

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

Fakulta elektrotechniky
a komunikačních technologií

BAKALÁŘSKÁ PRÁCE



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV RADIOELEKTRONIKY

DEPARTMENT OF RADIO ELECTRONICS

KVALITA MULTIMEDIÁLNÍHO OBSAHU U BRÝLÍ PRO VIRTUÁLNÍ REALITU

MULTIMEDIA QUALITY ASSESSMENT FOR HEADSET WITH VIRTUAL REALITY

BAKALÁŘSKÁ PRÁCE

BACHELOR'S THESIS

AUTOR PRÁCE

AUTHOR

Robert Majnuš

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Jan Kufa, Ph.D.

BRNO 2021



Bakalářská práce

bakalářský studijní program **Elektronika a komunikační technologie**

Ústav radioelektroniky

Student: Robert Majnuš

ID: 203285

Ročník: 3

Akademický rok: 2020/21

NÁZEV TÉMATU:

Kvalita multimediálního obsahu u brýlí pro virtuální realitu

POKYNY PRO VYPRACOVÁNÍ:

V teoretické části práce prostudujte principy virtuální a rozšířené reality. Seznamte se s různými typy zobrazovacích možností pro virtuální realitu. Navrhněte subjektivní testy hodnocení kvality multimediálního obsahu pro brýle s virtuální realitou. Vytvořte podklady pro navržené subjektivní testy. Test by se měl zabývat jak kvalitou a parametry vstupních dat, tak také parametry zobrazovače. Mezi zkoumané vlastnosti by mělo například patřit: vliv rozlišení, přenosové rychlosti a snímkové frekvence videa, různé principy získávání panoramatických záběrů a jasu zobrazovače. Pro kódování videosekvencí použijte různé dostupné kódovací algoritmy, jako například H.264, H.265, AV1 a jejich vhodná rozšíření pro virtuální realitu. Navrhněte subjektivní test pro měření kvality na co nejširším počtu respondentů.

V experimentální části práce proveďte vlastní test subjektivní kvality videa na respondentech. Dále proveďte analýzu a podrobný rozbor výsledků pomocí vámi navrženého subjektivního testu. Sledujte vliv rozdílných parametrů a typu respondentů na subjektivní hodnocení kvality videa. Určete vliv použitého multimediálního obsahu na hodnocení respondentů. Stanovte doporučení pro tvorbu multimediálního obsahu pro virtuální realitu.

DOPORUČENÁ LITERATURA:

- [1] Augmented reality and virtual reality. New York, NY: Springer Berlin Heidelberg, 2019. ISBN 978-3030062453.
- [2] Virtual reality toolbox for use with MATLAB and Simulink: user's manual : version 1.0. Prague: Humusoft, 2000. ISBN 80-7080-373-8.

Termín zadání: 8.2.2021

Termín odevzdání: 27.5.2021

Vedoucí práce: Ing. Jan Kufa, Ph.D.

prof. Ing. Tomáš Kratochvíl, Ph.D.
předseda rady studijního programu

UPOZORNĚNÍ:

Autor bakalářské práce nesmí při vytváření bakalářské 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

Tato bakalářská práce se zabývá vytvářením multimediálního obsahu pro virtuální realitu a navržení subjektivního testu pro hodnocení těchto sekvencí. V úvodní teoretické části jsou informace o virtuální a rozšířené realitě a několik typů brýlí pro virtuální realitu, které jsou dnes dostupné na trhu. Druhá kapitola se zaměřuje na metody měření kvality a hodnocení. Jedná se o metriky hodnocení subjektivní a objektivní. Dále jsou popsány jednotlivé typy kodeků, které nalezneme u většiny sekvencí na internetových portálech. Poté je kapitola o tvorbě obsahu pro virtuální realitu. Zejména panoramatické, 3D a 360° videa a fotografie. V praktické části této bakalářské práce je popsán mnou navržený postup úpravy sekvencí pro jednotlivé kodeky, návrh a použití subjektivní metody včetně vyhodnocení.

Klíčová slova

Virtuální realita, subjektivní metoda, 360° video, kodek, headset

Abstract

This bachelor thesis deals with creating multimedia content for virtual reality and designing a subjective test to evaluate these sequences. In the introductory theoretical part there is information about virtual and augmented reality and several types of virtual reality glasses that are available on the market today. The second chapter focuses on methods of quality measurement and evaluation. These are subjective and objective evaluation metrics. The following describes the different types of codecs that are found in most of the sequences on the Internet portals. Then there is a chapter on creating content for virtual reality. Especially panoramic, 3D and 360 ° videos and photos. In the practical part of this semester work is described by me the procedure of editing sequences for individual codecs and design and use of subjective method including evaluation.

Keywords

Virtual reality, subjective method, 360 ° video, codec, headset

Bibliografická citace:

MAJNUŠ, Robert. Kvalita multimediálního obsahu u brýlí pro virtuální realitu. Brno, 2021. Dostupné také z: <https://www.vutbr.cz/studenti/zav-prace/detail/131688>. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav radioelektroniky. Vedoucí práce Jan Kufa.

Prohlášení autora o původnosti díla

Jméno a příjmení studenta: *Robert Majnuš*

VUT ID studenta: *203285*

Typ práce: *Bakalářská práce*

Akademický rok: *2020/21*

Téma závěrečné práce: *Kvalita multimediálního obsahu u brýlí pro virtuální realitu*

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné 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 závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si 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.

V Brně dne: 27. května 2021

podpis autora

Poděkování

Děkuji vedoucímu semestrální práce Ing. Jan Kufa, Ph.D. za účinnou metodickou, pedagogickou a odbornou pomoc a další cenné rady při zpracování mé bakalářské práce.

V Brně dne:

.....
podpis autora

Obsah

Úvod.....	11
1 Virtuální realita.....	12
1.1 Historie virtuální reality.....	12
1.2 Virtuální realita.....	12
1.3 Rozšířená realita.....	13
1.4 Základní typy systémů pro VR.....	13
1.5 Typy brýlí pro VR.....	13
1.5.1 Oculus Go.....	14
1.5.2 Oculus Rift S.....	15
1.5.3 Oculus Quest.....	15
1.5.4 Oculus Quest 2.....	16
1.5.5 HTC VIVE.....	17
1.5.6 Samsung Gear.....	17
1.5.7 Acer Windows mixed.....	18
1.5.8 Primax 8K.....	19
2 Hodnocení kvality videa.....	20
2.1 Subjektivní.....	20
2.1.1 DSIS (Double-Stimulus Impairment Scale).....	20
2.1.2 DSCQS (The Double-Dtimulus Continuous Quality-Scale).....	21
2.1.3 SS (Single-Stimulus) methods.....	22
2.1.4 SSCQE (Single Stimulus Continuous Quality Evaluation).....	22
2.1.5 PC (Pair Comparison method).....	23
2.1.6 ACR (Absolute Category Rating).....	23
2.2 Objektivní.....	24
2.2.1 MSE (Mean Squared Error).....	24
2.2.2 PSNR (Peak Signal-to-Noise Ratio).....	24
2.2.3 SSIM (Structural Similarity Index Measure).....	25
2.2.4 VMAF (Video Multi-method Assessment Fusion).....	26
2.3 Další objektivní metody hodnocení kvality videa.....	27
2.3.1 NC (Normalized Correlation).....	27
2.3.2 AD (Average Difference).....	27
2.3.3 SC (Structural Content).....	27
2.3.4 MD (Maximum Difference).....	28
2.3.5 NAE (Normalized Absolute Error).....	28
3 Kódování videoobsahu.....	29
3.1 H.265/HEVC.....	29
3.1.1 Stromová struktura.....	29
3.1.2 Entropické kódování.....	30
3.1.3 Interpretace.....	31
3.1.4 Intrapredikce.....	31
3.1.5 Rekonstrukční filtr.....	31
3.2 VP9.....	32
3.2.1 Dělení snímků.....	32
3.2.2 Entropické kódování.....	33

3.2.3	Intrapredikce.....	33
3.2.4	Interpredikce.....	34
3.2.5	Transformační kódování.....	34
3.2.6	Rekonstrukční filtr.....	34
3.3	AVI.....	35
3.3.1	Dělení snímků.....	35
3.3.2	Entropické kódování.....	35
3.3.3	Intrapredikce.....	36
4.	Tvorba obsahu pro VR.....	37
4.1	Fotografie.....	37
4.1.1	3D fotografie.....	37
4.1.2	Panoramatická fotografie.....	37
4.1.3	360 stupňová fotografie.....	38
4.2	Videosekvence.....	39
4.2.1	3D video.....	39
4.2.2	360/180 stupňové video.....	40
4.3	Rozlišení videosekvencí.....	40
4.4	Snímkový kmitočet (FPS).....	40
5	Návrh subjektivního testu a vyhodnocení.....	41
5.1	Příprava a volba videí.....	41
5.2	Volba subjektivní metody a vyhodnocení.....	44
5.3	Průběh subjektivních testů.....	45
6	Výsledky hodnocení kvality pomocí vybraných objektivních a subjektivních metod.....	46
6.1	Výsledek objektivní kvality pomocí PSNR.....	46
6.2	Výsledek objektivní kvality pomocí SSIM.....	49
6.4	Subjektivní testování kvality videosekvencí.....	54
6.5	Doporučení pro tvorbu videoobsahu ve VR.....	57
7	Závěr.....	58
	Použité zdroje a literatura.....	59
	Seznam symbolů, veličin a zkratk.....	63
	Seznam PŘÍLOH.....	65
	A Referenční a upravené videosekvence.....	65
	B Varianty a hodnocení videosekvencí respondenty.....	65

Seznam obrázků

1.1: Oculus Go s ovládačem [4]	14
1.2: Oculus Rift S s ovladači [5]	15
1.3: Oculus Quest s ovladači [6]	16
1.4: Oculus Quest 2 s ovladači [38]	16
1.4: HTC VIVE s ovladači [3]	17
1.5: Samsung Gear s ovládačem [3]	18
1.6: Acer Windows Mixed [3]	18
1.7: Primax 8K [33]	19
2.1: Struktura prezentace testované sekvence pro DSIS [1..0]	21
2.2: Struktura prezentace testované sekvence pro DSCQS [10]	22
2.3: Struktura prezentace testované sekvence pro PC [24]	23
2.4 Struktura prezentace testované sekvence pro ACR[24]	23
3.1: Velikosti pro PB [12]	30
3.2: Rozvržení CB a TB v CTB [12]	30
3.3: Rozložení bloků a možnosti jejich dělení [13]	33
3.4:Možnosti režimů intrapredikce [13]	33
3.5: Dělení bloků pro AV1. převzato z [14]	35
4.1: Ukázka skládání panoramaticke fotografie [16]	38
4.1: Ukázka skládání 360° fotografie za pomoci mobilní aplikace [19]	39
5.1 Obsah videosekvence City	42
5.2: Obsah videosekvence Forest	42
5.3:Obsah videosekvence Street	42
5.4: Průběh mnou navrženého měření pomoci metody ACR-HR [24]	45
6.1:Závislost PSNR na bitové rychlosti CITY videosekvence	47
6.2:Závislost PSNR na bitové rychlosti FOREST videosekvence	47
6.3:Závislost PSNR na bitové rychlosti STREET videosekvence	48
6.4:Závislost SSIM na bitové rychlosti CITY videosekvence	49
6.5:Závislost SSIM na bitové rychlosti FOREST videosekvence	50
6.6:Závislost SSIM na bitové rychlosti STREET videosekvence	50
6.7:Závislost VMAF score na bitové rychlosti CITY videosekvence	52
6.8:Závislost VMAF score na bitové rychlosti FOREST videosekvence	52
6.9:Závislost VMAF score na bitové rychlosti STREET videosekvence.....	53
6.10: Výsledky MOS pro CITY videosekvenci	55
6.11: Výsledky MOS pro FOREST videosekvenci.....	55
6.12: Výsledky MOS pro STREET videosekvenci	56

Seznam tabulek

5.1 Podrobnosti o použitých videosekvencích.....	41
5.2: Tabulka pro hodnocení jednotlivých video sekvencí.....	44
6.1 Hodnoty výsledků objektivní metody PSNR pro všechny videosekvence.....	48
6.2 Hodnoty výsledků objektivní metody SSIM pro všechny videosekvence.....	51
6.3 Hodnoty výsledků objektivní metody VMAF pro všechny videosekvence.....	53

Úvod

Cílem této bakalářské práce je vytvořit vhodný multimediální obsah pro virtuální realitu s využitím videí s vysokým rozlišením a programu pro jejich úpravu za pomoci změny datového toku a využití zdarma přístupných kodeků, které jsou dnes nejčastěji používané a nebo ty, které se teprve chystají pro budoucí využití na internetových portálech.

Podmětem pro zpracování této bakalářské práce na toto téma je spousta. V dnešní době je spousta možností, kde lze přístroj na virtuální realitu vyzkoušet. Za posledních pár let bylo vytvořeno spousta výrobků pro uživatelský zážitek z virtuální reality s co největším pohlcením. Dalo by se říci, že každá větší firma, která se zabývá vývojem displejů nebo mobilních telefonů se chce na této technologii nějakým způsobem podílet. Využití je celá škála od konstrukčních návrhů, lékařství, armády, sportu, ale i pro využití běžného uživatele na relax. Dnes už stačí jen nasadit headset na hlavu s virtuální prohlídkou za polárním kruhem a užít si ničím nerušený zážitek pozorováním krásných detailů. Z toho důvodu je důležité kvalitní návrh prostředí, aby byl pro uživatele co nejkvalitnější a ničím nerušený.

Dnes už ani pro většinu uživatelů není problém sehnat kvalitní obsah, který se může přibližovat k rozlišení 8K. Virtuální realita není sice ještě na takové úrovni přehrávání, ale v případě, že jednotlivé displeje o velikosti mobilního telefonu s velmi malou vzdáleností od oka obsahují 4K displeje, je zapotřebí, aby daný obsah byl co nejlépe vytvořen a oko si nevšimlo nedokonalostí.

1 VIRTUÁLNÍ REALITA

1.1 Historie virtuální reality

Pojem virtuální realita (VR) pochází už ze 40. let minulého století, kde tento pojem poprvé použil francouzský avantgardní spisovatel Antonin Artaud ve své knize, kde popsal iluzorní charakter postav a objektů. Z technického hlediska první se první zmínka datuje k roku 1838, kdy vědec Charles Wheatstone ukázal, že lidský mozek je schopen vnímat dvojrozměrné obrazy z každého oka jako jeden trojrozměrný. Načež sestrojil stereoskop, ve kterém promítal fotografie, které nabyly dojmu hloubky a ponoření do dané fotografie. V roce 1929 byl sestroyen vynálezcem Edwinem Albertem Linkem první letecký simulátor, který byl používán ke komerční účelům. V polovině padesátých let vyvinul kameraman Morton Heiling první VR přístroj, který nesl název „sensorama“, která působila na všechny smysly, nejen zrak a sluch.

V roce 1960 vytvořil „Telesphere Mask“, která je prvním představitelem brýlí s displejem. O rok později vyvinula společnost Philco Corporation přístroj s názvem „Headsigt“, který obsahoval pro každé oko jednu obrazovku s magnetickým systémem sledování pohybu. Tato technologie byla vytvořena primárně pro armádní účely. Roku 1968 vytvořil počítačový vědec Ivan Sutherland přístroj, který byl napojen na počítač. Jednalo se o velmi těžké zařízení, tudíž bylo nutné jej zavěsit na strop. V roce 1969 přichází poprvé počítačový odborník Myron Krueger s umělou realitou. Ale až v roce 1987 poprvé přichází s pojmenováním VR Jaron Lanier, který založil vizuální programovací laboratoř, kde pomocí svého výzkumu vyvinul vybavení pro VR. Byli první společností, která začala prodávat brýle i rukavice pro VR. První pokusy použít VR do herních automatů přišla v roce 1991 firmou Vurrality Group, kde hráči pomocí brýlí mohli hrát reálnými stereoskopickými 3D obrázky. Roku 1995 přišla firma Nintendo v úbec první přenosnou 3D herní konzoli se skutečnou 3D zobrazovací grafikou. Výroba byla nakonec ukončena z důvodu nedostatečné podpory softwaru. Systém fungoval jen ve 2 barvách a to černá a červená. Dokonce i filmografie přispěla k většímu zájmu společnosti o VR a to filmem Matrix z roku 1999. Začátkem 21. století vývoj VR výrazně vzrostl z důvodu zvyšování výkonů počítačové i mobilní technologie, kdy roku 2016 byly představeny brýle pro VR Oculus [1].

1.2 Virtuální realita

Virtuální realita je založena na interaktivní zpětné vazbě (jako jsou vizualizace, zvuk a haptika). Proto při počátku samého vývoje byl založen na senzorech polohy a orientace. Nyní máme k dispozici již cenově dostupné verze tohoto zařízení, což způsobilo velkou inovaci v tomto odvětví.

Pro výpočet obrazu polohy pozorovatele musíme nejprve znát postavení a orientaci úhlu pohledu. První VR headsety z roku 1968 byly vybaveny jednoduchým řešením pomocí rotačního systému pro detekci pohybů hlavy. Později tento systém byl nahrazen elektromagnetickou technologií, která byla v dané době převládající technologií. Dnes je nejčastěji rozšířena optická technologie [2].

1.3 Rozšířená realita

Rozšířená realita (AR) slouží jako prostředek k obohacení vnímání prostředí pomocí digitální informace týkající se tohoto daného prostředí. Jedná se o informace, které jsou zejména vizuální, občas se setkáme i se sluchovými nebo haptickými. Nejčastěji se používá pro vizualizaci AR syntetické obrazy pomocí brýlí, náhlavních souprav, video projektorů nebo i mobilních telefonů. První tři zařízení slouží na principu superponování informace o přirozeném vidění. Mobilní telefony nabízejí pouze vzdálené prohlížení.

Z hlediska použití se AR dá použít například při konstruování nebo při stavbách domů, kde díky vytvořené simulaci můžeme procházet pozemek s tím, že již vidíme vizualizaci budoucí budovy. Dále se dá použít v zařízeních GPS, kde nám ukazuje značení a uživatelé stačí pouze pohledy na danou značku. Použití je celá škála, své odvětví si najde v automobilovém, ale i herním průmyslu.

Z hlediska porovnání AR a VR můžeme říct, že VR slouží k plnému pohlcení do daného prostředí jako je například vychutnat si pohyb s tučňáky na Antarktidě. AR pouze doplňuje náš vjem prostředí, nejlepší využití je při projektech, kde si díky tomuto zobrazení můžeme udělat představu o realizaci dané věci [2].

1.4 Základní typy systémů pro VR

Pro sledování lze použít vysokofrekvenční infračervené kamery (až 250Hz). Kamery identifikují sadu markerů, které jsou umístěny v prostoru, od kterých se odrážejí infračervené paprsky. Tyto značky jsou z pravidla umístěny na části lidského těla, nejčastěji na ruce, ramena, lokty, kolena atd. nebo i na předměty, které chceme v prostoru sledovat.

Tento systém je založen na optické technologii. Má velice přesné sledování, robustnost a nízkou latenci. Povrch, který může být zachycen dosahuje velikosti až 100m². Nevýhodou je dlouhá instalace – kamery musí být velmi stabilně umístěny a připojeny k počítači, vyžadují kalibraci a musí být zajištěno, že každý objekt musí být viditelný pro několik kamer.

Dalším systémem, který se používá je založen na podobném principu jako systém výše. Ovšem vychází z používání kamery bez umístění markerů na uživateli nebo věci. Jejich přístup je založen na extrahování siluet z pořízeného snímku. Pomocí použití více kamer a jejich kombinací je možné zobrazit základní 3D model lidského těla. Může tedy sledovat pohyby lidského těla, aniž by na něm musely být umístěny receptory. Na druhou stranu se sledování musí odehrávat v omezeném prostoru s konstantním pozadím (nejčastěji zelené) za účelem získání nejpřesnějšího modelu. Přesnost a latence je poněkud horší než u prvního řešení [2].

1.5 Typy brýlí pro VR

Headsety pro VR můžeme rozdělit do tří základních kategorií, jako jsou brýle s vlastním procesorem (all-in-one) a tedy přenosné. Pevně připojeny k počítači a poslední s možností vložení smartphonu nebo tabletu, které jsou rovněž mobilní.

Brýle all-in-one znamená, že obsahují všechno důležité komponenty pro zobrazení VR reality bez přidání jakéhokoliv externího zařízení. Velkou výhodou je mobilita a pohyb bez

jakéhokoliv rušení, které může způsobit například připojené kabely. Nevýhodou je baterie, obvykle 2600mAh (3 hodiny bez nabíjení).

Brýle s připojením na PC využívají výkon počítače. Uživatel tedy potřebuje kabelové připojení. Zároveň jeho kvalita je omezena tímto výkonem. Od toho se dále odvíjí i cena, kde uživatel potřebuje vyšší výkon počítače a zároveň musí investovat do samotného headsetu. Na druhou stranu tato varianta nabízí nejlepší zážitek z těchto variant. Kvalita je velice dobrá a nabízí příjemné ponoření do VR.

Verze pro smartphony je nejlevnější, ale dlouhodobé užívání může způsobovat bolesti hlavy a závratě. Proto není pro uživatele příliš vhodná. Jedná se z pravidla o krabici, která je vyrobena z kartonu nebo plastu, kterou si uživatel nasadí na hlavu. Je vybavena dvěma čočkami, které přímo směřují na displej telefonu, ten je umístěn v konstrukci. Běžnému uživateli je tedy možné dopřát koncept VR, ale není to takový kvalitní zážitek jako u předchozích dvou variant [3].

1.5.1 Oculus Go

Společnost Oculus díky svému vývoji v odvětví virtuální reality způsobila velký zájem o tvorbu vlastního headsetu. Hlavně díky své cenové dostupnosti, která se u některých brýlí cenově pohybuje okolo 7 000 korun. Což pro veřejnost vytvořil cenově dostupnou technologii, která byla velmi dobře přijata a tím způsobil velký pokrok této technologie.

Zmíním-li se o brýlích Oculus Go (obr.1.1) můžeme zde najít dva LCD displeje s rozlišením 1280 x 1440 pro každé oko, frekvence 72 Hz a zorný úhel 100°. Nabízí již kvalitní zážitek a vtažení do virtuální reality a to bez použité externího připojení. Využívá technologii 3DOF, která umožňuje vnímat okolí. Nevýhodou ovšem je, že se uživatel nemůže plynule pohybovat. Hardware, který může nabídnout, je CPU Qualcomm Snapdragon 821, 3GB vnitřní paměti a také grafiku Adreno 530, která je součástí procesoru. Můžeme na HMD najít 3.5mm jack a micro USB pro připojení k počítači a nabíjení. Na trhu můžeme najít dvě varianty, které obsahují 32GB a 64GB úložného prostoru. Obsahuje také ovladač pro snadnou volbu v menu, který je připojen s brýlemi přes Bluetooth. Dnes je tato varianta jedna z levnějších a zaplatili bychom za ní přibližně něco kolem 6 500 korun [4].



Obr. 1.1: Oculus Go s ovládačem [4]

1.5.2 Oculus Rift S

Byl vydán v květnu roku 2019 (obr.1.2) a nabízí technologii displeje Fast-switch LCD a to s rozlišením 1280 x 1440 pro každé oko s frekvencí 80Hz. Oproti předchozím brýlím disponuje již technologií 6DOF díky čemuž umožňuje pohyb v prostoru a tedy větší uvěřitelnost a reálnost. Všechno je to způsobeno 5 kamerami, které jsou na brýlích v oblasti očí umístěny zvenčí (2 přední, 2 boční, 1 horní). Ovšem tyto brýle musí být připojeny k počítači pro jejich použití a zde vzniká problém s mobilitou a omezení kabelem. Výrobce je doporučen standart GeForce GTX 1060 a procesor alespoň i5-4590. Rozhraní, která najdeme na brýlích, jsou mini display port a jedno USB a jack 3.5mm pro připojení audia. Cena se pohybuje kolem 13 000 a v balení nalezneme také 2. generaci ovladačů pro levou a pravou ruku [5].



Obr. 1.2: Oculus Rift S s ovládači [5]

1.5.3 Oculus Quest

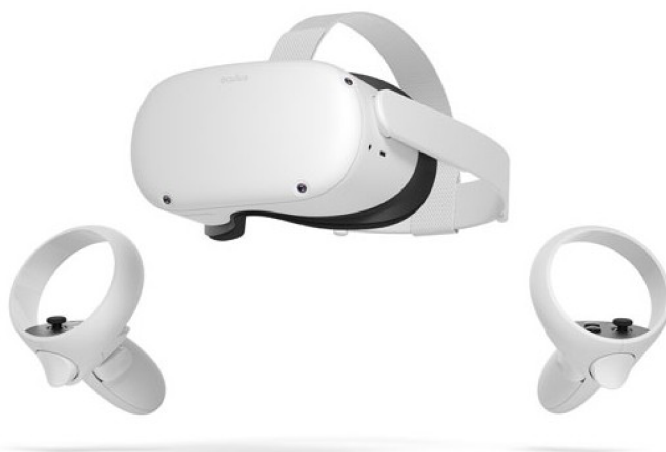
Dalším headsetem, který byl vyvinut společností Oculus je Oculus Quest. Disponuje OLED displejem s rozlišením 1440x1600 pro každé oko, obnovovací frekvencí 72Hz a zorné pole 95°. Jedná se o přenosný HMD tedy all-in-one (obr.1.3), se kterým můžeme fungovat bez nabíjení až 3 hodiny. Dále v něm najdeme procesor Qualcomm Snapdragon 835. Disponuje technologií stejnou jako Rift S neboli 6DOF a 4 kamerami. Samotná operační paměť v headsetu je 4GB, pro uložení najdeme paměti 64GB a 128GB. Celková cena s nižší ukládací paměti je okolo 14 500 korun. V balení jsou opět dva ovladače 2. generace, které jsou obdobné jak u předchozích brýlí [3, 6].



Obr. 1.3: Oculus Quest s ovládači [6]

1.5.4 Oculus Quest 2

Prozatím poslední headset vyvinut společností Oculus ve spolupráci s firmou Facebook. Nabízí vyšší rozlišení než předchozí model 1832x1920 pro každé oko s displejem Single Fast-Switch LCD. Dále pracuje se stejnou frekvencí 72Hz a snímání 6DOF. Změnu nalezneme v novém výkonnějším procesoru Qualcomm Snapdragon XR2, který disponuje třemi jádry. Operační paměť je také vyšší a to 6GB. Obsahuje také dva bezdrátové ovladače, stejné jako u Oculus Quest. Výhodou je nižší cena oproti předchozímu modelu a také provedení uložení 64 nebo 256GB [40].



Obr. 1.4: Oculus Quest 2 s ovládači [38]

1.5.5 HTC VIVE

Výrobce HTC ve spolupráci s Valve vytvořil tento typ brýlí (obr.1.4). Technologie pro sledování polohy a rotace brýlí využívá dvou senzorů, které používá technologii Lighthouse. Dokáže zaznamenávat polohu oblasti 4,6 x 4,6 metrů. Tuto technologii používají také ovladače, které mají taktéž stejné pole snímání. Dokáže zaznamenat pozici nehybného HMD (Head-mounted display) v kouli o rozměru poloměru 0,3mm. Tato hodnota odpovídá hranici vnímání člověkem.

Disponuje rozlišením 2160 x 1200 (1080 x 1200 pro každé oko) s frekvencí displeje 90Hz. Zorné pole je výrobcem definováno na 110°. Podle praktických testů odpovídá zorné pole pro horizontální 100° a pro vertikální 110°. Disponují konektorem pro připojení sluchátek. V roce 2019 se cena pohybovala kolem 21 500 korun. Pro použití je vyžadováno připojení k počítači. Díky tomu je potřeba připojovací kabel, který značně omezuje pohyb a také je zde nebezpečí zamotání uživatele do kabelu. Pro vykreslení virtuální reality pomocí těchto brýlí je zapotřebí výkonný počítač. Z toho důvodu jsou kladeny požadavky na hardware. Výrobcem jsou stanoveny doporučené prvky na procesor Intel Core i5-4590 a grafickou kartu NVIDIA GeForce GTX 1060 [7].



Obr. 1.4: HTC VIVE s ovládači [3]

1.5.6 Samsung Gear

Tento výrobek je vyvinut společností Samsung ve spolupráci s Oculus. Nejedná se v pravém slova smyslu o headset pro VR (obr.1.5). Nenajdeme na něm žádný displej. Tuto skutečnost nahrazují nové modely telefonů Samsung, zejména od modelů Galaxy S6 a vyšší. U takového telefonu už bývají kvalitní displeje pro virtuální realitu s rozlišením 2560 x 1440 pixelů, které zaručují standart kvality. To znamená, že zajišťují stabilní počet zobrazovaných snímků za vteřinu pro VR. Pochopitelně nemůžeme čekat stejnou kvalitu jako u brýlí, které jsou připojeny k počítači. Další nevýhodou je, že postrádá senzory na sledování pozic hlavy. Výhodou naopak je absence drátového připojení.

Obvykle se cena telefonů pohybuje v dnešní době od 4000 korun. Což je pro mnoho uživatelů již dostupná cena. Samotný MHD se na trhu pohybuje i s ovladačem za 4000 korun [7].



Obr. 1.5: Samsung Gear s ovládačem [3]

1.5.7 Acer Windows mixed

Firma Acer představila svůj headset pod jménem OJO 500 (obr.1.6). Díky vestavěnému kolečku je možnost měnit interpupilární vzdálenost a také díky mobilní aplikaci lze použít nástroje pro změnu vzdálenosti pro čistší obraz. Jedná se o nástroje Acer IPD Meter a software Acer IPD Indicator. Nabízí LCD displej s rozlišením 2880 x 1440 se zorným úhlem 100°. S pohledem na výbavu zde můžeme najít integrovaný audiosystém se zvukovody. V této části bychom mohli najít vestavěné reproduktory ve zvukovodu a vestavěného pole mikrofonů. Se zaměřením na vstupy narazíme na jeden HDMI 2.0 vstup a jedno připojení USB 3.0. Celý systém je připojen k počítači, tedy opět zůstává problém s kabelem a omezenou pohyblivostí. Výrobce je doporučena GPU alespoň GeForce GTX 1050. Cenu tohoto MHD bychom na českém trhu našli kolem 13 000 korun [8].



Obr. 1.6: Acer Windows Mixed [3]

1.5.8 Primax 8K

Tento MHD (obr.1.7) nabízí rozlišení až 3840 x 2160 pro každé oko. Jedná se o dva 4K CLPL displeje, které zobrazují 8K. Na takové rozlišení je potřeba displej o rozlišení 7680 x 4320, tedy se jedná pouze o marketingový tah společnosti Primax. Ovšem díky použití 4K displejů odpadá problém s tzv. ScreenDoorEffectem. Tento efekt znamená, že při malé vzdálenosti od oka jdou stále vidět mezery mezi pixely. Nabízí možnosti zobrazovacího pole 200°, což umožňuje velice kvalitní vtáhnutí do VR prostředí. Pracuje na frekvenci 75 nebo 90Hz. Díky možnosti připojení Brainwarp (asynchronní projekce) je možné vytáhnout frekvenci až na dvojnásobek těchto hodnot. Opět je vyžadováno připojení k počítači. Dle výrobce je doporučeno použití grafické karty alespoň GeForce GTX 980. Nabízí také řadu vstupů jako USB 3.0 a DP1.4. Obsahuje dva 3,5mm jacky na připojení audia. V tuto dobu se HMD na trhu vyskytuje kolem 27 000 korun [9].



Obr. 1.7: Primax 8K [33]

2 HODNOCENÍ KVALITY VIDEO

Pro hodnocení kvality multimediálního obsahu se používají dvě základní metody hodnocení, jedná se o metodu subjektivní a objektivní.

2.1 Subjektivní

Vychází z hodnocení kvality videa testovanou osobou, která podle vlastního uvážení rozhodne o závěru hodnocení. Tedy hodnocení je u každé osoby jiné.

Tato metoda vychází z doporučení ITU-R BT.500. Princip hodnocení vychází z výběru skupiny lidí (pozorovatelů), kteří podle svého zrakového vjemu hodnotí videa, která jim jsou promítána s různou změnou kvality. Vyhodnocení se poté stanovuje z definované stupnice z vybraného typu testů. Pro konečné hodnocení se provede průměr ze všech hodnot měření pro každé video.

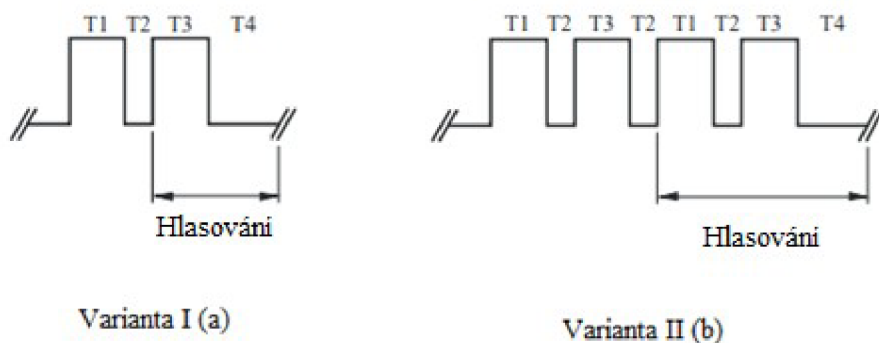
Výhodou takového měření je, že dokáže získat skutečnou kvalitu obrazu, jakou člověk vnímá svými smysly. Díky testování velkého množství lidí s větší délkou videa se zde objevuje vysoká časová a finanční náročnost hodnocení [10].

2.1.1 DSIS (Double-Stimulus Impairment Scale)

U této metody je testujícím osobám promítáno původní video a poté video zkreslené (změněno testovacím systémem). Pozorovatel tak může provést hodnocení stupně poškození kvality videa. Tato relace trvá až půl hodiny [10].

Struktura systému se dělí na dvě varianty prezentace:

- I. Referenční testovací obraz sekvence jsou uvedeny pouze jednou, jak je znázorněno na obr. 2.1. (a)
- II. Referenční testovací obraz dvakrát jako (b). Varianta II. Je tedy časově náročnější.



Obr. 2.1: Struktura prezentace testované sekvence pro DSIS [10]

T1 = 10s	Referenční obraz
T2 = 3s	Šedá úroveň
T3 = 10s	Testovaný obraz
T4 = 5 – 11s	Šedá úroveň

Pro hodnocení je stanovená stupnice 1 až 5, kde číslo 1 je nejvíce poškozený obraz [10].

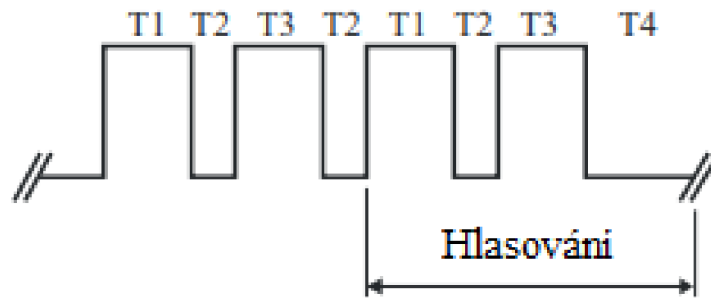
5	nepostřehnutelné
4	znatelné, ale ne nepříjemné
3	mírně otravné
2	nepříjemné
1	velmi nepříjemné

2.1.2 DSCQS (The Double-Dstimulus Continuous Quality-Scale)

Tato metoda je založena na promítání dvou stejných videí z čehož jedno má nepatrný rozdíl parametrů oproti druhému. Díky tomu, že změna je méně patrná musí mít pozorovatel také více času na srovnání obou videí. Celá relace opět trvá až půl hodiny. Na konci je poté zprůměrovaná hodnota pro každou testovací podmínku [10].

Můžeme tuto metodu měřit dvěma variantami:

- I. Pozorovatel může volně přepínat mezi videi, aby docílil vnímání změny kvality s každou změnou.
- II. Videi jsou spouštěna na dvou zařízeních současně a ve stejném čase. Po skončení je vyzván k hodnocení.



Obr. 2.2: Struktura prezentace testované sekvence pro DSCQS [10]

T1 = 10s	Testovací sekvence A
T2 = 3s	Šedá úroveň
T3 = 10s	Testovací sekvence B
T4 = 5 – 11s	Šedá úroveň

Normalizované měřítko stupnice hodnocení je v rozmezí 0 až 100.

2.1.3 SS (Single-Stimulus) methods

U této metody se nepoužívá původní, ještě neupravené, video. Pro testovací prezentaci se používá podobný cyklus jako u již zmiňované DSIS. Testovací sekvence jsou prezentovány třikrát, z nichž každá zahrnuje všechny sekvence. Existují dvě formy metody, v kontinuálním a numerickém měřítku. Díky tomu se dá tato metoda využít na vybrané typy poruch a s tím spojeny kombinace a vliv na hodnotící osobu [10].

2.1.4 SSCQE (Single Stimulus Continuous Quality Evaluation)

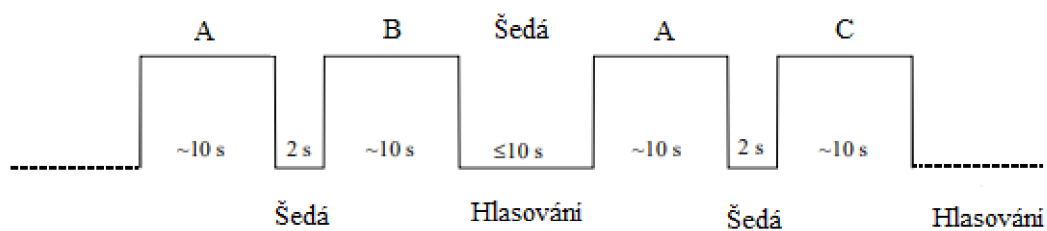
Průběh hodnocení opět nepoužívá původní, tedy referenční video. Pozorovateli je spouštěna pouze testovací sekvence, jejíž kvalita je kontinuálně měněna. Měření se provádí pomocí voliče v rozmezí 0 až 100. Podmínky, které musí být zachovány, jsou:

- volič odečítá hodnotu 2x za sekundu
- délka lineárního posunu by měla být 10cm

Každá testovací sekvence musí mít délku alespoň pět minut [10].

2.1.5 PC (Pair Comparison method)

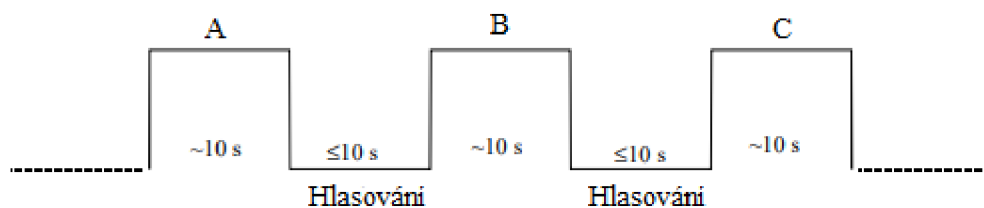
Pro metodu Pair comparios se nepoužívá referenční snímek. Pouze se porovnávají jednotlivé dvojice upravených sekvencí mezi sebou. V případě, že máme tři videosekvence X,Y,Z, které mají různou úpravu, tak se snažíme vytvořit všechny jejich kombinace. Tedy XY,XZ,YX,YZ,ZX,ZY. Po každých dvou po sobě jdoucích videí se provádí hodnocení, kde hodnotitel vybere podle něj lepší sekvenci. Průběh přehrávání vypadá takto: nejprve se přehraje video délky okolo 10 sekund, následně krátká pauza vyplněna šedým oknem s délkou dvou vteřin a následně druhá sekvence o obdobné délce 10 sekund jako první. Následně je delší pauza řádově desetisekundová pro zhodnocení videa. Takto se cyklus opakuje dokud se neprojdou všechny kombinace [24].



Obr. 2.3: Struktura prezentace testované sekvence pro PC [24]

2.1.6 ACR (Absolute Category Rating)

Metoda je postavena na prezentování testovacích sekvencí, jedna po druhé a následně jsou hodnoceny nezávisle na kategorii stupnice. Po každé prezentaci je hodnotitel požádán o vyhodnocení kvality sekvence [24].



Obr.2.4 Struktura prezentace testované sekvence pro ACR [24]

2.2 Objektivní

Objektivní hodnocení znamená, že daný obraz nebo sekvence je hodnocena nestranně. Pro každé měření musí být vždy stejný výsledek. Používá se z důvodu nahrazení zdlouhavého a drahého subjektivního hodnocení a zároveň, aby se přiblížilo ke stejnému výsledku.

Při použití těchto metod se vychází z hodnocení vzájemných rozdílů mezi originálním a testovaným videem. K tomuto vyhodnocení se používají matematické výpočty založené na maticovém porovnávání. Velkou výhodou je vyšší korelace výsledků oproti subjektivním metodám.

2.2.1 MSE (Mean Squared Error)

Jedná se o nejjednodušší a nejpoužívanější metodu objektivního měření kvality obrazu, která se využívá již několik desítek let. Její jednoduchost spořívá v tom, že každá sekvence může být vyhodnocena nezávisle, přičemž výpočty nezaberou moc času. Abychom mohli provést výpočet je zapotřebí originální obraz x a komprimovaný obraz \hat{x} , které mají stejnou šířku N_1 a výšku N_2 , poté můžeme dosadit do vzorce [11].

$$MSE = \frac{1}{N_1 N_2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} (x[n_1, n_2] - \hat{x}[n_1, n_2])^2 \quad (1.1)$$

Výsledná hodnota je poté rovna střední kvadratické chybě. V případě, že jsou obě zkoumaná videa shodná, výsledek se bude rovnat nule. Při zkoumání změn barevného obrazu RGB, provedeme výpočet všech tří barev a tuto hodnotu následně podělíme třemi [11].

2.2.2 PSNR (Peak Signal-to-Noise Ratio)

Vychází z předchozí metody MSE. Pouze výsledek je upraven do reálnějšího měřítka. Základním principem je vyjádření poměru s možnou maximální energií signálu a energií šumu. Z důvodu širokého dynamického spektra se u této metody výsledek zobrazuje v logaritmickém měřítku. Výšší hodnota PSNR značí kvalitnější obraz, pro 24bitový obraz je to hodnota mezi 30 až 50dB. V případě že se jedná o stejný obraz, není možné PSNR vypočítat, neboť MSE je v tomto případě nulové. Používá se výpočet [11].

$$PSNR = 10 \log_{10} \left(\frac{(2^B - 1)^2}{MSE} \right) \quad (1.2)$$

Kde písmeno B značí dynamický rozsah, v případě osmi bitů na kanál 255 [11].

2.2.3 SSIM (Structural Similarity Index Measure)

Předchozí metody nepočítají s psychovizuálním modelem lidského vnímání. Díky tomu vznikla metoda SSIM. Spočívá ve snaze vytvoření indexu, který je nejbližší lidskému vidění. Základním předpokladem pro tuto metodu je přirozený obraz vysoce strukturovaný a tedy zde nastává vysoká závislost pixelů, které přenášejí důležité informace o samotné struktuře objektů na daném snímku. Další nutnost je oddělit tyto strukturní informace od jasu, protože obě tyto proměnné jsou na sobě nezávislé. Musíme tedy rozdělit měření do tří srovnání: jas, kontrast a struktura.

Z počátku musíme porovnat hodnotu jasu obou signálů. Uvažujeme diskrétní signály, z toho poté můžeme vypočítat pomocí střední intenzity [11].

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i. \quad (1.3)$$

Pro porovnání funkce jasu $l(x,y)$ vyplyne funkce μ_x a μ_y .

Nyní díky tomuto výpočtu můžeme funkci jasu ze signálu odstranit. Pro odhadnutí kontrastu se využívá směrodatná odchylka (kvadratický průměr odchylky od aritmetického průměru). Na tento odhad použijeme vzorec [11]:

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{\frac{1}{2}} \quad (1.4)$$

Opět pro porovnání kontrastu $c(x,y)$ je použito porovnání σ_x a σ_y . Díky normalizaci svou směrodatnou odchylkou, může být provedeno porovnání struktury obrazů. Struktura $s(x,y)$ je na základě normalizovaných realizována.

$$\frac{(x-\mu)}{\sigma} \text{ a } \frac{(y-\mu)}{\sigma}$$

Nyní již máme všechny tři potřebné složky, které dále kombinujeme [11].

$$S(\mathbf{x}, \mathbf{y}) = f(l(\mathbf{x}, \mathbf{y}), c(\mathbf{x}, \mathbf{y}), s(\mathbf{x}, \mathbf{y})). \quad (1.5)$$

Pro porovnání jasu vycházíme z definovaného vzorce. [11]

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (1.6)$$

Konstanta C_1 je přidaná hodnota, která slouží pro zamezení nestability, neboť $\mu_x^2 + \mu_y^2$ se blíží nule [11].

$$C_1 = (K_1L)^2 \quad (1.7)$$

Kde hodnota K_1 je rovna hodnotě 0,01 a L je dynamický rozsah, který je obvykle pro 8bitů 255 kanálů.

Obdobně získáme i hodnotu pro porovnání kontrastu [11].

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (1.8)$$

Pro C_2 platí $C_2 = (K_2L)^2$ a pro hodnotu K_2 je obvykle dosazena hodnota 0,03.

Po odečtení normalizace signálu a jasu zjistíme strukturní informaci. Pomocí korelace získáme vzorec [11].

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (1.9)$$

Samotnou korelaci získáme použitím vzorce [11].

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y). \quad (1.10)$$

Nyní můžeme konečně díky kombinaci všech tří složek získat výpočet SSIM indexu [11].

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1.11)$$

Při vypočtení kvality obrazu počítáme z pravidla ve čtvercích o rozměrech 8x8 pixelů. Je to z důvodu výpočetní náročnosti u obrazu s velkým rozměrem. Nastávají i případy, kdy čtvercový tvar okna vytváří nežádoucí blokové artefakty. Díky tomu je možné použít i jiný tvar oken. Toto okno je postupně posouváno pixel po pixelu. Nakonec ze všech výsledků vypočítáme celkový průměr SSIM indexu. Tomuto postupu se říká MSSIM [11].

$$\text{MSSIM}(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{j=1}^M \text{SSIM}(\mathbf{x}_j, \mathbf{y}_j) \quad (1.12)$$

Kde \mathbf{X} a \mathbf{Y} značí původní popřípadě komprimovaný obraz. Hodnoty x_j a y_j nám značí obsah j -tého okna obrazu [11].

2.2.4 VMAF (Video Multi-method Assessment Fusion)

Metrika byla vytvořena firmou Netflix. Jedná se o metriku kvality videa, která má za úkol přiblížit hodnocení tak, jako by jej hodnotil skutečný člověk. Využívá několik zavedených metrik. Strojové učení a model HVS (Human visual system model), který je zaměřen právě na lidské vnímání obrazu. Tato analýza byla integrovaná do nástrojů třetích stran (například FFMPEG, MSU Video Quality). Výsledkem hodnocení je VMAF skóre, které je na stupnici od 0 do 100, kde 100 značí nejlepší kvalitu. Za “Dobré” skóre je považován výsledek 80 a více [36].

2.3 Další objektivní metody hodnocení kvality videa

Pro objektivní metody je i řada jiných metod hodnocení, které nemusí vycházet z referenčního snímku.

2.3.1 NC (Normalized Correlation)

Jedná se o vzájemný vztah dvou procesů popřípadě veličin, které představují dva různé obrazy s rozměry barvové informace $M \times N$ a to všechno v rámci jednoho kanálu. Když je zjištěna, že se nachází mezi jednotlivými veličinami korelace, můžeme dojít k závěru, že jsou s velkou pravděpodobností na sobě závislé. Nedá se ovšem pouze na základě korelace říci, zdali se jedná o případ příčiny a následku. Vzorec pro výpočet NC vypadá podle vzorce (1.12), kde $X_{i,j}$ a $Y_{i,j}$ jsou hodnoty pixelů pro první a druhý snímek [25].

$$NC = \frac{\sum_{i=1}^M \sum_{j=1}^N X_{i,j} Y_{i,j}}{\sqrt{\sum_{i=1}^M \sum_{j=1}^N X_{i,j}^2}} \quad (1.13)$$

2.3.2 AD (Average Difference)

V této metodě udává průměrný rozdíl, který se počítá z dvou jiných snímků s rozměry $M \times N$ a hodnotou rozdílů pixelů každého snímku $X_{i,j}$ a $Y_{i,j}$ [25].

$$AD = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (X_{i,j} - Y_{i,j}). \quad (1.14)$$

2.3.3 SC (Structural Content)

Využíváme zde poměr druhých mocnin a to pro hodnoty pixelů $X_{i,j}$ a $Y_{i,j}$. Kde $X_{i,j}$ je referenční obraz a $Y_{i,j}$ je testovaný obraz s rozměry $M \times N$ [25].

$$SC = \frac{\sum_{i=1}^M \sum_{j=1}^N X_{i,j}^2}{\sum_{i=1}^M \sum_{j=1}^N Y_{i,j}^2} \quad (1.15)$$

2.3.4 MD (Maximum Difference)

Metoda spočívá ve výpočtu rozdílů dvou největších hodnot pixelů $X_{i,j}$ a $Y_{i,j}$ pro dva různé obrazy [25].

$$MD = \text{Max}(|X_{i,j} - Y_{i,j}|) \quad (1.16)$$

2.3.5 NAE (Normalized Absolute Error)

Opět vychází z rozdílů dvou hodnot pixelů $X_{i,j}$ a $Y_{i,j}$ různých obrazů s rozměry $M \times N$, které jsou následně poděleny referenčním obrazem [25].

$$NAE = \frac{\sum_{i=1}^M \sum_{j=1}^N |X_{i,j} - Y_{i,j}|}{\sum_{i=1}^M \sum_{j=1}^N X_{i,j}} \quad (1.17)$$

3 KÓDOVÁNÍ VIDEOOBSAHU

Cílem vytváření kódování je především vytvoření vhodného obsahu pro přenos dat. Jedná se zejména o snížení přenosové rychlosti a tím i zmenšení šířky pásma, která je potřebná k přenosu. Tuto možnost provádí zařízení zvané kodek, jehož spojení slov je enkodér a dekodér. Funkce je taková, že enkodér umožňuje kompresi původního snímku nebo videa a dekodér následně převádí obsah zpět na zobrazovací formát [26]. Pro tuto funkci se využívá celá řada kodeků.

3.1 H.265/HEVC

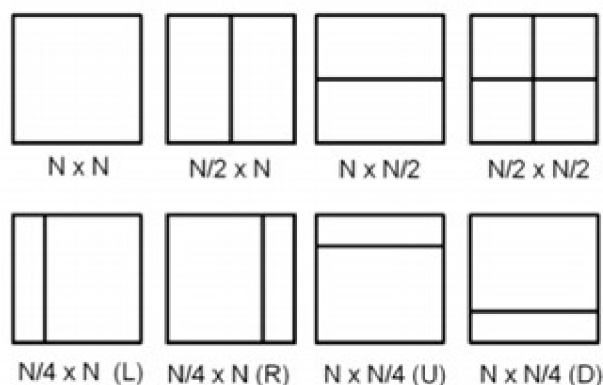
Jeden z možných nástrojů pro kompresi videí je právě HEVC (High Efficiency Video Coding). Za jeho sestrojením stojí skupina JCT-VC. Doporučení tohoto standartu je obsaženo v ITU T-REC-H.265. Můžeme o něm říct, že je nahrazením předchozího kodeku H.264. Je až o 50% účinnější, ovšem musíme počítat s vyšší náročností výpočtů. Využívá se převážně pro neprokládané videosekvence s vyšším rozlišením [12].

3.1.1 Stromová struktura

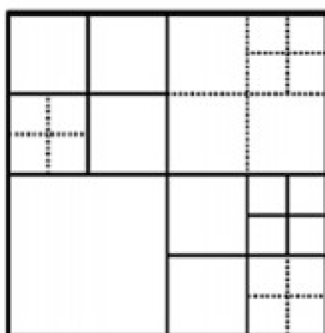
Oproti předchozím kodekům, které HEVC nahradil, spočívá jeho změna ve stromové struktuře. Díky tomu že je závislá na kódované jednotce CTU, tak se stává daleko více adaptivní. Ta se skládá z bloku jasového CTB a chrominanční blok CTB. Pomocí zvolení enkodéru můžeme měnit velikost těchto bloků. Jasový blok CTB má rozměr $N \times N$ vzorků, u chrominančního, je to poloviční hodnota tedy $N/2 \times N/2$ vzorků. N nabývá hodnot 16, 32 a 64. bloky CTB jsou dále rozděleny na kódované bloky CB. Jedná se opět o čtvercovou strukturu, která musí mít minimální velikost 8×8 pro jasové složky a alespoň 4×4 pro složky chrominanční. Jejich omezení v rámci maxima je stanoveno podle velikosti CTB. kódovaná jednotka CU je tvořena jedním jasovým a dvěma korespondujícími chrominančními CB bloky. Pro každou CU jednotku je přiřazen jeden ze dvou možných módů, inter nebo intra. Dále je rozděleno CU na predikční a transformační jednotku, tomu odpovídají bloky PB a TB.

Bloky PB se řadí do obdélníku a to s minimálním rozměrem 8×4 nebo 4×8 vzorků. Ve struktuře nadřazených CB bloků se dále nachází 1, 2 nebo 4 PB bloky. PU jednotka je tvořena obdobně jako kódované bloky CB. Při použití asymetrického dělení, můžeme použít pro blok CB s velikostí $N \times N$ rozdělení na dva jasové bloky PB. Přičemž první je o velikosti $(N/4) \times N$ nebo $N \times (N/4)$ a druhý musí být doplněn na danou velikost, tedy $(3N/4) \times N$ nebo $N \times (3N/4)$.

Dále se používá jednotka pro transformaci TU. Taktéž jsou rozděleny do bloků TB jako u předchozí jednotky. Jejich náplní jsou residua po predikci. Jejich velikost se odvíjí od velikosti CU a dále se větví do stromové struktury. Jedná se vždy o čtvercové bloky s velikostí 4×4 , 8×8 , 16×16 a 32×32 vzorků [12].



Obr. 3.1: Velikosti pro PB [12]



Obr. 3.2: Rozvržení CB a TB v CTB
(plná čára ohraničení CB, tečkovaná čára ohraničení TB) [12]

Pro transformační kodování se používají čtvercové TB bloky, jejich velikost volena pomocí enkodéru. Metoda vychází z diskrétní kosinovi transformace, která je základem pro 2D celočíselnou transformaci. Tuto transformaci využíváme zvláště pro každý jasový a chrominanční blok [12].

3.1.2 Entropické kódování

HEVC obsahuje pouze jednu metodu entropického kódování s názvem CABAC. Využívá obdobného algoritmu jako H.264/AVC. Jistým změny dokážou zvýšit přenosové rychlosti je také schopen přístupu k jednotlivým blokům již zmíněné stromové struktury a tím zvýšit efektivitu tohoto kódování [12].

3.1.3 Interpredikce

Interpredikce pracuje s dvěma vektory pohybu, které jsou přiřazeny k jednotce PU. Pohyby vektoru jsou voleny buď pro jednosměrnou, nebo obousměrnou predikci. Principem se jedná o to, že vektor nepřenáší celý, pouze jeho rozdíl oproti hodnotě předpokládané (predikované). Díky tomu je možné použít vyšší počet referenčních snímků a přitom zachovat koncept I, P, B. Predikce, která je čtvrt pixelová, se liší pro jednotlivé bloky jasové a chrominanční. Pro jasovou je zapotřebí použít filtr osmého řádu, zatímco pro chrominanční jen filtr čtvrtého řádu. Tomu odpovídají i velikosti bloků interpredikce a to 4 x 8 nebo 8 x 4 vzorků. Bohužel u těchto bloků je možné použít pouze jednosměrnou predikci. Díky tomu snížíme velikost dané paměti [12].

3.1.4 Intrapredikce

Intrapredikce je u HEVC prováděna predikcí sousedních bloků. Několik základních módů, které se dají pro predikci použít. Jedná se o módy vertikální, horizontální, úhlová, DC a planární. Pro dominantní vertikální predikci využíváme body na levé straně od kódovaného bloku, ty se nazývají hlavní pole. Body nad kódovaným blokem nazýváme boční pole. U horizontální je to opačný případ. Výpočet se poté provádí lineární interpolací všech bodů. Při použití větších velikostí bloků je použita rozšířená predikce, která obsahuje 35 módů, které jsou rozděleny podle úhlů [12].

3.1.5 Rekonstrukční filtr

Aplikace tohoto filtru pro kodek HEVC se dá popsat dvěma kroky. V prvním je použito pomocí filtru blokové potlačení struktury a ve druhém je využito SAO (Sample adaptive offset) filtru.

Filtr v prvním kroku slouží k omezení blokových znaků, které vznikají důsledkem kódování. Při povolení, je přiveden na všechna rozhraní mezi bloky a to vše v úrovni pro TU a PU. Filtr využívá mřížku 8 x 8 vzorků pro jasový i chrominanční signál. U předešlému kodeku H.264 je tato mřížka poloviční. HEVC je tedy oproti starší verzi méně náročný na výpočty. Nalezneme i podmínky filtru pro jasového i chrominanční signály. K filtraci hran jasového bloku dochází, když je pomocí intrapredikce získán alespoň jeden blok, mají rozdíl větší než jeden pixel při rozdílu mezi vektory pohybu bloků nebo jsou nenulové. Při filtraci jasového vzorku určujeme tři stupně – žádná, slabá a silná. Výběr je závislý na vlastnostech daných sousedních bloků. U chrominančních vzorků jsou stanoveny pouze dvě filtrace a to žádná nebo normální.

Použití filtru SAO se používá v úrovních CTB a to jen na vzorky, které splňují dané požadavky. Dá se aplikovat na vzorky, které již prošly prvním krokem. Tento filtr přidává vzorkům offset, který je stanoven enkodérem [12].

3.2 VP9

Tento kodek byl vyvinut společností Google Inc. s cílem kódovat sekvence ve vysoké kvalitě. Jedná se o nástupně předchozí verze VP8. Tento kodek vznikl na základě projektu WebM, kde jde o zdokonalování kodeku pomocí velké skupiny vývojářů, neboť se jedná o otevřený formát. V červnu roku 2013 byl tento kodek oficiálně vydán. Při kódování bychom tento kodek našli v multimediální knihovně pod názvem libvpx.

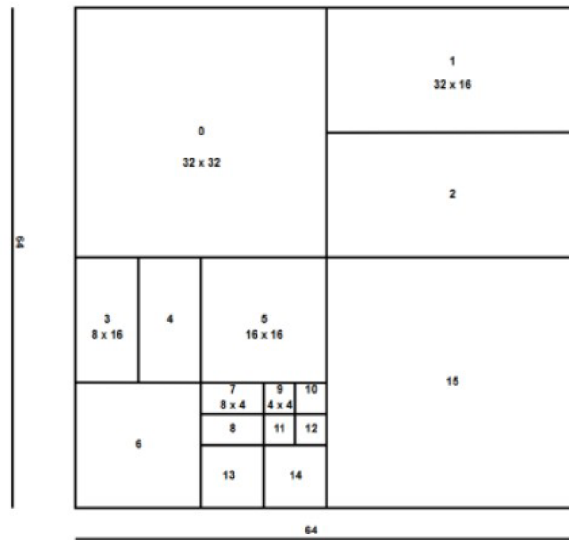
Důvod neustálého vývoje kodeku je zejména velký pokrok a náročnost uživatelů, kteří v dnešní době již vlastní zařízení s dobrou kvalitou obsahu. Hlavní uplatnění pro tento kodek nalezneme zejména na internetu, konkrétně ve webových aplikacích. Nejvíce je využíván internetovými video servery a aplikacemi používanými na videokonference. Taktéž tento kodek podporují všechny internetové prohlížeče. S tím je i spojená široká možnost užití na různých platformách jako jsou mobilní telefony, televize, počítače, smart televize atd.

Co se týče vysoké kvality videí, na které se tento kodek využívá, je tedy očekávané že podporuje základní standardy rozlišení, jako jsou 720p (1280 x 720), také full HD s rozlišením 1080p (1920 x 1080). Také už v dnešní době je aktualizován pro podporu Ultra HD videí, tedy 2160p (3840 x 2160).

Z počátku bylo podporováno pouze jeden typ barevného vzorkování YCbCr 4:2:0 s 8 bitovou hloubkou. Postupně se do kodeku vložily vlivem vývoje další tři nové profily, které rozšiřují možnosti vzorkování a bitovou hloubku. Profil 0, ten již zde byl zmíněn. Profil 1 je již rozšířen o podporu vzorkování YCbCr 4:2:2 a 4:4:4. Profily 2 a 3 jsou dále obohaceny o podporu 10 a 12 bitovou hloubku. Také tyto dva profily obsahují podporu alfa kanálu [13].

3.2.1 Dělení snímků

Oproti předchozí verzi prošel kodek VP9 zvýšením blokových oblastí jednotlivých sekvencí a tím se zvýšila i jeho účinnost. Využívá čtverců, které jsou až 64 x 64 vzorků. Tyto bloky můžeme zmenšovat až po velikost 4 x 4 vzorků. Takto malé bloky využíváme pro velkou jemnost predikce. Pro využití co nejvyšší účinnosti je zapotřebí, aby kódování a dekódování probíhalo paralelně. Díky tomuto faktu je VP9 vybaveno tzv. `frame_parallel_mode` [13].



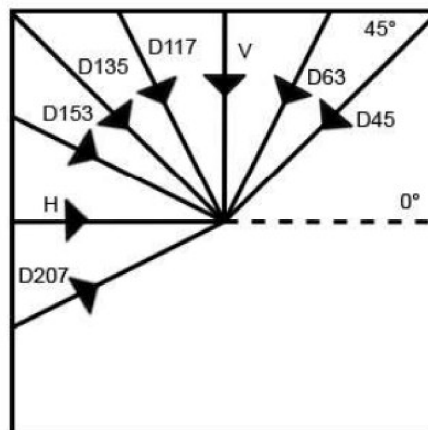
Obr. 3.3: Rozložení bloků a možnosti jejich dělení [13]

3.2.2 Entropické kódování

Využívá převážně aritmetického entropického kódování. Při kterém je vytvořen tzv. binární strom. Tento strom je tvořen z vybraného symbolu x prvkové abecedy, kdy počet hlavních uzlů je roven hodnotě $x - 1$. Každý vzniklý uzel obsahuje binární aritmetický kodér, který pracuje s 8 bitovou přesností [13].

3.2.3 Intrapredikce

Pro využití intrapredikce v kodeku VP9 existuje 10 různých módů, které jsou rozděleny podle úhlů a nabízí velikosti od 4×4 do 32×32 vzorků. Několik základních módů, které pro tuto predikci můžeme použít. Jedná se o vertikální (V), horizontální (H), true motion (TM) a 6 diagonálních predikcí, kde je označení D a poté číslo úhlu, který zaujímají [13].



Obr. 3.4: Možnosti režimů intrapredikce [13]

3.2.4 Interpredikce

Pro VP9 se dají použít až 4 módy určené pro iterpredikci s různou velikostí bloků od 4 x 4 do 64 x 64 vzorků. Při využití pohybové kompenzace je použit kodekem pouze jeden vektor. Ovšem můžeme spustit možnost složené predikce, která umožní použít dva vektory pohybu pro jeden blok. Výsledný vzorek se poté zprůměruje a tím získáme požadovanou hodnotu.

Vektor nám určí také o jaký referenční snímek se jedná. Může nastat jedna ze tří různých variant. Last frame, ten udává, jestli se jedná o již kódovaný snímek. Dále Goldenframe, ten nám dá informaci o tom, že se jedná o referenční snímek, který je uložen v bufferu. Poslední je Alterante reference frame, který se oproti předchozím případům nemusí vůbec zobrazit. Využívá se pouze pro zvýšení přesnosti predikce u referenčních snímků [13].

Predikce pohybových vektorů je stanovena seznamem, kde se nachází možné referenční pohybové vektory. Tento seznam je tvořen z referenčního obrazu, kde je až osm sousedících bloků, které jsou tvořeny těmito vektory. Pokud není seznam referenčních vektorů naplněn, provede se opět stejný proces. Ovšem už to není podmíněno referenčním obrazem [13].

Když už dojde k přiřazení bloku, můžeme provést kódování pomocí možných čtyř módů daného pohybového vektoru. Jednotlivé módy jsou NEW_MW, ten nám umožní přidat do již zmiňovaného seznamu přidat první zápis a to s možností s pohybovým vektorem delta, který je obsažen v datovém toku. Druhý mód NEAREST_MW slouží k prvnímu zápisu do seznamu, ale už bez možnosti úprav. S módem NEAR_MW se dá pracovat obdobně, ovšem je proveden pro druhý zápis seznamu opět bez dalších úprav. Čtvrtý a poslední je ZERO_MW, ten stanoví pohybový vektor na nulovou pozici (0,0) [10].

3.2.5 Transformační kódování

V kodeku VP9 nalezneme základní tři druhy transformace. Jedná se o asymetrickou diskrétní sinusovou transformaci (ADST), diskrétní kosinovou transformaci (DCT) a v poslední řadě Walsh-Hadamardovu transformaci (WHT). Jednotlivé transformace volíme podle režimu zvolené predikce. Pro využití transformace ADST najdeme uplatnění v iterpredikci, často s kombinací DCT, kde nastávají hybridní transformace v blocích o rozměru 4x4 až 16x16 vzorků. Transformaci DCT využíváme především pro intrapredikci a to pro bloky o velikosti 4x4 až 32x32 vzorků. Poslední transformaci WHT používáme pro bezztrátové kódování videí. Pro tuto možnost musí být velikost bloků 4x4 vzorků [13].

3.2.6 Rekonstrukční filtr

Vzhledem k tomu, že VP9 zvládá zpracovávat větší velikosti jednotlivých bloků, je zapotřebí, aby rekonstrukční filtr vyhladil ostré hrany v daném snímku. K tomu se využívá tzv. deblock filtr, který slouží k zahlazení hran bloků v celém obraze. Nejprve provádí tuto funkci pro směr vertikální a poté až horizontální. Najdeme velkou šířku filtrů o velikosti 2 až 16 vzorků.

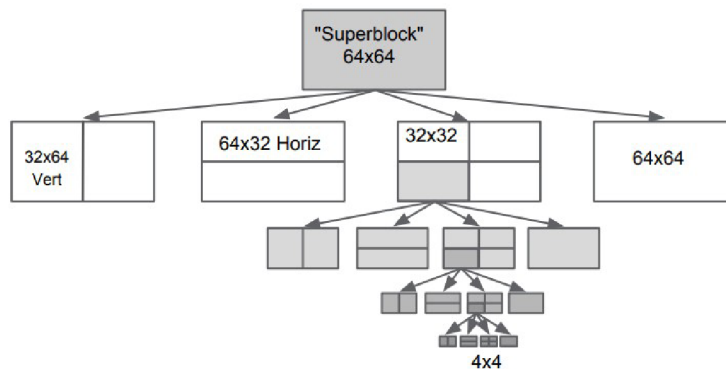
Rekonstrukční filtr obsahuje také možnost detekce plochosti snímku. Díky tomu kontroluje statické nebo ploché části obrazu a podle toho upravuje i velikost daného filtru [13].

3.3 AV1

AOMedia Video 1 (AV1) je zdokonalení předchozího kodeku VP9, jehož cílem je zvýšit účinnost až o 50%. Byl vytvořen společností Alliance for Open Media roku 2015. Hlavní myšlenkou této společnosti je vývoj kodeku, který bude bezplatný s širokým použitím a s nízkou velikostí. Oproti předchozím kodekům nezpůsobuje rozostření a velkou změnu detailů a to i při malých bitratech. Kodek je schopen využívat predikce barvonosné složky [15].

3.3.1 Dělení snímků

V případě dělení snímků je u AV1 tvořeno jedním takzvaným superblokem, který má rozměr až 64x64. Pro další dělení se používají další čtyři možnosti dělení bloků a to na obdobný blok 64x64 nebo s rozdělením na čtverec, kde každý blok velký 32x32. Dále se využívá vertikální rozdělení, zde jsou bloky rozpůleny daným směrem, tedy 32x64. Obdobná je funkce i pro horizontální rozdělení a to 64x32 bloků. Pro další dělení se postupuje u čtvercových bloků stále ke zmenšování a to až na velikost 4x4 bloků [14].



Obr. 3.5: Dělení bloků pro AV1 [14]

3.3.2 Entropické kódování

Využívá multi-symbolického aritmetického kódování. Kodek je již schopen využívat abecedu s více symboly. Oproti předchozím verzím nabízí až 15 bitovou pravděpodobnost a velikost abecedy je 16 symbolů. Encoder zapisuje od konce bufferu, přičemž pro jeden snímek je použit právě jeden buffer [14].

3.3.3 Intrapredikce

Pro AV1 se využívají směrové intraprediktory, které jsou schopny extrapolovat vedlejší pixely na základě zvoleného úhlu. AV1 disponuje osmi hlavními režimy pro volbu směru a také další tři režimy pro nesměrovou predikci. Mezi nimi jsou DC, Paeth prediktor a Smooth prediktor. Rozložení úhlu pro hlavní režimy je v rozmezí 45 až 203 stupňů. Pro hladký prediktor Smooth se využívají tři možnosti pro svislý nebo vodorovný směr, popřípadě využití průměru jejich aproximace. Označují se SMOOTH_V, SMOOTH_H a SMOOTH. Slouží k předpovědi bloku pomocí kvadratické interpolace. Path prediktor spočívá v kopírování pixelů. Nejprve dojde ke kopírování horní periferie vedlejšího bloku, poté z levé periferie vedlejšího bloku a nakonec v levém horním rohu. Tento pixel by se měl blížit k hodnotě P, která se počítá z: levý + horní - levý horní. Posledním prediktorem je barvonosná predikce, která se využívá pro modelování pixelů pomocí lineárních funkcí rekonstruovaných jasových pixelů [15].

4. TVORBA OBSAHU PRO VR

Pro tvorbu obsahu, který by mohl být implementován do virtuální reality je celá řada. Můžeme použít klasické fotografie, panoramata, 180° a 360° snímek a mnoho dalších. Totéž se dá použít i pro videa, z toho důvodu, že existuje řada možností, je důležité dbát na celý proces tvorby prostředí, které následně chceme pozorovat. Ať je řeč o samotném nahrání, popřípadě úprava kvality, střih, vyhlazování a tak všeobecně další procesy, které slouží k vytvoření co nejlepšího uživatelského zážitku.

4.1 Fotografie

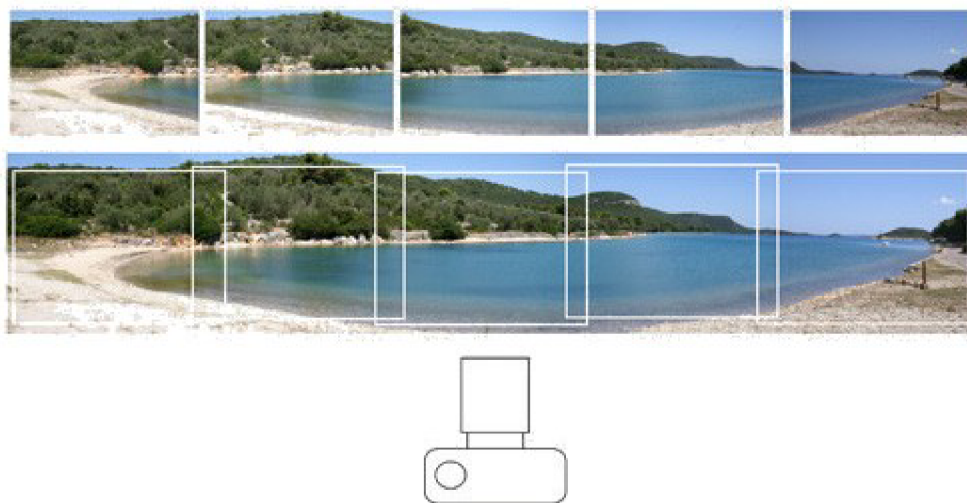
4.1.1 3D fotografie

Tvorba 3D fotografií byla, jak už je známo z historické části, již v 19. století. Pomocí dvou obrazů vytvořila takzvaná hloubka, která je základem pro 3D model. V dnešní době je práce s 3D snímky obdobná, ovšem již na toto zobrazení existují různé počítačové software, které jsou schopny upravit fotografii tak, aby se dosáhlo 3D efektu. Princip spočívá, jak již bylo zmíněno, ve dvou fotografiích, které jsou promítány pod mírně změněným úhlem. Takovou fotografii je možné pořídit fotoaparát a poté například v Adobe Photoshopu upravit a převést do třetí dimenze [17].

4.1.2 Panoramatická fotografie

Panoramatické fotografie vznikají ze skládání jednotlivých fotografických snímků. Toho lze dosáhnout mnoha různým použitím přístrojů od mobilního telefonu po panoramatické kamery a další.

Existují tři základní typy tvorby panoramat: horizontální, vertikální a kruhové. Jednotlivé typy se liší zobrazením, jak plyne z názvu. Pro pořízení horizontální fotografie je zapotřebí držet dané zařízení pro zachycení ve stejné výšce, aby linie probíhala rovnoběžně se spodní částí rámu. Jednotlivé fotografie by měli být překrývány o 10-50% s tím, že čím je vyšší pole překrytí, tím vyšší i komplexnost snímku. Při pořizování vertikálních snímků pro horizontální panorama dosáhneme vyšší rozlišení, než pro tvorbu pouze z horizontálních fotografií [16].



Obr. 4.1: Ukázka skládání panoramatické fotografie [16]

4.1.3 360 stupňová fotografie

Pro tvorbu takové fotografie je důležité zachytit celý prostor kolem dokola. Pro záznam se dá použít širokoúhlý objektiv, ovšem bude zapotřebí udělat několik desítek fotek pro celé zobrazení. Taktéž se dá použít například fotoaparát s rybím okem, zde už je zapotřebí menší množství záznamu, tedy od 3 do 6 snímků. Rybí oko má tu výhodu, že zvládá zachytit zejména snímky pořízeny z malé vzdálenosti od objektu a zároveň má velký úhel zachycení a to až 220° . Díky tomu je jednoduché spojení těchto fotografií, z důvodu dostatečného překrytí [18, 26].

V základu se jedná o panoramatickou fotografii. Jedním z nejkvalitnějších typů tvorby takového snímku je sférická 360° fotografie, kde je zachycen celý prostor i s pohledem dolů a nahoru. Další výhodou je, že nejde vidět fotograf ani stativ, na kterém je fotoaparát upevněn. Kvalita je v tomto případě velmi dobrá za použití vhodného nastavení fotoaparátu [18]. V dnešní době je i možné zachytit 360 stupňovou fotografii na svůj mobilní telefon pouze za použití aplikace. Tuto funkci umožňuje systém, který uživateli poskytne pole pro tvorbu fotografií a pouhým focením těchto bloků začne program spojovat jednotlivé snímky a tím vytvoří 360 stupňovou fotografii. Nevýhodou je zde systém spojování v místnostech, kde dochází k ne příliš dobrému spojení dvou sousedících snímků, zejména je to pozorovatelné u stropů popřípadě ostrých hran [19]. Dále se mohou snímky vytvářet pomocí fotoaparátů RICOH THETA nebo Insta360, které vytváří fotografii pomocí dvou čoček typu rybí oko, které mají zorné alespoň 180 stupňů a dokáží spojit statický objekt často bez známky překrytí těchto kamer [21].



Obr. 4.1: Ukázka skládání 360° fotografie za pomoci mobilní aplikace [19]

4.2 Videosekvence

Videa pro VR mohou znamenat pro spousty lidí různé pojmy. Může to být 360 video, které si za pomoci headsetu promítnete do VR a můžete se porozhlédnout kolem sebe. Ovšem pokud se jedná o video, není zde možnost sledování polohy. Dá se volně otáčet, ovšem pozice se hýbou pouze tehdy, když je pohyb zaznamenáván ve videu [20].

4.2.1 3D video

Pro tvorbu takového video obsahu je možné použít technologii, která zvládne pořídit několik 2-D snímků z různých pozic a na základě toho vygenerovat 3D model daného videa. Touto technologií disponuje například Photosynth. Díky zobrazení hloubky, je pro diváka video více pohlcující. Na zobrazení ve 3D se dá použít také 360 stupňová videa, kde je pohlčení ještě větší. Pro tvorbu 3D 360 videa je důležité zachycení z několika různých pozic s tím, že každá z nich má zachycené stejné místo z jiných úhlů. Takže je vytvořeno prostředí ze spousty vizuálně propojených scén. Poté je zapotřebí zpracovat tato videa a vygenerovat informace o jejich viditelnosti, v konkrétním místě včetně stínů a textur. Dále je zapotřebí všechny modely spojit dohromady a tím vytvořit model scény. Nakonec je zapotřebí v 3D modelovacím softwaru procházet celým tímto modelem a vyčistit všechny prvky, které mohou danou scénu kazit [23].

4.2.2 360/180 stupňové video

Video 360 poskytuje pohlcující zážitek z 360 stupňového videa. Je zde zapotřebí vlastnit všesměrovou kameru například Insta360 One X nebo Ricoh Theta a další [21].

Kamera funguje na principu duálního objektivu typu rybí oko se zorným polem 180 stupňů a buď je automaticky spojí do kamery a nebo nabízí software, který pomocí jediného kliknutí spojí tyto dva obrazy v jedno 360 stupňové video. Obdobně je vytvořeno i 180 stupňové video, ale použije se pouze pro jediný objektiv. Obvykle má každá čočka vyšší zorný úhel než 180 stupňů, to je z důvodu překrývání záběrů a zamaskování spojení.

360° kamery obsahují poměrně malé senzory, které jsou umístěny za každým objektivem. Díky malé velikosti senzorů a velkému úhlu snímání může nastat problém se zobrazením stínů. Tento problém ovšem řeší možnost použití doprovodných aplikací, které mohou modifikovat video pomocí filtrů a tím zlepšit danou expozici [22].

Nevýhodou použití těchto 360° kamer je finanční a časová náročnost. A to zejména z toho důvodu, že je zapotřebí kvalitní video obsah celého záznamu, který je problém zachytit na běžné zařízení. Z toho důvodu může být velkou výhodou využití 180° videí, které využívají menší zobrazovací pole a tím může být navýšena kvalita. omezí se časová náročnost a finance na pořízení přístroje [28].

4.3 Rozlišení videosekvencí

Rozlišení ve virtuální realitě je jeden z mnoha parametrů, který má vliv na lepší uživatelský zážitek (QoE). Ve virtuálních brýlích nalezneme dvě varianty displejů. První jsou dva displeje, a to každý pro jedno oko. Ve druhém případě je jeden displej rozdělen na dvě části, opět pro každé oko. Obvykle se vyjadřuje v klasickém rozlišení šířka x výška, ale je možnost pozorovat kvalitu v takzvaném ppi (počet pixelů na palec). Když se budeme bavit o kvalitě obsahu, který budeme do virtuální reality nahrávat. Zapotřebí je najít maximální možné rozlišení videa a snímkový kmitočet pro daný HMD. Například brýle Oculus Go a Quest dokáží dekódovat maximálním rozlišení 5760 x 2880 a při 30 snímcích za sekundu. Další důležitou věcí je, že pro kodek HEVC je omezená horizontální poloha do 4096 pixelů. Vertikálně může být tato hodnota přesažena. Z toho důvodu je důležité použít poměr stran 1:2 a to zejména při vysokém rozlišení videa [29].

4.4 Snímkový kmitočet (FPS)

Jedním z hlavních parametrů videa je i snímkový kmitočet. Jedná se o počet snímků, který je přehráván za jednu vteřinu. Obvykle se tato hodnota pohybuje kolem 24 snímků/s, což odpovídá přirozenému vnímání oka. Dále se používají také často videa s 30 a 60 snímků. Tento počet je zejména pro videa, kde je zaznamenáván pohyb. Jakožto přenos fotbalového utkání. Zde je spousta podnětu, které se hýbají, tak je důležité zajistit plynulost obrazu. Pro 60 snímků se dále najde použití například při tvorbě zpomaleného videa [31].

5 NÁVRH SUBJEKTIVNÍHO TESTU A VYHODNOCENÍ

5.1 Příprava a volba videí

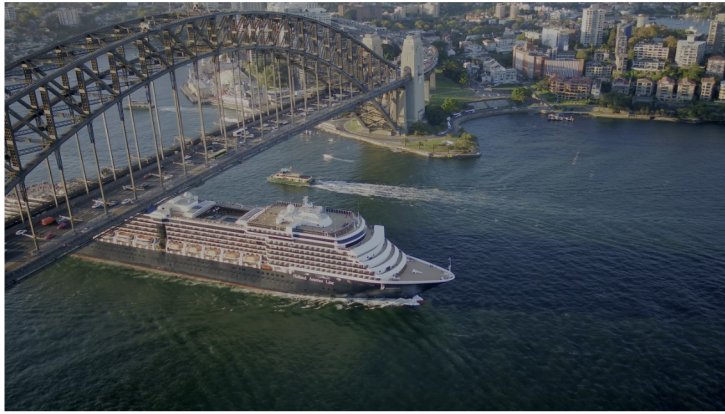
Při výběru vhodných videosekvencí bylo zapotřebí zvolit nejvyšší možné rozlišení, které podporuje headset, který při tomto testování budeme využívat. Oculus Go je schopen přehrávat videa až v rozlišení 5K (5760x2880), ale pro naše účely a dostupnost byla zvolena kvalita 4K, která se již hojně vyskytuje zdarma ke stažení na internetu. Dalším důležitým parametrem je datový tok, který při našem testování musí dosahovat alespoň 17Mb/s z toho důvodu, aby bylo možné videosekvenci upravovat a snižovat datový tok a tím docílit vhodného prostředí pro testování.

Pro testování byly zvoleny tři typy videí a to klasické 2D flat video, 3D video a 360° video, abych docílil různých pohledů diváka na videosekvenci, kterou Oculus Go dovede přehrát. Video pro účel testování bylo staženo z internetových portálů. Flat video označeno dále jako City bylo z portálu Playboard [33]. Video 3D označeno Forest bylo zasláno společností 3D-Taeuber [34]. Poslední 360° video dále Street bylo převzato z internetové galerie Kandao [32]. Videosekvence byly zkráceny na časovou stopu od 15 do 18 sekund, aby bylo možné všimnout si detailů a zároveň při testování nedocházelo k nudným pasážím dlouhých opakujících se sekvencí. Původní videa byla postupně zakódována do formátů, které se nejběžněji vyskytují s presetem v defaultním režimu (normal). Jedná se o VP9 (single-pass encoding mode) a HEVC (1-pass target bitrate) a jejich nejefektivnější následovník AV1 (AOMENC), který ještě není podporován velkou částí přehrávačů. Samotná videa již byly ve formátu VP9. Úpravy byly provedeny pomocí příkazového řádku s použitím programu FFmpeg [37]. Upravoval se formát i datový tok na hodnoty pro všechna videa 2, 6 a 15Mb/s.

Název	Rozlišení	Délka	Tok dat	formát	Typ videa	FPS
City	3840x2160	0:15	30,7Mb/s	VP9	flat	60
Forest	3840x2160	0:15	23,4Mb/s	VP9	3D	25
Street	3840x2160	0:18	26,3Mb/s	VP9	360°	30

Tab.5.1: Podrobnosti o použitých videosekvencích

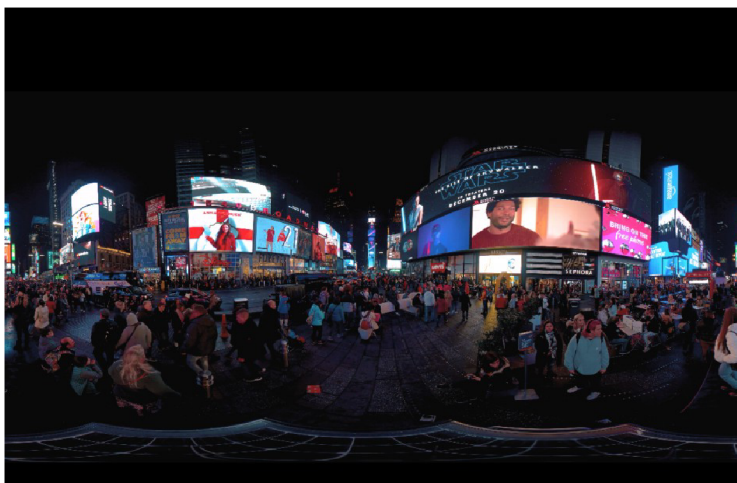
Při shánění vhodných videosekvencí docházelo k problémům zejména s volně dostupnými videi s vysokým bitratem a rozlišením 4K. U některých videí docházelo i k úplné nemožnosti přehrání v zařízení Oculus Go, kde přehrávač nebyl schopen toto video přehrát. U kódování se zvoleným bitratem, jsem se potýkal s problémem přesného nastavení. Datový tok se volil v rozmezí min-max, a to mělo za následek několik úprav jedné videosekvence, aby se docílilo správného toku u všech videí. Bohužel AV1, který byl nejefektivnější a nejvhodnější, provázal nejvíce problémů zejména s dlouhým kódováním a nemožností podpory u velké části volně dostupných přehrávačů. Výjimkou nebyl ani přehrávač v Oculus Go.



Obr.5.1: Obsah videosekvence City



Obr.5.2: Obsah videosekvence Forest



Obr.5.3: Obsah videosekvence Street

Střih videa probíhal za použití příkazu:

```
ffmpeg -i "d:\Users\PC\Desktop\Videa BP\3D\3D_Forest.mkv" -ss 00:01:11.0 -c copy -t 00:00:15.0 "d:\Users\PC\Desktop\Videa BP\3D\3D_Forest_cut.mkv"
```

kde *-i* je cesta k upravované sekvenci, *-ss* je doba, od které chceme sekvenci stříhat, *-c copy* kopíruje sekvenci, *-t* čas, který uběhne než dojde ke střihu

Dále jsem jednotlivé video obsahy musel upravit na jednotlivé typy kodeků. Pro mou práci jsem zkoumal ty, které jsou nejvíce používané a také ty, které se teprve začínají na různých streamovacích serverech používat. Jedná se o kodeky H.265/HEVC, VP9 a v poslední řadě AV1. Všechna tyto videa jsou upravována díky využití FFmpegu verze 4.3.2, který je volně ke stažení na internetu a obsahuje řadu funkcí již vložených.

Jednotlivé příkazy kodeků jsou:

HEVC

```
ffmpeg -i "d:\Users\PC\Desktop\Videa BP\3D\3Dlescut.mkv" -c:v libx265 -b:v 2M -minrate 1.5M -maxrate 2.5M -c:a aac "d:\Users\PC\Desktop\Videa BP\3D\Forest_2MHEVC.mkv"
```

VP9

```
ffmpeg -i "d:\Users\PC\Desktop\Videa BP\3D\3Dlescut.mkv" -c:v libvpx-vp9 -b:v 2M -minrate 1.5M -maxrate 2.5M -c:a aac "d:\Users\PC\Desktop\Videa BP\3D\Forest_2MVP9.mkv"
```

AV1

```
ffmpeg -i "d:\Users\PC\Desktop\Videa BP\3D\3Dlescut.mkv" -c:v libaom-av1 -b:v 2M -minrate 1.5M -maxrate 2.5M -c:a aac -strict experimental "d:\Users\PC\Desktop\Videa BP\3D\Fotest_2MAV1.mkv"
```

<i>-i</i>	cesta k souboru, který je upravován,
<i>-c:v</i>	volba kodeku (libx265 = HEVC; libvpx-vp9 = VP9; libaom-av1 = AV1)
<i>-b:v</i>	nastavení hodnoty bitratu (b/s)
<i>-minrate</i>	minimální datový tok (b/s)
<i>-maxrate</i>	maximální datový tok (b/s)
<i>-c:a</i>	nastavení hodnota kvality audia (aac = 128kb)

Aby bylo možné osobám, které se budou podílet na zkoumání kvality multimediálního obsahu, zhodnotit dané parametry, je zapotřebí změnit hodnoty bitratu, který ovlivňuje kvalitu videa. Rozhodl jsem se použít takové hodnoty, aby bylo možné určit rozdíly jednotlivých videí. Pro mou potřebu jsem si zvolil hodnoty 15, 6 a 2 Mb/s a to z toho důvodu, aby byla patrná jejich změna.

Změny se dají provádět přímým nastavením bitratu za použití *-c:v* a hodnota nebo nastavit maximální a minimální bitrate *-maxrate*, *-minrate*. Také se dá použít takzvaný faktor konstantní kvality *-crf*. Ten slouží k možnosti zachování videa oproti původnímu videu.

Může nabývat hodnot pro 8-bitů hodnoty 0-51, kde 0 je nejlepší kvalita přibližná co nejvíce k původní sekvenci a 51 je nejhorší možná kvalita. Podle hodnot bitů se mění také šířka daného nastavení.

5.2 Volba subjektivní metody a vyhodnocení

Pro volbu subjektivní metody jsem se rozhodl použít metodu ACR-HR se skrytou referencí. Je obohacena oproti ACR ještě o původní snímek, který je pro respondenty náhodně umístěn v měření. Tato metoda vychází, jak už bylo zmíněno v kapitole 2.6.1., z nahraných videí, která jsou postupně spouštěna za sebou s referenčním snímkem, a poté hned hodnocena v rámci šedého okna. Použití této metody je vhodné pro rychlejší provedení jednotlivých subjektivních testů. Respondenti budou hodnotit sekvence na stupnici od 0 do 100 z důvodu vyšší rozlišovací přesnosti. Pro takové měření je zapotřebí větší počet měření, tudíž pro můj návrh je vhodné použít hodnoty alespoň 15 subjektů podle ITU-R BT.500.

Pro vyhodnocení videí bude použit, jak již bylo zmíněno absolute category rating with hidden reference (ACR-HR). Pro vyhodnocení se v takovém případě dá použít takzvaná hodnota rozdílu kvality DMOS. Tím získáme rozdíl mezi testovanou a referenční sekvencí za pomoci použití vzorce[35]:

$$SA(KS) = V(KS) - V(REF) + 100 \quad (5.1)$$

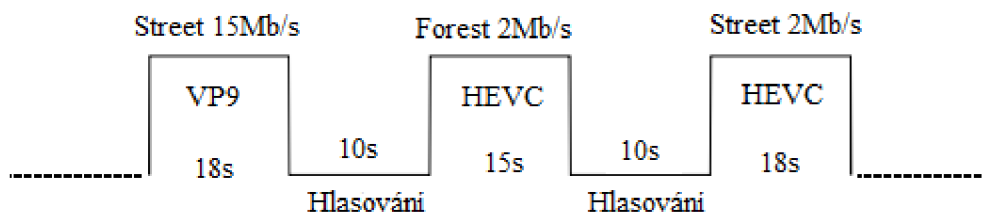
Kde, SA je skóre rozdílu, KS je označení pro kódovanou sekvenci, REF značí referenční snímek a V značí hodnotu zvolenou uživatelem podle tabulky 1.2. Při vyhodnocení značí Výbornou kvalitu SA = 100 a pro Velmi špatnou kvalitu SA = 0.

Může dojít i k situaci, kdy referenční snímek je hodnocen respondentem tak nízko, že výsledek DMOS, který počítá se vzájemným rozdílem hodnoceného videa, je zkreslený. V tomto případě je vhodnější použít MOS (Mean Opinion Score), který je počítán z průměru hodnocení jednotlivých videí. Opět se využívá stupnice 0-100. Tato metoda byla zvolena při vyhodnocení subjektivního testování.

Hodnocení v procentech	Slovní subjektivní hodnocení
100-80	Výborné
80-60	Dobré
60-40	Ucházející
40-20	Špatné
20-0	Velmi špatné

Tab.5.2: Tabulka pro hodnocení jednotlivých video sekvencí

5.3 Průběh subjektivních testů



Obr.5.4: Průběh mnou navrženého měření pomocí metody ACR-HR [24]

Měření započalo nejprve menším dotazem na daného respondenta ohledně jeho zkušeností s virtuální realitou a otázka na rozlišení zařízení, na kterém nejčastěji sleduje videa. Poté došlo k poučení respondenta o účelu jeho hodnocení a následně vysvětlení, jak videosekvence budou probíhat, kdy bude hrát video po dobu přibližně 15 sekund a následně bude prostor pro hlasování obdobně jako na obr.5.4. Poté byl respondentovi nasazen headset Oculus Go s optimálním usazením na hlavě a spuštěn test. Videosekvence byly spouštěny za sebou a já jsem zapisoval hodnoty hlášené testovanými osobami 0 až 100 do tabulkového programu. Videosekvence v celku trvaly okolo 13 minut, přičemž celý test i s poučením a vysvětlením se pohyboval okolo 20-25minut.

Z tohoto důvodu byly použity pouze tři videosekvence s třemi variantami datových toků, aby nebyl subjektivní test příliš časově náročný a mohlo by dojít ke zkreslení výsledků, které by ovlivnila soustředěnost respondentů.

6 VÝSLEDKY HODNOCENÍ KVALITY POMOCÍ VYBRANÝCH OBJEKTIVNÍCH A SUBJEKTIVNÍCH METOD

Cílem bakalářské práce je porovnat nejvhodnější multimediální obsah, který lze přehrávat ve virtuální realitě. Pro toto testování byly použita UHD (3840x2160) videa s datovým tokem 2Mb/s, 6Mb/s a 15Mb/s. Tato videa byla komprimovaná z originálních videí s vysokým datovým tokem několik desítek Mb/s. Následně byly zakódovány do tří kodeků. Které patří k nejvyskytovanějším v dnešní době. Jedná se o HEVC, VP9 a nejnovější AV1.

Tento obsah videosekvencí byl hodnocen podle subjektivní metody ACR-HR a objektivních metrik SSIM, PSNR a VMAF, které byly počítány programem FFMPEG verze 4.3.2 pomocí konzolových příkazů. Volby těchto metrik byly následující, PSNR je jeden z nejpoužívanějších metrik pro rekonstrukci obrazových kompresí kodeků. SSIM, který využívá predikci a oproti PSNR a odhaluje absolutní chyby. Poslední VMAF, který dokáže zachytit větší rozdíly oproti PSNR a umožňuje porovnávat kodeky v oblastech.

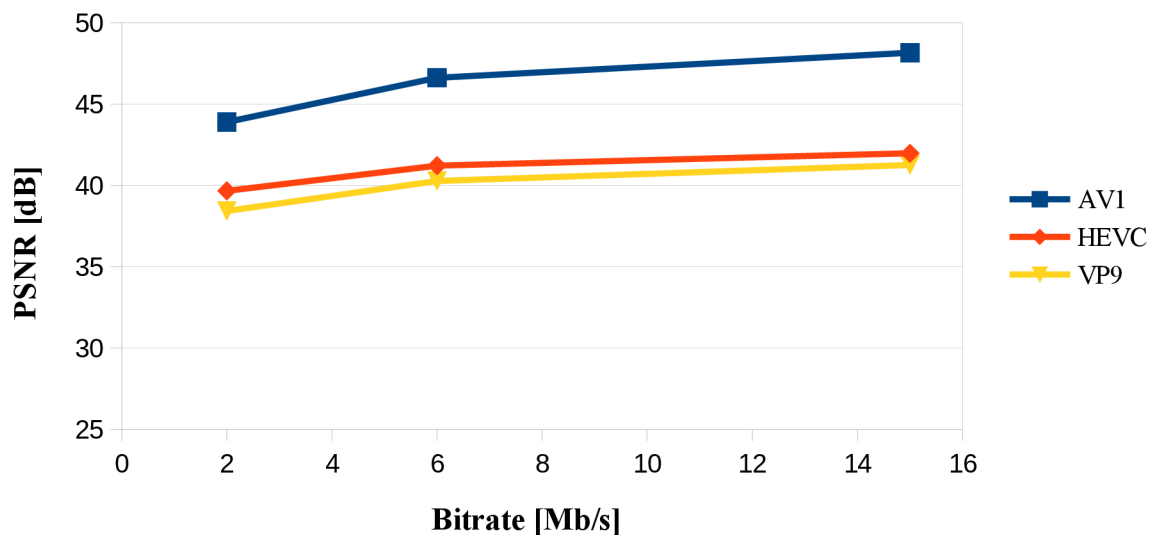
6.1 Výsledek objektivní kvality pomocí PSNR

Výsledkem PSNR (Peak Signal to Noise Ratio) je hodnota vyjádřena v decibelech [dB], kdy se hodnota pohybuje v našem případě od 28 dB do 48 dB. Čím je hodnota v decibelech vyšší tím je video i kvalitnější. Obvykle se tato hodnota pohybuje od 25 do 50 dB

Z výsledků lze vyčíst, že nejvhodnějším kodekem je AV1, který dosáhl nejvyšších hodnot u všech videí. Kodeky HEVC a AV9 jsou si v těchto výpočtech velmi podobné a jejich hodnoty jsou téměř totožné.

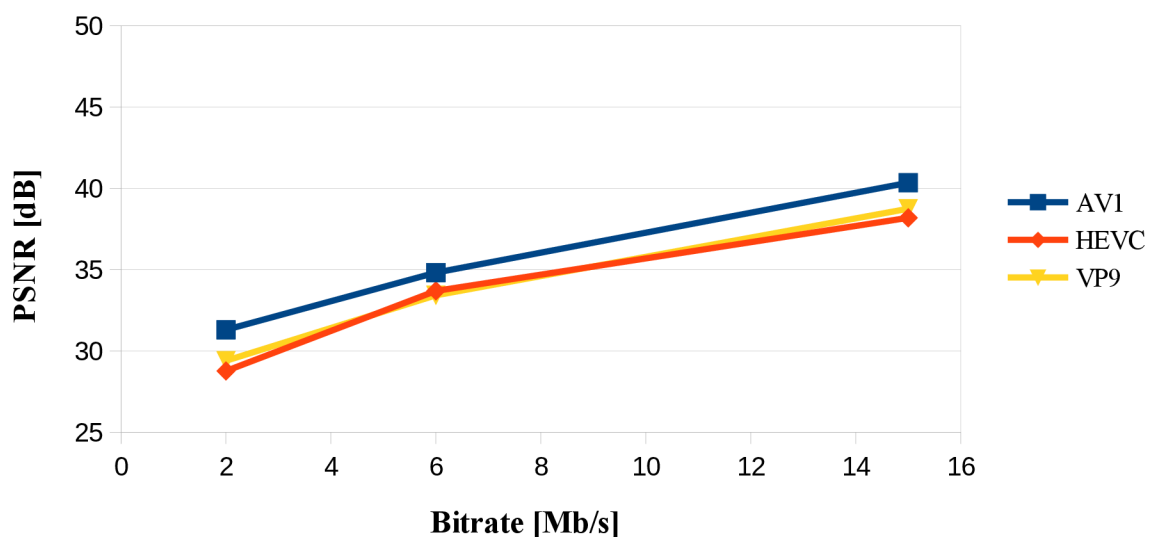
```
ffmpeg.exe -i videoToCompare.mp4 -i originalVideo.mp4 -lavfi  
psnr=stats_file=psnr_logfile.txt -f null -
```

Videosekvence CITY PSNR



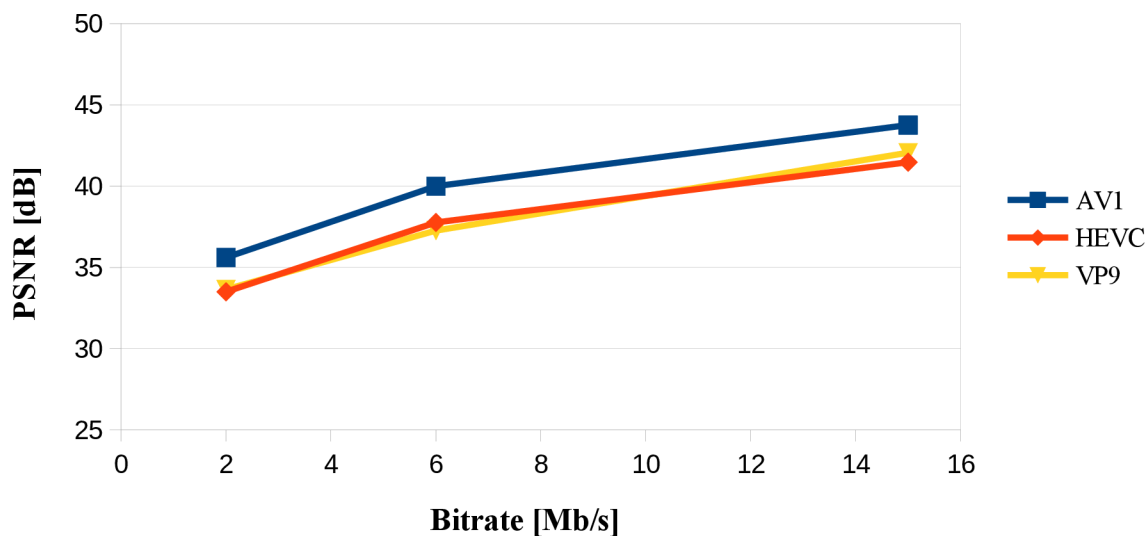
Obr 6.1: Závislost PSNR na bitové rychlosti CITY videosekvence

Videosekvence FOREST PSNR



Obr 6.2: Závislost PSNR na bitové rychlosti FOREST videosekvence

Videosekvence STREET PSNR



Obr 6.3: Závislost PSNR na bitové rychlosti STREET videosekvence

		Vypočítaná hodnota PSNR [dB]		
	Bitrate [Mb/s]	AV1	HEVC	VP9
Videosekvence CITY	2	35,606	33,499	33,645
	6	39,994	37,767	37,264
	15	43,756	41,471	42,053
Videosekvence FOREST	2	31,306	28,775	29,403
	6	34,823	33,702	33,431
	15	40,352	38,189	38,745
Videosekvence STREET	2	35,606	33,499	33,645
	6	39,994	37,767	37,264
	15	43,756	41,471	42,053

Tab 6.1: Hodnoty výsledků objektivní metody PSNR pro všechny videosekvence

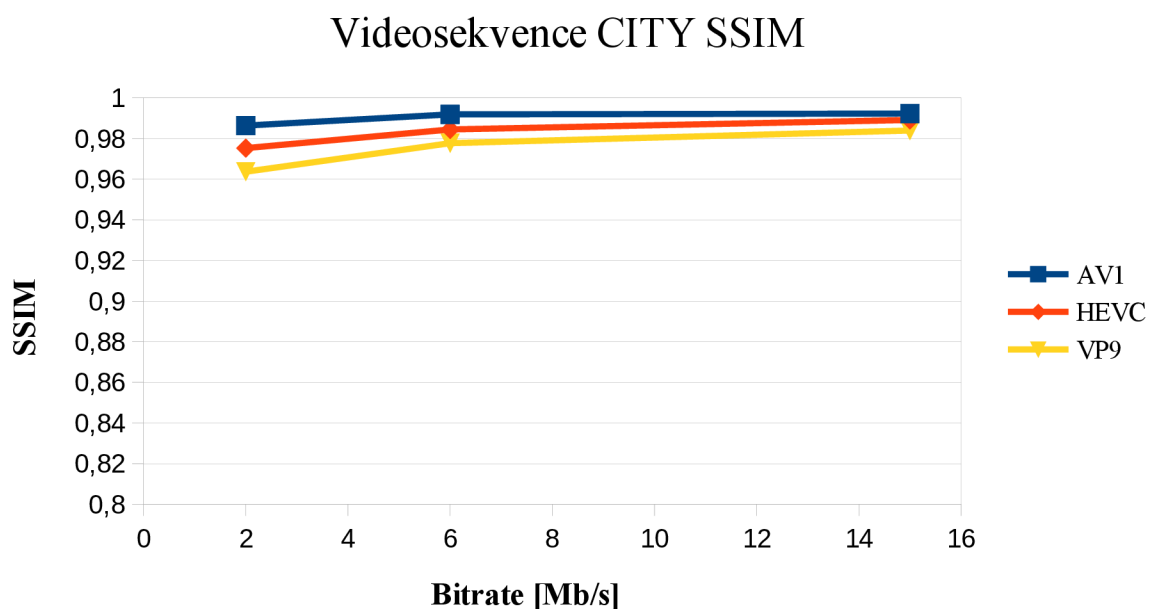
Z výsledků objektivní metody PSNR je patrné, že nejlepšími hodnotami dosáhl dle očekávání kodek AV1 a to ve všech změnách datového toku bez ohledu na typ videa. S porovnáním kodeku HEVC a VP9 vidíme, že jejich hodnoty jsou velmi podobné, ale při datových tocích 6 Mb/s dosahuje HEVC patrně lepších výsledků oproti VP9. Naopak u nízkého bitratu dosahuje lepších hodnot VP9 a to také při vysokých datových tocích 15 Mb/s, kde je už rozdíl patrnější kolem půl decibelu.

6.2 Výsledek objektivní kvality pomocí SSIM

Hodnoty SSIM se mohou pohybovat v rozmezí -1 (maximální rozdíl) až 1 (žádný rozdíl). Upravená videa jsou porovnávána s originálním nekomprimovaným videem. SSIM využívá strukturální zkreslení na základě rozdílu hodnot jasu. Metrika neporovnává hodnoty pixelů, ale obrazové prvky vnímané člověkem. Hodnotí se základní tři složky jas, kontrast a struktura.

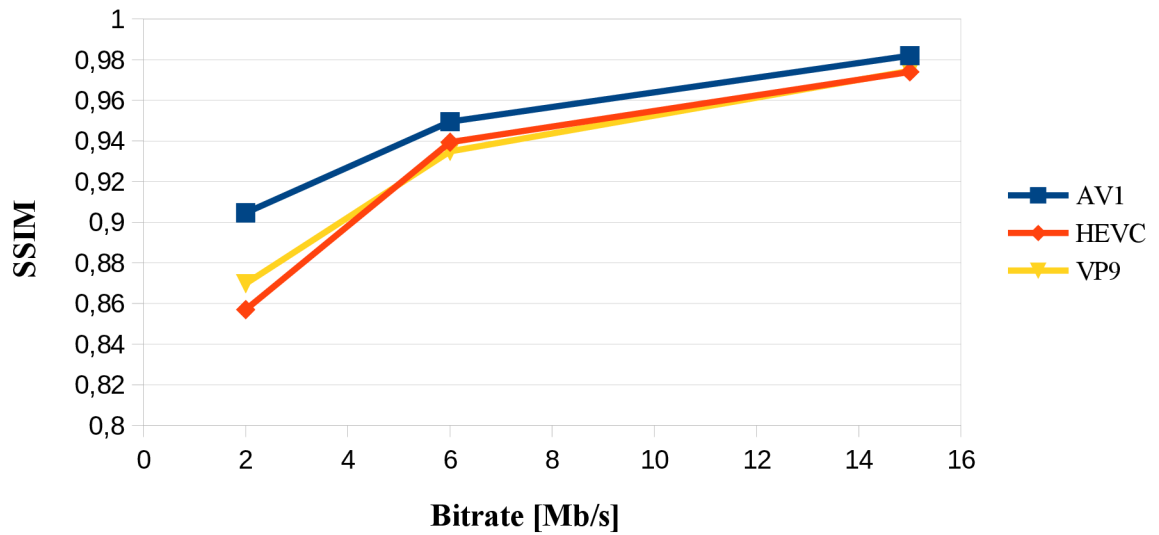
Můžeme opět říci, že kodek AV1 ve výpočtech SSIM opět vytváří nejlepší hodnoty této objektivní metody a teoreticky nejlepší prostředí pro diváka. Zbylé dva kodeky opět vychází velmi podobně a jsou mezi nimi jen malé rozdíly.

```
ffmpeg.exe -i videoToCompare.mp4 -i originalVideo.mp4 -lavfi  
ssim=stats_file=ssim_logfile.txt -f null -
```



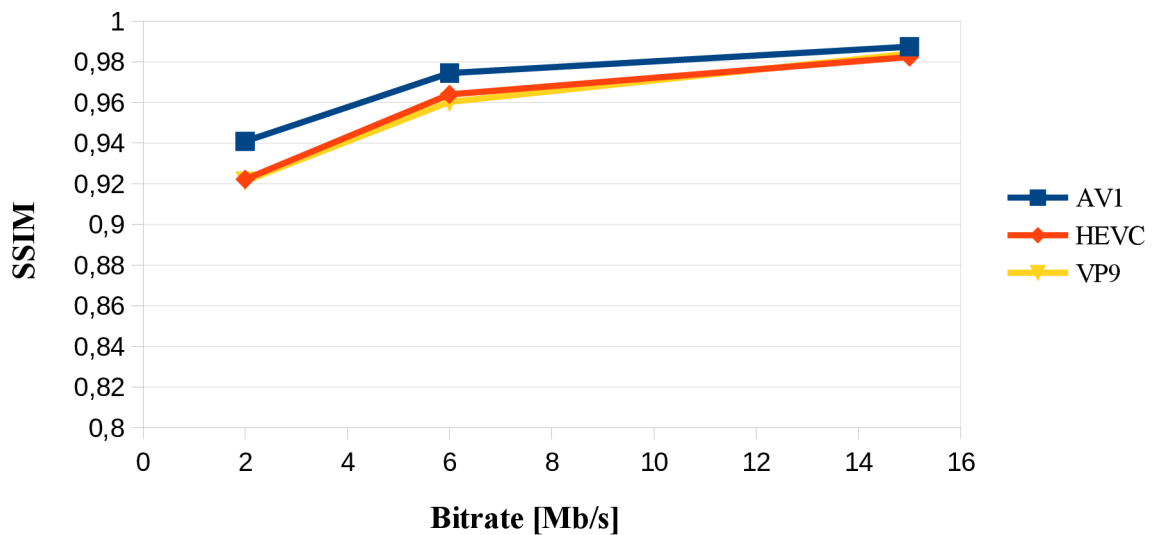
Obr 6.4: Závislost SSIM na bitové rychlosti CITY videosekvence

Videosekvence FOREST SSIM



Obr 6.5: Závislost SSIM na bitové rychlosti FOREST videosekvence

Videosekvence STREET SSIM



Obr 6.6: Závislost SSIM na bitové rychlosti STREET videosekvence

	Bitrate [Mb/s]	Vypočítaná hodnota SSIM		
		AV1	HEVC	VP9
Videosekvence CITY	2	0,986	0,975	0,963
	6	0,991	0,984	0,977
	15	0,992	0,989	0,983
Videosekvence FOREST	2	0,904	0,856	0,869
	6	0,949	0,939	0,934
	15	0,981	0,974	0,974
Videosekvence STREET	2	0,940	0,922	0,921
	6	0,974	0,963	0,960
	15	0,987	0,982	0,984

Tab 6.2: Hodnoty výsledků objektivní metody SSIM pro všechny videosekvence

Vyhodnocení objektivní metody SSIM dosahuje opět nejlepších hodnot ve všech případech kodek AV1. Největší rozdíly mezi jednotlivými kodeky jsou u nízkých datových toků. Hodnoty u 15 Mb/s už nedosahují takových rozdílů a například u videosekvence STREET jsou výsledky SSIM pro všechny kodeky téměř srovnatelné. Kodeky VP9 a HEVC jsou celkově v hodnocení SSIM velmi podobné a jejich výsledky se liší minimálně se změnou typu videosekvence. Nejlépe se hodnoty SSIM určovali z videa CITY, protože se jednalo o nejkvalitnější videosekvenci a tedy výsledky byly nejpřesnější.

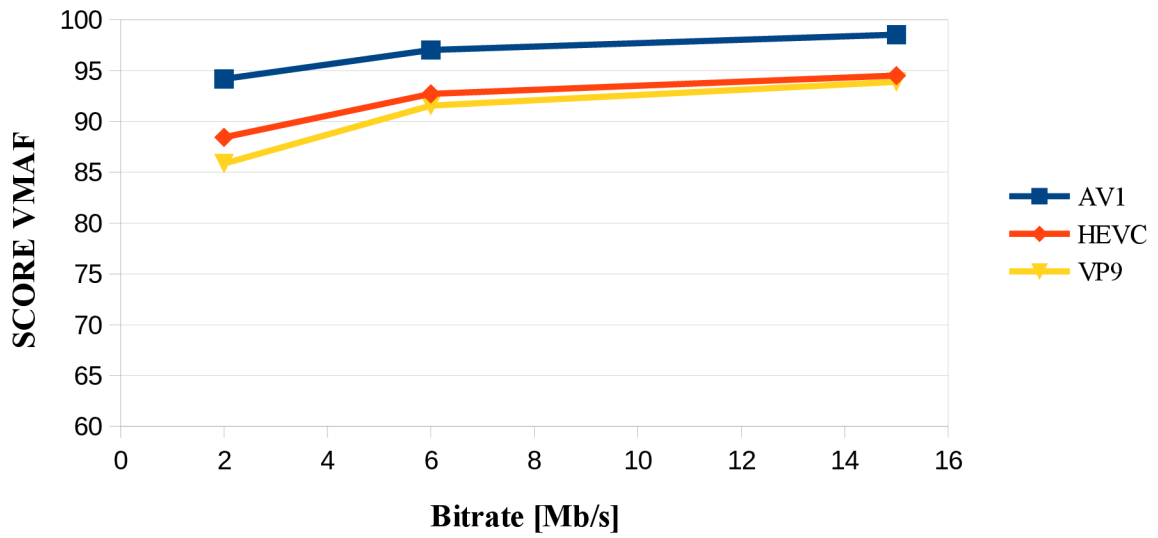
6.3 Výsledek objektivní kvality pomocí VMAF

Výsledkem objektivní metriky VMAF (Video Multi-method Assessment Fusion) je takzvané SCORE VMAF, které se pohybuje od 0 do 100, kde 0 označuje nejnižší kvalitu a 100 nejvyšší. Hodnoty byly počítány pomocí modelu pro VMAF 4K v0.6.1.

Nejlepší hodnot opět dosáhl kodek AV1, kdy se jeho hodnoty pohybují u vysokého datového toku okolo 98 VMAF SCORE.

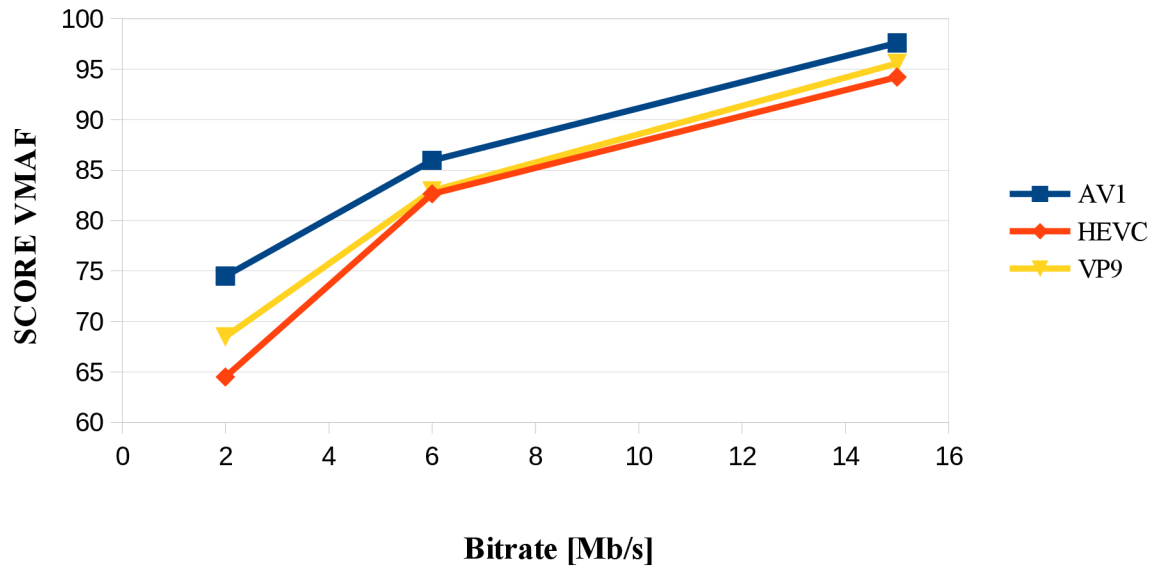
```
ffmpeg.exe -i videoToCompare.mp4 -i originalVideo.mp4 -lavfi
libvmaf="model_path=vmaf_4K_v0.6.1.pkl":log_path=vmaf_logfile.txt -f null -
```

Videosekvence CITY VMAF

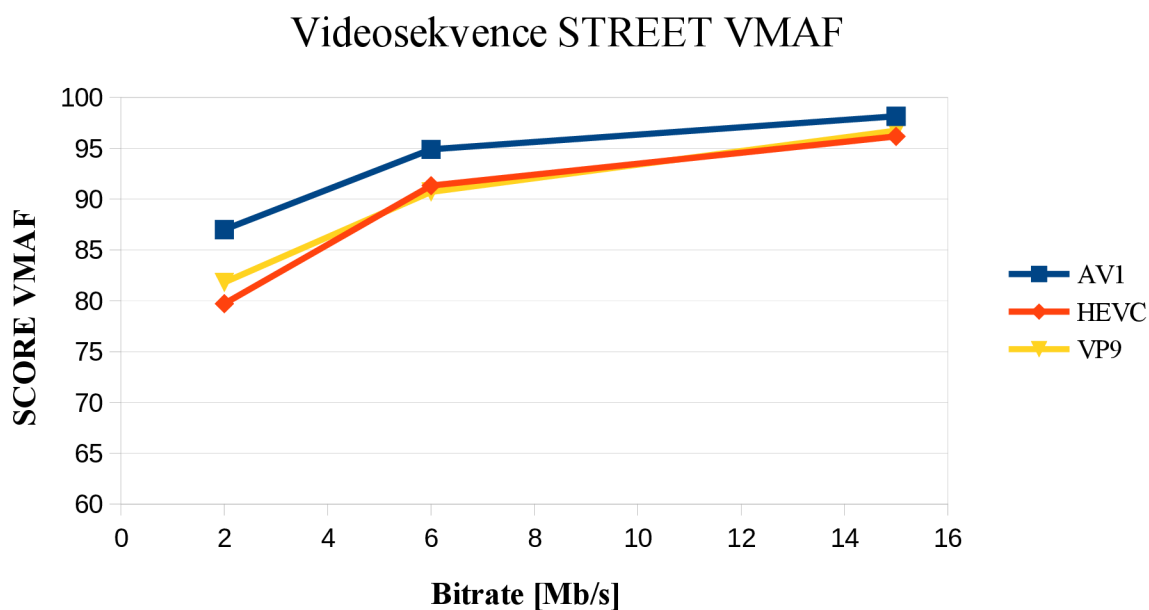


Obr 6.7: Závislost VMAF score na bitové rychlosti CITY videosekvence

Videosekvence FOREST VMAF



Obr 6.8: Závislost VMAF score na bitové rychlosti FOREST videosekvence



Obr 6.9: Závislost VMAF score na bitové rychlosti STREET videosekvence

	Bitrate [Mb/s]	Vypočítaná hodnota VMAF SCORE		
		AV1	HEVC	VP9
Videosekvence CITY	2	94,173	88,433	85,863
	6	97,032	92,712	91,563
	15	98,540	94,505	93,894
Videosekvence FOREST	2	74,496	64,503	68,478
	6	85,962	82,650	82,950
	15	97,605	94,211	95,575
Videosekvence STREET	2	87,013	79,724	81,809
	6	94,897	91,330	90,716
	15	98,150	96,170	96,740

Tab 6.3: Hodnoty výsledků objektivní metody VMAF pro všechny videosekvence

Metrika Netflixu pro 4K videa dosáhla nejlepších výsledků VMAF score pro kodek AV1. U nízkých bitratu dosahuje obrovských odchylek oproti HEVC a VP9 a to až necelých 10 VMAF score. Tento rozdíl se ztrácí se zvyšujícím se datovým tokem, ale i přes zvýšení

datového toku nemohou zbylé dva kodeky překonat nejnovější kodek AV1. Hodnoty HEVC dosahují lepších hodnot pouze u videa CITY, kde se opět se zvyšujícím datovým tokem rozdíl mezi kodeky zmenšuje. U videosekvencí FOREST a STREET dosahuje lepších výsledků kodek VP9, ale rozdíly už jsou menší.

6.4 Subjektivní testování kvality videosekvencí

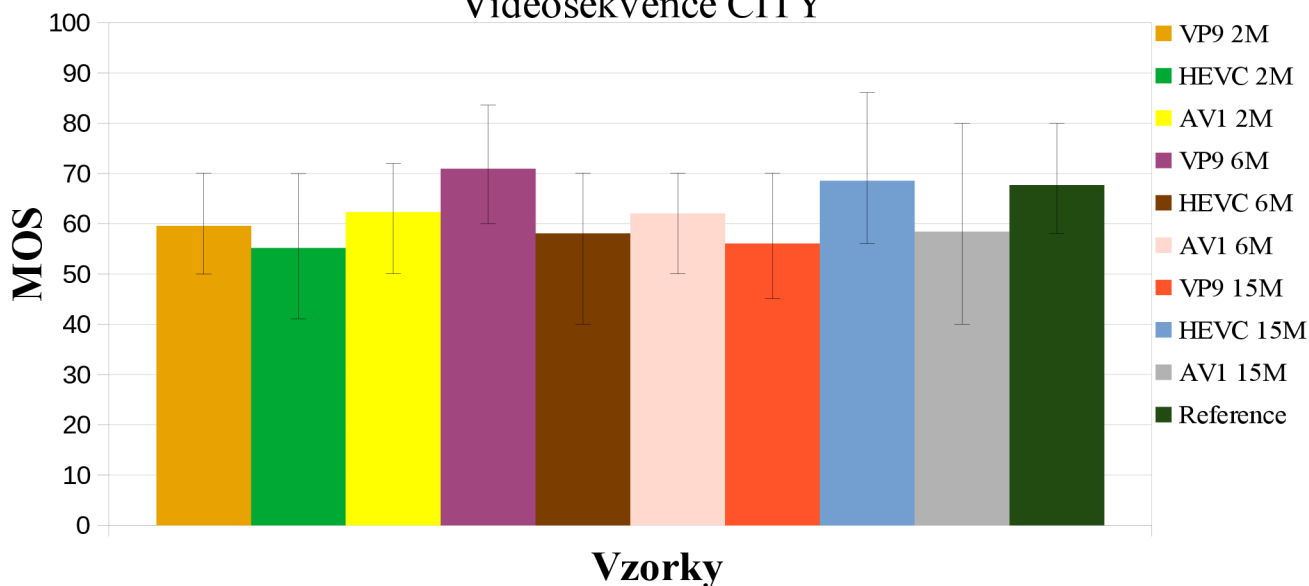
Bakalářská práce je zaměřena na tvorbu vhodného obsahu pro virtuální realitu a tedy nejpříjemnější zážitek ze samotného sledování. Z toho důvodu je zapotřebí využívat i subjektivního testování, které vytvoří reálný pohled osob na jednotlivá videa s různou kompresí a kodeky.

Subjektivní testování bylo prováděno s headsetem Oculus Go v programu SKYBOX VR Video Player vytvořený společností Source Studio, který byl volně dostupný ke stažení v knihovně Oculusu. Tohoto testování se účastnilo celkově 16 respondentů od 21 do 25 let (7 žen a 9 mužů), kteří byli testováni na jednu ze dvou variant kombinací upravovaných videí. Před testováním byli respondenti tázáni na otázky ohledně jejich kvality nejčastěji sledovaných videí a rozlišení zařízení, kterým si daná videa pouští. Téměř všichni testovaní neměli zkušenost s virtuálním headsetem, až na jednu výjimku. Podle dotazovaných nebyl nikdo, kdo by měl doma obrazovku nebo televizi o vyšším rozlišení, než Full HD (1920x1080).

Hodnotitelé byli seznámeni s požadavky na výsledky, kdy se hodnotí kvalita a jejich zážitek ze samotného videa. Respondenti následně postupně hlásily hodnoty, postupně za sebou, jak byly v dané variantě poskládány.

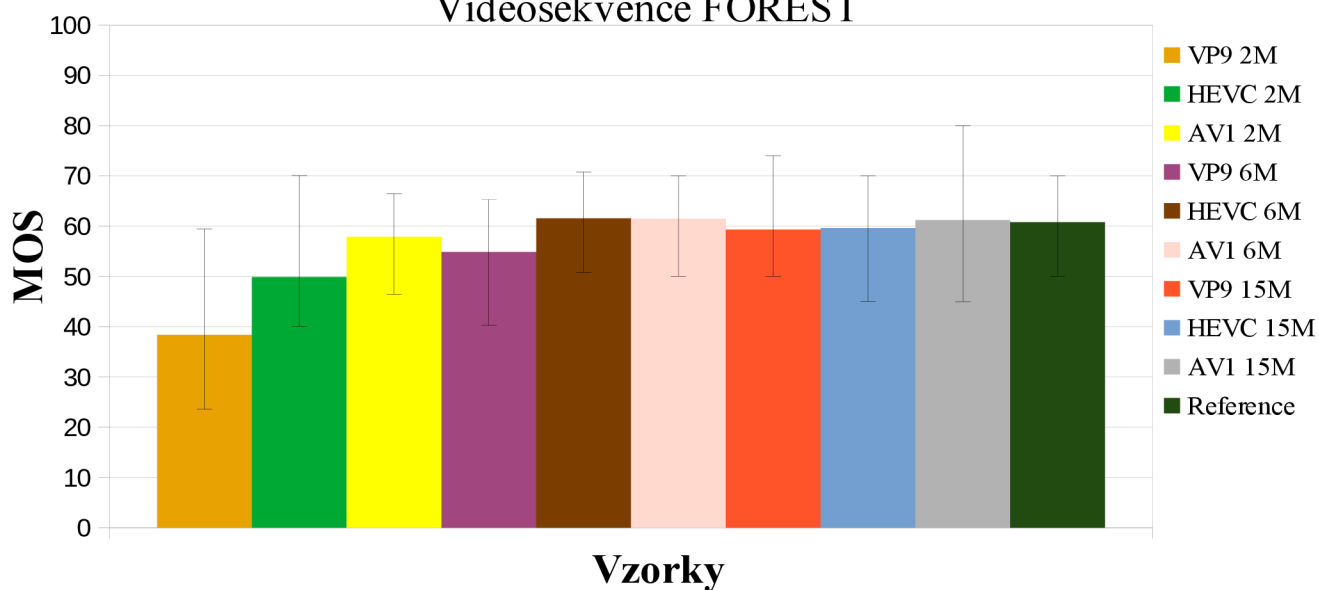
Při subjektivním testování došlo k problému v kodekem AV1, který přehrávač SKYBOX VR nemohl přehrát. Bohužel se jednalo o problém samotného VR setu, který tento kodek nepodporuje. Z toho důvodu bylo každé z videí, které bylo ve formátu AV1 překódováno pomocí kodeku HEVC a byl navýšen datový tok, aby kvalita odpovídala videu v AV1. Toto nastavení jsem prováděl pomocí objektivní metody PSNR a SSIM, kdy jsem měnil datový tok videa, abych dosáhl stejných hodnot kvality jako u AV1 [39]. Hodnoty navýšení u jednotlivých videí byly různé v rozmezí 1,5Mb/s – 7Mb/s.

Výsledek subjektivního hodnocení Videosekvence CITY

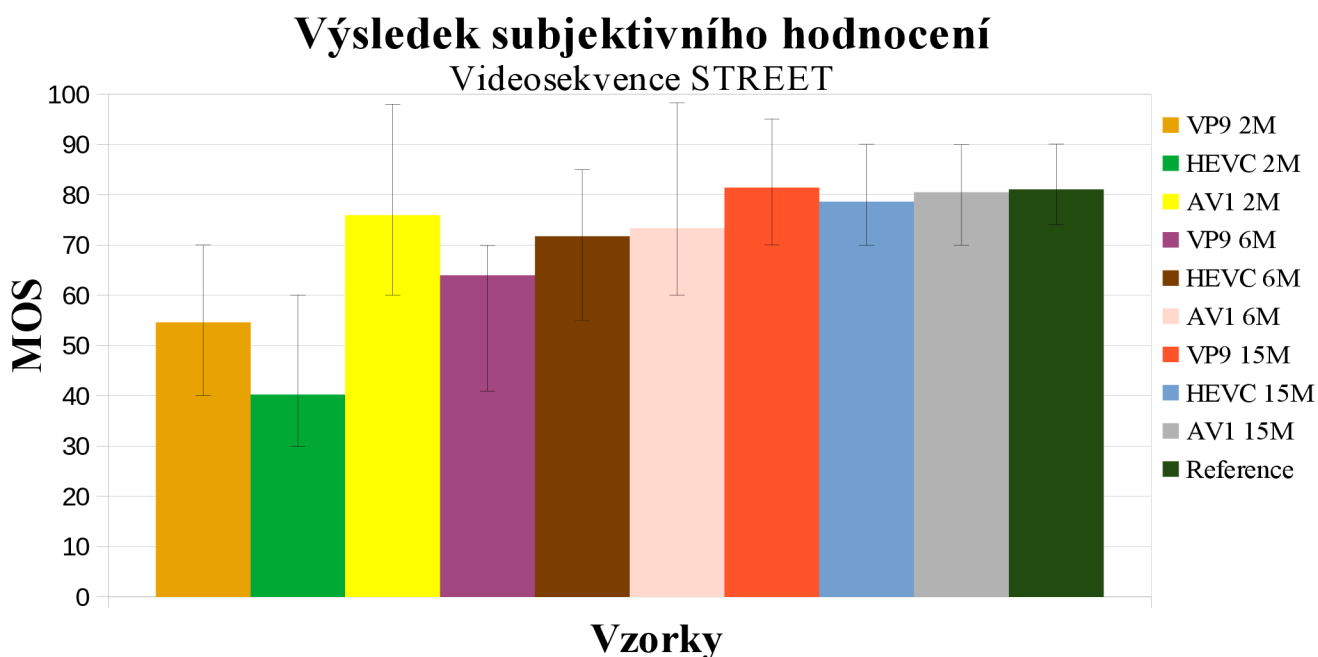


Obr 6.10: Výsledky MOS pro CITY videosekvenci

Výsledek subjektivního hodnocení Videosekvence FOREST



Obr 6.11: Výsledky MOS pro FOREST videosekvenci



Obr 6.12: Výsledky MOS pro STREET videosekvenci

Jak je již patrné z naměřených výsledků, tak referenční video nebylo hodnoceno podle očekávání. Respondenti dávali vysoká hodnocení i videím s nižší kvalitou obrazu. Bohužel to mělo za následek nemožnost použít hodnocení DMOS, neboť by výsledky byly zkreslené a nepřesné.

Video CITY bylo točeno za pomoci dronu a jedná se o dynamický záběr. Při testování je záběr na most, který při pohybu kamery začal v brýlích pro virtuální realitu vytvářet rušivé pruhy, které na respondenty negativně působily. Tento problém nastal u tenkých rovných liniích. Problém byl způsoben samotnými zvětšovacími skly brýlí, které vytvářely spektrální čáry a jeho odstranění nebylo možné. Tato vada se objevovala u více videí typu flat a to zejména u velmi světlých záběrů. Video FOREST je statické a k rušivým pruhům zde nedochází. Ovšem dochází při nízkých kvalitách k splývání zelené barvy, která ovlivňuje hodnocení respondenta. STREET videosekvence je opět statické, točeno z jednoho bodu a respondent pozoruje obraz kolem sebe bez pohybu.

Video CITY, které bylo také nejkvalitnější s nejvyšším snímkovým kmitočtem 60 dosahovalo při hodnocení velmi podobných hodnot pro většinu výsledků. Uživatelé si i při datovém toku 2Mb/s nebyli schopni všimnout změn v kvalitě. I když dostal kodek AV1 nejvyšší průměrné hodnocení, rozptýl VP9 a HEVC byl velmi podobný. Při zvýšení datového toku na hodnotu 6Mb/s dosáhl nejlepších výsledků kodek VP9. Dokonce jej respondenti hodnotili lépe, než referenční video ve stejném formátu. Tuto skutečnost mohl ovlivnit i rušivý záběr na most. Z toho jsem vyvodil závěr, že od hodnot 6Mb/s je již pro nezkušené oko hranice, kdy ve flat videu nevidí změny v obraze při zvyšujícím se datovém toku. Také se soustředil na videa, které neobsahují tenké světlé linie nejčastěji světlá obloha, odrazy světla od vodní hladiny nebo záběry na vysoké prosklené budovy.

Video STREET bylo hodnoceno velmi dobře a to i při nižších kvalitách. Vzhledem k tomu, že většina zúčastněných hodnotitelů viděli 360 stupňové video poprvé, tak jejich

hodnocení bylo už od 2Mb/s vysoké i když u videí v HEVC nerozeznávali ani obličej nebo byly vidět glitche (kostičkování obrazu). V tomto hodnocení opět dopadl nejlépe kodek AV1 a to i pro 6Mb/s, kde se vyrovnal referenčnímu videu. Zde už hodnotitelé viděli malé rozdíly a bylo pro ně jednodušší udávat vhodné hodnocení. Ovšem od hodnot 15Mb/s již respondenti nepoznali rozdíl mezi kodeky a byl tedy obraz pro všechny téměř totožný.

Video FOREST, které bylo ve 3D (snímek pro každé oko) působilo na respondenty velmi podobně pro všechny datové toky. I když u nízkých toků 2Mb/s je viditelné zhoršení hodnocení respondentů a to zejména při splývání zelených barev, které se na videosekvenci objevují. Hodnoty u formátů HEVC a AV1 jsou již velmi blízké referenčnímu snímku. Od 6Mb/s již vidíme téměř konstantní hodnotu MOS. Tedy můžeme z výsledků říci, že při 3D videu datového toku 6Mb/s nepozná respondent rozdíl. Dalším aspekt, který mohl dané měření ovlivnit byl fakt, že s 3D videem nemá spousta lidí zkušenosti a při promítání je samotný obraz spojen v ohnisku očí, které má každý člověk jinak. U headsetu Oculus Go jsou ještě dvě zvětšovací skla, která mohla ovlivnit hodnocení respondenta.

6.5 Doporučení pro tvorbu videoobsahu ve VR

Všechna doporučení jsou vyhodnocena ze subjektivní metody ACR-HR a objektivních metod PSNR, SSIM, VMAF z předchozích kapitol přehrávané přístrojem Oculus Go.

Testování subjektivní metody se zúčastnili zejména respondenti, kteří nemají k dispozici kvalitní televizi na přehrávání videí ve 2K/4K/8K, anebo nepřišli do styku s virtuální realitou. To mělo za následek jejich očekávání, s tím spojeno hodnocení videosekvencí. Můžeme tedy shrnout, že pro tvorbu flat videa v rozlišení 4K nebyl pro většinu respondentů patrný úbytek kvality u nízkých datových toků, a tedy je dostačující využít hodnot kolem 2Mb/s. U testování 3D videa můžeme stanovit hodnotu alespoň 6Mb/s, aby nebyl patrný rozdíl. Tímto krokem ušetříme velikost souboru a není zapotřebí zanechávat vysoký datový tok. Nakonec doporučení pro 360 stupňové video, které potřebuje co nejvyšší datový tok, protože i nezkušený pozorovatel si všimne nedokonalostí v takto velkém snímku. Zde se podobné hodnocení pohybuje od datového toku 15Mb/s, a tedy se zvyšujícím se datovým tokem přestane respondent vnímat zlepšení obrazu.

U použitého kodeku jsou nevhodnější výsledky u AV1, ale jak již bylo zmíněno, tak daný formát není nativně podporován přehrávačem Oculus Go. A tudíž není vhodný z hlediska implementace. Ve srovnání použití HEVC nebo VP9 nebyl pozorován výrazný rozdíl, neboť výsledky vyšly dost podobně.

7 ZÁVĚR

Cílem bakalářské práce bylo vytvořit vhodná doporučení pro tvorbu multimediálního obsahu do virtuální reality.

V první části byly popsány typy brýlí pro virtuální realitu, které jsou volně dostupné na českém trhu. Subjektivní a objektivní metody hodnocení kvality multimediálního obsahu. Byly uvedeny kodeky na kompresi videí VP9, HEVC a AV1 a jejich postup při kódování komprese. Dále byly zmíněny formy videí a fotografií a jakých technologií se pro dané medium užívá pro jejich získání.

V práci je popsán postup získávání videí s vysokým rozlišením 4K a datovým tokem přes 23 Mb/s. Volba tří různých typů videí konkrétně flat (CITY), 3D (FOREST) a 360 stupňové (STREET), aby bylo možné porovnat nejvhodnější obsah. Následná úprava těchto videí s použitím programu FFmpeg, zde se ořezala jejich délka na rozmezí 15-18s, datový tok 2, 6, 15 Mb/s a formát na VP9, HEVC a AV1. Je zde popsána volba a navržení subjektivní metody ACR-HR, její použití pro kódování videa a zpracování výsledků.

V další části jsem se zaměřil na testování videosekvencí formou mnou zvolených objektivních metod PSNR, SSIM a metodu netflixu VMAF, které byly graficky a tabulkově zpracovány. Celý proces hodnocení byl prováděn v programu FFmpeg. Upravené videosekvence byly implementovány do virtuálních brýlí Oculus Go. Testu se zúčastnilo 16 osob, které hodnotily videa 0-100 podle subjektivní metody ACR-HR. Výsledky byly zaznamenány a zpracovány v tabulkovém programu. Výsledky pomocí MOS byly zpracovány a provedeno vyhodnocení. Nakonec byly popsány doporučení pro tvorbu videoobsahu.

Nejlepších výsledků objektivních metod dosahoval s přehledem kodek AV1. Kodeky HEVC a VP9 se lišily minimálně pro různé metody nebo datové toky, ale i tak dosahovaly podobných hodnot. Pro subjektivní testy byla nejlépe hodnocena videosekvence STREET a byly zde i pozorovány největší odchylky se změnou datového toku. Videosekvence CITY a FOREST dosahují velmi podobných hodnot a to i pro různé datové toky a změnu kodeku.

POUŽITÉ ZDROJE A LITERATURA:

- [1] Virtuální realita — od počátku po současnost. EDTECH KISK [online]. Brno: Masarykova Univerzita, 2019 [cit. 2019-12-12]. Dostupné z: <https://medium.com/edtech-kisk/virtu%C3%A1ln%C3%AD-realita-od-po%C4%8D%C3%A1tku-po-sou%C4%8Dasnost-5d3c31d2fd92>
- [2] Virtual Reality and Augmented Reality. Londýn: ISTE Ltd., Wiley, 2018. ISBN 978-1-78630-105-5.
- [3] VR headsets and ar glasses – product roundups, rankings and lists of the best products. ANIWAA [online]. -: Steve Noble, Ludivine Cherdo, 2019 [cit. 2019-12-12]. Dostupné z: <https://www.aniwaa.com/best-of/vr-ar/>
- [4] Go Overwiev. Oculus [online]. -: Oculus, 2018 [cit. 2019-12-12]. Dostupné z: <https://www.oculus.com/go/>
- [5] Rift S Overwiev. Oculus [online]. -: Oculus, 2019 [cit. 2019-12-12]. Dostupné z: <https://www.oculus.com/rift-s/>
- [6] Quest Overwiev. Oculus [online]. -: Oculus, 2019 [cit. 2019-12-12]. Dostupné z: <https://www.oculus.com/quest/>
- [7] Tvorba a demonstrace 3D modelů pro VR [online]. Brno, 2017 [cit. 2019-12-12]. Dostupné z: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/69546/20131.pdf?sequence=2&isAllowed=y>. Diplomová práce. VUT Brno.
- [8] Acer představil Windows Mixed Reality headset pro čistotné. FZONE [online]. -: Martin Fajmon, 2018 [cit. 2019-12-12]. Dostupné z: <https://fzone.cz/clanky/acer-predstavil-windows-mixed-reality-headset-pro-cistotne-354>
- [9] Pimax 8K VR Headset 2. generace. PC HRAC.CZ [online]. -: JIBRO, 2017 [cit. 2019-12-12]. Dostupné z: <http://www.pchrac.cz/pimax-8k-vr-headset-2-generace/>
- [10] Methodologies for the subjective assessment of the quality of television images. ITU-R [online]. -: ITU, 2019 [cit. 2019-12-12]. Dostupné z: https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-14-201910-I!!PDF-E.pdf
- [11] Moderní kompresní formát digitálního obrazu [online]. Brno, 2014 [cit. 2019-12-12]. Dostupné z: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/34257/Moderni%20Kompresni%20Formaty%20Digitalniho%20obrazu.pdf?sequence=2&isAllowed=y>. Bakalářská práce. VUT Brno.
- [12] Měření kvality pro HEVC [online]. Brno, 2014 [cit. 2019-12-12]. Dostupné z: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/34257/Moderni%20Kompresni>

%20Formaty%20Digitalniho%20obrazu.pdf?sequence=2&isAllowed=y. Diplomová práce. VUT Brno.

[13] Srovnání standardů VP9 a H.265 pro kódování videosekvencí [online]. Brno, 2017 [cit. 2019-12-12]. Dostupné z: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/34257/Moderni%20Kompresni%20Formaty%20Digitalniho%20obrazu.pdf?sequence=2&isAllowed=y>. Bakalářská práce. VUT Brno.

[14] AOM - AV1. PARIS VIDEO TECH [online]. -: Pascal Massimino - Google, 2017 [cit. 2019-12-12]. Dostupné z: <https://parisvideotech.com/wp-content/uploads/2017/07/AOM-AV1-Video-Tech-meet-up.pdf>

[15] Srovnání moderních formátů pro ztrátovou kompresi obrazu [online]. Brno, 2017 [cit. 2019-12-12]. Dostupné z: <https://dspace.vutbr.cz/xmlui/handle/11012/180159>. Bakalářská práce. VUT Brno.

[16] Kadrowanie zdjęć panoramicznych. Optyczne [online]. Pruszcz: Anna Benicewicz-Miazga, Ernest Klauzinski, Anna Góra, 2011 [cit. 2019-12-12]. Dostupné z: https://www.optyczne.pl/58.1-poradnik-Kadrowanie_zdj%C4%99%C4%87_panoramicznych_Wst%C4%99p.html

[17] 3D Photography: A Beginner's Guide. Format [online]. -: Format team, 2018 [cit. 2019-12-12]. Dostupné z: <https://www.format.com/magazine/resources/photography/3d-photos>

[18] Jak fotit magická 360° panoramata a virtuální prohlídky. Fototypy [online]. -: JANVRSINSKY, 2012 [cit. 2019-12-12]. Dostupné z: <http://www.fototypy.cz/2012/06/08/jak-fotit-magicka-360-panoramata-a-virtualni-prohlidky/>

[19] How to Shoot 360-Degree Photos With Your Smartphone. GIZMODO [online]. -: David Nield, 2016 [cit. 2019-12-12]. Dostupné z: <https://gizmodo.com/how-to-shoot-360-degree-photos-with-your-smartphone-1783077083>

[20] What is a VR video? Quora [online]. -: Emory Craig, 2017 [cit. 2019-12-12]. Dostupné z: <https://www.quora.com/What-is-a-VR-video>
Ukončit editaci Zkopírovat citaci

[21] 360 Video. Digital bodies [online]. -: Emory Craig, 2019 [cit. 2019-12-12]. Dostupné z: <https://www.digitalbodies.net/360-video/>

[22] How a 360 camera works. Camera Jabber [online]. -: Jeff Meyer, 2018 [cit. 2019-12-12]. Dostupné z: <https://camerajabber.com/how-a-360-camera-works/>

[23] How do you convert 3D video to a virtual reality experience? Quora [online]. -: John L. Miller, 2016 [cit. 2019-12-12]. Dostupné z: <https://www.quora.com/How-do-you-convert-3D-video-to-a-virtual-reality-experience>

[24] Subjective video quality assessment methods for multimedia applications. <https://www.itu.int> [online]. ITU, 2008, 6. dubna 2008 [cit. 2019-12-09].

Dostupné z: https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-P.910-200804-I!!PDF-E&type=items

[25] Objektivní hodnocení kvality videa v prostředí [online]. Brno, 2011 [cit. 2019-12-12]. Dostupné z: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/6607/Objektivni%20hodnoceni%20kvality%20videa%20v%20prostredi%20MATLAB.pdf?sequence=1&isAllowed=y>. Bakalářská práce. VUT Brno.

[26] Srovnání výkonnosti nejnovějších standardů pro kódování [online]. Brno, 2012 [cit. 2019-12-12]. Dostupné z: https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/21360/Bakalarska_prace.pdf?sequence=1&isAllowed=y. Bakalářská práce. VUT Brno.

[27] Buying and Using a Fisheye Lens. Jossops [online]. -: Mark Higham, 2016 [cit. 2019-12-12]. Dostupné z: <https://www.jessops.com/c/advice/bg/fisheye-lenses>

[28] What are the advantages and disadvantages of 360 degree videos? Quora [online]. -: Guna, 2017 [cit. 2019-12-12]. Dostupné z: <https://www.quora.com/What-are-the-advantages-and-disadvantages-of-360-degree-videos>

[29] Encoding High-Resolution 360 and 180 Video for Oculus Quest and Oculus Go. Oculus Creators [online]. -: Eric Cheng, 2019 [cit. 2019-12-15]. Dostupné z: <https://creator.oculus.com/blog/encoding-high-resolution-360-and-180-video-for-oculus-go/>

[30] Pimax opens preorders for its very expensive 8K and 5K VR headsets. The Verge [online]. -: Chris Welch, 2018 [cit. 2019-12-12]. Dostupné z: <https://www.theverge.com/2018/10/24/18019254/pimax-8k-5k-vr-headset-preorders-now-available-features-price>

[31] Frame Rate: A Beginner's Guide. TechSmith [online]. -: DOUG BRUNNER, 2017 [cit. 2019-12-15]. Dostupné z: <https://www.techsmith.com/blog/frame-rate-beginners-guide/>

[32] 10-bit Video (QooCam 8K- 8k 30fps demo). KanDao [online]. -: KanDao, 2019 [cit. 2019-12-15]. Dostupné z: <https://www.kandaovr.com/gallery/>

[33] High Quality Aerial Video. *Playboard* [online]. Soul: 8K HDR CHANNEL, 2017, 20.11.2019 [cit. 2021-5-25]. Dostupné z: <https://playboard.co/en/video/RGoPU-OLQHE>

[34] 3D 4K Video 2160p: May forest walk. *3d-taeuber* [online]. Passau: 8K HDR CHANNEL, 2015, 2015 [cit. 2021-5-25]. Dostupné z: <http://3d-taeuber.de/index.html>

[35] Platforma pro subjektivní hodnocení video sekvencí [online]. Brno, 2016 [cit. 2019-12-16]. Dostupné z: <https://dspace.vutbr.cz/xmlui/bitstream/handle/11012/61590/Platforma%20pro%20subjektivni%20hodnoceni%20videosekvenci.pdf?sequence=2&isAllowed=y>. Bakalářská práce. VUT Brno.

[36] VMAF: The Journey Continues. *THE NETFLIX TECH BLOG*. California: Netflix Technology Blog, 2018, 25.10.2018, s. 1. Dostupné také z: <https://netflixtechblog.com/vmaf-the-journey-continues-44b51ee9ed12>

[37] VMAF: The Journey Continues. *FFmpeg*. France: FFmpeg, 2013, 20.02.2021, Dostupné také z: <https://www.ffmpeg.org/>

[38] Oculus Quest 2. *Oculus for developers*. Seattle: Facebook, 2020, 25.10.2018, , 1. Dostupné také z: <https://developer.oculus.com/learn/oculus-device-specs/>

[39] O. Zach, M. Seufert, M. Hirth, M. Slanina and P. Tran-Gia, "On use of crowdsourcing for H.264/AVC and H.265/HEVC video quality evaluation," 2017 27th International Conference Radioelektronika (RADIOELEKTRONIKA), Brno, Czech Republic, 2017, pp. 1-6, doi: 10.1109/RADIOELEK.2017.7937581.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

VR	Virtual reality
AR	Augmented Reality
3D	Trojrozměrný
HMD	Head mounted display
LCD	Liquid crystal display
CPU	Procesor
USB	Universal Serial Bus
3DOF	3 degree of freedom
6DOF	6 degree of freedom
OLED	Organic light-emitting diode
DSIS	Double-Stimulus Impairment Scale
DSCQS	The Double-Dstimulus Continuous Quality-Scale
SS	Single-Stimulus
SSCQE	Single Stimulus Continuous Quality Evaluation
PC	Pair Comparison method
ACR	Absolute Category Rating
ACR-HR	Absolute Category Rating hidden reference
MSE	mean square error
PSNR	peak signal-to-noise ratio
SSIM	Structural Similarity Index Measure
MSSIM	Mean Structural Similarity Index
VMAF	Video Multi-method Assessment Fusion
NC	Normalized Correlation
AD	Average Difference
SC	Structural Content
MD	Maximum Difference
NAE	Normalized Absolute Error
HEVC	High Efficiency Video Coding
H.264	High 264
Cr	Chrominanční složka barevného prostoru Y Cr Cb
Cb	Chrominanční složka barevného prostoru Y Cr Cb
JCT-VC	Joint Collaborative Team on Video Coding
ITU	International Telecommunication Union
CTU	Coding tree unit
CTB	Coding tree block
CU	Coding unit
CB	Coding block
PB	Prediction block
TB	Transform block
PU	Prediction unit
SAO	Sample adaptive offset
TU	Transform unit
AV1	AOMedia Video 1
V	Vertikal
H	Horizontal

TM	Thuemotion
ADTS	Asymetricka diskretní sinusova transformace
DCT	diskretní kosinova transformace
WHT	Walsh-Hadamardova transformace
DMOS	Differential Mean Opinion Score
MOS	Mean Opinion Score

SEZNAM PŘÍLOH

A Referenční a upravené videosekvence

B Varianty a hodnocení videosekvencí respondenty