní znalosti o testovaném systému. Nicméně je doporučeno během prvotních testování systémů použít malou skupinu odborníků (4-8) k získání počátečních výsledků. Před vlastním testováním by pozorovatelé měli být seznámeni s průběhem experimentu a měly by jim být pokud možno nezaujatě sděleny pokyny k hodnocení. Dále je vhodné na začátku experimentu použít několik tréninkových videí, které následně nebudou zahrnuty do finálního výsledku testu. Pořadí přehrávání jednotlivých sekvencí by mělo být náhodné, aby se zabránilo vzájemnému ovlivňování sekvencí a vlivu únavy. Podmínky pro pozorování jsou přesně definovány v doporučení ITU-R BT.500 a ITU-T P.910. Při vyhodnocení je pro každý zkoumaný parametr provedeno statistické vyhodnocení, které bere v úvahu 95% interval spolehlivosti. Výsledek testu musí obsahovat detaily i použitých sekvencí, podmínky pozorování a typ použité metody.[12],[13],[14]

Subjektivní hodnocení se obecně dělí na metody, které využívají dvě sekvence, kdy je pozorovateli promítnut referenční i testovaný snímek a je známo, který je který. Obě sekvence mohou být promítány na zobrazovači zároveň nebo po sobě. Příkladem jsou metody DSIS, DSCQS a SDSCE popsané dále. Druhou skupinou jsou metody používající jednu sekvenci. Není tedy zapotřebí mít k dispozici referenční snímek, ten však může být zahrnut mezi testované sekvence. Jako příklad jsou uvedeny metody SSCQE, ACR a ACR-HR. Jednotlivé metody jsou popsány v následujícím textu.

1.2.1 Metody využívající dvě sekvence

Metoda double-stimulus impairment scale (DSIS) je vhodná především k popisu vlastností nového systému nebo ke zhodnocení vlivu přenosu na signál. Pozorovateli je prvně promítnut referenční snímek, po kterém následuje testovaný snímek. Časový průběh je znázorněn na obr. 1-1. Pozorovatel pak vyhodnotí míru rušení či chybovosti na pětibodové stupnici. Její podoba ne naznačena na obr. 1-2.[12]



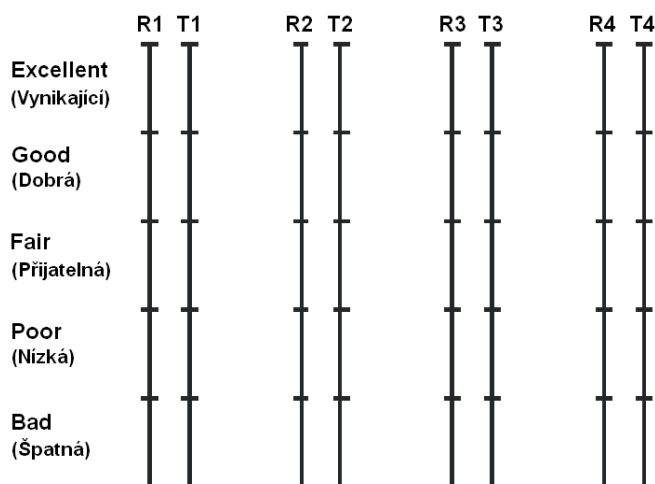
Obrázek 1-1 Časová struktura DSIS testu [12].

- 5 **imperceptible**
(nepozorovatelné)
- 4 **perceptible, but not annoying**
(pozorovatelné, ale neobtěžující)
- 3 **slightly annoying**
(mírně obtěžující)
- 2 **annoying**
(obtěžující)
- 1 **very annoying**
(velmi obtěžující)

Obrázek 1-2 Hodnotící stupnice pro DSIS metodu [12].

Double-stimulus Continuous Quality-scale (DSCQS) je metoda vycházející z DSIS. Hlavní rozdíl spočívá v tom, že pozorovatel hodnotí kvalitu obou sekvencí na kontinuální stupnici (případně na stupnici od 0 do 100), jak je naznačeno na obr. 1-3. Testování probíhá podle obr. 1-1 b), nebo je ponecháno na hodnotiteli, kolikrát bude dvojice sekvencí přehrána. Nemusí být explicitně sděleno, který snímek je referenční. Jako výstup k dalšímu zpracování je pak brán rozdíl hodnocení referenční a testované sekvence, který je normalizován na hodnoty od 0 do 100. [12]

Simultaneous Double Stimulus for Continuous Evaluation (SDSCE) je metoda, při které jsou zároveň promítány jak referenční, tak i testovaný snímek. Je vhodná pro měření delších videosekvencí, kde dochází ke změnám časové a prostorové aktivity a jejichž kvalita se mění s časem. Pozorovatel během celého testu kontinuálně hodnotí rozdíly mezi oběma sekvencemi na stupnici od 0 (velký rozdíl) do 100 (žádný rozdíl). [12]

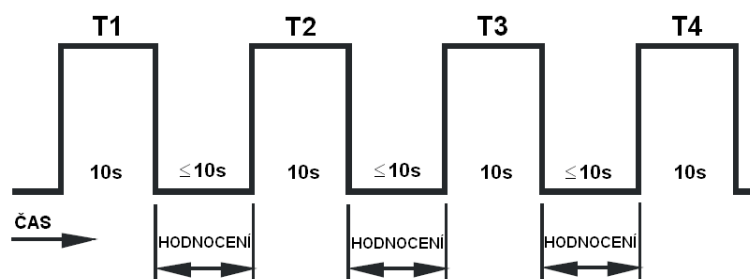


Obrázek 1-3 Hodnotící stupnice pro DSCQS metodu [12].

1.2.2 Metody využívající jednu sekvenci

Stimulus Continuous Quality Evaluation (SSCQE) analogická s metodou SDSCE s tím rozdílem, že pozorovatel nemá k dispozici referenční sekvenci. Měření tedy simuluje podmínky reálných systémů s přenosem (digitální televize). [12]

Absolute Category Rating (ACR) je metoda, při níž jsou testované sekvence prezentovány jedna za druhou, jak je znázorněno na obr. 1-4. Po přehrání každé sekvence následuje časový interval o délce přibližně 10 vteřin, kdy pozorovatel ohodnotí kvalitu. Doporučená škála hodnocení je vidět na obr. 1-5. V případě potřeby jemnějšího rozlišení je možno použít stupnici s devíti úrovněmi. Dále je možné použít kontinuální stupnice k hodnocení jiných aspektů než je kvalita (např. barva, kontrast, plynulost pohybu, atd.). Vzhledem k tomu, že není potřeba promítání referenčních sekvencí, je ACR relativně rychlá a snadná na implementaci. Tím, že pozorovatel nehodnotí rozdíly, ale celkovou kvalitu obrazu, je tato metoda vhodná například k hodnocení degradace obrazu při kompresi. Metoda ACR není vhodná při analýze chyb vyskytujících se na začátku a konci videosekvence, jelikož hodnotitel nezná referenční videosekvenci a nemusí tyto chyby zaznamenat. [14]



Obrázek 1-4 Časová struktura ACR testu [14].

- 5 **Excellent**
(Vynikající)
- 4 **Good**
(Dobrá)
- 3 **Fair**
(Přijatelná)
- 2 **Poor**
(Nízká)
- 1 **Bad**
(Špatná)

Obrázek 1-5 Hodnotící stupnice pro ACR metodu [14].

Absolute Category Rating with Hidden Reference (ACR-HR), jak název napovídá, jedná se o metodu vycházející z metody ACR. Hlavní rozdíl spočívá v tom, že mezi prezentovanými sekvencemi se nachází sekvence referenční. Pozorovatel však o její přítomnosti neví. Kvalita referenční sekvence by měla na takové úrovni, aby jí odborník mohl označit jako vynikající nebo alespoň dobrou (na pětibodové stupnici). Průběh testu a stupnice hodnocení je stejná jako u ACR (viz obr. 1-5 a obr. 1-6) [14],

$$DV(PVS) = V(PVS) - V(REF) + 5, (14)$$

kde $V(REF)$ je hodnocení referenční sekvence, $V(PVS)$ je hodnocení testované sekvence a $DV(PVS)$ je výsledná diferenční hodnota testované sekvence. Po aplikaci tohoto vzorce $DV=5$ značí vynikající kvalitu a $DV=1$ špatnou kvalitu. Aby se zabránilo případu, kdy by výsledek DV vycházel větší než pět (testovaná sekvence by se jevila kvalitnější než reference) je využit následující vzorec: [14]

$$crushed_DV = (7 * DV) / (2 + DV) \quad \text{pokud } DV > 5, (15)$$

kde DV je diferenční skóre testované sekvence a $crushed_DV$ je snížená hodnota.

Metoda ARC-HR díky skryté referenci snižuje vliv kvality referenčního videa (rozdíl v použité kameře u různých zdrojových sekvencí), vliv vlastností použitého zobrazovače a alespoň částečně eliminuje zaujetí pozorovatele vůči scéně. Díky menší časové náročnosti je vhodná k použití při testování velké sady sekvencí. [14]

2 H.265/HEVC

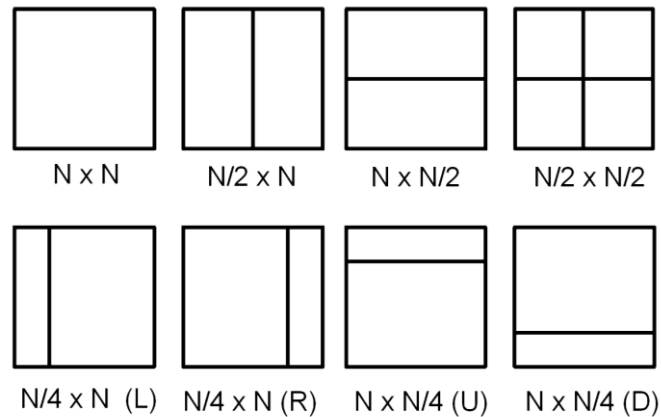
H.265/HEVC (High Efficiency Video Coding) je nový standard pro kompresi videa. Byl standardizován v dubnu 2013 v doporučení ITU T-REC-H.265 a v ISO/IEC 23008-2. Byl vyvinut skupinou JCT-VC (Joint Collaborative Team on Video Coding) a je považován za nástupce kodeku H.264/MPEG-4 AVC. Jeho cílem je vylepšení účinnosti kódování alespoň o 50 % oproti H.264, ovšem na úkor výpočetní náročnosti. Má široké zaměření – od mobilních aplikací, přes internetové služby v reálném čase, medicínské obrazy, televizní vysílání až po domácí a digitální kina ve vysokém a ultra vysokém rozlišení. Zaměřuje se převážně na neprokládané videosekvence ve vysokém rozlišení, o velké snímkové rychlosti a ve formátu 8-bit 420. Celý standard se zatím neustále rozvíjí a je průběžně doplňován o nové nástroje. [15]

2.1 Stromová struktura

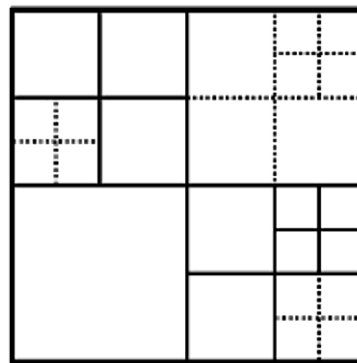
HEVC navazuje na postupy předchozích MPEG standardů. Výrazná změna však spočívá v základní stromové struktuře bloků, která je u HEVC daleko více adaptivní. Navíc není založena na makroblocích, ale na kódované jednotce CTU. Tato jednotka zahrnuje jeden jasový blok CTB a korespondující chrominanční bloky CTB. Tyto bloky mohou mít různou velikost, která je volena enkodérem. Jasový blok CTB má velikost $N \times N$ vzorků, jeden chrominanční CTB blok má velikost $(N/2 \times N/2)$ vzorků, kde N může být 16, 32 nebo 64. CTB bloky se dále pomocí hierarchické struktury dělí na kódované bloky CB. Ty jsou opět čtvercové s minimální velikostí 8×8 pro jasové a 4×4 pro chrominanční složky. Jejich maximální velikost je stejná jako velikost CTB. Jeden jasový a dva korespondující chrominanční CB bloky společně s přidruženou syntaxí vytvářejí kódovanou jednotku CU. Každé jednotce CU je přiřazen buď inter nebo intra predikční mód. CU je poté pomocí stromové struktury rozdělena na predikční jednotky PU a jednotky pro transformaci TU a jim odpovídající bloky PB a TB.

Predikční bloky PB mohou být obdélníkové o minimálních rozměrech 4×8 nebo 8×4 vzorků. V nadřazeném CB bloku mohou být 1, 2 nebo 4 PB bloky. Obdobně jako u kódovaných bloků CB, PB bloky vytvářejí predikční jednotky PU. Pokud je použito asymetrické dělení (AMP) může být jeden jasový CB blok o velikosti $N \times N$ rozdělen na dva jasové bloky PB. Jeden o velikosti $N \times (N/4)$ nebo $(N/4) \times N$ a druhý o velikosti $N \times (3N/4)$ nebo $(3N/4) \times N$. Možnosti mapování PB do CB jsou zobrazeny na obr. 2-1.

Dalším používaným typem jsou jednotky pro transformaci TU. Ty se obdobně jako předchozí jednotky skládají z bloků TB. Jsou tvořeny residui po predikci. Vycházejí z velikosti CU a mohou se dále stromově větvit. Bloky jsou vždy čtvercové a mohou mít velikosti 4×4 , 8×8 , 16×16 a 32×32 vzorků. Příklad větvení je na obr. 2-2. [16], [17]



Obrázek 2-1 Možné velikosti PB [17].



Obrázek 2-2 Rozdělení CTB do CB a TB (plnou čarou hranice CB, tečkované hranice TB), převzato z [17].

2.2 Transformační kódování

Transformační kódování se provádí na TB blocích. Ty jsou z toho důvodu vždy čtvercové. Jejich maximální a minimální velikost volí enkodér. Základ metody tvoří 2-D celočíselná transformace vycházející z diskretní kosinové transformace (DCT). Ta se provádí zvlášť na jasových a zvlášť na každém chrominančním bloku. Pro jasové bloky 4 x 4 vzniklé intrapredikcí je použita celočíselná transformace odvozená z diskretní sinové transformace DST. [16]

2.3 Entropické kódování

Pro HEVC je specifikována pouze jedna metoda entropického kódování a tou je CABAC (Context-adaptive binary arithmetic coding). Její algoritmus je téměř totožný s algoritmem, který využívá kodér H.264/AVC. Dochází jen k několika změnám, které umožňují dosažení vyšší přenosové rychlosti. Dále je v rozšířené míře umožněno pracovat a přistupovat k jednotlivým blokům stromové struktury tak, aby entropické kódování bylo co nejefektivnější. [17]

2.4 Intrapredikce

Intrapredikce u HEVC je podobná intrapredikci u AVC. Vzorky jsou predikovány pomocí sousedních bloků. Základní módy pro predikce jsou horizontální, vertikální, planární, DC a úhlový. Při dominantní horizontální predikci se využití body nad kódovaným blokem nazývají hlavní pole a body nalevo od kódovaného bloku se nazývají boční pole. Při dominantní vertikální predikci je tomu naopak. Pro výpočet se pak používá lineární interpolace všech dostupných bodů. Pro bloky větších velikostí je pak predikce rozšířena na celkem 35 módů (ty jsou definovány úhlem). V DC módu je k predikce stejnosměrné složky použita střední hodnota vzorků z bloku nad a vedle právě předikovaného vzorku. [3], [18].

2.5 Interpredikce

Při interpredikci je každé PU přiřazen jeden, respektive dva vektory pohybu pro jednosměrnou nebo obousměrnou predikci. Samotný vektor se pak nemusí přenášet a přenáší se pouze jeho rozdíl oproti předpovězené (predikované) hodnotě. Koncept snímků I, P, B zůstává zachován a je rozšířen o možnosti používání více referenčních snímků. Celkově je možno nastavení vzájemných vztahů snímků v rámci GOP struktury. Predikce je $\frac{1}{4}$ pixleová a je různá pro jasové a pro chrominanční bloky. Pro jasové bloky je použit filtr osmého řádu a pro chrominanční filtr čtvrtého řádu. Minimální velikost bloku pro interpredikci je 4×8 nebo 8×4 vzorků. U těchto bloků je navíc povolena jen jednosměrná predikce. Tím dojde ke snížení velikosti potřebné paměti. [3], [17].

2.6 Rekonstrukční filtr

Rekonstrukční filtr u kodeku HEVC je proveden ve dvou krocích. V prvním kroku je aplikován filtr pro potlačení blokové struktury, ve druhém pak filtr SAO (Sample adaptive offset).

Filtr pro potlačení blokové struktury má za úkol redukovat blokové artefakty vzniklé při kódování. Pokud je povolen, je aplikován na všech rozhraních mezi bloky na úrovni PU nebo TU. Narozdíl od kodeku H.264/AVC, kde je tento filtr aplikován nad mřížkou 4×4 vzorků, je u kodeku H.265/HEVC nad mřížkou 8×8 vzorků jak pro chrominanční, tak pro jasový signál. Tímto přístupem dochází ke snížení náročnosti výpočtu při zachování dostatečné obrazové kvality. Hrany jasových bloků se filtrují v následujících případech: alespoň jeden z bloků je získán pomocí intrapredikce, mají nenulové koeficienty nebo rozdíly mezi vektory pohybu bloků jsou větší než jeden pixel. Pro jasové vzorky jsou definovány tři stupně filtrace – žádná, slabá a silná, přičemž výběr závisí na vlastnostech sousedních bloků. Pro chrominanční vzorky jsou definovány dva stupně filtrace – žádná a normální. [17]

Filtr SAO je aplikován pouze na vzorky, které splňují určité požadavky. Algoritmus H.264 obdobný filtr nepoužívá. Používá se na vzorky, které již prošly prvním krokem filtrace. Provádí se nad oblastí, která je specifikována typem CTB. Samotný filtr pak přidává ke vzorkům offset, který je definován enkodérem. [17]

3 DATABÁZE VIDEOSEKVENČÍ

3.1 Zdrojové sekvence

Pro potřeby testování byla vytvořena databáze videosekvencí komprimovaných pomocí H.265/HEVC. Jako kodér byla použita referenční implementace HM-12.0+Rext-4.1 dostupná z [19]. Bylo použito deset zdrojových sekvencí, všechny o délce 250 snímků. Při rychlosti 25 snímků/s tedy měly sekvence dobu trvání 10 vteřin. Tato doba byla zvolena tak, aby odpovídala hodnotě běžně užívané při subjektivních testech. Formát byl zvolen YUV 4:2:0, progresivní. Jako prvních pět zdrojových sekvencí byly vybrány základní testovací sekvence dostupné z [20]. Ty již byly v nekomprimované podobě v potřebném formátu. Jediná úprava spočívala ve vystřížení potřebné délky (10s). K tomu jsme použili program `ffmpeg.exe`, který je součástí balíku programů FFmpeg, [21] s parametrem `-vframes 250`. Všechny tyto sekvence měly stejné rozlišení, a to 1980x1080 pixelů, tedy poměr stran 16:9. Scény v nich byly souvislé a neobsahovaly střih. Vzhledem k tomu, že snímky používané k reálné kompresi a přenosu obsahují střih a dochází u nich ke změně scény, není použití pouze těchto testových sekvencí vhodné. Proto byla vytvořena druhá skupina pěti sekvencí z běžné filmové produkce. Použití takto krátkých ukázek pro testovací účely je i u komerčních filmů povoleno, a proto nám nic nebránilo v jejich využití. Byl vybrán snímek *Skyfall* (z roku 2013), snímek *Knight and Day* (z roku 2010), snímek *2001: A Space Odyssey* (rok 1968) a záznam finále běhu mužů na 200 m z Olympijských her 2012 v Londýně. Tyto videosekvence byly již kódovány kodekem H.264, ovšem s poměrně vysokou bitovou rychlostí a dostačující kvalitou (nebyly patrné žádné artefakty či chyby). Navíc slouží jako referenční u subjektivního i objektivního hodnocení a případné nedostatky jsou tedy odečteny. Vlastnosti původních sekvencí jsou uvedeny v tabulce 3-1.

tabulka 3-1 Parametry původních sekvencí.

Název	Formát	Bitová rychlost	Rozlišení
Skyfall	MPEG-4 AVC (part 10)	33369 kb/s	1920x800
Knight and day	MPEG-4 AVC (part 10)	22292 kb/s	1920x800
Bolt	MPEG-4 AVC (part 10)	4160 kb/s	1920x1080
2001	MPEG-4 AVC (part 10)	8037 kb/s	1920x864

Pomocí programu `ffmpeg` bylo z těchto původních sekvencí vybráno několik desetivteřinových úseků. Ty byly převedeny do nekomprimovaného formátu YUV 4:2:0 progressive s rychlostí 25 snímků/s. Rozlišení zůstalo zachováno, pouze byly odstraněny černé pruhy v dolní a horní části obrazu. Syntaxe použitá v programu `ffmpeg` byla následující:

```
ffmpeg.exe -i skyfall.m2ts -ss 00:05:26 -vf crop=1920:800 -f rawvideo -vcodec rawvideo -pix_fmt yuv420p -s 1920x800 -r 25 -vframes 250 SKYFALL_2.yuv
```

Parametr `-i` specifikuje vstupní sekvenci, `-ss` udává čas od začátku sekvence, ve kterém se spustí kódování, `-vf crop` specifikuje vybranou oblast, která bude vyříznuta, `-f`, `-vcodec` a `-pix_fmt` nastavují formát výstupní sekvence, `-s` nastavuje výstupní rozlišení, `-r` specifikuje snímkovou rychlost a `-vframes` nastavuje počet kódovaných snímků

Tím bylo vytvořeno deset zdrojových sekvencí, jejich vlastnosti a popis jsou uvedeny v tabulce 3-2. Pro potřeby subjektivních testů (více informací v kapitole 5) bylo nutno tato zdrojová videa převést na nižší rozlišení. K tomu byl opět použit program FFmpeg. Zvolené rozlišení bylo SD (480p) a HD (720p), přičemž konkrétní rozlišení jednotlivých sekvencí je uvedeno v tabulce 3-3. Tato rozlišení byla volena tak, aby odpovídala rozlišení displeje zařízení, na nichž bude test prováděn. Všechny ostatní parametry zůstaly i po překódování zachovány.

Zdrojové sekvence byly zvoleny tak, aby obsahovaly jak téměř statické snímky, tak i snímky s rychlým pohybem a stříhem. Dále byly brány v potaz vlastnosti převládajících textur v obraze (např. jejich členitost, jejich změna v čase a náročnost na kódování) a celkové zastoupení barev ve scénách. Základní popis vybraných sekvencí je uveden v tabulce 3-2 a náhledy jsou k dispozici v příloze A.

tabulka 3-2 Parametry a popis full HD zdrojových sekvencí.

Název	Rozlišení	Popis
Sunflower_1080p	1920x1080	včela na slunečnici, téměř konstantní scéna, pouze pohyb hmyzu
Tractor_1080p	1920x1080	traktor na poli, pomalá změna scény
Park_Joy_1080p	1920x1080	lidé v parku, kontinuální pohyb celé scény
Ducks_1080p	1920x1080	vzlétající kachny, pohyb vodní hladiny
Crowd_Run_1080p	1920x1080	běžecké závody v parku, velká spektrální aktivita, bez stříhu
Skyfall_1_1080p	1920x 800	motorky jedoucí po střechách, rychlý pohyb v částech scény
Skyfall_2_1080p	1920x800	přestřelka na tržnici, rychlé stříhy, velké detaily scény, blízké záběry na postavy
Knight_1080p	1920x800	automobilová honička uprostřed města, dynamické změny scén
2001_1080p	1920x864	opice mlátící kostí do kosterních pozůstatků, statické pozadí, pomalý pohyb v popředí
Bolt_1080p	1920x1080	záznam sprintu, velká barevnost, rychlý pohyb

tabulka 3-3 Rozlišení SD a HD zdrojových sekvencí.

Název	Rozlišení	
	SD	HD
Sunflower	848x480	1280x720
Tractor	848x480	1280x720
Park_Joy	848x480	1280x720
Ducks_take_off	848x480	1280x720
Crowd_Run	848x480	1280x720
Skyfall_1	1152x480	1728x720
Skyfall_2	1152x480	1728x720
Knight_and_day	1152x480	1728x720
2001	1056x480	1600x720
Bolt	848x480	1280x720

3.2 Časová a prostorová aktivita

Pro další popis a zkoumání jednotlivých sekvencí je výhodné mít v číselné podobě informace o časové a prostorové aktivitě jednotlivých snímků. Pro samotné testování je pak vhodné, pokud zvolená databáze obsahuje videa s různorodou velikostí časové i prostorové aktivity.

Hodnoty pro zdrojové sekvence 1080p použité v této práci jsou vyneseny do grafu na obr. 3-1. Je vidět, že parametry zvolených videosekvencí jsou rozprostřeny napříč celou plochou grafu. Tím jsou zajištěny různé požadavky přímo určující dosažitelný kompresní poměr, a které jsou rozhodující zejména při požadavku na nízkou bitovou rychlost. Samotný výpočet byl proveden v programu Matlab podle předpisu z doporučení ITU-T P.910 [14]. Informace o časové a spektrální aktivitě jsou počítány zvlášť pro každý snímek sekvence. Výsledná hodnota vybraného parametru je potom maximum z hodnot všech snímků.

Výpočet prostorové aktivity SI je založen na filtraci jasové složky snímků pomocí Sobelova filtru. Z takto upraveného obrazu je pak počítána směrodatná odchylka hodnot pixelů. Z hodnot pro všechny snímky se pak vybere maximum. Tento postup lze popsat následujícím vzorcem:

$$SI = \max_{time} \left\{ std_{space} [Sobel(F_n)] \right\}, \quad (16)$$

kde $Sobel(F_n)$ je snímek v čase n po aplikaci filtru, std_{space} je směrodatná odchylka tohoto snímku a \max_{time} je maximální hodnota za uvažovaný čas.

Informace o časové aktivitě TI je založena na výpočtu rozdílů mezi jasovými pixely na stejném místě v prostoru v po sobě jdoucích snímcích. Toto lze popsat pomocí rovnice (17).

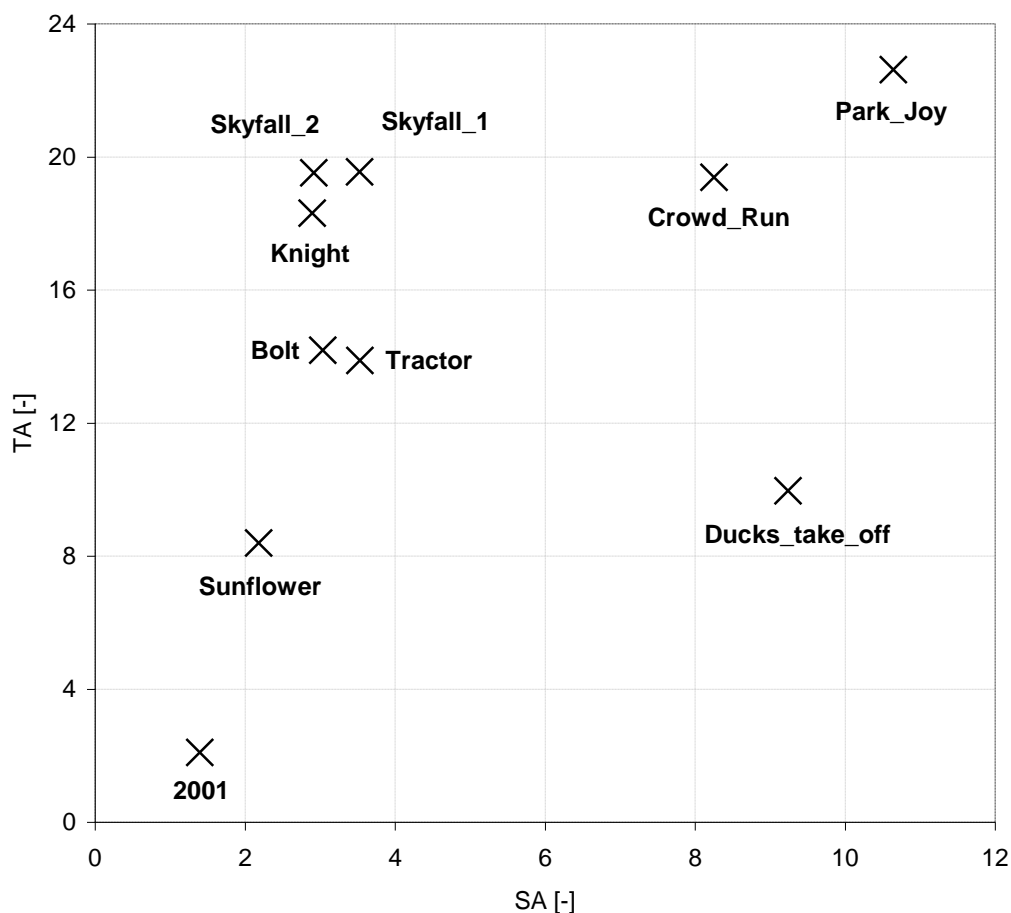
$$M_n(i, j) = F_n(i, j) - F_{n-1}(i, j), \quad (17)$$

kde $F_n(i, j)$ je hodnota pixelu v i -tém řádku a j -tém sloupci ve snímku n .

Výsledná hodnota se poté počítá obdobně jako spektrální aktivita, a to následujícím vzorcem:

$$TI = \max_{time} \left\{ std_{space} [M_n(i, j)] \right\}, \quad (18)$$

std_{space} je směrodatná odchylka rozdílu pixelů mezi snímky a \max_{time} je maximální hodnota za uvažovaný čas.



Obrázek 3-1 Spektrální a temporální aktivita použitých zdrojových sekvencí.

3.3 Kódování sekvencí

Kódování full HD, HD i SD sekvencí probíhalo na serverech s operačním systémem CentOS linux s následujícími parametry:

BL465c G5 2x Opteron 2384 4x2.7GHz, 32GB, CentOS release 6.4

BL465c G6 2x Opteron 2435 6x2.6GHz, 32GB, CentOS release 6.4

BL465c G6 2x Opteron 2435 6x2.6GHz, 32GB, CentOS release 6.4

Pro výpočet byl po zkompilování HM-12.0 použit enkodér `TappEncoderStatic`, ovládaný z příkazové řádky. Vzhledem k velkému množství parametrů byly tyto zadávány v textovém konfiguračním souboru. Příklad celého konfiguračního souboru je uveden v příloze B. Syntaxe pro enkodér byla následující:

```
TappEncoderStatic -c konfiguračni_soubor.txt
```

K vytvoření požadovaných výstupních sekvencí bylo využito zadání cílové bitové rychlosti. Tyto rychlosti byly voleny tak, aby v obraze byly patrné různé stupně zkreslení. Hodnoty PSNR získaných sekvencí se pohybovaly od 20 dB do 45 dB, čímž byla zajištěna vhodnost k použití pro subjektivní testy. Kritičtěji byla volena bitová rychlost pro HD (720p) a SD (480p) sekvence z důvodu dalšího využití k již zmíněným subjektivním testům

U full HD (1080p) sekvencí byl na základě testování zvolen rozsah rychlostí od 250 kb/s do 2,5 Mb/s pro sekvence s vyšší obtížností kódování a rozsah od 50 kb/s do 750 kb/s pro sekvence s nižší obtížností kódování. U HD (720p) sekvencí se bitová rychlost pohybovala od 900 kb/s do 3,1 Mb/s pro sekvence s vyšší obtížností kódování, od 300 kb/s do 1,7 Mb/s pro sekvence se střední obtížností kódování a od 100 kb/s do 700 kb/s pro sekvence s nízkou obtížností kódování. U SD (480p) sekvencí to pak byla rychlost v rozsahu 50 kb/s až 1,3 Mb/s respektive 25 kb/s až 250 kb/s.

Řízení bitové rychlosti se volí pomocí položky RateControl v konfiguračním souboru nastavením její hodnoty na hodnotu 1. Požadovaná bitová rychlost, udaná v bitech za sekundu, byla zapsána do položky TargetBitrate. Profil byl nastaven „Main“ u všech třech typů sekvencí. Level byl nastaven 4.1 pro full HD a HD sekvence. Pro sekvence SD byl zvolen Level 3. Velikost maximální jednotky pro kódování (CU) byla nastavena 64x64 pixelů. Deblock filtr byl povolen a vzdálenost pro hledání vektorů pohybu byla nastavena na 64 pixelů u full HD respektive 32 pixelů u HD a SD.

Enkodér dovoluje rozsáhlé možnosti nastavení skupiny snímků. Mimo jiné umožňuje nastavení konkrétních referencí pro jednotlivé snímky, velikost posuvu kvantizačního parametru a vlastnosti vyhlazovacího filtru pro každý snímek zvlášť. Experimentálně se nepodařilo vytvořit strukturu, která by byla stabilní a dosahovala dobrých výsledků. Byly tedy zvoleny parametry přiložené k samotnému balíku programů o celkové délce 16 snímků mezi po sobě následujícími I snímky. Mezi jednotlivými I snímky byla dvakrát po sobě použita následující struktura:

```

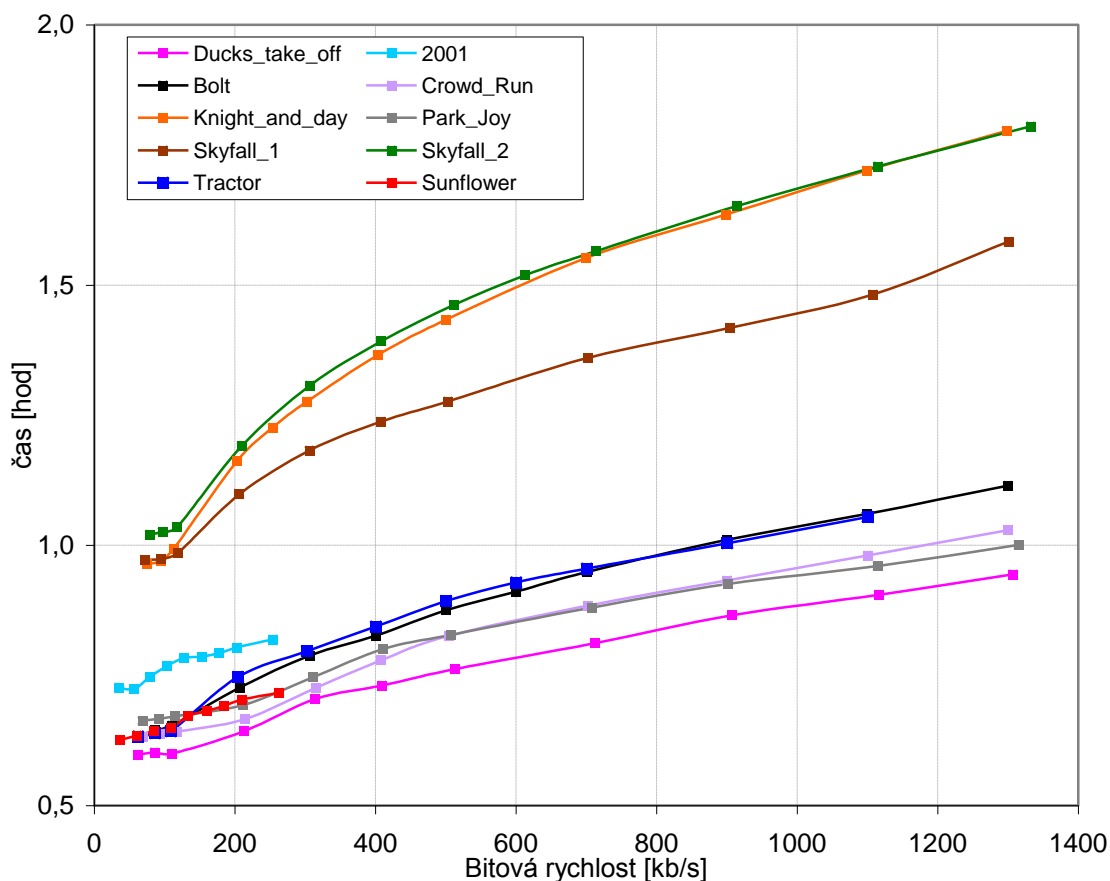
Frame1: B 8 1 0.442 0 0 0 4 4 -8 -10 -12 -16 0
Frame2: B 4 2 0.3536 0 0 0 2 3 -4 -6 4 1 4 5 11001
Frame3: B 2 3 0.3536 0 0 0 2 4 -2 -4 2 6 1 2 4 1111
Frame4: B 1 4 0.68 0 0 0 2 4 -1 1 3 7 1 1 5 10111
Frame5: B 3 4 0.68 0 0 0 2 4 -1 -3 1 5 1 -2 5 11110
Frame6: B 6 3 0.3536 0 0 0 2 4 -2 -4 -6 2 1 -3 5 11110
Frame7: B 5 4 0.68 0 0 0 2 4 -1 -5 1 3 1 1 5 10111
Frame8: B 7 4 0.68 0 0 0 2 4 -1 -3 -7 1 1 -2 5 11110

```

Výstupem enkodéru bylo vždy požadované zakódované video v binární podobě a dále také rekonstruovaný YUV 4:2:0 snímek. Díky tomu nebylo zapotřebí výstupní videa samostatně dekódovat a pro další použití při výpočtu objektivních metrik mohla být použita zmíněná YUV sekvence. Terminálový výstup enkodéru byl pak ukládán do textového souboru.

3.4 Zjištěné vlastnosti

Čas potřebný pro kódování se pohyboval v rozsahu od 2,5 hodiny do 6 hodin u full HD sekvencí, v rozmezí od 1,5 hodiny do téměř 4 hodin u HD sekvencí a v rozmezí od 35 minut do 110 minut u SD sekvencí. S vyšší bitovou rychlostí potřebný čas roste, to stejné platilo i pro obtížnost kódování. Na obr. 3-2 je zobrazena časová náročnost kódování SD sekvencí. Z ní je patrné, že potřebná doba roste přibližně lineárně se vzrůstající bitovou rychlostí komprimované videosekvence a trend křivek není závislý na obsahu sekvence. Na časové náročnosti se tedy výrazně projevuje velikost (rozlišení) použitého videa (viz sekvence Skyfall, Knight and day a 2001). Křivky pro full HD a HD videa byly obdobné.



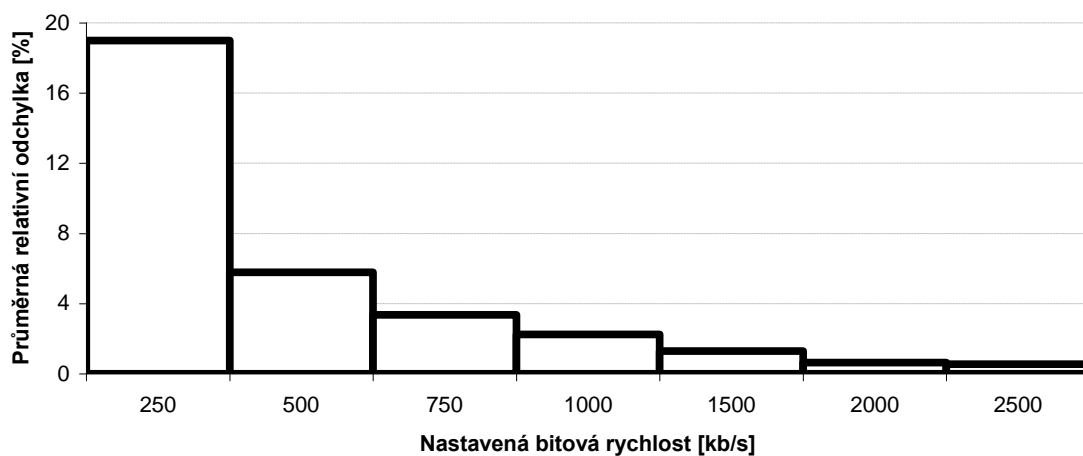
Obrázek 3-2 Časová náročnost kódování sekvencí v SD rozlišení.

Dosažená bitová rychlost se od nastavené výrazně lišila pro nižší bitové rychlosti (vyšší stupeň komprese). Na nich odchylka dosahovala průměrných hodnot kolem 30 % (SD), 12 % (HD) a 19 % (full HD). Tento rozdíl je způsoben tím, že při velmi nízkých bitových rychlostech byly na kódér kladeny nepřiměřené požadavky na kompresi. Neschopnost kódéru se s tímto vypořádat se projevila mimo jiné právě na tom, že nedokázal dodržet předepsanou rychlost. S rostoucí bitovou rychlostí však klesal stupeň komprese a kódér zvládal data přiměřeně zpracovat. Jak je vidět na obr 3-3 až 3-5, relativní odchylka nastavené a dosažené bitové rychlosti exponenciálně klesala. Při vyšších rychlostech byla její hodnota zanedbatelná, kódér neměl problém požadovanou rychlost dodržet.

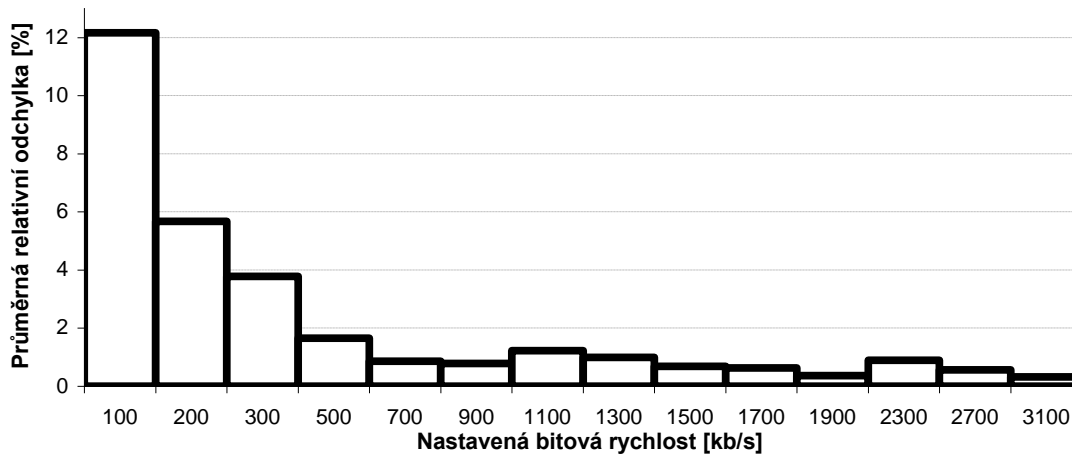
Výpočet průměrné relativní odchylky byl proveden podle následujícího vzorce:

$$\bar{\delta} = \sum_i \left(\frac{P - R_i}{P} \cdot 100 \right) / N \quad [\%], \quad (19)$$

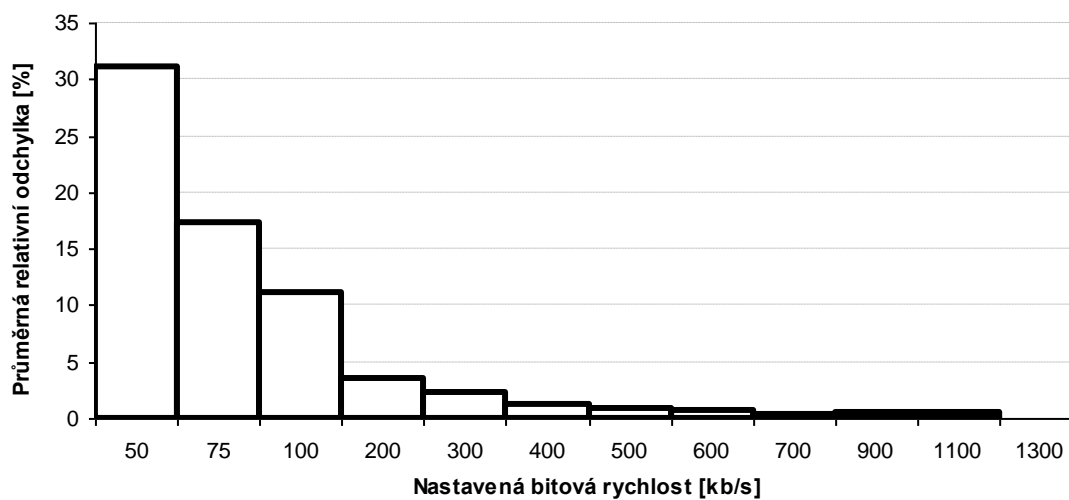
kde P je požadovaná rychlost, R_i je reálná rychlost i -tého videa a N je celkový počet videí s požadovanou rychlostí.



Obrázek 3-3 Průměrná relativní odchylka bitové rychlosti u full HD sekvencí.

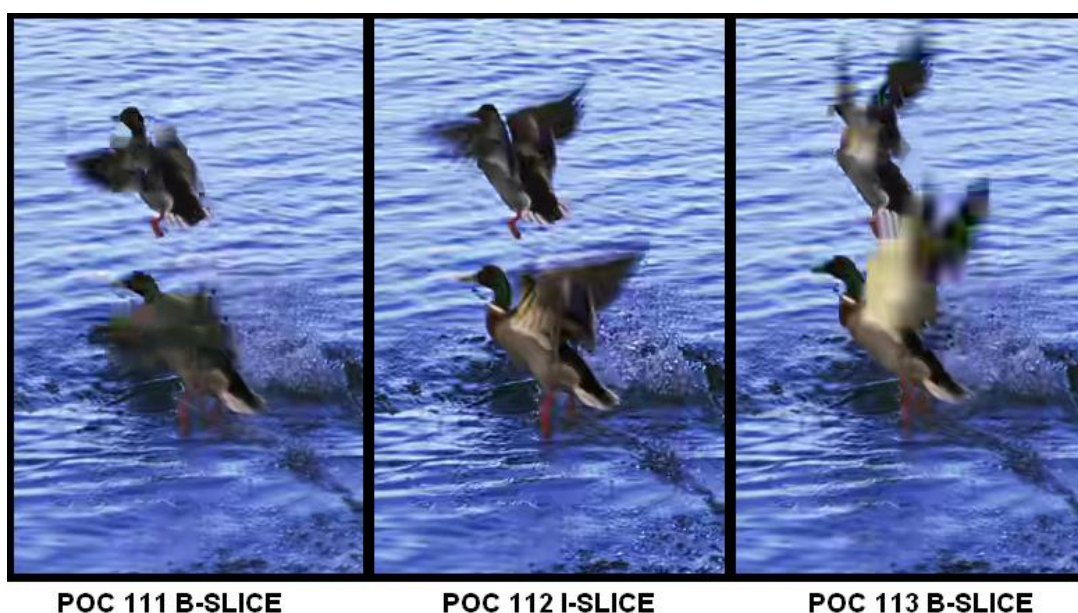


Obrázek 3-4 Průměrná relativní odchylka bitové rychlosti u HD sekvencí.



Obrázek 3-5 Průměrná relativní odchylka bitové rychlosti u SD sekvencí.

U kódovaných videosekvencí bylo při vyšším kompresním poměru patrné, že kodér protěžuje I snímky oproti snímkům P a B. Kvalita I snímků byla vyšší a jejich velikost (v bitech) oproti P a B byla o dva až tři řády vyšší. Toto zkreslení bylo nejvíce výrazné a rušivé u sekvence Ducks_take_off. Tato sekvence dosahuje vysoké prostorové aktivity (viz obr. 3-1) a převažuje v ní textura vodní hladiny s jemnými detaily, které jsou pro kodér obtížně zpracovatelné. Tři po sobě jsoucí snímky ze sekvence Ducks_take_off_480p_500kbps jsou znázorněny na obr. 3-6. Na něm je vidět, že pro I snímek je obraz velmi ostrý, kdežto v předcházejícím a následujícím snímku jsou velmi výrazné blokové artefakty. Dále je patrné, že u B snímků dochází ke značnému potlačení detailů. Tato skoková změna kvality se vyskytovala u všech uvažovaných rozlišení a působila na první pohled velmi rušivě.



Obrázek 3-6 Tři po sobě jdoucí snímky sekvence Ducks_take_off_480p_500kbps.

Při kódování sekvence Skyfall_1 docházelo k velmi výraznému poklesu obrazové kvality na konci sekvence. Jednalo se o zhruba 25 snímků, tedy poslední vteřinu videa. Ve zdrojové sekvenci dochází v tomto místě k rychlému pohybu, kdy část obrazu působí rozmazaně (viz. obr. 3-7). Kodér nezvládal tuto část zpracovat v odpovídající kvalitě a výsledek působil velmi rušivě. I snímek na pozici 224 měl zkreslení, odpovídající požadované bitové rychlosti. Následující I snímek na pozici 240 však již vykazoval mnohem větší zkreslení. Pro popis tohoto velkého skoku kvality bylo vypočteno PSNR sekvence obsahující snímky 1-224 a PSNR ze snímků 225-250. Hodnoty PSNR z první části sekvence byly výrazně vyšší než hodnoty druhé části (poslední 1 s videa). Rozdíl těchto dvou hodnot společně s hodnotou PSNR celé sekvence je uveden v tabulce 3-4. Z té je patrné, že při nižších bitových rychlostech, tedy při nižším kompresním poměru, není rozdíl tak výrazný. Při vyšších bitových rychlostech, kdy se celková hodnota PSNR pohybovala kolem 37 dB, byl rozdíl v kvalitě mezi první částí a druhou částí sekvence přibližně 8 dB. Jak tato hodnota napovídá, jednalo se o výrazný skok v obrazové kvalitě. Pro názornost je na obr. 3-7 zobrazen snímek na pozici 246 zdrojového videa v rozlišení 480p a jeho kódovaného protějšku s bitovou rychlostí 1,3 Mb/s a celkovou hodnotou PSNR 37,97 dB. Z něj je vidět, že i pro relativně velký bitový tok nebyl kodér schopen tuto část sekvence

zpracovat s odpovídající kvalitou.

tabulka 3-4 Rozdíly mezi PSNR mezi snímky 1-224 a snímky 225-250 sekvence Skyfall_1.

480p								
Bitová rychlost	[kb/s]	50	100	300	500	900	1100	1300
PSNR celé sekvence	[dB]	25,17	26,49	31,50	33,57	36,23	37,13	37,97
Rozdíl PSNR	[dB]	-0,53	0,85	4,04	5,79	8,26	8,74	8,84
720p								
Bitová rychlost	[kb/s]	300	500	700	900	1100	1300	1700
PSNR celé sekvence	[dB]	29,96	32,54	34,06	35,02	35,89	36,61	37,62
Rozdíl PSNR	[dB]	2,36	4,22	5,41	6,61	7,45	7,61	8,44
1080p								
Bitová rychlost	[kb/s]	250	500	750	1000	1500	2000	2500
PSNR celé sekvence	[dB]	29,33	32,49	34,31	35,44	36,85	37,84	38,55
Rozdíl PSNR	[dB]	1,23	4,05	5,22	6,17	7,14	8,14	8,95



Obrázek 3-7 Rozdíl zdrojové a kódované sekvence (1300kbps).

3.5 Vytvořená databáze

Výsledná vytvořená databáze obsahuje použitá zdrojová videa v rozlišení 480p, 720p a 1080p. Ty jsou k dispozici v nekomprimovaném formátu YUV 4:2:0. Dále databáze obsahuje náhledy těchto videí ve formátu H.264. Ty byly vytvořeny programem FFmpeg. Syntaxe byla následující:

```
ffmpeg -s 848x480 -i zdrojové_video.yuv -c:v libx264 -preset ultrafast -qp 1 h.264_zdrojové_video.mp4
```

Jak je vidět, kvantizační parametr byl nastaven na hodnotu 1, což znamená téměř bezztrátovou kompresi.

Toto nastavení bylo zvoleno tak, aby docházelo k co nejmenším degradacím způsobeným kóděm H.264. Dále byla zjištěna hodnota PSNR takto vytvořených sekvencí, která je rovněž v databázi k dispozici. Ta se pohybovala v rozmezí od 57 dB do 59 dB, což je hodnota dostatečná k tomu, aby tyto zdrojové sekvence mohly být použity jako referenční sekvence při subjektivních testech (více v kapitole 5). K dispozici je rovněž informace o hodnotě prostorové a časové aktivity všech zdrojových sekvencí v rozlišení 1080p. Pro rozlišení 480p a 720p se tyto hodnoty liší pouze nepatrně a proto není potřeba je udávat.

Hlavní část vytvořené databáze pak tvoří samotné komprimované soubory. Ty jsou rozděleny podle rozlišení do tří skupin (480p, 720p a 1080p). Každá z těchto skupin obsahuje sekvence ze všech deseti zdrojových sekvencí s různým stupněm komprese, která byla řízena pomocí bitové rychlosti. Pro rozlišení 480p je to 9-12 snímků pro každé zdrojové video. Celkem tedy 109 SD sekvencí. Pro rozlišení 720p bylo vytvořeno 5-8 snímků pro každé zdrojové video. Výsledný počet HD snímků byl 66. Pro poslední použité rozlišení 1080p bylo vytvořeno 7 videí pro každou z deseti zdrojových sekvencí. Celkem tedy 70 full HD videí. Celá databáze pak obsahuje 245 sekvencí s různým rozlišením a různým stupněm komprese.

Dále jsou v databázi k dispozici rekonstruovaná videa v nekomprimovaném formátu YUV 4:2:0 pro všechny kódované sekvence. Pro jejich zobrazení tedy není třeba mít k dispozici H.265 dekodér, což může být pro další testování výhodou.

Pro všechny sekvence je rovněž k dispozici konfigurační soubor v textovém formátu, který byl použit pro samotné kódování. Dále je k dispozici, rovněž v textovém formátu výstup enkodéru, který obsahuje parametry kódovaného videa, parametry pro kódování, celkový čas kódování, celkovou hodnotu PSNR pro jasovou a obě chrominanční složky a zároveň také informace o každém snímku sekvence. Tyto informace obsahují mimo jiné typ snímku, velikost kvantizačního parametru, pořadí kódování, pořadí přehrávání, čísla použitých referenčních snímků a PSNR pro jasovou a každou z chrominančních složek.

Dále jsou k dispozici H.264 protějšky sekvencí, které byly použity při subjektivních testech společně s jejich hodnotou PSNR (více v kapitole 5).

Databáze obsahuje hodnoty, které byly výstupem použitých objektivních metrik (viz kapitola 4). Tyto hodnoty jsou uloženy v textovém souboru u každé kódované sekvence. Pro metriky SSIM a PSNR je k dispozici jejich výstupní hodnota pro každý frame každé kódované sekvence. U metriky VQM je k dispozici jak výsledná číselná hodnota, tak i hodnoty jednotlivých složek popsanych v kapitole 1.1.3.

Hodnoty těchto metrik pro celý snímek jsou pak přehledně zobrazeny v tabulce v souboru `objective_tests.xls`, který je součástí databáze. V něm jsou rovněž k dispozici údaje o době kódování, požadované bitové rychlosti, reálné bitové rychlosti a relativní odchylce bitové rychlosti pro každé kódované video. Soubor rovněž obsahuje grafické znázornění zjištěných hodnot.

V souborech `480p_subjective_test.xls` a `720p_subjective_test.xls` jsou zaznamenány výsledky subjektivních testů opět s grafickým znázorněním. Celková velikost takto vytvořené databáze je 122 GB.

4 APLIKACE OBJEKTIVNÍCH METRIK

Na vytvořenou databázi byly aplikovány vybrané metriky objektivního hodnocení. Vzhledem k tomu že, byly k dispozici i referenční videosekvence, byly zvoleny dvě plně referenční metriky a to PSNR a SSIM, dále byla použita částečně referenční metoda VQM. Výpočty byly prováděny na stroji s těmito parametry: procesor Intel Core2Duo T9300, 2,5 GHz, RAM 3 GB 667 MHz, operační systém MS Windows 7 Professional 64-bit. Jednotlivé postupy jsou popsány v následující kapitole.

4.1 PSNR

Pro výpočet PSNR byl použit program EvalVid 2.7 psnr.exe [22], který se obsluhuje z příkazové řádky. Program počítá hodnotu PSNR pro jasovou složku každého snímku sekvence. Z těchto hodnot pak vytvoří průměr. Jednotlivé hodnoty byly ukládány do textového souboru. Použitá syntaxe byla následující:

```
psnr šířka_obrazu výška_obrazu 420 referenční_sekvence.yuv  
testovaná_sekvence.yuv > PSNR\psnr_testovaná_sekvence.txt
```

Výpočet pro sekvence 480p trval jednotky vteřin (0 s -15 s), pro sekvence 1080p desítky vteřin (20 s – 60 s) a pro sekvence 720p také jednotky až desítky vteřin (5 s – 40 s). Závislosti velikosti PSNR na bitové rychlosti jsou vidět na obr. 4-1 pro full HD sekvence, na obr. 4-2 pro HD sekvence a na obr. 4-3 pro SD sekvence. Podle průběhů lze snímky rozdělit do tří skupin podle hodnot PSNR. U první skupiny se výsledky pohybují v rozmezí 20-30 dB pro SD, 24-30 dB pro HD a 20-28 dB pro full HD sekvence. Kvalita je tedy nižší než u ostatních, jedná se o sekvence Crowd_Run, Park_Joy a Ducks_take_off. V druhé skupině se hodnoty pohybovaly v rozmezí 23-40 dB pro SD, 29-38 dB pro HD a 25-40 dB pro full HD sekvence. To naznačuje vyšší kvalitu než u první skupiny. Patří sem snímky Tractor, Bolt, Skyfall_2, Skyfall_1, Knight_and_Day. Poslední skupina obsahuje snímky s nízkou obtížností kódování a to Sunflower a 2001. Výsledné hodnoty se pak pohybují v rozmezí 26-40 dB pro SD, 29-40 pro HD a 28-42 dB pro full HD sekvence. Toto rozdělení přibližně odpovídá pozici snímků při rozložení časově-prostorové aktivity (obr. 3-1). Je tedy vidět, že tyto hodnoty ve velké míře odpovídají obtížnosti kódování.

Další možnost popisu kvality sekvencí pomocí hodnot PSNR je jejich převedení na hodnoty v rozmezí 1 až 5. Tento výpočet se provede pro každý snímek ze sekvence. K odpovídající hodnotě PSNR uvažovaného snímku je podle tabulky 5-1 [22] přiřazena celočíselná hodnota od jedné do pěti. Výsledná hodnota je pak průměr z celé videosekvence. Výhoda tohoto postupu spočívá v přímém porovnání s hodnotami získanými subjektivním hodnocením, a to hodnotami MOS. Ty se také pohybují na stupnici od jedné do pěti a proto je možné na první pohled tyto metody zběžně porovnat.

K tomuto přepočtu hodnot PSNR byl použit program EvalVid 2.7 MOS.exe s následující syntaxí:

```
mos PSNR PSNR\psnr_testovaná_sekvence.txt >  
MOS\mos_testovaná_sekvence.txt
```

Výpočet trval jednotky sekund. Změřené závislosti přibližně odpovídaly průběhům

křivek PSNR. I přes silnou degradaci výstupních sekvencí při nízkých bitových rychlostech hodnoty toho přepočtu neklesly pod 1,5. U sekvencí s nízkou obtížností kódování se při vysokých bitových rychlostech hodnoty blížily číslu 5. Jednotlivé grafické závislosti společně s odpovídajícími hodnotami jsou zaznamenány v databázi v souboru objective_tests.xls.

tabulka 4-1 Převodní hodnoty PSNR.

PSNR [dB]	
> 37	5 (Excellent)
31 – 37	4 (Good)
25 – 31	3 (Fair)
20 – 25	2 (Poor)
< 20	1 (Bad)

4.2 SSIM

Pro výpočet SSIM byl také použit program EvalVid 2.7 psnr.exe. Pomocí parametru `ssim` přidaného do syntaxe byla vypočítána hodnota SSIM pro každý snímek sekvence. Výstupem byl opět průměr. Použitý příkaz byl následující:

```
psnr šířka_obrazu výška_obrazu 420 referenční_sekvence.yuv
testovaná_sekvence.yuv ssim > SSIM\ssim_testovaná_sekvence.txt
```

Výpočet pro sekvence 480p trval jednotky vteřin (0 s-15 s), pro sekvence 1080p pak desítky vteřin (20 s-60 s). Závislosti velikosti SSIM na bitové rychlosti jsou vidět na obr. 4-4 pro full HD sekvence a na obr. 4-5 pro SD sekvence. Opět jsou zde patrné tři skupiny snímků shodné se skupinami viditelnými u PSNR. U první se hodnoty pohybují v rozmezí 0,46-0,78 pro full HD sekvence, 0,68-0,88 pro HD a 0,43-0,89 pro SD sekvence. V druhé skupině se hodnoty pohybují v rozmezí 0,78-0,98 pro full HD, 0,82-0,96 pro HD a 0,55-0,98 pro SD sekvence. Skupina s nejjednodušším kódováním dosahovala hodnot 0,82-0,98 pro full HD sekvence, 0,83-0,98 pro HD a 0,69-0,96 pro SD sekvence a to i přesto, že byly zvoleny nižší bitové rychlosti.

4.3 General VQM

Poslední metodou použitou k měření objektivní kvality byla VQM_G. K výpočtu byl použit program CVQM verze 30 dostupný z [23]. Písmeno C v názvu označuje command line – tedy ovládání z příkazové řádky. Program obsahuje rozsáhlé možnosti kalibrace. Umožňuje kalibraci dle ITU-T J.144, ITU-R BT.1683, ITU-T J.340, ITU-T J.244 a kalibraci manuální. Vzhledem k tomu, že při našich testech nebyl brán v úvahu vliv přenosové cesty, nebyla kalibrace nutná, a proto nebyla při výpočtech prováděna. Program obsahuje jak model General, tak i model Low Bandwidth (více v kapitole 1.1.3), oba v základní a rychlé variantě. V našem měření byl použit základní model General. Vstupní videosekvence mohou být jak prokládané, tak i neprokládané. Musí být v nekomprimovaném AVI souboru ve formátu RGB nebo UYVY. Vzhledem k tomu, že námi používané sekvence byly v souboru YUV ve formátu YUV420, byl k převodu použit program `ffmpeg.exe`, který je součástí balíku programů FFmpeg, [21].

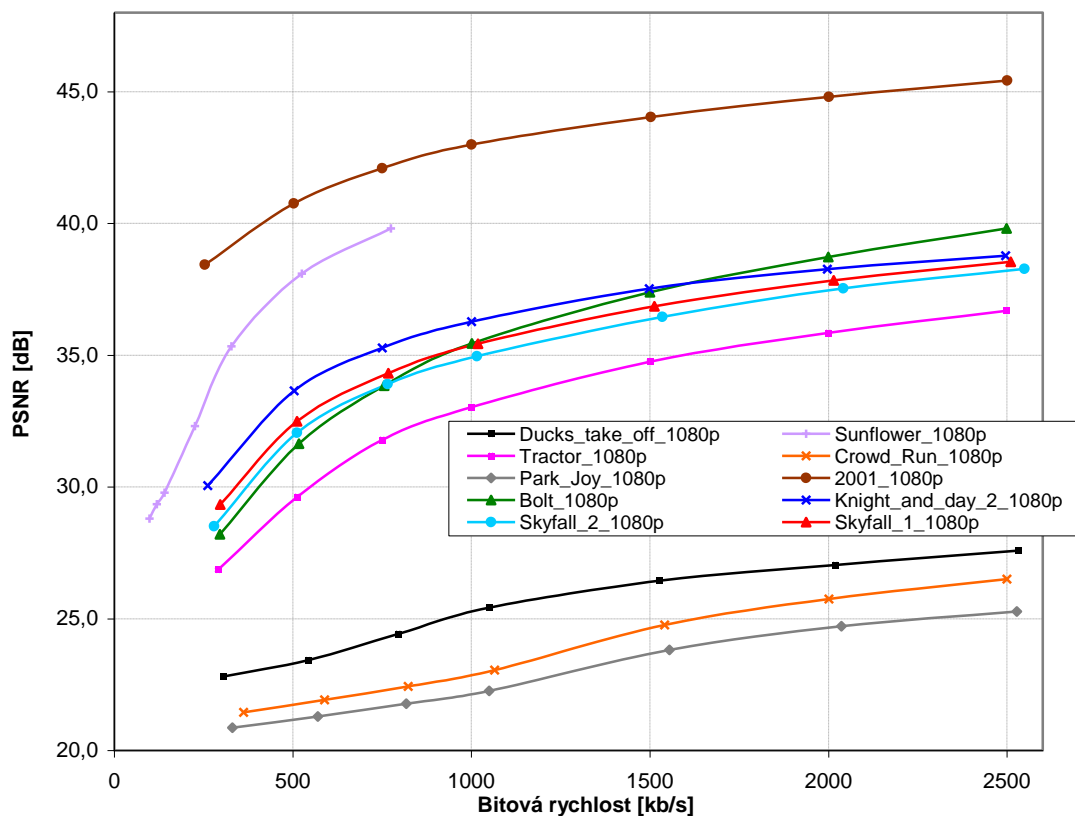
Syntaxe pro naše použití byla následující:

```
ffmpeg -pix_fmt yuv420p -s 1920x1080 -r 25 -i
testovaná_sekvence.yuv -vcodec rawvideo -pix_fmt uyvy422
testovaná_sekvence.avi
```

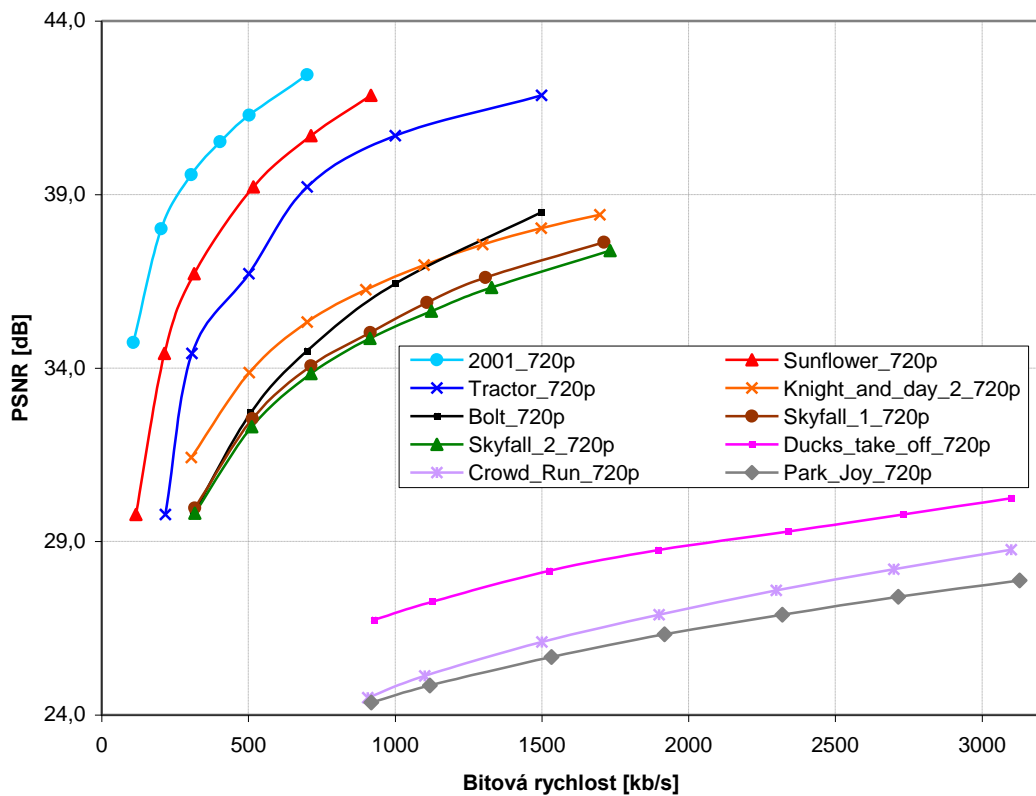
Po vytvoření sekvencí v potřebném formátu byl proveden výpočet pomocí příkazu:

```
cvqm ,referenční_sekvence.avi` ,testovaná_sekvence.avi`
,progressive` ,none` ,general`
```

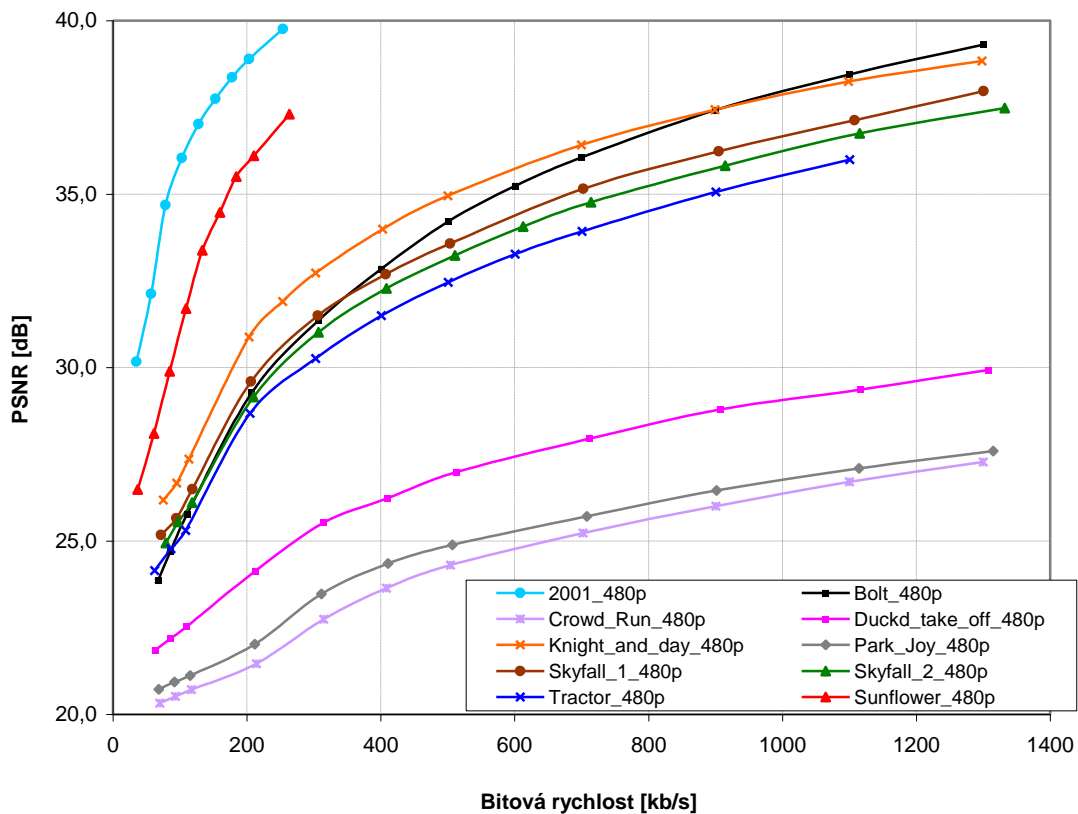
Výpočet trval přibližně dvě až tři minuty pro každou SD sekvenci, šest až osm minut pro HD sekvenci a devět až deset minut pro každou full HD sekvenci. Výstup byl uložen do textového souboru s názvem `testovaná_sekvence_model.txt`. Obsahoval jak výpis všech sedmi parametrů (viz kapitola 1.1.3) pro jednotlivá videa, tak i výslednou hodnotu, která bude využita k dalšímu zpracování. Oproti předešlým metodám není u CVQM tak výrazné rozdělení sekvencí do tří skupin. Grafické závislosti pro skupiny videí 480p a 1080p jsou zobrazeny na obr. 4-6 a 4-7. Z nich je patrné, že pro trend všech křivek je přibližně stejný a rozmístění samotných křivek je koncentrovanější. Toto naznačuje, že metoda VQM je pro popis kvality H.265 sekvencí vhodnější než předcházející metody PSNR a SSIM.



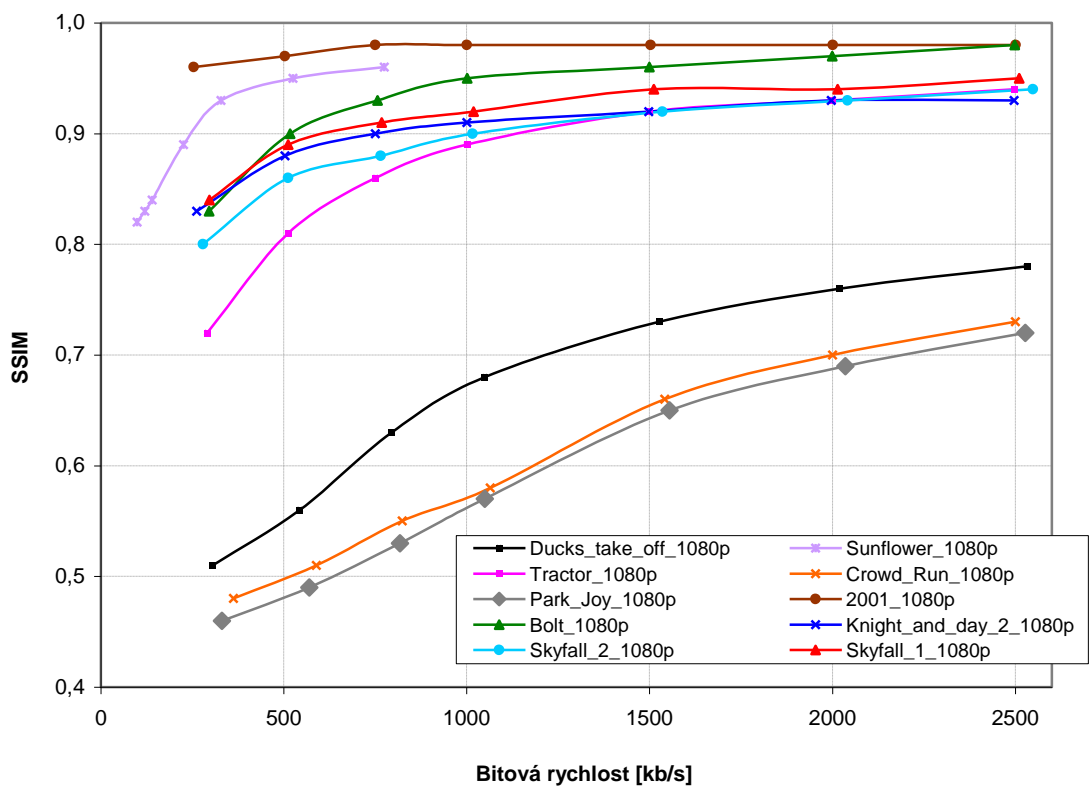
Obrázek 4-1 Závislost PSNR na bitové rychlosti pro full HD sekvence.



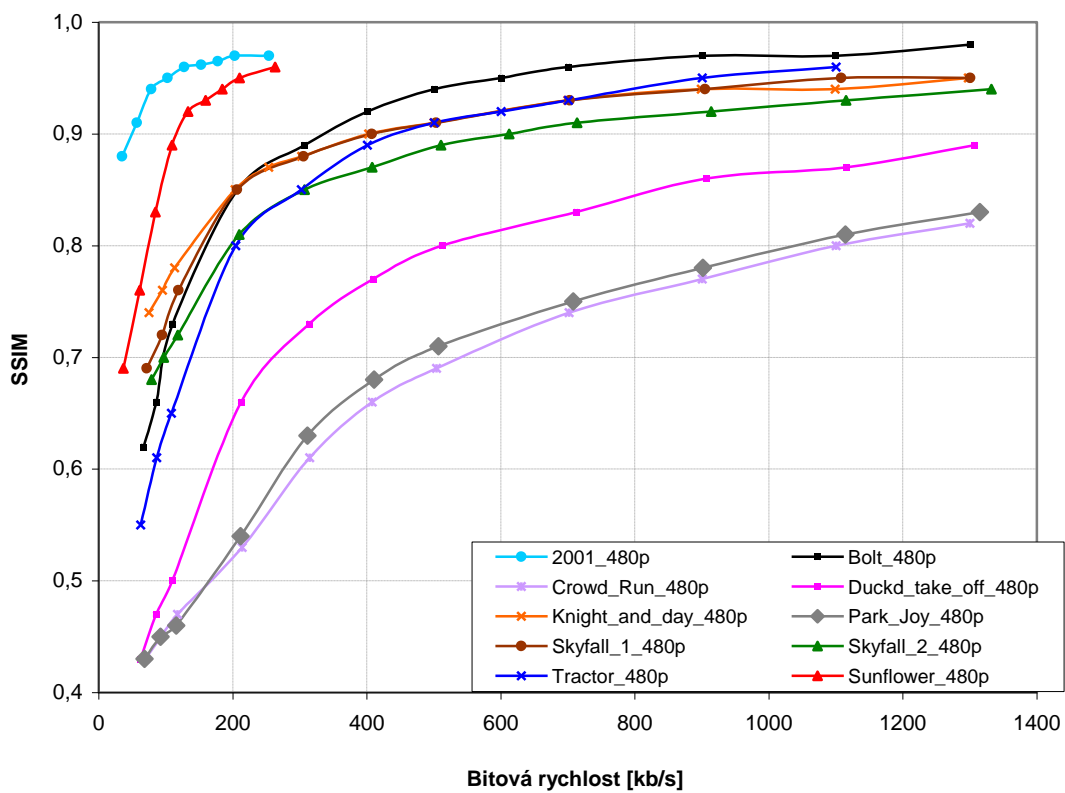
Obrázek 4-2 Závislost PSNR na bitové rychlosti pro HD sekvence



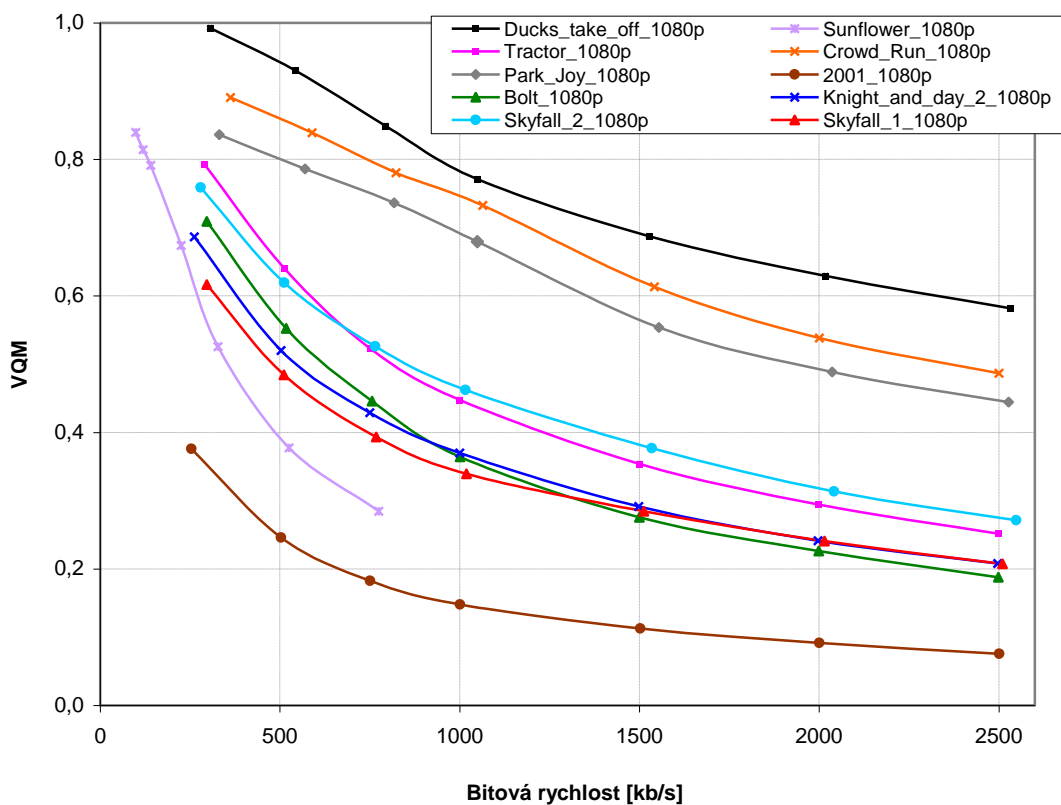
Obrázek 4-3 Závislost PSNR na bitové rychlosti pro SD sekvence.



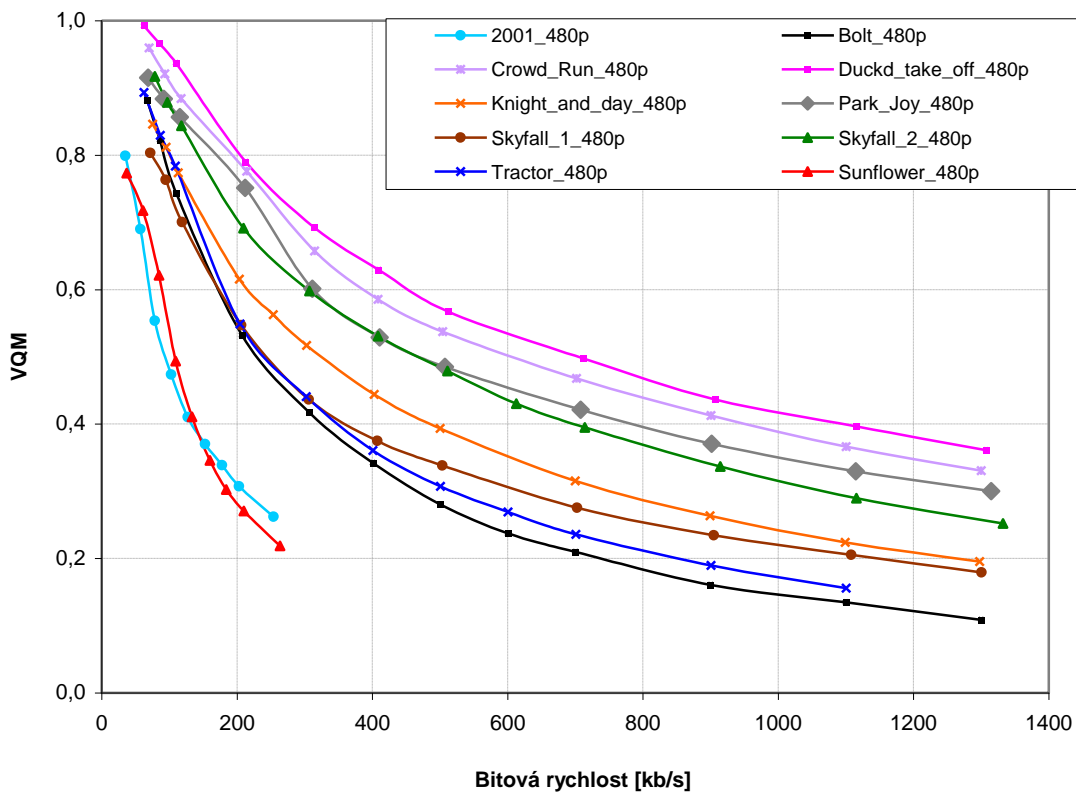
Obrázek 4-4 Závislost SSIM na bitové rychlosti pro full HD sekvence.



Obrázek 4-5 Závislost SSIM na bitové rychlosti pro SD sekvence.



Obrázek 4-6 Závislost VQM na bitové rychlosti pro full HD sekvence.



Obrázek 4-7 Závislost VQM na bitové rychlosti pro SD sekvence.

5 SUBJEKTIVNÍ HODNOCENÍ

5.1 Metoda sběru dat

Pro sběr dat byla použita metoda, při které jsou získány informace od velké skupiny pozorovatelů při běžných (ne laboratorních) podmínkách. Motivací k tomuto bylo, že se dnes video konzumuje v různých prostředích, a tomu jsme také přizpůsobili testovací metodiku. Dle zdroje [24] je tato metoda rychlá, spolehlivá a poskytuje dobré výsledky v porovnání s klasickými metodami. Jeden z hlavních důvodů, proč byla zvolena právě tato metoda spočívá ve snadném získání rozmanitého a tím i statisticky vypovídajícího panelu hodnotitelů. Při klasických metodách je potřeba zajistit přítomnost pozorovatele, či skupiny pozorovatelů v laboratoři případně jiném vhodném prostředí. To představuje problém převážně kvůli časové náročnosti a pro určité skupiny (např. podle věku) je obtížné sehnat hodnotitele ochotné zúčastnit se sezení. Zvolením této metody je možné tento problém částečně eliminovat díky tomu, že samotné hodnocení lze přesunout do prostředí, které je pro pozorovatele dostupné.

Samotná data pak mohou být sbírána dvěma způsoby. První přístup spočívá v použití webové platformy, kdy hodnocení probíhá na zařízení pozorovatele. Samotná videa jsou pak uložena na serveru, ke kterému musí mít aplikace přístup. Tím však mohou vzniknout problémy spojené s přenosem dat (např. nedostatečná rychlost, zpoždění, ztráta dat). Další nevýhodou tohoto přístupu je fakt, že každý hodnotitel má k dispozici jiný zobrazovač, tedy výrazně jiné podmínky hodnocení. Druhý přístup sběru dat tyto problémy odstraňuje tím, že samotné hodnocení probíhá na přenosném zařízení experimentátora. Ten také provede nastavení zmíněného zařízení, zná jeho parametry a má přehled o jeho chování. Při tomto způsobu může experimentátor rovněž alespoň částečně ovlivnit prostředí hodnocení. Vzhledem ke zmíněným výhodám byl pro potřeby této práce zvolen právě tento přístup.

Pro samotná hodnocení byla použita dvě přenosná zařízení. Prvním byl mobilní telefon Samsung S8530 Wave II. U něj byl původní systém Bada OS, v1.2 doplněn systémem Android Jellybean 4.3.1. z důvodu kompatibility s aplikací použitou k přehrávání sekvencí a ke sběru dat. Druhé zařízení bylo Asus Padfone společně s dokovací stanicí Asus Padfone Station. Vybrané parametry obou zařízení jsou uvedeny v tabulce 5-1.

tabulka 5-1 Parametry použitých přenosných zařízení.

	Asus Padfone Station	Samsung S8530 Wave II
Operační systém	Android 4.1.1	Android Jellybean 4.3.1.
CPU	Qualcomm Snapdragon S4 1,5 Ghz	Cortex-A8 1 GHz
GPU	Adreno 225	PowerVR SGX540
Paměť	1 GB LPDDR2	512MB
Displej	10,1" Super AMOLED, 1280 × 800	3.7" Super Clear LCD, 800 x 480

S přihlédnutím k velikosti rozlišení displeje (viz. tabulka 5-1) byla pro telefon Samsung vybrána skupina sekvencí s SD rozlišením. Pro telefon Asus bylo použito jak rozlišení SD, tak i HD. Displejem s full HD rozlišením ani jeden z dostupných přístrojů nedisponoval, proto bylo od testování těchto sekvencí opuštěno. Z vytvořené databáze byly použity sekvence s různou kvalitou, zahrnující všech deset zdrojových videosekvencí. Konkrétní rozlišení jsou uvedena v tabulce 3-3. Vzhledem k tomu, že ani jedno z obou zařízení neumožňovalo plynulé přehrávání komprimovaných videosekvencí standardu H.265 ani nekomprimovaných videosekvencí ve formátu YUV, bylo nutné sekvence překódovat do formátu H.264. Z databáze H.265 byly vybrané sekvence v rekonstruovaném formátu yuv pomocí programu Ffmpeg převedeny právě do formátu H.264. Syntaxe příkazu byla následující:

```
ffmpeg -s 848x480 -i h.265_rekonstuoovane_video.yuv -c:v libx264
-preset ultrafast -qp 1 h.264_video.mp4
```

Jak je vidět, stupeň komprese byl volen velmi malý (QP 1) tak, aby nedošlo k vytvoření dodatečných artefaktů. Při bezztrátovém kódování (QP 0) nepostačovaly hardwarové prostředky mobilních zařízení k plynulému přehrávání. Velikost QP 1 byla tedy zvolena jako kompromis. I přes to byla celková kvalita velmi dobrá. Hodnoty PSNR mezi H.265 sekvencemi a jejich H.264 protějšky se blíží k hodnotě 60 dB. Dodatečně vnesené chyby by tedy neměly být pro hodnotitele pozorovatelné. V tabulce 5-2 jsou uvedeny konkrétní hodnoty nejmenšího a největšího PSNR, které bylo počítáno pro všechny sekvence, které se uplatnily při subjektivním testu. Větší rozdíly mezi sekvencemi H.265 a H.264 byly u vyšších bitových rychlostí, tedy i u vyšší kvality.

tabulka 5-2 Hodnoty PSNR mezi H.265 a H.264 sekvencemi.

Název sekvence	PSNR [dB]			
	480p		720p	
	min	max	min	max
2001	58,95	59,18	59,12	59,71
Crowd_Run	57,97	58,13	58,02	58,17
Ducks_take_off	57,84	58,01	57,87	58,00
Park_Joy	58,11	58,52	58,16	58,46
Tractor	58,08	58,22	58,27	58,43
Sunflower	58,29	58,38	58,40	58,61
Bolt	58,73	59,15	59,11	59,58
Knight_and_day	59,49	60,26	59,99	60,96
Skyfall_1	59,55	59,87	59,90	60,39
Skyfall_2	59,16	59,87	59,48	60,38

Obě testovací zařízení byla následně upravena tak, aby na nich běžely jen nezbytně nutné procesy a aby bylo k dispozici co nejvíce volných systémových prostředků. Bylo ověřeno, že přehrávání požadovaných sekvencí je plynulé a ze strany zařízení tedy nebude vznikat dodatečné rušení.

5.2 Metoda hodnocení

S přihlédnutím ke zvolenému způsobu sběru dat byla zvolena testovací metoda ARC-HR. Hlavními kritérii pro výběr byla jednoduchost a nízká časová náročnost pro hodnotitele, což nejlépe splňuje metoda ARC. Vzhledem k tomu, že k dispozici byly i referenční sekvence, bylo vhodné zvolit variantu ARC-HR. Jak název této metody napovídá, mezi hodnocenými snímky se vyskytuje i referenční. Hodnotitel ale nemá o jeho přítomnosti informaci. Tímto postupem je alespoň částečně eliminován vliv původní sekvence

Struktura testu byla následující: Pozorovateli byl položen dotazník. Ten obsahoval položky věková kategorie, pohlaví, typ prostředí, ve kterém se provádí test a dotaz, zda má pozorovatel znalosti ohledně hodnocení videosekvencí. Poté uživatel spustil přehrávání tréninkové sekvence s dobou trvání 10 vteřin. Po skončení mu byly nabídnuty možnosti video zopakovat nebo jej ohodnotit na pětibodové stupnici podle obr. 2-6. Vzhledem k tomu, že hodnocení prováděl pouze jeden uživatel, nebyl čas na jeho hodnocení nijak omezen. Po tréninkové sekvenci následoval první testovaný snímek. Pořadí přehrávání jednotlivých sekvencí bylo pseudonáhodné (viz kapitola 5-3).

5.3 Sada testovacích sekvencí

Byly vytvořeny dvě testovací sady. První obsahovala videa v rozlišení 480p a byla testovaná jak na přístroji Samsung S8530 Wave II, tak i na Asus Padfone Station. Druhá testovací sada obsahovala videa v rozlišení 720p a vzhledem k parametrům použitých zařízení byla testována pouze na přístroji Asus Padfone Station. Každá z testovacích sad obsahovala 5 videosekvencí různé kvality od každé ze zdrojových sekvencí, tedy 50 sekvencí celkem. Společně s dalšími 10 referenčními a jednou tréninkovou sekvencí je to tedy celkem 61 videí. Při době trvání jedné sekvence 10 s je celková doba potřebná k přehrávání videí 610s. Pokud budeme uvažovat čas potřebný k ohodnocení jedné sekvence v rozmezí o pěti do deseti vteřin, dostáváme se na čas 305 až 610 vteřin. Společně s dobou přehrávání je pak celková doba trvání testu v rozmezí 915 až 1220 sekund. To odpovídá přibližně 15 až 20 minutám. Tato doba by se měla pro většinu hodnotitelů jevit jako přijatelná a neměl by být problém během ní udržet pozornost

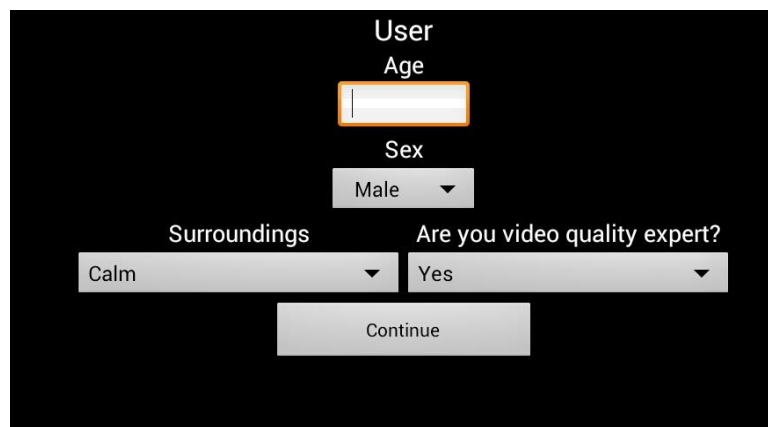
Samotné sestavení testovací sady probíhalo v několika krocích. Prvotně byly z vytvořené databáze předvybrány sekvence na základě hodnot PSNR. Ty byly vybrány tak, aby rozsah jejich kvality byl co možná největší. Tyto předvybrané sekvence byly promítnuty skupině tří hodnotitelů, kteří již v minulosti obdobné hodnocení prováděli. Od těchto pozorovatelů byly získány hodnoty MOS a zároveň také slovní komentář ke zhlédnutým sekvencím. Na základě těchto informací byly vybrány snímky tak, aby snímky od každé zdrojové sekvence obsahovaly hodnoty MOS jedna až pět. Ve druhé fázi byla takto vytvořená testovací sada promítnuta skupině pěti hodnotitelů, kteří neměli žádné předcházející zkušenosti s hodnocením kvality videí. Opět byly vypočítány hodnoty MOS jednotlivých sekvencí. U několika sekvencí bylo poté zjištěno, že zastoupení MOS není rovnoměrné a došlo tedy ještě k pozměnění těchto sekvencí. Tento postup v nezměněné podobě byl proveden jak pro sekvence v rozlišení 480p, tak i pro sekvence v rozlišení 720p. Takto získané testovací sady byly poté použity k samotnému subjektivnímu testu.

5.4 Program pro sběr dat

Pro samotné subjektivní testování sekvencí bylo potřeba vytvořit program, který bude přehrávat videa a zároveň bude sbírat hodnocení, které bude ve vhodné formě uloženo do paměti testovacího zařízení. Jak bylo zmíněno v kapitole 5-1, obě zařízení na kterých bude test prováděn, obsahují operační systém Android. Program pro testování byl tedy v jazyce Java.

Na základě volné licence byl upraven program Buscamedia Android App, který je dílem pracovníků univerzity ve španělském Madridu. K úpravě bylo využito prostředí Eclipse Standard 4.3.2 společně s balíkem Android Developer Tools. Ten obsahoval prostředky a funkce specifické právě pro Android aplikace. Vývojové prostředí rovněž umožňovalo debugování na reálném zařízení v reálném čase přes USB rozhraní, což se při úpravách ukázalo jako velká výhoda. Úpravy spočívaly v přidání pseudonáhodného pořadí přehrávání videosekvencí, vytvoření dotazníku. Dále bylo upraveno ukládání dat do databáze, přistupování k jednotlivým sekvencím a celkové grafické rozhraní.

Samotný program po spuštění zobrazí úvodní dialog, na kterém hodnotitel stisknutím tlačítka start spustí samotný test. Na následující obrazovce je poté dotazník hodnotitele. První položkou je věk, do kterého je potřeba zadat celé číslo v rozsahu 10 až 100. Další položkou je pohlaví. Následuje položka s výběrem prostředí, ve kterém je test prováděn. K dispozici je volba klidné prostředí, částečně rušivé a velmi rušivé prostředí. Poslední položkou je dotaz, zda se hodnotitel považuje za experta na kvalitu videa. Grafické rozložení dotazníku je zobrazeno na obr. 5-1.



The image shows a survey form titled "User" on a black background. The form contains the following elements:

- Age:** A text input field with a white border.
- Sex:** A dropdown menu with "Male" selected.
- Surroundings:** A dropdown menu with "Calm" selected.
- Are you video quality expert?:** A dropdown menu with "Yes" selected.
- Continue:** A button at the bottom center of the form.

Obrázek 5-1 Dotazník v programu pro testování.

Po dotazníku následuje samotné testování. Přehrávání sekvence spouští sám hodnotitel tlačítkem Start. Má tedy k dispozici kontrolu nad přehráváním a sám rozhoduje, ve které chvíli přehrávání začne. Po každé sekvenci se objeví obrazovka pro hodnocení. Pozorovatel má možnost ohodnotit právě přehranou sekvenci, případně tuto sekvenci jednou zopakovat. Pro samotné hodnocení je k dispozici pět hvězdiček. Každá hvězdička představuje jeden bod hodnocení podle stupnice na obr. 1-5. Slovní hodnocení odpovídající těmto bodům je v textové podobě zobrazeno nad hvězdičkami. Jedna hvězda představuje špatnou (bad) kvalitu, kdežto pět hvězd představuje vynikající (excellent) kvalitu. Grafické znázornění obrazovky pro hodnocení je uvedeno na obr. 5-2. Po ohodnocení se hodnotiteli opět objeví obrazovka s tlačítkem, které spouští další testovanou sekvenci. Po přehrávání a ohodnocení všech videí z testované sady se zobrazí obrazovka s poděkováním a tlačítkem pro ukončení testu.



Obrázek 5-2 Obrazovka pro hodnocení.

Pořadí přehrávání videí je pseudonáhodné. Po startu programu je vytvořeno pole *RANDOM_NUMBERS* s čísly od jedné do celkového počtu sekvencí v testované sadě. Každému číslu pak odpovídá právě jedna sekvence. Toto pole s čísly je pak po každém stisknutí tlačítka start na úvodní obrazovce náhodně přeskupeno. Přeskupení je provedeno podle následující syntaxe:

```
Collections.shuffle(RANDOM_NUMBERS, new  
Random(System.currentTimeMillis()));
```

Při samotném přehrávání se pak berou hodnoty z tohoto pole od prvního až po poslední místo. Pro dalšího pozorovatele je pole opět náhodně přeskupeno. Tímto postupem je zajištěno, že každý hodnotitel má unikátní pořadí sekvencí.

Hodnocení od pozorovatelů je již v průběhu testu ukládáno do paměti zařízení ve formě SQL databáze. V ní jsou uložena jak samotná hodnocení, tak i informace z dotazníku a čas startu a konce testu.

Testovaná videa byla uložena v paměti obou zařízení a vzhledem k tomu, že původní program obsahoval konkrétní názvy sekvencí, musel být upraven pro použití pro konkrétní testovací sadu.

5.5 Složení panelu hodnotitelů

Hodnocení se celkově zúčastnilo 69 respondentů. Z toho 22 účastníků hodnotilo testovací sadu s rozlišením 720p a 47 účastníků hodnotilo sadu s rozlišením 480p. Ze sady 480p bylo 27 hodnocení provedeno na zařízení Asus Padfone Station a 20 hodnocení na zařízení Samsung S8530 Wave II. Panel respondentů byl vybírán tak, aby byl co nejvíce rozmanitý. Procentuální zastoupení jednotlivých skupin je uvedeno v tabulce 5-3. Jsou zde hodnoty jak celkové tak i hodnoty pro dílčí testované soubory. Je vidět, že z hlediska pohlaví jsou obě skupiny víceméně rovnoměrně zastoupeny. Z věkových kategorií byla nejpočetnější kategorie v rozmezí 10 až 29 let, které měla u všech sad zastoupení větší než 50%. Zastoupení ostatních věkových kategorií bylo v rozsahu 10 až 26%. Nejstaršímu účastníku testu bylo 75 let, nejmladšímu pak 15 let, průměrný věk hodnotitelů byl 33 let.

Pro získání co nejspolehlivějších výsledků testů bylo prostředí předem vybíráno tak, aby bylo co nejméně rušivé. Žádný z respondentů proto v dotazníku ne zvolil položku velmi rušivé. Položku klidné zvolilo 61 % respondentů a položku částečně rušivé 39 % respondentů. Počet hodnotitelů, kteří uvedli, že se považují za experty na kvalitu videa byl 10 %. Většina hodnotitelů tedy neměla s podobným hodnocením předcházející zkušenosti.

Průměrný čas potřebný k testu byl 16 minut a 45 vteřin. Maximální čas byl potom 27:22 a minimální 13:36. Pouze 3 respondentům připadal čas neadekvátní, z nich jeden hodnocení nedokončil a jeho výsledky nebyly pro další zpracování brány v potaz.

tabulka 5-3 Procentuální zastoupení jednotlivých skupin respondentů.

	celkem	720p	480p		
			celkem	Samsung	Asus
Pohlaví					
muži	51%	41%	55%	55%	56%
ženy	49%	59%	45%	45%	44%
Věk					
10-29	64%	64%	64%	75%	56%
30-49	17%	9%	21%	15%	25%
50+	19%	27%	15%	10%	19%
Prostředí					
klidné	61%	59%	62%	60%	63%
částečně rušivé	39%	41%	38%	40%	37%
velmi rušivé	0%	0%	0%	0%	0%
Expert					
ano	10%	14%	9%	0%	15%
ne	90%	86%	91%	100%	85%

Poučení hodnotitelů před testem probíhalo podle doporučení ITU-T P.910, přičemž některým hodnotitelům převážně vyššího věku bylo nutno specifikovat princip bezreferenčního hodnocení a zdůraznit, že mají hodnotit kvalitu přehrávané sekvence.

5.6 Zpracování dat

Získaná data byla zkopírována z obou testovacích zařízení a pro snazší zpracování byla převedena z formátu SQL databáze do souboru XLS. Pro převod byl použit program SQLite Administrator. Data byla následně zpracována v programu Matlab.

Na získané hodnoty od všech hodnotitelů byl aplikován test, který detekoval pozorovatele, jejichž hodnoty se ve velké míře lišily od hodnot ostatních pozorovatelů. Výsledky těchto hodnotitelů totiž nemusí být vypovídající a mohou negativně ovlivnit celkový výsledek testu.

Rozsah hodnocení s_{ij} sekvence i od hodnotitele j , které bylo považováno za vyhovující, byl vypočten podle následujícího vzorce [25]:

$$s_{ij} > q_3 + 1,5(q_3 - q_1) \vee s_{ij} < q_1 - 1,5(q_3 - q_1), \quad (20)$$

kde q_1 a q_3 jsou dolní (25-ti procentní) a horní (75-ti procentní) kvantily normálního rozložení.

Tento rozsah přibližně odpovídá $\pm 2,7$ násobku směrodatné odchylky, tedy 99,3 % pokrytí normálního rozložení. Hodnotitel by byl pokládán za nevyhovujícího, pokud by více jak 20 % jeho hodnocení nespadlo do rozsahu definovaného vzorcem (20). Z celkového počtu 47 hodnotitelů bylo zaznamenáno 13 s 2 % hodnot mimo rozsah, 3 se 4 % hodnot mimo rozsah a 1 hodnotitel s 8 % hodnot mimo rozsah. Jak je z těchto čísel patrné, při tomto subjektivním testování nebyl žádný hodnotitel detekován jako nevyhovující.

Vzhledem k tomu, že data obsahovala hodnocení i pro referenční sekvence, byly jednotlivé hodnoty přepočítány na diferenční skóre podle vzorce (14). Tím došlo k částečné eliminaci vlivu referenčního videa. V případě, že pozorovatel ohodnotil referenční sekvenci nízkou hodnotou (například 2) a testovanou sekvenci vysokou hodnotou (například 5) vzorec (14) způsobil, že výsledná hodnota diferenčního skóre byla 7 (3-5+5). Aby nedocházelo k nepřiměřenému ovlivnění celkových výsledků takto vysokými hodnotami, byl dále aplikován vzorec (15). Tím bylo vypočteno zalimitované diferenční skóre, jehož maximální hodnota byla 5,44. Jak je vidět, toto číslo je blízké hodnotě pět (tedy maximální hodnotě na stupnici dle obr. 1-5) a vliv na celkový výsledek je přiměřenější.

Z takto získaných hodnot již mohl být vypočten diferenční MOS (DMOS), který byl použit pro následující statistické zpracování. Samotný výpočet probíhal podle následujícího vzorce:

$$DMOS_i = \frac{\sum_j^N s_{ij}}{N}, \quad (21)$$

kde N je počet platných hodnotitelů a s_{ij} je zalimitované diferenční skóre videa i od hodnotitele j .

Vzhledem k tomu, že vypočtené hodnoty DMOS byly získány pouze od vzorku populace (tedy od hodnotitelů účastnících se experimentu), bylo potřeba popsat vztah těchto hodnot k teoretickým hodnotám celé populace. Vhodným nástrojem k tomuto je výpočet intervalů spolehlivosti (CI) k odhadnuté střední hodnotě. Byla zvolena hladina významnosti $\alpha = 0,05$. To odpovídá 95% šanci, že reálná střední hodnota se bude pohybovat právě v tomto intervalu. Vzhledem k tomu, že rozsah souboru byl malý (počet hodnotitelů byl 22 a 47 pro první respektive pro druhou sadu testovacích sekvencí), bylo pro výpočet zvoleno Studentovo rozložení. To je oproti normálnímu rozložení kritičtější, což je pro takto malý soubor výhodou.

Samotný výpočet lze popsat následujícím vzorcem [25]:

$$CI_i = t_{(1-\alpha/2, N-1)} \frac{\sigma_i}{\sqrt{N}}, \quad (22)$$

kde $t_{(1-\alpha/2, N-1)}$ je kvantil Studentova rozložení se stupněm volnosti $N-1$ na požadované hladině významnosti α , N je počet platných hodnotitelů a σ_i je směrodatná odchylka zalimitovaného diferenčního skóre sekvence i od všech hodnotitelů.

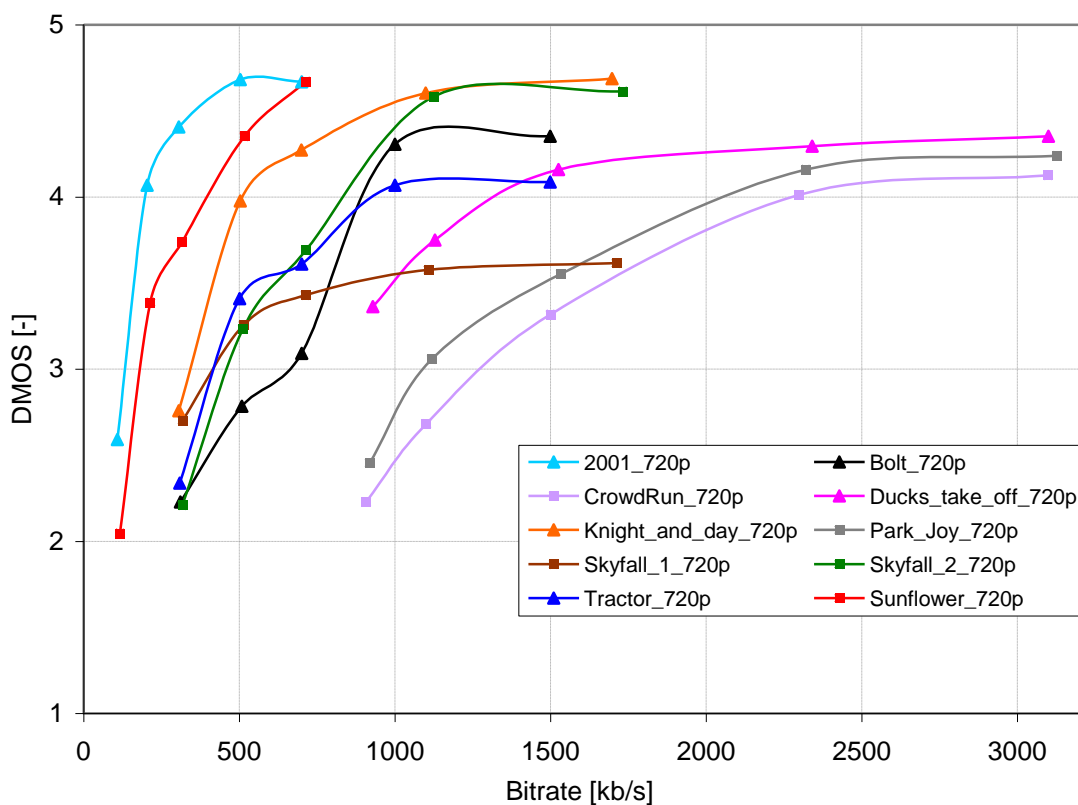
Pro testovací sadu 480p byly hodnoty DMOS a CI počítány zvlášť pro data získaná ze zařízení Samsung S8530 Wave II a ze zařízení Asus Padfone Station. Zároveň byly vypočteny hodnoty pro sloučená data z obou zařízení. Pro každou z položek dotazníku (věk, pohlaví, prostředí a zkušenost s testováním) byl také proveden samostatný výpočet DMOS společně se zjištěním patřičných intervalů spolehlivosti, opět s hladinou významnosti $\alpha = 0,05$. Jednotlivé výsledky společně s jejich grafickou reprezentací jsou součástí vytvořené databáze a jsou uloženy v souborech 720p_subjective_test.xls, 480p_subjective_test.xls, 480p_padfone_subjective_test.xls a 480p_samsung_subjective_test.xls.

5.7 Výsledky hodnocení

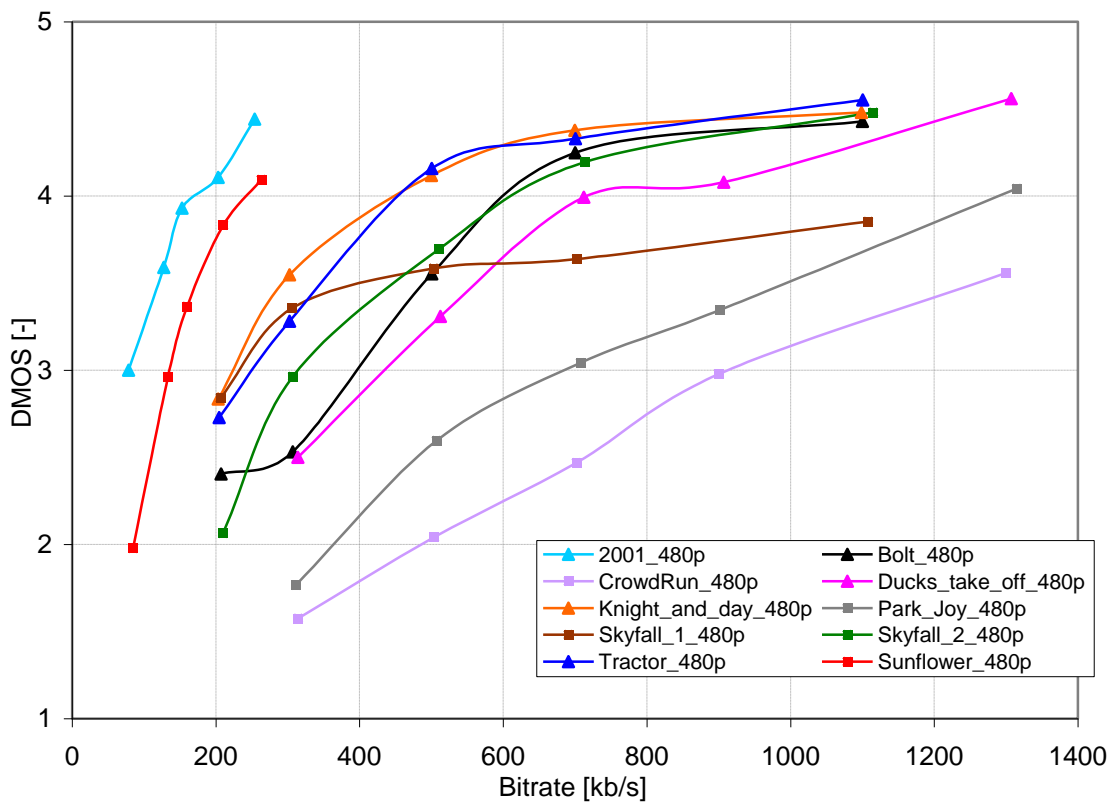
Hodnoty DMOS získané ze subjektivních testů byly graficky vyneseny v závislosti na velikosti bitové rychlosti a hodnot, objektivních metrik odpovídajících sekvencí. Jednotlivé body patřící k jedné zdrojové sekvenci byly spojeny křivkou tak, aby byl na první pohled patrný její trend. Tyto výsledky jsou na obr. 5-3 až 5-12.

5.7.1 Bitová rychlost

Na obr. 5-3 a 5-4 je znázorněna závislost DMOS na bitové rychlosti pro sekvence od jednotlivých zdrojových videí. Jsou zde na první pohled patrné sekvence s nízkou časovou a prostorovou aktivitou (2001 a Sunflower). U nich je dosahováno vysokých hodnot DMOS i pro nízké bitové rychlosti. Naopak pro sekvence s vysokou spektrální a časovou aktivitou (Park_Joy, Croew_Run a Duckds_take_off) je vidět, že i pro vysokou bitovou rychlost sekvence nedosahují tak dobrých výsledků. Dále lze pozorovat, že u sekvence Skyfall_1, která má podobné parametry jako sekvence Skyfall_2 a Knight_and_day od určité bitové rychlosti dochází k menšímu přírůstku hodnot DMOS než u ostatních sekvencí. Toto je podrobněji popsáno v kapitole 5.7.8.



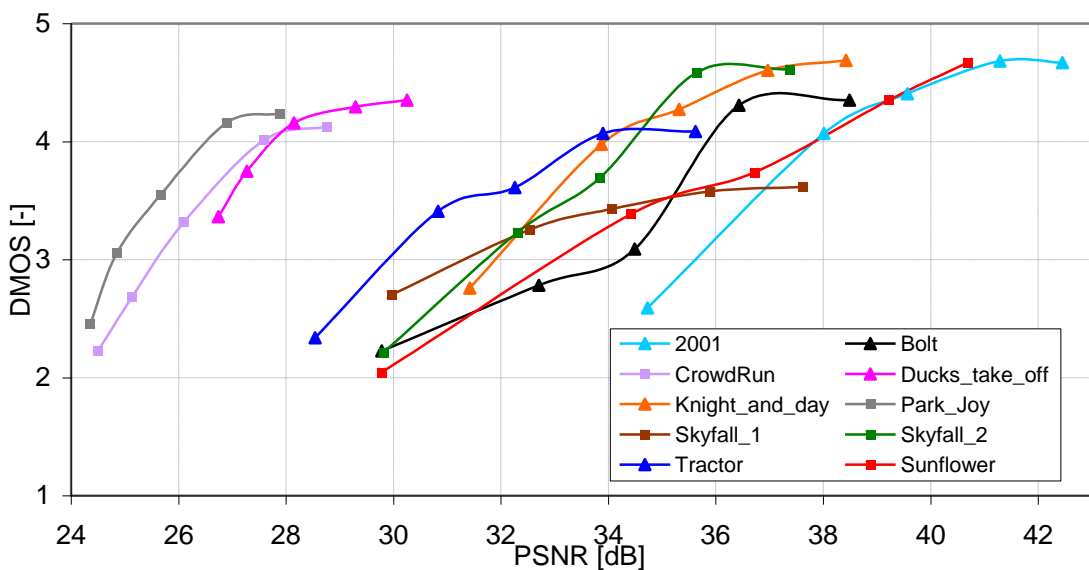
Obrázek 5-3 DMOS v závislosti na bitové rychlosti pro rozlišení 720p.



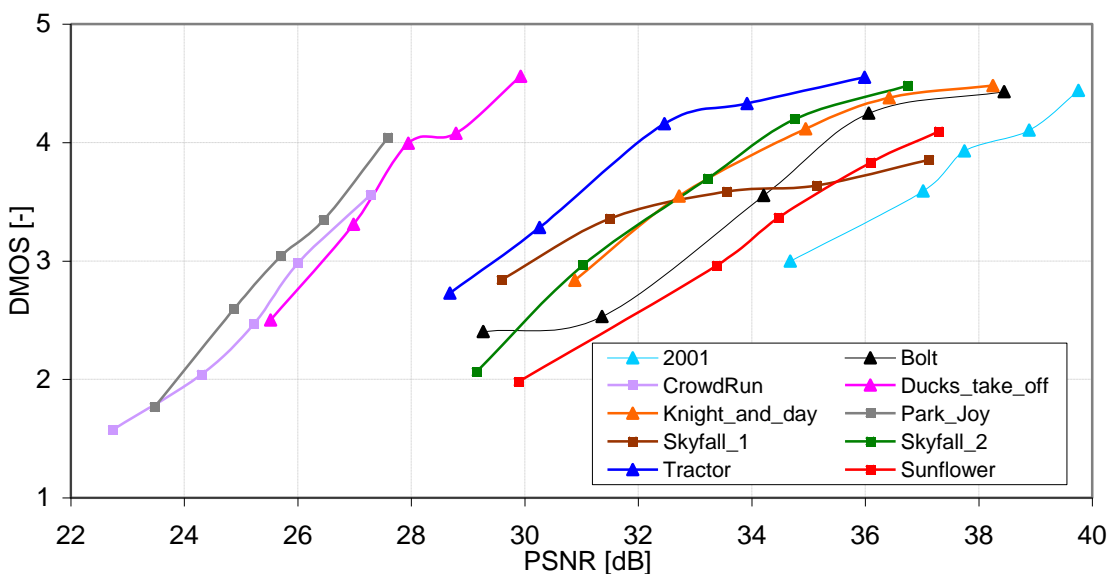
Obrázek 5-4 DMOS v závislosti na bitové rychlosti pro rozlišení 480p.

5.7.2 PSNR

Z obrázků 5-5 a 5-6 je patrné, že trend křivek je velmi podobný, celkový rozptyl jednotlivých křivek je však poměrně velký. Na první pohled není rozdíl mezi sekvencemi 480p a 720p nijak zvlášť výrazný. Sekvence s vysokou prostorovou i časovou aktivitou dosahují dobrých hodnot DMOS při daleko nižších hodnotách PSNR než ostatní sekvence. Rozdíl je nejmarkantnější u sekvencí Ducks_take_off, Park_Joy a Crowd_Run. Toto je pravděpodobně způsobeno potlačením detailů, které nejsou pro lidský zrak při rozhodování o kvalitě tak důležité. Metrika PSNR však tyto rozdíly zachycuje. Naopak u sekvencí s nízkou časovou a prostorovou aktivitou (2001 a Sunflower) nejsou při vysokých hodnotách PSNR hodnoty DMOS až o tolik lepší. Tato vlastnost společně s přihlédnutím k celkovému rozptylu jednotlivých křivek naznačuje, že metrika PSNR nemusí postihovat kvalitu kodéru H.265 v odpovídající míře.



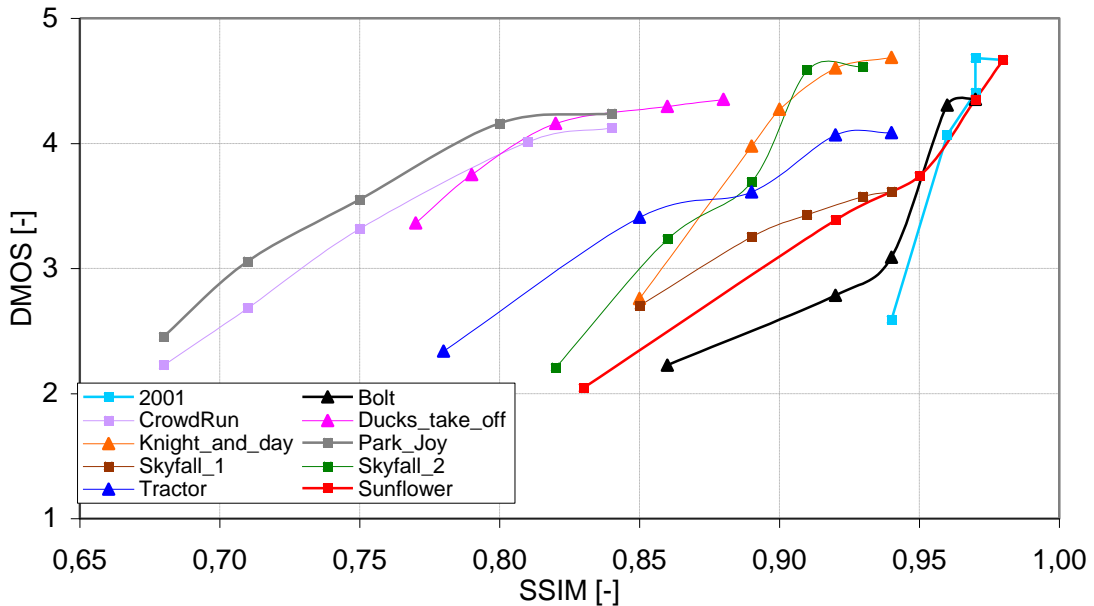
Obrázek 5-5 DMOS v závislosti na PSNR pro rozlišení 720p.



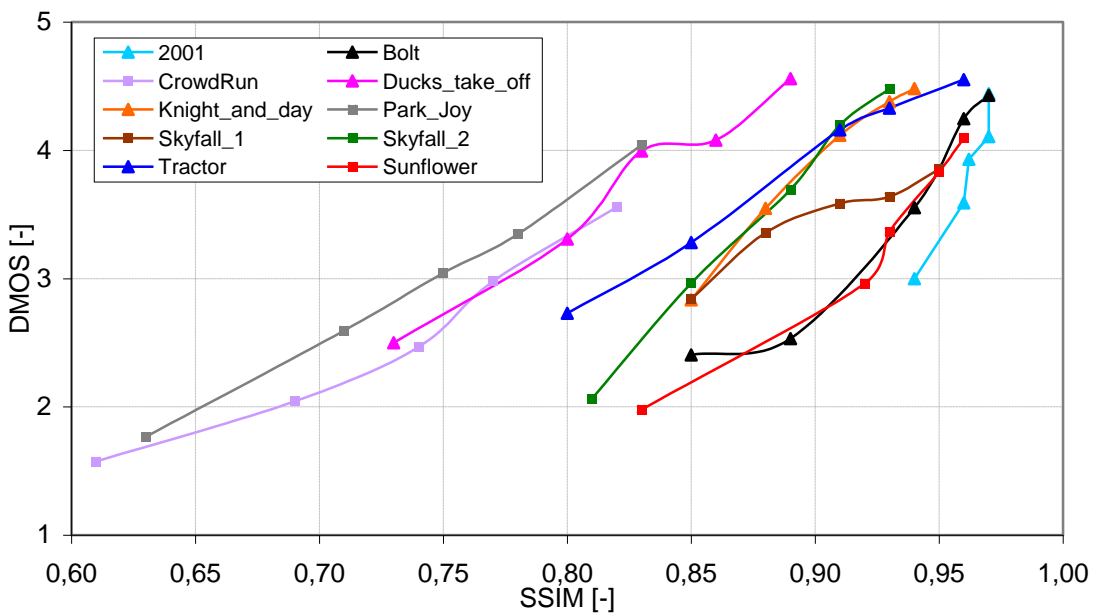
Obrázek 5-6 DMOS v závislosti na PSNR pro rozlišení 480p.

5.7.3 SSIM

Na obr. 5-7 a 5-7 je zobrazena závislost hodnot DMOS na SSIM. Je patrné, že rozložení jednotlivých křivek je koncentrovanější. Opět je však možno rozlišit sekvence s vysokou spektrální a časovou aktivitou a sekvence s nízkou časovou a spektrální aktivitou. Rozdíl mezi sekvencemi 480p a 720p opět není na první pohled nějak zvlášť výrazný a trendy křivek jsou u obou sad podobné. Stejně jako u předcházející metriky se i u metriky SSIM grafické závislosti napovídají, že tato metoda ne vždy plně koresponduje s hodnotami DMOS pro H.265 videosekvence.



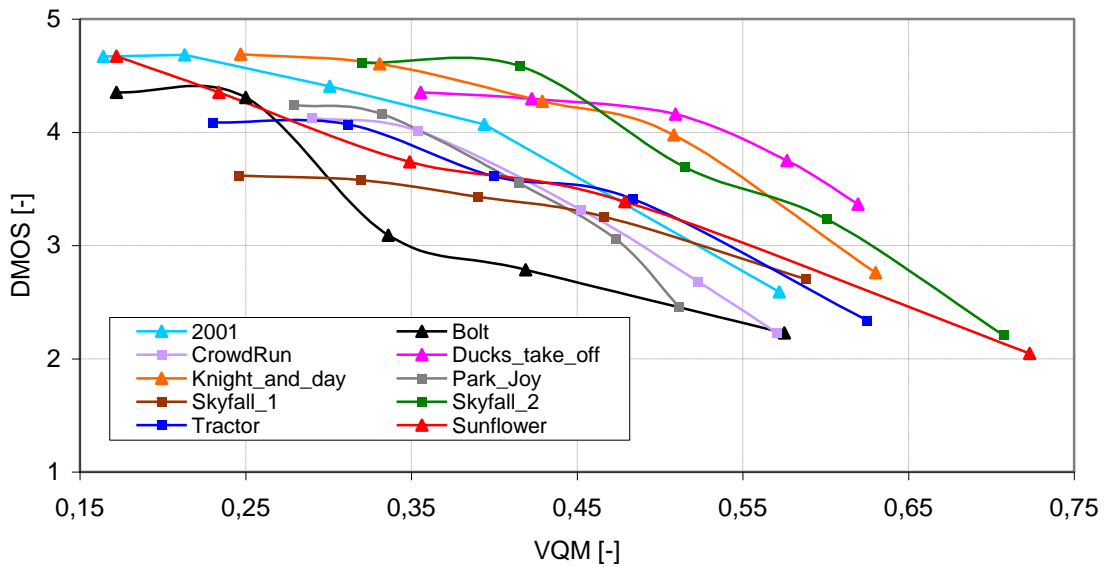
Obrázek 5-7 DMOS v závislosti na SSIM pro rozlišení 720p.



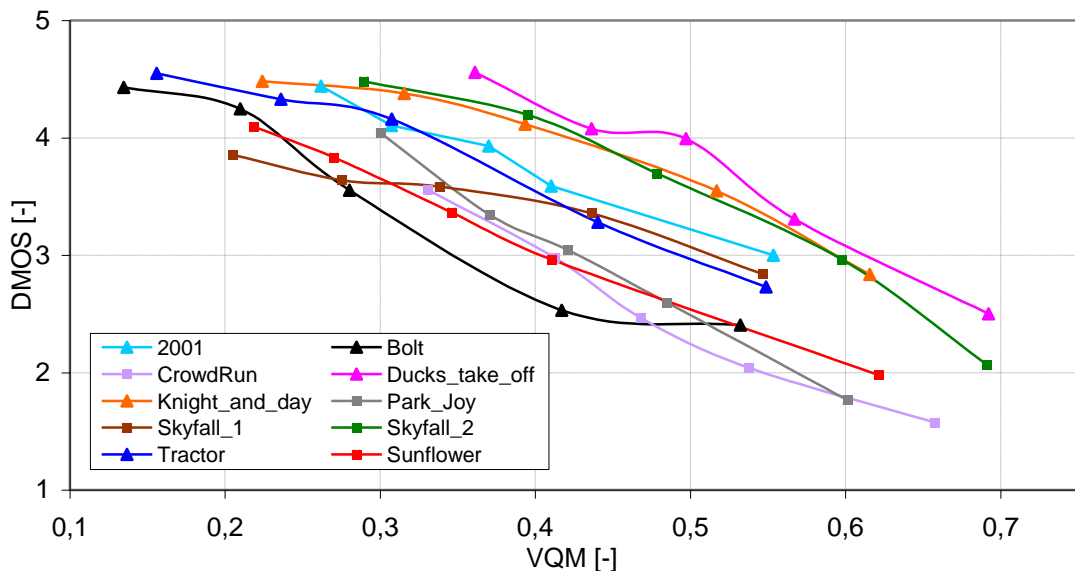
Obrázek 5-8 DMOS v závislosti na SSIM pro rozlišení 480p.

5.7.4 VQM

Na obr. 5-9 a 5-10 je znázorněna závislost DMOS všech testovaných sekvencí na hodnotách získaných metrikou VQM. Jak je zmíněno v kapitole 1.1.3, hodnota 0 odpovídá kvalitě shodné s referencí a hodnota 1 odpovídá velmi nízké kvalitě. To má za následek opačný trend křivek, než je tomu u metrik SSIM a PSNR. Oproti těmto metrikám je také z obrázků na první pohled patrné, že rozptyl jednotlivých křivek je poměrně malý. Nelze podle jejich rozložení jednoznačně určit, které sekvence dosahují vysokých či nízkých hodnot spektrální a časové aktivity. Trend všech křivek je také víceméně stejný. Tyto vlastnosti naznačují, že tato metoda by mohla být pro hodnocení kvality H.265 videosekvencí výhodnější než metody PSNR a SSIM.



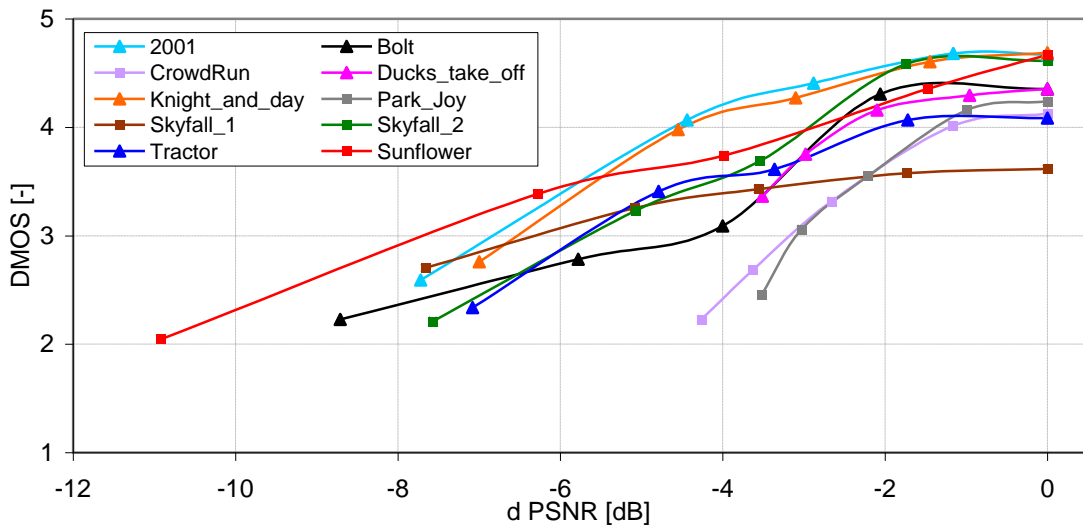
Obrázek 5-9 DMOS v závislosti na VQM pro rozlišení 720p.



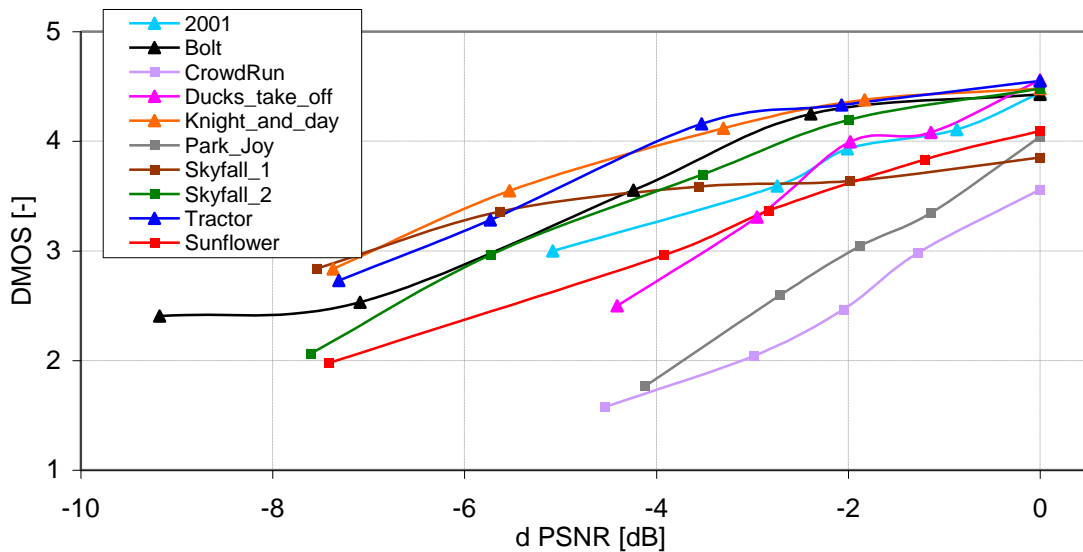
Obrázek 5-10 DMOS v závislosti na VQM pro rozlišení 480p.

5.7.5 Diferenční stupnice

Vzhledem k tomu, že rozložení křivek pro jednotlivá videa u metriky PSNR a SSIM bylo značně široké, byla k popisu použita ještě diferenční stupnice jejich hodnot. Tento postup předpokládá, že kodér je optimalizován pro volbu kvality tak, aby odpovídala subjektivnímu vjemu. Tedy, že kodér se bude snažit dodržet vnímanou kvalitu pro různé hodnoty PSNR různých sekvencí. [26] Při tomto výpočtu je tedy brán rozdíl mezi maximální hodnotou PSNR a hodnotou odpovídající zkoumanému bodu pro jeden typ sekvence. Výsledky se potom pohybují od nuly do hodnoty přibližně -10dB. Video s nejvyšší kvalitou má tedy o 10 dB vyšší hodnotu PSNR než video s nejnižší kvalitou. Tímto přístupem byl snížen rozptýl jednotlivých křivek. Jejich trend však zůstal zachován, jak je vidět na obr. 5-11 a 5-12. Tento postup byl aplikován na všechny tři objektivní testy. U metriky PSNR a SSIM došlo k výraznému zlepšení, u metriky VQM nebylo toto zlepšení pozorováno.



Obrázek 5-11 DMOS v závislosti na diferenčním PSNR pro rozlišení 720p.



Obrázek 5-12 DMOS v závislosti na diferenčním PSNR pro rozlišení 480p.

5.7.6 Korelační koeficienty

Pro výpočet korelačních koeficientů byly použity dvě metody. Byla to Pearsonova metoda PLCC (Pearson's Linear Correlation Coefficient) a Spearmanova metoda SROCC (Spearman's Rank Correlation Coefficient). Korelace byla počítána mezi hodnotami DMOS a výsledky objektivních metrik v normálním i v diferenčním měřítku. Výpočet byl proveden v programu Matlab pomocí funkce *corr* s parametrem *'Pearson'* pro výpočet PLCC a s parametrem *'Spearman'* pro výpočet SROCC. Koeficienty byly zjištěny jak pro sadu 720p tak i pro sadu 480p. U té byl proveden výpočet jak pro celkové hodnoty, tak i pro hodnoty získané z jednotlivých zařízení. Výsledky jsou uvedeny v tabulce 5-4. Vyšší korelace bylo dosaženo pro SD rozlišení. Dále je vidět, že korelace je lepší pro zařízení s menším displejem. Vzhledem k tomu, že s rostoucí hodnotou DMOS kvalita roste, kdežto s rostoucí hodnotou VQM kvalita klesá, vycházely korelační koeficienty záporně. U této metriky je potom do výsledku uvažována absolutní hodnota. Při porovnávání vybraných objektivních metod se jako nejvhodnější pro popis kvality sekvencí kódovaných kodekem H.265 jeví metoda VQM. Její korelační koeficienty s hodnotami DMOS získanými subjektivními testy se pohybují okolo hodnoty 0,8, což naznačuje dobrou shodu. Metody PSNR a SSIM se jeví jako slabší. Jejich korelační koeficienty se pohybují mezi přibližně 0,6 a 0,7, přičemž pro metriku PSNR jsou nepatrně nižší. Při předpokladu možnosti využití diferenčních hodnot těchto metrik však jejich korelační koeficient vzroste. U metriky VQM k tomuto nárůstu při použití difference nedochází.

tabulka 5-4 Hodnoty korelačních koeficientů SROCC.

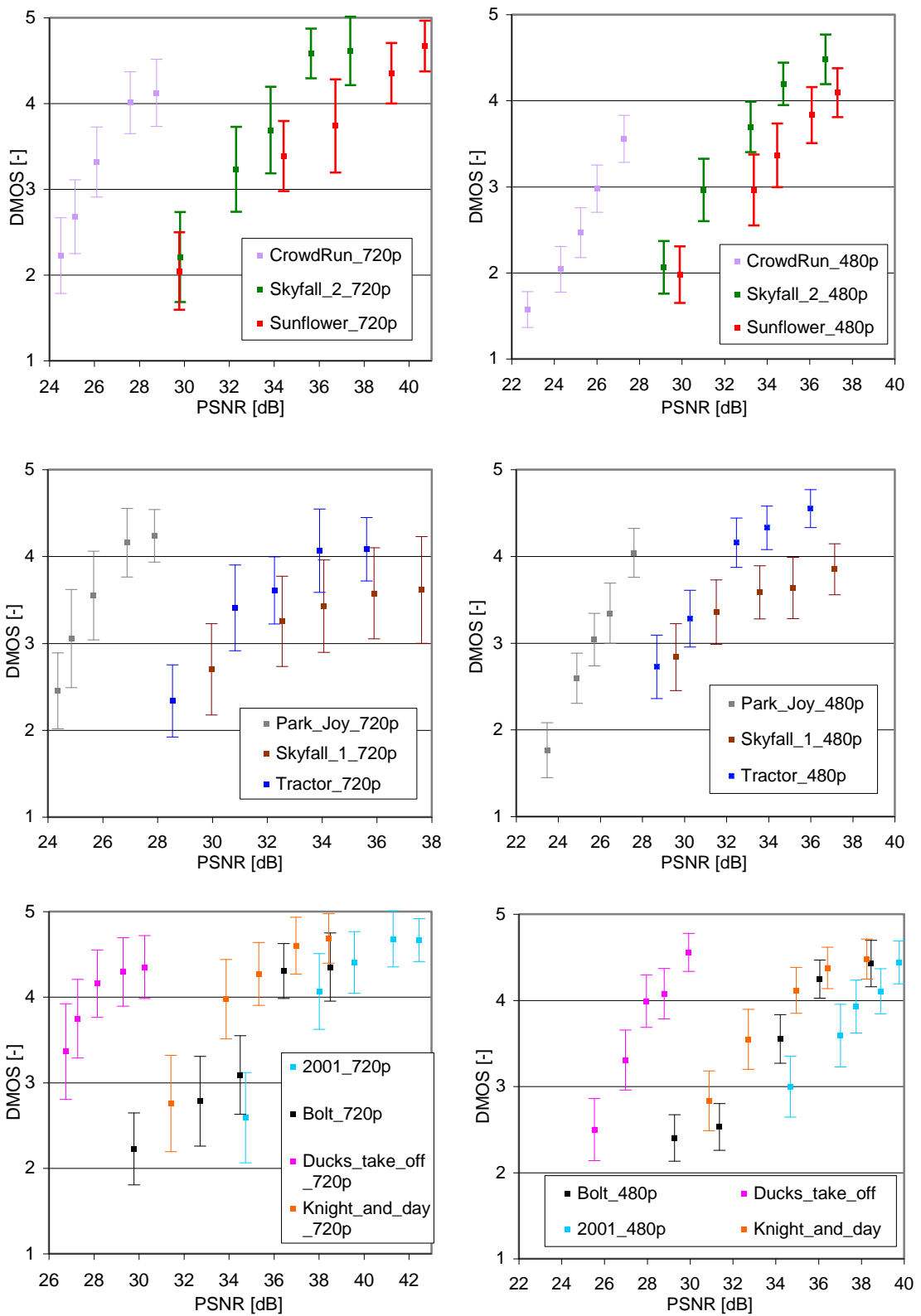
		PSNR	d PSNR	SSIM	d SSIM	VQM	d VQM
720p		0,58	0,83	0,56	0,83	0,79	0,85
480p	Celkem	0,72	0,69	0,77	0,88	0,81	0,79
	Samsung	0,75	0,66	0,78	0,86	0,81	0,77
	Asus	0,64	0,70	0,72	0,86	0,77	0,79

tabulka 5-5 Hodnoty korelačních koeficientů PLCC.

		PSNR	dPSNR	SSIM	dSSIM	VQM	dVQM
720p		0,60	0,82	0,58	0,82	0,78	0,82
480p	Celkem	0,71	0,73	0,72	0,87	0,81	0,80
	Samsung	0,74	0,69	0,75	0,86	0,81	0,77
	Asus	0,61	0,73	0,65	0,83	0,77	0,80

5.7.7 Intervaly spolehlivosti CI

Na obr. 5-13 jsou zachyceny intervaly spolehlivosti pro hodnoty DMOS pro všechny hodnocené sekvence z obou testovaných sad. Pro sadu o rozlišení 720p jsou tyto intervaly větší. To je částečně způsobeno menším souborem hodnotitelů pro toto rozlišení. Pro sadu 480p, kde byl počet respondentů více než dvojnásobný, jsou tyto intervaly na první pohled užší. Do hodnot CI se také mohly negativně promítnout změny podmínek pozorování, kdy byl hodnotitel v danou chvíli vyrušen. Další zkreslení také mohlo být způsobeno tím, že hodnotitel ať už omylem nebo úmyslně v testovacím programu zvolil jinou hodnotu kvality než pozoroval. Celkově ale intervaly nenabývají hodnot, které by se daly označit jako nevyhovující.



Obrázek 5-13 Velikosti CI hodnot DMOS (vlevo 720p, vpravo 480p).

5.7.8 Další zjištěné vlastnosti

Bylo zjištěno, že na subjektivní hodnocení má velký vliv i kvalita několika posledních vteřin sekvence. U snímku Skyfall_1 docházelo v poslední sekundě ke značné degradaci kvality (více v kapitole 3-4). Objektívni metriky jsou ve většině aplikací počítány přes celý časový úsek a na výsledné hodnotě se tedy několik posledních vteřin příliš neprojeví. Z výsledků subjektivních hodnocení je však jasné, že na hodnotitele má tato změna kvality velký vliv. Jak je vidět na grafech 5-3 až 5-12, od určité fáze dochází k minimálnímu nárůstu hodnoty DMOS u této sekvence. Jinak řečeno, i při vysoké bitové rychlosti a relativně vysoké kvalitě prvních devíti vteřin videa bude pozorovatel velice ovlivněn onou poslední vteřinou se sníženou kvalitou. Toto se pak negativně projeví na jeho hodnocení.

Další vlastnost, která byla pozorována převážně u sekvence Ducks_take_off, a to velká změna kvality mezi snímky I a B jednotlivých sekvencí, nebyla pozorovateli zaznamenána. Tento problém tedy nebyl tak rušivý jak se na první pohled zdál, a na celkovém hodnocení nebyl patrný.

Při zpracování dat z testovacího souboru 480p pro obě zařízení bylo zjištěno, že hodnoty DMOS jsou vyšší u přístroje ASUS než u přístroje Samsung. Rozdíly byly nejvyšší u sekvencí Sunflower a Ducks_take_off, ve kterých se v centru obrazu pohybuje výrazný objekt. Tyto rozdíly místy dosahovaly hodnoty, která byla nepatrně větší než jedna, tedy značný rozdíl hodnocené kvality. Nejmenší rozdíly (místy téměř nulové) byly u akčních sekvencí se střihem, jmenovitě Skyfall_1, Skyfall_2, Knight_and_Day a Bolt. U ostatních sekvencí se tyto rozdíly pohybovaly v řádu desetin. Hlavním důvodem pro jiné hodnocení byl rozdíl ve velikosti obrazovky obou zařízení. Pro větší zobrazovač byly v obraze zřetelnější detaily a hodnocení tedy bylo daleko víc kritické. Grafické zobrazení této vlastnosti je pro vybrané sekvence uvedeno v příloze C.

Dále bylo zjištěno, že ženy hodnotí převážnou část videosekvencí méně kriticky. U akčních sekvencí Skyfall a Knight_and_day nebyl rozdíl tak výrazný a v některých případech bylo hodnocení mužů kritičtější. Celkové rozdíly mezi hodnocením mužů a žen však nebyly zvlášť výrazné a hodnoty DMOS se lišili maximálně o několik desetin. Příklad je rovněž v příloze C.

Z hlediska věku byly rozdíly výraznější. Dle předpokladu nejkritičtější hodnocení bylo u skupiny s věkovým rozsahem 10-29 let. Tato skupina rovněž dosahovala nejnižšího rozptylu při jednotlivých hodnocení. Respondenti nad 50 let věku hodnotili kvalitu ze všech skupin nejlépe. V některých případech byl DMOS této skupiny o více než jedna vyšší oproti celkovému DMOS. Hodnocení této skupiny pak také vykazovalo nejvyšší rozptyl. Příklad pro vybrané sekvence je uveden v příloze C.

Dle předpokladu respondenti, kteří uvedli, že se považují za experty na kvalitu videa, hodnotili sekvence daleko přísněji. Tato skupina má totiž znalosti o vlastnostech chyb přítomných v kódovaných videosekvencích a proto se na ně zaměřuje. Počet hodnotitelů ze skupiny expert byl však malý, což znemožnilo další zpracování, které by mělo statisticky vypovídající hodnotu.

Z hlediska vlivu okolí bylo pro sekvence 480p v klidném prostředí hodnocení méně kritické. U sekvencí 720p nebyl pozorován jednoznačný trend a obecné závěry tedy nelze určit.

ZÁVĚR

V práci byly popsány základní postupy při hodnocení kvality videa. Byly uvedeny nejpoužívanější metody objektivního hodnocení kvality, konkrétně metriky s plnou nebo alespoň částečnou referencí. Byly nastíněny rozdíly mezi metodami využívající model lidského vidění (HVS) a pixelově-orientovanými metrikami. Dále byly popsány nejdůležitější metody subjektivního hodnocení kvality. Byla zmíněna jejich vhodnost pro použití v konkrétních aplikacích. Práce obsahuje základní popis komprimačního standardu H.265/HEVC a jeho nejdůležitějších nástrojů.

V práci je popsán postup komprimace videosekvencí ve vysokém a standardním rozlišení s využitím referenční implementace HM-12 standardu HEVC. Jsou popsány nejdůležitější parametry použité pro kompresi. Je zmíněna časová náročnost kódování a dále pak schopnost enkodéru dodržet předepsanou bitovou rychlost. Byla vytvořena databáze videosekvencí, která celkem obsahuje 245 snímků, 70 ve full HD rozlišení, 66 v HD rozlišení a 110 v SD rozlišení. Kapitola 3 obsahuje popis této databáze. Na ní byly aplikovány vybrané objektivní testy. Byly použity metody SSIM, PSNR, General VQM a DMOS mapovaný na základě hodnot PSNR.

Dále byl navržen postup pro sběr dat subjektivního hodnocení kvality pro SD a HD sekvence. Byla zvolena metoda ACR-HR. Ta jak název napovídá, obsahuje skrytou referenci, čím je eliminováno zaujetí pozorovatele vůči sekvenci. Tato metoda je časově méně náročná, a tím pádem vhodná ke zvolené implementaci. Byly sestaveny dvě sady videí pro testování, na které byla tato metoda implementována. Data byla statisticky zpracována a porovnána s výsledky objektivních metod.

K porovnání byly použity jak grafické závislosti, tak i vybrané metody korelace. Ze získaných hodnot vyplývá, že subjektivním hodnocením nejlépe odpovídá metrika VQM. Metriky PSNR a SSIM dosahovaly poněkud horších výsledků. Ke zlepšení došlo v případě uvažování diferenční závislosti, kdy se korelační koeficient těchto metrik blížil hodnotám dosahovaným u VQM.

LITERATURA

- [1] WINKLER, Stefan. *Digital video quality: vision models and metrics*. Chichester: John Wiley, 2005, xiv, 175 s. ISBN 04-700-2404-6.
- [2] RICHARDSON, Iain E. *Video codec design: developing image and video compression systems*. Chichester: Wiley, 2002, x, 303 s. ISBN 04-714-8553-5.
- [3] KLEJMOVÁ, E. *Srovnání výkonnosti nejnovějších standardů pro kódování videosekvencí*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2012. 65 s. Vedoucí bakalářské práce Ing. Martin Slanina, Ph.D.
- [4] WANG, Z., A.C. BOVIK, H.R. SHEIKH a E.P. SIMONCELLI. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing* [online]. 2004, vol. 13, issue 4, s. 600-612 [cit. 2014-05-20]. DOI: 10.1109/TIP.2003.819861. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1284395>
- [5] ANSI T1.801.03. American National Standard for Telecommunications – Digital transport of one-way video signals – Parameters for objective performance assessment. American National Standards Institute. 2003
- [6] Rec. ITU-T J.144. Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference. Geneva: ITU, 2004
- [7] Rec. ITU-T J.143. User requirements for objective perceptual video quality measurements in digital cable television. Geneva: ITU, 2000
- [8] PINSON, M.H. a S. WOLF. A New Standardized method for Objectively Measuring Video Quality. *IEEE Transactions on Broadcasting* [online]. 2004, vol. 50, issue 3, s. 312-322 [cit. 2014-05-20]. DOI: 10.1109/TBC.2004.834028. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1337089>
- [9] Rec. BT.1683-0. Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference. Geneva: ITU, 2004
- [10] WOLF, Stephen a Margaret PINSON. Video Quality Measurement Techniques. In: [online]. June 2002 s. 113 [cit. 2013-12-14]. Dostupné z: <http://www.its.bldrdoc.gov/publications/2423.aspx>
- [11] PINSON, Margaret a Stephen WOLF. Low Bandwidth Reduced Reference Video Quality Monitoring System. [online]. January 2005 s. 4 [cit. 2013-12-14]. Dostupné z: <http://www.its.bldrdoc.gov/publications/2575.aspx>
- [12] Rec. ITU-R BT.500-13. Methodology for the subjective assessment of the quality of television Picture. Geneva: ITU, 2012
- [13] Rec. ITU-R BT.710-4. Subjective assessment methods for image quality in high-definition television. Geneva: ITU, 1998
- [14] Rec. ITU-R P.910. Subjective video quality assessment methods for multimedia applications. Geneva: ITU, 2008
- [15] Rec. ITU-T H.265. High efficiency video coding. Geneva: ITU, 2013.

- [16] OHM, Jens-Rainer, Gary J. SULLIVAN, Heiko SCHWARZ, Thiow Keng TAN a Thomas WIEGAND. Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC). *IEEE Transactions on Circuits and Systems for Video Technology* [online]. 2012, vol. 22, issue 12, s. 1669-1684 [cit. 2014-05-20]. DOI: 10.1109/TCSVT.2012.2221192. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6317156>
- [17] SULLIVAN, Gary J., Jens-Rainer OHM, Woo-Jin HAN, Thomas WIEGAND a Thomas WIEGAND. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Transactions on Circuits and Systems for Video Technology* [online]. 2012, vol. 22, issue 12, s. 1649-1668 [cit. 2014-05-20]. DOI: 10.1109/TCSVT.2012.2221191. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6316136>
- [18] BOSSEN, Frank, Benjamin BROSS, Karsten SUHRING, David FLYNN a Thomas WIEGAND. HEVC Complexity and Implementation Analysis. *IEEE Transactions on Circuits and Systems for Video Technology* [online]. 2012, vol. 22, issue 12, s. 1685-1696 [cit. 2014-05-20]. DOI: 10.1109/TCSVT.2012.2221255. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6317152>
- [19] Svn_HEVCSoftware – Revision 3759. [online]. [cit. 2013-12-14]. Dostupné z: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/branches/HM-12.0-dev/
- [20] Xiph.org Video Test Media [derf's collection]. *Xiph.Org Foundation* [online]. [cit. 2013-12-16]. Dostupné z: <http://media.xiph.org/video/derf/>
- [21] SCHWARZ, K. *Ffmpeg W64 Static build version git-f618cb1 (2013-12-16)* [online]. 2013. [cit. 2013-12-14]. Dostupný z : <http://ffmpeg.zeranoe.com/builds>.
- [22] KLAUE, Jirka. EvalVid – A Video Quality Evaluation Tool-set. *Telecommunication Networks GROUP*[online]. [cit. 2013-12-14]. Dostupné z: <http://www2.tkn.tu-berlin.de/research/evalvid/>
- [23] Video Quality Metric (VQM) Software. *Institute for Telecommunication Sciences* [online]. [cit. 2013-12-14]. Dostupné z: <http://www.its.bldrdoc.gov/resources/video-quality-research/software.aspx>
- [24] FIGUEROLA SALAS, Óscar, Velibor ADZIC, Akash SHAH, Hari KALVA a Thomas WIEGAND. Assessing internet video quality Using crowdsourcing. *Proceedings of the 2nd ACM international workshop on Crowdsourcing for multimedia – CrowdMM ,13* [online]. New York, New York, USA: ACM Press, 2013, vol. 22, issue 12, s. 23-28 [cit. 2014-05-20]. DOI: 10.1145/2506364.2506366. Dostupné z: <http://dl.acm.org/citation.cfm?doid=2506364.2506366>
- [25] HANHART, Philippe, Martin RERABEK, Francesca DE SIMONE, Touradj EBRAHIMI a Andrew G. TESCHER. Subjective quality evaluation of the upcoming HEVC video compression standard. *Proceedings of the 2nd ACM international workshop on Crowdsourcing for multimedia – CrowdMM ,13* [online]. New York, New York, USA: ACM Press, 2013, vol. 22, issue 12, 84990V- [cit. 2014-05-20]. DOI: 10.1117/12.946036. Dostupné z: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.946036>
- [26] SEUFERT, Michael, Martin SLANINA, Sebastian EGGER, Meik KOTTKAMP a Andrew G. TESCHER. “To pool or not to pool: A comparison of temporal pooling methods for HTTP adaptive video streaming. *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)* [online]. New York, New York, USA: IEEE, 2013, vol. 22, issue 12, s. 52-57 [cit. 2014-05-20]. DOI: 10.1109/QoMEX.2013.6603210. Dostupné z: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6603210>

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

ACR	Absolute Category Rating
ACR-HR	Absolute Category Rating with Hidden Reference
ANSI	American National Standards Institute
AVC	Advanced Video Coding
AVC	Advanced Video Coding
CABAC	Context-adaptive Binary Arithmetic Coding
CB	Coding Block
CI	Confidence Interval
CTB	Coding Tree Block
CTU	Coding Tree Unit
CU	Coding Unit
DCT	Discrete Cosine Transform
DSCQS	Double-Stimulus Continuous Quality-scale
DSIS	Double-stimulus Impairment Scale
DST	Discrete Sine Transform
GOP	Group of Pictures
HEVC	High Efficiency Video Coding
HVS	Human Visual System
ITU	International Telecommunication Union
JCT-VC	Joint Collaborative Team on Video Coding
MOS	Mean Opinion Score
MPEG	Motion Picture Expert Group
MPEG	Moving Picture Experts Group
MSE	Mean Squared Error
NTIA	National Telecommunications and Information Administration
PB	Prediction Block
PLCC	Pearson's Linear Correlation Coefficient
PSNR	Peak Signal-to-noise Ratio
PU	Prediction Unit
QP	Quantization Parameter
SAO	Sample Adaptive Offset
SDSCE	Simultaneous Double Stimulus for Continuous Evaluation
SROCC	Spearman's Rank Correlation Coefficient
SSCQE	Stimulus Continuous Quality Evaluation

SEZNAM PŘÍLOH

A **Náhledy použitých videosekvencí**

B **Příklad konfiguračního souboru**

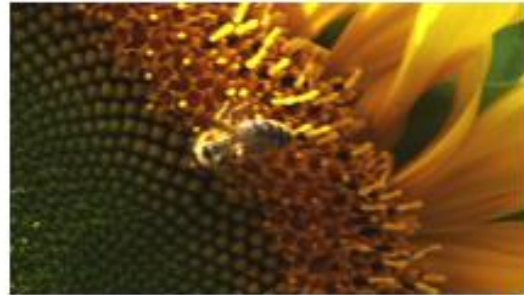
C **Doplňkové graf**

D **DVD**

A NÁHLEDY POUŽITÝCH VIDEOSEKVENCÍ



Tractor



Sunflower



Ducks_take_off



Crowd_Run



Park_Joy



Bolt



2001



Skyfall_1



Skyfall_2



Knight_and_Day

B PŘÍKLAD KONFIGURAČNÍHO SOUBORU

```
##### File I/O #####
InputFile          :
/home/xklejm00/zdrojove_sekvence/2001_1056x480p25_250f.yuv
InputBitDepth      : 8          # Input bitdepth
FrameRate          : 25         # Frame Rate per second
FrameSkip          : 0          # Number of frames to be
skipped in input
SourceWidth        : 1056       # Input  frame width
SourceHeight       : 480        # Input  frame height
FramesToBeEncoded  : 250        # Number of frames to be
coded
##### File I/O #####
BitstreamFile      : /home/xklejm00/2001/2001_200.bin
ReconFile          : /home/xklejm00/2001/2001_200_rec.yuv

##### Profile #####
Profile            : main
Level              : 3
##### Unit definition #####
MaxCUWidth         : 64         # Maximum coding unit width in
pixel
MaxCUHeight        : 64         # Maximum coding unit height in
pixel
MaxPartitionDepth  : 4          # Maximum coding unit
depth
QuadtreeTULog2MaxSize : 5      # Log2 of maximum transform size
for
                                # quadtree-based TU coding
(2...6)
QuadtreeTULog2MinSize : 2      # Log2 of minimum transform size
for
                                # quadtree-based TU coding
(2...6)
QuadtreeTUMaxDepthInter : 3
QuadtreeTUMaxDepthIntra : 3

##### Coding Structure #####
IntraPeriod        : 16        #Period of I-Frame ( -1 = only
first)
DecodingRefreshType : 0        # Random Accesss 0:none, 1:CDR, 2:IDR
GOPSize            : 8         # GOP Size (number of B slice =
GOPSize-1)
# Type POC Qpoffset Qpfactor tcOffsetDiv2 betaOffsetDiv2 temporal_id
#ref_pics_active #ref_pics reference pictures      predict deltaRPS
#ref_idcs reference idcs
Frame1: B  8  1  0.442    0  0  0  4  4  -8 -10 -12 -16  0
Frame2: B  4  2  0.3536   0  0  0  2  3  -4 -6  4      1  4  5  1  1
0  0  1
Frame3: B  2  3  0.3536   0  0  0  2  4  -2 -4  2  6      1  2  4  1  1
1  1
Frame4: B  1  4  0.68      0  0  0  2  4  -1  1  3  7      1  1  5  1  0
1  1  1
```



```

Frame5: B 3 4 0.68      0 0 0 2 4 -1 -3 1 5      1 -2 5 1 1
1 1 0
Frame6: B 6 3 0.3536  0 0 0 2 4 -2 -4 -6 2      1 -3 5 1 1
1 1 0
Frame7: B 5 4 0.68      0 0 0 2 4 -1 -5 1 3      1 1 5 1 0
1 1 1
Frame8: B 7 4 0.68      0 0 0 2 4 -1 -3 -7 1      1 -2 5 1 1
1 1 0

#===== Motion Search =====
FastSearch                : 1                # 0:Full search 1:TZ
search
SearchRange                : 32              # (0: Search range is a Full
frame)
BipredSearchRange         : 4              # Search range for bi-prediction
refinement
HadamardME                : 1              # Use of hadamard measure for
fractional ME
FEN                        : 1              # Fast encoder decision
FDM                        : 1              # Fast Decision for Merge RD
cost

#===== Quantization =====
QP                        : 32              # Quantization
parameter(0-51)
MaxDeltaQP                : 0              # CU-based multi-QP
optimization
MaxCuDQPDepth            : 0              # Max depth of a minimum CuDQP P
DeltaQpRD                 : 0              # Slice-based multi-QP
optimization
RDOQ                      : 1              # RDOQ
RDOQTS                   : 1              # RDOQ for transform skip

#===== Deblock Filter =====
DeblockingFilterControlPresent: 0          # Dbl control params
present
LoopFilterOffsetInPPS      : 0            # Dbl params: 0=varying
params in SliceHeader, param = base_param + GOP_offset_param;
1=constant params in PPS, param = base_param)
LoopFilterDisable         : 0            # Disable deblocking
filter
LoopFilterBetaOffset_div2  : 0            # base_param: -6 ~ 6
LoopFilterTcOffset_div2   : 0            # base_param: -6 ~ 6
DeblockingFilterMetric     : 0            # blockiness metric

#===== Misc. =====
InternalBitDepth           : 8            # codec operating bit-
depth

#===== Coding Tools =====
SAO                       : 1            # Sample adaptive offset (0: OFF, 1: ON)
AMP                       : 1            # Asymmetric motion partitions (0: OFF,
1: ON)
TransformSkip             : 1            # Transform skipping (0: OFF, 1: ON)
TransformSkipFast        : 1            # Fast Transform skipping (0: OFF, 1: ON)
SAOLcuBoundary           : 0            # SAOLcuBoundary using non-deblocked
pixels

#===== Slices =====

```

```

SliceMode                : 0
SliceArgument            : 1500          # Argument for
'SliceMode'.
LFCrossSliceBoundaryFlag : 1            # In-loop filtering

#===== PCM =====
PCMEnabledFlag           : 0             # 0: No PCM mode
PCMLog2MaxSize           : 5            # Log2 of maximum PCM block
size.
PCMLog2MinSize           : 3            # Log2 of minimum PCM block
size.
PCMInputBitDepthFlag     : 1            # 0: PCM bit-depth is
internal bit-depth. 1: PCM bit-depth is input bit-depth.
PCMFilterDisableFlag     : 0            # 0: Enable loop filtering on
I_PCM samples. 1: Disable loop filtering on I_PCM samples.

#===== Tiles =====
UniformSpacingIdc        : 0             # 0: the column
boundaries are indicated by ColumnWidth array, the row boundaries are
indicated by RowHeight array
                                                                    # 1: the column
and row boundaries are distributed uniformly
NumTileColumnsMinus1    : 0             # Number of
columns in a picture minus 1
ColumnWidthArray         : 2 3          # Array
containing ColumnWidth values in units of LCU (from left to right in
picture)
NumTileRowsMinus1       : 0             # Number of
rows in a picture minus 1
RowHeightArray           : 2            # Array
containing RowHeight values in units of LCU (from top to bottom in
picture)

LFCrossTileBoundaryFlag : 1             # In-loop
filtering is across or not across tile boundary.
                                                                    # 0: not across,
1: across

#===== WaveFront =====
WaveFrontSynchro         : 0             # 0: No
WaveFront ynsynchronization (WaveFrontSubstreams must be 1 in this
case).
                                                                    # >0: WaveFront
synchronises with the LCU above and to the right by this many LCUs.

#===== Quantization Matrix =====
ScalingList               : 0             # ScalingList 0
: off, 1 : default, 2 : file read
ScalingListFile           : scaling_list.txt # Scaling List
file name. If file is not exist, use Default Matrix.

#===== Lossless =====
TransquantBypassEnableFlag: 0 # Value of PPS flag.
CUTransquantBypassFlagValue: 0 # Constant lossless-value signaling per
CU, if TransquantBypassEnableFlag is 1.

#===== Rate Control =====
RateControl               : 1             # Rate control:
enable rate control

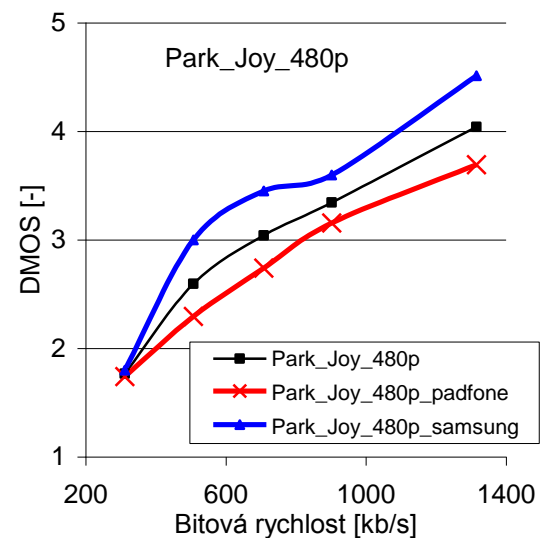
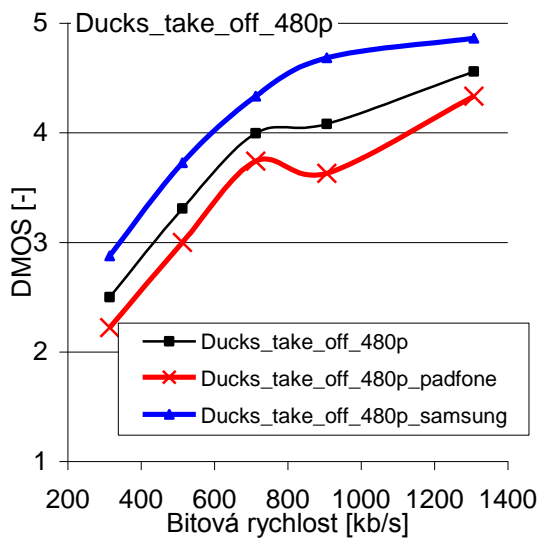
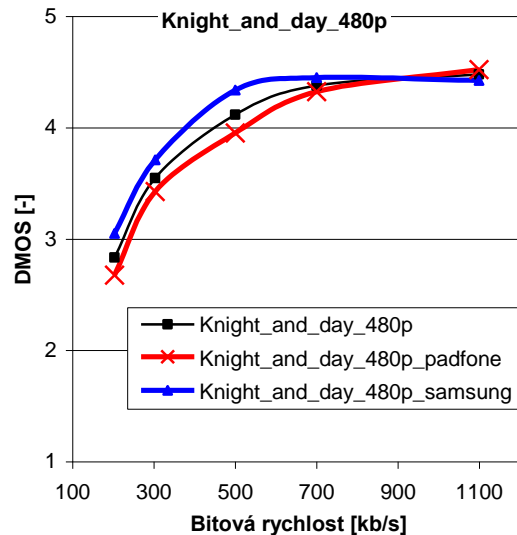
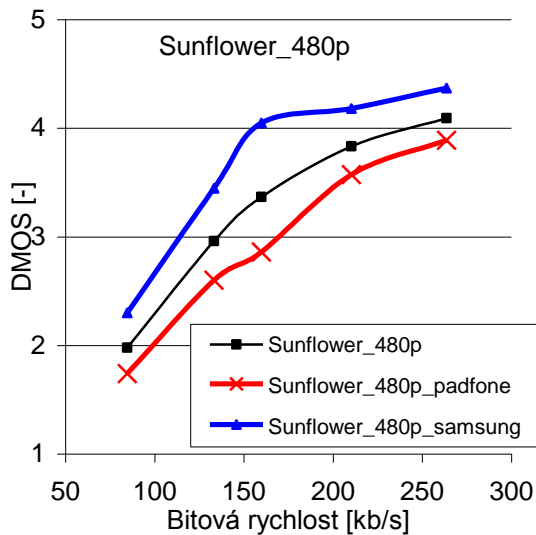
```

```
TargetBitrate           : 25000           # Rate control:
target bitrate, in bps
KeepHierarchicalBit     : 2               # Rate control:
0: equal bit allocation; 1: fixed ratio bit allocation; 2: adaptive
ratio bit allocation
LCULevelRateControl     : 1               # Rate control:
1: LCU level RC; 0: picture level RC
RCLCUSeparateModel      : 1               # Rate control:
use LCU level separate R-lambda model
InitialQP                : 0               # Rate control:
initial QP
RCForceIntraQP          : 0               # Rate control:
force intra QP to be equal to initial QP

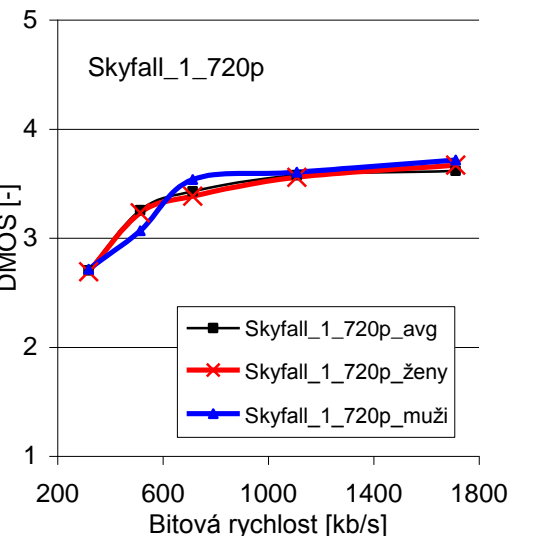
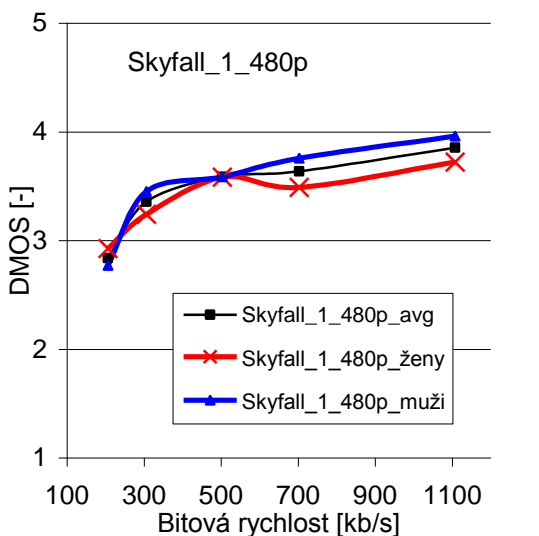
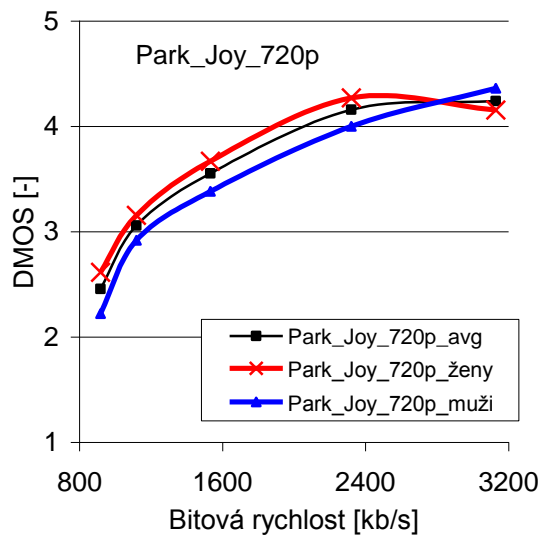
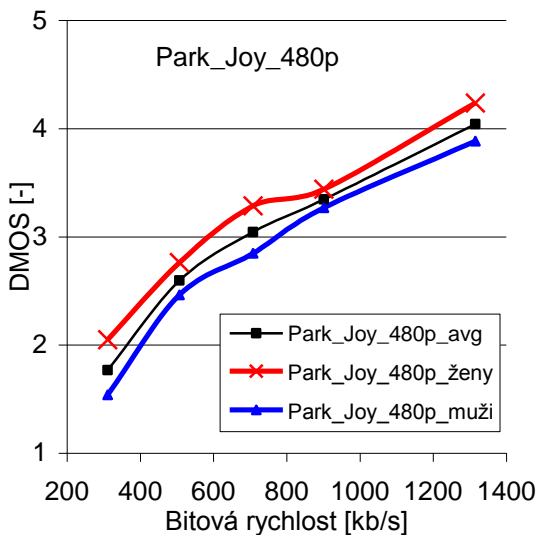
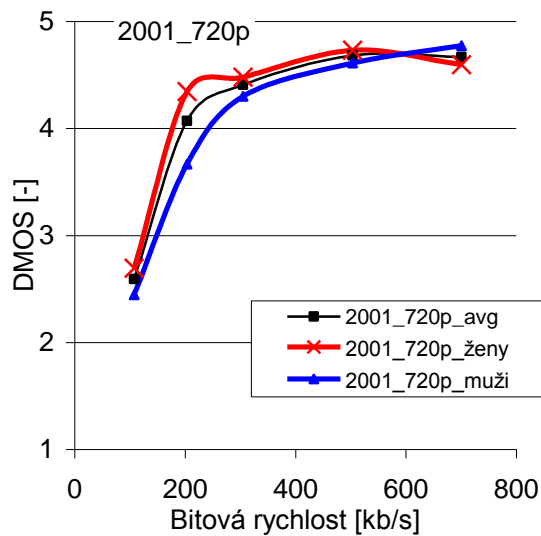
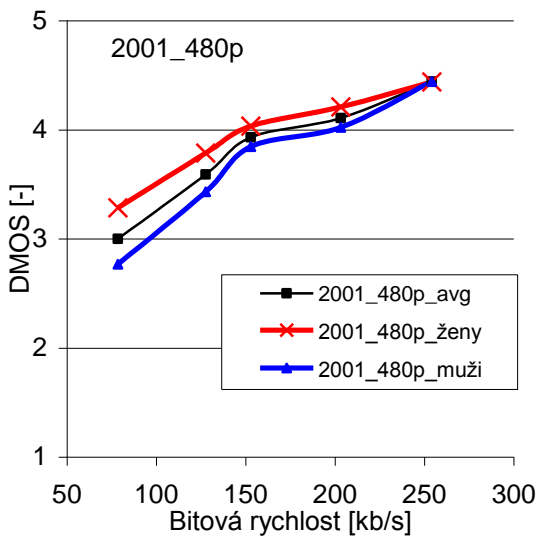
### DO NOT ADD ANYTHING BELOW THIS LINE ###
### DO NOT DELETE THE EMPTY LINE BELOW ###
```

C DOPLŇKOVÉ GRAFY

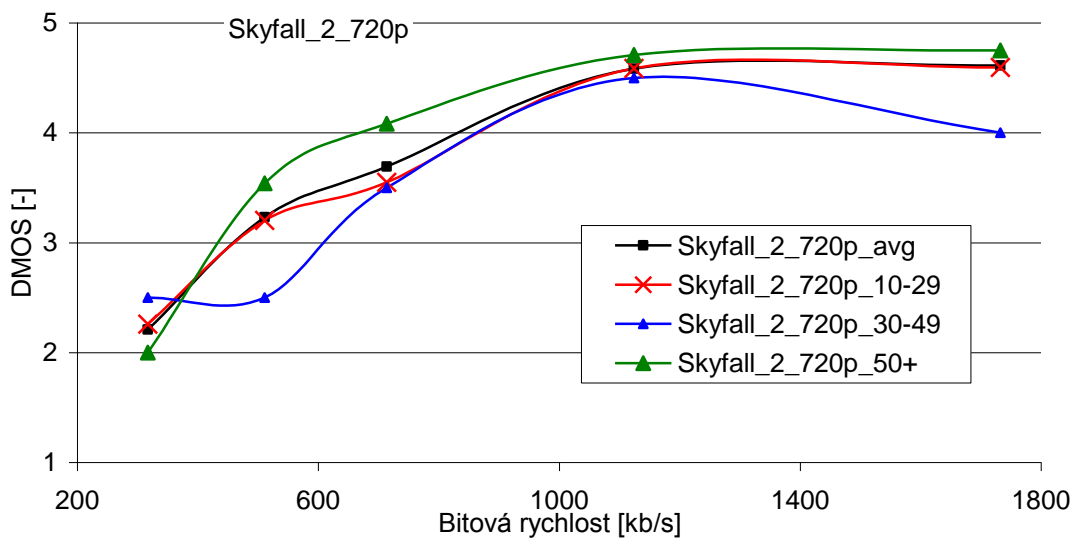
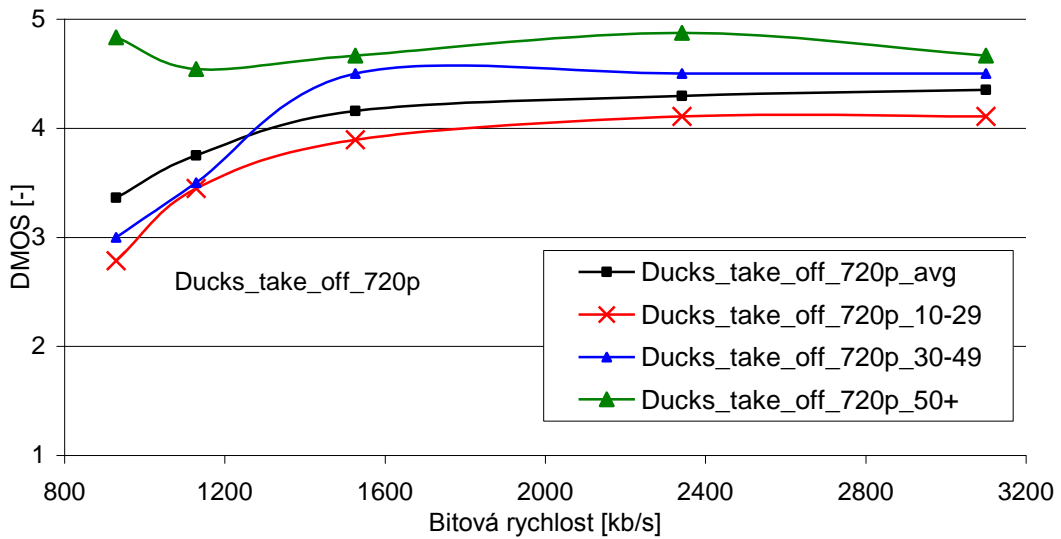
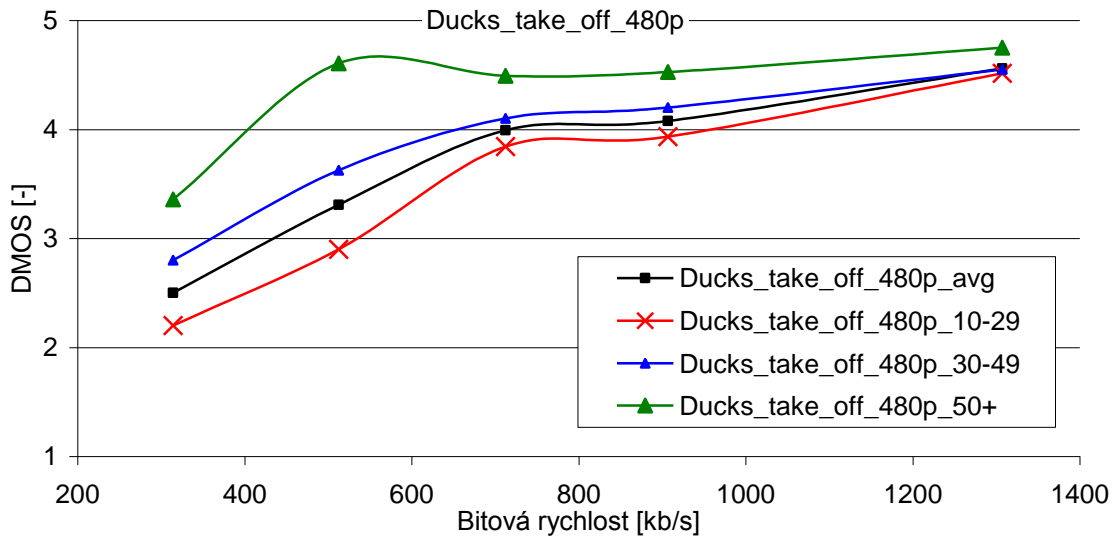
Rozdíly mezi hodnotami DMOS sekvenční 480p ze zařízení Samsung a Asus pro vybrané sekvence



Rozdíly mezi hodnocením kvality u mužů a žen pro vybrané sekvence



Rozdíly mezi hodnocením kvality jednotlivých věkových skupin pro vybrané sekvence



D DVD

- **H.265 sekvence** (H265_sekvence\...)
- **Zdrojové sekvence 1080p** (zdrojove_sekvence\...)
- **Výsledky subjektivních a objektivních testů** (vysledky\...)
- **Aplikace pro Android** (android_app\...)